Dynamic Programming for Multichannel Blind Seismic Deconvolution

Alon Heimer and Israel Cohen, Technion – Israel Institute of Technology, and Anthony A. Vassiliou, GeoEnergy*

SUMMARY

We present an algorithm for multichannel blind deconvolution of seismic signals, which exploits lateral continuity of earth layers by dynamic programming approach. We assume that reflectors in consecutive channels, related to distinct layers, form continuous paths across channels. We introduce a quality measure for evaluating the quality of a continuous path, and iteratively apply dynamic programming to find the best continuous paths. The improved performance of the proposed algorithm and its robustness to noise, compared to a competitive algorithm, are demonstrated using simulated and real seismic data examples.

INTRODUCTION

Blind deconvolution is generally an ill-posed problem, and requires some a priori information about the channels or the wavelet. The reflectivity sequence is often modeled as a Bernoulli-Gaussian random sequence, and second-order statistics may be used to partially reconstruct the input signal. Several methods based on high-order statistics have been developed (Mendel, 1991; Lazear, 1993), which require very long data to properly estimate the output statistics. Alternatively, the wavelet can be modeled as an autoregressive moving-average (ARMA) process, and a maximum likelihood estimator for the reflectivity can be derived (Mendel et al., 1981).

Multichannel blind deconvolution is often more advantageous and more robust than single-channel blind deconvolution (see Tong and Perreau (1998) and references therein, Xu et al. (1995); Luo and Li (1998)). Sparsity of the reflectivity sequences may be used to cope with the ill-posed nature of the basic blind deconvolution problem (Kaaresen and Taxt, 1998; Jeffs, 1998), and to improve the performance of non-blind deconvolution methods (Perros-Meilhac et al., 2001). Channel sparsity has been used in Kaaresen and Taxt (1998), together with the assumption of short wavelet, to formulate an efficient channel estimation method suitable for relatively short traces (see also Kaaresen (1998)). Lateral continuity of the reflectors across channels is also used to further improve the channel estimates. Idier and Goussard (1993) model the 2-D structure of the underground reflectivity as a Markov-Bernoulli random field, and impose lateral continuity to generate deconvolution results that are far superior to those obtainable by single-channel deconvolution methods.

In this paper, which summarizes the results in (Heimer and Cohen, 2007), lateral continuity of reflectors across channels is combined with the blind deconvolution algorithm of Kaaresen and Taxt (1998). We employ dynamic programming (Amini et al., 1990; Buckley and Yang, 1997) to find the shortest continuous paths of reflectors across channels, and develop an improved multichannel blind deconvolution algorithm for seismic signals, which exploits the lateral continuity of earth layers. Rather than measuring the increase in the fit to the data each single reflector yields, versus the decrease in sparsity of the channel estimates, we measure the increase in the fit to the data obtained by a complete continuous path of reflectors, versus the decrease in the sparsity of paths. The increase in the fit to the data achieved by a continuous path of reflectors is calculated as the sum of contributions to the fit from all reflectors in that path. The improved performance of the proposed algorithm and its robustness to noise, compared to the blind deconvolution algorithm of Kaaresen and Taxt, are demonstrated by using simulated and real seismic data examples.

SIGNAL MODEL

We assume M received signals (traces), each generated by a single input signal $h[n]$ passing through a channel $x^{(m)}[n]$ and corrupted by additive uncorrelated noise $e^{(m)}[n]$. The output signal of channel m can be written as

$$z^{(m)}[n] = \sum_{k=0}^{K-1} h[k]\, x^{(m)}[n-k] + e^{(m)}[n] \tag{1}$$

where $m = 1, \ldots, M$, and $n = 1, \ldots, N$. We assume the following:

1. All channels are excited by the same wavelet h.
2. The wavelet h has a finite support of length K, which is shorter than the channel.
3. Each channel is sparse, i.e., the number of non-zero elements (reflectors) in a channel is small relative to the channel's length.
4. The element in a channel are independent and identically distributed with zero-mean Bernoulli-Gaussian distribution.
5. Reflectors in consecutive channels form continuous paths across channels.
6. The noise $e^{(m)}[n]$ is white, Gaussian, and independent of $h[n]$ and $x^{(m)}[n]$.

The third assumption makes it useful to write (1) in the following way:

$$z^{(m)}[n] = \sum_{p=1}^{P} h[n-n_{m,p}]\, a_{m,p} + e^{(m)}[n] \tag{2}$$

where $n_{m,p}$ is the discrete time of reflection p in channel m, and $a_{m,p}$ is its amplitude. The matrix representation of (2) is given by

$$\mathbf{z}^{(m)} = \mathbf{H}^{(m)}\mathbf{a}^{(m)} + \mathbf{e}^{(m)} \tag{3}$$

where $\mathbf{z}^{(m)} = \begin{bmatrix} z^{(m)}[0] & z^{(m)}[1] & \ldots & z^{(m)}[N] \end{bmatrix}^T$, $\mathbf{H}^{(m)}$ is a matrix with P columns as the number of reflections in channel m, $\mathbf{H}^{(m)}_{np} = h[n-n_{m,p}]$, and $\mathbf{e}^{(m)} = \begin{bmatrix} e^{(m)}[0] & e^{(m)}[1] & \ldots & e^{(m)}[N] \end{bmatrix}^T$. For later use we adopt the notation in Kaaresen and Taxt (1998) and define the matrix $\mathbf{H}^{w}$ as the columns of $\mathbf{H}$ corresponding to reflectors inside a certain time window w, $\mathbf{a}^{w}$ represents the amplitudes of the reflectors inside the window, $\mathbf{H}^{\overline{w}}$ contains the columns of $\mathbf{H}$ corresponding to reflectors outside the window, and $\mathbf{a}^{\overline{w}}$ represents the amplitudes of the reflectors outside the window (the channel index m is omitted for convenience).

DYNAMIC PROGRAMMING FOR FINDING THE SHORTEST CONTINUOUS PATH

Dynamic programming is an effective way to find a global minimum in some nonconvex optimization problems. We now briefly describe the problem of finding the shortest continuous path across an image, and the solution by dynamic programming (Buckley and Yang, 1997).

Problem Formulation

Assume we have a gray level image of size $N \times M$. The problem is to find a "path", i.e., a sequence of M pixels $\{(n_m, m)\}_{m=1}^{M}$ one in each column, such that the following two conditions are satisfied:

1. $|n_m - n_{m+1}| \leq d$ for all $m = 1, 2, ..., M-1$, where d is some small positive integer constant. We refer to this condition as the "continuity condition", and to a path for which this condition is satisfied as a continuous path.

2. The path is the "shortest" among all continuous paths, in the sense that a certain length measure is minimized for this path among all the continuous paths.

A common length measure for a given path is the sum of gray levels of the pixels, which the path passes through. In this work we are interested in a length criterion that measures not only the total intensity of the path (as represented by the sum of gray levels) but also the degree to which the gray levels along the path have the same sign (since all reflectors representing the same underground layer have the same sign). One way to achieve that is to use a length criterion that combines the total variation of the path and the sum of gray levels. The total variation is defined as

$$TV = \sum_{m=1}^{M-1} |p(n_m, m) - p(n_{m+1}, m+1)|$$

where $p(n,m)$ is the gray level of pixel (n,m). The length measure in this case is $S = \mu\, TV + \sum_{m=1}^{M} p(n_m, m)$, where μ is some weight that controls the tradeoff between sum of gray levels and total variation along the path.

Extraction of Shortest Path

We now describe the dynamic programming algorithm using the sum of gray levels as a length measure. Let (n,m) be the coordinates of a pixel in row n and column m, let $P_{n,m} = \{(n_1,1),(n_2,2),...,(n,m)\}$ denote a continuous path starting at column 1 and ending at pixel (n,m), and let $p(n,m)$ represent the value of the pixel (n,m). Denote by $P^o_{n,m}$ the continuous path that starts at column 1 and ends at pixel (n,m) with minimum sum of pixels values, and by $S^o_{n,m}$ the sum of pixels values along the path $P^o_{n,m}$. The algorithm starts with the following initialization:

$$P^o_{n,1} = (n,1) \quad \text{for } n = 1,2,\ldots,N \qquad (4)$$
$$S^o_{n,1} = p(n,1) \quad \text{for } n = 1,2,\ldots,N\,. \qquad (5)$$

Then for each column $m = 2,\ldots,M$ compute

$$S^o_{n,m} = p(n,m) + \min_{n-d \leq k \leq n+d} S^o_{k,m-1} \quad \text{for } n = 1,2,\ldots,N \qquad (6)$$
$$P^o_{n,m} = \left(P^o_{k^o,m-1} \vee (n,m)\right) \quad \text{for } n = 1,2,\ldots,N \qquad (7)$$

where k^o denotes the value of k achieving the minimum in (6), and the sign $\vee$ means adding the pixel on its right to the end of the path. After the column M is processed, we obtain the optimal path P_{opt} by

$$n^o = \arg\min_n S^o_{n,M} \qquad (8)$$
$$P_{opt} = P^o_{n^o,M}\,. \qquad (9)$$

MULTICHANNEL BLIND IDENTIFICATION

We assume that any nonzero element in a channel is a member of a path of nonzero elements across the channels, which satisfies the continuity condition (Condition 1). In the basic algorithm, a decision whether or not to add a reflector to a channel estimate, is made according to the increase in the fit to the data compared to the decrease in the sparsity measure. In the multichannel version of the algorithm, a measure of local continuity is added to the decision criterion, which encourages reflectors that are members of a local continuous sequence. The approach presented here allows the process to add, not single reflectors, but complete continuous paths of reflectors to the image of channel estimates. Rather than measuring the increase in the fit to the data each single reflector yields, versus the decrease in sparsity of the channel estimate, we measure the increase in the fit to the data obtained by a complete continuous path of reflectors, versus the decrease in sparsity of paths, i.e., number of continuous reflector paths. The increase in the fit to the data achieved by a path of reflectors is calculated as the sum of contributions to the fit from all reflectors in the path.

Let us assume a certain continuity parameter d for which the paths of reflectors in the true channels are continuous. The steps of the new algorithm are as follows:

Step 1 - Initialization:
The initialization is carried out by finding several continuous paths (with parameter d) along which the sum of gray levels is maximal or minimal among all continuous paths in the same region of the data:

1.1) Use dynamic programming to find the continuous path of minimal length in the original image.

1.2) Similarly, find the path of maximal length (can be done by finding the path of minimal length in the image multiplied by -1).

1.3) Choose the path from the above two steps for which the absolute value of the length is maximal and add it to the channel estimate.

1.4) Let $S \triangleq \{(n,m) : |n - n_m| < D\}$ be a strip of a certain predetermined width D around the path $\{(n_m,m)\}_{m=1}^{M}$ found in the previous step. Then, for finding the next extremum path, we replace the values of the true data in that strip by zeros, and subsequently extract the extremum path (the received signal z is randomly distributed with zero mean, and therefore replacing the data values with zeros ensures that this strip will not be a part of the following paths in the initialization step, since zeros are not contributing to an extremum value).

1.5) If the number of paths found so far is less than a predetermined number, go to step 1.1.

At the end of the initialization step, we replace all the zero values, inserted in Step 1.4, back to their original values.

Step 2 - Wavelet Estimation:
After initialization of the channel estimate, the wavelet estimate is obtained in the same manner as in the basic algorithm, which is simply the least squares fit to the data given the current channel estimate and the received signals.

Step 3 - Channel Estimation:
The channel estimates are updated by examining and comparing a few alternatives that include adding a new path of reflectors, and removing or translating an existing path. Since the number of possible paths is much larger than the number of possible single reflectors, we use again dynamic programming to find the continuous path that maximally increases the fit to the data. That path only is examined as a candidate for inclusion or substitution of other paths. In order to find a path of reflectors that maximally increases the fit to the data, given all existing reflectors, we execute the following two steps procedure:

Step 3.1: For each channel, calculate the contribution of adding a reflector at each possible location at that channel. The "Quality" of placing a new reflector at time t is calculated as follows:

3.1.1) Temporarily place a new reflector at time t.

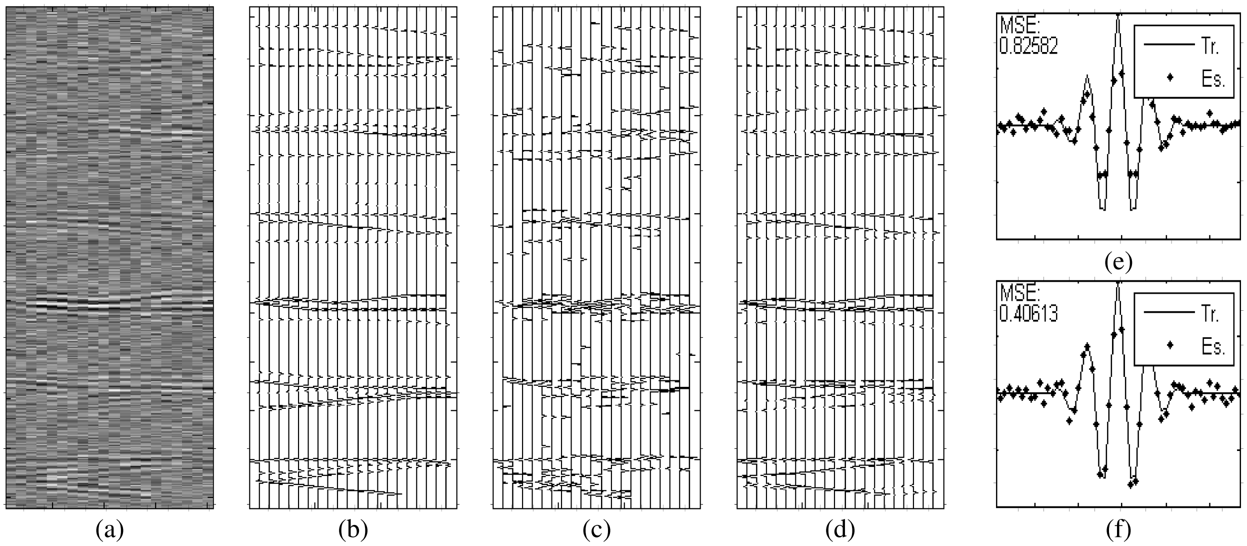

Figure 1: Results of multichannel blind deconvolution obtained by using the basic and new algorithms in a high noise situation. (a) Received noisy traces (SNR $= -10$ dB); (b) True channels; (c) channels estimates obtained by using the basic algorithm; (d) channels estimates obtained by using the proposed algorithm; (e) wavelet estimate obtained by using the basic algorithm; (f) wavelet estimate obtained by using the proposed algorithm.

3.1.2) Define a window centered at time t.

3.1.3) Calculate the contribution to the data made by existing reflectors outside the window, using the current wavelet estimate, and subtract it from the data.

3.1.4) Find the amplitudes of reflectors inside the window giving the best fit to the remainder of the received signal z_r , given their locations, and the current wavelet estimate.

3.1.5) Calculate a new approximation to the entire received signal $\mathbf{z}$ using existing reflectors outside the window, and the new reflectors inside the window from (3.1.4):

$$\widehat{\mathbf{z}} = \widehat{\mathbf{H}}^{\overline{w}} \widehat{\mathbf{a}}^{\overline{w}} + \widehat{\mathbf{H}}^{w} \widehat{\mathbf{a}}^{w}. \tag{10}$$

3.1.6) Calculate the difference between the data and its new approximation:

$$\mathbf{z}_e = \mathbf{z} - \widehat{\mathbf{z}}. \tag{11}$$

3.1.7) Our general goal is to minimize $\|\mathbf{z}_e\|^2$, but we would also like all the reflectors along the best path to have the same sign. Hence, we define the "quality" of adding a reflector at time t as the sign of this reflector (calculated in Step 3.1.4) multiplied by some positive function whose maximization is equivalent to minimizing $\|\mathbf{z}_e\|^2$, e.g.,

$$Q(t) = sign\left[\widehat{\mathbf{a}}^{w}(t)\right]\left(L - \mathbf{z}_e^T \mathbf{z}_e\right) \tag{12}$$

where L is some large constant (selected so that it is greater than $\mathbf{z}_e^T \mathbf{z}_e$ throughout the algorithm). The definition of the quality in (12) allows us to seek the path with maximum qualities (path with positive reflectors mostly increasing the fit to the data) and path of minimum qualities (path with negative reflectors mostly increasing the fit to the data).

Step (3.1) generates an image of "qualities" which we will refer to as the "qualities image", representing the improvement of the fit to the data achieved by placing a new reflector at each possible location at each channel given the existing reflectors.

Step 3.2: Now we employ dynamic programming to find two continuous paths (with parameter d) in the qualities image. The first path is characterized by minimal sum of qualities, while the second path is associated with maximal sum of qualities. Out of these two paths we select the one with larger absolute value of quality sums, since this path maximally increases the fit to the data. The total contribution of the best path is $\left|S^o_{n^o,M}\right|$. After finding the best path, given the existing reflectors and the current wavelet estimate, a possible option is to continue with the same procedure as described in Kaaresen and Taxt (1998), i.e., calculate a quality criterion combining the increase in the fit to the data due to the additional path and the decrease in the sparsity of paths. Here we use a slightly different approach, assuming some a priori knowledge of the reflectors probability density, or the expected number of reflections in the data. As long as the channel estimate includes less than this expected number of reflections, we add new paths. Each additional path is inserted into a FIFO list, and when we reach the desired number of reflectors, we sequentially perform the following steps:

3.2.1) Remove the next existing path from the beginning of the list and from the channel estimate.

3.2.2) Find a new best path.

3.2.3) Add the new path to the end of the list and to the channel estimate.

This procedure is similar to the one described in Kaaresen and Taxt (1998), since we actually move each path to a different better location. If a path current location is optimal, then the path is selected again at the same location. The advantage of this procedure is that tuning of a sparsity parameter is unnecessary until the channel estimate has the desired number of reflectors. Tuning of a sparsity parameter depends on the SNR, and involves several executions of the whole process.

Steps 2 and 3 are iterated until a certain stopping criterion is satisfied. In the production of the results in this work the stopping criterion was a predetermined number of iterations.

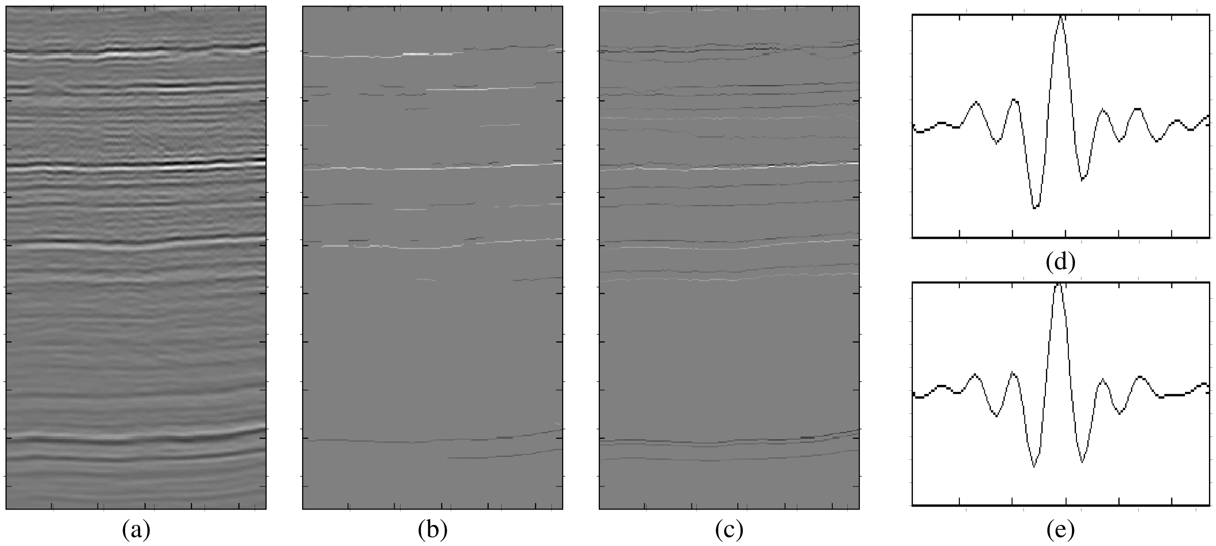

Figure 2: (a) Real seismic data containing 110 traces of 551 samples long; (b) channels estimates obtained by using the basic algorithm; (c) channels estimates obtained by using the proposed algorithm; (d) wavelet estimate obtained by using the basic algorithm; (e) wavelet estimate obtained by using the proposed algorithm.

EXPERIMENTAL RESULTS

Simulated Data

Twenty traces of 511 samples length are generated by convolving a certain wavelet with a pattern of reflectors. The reflectors, shown in Figure 1(b), are generated with the following properties:

1. The reflectors at each channel are Bernoulli distributed according to a selected density (probability of 5% for a reflector at any certain sample was used throughout all simulations).

2. The reflectors amplitudes in each channel are Gaussian distributed.

3. The reflectors are located along continuous paths across the channels (although a path does not have to start at the first channel and end at the last).

4. The progress of the path across channels is a Markov chain of locations, where the location of a reflector in channel m on a certain path is the linear continuation of the reflectors locations in channels $m-1$ and $m-2$, with some probability of moving one sample up or down from that location. The reflectors amplitudes along a path are also a Markov chain, with the restriction that the reflectors signs remain the same along the path.

The wavelet that is used in our simulations is shown in solid line in Figure 1(e). The traces are corrupted by high level white Gaussian noise (SNR $= -10$ dB). The SNR is defined by SNR $\triangleq \left(\sigma_a^2/\sigma_e^2\right)\sum_k h^2[k]$, where σ_a is the standard deviation of the reflectors amplitudes, and σ_e is the standard deviation of the noise. A comparison is made between the basic multichannel algorithm proposed by Kaaresen and Taxt (1998) (with consideration of local continuities of reflectors), and the new version proposed in this work. The results are presented in Figure 1.

It is clear from the results that the proposed algorithm recovers the true reflectors more accurately than the basic algorithm. Furthermore, the wavelet estimate obtained by the proposed algorithm is characterized by lower MSE (MSE $= \left\|\mathbf{h}-\widehat{\mathbf{h}}\right\|_2$). A comparison of Figures 1(c) and (b) shows that the basic algorithm generates false reflectors in some areas, and misdetects true reflectors in other areas. Hence, the improvement by using the proposed algorithm is not related to the specific choice of sparsity parameter in the basic algorithm, since a larger sparsity parameter would cause additional true reflectors to disappear, and a smaller sparsity parameter would lead to additional false reflectors.

Real Data

Figure 2(a) shows real seismic data (courtesy of GeoEnergy Inc., Texas) containing 110 traces of 551 samples long. The channel and wavelet estimates obtained by using the basic and proposed algorithms are presented in Figure 2. Since the true layer structure is unknown, one can only appreciate the continuous nature of the channel estimates obtained by using the proposed algorithm.

CONCLUSION

We have presented an improved algorithm for multichannel blind deconvolution in seismic applications, where reflectors in channels are sparse and laterally continuous. The improved performance, compared to that obtained by an existing algorithm, is achieved by combining the existing approach with a dynamic programming method for finding continuous lines in images. We have demonstrated the robustness of the proposed algorithm to high noise level, and the mechanism that enables excluding local maxima of the quality measure ℓ. In return, the proposed algorithm is characterized by higher computational complexity and slower convergence rate than the existing algorithm. In some applications the reflectors paths may vary more rapidly between channels, which necessitates increasing the parameter d. However, increasing the parameter d relaxes the continuity constraint and accordingly may reduce the benefits anticipated from the proposed approach.

REFERENCES

Amini, A. A., T. Weymouth, and R. Jain, 1990, Using dynamic programming for solving variational problems in vision: IEEE Transactions on Pattern Analysis and Machine Intelligence, **12**, 855–867.

Buckley, M. and J. Yang, 1997, Regularized shortest-path extraction: Pattern Recognition Letters, **18**, 621–629.

Heimer, A. and I. Cohen, 2007, Multichannel blind seismic deconvolution using dynamic programming: submitted to Signal Processing.

Idier, J. and Y. Goussard, 1993, Multichannel seismic deconvolution: IEEE Transactions on Geoscience and Remote Sensing, **31**, 961–979.

Jeffs, B. D., 1998, Sparse inverse solution methods for signal and image processing applications: Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, **3**, 1885–1888.

Kaaresen, K. F., 1998, Evaluation and applications of the iterated window maximization method for sparse deconvolution: IEEE Transactions on Signal Processing, **46**, 609–624.

Kaaresen, K. F. and T. Taxt, 1998, Multichannel blind deconvolution of seismic signals: Geophysics, **63**, 2093–2107.

Lazear, G. D., 1993, Mixed-phase wavelet estimation using fourth order cumulants: Geophysics, **58**, 1042–1051.

Luo, H. and Y. Li, 1998, The application of blind channel identification techniques to prestack seismic deconvolution: Proceedings of the IEEE, **86**, 2082–2089.

Mendel, J. M., 1991, Tutorial on high-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications: Proceedings of the IEEE, **79**, 278–304.

Mendel, J. M., J. Kormylo, F. Aminzadeh, and J. S. Lee, 1981, A novel approach to seismic signal processing and modeling: Geophysics, **46**, 1398–1414.

Perros-Meilhac, L., E. Moulines, K. Abed-Meraim, and P. Chevalier, 2001, Blind identification of multipath channels: A parametric subspace approach.: IEEE Transactions on Signal Processing, **49**, 1468–1480.

Tong, L. and S. Perreau, 1998, Multichannel blind identification: From subspace to maximum likelihood methods: Proceedings of the IEEE, **86**, 1951–1968.

Xu, G., H. Liu, L. Tong, and T. Kailath, 1995, A least-squares approach to blind channel identification: IEEE Transactions on signal processing, **43**.