V SIMPÓSIO BRASILEIRO DE GEOFÍSICA
V SIMBGF
SBGf

Multiple attenuation via Blind Source Separation methods

Camilo P. F. Muñoz, Carlos A. da Costa Filho, Leonardo T. Duarte, João M. T. Romano, Martin Tygel, UNICAMP*

ABSTRACT

Multiple elimination is a very important step in seismic data processing. It is commonly undertaken by methods using least squares (LS) filters to perform adaptive subtraction. However, these techniques usually cause some collateral damage on primaries that intersect with multiples. This is due to the LS filters whose optimization criterion relies on removing as much energy as possible around the predicted multiples. In recent works, Blind Source Separation (BSS) methods have been applied to geophysics and have shown interesting results for multiple extraction. These techniques are able to identify and separate primaries from multiples without any adaptive subtraction, hence minimizing error when primaries and multiples overlap. In view of these initial results, we present in this paper a study on the application of BSS techniques to the problem of multiple extraction. Our study encompasses the analysis of different BSS methods and their application to a number of scenarios, considering both synthetic and real data.

INTRODUCTION

Many popular algorithms in seismic processing, such as various migration and inversion schemes require as input seismic data without multiple reflections. Very often, though, field-measured seismic data present multiples that interfere with these algorithms and thus constitute a recurrent problem in seismic data processing. The phenomenon can be quite disturbing, especially for shallow water marine data. Several techniques have been applied to try to attenuate some or all of these multiples. For instance, Verschuur (2006) describes methods based on moveout discrimination to separate primaries and multiples in different domains by applying suitable transforms. Popular transforms for these methods include the parabolic Radon transform, the f–k transform and the linear Radon transform (τ–p transform). Hence, in these domains it may be possible to mute regions associated to multiples and then revert to the original domain, preserving the primaries. However, it is recognized that multiples still leak to the unmuted regions and additional filtering is necessary to improve results with these methods.

A second class of methods is based on a two-folded procedure composed of prediction and extraction of multiples. Currently, prediction can be performed in shallow water by predictive deconvolution. Multiples in data with small offsets (offset $\approx {}^{1}/_{10}$ of depth) may be estimated by this technique based on a calculated time shift of data. But more generally, for data with good coverage, surface related multiple elimination (SRME) prediction algorithm is the most precise prediction method (Verschuur et al., 1992).

Once predicted, multiples must be extracted from the original data. Usually this step is undertaken by least squares (LS) filters. Using a minimal energy criterion, SRME prediction is fitted to original data and then adaptively removed. However regions where overlapping of primaries and multiples occur are frequently blurred or totally erased during the extraction phase of processing.

A more recent approach to the problem of multiple attenuation has been introduced by Lu (2006) then developed by Kaplan and Innanen (2008) and Donno (2011) in the past recent years. These approaches suggests that the extraction may be performed by Blind Source Separation (BSS) algorithms with no further need for adaptive subtraction. Results have shown that these methods present better results in primary and multiple overlapping regions (see, for instance, Kaplan and Innanen, 2008; Donno, 2011). Still, the works on this subject have mainly focused on BSS methods based on Independent Component Analysis (ICA). Although this class of methods is certainly the most popular in BSS, recent advances in the area have shown that better performance may be achieved by exploiting prior information that are not taken into account by ICA techniques. In particular, much attention has been paid to methods that explore the sparseness of the desired signals (Comon and Jutten, 2010).

In the search of a better understanding of the application of BSS methods to the problem of multiple extraction, we provide in this paper a set of experiments considering different classes of solutions and their application to different scenarios. Our major goal is to identify which solutions are better suited to perform multiple extraction in the scenarios considered in our study. Moreover, we also aim at analyzing the influence of some practical parameters, such as the size of the windows in which BSS is performed.

METHODOLOGY/INVESTIGATED PROBLEM

BSS methods allow us to retrieving a set of desired signals, often called sources, based only on input signals that correspond to mixtures of the original sources. When the mixing process is modeled by a linear instantaneous system, the mixed signals, which are represented by the vector $\mathbf{x}$, are given by:

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{A}$ is the mixing matrix and $\mathbf{s}$ denotes the vector of sources.

With very few hypotheses about the mixing process and the sources, it is possible to retrieve the sources from the observed mixtures. For instance, a popular approach to perform BSS in linear models is based on independent component analysis (ICA), which works under the assumption that the sources can be modeled as statistically independent random variables. Given that the mixing process makes the observed signals statistically dependent, the idea behind ICA is to adjust a separating matrix $\mathbf{W}$ so that $\mathbf{y} = \mathbf{W}\mathbf{x}$, which corresponds to the retrieved sources, be as independent as possible (Romano et al., 2011; Comon and Jutten, 2010).

In geophysics, more recent applications of ICA have proved to be competitive with least squares adaptive subtraction. The main asset in BSS-based solutions is the capability to identify and retrieve multiples without minimizing an energy criterion. In the problem of multiple extraction, two sources are modeled to take into account the primaries and multiples respectively, that is:

$$\mathbf{s} = \begin{bmatrix} \mathbf{d}_0^T \\ \mathbf{d}_1^T \end{bmatrix} \tag{2}$$

where $\mathbf{d}_0^T$ and $\mathbf{d}_1^T$ denote the signals associated with the primaries and multiples, respectively. Typically, BSS methods are not able to predict the order in which sources will be found; $\mathbf{d}_0$ and $\mathbf{d}_1$ may commute places inside $\mathbf{s}$. Thus, it is important to have visual or numerical criterion such as correlation between matrices to properly identify $\mathbf{d}_0$ and $\mathbf{d}_1$.

Concerning the mixing process, the approach proposed by (Donno, 2011) considers two mixtures. The first is given by the data (typically a common-shot gather), which is clearly a mixture of primaries and multiples. The second mixture is obtained by multiple prediction methods. The rationale behind this approach is that prediction methods do not provide a perfect multiple estimation. Therefore, the obtained signal can also be regarded as a mixture of primaries and multiples. In mathematical terms, the mixtures are thus given by

$$\mathbf{x} = \begin{bmatrix} \mathbf{d}^T \\ \mathbf{m}_e^T \end{bmatrix}. \tag{3}$$

where $\mathbf{d}$ corresponds to the data and $\mathbf{m}_e$ to a multiple estimation, which can be obtained by different methods such as SRME and predictive deconvolution.

Both the data $\mathbf{d}$ and the multiple estimation $\mathbf{m}_e$ must be vectorized since the mixing model proposed by (Donno, 2011) treats seismic data in 2D time-space windows instead of in a trace by trace basis. Therefore 2D data must be properly converted to 1D vector in order to preserve mixture and sources dimensions.

Besides ICA, the problem of BSS can be dealt with by alternative approaches based on information other than statistical independence. For instance, in second-order methods (Romano et al., 2011), separation is conducted by taking into account the correlation structure of the data for different time offsets. Therefore, it is more suitable to be used with time structured data, such as seismic traces. In our investigation, we shall consider the most popular second-order method: the SOBI algorithm. This technique is based on a joint diagonalization procedure of several correlation matrices.

Another recent approach to the problem of BSS considers the case in which the sources are sparse. The sparsity here can take place in time or in other domains (frequency, time-frequency, etc) — seismic data are often sparse in time. Among the different methods to retrieve sparse sources, the recent approach proposed by (Duarte et al., 2011) is able to separate sources that have a different degree of sparsity. In the problem tackled in this paper, the sources have indeed different degrees of sparsity, since traces containing only primaries are clearly sparser than traces containing multiples. The method proposed in (Duarte et al., 2011) is based on the minimization of a smoothed version of the ℓ_0 norm, which, roughly speaking, measures the sparsity of a given signal.

The three BSS strategies discussed above (ICA, SOBI algorithm, and sparsity-based ℓ_0 norm method) will be investigated in our work. We also conduct two new studies. First, we consider a different mixing model with estimated primaries instead of usual $\mathbf{m}_e$ estimated multiples. Secondly we will search for better efficiency of algorithms by concatenating ICA techniques with ℓ_0 norm separation. We expect to profit from different data structures and optimization processes to obtain better source separation.

Other aspects of separation that were not discussed previously will be studied. For instance, 2D window size choice has proved to have influence over the final separation result. We also investigated separation with other estimators such as non-deconvolved SRME and predictive deconvolution for large and small offsets.

RESULTS

We here consider several scenarios with the aim of assessing the performance of the separation algorithms discussed above. We perform tests considering synthetic data as well as real data. The synthetic datasets allow us to use both visual criteria and numeric measurements based on the normalized L^2 norm of the error between the actual and retrieved signals for evaluating the performance of the algorithms.

In tests 1, 2, and 3, we compare three separation algorithms. The first one, SOBI, was chosen for its ability to separate time structured data. FastICA, on the other hand, separates sources according to a nongaussianity criterion, while the method based on ℓ_0 norm minimization (for short, referred as ℓ_0) separates sources based on their sparsity.

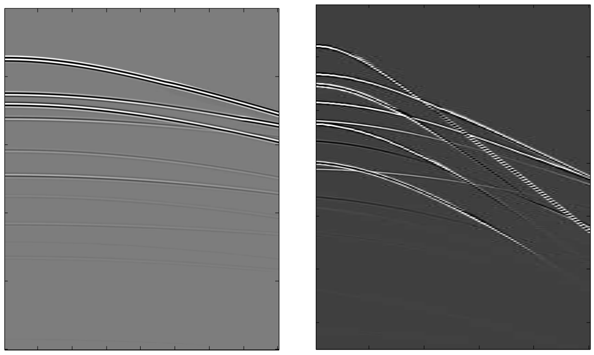
Test 1 intend to identify which criteria would be more efficient on seismic data and we consider two synthetic datasets in this case. The first dataset (Data1) was generated with the help of the NORSAR-2D platform over a four horizontal reflectors model with constant velocity. Data2 model was generated in a similar manner to that of Data1; however, the simpler geological structure of Data1 was replaced by a more elaborate model with curved reflectors that formed both shallow and deep water regions. Figure 1 shows a common shot of each one of these datasets.

Table 1: L^2 error of the estimated primaries using true multiples $\mathbf{m}$ and true primaries $\mathbf{p}$.

	Data1		Data2	
	$\mathbf{m}$	$\mathbf{p}$	$\mathbf{m}$	$\mathbf{p}$
SOBI	0.75%	0.75%	0.13%	0.13%
FastICA	8.98%	8.97%	0.00%	0.00%
ℓ_0	0.75%	0.36%	0.00%	0.00%

As mentioned before, the mixtures comprises the fully vectorized data $\mathbf{d}$ and the vectorized multiple estimation $\mathbf{m}_e$. Vectorized exact synthetic multiples are used to act as perfect estimators. Thereby, we assure that all possible estimation bias that could be introduced by multiples estimation techniques is removed. Table 1 shows that normalized L^2 norm error measurements have been similar on Data1 for SOBI and ℓ_0 norm algorithms, while FastICA introduces a larger error, of approximately 9%, to multiple separation.

We then reproduced the same test on Data2. According to Table 1, the smallest error for Data2 was obtained for FastICA algorithm instead of SOBI.



(a) Data1 dataset. (b) Data2 dataset.

Figure 1

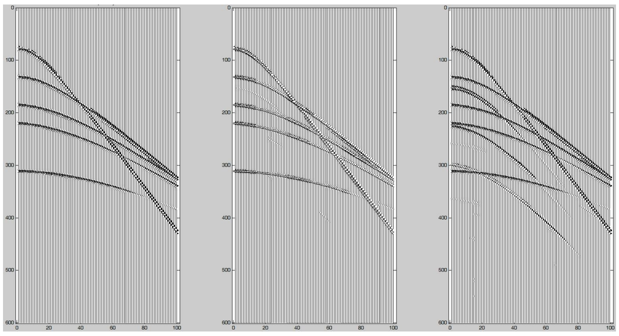


Figure 2: On the left are the true primaries, on the center the primaries recovered by FastICA and on the right those recovered by SOBI.

In order to illustrate visually the results presented in Table 1, we show in Figure 2. At the center, we see that FastICA was able to separate these same primaries from multiples. However, the estimation provided by the SOBI algorithm was corrupted by strong residual multiples.

Differences in algorithm performances for Data1 and Data2 indicate that there is no best *a priori* choice of a more efficient separation algorithm on seismic data.

Test 2 aims at extending the approach proposed by (Donno, 2011). More precisely, we intend to verify if it is possible to perform source separation by using primaries as estimators in the initial mixture, i.e., $\mathbf{m}_e$ is replaced by estimated primaries.

The test was performed on Data1 as well as on Data2. To avoid errors due to changes on estimation of multiples and primaries techniques that could be misleading, we considered the true primaries as estimator. The normalized L^2 norm measurements, shown on Table 2, indicate that separation is as efficient with primaries as input as it is with multiples as input. Equal results were obtained for all three algorithms in both data sets presented. It is important to notice that small variations in measurements may be neglected due to small differences that might occur during the optimization process of the algorithms. Higher variations can be seen in the results provided by the method based on the ℓ_0 norm. This is possibly due to the evolutionary algorithm in this method — indeed, evolutionary algorithms may converge to solutions with a wider fitness range according to the number of runs executed.

In **Test 3**, we analyze if better performance can be achieved by sequentially applying different separation algorithms.

Let us take for example the SOBI algorithm associated with the ℓ_0 norm. One could expect to separate sources using time properties of the data with SOBI. Then, by introducing the output of the SOBI algorithm on a new ℓ_0 norm run, it could be possible to enhance the separation by taking into account the sparsity of the data. Multiples obtained as an output of the SOBI separation were introduced as estimator for multiples in the input mixture of the ℓ_0 norm algorithm. This test was made in both Data1 and Data2. After that, it was reproduced in the inverted order using first the ℓ_0 norm, and applying its output to the SOBI input mixture.

Results shown in Table 2 indicate that the final output will never attain a lower error than that of the best fitted algorithm. For instance, for the ℓ_0 norm concatenated with SOBI on the Data1 set, the measurements overall presented a 0.75% error, which is equal to the result obtained by applying the SOBI algorithm alone. In this case, the concatenation of two algorithms gave the same result as for the best fitted algorithm. In the inverted order, SOBI separation first and ℓ_0 norm afterwards, the result changes and we obtain a final error of 0.77%, which is a slightly worse result than applying only the SOBI algorithm. Table 2 shows other examples of this behavior.

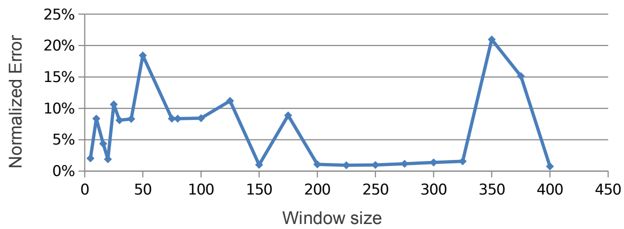
Table 2: L^2 error of the estimated primaries using different algorithms.

	Data1	Data2
SOBI + ℓ_0	0.77%	0.05%
FastICA + ℓ_0	8.55%	0.10%
ℓ_0 + SOBI	0.75%	0.13%
ℓ_0 + FastICA	8.97%	0.00%

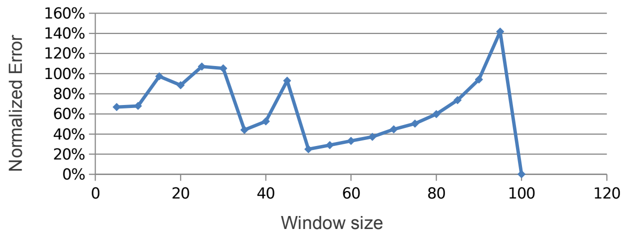
Test 4 intend to analyze window sizes as a parameter for 2D windowed separation. Previous articles that studied BSS applications on geophysical data discussed different forms of performing ICA on data. At first it was applied on a trace-by-trace basis in Kaplan and Innanen (2008), but it has been more recently shown that separation is more efficient if executed on 2D data windows (Donno, 2011). However, these 2D windows introduce a new parameter, namely the window size, that should be well adjusted in order to efficiently increase the ICA based separation.

We have tested several window sizes in both Data1 and Data2. Hence, we obtained normalized L^2 norm curves that describe the evolution of error according to the 2D window size choice. Separation was performed by the SOBI algorithm for both datasets. According to Figure 3, the error evolution for increasing window size is quite unstable. On the other hand, it seems that the fundamental shape of the curve is conserved for different datasets. We can also note that the best global result was always obtained for one single window that performs the BSS on all the data at once.

Figure 3 gives additional information for window sizes that are near the superior half data size. These window sizes have small L^2 norm error values. They seem to form the most stable region where window sizes may be chosen for processing without having to consider the entire data. This might be of use for very large data sets that must be read from disk in parcels.



(a) Data1



(b) Data2

Figure 3: Normalized L^2 errors for differently sized windows in Data1 and Data2 datasets.

In **Test 5**, we consider a more realistic situation, in which a perfect estimation of the multiples is not available. Indeed, in the application of the studied methods to real data, estimation of multiples will become a main issue to obtain a good separation and elimination of multiples. To investigate this issue, we have considered a new synthetic dataset (Data3) generated via finite differences methods on wave equations over a 1D model. It consists of a simple model with a single primary and the first two multiples.

Estimation tests were ran on four different scenarios. First deconvolved SRME (dSRME) was used as estimator to compare ICA efficiency with usual SRME adaptive subtraction performed by an LS filter. In the second scenario, we used the same SRME estimation but without the deconvolution step that is used to account for the source signature (ndSRME). Finally, two tests were made for predictive deconvolution (PD and soPD, respectively) using all available offsets and then only smaller offsets (largest offset ≈ depth). The tests were made with a single windowed FastICA algorithm.

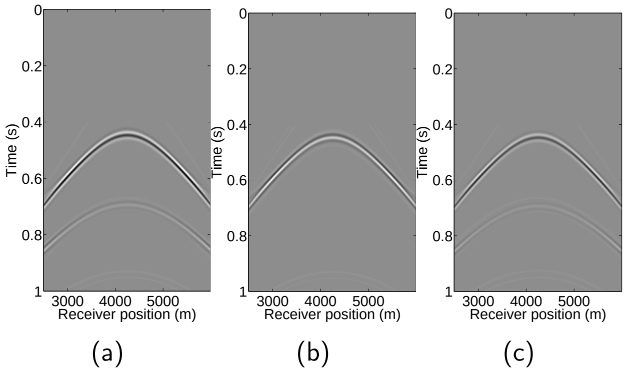


Figure 4: The original common shot record (a), the primaries using deconvolved (b) and nondeconvolved (c) SRME data.

For a more precise comparison of extraction with ICA using deconvolved SRME estimator and least squares adaptive multiple subtraction, several filter sizes were tested. As illustrated in Table 3, the best results were obtained using a finite impulse response (FIR) filter with 105 taps (1.5 times the source wavelet length). In this case, ICA presents a slightly smaller error of 3.41% against 3.53% for the LS subtraction.

Table 3: L^2 error of the estimated primaries.

	Number of filter points				
	1	35	70	105	140
ICA	23.16%	3.76%	3.70%	3.41%	6.85%
LS	23.83%	3.78%	3.77%	3.53%	7.10%

Furthermore, according to Table 4 the error increases considerably if there is no wavelet amplitude adaptation before executing the ICA algorithm. With 24.83% error extraction simply attenuates multiples intensities without actually eliminating them (see Figure 4). The same order of error occurs when predictive deconvolution is applied to any offset. However for smaller offsets, predictive deconvolution exceeds all other estimators for ICA separation with a small error of 3.38%. This may be applied to shallow water multiple elimination with little computational cost and good results.

Table 4: L^2 errors using different estimators in comparison to standard LS subtraction.

Subtraction	Estimator used by FastICA			
LS	dSRME	ndSRME	PD	soPD
3.53%	3.41%	24.83%	27.40%	3.38%

To conclude our work, in search of better understanding of BSS methods applied to geophysics, we consider **Test 6**, in which we study multiple extraction on real marine data of the Jequitinhonha Bay, kindly ceded by PETROBRAS. The estimated multiples were given by SRME. FastICA, SOBI and ℓ_0 norm algorithms were used in this study to compare the efficiency of multiple extraction on this new dataset. In this case, only visual criteria can be used to evaluate the algorithms performance since we lack information about exact multiples.

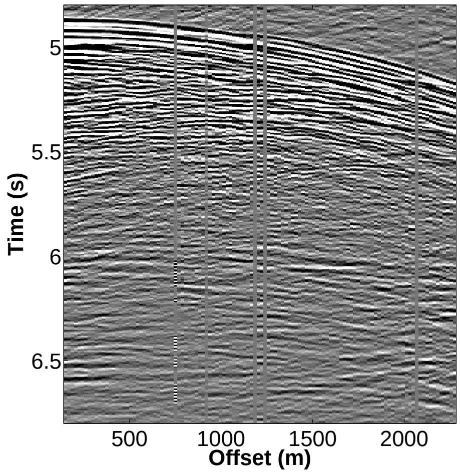
Applied to this data, SOBI algorithm and FastICA seem to have given better separation results. Figure 5 shows a comparison of the best fitted separation algorithm and LS filters extraction. Though largely similar, the SOBI method seems to have given a better result.

DISCUSSION AND CONCLUSIONS

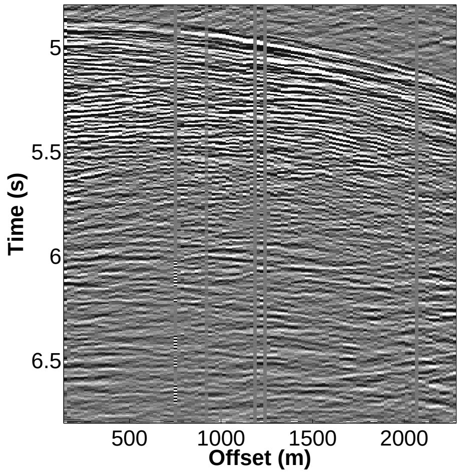
In this work, we investigated the application of source separation methods to the problem of multiple attenuation. This approach has been already exploited in previous works, in which source separation was conducted by means of methods on the nongaussianity maximization approach. In our study, though, we also considered methods based on other types of information, such as temporal structure and sparsity.

Overall, in the performed tests, the algorithms under study (FastICA, SOBI, and method based on the ℓ_0 norm minimization) provided similar results. We have also shown that primaries may be used as estimators instead of multiples with no damage to results. This may be at use in cases where multiple estimation is more complicated then estimating primaries. In addition, Test 3 indicated that sequentially applying separation techniques does not improve the global identification of multiples. We also provided some results to analyze the influence of the window size on the separation performace. All these tests allowed us to consider different settings in order to search for new ways to enhance the performance of BSS algorithms in the context of multiple extraction.

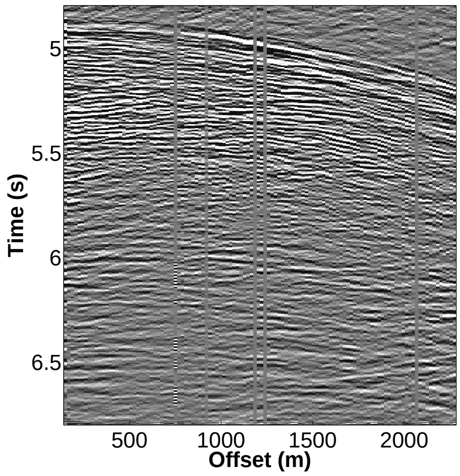
In this study, we also observed that a proper estimation of the primary remains the most important aspect of multiple extraction with separation methods. Indeed, best results were obtained for deconvolved SRME and predictive deconvolution restricted to offsets smaller or



(a) Original common shot record.



(b) Primaries by LS subtraction.



(c) Primaries by SOBI.

Figure 5: Jequitinhonha Bay dataset and two demultipled common shots. The data has been clipped and time cropped.

equal to seafloor depth. Further studies might be developed in search for better estimation on primaries and multiples overlapping regions. This might highly increase BSS ability to extract multiples even though results on real data show these techniques are already competitive with usual extraction techniques.

ACKNOWLEDGMENTS

The authors would like to thank PETROBRAS for partially funding the work and for conceding the Jequitinhonha dataset. We are also indebted to thank Yakov Nae, for providing the Data1 and Data2 datasets, and Everton Nadalin for his advice and cooperation.

REFERENCES

Comon, P., and Jutten, C., 2010, Handbook of Blind Source Separation: Independent Component Analysis and Applications: Academic Press.

Donno, D., 2011, Improving multiple removal using least-squares dip filters and independent component analysis: Geophysics, **76**, no. 5, V91–V104.

Duarte, L. T., Suyama, R., Attux, R. R. F., Romano, J. M. T., and Jutten, C., 2011, Blind extraction of sparse components based on ℓ_0-norm minimization: Blind extraction of sparse components based on ℓ_0-norm minimization:, Proc. of the IEEE Workshop on Statistical Signal Processing (SSP2011), 617–620.

Kaplan, S. T., and Innanen, K. A., 2008, Adaptive separation of free-surface multiples through independent component analysis: Geophysics, **73**, no. 3, V29–V36.

Lu, W., 2006, Adaptive multiple subtraction using independent component analysis: Geophysics, **71**, no. 5, S179–S184.

Romano, J. M. T., Attux, R., Cavalcante, C., and Suyama, R., 2011, Unsupervised signal processing: channel equalization and source separation: CRC Press.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**.

Verschuur, E., 2006, Seismic multiple removal techniques: past, present and future:, Education Tour Series EAGE Publications bv.