Multichannel Sparse Spike Inversion

Deborah Pereg and Israel Cohen, Technion – Israel Institute of Technology, and Anthony A. Vassiliou, GeoEnergy

Abstract. In this paper, we address the problem of sparse multichannel seismic deconvolution. We introduce Multichannel Sparse Spike Inversion (MSSI) as an iterative procedure, which deconvolves the seismic data and recovers the earth two-dimensional (2D) reflectivity image, while taking into consideration the relations between spatially-neighboring traces. We demonstrate the improved performance of the proposed algorithm and its robustness to noise, compared to competitive single-channel algorithm through simulations and real seismic data examples.

Keywords: Deconvolution; Optimization; Reflectivity; Signal Processing; Sparse.

1. Introduction

In the field of signal processing, it is often necessary to recover an input signal from its filtered version. The operation of deconvolution is ideally set to achieve this goal, and to undo the operation of a linear time invariant system performed on the input signal. In the seismic setting, a short-duration acoustic pulse is transmitted from the earth surface. The reflected pulses from the ground are then received by a sensor array [1]. Our goal is to reveal the ground layer's structure hidden in each of the received seismic traces.

We assume that the short seismic pulse (the wavelet) is known and approximately time-invariant. This assumption is common in seismic data processing [2–6]. Even under this assumption, the inversion process is often unstable. The seismic wavelet is bandlimited, and the seismic trace might be noisy. Due to this instability, there are many possible reflectivity series that could fit the same measured seismic traces. The objective of our work is to find the best estimate of the reflectivity. We assume the reflectivity is sparse. Hence, its extraction could be done by sparse inversion techniques.

In previous works, the solution to the multichannel deconvolution problem involves separation of the seismic data into independent vertical one-dimensional (1D) deconvolution problems, where each reflectivity channel is estimated apart from the other channels [1–5, 7–10]. The wavelet is taken to be a 1D column signal, and each 1D reflectivity column appears in the vertical direction as a sparse spike train. Some of these methods describe the reflectivity and the noise as two independent stochastic processes with known second-order statistics. Berkhout [1] tried to solve the seismic blind deconvolution problem by assuming that the reflectivity is a white sequence and that the seismic wavelet is a minimum phase signal. Many attempts have been made to avoid the minimum phase assumption. Some of these methods are blind, meaning that both the reflectivity and the wavelet are unknown. Homomorphic deconvolution [11], implemented in exploration seismology by Ulrych [7], was first developed for restoring reverberated and resonated sound and speech. It was also implemented for the case of blurred images [11]. In homomorphic deconvolution, we find the log amplitude of the distorting system in the frequency domain. Then we can restore the signal of interest by simply subtracting the log amplitude of the distorting system from the log amplitude of the observation signal in the frequency domain. Minimum Entropy Deconvolution (MED) [8] and Maximum Kurtosis Adaptive Filtering [10], try to find a deconvolution filter, by optimization of a sparsity cost function. The struggling point of these methods is that they are suboptimal and produce unstable results due to the certain shortcomings. Homomorphic deconvolution is unable to correct the unknown phase distortions and tend to be highly sensitive to noise. MED and Maximum Kurtosis Adaptive Filtering are sensitive to noise and greatly influenced by the assumed length of the deconvolution filter, in addition to their inclination to cancel small reflectivity spikes.

Sparse seismic inversion methods have managed to produce stable reflectivity solutions, see e.g., [3, 5, 9, 12] use Matching Pursuit Decomposition (MPD) to decompose

the seismic trace into reflectivity patterns. In order to increase the lateral resolution beyond the resolution that could be achieved by wavelet inverse filtering, they depend on a-priori knowledge. A starting model is built according to this prior information. Unfortunately, the starting model can be inaccurate due to lateral variations in the waveform interference path, in the propagation rate or in the earth layers' impedances. Also, the myopia limitation of the MPD method becomes apparent when the dictionary is non-orthogonal. Basis Pursuit Decomposition (BPD) [13] is more advantageous. Originally developed as a compressive sensing technique. BPD utilizes an l_1 norm optimization and finds a single global solution in a computationally more efficient way. Moreover, it performs well even when dictionary elements are non orthogonal.

Other important methods are Sparse Spike Inversion (SSI) [2] and Basis Pursuit Inversion (BPI) [4]. SSI and BPI recover each column of the reflectivity by solving a simple Basis Pursuit Denoising (BPDN) problem [14]. These methods perform very well under sufficiently high signal-to-noise ratio (SNR). Dosal and Mallat [15] provide a lower bound on the minimum distance between spikes that can be recovered by ℓ_1 penalized deconvolution. However, one of the main disadvantages of these methods is that they ignore the correlation between adjacent traces. This correlation emerges from the natural assumption that the earth layers are horizontally structured. We refer the reader to [6] for a full comparison between SSI vs. BPI.

Obviously, utilization of 1D restoration methods in the case of 2D seismic data is not optimal. Single-channel methods do not exploit the relations between spatially near traces. Thus, multichannel deconvolution is more robust. Zhang et al. [16] suggest to extend the BPI method to a multi-trace process with spatial regularization added in order to enhance lateral continuity and vertical resolution. Two variations of multichannel Bayesian deconvolution methods are suggested by Idier and Goussard [17]. Their approach is based on two Markov-Bernoulli-Gaussian reflectivity models (MBG I and II). The first model is a 2D extension of the 1D Bernoulli-Gaussian (BG) representation. Mendel et al. [18, 19] use this 1D BG model in their Maximum-likelihood algorithm to estimate the reflectivity and the wavelet. The second model (MBG II) is more adapted to the physical and geometrical characteristics of the earth layers' acoustic impendances. The deconvolution is performed by a suboptimal Maximum a-posteriori (MAP) estimator. Then, they use a method similar to the single most likely replacement (SMLR) algorithm [18] to iteratively recover each reflectivity column from the corresponding observed seismic trace and the preceding estimated reflectivity column. Kaaresen and Taxt [20] also propose a multichannel version of their single-channel blind deconvolution algorithm. The procedure repeats two stages: first, the wavelet is estimated by least-squares fit, and then the reflectivity is estimated by the iterated window maximization algorithm [21]. The algorithm produces better channel estimates since it updates more than one reflector in one trace at once, and also encourages lateral smoothness of the reflectors. However, these methods rely on a parametric model that leads to a nonconvex optimization problem. Usually, it is very difficult to find a global optimal solution to this kind of problems. The solution is

normally found by searching for correct reflectivity spikes' locations, within a limited number of potential reflectivity sequences (as in the SMLR algorithm mentioned above [19]). This way, an optimal solution is achieved at the expense of heavy computational burden and an extended search.

Heimer, Cohen, and Vassiliou [22], [23] also propose a multichannel blind deconvolution. They integrate the algorithm of Kaaresen and Taxt [20] with dynamic programming [24], [25]. Valid reflectivity states and transitions between reflector arrangements of spatially-neighboring traces are defined. Then, the sequences of reflectors that are legally concatenated to other reflectors by valid transitions are extracted. Heimer and Cohen [26] also propose a method based on the Markov-Bernoulli random field (MBRF) modeling. The Viterbi algorithm [27] is applied to the search of the most likely sequences of reflectors concatenated across the traces by legal transitions.

Ram, Cohen, and Raz [28] also propose two multichannel blind deconvolution algorithms for the restoration of 2D seismic data. Both algorithms are based on the Markov-Bernoulli-Gaussian I (MBG I) reflectivity model. In the first algorithm, each reflectivity channel is estimated from the corresponding observed seismic trace, while taking into consideration the estimate of the previous reflectivity channel. The procedure is carried out using a slightly modified maximum posterior mode (MPM) algorithm [29]. The second algorithm considers estimates of both the previous and following neighboring columns.

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

The remainder of the paper is organized as follows. In Section 2, we review the basic theory of the seismic deconvolution problem. In Section 3, we introduce our algorithm. In Section 4, we present simulation and real data results. Finally, in Section 5, we conclude and discuss further research.

2. Problem Formulation

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

$$s(t) = w(t) * r(t) + n(t) \tag{1}$$

where $w(t)$ is the seismic wavelet, $r(t)$ is the reflectivity series, and $n(t)$ is the noise. The symbol $*$ denotes one-dimensional linear convolution operation. This model assumes that the earth structure is stratified. It consists of planar horizontal layers of constant impedance and reflections are generated at impedance discontinuities, i.e.,

at the boundaries between adjacent layers. Each 1D seismic trace is a convolution of the seismic wavelet and the reflectivity pattern. All channels are excited by the same wavelet $w(t)$. The support of the wavelet is finite and shorter than the channel's length.

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

A seismic trace consists of a linear combination of $w(t)$ and its time shifts, corresponding to the non-zero reflectors in $r(t)$. The discrete convolution (1) can be written in matrix-vector form as

$$s_{N\times 1} = W_{N\times M} r_{M\times 1} + n_{N\times 1} \tag{2}$$

where $W_{N\times M} \in \mathbb{R}^{N\times M}$, represents the dictionary.

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

$$\min \|r_{M\times 1}\|_0 \text{ subject to } \|s_{N\times 1} - W_{N\times M} r_{M\times 1}\|_2^2 < \varepsilon\,. \tag{3}$$

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

$$\min_{r_{M\times 1}} \frac{1}{2} \|s_{N\times 1} - W_{N\times M} r_{M\times 1}\|_2^2 + \lambda \|r_{M\times 1}\|_1\,. \tag{4}$$

The optimization problem as defined in (4) is called Least Absolute Shrinkage and Selection Operator (LASSO) [30]. The l_1 penalty in similar problems is used in order to promote a sparse solution $r_{M\times 1}$ [13, 14].

On the other hand, the BPI method, proposed by Zhang and Castagna [4], apply "dipole decomposition", i.e., each pair of neighboring impulses in the reflectivity sequence is represented as a linear combination of even and odd impulse pairs. Each even and odd pair corresponds to the top and base reflector of a layer. Since the layer thickness is unknown, the dictionary comprises all possible thicknesses up to a maximum layer time-thickness.

3. Multichannel Sparse Spike Inversion (MSSI)

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

Assume J adjacent columns, and for simplicity assume that J is odd. Denote the current column, which we wish to estimate, by $\mathbf{r}_i$, and a previous or subsequent column by $\mathbf{r}_{i+k}$ where $-\frac{J-1}{2} \le k \le \frac{J-1}{2}$. We estimate each reflectivity column from the corresponding observed seismic trace $\mathbf{s}_i$, taking into consideration the current estimate of $\frac{J-1}{2}$ preceding reflectivity columns, and of $\frac{J-1}{2}$ subsequent reflectivity columns. Out

of J estimated columns only the middle reflectivity column is kept. The estimates of the other $J-1$ columns are discarded. If we wish to use only the subsequent column (i.e. $J=2$), we keep the first reflectivity column, and discard the subsequent column (in this case $-\frac{J}{2} < k \leq \frac{J}{2}$).

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

$$\min_{\mathbf{r}_i,...,\mathbf{r}_{i\pm\frac{J-1}{2}}} \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \frac{1}{2} \left\| \mathbf{s}_{i+k} - \mathbf{W}\mathbf{r}_{i+k} \right\|_2^2 + \lambda_0 \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \left\| \mathbf{r}_{i+k} \right\|_1 + \sum_{k=-\frac{J-1}{2},k\neq 0}^{\frac{J-1}{2}} \frac{1}{2} \lambda_k \left\| \mathbf{r}_i - \mathbf{H_k}\mathbf{r}_{i+k} \right\|_2^2. \quad (5)$$

Where $\mathbf{r}_i, ..., \mathbf{r}_{i\pm\frac{J-1}{2}}$ are J reflectivity columns, and $\mathbf{s}_i, ..., \mathbf{s}_{i\pm\frac{J-1}{2}}$ are J corresponding seismic traces. $\mathbf{W}$ is the convolution matrix formed by the seismic discrete wavelet $\mathbf{w}$, assumed to be known. The tradeoff parameter λ_0 controls the balance between the reflectivity sparseness and the least-squares error. The tradeoff parameters λ_k promote smoothness of the reflectivity in the horizontal direction. $\mathbf{H_k}$ is the convolution matrix of a Low-pass filter. We can choose $\mathbf{H_k}$ to be the convolution matrix of a Hamming window or an Averaging filter. Hence, $\mathbf{H_k}$ controls the smoothness as it reduces the penalty for layer boundaries whose orientation is diagonally descending, horizontal, and diagonally ascending. The size of the smoothing filter controls the desired smoothness of the resultant reflectivity image. This way, the minimization is performed by taking into account the distances between each reflectivity column and the preceding and subsequent reflectivity columns.

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

$$(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_{i\pm1}, ..., \hat{\mathbf{r}}_{i\pm\frac{J-1}{2}}) = \min_{\mathbf{r}_i,\mathbf{r}_{i\pm1},...,\mathbf{r}_{i\pm\frac{J-1}{2}}} J(\mathbf{r}_i, \mathbf{r}_{i\pm1}, ..., \mathbf{r}_{i\pm\frac{J-1}{2}}), \quad (6)$$

$$s.t. \quad \mathbf{r}_i^T\mathbf{r}_i = \ldots = \mathbf{r}_{i\pm\frac{J-1}{2}}^T\mathbf{r}_{i\pm\frac{J-1}{2}} = 1$$

where

$$J(\mathbf{r}_i, \mathbf{r}_{i\pm1}, ..., \mathbf{r}_{i\pm\frac{J-1}{2}}) = \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \frac{1}{2} \left\| \mathbf{s}_{i+k} - W\mathbf{r}_{i+k} \right\|_2^2 + \lambda_0 \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \mathcal{R}_\epsilon(\mathbf{r}_{i+k}) + \sum_{k=-\frac{J-1}{2},k\neq 0}^{\frac{J-1}{2}} \frac{1}{2} \lambda_k \left\| \mathbf{r}_i - \mathbf{H_k}\mathbf{r}_{i+k} \right\|_2^2 \quad (7)$$

and

$$\mathcal{R}_\epsilon(\mathbf{r}) = \sum_j (\sqrt{r_j^2 + \epsilon^2} - \epsilon). \quad (8)$$

For small ϵ, such as $\epsilon = 0.01$ [31], the regularization parameter $\mathcal{R}_\epsilon(\mathbf{r})$ is a smoothed ℓ_1 norm approximation that promotes sparsity of the solution (also called hybrid $\ell_1 - \ell_2$ or hyperbolic penalty [32]). $\mathcal{R}_\epsilon(\mathbf{r})$ is also used for seismic blind deconvolution in [31]. The use of the hybrid $\ell_1 - \ell_2$ norm, which is differentiable, rather than the ℓ_1 norm, enables to use simple optimization techniques such as steepest descent method.

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

$$\mathcal{L}(\mathbf{r}_i, \mathbf{r}_{i\pm1}, ..., \mathbf{r}_{i\pm\frac{J-1}{2}}) = J(\mathbf{r}_i, \mathbf{r}_{i\pm1}, ..., \mathbf{r}_{i\pm\frac{J-1}{2}}) - \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \frac{\eta_k}{2}(\mathbf{r}_{i+k}^T \mathbf{r}_{i+k} - 1) \tag{9}$$

with Lagrange multipliers given by the scalars η_k . The minimization must satisfy

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_{i+k}} = \mathbf{g}_{i+k} - \eta_{i+k}\mathbf{r}_{i+k} = 0, \qquad -\frac{J-1}{2} \leq k \leq -\frac{J-1}{2} \tag{10}$$

where $\mathbf{g}_{i+k} = \frac{\partial J}{\partial \mathbf{r}_{i+k}}$.
Multiplying (10) by $\mathbf{r}_{i+k}^T$ and using the constraint $\mathbf{r}_{i+k}^T \mathbf{r}_{i+k} = 1$ yields

$$\eta_{i+k} = \mathbf{r}_{i+k}^T \mathbf{g}_{i+k}. \tag{11}$$

Then, the projection of the gradient on the unit sphere can be expressed via

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_{i+k}} = \mathbf{g}_{i+k} - \mathbf{r}_{i+k}^T \mathbf{g}_{i+k} \mathbf{r}_{i+k}. \tag{12}$$

The classical update rule of steepest descent algorithm is given by

$$\mathbf{r}_{i+k,l+1} = \mathbf{r}_{i+k,l} - \mu_l \mathbf{h}_{i+k,l}$$

with a normalized gradient

$$\mathbf{h}_{i+k,l} = \frac{\partial \mathcal{L}}{\partial \mathbf{r}_{i+k}} / |\frac{\partial \mathcal{L}}{\partial \mathbf{r}_{i+k}}|.$$

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

As in [31], it should be mentioned that we must initialize the steepest-descent algorithm by a solution that is close to the final reflectivity. Since the data is structurally close to the true sparse reflectivity, we can use it as an initial solution. Practically, this choice is advantageous and resolves into a sparse estimate of the reflectivity.

4. Experimental results

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

Synthetic Data

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

For $J = 2$ the above optimization problem reduces to:

$$\min_{\mathbf{r}_i, \mathbf{r}_{i+1}} \frac{1}{2} \|\mathbf{s}_i - \mathbf{W}\mathbf{r}_i\|_2^2 + \frac{1}{2} \|\mathbf{s}_{i+1} - \mathbf{W}\mathbf{r}_{i+1}\|_2^2 + \lambda_0 \{\|\mathbf{r}_i\|_1 + \|\mathbf{r}_{i+1}\|_1\} + \lambda_1 \frac{1}{2} \|\mathbf{r}_i - \mathbf{H_1}\mathbf{r}_{i+1}\|_2^2 . \tag{13}$$

For $J = 3$ the above optimization problem is:

$$\begin{aligned} &\min_{\mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}} \frac{1}{2} \|\mathbf{s}_{i-1} - \mathbf{W}\mathbf{r}_{i-1}\|_2^2 + \frac{1}{2} \|\mathbf{s}_i - \mathbf{W}\mathbf{r}_i\|_2^2 + \frac{1}{2} \|\mathbf{s}_{i+1} - W\mathbf{r}_{i+1}\|_2^2 \\ &+\lambda_0 \{\|\mathbf{r}_{i-1}\|_1 + \|\mathbf{r}_i\|_1 + \|\mathbf{r}_{i+1}\|_1\} \\ &+\lambda_{-1} \frac{1}{2} \|\mathbf{r}_i - \mathbf{H_1}\mathbf{r}_{i-1}\|_2^2 + \lambda_1 \frac{1}{2} \left\|\mathbf{r}_i - \mathbf{H_{-1}}\mathbf{r}_{i+1}\right\|_2^2 . \end{aligned} \tag{14}$$

These schemes were tested for different values of λ_0, λ_1 and λ_{-1}, with SNR = 10 dB and 5 dB. As was mentioned before, the tradeoff parameter λ_0 balances the reflectivity sparseness and the minimization of the residual term. Increasing λ_0 decreases the sparsity of the solution, whereas decreasing λ_0 may lead to noise amplification. The tradeoff parameters $\lambda_{\pm 1}$ promote smoothness of the reflectivity in the horizontal direction.

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

In MSSI-2, the minimization is performed by taking into account the distance between each reflectivity column and the subsequent reflectivity column. In each step, we estimate two adjacent columns simultaneously. Even though two reflectivity columns estimates were obtained, we keep only the current reflectivity column estimate. The estimate of the subsequent column is discarded, since this column will be estimated with its subsequent column in the next step. In MSSI-3, the minimization is performed by taking into account the distances between each reflectivity column and both the preceding and subsequent reflectivity columns. In each step, we estimate three adjacent columns simultaneously. Out of the three obtained estimates, only the middle reflectivity column is kept. The estimates of the preceding and the subsequent columns are discarded.

With different experiments, we concluded that the best results are achieved when

$H_{\pm 1}$ is a convolution matrix of a 3 taps averaging filter,

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

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

To analyze the stability of the method under different levels of noise, we generated 20 different realizations of 2D reflectivities of size 76×98, one example is shown in Fig. 1(a). We then convolved it with a 27 samples long Ricker wavelet and added white Gaussian noise with SNRs of 10 dB and 5 dB. Two of the realizations with SNRs of 10 dB and 5 dB are shown in Fig. 1(b) and Fig. 1(c), respectively. The seismic wavelet is shown in Fig. 1(d).

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

$$\rho_{\hat{\mathbf{r}}\mathbf{r}} = \frac{\hat{\mathbf{r}}^T \mathbf{r}}{\|\hat{\mathbf{r}}\|_2 \|\mathbf{r}\|_2} \tag{15}$$

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

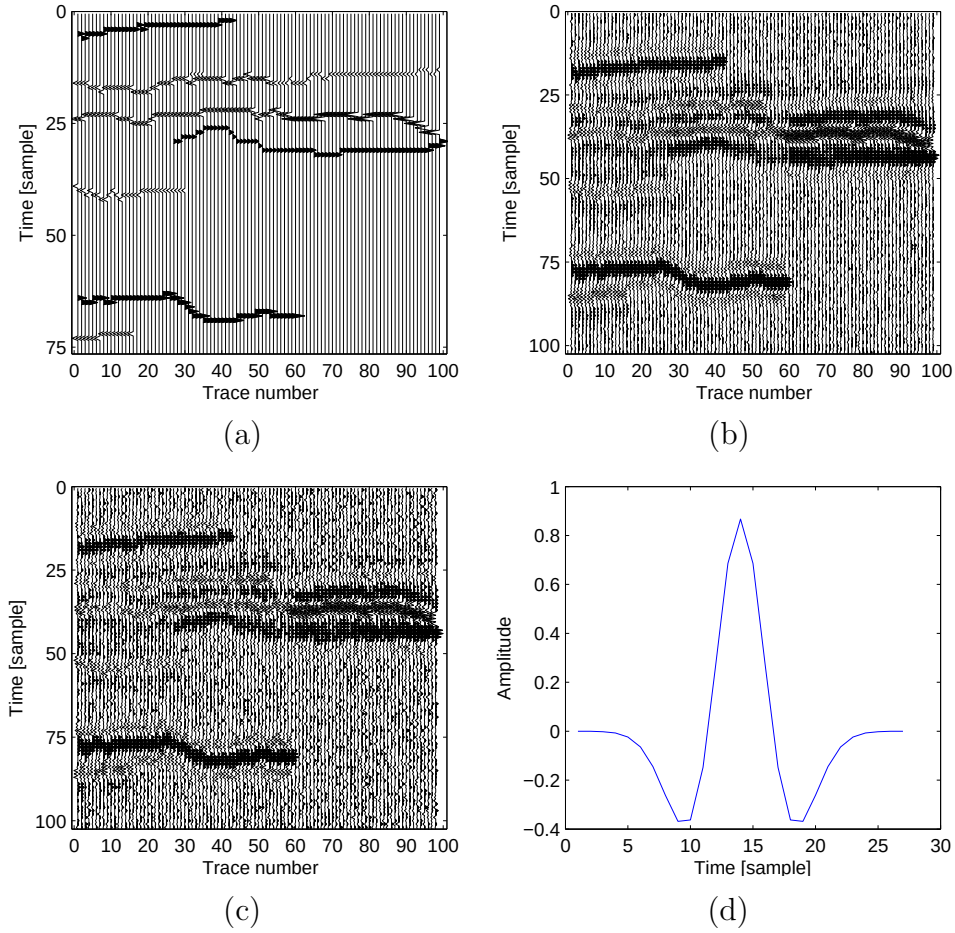

Figure 1. Synthetic reflectivity, wavelet and data sets: (a) Synthetic 2D reflectivity section; (b) 2D seismic data (SNR = 10 dB); (c) 2D seismic data (SNR = 5 dB); (d) Wavelet.

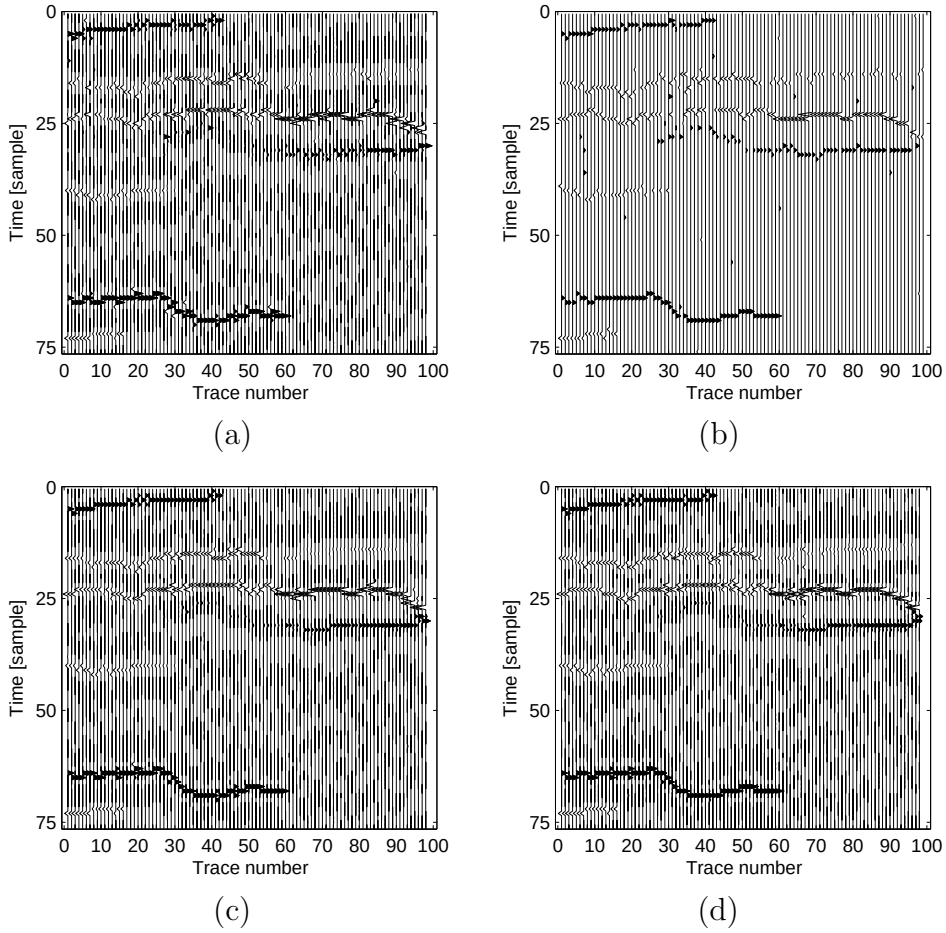

Figure 2. Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 10 dB; (b) MC-II deconvolution results for SNR = 10 dB; (c) MSSI-2 results for SNR = 10 dB; (d) MSSI-3 results for SNR = 10 dB.

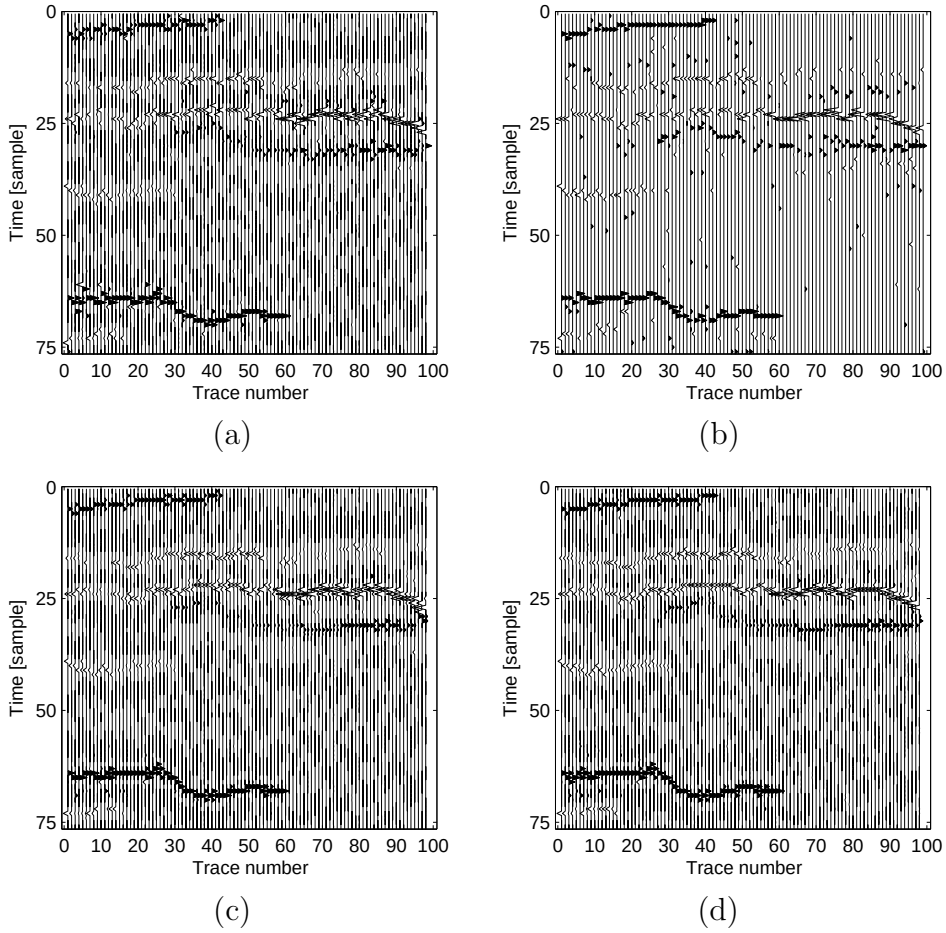

Figure 3. Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 5 dB; (b) MC-II deconvolution results for SNR = 5 dB; (c) MSSI-2 results for SNR = 5 dB; (d) MSSI-3 results for SNR = 5 dB.

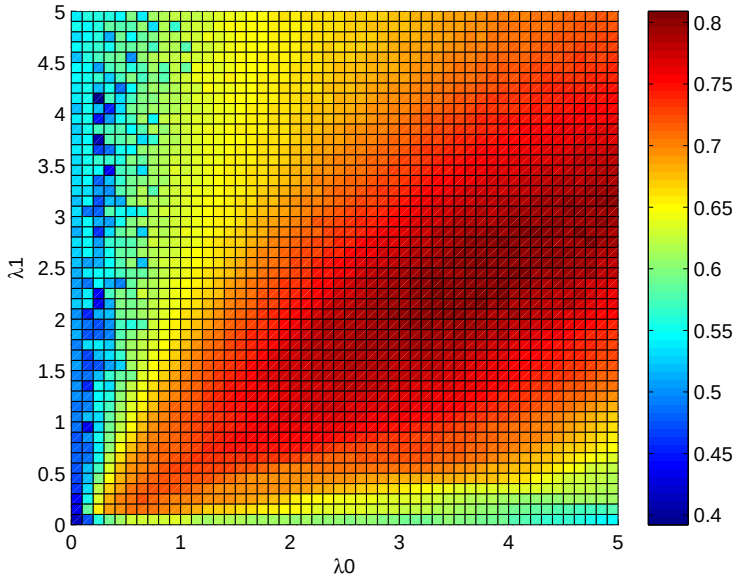

Figure 4. Correlation coefficient vs. deconvolution parameters λ_1 and λ_0 for synthetic 2D data deconvolution (SNR = 5 dB).

We compare our results to a single channel deconvolution (SSI) and to the multichannel deconvolution algorithm described in [28] (MC-II). The estimated reflectivities, obtained by SSI, MC-II, and by MSSI, for the seismic data with SNR of 10 dB, are shown in Fig. 2. The SSI method was implemented by simply assigning $\lambda_{\pm 1}$ to zero, using the best estimated λ_0 for SSI. For this example, the correlation coefficients between the original reflectivity and the estimated reflectivity, with our method, are $\rho = 0.88$ and $\rho = 0.9$ for $J = 2$ and $J = 3$, respectively, whereas the correlation coefficient achieved by single-channel deconvolution is $\rho = 0.78$, and $\rho = 0.86$ for MC-II. The best results in terms of correlation coefficients were achieved with $\lambda_0 = 3.1$ and $\lambda_1 = 0.9$ for the two channel implementation, with $\lambda_0 = 2.8, \lambda_{-1} = 1$ and $\lambda_1 = 0.7$ for the three-channel implementation, and with $\lambda_0 = 2.8$ for SSI. Practically, the values of λ_0 and $\lambda_{\pm 1}$ are data dependent and determined empirically. The best result is not necessarily achieved by setting $\lambda_1 = \lambda_{-1}$.

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

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

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

The series of synthetic tests that we have performed during our research indicate that the optimal correlation can be achieved using different λ_0 and $\lambda_{\pm 1}$ values, depending on the channel characteristics: the number of reflectors, the layers' thicknesses (distances between reflectors), the channel sparsity, and the signal-to-noise ratio (SNR). It is recommended that λ_1 and λ_{-1} values will not be too large so as to avoid over-smoothing of the estimated reflectivity. The parameters can be chosen by inspecting the correlation coefficient of a few columns.

Fig. 4 presents the correlation coefficient values as a function of λ_0 and λ_1, for 10 columns of the seismic data with SNR of 5 dB, depicted in Fig. 1(c). As can be seen, there is an area of values that gives the best results. This implies that the user does not have to know the exact value of the regularization parameters in order to get a good recovery. The correlation coefficients for $\lambda_1 = 0$, which represent the single-channel scores are significantly smaller than the values achieved by a non-zero value of λ_1. This implies that the MSSI outperforms the single-channel method (SSI).

The average processing times of a data set of size 76×98 on Intel(R)Core(TM)i5-4430 CPU @3GHz, by Matlab implementations of the single-channel and the proposed algorithms - MSSI-2 and MSSI-3 are 1.18, 1.41 and 1.57 minutes, respectively.

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

the improvement is getting smaller as the SNR increases, meaning that all algorithms perform better when the noise level is lower.

Real Data

We applied the proposed deconvolution scheme, to real seismic data from a small land 3D survey in North America (courtesy of GeoEnergy Inc., TX) of size 350×200, shown in Fig. 5(a).The assumed wavelet is shown in Fig. 5(b). The reflectivity sections obtained by single-channel deconvolution, by MC-II, by MSSI-2 and by MSSI-3 are shown in Fig.6(a),(b),(c) and (d), respectively. The seismic data reconstructed as a convolution between the estimated reflectivity and a given wavelet, are shown in Fig. 6(e),(f),(g) and (h). Visually comparing these reflectivity sections, it can be seen that the layer boundaries in the estimates obtained by MSSI are more continuous and smooth than the layer boundaries in the single-channel deconvolution estimates. Moreover, MSSI also detects parts of the layers that the single-channel deconvolution misses. It can also be seen that the reconstructed seismic data obtained by MSSI is more accurate than the one obtained by SSI. Since the ground truth is unknown, to asses the performance of the methods, we calculate the correlation coefficient between the reconstructed data to a noise-free seismic data. The obtained correlation between the original and reconstructed seismic data for MSSI is $\rho_{s,\hat{s}} = 0.9$ when $\lambda_0 = 9$ and $\lambda_1 = 30$ for MSSI-2, and $\rho_{s,\hat{s}} = 0.91$ when $\lambda_0 = 9$ and $\lambda_{\pm 1} = 28$ for MSSI-3. Whereas for SSI we get $\rho_{s,\hat{s}} = 0.89$ when $\lambda_0 = 5$, and for MC-II we have $\rho_{s,\hat{s}} = 0.76$. The parameters for all methods were chosen to best fit the observed data using the correlation of a few columns. The estimates produced by MSSI-2 and MSSI-3 are quite close, though the latter manages to recover a slightly more continuous image.

As mentioned before, experimental results show that the best results are achieved when $H_{\pm 1}$ is a convolution matrix of a 3 taps averaging filter, which means that we assume that spikes of two neighboring traces are close by less than 3 samples. This hypothesis might seem very restrictive in the case of real data. However, $H_{\pm 1}$ as a convolution matrix of a 3 taps only averaging filter outperforms other filter choices, for the reason that this choice balances between the ability to detect layers' discontinuities and more complex structure and at the same time also to create a smooth reflectivity image. This is a great advantage compared to other existing methods. Choosing a longer filter causes over-smoothing of the recovered reflectivity and blurring of natural breaks in the earth structure. Choosing the lateral derivative instead of the third term as defined in (7) would encourage horizontal lines ignoring the subsurface curves structure.

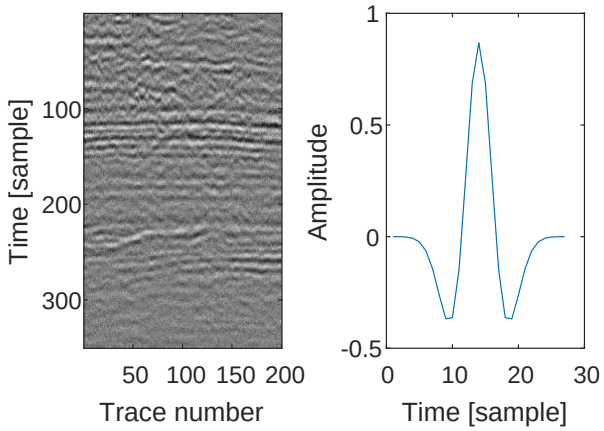

Figure 5. Real data and assumed wavelet: (a) Real seismic data (SNR = 5 dB); (b) wavelet.

5. Conclusions

We have presented a multichannel deconvolution algorithm in seismic applications. The algorithm both promotes sparsity of the solution and also takes into consideration the spatial dependency between neighboring traces in the deconvolution process. We have demonstrated that our deconvolution results are visually superior, compared to a single-channel deconvolution algorithm, for synthetic and real data, under sufficiently high SNR. Our second implementation (MSSI-3) performs better, on both synthetic and real data. The reason for that is that MSSI-3 takes into account more information from neighboring traces in the deconvolution process of each trace, compared to the first implementation (MSSI-2) that uses information from only one neighboring trace. The improved performance of the proposed algorithm compared to the single-channel algorithm was also apparent in qualitative assessment. It also shows that the second implementation's results are more accurate.

The choice of regularization parameters is still an open problem. The use of a too small λ_0 could result in an increased resolution of the estimated reflectivity which is not necessarily real. In addition, one needs to find the correct balance between all regularization parameters. It should also be mentioned that in this study we used a time-spatial-invariant known wavelet for simplicity. In practice, a time and spatial varying wavelet could produce better results, taking into account wave propagation effects, such as attenuation and dispersion.

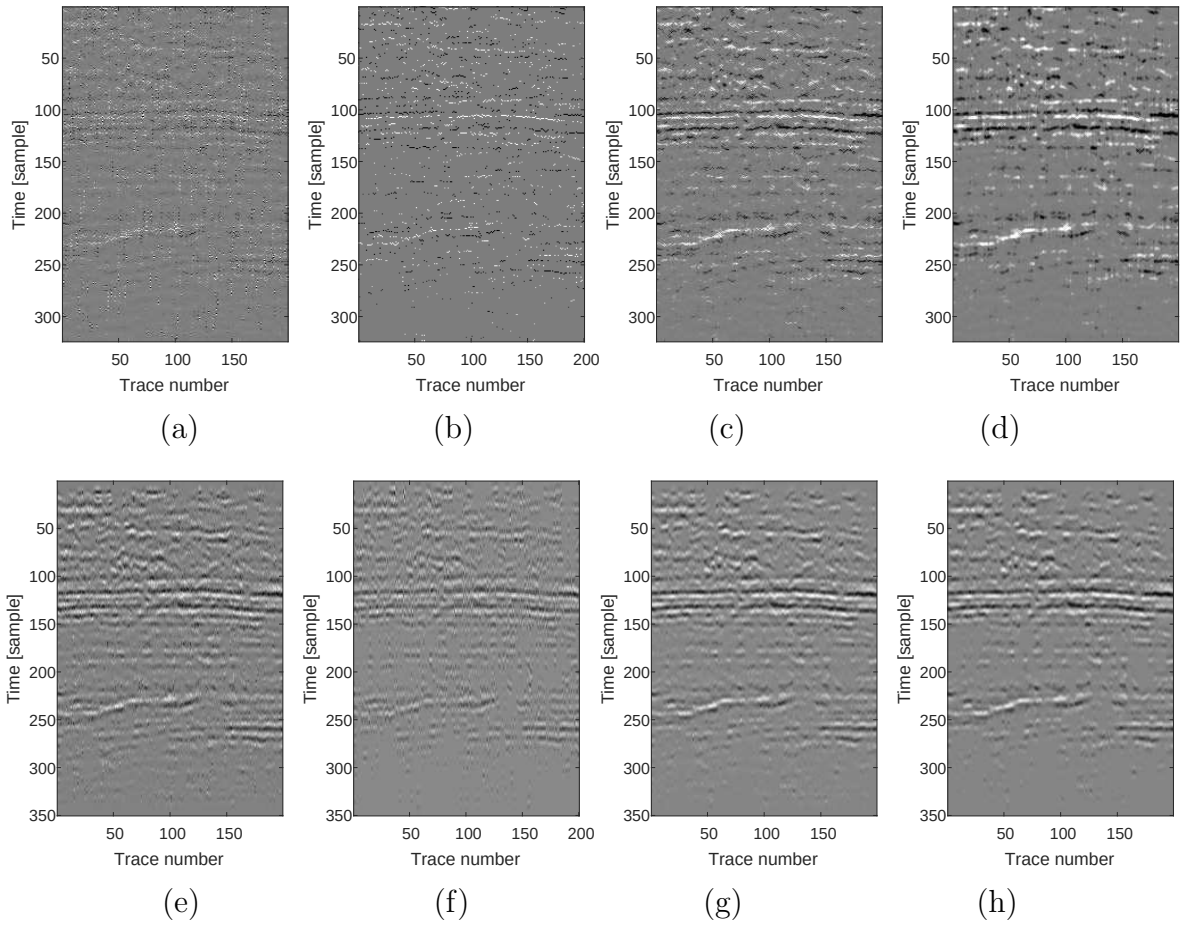

Figure 6. Real data deconvolution results: (a) Single-channel estimated reflectivity; (b) MC-II estimated reflectivity ;(c) MSSI-2 estimated reflectivity; (d) MSSI-3 estimated reflectivity; (e) Single-channel reconstructed data; (f) MC-II reconstructed data; (g) MSSI-2 reconstructed data; (h) MSSI-3 reconstructed data.

References

[1] Berkhout A 1986 *Proceedings of the IEEE* **74** 1133–1159
[2] Taylor H L, Banks S C and McCoy J F 1979 *Geophysics* **44** 39–52
[3] Nguyen T and Castagna J 2010 *Journal of Seismic Exploration* **19** 303–320
[4] Zhang R and Castagna J 2011 *Geophysics* **76** 147–158
[5] Gholami A and Sacchi M D 2012 *IEEE Trans. Geosci. Remote Sens.* **50** 4105–4116
[6] Rozenberg I, Cohen I and Vassiliou A 2014 Performance measures for sparse spike inversion vs. basis pursuit inversion *Proceedings of the 28th IEEE Convention of Electrical and Electronics Engineering in Israel* vol 1 pp 1–5
[7] Ulrych T 1971 *Geophysics* **36** 650–660
[8] Wiggins R 1978 *Geoexploration* **16** 21–35
[9] Riel P V and Berkhout A J 1985 *Geophysics* **50** 1440–1455
[10] Pereg D and Ben-Zvi D 2016 Blind deconvolution via maximum kurtosis adaptive filtering *Proceedings of 2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)* pp 1–5
[11] Stockham T, JR, Cannon T and Ingebretsen R 1975 *Proceedings of the IEEE* **63** 678–692
[12] Pham M Q, Duval L, Chaux C and Pesquet J C 2014 *IEEE Trans. Signal Process.* **62** 4256–4269
[13] Chen S S, Donoho D L and Saunders M A 2001 *SIAM Rev.* **43** 129–159
[14] Elad M 2010 *Sparse and Redundant Representations* (Springer)
[15] Dossal C and Mallat S 2005 Sparse spike deconvolution with minimum scale *Signal Processing with Adaptive Sparse Structured Representations (SPARS)* pp 123–126
[16] Zhang R, Sen M K and Srinivasan S 2013 *Journal of Geophysics an Engineering* **10** 035012
[17] Idier J and Goussard Y 1993 *IEEE Trans. Geoscience, Remote sensing* **31** 961–979
[18] Mendel J, Kormylo J, Aminzadeh F, Lee J and FHabibi-Ashrafi 1981 *Geophysics* **46** 1398–1414
[19] Kormylo J and Mendel J 1982 *IEEE Trans. Information Theory* **28** 482–488
[20] Kaaresen K and Taxt T 1998 *Geophysics* **63** 2093–2107
[21] Kaaresen K 1997 *IEEE Trans. Signal Processing* **45** 1173–1183
[22] Heimer A, Cohen I and Vassiliou A 2007 Dynamic programming for multichannel blind seismic deconvolution *Proc. Society of Exploration Geophysicist International Conference, Exposition and 77th Annual Meeting* (San Antonio) pp 1845–1849
[23] Heimer A and Cohen I 2008 *Signal Processing* **88** 1839–1851
[24] Amini A, Weymouth T and Jain R 1990 *IEEE Trans. Pattern Anal. Mach. Intell.* **12** 855–867
[25] Buckle M and Yang J 1997 *Pattern Recognition Lett. 18* 621–629
[26] Heimer A and Cohen I 2009 *IEEE Trans. Geoscience and Remote Sensing* **47** 2047–2058
[27] Forney D 1973 The viterbi algorithm *Proc. IEEE* pp 268–278
[28] Ram I, ICohen and Raz S 2010 *IEEE Trans. Signal Processing* **58** 2757–2769
[29] Chalmond B 1989 *Pattern Recognition* **22** 747–761
[30] Tibshirani R J 2013 *Electronic Journal of Statistics* **7** 1456–1490
[31] Kazemi N and Sacchi M 2014 *Geophysics* **79** 143–152
[32] Repettil A, Pham M Q, Duval L, Chouzenoux E and Pesquet J C 2015 *IEEE Signal Process. Lett.* **22** 539–543