ANDRÉ KAZUO TAKAHATA

UNIDIMENSIONAL AND BIDIMENSIONAL SEISMIC DECONVOLUTION

DECONVOLUÇÃO SÍSMICA UNIDIMENSIONAL E BIDIMENSIONAL

CAMPINAS
2014

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

ANDRÉ KAZUO TAKAHATA

UNIDIMENSIONAL AND BIDIMENSIONAL SEISMIC DECONVOLUTION

Advisor: Prof. Dr. Renato da Rocha Lopes

DECONVOLUÇÃO SÍSMICA UNIDIMENSIONAL E BIDIMENSIONAL

Doctorate thesis submitted to the Electrical Engineering Graduate Program of the School of Electrical and Computer Engineering of the University of Campinas to obtain the Ph.D. grade in Electrical Engineering, in the area of Telecommunications and Telematics.

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas para obtenção do título de Doutor em Engenharia Elétrica, na área de Telecomunicações e Telemática.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE
DEFENDIDA PELO ALUNO ANDRÉ KAZUO TAKAHATA
E ORIENTADO PELO PROF. DR. RENATO DA ROCHA LOPES

Assinatura do Orientador

_____

CAMPINAS
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

T139u

Takahata, André Kazuo, 1982-
Unidimensional and bidimensional seismic deconvolution / André Kazuo Takahata. – Campinas, SP : [s.n.], 2014.

Orientador: Renato da Rocha Lopes.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Processamento digital de sinais. I. Lopes, Renato da Rocha,1972-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Deconvolução sísmica unidimensional e bidimensional
Palavras-chave em inglês:
Signal processing - Digital techniques
Área de concentração: Telecomunicações e Telemática
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Renato da Rocha Lopes [Orientador]
Ricardo Tokio Higuti
Ricardo Suyama
Wu Shin-Ting
Romis Ribeiro de Faissol Attux
Data de defesa: 28-02-2014
Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: André Kazuo Takahata

Data da Defesa: 28 de fevereiro de 2014

Título da Tese: "Unidimensional and Bidimensional Seismic Deconvolution"

Prof. Dr. Renato da Rocha Lopes (Presidente): _____

Prof. Dr. Ricardo Tokio Higuti: _____

Prof. Dr. Ricardo Suyama: _____

Profa. Dra. Wu Shin-Ting: _____

Prof. Dr. Romis Ribeiro de Faissol Attux: _____

ABSTRACT

In this work, we consider signal processing techniques that aim to improve the resolution of images of the subsurface of the Earth generated from seismic data. One such technique is unidimensional deconvolution, which aims to eliminate distortions caused by limitations in the seismic source frequency band, as well as distorting effects caused by frequency components absorption and phase changes during seismic propagation. We analyze both supervised methods, in which reference signals are used in addition to the seismic measurements to determine the decovolution filter, as well as unsupervised methods, in which only the seismic measurements are used. Particularly, we analyze Wiener filtering and least squares methods on the supervised case. As for the unsupervised algorithms, we discuss the hypotheses that underlie these methods, which are based on the statistics of the reflectivity of the subsurface and the phase spectrum of the wavelet pulse. We analyze especially the use of the prediction error filter, which uses second order statistics (SOS) and requires a minimum phase wavelet, and we show our contribution on a method that uses higher order statistics (HOS) called banded independent component analysis (B-ICA), which does not requires that the wavelet be minimum phase. We also present a case study using log data measured in a borehole and seismic data in order to illustrate our analysis. In bidimensional deconvolution, we consider, besides the seismic source distortions considered in the 1D approach, distortions in seismic imaging caused by the acquisition geometry and velocity model complexity associated with the geological structure of the subsurface. These distortions can be quantified in seismic images created through the technique called prestack depth migration (PSDM) using a 2D convolution model between the reflectivity of the subsurface and the so-called resolution function. Under appropriate hypotheses, the resolution function can be seen as a point spread function (PSF). Thus, the objective of 2D deconvolution is to attenuate the effect of these PSFs. In this work, we review the basic aspects of the 2D convolutional model and PSF estimation, as well as the imaging process, and we show our contribution on 2D deconvolution using an inverse filtering approach.

Keywords: Signal processing - Digital techniques;

RESUMO

Neste trabalho consideramos técnicas de processamento de sinais que têm como objetivo aumentar a resolução de imagens da subsuperfície geradas com dados sísmicos. Uma das técnicas consideradas é a deconvolução unidimensional, que tem como finalidade eliminar distorções causadas pelas limitações em banda de frequência da fonte sísmica, bem como pela absorção de componentes e distorções de fase ocorridas durante a propagação da onda sísmica. Nesta tese, analisamos tanto métodos chamados supervisionados, em que estão disponíveis medidas adicionais às medidas sísmicas, que podem guiar o processo de deconvolução, quanto os métodos não supervisionados, em que apenas as medidas sísmicas são consideradas. Em particular, tratamos dos métodos de filtragem de Wiener e mínimos quadrados para os métodos supervisionados. Em relação aos métodos não supervisionados, discutimos as hipóteses para o funcionamento dos métodos envolvendo as estatísticas referentes à refletividade de subsuperfície e do espectro de fase do pulso sísmico. Em particular, analisamos principalmente o uso do filtro de erro de predição, que utiliza estatísticas de segunda ordem (SOS) e requer um pulso de fase mínima, e mostramos nossa contribuição sobre um método que utiliza estatísticas de ordem superior (HOS) chamado de "*banded independent component analysis*" (B-ICA) e que não exige que o pulso seja de fase mínima. Por fim, realizamos um estudo de caso envolvendo dados obtidos em um poço e dados sísmicos com fim de ilustrar nossa análise. Na deconvolução bidimensional são tratadas, além das distorções pela fonte sísmica consideradas na abordagem unidimensional, distorções causadas pela geometria de aquisição de dados e de variações de velocidade de propagação sísmica causadas por complexidades geológicas. Tais distorções podem ser quantificadas em imagens sísmicas obtidas pela técnica de imageamento chamada migração em profundidade pré-empilhamento (PSDM) por meio de uma relação de convolução bidimensional entre a refletividade da subsuperfície e uma função de resolução. Sob hipóteses adequadas, a função de resolução pode ser modelada como uma função de espalhamento pontual (PSF) e a deconvolução bidimensional, portanto, consiste em atenuar o efeito dessas PSFs. Neste trabalho revisamos os aspectos básicos desta modelagem e da estimação das PSFs, bem como do processo de imageamento, e mostramos a nossa contribuição para a deconvolução bidimensional por meio de um método de filtragem inversa. .

Palavras-chave: Processamento digital de sinais;

CONTENTS

To my parents,

Olga and Yuji

and my sister,

Eliza

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Prof. Dr. Renato da Rocha Lopes for his guidance and friendship, especially during difficult times throughout the period of my studies. His vast knowledge in many areas, patience and dedication were fundamental for the execution of the research and for the writing of this thesis. I feel very lucky that I had him as the advisor of my studies, as I had an excellent experience during my PhD course.

I thank my co-advisor during my exchange program (PDSE), Prof. Dr. Leiv Gelius, for kindly receiving me in the University of Oslo (UiO) in Norway. His guidance and immense scientific knowledge, especially in geophysics, were essential for the completion of this thesis.

I acknowledge the members of my thesis defense committee, Prof. Dr. Ricardo Higuti, Prof. Dr. Ricardo Suyama, Prof. Dr. Wu Shin-Ting and Prof. Dr. Romis Attux for their observations which greatly contributed for the improvement of the thesis.

I must also thank the friends of DSPCom and CEPETRO, as well as all friends I met at UNICAMP and UiO. I had a privilege to work with many brilliant minds, whose contributions were fundamental to the success of this thesis. Besides scientific collaboration, I must thank them for their friendship and support, as well as nice moments of leisure and debates. In particular, I acknowledge Prof. Dr. João Marcos Romano and Prof. Dr. Romis Attux for their kind advices and guidance during my stay at DSPCom. I also acknowledge Prof. Dr. Martin Tygel for his support throughout the thesis. I also thank Prof. Dr. Anne Solberg and Prof. Dr. Isabelle Lecomte for their support during my stay in Norway.

I thank all my family and friends, specially my parents, Olga and Yuji and my sister Eliza for their permanent support and encouragement.

I thank my sweet and beautiful girlfriend Adâmara for her support, partnership, patience and the nice moments we shared these past years.

I would like to thank CAPES and PETROBRAS for their financial support.

I would like to thank the UNICAMP, specially the staff of FEEC who allowed the conclusion of this work.

TABLE OF ACRONYMS

2DFT	2D Fourier Transform
ANP	*Agência Nacional do Petróleo, Gás Natural e Biocombustíveis* (Brazilian National Agency of Petroleum, Natural Gas and Biofuels)
BGR	Benveniste-Goursat-Ruget
B-ICA	Banded Independent Component Analysis
BSS	Blind Source Separation
CMP	Common-Midpoint
CO	Common Offset
CRS	Common Reflection Surface
CS	Common Shot
HOS	Higher-Order Statistics
i.i.d.	Independent and Identically Distributed
ICA	Independent Component Analysis
LPF	Low Pass Filter
LS	Least Squares
LSM	Least Squares Migration
LTI	Linear Time Invariant
MD	Migration Deconvolution
MED	Minimum Entropy Deconvolution
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NMO	Normal Moveout
PCA	Principal Component Analysis
pdf	Probability Density Function
PEF	Prediction-Error Filter
PSDM	Prestack Depth Migration
PSF	Point Spread Function
SCA	Sparse Component Analysis
SNR	Signal-to-Noise Ratio
SOS	Second-Order Statistics

SW	Shalvi-Weisnstein
WSS	Wide-Sense Stationary
ZO	Zero offset

"Certas palavras tem ardimentos; outras não.
A palavra jacaré fere a voz.
É como descer arranhado pelas escarpas de
um serrote.
É nome com verdasco de lodo no couro.
Além disso é agríope (que tem olho medonho).
Já a palavra garça tem para nós um sombreamento
de silêncios...
E o azul seleciona ela!"

(Manoel de Barros)

1 INTRODUCTION

An increase in the number of discoveries of hydrocarbon reserves has been observed in Brazil in the last years. The potential of hydrocarbon exploration in the country can be exemplified by the success of the 11th bidding round of oil and gas exploration areas made by the national regulator (ANP) in 2013. In this round, 12 domestic and 18 foreign companies invested 2.8 billion reais (US$1.4 billion) for exploration rights, setting a new national record [ANP, 2013; THE ECONOMIST, 2013]. Nevertheless, the exploration of these resources is far from trivial, as illustrated by the case of OGX, an oil company owned by one of the richest men in the country, which lost nearly 90% of its value during 2013. This fact is partly explained by the fact that this company failed to meet its production targets at its wells located in offshore Brazil [THE ECONOMIST, 2013b; FOLHA DE S.PAULO, 2013]. The difficulties in hydrocarbon exploration stem from the fact that, especially in offshore Brazil, the reservoirs are buried below many kilometers of rocks and sea water. This fact generates challenges in all steps of hydrocarbon exploration process, from the discovery of promising areas to the extraction of oil from deep waters, which require constant research and innovation in many fields.

In this context, this thesis is focused on digital signal processing theory, in particular to its use to enhance images obtained by the seismic method. The seismic method [YILMAZ, 2001] has applications especially on the discovery of new oil and gas fields and on monitoring reservoirs that are already producing. In this chapter, we provide brief rudiments of the seismic method in Section 1.1 in order to familiarize the reader with the terms that are used throughout the thesis. Then, in Section 1.2 we will present the contents and organization of this document and in the end, in Section 1.3, we will list the works published by the author as a PhD candidate.

1.1 ELEMENTS OF SEISMIC IMAGING

In this subsection we will describe some rudiments of seismic imaging, with the aim of familiarizing the reader of this work to the terms of the area. For a formal introduction to the field, we ask the reader to refer to books in geophysics such as [YILMAZ, 2001].

The seismic method is the main tool used by geologists and geophysicists to obtain information about the subsurface [YILMAZ, 2001]. In this technique, seismic waves are generated in a controlled fashion in the surface by seismic sources, such as explosive loads or vibroseis, in the case of terrestrial acquisitions, or airguns, in the case of marine acquisitions. The reflected waves are measured by a set of receivers called geophones or hydrophones, depending on the environment in which the seismic acquisition is made. An example in the marine case is shown in Figure 1. The obtained data is later processed, often with intense human labor and computational cost in order to produce a bi or tridimensional image, such as the one shown in Figure 2, which is then interpreted by a geologist. The main steps of a typical workflow in marine seismic data acquisition and processing will be described with more detail in the following. The marine case is focused here both because of its relevance on the Brazilian scenario, where many of the known reserves are located in the coastal regions of the country, and because the terrestrial case imposes the consideration of more issues that are not the scope of this thesis such as the effect of topography and the existence of more interfering waves, such as the *ground-roll*, a surface wave caused by boundary effects between the soil and the air [SHERIFF, 2002]. However, we emphasize that the techniques considered here can be used in both marine and terrestrial case, once these issues are overcome [YILMAZ, 2001].

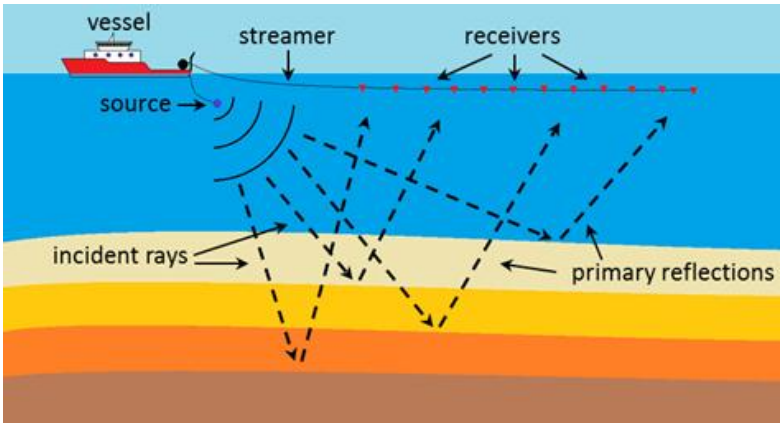


Figure 1: Diagram showing marine seismic data acquisition. A vessel tows the seismic source, the air gun, and a streamer, which is a cable that contains a set of receivers, the hydrophones. The source emits a seismic wave, which follow the paths indicated by the incident rays. These rays are reflected by the structures of the subsurface and the reflections return to the surface following the reflected rays. In this picture, we only show the primary reflections, which are the waves that suffered only one reflection in the path between the source and the sensors.

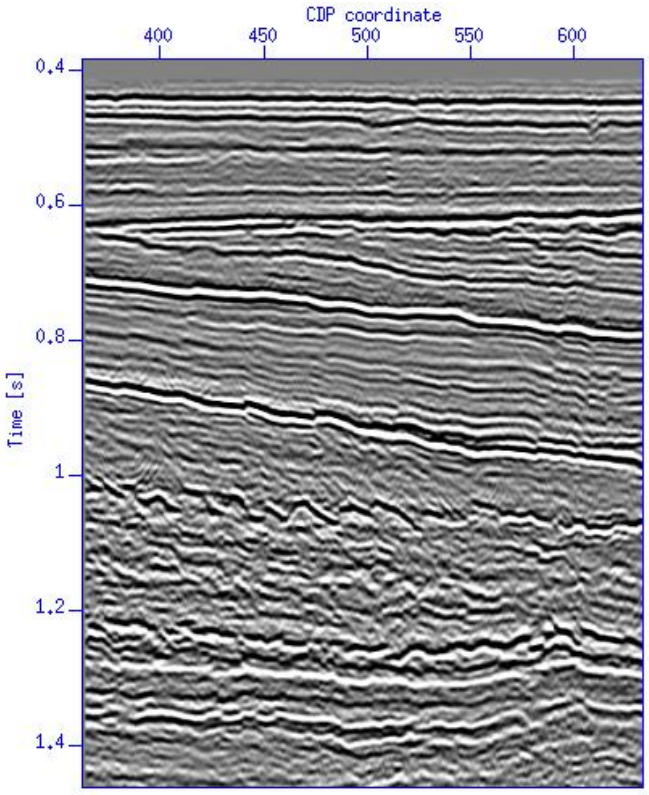


Figure 2: The seismic data, acquired as described in Figure 1, goes through intensive data processing to produce an image, such as the one shown here. This type of image is used by trained interpreters in order to create a geological model of the subsurface.

1.1.1 Data acquisition and illumination

Figure 1 shows a diagram on how seismic data is acquired in the marine environment. As described in the diagram, the vessel tows a seismic source, typically an air gun, a device that is able to produce a strong pressure wave in the water. Also, the vessel tows the streamer, a cable that contains a set of sensors called hydrophones. This streamer is usually 10 to 12 kilometers long in modern acquisitions. The vessel navigates over the area of interest shooting the source several times, usually in constant intervals of space so that regular spatial sampling is obtained. If the acquisition is made along a single straight line with only one streamer, the resulting data allows the production of a bi-dimensional image of the subsurface and thus this type of acquisition is called 2D acquisition. On the other hand, if the acquisition is made over a 2D area on the sur-

face, this allows one to produce a tri-dimensional image of the subsurface and this is thus called 3D acquisition. The acquisition geometry is important to determine how the features of the subsurface are imaged as it determines how the seismic waves interact with the subsurface, i.e., the *illumination*[1]. For example, Figure 3 shows the result of a simulated seismic acquisition with the use of ray tracing, which uses ray theory [BLEISTEIN, 1984; ČERVENý, 2001] in order to trace the raypaths given a velocity model. The red triangles show the sources positions and the black lines the reflected rays that reached the receivers. The left hand side of the salt dome, represented as the brown region, has not been illuminated by this survey. As a consequence, this region will not be imaged if this dataset is used.

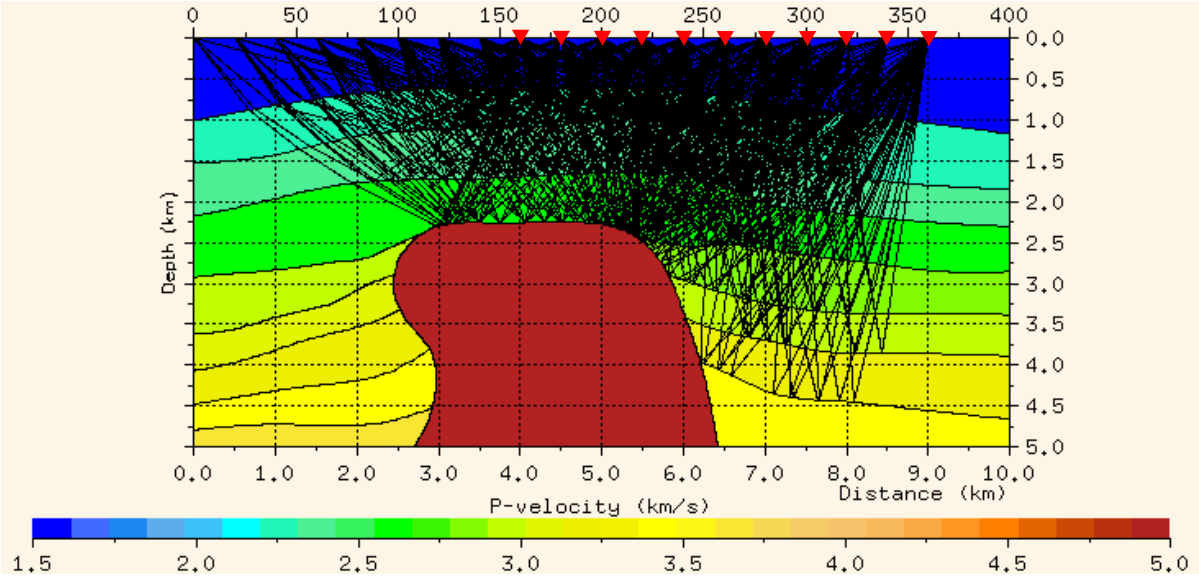


Figure 3: Result of ray tracing. The red triangles show the sources positions and the black lines the reflected rays that reached the receivers. The left hand side of the salt dome, in brown, has not been illuminated (c.f. footnote) by the survey.

1.1.2 Data preprocessing

The acquired data consist of a set of seismic traces or simply *traces*, which are the individual readings of a single receiver corresponding to one shot. In the initial steps, called *prepro-*

[1] In this case the term illumination is used as an analogy to the phenomenon of visible light striking a surface. In our context illumination indicates seismic waves reaching a reflector and thus being available for reflection. [SHERIFF, 2002].

cessing, these traces are conditioned so that they can be further processed. In this step, the information about the geometry of the acquisition is placed on the headers of the files which contain the seismic data and bad traces, due to e.g. malfunctioning receivers, are edited, usually manually as they are identifiable by visual inspection. It is important to notice that many signal processing challenges are posed during the preprocessing of seismic data. Firstly, the information recorded at the receivers is often a mixture of different waves that need to be identified and separated. In most cases, seismic imaging techniques require the enhancement of the primary reflections of the compression wave (P-wave), as the ones shown in Figure 1, and the suppression of other waves through dedicated seismic signal processing techniques such as the ones described in [YILMAZ, 2001]. Examples of these undesired factors include *swell noise*, which is caused by sea waves, *direct waves*, which are waves that travel directly from the source to the receivers, without reaching the subsurface, and multiple reflection waves or simply *multiples*, which are waves that suffer multiple reflections on the way between the source and the receiver, such as in the example shown in Figure 4. Secondly, distortions that happened during the propagation of the seismic wave must also be compensated. These procedures include the compensation of wave-energy decay due to absorption and geometrical divergence of the wavefront. *Ghosts*, which are caused by secondary reflections at the surface that happen shortly after the seismic pulse leaves the source or before reach the receiver, interfere with the primaries as shown in Figure 5, creating notches in the amplitude spectrum of the seismic data. A process called *deghosting* is used in order to eliminate this effect. Limitations in the frequency bandwidth of the seismic pulse and losses in frequency components and distortions in the phase also degrade the seismic image by causing the loss of resolution. *Deconvolution* techniques are used in order to try attenuate these effects and will be described with more detail in chapters 2 and 3, as this process plays a central role in this thesis. In fact, it will be seen that deconvolution techniques can also be used in later steps of seismic processing.

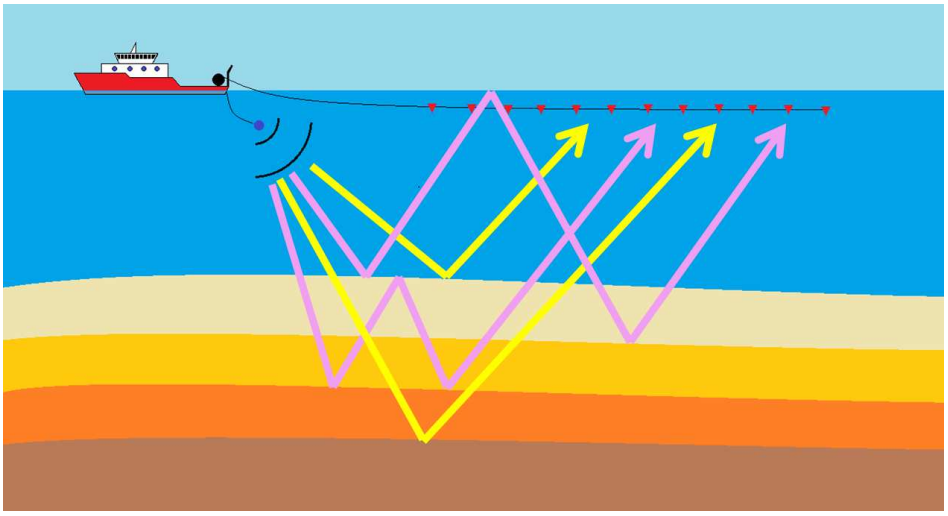


Figure 4: Yellow lines represent primary reflections, while the pink lines represent the multiple reflections. The primaries only suffer one reflection in the upward direction, while the multiples are reflected downwards at least once [VERSCHUUR, 2006].

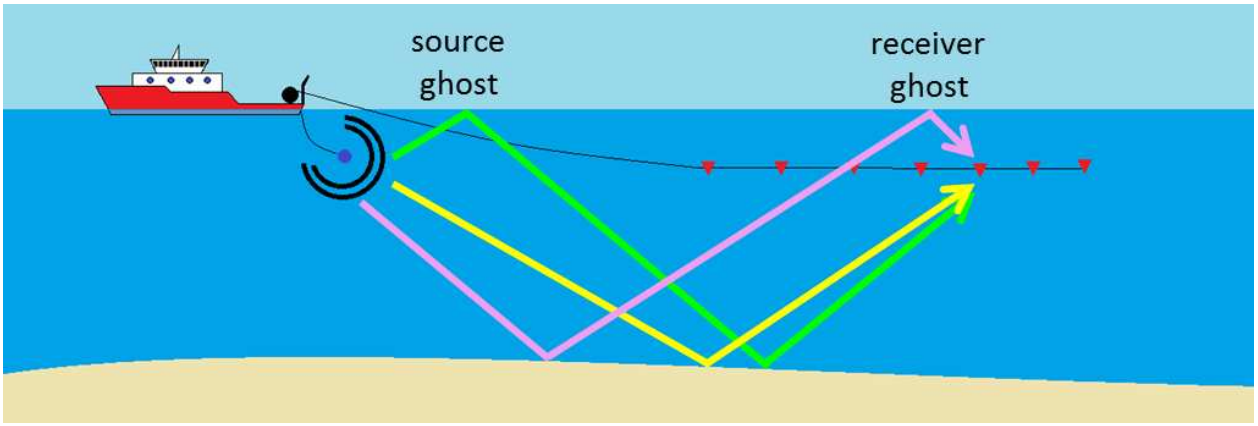


Figure 5: The green line represents a source ghost, while the pink line represents a receiver ghost. These rays interfere with the primary reflection, represented in yellow.

1.1.3 Trace sorting and the common-midpoint (CMP) method

One of the reasons that motivate the use of several receivers located at different positions relatively to the source is that this results in redundancy in the data so that the signal to noise ratio of the resulting image may be enhanced. One technique widely used in this sense is the CMP method [YILMAZ, 2001]. As shown in Figure 6, the actual seismic data acquisition is made in shot-receiver (s, g) coordinates. A set of traces corresponding to a single shot is called a *Common Shot* (CS) gather of traces. In seismic signal processing, it is common to sort and gather the traces accordingly to other types of geometries. Some of these geometries use the midpoint, y, and *half-offset* [2], ρ, between the source and receiver, which form a new coordinate system, (y, ρ), defined in terms of (s, g) as

$$y = \frac{s+g}{2}, \rho = \frac{g-s}{2}.$$

As shown in Figure 7 and Figure 8, groups of traces that have common value of midpoint coordinates or half-offset are called, respectively, *Common Midpoint* (CMP) gather and *Common Offset* (CO) gather. It can be seen in Figure 7 that if a homogeneous medium and a single horizontal reflector is considered, all the reflected rays in a CMP geometry emerge from the same point in depth. The corresponding CMP gather is shown in Figure 9. The position of the seismic pulse or *wavelet* in each trace is determined by the *traveltime*, needed for the seismic wave to travel from the respective source to the reflector and from the reflector back to the respective receiver. The traveltime in the CMP gather can be derived from the Pythagorean theorem and is given by the following expression for the receiver located at the position ρ_i:

$$t^2(\rho_i) = t_0^2 + \frac{4\rho_i^2}{C^2}, \quad (1)$$

where C is the propagation velocity of the seismic wave and t_0 is the two-way *Zero Offset* (ZO) traveltime, i.e., the traveltime corresponding to a fictitious case in which the positions of the source and receiver coincide. The increase of the traveltime caused by the increase of offset is called *Normal Moveout* (NMO) [SHERIFF, 2002]. It is important to notice from the relationship

[2] We use ρ instead of the usual h to denote half-offset in order to avoid confusion with $h(t)$, which indicates the seismic signature in this Thesis.

between t and ρ_i that (1) describes a hyperbolic shape, as shown in Figure 9 and is only exact for the case of a single layer in a homogeneous model, as depicted in Figure 7. If more layers with different *intervalar velocities* (i.e, wave propagation velocities) are added, such as in Figure 1, then the traveltimes for the resulting reflections will no longer be perfectly described by a hyperbola. Nevertheless, (1) can be considered a fair approximation for the real traveltimes [YILMAZ, 2001] for small values of ρ_i. The value of C associated with small offsets is called *NMO velocity*. In practice, the velocity is not known, and is estimated by a procedure called *velocity analysis*. Given a CMP gather, this procedure consists in sweeping values of C in order to find the respective curve that maximize the coherency of the data for each value of t_0.The rationale for this is that, if indeed there is a reflection at time t_0. with velocity C, then all the traces should contain the same information, pertaining to the same reflection. One common way to determine coherency of the data on a given curve is through the use of a second-order energy measure called *semblance* [YILMAZ, 2001]. The values of C obtained by this method are called *stacking velocities.*

After the stacking velocities are determined, a process called *NMO correction* is applied to the CMP gather in order to eliminate the normal moveout effect so that the resulting traces simulate a ZO trace. As an example, it is displayed in Figure 10 a noisy version of the CMP gather in Figure 9 after NMO correction. It is possible to observe that the hyperbolic event in Figure 9 has been horizontalized, i.e., all traces present the seismic event at $t = t_0$, as in the ZO trace. Then, after this step, the resulting traces are summed or *stacked*, so that the horizontalized event, which represents the signal of interest, is enhanced and the noise is cancelled out. Thus, the main objective of the CMP is to allow the production of a set of stacked traces that not only simulates ZO traces but also have an improved signal to noise ratio (SNR). This improvement actually happens if the traveltimes of the actual reflections are close enough to (1), i.e., the geology model is reasonably close to a horizontally layered model, without horizontal variation of velocity. It is worthwhile to mention that there are other stacking methods, such as the Common Reflection Surface (CRS) method, which enables one to use sets of CMP gathers with neighboring midpoints in order to combine even more traces with the objective of increasing the SNR of the stacked trace [HUBRAL *et al.*, 1998].

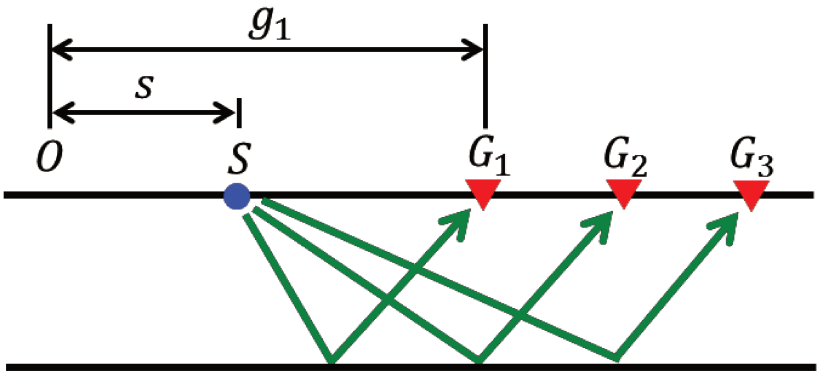


Figure 6: A common shot (CS) gather. This represents the geometry used when seismic data is acquired. A seismic source located in coordinate s produces seismic waves and the reflections are captured at the receivers located at coordinates g_i.

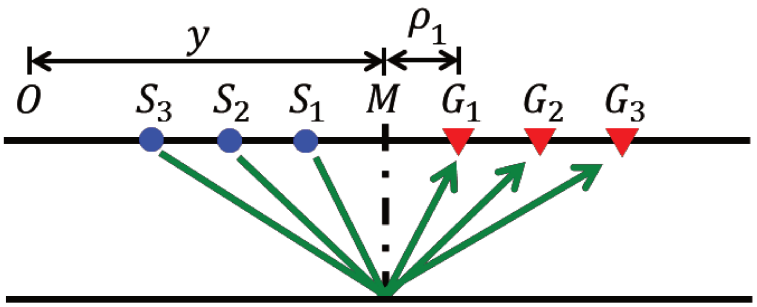


Figure 7: A common midpoint (CMP) gather. In this type of gather, the traces are sorted in sets where all traces share the same midpoint between the respective source and receiver (M). The coordinate of this common midpoint is represented by y and the half distance between a given source, S_i, and a receiver, G_i, (half offset) is represented by ρ_i.

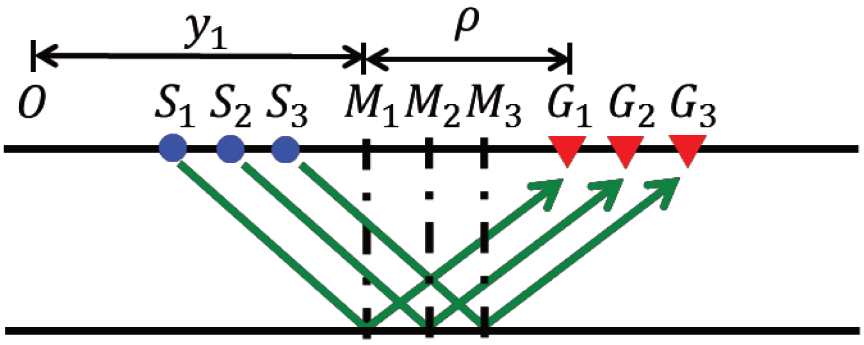


Figure 8: A common offset (CO) gather. In this type of gather, the traces are sorted in sets where all traces share the same offset (or half offset, represented by ρ) and each trace correspond to a different midpoint coordinate, represented by y_i.

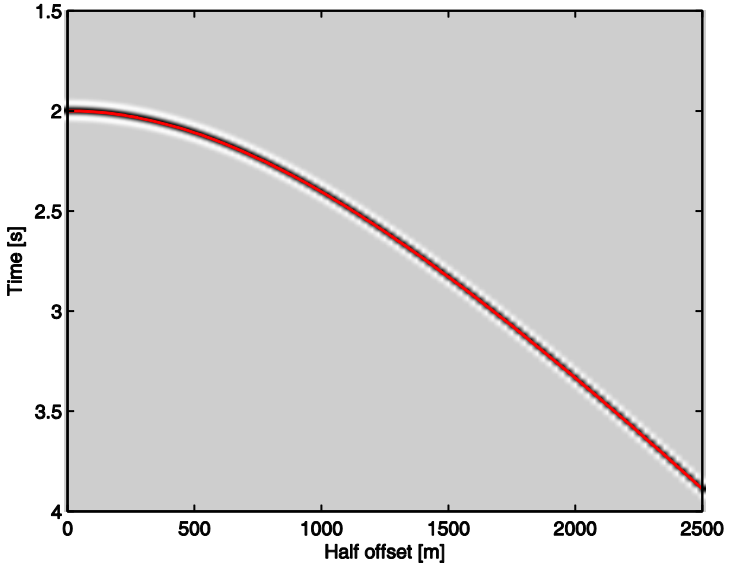


Figure 9: Each column of this grayscale Figure displays a trace of a CMP gather sorted accordingly to the offsets of its traces. These traces were obtained by a simulating data acquisition on a model with a single horizontal reflector located at the depth of $z = 1500\text{m}$ in a homogeneous medium, whose seismic wave propagation velocity is $C = 1500\text{m/s}$. The position of the seismic wavelet on the trace is determined by the traveltime needed for the seismic wave to travel from the respective source to the reflector and from the reflector back to the respective receiver. The Figure shows that the position of the wavelets matches the traveltime curve described by (1) and depicted in red in the Figure.

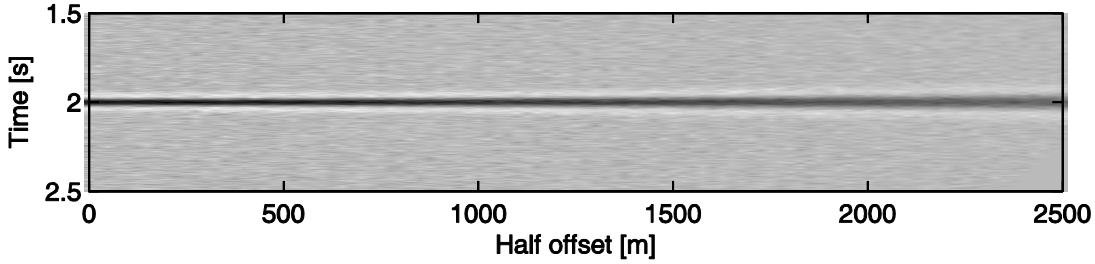


Figure 10: Noisy version of the CMP section of Figure 9 after NMO correction.

1.1.4 Seismic migration

As previously seen, the ZO traces contain seismic events of the hypothetical case in which the position of the sources and receivers coincide, and the stacking process produces images close to these traces under favorable conditions. However, seismic sections formed by putting these

traces side by side will often not represent accurately the position of the geological features. The reason for that is because these images assume that each trace contains only reflections that originate vertically below the source/receiver coordinate, which is seldom true. In Figure 11, where we assume a homogenous model, we observe that a time recording of a reflection may have origin in any point of the curve, which is a circumference with center at the source/receiver point, called *isochrone* [SHERIFF, 2002]. Once a good estimate of the velocity model of the subsurface is obtained, either through the velocity analysis prior to NMO correction or more advanced techniques, such as seismic tomography (e.g. [JONES, 2010]), this type of ambiguity is solved by a *migration* algorithm. Migration algorithms are a family of algorithms whose objective is to correctly position the geological features in the resulting seismic section (e.g. [YILMAZ, 2001; MOUSA, 2012]). As pointed out in [SHERIFF, 2002], migration can be accomplished by different approaches, such as integration over diffraction curves (Kirchhoff type of migration) or by using methods such as numerical phase shifting or finite differences in order to perform this inversion. Migration methods that are applied to stacked sections are called post-stack migration methods. On the other hand, complex geology may prevent one to use the CMP method successfully as this method assumes the existence of structures with horizontal layers and little horizontal velocity variation. If these conditions are not met, such as in areas where there are salt domes, then the use of the so called prestack migration methods may be required, as they operate directly on the preprocessed data and stacking is not used [YILMAZ, 2001]. A further discussion on migration is made in Chapter 3, where we will show how a 2D deconvolution may be used to improve a migrated image.

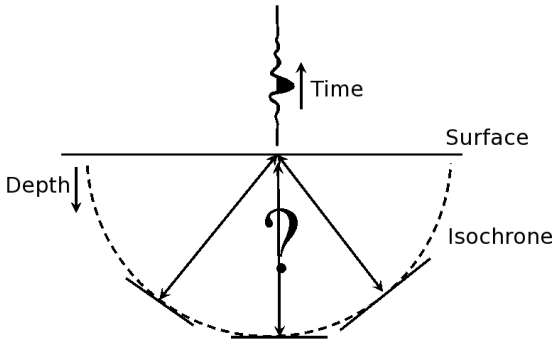


Figure 11: A seismic event recorded in time may be located in any point along the isochrone curve. The role of migration is to eliminate this ambiguity and properly locate the position of this event.

1.2 CONTENTS AND ORGANIZATION

In this work, we present analyses and contributions to the use of digital signal processing tools in unidimensional and bidimensional seismic deconvolution. As will be seen on the following chapters, these problems are motivated by the fact that the observations obtained in the surface, the seismic traces, always provide distorted or incomplete information about the subsurface.

If a single trace is analyzed, i.e., a unidimensional approach is adopted, the seismic experiment can be seen as an attempt to capture the impulsive response of the earth. In practice, the seismic trace is a distorted version of this ideal impulsive response, as the bandwidth limitations of the seismic source and distortions of frequency components caused by absorption and phase changes during the seismic wave propagation, in addition to noise, corrupt and blur this ideal impulsive response [NEELAMANI, 2008]. Thus, the deconvolution problem is used to estimate this true impulsive response from this distorted observation in order to enhance the final quality of the resulting seismic image. The choice of methods to solve this problem depends on the available data and on some *a priori* information about the problem. By using the terminology of digital signal processing literature (e.g. [ROMANO *et al.*, 2010]), if some information in addition to the seismic traces is available, such as some direct measurements of the wavelet, as proposed in [ZIOLKOWSKI, 1991], or the structure of the earth through a well log, as in [EDGAR and VAN DER BANN, 2011], then *supervised* type of signal processing techniques, such as the inversion methods described in [YILMAZ, 2001], may be employed to perform deconvolution. However, often, this extra information is not available or is inaccurate due to non-stationarities or corruption by noise, making the use of supervised methods not possible. This leads to the need for *unsupervised* methods. One of the first unsupervised methods was proposed in [ROBINSON, 1954], in which a white reflectivity series and a minimum phase wavelet were considered enabling the use of the Wiener theory on prediction error filtering [WIENER, 1949]. As shown in, e.g., [ROMANO *et al.*, 2010], this technique is constrained to minimum phase wavelets by the fact that it only uses second order statistics (SOS), which do not carry phase information. In order to overcome this, unsupervised methods use higher order statistics (HOS), such as the ones proposed in [WIGGINS, 1978; LAZEAR, 1993; MISRA and SACCHI, 2007] and many others.

In Chapter 2, we review the basics of supervised and unsupervised signal processing. In particular, we analyze the method proposed in [KAPLAN and ULRYCH, 2003; KAPLAN, 2003]

which uses a variation of independent component analysis (ICA), an important technique in unsupervised signal processing theory [HYVÄRINEN *et al.*, 2001; COMON and JUTTEN, 2010; ROMANO *et al.*, 2010], called Banded ICA (B-ICA). In this analysis, we show our contribution to the use of the method [TAKAHATA *et al.*, 2012], where we propose the use of B-ICA as a wavelet estimation procedure followed by the calculation of a least squares (LS) inverse filter. We extend the contents of this publication to a case study where we assess our improvement by comparing it to the original B-ICA algorithm and the use of prediction error filters (PEFs) for deconvolution. The choice of the latter was motivated by the fact that it was the first unsupervised methods for deconvolution and also by the fact that it is still one of the standard algorithms in the industry [YILMAZ, 2001]. In this case, we used both synthetic and processed (migrated) field data in order to perform the tests. A contribution of this work in this context was the use of well log data, which are detailed records of the geological formations obtained through direct measurements made in a borehole. This allowed us to compute the wavelet and its inverse in a deterministic manner, as well as to compute more realistic synthetic data. In other words, the existence of this well log allowed us to estimate the reflectivity more directly, thus providing a benchmark for the deconvolution algorithms.

The unidimensional approaches are constrained by the fact that only one measurement is considered at a time and the use of information of neighboring traces is inexistent or very limited. In the bidimensional approach, in addition to the distorting effects of the seismic wavelet, considered in the 1D case, the effects of the acquisition geometry and velocity variations caused by geologic complexities are taken into account. In [LECOMTE and GELIUS, 1998; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002] it is shown that limited frequency band of seismic data, data acquisition geometry constrained to a finite region in the surface and the pattern followed by the propagation of the seismic waves cause the seismic images, more precisely the ones obtained by prestack depth migration (PSDM), to be blurred versions of the actual structure of the subsurface. In these works, this distortion is quantified by the *resolution function*. Under proper assumptions, these resolution functions can be interpreted as point spread functions (PSFs), which have been used in image processing for quantifying distortions (e.g., [BANHAM and KATSAGGELOS, 1997]). The PSDM images are then modeled as the result of the bidimensional convolution between the PSFs and the original structure of the subsurface. Thus, the role of deconvolution is to eliminate, or at least to attenuate, the effect of these PSFs in order to improve

the PSDM section so that the process results in a section which is closer to the actual geological structure. A direct inversion approach was adopted in [GELIUS *et al.*, 2002; SJOEBERG *et al.*, 2003], but only small portions around the vertical central axis of the PSFs (pseudo 2D) were used in these cases, as the whole PSFs would lead to instabilities in the inversion.

Related works in enhancing PSDM sections include the methods derived from least-squares migration (LSM) [NEMETH *et al.*, 1999] and migration deconvolution (MD) [HU *et al.*, 2001; YU *et al.*, 2006]. These methods consider the whole modeling and migration operators and may impose a heavy computation burden or create the need for simplifying schemes, as opposed to the bidimensional deconvolution method, where the PSFs can be calculated with relative ease with the use of ray tracing methods [GELIUS *et al.*, 2002; LECOMTE, 2008]. These methods based on LSM and MD are out of the scope of this thesis and further discussion may be found in [YU and SCHUSTER, 2003; TANG, 2009] .

In Chapter 3, we review, following [LECOMTE and GELIUS, 1998; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002; LECOMTE, 2008], the bidimensional convolutional model and the fundamentals of PSDM and PSF estimation. Next, we show in more detail the solution for the 2D deconvoltution problem proposed in [TAKAHATA *et al.*, 2013]. In this work, we proposed the use of a LS inverse filtering (or Wiener filtering) method instead of direct inversion used previously in literature. Even though, as shown in, e.g., [BANHAM and KATSAGGELOS, 1997], this approach has been long a common practice in image processing, our contribution is to propose its use in this particular context, which also motivated new analysis that gave new insights into the problem. In particular, an analysis of the 2D Fourier transform of the PSF showed null elements that brought instability to the inverse filtering calculation, which are probably also one main source of instability in the inversion approaches in the literature. In this chapter, we also discuss a refinement addressing this problem in the inverse filtering approach, allowing the deconvolution with the whole PSF, as opposed to previous approaches and we show our results in controlled and field data.

In Chapter 4, we present our final conclusions and perspectives.

1.3 PUBLISHED WORK

Journal Publications

GELIUS, L.J.; TYGEL, M.; TAKAHATA, A.K.; ASGEDOM, E.G.; SERRANO, D.R. High-Resolution Imaging of Diffractions - A window-steered MUSIC approach. **Geophysics**, v. 78, n. 6, p. S255-S264, 2013.

TAKAHATA, A. K.; NADALIN, E. Z.; FERRARI, R.; DUARTE, L. T.; SUYAMA, R.; LOPES, R. R.; ROMANO, J. M. T.; TYGEL, M. Unsupervised Processing of Geophysical Signals: A Review of Some Key Aspects of Blind Deconvolution and Blind Source Separation**. IEEE Signal Processing Magazine**, v. 29, n. 4, p. 27-35, 2012.

NADALIN, E. Z. ; TAKAHATA, A. K. ; DUARTE, L. T.; SUYAMA, R.; ATTUX, R. Blind Extraction of the Sparsest Component. **Lecture Notes in Computer Science**, v. 6365, p. 394-401, 2010.

Conference Publications

COVRE, M.R.; BARROS, T.; TAKAHATA, A.K.; LOPES, R. R**. Imageamento de Difrações Sísmicas Baseado em Métodos de Alta-Resolução.** In: Anais do XXXI Simpósio Brasileiro de Telecomunicações, Fortaleza, CE, 2013. (In Portuguese).

TAKAHATA, A.K.; GELIUS, L.J.; LOPES, R.R.; TYGEL, M.; LECOMTE, I. **2D Spiking Deconvolution Approach to Resolution Enhancement of Prestack Depth Migrated Seismic Images.** In: 75th EAGE Conference and Exhibition incorporating SPE EUROPEC 2013, London, UK, 2013.

TAKAHATA, A. K. ; AGULHARI, C. M. ; DUARTE, L. T. ; LOPES, R. R. ; PERES, P. L. D. ; ROMANO, J. M. T . **Métodos de Compressed Sensing para Compressão de Sinais ECG.** In:

XXII Congresso Brasileiro de Engenharia Biomédica, p. 640-643, Tiradentes, MG, 2010. (In Portuguese).

2 UNIDIMENSIONAL SEISMIC DECONVOLUTION

In this chapter, we review and extend the topics approached in [TAKAHATA *et al.*, 2012], especially regarding unsupervised deconvolution. In Section 2.1, we review the basic connections between the structure of the earth and its impulsive response, the reflectivity, and also between the reflectivity and the seismic traces through the point of view of the convolutional model. In Section 2.2, we review the concepts underlying supervised signal processing for seismic deconvolution and wavelet estimation through an analogy to supervised signal processing in data communication channel estimation and equalization. After that, in Section 2.3, we analyze unsupervised signal processing techniques for seismic deconvolution and wavelet estimation. Initially, a study on deconvolution with the use of the prediction error filtering is made and its limitations due to the use of second order statistics (SOS) are discussed. Next, we discuss the use of higher order statistics (HOS) in order to overcome these limitations. After this, we review independent component analysis (ICA) and blind source separation (BSS), which are also major topics in unsupervised signal processing. In the sequence, we explore an approach, initially proposed in [KAPLAN and ULRYCH, 2003; KAPLAN, 2003], that combines these topics, ICA and seismic deconvolution. In this technique, a variant of ICA called banded ICA (B-ICA) is used for seismic deconvolution and wavelet estimation. We first describe the method and then we show our contribution in [TAKAHATA *et al.*, 2012], where we proposed the use of B-ICA for wavelet estimation in order to calculate a supervised deconvolution filter. Afterwards, we extend this contribution in the Section 2.4 in a form of a case study. In this section, we characterize a reflectivity series estimated from logs obtained in a borehole and we present the results of the use of PEF and B-ICA based methods in synthetic data. After that, we present the results of PEF, B-ICA and a supervised method to a set of migrated traces. At the end, we present our conclusions in Section 2.5.

2.1 THE CONVOLUTIONAL MODEL

As shown in Figure 12, reflection happens when a seismic wave reaches the boundary between layers with different acoustic impedances. The acoustic impedances are determined by

propagation velocity and density. However, other phenomena, such as transmission and conversion, happen simultaneously to reflection. Thus, the determination of reflection coefficients, i.e., the ratio of amplitudes between the reflected and incident waves, may be complicated in a general case, being described by the Zoepprits Equations [SHERIFF, 2002]. However, phenomena other than reflection can be ignored if it is assumed that [YILMAZ, 2001]:

- The subsurface consists in a set of horizontal layers with velocity variations depending only on depth.
- The seismic wave is a plane wave which propagates in the vertical direction and strikes the reflectors at normal incidence.

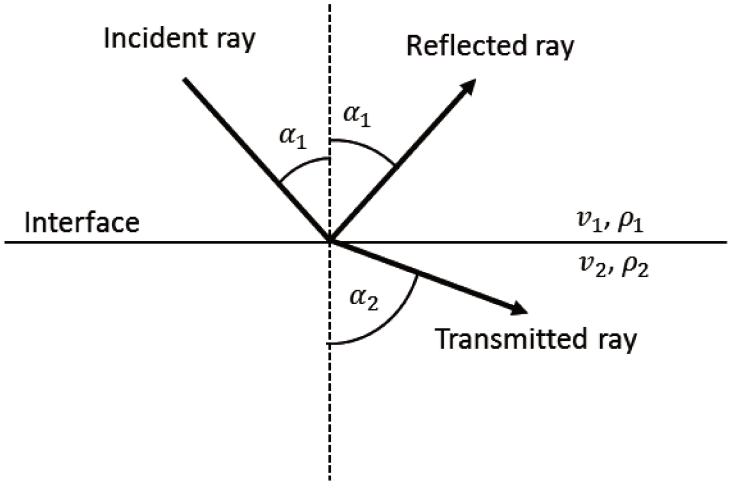


Figure 12: When a seismic wave reaches an interface, reflection and transmission are observed following the path of the rays depicted in this image.

The first assumption is met in cases where there is no strong lateral velocity or density variation, while the second assumption is approximated if we consider depths larger than the receiver array aperture, so that the angle of reflection is small and its effects are negligible. Then, the reflection coefficient *s* can be calculated as

$$s = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \tag{2}$$

where the acoustic impedance of the i-th layer is defined as $Z_i = v_i \rho_i$, where v_i and ρ_i are, respectively, the seismic wave propagation velocity and the density of this layer. This is analogous to the impedance in transmission lines, where the contrasts of impedance cause the reflection of

electromagnetic waves along the line. Figure 14 shows that the properties of each layer of the geological section determine their acoustic impedance. The associated reflectivity series $s(z)$, where z is the depth coordinate, can then be derived with the use of (2). These values may be mapped onto the time domain so that a reflectivity function $s(t)$ describes the reflection events in time, represented by t. If only the primary reflections are considered and we assume the existence of a continuous function of impedance, $z(t)$, then we have, by applying (2) to infinitesimal sized layers [SENGBUSH *et al.*, 1961] the reflectivity function in the time domain:

$$s(t) = \frac{d}{dt}\log(z(t)). \tag{3}$$

From the point of view of unidimensional deconvolution, the seismic experiment may be interpreted as a way to identify the interior of the earth through the estimation of its impulsive response, where the primary reflections, in an idealized situation, correspond to the reflectivity function, $s(t)$. Thus, one of the approaches is to use a seismic source, such as an air-gun or an explosive, to generate a seismic pulse that resembles an ideal impulse as the input of the system to be identified, as shown in Figure 13. Each type of seismic source generates its own type of signal called seismic signature or wavelet, which is denoted as $h(t)$.

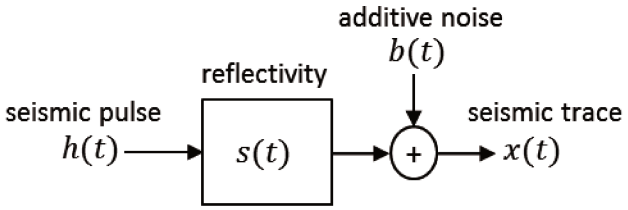


Figure 13: The seismic experiment may be interpreted as an attempt to obtain the impulsive response of the earth. The seismic pulse, $h(t)$, which is an approximation of an ideal impulse is the input of the system to be identified, represented by the reflectivity, $s(t)$. The seismic trace, $x(t)$, is the approximate impulsive response of the earth corrupted by additive noise.

Thus, given the seismic source or wavelet, $h(t)$, and the seismic reflectivity, $s(t)$, if we consider that the earth behaves, at least approximately as an linear and time invariant system we may model the seismic trace as

$$x(t) = s(t) * h(t) + b(t). \tag{4}$$

where $b(t)$ is the additive noise and the convolution is defined as:

$$s(t) * h(t) = \int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau. \quad (5)$$

In (5), we can interpret that the convolution operation actually models the seismic trace as superposition of delayed versions of the seismic signature, $h(t-\tau)$, weighted by the amplitude of the reflectivity $s(\tau)$. This is only valid if we assume that the wavelet does not vary along the trace. However, this assumption is often not true if the whole measurement length is considered, as the subsurface distorts the seismic wavelet, e.g., by absorbing the high frequency contents, as the seismic waves propagate. Nevertheless, this assumption can be approximated if short time windows are considered [ROBINSON and TREITEL, 1980]. Further, the convolution models a linear distortion imposed on the reflectivity by the wavelet, as a perfect recovery of the reflectivity would only be possible if $h(t) = \delta(t)$, where $\delta(t)$ is the Dirac delta distribution, and in the absence of noise, since, in this case

$$s(t) * \delta(t) = s(t). \quad (6)$$

This is not possible in real situations, as it would require an infinite frequency bandwidth associated with the seismic signature. Thus, $h(t)$ also models the restriction in the bandwidth of the data imposed by limitations in the seismic source and attenuations on the seismic waves along the path between the source and receiver. Therefore, the role of the deconvolution is to attenuate as much as possible the distorting effect of the wavelet on the seismic trace $x(t)$ in order to obtain the best approximation of the reflectivity function.

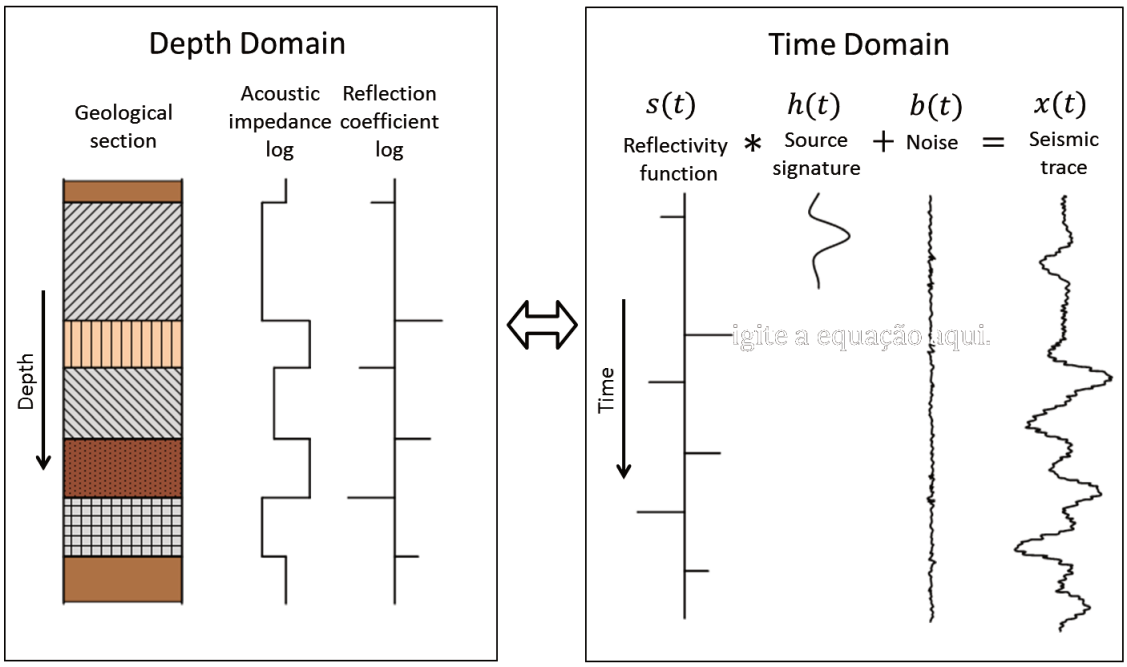


Figure 14: The reflection of seismic waves is caused by the difference of acoustic impedance between different geological layers [TAKAHATA *et al.*, 2013].

Before ending this subsection, it is important to note that, as shown, for example, in [ROBINSON and OSMAN, 1996; YILMAZ, 2001], seismic deconvolution and many other seismic signal processing procedures use digital signal processing techniques. This implies the use of samples of band limited signals, which is appropriate in the case of seismic signals, as the signals of interest are seldom above 100 Hz. If now we consider that the seismic signal has been sampled at an appropriate rate (typically one to four milliseconds of sampling interval), then we can express (4) as

$$x(n) = s(n) * h(n) + b(n)\,, \tag{7}$$

as depicted in Figure 15, where $x(n)$ represents the sample corresponding to $x(n \cdot \Delta t)$, Δt corresponds to the sampling period and the convolution in the discrete domain is defined as:

$$s(n) * h(n) = \sum_{k=0}^{N_h - 1} h(k)s(n-k), \tag{8}$$

admitting that the wavelet, $h(n)$, can be modeled as a linear and time invariant (LTI) system whose impulse response can be described with N_h coefficients.

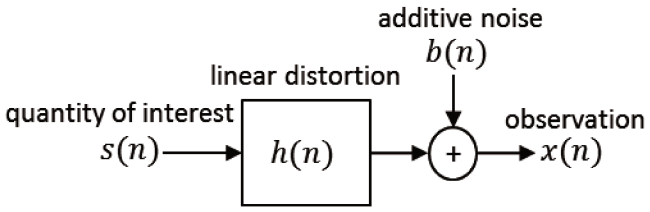


Figure 15: The quantity of interest can be estimated only through an observation which is corrupted by a linear distortion and additive noise.

2.2 SUPERVISED DECONVOLUTION AND WAVELET ESTIMATION

The convolution model described in (7) is illustrated in Figure 15. In some situations, it is possible to estimate this wavelet directly, such as in marine or vibroseis land acquisition cases. In this case, the deconvolution can be performed in a supervised fashion and is called *deterministic deconvolution* [NEELAMANI, 2008]. Also, well logs information can be combined with seismic data in order to estimate the seismic wavelets [EDGAR and VAN DER BANN, 2011]. Under these circumstances, the so called supervised linear filtering techniques described, for example, in [HAYKIN, 2001; YILMAZ, 2001; ROMANO *et al.*, 2010] can be used to perform these tasks.

Supervised techniques are based on calculating a linear filter, with impulse response $w(n)$, given the information of the input signal, $x(n)$, which corresponds to the seismic trace, and a reference signal, $d(n)$, as depicted in Figure 16. The reference signal is used to guide the estimation of the filter weights $w(n)$, and thus the name "supervised". Then, the problem of supervised linear filtering is to perform this estimation so as to minimize some criterion based on the difference between this reference and the filter output, $y(n)$, i.e., the error signal

$$e(n) = d(n) - y(n). \tag{9}$$

Among the many possible criteria, the Wiener type of criterion aims to minimize the mean squared value of the error signal (MSE), under the hypothesis that $d(n)$ and $x(n)$ are jointly wide sense stationary random signals. Another criterion is the least squares (LS) type of criterion, which aims to minimize the MSE if a deterministic signal or a single realization of a random process is considered (e.g., [ROMANO *et al.*, 2010; YILMAZ, 2001]).

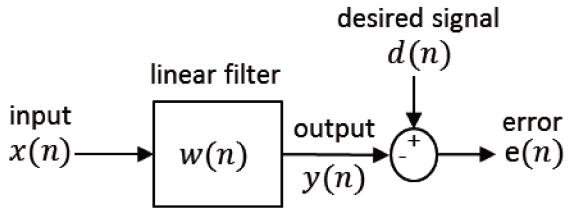


Figure 16: Supervised filtering block diagram. The linear filter, $w(n)$, is calculated so that, given its input, $x(n)$, the output, $y(n)$, minimizes the error signal, $e(n)$, relative to a desired or guiding signal, $d(n)$, accordingly to a given criterion such as LS or Wiener criterion.

As shown in [ROMANO *et al.*, 2010], the LS type of criterion for parameter estimation minimizes the sum of the squared error between the modeled data and the observed data. To compute the optimal filter in this context, we first describe the filter output, which is given by the convolution

$$y(n) = \sum_{m=0}^{N_w-1} w(m)x(n-m), \tag{10}$$

as a system of linear equations. Consider that the number of available samples is finite so that

$$x(n) = \begin{cases} x(n) \in \mathbb{R} \ \ if \ n \in \{0, 1, \dots, N_x - 1\} \\ 0 \text{ otherwise} \end{cases},$$

$$w(n) = \begin{cases} w(n) \in \mathbb{R} \ \ if \ m \in \{0, 1, \dots, N_w - 1\} \\ 0 \text{ otherwise} \end{cases},$$

$$y(n), d(n), e(n) = \begin{cases} y(n), d(n), e(n) \in \mathbb{R} \ \ if \ n \in \{0, 1, .., N_x + N_w - 1\} \\ 0 \text{ otherwise} \end{cases}.$$

Now, we define the vector

$$\mathbf{w} = [w(0) \quad w(1) \quad \dots \quad w(N_w - 1)]^T$$

and

$$\mathbf{x}(n) = [x(n) \quad x(n-1) \quad \dots \quad x(n - N_w + 1)]^T, n = 0, \dots, N_x + N_w - 1. \tag{11}$$

From (10), then

$$y(n) = (\mathbf{x}(n))^T \mathbf{w}. \tag{12}$$

If we define

$$\mathbf{y} = [y(0) \quad y(1) \quad \dots \quad y(N_x + N_w - 1)]^T,$$

then, we can write

$$\mathbf{y} = \mathbf{X}_{LS}\mathbf{w}. \tag{13}$$

where $\mathbf{X}_{LS}$ is a $(N_x + N_w - 1) \times N_w$ Toeplitz matrix so that

$$\mathbf{X}_{LS} = \begin{bmatrix} \mathbf{x}^T(0) \\ \mathbf{x}^T(1) \\ \vdots \\ \mathbf{x}^T(N_x + N_w - 1) \end{bmatrix} = \begin{bmatrix} x(0) & 0 & \dots & 0 \\ x(1) & x(0) & \ddots & 0 \\ \vdots & x(1) & \ddots & \vdots \\ x(N_x - 1) & \vdots & \ddots & 0 \\ 0 & x(N_x - 1) & \ddots & x(0) \\ 0 & 0 & \ddots & x(1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x(N_x - 1) \end{bmatrix}. \quad (14)$$

If we define

$$\mathbf{d} = [d(0) \quad d(1) \quad \dots \quad d(N_x + N_w - 1)]^T, \quad (15)$$
$$\mathbf{e} = [e(0) \quad e(1) \quad \dots \quad e(N_x + N_w - 1)]^T,$$

then, the optimal filter $\mathbf{w}$ accordingly to the LS criterion is the one that minimizes the square norm of the error function, i.e:

$$J_{LS}(\mathbf{w}) = \mathbf{e}^T\mathbf{e} = (\mathbf{d} - \mathbf{y})^T(\mathbf{d} - \mathbf{y}) = (\mathbf{d} - \mathbf{X}_{LS}\mathbf{w})^T(\mathbf{d} - \mathbf{X}_{LS}\mathbf{w}).$$

The well-known solution (e.g., [ROMANO *et al.*, 2010])[3] for finding the LS solution of the filter is then:

$$\mathbf{w}_{LS} = ((\mathbf{X}_{LS})^T\mathbf{X}_{LS})^{-1}(\mathbf{X}_{LS})^T\mathbf{d}, \quad (16)$$

Later, the works of [WIENER and HOPF, 1931; KOLMOGOROV, 1939; LEVINSON, 1947; WIENER, 1949; DURBIN, 1960] established the Wiener filter theory, where the minimum mean square error (MMSE) criterion was adopted in order to estimate the optimum filter. This criterion assumes that $x(n)$ and $d(n)$ are jointly wide sense stationary random variables. The MSE is then defined using (9) and (12) as

$$J_{MSE}(\mathbf{w}) = E[e^2(n)] = E[(d(n) - y(n))^2] = E[(d(n) - \mathbf{w}^T\mathbf{x}(n))^2],$$

where $E[\cdot]$ is the expectation operator. By using the definition of autocorrelation function as

$$r(i - k) = E[x(n - i)x(n - k)],$$

and of cross-correlation as

$$p(-k) = E[x(n - k)d(n)],$$

after some algebraic manipulations, we have

$$J_{MSE}(\mathbf{w}) = \sigma_d^2 - 2\mathbf{w}^T\mathbf{p} + \mathbf{w}^T\mathbf{R}\mathbf{w}, \quad (17)$$

where σ_d^2 is the variance of $d(n)$, $\mathbf{R}$ is the $N_w \times N_w$ autocorrelation matrix of $x(n)$ given by

$$\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}(n)^T], \quad (18)$$

[3] The LS method was developed originally in [GAUSS, 1809]. [ROMANO *et al.*, 2010] is indicated as a modern approach to this method.

and $\mathbf{p}$ is the $N_w \times 1$ cross-correlation vector between $x(n)$ and $d(n)$ given by

$$\mathbf{p} = E[\mathbf{x}(n)d(n)], \tag{19}$$

or in expanded form:

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & r(2) & \dots & r(M-1) \\ r(1) & r(0) & r(1) & \dots & r(M-2) \\ r(2) & r(1) & r(0) & \dots & r(M-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r(M-1) & r(M-2) & r(M-3) & \dots & r(0) \end{bmatrix}, \tag{20}$$

$$\mathbf{p} = [p(0) \quad p(-1) \quad p(-2) \quad \dots \quad p(1-M)]^T, \tag{21}$$

The minimum value of $J_{MSE}(\mathbf{w})$ can be obtained as

$$\mathbf{R}\mathbf{w} = \mathbf{p}, \tag{22}$$

and thus the Wiener solution is

$$\mathbf{w}_W = \mathbf{R}^{-1}\mathbf{p}. \tag{23}$$

By substituting (23) into (17) we obtain the optimal value of J_{MSE} accordingly to the MMSE criterion:

$$J_{MSE}(\mathbf{w}_W) = \sigma_d^2 - \mathbf{p}^T\mathbf{w}_W = \sigma_d^2 - \mathbf{p}^T\mathbf{R}^{-1}\mathbf{p}, \tag{24}$$

The choice of the input and of the desired signals depends on the objective of the estimation of the linear filter. One of these applications consists in modeling the linear distortion, i.e. the seismic signature, as shown in Figure 17. In this case, the input of the linear filter may be the reflectivity function estimated from a well log, $\tilde{s}(n)$, and supervised filtering techniques are used to estimate a model of the wavelet, $w(n)$,which will produce a filter output, i.e., a synthetic trace, $y(n)$, that will be as close as possible to the measured trace, $d(n)$. In order to obtain an accurate wavelet, the reflectivity must comply with the persistent excitation conditions, i.e., it must excite all modes of the linear distortion [ROMANO *et al.*, 2010]. Usually, this requirement is fulfilled as the reflectivity may be approximated to a white signal [YILMAZ, 2001]. A white signal is defined as a stochastic process whose autocorrelation is

$$r(k) = \begin{cases} \sigma_s^2 \text{ if k} = 0 \\ 0 \text{ otherwise} \end{cases}, \tag{25}$$

or $r(k) = \sigma_s^2\delta(k)$, where σ_s^2 is the variance of the signal. Also, it is important to note that as in any type of estimation, there may be mismatches between the estimated reflectivity from the well data, $\tilde{s}(n)$, and the actual reflectivity, $s(n)$, and thus they are shown as different signals in Figure 17. Nonetheless, this issue is outside the scope of this work as we assume that the estimated re-

flectivity is close enough to the actual one. On the other hand, we highlight that it is always necessary to inspect the quality of the results in order to verify if this assumption is reasonable.

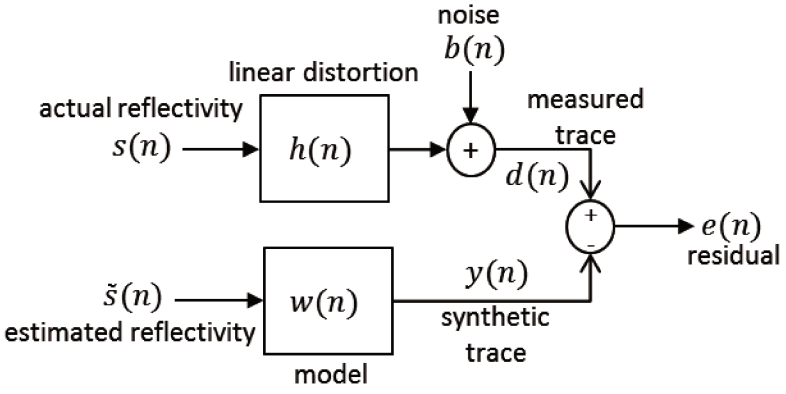


Figure 17: Block diagram for linear distortion identification.

Another application of supervised linear filtering is on supervised deconvolution, as illustrated in Figure 18. This can be seen as the dual of the wavelet identification case, as an approximation of the inverse filter of the wavelet is calculated. In this case, the deconvolution filter coefficients, $w(n)$, are estimated in order to eliminate the effect of the linear distortion, $h(n)$, from the observation, the seismic trace, $x(n)$, so that the output, $y(n)$, is as close as possible to the reflectivity, $s(n)$. Thus, in this case, it is assumed that there is a well log so that an estimate of the reflectivity, $\tilde{s}(n)$, can be calculated and used as the desired signal, i.e., $d(n) = \tilde{s}(n)$.

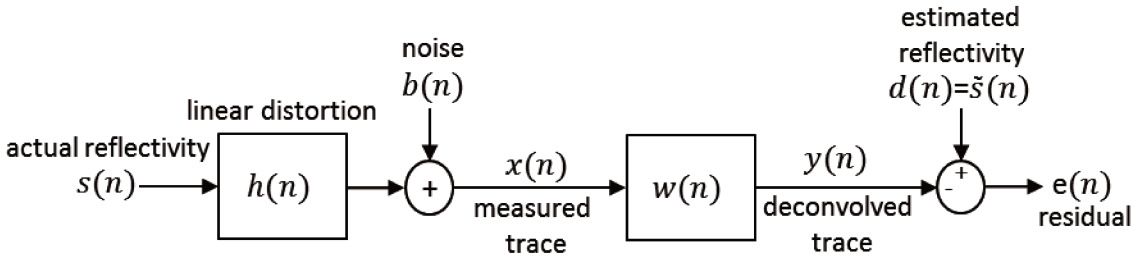


Figure 18: Block diagram showing the estimation setup for the deconvolution filter.

A complementary approach to deconvolution arises when the wavelet is somehow measured or estimated, e.g., with the use of the well, as mentioned previously. Assume that the wavelet is given by a time limited function so that

$$h(n) = \begin{cases} h(n) \in \mathbb{R} \;\; if \; k \in \{0, 1, .., N_h - 1\} \\ 0 \text{ otherwise} \end{cases}.$$

If we do not consider the noise element in Figure 18, we have

$$y(n) = x(n) * w(n) = \big(s(n) * h(n)\big) * w(n) = s(n) * (h(n) * w(n)). \quad (26)$$

This deconvolution approach, also called *spiking deconvolution*, consists in designing a linear filter $w_{LS}(n)$ of $(N_w - 1)$-th order with the LS criterion, so that

$$h(n) * w_{LS}(n) \approx \delta(n), \quad (27)$$

where $\delta(n)$ is the Kronecker delta or zero-lag spike, defined as

$$\delta(n) = \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ otherwise} \end{cases}.$$

Thus, the spiking deconvolution filter, $w_{LS}(n)$, which satisfies the condition in (27), may be used to attenuate the effect of the wavelet in the seismic trace, as if we substitute $w_{LS}(n)$ in (26) we get:

$$y(n) = x(n) * w_{LS}(n) = s(n) * (h(n) * w_{LS}(n)\ ) \approx s(n) * \delta(n) = s(n).$$

Qualitatively, we try to reshape the wavelet of the measured seismic trace with the use of $w_{LS}(n)$ so that the original wide wavelet is transformed to a narrower one, ideally a spike, so that the resolution of the final trace is enhanced, and in the ideal case the resulting trace becomes the reflectivity.

Thus, as shown in Figure 19, in this case, the goal is to calculate a linear filter so that the output, the processed wavelet, is as close as possible to the desired wavelet. Therefore, in order to obtain $w_{LS}(n)$, as described in (27), $\delta(n)$ is set as the desired wavelet. Thus we replace the variables depicted in Figure 19 in (16) and then we obtain[4]

$$\mathbf{w}_{LS} = ((\mathbf{H}_{LS})^T \mathbf{H}_{LS})^{-1} (\mathbf{H}_{LS})^T \boldsymbol{\delta}, \quad (28)$$

where $\mathbf{H}_{LS}$ is a $(N_h + N_w - 1) \times N_w$ Toeplitz matrix so that

$$[\,\mathbf{H}_{LS}]_{i,j} = h(i-j), i = 0, \dots, N_h + N_w - 1, j = 0, \dots, N_w - 1, \quad (29)$$

and

[4] In our case, we use LS in the sense of finding a vector $\mathbf{w}_{LS}$, which minimizes $\|\boldsymbol{\delta} - \mathbf{H}_{LS}\mathbf{w}_{LS}\|_2$ as in [YILMAZ, 2001], although the classical definition [GAUSS, 1809] of LS is the fitting of an analytical function to a set of data so that the sum of the squares of the deviations of the data points to the curve described by the function is minimized [SHERIFF, 2002]. In fact, as shown in Appendix A, if the reflectivity, $s(n)$, is a white signal and the noise is negligible, then $\mathbf{w}_{LS}$ is equivalent to the Wiener filter $\mathbf{w}_W$, which will minimize the mean square error of the reflectivity estimation in the filter output.

$$\mathbf{H}_{LS} = \begin{bmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \ddots & 0 \\ \vdots & h(1) & \ddots & \vdots \\ h(N_h-1) & \vdots & \ddots & 0 \\ 0 & h(N_h-1) & \ddots & h(0) \\ 0 & 0 & \ddots & h(1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h(N_h-1) \end{bmatrix}. \tag{30}$$

Also, $\boldsymbol{\delta}$ is a $(N_h + N_w - 1) \times 1$ vector so that

$$\boldsymbol{\delta} = [1 \quad 0 \quad 0 \quad \dots \quad 0]^T.$$

In fact, if we consider the reflectivity, $s(n)$, to be a white signal and we neglect the noise, then, as shown in Appendix A, the Wiener filter, $w_W(n)$, is actually the same as the LS filter, $w_{LS}(n)$. This means that, under these conditions, the output of the LS filter will be, actually, the as similar as possible, in the MMSE sense, to the reflectivity.

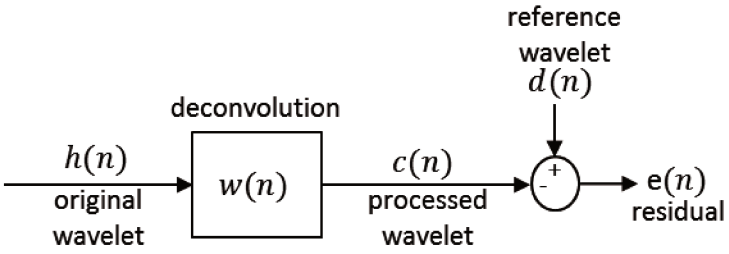


Figure 19: Block diagram for wavelet processing.

Also, if we consider a unit variance for $s(n)$, then we have, from the results shown in Appendix A, that the expressions of autocorrelation and cross-correlation in terms of $\mathbf{H}_{LS}$ are given by:

$$\mathbf{R} = (\mathbf{H}_{LS})^T \mathbf{H}_{LS},$$
$$\mathbf{p} = (\mathbf{H}_{LS})^T \boldsymbol{\delta}.$$

From (30), we have then

$$\mathbf{p} = [h(0) \quad 0 \quad \dots \quad 0]^{\boldsymbol{T}}, \tag{31}$$

and from (20) and (22), the expression for the LS filter in expanded form [YILMAZ, 2001]:

$$\begin{bmatrix} r(0) & r(1) & r(2) & \dots & r(N_w-1) \\ r(1) & r(0) & r(1) & \dots & r(N_w-2) \\ r(2) & r(1) & r(0) & \dots & r(N_w-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r(N_w-1) & r(N_w-2) & r(N_w-3) & \dots & r(0) \end{bmatrix} \begin{bmatrix} w_{LS}(0) \\ w_{LS}(1) \\ w_{LS}(2) \\ \vdots \\ w_{LS}(N_w-1) \end{bmatrix} = \begin{bmatrix} h(0) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (32)$$

Finally, as described in [YILMAZ, 2001], other types of signals are used as desired output wavelet instead of $\boldsymbol{\delta}$ in (28). One of them is the spike at an arbitrary lag, as this may lead to a smaller error if compared with the use of a zero lag spike. This is called *delayed-spike deconvolution*. Another type of desired wavelet may be the zero phase wavelet, i.e., a wavelet whose phase spectrum is constant at value equal to zero. This wavelet has even symmetry around the origin in time and thus is non-causal. Even though this type of wavelet is not physically realizable, it is often used in seismic interpretation, as the peaks and troughs of a seismogram containing this type of wavelet correspond to reflectors with positive and negative reflectivity, depending on the adopted convention [SIMM and WHITE, 2002]. If causality is to be kept, then a delayed version of the zero phase wavelet may be considered. An arbitrary shape wavelet may also be another option. This actually defines *pulse shaping*, in which the original wavelet is shaped into another desired shape.

2.2.1 Supervised deconvolution in data communication channel equalization

The convolutional model described in Figure 15 is also used to describe problems in other areas where digital signal processing is used, such as audio signal processing [ZÖLZER, U., 2008], dynamic systems analysis [GEROMEL and PALHARES, 2004] and telecommunications [ROMANO *et al.*, 2010]. Thus, many problems in these fields based on the convolutional model are analogous to the seismic case. In particular, we highlight the application in telecommunications, as some fundamental results are applicable in both areas. In telecommunications, the convolutional model is used to model the transmission of data through a linear distorting channel. As shown in Figure 20, in this case, the quantity of interest, that is analogous to the seismic reflectivity, is the message sent from the transmitter through a wireless channel, which is analogous to the wavelet. The receiver gets a superposition of delayed signals weighted by different attenuations caused by multipath propagation. This causes the received signal to be a linearly distorted version

of the transmitted one, as in Figure 15, and thus it can be considered analogous to the seismic trace. This phenomenon is called intersymbol interference (IIS) [ROMANO *et al.*, 2010]. In order to eliminate or attenuate the distorting effects which cause IIS, a procedure similar to seismic deconvolution, called channel equalization, is used.

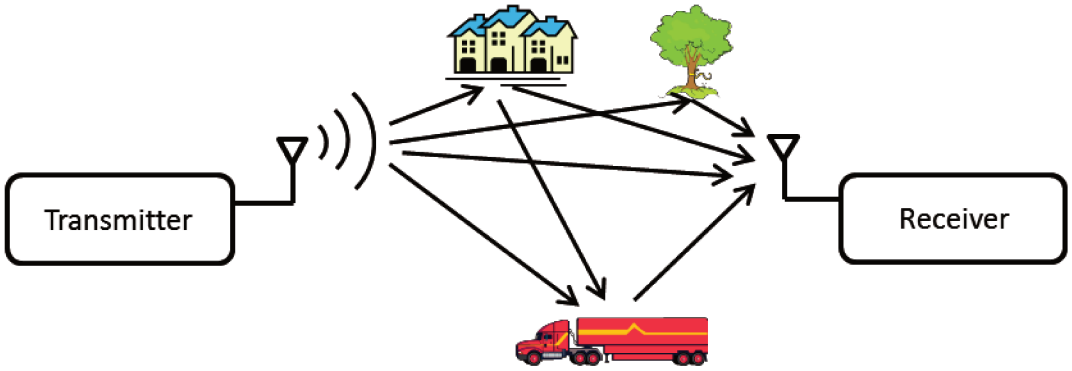


Figure 20: The transmitted data may travel through many paths to the receiver.

In some cases, it is possible to estimate the communications channel or its inverse by sending blocks of bits (assuming digital communications) that are known *a priori* to both the transmitter and receiver sides, called training bits. These training bits can be used to estimate the linear distortion caused by the transmission channel, analogously to the case described in Figure 17. The difference is that, in this case, the match is perfect between the training bits sent from the transmitter side and the training bits used for channel estimation, while, in the seismic, case this relationship may not be completely true, as previously discussed.

Another application of supervised signal processing is in channel equalization or channel deconvolution. This is analogous to the seismic case shown in Figure 18. As in the case of channel estimation, the training bits are transmitted through the distorting channel. At the receiver side, the autocorrelation of the received signal and the cross-correlation between the received signal and the trainings bits are estimated. This information is used to compute the optimum Wiener filter as described in (23), so that the filter output is close to the training bits, which guide the filter design. This is analogous to the use of an estimated reflectivity log order to estimate a deconvolution filter in the seismic case.

It is interesting to notice that, especially in mobile devices, such as in the scenario described in Figure 20, the channel is not time invariant. In this case, the position of the receivers and the scatterers, such as the truck in Figure 20, change, modifying the reflection pattern of the

transmitted electromagnetic waves and changing the channel coefficients. Thus, it is common to interleave training data with user data, which is, of course, not known by the receiver at transmission time, in order to assure that the equalizer is adjusted from time to time. In fact, adaptive algorithms, such as the least mean square (LMS) algorithm [HAYKIN, 2001; ROMANO *et al.*, 2010], allow the dynamic estimation of the equalizers by using the error information to adapt the filter to the changes of the channel in time.

2.3 UNSUPERVISED SIGNAL PROCESSING

In the previous section, we estimated the linear distortion or its inverse with the aid of a guiding signal, such as the estimate of the reflectivity function obtained from a well log in geophysical signal processing or the training bits in telecommunications, in order to eliminate through deconvolution the distortions effects in the measured signals. Instead of directly computing the inverse, another possibility was the use of estimates of the wavelet in the seismic case to calculate deconvolution filters using the LS criterion. These methods are called supervised signal processing, as extra information, besides the measured data, is used to guide the estimation of the quantities of interest.

However, in some cases, the use of guiding signals is not desirable, nor feasible. In the telecommunications case sending training data means that some part of the transmission time is not occupied with user data, which lowers the effective user data transmission rate. In the geophysical processing case, direct wavelet measurements or well data are not always available. In the absence of these guiding signals, unsupervised or blind techniques, which rely exclusively on the measured data and on a minimal amount of hypotheses concerning the signal of interest and on the linear distortion, become an option [ROMANO *et al.*, 2010].

Nevertheless, blindly estimating the system input and its distortion is an ill-posed problem, as, for instance, in the seismic case, infinite combinations of wavelets and reflectivity may result in the same trace. Thus, one of the main challenges in this case is to establish *a priori* hypotheses about the structures of the wavelet and the reflectivity in order to make the problem tractable. These hypothesis often concern statistical properties of the reflectivity, and thus unsupervised techniques in seismic deconvolution receive the name *statistical deconvolution*, in oppo-

sition to the supervised approach, which, as already mentioned, receives the name *of deterministic deconvolution* [NEELAMANI, 2008].

2.3.1 Predictive deconvolution

The use of predictive deconvoluiton for seismic signals was initially proposed by E.A. Robinson in his PhD thesis [ROBINSON, 1954], based on the work of Wiener [WIENER, 1949]. Curiously, as later stated in [ROBINSON and OSMAN, 1996], "deconvolution was the first truly digital signal processing method". As stated in, e.g., [ROMANO *et al.*, 2010], the problem of prediction consists in estimating future values of a time series from past and present information. If the signal to be predicted is a wide sense stationary random process, then a linear filter, such as the one illustrated in Figure 21, can be used to minimize the prediction error in the MMSE sense. In this case, the input itself, $x(n)$, plays the role of the desired signal and its delayed version, whose lag is l, given by $x(n-l)$, is used as the input of the linear filter. The prediction error is given by

$$e(n) = x(n) - y(n), \tag{33}$$

where

$$y(n) = x(n-l) * w(n). \tag{34}$$

is the predicted time series. The coefficients of the $(M-1)$-th order filter $w(n)$ which minimize $E[e^2(n)]$ are obtained by solving the Wiener-Hopf equations described in (22). However, in this case, the cross-correlation vector is replaced by the autocorrelation coefficients of $x(n)$, given by

$$\mathbf{r}(l) = E[\mathbf{x}(n-l)x(n)] = [r(l) \quad r(l+1) \quad \cdots \quad r(l+M-1)]^T, \tag{35}$$

so that the Wiener-Hopf equations become

$$\mathbf{R}\mathbf{w}_P = \mathbf{r}(l). \tag{36}$$

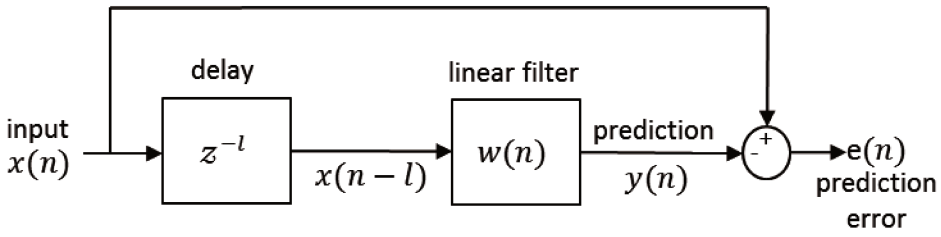


Figure 21: Block diagram of the prediction error filter with lag l.

In particular, the unit-lag predictor with $l = 1$ is of interest as it can be used in the context of unsupervised deconvolution. In order to explain this, we must first define a minimum phase system. If we define the Z-transform of a linear system with impulse response $h(n)$ as

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n}, \tag{37}$$

then it will be associated with a minimum phase system if all of its zeros and poles are contained in the unit circle[5]. On the other hand, non-minimum phase systems are the ones whose zeros are located exclusively outside the unit circle (maximum phase systems) or the ones which have zeros both inside and outside the unit circle (mixed phase systems). In fact, if we consider the definition of the Fourier transform of a discrete signal, $h(n)$, as

$$H(f) = \sum_{n=0}^{\infty} h(n) e^{-j2\pi f n}, \tag{38}$$

and its magnitude spectrum as $|H(f)|$, an arbitrary stable and causal system, represented by $H(z)$, may be factored, as in [OPPENHEIM and SCHAFER, 1989], so that

$$H(z) = H_{min}(z) H_{ap}(z), \tag{39}$$

where $H_{min}(z)$ is related to a minimum phase response $h_{min}(n)$, such that

$$|H(f)| = |H_{min}(f)|, \tag{40}$$

and $H_{ap}(z)$ is a causal and stable all pass function, i.e., a function that engenders a frequency response is given by

$$\left|H_{ap}(f)\right| = 1.$$

[5] In seismics z is associated to unit delay instead of z^{-1}, and thus a minimum phase system under that definition has zeroes outside the unit circle (e.g. [SHERIFF, 2002]).

As pointed out in [OPPENHEIM and SCHAFER, 1989], the name "minimum phase" stems from the minimum phase-lag property of minimum phase systems. From (39), the phase spectrum of $H(z)$ is given by

$$\arg[H(f)] = \arg[H_{min}(f)] + \arg\left[H_{ap}(f)\right].$$

As the continuous phase of all pass systems is negative for all normalized frequencies in the interval $0 \leq f \leq 0.5$, the phase-lag, i.e., the negative of the phase, is always greater in this interval in a non-minimum phase system if compared to the phase lag of the minimum phase system which has the same amplitude spectrum. This allows the phase to be uniquely determined from the amplitude spectrum in minimum phase systems.

Also, among many other properties, minimum phase systems have the minimum energy delay [OPPENHEIM and SCHAFER, 1989] (or simply minimum delay [ROBINSON and TREITEL, 1980]) property. To understand this property, we first define the partial energy of an arbitrary phase system, $h(n)$, as

$$E(n) = \sum_{m=0}^{n} |h(m)|^2.$$

Then, for a minimum phase system, $h_{min}(n)$,

$$\sum_{m=0}^{n} |h(m)|^2 \leq \sum_{m=0}^{n} |h_{min}(m)|^2,$$

for all stable and causal with arbitrary phase $h(n)$ with the same amplitude spectrum as $h_{min}(n)$. This is interesting in the seismic case, as seismic sources try to mimic impulsive sources, so that the energy is located in the beginning of the wavelet, and thus, in some cases, it is reasonable to assume that the wavelet is minimum phase.

The characterization of minimum phase systems is important, as the use of prediction for deconvolution is based on these two hypotheses [ROBINSON and TREITEL, 1980]: Firstly, the reflectivity function may be modeled by as a white wide sense stationary (WSS) random process. Secondly, the seismic wavelet can be represented as the impulse response to an all-pole minimum phase system As described, for example, in [HAYKIN, 2001], under these hypotheses, the reflectivity can be recovered in a unsupervised fashion with the use of a unit-lag PEF, which can be defined as a filter whose output is the prediction error, represented by $e(n)$ in Figure 21. In this case, the input of the filter is represented by $x(n)$ and $l = 1$. Also, the number of coefficients of

$w(n)$ must be equal to or greater than the number of poles of the wavelet. In fact, the unit-lag PEF is related to spiking deconvolution, as it is actually equivalent to a zero lag spiking deconvolution filter, except for a scale factor, as shown, for example, in [YILMAZ, 2001].

Prediction error filtering is a relatively simple and robust way to perform deconvolution, as only second order statistics (SOS) of the seismic trace are used to compute the filter coefficients, as shown in (36). However, its applicability is constrained by its underlying hypothesis. First, the signal of interest, in our case the reflectivity, is considered to be made as a set of uncorrelated samples (white random process) as the PEF can be shown to be a *whitening filter* [ROMANO *et al.*, 2010], i.e.,

$$\mathrm{E}[e(n)e(n-k)] = 0, k \geq 1,$$

for a PEF with enough number of coefficients[6].

This shows that the prediction error will be a set of uncorrelated samples. This also indicates that the PEF will introduce distortions to the output if the samples of the signal of interest are correlated to each other. The second limitation of the method is on the minimum phase wavelet hypothesis. If the wavelet is not minimum phase, then the PEF still will yield an uncorrelated series of samples. However, this result will not correspond to the actual reflectivity.

Now, we show an example of on how the use of prediction error filtering for deconvolution results in different outcomes in the cases of minimum and mixed phase wavelets. The white random signal $s(n)$ in Figure 23 is convolved with the minimum phase wavelet, $h_{min}(n)$, and the mixed phase wavelet, $h_{mix}(n)$, depicted in Figure 22, resulting, respectively, in $x_{min}(n)$ and $x_{mix}(n)$, as in Figure 23. The autocorrelation function and the PEF are calculated for both resulting signals as shown in Figure 24. It is observed that the results in both cases are the same. This stems from the fact that autocorrelation functions are not affected by the phase of $h(n)$. This can be shown by the fact that if

$$x(n) = h(n) * s(n),$$

then, as in [MENDEL, 1991], using the definition of $H(f)$ in (38) we have:

$$S_x(f) = \sigma_s^2 |H(f)|^2, \tag{41}$$

where σ_s^2 is the variance of $s(n)$ and $S_x(f)$ is the *power spectral density* of $x(n)$ defined as:

[6] The whitening property is attained for a finite length PEF only in the case in which the input is a signal which can be modeled as an autoregressive process [HAYKIN, 2001]. Otherwise, the whitening property would be attained only asymptotically accordingly to the length of the PEF.

$$S_x(f) = \sum_{l=-\infty}^{\infty} r(l)e^{-j2\pi fl}.$$

Equation (41) shows that $S_x(f)$ can be interpreted as the Fourier transform of the autocorrelation function of $x(n)$, $r(l)$. Thus the fact that the autocorrelation of $x(n)$ does not carry the phase information of $h(n)$ comes from the fact that, as shown in (41), $S_x(f)$ is a function of $|H(f)|^2$ which, in turn, is the power spectrum of $h(n)$ and is not affected by the phase information of $h(n)$. As a consequence, as shown in Figure 24(c) and Figure 24(d), the resulting PEFs are the same for the minimum phase and mixed phase cases.

The PEFs are applied to the outputs of the minimum and mixed phase systems, resulting, respectively, in $y_{min}(n)$ and $y_{mix}(n)$, as shown in Figure 25. The original signal $s(n)$ is almost perfectly recovered in the minimum phase case, but only a distorted signal is obtained in the mixed phase case. If we compare the Z-plane plots of the PEF in Figure 24 and of the minimum phase system in Figure 22, the zeros of the PEF are at the same position of the poles, which allows the PEF to cancel the effect of the minimum phase system. From the decomposition in (39) we observe that, in the mixed phase system, the PEF is only able to cancel out the minimum phase part, but the all-pass component is left, which causes the residual distortion observed in $y_{mix}(n)$ in Figure 25. In the following sessions, some approaches to overcome the limitations imposed by the minimum phase hypothesis in the use of PEFs for unsupervised (also known as blind) deconvolution will be discussed.

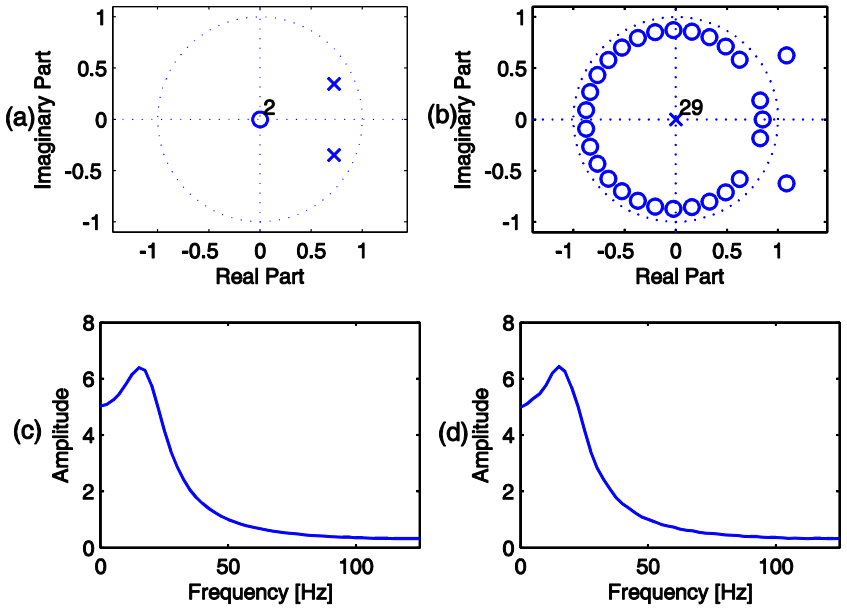


Figure 22: (a) Zero-pole plot of the minimum phase all-pole system. (b) Zero-pole plot of the mixed phase system. (c) Amplitude spectrum of the frequency response of the minimum phase all-pole system. (d) Amplitude spectrum of the frequency response of the mixed phase system.

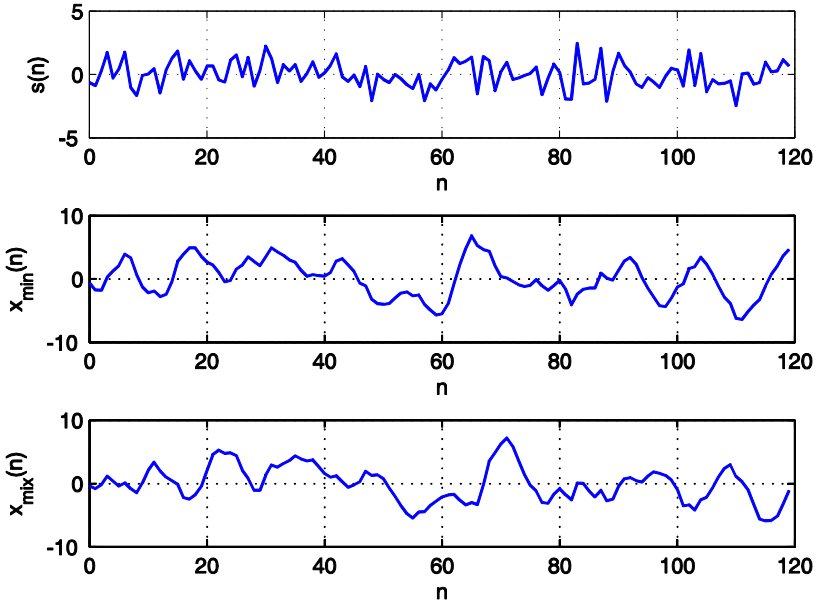


Figure 23: $s(n)$: white random signal. $x_{min}(n)$: Output from a minimum phase system shown in Figure 22. $x_{mix}(n)$: Output from a mixed phase system shown in Figure 22.The two systems have the frequency response with the same amplitude spectrum.

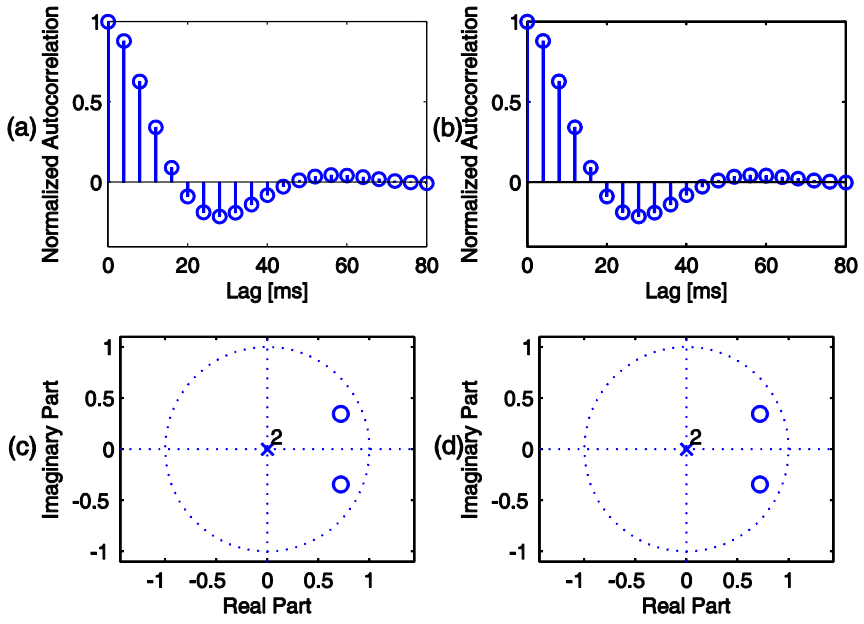


Figure 24: (a) and (b) Autocorrelation functions of the minimum phase and mixed phase systems. (c) and (d) Zero-pole plot of the estimated PEFs from the outputs of the minimum phase and mixed phase systems when the inputs were white signals.

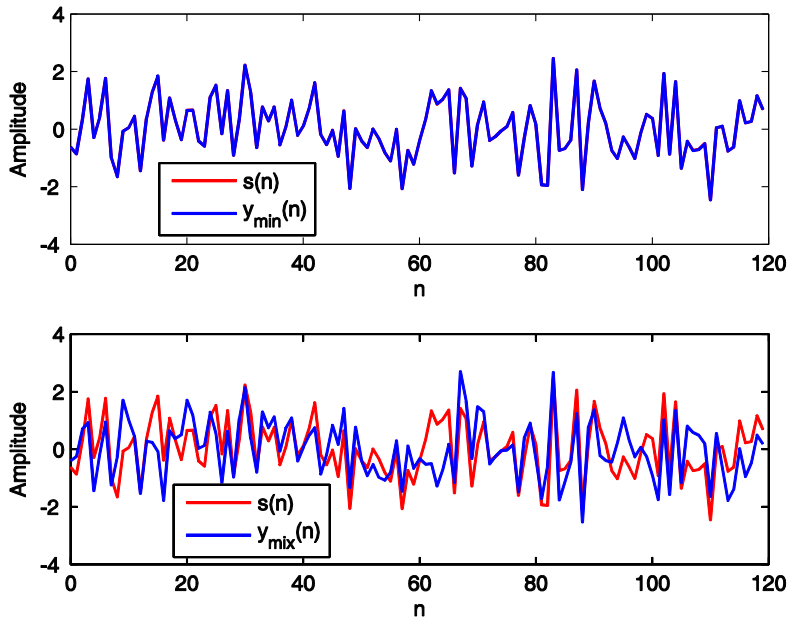


Figure 25: $s(n)$: white random signal. $y_{min}(n)$: output from PEF for the minimum phase system case. $y_{mix}(n)$: output from PEF for the mixed phase system case.

2.3.2 Mixed phase wavelet deconvolution and higher order statistics

As explained in the Subsection 2.3.1, the use of PEF for deconvolution is constrained by the fact that the phase information of the wavelet is not carried by the second order statistics. In order to overcome this, many techniques based on the use of higher order statistics (HOS), which carry phase information, were proposed. One of the first techniques in this sense was the minimum entropy deconvolution (MED) proposed in [WIGGINS, 1978]. In this work, no assumptions about the phase of the wavelet, $h(n)$, are made. Instead, the reflectivity, $s(n)$, is assumed to be sparse, i.e., to be composed of few large spikes. Thus, the method consists in obtaining a fixed length deconvolution filter, $w(n), n = 0, \cdots, N_w - 1$, from the information provided by a set of N_t traces, $x_i(n), i = 0, \dots, N_t - 1$, that produces a "simplest" set of outputs, $y_i(n), i = 1, \dots, N_t - 1$, or in other words, a set of outputs with the minimum entropy. In this case, it is assumed that all N_t traces have the same wavelet. Given that the outputs are obtained by

$$y_i(n) = w(n) * x_i(n),$$

Wiggins proposed the use of an optimization procedure to search for the values of $w(n)$ that maximize the varimax norm of the output, which is computed as

$$V_y = \sum_{i=0}^{N_t-1} K_{y_i},$$

where K_{y_i}, an estimate of the normalized kurtosis of $y_i(n)$, is given by

$$K_{y_i} = \frac{\sum_{n=0}^{N-1} y_i^4(n)}{(\sum_{n=0}^{N-1} y_i^2(n))^2}. \quad (42)$$

After further developments introduced in [OOE and ULRYCH, 1979; CLAERBOUT, 1978; GODFREY, 1978; GRAY, 1979], a theoretical framework was later provided in [DONOHO, 1981] for the single trace case, where the trace is assumed to be an independent and identically distributed (i.i.d.) non-Gaussian random process. Initially, the equivalence of two random variables with finite variance is defined. Under this definition, the random variables S and X are equivalent, expressed as $S \doteq X$, if S has the same probability distribution as $hX + c$, for some constants c and $h \neq 0$. Next, the operator $\succcurlyeq$ is defined such that $S \succcurlyeq X$ indicates that

$$X \doteq \sum_i h_i S_i,$$

where $\sum_i h_i^2 < \infty$ and S_i are independent random variables with the same distribution. The operator $\succcurlyeq$ has the properties of transitivity, which implies that

$$\text{if } S \succcurlyeq X \text{ and } X \succcurlyeq B \text{ then } S \succcurlyeq B,$$

and asymmetry, which implies that

$$\text{if } S \succcurlyeq X \text{ and } X \succcurlyeq S, \text{ then } S \doteq X.$$

In fact, one important observation about the second property is made in [DONOHO, 1981], which states that B is a Gaussian random variable if and only if B has finite variance and

$$B \doteq \sum_i h_i B_i, \tag{43}$$

where the summation represents a linear combination of at least two random variables B_i. From this, if B has Gaussian distribution, then $B \succ\!\!\!\!\!\!\cdot\ S$ if S is a non-Gaussian random variable, as the linear combination of two independent Gaussian random variables can only be Gaussian. In this case, the operator $\succ\!\!\!\!\!\!\cdot$ indicates that S cannot be written as a linear combination of random variables with the same distribution as B. Also, the central limit theorem states that $X_N = \frac{1}{\sqrt{N}} \sum_{i=0}^{N} X_i$ converges to a Gaussian random variable B for a large value of N when X has finite variance and zero mean. As $X \succcurlyeq X_N$ and $X_N \to B$, then it is taken as definition that $X \succcurlyeq B$ for all random variables X with an arbitrary probability distribution and finite variance. Thus, if $s_i(n)$ represents a single reflectivity function that is modeled as an i.i.d. random process with arbitrary distribution, $x_i(n)$ is the respective measured seismic trace and $b(n)$ a i.i.d. random process with Gaussian distribution, we have that

$$s_i(n) \succcurlyeq x_i(n) \succcurlyeq b(n), \tag{44}$$

The first $\succcurlyeq$ sign comes from the fact that

$$x_i(n) = h(n) * s_i(n), \tag{45}$$

which implies that $x_i(n)$ is a linear combination of delayed samples of $s_i(n)$, which are independent realizations of the same random variable, weighted by the coefficients of $h(n)$, and thus by definition we have that

$$s_i(n) \succcurlyeq x_i(n). \tag{46}$$

The second $\succcurlyeq$ sign in (44) is justified by the fact that $b(n)$ is a Gaussian distributed variable. (44) states that a linear distortion on $s_i(n)$ results in a signal $x(n)$ which is closer to a Gaussian. This fact allows a more general approach to MED. Since the objective of the deconvolution is to produce an output, $y_i(n)$, which is closer to the reflectivity, $s_i(n)$, than the seismic trace,

$x_i(n)$, the approach proposed by Donoho consists in using an objective function which discriminates how Gaussian a random variable is, i.e., an objective function $O[\cdot]$ such that

$$\text{if } y_i(n) \succcurlyeq x_i(n) \text{ then } \; O[y_i(n)] \geq O[x_i(n)]. \tag{47}$$

Thus, the MED framework proposed in [DONOHO, 1981] consists in finding an optimum deconvolution filter, $w(n)$, solving.

$$w(n) = \underset{\widetilde{w}(n)}{\operatorname{argmax}}\, O[y_i(n)] = \underset{\widetilde{w}(n)}{\operatorname{argmax}}\, O[\widetilde{w}(n) * x_i(n)]. \tag{48}$$

It is interesting to notice that this optimization makes sense only under the hypothesis that $s_i(n)$ is non-Gaussian. If this assumption is false, then $y_i(n)$ will have a Gaussian distribution for all $\widetilde{w}(n)$ such that $0 < \sum_{n=0}^{N_w-1} \widetilde{w}^2(n) < \infty$, and thus $O[\cdot]$ will not be useful to discriminate how $y_i(n)$ is close to the actual reflectivity $s_i(n)$.

Next, it is shown in [DONOHO, 1981] that several measures are consistent with the condition described in (47). The first of them are based on standardized cumulants. The cumulants of a random variable X are obtained from the cumulant generating function (e.g. [ROMANO *et al.*, 2010]):

$$\Upsilon_X(\omega) = \ln\{E[\exp(j\omega X)]\},$$

where $E[\cdot]$ denotes expectation. The m-th cumulant of X is then defined as

$$C_X^m = (-j)^{-m} \left. \frac{\partial^m \Upsilon_X(\omega)}{\partial \omega^m} \right|_{\omega=0},$$

and the corresponding standardized cumulant is defined as

$$\kappa_X^m = \frac{C_X^m}{(C_X^2)^{m/2}}.$$

It is shown that

$$\text{if } S \succcurlyeq X \text{ then } \; |\kappa_S^m| \geq |\kappa_X^m|, m > 2.$$

The second cumulant is simply the variance and thus, as discussed previously, it does not carry the information about the phase. Therefore, only cumulants with $m > 2$ may be considered for MED. The standardized cumulant of $y_i(n)$ with $m = 4$ is called kurtosis and is estimated by

$$\left|\kappa_{y_i}^4\right| = \left|K_{y_i} - 3\right|,$$

where K_{y_i} is defined in (42) if the reflectivity, $s_i(n)$, and thus also $y_i(n)$, have zero mean. From this result, it is concluded in [DONOHO, 1981] that maximizing K_{y_i} as initially advocated in

[WIGGINS, 1978] is consistent with (47) if the reflectivity respects $K_{s_i} > 3$, which is a quantitative way to express the sparsity hypothesis in the original work. If $K_{s_i} < 3$, then the deconvolution should be done by minimizing the varimax norm.

A more direct link to Shannon differential entropy, defined for X as

$$H_X = -\int_{-\infty}^{\infty} p_X(x) \ln\big(p_X(x)\big)\, dx,$$

is also provided in [DONOHO, 1981] in the case where the wavelet is normalized as $\sum_{n=0}^{N-1} h^2(n) = 1$. Then it is possible to show that [BERCHER and VIGNAT, 2000]:

$$-H_{s_i} \geq -H_{x_i},$$

for $x_i(n)$ and $s_i(n)$ expressed as in (45). This shows that the minimization of Shannon entropy is also consistent with (47) and (48). Thus all deconvolution procedures consistent with the principle described in (47), i.e., based on a optimization procedure as described in (48) truly work by reducing the entropy of the output and hence they receive the name of MED-type deconvolution. A further discussion can be found in [WALDEN, 1985].

Also, in the context of data communications theory, two other theorems give support to the role of HOS in unsupervised deconvolution. The first one is the Benveniste-Goursat-Rouget (BGR) theorem [BENVENISTE *et al.*, 1980]. As in the case studied in [DONOHO, 1981], a signal of interest, $s(n)$, modeled as an i.i.d. non-Gaussian random process, is observed through an LTI system, which introduces distortions on the original signal, producing a series of observations, $x(n)$. The signal of interest is estimated with the use of another LTI system, i.e., the deconvolution filter. The theorem states that if one is able to design a filter such that the probability density function (pdf) of the signal at its output, $y(n)$, has the same distribution as $s(n)$, then the deconvolution will be perfect up to a delay, d, and a complex unit-magnitude gain, α, which can be represented as

$$y(n) = \alpha s(n - d). \tag{49}$$

In fact, ambiguity of amplitude and delay also happens in MED-type of deconvolution, as described by [DONOHO, 1981], because the functions chosen for the optimization in (48) intend to quantify how far the distribution of a random variable is away from the Gaussian distribution, as described in (47), and, as such, these functions are invariant to scale and delay.

The second theorem was proposed in [SHALVI and WEINSTEIN, 1990], which showed that the pdf matching requirement of the BGR theorem was excessively rigorous. The Shalvi-

Weinstein (SW) theorem states that under the same conditions as the BGR theorem, it is sufficient to obtain $E[y^2(n)] = E[s^2(n)]$ and a match between a cumulant of order higher than two of $y(n)$ and $s(n)$ so that the deconvolution is perfect up to the conditions described in (49). This theorem thus greatly simplifies the requirements for blind deconvolution, as the full pdf matching requirement in the BGR theorem implies the matching of statistics of all orders while accordingly to the SW theorem it suffices to perform power normalization and then a match of a single HOS. As in the discussion in [TUGNAIT, 1992], the SW approach is based on same principles proposed by Wiggins and Donoho. Thus all these works complement each other in the sense of supporting the use of HOS for unsupervised deconvolution.

Other approaches that involve the use of HOS, in particular cumulants, for unsupervised deconvolution include the fourth order cumulant matching approach developed by [TUGNAIT, 1987] in the context of linear system identification. It was first applied in real seismic data in [LAZEAR, 1993] and further analyzed and developed in works such as [HARGREAVES, 1994; VELIS and ULRYCH, 1996; MISRA and SACCHI, 2007; MISRA and CHOPRA, 2010]. Further discussion concerning the use of HOS based in cumulants for deconvolution and wavelet estimation may be found, for example, in [NIKIAS and RAGHUVEER, 1987; MENDEL, 1991; NIKIAS and MENDEL, 1993; SACCHI and ULRYCH, 2000].

2.3.3 Independent component analysis and blind source separation

Blind source separation (BSS) is another important application of unsupervised signal processing described, for example, in [HYVÄRINEN *et al.*, 2001; COMON and JUTTEN, 2010; ROMANO *et al.*, 2010]. It is related to the unsupervised (or blind) deconvolution problem, in the sense that this problem also consists in estimating a set of quantities of interest from information obtained from some observations at the output of an unknown distorting system and few statistical assumptions. The difference is that in BSS the distorting system has multiple inputs and multiple outputs. Thus, as shown in Figure 26, in this class of problems, *N* different sources generate initially a set of signals at each sample time n, called snapshots, represented by the vector

$$\mathbf{s}(n) = [s_1(n) \quad s_2(n) \quad \dots \quad s_N(n)]^T.$$

These signals are observed through a system that mixes them and distorts them. A set of M sensors capture the outputs of this system, forming the set of observations or mixtures represented by

$$\mathbf{x}(n) = [x_1(n) \quad x_2(n) \quad \dots \quad x_M(n)]^T.$$

Hence, the BSS problem consists in recovering the original signals, i.e., performing source separation only from the information brought by the mixtures, without *a priori* knowledge of the mixing system.

If we assume that the mixing system is linear, time invariant and memoryless[7], i.e., that the mixing system can be represented by a $M \times N$ matrix $\mathbf{A}$, then we may write

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), \tag{50}$$

where the elements of $\mathbf{s}(n)$ and $\mathbf{x}(n)$ correspond to the respective sources or mixtures.

If the matrix is square, i.e., $M = N$, so that the number of the sources and mixtures is the same, as well as invertible, then the BSS problem may be cast as the estimation of the inverse matrix, $\mathbf{A}^{-1}$, only from the mixtures given by the vector $\mathbf{x}(n)$. In [COMON, 1994], it was shown that if, analogously to the of unsupervised deconvolution, the source signals, $s_i(n)$, are mutually independent and non-Gaussian, then it is possible to find such matrix up to some ambiguities analogous to the gain and delay factors in unsupervised deconvolution, given by α and d in (49). The principle of Comon's approach stems from the Darmois-Skitovich theorem [KAGAN *et al.*, 1973], which, in fact, also underlies the theory of MED developed in [DONOHO, 1981]. To explain the theorem and its impact, let us consider two random variables y_1 and y_2 such that

$$y_1 = a_1 s_1 + a_2 s_2 + \cdots + a_N s_N$$
$$y_2 = b_1 s_1 + b_2 s_2 + \cdots + b_N s_N,$$

where $s_1, s_2, \dots, s_N$ are zero mean mutually independent variables and a_i and b_i are constants. The theorem states that if y_1 and y_2 are statistically independent and $a_i b_i \neq 0$ for more than one value of i, then this implies that s_i are Gaussian random variables for all $a_i b_i \neq 0$. Therefore, this means that independent variables cannot result from a mixture of non-Gaussian variables. Thus,

[7] In problems in fields such as telecommunications and audio processing, mixtures are made by the superposition of delayed and scaled versions of source signals and are hence called convolutive mixtures. If proper hypotheses are met, then techniques described in, e.g., [HYVÄRINEN *et al.*, 2001; COMON and JUTTEN, 2010] may be used to perform BSS in convolutive mixtures. However, these techniques are out of the scope of this work, as we used the technique called Banded ICA (B-ICA), proposed in [KAPLAN and ULRYCH, 2003; KAPLAN, 2003], which allows the use of the ICA for memoryless systems for performing deconvolution, as described in Subsection 2.3.4.

if the inputs of the mixing systems are mutually independent and non-Gaussian, separation can be obtained by calculating an $N \times N$ separation matrix $\mathbf{W}_{ICA}$ such that, when

$$\mathbf{y}(n) = \mathbf{W}_{ICA}\mathbf{x}(n),$$

we get a vector $\mathbf{y}(n)$, where

$$\mathbf{y}(n) = [y_1(n) \quad y_2(n) \quad \dots \quad y_N(n)]^T,$$

and all resulting signals $y_i(n)$ are mutually independent. In this case, the original sources are recovered up to a scale factor and permutations, as scaling and changing the order of the signals in $\mathbf{y}(n)$ will not change the fact that these signals are statistically independent. As this approach to BSS involves obtaining independent components, it is called independent component analysis (ICA).

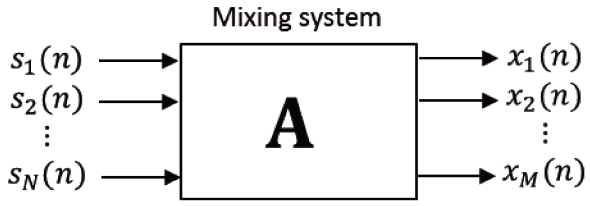


Figure 26: Mixing system

Before considering the methods to perform ICA, let us first consider a SOS based preprocessing step called whitening (e.g., [HYVÄRINEN *et al.*, 2001; ROMANO *et al.*, 2010]). As it will be shown in the following, this preprocessing step allows one to restrict the search of a separation matrix to the domain of orthogonal matrices, thus simplifying the task of ICA.

Whitening is analogous to the use of PEF in the context of deconvolution, in the sense that PEFs are also whitening filters and also use only SOS. In the whitening stage, correlation information is used to obtain an $N \times N$ matrix $\mathbf{W}_W$ so that

$$\mathbf{z}(n) = \mathbf{W}_W\mathbf{x}(n), \tag{51}$$

where

$$\boldsymbol{z}(n) = [z_1(n) \quad z_2(n) \quad \dots \quad z_N(n)]^T,$$

and the output signals $z_i(n)$ are uncorrelated and have unit variance, i.e.,

$$E\{z_i(n)z_j(n)\} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases},$$

or

$$\mathbf{R}_{\mathbf{z}} = E\{\mathbf{z}(n)\mathbf{z}^T(n)\} = \mathbf{I},$$

where $\mathbf{R_z}$ is the correlation matrix of $\mathbf{z}(n)$ and $\mathbf{I}$ is the identity matrix.

In order to obtain $\mathbf{W}_W$, let us first consider the eigendecomposition of the correlation matrix of $\mathbf{x}(n)$ given by

$$\mathbf{R_x} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\} = \mathbf{E}\mathbf{D}\mathbf{E}^T,$$

where

$$\mathbf{E} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_N]$$

is a matrix made with the orthogonal unit norm eigenvectors $\mathbf{e}_i$ of $\mathbf{R_x}$ and

$$\mathbf{\Lambda} = \mathrm{diag}([\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_N])$$

is the diagonal matrix that contains the respective eigenvalues.

Then, it is possible to show that

$$\mathbf{W}_W = \mathbf{\Lambda}^{-1/2}\mathbf{E}^T, \tag{52}$$

as substitution into (51), followed by the calculation of the correlation matrix of $\mathbf{z}(n)$, leads to

$$\mathbf{R_z} = E\{\mathbf{z}(n)\mathbf{z}^T(n)\} = \mathbf{W}_W E\{\mathbf{x}(n)\mathbf{x}^T(n)\}\mathbf{W}_W^T = \left(\mathbf{\Lambda}^{-1/2}\mathbf{E}^T\right)(\mathbf{E}\mathbf{\Lambda}\mathbf{E}^T)\left(\mathbf{E}\mathbf{\Lambda}^{-1/2}\right) = \mathbf{I}. \tag{53}$$

It is interesting to notice that this result is not unique. In order to show this, consider another matrix such that

$$\mathbf{W'}_W = \mathbf{U}\mathbf{W}_W,$$

where $\mathbf{U}$ is an orthogonal matrix, i.e., $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$. Then it is possible to show that $\mathbf{W'}_W$ is also a whitening matrix for all possible values of $\mathbf{U}$ because if

$$\mathbf{z}'(n) = \mathbf{W'}_W\mathbf{x}(n) = \mathbf{U}\mathbf{W}_W\mathbf{x}(n) = \mathbf{U}\mathbf{z}(n), \tag{54}$$

then we obtain, by using (53), that

$$\mathbf{R_{z'}} = E\{\mathbf{z}'(n)\mathbf{z}'^T(n)\} = \mathbf{U}E\{\mathbf{z}(n)\mathbf{z}^T(n)\}\mathbf{U}^T = \mathbf{U}\mathbf{R_z}\mathbf{U}^T = \mathbf{I}.$$

Thus, this shows that it is not enough to perform whitening, and hence only SOS, to perform source separation, as uncorrelated signals $\mathbf{z}'(n)$ may be the result of the linear combination of other uncorrelated signals, $\mathbf{z}(n)$, as shown in (54). Also, it is important to observe that this shows that ICA is not suited when BSS is performed for two or more Gaussian sources because uncorrelated Gaussian random variables are also mutually independent. This means that if there are more than one Gaussian sources, an independent component that has a Gaussian distribution may still be the linear combination of these Gaussian sources instead of an isolated source signal.

Nevertheless, following [HYVÄRINEN *et al.*, 2001], we may use the whitening as a preprocessing step for ICA. If we substitute (50) into (51), we obtain that

$$\mathbf{z}(n) = \mathbf{W}_W \mathbf{A}\mathbf{s}(n) = \mathbf{A}'\mathbf{s}(n), \tag{55}$$

where

$$\mathbf{A}' = \mathbf{W}_W \mathbf{A},$$

represents a residual mixing matrix. If we further substitute (55) into (53) and we assume, without loss of generality, that the signals from the sources have zero mean and unit variance, i.e., that the correlation matrix of $\mathbf{s}(n)$ is given by

$$\mathbf{R}_s = E\{\mathbf{s}(n)\mathbf{s}^T(n)\} = \mathbf{I},$$

we obtain that

$$E\{\mathbf{z}(n)\mathbf{z}^T(n)\} = \mathbf{A}'E\{\mathbf{s}(n)\mathbf{s}^T(n)\}\mathbf{A}'^T = \mathbf{A}'\mathbf{A}'^T = \mathbf{I},$$

which means that the residual mixing matrix is also an orthogonal matrix. Thus, as previously mentioned, if the whitening is used as a preprocessing step, ICA may be performed by just finding an orthogonal separating matrix that is able to eliminate the effect of the residual mixing matrix, which simplifies the ICA implementation.

A classic example that illustrates whitening [HYVÄRINEN *et al.*, 2001; ROMANO *et al.*, 2010] considers two independent sources $s_1(n)$ and $s_2(n)$, which are modeled as random variables with uniform distribution and zero mean and unit variance, so that their joint pdf is given by

$$p_s(s_1, s_2) = \begin{cases} 1/12 & \text{if } -\sqrt{3} \leq s_1, s_2 \leq \sqrt{3} \\ 0 & \text{otherwise} \end{cases}.$$

A 10000-sample realization of $s_1(n)$ and $s_2(n)$ is shown in Figure 27 in the form of a scatterplot, in which a small blue dot is placed at the coordinate given by $(s_1(n), s_2(n))$. Next, these signals are distorted by the mixing matrix, given by

$$\mathbf{A} = \begin{bmatrix} 8.5 & 0.5 \\ 9 & 2 \end{bmatrix}.$$

The scatter plot for the resulting mixtures $x_1(n)$ and $x_2(n)$ is shown in Figure 28. The linear distortion caused by the mixing matrix stretches and rotates the original square into a parallelogram. It is interesting to notice that, given the information about one of the mixtures, for example, $x_1(n)$, then it is possible to infer some information about $x_2(n)$. It is possible to show that the range of values attainable by $x_2(n)$ when $x_1(n) = 15$ is completely different when $x_1(n) = -15$. Thus, this shows that the mixtures are not independent. The result of whitening is shown in Figure 29. The square shape is recovered, but, as in Figure 28, it is possible to verify that the whitened signals are not independent due to a rotation of the square. This rotation is a geomet-

rical manifestation of the residual mixing matrix from the whitening process, as orthogonal matrices can also be geometrically interpreted as rotation matrices.

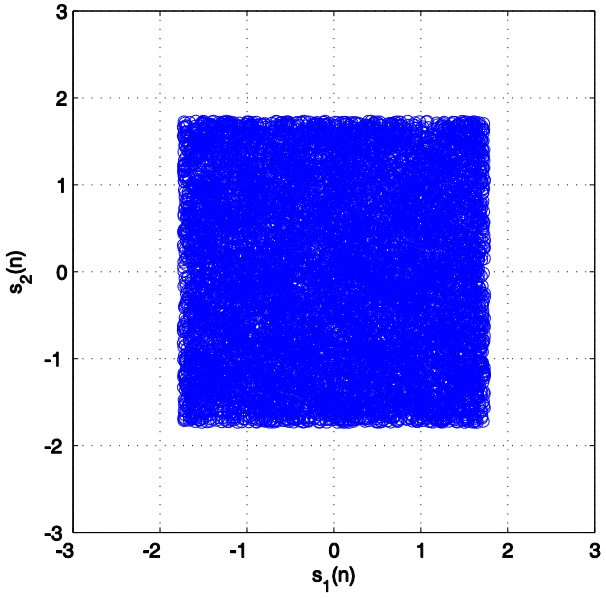


Figure 27: Scatter plot of the independent sources: $s_1(n)$ and $s_2(n)$

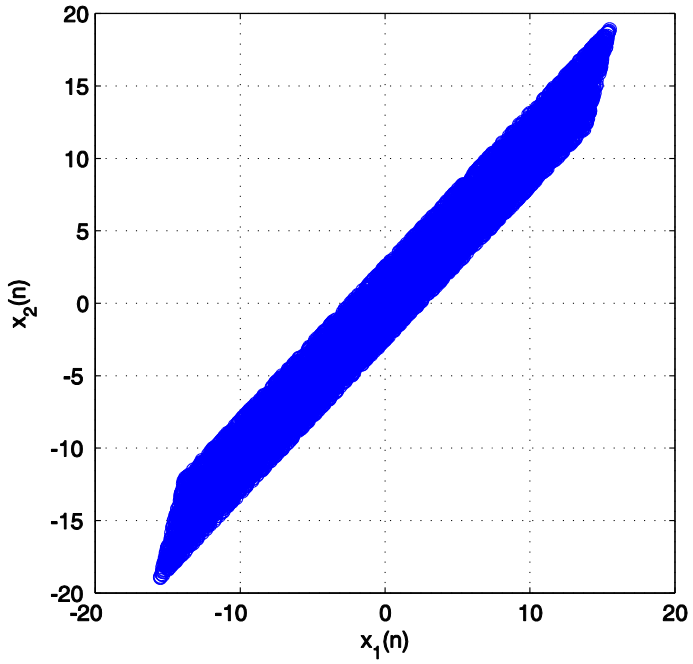


Figure 28: Scatter plot of the mixtures: $x_1(n)$ and $x_2(n)$

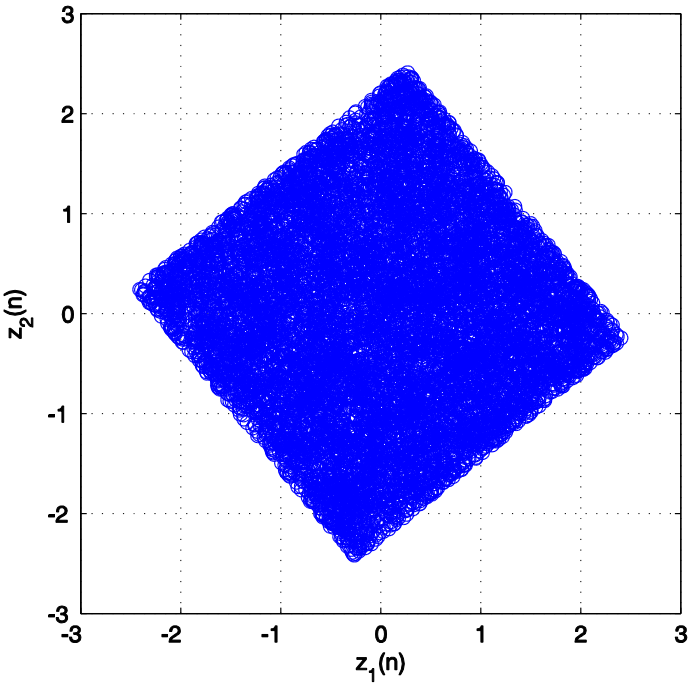


Figure 29: Scatter plot of the whitened outputs: $z_1(n)$ and $z_2(n)$

In order to overcome the limitations of SOS and recover the residual mixing matrix, ICA techniques using HOS were discussed in [COMON, 1994]. These techniques seek an orthogonal matrix $\mathbf{Q}$ so that

$$\mathbf{y}(n) = \mathbf{Q}\mathbf{z}(n) = \mathbf{Q}\mathbf{W}_W\mathbf{x}(n) = \mathbf{Q}\mathbf{A}'\mathbf{s}(n) = \mathbf{D}\mathbf{P}\mathbf{s}(n), \tag{56}$$

where $\mathbf{y}(n)$ represent the independent components and also an estimate of the independent sources up to a scale factor, represented by the diagonal matrix, **D**, and a permutation, represented by the permutation matrix, **P.**

2.3.4 Deconvolution based in B-ICA

In order to show an application of unsupervised signal processing in seismic signal processing we now focus on unsupervised deconvolution with a variation of ICA, called banded ICA (B-ICA), introduced in [KAPLAN and ULRYCH, 2003; KAPLAN, 2003]. As in [TAKAHATA *et al.*, 2012], we choose this method as it combines the two problems in unsupervised signal processing addressed so far, namely, unsupervised deconvolution and ICA.

As in (7), the convolutional model is given by

$$x(n) = h(n) * s(n),$$

where the noise was neglected and, as discussed previously, $x(n)$ represents the seismic trace, which is a version of the seismic reflectivity, $s(n)$, corrupted by the seismic wavelet, $h(n)$. In order to fit this model into the mixing model in (50), which underlies ICA, we first rewrite the convolution as a matrix operation, as in Section 2.2, so that

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{57}$$

where $\mathbf{x}$ and $\mathbf{s}$ are $N_s \times 1$ vectors given by

$$\mathbf{s} = [s(0) \quad s(1) \quad \dots \quad s(N_s + 1)]^T$$
$$\mathbf{x} = [x(0) \quad x(1) \quad \dots \quad x(N_s + 1)]^T$$

and $\mathbf{A}$ is a $N_s \times N_s$ convolution matrix described as

$$\mathbf{A} = \begin{bmatrix} h(0) & 0 & \dots & 0 & 0 & \dots & 0 \\ h(1) & h(0) & \dots & 0 & 0 & \ddots & 0 \\ \vdots & h(1) & \ddots & \vdots & \vdots & \ddots & 0 \\ h(N_h - 1) & \vdots & \ddots & h(0) & 0 & \ddots & 0 \\ \vdots & h(N_h - 1) & \ddots & h(1) & h(0) & \ddots & \vdots \\ 0 & \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & h(N_h - 1) & h(N_h - 2) & \dots & h(0) \end{bmatrix}$$
$$= [\mathbf{N}_0\mathbf{h} \quad \mathbf{N}_1\mathbf{h} \quad \dots \quad \mathbf{N}_{N_s-1}\mathbf{h}],$$

where the seismic wavelet is represented by an $N_h \times 1$ vector

$$\mathbf{h} = [h(0) \quad h(1) \quad \dots \quad h(N_h + 1)]^T,$$

and $\mathbf{N}_k$ is a zero padding matrix that such that

$$\mathbf{a}_k = \mathbf{N}_k\mathbf{h}, \tag{58}$$

where $\mathbf{a}_k$ is the k-th column of $\mathbf{A}$. By comparing (57) and (50), we observe that the trace and the reflectivity play the role of the receivers and sources and the convolution matrix is equivalent to the mixing matrix. However, we observe that (57) provides a single snapshot, which is not adequate for ICA as an ensemble of snapshots are needed to compute the statistics in the search of the independent components. As delayed inputs leads to outputs with the same delay in the convolution operation, delayed versions of the seismic trace and reflectivities can be considered

$$\mathbf{s}(n) = [s(n - N_s + 1) \quad \dots \quad s(n - 1) \quad s(n)]^T$$
$$\mathbf{x}(n) = [x(n - N_s + 1) \quad \dots \quad x(n - 1) \quad x(n)]^T,$$

where $n = 0, \dots, N_s - 1$ and $x(n) = 0$ and $s(n) = 0$ for $n < 0$. Thus, if replace $\mathbf{s}$ and $\mathbf{x}$ from (57) with the vectors defined by $\mathbf{s}(n)$ and $\mathbf{x}(n)$, we obtain

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), n = 0, \dots, N_s - 1,$$

which is the same as (50) for N_s snapshots corresponding to N_s sources and N_s receivers. If we consider the following $N_s \times N_s$ matrices

$$\mathbf{X} = [\mathbf{x}(0) \quad \mathbf{x}(1) \quad \dots \quad \mathbf{x}(N_s - 1)]$$
$$= \begin{bmatrix} 0 & 0 & \dots & 0 & x(0) \\ \vdots & \vdots & \iddots & x(0) & \vdots \\ 0 & x(0) & \iddots & \iddots & x(N_s - 2) \\ x(0) & x(1) & \dots & x(N_s - 2) & x(N_s - 1) \end{bmatrix},$$

and

$$\mathbf{S} = [\mathbf{s}(0) \quad \mathbf{s}(1) \quad \dots \quad \mathbf{s}(N_s - 1)\,]$$
$$= \begin{bmatrix} 0 & 0 & \dots & 0 & s(0) \\ \vdots & \vdots & \iddots & s(0) & \vdots \\ 0 & s(0) & \iddots & \iddots & s(N_s - 2) \\ s(0) & s(1) & \dots & s(N_s - 2) & s(N_s - 1) \end{bmatrix},$$

then we obtain that

$$\mathbf{X} = \mathbf{A}\mathbf{S}.$$

Thus, each source and receiver corresponds to a delayed version of, respectively, the seismic reflectivity and the seismic trace. If we assume that the reflectivity is composed of a sequence of non-Gaussian and i.i.d variables, as in the case of MED deconvolution [DONOHO, 1981], then ICA could, in principle, be used to obtain an independent component from the set of delayed versions of the traces, **X**, that would correspond to an original source that represents the original reflectivity or a delayed version of it. Nevertheless, as pointed out in [KAPLAN and ULRYCH, 2003], the use of plain ICA does not take into account the banded structure of the **A** matrix. Also, they point out that the search for independent components could be harmed by the fact that the first rows of matrices **X** and **S** contain few non-zero realizations, which would distort the computations for the statistics of the independent components.

In order to overcome these issues, a variant of ICA, the B-ICA, is proposed in [KAPLAN and ULRYCH, 2003] so that unsupervised deconvolution can be performed. The B-ICA consists of the following steps:

Step 1- Data Rearrangement:

The number of mixtures is reduced to $N'_s < N_s$by eliminating the first $(N'_s - N_s)$ rows of $\mathbf{S}$ and $\mathbf{X}$, thus obtaining

$$\mathbf{S}' = \begin{bmatrix} 0 & 0 & \cdots & 0 & s(0) & \cdots & s(N'_s - N_s) \\ \vdots & \vdots & \iddots & \iddots & \vdots & \iddots & \vdots \\ 0 & s(0) & \cdots & \cdots & s(N'_s - 2) & \cdots & s(N_s - 2) \\ s(0) & s(1) & \cdots & \cdots & s(N'_s - 1) & \cdots & s(N_s - 1) \end{bmatrix} \tag{59}$$

$$\mathbf{X}' = \begin{bmatrix} 0 & 0 & \cdots & 0 & x(0) & \cdots & x(N'_s - N_s) \\ \vdots & \vdots & \iddots & \iddots & \vdots & \iddots & \vdots \\ 0 & x(0) & \cdots & \cdots & x(N'_s - 2) & \cdots & x(N_s - 2) \\ x(0) & x(1) & \cdots & \cdots & x(N'_s - 1) & \cdots & x(N_s - 1) \end{bmatrix} \tag{60}$$

and the n-th column of $\mathbf{X}'$ is denoted by $\mathbf{x}'(n)$. Although a $N'_s \times N'_s$ reduced convolution matrix $\mathbf{A}'$ no longer provides an exact mapping between $\mathbf{S}'$ and $\mathbf{X}'$, accordingly to [KAPLAN and ULRYCH, 2003], the impact of this error is small and as this eliminates the excessive zeroes on the original mixtures, it improves the statistics of the mixtures and the performance of the ICA algorithm.

Step 2 – Data whitening

As discussed in the previous sub-section a whitening matrix $\mathbf{W}_W$ is calculated using SOS of the mixtures $\mathbf{x}'(n)$, as described in (52), and thus we decorrelate and equalize the variance of the output signals as in (51), resulting in

$$\mathbf{z}'(n) = \mathbf{W}_W \mathbf{x}'(n).$$

Step 3 – Banded structure reinforcement

As shown in (56), some ICA procedures work by obtaining an orthogonal matrix $\mathbf{Q}$ that produces a set of independent components such that.

$$\mathbf{y}'(n) = \mathbf{Q}\mathbf{z}'(n) = \mathbf{Q}\mathbf{W}_W \mathbf{x}'(n). \tag{61}$$

Let us now consider an auxiliary matrix $\mathbf{V}$ so that

$$\mathbf{V}\mathbf{y}'(n) = \mathbf{x}'(n). \tag{62}$$

By comparing (61) and (62) we obtain

$$\mathbf{V} = (\mathbf{Q}\mathbf{W}_W)^{-1} = \mathbf{W}_W^{-1}\mathbf{Q}^T,$$

from which,

$$\mathbf{Q}^{T} = \mathbf{W}_{W}\mathbf{V}.$$

Thus,

$$\mathbf{q}_k = \mathbf{W}_{W}\mathbf{v}_k,$$

where $\mathbf{q}_k$ is the k-th row of $\mathbf{Q}$ and $\mathbf{v}_k$ is the k-th column of $\mathbf{V}$. As the independent components $\mathbf{y}'(n)$ represents the sources up to a scale factor and permutation, then it is reasonable to assume that $\mathbf{V}$ and the mixing matrix $\mathbf{A}$ have similar structures, as (61) also represents a mapping between sources and mixtures. Thus, if we consider the expression for the k-th column of $\mathbf{A}'$, similarly to (58), we may assume

$$\mathbf{v}_k = \mathbf{N'}_k\mathbf{h},$$

where $\mathbf{N'}_k$ is the zero padding matrix that maps $\mathbf{h}$ into the k-th column of $\mathbf{A}'$.

Now, if we analyze the case for the k-th independent component we obtain from (61)

$$\tilde{y'}_k(n) = \mathbf{q}_k^{T}\mathbf{z}'(n) = (\mathbf{W}_{W}\mathbf{N'}_k\mathbf{h})^{T}\mathbf{z}'(n) = \mathbf{h}^{T}\mathbf{N'}_k^{T}\mathbf{W}_{W}^{T}\mathbf{z}'(n) = \mathbf{h}^{T}\tilde{\mathbf{x}}(n), \qquad (63)$$

where $\tilde{\mathbf{x}}(n)$ is a new set of N_h new mixtures given by

$$\tilde{\mathbf{x}}(n) = \mathbf{N'}_k^{T}\mathbf{W}_{W}^{T}\mathbf{z}'(n).$$

As advocated by [KAPLAN and ULRYCH, 2003], by obtaining this new set of mixtures, the banded structure of the mixture matrix is reinforced. As shown in (63), the wavelet, given by $\mathbf{h}$, is now related to the mixtures and to the reflectivity estimation explicitly. Hence, the new set of mixtures, $\tilde{\mathbf{x}}(n)$, produced during this step, are further processed on the following steps.

Step 4 – ICA

As $\mathbf{h}$ and $y'_i(n)$ are still unknown, an unsupervised procedure, in the case ICA, is used to estimate them. As $\tilde{\mathbf{x}}(n)$ is not necessarily made of decorrelated signals, a new whitening procedure is applied for pre-processing, as in (51), resulting in

$$\tilde{\mathbf{z}}(n) = \widetilde{\mathbf{W}}_{W}\tilde{\mathbf{x}}(n),$$

where $\widetilde{\mathbf{W}}_{W}$ is the $N_h \times N_h$ whitening matrix obtained from $\tilde{\mathbf{x}}(n)$. Next, an $N_h \times N_h$ orthonormal matrix $\widetilde{\mathbf{Q}}$ is calculated with the use of an ICA criterion, resulting in the set of independent components given by

$$\tilde{\mathbf{y}}(n) = \widetilde{\mathbf{Q}}\tilde{\mathbf{z}}(n) = \widetilde{\mathbf{Q}}\widetilde{\mathbf{W}}_{W}\tilde{\mathbf{x}}(n) = \widetilde{\mathbf{W}}_{ICA}\tilde{\mathbf{x}}(n), \qquad (64)$$

where the separating matrix estimated with ICA is given by

$$\widetilde{\mathbf{W}}_{ICA} = \widetilde{\mathbf{Q}}\widetilde{\mathbf{W}}_{W}.$$

In our case, we used the FastICA algorithm, for performing ICA, as described in [HYVÄRINEN *et al.*, 2001]. However, other methods such as the negentropy based ICA methods used in the original implementation of B-ICA in [KAPLAN and ULRYCH, 2003; KAPLAN, 2003] are also possible. The discussion about the implication of the choice of the ICA method is out of the scope of this work and is regarded as a future topic of research.

Step 5 – Wavelet and reflectivity estimation

From (64), we observe that the ICA produces actually a collection of N_h candidates for wavelet and reflectivity estimation, given according to (63) by

$$\tilde{y}'_i(n) = \tilde{\mathbf{h}}_i^T \tilde{\mathbf{x}}(n),$$

where $\tilde{\mathbf{h}}_i^T$ corresponds to the i-th row of $\widetilde{\mathbf{W}}_{ICA}$. From $\mathbf{N}'_k$ in Step 3, as in [KAPLAN and ULRYCH, 2003], we choose the i-th wavelet that best predicts the actual trace by solving

$$i^* = \arg\min_i(\|\mathbf{x}'_k - c_i\hat{\mathbf{x}}'_i\|),$$

where

$$\mathbf{x}'_k = [x'(k - N'_s + 1) \quad x'(k - N'_s + 2) \quad \dots \quad x'(k - N'_s + N_s)],$$

such that, $x'(n) = 0$ if $n < 0$, correspond to the k-th mixture in $\tilde{\mathbf{x}}'(n)$ and $\hat{\mathbf{x}}'_i$ is given by

$$\hat{\mathbf{x}}'_i = [\hat{x}'_i(0) \quad \hat{x}'_i(1) \quad \dots \quad \hat{x}'_i(N_s)],$$

in which

$$\hat{x}'_i(n) = h_i(n) * y(n).$$

Also,

$$c_i = \frac{\mathbf{x}'^T_k \hat{\mathbf{x}}'_i}{\|\hat{\mathbf{x}}'_i\|^2},$$

as this minimizes $\|\mathbf{x}'_k - c_i\hat{\mathbf{x}}'_i\|$ for a given pair of $\mathbf{x}'_k$ and $\hat{\mathbf{x}}'_i$.

Additional step 6 – Wavelet shaping

The optimum estimate of the wavelet, $\tilde{\mathbf{h}}_*$ can then be used for estimating a wavelet shaping filter, such as the one described in (28), in order to convert it to a zero lag spike or a lagged spike. If we admit that the wavelet does not vary in the neighboring traces, this filter can also be

used to deconvolve a gather of traces. This step is a contribution of this work [TAKAHATA *et al.*, 2012].

The additional Step 6 was motivated in [TAKAHATA *et al.*, 2012] by the fact that, as shown in Section 2.4, the use of the B-ICA method (Steps 1 to 5) leads to a deconvolved trace that may be delayed if compared to the reflectivity, which prevents the method to be directly applied to seismic sections as the one displayed in Figure 30. As pointed out in [KAPLAN and ULRYCH, 2003], this delay is caused by the way the input samples are ordered in the input of the algorithm in (60) in Step 1. In the unsupervised scenario, this delay is unknown, as the delay is related to source ordering, as pointed out in (59) and the fact that ICA is not able to determine the source order due to a permutation ambiguity as indicated by the permutation matrix **P** in (56) in Subsection 2.3.3. This becomes an issue when deconvolving the traces of a seismic section because if each trace of the section is deconvolved with the B-ICA method individually, then the result will become a set of deconvolved traces where each trace has a different delay, which results in the loss of the lateral continuity of the seismic events. The use of Step 6 involves the choice of a representative trace in the seismic section so that B-ICA can be used for wavelet estimation. The Step 6, then, enables one to derive a deconvolution filter from the estimated wavelet. Then, the deconvolution of a seismic section becomes possible by applying the deconvolution filter to the traces of the seismic section under the hypothesis that the wavelet varies little within neighboring traces. In this case a delay will also be observed, but as all traces will have the same delay, the lateral continuity of the seismic events is preserved. A complementary method for estimating the delay, such as maximizing the correlation with the input traces, then becomes necessary. Later, as will be shown in Section 2.4, we observed that the use of Step 6 resulted in a superior result if compared to the deconvolved trace produced by the original method, as Step 6 seems to introduce some robustness to non-idealities of the reflectivity regarding the i.i.d. hypothesis.

2.4 RESULTS

After exploring the main topics in the theory of convolutional modeling in the seismic setting and in supervised and unsupervised seismic deconvolution and wavelet estimation, we will present results based on field and synthetic data in this subsection. Initially, we will characterize the data from logs obtained from measurements made in a borehole as displayed in Figure 30. Next, we will present a study on synthetic traces where we try to assess the effect of the phase of the wavelet and the statistical properties of the reflectivity on the unsupervised deconvolution algorithms previously mentioned. In the end, we apply both supervised and unsupervised deconvolution techniques to a subset of data from the migrated section shown in Figure 30, which was obtained by processing a field data.

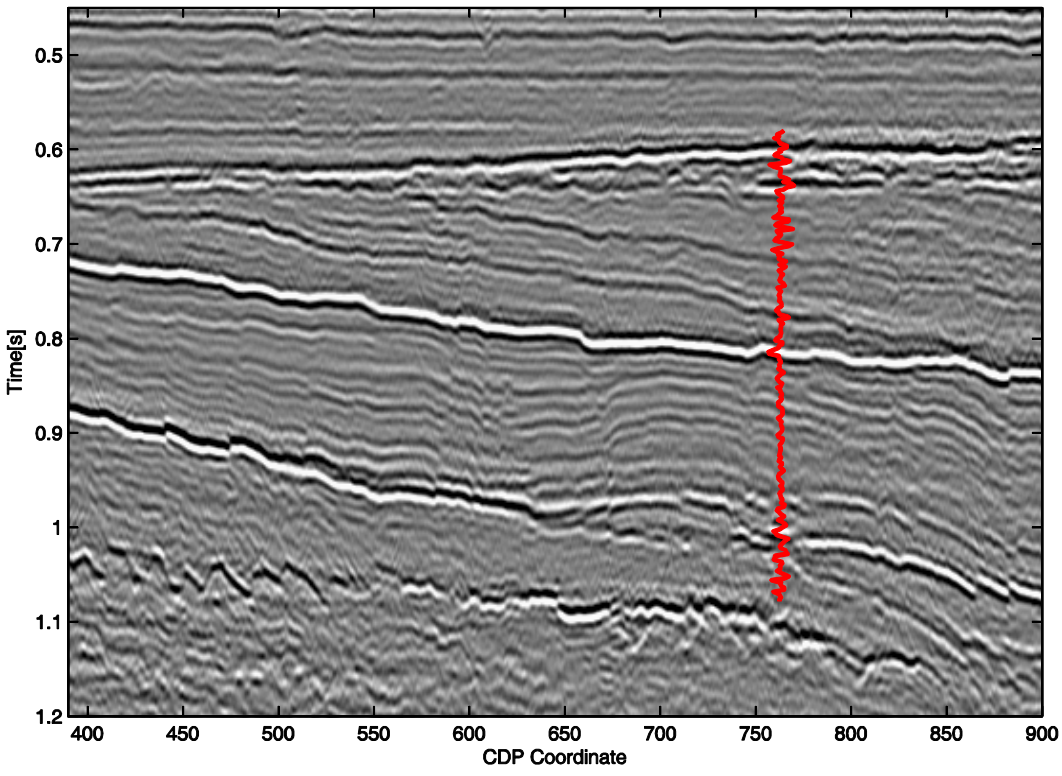


Figure 30: 2D time migrated seismic section. At the well position, the reflectivity function obtained from the well log is shown in red.

2.4.1 Reflectivity characterization

As fundamental hypotheses of unsupervised deconvolution are based on the statistics of the reflectivity function, in this section, we characterize a reflectivity function calculated from data collected at a well in offshore Norway. The corresponding position of the well at a 2D time migrated seismic section is indicated in Figure 30 as a red curve. This curve actually shows the reflectivity function calculated from a section of the information from the well log ranging from 0.58s to 1.08s. The actual well begins in a shallower part and reaches deeper parts of the well.

A well log is obtained by lowering a probe down the borehole which measures one or more characteristics of the local geology, which include resistivity, P-wave propagation velocity (sonic log), density, neutron porosity, etc. [SHERIFF, 2002]. One characteristic of the well log is that the resolution is very high if compared to the seismic log. The sampling interval of the well log is usually in meters or even in centimeters, while the resolution of seismic data is usually given in tens of meters. For instance, let us consider the typical case of a wavelet with dominant frequency, f_0, of 30Hz travelling through a medium with a velocity of $c = 2000$m/s. Then the dominant wavelength is

$$\lambda_0 = \frac{c}{f_0} = 66.67\text{m}.$$

The actual vertical resolution limit, i.e., the minimum vertical separation between two reflectors so that these reflectors are distinguishable in seismic data, can be approximated by $\lambda_0/4$ [KALLWEIT and WOOD, 1982]. In this case the vertical resolution limit is 16.67m.

Another difference of the well log and the seismic log is that the first one is composed of measurements as a function of depth, while the second one is composed by measurements as a function of time. The procedure used to match both measurements is called well tying [WHITE and SIMM, 2003]. In this procedure, the well data are pre-processed and a reflectivity series in time is produced from the sonic and density logs using the equations described in Section 2.1. Later, a wavelet is calculated with procedures similar to the one described in Section 2.2, and a synthetic seismogram is produced. The synthetic seismogram is then compared to the seismic data and further refinements on the wavelet estimation and on the conversion from depth to time domains are executed. These refinements are made so that the events in the synthetic trace tie to the seismic data. Migrated seismic sections, as shown in Figure 30, are normally used, as migra-

tion algorithms eliminate dip ambiguities and places reflectors and scatterers on the correct positions, in the ideal case, as discussed in Chapter 1. The logs used here were previously tied to the seismic data with the use of professional software (CGG-Geoview).

Initially we consider a reflectivity function calculated from a very highly sampled density and sonic logs as shown in Figure 31, in which the adopted sampling rate is 0.2ms. In Figure 32, the power spectral density function indicates that most frequency content of the reflectivity function is contained between 250Hz and 1000Hz approximately. However, it is important to notice that these frequencies are not present on the seismic data, as shown in Figure 33, as they are absorbed by the earth. The synthetic seismogram, after proper resampling to the same rate as the seismic trace (2 ms) and estimating the wavelet, is shown in Figure 34(b). The Pearson correlation to the seismic trace is 84%.

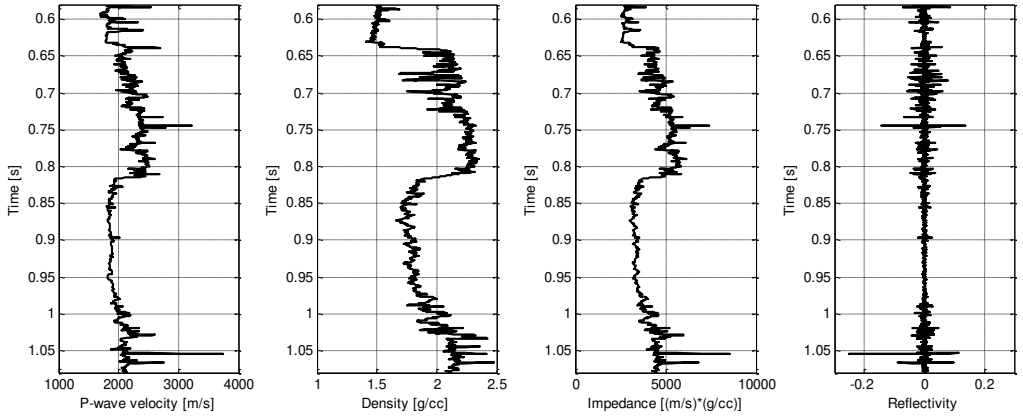


Figure 31: From left to right: Sonic log, density log, calculated impedance and calculated reflectivity. The sampling rate is 0.2ms.

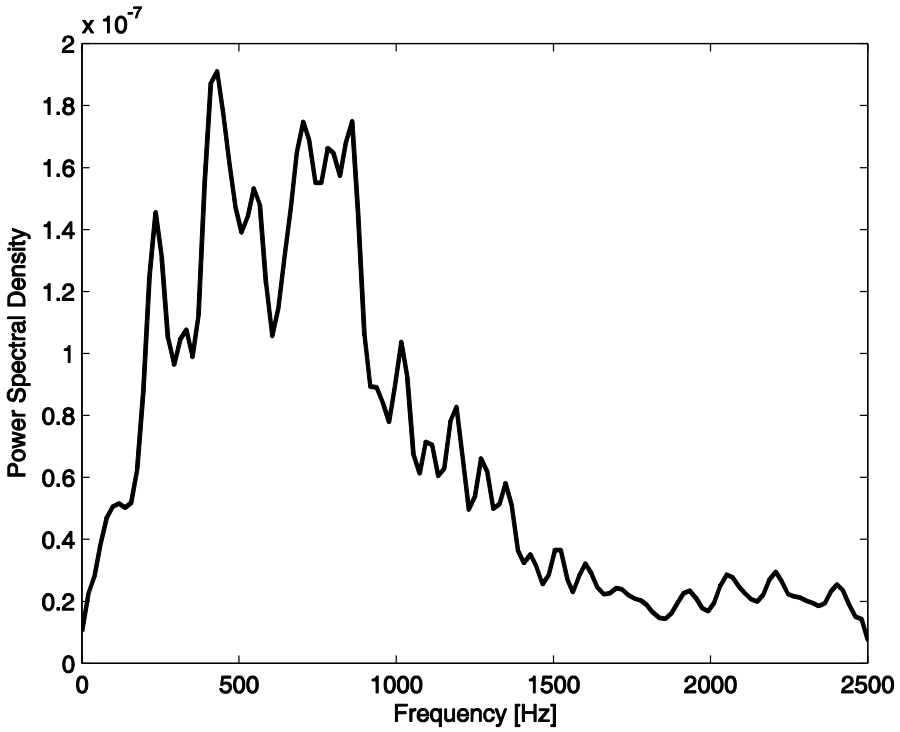


Figure 32: Power Spectral Density spectrum of the reflectivity sampled at 0.2ms.

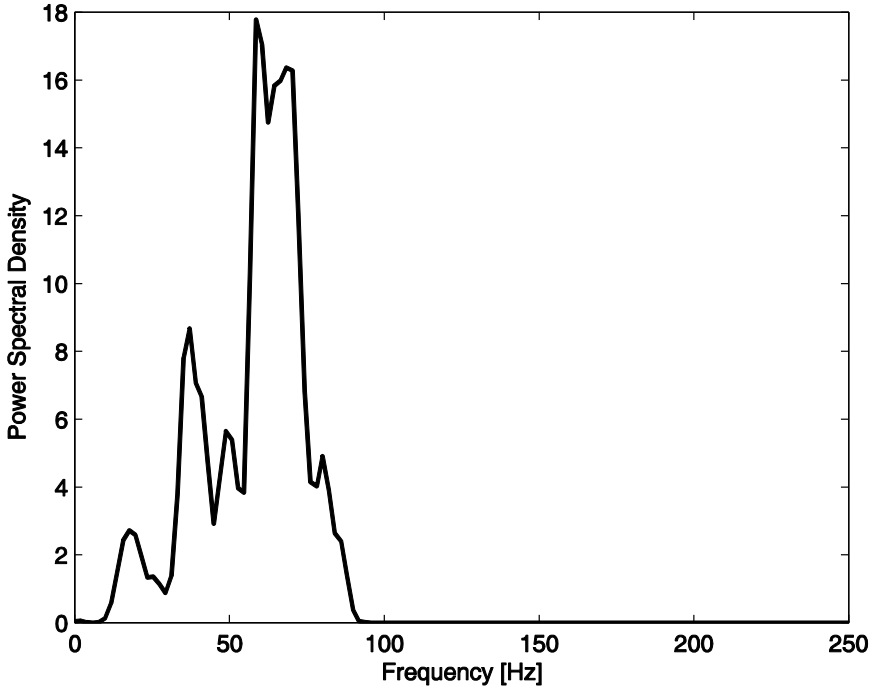


Figure 33: Power spectral density spectrum of the seismic trace.

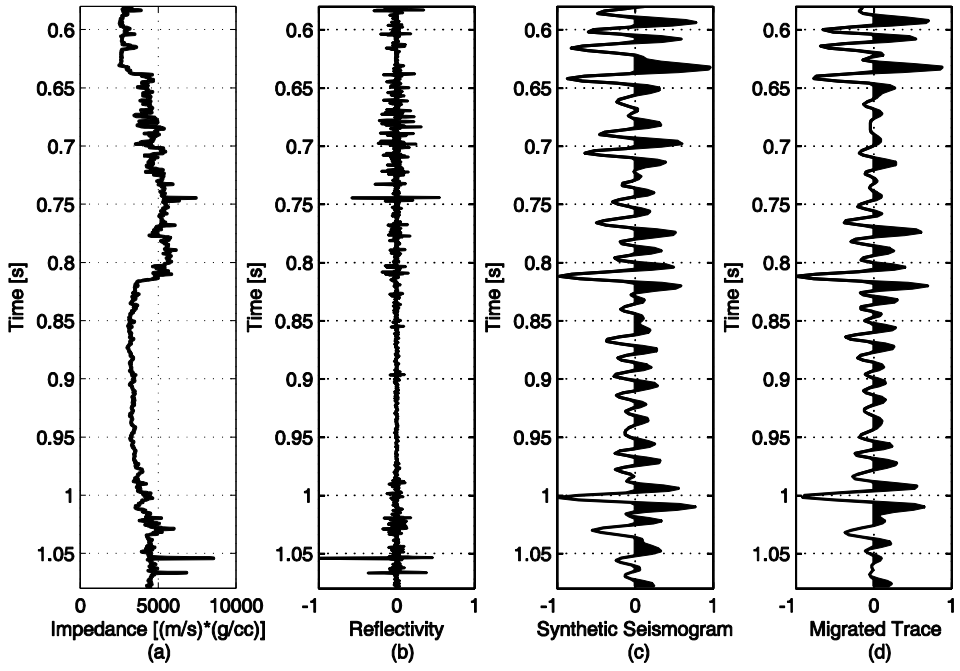


Figure 34: (a) and (b): Impedance and reflectivity logs at 0.2ms. (c) Synthetic seismogram. (d) Migrated trace.

In Figure 35, the histogram of the reflectivity samples, sampled at 0.2ms, are shown. We verify that it has a leptokurtic distribution [COMON and JUTTEN, 2010], as the estimated kurtosis (or fourth cumulant) is 70,0. We observed that the reflectivity pdf is well approximated by a generalized Gaussian distribution given by

$$p(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left[-(|x-\mu|/\alpha)^{\beta}\right],$$

where Γ, α, μ and β denote, respectively, the gamma function, the scale factor, the mean and the shape factor. Note that, $\beta = 2$ corresponds to the Gaussian distribution and $\beta = 1$ corresponds to the Laplacian distribution. As the kurtosis, γ is linked to β by

$$\gamma = \frac{\Gamma(5/\beta)\Gamma(1/\beta)}{\Gamma(3/\beta)^2} - 3,$$

the β estimated by solving numerically this equation is 0.37. The respective curve for the generalized Gaussian distribution is plotted in Figure 35.

In Figure 36, the autocorrelation function of the reflectivity sampled at 0.2ms is shown. The 95% confidence interval is indicated by the red lines for the hypothesis that the autocorrela-

tion is null for a given lag. This Figure shows that the samples are correlated, especially for small lags, although these lags are smaller than the sampling rate of the seismic data, which is 0.2ms.

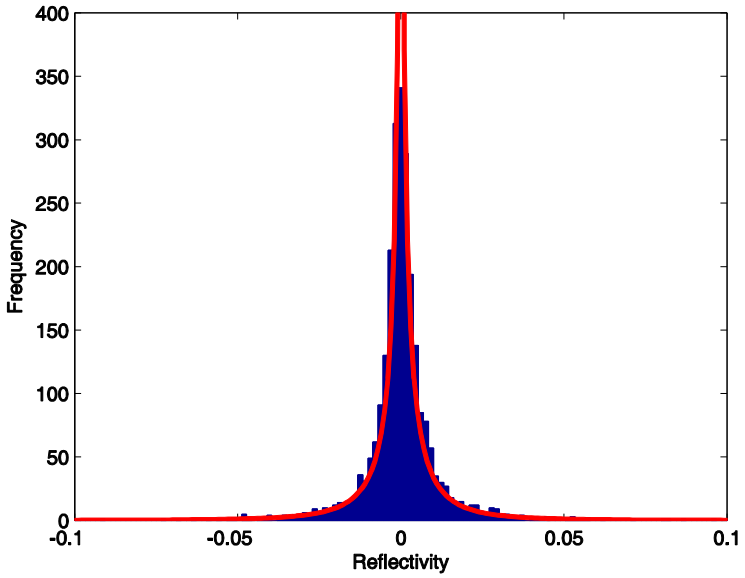


Figure 35: Histogram of the reflectivity samples at 0.2ms. In red, the generalized Gaussian distribution curve for corresponding estimated shape factor.

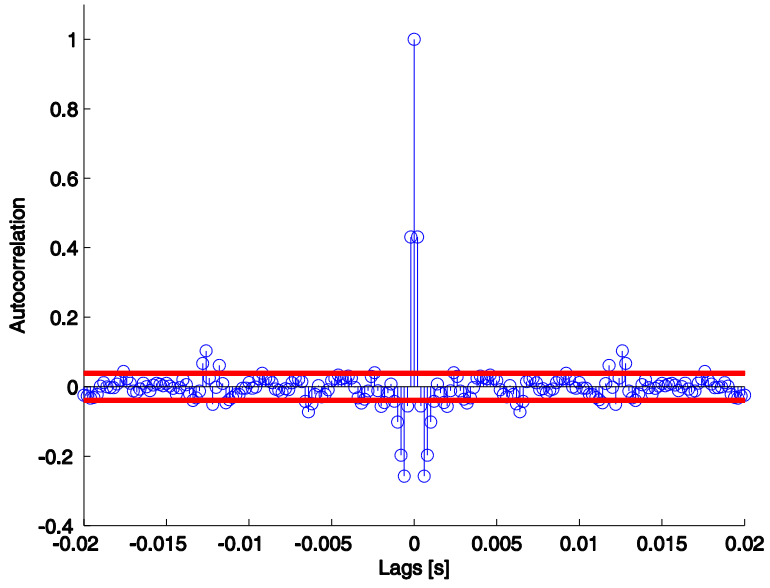


Figure 36: Autocorrelation function of the reflectivity sampled at 0.2ms. The red lines show the 95% confidence interval.

Next, a reflectivity log is calculated from sonic and density logs resampled at the same rate as the seismic data, 2ms, as shown in Figure 37. If we compare the resultant reflectivity in Figure 38(b) with the seismic data in Figure 38(d), we verify that it is visually closer to the seismic data, if compared to the finer sampling case in Figure 34(b), as the quantities are sampled in the same scale. The synthetic seismogram is shown in Figure 38(c), and it also has a correlation of 84% with the migrated trace. By visual inspection, if we compare the Figure 38(c) and Figure 38(d), we observe that features such as the negative peaks just after 0.8s and 1.0s are well reproduced. However, the spikes between 0.65s and 0.7s in the reflectivity graph are not present in the seismic log. By observing this region in Figure 30, we observe that this region has been somehow blurred in the migrated image, which may indicate that imaging was not completely successful in this part of the data.

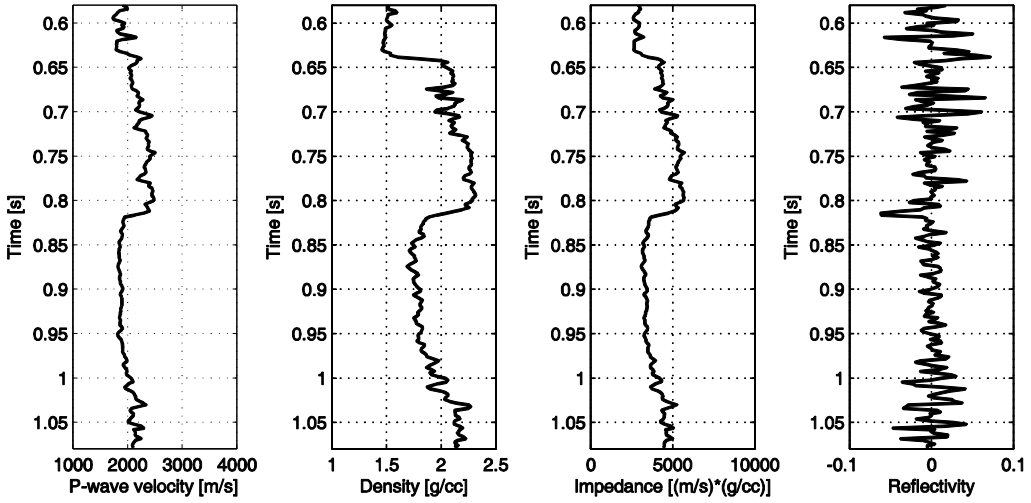


Figure 37: From left to right: Sonic log, density log, calculated impedance and calculated reflectivity. The sampling rate is 2ms.

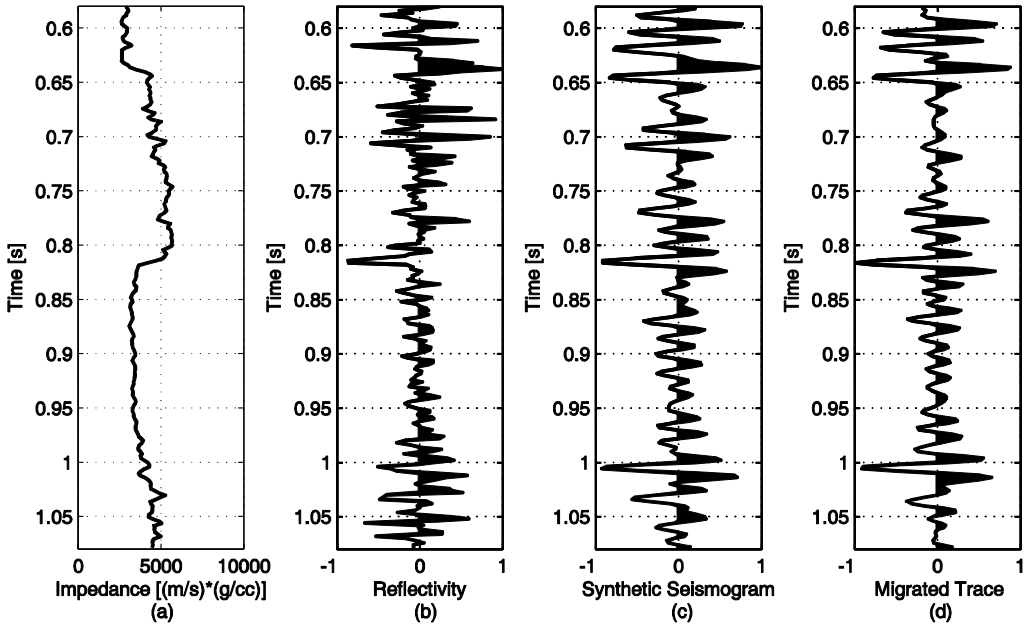


Figure 38: (a) Seismic impedance. (b) Seismic reflectivity. (c) Synthetic seismogram. (d) Migrated trace.

The histogram of the amplitudes of the reflectivity at sampling rate of 2ms is shown in Figure 39. The histogram is coarser, as in this case there are 10 times fewer samples than in Figure 35. In this case, the associated kurtosis is 4.93, and the corresponding β is 1.16, so this is also a leptokurtic distribution. This lower kurtosis can be associated to the fact that the reflectivity is smoother if compared to its version sampled at 0.2ms, which becomes clear if we consider the power spectral density spectrum of the coarser sampled version in Figure 40. It is verified that the high frequency content in Figure 32 is no longer present.

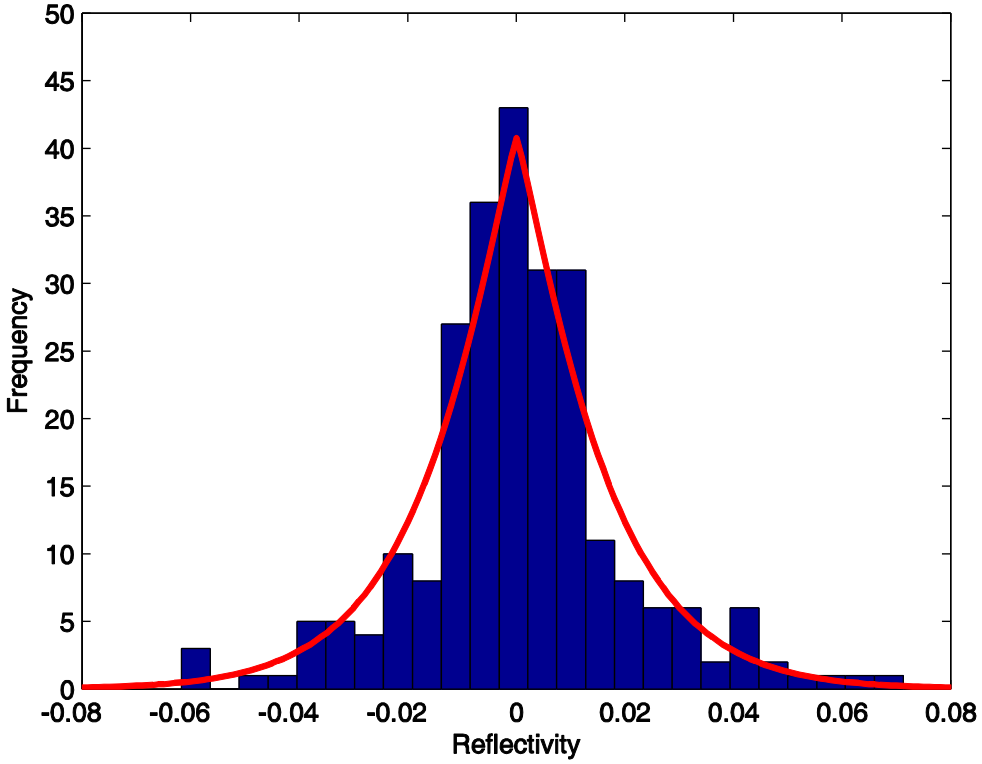


Figure 39: Histogram of the reflectivity samples at sampling rate of 2ms. In red the generalized Gaussian distribution curve for corresponding estimated shape factor.

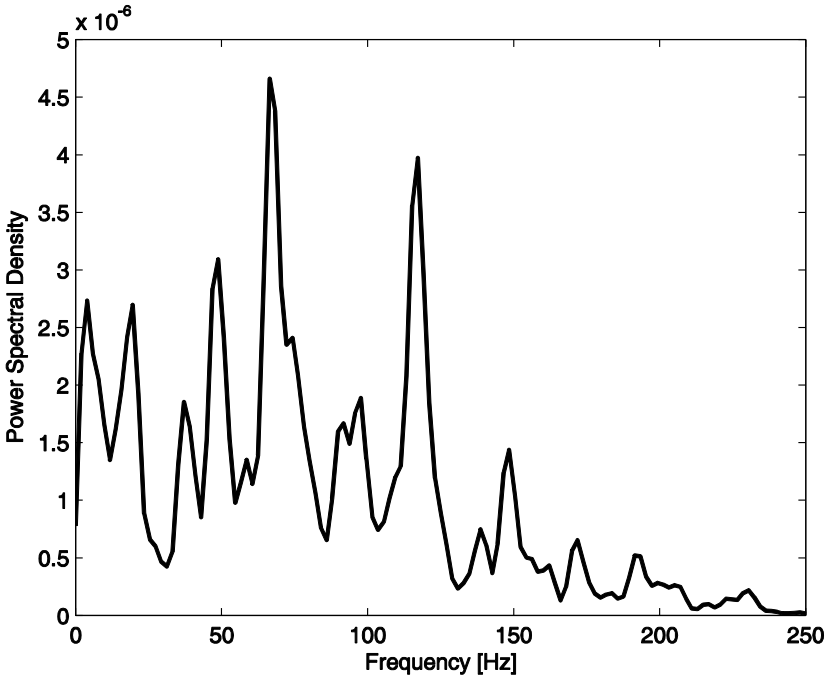


Figure 40: Power spectral density spectrum of the reflectivity sampled at 2ms

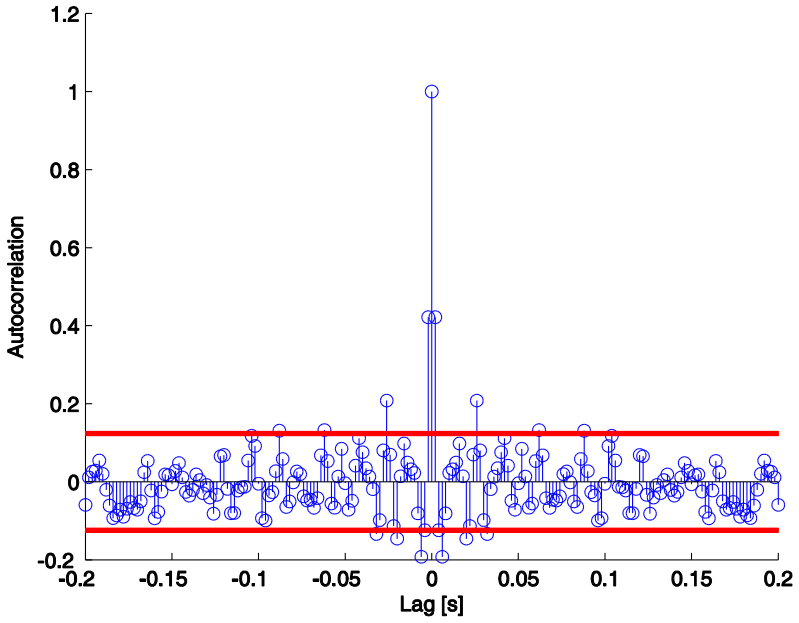


Figure 41: Autocorrelation function of the reflectivity sampled at 2ms. The red lines show the 95% confidence interval.

In Figure 41, the autocorrelation function is shown for the reflectivity sampled at 2ms. We observe that there is some correlation between the reflectivity samples, as some values of the autocorrelation function are outside of the 95% confidence interval. This is corroborated by the power spectral density spectrum in Figure 40, as it does not have a flat response as expected if the reflectivity were ideally white.

2.4.2 Unsupervised deconvolution in synthetic data

We compare now the results of the unsupervised deconvolution algorithms presented in Subsection 2.3.4, which are spiking deconvolution with the use of prediction error filtering, B-ICA and wavelet estimation with B-ICA followed by an LS inverse filter, as described by the additional Step 6.

Initially, a synthetic white reflectivity with i.i.d. random samples, chosen accordingly to a Bernoulli-Gaussin distribution, is considered. In this case, a sample has a probability given by p of being a non-zero variable. When this is the case, the sample has its amplitude given by a Gaussian distribution with zero mean, in our case, and variance equal to σ^2. In this case, we chose $p = 0.2$ and $\sigma^2 = 1$, and a synthetic trace thus obtained is shown in Figure 42. In the first test, we applied the unsupervised deconvolution algorithms to a synthetic trace produced with the use of the minimum phase Berlage wavelet [ALDRIDGE, 1990] shown in Figure 43. The amplitude spectrum associated with this wavelet is shown in Figure 44. The results of unsupervised deconvolution are shown in Figure 42 and the wavelet estimated with B-ICA is shown in Figure 43. It is possible to verify by inspection that the unsupervised methods are able to recover the original reflectivity from the seismic trace. However, in the B-ICA case, a delay is verified on the result. As pointed out by [KAPLAN and ULRYCH, 2003], this is associated with the nature of the B-ICA, which uses delayed versions of the seismic trace as its inputs. In fact, this delay is sometimes also observed in the B-ICA+LS approach, as the LS inverse filter allows smaller error if the estimated wavelet is shaped into a lagged spike instead of a zero-lag spike as in [YILMAZ, 2001]. Also, as shown in Figure 45, the candidate wavelets produced in Step 5 of B-ICA are likely to be shifted versions of the original wavelet, which is also associated with this observed delay. Thus, in the results of B-ICA+LS approach considered here, a slightly larger trace was decon-

volved with the LS approach and a displacement was applied in order to maximize the correlation with the original reflectivity. In Table 1, the Pearson correlation between the deconvolved traces and the reflectivity indicates that the unsupervised methods were able to recover the original reflectivity as the resultant Pearson correlation is above 90%. It is important to observe that the B-ICA results considered in this table had their delay compensated by correlation maximization prior to the measurement of their correlation.

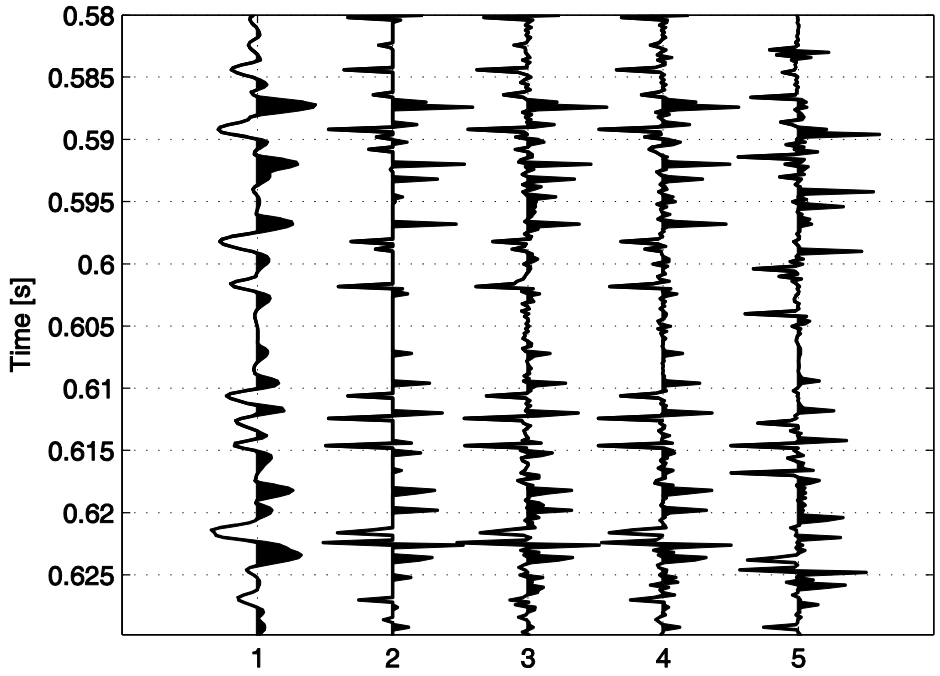


Figure 42: 1-Synthetic trace (minimum phase wavelet). 2- Synthetic Reflectivity. Deconvolved traces: 3-PEF. 4-B-ICA+LS inverse filter. 5-B-ICA.

Table 1: Pearson correlation between the deconvolved traces and the reflectivity.

	Synthetic reflectivity		Reflectivity from logs	
	Minimum phase	Mixed phase	Minimum phase	Mixed phase
Synthetic trace	46%	44%	64%	67%
PEF	95%	60%	90%	48%
B-ICA+LS	96%	93%	84%	66%
B-ICA	93%	81%	67%	58%

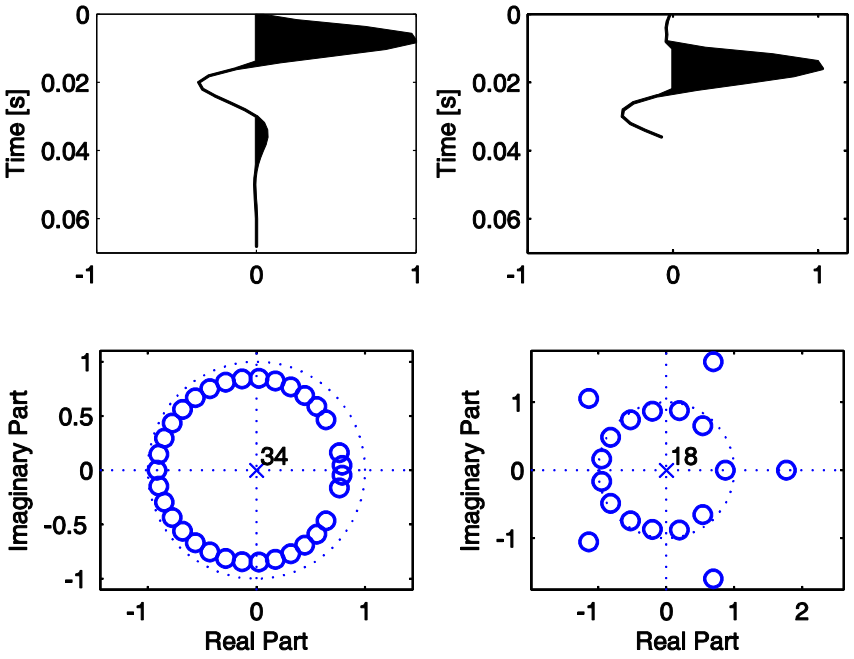


Figure 43: Left: Original minimum phase wavelet and respective zero-pole plane plot. Right: Estimated wavelet with B-ICA and respective zero-pole plane plot (synthetic reflectivity).

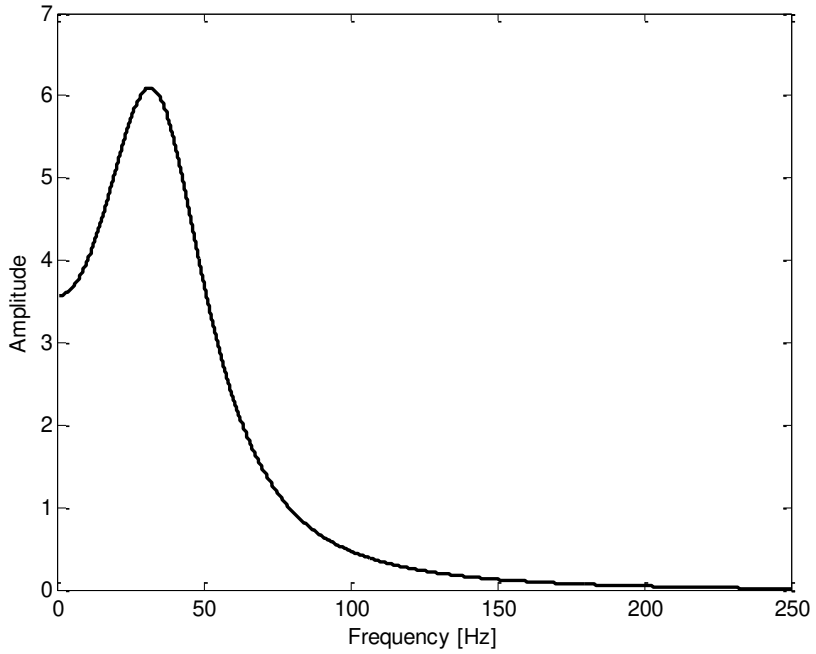


Figure 44: Amplitude spectrum of the minimum phase wavelet.

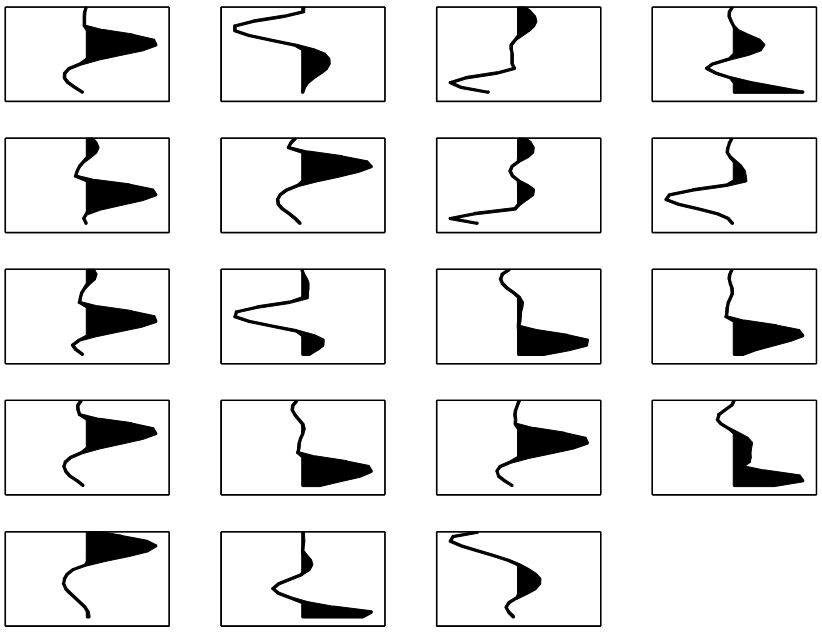


Figure 45: Candidate wavelets produced in Step 5 of B-ICA.

In the next test, we used the same random reflectivity, but we used the mixed phase wavelet depicted in Figure 46 and Figure 47. This wavelet was estimated from the reflectivity and the seismic trace shown in Figure 38 using the Wiener method described in Section 2.2. The results are shown in Figure 48, and it is possible to observe that the B-ICA methods were able to recover most of the features of the original reflectivity, while the result of prediction error filtering is completely distorted. This also is shown in Table 1, in which the B-ICA methods present high correlation if compared to the prediction error filtering. In this case, the B-ICA+LS approach produced a slightly better result than the plain B-ICA method. Also, in Figure 46, it is possible to observe that B-ICA was able to estimate the original mixed phase wavelet.

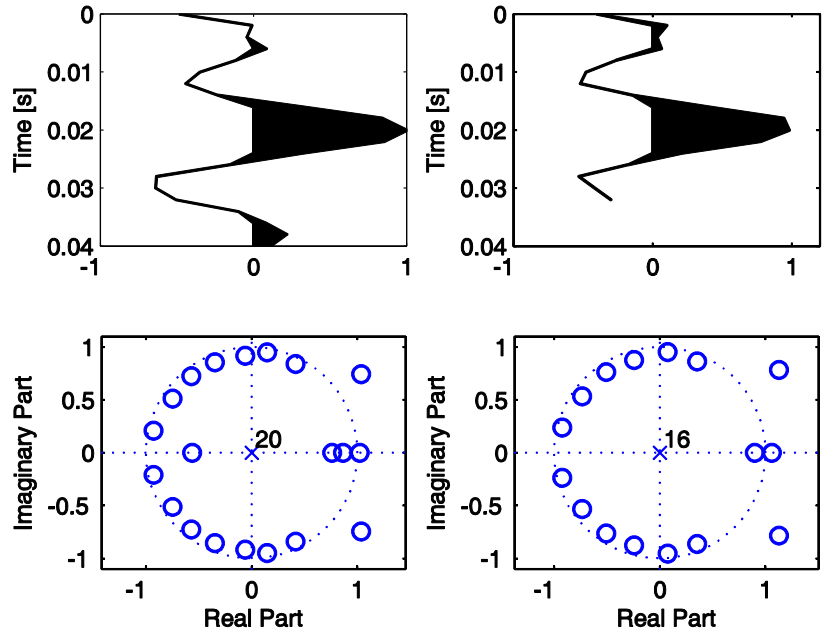


Figure 46: Left: Original mixed phase wavelet and respective zero-pole plane plot. Right: Estimated wavelet with B-ICA and respective zero-pole plane plot (synthetic reflectivity).

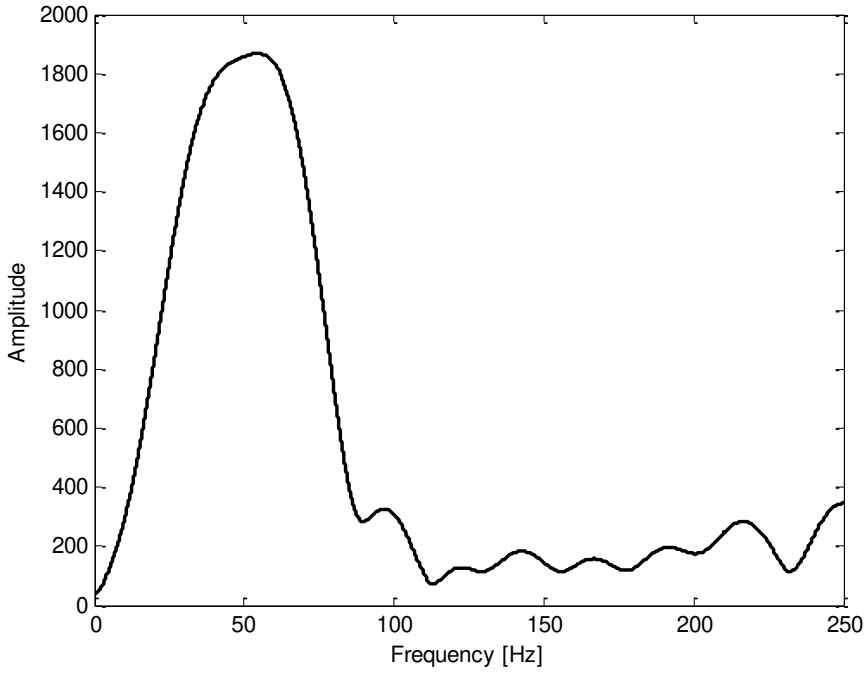


Figure 47: Amplitude spectrum of the mixed phase wavelet.

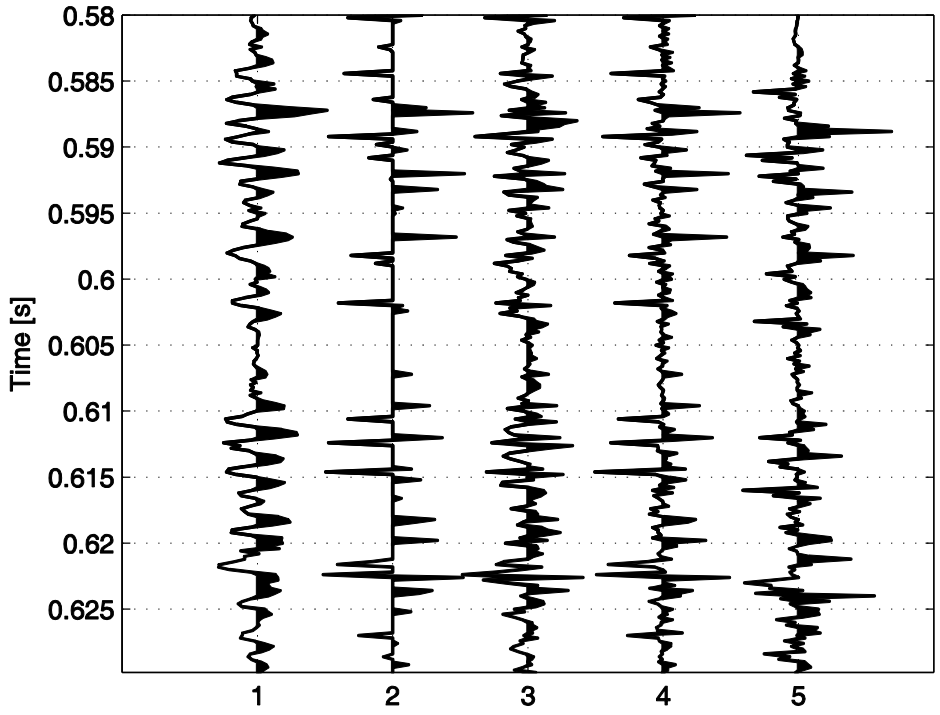


Figure 48: 1-Synthetic trace (mixed phase wavelet). 2- Synthetic reflectivity. Deconvolved traces: 3-PEF. 4-B-ICA+LS inverse filter. 5-B-ICA.

Following the tests with an ideally i.i.d. random reflectivity, we now performed tests using the reflectivity estimated from the well logs shown in Figure 38, so that the reflectivities are no longer ideally i.i.d. The results for the minimum phase case are shown in Figure 49. As also shown in Table 1, the plain B-ICA has a poor performance, but B-ICA+LS and the use of PEF are able to obtain reasonable results. However, in this case, the use of PEF has a slightly better result. The reason for that may be the fact that the use of PEF only assumes a white reflectivity, while the B-ICA method requires a stronger i.i.d. condition. Thus, this may provide a larger robustness to the use of PEF when the reflectivity is not completely white, as in this case. Nevertheless, as shown in Figure 50, the shape of the wavelet estimated by B-ICA is close to the original one, which explains the reasonable performance of the B-ICA+LS approach.

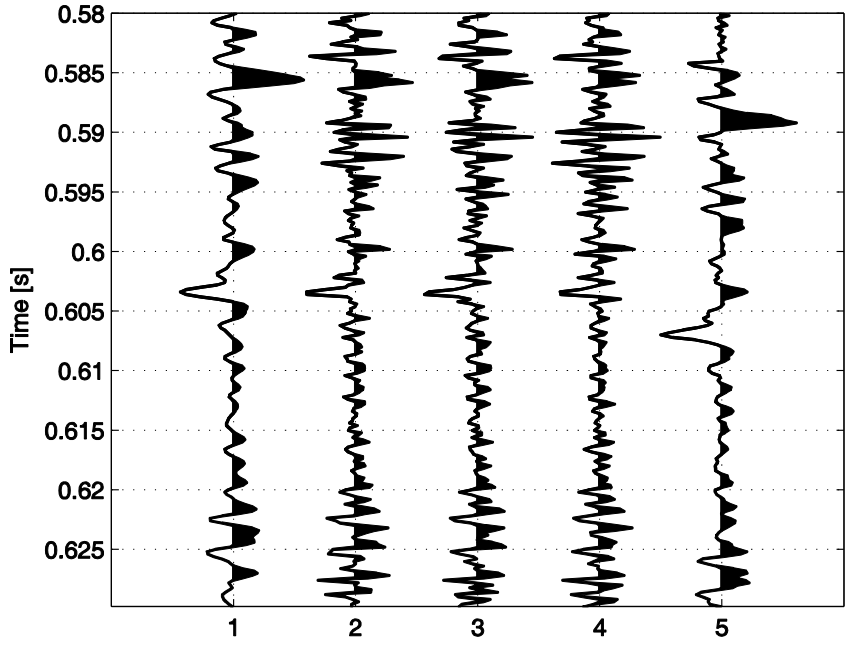


Figure 49: 1-Synthetic trace (minimum phase wavelet). 2- Reflectivity estimated from logs. Deconvolved traces: 3-PEF. 4-B-ICA+LS inverse filter. 5-B-ICA.

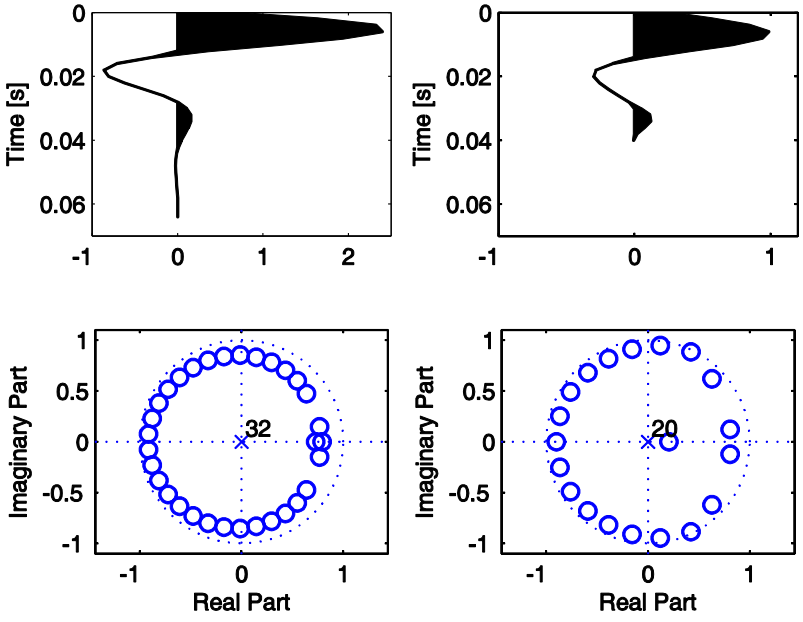


Figure 50: Left: original minimum phase wavelet and respective zero-pole plane plot. Right: Estimated wavelet with B-ICA and respective zero-pole plane plot (reflectivity estimated from logs).

As a final test, a synthetic trace is produced from the reflectivity estimated from the well logs and the mixed phase wavelet previously used. The results of the unsupervised deconvolution procedures are described in Figure 51. In Table 1, it is possible to verify that the deconvolution procedures were not able to improve the Pearson coherency of the outputs if compared to the input seismic trace. However, by observing Figure 51, it was verified that the result of B-ICA+LS provided an increase in resolution but it also introduced a high frequency noise. This noisy data was, then, filtered with a low pass filter (LPF) generated with the Parks-McClellan algorithm (e.g. [OPPENHEIM and SCHAFER, 1989]) and the result is shown in Figure 53 and in Table 2. The LPF has passband whose highest frequency is $f_p = 110Hz$ and stopband whose lowest frequency is $f_s = 120Hz$. Also, the estimated wavelet, shown in Figure 52 had a reasonably close shape to the original one. The use of LPF allowed an improvement in the Pearson's correlation coefficient of both PEF and B-ICA+LS cases, while keeping B-ICA+LS as the method that produced the output with the largest correlation to the original reflectivity.

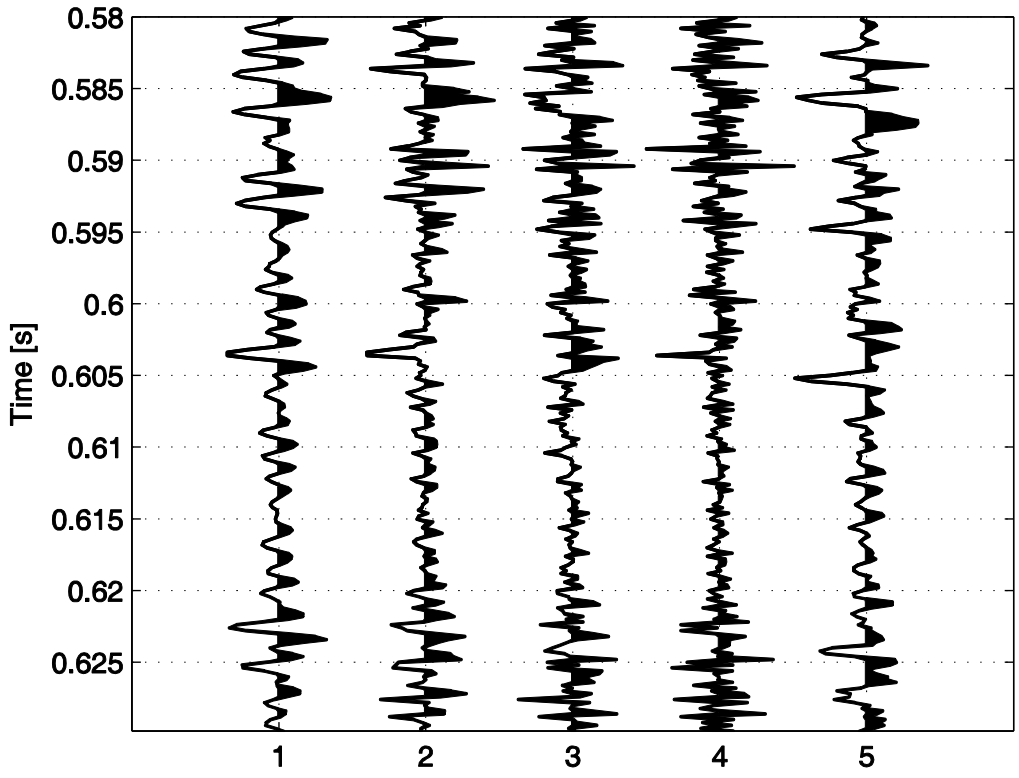


Figure 51: 1-Synthetic trace (mixed phase wavelet). 2- Reflectivity estimated from logs. Deconvolved traces: 3-PEF. 4-B-ICA+LS inverse filter. 5-B-ICA.

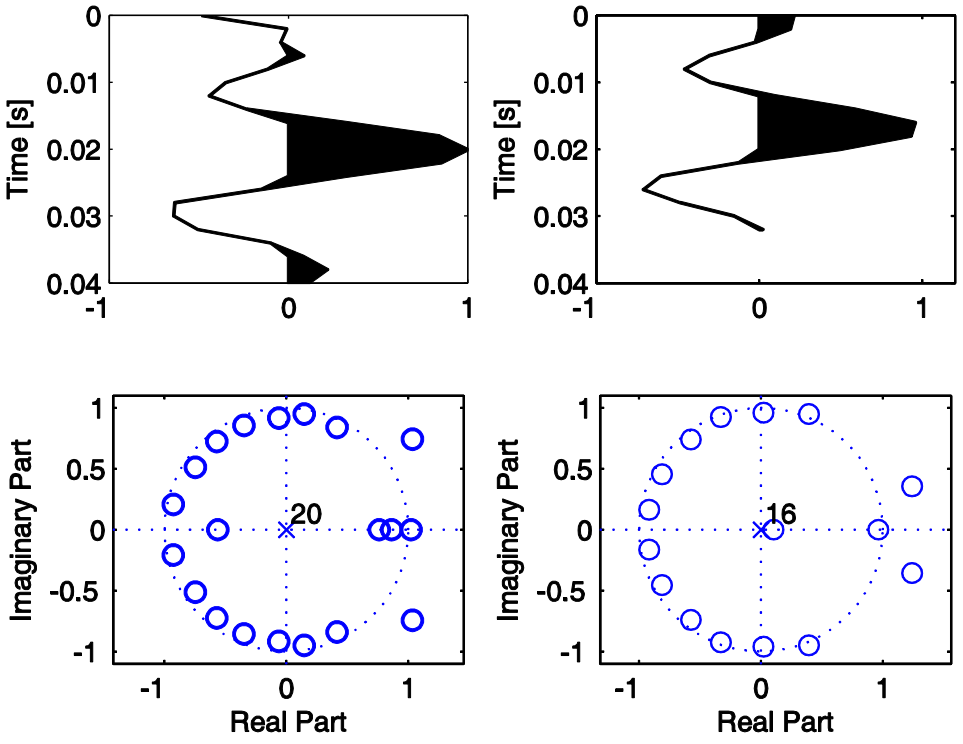


Figure 52: Left: original mixed phase wavelet and respective zero-pole plane plot. Right: Estimated wavelet with B-ICA and respective zero-pole plane plot (reflectivity estimated from logs).

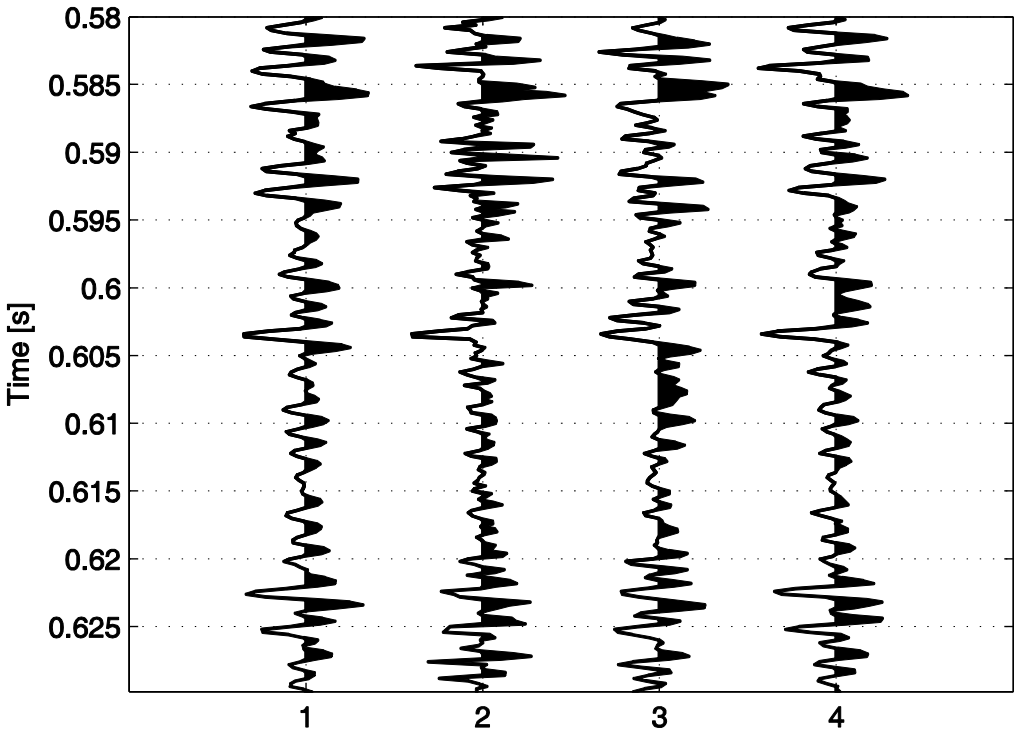


Figure 53: 1-Synthetic trace (mixed phase wavelet). 2- Reflectivity estimated from logs. Deconvolved traces: 3-PEF+LPF. 4-B-ICA+LS inverse filter+LPF.

Table 2: Pearson correlation between the deconvolved traces and the reflectivity comparing the use of LPF in the case where the synthetic trace is calculated from the mixed phase and the reflectivity estimated from the well logs.

	Without LPF	With LPF
PEF	48%	67%
B-ICA+LS	66%	76%

2.4.3 Unsupervised deconvolution in migrated section from field data

In the sequence of the tests, we applied the deconvolution algorithms to the trace of the migrated section corresponding to the well location shown in Figure 30. The results are shown in Figure 54 and in Table 3, where the Pearson's correlations of the outputs of the deconvolution algorithms and the reflectivity estimated from the well logs are displayed. In the tests with field data, we used a LPF whose passband has highest frequency of $f_p = 100Hz$ and stopband has lowest frequency of $f_s = 110Hz$, where the use of LPF is indicated. As a reference, a supervised LS deconvolution filter, as described in Figure 18 in subsection 2.2, was also calculated from the migrated trace and the reflectivity estimated from the well logs. As shown in Table 3, the supervised LS deconvolution filter produced the output with the largest correlation, which was expected, as the supervised method takes into account the estimated reflectivity information in order to obtain the filter, in contrast to the unsupervised methods, which only use the information obtained from the seismic trace. Despite this difference, if we compare the supervised and B-ICA + LS inverse filter + LPF outputs in Figure 54, we observe that both outputs present similar features, except for the use of PEF. Also, Table 3 shows that the PEF+LPF approach presents a low correlation, in fact lower than the original trace, which indicates that this approach introduces distortion instead of improvement. The main reason for this is that the wavelet in this case is non-minimum phase. Also, the regular B-ICA also presents a low correlation, while the B-ICA + LS inverse filter + LPF approach present a slight increase of correlation if compared to the original trace.

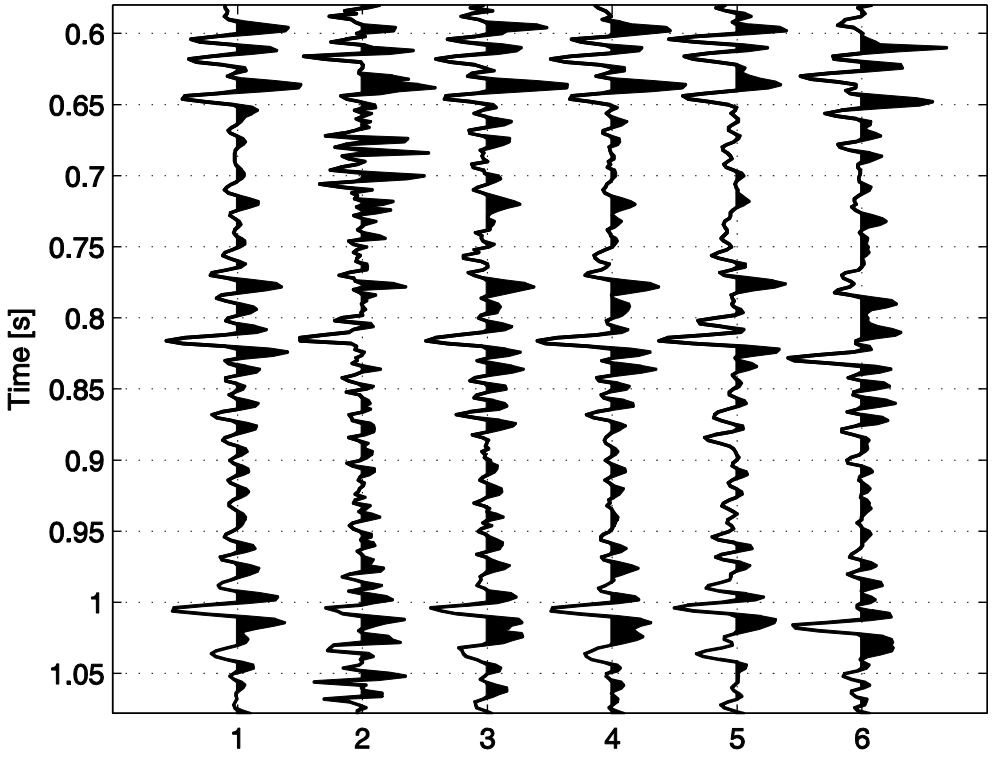


Figure 54: 1 – Migrated trace. 2 - Reflectivity estimated from well logs. Deconvolved traces: 3 – Supervised LS deconvolution. 4 – B-ICA+LS inverse filter+LPF. 5 - PEF+LPF. 6 – B-ICA.

Table 3: Pearson correlation between the deconvolved traces and the reflectivity when the input is the migrated trace shown in Figure 54.

	Minimum phase
Migrated trace	53%
Supervised LS	60%
B-ICA+LS+LPF	55%
PEF+LPF	47%
B-ICA	48%

Finally, the devonvolution algorithms were tested on a subset of traces of the migrated section of Figure 30 and the results are displayed in Figure 55. We chose this set of data, as we only had access to the result of this fully processed seismic section, which was ready for seismic interpretation. It is also important to notice that we did not have access to raw data, so we could not interfere with the processing of the migrated image. As in the single trace case, a supervised LS deconvolution filter was calculated from the reflectivity estimated from the well logs and the trace corresponding to the position of the well and the same filter was applied to all traces of the subset assuming that the wavelet does not vary significantly between the traces. A similar approach was used for performing deconvolution with B-ICA + LS inverse filter + LPF. The LS deconvolution filter was calculated using the migrated trace corresponding to the well position and the same filter was applied to all traces of the subset. In the PEF + LPF case, a deconvolution filter was calculated for each trace of the subset and thus each trace was deconvolved with its own filter. It is possible to observe that the PEF+LS approach distorted the image by, e.g., destroying the lateral continuity of some reflectors pointed by the yellow arrow in Figure 55(d) and in Figure 56(b). The B-ICA + LS inverse filter + LPF enhanced some regions of the image as in the one indicated by the yellow and green arrows in Figure 55(c), where some reflectors that were hidden or blurred in the original section become visible, as displayed in the zoomed versions in Figure 57 and Figure 58. It was also noticed that lateral continuity was not harmed by the use of the method as in the PEF+LS approach. By comparing the output of the unsupervised method to the supervised LS method, we observe that the images are comparable, but with some differences, such as the reflector pointed by the yellow arrow in Figure 55(b) and Figure 59(a), which is absent in the result of the unsupervised method.

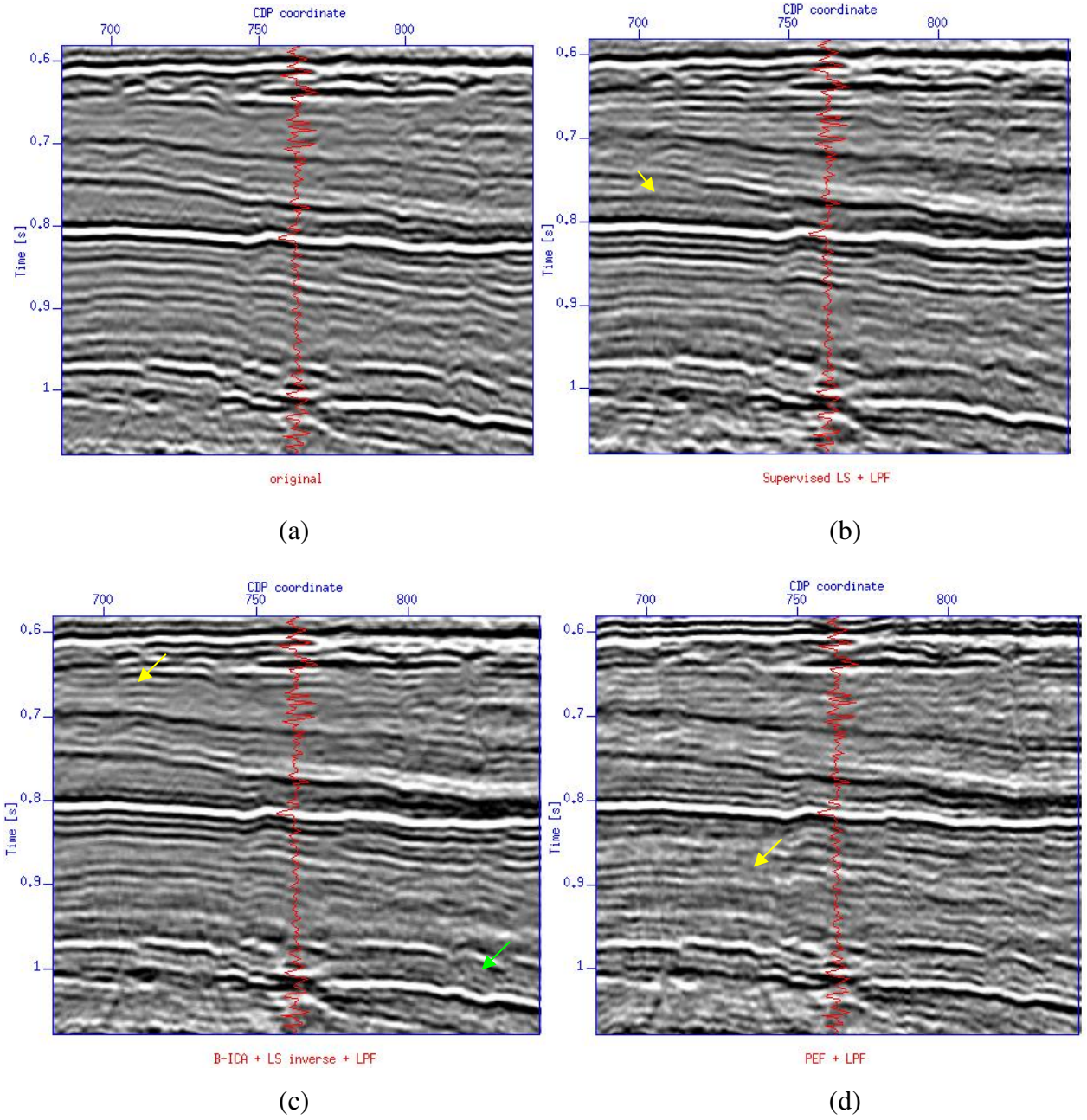
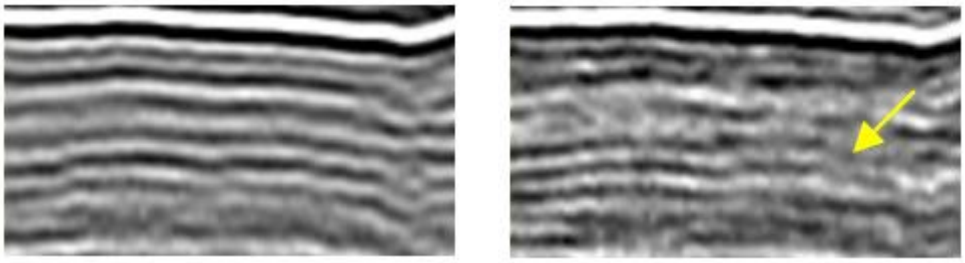
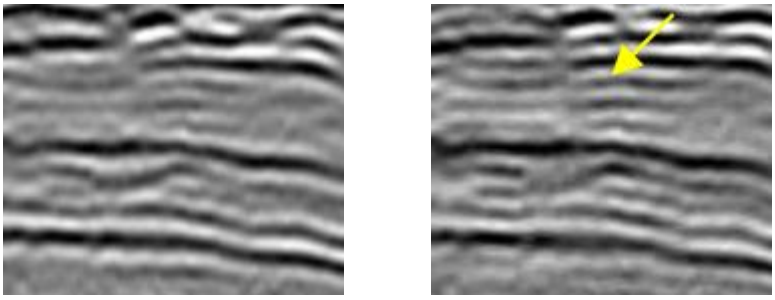


Figure 55: (a) Original traces. (b) Output of supervised LS deconvolution. The deconvolution filter was calculated using the reflectivity estimate from the well log and the trace corresponding to its position and was applied to all traces. (c) Deconvolution using B-ICA + LS inverse filter + LPF approach. The deconvolution filter was calculated using the trace corresponding to the well position and was applied to all traces. (d) Deconvolution using PEF + LPF. The deconvolution filter was calculated separately for each trace and then was applied to the respective trace.



(a) (b)

Figure 56: Zoomed versions of Figure 55(a) and Figure 55(d). (a) Shows the original traces, while (b) shows the result of deconvolution using PEF+LPF. The lateral continuity of some reflectors is lost, as shown in the region pointed by the yellow arrow in (b).



(a) (b)

Figure 57: Zoomed versions of Figure 55(a) and Figure 55(c). (a) shows the original traces, while (b) shows the result of deconvolution using B-ICA+LS inverse filter+LPF approach. Some reflectors that were hidden or weak in (a) where enhanced in (b) as pointed by the yellow arrow.

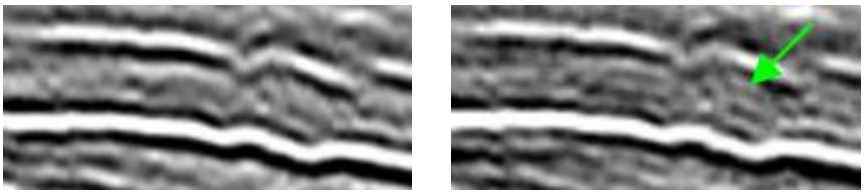
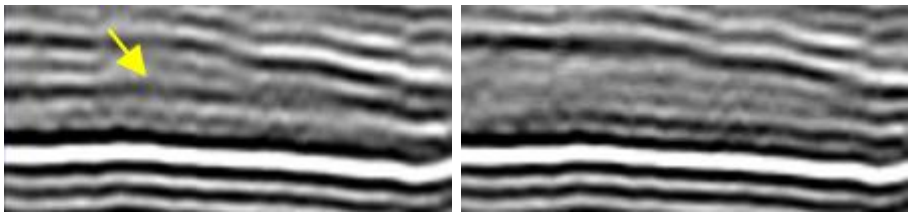


Figure 58: Zoomed versions of Figure 55(a) and Figure 55(c). (a) shows the original traces, while (b) shows the result of deconvolution using B-ICA+LS inverse filter+LPF approach. Some reflectors that were hidden or weak in (a) where enhanced in (b) as pointed by the green arrow.



(a) (b)

Figure 59: Zoomed versions of Figure 55(b) and Figure 55(c). (a) shows the result of deconvolution using the supervised LS approach, while (b) shows the result of deconvolution using B-ICA+LS inverse filter+LPF. The reflector pointed by the yellow arrow in (a) is not shown in (b).

2.5 CONCLUSION

We started this chapter by analyzing the basics of supervised and unsupervised deconvolution and wavelet estimation tehcniques in the unidimensional context. Supervised techniques using LS or Wiener filtering were initially reviewed and were shown to be useful for benchmarking the unsupervised techniques discussed in this work, namely the use of PEF for spiking deconvolution and the B-ICA and B-ICA+LS inverse filtering approaches, as described in later parts of the conclusion. Next, we reviewed the main theoretical aspects of unsupervised deconvolution, in special the use of PEF for spiking deconvolution [ROBINSON, 1954], which only works for minimum phase, and white reflectivity and the theoretical framework presented in [DONOHO, 1981], which explains methods such as MED [WIGGINS, 1978], that use HOS in order enable unsupervised deconvolution for arbitrary phase wavelets under the hypothesis that the samples of the reflectivity can be assumed to be a realization of a i.i.d. non-Gaussian process.

As in [TAKAHATA *et al.*, 2012], we focused on the unsupervised technique proposed in [KAPLAN and ULRYCH, 2003], which uses a variant of ICA, a technique used in BSS. This variant, called B-ICA, allows one to perform seismic deconvolution and wavelet estimation in a way that connects unsupervised deconvolution and ICA. As in the classical deconvolution algorithms for non-minimum phase wavelets, this technique also allows the deconvolution of arbitrary phase wavelets and requires an i.i.d. non-Gaussian reflectivity, as these are also hypotheses that underlie ICA techniques. Our contribution to the use of B-ICA, presented in [TAKAHATA *et al.*, 2012], was to add an additional step in which we use the wavelet estimated by B-ICA to

calculate an LS inverse deconvolution filter instead of using directly the deconvolved trace estimated by the B-ICA method.

Following this analysis, we extended the result in [TAKAHATA *et al.*, 2012] by performing a case study on the method using migrated data and well logs. We started with the characterization of a reflectivity function using logs from measurements made in a borehole. We performed a statistical analysis considering two sampling rates: 0.2ms, representing a closer rate to the log acquisition in the borehole, and 2ms, representing the sampling rate of seismic data. Both analyses showed that the distribution of amplitudes was non-Gaussian and obeyed a leptokurtic distribution. It was also observed that the autocorrelation had significant values especially for small lags, although with small values if compared to the value at zero lag. This indicated that the reflectivity was not an ideal white signal and thus also not i.i.d. Therefore, both the whiteness requirement of the PEF-based method and the i.i.d. requirement of the B-ICA method were not met, at least not perfectly, in this situation. Next, a segment of the trace of the migrated section corresponding to the location of the well was used to estimate the wavelet with a supervised method, in which the reflectivity from the well log was used as a training signal. It was verified that the wavelet was mixed phase in the chosen segment, as opposed to the minimum phase assumption for deconvolution using PEF.

This information was used to create synthetic traces with different characteristics with respect to the statistics of the reflectivity and the phase of the wavelet, where some traces met the hypothesis of the unsupervised deconvolution algorithms and others corresponded to more realistic cases. The tested algorithms were the unit lag PEF, classically used for seismic deconvolution, deconvolution with B-ICA, as proposed in [KAPLAN and ULRYCH, 2003], and the B-ICA for wavelet estimation followed by the calculation of an inverse LS filter as proposed in [TAKAHATA *et al.*, 2012]. The initial tests showed that the B-ICA-based methods are effective for both mixed and minimum phase wavelets, if the i.i.d. requirement is respected, as opposed to the use of PEF, which failed in the mixed phase case. The remaining tests were conducted with the non-ideally i.i.d. reflectivity, which was estimated from the well logs. The results showed that the use of B-ICA for reflectivity estimation was not effective in this case, suggesting that the method is sensitive to non-idealities in the i.i.d. hypothesis. On the other hand, the B-ICA+LS inverse filter was effective in the minimum and mixed phase cases (after low pass filtering in the latter), suggesting that the use of B-ICA for wavelet estimation has some robustness even if the

reflectivity is not perfectly white. A certain degree of robustness was also observed in the use of PEF in the minimum phase wavelet case. The reason for this may stem from the fact that the use of PEF requires only whiteness, which is a weaker requirement than independence. As a final observation on the deconvolution of the synthetic traces, as delayed versions of the trace are used as inputs of the B-ICA method, a delay was also noticed when using the methods based on the B-ICA. This is inherent to the method, as noted by [KAPLAN and ULRYCH, 2003]. In this work, the reported correlations were calculated after compensating the delay, which was done by finding the delay that maximized the correlation between the original and the deconvolved traces.

In the sequence, we tested the deconvolution algorithms on a migrated trace corresponding to the well location. In this test, a supervised LS inverse filter was also calculated in order to provide a reference. Similarly to the case with the synthetic trace produced from the convolution of the reflectivity estimated from the well logs and the mixed phase wavelet, the B-ICA+LS+LPF produced the best result and qualitatively was able to reproduce almost all the features obtained with the use of the supervised LS filter, which reinforces the fact that this approach has some robustness to the fact the reflectivity is not perfectly i.i.d., which is not shared with the deconvolution method using pure B-ICA. Also, ICA+LS+LPF approach outperformed the use of PEF, as the hypotheses of the method do not consider the phase of the wavelet, in opposition to the use of PEF.

We also applied the deconvolution methods to a sub-region of a migrated section neighboring the well position. In these tests, we assumed that the wavelet did not vary much from trace to trace. Firstly, we observed that the use of PEF was not adequate as it destroyed horizontal continuity observed in some reflectors of the original section. The main reason for this is that the wavelet is non-minimum phase as processed data was used. This happens because, even if the original wavelet is minimum phase, regular seismic processing procedures that were used in this dataset, such as band-pass filtering, stacking and migration are usually not designed to keep this character of the wavelet phase. We also observed that the supervised LS filter and B-ICA + LS inverse approaches could enhance some parts of the image, making some reflectors more visible.

From the results in synthetic and migrated traces from field data, we confirmed that the use of unit lag PEF for deconvolution is effective only if the wavelet is minimum phase and otherwise it may harm the result. Moreover, the sole use of B-ICA for deconvolution, as in [KAPLAN and ULRYCH, 2003], only succeeds if the reflectivity is i.i.d., independently of the

phase of the wavelet. However, B-ICA seems not to be robust for deconvolution if the reflectivity is not ideally i.i.d. On the other hand, it seems that B-ICA has a certain degree of robustness in wavelet estimation, as its wavelet estimates were successfully used for calculating inverse LS filters for deconvolution as in [TAKAHATA *et al.*, 2012].

3 BIDIMENSIONAL DECONVOLUTION

The deconvolution techniques described so far can be called unidimensional techniques, as only vertical information (either time or depth) is used to calculate the deconvolution filter or the seismic wavelet, which actually can be seen as a way to quantify vertical linear distortion. In these techniques, no information about how the data of neighboring traces relate to each other, i.e., no horizontal information is used. In opposition to the unidimensional approach, bidimensional deconvolution uses a bidimensional linear distortion model and aims to enhance both vertical and horizontal resolutions. In our case, we use the concept of resolution function adopted in works such as [LECOMTE and GELIUS, 1998; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002] in order to quantify the distortion in both vertical and horizontal directions of prestack depth migrated (PSDM) images. Besides the frequency information used in the 1D case, limitations in acquisition geometry and complexities in the geology are also considered. As advocated in [GELIUS and LECOMTE, 2000], under proper assumptions, the resolution function describes the so called point spread function (PSF). The concept of PSFs has been used widely in the image processing community in order to quantify the amount of image degradation caused by blurring (for a complete review, see, for example, [BANHAM and KATSAGGELOS, 1997]). In [LECOMTE, 2008], an interesting example on the deblurring of images produced by the Hubble Space Telescope is described. In this case, stars, which should appear as dots in the resulting image, were blurred by the intrinsic 2D impulse response of the telescope mirror, i.e., the PSF. The major problem was that imperfections in the mirror resulted in space-variant PSFs, and thus standard algorithms were unable to properly sharpen the resulting image. The problem was solved by measuring the PSFs all over the mirror surface with an optical method, and then using a space-variant deconvolution method.

Within seismic processing, a PSDM image can be modelled as the convolution between the PSF and the actual model quantities [GELIUS *et al.*, 2002]. The key point is that these PSFs can be calculated with relatively low computational effort with the use of ray tracing. This feature has been used to simulate efficiently 2D and 3D PSDM images with the method called simulated prestack local imaging (SimPLI) described, for example in [LECOMTE and POCHON-GUERIN, 2005; LECOMTE, 2008]. The concept of PSF can also be used to enhance the resolution of PSDM images through 2D deconvolution methods. This approach has been explored in

works such as [GELIUS *et al.*, 2002; SJOEBERG *et al.*, 2003]. However, these works only consider the use of a smaller subsection along the vertical direction of each PSF, which is close to the 1D approach, where only the central vertical line of the PSF is used. This imposes limits on enhancing horizontal resolution, and can be seen as a pseudo 2D approach, in fact. Other approaches for PSDM sections enhancement based on techniques such as least-squares migration [NEMETH *et al.*, 1999] and migration deconvolution [HU *et al.*, 2001] form another branch of research. These methods explicitly take into consideration the modelling and the migration operator, and require a larger computational effort or simplifying schemes, such as considering only the main diagonal of the so called Hessian matrix. The interested reader may find an interesting discussion, for example, in [TANG, 2009].

In this chapter, we show in detail and extend the results obtained in [TAKAHATA *et al.*, 2013] where we proposed the method and showed that it is able to increase the resolution of seismic images with the use of the entire PSF instead of only using the central vertical part of the PSF as in previous works [GELIUS *et al.*, 2002; SJOEBERG *et al.*, 2003]. In Subsections 3.1, 3.2 and 3.3 we review key topics on PSDM and the use of 2D deconvolution based in works such as [GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002; LECOMTE, 2008]. In Subsection 3.4, we introduce our main contribution, as proposed in [TAKAHATA *et al.*, 2013], which is the use of a 2D filtering approach based on 2D spiking deconvolution in the context of enhancement of resolution of PSDM images with the use of PSFs. This contrasts with other works in literature, such as [SJOEBERG *et al.*, 2003], in which an inversion approach was used. In this case, we admit that the PSF is somehow estimated and we calculate the deconvolution filter in a supervised manner. It is important to notice that our contribution is not on the concept of spiking deconvolution, which, as shown in the previous chapter, is an already well-established concept, but rather, on its application in the present context. Finally, in Subsection 3.5 we present numerical results obtained on controlled and field data, and in Subsection 3.6 we present our conclusions and we discuss the potentials and the limitations of the method.

3.1 THE 2D CONVOLUTIVE MODEL: THE RESOLUTION FUNCTION AND THE POINT SPREAD FUNCTION

Let $s(\mathbf{r})$ be a quantity that characterizes a geological model, such as the reflectivity, at a given point $\mathbf{r}$. As described in [GELIUS *et al.*, 2002], its resolution is controlled by the resolution function which can be derived with the aid of a Fourier vector, called scattering wavenumber. This vector can be calculated with the use of ray tracing with relative ease, as we will describe with more detail later in this chapter. In the aforementioned work, both a smooth acoustic or scalar velocity model of the subsurface (as shown in Figure 60) and local reaction are assumed. This means that the scattering or reflection at a given model point $\mathbf{r}$ is only caused by interactions within a surrounding small region, being negligible the interactions with other parts of the model. A local plane-wave contribution is also assumed (far field assumption), as pointed out in [GELIUS and LECOMTE, 2000]. Given these assumptions, the spatial Fourier transform of $s(\mathbf{r})$ is defined as

$$S(\mathbf{K}) = \int_{\Omega} s(\mathbf{r}')\exp(-j\mathbf{K}\cdot\mathbf{r}')d\mathbf{r}', \tag{65}$$

Here, $\mathbf{K}$ denotes the Fourier vector, or the scattering wavenumber vector, as mentioned before, at the center point $\mathbf{r}$ and $\mathbf{r}'$ is a position vector that denotes the points within a small region Ω around $\mathbf{r}$ [GELIUS *et al.*, 2002]. As shown in Figure 60, and following [GELIUS *et al.*, 2002], this vector can be linked to the seismic survey geometry by the relationship

$$\mathbf{K} = -\omega\nabla\tau(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}) = -\omega\left[\nabla\tau_s(\mathbf{r}, \mathbf{r}_s) + \nabla\tau_g(\mathbf{r}_g, \mathbf{r})\right] = -\mathbf{k}_s(\mathbf{r}, \mathbf{r}_s) + \mathbf{k}_g(\mathbf{r}_g, \mathbf{r}). \tag{66}$$

In the above equation, $\mathbf{r}_g$ and $\mathbf{r}_s$ denote, respectively, the positions of the receiver and source, ω is the frequency of the source signature, $\tau(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$ is the total traveltime and $\tau_s(\mathbf{r}, \mathbf{r}_s)$ and $\tau_g(\mathbf{r}_g, \mathbf{r})$ are, respectively, the traveltimes along the rays from the source at $\mathbf{r}_s$ to the model point at $\mathbf{r}$ and from the model point to the receiver at $\mathbf{r}_g$. In addition to that, the wavenumber vectors $\mathbf{k}_s$ and $\mathbf{k}_g$ represent the local directions of the Green's functions associated to the incident ray that is generated in $\mathbf{r}_g$ and the scattered ray that is received at $\mathbf{r}_s$, respectively. These Green's functions must be calculated in the background model. In case of complete coverage in the Fourier space, it follows from (65) that $s(\mathbf{r})$ can be obtained by the inverse Fourier transform of $S(\mathbf{K})$, i.e.,

$$s(\mathbf{r}) = \int S(\mathbf{K})\exp(j\mathbf{K}\cdot\mathbf{r})\mathrm{d}\mathbf{K}$$

Unfortunately, as can be inferred from Figure 60, the range of directions of $\mathbf{K}$ is constrained by the limited number of source and receiver positions at the surface, i.e., a 360^{o} coverage is, in practice, not possible. Moreover, from (66), we see that the bandwidth limitations imposed by the source signature, in combination with attenuation effects, impose a constraint on the length of $\mathbf{K}$, which is proportional to ω. If we describe these actual band limitations in frequency and direction by $H(\mathbf{K})$, the actual estimated model parameter, $x(\mathbf{r})$, can be expressed as

$$x(\mathbf{r}) = \int H(\mathbf{K})S(\mathbf{K})\exp(j\mathbf{K}\cdot\mathbf{r})\,\mathrm{d}\mathbf{K}. \tag{67}$$

Substituting (65) in (67), we have:

$$x(\mathbf{r}) = \int_{\Omega} s(\mathbf{r}')\left[\int H(\mathbf{K})\exp(j\mathbf{K}\cdot(\mathbf{r}-\mathbf{r}'))\,\mathrm{d}\mathbf{K}\right]d\mathbf{r}' = \int_{\Omega} s(\mathbf{r}')h(\mathbf{r}-\mathbf{r}')d\mathbf{r}', \tag{68}$$

where $h(\mathbf{r})$ is defined by the inner integral (between brackets) in the above equation. In fact, $h(\mathbf{r})$ is the resolution function, which describes the distortions of the estimated model. Moreover, inspection of (68) shows that the relationship between $s(\mathbf{r})$ and $x(\mathbf{r})$ is described by a 2D convolution between the actual model quantity, $s(\mathbf{r})$ and $x(\mathbf{r})$. As $h(\mathbf{r})$ can be seen as the impulse response of a linear system at the point $\mathbf{r}$, i.e., a measure of how a point scatterer blurs or spreads, it is also called Point Spread Function (PSF). This applies for Born scattering [MILLER *et al.*, 1987; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002], in which it is assumed that every discontinuity of the velocity model can be approximated by a point scatterer.

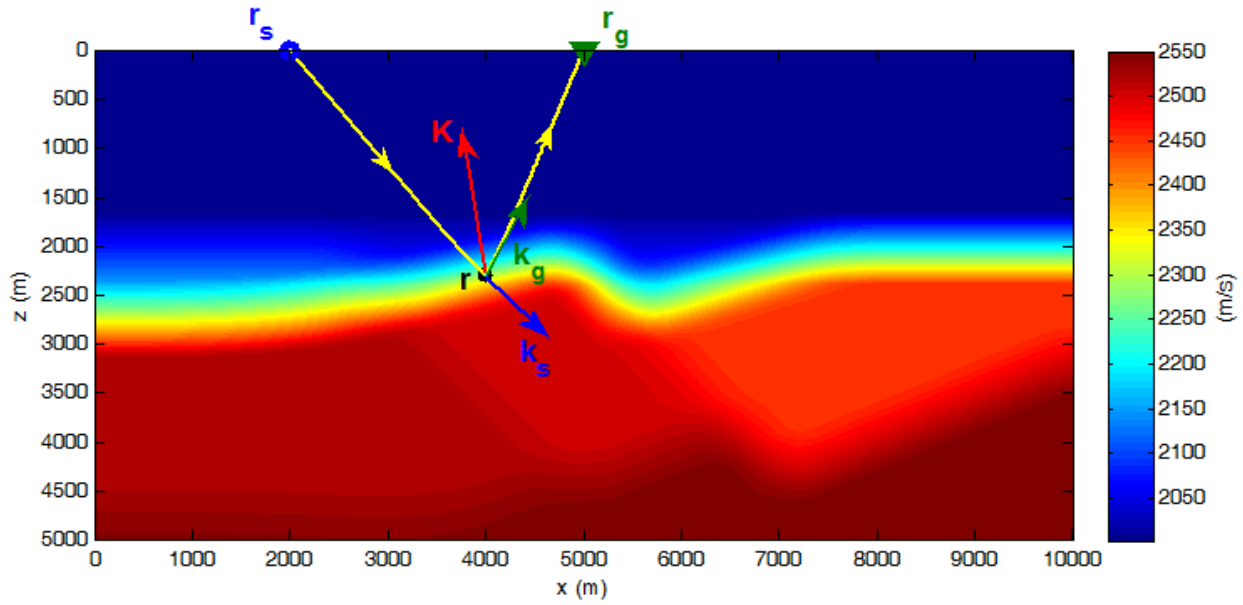


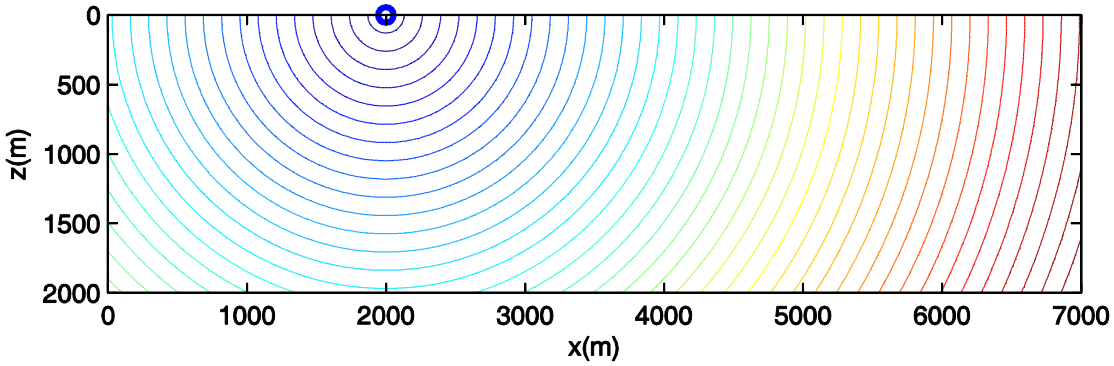
Figure 60: A smooth velocity field corresponding to a simple fault system and a point scatterer located at $\mathbf{r}$. The source and receiver are located, respectively, at $\mathbf{r}_s$ and $\mathbf{r}_g$. The local directions of the Green's functions of the incident and scattered wavefields are given by $\mathbf{k}_s$ and $\mathbf{k}_g$ and $\mathbf{K}$ is the corresponding scattering wavenumber.

3.2 PRESTACK DEPTH MIGRATION AND POINT SPREAD FUNCTION

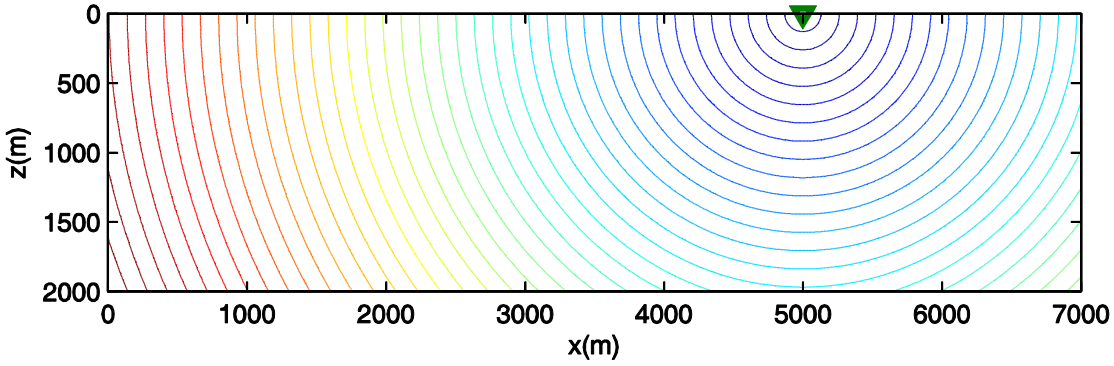
The aim of depth migration is to locate in depth the information recorded in time. In order to do so, it relies on the velocity field and the traveltimes, representing the link between the time and depth domains. For illustration purposes, a homogeneous model considering a single point scatterer is considered. Figure 61(a) and Figure 61(b) show the one way traveltimes from a source located at $\mathbf{r}_s$ and a receiver at $\mathbf{r}_g$. Such traveltimes are denoted, respectively, by $\tau_s(\mathbf{r}, \mathbf{r}_s)$ and $\tau_g(\mathbf{r}_g, \mathbf{r})$. Note that the curves with constant traveltime values represent the wavefronts. Figure 61(c) shows the scattering traveltimes,

$$\tau(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}) = \tau_s(\mathbf{r}, \mathbf{r}_s) + \tau_g(\mathbf{r}_g, \mathbf{r}),$$

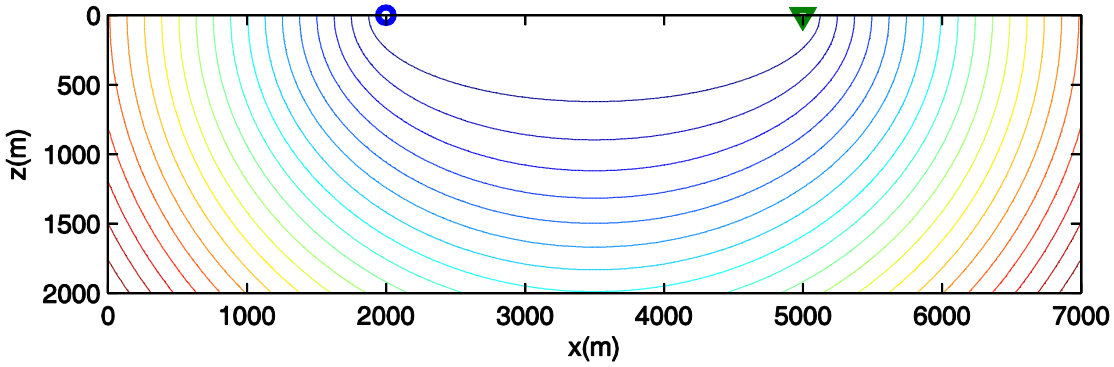
which link depth and time when considering the backscattered energy emitted at $\mathbf{r}_s$ and received at $\mathbf{r}_g$. The points of the image where $\tau(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$ is constant form the *scattering isochrones*. In a homogeneous medium, they have an elliptical shape, with foci located at $\mathbf{r}_s$ and $\mathbf{r}_g$. If we consider a single trace and place the point scatterer as in Figure 62(a) and Figure 62(b), where it is indicated as a yellow star, the migrated image will follow the isochrone curve which contains the respective scatterer. In Figure 62(a), the source and receiver coincide (zero offset, ZO), and are just above the point scatterer. The resulting image is a circle (an ellipse with coinciding foci) with constant width. In Figure 62(b), an offset exists between the source and receiver, so the migrated image has an elliptical shape, wider than in the ZO case and also with a varying width. These images are examples of impulse responses of the PSDM algorithm. Essentially, this means that if a seismic event is received at a given traveltime in the trace, the event may have occurred at any point along the respective isochrone.



(a) Isotraveltimes from source.

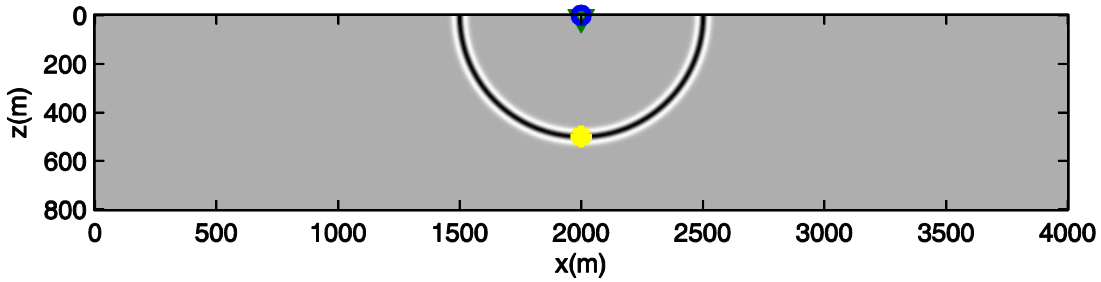


(b) Isotraveltimes from receiver.

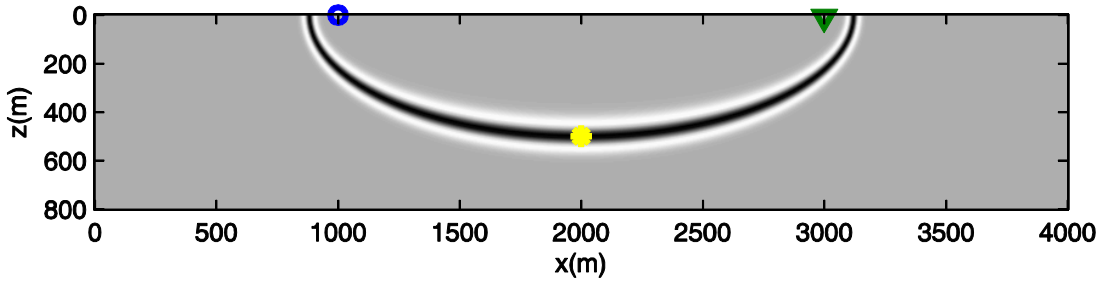


(c) Isochrones.

Figure 61: Considering an homogeneous model: (a) show the contour lines of the traveltimes from the source (blue circle) to the points of the model. (b) shows the contour lines of the traveltimes from the points of the model to the reciver (green triangle). As a homogeneous model is considered, these lines form circumnferences with center in the source and receiver, respectively. (c) shows the countour lines of the scattering traveltime, which describes the total traveltime from the source, to the model points and back to the receiver. These lines, called isochrones, form ellipses with foci located at the source and receiver.



(a) Impulse response of PSDM for a single ZO trace.

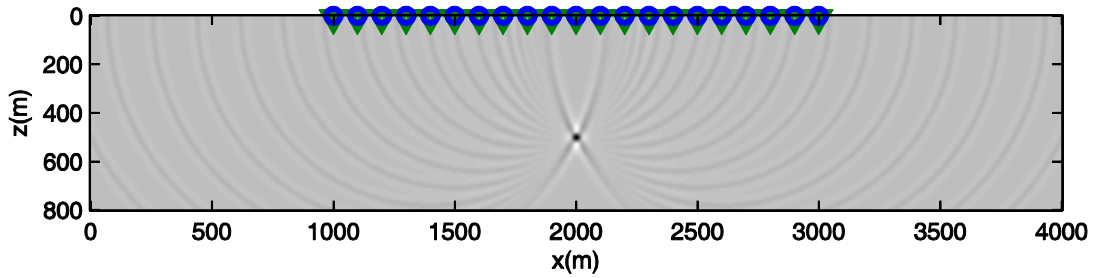


(b) Impulse response of PSDM for a single trace whose offset is 2000m.

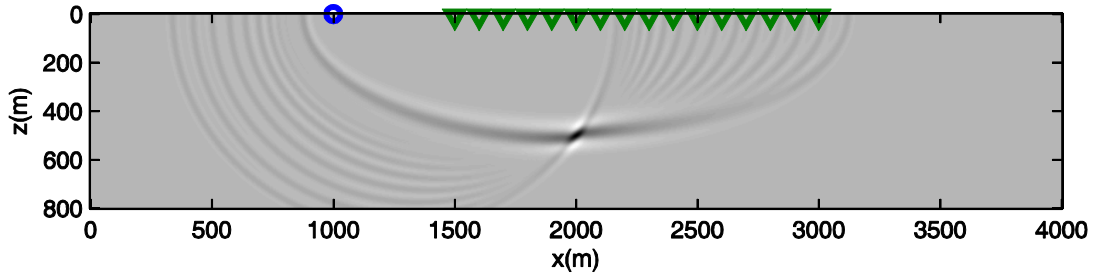
Figure 62: Impulse response of the PSDM algorithm in a homogeneous medium. The yellow star corresponds scatterer position. Blue circles and green triangles correspond, respectively, to the source and receiver.

When several traces are available, each one gives rise to its own isochrone. The resulting migrated image is obtained by adding these individual isochrones. As an example, in Figure 63(a), there is the result of the sum of different impulse responses calculated for different ZO traces for a single scatterer. Observe that there is a constructive interference at the point-scatterer location, while destructive interference takes place at other points, thus eliminating the ambiguity found in Figure 62. However, the resulting image is not a perfect impulse, but rather a blurred version of it, i.e., the PSF. Thus, the PSF limits the ability to resolve the position of the scatterer. Figure 63(b) is the corresponding result for a common-shot gather. By comparing these two images, we observe that the resulting patterns around the scatterer position are different, in spite of the fact that the same scatterer and pulse were used. Thus, the use of two different acquisition geometries produced two different PSFs. As previously discussed, the migrated image should show only the point scatterer. However, as mentioned before, distortion is caused by the use of a finite number of traces, which implies limited aperture, and also by limited frequency bandwidth.

From the foregoing discussion, we see that it may be possible to estimate a PSF by migrating the measured data associated with a point scatterer. However, this is very costly, since it involves the use of prestack synthetic data followed by migration. According to [LECOMTE, 2008], however, such procedure is the only alternative when it comes to wave-type migration. For migration methods of Kirchhoff type, in which a local interaction assumption is considered, a scattering wavenumber vector approach, as introduced in (66), can be applied with much less cost upon the use of ray tracing.



(a) Result of PSDM for a ZO gather.



(b) Result of PSDM for a common-shot gather.

Figure 63: Migrated image of complete trace gathers. The blue circles and the green triangles correspond, respectively, to the sources and receivers. The velocity model is homogeneous.

To show this claim, we start by considering the illumination vector, $\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$, which is defined in [LECOMTE, 2008] as

$$\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}) = \mathbf{p}_g(\mathbf{r}_g, \mathbf{r}) - \mathbf{p}_s(\mathbf{r}, \mathbf{r}_s) = \frac{1}{c(\mathbf{r})}\left[\hat{\mathbf{u}}_g(\mathbf{r}_g, \mathbf{r}) - \hat{\mathbf{u}}_s(\mathbf{r}, \mathbf{r}_s)\right], \tag{69}$$

where $\hat{\boldsymbol{u}}_s$ and $\hat{\boldsymbol{u}}_g$, are unit vectors, evaluated at the image point, $\mathbf{r}$, that are perpendicular, respectively, to the wavefront incident from the source at $\mathbf{r}_s$ and to the one scattered and received at $\mathbf{r}_g$.

As in (69), $c(\mathbf{r})$ is the velocity at the image point $\mathbf{r}$, then $\mathbf{p}_s(\mathbf{r}, \mathbf{r}_s)$ and $\mathbf{p}_g(\mathbf{r}_g, \mathbf{r})$ have the magnitude corresponding to the reciprocal of the velocity at $\mathbf{r}$ and directions corresponding to $\hat{\boldsymbol{u}}_s$ and $\hat{\boldsymbol{u}}_g$ and thus are defined as the slowness vectors associated with the respective wavefields. From the eikonal equation, we also have:

$$\nabla\tau_s(\mathbf{r}, \mathbf{r}_s) = \mathbf{p}_s(\mathbf{r}, \mathbf{r}_s) \text{ and } \nabla\tau_g(\mathbf{r}_g, \mathbf{r}) = -\mathbf{p}_g(\mathbf{r}_g, \mathbf{r}). \tag{70}$$

Therefore, by substituting (70) in (69) and comparing the result with (66), the relationship between the illumination vector and the scattering wavenumber vector is found to be

$$\mathbf{K}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}) = \omega\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}). \tag{71}$$

Figure 64 sketches how the slowness vectors can be easily computed with the use of ray tracing, since they are tangent to the ray paths which meet at the image point. It also depicts the resulting illumination vector $\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$. If a single temporal frequency ω is considered, (71) represents the 2D Fourier Transform (2DFT) of a monochromatic plane wave (a point in the wavenumber domain). If a pulse with limited bandwidth, as shown in Figure 65, is considered, all frequencies are mapped to the wavenumber domain along the scattering wavenumber vectors according to (71) and shown in Figure 66. Figure 67 shows the Fourier amplitudes of the wavelet along the scattering wavenumber vector indicated in red in Figure 66. The amplitudes are symmetric with respect to the origin, since both positive and negative frequencies are considered here. This spectrum represents a band-limited plane wave in the space domain as the one displayed in Figure 68.

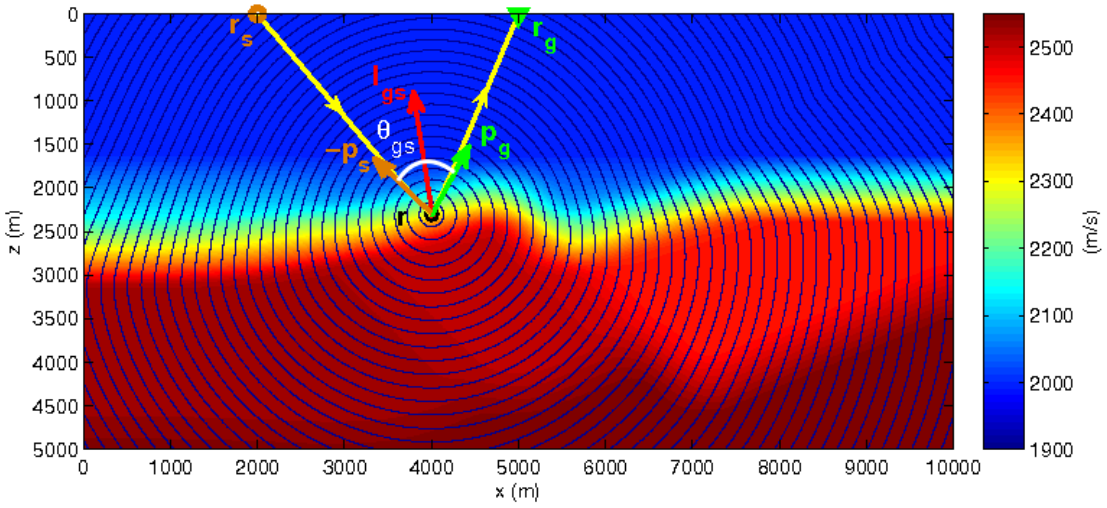


Figure 64: Same setting as in Figure 60: $\mathbf{I}$ is the illumination vector and θ is the opening angle between $-\mathbf{p}_s$ and $\mathbf{p}_g$. The first arriving wavefronts from a secondary source at the point scatterer are displayed.

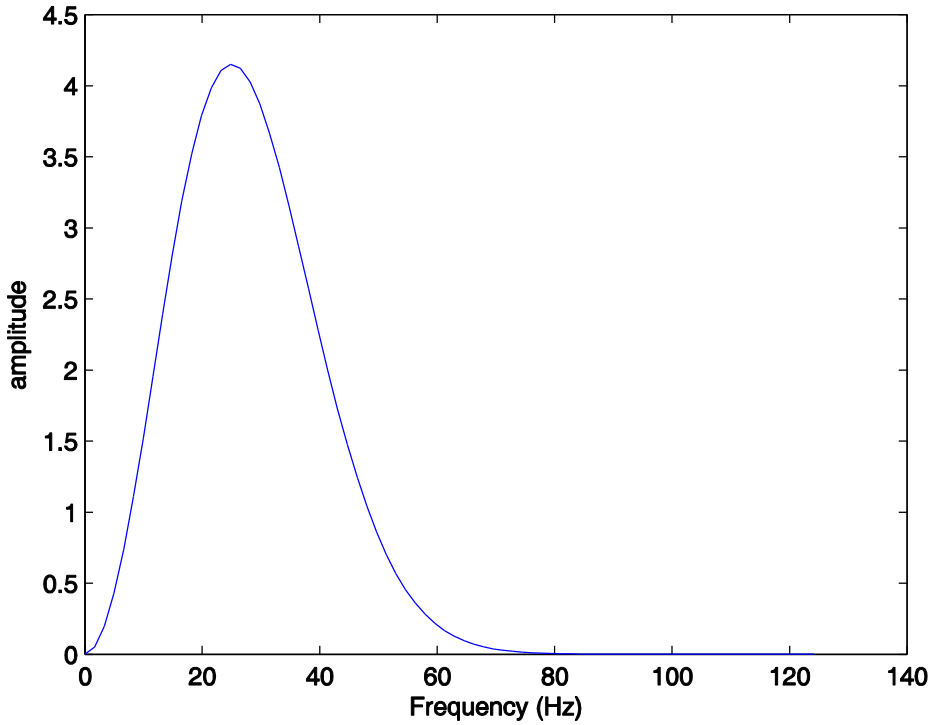


Figure 65: The amplitude spectrum of a Ricker wavelet [RICKER, 1944; HOSKEN, 1988] with central frequency of 25Hz .

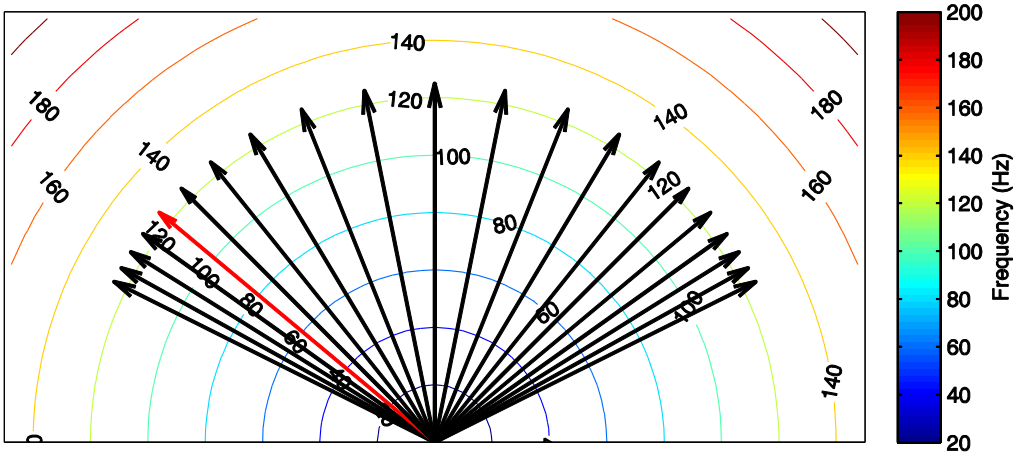


Figure 66: Scattering wavenumber vectors for the survey in Figure 63(a). The sampling rate is 4ms, thus the maximum length of the vectors corresponds to the Nyquist frequency, 125Hz.

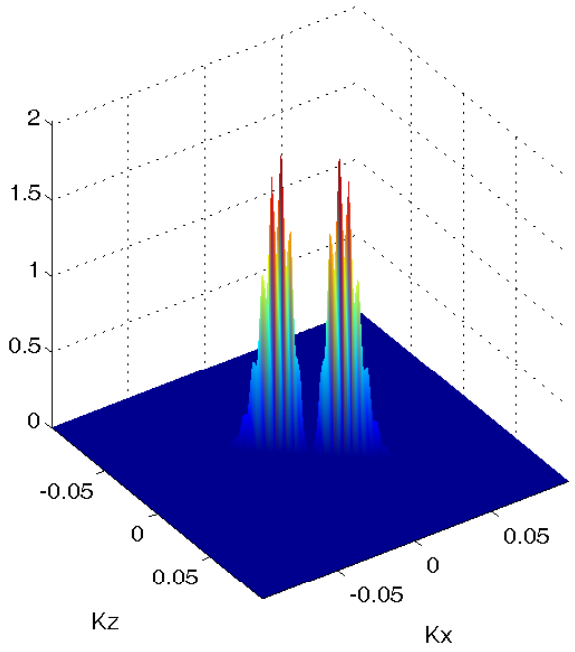


Figure 67: Fourier amplitudes of the Ricker wavelet of Figure 65 mapped onto the scattering wavenumber vector displayed in red in Figure 66. Some imperfections in the mapping caused the ripples in the amplitude spectrum.

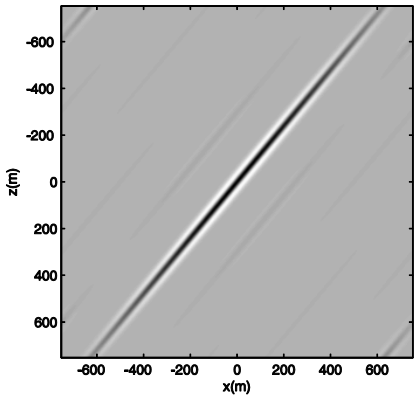


Figure 68: Band limited plane wave corresponding to the spectrum in Figure 67. The ripples in the amplitude spectrum in Figure 67 produce the artifacts shown as weak diagonal lines parallel to the main plane wave.

3.3 2D PSF ESTIMATION

As (71) and also the discussion in the previous section suggest, the PSF at a point of a PSDM image may be calculated, whenever a velocity model in depth and a pulse amplitude spectrum are given. Namely, in algorithmic form, we have:

Step 1: Select an image point.

Step 2: For each trace (source-receiver pair) do:

Step 2.1: Estimate the scattering isochrones and the illumination vector.

Step 2.2: Map the amplitude spectrum in the wavenumber domain according to (71) and Figure 65-Figure 67 to obtain the corresponding scattering wavenumber vector.

Step 3: Since the PSDM image represents the superposition of impulse responses (such as the ones shown in Figure 63(a) and Figure 63(b)), the scattering wavenumber vectors corresponding to each trace are to be summed. This result gives the PSF in the wavenumber domain.

Step 4: Transform the PSF from the wavenumber domain to the space domain as in Figure 68.

In our implementation, the isochrone estimation in Step 2.1 was done with the aid of the software package FDTIMES [PODVIN, 2006] developed in [PODVIN and LECOMTE, 1991]. Given a velocity field and a source location, the package provides the traveltime of the first arriving wavefront for all points of the model, as shown in Figure 64. In this figure, a secondary source located in the subsurface of the velocity model presented in Figure 60 was simulated. The rays to the source and receiver at the surface were traced using the traveltime gradients. Such rays are perpendicular to the wavefronts. The package was also used in the homogeneous model (C=1500m/s) analyzed in Figure 61. The scattering traveltimes, $\tau(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$, in Figure 61(c) were obtained as the sum of the source and receiver traveltimes to the image point (see Figure 61(a) and Figure 61(b)). As an illustration, Figure 69(a) and Figure 69(b) show several PSFs calculated for points along the isochrones (impulse responses) in Figure 62(a) and Figure 62(b). By comparing the PSFs with the migrated images, it is verified that when a single trace and a single event are considered, the PSFs consist in plane approximations that are tangent to the pattern of the respective migrated image. Also, it is observed that the width of the PSFs follow the width of the resulting pattern of the migrated images.

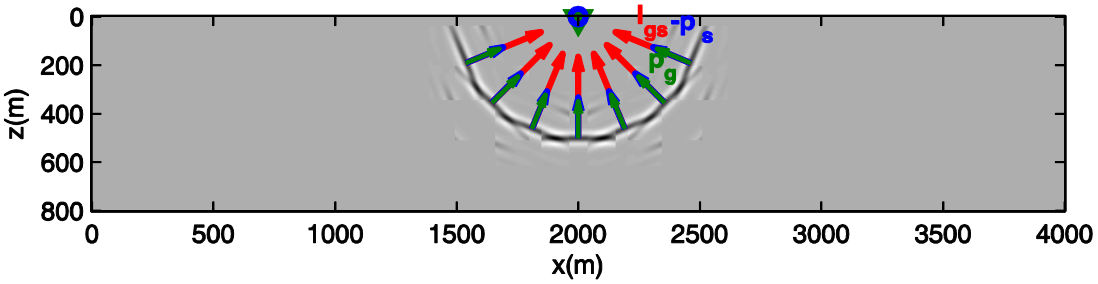
The result of variation in direction and width, i.e., resolution, can be explained with the use of the illumination vectors, **I**, described in (69) and plotted, together with the slowness vectors, $-\mathbf{p}_s$ and $\mathbf{p}_g$, in Figure 69(a) and Figure 69(b), It is possible to verify that the PSFs are perpendicular to the illumination vectors, as the **I** vector is related to the representation of the PSF in the wavenumber (**K**) domain as shown in (71) [LECOMTE, 2008]. The variation of resolution of the PSFs along the isochrones can be analyzed, if we recast (69) in the form

$$\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}) = \frac{2\cos(\theta/2)}{c(\mathbf{r})}\hat{\mathbf{u}}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r}), \tag{72}$$

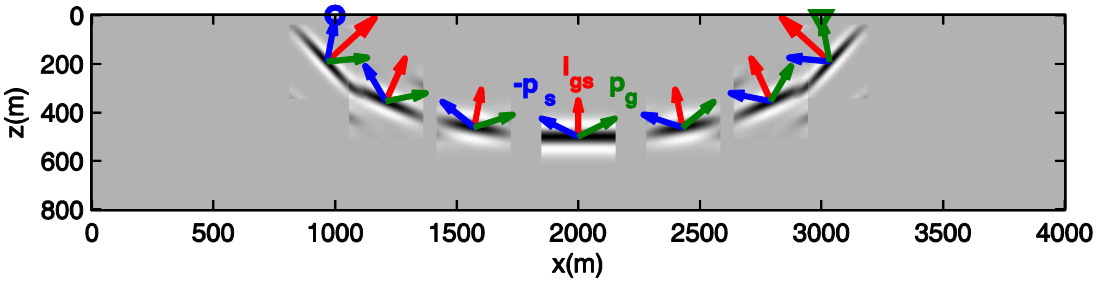
where θ is the opening angle between $-\mathbf{p}_s$ and $\mathbf{p}_g$ and $\hat{\mathbf{u}}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$ is the unit vector in the direction of the illumination vector [LECOMTE, 2008]. It is interesting to notice that $\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$ has a maximum length of $2/c(\mathbf{r})$ when $\theta = 0$, namely in case of normal-incidence backscattering. This can be observed in Figure 69(a), where the PSFs were calculated along an isochrone corresponding to the zero offset, single trace, scenario presented in Figure 62(a). In this case, $\theta = 0$ at all points of the isochrone and the size of the illumination vector, plotted in red in Figure 69(a), is constant. As a consequence, the thickness of the migrated trace in Figure 69(a) is also constant along the isochrone. As indicated by (72), the size of $\mathbf{I}(\mathbf{r}_g, \mathbf{r}_s, \mathbf{r})$ decreases as θ increases. This

can be observed in Figure 69(b) where the PSFs were calculated along the isochrone corresponding to the non-zero offset scenario presented in Figure 62(b). In this case, the angle is wider at the bottom of the isochrone. Thus, the illumination vector is smaller in this region and, as shown in Figure 62(b) and Figure 69(b), the resolution is better in the regions closer to the receiver or the source and it becomes worse near the bottom. In fact, if we consider contribution of the illumination vector to the scattering wavenumber vector, as indicated in (71), we verify that having a small illumination vector affects the resolution of the seismic image as much as having a small frequency bandwidth in the source signature.

As a next illustration, we compute the illumination vectors for the acquisition geometries in Figure 63(a) and in Figure 63(b). These are displayed, at the point-scatterer position, in Figure 70(a) and in Figure 70(b). For the ZO case of Figure 70(a), we have $\theta = 0$ and the illumination vector has the maximum length for all traces. For the common-shot situation of Figure 70 (b), the size of the illumination vector decreases as the offset and θ increase, indicating a poorer resolution at the larger offsets. We clearly see an improved resolution for the ZO case when compared to the CS counterpart whose PSFs are shown in Figure 71(a) and Figure 71(b), respectively. The same comparison can be made when we consider the blur patterns observed in the actual PSDM images in in Figure 63(a) and in Figure 63(b). Note that the PSFs in Figure 71 are good approximations of the blurring patterns around the scatterer in Figure 63.

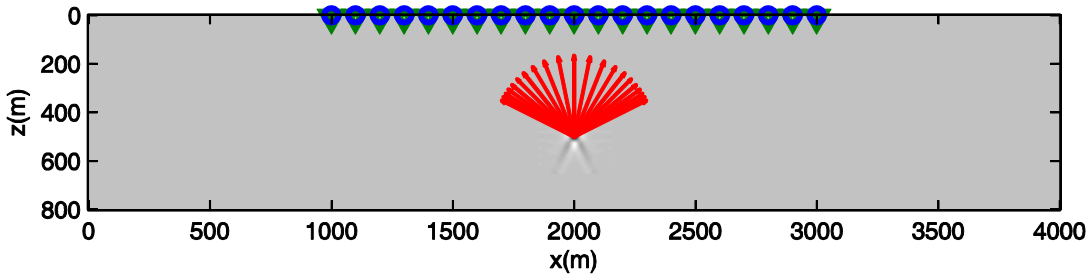


(a) PSFs along an isochron corresponding to a migrated ZO trace.

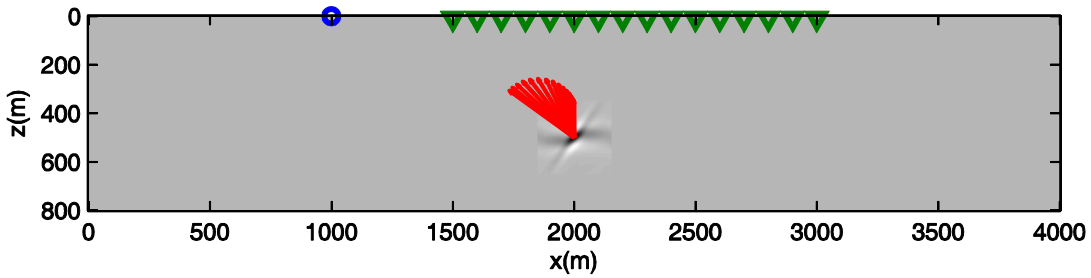


(b) PSFs along an isochron corresponding to a migrated trace with offset 2000m

Figure 69: PSFs calculated along two migrated traces corresponding to the acquisition geometries in Figure 62(a) and in Figure 62(b). The blue, red and green vectors correspond respectively to $-\mathbf{p}_s$, $\mathbf{I}$ and $\mathbf{p}_g$.

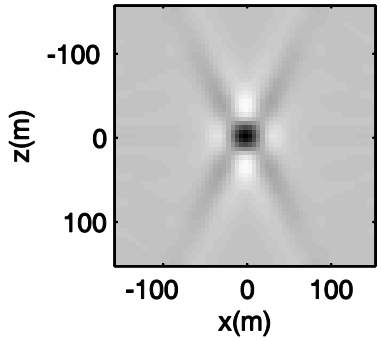


(a) Illumination vectors calculated at the scatterer for a ZO geometry

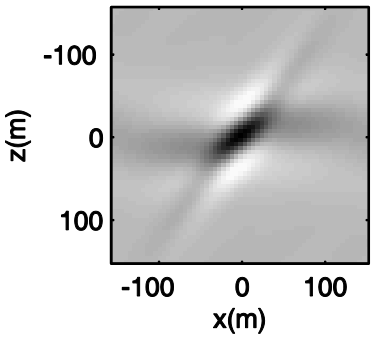


(b) Illumination vectors calculated at the scatterer for a common-shot geometry

Figure 70: Illumination vectors for acquisition geometries of Figure 63(a) and of Figure 63(b).



(a) PSF corresponding to Figure 70(a).



(b) PSF corresponding to Figure 70(b).

Figure 71: PSFs calculated at the scatterer for the acquisition geometries: (a) ZO of Figure 63(a) and (b) CS of Figure 63 (b).

3.4 2D DECONVOLUTION

In this section, we show a method to improve the resolution of an image by mitigating the effects of the PSF. The 2D spiking deconvolution filter is analogous to the 1D case in Chapter 2. In the 1D case, a deconvolution filter is calculated in order to transform the source signature into a narrow wavelet, ideally a spike, in the resulting trace, thus enhancing the vertical resolution. In the 2D case, a wide PSF (Figure 72(a)) is filtered with a 2D filter (Figure 72(b)) to produce a narrower PSF (Figure 72(c)). As in the 1D case, the ideal 2D spiking deconvolution filter is the one that transforms a PSF into a 2D spike (Figure 72(d)), defined as

$$\delta(i,j) = \begin{cases} 1 \text{ if } i = 0 \text{ and } j = 0 \\ 0 \text{ otherwise} \end{cases}, i,j \in \mathbb{N}.$$

So let us define the following signals:

- The sampled versions of the PSDM (blurred) image ($2M+1 \times 2N+1$ image):

$$x(m,n) = \begin{cases} x(m,n) \in \mathbb{R} \text{ if } m = -M, -M+1, \dots, M, \quad n = -N, -N+1, \dots, N \\ 0 \text{ otherwise} \end{cases}, m,n \in \mathbb{N}$$

- PSF ($2K+1 \times 2L+1$ image):

$$h(k,l) = \begin{cases} h(k,l) \in \mathbb{R} \text{ if } k = -K, -K+1, \dots, K, \quad l = -L, -L+1, \dots, L \\ 0 \text{ otherwise} \end{cases}, k,l \in \mathbb{N}$$

- 2D deconvolution filter ($2P+1 \times 2Q+1$ image):

$$w(p,q) = \begin{cases} w(p,q) \in \mathbb{R} \text{ if } p = -P, -P+1, \dots P, \quad q = -Q, -Q+1, \dots, Q \\ 0 \text{ otherwise} \end{cases}, p,q \in \mathbb{N}$$

- Filter output ($2M+2P+1 \times 2N+2Q+1$ image):

$$c(i,j) = \begin{cases} c(i,j) \in \mathbb{R} \text{ if } i = -M-P, -M-P+1, \dots, M+P, \quad j = -N-Q, -N-Q+1 \dots, N+Q \\ 0 \text{ otherwise} \end{cases},$$

$$i,j \in \mathbb{N}$$

As $c(i,j)$ is the output of the deconvolution filter, it is given by

$$c(i,j) = \sum_{p=-P}^{P} \sum_{q=-Q}^{Q} w(p,q)x(i-p, j-q) = w(i,j) ** x(i,j), \tag{73}$$

where the blurred image, $x(i,j)$, is the input, $w(p,q)$ is the 2D spiking deconvolution filter and the symbol ** denotes the 2D deconvolution.

In the context of finding the 2D spiking deconvolution filter, we first consider an input image generated by the blurring of the single spike in Figure 72(d). The resulting figure is thus the single PSF shown in Figure 72(a). In this case, we have that $x(i,j) = h(i,j)$. The goal then is to find a 2D filter, as the one in Figure 72(b), that outputs an image which is similar to a spike (Figure 72(c)) when the input is $h(i,j)$, i.e.,

$$c^h(i,j) = w(i,j) ** h(i,j) \approx \delta(i,j). \tag{74}$$

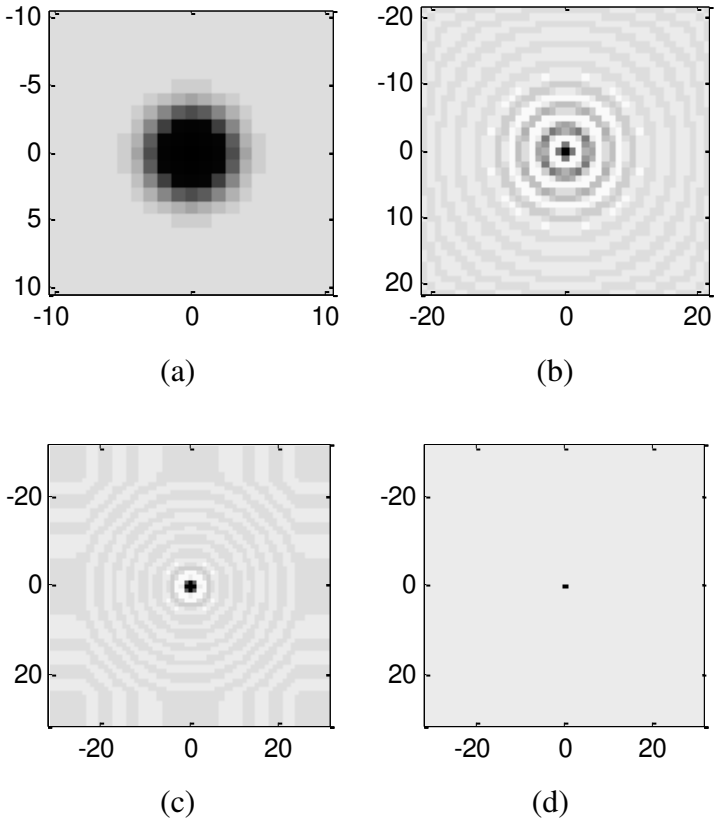


Figure 72: (a) PSF. (b) 2D deconvolution filter. (c) 2D Deconvolved PSF. (d) Ideal spike.

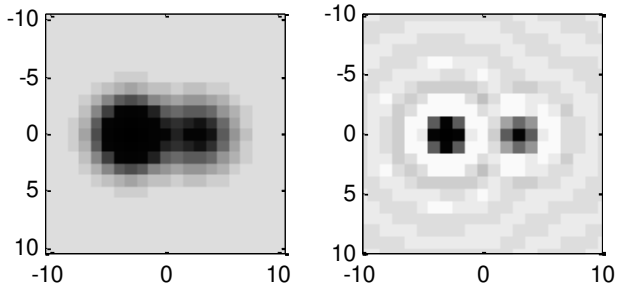
In order to analyze the effect of this deconvolution filter in a general image, we start by considering an original image, $s(i,j)$, that is blurred by a linear and space invariant PSF, $h(i,j)$, resulting in a blurred version, $x(i,j)$, such that:

$$x(i,j) = s(i,j) ** h(i,j).$$

The 2D spiking deconvolution filter used in (74) can then be used to restore $s(i,j)$ from $x(i,j)$, or at least to mitigate the effect of the PSF, as

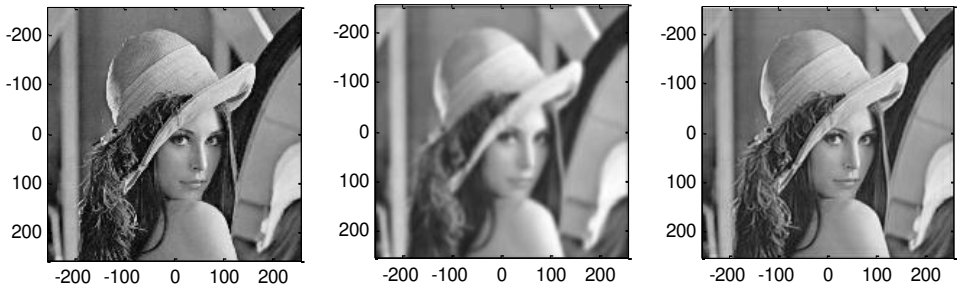
$$x(i,j) ** w(i,j) = s(i,j) ** h(i,j) ** w(i,j) \approx s(i,j) ** \delta(i,j) = s(i,j).$$

In fact, an actual blurred image can be seen as a superposition of several PSFs, each one of them around one pixel of the original image, weighted by its amplitude. The 2D sipky deconvolution filter works by enhancing the resolution of each of these shifted and scaled PSFs. As an example, we consider a scenario where there are two blurred spikes, as shown in Figure 73(a). By performing the 2D deconvolution, the two points become distinguishable, as shown in Figure 73(b). A more general case is shown in Figure 74, where the original Lena image in Figure 74(a) is blurred with a PSF, as shown in Figure 74(b).The blurred image is deconvolved and the result shown in Figure 74(c) is very close to the original one.



(a) (b)

Figure 73**:** (a) Two separate points are blurred by a PSF. (b) 2D Deconvolved image, the points are now cleary distinguishable.



(a) (b) (c)

Figure 74: (a) Original Lena image. (b) Image blurred with Gaussian PSF. (c) 2D Deconvolved image.

3.4.1 Obtaining the the 2D spiking deconvolution filter

The least squares criterion is used to calculate the optimal 2D spiking deconvolution filter similarly to the case of 1D deconvolution explored in the previous chapter. In order to do so, we define the representation of a 2D image by a 1D vector with the use of the lexicographical ordering as in [SJOEBERG *et al.*, 2003]. An example for $K = 1$ and $L = 1$, is given:

$$\mathbf{h} = \text{lex}\left\{\begin{bmatrix} h(-1,-1) & h(-1,0) & h(-1,1) \\ h(0,-1) & h(0,0) & h(0,1) \\ h(1,-1) & h(1,0) & h(1,1) \end{bmatrix}\right\} = \begin{bmatrix} h(-1,-1) \\ h(-1,0) \\ h(-1,1) \\ h(0,-1) \\ h(0,0) \\ h(0,1) \\ h(1,-1) \\ h(1,0) \\ h(1,1) \end{bmatrix}$$

In order to illustrate this, Figure 75(a) shows, for $K = 1$ and $L = 1$, where each color represents a different number. Figure 75(a) shows $h(k,l)$, and Figure 75(b) shows its lexicographical ordering representation. Figure 75(c) shows its flipped version, $h(-k,-l)$. A flipped and shifted version, $h(-1-k,-1-l)$ and its lexicographic representation, denoted by $\mathbf{h}^{-1,-1}$, are shown in Figure 75(d) and in Figure 75(e) respectively. The white spaces represent null elements.

Now, if we analyze the expression for $c^h(i,j)$, (74), rewritten as:

$$c^h(i,j) = \sum_{k=-P}^{P} \sum_{l=-Q}^{Q} w(k,l)h(i-k,j-l), \tag{75}$$

we verify that $h(i-k,j-l)$ is a flipped and shifted version of $h(k,l)$, as shown in Figure 75(c) for $i = j = 0$ and Figure 75(d) for $i = j = -1$. By denoting $\mathbf{h}^{i,j}$ as the lexicographic representation of $h(i-k,j-l)$, as the in the example in Figure 75(e), and $\mathbf{w}$ as the lexicographic representation of $w(k,l)$, we have, by substituting in (75), that

$$c^h(i,j) = (\mathbf{h}^{i,j})^T\mathbf{w}$$

If we write $c^h(i,j)$ in the lexicographical order, we can obtain a convolution matrix $\mathbf{H}$ whose rows correspond to the values of $(\mathbf{h}^{i,j})^T$, so that

$$\mathbf{c}^h = \begin{bmatrix} c(-M-P,-N-Q) \\ c(-M-P,-N-Q+1) \\ \vdots \\ c(M+P,N+Q) \end{bmatrix} = \begin{bmatrix} (\mathbf{h}^{-M-P,-N-Q})^T \\ (\mathbf{h}^{-M-P,-N-Q+1})^T \\ \vdots \\ (\mathbf{h}^{M+P,N+Q})^T \end{bmatrix} \mathbf{w} = \mathbf{H}\mathbf{w}.. \quad (76)$$

An example of a convolution matrix is given in Figure 76 for $P = 1, Q = 1, K = 1$ and $L = 1$, and $h(k,l)$ shown in Figure 75, where it is possible to verify that the seventh row of the matrix corresponds to $\mathbf{h}^{-1,-1}$ shown in Figure 75(e).

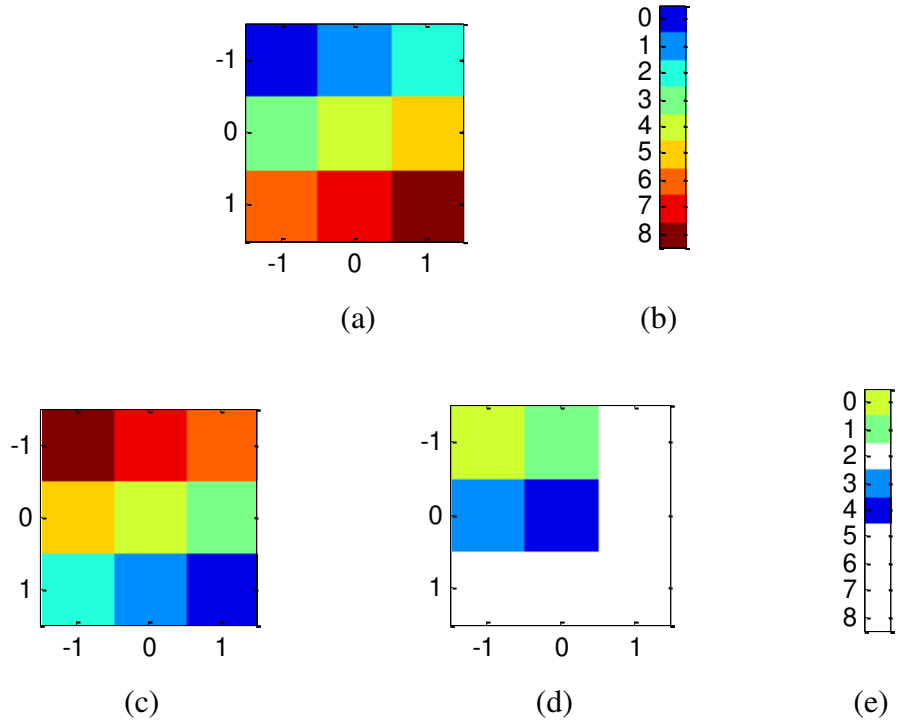


Figure 75: Illustrative example for $K = 1$ and $L = 1$. (a) $h(k,l)$. (b) lexicographic representation of $h(k,l)$. (c) $h(-k,-l)$. (d) $h(-1-k,-1-l)$. (e) $\mathbf{h}^{-1,-1}$.

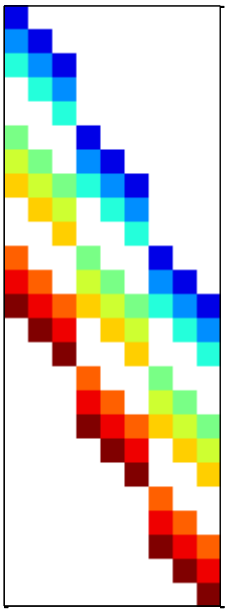


Figure 76**:** Illustrative example of **H** for P=1, Q=1, K=1 and L=1.

If we denote $\boldsymbol{\delta}$ as the lexicographical representation of $\delta(i,j)$, then, in order to obtain $\mathbf{c}^h \approx \boldsymbol{\delta}$**,** we need to solve the following optimization problem:

$$\min_{\mathbf{w}} \|\mathbf{H}\mathbf{w} - \boldsymbol{\delta}\|_2^2 + \lambda \|\mathbf{w}\|_2^2$$

whose solution is the optimum filter, according to the regularized least squares criterion, which is given by:

$$\mathbf{w} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{J})^{-1}\mathbf{H}^T\,\boldsymbol{\delta}$$

where **J** is the identity matrix[8] and λ is the regularization term introduced to avoid amplification of spectral notches.

[8] We use **J** instead of **I**, as in the usual notation for the identity matrix, in order to avoid confusion with the illumination vector, which is denoted by **I.**

3.5 RESULTS

3.5.1 Homogeneous medium

In this section, we show some results of applying the proposed method to a homogeneous model. Initially, a single trace configuration with offset equal to 2000m and a single scatterer is considered. The velocity of the medium is C = 2000m/s and a Ricker wavelet with center frequency of 25Hz is used [RICKER, 1944; HOSKEN, 1988]. Figure 77(a) shows the result of diffraction-stack migration [FRENCH, 1974] of this trace. The migrated sections obtained with this method actually indicate the probability of existence of a seismic event at each sampled point of a region of the subsurface. This is done by summing (i.e., stacking) the amplitudes of all traces at times where the response of a hypothetical scatterer located at the chosen point could be recorded during the seismic experiment. High absolute values indicate that this given point is highly likely to be an actual scatterer. In Figure 77(a), we see that the migration of a single trace results in an isochrone curve as in Figure 62, as all points of the ischrone are equally likely to be the origin of the recorded energy. Three possible scatterer positions are displayed as red crosses. Figure 77(b)-Figure 77(d) show the results of 2D deconvolution for the PSFs estimated at each cross. Note that the whole image is deconvolved with a single 2D deconvolution filter. The corresponding PSFs are shown in Figure 78(a)- Figure 78(c). It is observed that the resolution is increased near the scatterers, but the remaining part of the isochrone gets distorted. This is due to the fact that the PSF changes along the isochrone, as shown in Figure 69(a) and Figure 69(b). The assumption of a spatially invariant PSF is, thus, not globally correct. On the other hand, if a small target region is considered and the velocity field is smooth enough, it becomes reasonable to assume that the PSF is locally invariant in space. As an example, consider a scenario with two nearby scatterers. They are located at $(-20, 2000)^T$ and $(30, 2000)^T$ (coordinate system $(x, z)^T$). A common-shot gather (cf. Figure 79) with source at the origin and the receivers ranging from $(-1600, 0)^T$ to $(1600, 0)^T$ (10m spacing), is migrated by diffraction stacking. The result is shown in Figure 80(a), in which we see that it is not possible to discriminate the two scatterers. However, the use of 2D deconvolution, as shown in Figure 80(c), allows to discriminate the two scatterers. Thus, the use of 2D deconvolution enables increased lateral resolution. The 1D deconvolution was also

tested by considering just the central vertical part of the PSF, as in [GELIUS *et al.*, 2002]. The result is displayed in Figure 80(b). In this case, the lateral resolution is slightly increased, but the 2D deconvolution provides a better separation between the scatterers, as expected. The above 2D deconvolution result is comparable to the one presented in [GELIUS *et al.*, 2013], shown in Figure 80(d), in which a high resolution method called multiple signal classification (MUSIC) was used for the same data. The main difference is that MUSIC is rather a localization technique and its use implies the loss of amplitude information. That does not happen with the deconvolution technique. On the other hand, the result provided by MUSIC is more accurate in the sense that it does not show the migration artifacts and the ringing observed in the deconvolved image.

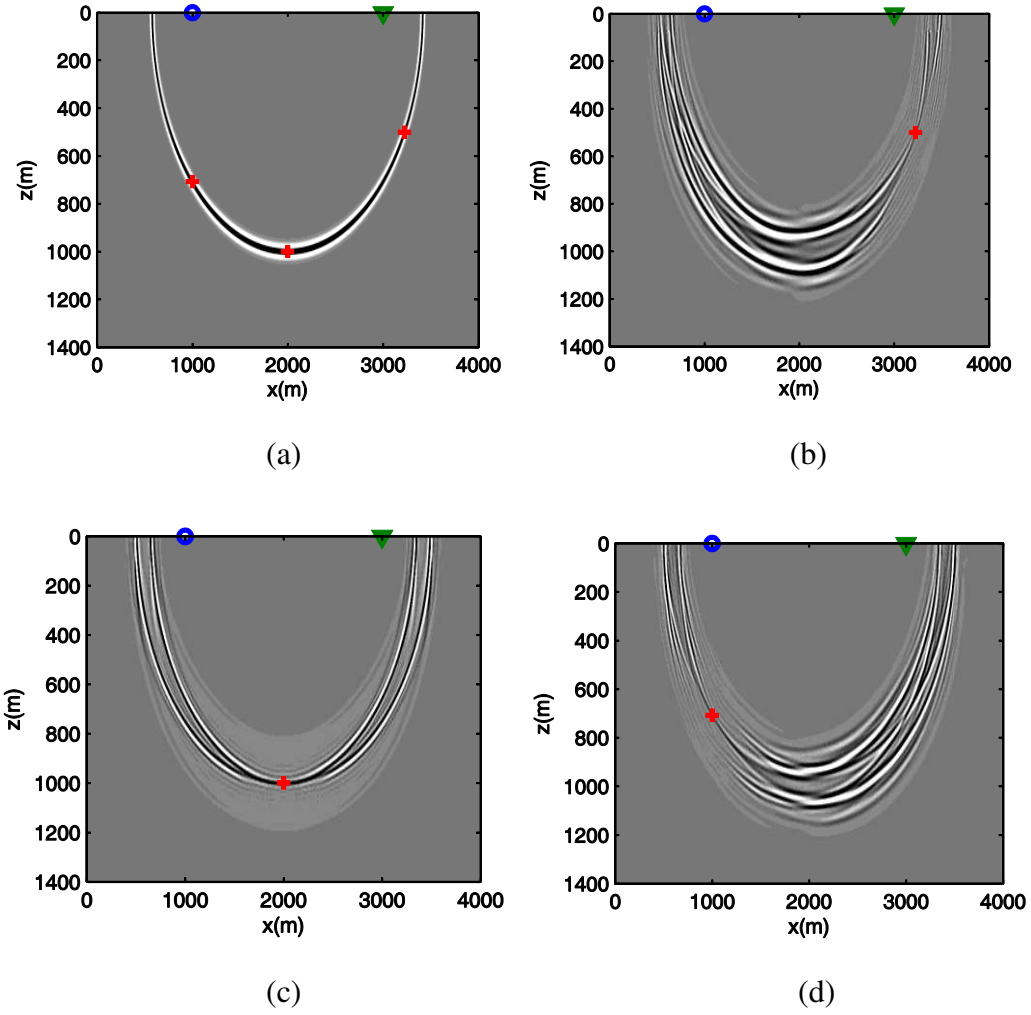
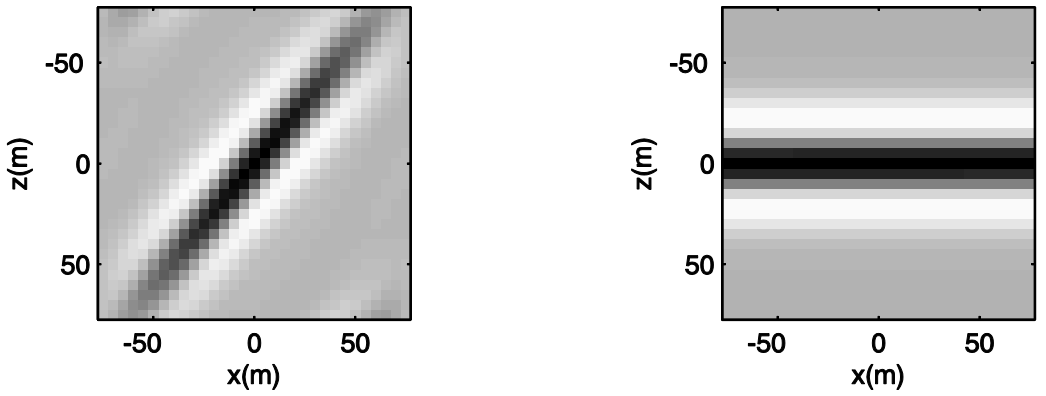
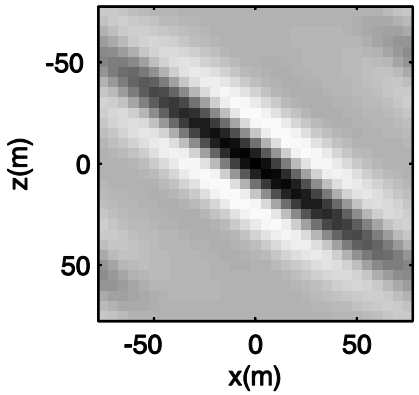


Figure 77 (a) Migrated single trace with offset 2000m. Red crosses represent three possible locations for scatterers; (b)-(d) Results of 2D deconvolution with PSFs estimated at each red cross.



(a)

(b)



(c)

Figure 78: PSFs used to obtain the 2D deconvolution results shown in, respectively: (a) Figure 77(b); (b) Figure 77(c) and (c) Figure 77(d).

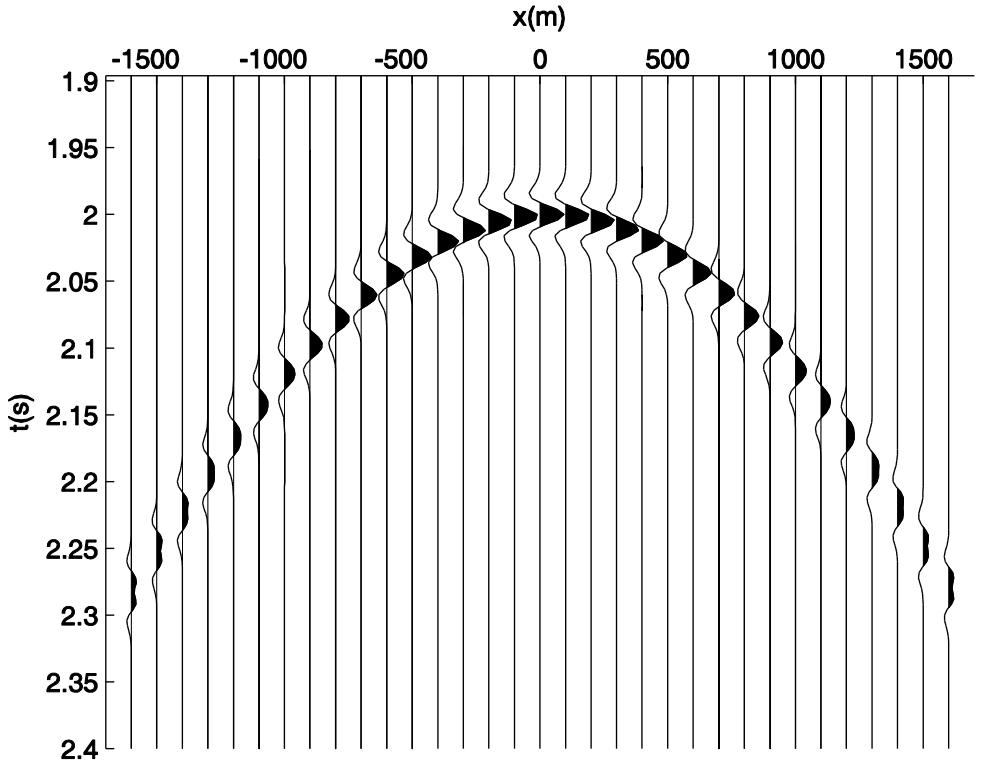
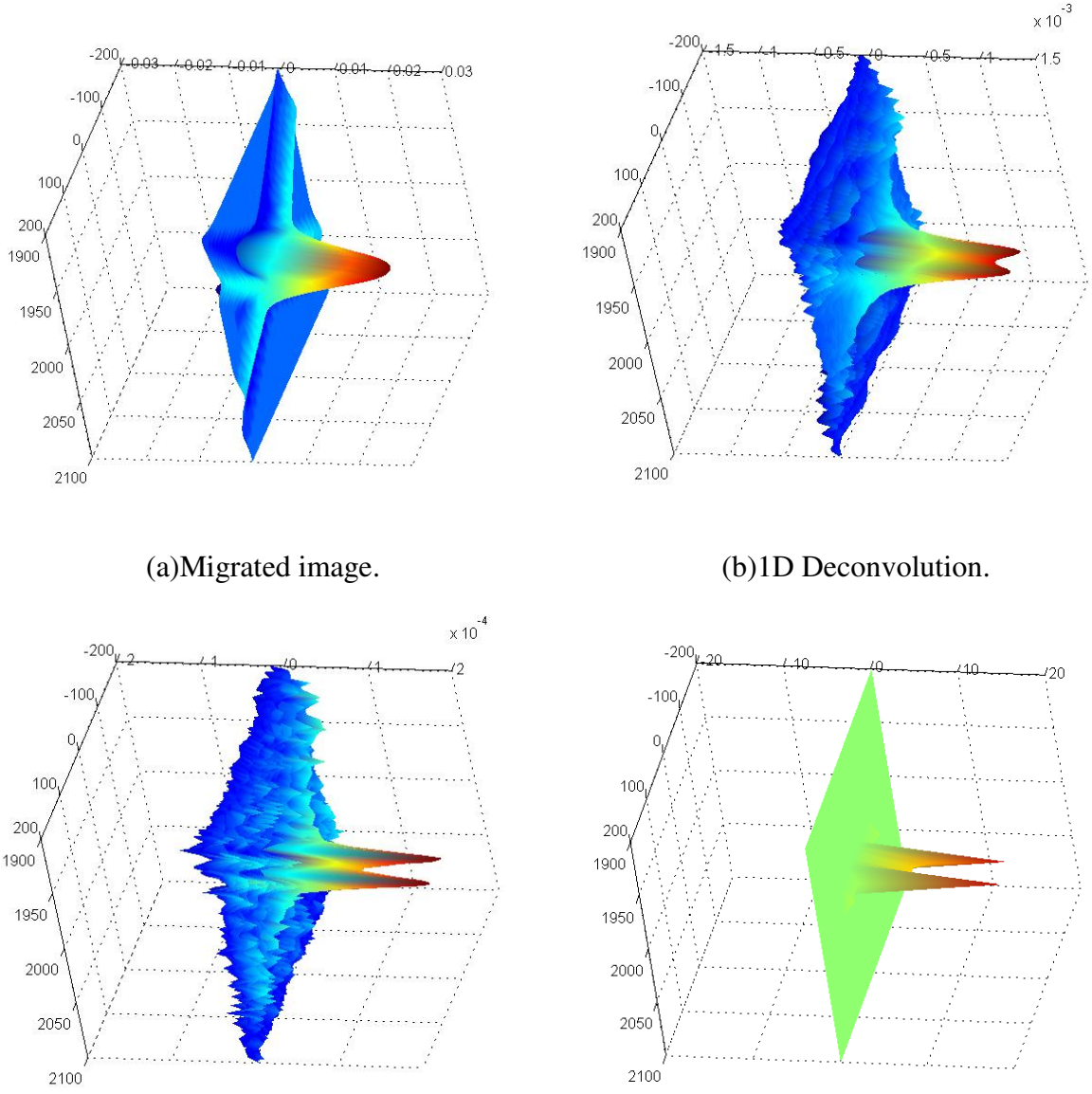


Figure 79: Common-shot gather of two nearby scatterers located at (-20, 2000)^T and (30, 2000)^T. The velocity is C = 2000m/s.



(a)Migrated image.

(b)1D Deconvolution.

(c) 2D Deconvolution.

(d) MUSIC imaging.

Figure 80: Migrated, deconvolved and MUSIC images of two nearby scatterers

3.5.2 Fault system

A fault model with layers of different velocities, as shown in Figure 81, is now considered. This model represents the original velocities before smoothing, as shown in Figure 60 and Figure 64. A CO section (4000m), whose acquisition geometry is described in Figure 81, was obtained using 2D ray tracing (NORSAR software package). In the figure, we show a small target area that has been selected for later migration. The PSF, shown in Figure 82 is calculated at the position indicated by a black cross in Figure 81.

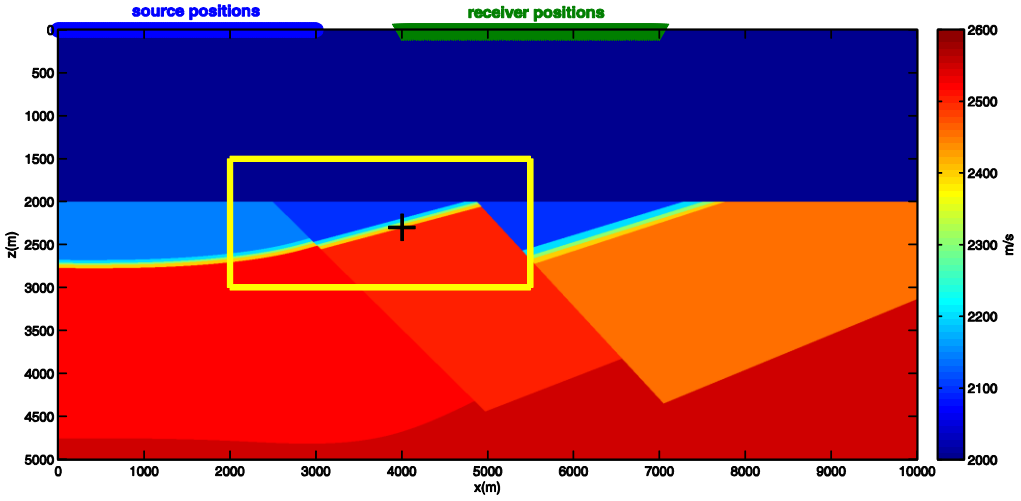


Figure 81: Fault system model: The selected target is highlighted with a yellow box. The black cross describes the point where the PSF for the 2D deconvolution is calculated.

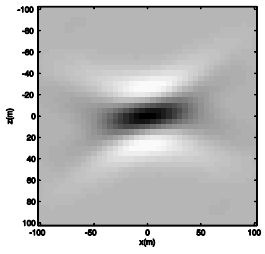
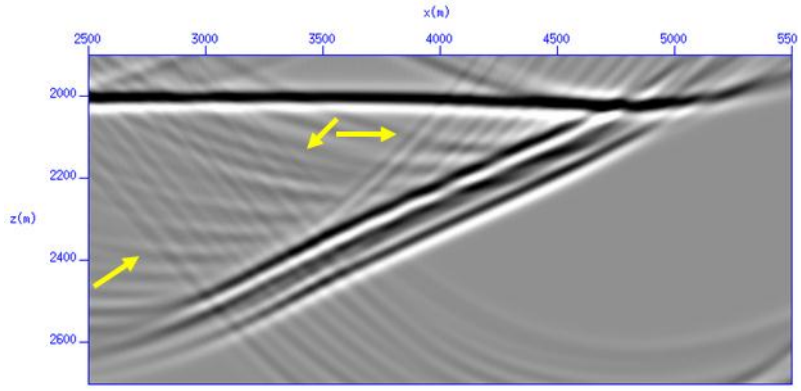


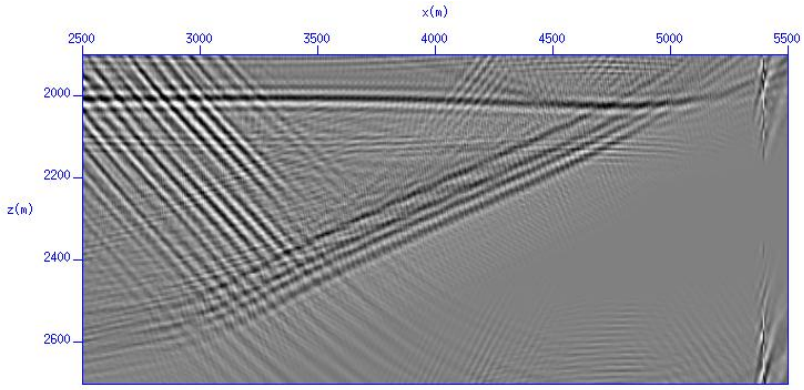
Figure 82: The PSF calculated at the position indicated by the black cross.

The diffraction-stack migration of the CO data corresponding to the small rectangle in Figure 81 is shown in Figure 83(a). Observe that, of the four inclined reflectors displayed in Fig-

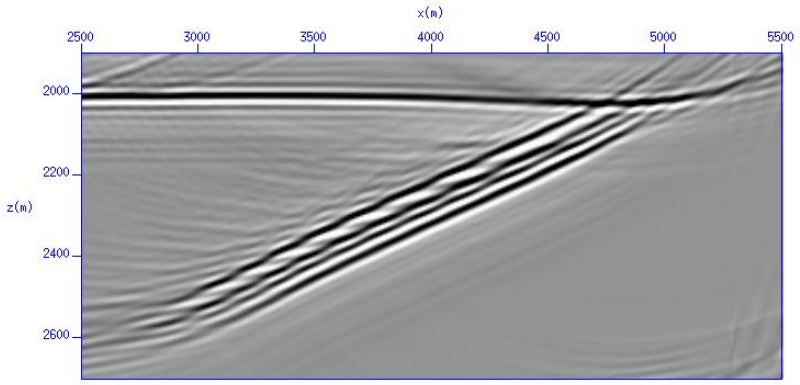
ure 84, only three are visible in the migrated image. The 2D deconvolution method was then applied to the migrated image. The result is shown in Figure 83(b). As can be seen in this figure, all four reflectors are now resolved. However, it is also observed that the original migration artifacts in Figure 83(a), such as the light colored curves pointed by the yellow arrows, known as migration smiles [SHERIFF, 2002], have been somewhat amplified. Also, some ringing is associated with the reflectors. This can be partly explained by observing the 2D Fourier Transform (2DFT) amplitude spectrum of the PSF in Figure 85(a). As the 2D spiking deconvolution filter seeks the inversion of the PSF in the mean squares sense, it enhances the low amplitude frequencies of the PSF. Since the PSF has null wavenumber components at the points which are not covered by the scattering wavenumber vectors (determined by the acquisition geometry, velocity model and the seismic frequency band), a high gain is associated with these directions. As these frequency components are associated with non-existing illumination directions and temporal frequencies of the actual data, they contain mostly noise, so their enhancement leads to the undesired effects seen in Figure 83(b). This is supported by the comparison between the 2DFT amplitude spectra of the migrated image and the deconvolved image, shown respectively in Figure 86(a) and Figure 86(b). For the migrated image, most of the energy is located around the origin and within the limits of the "band" associated with the 2DFT amplitude spectrum of the PSF seen in Figure 85(a). On the other hand, the 2DFT amplitude spectrum of the deconvolved image shows high amplitudes along a band around the $\mathbf{k}_x$ and $\mathbf{k}_y$ axes which continue to the edges of the spectra, and indicates the presence of high frequency artifacts, as seen in Figure 83(b). In order to eliminate these components, a 2D filter calculated from a tapered wavenumber mask displayed in Figure 85(b) was used. The 2DFT amplitude spectrum in Figure 85(b) was transformed to the space domain and the resulting filter was applied to the deconvolved image. The result is displayed in Figure 83(c). It is possible to observe in the resulting image that the artifacts and the high frequency ringing that were dimming the reflectors have now been attenuated. Also, in the resulting amplitude spectrum, in Figure 86(c), the high wavenumber components, i.e., the components outside the wavenumber band of the PSF have been attenuated. Nevertheless, we observe that the artifacts are still stronger in the deconvolved section. This may be avoided with use of more advanced regularization techniques in the deconvolution process, which is the object of current research.



(a) Migrated image.



(b) 2D deconvolution.



(c) 2D deconvolution followed by 2D filtering.

Figure 83: (a) shows the result of the migration of a CO section. Only three reflectors out of the four reflectors in Figure 84 are shown. The light colored curves pointed by the yellow arrows are artifacts produced by the migration algorithm known as migration smiles [SHERIFF, 2002]. (b) Shows the result of the use of the 2D deconvolution filter to the image in (a). The four reflectors are now resolved, but the migration artifacts from Figure (a) have been amplified. (c) Shows the result after 2D filtering with the tapered wavenumber mask. The artifacts have been attenuated.

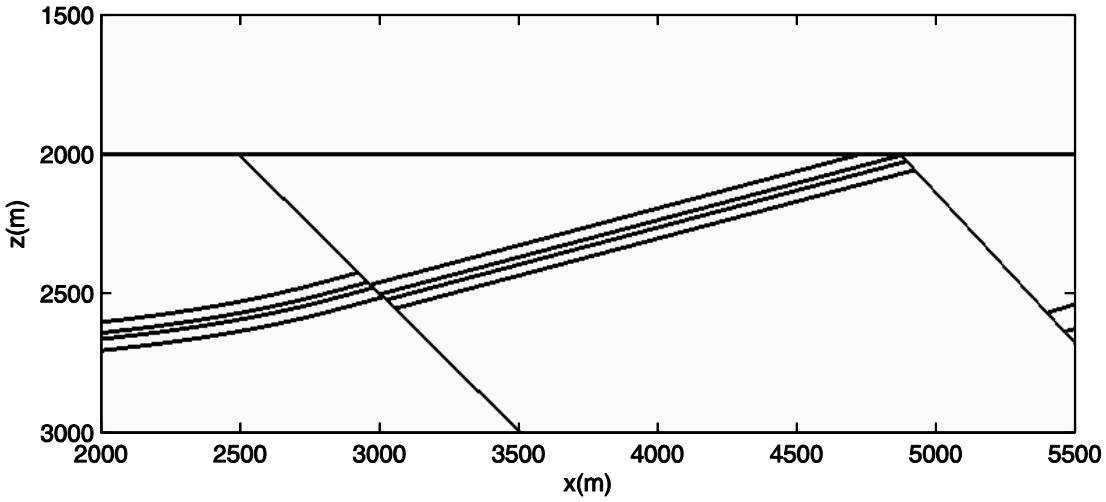
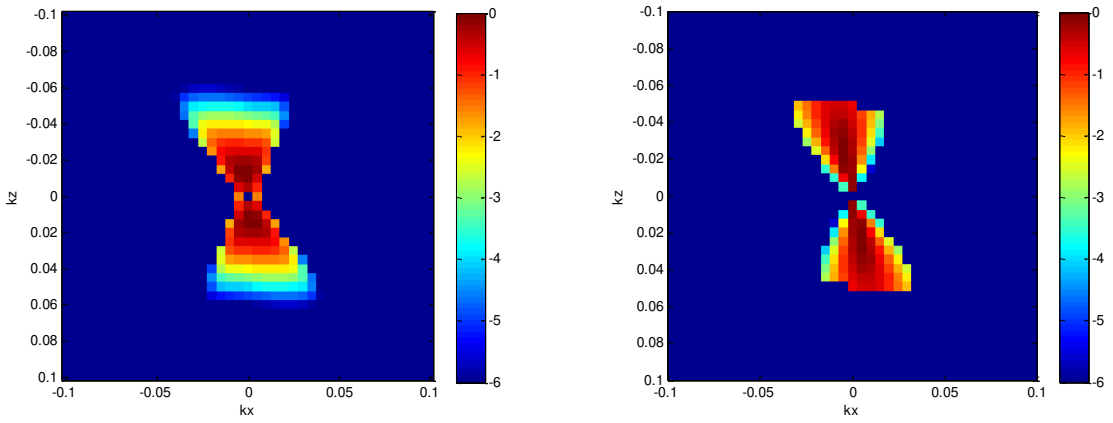


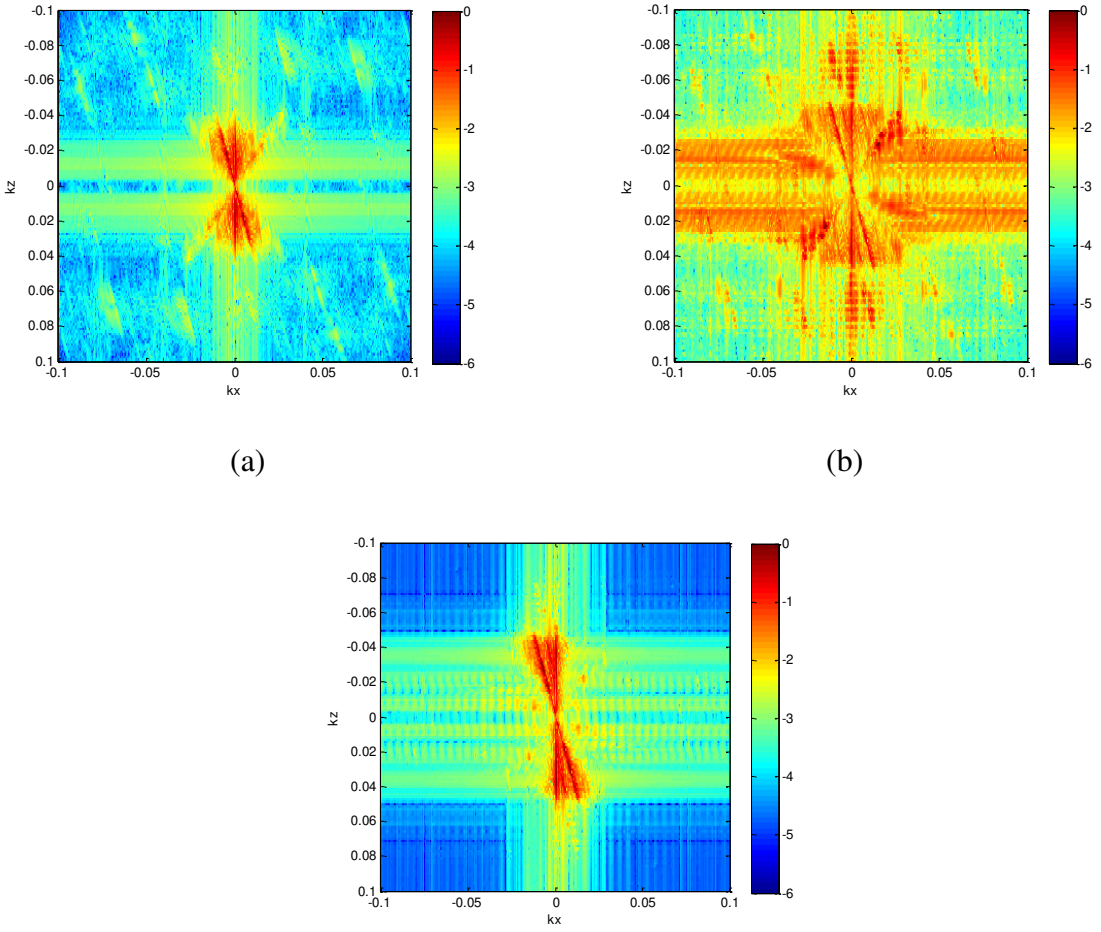
Figure 84: The interfaces in the selected target.



(a) Amplitude spectrum of the PSF.

(b) Tapered wavenumber mask.

Figure 85: (a) 2DFT amplitude spectrum of the PSF used for 2D deconvolution and (b) the tapered wavenumber mask used to filter the deconvolved image.



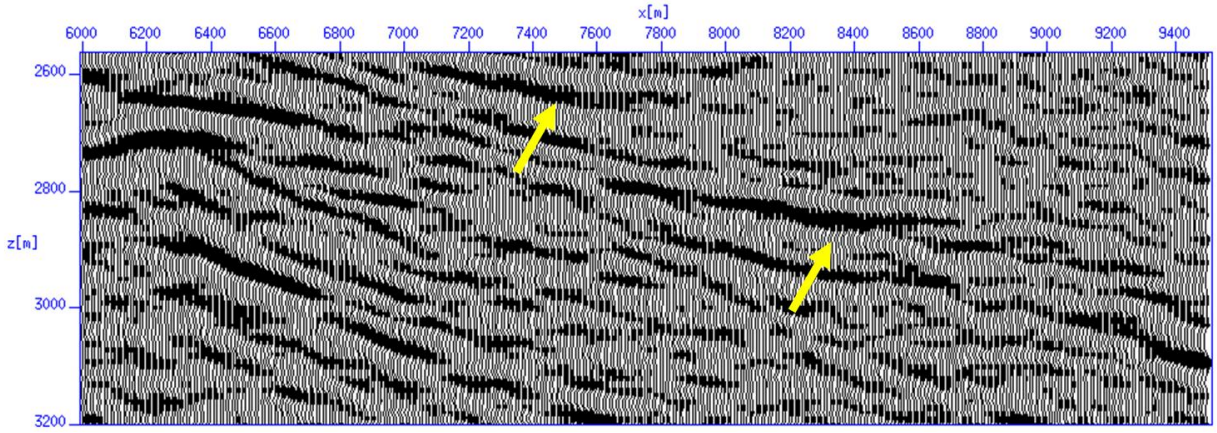
(a) (b)

(c)

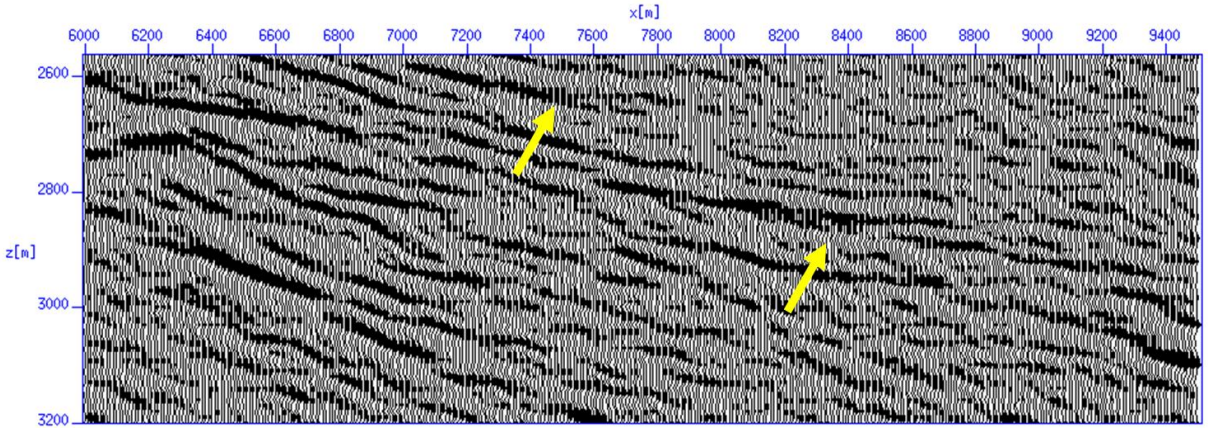
Figure 86: 2DFT amplitude spectrum of the migrated and deconvolved images: (a) Migrated image; (b) After the 2D deconvolution and (c) After the 2D deconvolution followed by 2D filtering.

3.5.3 Field Data Example

In this subsection, we show a preliminary result of the use of the 2D deconvolution filter followed by the tapered wavenumber mask in seismic sections obtained from processed field data. In Figure 87(a), there is a section obtained from a migration of a 1650m CO section. In Figure 87(b) we show the 2D deconvolved section. We observe that the resolution at the wedges pointed by the yellow arrows was enhanced as it became possible to differentiate the upper and lower reflectors in a larger portion of these wedges.



(a) migrated section



(b) 2D deconvolved section

Figure 87: 2D deconvolution of a 2D CO (1650m) migrated section.

3.6 CONCLUSION

In this chapter, we presented the results from [TAKAHATA *et al.*, 2013] in more detail and showed our contribution to the enhancement of resolution of PSDM images which consisted in the use of 2D filtering approach derived from the PSFs with the use of a 2D spiking deconvolution approach. We started by reviewing, following [LECOMTE and GELIUS, 1998; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002], how the concept of resolution function can be used to describe how limitations in the acquisition geometry and frequency bandwidth and complexities in the velocity model affect the resolution of seismic sections generated by PSDM. A seismic section obtained by PSDM can be seen as a version of the true model of the subsurface that has been distorted or blurred by the resolution function. Further, under proper assumptions, these resolution functions can be interpreted as PSFs as used in the context of image processing.

Next, we described and implemented a method from the referred literature for calculating these PSFs with the use of illumination vectors and scattering wavenumber vectors, using a ray-tracing algorithm. Also, we implemented the diffraction stack algorithm with the same ray-tracing algorithm in order to produce PSDM sections. In order to test these algorithms, we performed tests in a single scatterer scenario. In these tests, we observed that the PSFs successfully approximated the distortions caused by the PSDM. In the single trace configuration case, we observed that the PSFs approximated successfully the impulse response produced by the PSDM algorithm within small regions selected around points of the respective isochrone. It was observed that the space-varying PSFs could follow the changes of direction and resolution caused by the change of illumination along the isochrone. In the tests with gathers of traces, it was observed that the PSFs could successfully approximate the result of the PSDM around the position of the point scatterer.

After these tests, we presented our main contribution, in which we proposed to use a filtering approach to perform 2D spiking deconvolution, in contrast to previous works in literature such as [GELIUS *et al.*, 2002; SJOEBERG *et al.*, 2003], where an inversion approach was used. It is important to notice that, in these previous works, only a small section around the vertical section (pseudo 2D) of the PSF was used, as, otherwise, the method would become unstable. Thus, our first contribution allowed the use of the whole PSF for deconvolution. The 2D spiking deconvolution technique was applied to various types of migrated synthetic data. It was shown

for a homogeneous medium case that the use of 2D deconvolution improves the lateral resolution of migrated images. In fact, the results were comparable to those obtained by methods of high resolution such as MUSIC when it comes to separation of two nearby point scatterers (diffraction limited case) [GELIUS *et al.*, 2013]. The results for a layered fault model demonstrated that the method is also able to enhance the resolution of reflectors. However, we observed that already existing migration artifacts can possibly be further enhanced by the deconvolution. These artifacts arise if null components of the 2D spectrum of the PSF, which correspond to frequencies and illumination directions that do not exist on the data, are not considered. Thus, our second contribution to the method was a refinement step with the implementation of a 2D filter based on a 2D tapered wavenumber mask to attenuate the effect of the artifacts. This approach was also applied to migrated field data, in which the resolution of the image was enhanced.

4 CONCLUSIONS AND PERSPECTIVES

In this thesis, we explored the deconvolution problem from the unidimensional and bidimensional perspectives. In Chapter 2, we focused on the unidimensional problem. We reviewed the basic theory of supervised deconvolution and wavelet estimation exploring LS and Wiener filtering and also the theoretical aspects of unsupervised deconvolution, especially regarding the well-established prediction error filtering [ROBINSON, 1954], which uses SOS and assumes white reflectivity and minimum phase wavelet. We also described a method that uses HOS called B-ICA [KAPLAN and ULRYCH, 2003; KAPLAN, 2003], which assumes a non-Gaussian i.i.d. random reflectivity and does not impose any requirements on the phase of the wavelet. B-ICA provides both a wavelet estimate and a deconvolved reflectivity. Our contribution was to show that better performance can be achieved if, instead of using the deconvolved output of B-ICA, we use its wavelet estimate to calculate an inverse filter, as shown in [TAKAHATA *et al.*, 2012].

After the analysis of the theoretical aspects of deconvolution techniques, we made a case study using logs obtained from well measurements and migrated field data. From the statistical characterization of the reflectivity estimated from the well measurements, we concluded that its amplitudes have a non-Gaussian distribution, and the reflectivity is not perfectly white (and thus also not perfectly i.i.d.) as required by the unsupervised algorithms. The use of supervised methods allowed us to verify that the wavelet was mixed phase in the chosen time window, as opposed to the minimum phase assumption for prediction error filtering. This information was used to create different synthetic scenarios regarding the phase of the wavelet and the statistics of the wavelet, allowing us to create traces that met the hypothesis of the deconvolution algorithms and also traces closer to real situations. The well data was also useful to validate the outputs of the unsupervised deconvolution performed on the trace of the migrated section corresponding to the well position.

Tests with this set of synthetic data and traces from the migrated field data confirmed that the use of PEF is ineffective when the wavelet is mixed phase, but we also observed that it has some robustness for non-idealities of the reflectivity statistics when the wavelet is minimum phase. On the other hand, the B-ICA was effective while deconvolving minimum and mixed phase wavelets for i.i.d. reflectivities, but it was ineffective for the non-i.i.d. realistic reflectivity. However, the wavelet estimation with the use of B-ICA seemed to possess some robustness to

this non-ideality as the use of additional step proposed by us performed well in all synthetic scenarios, as well as in the migrated field data.

Next steps of research could involve further tests on the robustness of the B-ICA to the existence of correlation between the samples of the reflectivity, as results in 2.4 showed that an actual reflectivity is not perfectly white. Also, more detailed analysis on the effects of the data rearrangement in Step 1 of the algorithm and on the choice of the wavelet and reflectivity pair in Step 5 could deepen the understanding of the method. Moreover, understanding the implications of the choice of the technique used to perform the ICA step (Step 4) on the result of deconvolution could also lead to improvements. Furthermore, the use of sparse component analysis (SCA), e.g., [GRIBONVAL and LESAGE, 2006], instead of ICA could also lead to new insights into the approach of joining BSS based approaches and deconvolution. In addition, tests in data produced in different stages of processing should be considered. Finally, tests in data acquired from different geological sites, including data acquired in Brazil, should be considered.

In Chapter 3, we focused on bidimensional deconvolution. As opposed to the unidimensional approach adopted in Chapter 2, where the information of a single trace or very limited information on neighboring traces may be used and the resolution only on the vertical direction is considered, the bidimensional approach considers the quantification of distortions, also in the horizontal direction, i.e., between neighboring traces, and allows the enhancement of both lateral and vertical resolution of seismic images. Following the approach in [LECOMTE and GELIUS, 1998; GELIUS and LECOMTE, 2000; GELIUS *et al.*, 2002], we explored the concepts of resolution function, scattering wavenumber vector and illumination vector to calculate PSFs, which models the blur in PSDM sections caused by the limitations in the geometry acquisition and the geological complexities, in addition to the band limitation of the wavelet, which is the only factor considered in the unidimensional approach.

Our contribution to the use of PSFs in bidimensional deconvolution was to model and use an LS spiking deconvolution approach [TAKAHATA *et al.*, 2013], as opposed to the use of direct inversion approaches such as in [GELIUS *et al.*, 2002; SJOEBERG *et al.*, 2003]. These previous approaches were limited by the fact the use of the whole PSF would lead to instabilities and thus a pseudo 2D approach was used where only the central region around the vertical axis of the PSF was used. In fact, it was also observed that a naïve use of the spiking deconvolution approach would also lead to instabilities. An analysis on the 2DFT amplitude spectrum of the PSF

allowed the discovery of spectral nulls caused by inexistent frequencies and illumination directions in the data, and that these components were enhancing noise on the deconvolved image. A refinement with a tapered wavenumber mask filter was then proposed in order to eliminate these spurious components. Tests in controlled and field data showed that this method was able to increase the resolution of PSDM sections, especially in the horizontal direction.

Further improvements in the method will consist in enhancing the ability to increase the resolution of PSDM sections while mitigating artifacts. In order to achieve this, regularization methods for the inversion approach could be elaborated based on a further analysis on the null components of the 2D spectrum of the PSFs. Moreover, regularization methods exploring sparsity in the Wavelet domain [BELGE *et al.*, 2000] or in the Curvelet domain [KUMAR and HERRMANN, 2008] could also be explored. Also, in this thesis, it was assumed that the PSFs were locally space invariant. This is actually an approximation, valid for small areas with similar illumination. However, larger areas or more complex geology may imply in changes of illumination in the area of interest, which, as consequence, lead to variations in the PSFs. Thus, adaptive filtering methods or inverse methods which take these variations into account should also be analyzed. Finally, the PSFs method assumes that the seismic image is formed by the response of point scatterers. As seismic data contains predominantly reflections, a reflector spread function, as advocated in [GELIUS *et al.*, 2002], instead of the PSF, may also improve the results of deconvolution. Finally, an extension for processing 3D data should be considered. 3D seismic data may be obtained by performing the acquisition over a 2D area at the surface instead of a single line as in the 2D case. The processing of this type of data leads to 3D data volumes which can be used to produce 3D models of the earth [YILMAZ, 2001]. The extension of the method would then start with the implementation of a method to estimate 3D PSFs by considering 3D illumination vectors by using the source and receiver positions in the acquisition surface and using 3D ray tracing in a three-dimensional velocity model. After that, the LS spiking deconvolution filtering approach may be applied by using a 3D convolution matrix that is similar to the 2D matrix presented in (76). The use of 3D deconvolution would then allow the increase of resolution in two horizontal directions, which could enhance the ability to visualize features in horizontal sections of seismic volumes.

5 REFERENCES

ALDRIDGE, D. F. Short Note: The Berlage Wavelet. **Geophysics**, v. 55, n. 11, p. 1508-1511, 1990.

ANP. ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. **Notas à Imprensa 2013 - Maio**, 14 May 2013. Available at: <http://www.anp.gov.br/?pg=65961>. Accessed on: 27 August 2013.

BANHAM, M. R.; KATSAGGELOS, A. K. Digital Image Restoration, v. 14, n. 2, p. 24-41, 1997.

BELGE, M.; KILMER, M. E.; MILLER, E. L. Wavelet Domain Image Restoration with Adaptive Edge-Preserving Regularization. **IEEE Transactions on Image Processing**, v. 9, n. 4, p. 597-608, 2000.

BENVENISTE, A.; GOURSAT, M.; RUGET, G. Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications. **IEEE Transactions on Automatic Control**, v. 25, n. 3, p. 385 - 399, 1980.

BERCHER, J.-F.; VIGNAT, C. Estimating the Entropy of a Signal with Applications. **IEEE Transactions on Signal Processing**, v. 48, n. 6, p. 1687-1694, 2000.

BLEISTEIN, N. **Mathematical Methods for Wave Phenomena**. Orlando, FL: Academic Press, 1984.

ČERVENý, V. **Seismic Ray Theory**. Cambridge, UK: Cambridge University Press, 2001.

CLAERBOUT, J. Minimum Information Deconvolution. **Stanford Exploration Project**, v. 15, p. 109-122, 1978.

COMON, P. Independent Component Analysis, a new concept? **Signal Processing**, v. 36, p. 287-314, 1994.

COMON, P.; JUTTEN, C. (Eds.). **Handbook of Blind Source Separation:** Independent Component Analysis and Applications. Burlington: Academic Press, 2010.

DONOHO, D. On Minimum Entropy Deconvolution. In: FINDLEY, D. **Applied Time Series Analysis II**. [S.l.]: Academic Press, 1981. p. 565-608.

DURBIN, J. The Fitting of Time Series Models. **Review of the International Statistical Institute**, v. 28, p. 233-244, 1960.

EDGAR, J. A.; VAN DER BANN, M. How Reliable is Statistical Wavelet Estimation? **Geophysics**, v. 76, n. 4, p. V59-V68, July-August 2011.

FOLHA DE S.PAULO. Folha de S. Paulo. **Derretimento da OGX provoca mudanças no Ibovespa**, 10 August 2013. Available at: <http://www1.folha.uol.com.br/mercado/2013/08/1324630-derretimento-da-ogx-provoca-mudancas-no-indice-bovespa.shtml>. Accessed on: 27 August 2013.

FRENCH, W. S. Two Dimensional and Three Dimensional Migration of Model-Experiment Reflection Profiles. **Geophysics**, v. 39, n. 3, p. 265-277, 1974.

GAUSS, C. F. **Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium**. Hamburg: F. Perthes & I.H. Besser, 1809.

GELIUS, L.-J.; LECOMTE, I. The Resolution Function in Linearized Born and Kirchhoff Inversion. In: HANSEN, P.C.; JACOBSEN, B. H.; MOSEGAARD, K. **Lecture Notes in Earth Sciences:** Methods and Applications of Inversion: Springer Berlin Heidelberg, v. 92, 2000. p. 129-141.

GELIUS, L.-J.; LECOMTE, I.; TABTI, H. Analysis of the Resolution Function in Seismic Prestack Imaging. **Geophysical Prospecting**, v. 50, p. 505-515, 2002.

GELIUS, L.-J.; TYGEL, M.; TAKAHATA, A. K.; ASGEDOM, E. G.; SERRANO, D. R. High-Resolution Imaging of Diffractions — A Window-Steered MUSIC Approach. **Geophysics**, v. 78, n. 6, p. S255-S264, 2013.

GEROMEL, J. C.; PALHARES, A. G. B. **Análise Linear de Sistemas Dinâmicos:** Teoria, Ensaios Práicos e Exercícios. São Paulo, SP: Edgard Blücher, 2004.

GODFREY, B. An Information Theory Approach to Deconvolution. **Stanford Exploration Project**, v. 15, p. 157-182, 1978.

GRAY, W. C. **Variable Norm Deconvolution: PhD thesis**. Stanford University. 1979.

GRIBONVAL, R.; LESAGE, S. **A Survey of Sparse Component Analysis for Blind Source Separation:** Principles, Perspectives and New Challenges. ESANN'2006 Proceedings - European Symposium on Artificial Neural Networks. Bruges, Belgium: 2006. p. 323-330.

HARGREAVES, N. **Wavelet Estimation via Fourth-Order Cumulants**. SEG Technical Program Expanded Abstracts 1994. p. 1588-1590.

HAYKIN, S. O. **Adaptive Filter Theory**. 4th. ed. Englewood Cliffs, NJ: Prentice Hall, 2001.

HOSKEN, J. W. J. Ricker Wavelets in Their Various Guises. **First Break**, v. 6, n. 1, p. 24-33, 1988.

HU, J.; SCHUSTER, G. T.; VALASEK, P. A. Poststack Migration Deconvolution. **Geophysics**, v. 66, n. 3, 2001.

HUBRAL, P.; HÖCHT, G.; JÄGER, R. **An Introduction to the Common Reflection Surface Stack**. Extended abstracts book: EAGE 60th Conference and Technical Exhibition. Leipzig, Germany: 1998.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. **Independent Component Analysis**. New York, NY: John Wiley & Sons, 2001.

JONES, I. F. Tutorial: Velocity Estimation via Ray-Based Tomography. **First Break**, v. 28, n. 2, p. 45-52, 2010.

KAGAN, A. M.; LINNIK, Y. V.; RAO, C. R. **Characterization Problems in Mathematics**. New York, NY: John Wiley and Sons, 1973.

KALLWEIT, R. S.; WOOD, L. C. The Limits of Resolution of Zero-Phase Wavelets. **Geophysics**, v. 47, n. 4, p. 1035-1046, 1982.

KAPLAN, S. T. **Principal and Independent Component Analysis for Seismic Data, M.Sc. Dissertation**. Department of Earth and Ocean Sciences, The Faculty of Graduate Studies, UBC. Vancouver, BC. 2003.

KAPLAN, S. T.; ULRYCH, T. J. **Blind Deconvolution and ICA with a Banded Mixing Matrix**. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation. Nara: 2003.

KOLMOGOROV, A. N. Sur l'Interpretation et l'Extrapolation des Suites Stationnaires. **Comptes Rendus de l'Académie des Sciences**, v. 208, p. 2043-2045, 1939.

KUMAR, V.; HERRMANN, F. J. **Deconvolution with Curvelet-Domain Sparsity**. SEG Technical Program Expanded Abstracts 2008. Las Vegas, NV: 2008. p. 1996-2000.

LAZEAR, G. D. Mixed-Phase Wavelet Estimation Using Fourth-Order Cumulants. **Geophysics**, v. 58, n. 7, p. 1042-1051, 1993.

LECOMTE, I. Resolution and Illumination Analyses in PSDM: A Ray-Based Approach. **The Leading Edge**, May 2008. 650-663.

LECOMTE, I.; GELIUS, L.-J. **Have a look at the resolution of prestack depth migration for any model, survey and wavefields.** SEG Technical Program Expanded Abstracts. 1998. p. 1112-1115.

LECOMTE, I.; POCHON-GUERIN, L. **Simulated 2D/3D PSDM images with a fast, robust, and flexible FFT-based filtering approach**. SEG Technical Program Expanded Abstracts 2005. Houston, TX. 2005. p. 1810-1813.

LEVINSON, N. The Wiener RMS Error Criterion in Filter Design and Prediction. **Journal of Mathematical Physics**, v. 25, p. 261-278, 1947.

MENDEL, J. M. Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications. **Proceedings of the IEEE** , v. 79, n. 3, p. 278 - 305, 1991.

MILLER, D.; ORISTAGLIO, M.; BEYLKIN, G. A New Slant on Seismic Imaging: Migration and Integral Geometry. **Geophysics**, v. 52, n. 7, p. 943-964, 1987.

MISRA, S.; CHOPRA, S. Phase Stability via Nonlinear Optimization: A Case Study. **The Leading Edge**, v. 29, n. 11, p. 1338-1343, 2010.

MISRA, S.; SACCHI, M. D. Non-Minimum Phase Wavelet Estimation by Non-Linear Optimization. **Geophysical Prospecting**, v. 55, n. 2, p. 223-234, 2007.

MOUSA, W. A. Seismic Migration: A Digital Filtering Process Reducing Oil Exploration Risks. **IEEE Signal Processing Magazine**, v. 29, n. 3, p. 117-123, 2012.

NEELAMANI, R. Deconvolution and Optimal Filtering in Seismology. In: HAVELOCK, D.; KUWANO, S.; VORLÄNDER, M. **Handbook of Signal Processing in Acoustics**. New York: Springer, v. 2, 2008. Cap. 87, p. 1571-1583.

NEMETH, T.; WU, C.; SCHUSTER, G. T. Least-Squares Migration of Incomplete Reflection Data. **Geophysics**, v. 64, n. 1, 1999.

NIKIAS, C. L.; MENDEL, J. M. Signal Processing with Higher-Order Spectra. **IEEE Signal Processing Magazine**, v. 10, n. 3, p. 10-37, 1993.

NIKIAS, C. L.; RAGHUVEER, M. R. Bispectrum Estimation: A Digital Signal Processing Framework. **Proceedings of the IEEE**, v. 75, n. 7, p. 869-891, 1987.

OOE, M.; ULRYCH, T. J. Minimum Entropy Deconvolution With an Exponential Transformation. **Geophysical Prospecting**, v. 27, n. 2, p. 458-473, 1979.

OPPENHEIM, A. V.; SCHAFER, R. W. **Discrete-Time Signal Processing**. Englewood Cliffs, NJ: Prentice-Hall, 1989.

PODVIN, P. Ecole des Mines de Paris. **FDTIMES:** Computation of seismic travel times by finite-differences, 2006. Available at: <http://www.geophy.mines-paristech.fr/soft/fdtimes/html_doc/index.html>. Accessed on: 01 February 2014.

PODVIN, P.; LECOMTE, I. Finite Difference Computation of Traveltimes in Very Contrasted Velocity Models: a Massively Parallel Approach and its Associated Tools. **Geophysical Journal International**, v. 105, n. 1, p. 271-284, 1991.

RICKER, N. Wavelet Functions and Their Polynomials. **Geophysics**, v. 9, n. 3, p. 314–323, 1944.

ROBINSON, E. A. **Predictive Decomposition of Time Series with Applications to Seismic Exploration**. Department of Geology and Geophysics, MIT. Cambridge. 1954.

ROBINSON, E. A.; OSMAN, O. M. Introduction to deconvolution. In: E.A. ROBINSON, O. M. O. **Deconvolution 2**. Tulsa, OK: Society of Exploration Geophysicists, 1996. Cap. 1, p. 726.

ROBINSON, E. A.; TREITEL,. **Geophysical Signal Analysis**. Englewood Cliffs, NJ: Prentice Hall, 1980.

ROMANO, J. M. T.; ATTUX, R.; CAVALCANTE, C. C.; SUYAMA, R. **Unsupervised Signal Processing**. Boca Raton: CRC Press, 2010.

SACCHI, M. D.; ULRYCH, T. J. Nonminimum-Phase Wavelet Estimation Using Higher Order Statistics. **The Leading Edge**, v. 19, n. 1, p. 80-83, 2000.

SENGBUSH, R. L.; LAWRENCE, P. L.; MCDONAL, F. J. Interpretation of Synthetic Seismograms. **Geophysics**, v. 26, n. 2, p. 138-157, April 1961.

SHALVI, O.; WEINSTEIN, E. New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels). **IEEE Transactions on Information Theory**, v. 36, n. 2, p. 312-321, 1990.

SHERIFF, R. E. **Encyclopedic Dictionary of Applied Geophysics**. 4th. ed. Tulsa, OK: Society of Exploration Geophysicists, 2002.

SIMM, R.; WHITE, R. Phase, Polarity and the Interpreter's Wavelet. **First Break**, v. 20, n. 5, 2002.

SJOEBERG, T. A.; GELIUS, L.-J.; LECOMTE, I. **2-D deconvolution of seismic image blur**. SEG Technical Program Expanded Abstracts. Dallas, TX. 2003. p. 1055-1058.

TAKAHATA, A. K.; GELIUS, L.-J.; LOPES, R. R.; TYGEL, M.; LECOMTE, I. **2D Spiking Deconvolution Approach to Resolution Enhancement of Prestack Depth Migrated Seismic Images**. Extended abstracts book: 75th EAGE Conference and Technical Exhibition. London, UK: [s.n.]. 2013. p. Th_06_07.

TAKAHATA, A. K.; NADALIN, E. Z.; FERRARI, R.; DUARTE, L. T.; SUYAMA, R.; ROMANO, J. M. T.; TYGEL, M. Unsupervised Processing of Geophysical Signals. **IEEE Signal Processing Magazine**, v. 29, n. 4, p. 27-35, July 2012.

TANG, Y. Target-Oriented Wave-Equation Least-Squares Migration/Inversion with Phase-Encoded Hessian. **Geophysics**, v. 74, n. 6, 2009.

THE ECONOMIST. The Economist. **Brazil's oilfields, Back in business**, 18 May 2013. Available at: <http://www.economist.com/news/business/21578095-strong-bidding-exploration-rights-ends-industrys-long-dry-spell-back-business>. Accessed on: 27 August 2013.

THE ECONOMIST. The Economist. **The empire of Eike Batista, Eike's breaky heart**, 29 June 2013b. Available at: <http://www.economist.com/news/business/21580160-can-brazils-best-known-businessman-save-his-crumbling-empire-eikes-breaky-heart>. Acessed on: 31 January 2014.

TUGNAIT, J. K. Identification of Linear, Stochastic Systems via Second- and Fourth-Order Cumulant Matching. **IEEE Transactions on Information Theory**, v. IT-33, n. 3, p. 393-407, 1987.

TUGNAIT, J. K. Comments on “New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)". **IEEE Transactions on Information Theory**, v. 38, n. 1, p. 210-213, 1992.

VELIS, D. R.; ULRYCH, T. J. Simulated Annealing Wavelet Estimation via Fourth-Order Cumulant Matching. **Geophysics**, v. 61, n. 6, p. 1939-1948, 1996.

VERSCHUUR, D. J. **Seismic Multiple Removal Techniques:** Past, Present and Future. Houten, Netherlands: EAGE Publications, 2006.

WALDEN, A. T. Non-Gaussian Reflectivity, Entropy, and Deconvolution. **Geophysics**, v. 50, n. 12, p. 2862-2888, 1985.

WHITE, R.; SIMM, R. Tutorial: Good Practice in Well Ties. **First Break**, v. 21, n. 10, p. 75 - 83, 2003.

WIENER, N. **Extrapolation, Interpolation, and Smoothing of Stationary Time Series**. New York, NY: Wiley, 1949.

WIENER, N.; HOPF, E. **Über eine Klasse Singulärer Integralgleichungen**. Berlin, Germany: Sitz. Berlin. Akad. Wiss., 1931. 696-706 p.

WIGGINS, R. A. Minimum Entropy Deconvolution. **Geoexploration**, v. 16, n. 1-2, p. 21–35, 1978.

YILMAZ, Ö. **Seismic Data Analysis**. Society of Exploration Geophysicists, 2001.

YU, J.; HU, J.; SCHUSTER, G. T.; ESTILL, R. Prestack Migration Deconvolution. **Geophysics**, v. 71, n. 2, p. S53-S62, 2006.

YU, J.; SCHUSTER, G. T. **Migration Deconvolution vs. Least Squares Migration**. SEG Technical Program Expanded Abstracts 2003. Dallas, TX. 2003. p. 1047-1050.

ZIOLKOWSKI, A. Why Don't We Measure Seismic Signatures? **Geophysics**, v. 56, n. 2, p. 190-201, 1991.

ZÖLZER, U. **Digital Audio Signal Processing**. Chichester, UK: John Wiley & Sons, 2008.

APPENDIX A: PROOF OF EQUIVALENCE BETWEEN LS AND WIENER FILTERS FOR WHITE REFLECTIVITY

Here we show that, as mentioned in Chapter 2, if we consider the reflectivity, $s(n)$, to be a white signal and we neglect the noise, then, the Wiener filter, $w_W(n)$, is actually the same as the zero lag spiking deconvolution LS filter, $w_{LS}(n)$.

From (7), neglecting the noise, and (8), we have that

$$x(n-m) = \sum_{q=0}^{N_h-1} h(k)s(n-m-q). \tag{77}$$

If we make a variable change so that

$$q' = q + m,$$

then we have

$$x(n-m) = \sum_{q'=m}^{N_h+m-1} h(q'-m)s(n-q') \ .$$

As $h(q'-m) = 0$ if $q' < m$ or $q' > Q + m - 1$, then we have

$$x(n-m) = \sum_{q'=0}^{N_h+N_w-1} h(q'-m)s(n-q') \ .$$

Thus, if we define as

$$\mathbf{s}(n) = [s(n) \quad s(n-1) \quad \ldots \quad s(n-N_h-N_w+1)]^T$$
$$\mathbf{h}(m) = [h(-m) \quad h(1-m) \quad \ldots \quad h(N_h+N_w-1-m)]^T,$$

then we can express $x(n-m)$ as a product of these vectors so that

$$x(n-m) = \mathbf{h}^T(m)\mathbf{s}(n).$$

We can substitute this into (11) and thus we have

$$\mathbf{x}(n) = \mathbf{H}_W\mathbf{s}(n), \tag{78}$$

where $\mathbf{H}_W$ is a $N_w \times (N_h + N_w - 1)$ Toeplitz matrix so that

$$\mathbf{H}_W = \begin{bmatrix} \mathbf{h}^T(0) \\ \mathbf{h}^T(1) \\ \vdots \\ \mathbf{h}^T(M-1) \end{bmatrix}.$$

As the sample in the *m*-th line and *q'*-th column of $\mathbf{H}_W$ are, respectively, linked to the *m*-th delay of *x*(*n*) and *q'*-th delay of *s*(*n*), then, from (10), we obtain that

$$[\,\mathbf{H}_W]_{i,j} = h(j-i),\, i = 0, \ldots, N_w - 1, j = 0, \ldots, N_h + N_w - 1, \tag{79}$$

where $[\mathbf{H}_W]_{i,j}$ is the element of $\mathbf{H}_W$ located at the *i*-th line and *j*-th column of the matrix,

$$\mathbf{H}_W = \begin{bmatrix} h(0) & h(1) & \cdots & h(N_h-1) & 0 & 0 & \ldots & 0 \\ 0 & h(0) & h(1) & \cdots & h(N_h-1) & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & h(0) & h(1) & \ldots & h(N_h-1) \end{bmatrix}, \tag{80}$$

From (78) and (18), we now calculate the autocorrelation matrix:

$$\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}(n)^T] = E[\mathbf{H}_W\mathbf{s}(n)\mathbf{s}(n)^T\mathbf{H}_w^T] = \sigma_S^2\mathbf{H}_W\mathbf{H}_w^T, \tag{81}$$

where σ_S^2 is the variance of $s(n)$, as $s(n)$ is a white signal and therefore $E[\mathbf{s}(n)\mathbf{s}(n)^T] = \sigma_S^2\mathbf{I}$, as consequence from (25). Also from (25), it's possible to verify that

$$E[\mathbf{s}(n)s(n)] = E\{[s(n) \quad s(n-1) \quad \ldots \quad s(n-Q-M+1)]^T \cdot s(n)\} = \boldsymbol{\delta},$$

where $\boldsymbol{\delta}$ is a $(N_h + N_w - 1) \times 1$ vector so that

$$\boldsymbol{\delta} = [1 \quad 0 \quad 0 \quad \ldots \quad 0]^T.$$

Thus, as the desired signal is $s(n)$, the cross-correlation is given by

$$\mathbf{p} = E[\mathbf{x}(n)s(n)] = E[\mathbf{H}_W\mathbf{s}(n)s(n)] = \sigma_S^2\mathbf{H}_W\boldsymbol{\delta}. \tag{82}$$

By substituting (81) and (82) in (23) we have then:

$$\mathbf{w}_W = (\mathbf{H}_W\mathbf{H}_w^T)^{-1}\mathbf{H}_W\boldsymbol{\delta}. \tag{83}$$

By comparing (29) and (30) to (79) and (80), then we observe that

$$\mathbf{H}_W = (\mathbf{H}_{LS})^T. \tag{84}$$

Thus, if we compare (28) and (83), we have, as a consequence in the case of white reflectivity, that actually

$$\mathbf{w}_W = \mathbf{w}_{LS}.$$

Therefore, this shows that if the seismic trace is processed with the LS filter, the error of the output in respect to the reflectivity will be minimal in the MMSE sense if the reflectivity is white.

Also, if we consider a unit variance for $s(n)$, then we have from (81) and (84) the expressions of autocorrelation and cross-correlation in terms of $\mathbf{H}_{LS}$:

$$\mathbf{R} = (\mathbf{H}_{LS})^T\mathbf{H}_{LS},$$

$$\mathbf{p} = (\mathbf{H}_{LS})^T\,\boldsymbol{\delta}.$$