Multichannel Semi-blind Sparse Deconvolution of Seismic Signals

Merabi Mirel and Israel Cohen, Technion - Israel Institute of Technology, and Anthony A. Vassiliou, GeoEnergy

SUMMARY

Seismic deconvolution is associated with recovering the reflectivity series from a seismic signal when the wavelet is known. In this paper, we address the problem of multichannel semi-blind seismic deconvolution, where the wavelet is unknown and 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 ℓ_2 minimization, exploiting the fact that all channels share the same wavelet.

INTRODUCTION

Deconvolution and kernel estimation 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. 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 (Mendel, 2013). 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.

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

$$s[n] = r[n] \circledast w[n] + v[n] \tag{1}$$

where $\circledast$ 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 channels, all share the same wavelet and the noise in the channels are statistically independent and identically distributed. Denoting by i the channel index, we get the following set of input-output relations:

$$s_i[n] = r_i[n] \circledast w[n] + v_i[n],\ 1 \leqslant i \leqslant N. \tag{2}$$

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

$$\mathbf{s}_i = \mathbf{W}\mathbf{r}_i + \mathbf{v}_i,\ 1 \leqslant i \leqslant N \tag{3}$$

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

The goal of the basic problem is to recover $\mathbf{r}_i$ from $\mathbf{s}_i$ while assuming full knowledge of $\mathbf{W}$ and σ_v. This problem has been widely investigated and is called the deconvolution problem. A wide variety of solutions have been proposed to solve this problem, depending on the model of the signal $\mathbf{r}_i$. In seismic deconvolution, assuming a sparse model for $\mathbf{r}_i$, sparse deconvolution methods have been proposed. One of them is Basis Pursuit Denoising (Elad, 2010; Gill et al., 2011; Lu and Vaswani, 2010; Dai and Pelckmans, 2012), which is an approach that solves the following optimization problem:

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

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

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

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

where ε controls the above-mentioned trade-off.

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

MULTICHANNEL SEMI-BLIND DECONVOLUTION

Our purpose is to establish a general method for the Semi-Blind Deconvolution problem and to specifically analyze a case 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 and choose to work with the ℓ_2 minimization for wavelet recovery. This method was adapted to best fit our problem. 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 ℓ_2 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.

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 a specific case. Denote by $\mathbf{w} \in \mathbb{R}^{N_w}$ the vector representation of the signal $w[n]$, $1 \leqslant n \leqslant N_w$. The additive noise to the wavelet is denoted by $\mathbf{w_v} \in \mathbb{R}^{N_w}$ 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,

$$\mathbf{W} = \mathbf{W'} - \mathbf{W_v}. \tag{6}$$

Substituting (6) into (3) we get,

$$\mathbf{s}_i = \mathbf{W}\mathbf{r}_i + \mathbf{v}_i = (\mathbf{W'} - \mathbf{W_v})\mathbf{r}_i + \mathbf{v}_i = \mathbf{W'}\mathbf{r}_i + \mathbf{v}_i - \mathbf{W_v}\mathbf{r}_i. \tag{7}$$

Looking at this relation,

$$\mathbf{s}_i = \mathbf{W'}\mathbf{r}_i + \mathbf{v}_i - \mathbf{W_v}\mathbf{r}_i \tag{8}$$

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

$$\mathbf{v}'_i = \mathbf{v}_i - \mathbf{W_v}\mathbf{r}_i. \tag{9}$$

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. The literature does not prove, nor imply, a generic analysis for choosing the trade-off parameter, ε, but it is clear that this parameter has a strong relation to the standard deviation of a more general definition of a term that accounts for all noise elements in the problem. In our case we first define a new term, $\bar{v}'_i$, as the sum of the noise elements:

$$\bar{v}'_i = \sum_{n=1}^{N_r+N_w-1} v'_i[n] \tag{10}$$

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

$$\varepsilon_i = \sigma_{\bar{v}'_i} = \sqrt{E(\bar{v}'_i - E\bar{v}'_i)^2} \tag{11}$$

where E is the expectation operator.

For the second step, we assume 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 (Elad, 2010; Tosic and Frossard, 2011; Skretting and Engan, 2010; Lloyd, 1982).

The seismic data can be considered a linear combination of the columns of $\mathbf{W}$, the dictionary, where the reflectivity series can be treated as the coefficients. This makes sense because the columns of $\mathbf{W}$ are shifted versions of $\mathbf{w}$. With that in mind, finding the wavelet when the reflectivity series is known can be treated by methods from the field of dictionary learning, as the purpose of this stage is to update and learn $\mathbf{W}$ (defined directly by $\mathbf{w}$). We use a method of dictionary update based on the Signature Dictionary as described in (Elad, 2010). 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 $\{\mathbf{s}_i\}_{i=1}^N$ and $\{\mathbf{r}_i\}_{i=1}^N$ are known and the k-th column of $\mathbf{W}$, denoted by $\mathbf{w}_k$, is a shifted version of $\mathbf{w}$, i.e., $\mathbf{w}_k = \mathbf{R}_k\mathbf{w}$ where

$$\mathbf{R}_k = \begin{pmatrix} \mathbf{0}_{k-1\times N_w} \\ \mathbf{I}_{N_w\times N_w} \\ \mathbf{0}_{N_r-k\times N_w} \end{pmatrix}. \tag{12}$$

This minimization problem was solved in (Elad, 2010) to obtain the optimal $\mathbf{w}$, although solved for different $\mathbf{R}_k$ matrices. Accordingly we get the following solution:

$$\mathbf{w}^{\text{opt}} = \left(\sum_{k=1}^{N_r} \sum_{j=1}^{N_r} \left[\sum_{i=1}^{N} r_i[k] r_i[j] \right] \mathbf{R}_k^T \mathbf{R}_j \right)^{-1} \sum_{i=1}^{N} \sum_{k=1}^{N_r} r_i[k] \mathbf{R}_k^T \mathbf{s}_i \tag{13}$$

from which we can update the matrix $\mathbf{W}$.

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

Now we proceed with analyzing a specific case of wavelet uncertainty: AWGN contamination of the wavelet. The model we are assuming is as follows:

$$\mathbf{w'} = \mathbf{w} + \mathbf{w_v} \tag{14}$$

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

$$w_v[k] \sim \mathcal{N}(0, \sigma_w^2), \ \ 1 \leqslant k \leqslant N_w. \tag{15}$$

In addition we assume that $\mathbf{w_v}$ and $\{\mathbf{v}_i\}_{i=1}^{N}$ are statistically independent. We recall that our purpose is to choose ε_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.

Recall that

$$\mathbf{W_v} = \begin{pmatrix} w_v[1] & 0 & \cdots & 0 \\ w_v[2] & w_v[1] & & 0 \\ \vdots & w_v[2] & \ddots & \vdots \\ w_v[N_w] & \vdots & & 0 \\ 0 & w_v[N_w] & & w_v[1] \\ 0 & 0 & \ddots & w_v[2] \\ \vdots & 0 & \vdots & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_v[N_w] \end{pmatrix}.$$

Substituting this into (9) we obtain

$$v_i'[k] = v_i[k] - \sum_{j=1}^{k} w_v[k-j+1] r_i[j], \;\; 1 \leqslant k \leqslant N_r + N_w - 1. \tag{16}$$

Substituting this into (10) we get,

$$\bar{v}_i' = \sum_{n=1}^{N_r+N_w-1} v_i[n] - \sum_{j=1}^{N_w} \left(\sum_{k=1}^{N_r} r_i[k] \right) w_v[j]. \tag{17}$$

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

$$\begin{aligned} \sigma_{\bar{v}_i'}^2 &= \sum_{n=1}^{N_r+N_w-1} \sigma_{\mathbf{v}_i[l]}^2 + \sum_{j=1}^{N_w} \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 \qquad (18) \\ &= (N_r + N_w - 1)\sigma_v^2 + N_w \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 \end{aligned}$$

and now we can obtain ε_i from (11):

$$\varepsilon_i = \sqrt{(N_r + N_w - 1)\sigma_v^2 + N_w \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2}. \tag{19}$$

Notice the dependence on σ_v and σ_w, which varies from iteration to iteration. We show an easy way to update σ_w in each iteration to best fit ε to the current iteration. Updating σ_v is not trivial and has no analytical solution to date, so in our system we assume σ_v remains constant from one iteration to the next. To update σ_w we analyze the current wavelet that was recovered from the last iteration and the initial wavelet that was given to us. The initial wavelet, $\mathbf{w}_{\text{init}}$ and the current wavelet, $\mathbf{w}_{\text{curr}}$, can be modeled as,

$$\begin{aligned} \mathbf{w}_{\text{init}} &= \mathbf{w} + \mathbf{w_v} \\ \mathbf{w}_{\text{curr}} &= \mathbf{w} + \mathbf{w}'_{\mathbf{v}} \end{aligned} \tag{20}$$

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

$$\begin{aligned} \|\mathbf{w}_{\text{init}} - \mathbf{w}_{\text{curr}}\|_2^2 &= \left\|\mathbf{w} + \mathbf{w_v} - \left(\mathbf{w} + \mathbf{w}_{\mathbf{v}}'\right)\right\|_2^2 \qquad (21) \\ &= N_w \left[\sigma_w^2 + \left(\sigma_w'\right)^2 \right]. \end{aligned}$$

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

$$\sigma_w' = \sqrt{\frac{1}{N_w} \|\mathbf{w}_{\text{init}} - \mathbf{w}_{\text{curr}}\|_2^2 - \sigma_w^2}. \tag{22}$$

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

RESULTS AND DISCUSSION

In this section we describe the experimental results obtained from testing the performances of Multichannel Semi-blind Sparse Deconvolution (MSSD). We focus on synthetic data. Synthetic reflectivity sequences were created using the model presented in (Idier and Goussard, 1993) with signal-to-noise ratio (SNR) varied in the range $0 - -20$ dB. Fifty channels were created using the Ricker wavelet. We observed the results in terms of the correlation between the recovered reflectivity and the original reflectivity, and compared them to the case where we assume a fixed wavelet. The fixed wavelet was tested with the SSI (Dossal and Mallat, 2005; Taylor et al., 1979; Oldenburg et al., 1986) and SMBD (Kazemi and Sacchi, 2014) algorithms. The SSI represented the non-blind directive and the SMBD represented the blind directive in our comparison. The comparison showed without doubt that MSSD outperformed both SSI and SMBD.

In the next figures we present some results related to an example of SNR = 20 dB and $\sigma_w = 0.15$.

In Figure 1 we show the seismic data as was received in the seismic sensors. The seismic data is a result of the convolution between the true reflectivity and the true wavelet. Figure 2 examines the reflectivity recovered using MSSD. Even without a quantitative measure we can see that the signal outlines are recovered nicely using MSSD. At certain points we can see that MSSD has created discontinuities, for example near channel number 13 at time 25 and channel number 20 at time 170. 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.

Figure 3 examines a "low-level" effect of the MSSD recovery by focusing on a specific channel. We notice the difficulty in recovering adjacent impulses, such as samples 64 and 70. This is because the wavelet is wide (in the time domain) so it is difficult to distinguish between two close impulses.

Figure 4 examines the estimation of the wavelet. We recall that MSSD is a two-stage algorithm, where the second stage is wavelet recovery. Here we can see the initial wavelet that was provided at the beginning and the estimated wavelet, both compared to the true wavelet. We can see that the estimation of the wavelet is almost perfect.

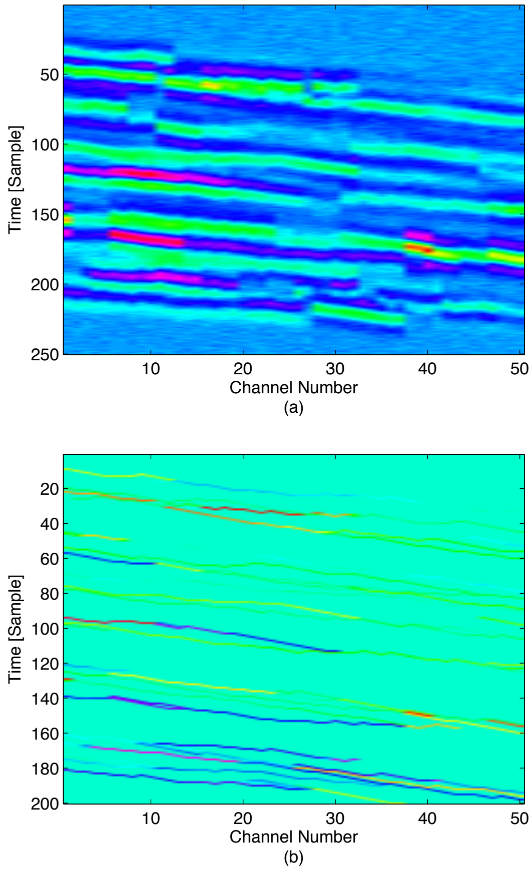

Figure 1: $SNR = 20dB$, $\sigma_w = 0.15$, (a) Seismic Data (b) true (original) reflectivity.

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 one case in which we analytically calculated the trade-off parameter. We compared our method to blind and non-blind methods. The results clearly show the advantage and logic behind MSSD. Also, we presented a more detailed example and discussed some effects of our method. 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.

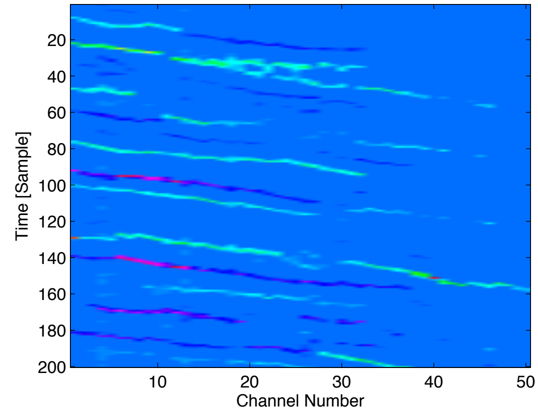

Figure 2: $SNR = 20dB$, $\sigma_w = 0.15$: Recovered reflectivity using MSSD.

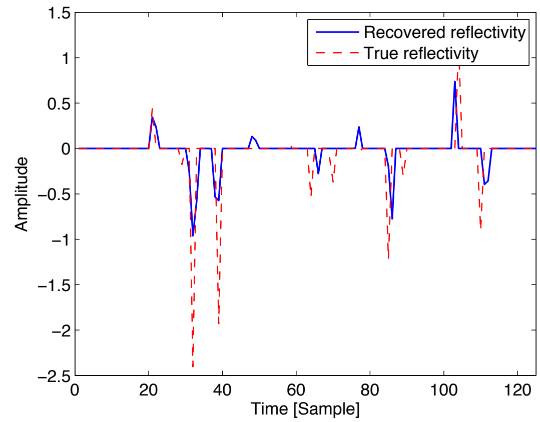

Figure 3: $SNR = 20dB$, $\sigma_w = 0.15$: MSSD recovered reflectivity compared to the original reflectivity, within a specific channel.

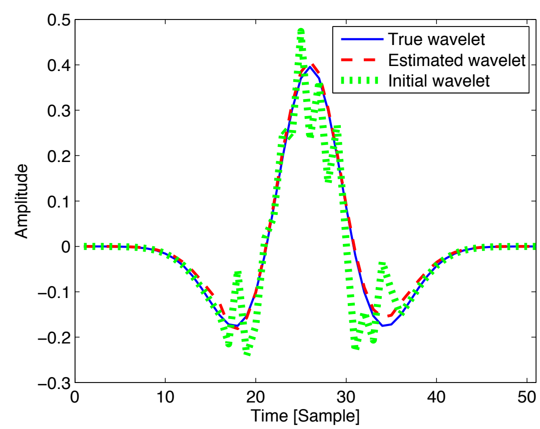

Figure 4: Wavelet estimation, $SNR = 20dB$, $\sigma_w = 0.15$.