

## Sparse multichannel blind deconvolution

Nasser Kazemi<sup>1</sup> and Mauricio D. Sacchi<sup>1</sup>

### ABSTRACT

We developed a sparse multichannel blind deconvolution (SMBD) method. The method is a modification of the multichannel blind deconvolution technique often called Euclid deconvolution, in which the multichannel impulse response of the earth is estimated by solving an homogeneous system of equations. Classical Euclid deconvolution is unstable in the presence of noise and requires the correct estimation of the length of the seismic wavelet. The proposed method, on the other hand, can tolerate moderate levels of noise and does not require a priori knowledge of the length of the wavelet. SMBD solves the homogeneous system of equations arising in Euclid deconvolution by imposing sparsity on the unknown multichannel impulse response. Trivial solutions to the aforementioned homogeneous system of equations are avoided by seeking sparse solutions on the unit sphere. We tested SMBD with synthetic and real data examples. Synthetic examples were used to judge the viability of the method in terms of noise. We found that SMBD gives reasonable estimates of the wavelet and reflectivity series for  $S/N \geq 4$ . The results clearly deteriorated when we tried to work on data that were severely contaminated by noise. A real marine data set was also used to test SMBD. In this case, the estimated wavelet was compared with a wavelet estimated by averaging first breaks. The estimated wavelet showed a noticeable resemblance to the average first break with normalized correlation coefficient of 0.92.

### INTRODUCTION

Deconvolution is an important and recurrent topic in seismic data processing. Many signals and images can be represented via the convolution of an unknown signal of interest and a blurring kernel. In general, the process that permits to remove the effects of the blurring kernel on the observed signal or image is called *deconvolution*.

When the signal of interest and the blurring kernel are unknown, the aforementioned process is denominated as *blind deconvolution* (Shalvi and Weinstein, 1990). In seismic data processing, the signal of interest is the impulse response of the earth and the blurring kernel is the seismic wavelet. In general, the seismic wavelet is considered unknown and, therefore, exploration seismologists are often faced with a typical blind deconvolution problem (Ulrych et al., 1995).

In seismic data processing, deconvolution is a part of early efforts to enhance resolution of seismic data and an important component of the transition from analog analysis of seismic data to digital-seismic data processing (Robinson, 1967). Early work using linear prediction theory solves the seismic blind deconvolution problem by making two fundamental assumptions. The earth's reflectivity series (impulse response of the earth) is a white sequence, and the seismic wavelet is a minimum phase sequence (Robinson and Treitel, 1964). These two assumptions permit estimation of a causal and stable inverse filter that is applied to the data to estimate the impulse response of the earth. Many deconvolution methods have been proposed to overcome the minimum phase assumption. Two early attempts are homomorphic deconvolution based on the work by Oppenheim and Schaffer (1968) and Oppenheim et al. (1976) and implemented for the first time in exploration seismology by Ulrych (1971). A comprehensive theoretical and practical study of the application of homomorphic systems to deconvolution and suppression of air gun reverberations is provided in Buhl et al. (1974) and Stoffa et al. (1974). Similarly, practical methods to implement homomorphic blind deconvolution on real data have been proposed by Otis and Smith (1977) and, more recently, by Herrera and van der Baan (2012). The minimum entropy deconvolution (MED) algorithm is another method that avoids the minimum phase assumption (Wiggins, 1978). MED assumes that the reflectivity is a sparse sequence. The MED algorithm estimates a nonminimum phase filter by maximizing a measure of sparsity of the seismic trace (Donoho, 1981). The measure of sparsity is the varimax norm that is also equivalent to an estimate of kurtosis (Longbottom et al., 1988; White, 1988). The maximization of the varimax norm, or the equivalent kurtosis, is analogous to minimization of a measure of entropy.

Manuscript received by the Editor 15 December 2013; revised manuscript received 24 March 2014; published online 24 June 2014.

<sup>1</sup>University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: kazemino@ualberta.ca; msacchi@ualberta.ca.

© 2014 Society of Exploration Geophysicists. All rights reserved.

The latter is coincident with Wiggins's interpretation of minimum disorder or minimum entropy as a synonym of sparsity (Sacchi et al., 1994). The convolution of the estimated MED inverse filter with the seismogram yields the reflectivity and the inverse of the MED filter is an estimator of the seismic source wavelet. Practical algorithms inspired in the MED method have been proposed for seismic data dephasing. Maximum kurtosis phase correction, for instance, can be used to find a phase rotation term that maximizes the kurtosis of the seismic trace (Levy and Oldenburg, 1987). It can be shown that, under ideal conditions, the phase rotation that maximizes the kurtosis of the seismic trace also dephases the seismic wavelet (Longbottom et al., 1988; Cambois and Hargreaves, 1994).

The homomorphic deconvolution, MED, and maximum kurtosis phase estimation methods suffer from a variety of shortcomings. For instance, homomorphic deconvolution is inclined to instability due to phase unwrapping and by its inherent inability to incorporate an additive noise term into its formulation. MED deconvolution often tends to annihilate small reflection coefficients (Ooe and Ulrych, 1979; Walden, 1985) and can become unstable in the presence of noise and highly sensitive to operator length (Nickerson et al., 1986). Maximum kurtosis phase estimation is sensitive to the bandwidth of the wavelet (White, 1988; Xu et al., 2012).

The history of seismic deconvolution is populated by interesting statistical methods for blind deconvolution. These methods, however, often only work under ideal signal conditions. For instance, an important excitement was generated by methods based on fourth-order cumulant matching (Hargreaves, 1994; Velis and Ulrych, 1996) and homomorphic deconvolution via fourth-order cumulants (Sacchi et al., 1996). Fourth-order cumulants are attractive because they can be computed from the seismic trace and they do preserve the phase of the wavelet when the reflectivity is a sparse sequence with nonvanishing kurtosis (Sacchi and Ulrych, 2000). However, the conditions for robust wavelet estimation required by cumulant-based methods are not often satisfied by real seismic data (Stogioglou et al., 1996).

Last, we also mention parametric methods based on maximum likelihood estimation. These methods attack the blind deconvolution via a two-stage procedure. First, the wavelet is assumed to be known and the reflectivity is estimated by maximizing likelihood. Then, the reflectivity is fixed and the likelihood function is maximized with respect to the wavelet (Mendel, 1983; Kaaresen and Taxy, 1998; Canadas, 2002).

In this paper, we studied a multichannel blind deconvolution algorithm often called *Euclid's deconvolution*. The method was first discussed in the geophysical literature by Rietsch (1997a) and tested with real data examples in Rietsch (1997b). The method has been previously investigated by Xu et al. (1995) for blind channel estimation in communication systems. The method has also been used to improve speech recognition by Liu and Malvar (2001). The idea can be summarized as finding common factors of the  $z$ -transform of the source wavelet embedded in a group of seismograms with different reflectivity sequences. The problem leads to the estimation of the multichannel seismic reflectivity via the solution of homogeneous system of equations (Rietsch, 1997a; Mazzucchelli and Spagnolini, 2001). In the ideal case, the eigenvector associated to the minimum nonzero eigenvalue of the homogenous system of equations is an estimator of the multichannel reflectivity. However, small level of noise impinges on the correct identification of the

eigenvector associated to the impulse response of the earth. This problem is examined in detail by Rietsch (1997a, 1997b).

A variant of Euclid deconvolution is used by Royer et al. (2012) to separate source signatures from propagation effects in teleseismic data. Similar approaches to Euclid deconvolution were also proposed for image restoration (Harikumar and Bresler, 1999; Sroubek and Milanfar, 2012).

Our main contribution is an improvement to Euclid deconvolution to make it applicable to real data processing. The proposed method can tolerate moderate amounts of noise and does not require a priori knowledge of the length of the seismic wavelet. In the proposed method, the homogeneous system of equation is satisfied by a sparse solution (sparse impulse responses). In other words, we are assuming a reflectivity sequence that is sparse. The problem leads to a nonquadratic minimizing technique, where the solution must be constrained to lie on the unit sphere. We discuss a steepest descent method that permits to obtain accurate estimates of the seismic reflectivity and wavelet in the presence of a moderate amount of noise.

## THEORY

### Multichannel blind deconvolution

The earth's impulse response can be modeled as a linear system (Robinson and Treitel, 1980). The input-output relationship for this system, assuming a stationary source wavelet and no noise, can be written as follows:

$$d_j[n] = \sum_k w[n-k]r_j[k], \quad j = 1 \dots J, \quad (1)$$

where the multichannel seismic data is given by  $\mathbf{d}_j = (d_j[1], d_j[2], \dots, d_j[N])^T$ . Similarly, the impulse response for channel  $j$  is given by  $\mathbf{r}_j = (r_j[1], r_j[2], \dots, r_j[M])^T$ , and the seismic source wavelet via the vector  $\mathbf{w} = (w[1], w[2], \dots, w[L])^T$ . We stress that  $N = M + L - 1$ . We also remind the readers that convolution can be represented via the  $z$ -transform as follows

$$D_j(z) = W(z)R_j(z), \quad j = 1 \dots J. \quad (2)$$

By virtue of equation 2, it is easy to show that

$$D_p(z)R_q(z) - D_q(z)R_p(z) = 0, \quad \forall p, q. \quad (3)$$

The latter can be rewritten in matrix-vector form as follows:

$$\mathbf{D}_p \mathbf{r}_q - \mathbf{D}_q \mathbf{r}_p = \mathbf{0}, \quad (4)$$

where  $\mathbf{D}_p$  and  $\mathbf{D}_q$  in equation 4 represent the convolution matrices of channels  $p$  and  $q$ , respectively. The combination of all possible equations leads to the following homogeneous system of equations:

$$\mathbf{A} \mathbf{x} = \mathbf{0}, \quad (5)$$

where



we prefer to use an educated step that was derived from Rodrigues' rotation formula (Murray et al., 1994):

$$\mathbf{x}_{k+1} = \cos(\theta_k)\mathbf{x}_k + \sin(\theta_k)\mathbf{h}_k. \quad (19)$$

It is easy to show via a few algebraic manipulations that with this expression the updated solution  $\mathbf{x}_{k+1}$  is on the unit sphere (see Appendix A). Notice that by choosing a small angle  $\theta_k < 0$ , we obtain the desired small step in the direction of steepest descent because  $\sin(\theta_k) \approx \theta_k < 0$ . In equation 19, the effect of the term  $\cos(\theta_k)$  is to shrink the current position  $\cos(\theta_k)\mathbf{x}_k$  in a way that the updated solution lies on the unit sphere. This is illustrated by Figure 2b. Our problem has been reduced to a 1D minimization on the sphere. The minimization is carried out by a simple update of  $\theta_k$  that guarantees that  $J(\mathbf{x}_k) > J(\mathbf{x}_{k+1})$  (see Appendix B).

It is important to point out that the steepest-descent algorithm must be initialized by a solution that is close to the final reflectivity. This is because the problem is nonlinear and multimodal. Our test indicates that starting with the data as an initial solution always leads to sparse estimate of the reflectivity. The last assessment is supported because the data are structurally close to the true sparse reflectivity.

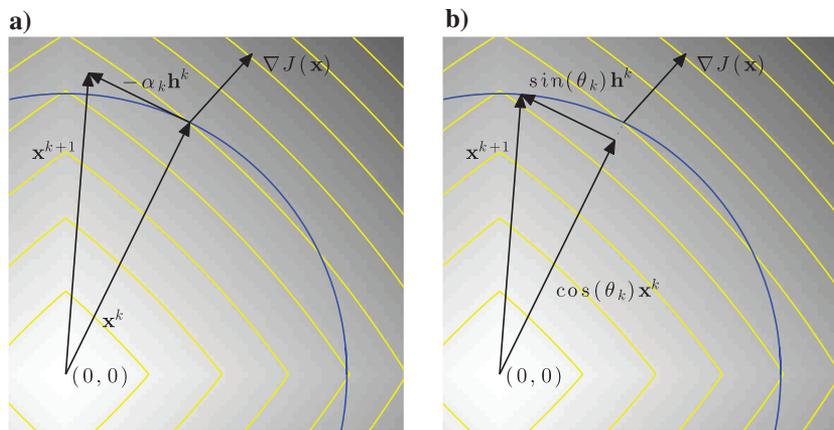


Figure 2. Cartoon representation of the classical steepest descent and proposed approach. (a) Steepest descent algorithm via classical update rule  $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{h}_k$ ,  $\alpha_k$  is the step length and (b) proposed steepest descent algorithm via the update rule  $\mathbf{x}_{k+1} = \cos(\theta_k)\mathbf{x}_k + \sin(\theta_k)\mathbf{h}_k$ , the step length is given by  $\sin(\theta_k)$ . In this case, the updated position is guaranteed to lie on the unit sphere.

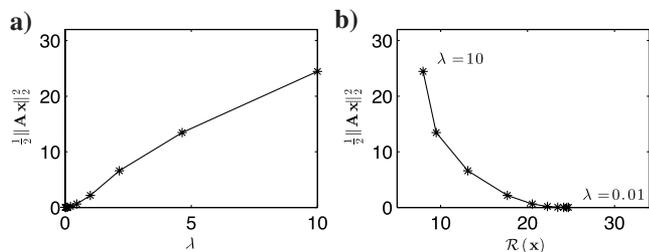


Figure 3. Exploring the importance of trade-off parameter  $\lambda$  in the SMBD method using synthetic data with  $S/N = 100$ . (a) Trade-off parameter versus the  $l_2$  norm of homogeneous system of equations and (b) trade-off curve using different values of regularization parameter.

### EXAMPLES

To examine the performance of the proposed method, we introduce two figures of merit for the estimated wavelet and the estimated reflectivity series. For instance, if a true generic signal is denoted by  $\mathbf{y}_0$  and the estimated signal by  $\mathbf{y}$ , we define the quality of the reconstruction,  $Q$ , as follows:

$$Q = 10 \log \frac{\|\mathbf{y}_0\|_2^2}{\|\mathbf{y}_0 - \mathbf{y}\|_2^2}. \quad (20)$$

Our second figure of merit is the normalized correlation coefficient (NCC):

$$\text{NCC} = \frac{\mathbf{y}^T \mathbf{y}_0}{\|\mathbf{y}_0\|_2 \|\mathbf{y}\|_2}. \quad (21)$$

We stress that blind deconvolution algorithms can only determine unscaled versions of the seismic wavelet and reflectivity series. Therefore, all our estimators must be rescaled prior to computing  $Q$ . For this purpose, the estimated signal  $\mathbf{y}$  is multiplied by a scalar  $\alpha$  such that  $\|\mathbf{y}_0 - \alpha \mathbf{y}\|_2^2$  is minimum. It is clear that  $\mathbf{y}$  in equation 20 is replaced by  $\alpha \mathbf{y}$ . It is also clear that the aforementioned scaling is not needed for the computation of the NCC.

We will denote  $\text{NCC}_w$  and  $\text{NCC}_x$  as the normalized correlation coefficients for the wavelet and the reflectivity, respectively. Similarly, we will use  $Q_w$  and  $Q_x$  to indicate the quality of the reconstruction of the wavelet and reflectivity, respectively.

### Simulations

To test the method, we first run a synthetic example with high signal-to-noise ratio ( $S/N = 100$ ). We run the algorithm for different values of the parameter  $\lambda$  to explore the trade-off curve. The latter is portrayed in Figure 3. With different experiments, we concluded that  $\lambda = 4$  yields the best result for different levels of  $S/N$ . The true sparse reflectivity is displayed in Figure 4a in conjunction with the data (Figure 4b). The sampling interval for this exercise was chosen  $\Delta t = 2$  ms. We also used a Ricker wavelet of central frequency 40 Hz and a  $50^\circ$  phase rotation. The estimated reflectivity for this example is shown

in Figure 4c. The estimated reflectivity and the seismic traces were used to estimate the seismic wavelet via multichannel frequency domain deconvolution. The estimated wavelet and the true input wavelet are shown side by side in Figure 4d. In this example, the quality of reconstruction for the wavelet is  $Q_w = 14$  dB. Similarly, the quality of the reconstruction for the reflectivity is  $Q_x = 5$  dB. Normalized correlation coefficients for the wavelet and reflectivity are given by  $\text{NCC}_w = 0.96$  and  $\text{NCC}_x = 0.82$ , respectively.

We rerun the synthetic example with additive noise  $S/N = 4$ . In this case, we obtain a reconstruction quality of  $Q_w = 13$  dB and a normalized correlation coefficient  $\text{NCC}_w = 0.89$ . Similarly, we obtain  $Q_x = 3.8$  dB and  $\text{NCC}_x = 0.75$  for the reflectivity series. The results for this simulation are shown in Figure 5. Finally, we also

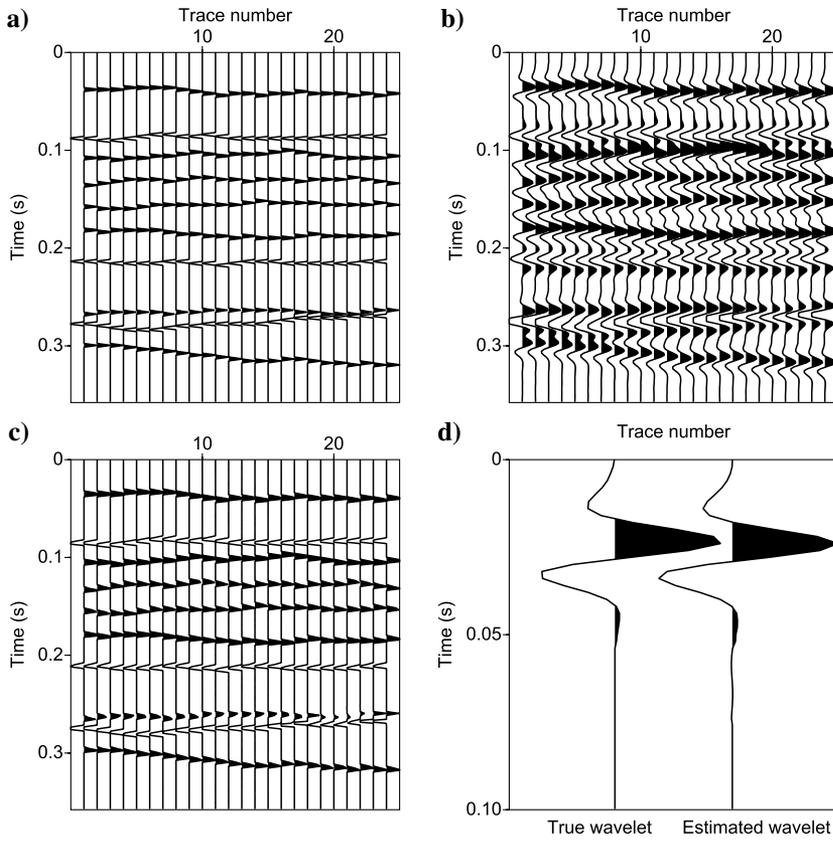


Figure 4. Performance of the SMBD method using synthetic data with  $S/N = 100$ . (a) True synthetic reflectivity sequences, (b) seismic traces with  $S/N = 100$ , (c) estimated sparse reflectivity sequences, and (d) true and estimated wavelets.

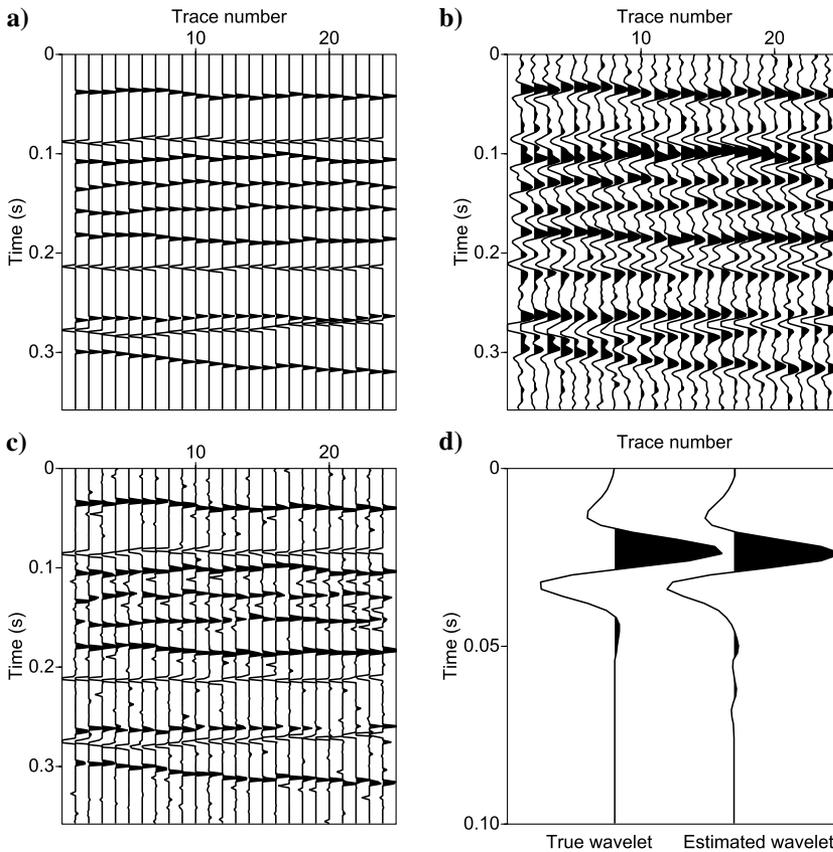


Figure 5. Performance of the SMBD method using synthetic data with  $S/N = 4$ . (a) True synthetic reflectivity sequences, (b) seismic traces with  $S/N = 4$ , (c) estimated sparse reflectivity sequences, and (d) true and estimated wavelets.

provide convergence curves for the algorithm for  $S/N = 2, 4,$  and  $100$  in Figure 6.

To analyze the stability of the sparse multichannel blind deconvolution (SMBD) method under different levels of noise, we run a Monte Carlo simulation with 20 different realizations of noise and seismic reflectivity for a given  $S/N$  and parameter  $\lambda$ . Each realization of the reflectivity is similar to Figure 4a. We have been careful

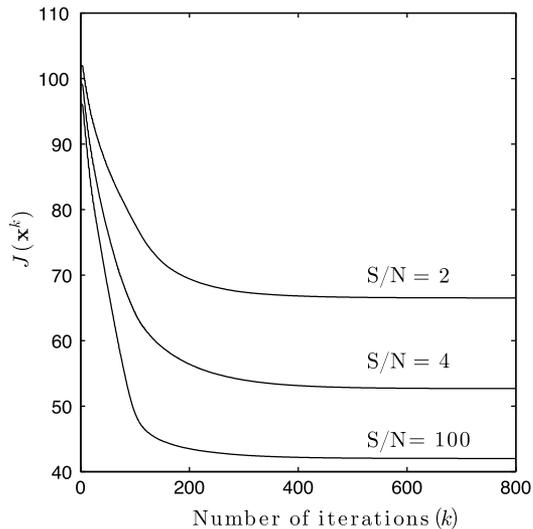


Figure 6. Convergence behaviour of SMBD method using synthetic data with different levels of noise.

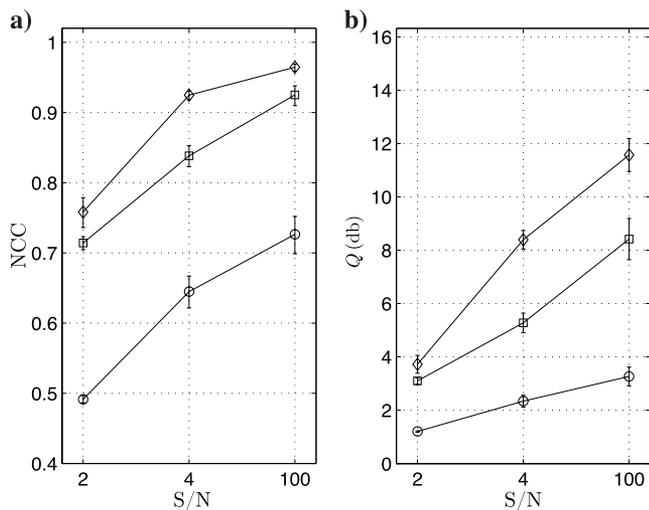


Figure 7. (a) Mean and standard error of normalized correlation coefficients versus  $S/N$  and (b) mean and standard error of the quality of the reconstruction versus  $S/N$ . These results were obtained by running SMBD on 20 realizations of reflectivity models that are similar to the reflectivity shown in Figure 4a. Diamonds ( $\diamond$ ) are used to indicate  $NCC_w$  and  $Q_w$  values for the estimated seismic wavelet. Similarly, circles ( $\circ$ ) are used to indicate the  $NCC_x$  and  $Q_x$  for the estimated reflectivity. Squares ( $\square$ ) are used to indicate the  $NCC_x$  and  $Q_x$  computed after applying an Ormsby trapezoidal filter to the true reflectivity and the estimated reflectivity. The trade-off parameter of the SMBD method is  $\lambda = 1$ . Note that the  $S/N$  values are not linearly spaced.

in producing realizations with 10 reflectors each and similar temporal and spatial variability. This was done by taking a random reflectivity composed of 10 impulses of random amplitude for the first trace and randomly perturbing the times to generate the reflectivity of the remaining traces. The amplitude of the reflection coefficients for a given reflector were also allowed to vary in space by a very small amount. The 20 realizations were used to estimate 20 wavelets and 20 multichannel reflectivity estimators that were used to estimate averages and standard errors for  $NCC_w$ ,  $NCC_x$ ,  $Q_w$ , and  $Q_x$ .

At this point, it is important to mention that the computation of the standard error of the normalized correlation coefficients  $NCC_w$  and  $NCC_x$  requires special attention. Normalized correlation coefficients are bounded by unity and, therefore, they are not normally distributed (Weatherburn, 1949). To compute the mean and standard error of the normalized correlation coefficient, we first apply the Fisher's transform to create a new variable  $F = \frac{1}{2} \ln \left[ \frac{1 + (NCC)}{1 - (NCC)} \right]$  that is distributed almost normally. We can

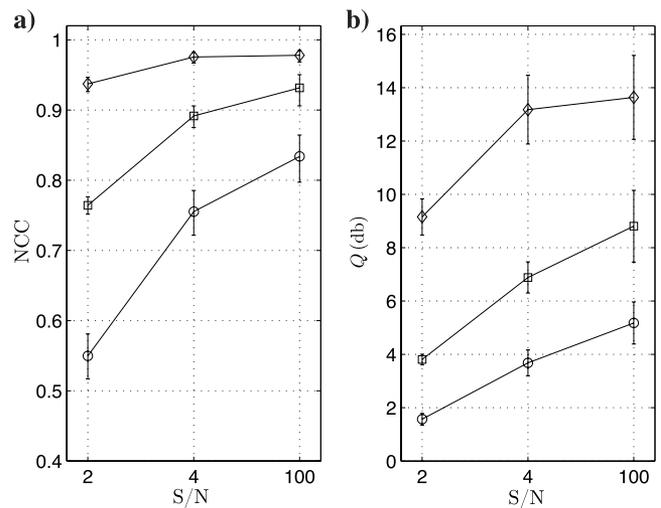


Figure 8. Similar to Figure 7 but with  $\lambda = 4$ .

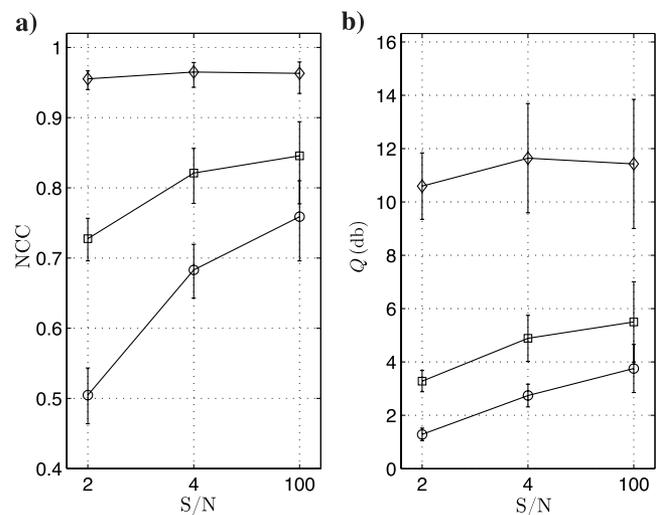


Figure 9. Similar to Figure 7 but with  $\lambda = 10$ .

now compute the mean and standard error of the variable  $F$ , which are then inverse transformed to obtain the desired mean and standard error of the normalized correlation coefficients (VanDecar and Crosson, 1990; Herrera and van der Baan, 2012).

The average for 20 realizations of our two figures of merit and their standard deviations are shown in Figures 7–9 for  $\lambda = 1, 4$ , and 10, respectively. In these figures, the diamonds ( $\diamond$ ) are used to indicate  $NCC_w$  and  $Q_w$  values for the estimated seismic wavelet. Similarly, circles ( $\circ$ ) are used to indicate the  $NCC_x$  and  $Q_x$  values for the estimated seismic reflectivity. We observe that the proposed deconvolution scheme performs better at estimating the wavelet than the reflectivity. The reflectivity is a full band sequence with

low and high frequencies annihilated by the bandpass character of the seismic wavelet. Therefore, recovery of frequencies outside the natural band imposed by the seismic wavelet fully relies on the sparse reflectivity assumption (Sacchi et al., 1994). Our synthetic data were sampled at 2 ms and, therefore, hoping to recover a full band reflectivity with spectral contributions from 0 to 250 Hz (Nyquist frequency) is definitely an unworkable cause. The latter is reflected by the low values of  $Q_x$  in comparison to those of  $Q_w$ . To gain critical understanding of the limits of our algorithm, we also compute  $Q_x$  values for bandpassed versions of the true and estimated reflectivity series. In this case, we use an Ormsby trapezoidal filter (Sheriff, 2002) defined by four corner frequencies

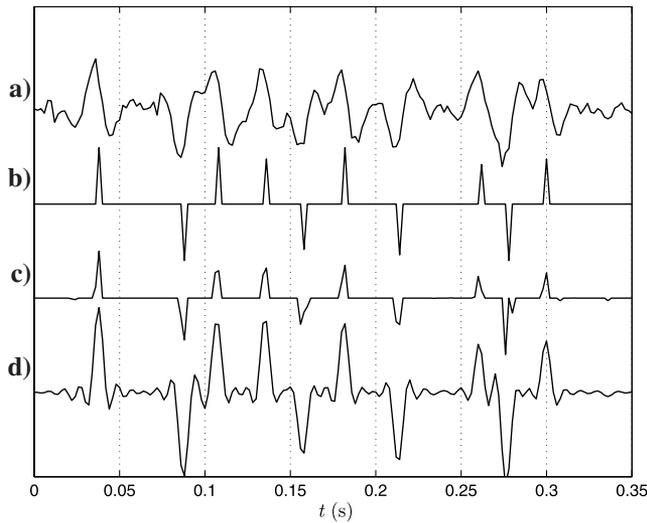


Figure 10. (a) First trace from one realization of multichannel data similar to Figure 5b, (b) true reflectivity series, (c) estimated reflectivity via SMBD, and (d) estimated reflectivity after applying an Ormsby trapezoidal filter with corner frequencies 0, 1, 100, and 125 Hz. The simulation corresponds to values  $S/N = 4$  and  $\lambda = 4$ .

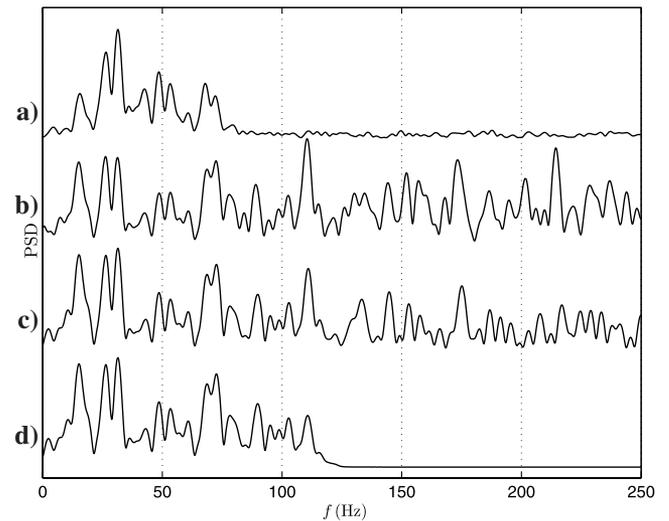


Figure 11. Power spectral density of the data portrayed in Figure 10. (a) Seismic trace, (b) true reflectivity series, (c) estimated reflectivity via SMBD, and (d) estimated reflectivity after applying an Ormsby trapezoidal filter with corner frequencies 0, 1, 100, and 125 Hz.

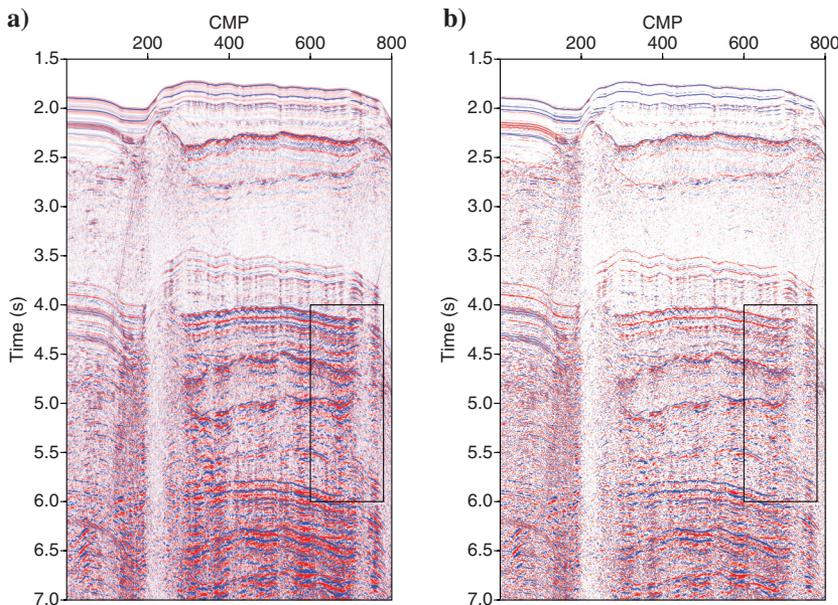


Figure 12. (a) Near-offset section of data set from the Gulf of Mexico and (b) estimated sparse reflectivity.

0, 1, 100, and 125 Hz to restrict the bandwidth of the true reflectivity and estimated reflectivity before computing  $Q_x$  and  $NCC_x$ . The results are also included in Figures 7–9 and are indicated with squares ( $\square$ ). One can observe an improvement in  $NCC_x$  and  $Q_x$  when high frequencies are excluded from the recovered reflectivity.

Finally, Figures 10 and 11 portray the resulting reflectivity inversion in time and frequency domain for one realization and for the first seismic trace of a group of 24 traces. The example also corresponds to a parameter  $\lambda = 4$  and  $S/N = 4$ . Figure 10a presents the seismic trace, Figure 10b shows the true reflectivity series, Figure 10c is the full band reflectivity estimated by SMBD, and Figure 10d is the estimated reflectivity after applying the Ormsby trapezoidal filter. Figure 11 shows, in the same order, the power spectral density of the signals portrayed in Figure 10. It is evident from Figure 11 that the unfiltered spectra match well with the true spectra up to approximately 100 Hz.

### Real data example

Unfortunately, we do not have an automatic way of estimating the trade-off parameter  $\lambda$  needed by SMBD. The simulations in the previous section were used to explore the behavior of the algorithm in terms of the parameter  $\lambda$ . Based on the analysis presented in Figures 7–9, we have selected  $\lambda = 4$  because it provided the best reconstruction of the wavelet and reflectivity for a moderate level of noise ( $S/N = 4$ ).

Our real data test uses the Gulf of Mexico, Mississippi Canyon data set. These data have been extensively used for testing multiple suppression algorithms (see, e.g., [Verschuur and Prein, 1999](#)). SMBD was run on the near-offset section of the Mississippi Canyon data set. The input data and the estimated sparse impulse response are shown in Figure 12a and 12b, respectively. We also show details of the seismic sections before and after deconvolution in Figure 13a and 13b. Notice that Figure 13b is the resulting sparse impulse response estimated by SMBD. This data set is contaminated by

multiples. Therefore, our blind deconvolution algorithm was used to estimate the full impulse response including multiples rather than the primary only impulse response (reflectivity). We have run SMBD in windows of 1 s in time and 200 traces with 10% overlap in time and space. The windows were patched back together to produce Figure 12b.

We used the estimated impulse response of the whole near-offset section to estimate the wavelet via a multichannel frequency domain deconvolution. Wavelets computed from individual windows were similar and this is why we have decided to compute one wavelet for the whole near-offset section. The seismic wavelet is portrayed in Figure 14. We also displayed the estimated wavelet obtained by aligning and averaging the first break. The resemblance of the two wavelets is remarkable, with normalized correlation coefficient  $NCC_w = 0.92$ .

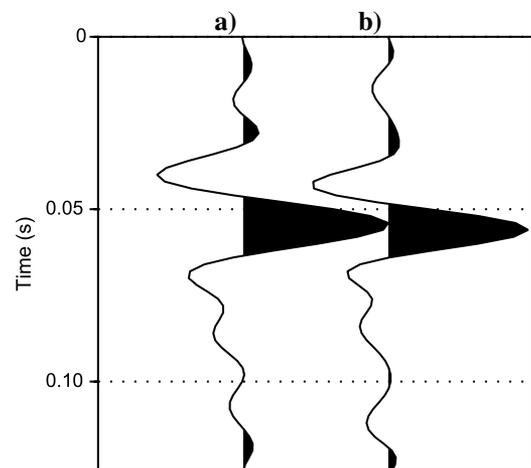
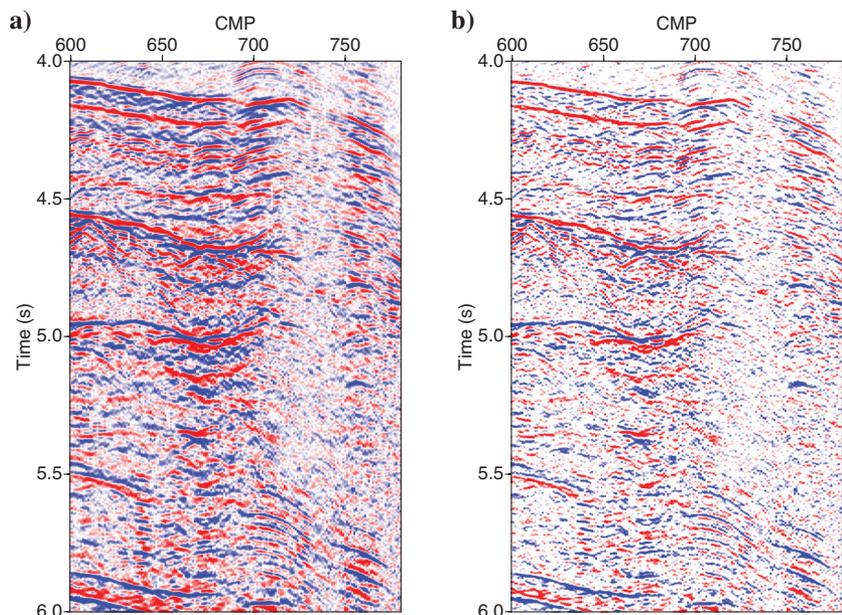


Figure 14. Estimated wavelet for Gulf of Mexico data set. (a) Estimated wavelet using the SMBD method and (b) estimated wavelet obtained by averaging the first break after alignment.

Figure 13. A zoom into the black rectangles marked in Figure 12. (a) Before deconvolution and (b) after deconvolution.



## CONCLUSIONS

We have presented an algorithm inspired by Euclid deconvolution that permits us to estimate the seismic reflectivity without a priori knowledge of the seismic wavelet. The seismic wavelet is computed as a by-product of the process via a multichannel frequency domain deconvolution between traces and estimated reflectivity sequences. The core of the algorithm is the estimation of the reflectivity by the solution of the multichannel homogeneous system of equations with sparsity constraints.

An optimization problem that uses the method of steepest descent was developed. To avoid trivial solutions, the reflectivity vector was constrained to have unit norm. The latter leads to a constrained optimization problem, where one attempts to estimate a sparse signal that fits a multichannel homogeneous system of equations and, in addition, the signal lies on the unit sphere. This optimization problem was solved by the method of steepest descent with an update rule that keeps current estimates of the sparse reflectivity on the unit sphere. The method is stable under a variety of noise levels and for different values of the trade-off parameter  $\lambda$ . We stress that like in every deconvolution scenario, the method works well when it honors certain assumptions. In this case, not only the wavelet needs to be stationary for all traces but the reflectivity must be sparse.

We used synthetic and real data examples to evaluate the method. Synthetic examples permitted us to assess the viability of the method in terms of noise. The method gives reasonable estimates of wavelet and reflectivity series with  $S/N = 4$  and higher. We have obtained workable results for  $S/N = 2$ . However, the results of the method clearly deteriorate when we try to push it to work on data severely contaminated with noise. We have observed that the quality of the estimated wavelet is superior to the quality of the estimated reflectivity for the same  $S/N$ .

We also applied the method to near-offset section of Gulf of Mexico data set. Our estimated wavelet has a remarkable semblance with the wavelet estimated by aligning and stacking first breaks.

We emphasize that the SMBD method does not consider coherent noise in the convolutional model. In this regard, we believe that its application to onshore data will require extensive preconditioning to remove coherent noise. This is likely true for all blind deconvolution methods.

## ACKNOWLEDGMENTS

We thank the sponsors of the Signal Analysis and Imaging Group at the University of Alberta. The authors also thank R. H. Herrera for sharing with us interesting ideas about blind deconvolution and WesternGeco for providing the Gulf of Mexico, Mississippi Canyon data set.

## APPENDIX A

### CURVILINEAR LINE SEARCH ON THE UNIT SPHERE

This section proves that curvilinear line search on the unit sphere is equal to equation 19. In our problem, we need a rotation matrix that preserves the sparse solution on the unit sphere. This is equivalent to have a rotation matrix  $\mathcal{R}_k$ :

$$\mathbf{x}_{k+1} = \mathcal{R}_k \mathbf{x}_k, \quad (\text{A-1})$$

such that  $\|\mathbf{x}_{k+1}\|_2 = \|\mathbf{x}_k\|_2 = 1$ . Using angle axis representation of rotation matrix via aka Rodrigues' formula, we have

$$\mathcal{R}_k \mathbf{x}_k = \mathbf{x}_k + \sin(\theta_k)(\mathbf{r}_k \times \mathbf{x}_k) + (1 - \cos(\theta_k))(\mathbf{r}_k \mathbf{r}_k^T - \mathbf{I})\mathbf{x}_k, \quad (\text{A-2})$$

where  $\mathbf{r}_k$  is rotation axis,  $\theta_k$  is rotation angle,  $\mathbf{I}$  is identity matrix, and  $k$  is iteration. We need to choose  $\mathbf{r}_k$  in such a way that it lets the search go in the direction of the projection of the gradient into the tangent plane on the sphere. Hence, the only option is  $\mathbf{r}_k = \mathbf{x}_k \times \mathbf{h}_k$  or in other words,

$$\mathbf{h}_k = \mathbf{r}_k \times \mathbf{x}_k, \quad (\text{A-3})$$

where  $\mathbf{h}_k$  is the normalized projected gradient on the sphere. Note that the length of  $\mathbf{r}_k$ ,  $\mathbf{x}_k$ , and  $\mathbf{h}_k$  are equal to one. Now it is easy to show that the rotation axis is orthogonal to the current solution. Hence, we have

$$\mathbf{r}_k^T \mathbf{x}_k = 0. \quad (\text{A-4})$$

Inserting equations A-3 and A-4 into equation A-2 yields

$$\mathcal{R}_k \mathbf{x}_k = \sin(\theta_k)\mathbf{h}_k + \cos(\theta_k)\mathbf{x}_k. \quad (\text{A-5})$$

## APPENDIX B

### CONVERGENCE BEHAVIOR OF SMBD ALGORITHM

This section proves the convergence of the proposed steepest descent method. For  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{m}_k$  as long as  $-\mathbf{g}_k^T \mathbf{m}_k > 0$  it is gradient descent.  $\mathbf{g}_k$  is the gradient of the cost function at iteration  $k$ . In a gradient descent, the step size should be small enough to guarantee the convergence. We will use the same concept to prove the convergence of the proposed technique.

Again, consider the angle axis representation of rotation matrix via aka Rodrigues' formula:

$$\mathcal{R}_k \mathbf{x}_k = \mathbf{x}_k + \sin(\theta_k)(\mathbf{r}_k \times \mathbf{x}_k) + (1 - \cos(\theta_k))(\mathbf{r}_k \mathbf{r}_k^T - \mathbf{I})\mathbf{x}_k, \quad (\text{B-1})$$

by analogy we have

$$\mathbf{m}_k = \sin(\theta_k)(\mathbf{r}_k \times \mathbf{x}_k) + (1 - \cos(\theta_k))(\mathbf{r}_k \mathbf{r}_k^T - \mathbf{I})\mathbf{x}_k. \quad (\text{B-2})$$

It is very interesting that unlike conventional steepest descent, the evolving direction depends on the step size  $\theta_k$ . Now, we need to check if  $-\mathbf{g}_k^T \mathbf{m}_k > 0$ . We should emphasize that the steepest descent is valid only for sufficiently small step sizes. Assuming small angles, we have  $\sin(\theta) \approx \theta$  and  $(1 - \cos(\theta)) \approx 0$ . Hence, equation B-2 simplifies to

$$\mathbf{m}_k \approx \theta_k(\mathbf{r}_k \times \mathbf{x}_k) = \theta_k \mathbf{h}_k. \quad (\text{B-3})$$

Obviously,  $-\mathbf{g}_k^T \theta_k \mathbf{h}_k > 0$ . Hence, it is gradient descent. It is worth mentioning that to satisfy the constraint, we used  $\mathbf{m}_k =$

$\sin(\theta_k)(\mathbf{r}_k \times \mathbf{x}_k) + (1 - \cos(\theta_k))(\mathbf{r}_k \mathbf{r}_k^T - \mathbf{I})\mathbf{x}_k$ , but to prove the convergence of the proposed steepest descent, we have  $\mathbf{m}_k \approx \theta_k \mathbf{h}_k$ .

## REFERENCES

- Bube, K., and R. Langan, 1997, Hybrid minimization with applications to tomography: *Geophysics*, **62**, 1183–1195, doi: [10.1190/1.1444219](https://doi.org/10.1190/1.1444219).
- Buhl, P., P. Stoffa, and G. Bryan, 1974, The application of homomorphic deconvolution to shallow water marine seismology, Part I: Real data: *Geophysics*, **39**, 417–426, doi: [10.1190/1.1440439](https://doi.org/10.1190/1.1440439).
- Cambois, G., and N. Hargreaves, 1994, Zero-phase conversion of marine data using one parameter phase filters and kurtosis maximization: 64th Annual International Meeting, SEG, Expanded Abstracts, 1591–1594.
- Canadas, G., 2002, A mathematical framework for blind deconvolution inverse problems: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2202–2205.
- Donoho, D., 1981, On minimum entropy deconvolution: Academic Press, vol. 109, 2831–2839.
- Hargreaves, N., 1994, Wavelet estimation via fourth-order cumulants: 64th Annual International Meeting, SEG, Expanded Abstracts, 1588–1590.
- Harikumar, G., and Y. Bresler, 1999, Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms: *IEEE Transactions on Image Processing*, **8**, 202–219, doi: [10.1109/83.743855](https://doi.org/10.1109/83.743855).
- Herrera, R. H., and M. van der Baan, 2012, Short-time homomorphic wavelet estimation: *Journal of Geophysics and Engineering*, **9**, 674–680, doi: [10.1088/1742-2132/9/6/674](https://doi.org/10.1088/1742-2132/9/6/674).
- Kaarensen, K., and T. Taxt, 1998, Multichannel blind deconvolution of seismic signals: *Geophysics*, **63**, 2093–2107, doi: [10.1190/1.1444503](https://doi.org/10.1190/1.1444503).
- Lee, S. I., H. Lee, P. Abbeel, and A. Y. Ng, 2006, Efficient L1 regularized logistic regression: *AAAI*, 401–408.
- Levy, S., and D. Oldenburg, 1987, Automatic phase correction of common midpoint stacked data: *Geophysics*, **52**, 51–59, doi: [10.1190/1.1442240](https://doi.org/10.1190/1.1442240).
- Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: *Geophysics*, **77**, no. 1, V1–V9, doi: [10.1190/geo2011-0099.1](https://doi.org/10.1190/geo2011-0099.1).
- Liu, J., and H. S. Malvar, 2001, Blind deconvolution of reverberated speech signals via regularization: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **5**, 3037–3040, doi: [10.1109/ICASSP.2001.940298](https://doi.org/10.1109/ICASSP.2001.940298).
- Longbottom, J., A. T. Walden, and R. E. White, 1988, Principles and application of maximum kurtosis phase estimation: *Geophysical Prospecting*, **36**, 115–138, doi: [10.1111/j.1365-2478.1988.tb02155.x](https://doi.org/10.1111/j.1365-2478.1988.tb02155.x).
- Mazzucchelli, P., and U. Spagnolini, 2001, Least squares multichannel deconvolution: 63rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, P-169.
- Mendel, J. M., 1983, Optimal seismic deconvolution: An estimation based approach: Academic Press.
- Murray, R. M., Z. Li, and S. S. Sastry, 1994, A mathematical introduction to robotic manipulation: CRC Press.
- Nickerson, W. A., T. Matsuoka, and T. J. Ulrych, 1986, Optimum-lag minimum-entropy deconvolution: 56th Annual International Meeting, SEG, Expanded Abstracts, 519–522.
- Ooe, M., and T. J. Ulrych, 1979, Minimum entropy deconvolution with an exponential transformation: *Geophysical Prospecting*, **27**, 458–473, doi: [10.1111/j.1365-2478.1979.tb00979.x](https://doi.org/10.1111/j.1365-2478.1979.tb00979.x).
- Oppenheim, A., G. Kopec, and J. Tribolet, 1976, Signal analysis by homomorphic prediction: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **24**, 327–332, doi: [10.1109/TASSP.1976.1162828](https://doi.org/10.1109/TASSP.1976.1162828).
- Oppenheim, A., and R. Schaffer, 1968, Homomorphic analysis of speech: *IEEE Transactions on Audio and Electroacoustics*, **16**, 221–226, doi: [10.1109/TAU.1968.1161965](https://doi.org/10.1109/TAU.1968.1161965).
- Otis, R., and R. Smith, 1977, Homomorphic deconvolution by log spectral averaging: *Geophysics*, **42**, 1146–1157, doi: [10.1190/1.1440780](https://doi.org/10.1190/1.1440780).
- Rietsch, E., 1997a, Euclid and the art of wavelet estimation, Part I: Basic algorithm for noise-free data: *Geophysics*, **62**, 1931–1938, doi: [10.1190/1.1444293](https://doi.org/10.1190/1.1444293).
- Rietsch, E., 1997b, Euclid and the art of wavelet estimation, Part II: Robust algorithm and field-data examples: *Geophysics*, **62**, 1939–1946, doi: [10.1190/1.1444294](https://doi.org/10.1190/1.1444294).
- Robinson, E., 1967, Predictive decomposition of time series with application to seismic exploration: *Geophysics*, **32**, 418–484, doi: [10.1190/1.1439873](https://doi.org/10.1190/1.1439873).
- Robinson, E., and S. Treitel, 1964, Principles of digital filtering: *Geophysics*, **29**, 395–404, doi: [10.1190/1.1439370](https://doi.org/10.1190/1.1439370).
- Robinson, E. A., and S. Treitel, 1980, *Geophysical signal analysis*: Prentice-Hall, Inc.
- Royer, A., M. Bostock, and E. Haber, 2012, Blind deconvolution of seismograms regularized via minimum support: *Inverse Problems*, **28**, 125010–125027, doi: [10.1088/0266-5611/28/12/125010](https://doi.org/10.1088/0266-5611/28/12/125010).
- Sacchi, M., and T. Ulrych, 2000, Nonminimum-phase wavelet estimation using higher order statistics: *The Leading Edge*, **19**, 80–83, doi: [10.1190/1.1438466](https://doi.org/10.1190/1.1438466).
- Sacchi, M., D. Velis, and A. Comínguez, 1994, Minimum entropy deconvolution with frequency-domain constraints: *Geophysics*, **59**, 938–945, doi: [10.1190/1.1443653](https://doi.org/10.1190/1.1443653).
- Sacchi, M., D. R. Velis, and T. J. Ulrych, 1996, Wavelets via polycepstra: 66th Annual International Meeting, SEG, Expanded Abstracts, 1583–1586.
- Schmidt, M., G. Fung, and R. Rosales, 2007, Fast optimization methods for L1 regularization: A comparative study and two new approaches, in Joost N Kok, Jasek Koronacki, Raomon Lopez Mantaras, Stan Matwin, Dunja Mladenić, and Andrzej Skowron, eds., *Machine Learning: ECML 2007*: Springer, Lecture Notes in Computer Science, vol. 4701, 286–297, doi: [10.1007/978-3-540-74958-5\\_28](https://doi.org/10.1007/978-3-540-74958-5_28).
- Shalvi, O., and E. Weinstein, 1990, New criteria for blind deconvolution of nonminimum phase systems (channels): *IEEE Transactions on Information Theory*, **36**, 312–321, doi: [10.1109/18.52478](https://doi.org/10.1109/18.52478).
- Sheriff, R. E., 2002, *Encyclopedic dictionary of applied geophysics*: SEG, 4th ed.
- Sroubek, F., and P. Milanfar, 2012, Robust multichannel blind deconvolution via fast alternating minimization: *IEEE Transactions on Image Processing*, **21**, 1687–1700, doi: [10.1109/TIP.2011.2175740](https://doi.org/10.1109/TIP.2011.2175740).
- Stoffa, P., P. Buhl, and G. Bryan, 1974, The application of homomorphic deconvolution to shallow water marine seismology, Part I: Models: *Geophysics*, **39**, 401–416, doi: [10.1190/1.1440438](https://doi.org/10.1190/1.1440438).
- Stogioglou, A., S. McLaughlin, and A. Ziolkowski, 1996, Asymptotic performance analysis for fourth-order cumulant based deconvolution: 66th Annual International Meeting, SEG, Expanded Abstracts, 1587–1590.
- Ulrych, T., D. Velis, and M. D. Sacchi, 1995, Wavelet estimation revisited: *The Leading Edge*, **14**, 1139–1143, doi: [10.1190/1.1437089](https://doi.org/10.1190/1.1437089).
- Ulrych, T. J., 1971, Application of homomorphic deconvolution to seismology: *Geophysics*, **36**, 650–660, doi: [10.1190/1.1440202](https://doi.org/10.1190/1.1440202).
- VanDecar, J. C., and R. D. Crosson, 1990, Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares: *Bulletin of the Seismological Society of America*, **80**, 150–169.
- Velis, D., and T. Ulrych, 1996, Simulated annealing wavelet estimation via fourth-order cumulant matching: *Geophysics*, **61**, 1939–1948, doi: [10.1190/1.1444109](https://doi.org/10.1190/1.1444109).
- Verschuur, D. J., and R. J. Prein, 1999, Multiple removal results from Delft University: *The Leading Edge*, **18**, 86–91, doi: [10.1190/1.1438164](https://doi.org/10.1190/1.1438164).
- Walden, A. T., 1985, Non-Gaussian reflectivity, entropy, and deconvolution: *Geophysics*, **50**, 2862–2888, doi: [10.1190/1.1441905](https://doi.org/10.1190/1.1441905).
- Weatherburn, C. E., 1949, *A first course in mathematical statistics*: Cambridge University Press.
- White, R. E., 1988, Maximum kurtosis phase correction: *Geophysical Journal*, **95**, 371–389, doi: [10.1111/j.1365-246X.1988.tb00475.x](https://doi.org/10.1111/j.1365-246X.1988.tb00475.x).
- Wiggins, R. A., 1978, Minimum entropy deconvolution: *Geophysical Journal*, **16**, 21–35, doi: [10.1016/0016-7142\(78\)90005-4](https://doi.org/10.1016/0016-7142(78)90005-4).
- Xu, G., H. Liu, L. Tong, and T. Kailath, 1995, A least-squares approach to blind channel identification: *IEEE Transactions on Signal Processing*, **43**, 2982–2993, doi: [10.1109/78.476442](https://doi.org/10.1109/78.476442).
- Xu, Y., P. Thore, and S. Duchenne, 2012, The reliability of the kurtosis-based wavelet estimation: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-1221.1](https://doi.org/10.1190/segam2012-1221.1).