



# Multichannel Semi-blind Deconvolution (MSBD) of seismic signals



Merabi Mirel\*, Israel Cohen

Technion – Israel Institute of Technology, 3200003 Haifa, Israel

## ARTICLE INFO

### Keywords:

Seismic deconvolution  
Semi-blind  
Wavelet estimation  
Reflectivity recovery  
Sparse deconvolution  
Multichannel deconvolution

## ABSTRACT

Seismic deconvolution is a general problem associated with recovering the reflectivity series from a seismic signal when the wavelet is known. In this paper, we solve the problem of semi-blind seismic deconvolution, where the wavelet is known up to some error. The Multichannel Semi-blind Deconvolution (MSBD) model was developed for cases where there is some uncertainty in the assumed wavelet. We present a novel, two-stage iterative algorithm that recovers both the reflectivity and the wavelet. While the reflectivity series is recovered using sparse modeling of the signal, the wavelet is recovered using L2 minimization, exploiting the fact that all channels share the same wavelet. The L2 minimization solution is revised to suit the multichannel case. An analysis is made for each wavelet uncertainty according to the parameters of the respective recovery method. We show that our algorithm outperforms the straightforward method of assuming the initial wavelet. As a side result, we also show that the final estimated wavelet fits the true wavelet better than the initial one.

## 1. Introduction

The deconvolution problem and the kernel estimation problem are two problems common to many fields, including engineering, physics and others. Different approaches for solving the problems can be found in the literature depending on the specific problem, the a priori knowledge and the different assumptions made about the signals in the problem. The basic idea is that a signal goes through a linear system (defined by the kernel), the output of the linear system is contaminated by some noise and the goal is to recover the kernel and the input signal. Kernel estimation problems assume to know the signal and aim to find the kernel, while deconvolution problems assume to know the kernel and aim to find the signal.

Our discussion is on seismic signals. An interesting way of modeling can be as follows: A series of impulses are generated in the underground layers of the earth. This series goes through the earth until it is received on the surface by an array of seismic sensors. The kernel defining the channel traversed by the impulse series is called the wavelet, which is defined by the seismic source. This kind of modeling in the literature is used often as a convenient approach to seismic modeling, for example in [1]. The recorded data, in the form of seismic traces, are analyzed, and interesting parameters are extracted to improve understanding of the layer structure, channel modeling in that particular area and so on. In some cases it is also common to transmit a very short (in the time domain) pulse from the surface, let it traverse the earth channel, reflect off one of the layers and return to the surface.

Different approaches and models of the same problem can lead to different performances. When developing a solution for a problem in this field, one of the first things to determine is whether the solution is going to be based on a stochastic or deterministic model as well as defining the parameters. Our method will focus on deterministic modeling of the problem, but first we review different methods from both disciplines with greater emphasis on deterministic modeling methods.

First we briefly describe the different stochastic modeling methods. Kormylo and Mendel (1982) estimate the wavelet using ARMA (Auto-Regressive Moving Average) and SMLR (Single Most Likely Replacement) for the reflectivity estimation while assuming a BG (Bernoulli–Gaussian) model [2]. Both algorithms use second-order statistics. Kaaresen and Taxt (1998) also assume the BG model but they use the IWM (Iterated Window Maximization) algorithm for reflectivity estimation and the Least-Squares method for wavelet estimation [3]. Heimer et al. (2009) proposed a blind multichannel deconvolution method based on the statistical properties of the signal [4]. Specifically, it is based on the Markov–Bernoulli random field modeling. Their method accounts for layer discontinuities resulting from splitting, merging, starting or terminating layers within the region of interest. Ram et al. (2010) also propose a method based on the statistical properties of the signal [5], where the spatial dependency between neighboring traces is exploited by a priori assuming 2D reflectivity. The algorithm is based on the MBG (Markov Bernoulli–Gaussian) reflectivity model.

Methods using higher order statistics are also very important to

\* Corresponding author.

E-mail addresses: [merabi.mirel@gmail.com](mailto:merabi.mirel@gmail.com) (M. Mirel), [icohen@ee.technion.ac.il](mailto:icohen@ee.technion.ac.il) (I. Cohen).

mention. For example, Velis and Ulrych (1999) propose a fourth-order cumulant matching method [6]. They use the mean-square error as a measure for matching between the trace cumulant and the wavelet moment. Van der Baan (2009) also uses a higher order statistics for wavelet estimation and deconvolution [7]. He maximizes the kurtosis measure for phase estimation in cases of phase mismatches. Menanno and Mazzotti (2012) use the vectorial nature of the wavefield [8] and exploit correlation between different channels in the multichannel case in a method of quaternion deconvolution. Their method is an extension of the Wiener filter and uses high order statistics.

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

A modification of the Euclid Deconvolution is presented by Kazemi and Sacchi (2014), who developed the SMBD (Sparse Multichannel Blind Deconvolution) method [9]. They exploit the sparsity property in the reflectivity series to make the Euclid Deconvolution more robust to noise and generally improve the method. In practice, they define an L2 minimization problem with a regularization term and a requirement on the sparsity of the signal.

Repetti et al. (2015) [10] present in a new study a different approach to the L1/L2 minimization problem for the blind deconvolution case. They propose a new penalty, Smoothed One-Over-Two (SOOT), based on a smooth approximation of the L1/L2 function. The SOOT penalty enables them to avoid the problems raised from solving non-convex and non-smooth optimization problems. They develop a proximal-based algorithm to solve the minimization problem and derive theoretical convergence. Furthermore, there are strong connections between blind deconvolution and blind compressed sensing, which presents the basic idea of recovering a sparse signal from a small number of linear measurements. Blind compressed sensing is widely studied and discussed by Gleichman and Eldar (2011) [11]. They suggest different constraints on the sparsity basis which allow them to guarantee a unique solution, while avoiding the need of prior knowledge of the sparsity basis, which is essential for the recovery process. They introduce a general sampling and reconstruction process which can be suited for all the signals that have a sparse representation and are under the conditions and restraints presented in their work. Another interesting work on blind compressed sensing is introduced by Rosenbaum and Tsybakov (2010) [12], which also account for matrix uncertainty. In their work they suggest new estimators, since they find previous estimators as unstable. Their main conclusion of choosing smartly the regularization parameter as a key to success recovery will support our findings on the regularization parameters later on.

Nguyen and Castagna (2010) presented a method that exploits the sparse properties of the reflectivity series using Matching Pursuit Decomposition (MPD) [13]. MPD involves a few steps and eventually decomposes the seismic data into a superposition of wavelet atoms generated from the locations, amplitudes and scaling (physically translated into different center frequencies) of a base wavelet form. The method correlates a wavelet dictionary with the data and marks the parameters iteratively, recording the best-fit wavelet in each iteration. It is very important to keep the wavelet dictionary orthogonal in this

case. MPD was also used by Wang [14,15].

Zhang and Castagna (2011) presented a method based on Basis Pursuit Decomposition (BPD) for seismic inversion [16]. They used an algorithm presented by Chen et al. (2001) [17,18] to solve the Basis Pursuit problem, which is an L1 optimization problem. They also utilized a special dictionary form. The special form results from the Dipole Decomposition process on the reflectivity series. Dipole Decomposition is a method that decomposes the reflectivity series into a summation of even and odd impulse pairs, weighted by different amplitudes. The BPD method is presented by Bork and Wood (2001) [19]. A brief summary of the method can be found in [16]. The basis pursuit algorithm is an L1 optimization problem that can be solved in different ways. The great advantage of BPD over MPD comes from the process used to solve the problems. MPD is an iterative process that extracts the best-fit wavelet atom, subtracts it from the data, finds the next best-fit atom, etc. The solution therefore depends on the order of the wavelet atoms in the wavelet dictionary, so different ordering of the same wavelet dictionary can lead to different solutions. Conversely, BPD is not affected by the ordering of the wavelet dictionary and obtains one solution for all the different combinations that eventually construct the same wavelet dictionary. BPD has more advantages over MPD, including interference handling, computational efficiency, and good stability even when the wavelet dictionary is not orthogonal. More information and analytical developments about Basis Pursuit and Matching Pursuit are widely discussed in [20].

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

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

As is shown later on in greater detail, we define the problem as a Multichannel Semi-blind Deconvolution problem. This means that we assume to have an array of sensors located close enough to each other, to get some correlation between different seismic traces, and we also assume to know a noisy version of the kernel. We combine methods of kernel estimation and deconvolution to solve this problem.

In the next sections of this paper we formulate the problem mathematically, step by step, present the classic well-known solutions for the generic problems and adapt these solutions to fit our specific problem. Finally, we use examples to test our method, compare it with

the classic solutions and discuss the results.

## 2. Problem formulation

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

$$s[n] = r[n]*w[n] + v[n] \quad (1)$$

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

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

$$s_i[n] = r_i[n]*w[n] + v_i[n], \quad 1 \leq i \leq N. \quad (2)$$

We can write (2) in the following vector–matrix form:

$$s_i = \mathbf{W}r_i + v_i, \quad 1 \leq i \leq N \quad (3)$$

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

The goal of the basic problem is to recover  $r_i$  from  $s_i$  while assuming full knowledge of  $\mathbf{W}$  and  $\sigma_v$ . This problem has been widely investigated and is called the deconvolution problem. A wide variety of solutions have been proposed to solve this problem, depending on the model of the signal  $r_i$ . In seismic deconvolution, assuming a sparse model for  $r_i$ , sparse deconvolution methods have been proposed. One of them is Basis Pursuit Denoising (BPDN) [20,26–28], which is an approach that solves the following optimization problem:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (4)$$

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

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

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s. t. } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 < \epsilon \quad (5)$$

where  $\epsilon$  controls the abovementioned trade-off.

In our case, we do not have full knowledge of  $\mathbf{A}$ . We have  $\mathbf{A}'$ , which is a noisy version of  $\mathbf{A}$  and holds the relation  $\mathbf{A} = \mathbf{A}' - \mathbf{A}_v$ .  $\mathbf{A}_v$  represents the uncertainty in  $\mathbf{A}$  and later on we address the problem with a specific definition of  $\mathbf{A}_v$ .

## 3. Multichannel Semi-blind Deconvolution – MSBD

Our purpose is to establish a general method for the Semi-blind Deconvolution problem and to specifically analyze two different cases of wavelet uncertainty as shown later. First, we introduce the general method. The method relies on the different modeling of each of the recovered signals. We know nothing about the reflectivity signal besides the fact that it is sparse, however we assume to know the wavelet up to some level of noise. The noise can be additive to the wavelet signal or intrinsic to one of the parameters that form the wavelet model, or any other noise that can be mathematically formulated. We assume non-sparse representation of the wavelet signal. This means that we use the standard deconvolution methods for the wavelet. In our case, we choose to work with the L2 minimiza-

tion method. This method was adapted to best fit our problem. The adaption that was made and the full form of the deconvolution are shown later. The method we propose is an iterative method, with two steps in each iteration, as follows:

1. Assume to know the wavelet and use the sparse deconvolution method to recover the reflectivity signal.
2. Assume to know the reflectivity and use the L2 minimization method to recover the wavelet.

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

We now introduce the general method of choosing the trade-off parameter when assuming additive noise to the wavelet and later on we demonstrate this method for specific cases. Denote by  $\mathbf{w} \in \mathbb{R}^{N_w}$  the vector representation of the signal  $w[n]$ ,  $1 \leq n \leq N_w$ . The additive noise to the wavelet is denoted by  $\mathbf{w}_v \in \mathbb{R}^{N_w}$  and the corresponding convolution matrix is  $\mathbf{W}_v \in \mathbb{R}^{(N_r+N_w-1) \times N_r}$ . In the same way we denote the initial wavelet we are given and its corresponding convolution matrix as  $\mathbf{w}' \in \mathbb{R}^{N_w}$  and  $\mathbf{W}' \in \mathbb{R}^{(N_r+N_w-1) \times N_r}$  so we get the relation,

$$\mathbf{W} = \mathbf{W}' - \mathbf{W}_v. \quad (6)$$

Substituting (6) into (3) we get,

$$s_i = \mathbf{W}r_i + v_i = (\mathbf{W}' - \mathbf{W}_v)r_i + v_i = \mathbf{W}'r_i + v_i - \mathbf{W}_v r_i. \quad (7)$$

Looking at this relation,

$$s_i = \mathbf{W}'r_i + v_i - \mathbf{W}_v r_i \quad (8)$$

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

$$v'_i = v_i - \mathbf{W}_v r_i. \quad (9)$$

For the first step we assume to know the wavelet and recover the reflectivity series. As mentioned before, this is done by applying the BPDN solution to our problem. As we see in (5), we have to choose the trade-off parameter,  $\epsilon$ , wisely. The literature does not prove, nor imply, a generic analysis for choosing this parameter, but as mentioned above (Section 2),  $\epsilon$  has a strong relation to the total standard deviation of the noise in the problem. In our case we decided to define the total standard deviation of the noise in two forms. Each form has its advantages and disadvantages.

The first form is as follows:

$$V_i = \sum_{j=1}^{N_r+N_w-1} v'_j[n] \quad (10)$$

$$\epsilon_i = \sigma_{V_i}. \quad (11)$$

The second form is:

$$\epsilon_i = \sqrt{\sum_{j=1}^{N_r+N_w-1} \sigma_{v'_j}^2}. \quad (12)$$

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

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

For the second step, we assume to know the reflectivity series and aim to recover the wavelet. Unlike the first step, here we cannot apply BPDN, or any other sparse deconvolution method for that matter. The simple reason is that the wavelet is not a sparse signal. We look at this problem from another point of view, dictionary learning. Dictionary learning is a broad field that can provide many insights on how to update the wavelet. Several ideas were tested according to [20,29–31].

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

Specifically, we would like to minimize the  $l_2$  expression  $\sum_{i=1}^N \|\mathbf{s}_i - \mathbf{W}\mathbf{r}_i\|_2^2$ , where  $\{\mathbf{s}_i\}_{i=1}^N$ ,  $\{\mathbf{r}_i\}_{i=1}^N$  are known and  $\mathbf{W}$  has the special form of  $\mathbf{W}_k = \mathbf{R}_k\mathbf{w}$ , where  $\mathbf{W}_k$  is the  $k$ -th column of  $\mathbf{W}$  and  $\mathbf{R}_k$  is a matrix that fits  $\mathbf{w}$  into a zero-vector from its  $k$ -th element, i.e.,  $\mathbf{R}_k$  looks like this:

$$\mathbf{R}_k = \begin{pmatrix} \mathbf{0}_{k-1 \times N_w} \\ \mathbf{1}_{N_w \times N_w} \\ \mathbf{0}_{N_r - k \times N_w} \end{pmatrix}. \quad (13)$$

This minimization problem was solved in [20] to obtain the optimal  $\mathbf{w}$ , although solved for different  $\mathbf{R}_k$  matrices. Accordingly we get the following solution:

$$\mathbf{w}^{\text{opt}} = \left( \sum_{k=1}^{N_r} \sum_{j=1}^{N_r} \left[ \sum_{l=1}^N \mathbf{r}_l[k] \mathbf{r}_l[j] \right] \mathbf{R}_k^T \mathbf{R}_j \right)^{-1} \sum_{i=1}^N \sum_{k=1}^{N_r} \mathbf{r}_i[k] \mathbf{R}_k^T \mathbf{s}_i \quad (14)$$

and now we can update the matrix  $\mathbf{V}$ .

This step is common to all different kinds of uncertainties in the wavelet since the true wavelet model and its connection to the seismic data are not affected by wavelet uncertainty.

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

### 3.1. Wavelet AWGN

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

$$\mathbf{w}' = \mathbf{w} + \mathbf{w}_v \quad (15)$$

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

$$\text{i. i. d } \mathbf{w}_v[k] \sim \mathcal{N}(0, \sigma_w^2), \quad 1 \leq k \leq N_w. \quad (16)$$

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

We recall that our purpose is to choose  $\epsilon_i$  according to (10) and (11). In this case we have the exact form as in (6) so no further adaptations need to be made to fit the proposed model and method.

First let us examine the general element in  $\mathbf{v}'_i$  as defined in (9). To do this we recall that

$$\mathbf{w}_v = [w_{v;1}, w_{v;2}, \dots, w_{v;N_w}]^T \quad \text{and} \quad \mathbf{W}_v = \begin{pmatrix} w_{v;1} & 0 & \dots & 0 \\ w_{v;2} & w_{v;1} & & 0 \\ \vdots & w_{v;2} & \ddots & \vdots \\ w_{v;N_w} & \vdots & & 0 \\ 0 & w_{v;N_w} & & w_{v;1} \\ 0 & 0 & \ddots & w_{v;2} \\ \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{v;N_w} \end{pmatrix}.$$

Substituting this into (9) we get,

$$\mathbf{v}'_i = \mathbf{v}_i - \mathbf{W}_v \mathbf{r}_i = \begin{bmatrix} \mathbf{v}_i[1] \\ \mathbf{v}_i[2] \\ \vdots \\ \mathbf{v}_i[N_r + N_w - 1] \end{bmatrix} - \begin{pmatrix} w_{v;1} & 0 & \dots & 0 \\ w_{v;2} & w_{v;1} & & 0 \\ \vdots & w_{v;2} & \ddots & \vdots \\ w_{v;N_w} & \vdots & & 0 \\ 0 & w_{v;N_w} & & w_{v;1} \\ 0 & 0 & \ddots & w_{v;2} \\ \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{v;N_w} \end{pmatrix} \begin{bmatrix} \mathbf{r}_i[1] \\ \mathbf{r}_i[2] \\ \vdots \\ \mathbf{r}_i[N_r] \end{bmatrix}. \quad (17)$$

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

$$\mathbf{v}'_i[k] = \mathbf{v}_i[k] - \sum_{j=1}^k w_{v;k-j+1} \mathbf{r}_i[j], \quad 1 \leq k \leq N_r + N_w - 1. \quad (18)$$

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

Substituting this into (10) we get,

$$\begin{aligned} \mathbf{V}_i &= \sum_{l=1}^{N_r+N_w-1} \mathbf{v}'_i[l] = \sum_{l=1}^{N_r+N_w-1} \left( \mathbf{v}_i[l] - \sum_{j=1}^l w_{v;l-j+1} \mathbf{r}_i[j] \right) \\ &= \sum_{l=1}^{N_r+N_w-1} \mathbf{v}_i[l] - \sum_{l=1}^{N_r+N_w-1} \sum_{j=1}^l w_{v;l-j+1} \mathbf{r}_i[j] = \sum_{l=1}^{N_r+N_w-1} \mathbf{v}_i[l] \\ &\quad - \sum_{j=1}^{N_w} \left( \sum_{k=1}^{N_r} \mathbf{r}_i[k] \right) w_{v;j}. \end{aligned} \quad (19)$$

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

$$\begin{aligned} \sigma_{\mathbf{V}_i}^2 &= \sum_{l=1}^{N_r+N_w-1} \sigma_{\mathbf{v}_i[l]}^2 + \sum_{j=1}^{N_w} \left( \sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 = (N_r + N_w - 1) \sigma_w^2 \\ &\quad + N_w \left( \sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 \end{aligned} \quad (20)$$

and now we can obtain  $\epsilon_i$  from (11):

$$\epsilon_i = \sigma_{\mathbf{V}_i} = \sqrt{(N_r + N_w - 1) \sigma_w^2 + N_w \left( \sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2}. \quad (21)$$

Notice the dependence on  $\sigma_w$  and  $\sigma_{w'}$ , which changes from iteration to iteration. We show an easy way to update  $\sigma_w$  in each iteration to best fit  $\epsilon$  to the current iteration. Updating  $\sigma_w$  is not trivial and has no analytical solution to date, so in our system we assume  $\sigma_w$  stays constant from one iteration to the next.

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

$$\mathbf{w}_{init} = \mathbf{w} + \mathbf{w}_v \mathbf{w}_{curr} = \mathbf{w} + \mathbf{w}'_v \quad (22)$$

where we assume that  $\mathbf{w}_v[k] \sim \mathcal{N}(0, \sigma_w^2)$ ,  $1 \leq k \leq N_w$ , and

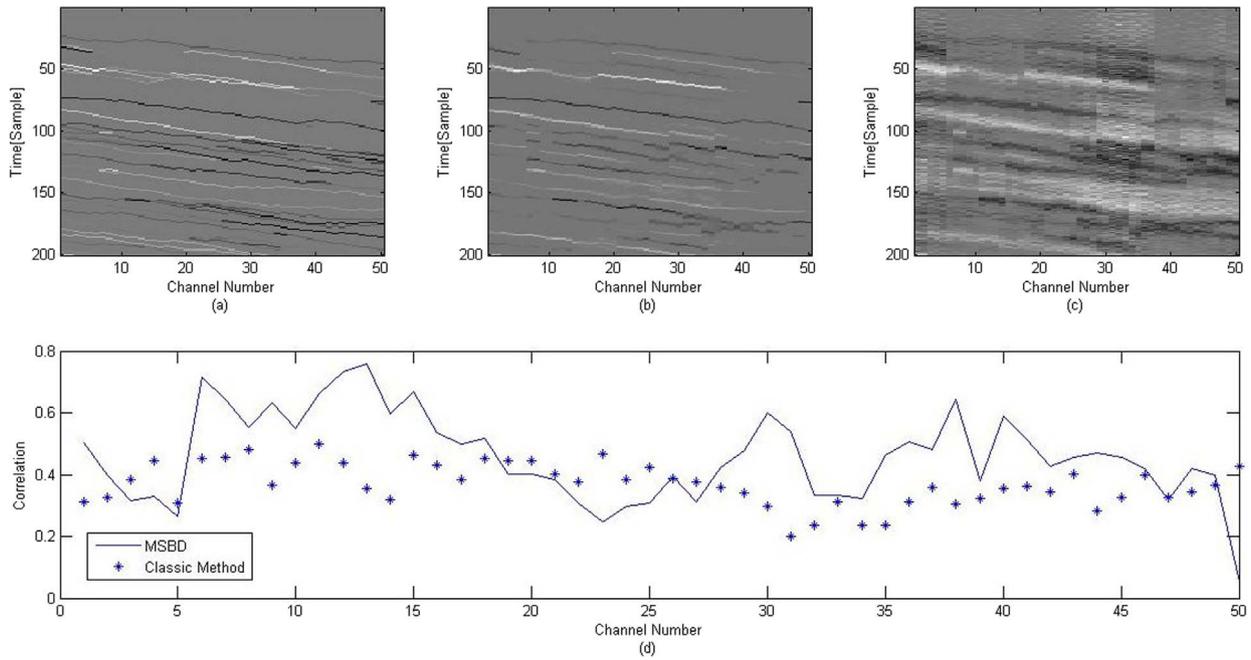


Fig. 1. True and recovered reflectivity and the correlations between them: (a) true (original) reflectivity, (b) recovered reflectivity using MSBD, (c) recovered reflectivity using the classic method, and (d) correlation between true and recovered reflectivity as a function of channel number.

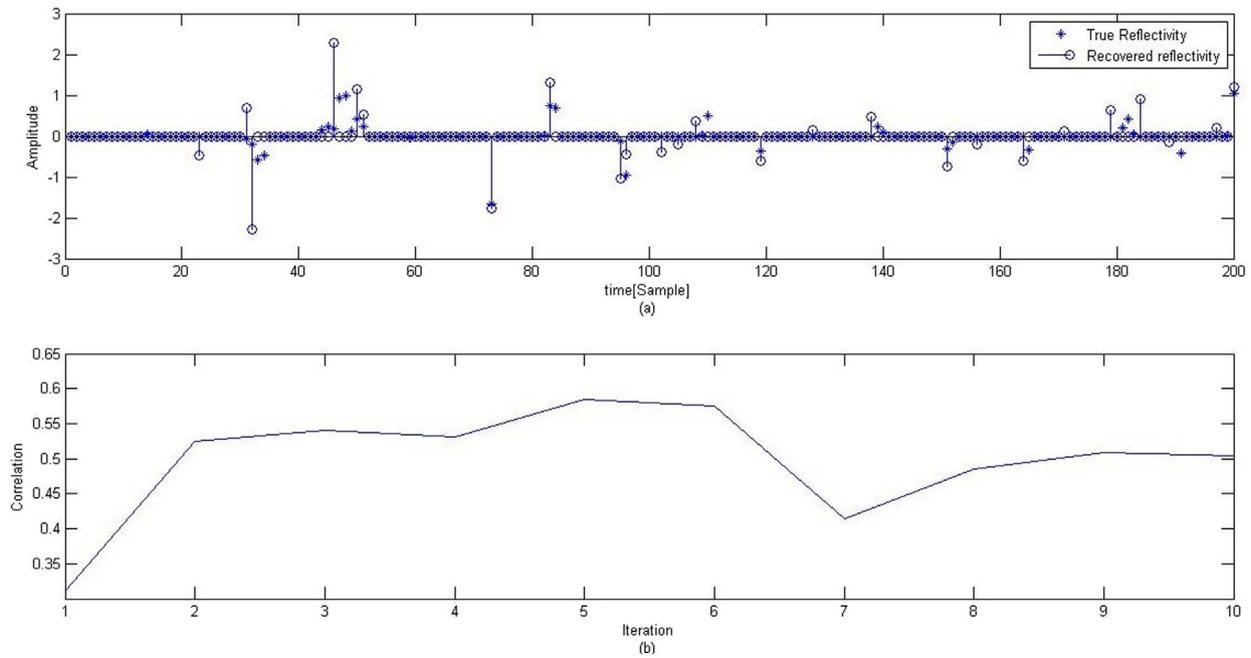


Fig. 2. (a) MSBD recovered reflectivity compared to the original reflectivity, within a specific channel. (b) The correlation between recovered and original reflectivity within a specific channel, as a function of iteration.

$w'_w[k] \sim \mathcal{N}(0, \sigma_w^2)$ ,  $1 \leq k \leq N_w$ . We know  $\sigma_w$  and aim to find  $\sigma'_w$ . We recall that we hold  $w_{mit}$  and  $w_{curr}$  fixed, and analyze the following term:

$$\|w_{mit} - w_{curr}\|_2^2 = \|w + w_v - (w + w'_v)\|_2^2 = N_w[\sigma_w^2 + (\sigma'_w)^2]. \quad (23)$$

It is easy to see that we can extract  $\sigma'_w$ ,

$$\sigma'_w = \sqrt{\frac{1}{N_w} \|w_{mit} - w_{curr}\|_2^2 - \sigma_w^2}. \quad (24)$$

Now we can update  $\sigma_w$  at the beginning of each iteration.

### 3.2. Wavelet parametric change

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

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

$$w(t; f) = (1 - 2\pi^2 f^2 t^2) e^{-\pi^2 f^2 t^2} \quad (25)$$

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

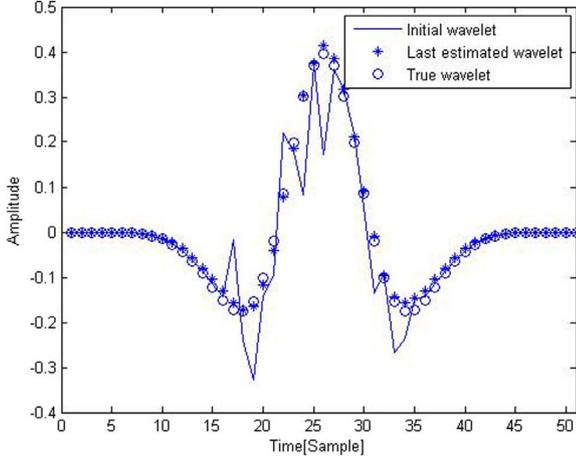


Fig. 3. Wavelet estimation.

defined by the parameter  $f$ , where  $f \neq f_0$ .

In order to simplify the later analysis of  $\epsilon$ , we present the change in  $f$  as an additive term to the true wavelet. To do this, we need to represent  $w(t; f)$  as a Taylor series. We develop the series up to three elements. The first two derivatives are:

$$\begin{aligned} \frac{\partial w(t; f)}{\partial f} &= -4\pi^2 f t^2 e^{-\pi^2 f^2 t^2} - 2\pi^2 f t^2 (1 - 2\pi^2 f^2 t^2) e^{-\pi^2 f^2 t^2} \\ &= 2\pi^2 f t^2 (2\pi^2 f^2 t^2 - 3) e^{-\pi^2 f^2 t^2} = (4\pi^4 f^3 t^4 - 6\pi^2 f t^2) e^{-\pi^2 f^2 t^2} \\ \frac{\partial^2 w(t; f)}{\partial f^2} &= (12\pi^4 f^2 t^4 - 6\pi^2 t^2) e^{-\pi^2 f^2 t^2} \\ &\quad - 2\pi^2 f t^2 (4\pi^4 f^3 t^4 - 6\pi^2 f t^2) e^{-\pi^2 f^2 t^2} \\ &= (24\pi^4 f^2 t^4 - 8\pi^6 f^4 t^6 - 6\pi^2 t^2) e^{-\pi^2 f^2 t^2}. \end{aligned} \quad (26)$$

So we can approximately write that:

$$w(t; f) \approx w(t; f_0) + \left. \frac{\partial w(t; f)}{\partial f} \right|_{f=f_0} (f - f_0) + \frac{1}{2} \left. \frac{\partial^2 w(t; f)}{\partial f^2} \right|_{f=f_0} (f - f_0)^2. \quad (27)$$

Putting this into a vector-matrix form in order to maintain the relation in (1) we get,

$$\mathbf{W} = \mathbf{W}_0 + \mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2 \quad (28)$$

where  $\mathbf{W}_0$ ,  $\mathbf{DW}_1$  and  $\mathbf{DW}_2$  are the convolution matrices of  $w(t; f_0)$ ,  $\frac{\partial w(t; f)}{\partial f}$  and  $\frac{\partial^2 w(t; f)}{\partial f^2}$ , respectively, and  $\Delta f = f - f_0$ . Applying this notation to (6)

we can see that  $\mathbf{W}' = \mathbf{W}_0$  and  $\mathbf{W}_v = -(\mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2)$ . Following this we can write (9) as,

$$\mathbf{v}'_i = \mathbf{v}_i + \left( \mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2 \right) \mathbf{r}_i. \quad (29)$$

Now we assume that  $\Delta f \sim \mathcal{N}(0, \sigma_f^2)$  and analyze (9):

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

$1 \leq k \leq N_r + N_w - 1$ .

Table 1

Correlations between recovered and original reflectivity for wavelet AWGN.

$\sigma_w$	SNR														
	0 dB			5 dB			10 dB			15 dB			20 dB		
	SSI	SMBD	MSBD												
0.05	0.4134	0.437	<b>0.4606</b>	0.5738	0.5818	<b>0.6695</b>	0.6234	0.5639	<b>0.6972</b>	0.5748	0.6281	<b>0.8231</b>	0.6724	0.6914	<b>0.8573</b>
0.1	0.4044	0.4297	<b>0.4545</b>	0.5202	0.4941	<b>0.7882</b>	0.4131	0.4193	<b>0.5274</b>	0.4561	0.5734	<b>0.7702</b>	0.2985	0.3761	<b>0.8817</b>
0.15	0.3903	0.4719	<b>0.5258</b>	0.4604	0.3184	<b>0.5974</b>	0.3832	0.4859	<b>0.8174</b>	0.3896	0.5379	<b>0.7768</b>	0.3210	0.5562	<b>0.7710</b>
0.2	0.3008	0.3717	<b>0.4832</b>	0.4258	0.4912	<b>0.5755</b>	0.3155	0.5804	<b>0.6574</b>	0.4530	0.6031	<b>0.7531</b>	0.3082	0.7015	<b>0.7239</b>

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

We would like to use (12); to do so we must first examine the variance of  $\mathbf{v}'_i[k]$ :

$$\begin{aligned} E[\mathbf{v}'_i[k]] &= \frac{1}{2} (\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i)[k] \sigma_f^2 E[(\mathbf{v}'_i[k])^2] = \sigma_v^2 + (\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^2 \\ &\quad + \frac{3}{4} (\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^4 \sigma_{v_i}^2[k] = E[(\mathbf{v}'_i[k])^2] - (E[\mathbf{v}'_i[k]])^2 \\ &= \sigma_v^2 + (\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^2 + \frac{1}{2} (\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^4 \end{aligned} \quad (30)$$

and we conclude:

$$\begin{aligned} \epsilon_i &= \sqrt{\sum_{n=1}^{N_r} \sigma_{v_i}^2} = \sqrt{\sum_{k=1}^{N_r} \left( \sigma_v^2 + (\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^2 + \frac{1}{2} (\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i)^2[k] \sigma_f^4 \right)} \\ &= \sqrt{N_r \sigma_v^2 + \sigma_f^2 \sum_{k=1}^{N_r} (\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i)^2[k] + \frac{1}{2} \sigma_f^4 \sum_{k=1}^{N_r} (\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i)^2[k]}. \end{aligned} \quad (31)$$

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

#### 4. Results and discussion

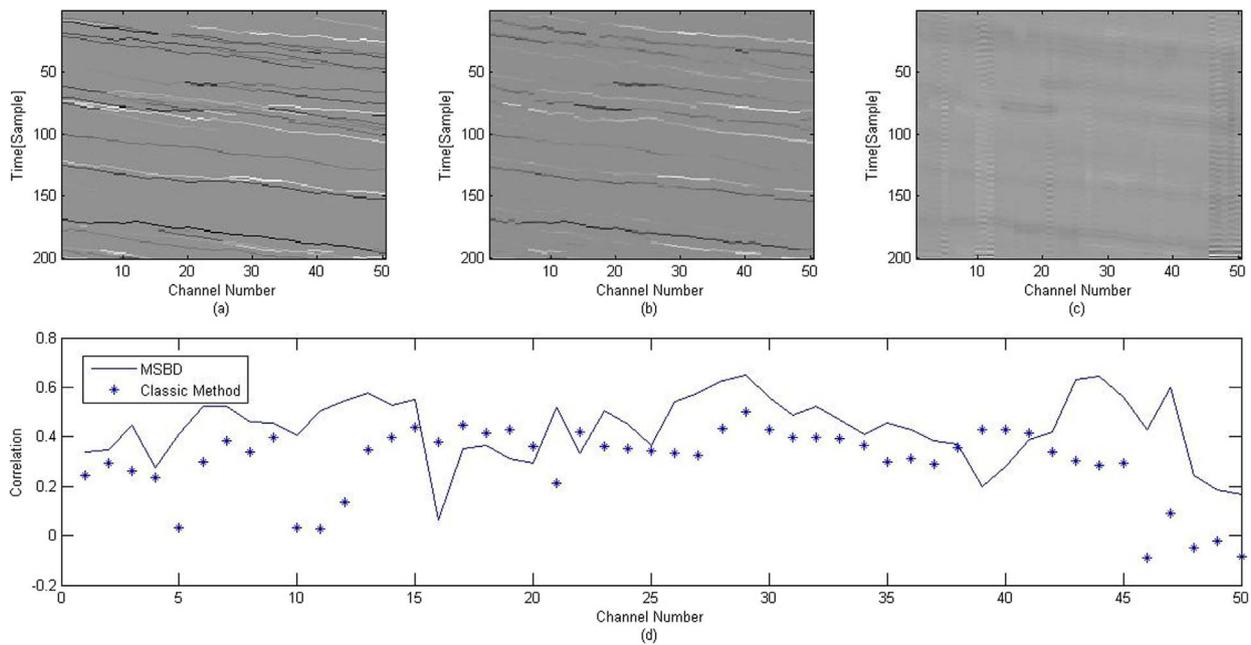
In this section we present experimental results obtained from testing the performances of MSBD. We will focus on synthetic data and we will also present a real data example. The synthetic reflectivity sequences were created using the model presented in [32] with SNR varied in the range 0–20 dB. Fifty channels were used and the wavelet was created using the Ricker wavelet.

In the first case, wavelet AWGN, the standard deviation of the wavelet AWGN,  $\sigma_w$ , was varied in the range 0.05–0.2. Twenty iterations were performed.

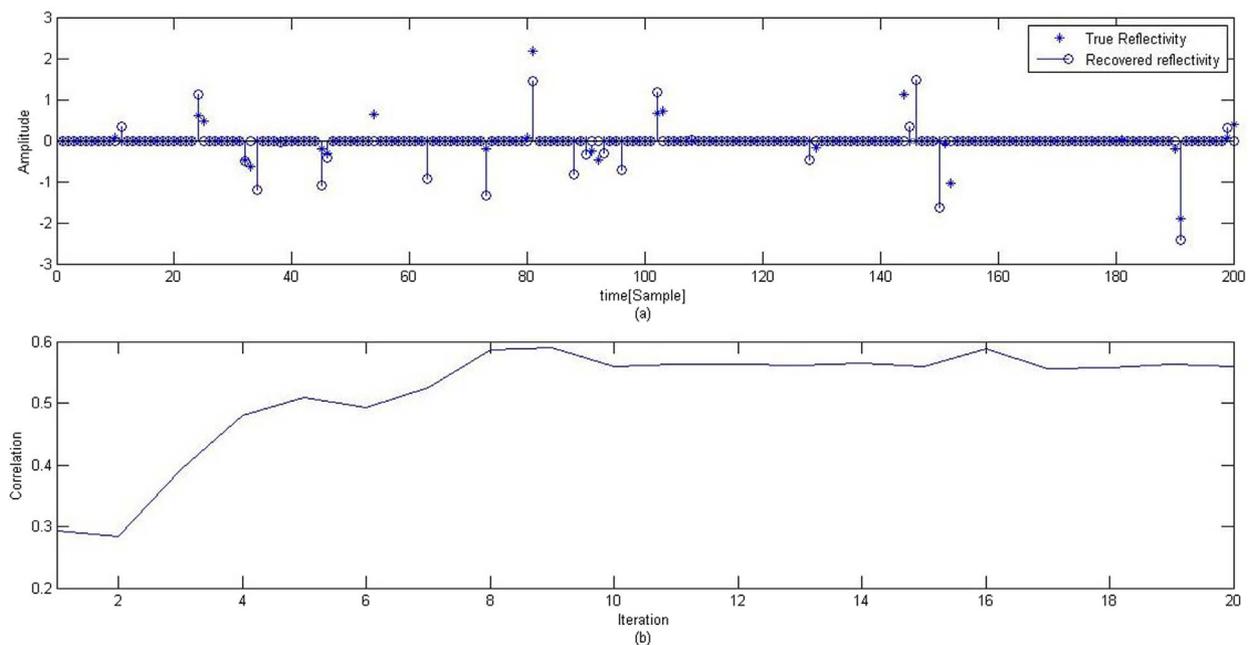
We present the results in terms of the correlation between the recovered reflectivity and the original reflectivity, and compare this to the case where we assume a fixed wavelet. The fixed wavelet is tested with the SSI and SMBD algorithms that were presented in the Introduction section. The SSI will represent the non-blind directive and the MSBD will represent the blind directive in our comparison. All the results of those algorithms are presented, although when we dive into figures and detailed explanation of the results, we will compare the MSBD only to SSI, just to not burden the reader. We will sometime address the SSI as the “classic method”.

In the next figures we present some graphs and results related to an example of SNR=20 dB and  $\sigma_w = 0.1$ .

First we examine in Fig. 1 the true (original) and recovered reflectivity series from both (MSBD and SSI) methods. A visual inspection of the recovered reflectivity graphs indicates a major improvement with MSBD. Even without a quantitative measure we can see that the signal outlines are recovered nicely using MSBD. At certain points we can see that MSBD has created discontinuities, for



**Fig. 4.** True and recovered reflectivity and their correlation: (a) true (original) reflectivity, (b) recovered reflectivity using MSBD, (c) recovered reflectivity using the classic method, and (d) correlation between recovered and true reflectivity as a function of channel number.



**Fig. 5.** (a) MSBD recovered reflectivity compared to the original reflectivity, within a specific channel. (b) The correlation between recovered and true reflectivity in a specific channel as a function of iteration.

example near channel number 7, at time 100 and time 125. These discontinuities are due to the fact that each channel is recovered irrespective of its neighbors; the neighbors are taken into account only at the stage of wavelet estimation. The correlation between recovered and true reflectivity for each channel (see Fig. 1(d)) provides a quantitative measure for testing the quality of the reflectivity recovery. It is clear that in most channels MSBD outperforms the SSI method.

Now we examine more “low-level” effects of the recovery by MSBD by focusing on recovery within a specific channel, in this case channel 1 (Fig. 2). When we examine the recovered and true reflectivity series (Fig. 2(a)), the first effect we notice is the difficulty in recovering adjacent impulses, such as samples 52 and 53. This is because the wavelet is wide (in the time domain) so it is difficult to distinguish

between two close impulses. If we look at sample 75 which is far from the previous and next impulses we note almost perfect recovery by MSBD. Now we examine the correlation between recovered and true reflectivity in this same channel as a function of iteration (Fig. 2(b)). In general, the correlation tends to increase but there are parts where it decreases. Also, we can clearly see that the final correlation is not the highest among all iterations. It will be interesting to check an algorithm that finds the best moment to stop the iterations. Of course this kind of algorithm will have to take into account all the channels and not only a specific one.

Finally, we examine the estimation of the wavelet (Fig. 3). We recall that MSBD is a two-stage algorithm, where the second stage is wavelet recovery. Here we can see the initial wavelet that was provided at the

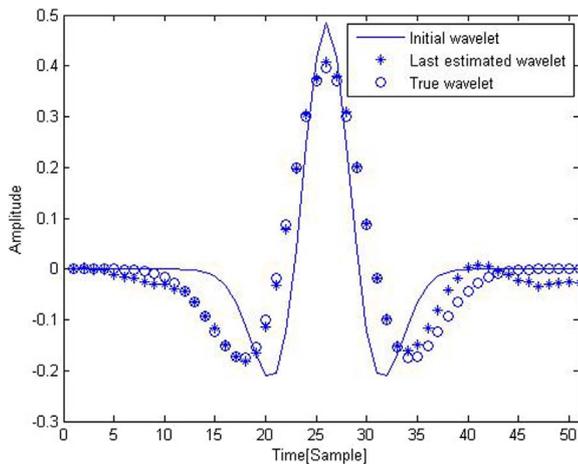


Fig. 6. Wavelet estimation.

beginning, and the estimated wavelet at the last iteration, both compared to the true wavelet from which the seismic data were created. We can see that the estimation of the wavelet is almost perfect.

The mean correlations between the recovered and original reflectivity measured across all channels, for all tested cases, i.e., SNR varies from 0 to 20 dB and  $\sigma_w$  varies from 0.05 to 0.2, are presented in Table 1. Separate correlations for the SSI, SMBD and MSBD are presented for each value of SNR and  $\sigma_w$ .

A few interesting points arise from these results. First, we see that the mean correlation of MSBD outperforms the mean correlation of SSI and SMBD methods. This demonstrates that the great potential improvement MSBD has to offer for seismic deconvolution. Second, we would have expected to see a fixed trend of improvement in the correlation with increasing SNR and decreasing  $\sigma_w$ . For SNR this assumption held true, besides three exceptions:  $\sigma_w = 0.1$  and SNR = 5 dB,  $\sigma_w = 0.15$  and SNR = 10 dB,  $\sigma_w = 0.2$  and SNR = 15 dB, however for  $\sigma_w$  this assumption was invalid. This is due to the fact that recovery of the reflectivity is more related to the relative changes in adjacent samples of the wavelet than to the standard deviation of the noise in the signal.

Until now we saw how MSBD handles the case of wavelet AWGN; now we present its performance under wavelet parametric change as described above (Section 3.2).

This reflectivity series was created in the same way as for the wavelet AWGN, according to the same element and with the same parameters. Likewise, the SNR of the seismic data varied from 0 to 20 dB, the wavelet was the Ricker Wavelet with  $f_0 = 2$  and  $\sigma_f$  varied from 0.5 to 1.2.

We illustrate the main results using an example with SNR=10 dB and  $\sigma_f = 1$ , again using 20 iterations.

First we examine the true reflectivity series (Fig. 4(a)), the recovered reflectivity series after applying MSBD (Fig. 4(b)) and the recovered reflectivity series using the SSI method where we assume a fix wavelet throughout the process (Fig. 4(c)). Once again, we can

visually see the superiority of MSBD and the improvement it offers, even without a quantitative measure. Using the SSI recovery, the outlines of the reflectivity signals are visible, but very unclear and contain many discontinuities. In contrast, the MSBD recovery series shows clear reflectivity outlines recovered with high correlation to the true reflectivity. Surprisingly, there are far fewer discontinuities here compared to the wavelet AWGN (Fig. 1(b)), even though neighboring channels were not taken into account when recovering a specific channel. This could be due to better wavelet estimation as is shown in Fig. 6. Visual inspection of the correlation between recovered and true reflectivity for each channel (Fig. 4(d)) indicates that MSBD clearly outperforms the SSI method. The sharp increases and decreases in the correlation measure are due to the correlation calculation. Even if a certain recovery method recovers the amplitudes of the trace perfectly, a shift of one sample in the impulse times can produce a very small correlation measure at certain points in time (if we assume no adjacent impulses then the correlation will be 0). A shift of one sample in the recovery is quite common, since the wavelet itself is not an impulse and has a certain width in the time domain. This is highly affected by the wavelet sample rate.

We now examine the recovery of a specific channel, in this case, channel 45 (Fig. 5). Comparison between the reflectivity series recovered by MSBD and the true reflectivity series (Fig. 5(a)) demonstrates the same effects observed for the wavelet AWGN in a specific channel (Fig. 2(a)). Now we examine the correlation between recovered and true reflectivity for the same channel as a function of iteration (Fig. 5(b)). Again, the correlation demonstrates a generally increasing trend, however it also decreases at some points along the process. As in the case of wavelet AWGN (Fig. 2(b)), the correlation at the final iteration is not the highest among all iterations.

Finally, we examine the estimated wavelet at the last iteration compared to the true wavelet and the initial wavelet that was provided to us (Fig. 6). We can see a very good estimation of the wavelet, even better than that for the wavelet AWGN in Fig. 3.

The mean correlations between the recovered and original reflectivity measured across all channels, for all tested cases, i.e., SNR varies from 0 to 20 dB and  $\sigma_f$  varies from 0.5 to 1.2, are presented in Table 2. Separate correlations for the SSI, SMBD and MSBD are presented for each value of SNR and  $\sigma_f$ .

In general, the correlation between the recovered and original reflectivity increases as SNR increases, as was found for wavelet AWGN (Table 1). However, it is not clear whether a decrease in  $\sigma_f$  leads to better recovery. This is because under wavelet parametric change the relative amplitude of adjacent samples of the wavelet is maintained while only the width of the wavelet and the absolute amplitudes change. In other words, the maximum amplitude of the wavelet will be maintained at the center of the wavelet, so the quality of the recovery is not necessarily related to  $\sigma_f$ .

A very interesting point we can easily see in both simulations is that we can clearly see that if we have some initial knowledge about the wavelet, the semi-blind (MSBD) method has more potential in recovering the reflectivity series with better performance than the blind (SMBD) or non-blind (SSI) directives. This makes sense because

Table 2  
Correlations between recovered and original reflectivity for wavelet parametric change.

$\sigma_f$	SNR														
	0 dB			5 dB			10 dB			15 dB			20 dB		
	SSI	SMBD	MSBD												
0.5	0.3942	0.3501	<b>0.5339</b>	0.2590	0.3971	<b>0.6110</b>	0.3201	0.4185	<b>0.9232</b>	0.7062	0.7205	<b>0.7721</b>	0.8947	0.8827	<b>0.8954</b>
0.8	0.5442	0.5903	<b>0.6057</b>	0.6673	0.7036	<b>0.7471</b>	0.4537	0.6407	<b>0.6956</b>	0.7020	0.7540	<b>0.8202</b>	0.5266	0.7152	<b>0.8112</b>
1	0.4278	0.5013	<b>0.5649</b>	0.3303	0.4155	<b>0.6991</b>	0.3540	0.5914	<b>0.6965</b>	0.5160	0.6510	<b>0.7686</b>	0.6621	0.7958	<b>0.7309</b>
1.2	0.4953	0.5183	<b>0.5359</b>	0.4271	0.5356	<b>0.6553</b>	0.5768	0.5608	<b>0.6790</b>	0.7272	0.7012	<b>0.7822</b>	0.5780	0.7825	<b>0.8028</b>

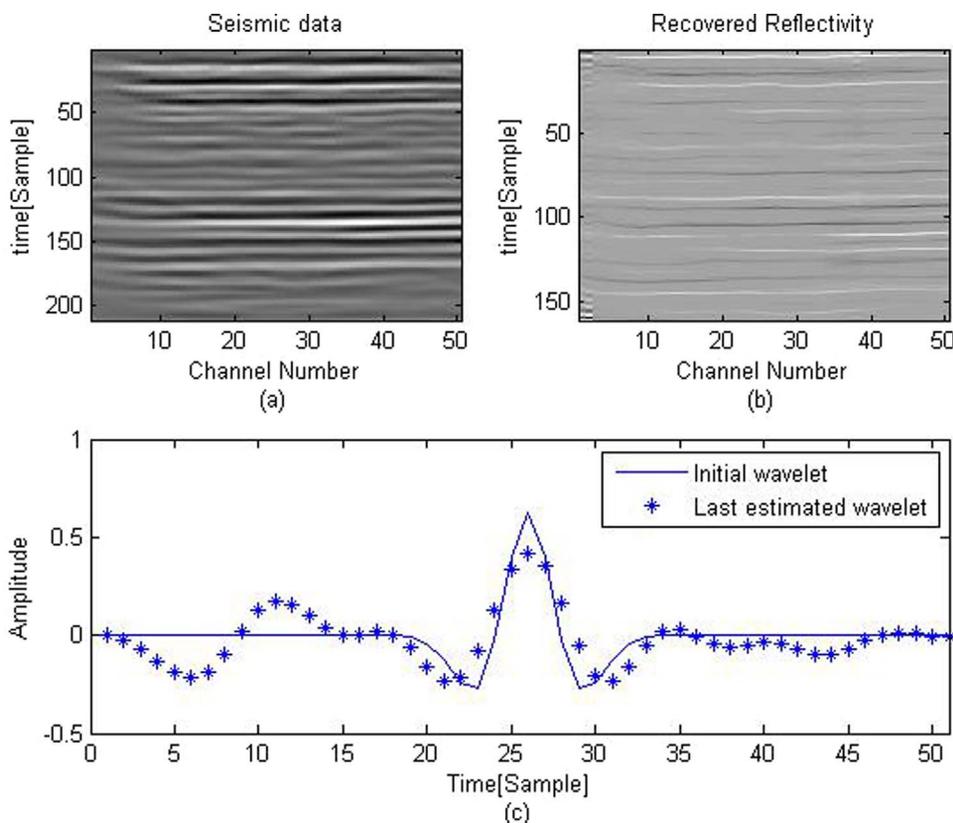


Fig. 7. Real data example: (a) seismic data, (b) recovered reflectivity using MSBD, and (c) wavelet.

non-blind methods assume that the wavelet is given and does not change – which is not the case, and blind methods assume nothing about the initial wavelet so they are not exploiting all the information given in the problem. This clearly states the importance of the method we propose.

In these simulations we used a private laptop with an i5 Intel core, 16 GB RAM. The mean computation time was 2.43 s for SSL, 5.52 s for SMBD and 37.12 s for MSBD with 20 iterations. The mean time was calculated from all of the abovementioned simulated data.

In the next figures we will see a real data example, courtesy of GeoEnergy Inc., Texas. In this example we did not have a priori information about the wavelet, so we assumed a Ricker wavelet with  $f = 1.5$ ,  $\sigma_f = 0.5$ . The SNR in this example is 2 dB. In Fig. 7 we can see the seismic data that was given to us and the recovered reflectivity using MSBD. Also, we can see the initial wavelet that was given to the algorithm and the optimal wavelet the algorithm found as optimal.

We can see the good recovery of MSBD. There are still discontinuities, for example in channel 38, around sample number 80. This discontinuities can be due to non-accurate assumptions on the noise in the setup. Still, we can clearly observe the un-blurring effect of the MSBD that has managed to distinguish between different spikes from near channels and from near samples in time.

## 5. Summary and conclusions

In this study we presented a new deconvolution method based on a two-stage iterative process that recovers the reflectivity series from the seismic data given a wavelet containing some kind of an uncertainty. We presented a general two-stage method, where one of the steps is fixed at the wavelet recovery stage, and the other is semi-fixed at the reflectivity recovery stage. The recovery of the reflectivity is semi-fixed because in general the method does not change from one type of signal to another; they all apply the BPDN solution for reflectivity recovery. The part in this stage that does change is the way we choose the trade-

off parameter in the BPDN solution.

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

## References

- [1] J. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*, 1983.
- [2] J. Kormylo, J. Mendel, Maximum likelihood detection and estimation of Bernoulli–Gaussian processes, *IEEE Trans. Inf. Theory* 28 (1982) 482–488.
- [3] K.F. Kaarensen, T. Taxt, Multichannel blind deconvolution of seismic signals, *Geophysics* 63 (1998) 2093–2107.
- [4] A. Heimer, I. Cohen, A.A. Vassiliou, Multichannel seismic modeling and inversion based on Markov–Bernoulli random field, *IEEE Trans. Geosci. Remote Sens.* 47 (2009) 2047–2058.
- [5] I. Ram, I. Cohen, S. Raz, Multichannel deconvolution of seismic signals using statistical MCMC methods, *IEEE Trans. Signal Process.* 58 (2010) 2757–2770.
- [6] D. Velis, T. Ulrych, Simulated annealing wavelet estimation via fourth-order cumulant matching, *Geophysics* (1999) 1939–1948.
- [7] M. Van der Baan, Time-varying wavelet estimation and deconvolution by kurtosis maximization, *Geophysics* (2008) V11–V18.
- [8] G. Menanno, A. Mazzotti, Deconvolution of multicomponent seismic data by means of quaternions: theory and preliminary results, *Geophys. Prospect.* (2012) 217–238.
- [9] N. Kazemi, D. Sacchi, Sparse multichannel blind deconvolution, *Geophysics* 79 (2014) 143–152.
- [10] A. Repetti, et al., Euclid in a taxicab: sparse blind deconvolution with smoothed l1/l2 regularization, *IEEE Signal Process. Lett.* (2015) 539–543.
- [11] S. Gleichman, Y. Eldar, Blind compressed sensing, *IEEE Trans. Inf. Theory* (2011) 6598–6975.
- [12] M. Rosenbaum, A. Tsybakov, Sparse recovery under matrix uncertainty, *Ann. Stat.* (2010) 2620–2651.
- [13] T. Nguyen, J. Castagna, High resolution reflectivity inversion, *J. Seism. Explor.* 19 (2010) 303–320.
- [14] Y. Wang, Seismic time-frequency spectral decomposition by matching pursuit, *Geophysics* 72 (2007) V13–V20.

- [15] Y. Wang, Multichannel matching pursuit for seismic trace decomposition, *Geophysics* 75 (2007) V61–V66.
- [16] R. Zhang, J. Castagna, Seismic sparse-layer reflectivity inversion using basis pursuit decomposition, *Geophysics* 76 (2011) 147–158.
- [17] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, *SIAM Rev.* 43 (2001) 129–159.
- [18] S.S. Chen, D.L. Donoho, Basis pursuit, *Signals Syst. Comput.* 1 (1994) 41–44.
- [19] J. Bork, L.C. Wood, Seismic interpretation of sonic logs, in: 71st Annual International Meeting, SEG, Expanded Abstracts, 2001, pp. 510–513.
- [20] M. Elad, Sparse and Redundant Representations, 2010.
- [21] R. Zhang, M.K. Sen, S. Srinivasan, Multi-trace basis pursuit inversion with spatial regularization, *J. Geophys. Eng.* 10 (2013).
- [22] R. Zhang, M.K. Sen, S. Srinivasan, A prestack basis pursuit seismic inversion, *Geophysics* 78 (2013) R1–R11.
- [23] C. Dossal, S. Mallat, Sparse spike deconvolution with minimum scale, in: *Signal Processing with Adaptive Sparse Structured Representations*, 2005, pp. 123–126
- [24] H.L. Taylor, S.C. Banks, J.F. McCoy, Deconvolution with the L1 norm, *Geophysics* 44 (1979) 39–52.
- [25] D.W. Oldenburg, S. Levy, K.J. Stinson, Inversion of band-limited reflection seismograms: theory and practice, *Proc. IEEE* 74 (1986) 487–497.
- [26] P.R. Gill, A. Wang, A. Molnar, The in-crowd algorithm for fast basis pursuit denoising, *IEEE Trans. Signal Process.* 59 (2011) 4595–4605.
- [27] W. Lu, N. Vaswani, Modified basis pursuit denoising (Modified-BPDN) for noisy compressive sensing with partially known support, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 3926–3929.
- [28] L. Dai, K. Pelckmans, An ellipsoid based, two-stage screening test for BPDN, in: *European Signal Processing Conference*, 2012, pp. 654–658
- [29] I. Tomic, P. Frossard, Dictionary learning – what is the right representation for my signal?, *IEEE Signal Process. Mag.* (2011) 27–38.
- [30] K. Skretting, K. Engan, Recursive least squares dictionary learning algorithm, *IEEE Trans. Signal Process.* 58 (2010) 2121–2130.
- [31] S.P. Lloyd, Least squares quantization in PCM, *IEEE Trans. Inf. Theory* IT-28 (1982) 129–137.
- [32] J. Idier, Y. Goussard, Multichannel seismic deconvolution, *IEEE Geosci. Remote Sens.* 31 (1993) 961–979.