André K. Takahata, Everton Z. Nadalin, Rafael Ferrari, Leonardo Tomazeli Duarte, Ricardo Suyama, Renato R. Lopes, João M.T. Romano, and Martin Tygel

Unsupervised Processing of Geophysical Signals

A review of some key aspects of blind deconvolution and blind source separation

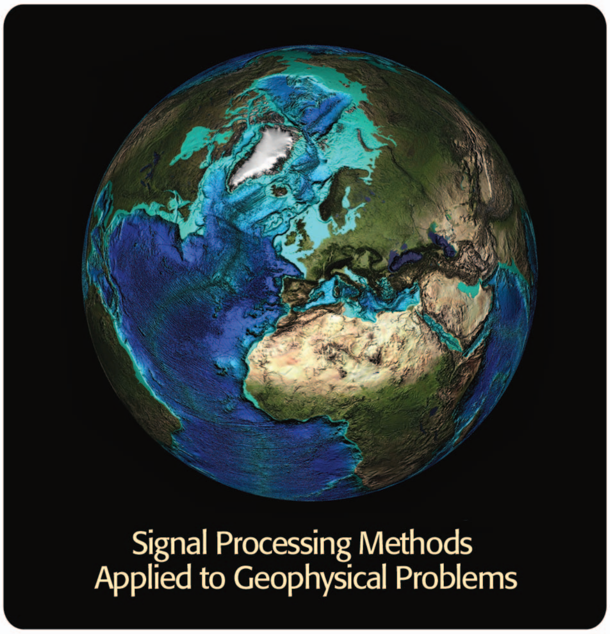


IMAGE COURTESY OF U.S. DEPARTMENT OF COMMERCE/NOAA/NESDIS/NATIONAL GEOPHYSICAL DATA CENTER

Unsupervised signal processing has been an exciting theme of research for at least three decades. It finds the potential application in practically all fields where well-established techniques of digital signal processing have been employed, including telecommunications; speech and audio processing; image, radar, and sonar; and biomedical signals. Among these classical problems, geophysical signal processing has played a prominent role in the development of unsupervised methods. In fact, the field of unsupervised processing can be said to have started with the early application of Wiener's theories to seismology.

Wiener filters involve second-order statistics (SOS), specifically correlation and power spectrum density, and are built over the theoretical framework of linear systems and Gaussian signals. In contrast, unsupervised filtering uses higher-order statistics (HOS), enabling the weakening of classical assumptions about the systems and signals under study. Thus, in the last few years, there has been an increasing interest in moving from correlation-based methods to unsupervised techniques in seismic signal processing.

This article reviews some key aspects of two important branches in unsupervised signal processing: blind deconvolution and blind source separation (BSS). It also gives an overview of their potential application in seismic processing, with an emphasis on seismic deconvolution. Finally, it presents illustrative results of the application, on both synthetic and real data, of a method for seismic deconvolution that combines techniques of blind deconvolution and blind source separation. Our implementation of this method contains some improvements over the original method in the literature described in [1].

INTRODUCTION

Reflection seismic plays a fundamental role in deriving information on the subsurface of an area under analysis. To achieve this, seismic waves are produced from sources such as dynamite or airguns. These waves are reflected on the subsurface, and sensor grids located in the surface measure the reflections. These measurements then undergo intense processing, involving significant

Digital Object Identifier 10.1109/MSP.2012.2189999
Date of publication: 12 June 2012

human and computational effort. The goal is to provide information on the subsurface to applications such as imaging and parameter estimation of geological structures that are relevant to exploration and monitoring of hydrocarbon reservoirs, assessing sites for CO_2 sequestration and nuclear waste deposition.

From a signal processing perspective, establishing the connection between the signal sources and the sensors is a fundamental and often challenging problem for two main reasons. First, the seismic source and the subsurface propagation are not ideal, which distorts the useful signals. Second, the output is a mixture of different waves, which must be identified and separated. In many cases, a suitable hypothesis consists in considering linear distortions and mixtures. Thus, a convolution relationship models the fact that the source signal and the propagation environment are not ideal, akin to a transmission through a linear channel, while the mixing process consists of a linear combination of the waves. To recover the useful information, channel deconvolution and source separation can be performed by different methods [2].

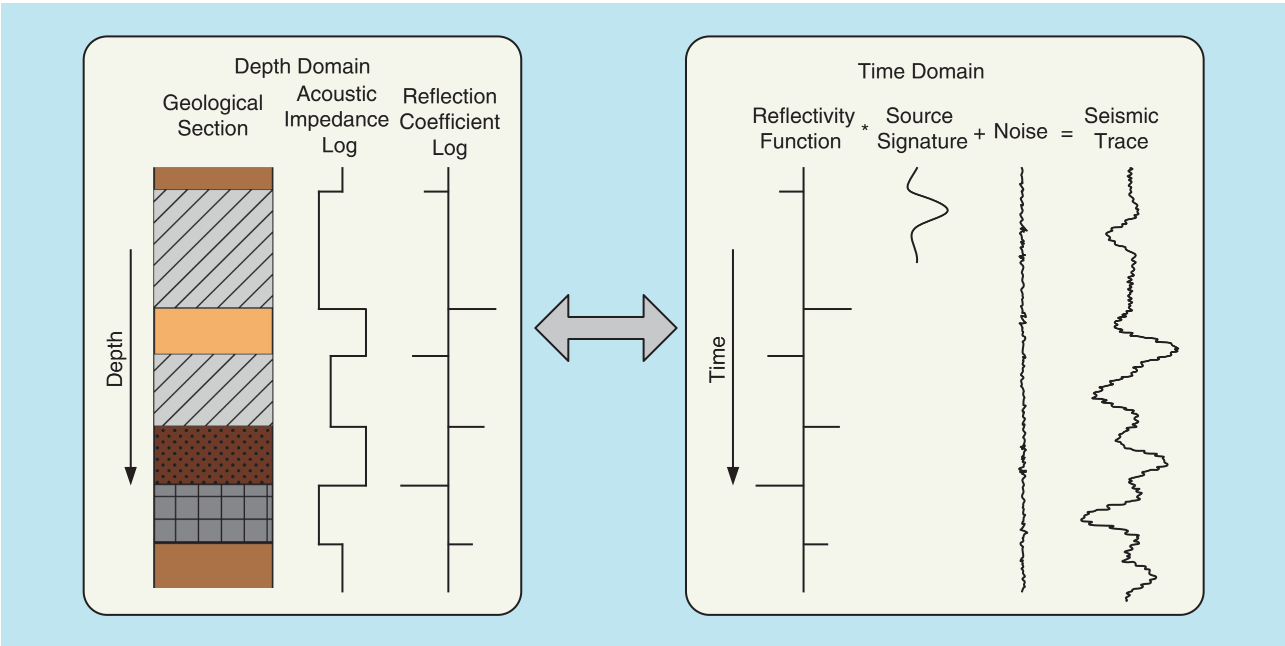
A fundamental question comes before the choice of a deconvolution/separation method: In addition to the measured data, what information is available on the input sources and the convolution/mixing system? In fact, if one of them is known or well estimated, the processing task is rather simplified and classical inversion methods, as the ones described in [3], can be used. In the literature of statistical signal processing, these are also referred to as supervised methods.

However, the use of supervised techniques may not be possible if the information on the input sources or the convolutive/mixing system is unavailable, or its estimation is undesirable. In any case, the lack of information requires the use of unsupervised, or blind, signal processing, which is the subject of this article. In the following, we describe unsupervised techniques and their application to seismic processing.

SEISMIC DECONVOLUTION

In reflection seismic data processing, deconvolution is used to estimate, from the measured data, the reflectivity function of the subsurface, which is the earth response to an ideal impulsive, seismic source. Mathematically, the measured samples, $x(n)$, are assumed to be the convolution of the reflectivity function, $\rho(n)$, with the source signature, $h(n)$. With the addition of a noise term, $b(n)$, the seismic trace is given by $x(n) = h(n) * \rho(n) + b(n)$, as shown in Figure 1. In some cases, such as in marine seismic exploration and vibroseis-based land exploration, the source signature can be estimated [4]. Under some conditions this information can be used to deconvolve the seismic trace in a supervised fashion, which is often referred to as deterministic deconvolution in seismology [4], [5]. It is relatively straightforward to develop both time-domain and frequency-domain techniques to perform supervised deconvolution. A common approach is to estimate the reflectivity function through linear filtering, using the Wiener-Levinson minimization of the mean-squared-error (MSE)[3].

When the source signature is not available, the estimation of the reflectivity function becomes an unsupervised task, which must be carried out based only on the observed samples and on a minimum amount of hypothesis about the source and the system. However, even in the absence of noise, this is an ill-posed problem, because multiple combinations of source signatures and reflectivities can result in the same seismic trace. The challenge is to



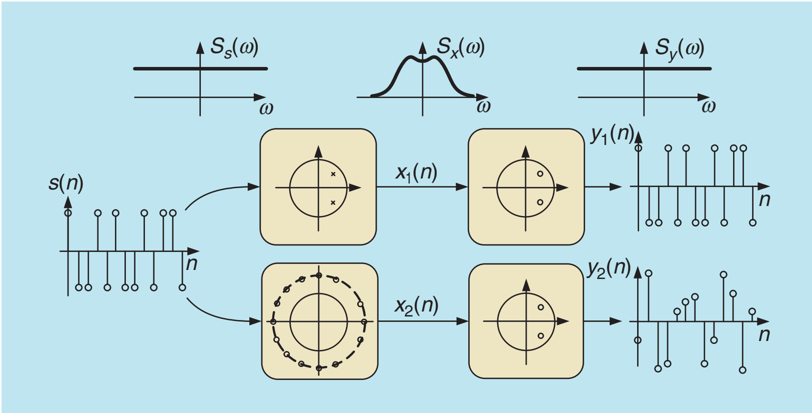
[FIG1] The properties of the subsoil layers, density and seismic velocity, determine their acoustic impedance. It is possible to calculate a reflectivity function of the geological section by considering the change of these properties at the boundaries between layers. In the convolutional model, the observed trace is the reflectivity function convolved with the source signature and corrupted by a random additive noise.

overcome this ambiguity by exploiting prior knowledge about the structure of the source signature and the reflectivity.

PREDICTIVE DECONVOLUTION

Interestingly, one of the first applications of unsupervised signal processing was in seismic deconvolution. This application was proposed in 1954 by Enders A. Robinson in his Ph.D. dissertation [6], in which he showed that the recently developed Wiener theory on prediction and filtering [7] could be used for predictive, unsupervised deconvolution. Robinson considered the following two ad hoc hypotheses [8]:

1) the seismic wavelet is the impulse response of an all-pole, minimum phase system
2) the impulse response of the layered earth model behaves like a decorrelated (white) signal, so that it has a flat frequency spectrum.



[FIG2] The white input signal $s(n)$ feeds a minimum-phase all-poles system and a maximum-phase all-zero systems both with the same magnitude response. Thus, the output signals $x_1(n)$ and $x_2(n)$ have the same power spectrum density $S_x(w)$. The minimization of the prediction mean squared error leads to the same prediction-error filter for both $x_1(n)$ and $x_2(n)$. However, in the first case the prediction-error filter exactly corresponds to the inverse of the minimum-phase all-poles system, while it compensates only the magnitude distortions in the second case. Hence $y_1(n)$ and $y_2(n)$ are both decorrelated signals but only $y_1(n)$ recovers the original source $s(n)$.

In other words, the observed signal can be modeled as an autoregressive process, so that the model parameters and the impulse response can be estimated by linear prediction [2].

Deconvolution employing prediction-error filters was quite appealing, since it uses only the SOS (correlation) of the measured signal, which simplifies the task significantly. However, the soundness of second-order methods is limited by the effectiveness of the underlying hypotheses. First, a prediction-error filter acts as a whitening filter, so that ideally it recovers an uncorrelated signal. In other words, if the signal to be recovered is correlated, a prediction-error filter necessarily introduces residual distortions. Second, and perhaps most importantly, if the seismic wavelet cannot be modeled as an all-pole, minimum phase filter, the prediction-error filter is still ideally able to recover a white reflectivity, but different from the actual one.

Figure 2 illustrates the fundamental limitation of linear predictive deconvolution by considering a white source to be recovered and two convolution systems whose frequency responses have the same magnitude but different phases. Since predictive deconvolution is based only on SOS, it does not succeed in recovering the input of the nonminimum phase system. In this scenario, the minimization of the mean squared prediction error leads to the same prediction coefficients for both systems. Nevertheless, only the signal recovered from a minimum-phase system corresponds to the original source. This, in fact, is a well-known limitation of SOS: the power spectrum density of the input and output signals of a linear time-invariant system are related by the magnitude of the frequency response of the system, and these quantities do not carry any phase information [2], [3].

FUNDAMENTAL THEOREMS

To overcome the limitations of predictive deconvolution, unsupervised techniques must go beyond correlation and power spectrum density and must instead deal with HOS. The benefits of HOS, such as their ability to recover inputs of nonminimum phase systems, are established by the two following results.

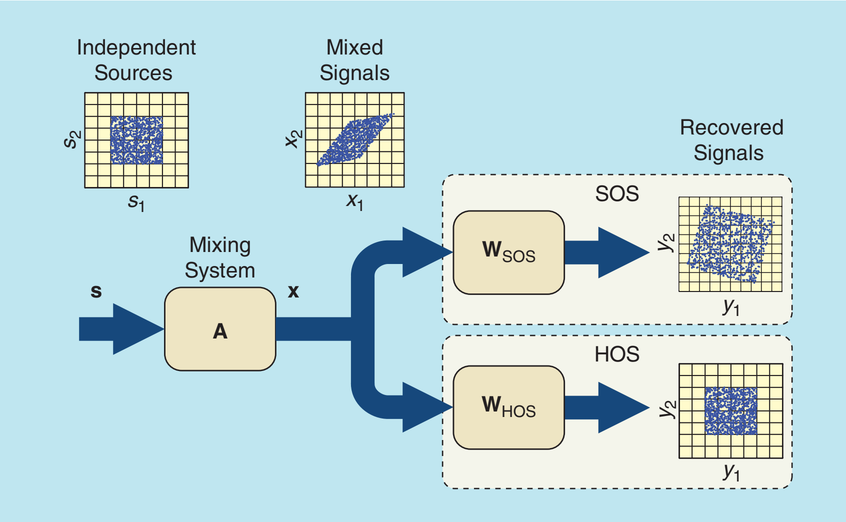
The Benveniste-Goursat-Ruget (BGR) theorem [9] considers a scenario in which a source signal, $s(n)$, composed of non-Gaussian independent and identically distributed (i.i.d.) samples is observed through a noiseless, linear time-invariant convolution system with output $x(n)$. An estimate of the source, $y(n)$, must be obtained from $x(n)$ and the knowledge of the probability density function (pdf) of $s(n)$. Under these conditions, the theorem states that, if the pdfs of $y(n)$ and $s(n)$ are the same, then correct estimation is attained, except for a delay, d, and a complex unit-magnitude gain, α, i.e., $y(n) = \alpha s(n-d)$.

This result is crucial, since it establishes the viability of obtaining an efficient deconvolution filter, with the sole aid of statistical properties of the source, and without any additional assumption about the input signal or the system impulse response. However, it does require more than simply working with SOS, since the recovery involves the pdfs of the signals.

The BGR theorem was the first solid theoretical justification for unsupervised deconvolution. However, a decade later, another important result demonstrated that the condition of matching pdfs of this theorem was excessively stringent.

The Shalvi-Weinstein (SW) theorem [10] considers the same scenario of the BGR theorem but shows if $E\{s(n)^2\} = E\{y(n)^2\}$ and a nonzero cumulant of an order higher than two of $s(n)$ and $y(n)$ are equal, then the recovered signal $y(n)$ corresponds to a delayed and possibly scaled version of the transmitted signal $s(n)$. Intuitively, this means that, after a sort of power normalization, an HOS can be used to find an effective deconvolution filter.

The importance of the SW theorem is notorious, since it assures that it is not necessary to match the pdfs, and hence match all the moments, of the two signals of interest to accomplish blind deconvolution: it suffices to take into account the second-order moment and, for instance, the fourth-order



[FIG3] The mixed signal, x, goes through a separation algorithm using SOS and HOS. In the first case, the scatter plot shows that the obtained signals are uncorrelated, but do not recover the source. By using HOS in the second case, one can perfectly recover *s*.

information in the kurtosis, provided it is nonzero. An illustrative example of application of HOS in geophysics can be found in [11].

BLIND SOURCE SEPARATION

BSS is another fundamental topic of unsupervised processing theory [2]. Similar to blind deconvolution, BSS assumes only some statistical knowledge about sources and system. However, in BSS, the system has multiple inputs and multiple outputs. The problem can be stated as follows: at sample time n, a set of N signals of interest, $\mathbf{s}(n) = [s_1(n)\ s_2(n)\ \ldots\ s_N(n)]^T$, is generated by the sources, and observed through an environment that mixes them. The mixtures are captured by a set of M sensors, providing the observed signals $\mathbf{x}(n) = [x_1(n)\ x_2(n)\ \ldots\ x_M(n)]^T$. The goal of source separation algorithms is to estimate all source signals based only on the observations, by means, for instance, of a separating system. If the mixing system is linear, memoryless, and time-invariant, it can be described by an $M \times N$ matrix $\mathbf{A}$. In this case, if the number of sources and sensors is equal and the mixing matrix is nonsingular, the sources can be recovered by a matrix $\mathbf{W}$, which represents the inverse of $\mathbf{A}$. The question is how to compute $\mathbf{W}$ based only on the observations and some statistical knowledge of the sources and system.

A first idea to adjust $\mathbf{W}$ relies on SOS, i.e., $\mathbf{W}$ is adjusted so as to decorrelate the recovered signals $\mathbf{y}(n) = [y_1(n)\ y_2(n)\ \ldots\ y_M(n)]^T$. This procedure is usually called whitening and is in a way similar to the predictive approach of unsupervised deconvolution. The whitening procedure can be performed by a classical data analysis technique called principal component analysis (PCA) [2]. However, it cannot guarantee a proper separation as shown in the following example.

Figure 3 depicts a case in which two mutually independent sources, each with a uniform distribution, are mixed by a linear instantaneous mixture. The top separating system consists of a whitener, and is based only on SOS. As can be observed from the scatter plot of its output, this system does not recover the original sources. In fact, there remains a rotation factor, which cannot be inferred from SOS. The solution to this problem, based on HOS, and usually referred to as BSS, will be discussed in the following.

The seminal work in [13] proposed the first effectively BSS technique by adjusting the separating matrix $\mathbf{W}$. This is the essence of the Hérault-Jutten algorithm, which produces a set of signals $\mathbf{y}(n)$ that are nonlinearly uncorrelated, rather than just uncorrelated. The use of nonlinear devices results in an implicit use of HOS, which explains the good performance of the algorithm. Since it was the first BSS method to use HOS, and due to its efficiency, the Hérault-Jutten algorithm was a major breakthrough in BSS.

The need for HOS in the BSS problem was definitively clarified by the introduction of the concept of independent component analysis (ICA), first formalized by Comon in [14]. In contrast with the whitening approach, the main idea in ICA is to adjust the matrix $\mathbf{W}$ so that the recovered signals become mutually independent. If there is at most one Guassian source and the mixing matrix $\mathbf{A}$ is invertible, then ICA will separate the source signals [14], as shown in Figure 3. In other words, under some assumptions, recovering independent signals implies source separation.

It is noteworthy that, although blind deconvolution and source separation approaches have originated independently and in somewhat distinct scientific communities, strong relationships can be found between these two unsupervised problems. In both cases, dealing with SOS is not sufficient and leads only to whitening. Also, both solutions present ambiguities: while blind deconvolution cannot determine a scaling factor and a delay, BSS is constrained by scaling and permutation ambiguities. The latter means that, even though the sources can be recovered, their order cannot be determined.

Among the wide range of unsupervised processing techniques that can be applied to geophysical signals, this article is more focused on deconvolution, so that we present now only a brief comment about two distinct applications where the use of BSS is quite interesting:

MULTIPLE ATTENUATION USING BSS

In general, several seismic processing techniques rely on the assumption that the signals in the seismic data contain only primary reflections, that is, measurements of seismic waves that were reflected only once. In practice, however, the seismic records may contain readings on waves that suffered multiple reflections on the path between the seismic source and receiver. This may happen, for instance, in marine records when a wave reflects on the sea bottom, then on the surface, then on the bottom again before reaching the receiver. These multiple reflections, simply called multiples, may interfere with the primaries, obscuring them. Also, most seismic signal processing techniques will treat multiples as primaries, leading to images that do not reflect the actual subsurface geology. Thus, multiple attenuation is an important part of seismic signal processing.

One important class of multiple attenuation algorithms operates in two stages [15]. First, the multiples are predicted [16], [17] and, then, these predicted multiples are somehow subtracted from the seismic data [18]. However, the success of these methods greatly depends on the quality of the predicted multiples. To improve the quality of multiple subtraction, several works have applied BSS techniques to multiple attenuation. In this case, primary and multiple reflections are considered different signal sources to be separated. For further details, the reader is referred to [19]–[22].

WAVE SEPARATION

Wave separation is another seismic processing problem that may benefit from BSS. This problem arises because a seismogram is composed of different sorts of waves, such as ground roll, direct waves, and noise. Thus, a fundamental task in seismic processing is the separation of these different types of seismic waves, leading to the extraction or the enhancement of the information of interest. This problem is another clear example of unsupervised task, since only a few properties of the desired waves are taken into account during the recovery process. Among the different strategies to perform wave separation, much attention has been given to methods based on the singular value decomposition (SVD).

SVD is an ubiquitous tool in all branches of signal processing and data analysis and can also be used in a seismogram decomposition. In the seismic literature [23], [24], such decomposition gives rise to an eigenimage, obtained from the left singular vectors and the right singular vectors, which are referred as to propagation vectors and normalized wavelets, respectively, since they give the time and amplitude dependence of each eigenimage. One of the most emblematic examples of SVD application in this context can be found in [25], in which the authors consider the problem of separating the downgoing and upgoing wavefields in vertical seismic profiling (VSP) interpretation. Besides, SVD has been used in other applications, such as ground roll attenuation [26] and signal-to-noise ratio (SNR) enhancement [27].

The application of SVD to wave separation is particularly useful to isolate laterally aligned events, e.g., a horizontal event in a simulated zero-offset section. This approach is closely related to PCA. In [24], an alternative approach based on ICA is proposed. By searching for independent (instead of orthogonal) vectors, the method proposed in [24] presents better results when dealing, for instance, with dipping events mixed with horizontal events.

ICA-BASED SEISMIC DECONVOLUTION

In this section, a particular application of unsupervised seismic processing is studied in more detail: the banded ICA (B-ICA), introduced in [1]. We focus on this method because it combines the two problems addressed so far in this article: blind deconvolution and BSS. To fit the ICA framework, B-ICA models the convolution as a linear system of equations, $\mathbf{x} = \mathbf{A}\mathbf{s}$. In this formulation, $\mathbf{x} = [x(0)\ x(1)\ \ldots\ x(N-1)]^T$ and $\mathbf{s} = [\rho(0)\ \rho(1)\ \ldots\ \rho(N-1)]^T$ represent the trace and the reflectivity, respectively. The $(N \times N)$ convolution matrix, $\mathbf{A}$, has a banded Toeplitz structure whose nonzero elements correspond to the wavelet coefficients, and can be written as

$$\mathbf{A} = \begin{bmatrix} h(0) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ h(1) & h(0) & \cdots & 0 & 0 & \cdot & 0 \\ \vdots & h(1) & \ddots & \vdots & \vdots & \ddots & 0 \\ h(N_h-1) & \vdots & \ddots & h(0) & 0 & \ddots & 0 \\ \vdots & h(N_h-1) & \ddots & h(1) & h(0) & \ddots & \vdots \\ 0 & \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & h(N_h-1) & h(N_h-2) & \cdots & h(0) \end{bmatrix},$$

$$= [\mathbf{N}_1\mathbf{h}\ \ \mathbf{N}_2\mathbf{h}\ \ \cdots\ \ \mathbf{N}_N\mathbf{h}],$$

where $\mathbf{h} = [h(0)\ h(1)\ \ldots\ h(N_h-1)]$ corresponds to the seismic wavelet with N_h samples, and $\mathbf{N}_i$ corresponds to a zero padding matrix that maps $\mathbf{h}$ to the column $\mathbf{a}_i$ of $\mathbf{A}$. Now, the model $\mathbf{x} = \mathbf{A}\mathbf{s}$ only provides a single snapshot of the mixture vector $\mathbf{x}$, which is not suited for ICA. To cope with this issue, delayed versions of the trace and the reflectivity are defined as $\mathbf{s}(n) = [\rho(n-N+1)\ \ldots\ \rho(n-1)\ \rho(n)]^T$ and $\mathbf{x}(n) = [x(n-N+1)\ \ldots\ x(n-1)\ x(n)]^T$ in which $x(n) = 0$ and $\rho(n) = 0$ for $n < 0$. This process generates N snapshots $\mathbf{s}(n)$ and $\mathbf{x}(n)$, for $n = 0\ \ldots\ N-1$. This way, each source and mixture corresponds to a delayed version of the reflectivity and the convolved signal respectively. Note that ICA requires that the source consists of sequences of non-Gaussian, i.i.d., variables [2], which we assume to be satisfied by the reflectivity. Based on this model, the following steps are taken to perform the ICA-based seismic deconvolution [1]:

STEP 1–DATA REARRANGEMENT

$M < N$ mixtures are obtained such that $\mathbf{x}(n) = [x(n-M=1)\ \ldots\ x(n-1)\ x(n)]^T$ and $\mathbf{s}(n) = [s(n-M+1)\ \ldots\ s(n-1)\ s(n)]^T$. In this case, the linear relationship between $\mathbf{x}(n)$ and $\mathbf{s}(n)$ is only approximated and the observation matrix is written as

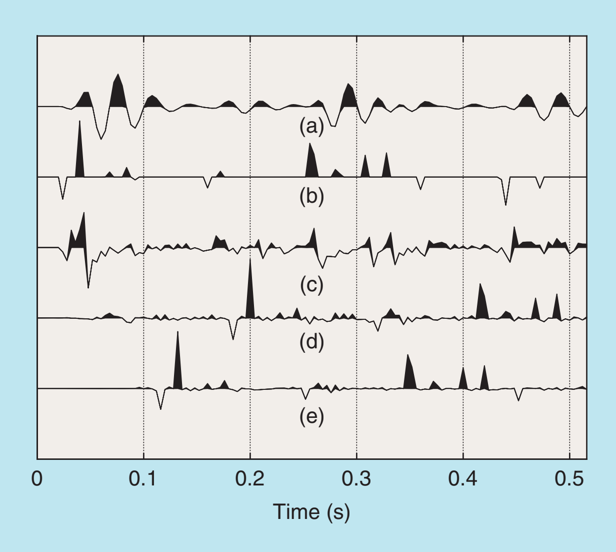
$$\mathbf{X} = \begin{bmatrix} \mathbf{x}^T(n-M+1) \\ \vdots \\ \mathbf{x}^T(n-1) \\ \mathbf{x}^T(n) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \cdots & 0 & x(0) & \cdots & x(N-M) \\ \vdots & \vdots & \iddots & \iddots & \vdots & \iddots & \vdots \\ 0 & x(0) & \cdots & \cdots & x(M-2) & \cdots & x(N-2) \\ x(0) & x(1) & \cdots & \cdots & x(M-1) & \cdots & x(N-1) \end{bmatrix}.$$

The first rows of the new data matrix have fewer zeros than in the full data matrix. As discussed in [1], this improves the statistical properties of the mixture matrix, which improves the performance of the ICA algorithm.

STEP 2–DATA WHITENING

A whitening processing is used to both decorrelate the signals and equalize their power. The whitening matrix, $\mathbf{W}_{\mathrm{SOS}}$, is calculated



[FIG4] **(a) Synthetic trace, (b) synthetic reflectivity, (c) linear prediction error filter, (d) B-ICA, and (e) B-ICA + Wiener filter.**

with the use of second-order statistics [2], such that a new set of mixtures $\mathbf{z}(n) = \mathbf{W}_{SOS}\,\mathbf{x}(n)$ is produced and $E\{\mathbf{z}(n)\mathbf{z}^T(n)\} = \mathbf{I}$, in which $\mathbf{I}$ is an $M \times M$ identity matrix.

STEP 3–DIMENSION REDUCTION

A linear transformation is carried out over $\mathbf{z}(n)$ such that a new set of N_h mixtures, $\tilde{\mathbf{x}}(n) = \mathbf{N}_k^T \mathbf{W}_{sos}^T \mathbf{z}(n)$ is obtained. As $\mathbf{N}_k$ is the matrix that maps the wavelet $\mathbf{h}$ to the kth column of the convolution matrix $\mathbf{A}$, this step reinforces the banded structure of the mixing matrix.

STEP 4–ICA

An ICA algorithm, such the ones described in [2], is applied on $\tilde{\mathbf{x}}(n)$. A separating $N_h \times N_h$ matrix, $\tilde{\mathbf{W}}_{HOS}$, is calculated such that a set of independent estimated sources $\tilde{\boldsymbol{y}}(n) = \tilde{\mathbf{W}}_{HOS}\,\tilde{\boldsymbol{x}}(n)$ is obtained. It is important to notice that $\tilde{\mathbf{x}}(n)$ is not necessarily whitened. Therefore, $\tilde{\mathbf{W}}_{HOS}$ can be calculated by $\tilde{\mathbf{W}}_{HOS} = \tilde{\mathbf{Q}}\tilde{\mathbf{W}}_{SOS}$, where $\tilde{\mathbf{Q}}$ is an orthonormal rotation matrix given by many ICA procedures, and $\tilde{\mathbf{W}}_{SOS}$ is calculated over $\tilde{\boldsymbol{x}}(n)$, similarly to Step 2.

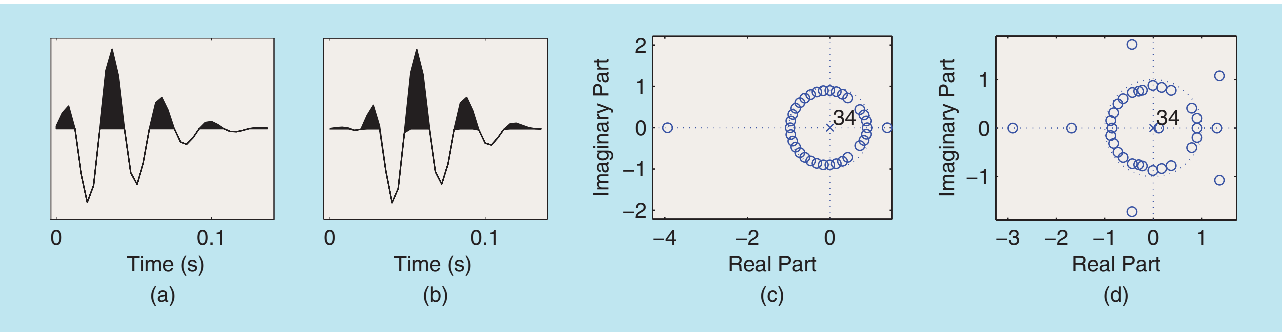
STEP 5–WAVELET AND REFLECTIVITY ESTIMATION

Among the estimated sources, $\tilde{y}_i(n) = \tilde{\mathbf{h}}_i^T \tilde{\mathbf{x}}(n)$, there is one such that $\tilde{y}_i(n)$ is proportional to the original reflectivity, $s(n)$, and that the corresponding row, $\tilde{\mathbf{h}}_i$, of $\tilde{\mathbf{W}}_{HOS}$ approximates the wavelet. Given the knowledge of the matrix $\mathbf{N}_k$ chosen in Step 3, the optimal $(y_*(n), \tilde{\mathbf{h}}_*)$ pair is chosen such that $i_* = \arg\min_i (\min_{c_i} \|\mathbf{x}_k - c_i(\tilde{\mathbf{h}}_i * \tilde{\boldsymbol{y}}_i)\|_2)$, in which $\mathbf{x}_k = [x(k-N)\; x(k-N+1) \ldots x(k-1)]^T$, where $x(n) = 0$ for $n < 0$, and $\tilde{\boldsymbol{y}}_i = [\tilde{y}_i(0)\; \tilde{y}_i(1) \ldots \tilde{y}_i(N-1)]$, as described in [1].

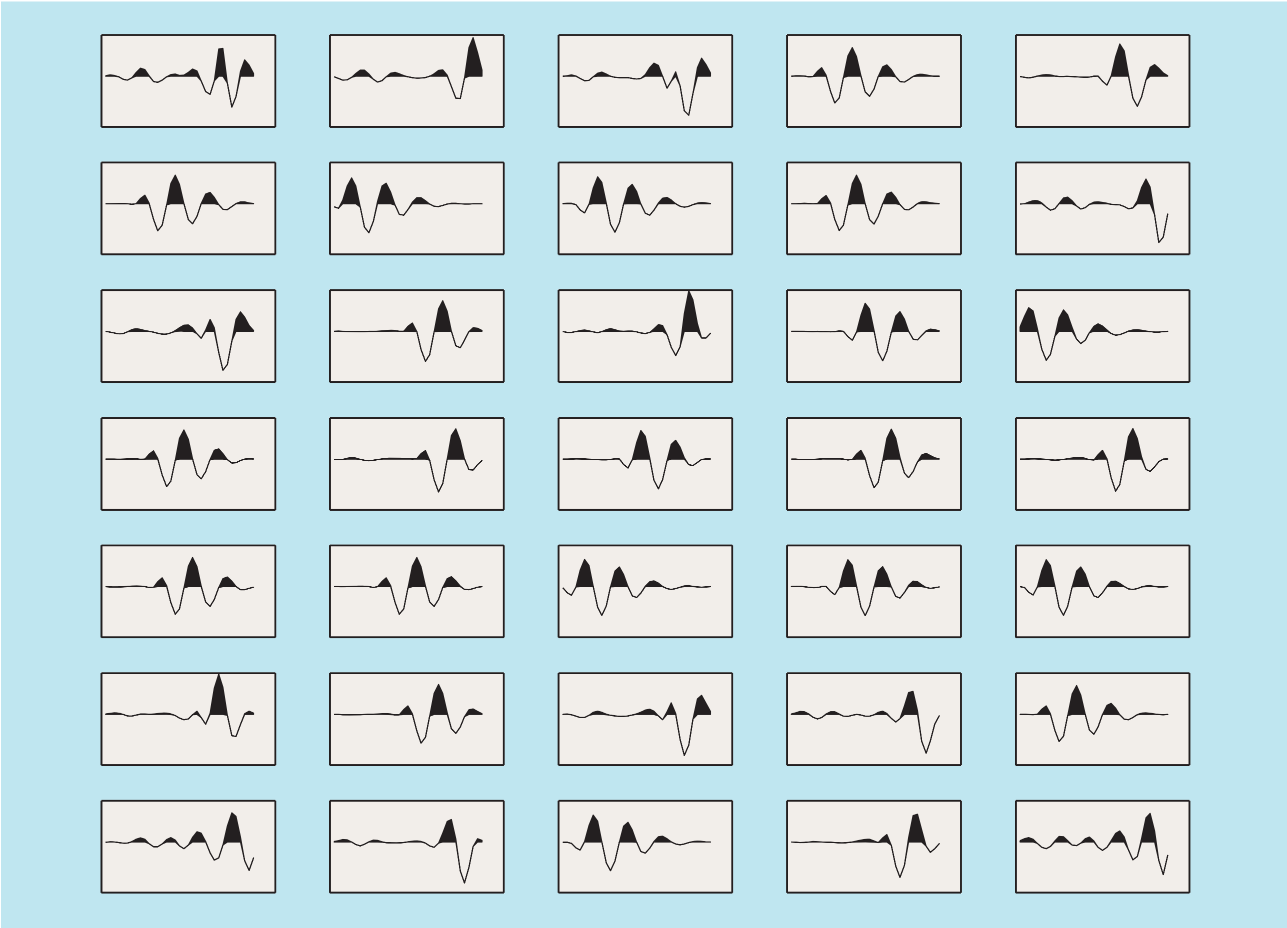
We now present some results of the application of B-ICA to seismic deconvolution. We begin with synthetic data. A noiseless synthetic trace is shown in Figure 4(a) and is generated by convolving a synthetic reflectivity function, which is formed by the random spike train shown in Figure 4(b) with a 35-points Berlage wavelet, shown in Figure 5(a). This wavelet is mixed phase as shown by the plot of its zeros in Figure 5(c). Figure 6 shows a set of 35 candidate wavelets obtained after applying the B-ICA with the Infomax ICA algorithm [12]. The estimated wavelet, chosen from the candidates according to the criterion in Step 5, is shown in Figure 5(b). As shown in Figure 5(d), this wavelet is also mixed phase. Also, the corresponding estimated reflectivity is shown in Figure 4(d).

It can be observed that the wavelet and the reflectivity are delayed by some samples when compared to their actual values. This uniform delay is unavoidable because the mixtures are formed by lagged versions of the convolved signal [1]. As shown in Figure 4(c), there is no delay if a linear prediction error filter is applied to the convolved signal. However, the result obtained with the prediction error filter is more distorted than the one obtained with B-ICA. After eliminating the delay of the latter by comparing it to the original reflectivity and normalizing the involved signals, the MSE is 1.59×10^{-3} for the prediction error filter and 1.78×10^{-4} for the B-ICA. This is due to the fact that a mixed-phase wavelet is used and the prediction error filter is not suited for this case, as discussed previously.

Let us now consider a scenario in which an entire trace gather is deconvolved. B-ICA could be applied for each trace, and the results could be collected in a convolved set of deconvolved traces. This is a standard procedure on the deconvolution literature. However, this is not appropriate in this case, since each deconvolved trace has a different unknown lag, which destroys the



[FIG5] **(a) Original wavelet, (b) wavelet estimated by B-ICA, (c) zero-pole plot for the original wavelet, and (d) zero-pole plot for the estimated wavelet.**



[FIG6] **Set of candidate wavelets.**

lateral coherence on the seismic data. However, since the use of B-ICA on a single trace provides a good wavelet estimate, we propose an additional step for the algorithm.

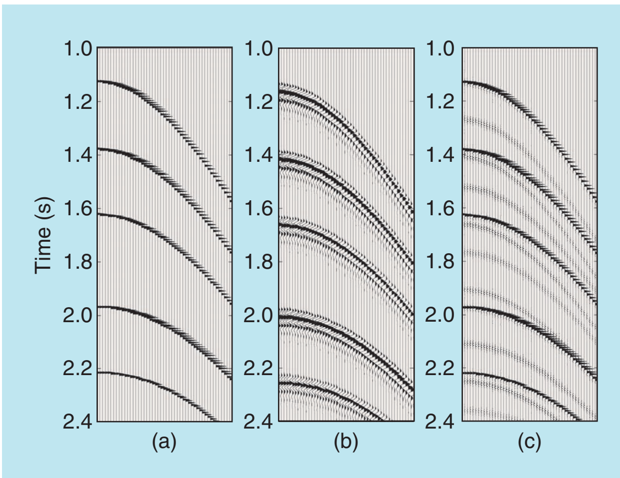
ADDITIONAL PROPOSED STEP–WIENER FILTERING

The estimated wavelet $\tilde{\mathbf{h}}_*$ of a given trace is used to calculate a Wiener filter for a desired spike output with lag greater than or equal to zero. The result is used to filter the traces of the gather to obtain an estimate of the reflectivity.

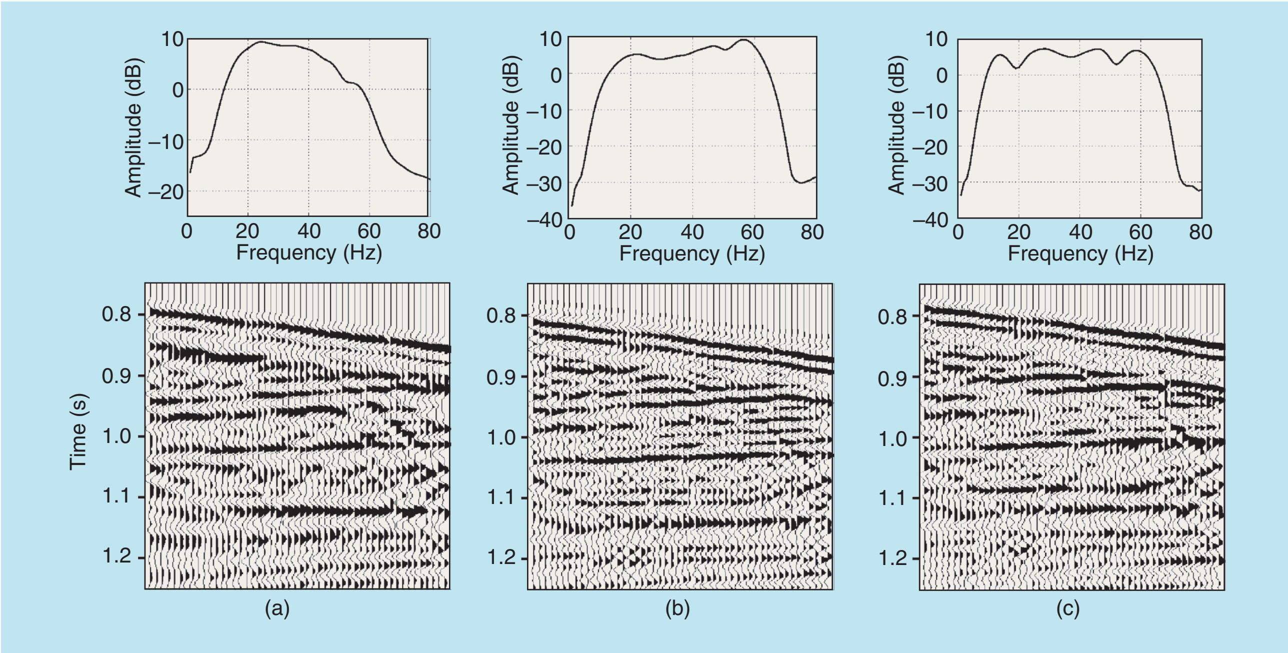
In this additional step, we assume that the difference between wavelets that generate neighboring traces is negligible. To test this approach, a 70-tap Wiener filter, calculated using the additional step, was applied to the trace presented in Figure 4(a). The result is shown in Figure 4(e). The observed delay is associated to the delay on the estimated wavelet and the fact that the wavelet is mixed phase. On the other hand, it was verified that the resulting MSE was 5.06×10^{-5}, better than that obtained by B-ICA alone.

To test the use of B-ICA followed by Wiener filtering in a gather, the synthetic reflectivity function of Figure 7(a) was generated, and then convolved with the Berlage wavelet to simulate a common shot gather, as shown in Figure 7(b). Figure 7(c) shows the deconvolved traces after applying B-ICA combined with the Wiener filter. The result almost perfectly recovers the reflectivity function.

We now discuss the application of the B-ICA with Wiener filtering to the stacked field data presented in Figure 8(a). This data was filtered with a band pass filter from 10 to 60 Hz. The result of B-ICA with Wiener filtering is shown in Figure 8(b), where it can be observed that B-ICA and Wiener filtering enhances the resolution of the reflectors in most parts of the seismic section. For



[FIG7] **(a) Synthetic reflectivity, (b) convolved signals, and (c) after using B-ICA and Wiener filter.**



[FIG8] (a) Original data, (b) data deconvolved with B-ICA and Wiener filter, and (c) data deconvolved with forward prediction error filter.

example, the reflectors located on the interval between 0.9s – 1.1s on the original data have better continuity on the deconvolved data. Note that the deconvolution introduced a slight delay on the traces. As in the case of Figure 4(e), this happens due to delays on the estimated wavelet caused by the method and the fact that the obtained wavelet is mixed phase. For comparison, the result of predictive deconvolution is shown in Figure 8(c). The prediction error filter has a lag of 40 ms. It can be observed that B-ICA and Wiener filtering performs better in the interval given previously. On the amplitude spectrums, we observe that both deconvolution methods succeed in flattening the spectrum over the band of the input data. It can be observed that B-ICA presents a slightly higher content on the higher frequencies.

The results presented in this section illustrate the potential of ICA to improve the quality of seismic deconvolution. However, many issues must be solved before it becomes a practical procedure. Besides the delay ambiguity problem, elements such as noise and wavelet variation in time must be incorporated to new algorithms. This branch of research has given rise to several works, like [28] and [29], and may be considered as an open and stimulating field.

SUMMARY

This article presented an overview of potential applications of unsupervised signal processing in geophysical signals. First, we reviewed the main theorems and theoretical aspects of unsupervised signal processing, showing how these techniques can provide good results in scenarios where methods based on SOS may fail. Then, we describe some potential applications of unsupervised techniques to different problems in seismic processing. In particular, we show the effectiveness of ICA-based seismic deconvolution and propose an improvement by combining it with an additional Wiener filtering for reflectivity estimation. It is worth pointing out that a great number of applications and techniques were not described in this article, such as multiple sensors processing, nonlinear predictive deconvolution, and automatic facies classification, to name a few. In other words, the application of unsupervised methods to seismic problems is a flourishing field still in its infancy. And certainly new unsupervised methods may be proposed in the context of seismic processing. Clearly, this is a topic with great potential for cross-fertilization between the signal processing and geophysics communities.

ACKNOWLEDGMENTS

This work has been partially supported by PETROBRAS, CAPES, and CNPq. We also thank the three anonymous reviewers for their careful reviews and important suggestions that significantly improved the quality of this article.

AUTHORS

André K. Takahata (kazuo@decom.fee.unicamp.br) received his B.S. (2005) and M.Sc. (2009) degrees in electrical engineering from the University of Campinas (UNICAMP), Brazil, where he has been a Ph.D. student in electrical engineering since 2009. His main research interest is in digital signal processing, especially its application to seismic data analysis.

Everton Z. Nadalin (nadalin@dca.fee.unicamp.br) received his B.S. (2004), M.Sc. (2007), and Ph.D. (2011) degrees in electrical engineering from UNICAMP. He has been a postdoctoral researcher at UNICAMP since 2011. His main research interest is in digital signal processing, especially its application to BSS and sparse representation.

Rafael Ferrari (rferrari@decom.fee.unicamp.br) received his B.S. (2001), M.Sc. (2005), and Ph.D. (2011) degrees in electrical engineering from UNICAMP, where he has been a postdoctoral researcher since 2011. His main research interest is in digital signal processing, especially its application to communication systems and to seismic data analysis.

Leonardo Tomazeli Duarte (leonardo.duarte@fca.unicamp.br) received the B.S. and the M.Sc. degrees in electrical engineering from UNICAMP in 2004 and 2006, respectively, and the Ph.D. degree from the Grenoble Institute of Technology (Grenoble INP), France, in 2009. He is currently an assistant professor at the School of Applied Sciences at UNICAMP. His research interests are mainly associated with the theory of unsupervised signal processing and include BSS, independent component analysis, and Bayesian methods. He is a Member of the IEEE.

Ricardo Suyama (ricardo.suyama@ufabc.edu.br) received his B.S. (2001), M.Sc. (2003), and Ph.D. (2007) degrees in electrical engineering from UNICAMP. He has been an assistant professor at the Universidade Federal do ABC (UFABC) since 2010. His research interests are in unsupervised signal processing with a focus on BSS.

Renato R. Lopes (rlopes@decom.fee.unicamp.br) received his B.S. (1995) and M.Sc. (1997) degrees in electrical engineering from UNICAMP. He received an M.Sc. degree in mathematics (2001) and a Ph.D. degree in electrical engineering (2003) from the Georgia Institute of Technology. He has been an assistant professor at the School of Electrical and Computer Engineering at UNICAMP since 2006. His main research interest is in digital signal processing, especially its application to communication systems and to seismic data analysis.

João M.T. Romano (romano@dmo.fee.unicamp.br) received his B.S. (1981) and M.Sc. (1984) degrees in electrical engineering from UNICAMP and the Ph.D. degree (1987) from the University of Paris-XI. In 1988, he joined the School of Electrical and Computer Engineering at UNICAMP, where he is now a professor. He was an invited professor at the University René Descartes, at CNAM-Paris and at ENS-Cachan (France). His research interests are in signal processing and adaptive filtering, with applications like blind equalization and source separation. From 2000 to 2004, he was the president of the Brazilian Communications Society. He is a Senior Member of the IEEE.

Martin Tygel (tygel@ime.unicamp.br) received his M.Sc. (1976) and Ph.D. (1979) degrees in mathematics from Stanford University. He taught at the Federal University of Bahia before joining UNICAMP in 1983, where he is now a professor. He was a Humboldt scholar (1985–1987) in Hannover (Germany), and a visiting professor at the Universities of Karlsruhe (Germany) and Trondheim (Norway). In 2002, he received the Conrad Schlumberger Award of the European Association of Geophysicists and Engineers. His research interest is in computational geophysics with a focus on seismic processing, imaging and inversion, related exploration, and monitoring of hydrocarbon reservoirs.

REFERENCES

[1] S. T. Kaplan and T. J. Ulrych, "Blind deconvolution and ICA with a banded mixing matrix," in *Proc. 4th Int. Symp. Independent Component Analysis and Blind Signal Separation,* Nara, Japan, 2003, pp. 591–596.

[2] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, *Unsupervised Signal Processing.* Boca Raton, FL: CRC Press, 2010.

[3] Ö. Yilmaz, *Seismic Data Analysis*. Tulsa, OK: Society of Exploration Geophysicists, 2001.

[4] J. A. Edgar and M. van der Bann, "How reliable is statistical wavelet estimation?" *Geophysics*, vol. 76, no. 4, pp. V59–V68, July–Aug. 2011.

[5] R. Neelamani, "Deconvolution and optimal filtering in seismology," in *Handbook of Signal Processing in Acoustics,* D. Havelock, S. Kuwano, and M. Vorländer, Eds. New York: Springer-Verlag, 2009, vol. 2, ch. 87, pp. 1571–1583.

[6] E. A. Robinson, "Predictive decomposition of time series with applications to seismic exploration," Ph.D. dissertation, Dept. Geol. Geophys., MIT, Cambridge, 1954.

[7] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York: Wiley, 1949.

[8] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1980.

[9] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications," *IEEE Trans. Automat. Contr.*, vol. 25, no. 3, pp. 385–399, June 1980.

[10] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, no. 2, pp. 312–321, Mar. 1990.

[11] S. Misra and M. D. Sacchi, "Non-minimum phase wavelet estimation by nonlinear optimization of all-pass operators," *Geophys. Prospect.*, vol. 55, no. 2, pp. 223–234, 2007.

[12] P. Comon and C. Jutten, Eds., *Handbook of Blind Source Separation*: *Independent Component Analysis and Applications*. Burlington, MA: Academic Press, 2010.

[13] J. Hérault, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *Proc. GRETSI*, Nice, France, 1985, pp. 1017–1022.

[14] P. Comon, "Independent Component Analysis, a new concept? " *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.

[15] A. B. Weglein, "Multiple attenuation: An overview of recent advances and the road ahead," *The Leading Edge*, vol. 18, no. 1, pp. 40–44, 1999.

[16] A. J. Berkhout and D. J. Verschuur, "Estimation of multiple scattering by iterative inversion—Part I: Theoretical considerations," *Geophysics*, vol. 62, no. 5, pp. 1586–1595, 1997.

[17] A. B. Weglein, F. V. Araújo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, "Inverse scattering series and seismic exploration," *Inverse Problems: Inst. Phys. Publishing*, vol. 19, no. 6, pp. R27–R83, 2003.

[18] A. Guitton and D. J. Verschuur, "Adaptive subtraction of multiples using the L1-norm," *Geophys. Prospect.*, vol. 52, no. 1, pp. 27–38, 2004.

[19] W. Lu, "Adaptive multiple subtraction using independent component analysis," *Geophysics*, vol. 71, no. 5, pp. 179–184, 2006.

[20] W. Lu and L. Liu, "Adaptive multiple subtraction based on constrained independent component analysis," *Geophysics*, vol. 74, no. 1, pp. V1–V7, 2009.

[21] S. T. Kaplan and K. A. Innanen, "Adaptive separation of free surface multiples through independent component analysis," *Geophysics*, vol. 73, no. 3, pp. V29–V36, 2008.

[22] B. Dragoset, E. Verschuur, I. Moore, and R. Bisley, "A perspective on 3D surface-related multiple elimination," *Geophysics*, vol. 75, no. 5, pp. 75A245–75A261, 2010.

[23] J. L. Mari, "Seismic wave separation by SVD and (F-K) combined filters," in *Proc. Extended Abstracts 2006 2nd Int. Symp. Communications, Control and Signal Processing,* Marrakech, Morocco, 2006.

[24] V. D. Vrabie, J. I. Mars, and J.-L. Lacoume, "Modified singular value decomposition by means of independent component analysis," *Signal Process.*, vol. 84, no. 3, pp. 645–652, 2004.

[25] S. L. M. Freire and T. J. Ulrych, "Application of singular value decomposition to vertical seismic profiling," *Geophysics*, vol. 53, no. 6, pp. 778–785, 1988.

[26] M. J. Porsani, M. G. Silva, P. E. M. Melo, and B. Ursin, "SVD filtering applied to ground-roll attenuation," *J. Geophys. Eng.*, vol. 7, no. 3, pp. 284–289, 2010.

[27] M. Bekara and M. Van der Baan, "Local singular value decomposition for signal enhancement of seismic data," *Geophysics*, vol. 72, no. 2, pp. V59–V65, 2007.

[28] A. Al-Qaisi, W. L. Woo, and S. S. Dlay, "Sparsity exploitation of mixing matrix and reflectivity sequence for ICA-blind seismic deconvolution," in *Proc. 3rd Int. Conf. Information and Communication Technologies: From Theory to Applications,* Damascus, Syria, 2008, pp. 1–6.

[29] G. Wei and L. Huaishan, "Frequency domain seismic blind deconvolution based on ICA," in *Proc. 2010 Int. Forum Information Technology and Applications,* Kunming, China, 2010, pp. 440–443.

[SP]