

Blind Submarine Seismic Deconvolution for Long Source Wavelets

Benayad Nsiri, *Associate Member, IEEE*, Thierry Chonavel, Jean-Marc Boucher, *Senior Member, IEEE*, and Hervé Nouzé

Abstract—In seismic deconvolution, blind approaches must be considered in situations where reflectivity sequence, source wavelet signal, and noise power level are unknown. In the presence of long source wavelets, strong interference among the reflectors contributions makes the wavelet estimation and deconvolution more complicated. In this paper, we solve this problem in a two-step approach. First, we estimate a moving average (MA) truncated version of the wavelet by means of a stochastic expectation-maximization (SEM) algorithm. Then, we use Prony's method to improve the wavelet estimation accuracy by fitting an autoregressive moving average (ARMA) model with the initial truncated wavelet. Moreover, a solution to the wavelet initialization problem in the SEM algorithm is also proposed. Simulation and real-data experiment results show the significant improvement brought by this approach.

Index Terms—Bernoulli-Gaussian (BG) process, blind deconvolution, Gibbs sampler, maximum likelihood (ML), maximum posterior mode (MPM), Monte Carlo Markov chains (MCMCs) methods, Prony algorithm, seismic deconvolution, stochastic expectation-maximization (SEM).

I. INTRODUCTION

THE aim of marine seismic exploration is to recover the geological structure of the sea bottom lithography from the analysis of reflected acoustic waves, originating in a seismic wavelet emitted by an acoustic source. The recorded seismic trace can be modeled as the convolution between a source wavelet and a reflectivity sequence, in the presence of additive noise [1]. It is often necessary to apply an inverse transform, called deconvolution, to obtain the reflectivity sequence and

to improve the detection accuracy of layers from seismic data records. If the source wavelet is known, reflectivity can be recovered quite easily from the seismic trace [1]. However, in marine seismic deconvolution, the source wavelet is not exactly known, at least for the following two reasons: acoustic sources such as air guns and sparkers cannot be sufficiently controlled in operational conditions, and the received signal is modified by the ghost phenomenon caused by multiple reflections on the free sea surface and high-frequency absorption during propagation. In practice, a simple way to account for these later phenomena consists in modeling them through distortion of the wavelet [2]. Then, the source wavelet is not exactly known and a blind deconvolution procedure must be applied to recover the unknown source wavelet and the reflectivity sequence [3].

If the wavelet is sufficiently short, as in high-resolution marine seismic exploration [2], where the studied seabed thickness is only about 100 m with a resolution of 1 m, blind deconvolution gives only small improvements compared to raw seismic traces. On the contrary, in some experiments, the emitted wavelet must have a very small frequency bandwidth to penetrate deeper in the soil. As a consequence, we get a long oscillating wavelet. In this case, the seismic image obtained by stacking up consecutive traces is blurred and blind deconvolution is really useful in producing quality images for geologists. In such situations, estimating directly the model parameters as in [4] generally yields a high variance of the wavelet estimator.

So far, many techniques have been proposed to solve the problem of blind deconvolution with respect to different criteria. They are based on the classical hypothesis of reflectivity whiteness. One early technique that copes with the presence of noise is based on Wiener filtering [1], [3]. It supplies a minimum phase wavelet from the second-order statistical content of the trace signal. Unfortunately, this approach is not satisfactory since the source wavelet is usually nonminimum phase. Thus, further statistical information upon the reflectivity sequence must be considered to recover the wavelet phase [5].

The sparse nature of the reflectivity sequence is not well described by a Gaussian model, as assumed when minimum-variance deconvolution [3] is applied. In addition, non-Gaussianity is mandatory for blind deconvolution of scalar signals convolved with a nonminimum phase filter. A main underlying idea in using higher order statistics is highlighting non-Gaussianity of the deconvolved sequence by maximizing contrast criteria [6], [7].

Many related solutions have been tested. In particular, wavelet spectrum envelope can be estimated by using an autoregressive moving average (ARMA) model of the wavelet, that is, a wavelet with rational transfer function. Its phase can

Manuscript received July 10, 2006; revised March 30, 2007; accepted April 16, 2007. This work was supported by IFREMER, Bretagne Region, France.

Associate Editor: R. Chapman.

B. Nsiri is with the Ain Chock Faculty of Sciences, Physics Department, University Hassan II, Casablanca 20101, Morocco, the Groupe Signaux, Communications et Multimédia, Laboratoire de Recherche en Informatique et Télécommunication (GSCM-LRIT), Rabat, Morocco, and the Ecole Nationale Supérieure des Télécommunications (ENST) de Bretagne, Brest 29200, France (e-mail: benayad.nsiri@enst-bretagne.fr).

T. Chonavel is with the Department of Signal and Communications, Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest 29200, France (e-mail: thierry.chonavel@enst-bretagne.fr).

J.-M. Boucher is with the Department of Signal and Communications, Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest 29200, France and the National Scientific Research Center Laboratory, Traitement Algorithmique et Matériel de la Communication, de l'Information et de la Connaissance, UMR 2872, Brest 29238, France (e-mail: jm.boucher@enst-bretagne.fr).

H. Nouzé is with the Marine Geosciences Department, Institut Français de Recherche pour L'Exploitation de la Mer, Bretagne 29280, France (e-mail: hnouze@ifremer.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JOE.2007.899408

be recovered by maximizing a kurtosis criterion [8] thanks to reflections through the unit circle of zeros of the moving average (MA) part, that is, the polynomial part of the wavelet. Alternatively, multispectral techniques can be used, for instance, by recovering the phase of the bispectrum of the trace after it has been whitened at the second order [9].

The main difficulty of these approaches is that they require a large amount of data to achieve reliable estimation. But, in general, seismic trace signals often contain less than 1000 samples, leading to limited performance with the aforementioned approaches.

Another way of applying the whiteness hypothesis for deconvolution consists in minimizing the mutual information rate [10]. It is related to the Kullback–Leibler divergence, which measures the independence degree of the deconvolved data sequence. Unfortunately, this method is not robust with respect to additive noise level, and its performance drops as the signal-to-noise ratio (SNR) decreases.

Instead of considering higher order statistics or general purpose information criteria, a nice way to fully account both for non-Gaussianity and sparseness of the reflectivity sequence consists in considering a prior probability model for the reflectivity sequence. In [11], Mendel introduced a Bernoulli–Gaussian (BG) model of the reflectivity sequence. The Bernoulli variable indicates the presence or absence of a reflector. This model has been extended to a mixture of two Gaussian distributions by associating the high-variance Gaussian distribution to strong reflectors [4], [11]. More precisely, conditional to the value (zero or one) of the Bernoulli variable, the sequence at the point under consideration has a Gaussian distribution with either low or high variance, the later case corresponding to the presence of a reflector. Recovering this reflectivity sequence from the data and the corresponding underlying state (strong or weak reflectivity) enables layers localization in the subsurface. The direct optimization of the corresponding likelihood criterion is unfeasible in practice and many approximate solutions to this problem have been proposed. In particular, the single most likely replacement (SMLR) [11] is a suboptimal solution to this problem. It works in the following two iterative steps: 1) an estimate of the wavelet is calculated from data and an estimated reflectivity sequence and 2) one reflector is updated (the one that achieves best posterior likelihood increase) with this wavelet. The SMLR works in real time, but it may converge to a local optimum. The iterated window maximization algorithm [12] looks very similar to the SMLR, but instead of modifying only one variable at each step, many variables are updated at the same time, leading to improved estimator.

The posterior mean estimator of the reflectivity sequence can be found by using Monte Carlo Markov chains (MCMCs) [13], at the expense of higher computational effort.

As far as unknown parameters such as the wavelet or possibly unknown parameters of the BG model or additive noise variance are concerned, a standard tool for estimating them is the expectation–maximization (EM) algorithm [14], [15] that maximizes the likelihood functional of the completed data model, that is the model that accounts for hidden variables corresponding here to the reflectivity and its underlying Bernoulli process. But, due to its deterministic structure, the EM converges to a local optimum.

This problem can be overcome by using stochastic versions of the EM algorithm, namely, the stochastic EM (SEM) algorithm [16], [17] or the stochastic approximation EM (SAEM) [18]. Finally, the estimation of the BG noise and source wavelet parameters can also be solved by means of MCMC techniques in a fully Bayesian framework. In this case, parameters' prior distributions are introduced [19]–[22]. A comparison of SEM and Bayesian approaches can be found in [4].

After parameter estimation, a further step is carried out to enable reflectivity sequence deconvolution. This is achieved by the maximum posterior mode (MPM) method [23], [24], which also involves MCMC simulation through Gibbs sampling [25]. Let us note that an alternative to the MPM for BG deconvolution is the closely related but suboptimal iterated conditional mode (ICM) algorithm [26].

SEM and Bayesian approaches have been applied mostly with short wavelets. Let us remark that the short wavelet case has been extended to multichannel deconvolution [27]. So far, the long wavelet case has not been considered often, despite its practical interest [28]. In fact, dealing with long unknown wavelets is a difficult task. In such situations, direct ARMA modeling of the wavelet generally appears to fail and estimating the parameters of a long MA model generally yields high variance of the wavelet estimator.

In this paper, we propose a new method to deal with long ARMA wavelet sources in reflectivity blind deconvolution that overcomes the aforementioned problems within the framework of classical blind seismic deconvolution techniques. More precisely, the following two-step approach is proposed. The first step yields a robust estimation of a truncated version of the wavelet by using the maximum-likelihood (ML) approach, via an SEM algorithm. Then, an improved wavelet estimation is achieved by fitting an ARMA model with the initial wavelet truncated version (MA model), using Prony's algorithm [29], [30].

In addition, we propose an efficient method to estimate the order of the initial MA wavelet, based on a tradeoff between bias and variance, and we introduce a new criterion to ensure an accurate wavelet impulse response initialization in the SEM procedure. These are important practical issues for achieving accurate wavelet estimation.

This paper is organized as follows. In Section II, we describe the data model and the standard SEM and MPM procedures for blind reflectivity deconvolution. In Section III, we describe our two-step improved wavelet estimation procedure. In Section IV, we address the problems of the initial MA wavelet order selection and of its initial choice in the SEM procedure. In Section V, we show on simulations and on real-data experiments the significant improvement brought by this approach. Finally, we summarize our work in Section VI.

II. DATA MODEL AND STANDARD PROCEDURES FOR BLIND REFLECTIVITY DECONVOLUTION

The blind submarine deconvolution aims at restoring the sequence of sea bottom reflectivity. The observed signal consists of a noisy version of the reflectivity sequence convolved with the source wavelet

$$y_k = \sum_{i=0}^L h_i r_{k-i} + w_k, \quad k = 1, \dots, N \quad (1)$$

where $\mathbf{h} = (h_k)_{k=0,L}$ is the wavelet finite-impulse response of length L , $\mathbf{r} = (r_k)_{k=1,N}$ is the reflectivity sequence, and $\mathbf{w} = (w_k)_{k=1,N}$ is the observation white noise sequence with variance σ_w^2 . The reflectivity process \mathbf{r} is described by means of a generalized BG process [11], characterized by an underlying state model $\mathbf{q} = (q_k)_{k=0,N}$ with $q_k = 1$ at high reflectivity points and $q_k = 0$ at low reflectivity points. Thus, r_k conditional to q_k is distributed according to a zero-mean Gaussian distribution with variance σ_1^2 if $q_k = 1$ and σ_0^2 if $q_k = 0$

$$r_k \sim \lambda \mathcal{N}(0, \sigma_1^2) + (1 - \lambda) \mathcal{N}(0, \sigma_0^2) \quad (2)$$

with $\lambda = P(q_k = 1)$ being the probability of having a reflector at position k , and $\sigma_1^2 \gg \sigma_0^2$. The BG distribution parameter vector will be denoted by $\boldsymbol{\theta} = (\mathbf{h}, \lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$. For sake of clarity, we recall derivations of the standard SEM and MPM procedures.

A. Maximum Likelihood

To solve the deconvolution problem, one needs to estimate $\boldsymbol{\theta}$. For that, we can use the log-likelihood criterion $\ln p(\mathbf{y}|\hat{\boldsymbol{\theta}})$ [3]. Unfortunately, direct maximization of the log likelihood is unfeasible. However, we are confronted with an incomplete data problem [14] where the incomplete data are given by $\mathbf{z} = (z_k)_{k=0,N}$, with $z_k = (q_k, r_k)$. Then, we are led to work with the log likelihood of the completed data model, defined by $L(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta}) = \ln p(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta})$. Since $\mathbf{z} = (\mathbf{q}, \mathbf{r})$, $p(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta})$ can be written as

$$p(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta}) = p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta})p(\mathbf{r}|\mathbf{q}, \boldsymbol{\theta})p(\mathbf{q}|\boldsymbol{\theta}). \quad (3)$$

Each component of this equation can be expressed easily. As \mathbf{q} is a vector of independent Bernoulli variables, $p(\mathbf{q}|\boldsymbol{\theta}) = \prod_{k=1}^N p(q_k|\boldsymbol{\theta}) = \prod_{k=1}^N \lambda^{q_k} (1 - \lambda)^{1-q_k}$. Furthermore, the variables r_k are independent conditional to the variables q_k . Accordingly

$$\begin{aligned} p(\mathbf{r}|\mathbf{q}, \boldsymbol{\theta}) &= \prod_{k=1}^N p(r_k|q_k, \boldsymbol{\theta}) \\ p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta}) &= \frac{1}{(2\pi\sigma_w^2)^{N/2}} \exp \left[-\frac{\sum_{k=1}^N (y_k - (h * r)_k)^2}{2\sigma_w^2} \right] \quad (4) \\ L(\mathbf{y}, \mathbf{z}|\boldsymbol{\theta}) &= \frac{-(\mathbf{y} - \mathbf{H}\mathbf{r})^T (\mathbf{y} - \mathbf{H}\mathbf{r})}{2\sigma_w^2} \\ &\quad - \frac{\mathbf{S}_1(\mathbf{z})}{2\sigma_1^2} - \frac{\mathbf{S}_0(\mathbf{z})}{2\sigma_0^2} - \left(\frac{N}{2}\right) \ln(\sigma_w^2) \\ &\quad + n_1(\mathbf{q}) \ln \left(\frac{\lambda}{\sigma_1}\right) + n_0(\mathbf{q}) \ln \left(\frac{(1-\lambda)}{\sigma_0}\right) \quad (5) \end{aligned}$$

where $(*)$ represents convolution, with

$$\begin{aligned} S_1(\mathbf{z}) &= \mathbf{r}\mathbf{Q}\mathbf{r}, & S_0(\mathbf{z}) &= \mathbf{r}(I - \mathbf{Q})\mathbf{r}, & \mathbf{Q} &= \text{diag}(\mathbf{q}) \\ n_1(\mathbf{q}) &= \mathbf{q}^T \mathbf{q}, & n_0(\mathbf{q}) &= N - n_1(\mathbf{q}), & \mathbf{H}\mathbf{r} &= \mathbf{R}\mathbf{h} \quad (6) \end{aligned}$$

and \mathbf{H} and \mathbf{R} are the convolution matrices associated with \mathbf{h} and \mathbf{r} , respectively.

B. Gibbs Sampler

When the complete data vector \mathbf{z} is known, with respect to $\boldsymbol{\theta}$, maximizing $p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta})$ is straightforward [see (5)]. In practice, \mathbf{z} is unknown, but it can be simulated using the Gibbs sampler [31], which consists of an iterative random simulation according to the probability distribution $p(z_k|\mathbf{y}, \mathbf{z}_{-k})$ ($k = 1, \dots, N$), where $\mathbf{z}_{-k} = (z_0, \dots, z_{k-1}, z_{k+1}, \dots, z_N)$. The motivation of the Gibbs sampling is that iterative drawing of entries \mathbf{x}_k , $k = 1, \dots, N$, of a vector $\mathbf{x} = (x_1, \dots, x_N)$ from $p(x_k|\mathbf{x}_{-k})$ amounts to simulating the whole distributions of \mathbf{x} [31] in a generally much simpler way. Note that $p(z_k|\mathbf{y}, \mathbf{z}_{-k}) = p(r_k|q_k, \mathbf{y}, \mathbf{z}_{-k})p(q_k|\mathbf{y}, \mathbf{z}_{-k})$, where $p(r_k|q_k, \mathbf{y}, \mathbf{z}_{-k}) \sim \mathcal{N}(m_{q_k}, \sigma_{q_k}^2)$ ($q_k = 0, 1$). The expressions of m_{q_k} , $\sigma_{q_k}^2$, and $p(q_k = 1|\mathbf{y}, \mathbf{z}_{-k}) = 1 - p(q_k = 0|\mathbf{y}, \mathbf{z}_{-k})$ can be found in [33].

For the simulation of \mathbf{z} , I iterations of the Gibbs sampler are implemented as follows [13], [25].

For $i = 1, \dots, I$ and $k = 1, \dots, N$, do the following:

- compute $p(q_k = 1|\mathbf{y}, \mathbf{z}_{-k})$ [see (1)];
- simulate $q_k^{(i)} \sim \mathcal{B}(p(q_k = 1|\mathbf{y}, \mathbf{z}_{-k}))$ ($\mathcal{B}(p)$ denotes the Bernoulli distribution with parameter p);
- simulate $r_k^{(i)} \sim \mathcal{N}(m_{q_k}, \sigma_{q_k}^2)$;
- update $z_k^{(i)}: z_k^{(i)} = (q_k, r_k)$.

C. SEM Algorithm

Since we can now simulate \mathbf{z} with the Gibbs sampler, the SEM algorithm can be used to estimate $\boldsymbol{\theta}$. The SEM algorithm is a stochastic version of the EM where the expectation step is replaced by simulation of one variable \mathbf{z} . Starting from an initial parameter $\boldsymbol{\theta}^{(0)}$, an iteration of SEM consists of the following steps, for iterations $i = 1, \dots, I$:

- SE step: simulate $\mathbf{z}^{(i)} \sim p(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta}^{(i)})$;
- M step: $\hat{\boldsymbol{\theta}}^{(i)} = \arg \max_{\boldsymbol{\theta}} p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta})$.

The SEM does not converge pointwise. It generates a Markov chain whose stationary distribution is more or less concentrated around the ML parameter estimator. A natural parameter estimate from an SEM sequence $(\boldsymbol{\theta}^{(i)})_{i=1, \dots, I}$ is the mean of the iterates value $(I - I_0)^{-1} \sum_{i=I_0+1}^I \boldsymbol{\theta}^{(i)}$ where the first I_0 burn-in iterates have been discarded. An alternative estimate would be the parameter $(\boldsymbol{\theta}^{(i)})$, $i \in \{I_0 + 1, \dots, I\}$, leading to the highest likelihood.

Leaving for clarity the iteration exponent $(\cdot)^{(i)}$, the SEM algorithm can be summarized as follows.

- Initialization: choose $\boldsymbol{\theta}^{(0)}$ and $\mathbf{z}^{(0)}$.
- For $i = 1, \dots, M$
 - E step: simulate $\mathbf{z} \sim p(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})$;
 - M step: calculate the parameters

$$\begin{aligned} \hat{\mathbf{h}} &= (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{y}, & \hat{\sigma}_w^2 &= N^{-1} \|\mathbf{y} - \mathbf{R}\hat{\mathbf{h}}\|_2^2 \\ \hat{\lambda} &= N^{-1} \mathbf{q}^T \mathbf{q}, & \hat{\sigma}_0^2 &= \frac{\mathbf{r}^T (I - \mathbf{Q}) \mathbf{r}}{N - \mathbf{q}^T \mathbf{q}}, & \hat{\sigma}_1^2 &= \frac{\mathbf{r}^T \mathbf{Q} \mathbf{r}}{\mathbf{q}^T \mathbf{q}} \quad (7) \end{aligned}$$

where $\|u\|_2 = u^T u$.

D. MPM Deconvolution

At the end of the SEM procedure, $\boldsymbol{\theta}$ is well estimated while reflectivity \mathbf{r} is not. The fact that the wavelet is accurately es-

timated while reflectivity parameters are not can be understood easily: The wavelet estimation benefits from multiple reflectors diversity, and since it has a smooth shape, its mean square or ML estimate is little affected by reflector position small shifts or close reflectors overlapping. On the contrary, the parameters of a single reflector do not benefit from averaging effects in their estimation. Similarly, parameters $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ represent noise and reflectivity statistics. Thus, they also all benefit more from statistical averaging since the record is long. Therefore, an additional deconvolution procedure based on the MPM algorithm is performed for recovering \mathbf{r} from knowledge of $\boldsymbol{\theta}$ [33].

As the $\boldsymbol{\theta}$ parameters, and, in particular, wavelet \mathbf{h} , are completely estimated by the SEM procedure, the deconvolution is no longer blind. But we still need to estimate the hidden variables in vector \mathbf{z} from the observation.

To estimate \mathbf{z} , let us remark that the posterior log likelihood of \mathbf{z} conditional to \mathbf{y} and $\boldsymbol{\theta}$ is

$$L(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta}) = -\frac{(\mathbf{y} - \mathbf{H}\mathbf{r})^T(\mathbf{y} - \mathbf{H}\mathbf{r})}{2\sigma_w^2} - \frac{\mathbf{r}^T\mathbf{Q}\mathbf{r}}{2\sigma_1^2} - \frac{\mathbf{r}^T(\mathbf{I} - \mathbf{Q})\mathbf{r}}{2\sigma_0^2} + f(q, \lambda, \sigma_1, \sigma_2) \quad (8)$$

where the function $f(q, \lambda, \sigma_1, \sigma_2)$ does not depend on \mathbf{r} . The principal separation [11] allows us to address the maximization of $L(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})$ in the following two steps:

- detection: $\hat{\mathbf{q}} = \arg \max_{\mathbf{q}} p(\mathbf{q}|\mathbf{y})$;
- estimation: $\hat{\mathbf{r}} = \arg \max_{\mathbf{r}} p(\mathbf{r}|\mathbf{y}, \hat{\mathbf{q}})$.

$\ln p(\mathbf{r}|\mathbf{y}, \hat{\mathbf{q}})$ is quadratic in \mathbf{r} , for fixed \mathbf{q} , leading to

$$\hat{\mathbf{r}}(\mathbf{q}) = \left[\mathbf{H}^T\mathbf{H} + \frac{\sigma_w^2}{\sigma_1^2} \left(\mathbf{Q} + \frac{\sigma_1^2}{\sigma_0^2}(\mathbf{I} - \mathbf{Q}) \right) \right]^{-1} \mathbf{H}\mathbf{y}. \quad (9)$$

The detection step is not so easy, because the vector \mathbf{q} has 2^N (N is the signal length) configurations. Testing all of them is not possible. Instead, the posterior mode of each individual q_k can be searched applying for the MPM [19], [24].

The MPM algorithm simply generates samples of \mathbf{z} drawn from $p(\mathbf{z}|\mathbf{y}, \boldsymbol{\theta})$ just as for the Gibbs sampler used in the SEM algorithm and described at the end of Section II-B. Then, from these samples $(q_k^{(i)}, r_k^{(i)})_{i=1, I}$, a decision is made upon each q_k , and r_k is estimated conditionally to q_k as follows.

- Detect \hat{q}_k

$$\hat{q}_k = \begin{cases} \frac{1}{I - I_0} \sum_{i=I_0+1}^I q_k^{(i)}, & \text{if } q_k > s \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

- Estimate \hat{r}_k

$$\hat{r}_k = \begin{cases} \frac{\sum_{i=I_0+1}^I q_k^{(i)} r_k^{(i)}}{\sum_{i=I_0+1}^I q_k^{(i)}}, & \text{if } \hat{q}_k = 1 \\ \frac{\sum_{i=I_0+1}^I (1 - q_k^{(i)}) r_k^{(i)}}{\sum_{i=I_0+1}^I (1 - q_k^{(i)})}, & \text{otherwise.} \end{cases} \quad (11)$$

For $s = 0.5$, \hat{q}_k is the posterior mode of q_k . Other values of s correspond to other Bayesian cost functions [34].

III. IMPROVED WAVELET ESTIMATION

In some seismic experiments, the wavelet impulse response \mathbf{h} is quite long. In such cases, the mean square error (MSE) of the estimator is quite large. In particular, the last coefficients of \mathbf{h} , which have small values, are poorly estimated. For this reason, searching for a vector \mathbf{h} with reduced length generally enables a good compromise between bias and variance properties of the estimator. However, performing the MPM deconvolution with a truncated wavelet will degrade detection and estimation performance for the reflectivity sequence.

The true wavelet can be described by an MA model of length L_0 [MA(L_0) model] with transfer function $h(z) = \sum_{l=0, L_0} h_l z^{-l}$. To avoid using a model with many parameters when L_0 is large, one could instead use an ARMA(p, q) model, that is a wavelet with transfer function in the form $(\sum_{l=0, q} b_l z^{-l}) / (1 + \sum_{k=1, p} a_k z^{-k})$. In general, one can take p and q much smaller than L_0 , in particular, when dealing with long oscillating wavelets where oscillations are well modeled by poles of the ARMA model. However, as discussed in the introduction, direct use of an ARMA model in the SEM procedure generally fails. Hence, we have developed the approach described hereafter.

First, we use the SEM procedure described in Section II with an MA(L) wavelet model, with L smaller than the true wavelet length L_0 [35]. Then, from this truncated estimate of the wavelet, with transfer function $h_{\text{MA}}(z) = \sum_{k=0, L} h_{\text{MA}, k} z^{-k}$, we look for an ARMA(p, q) wavelet with transfer function

$$h_{\text{ARMA}}(z) = \frac{\sum_{l=0, q} b_l z^{-l}}{1 + \sum_{k=1, p} a_k z^{-k}} = \sum_{k=0, \infty} h_{\text{ARMA}, k} z^{-k}. \quad (12)$$

Then, the idea is to choose coefficients $(a_k)_{k=1, p}$ and $(b_l)_{l=0, q}$ such that $h_{\text{ARMA}, k} \approx h_{\text{MA}, k}$ for $k = 0, \dots, L$. This can be achieved in a rather simple way by minimizing

$$J(\mathbf{a}, \mathbf{b}) = \sum_{l=0}^{L_{\max}} \left| \sum_{k=0}^p a_k h_{l-k} - b_l \right|^2 \quad (13)$$

where $b_l = 0$ for $l > q$. This method is known as Prony's method [29], [30]. Straightforward calculations show that the optimum of $J(\mathbf{a}, \mathbf{b})$ is $(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \arg \min_{\mathbf{a}, \mathbf{b}} J(\mathbf{a}, \mathbf{b})$, where

$$\hat{\mathbf{a}} = -(\mathbf{H}_1^H \mathbf{H}_1)^{-1} \mathbf{H}_1^H \mathbf{V}, \quad \hat{\mathbf{b}} = \mathbf{H}_0 [1 \quad \hat{\mathbf{a}}]^T \quad (14)$$

and

$$\mathbf{H}_0 = \begin{bmatrix} h_0 & 0 & \dots & 0 \\ h_1 & h_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ h_q & h_{q-1} & \dots & h_{q-p} \end{bmatrix}$$

$$\mathbf{H}_1 = \begin{bmatrix} h_q & h_{q-1} & \dots & h_{q-p} \\ h_{q-1} & h_{q-2} & \dots & \vdots \\ \vdots & \dots & \ddots & \vdots \\ h_{L_{\max}} & h_{L_{\max}-1} & \dots & h_{L_{\max}-p} \end{bmatrix}$$

$$\mathbf{V} = [h_{q+1}, \dots, h_{L_{\max}}]^T, \quad L_{\max} - q > p. \quad (15)$$

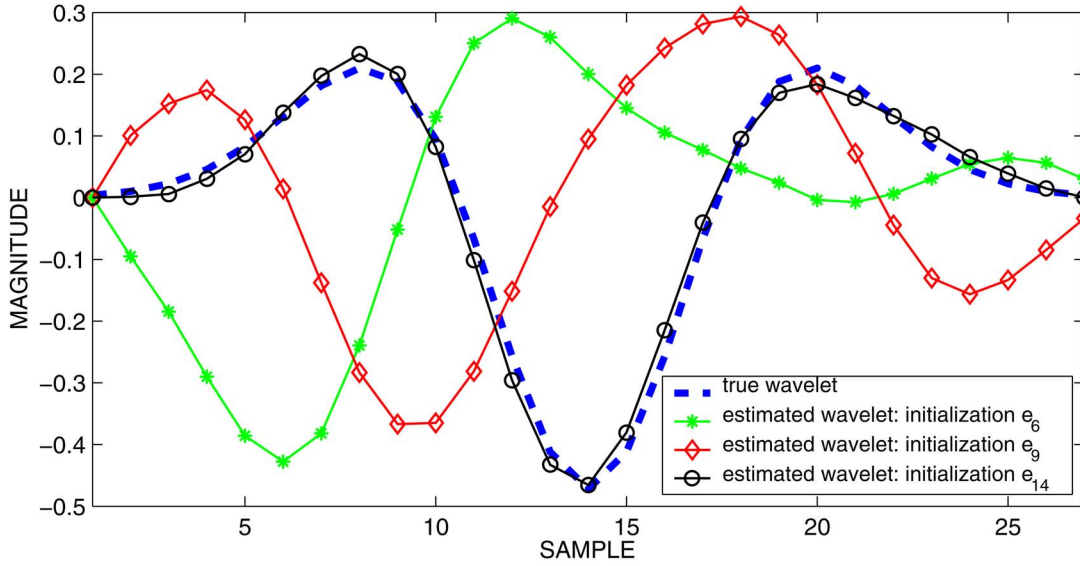


Fig. 1. Estimated wavelet for different initializations.

To estimate the respective orders p and q of the AR and MA parts, in [8], [36], we have proposed to use a Kurtosis maximization criterion. Although this criterion is efficient in many cases, it may not be very satisfactory in certain situations. Alternatively, we propose to estimate p and q by use of an exhaustive search. For this aim, we use a nested loop with p and q as indices. The optimum for (p, q) is obtained when the criterion $J(\hat{\mathbf{a}}, \hat{\mathbf{b}})$ is minimum. We will call SEM+P (P for Prony) the parameter estimation procedure that we have just proposed that consists in applying the SEM algorithm followed by an ARMA extension of the initial MA wavelet estimate.

IV. INITIAL MA MODEL ORDER SELECTION AND MAXIMUM POSITION SELECTION

We now consider two important practical issues for the SEM+P algorithm working properly. Note that these considerations can be useful for many other algorithms.

A. Initial MA Model Order Selection

In this section, we explain how the length L of the initial MA(L) wavelet can be selected. Our approach is based on an MSE criterion. The idea is to look for the order L , for which the estimated wavelet is as close as possible to the true wavelet \mathbf{h} . Unfortunately, the true wavelet itself is not known in practice. As the wavelet length is in some (possibly large) interval $[l_1, l_2]$, we consider all the wavelet estimates $(\hat{\mathbf{h}}_l)_{l=l_1, l_2}$ supplied by the SEM algorithm for these choices of l . The MSE criterion would lead to choosing $L = \arg \min_{l \in [l_1, l_2]} \|\hat{\mathbf{h}}_l - \mathbf{h}\|$. Since \mathbf{h} is unknown, we estimated it as

$$\bar{\mathbf{h}} = \frac{1}{(l_2 - l_1 - 1)} \sum_{l=l_1, l_2} \hat{\mathbf{h}}_l. \quad (16)$$

If we compare the functions MSE: $l \rightarrow \|\hat{\mathbf{h}}_l - \mathbf{h}\|^2$ and $\widehat{\text{MSE}}: l \rightarrow \|\hat{\mathbf{h}}_l - \bar{\mathbf{h}}\|^2$, we have checked on simulations that

they roughly behave in the same way (see Section V). Thus, a good choice for L is

$$L = \arg \min_{l \in [l_1, l_2]} \|\hat{\mathbf{h}}_l - \bar{\mathbf{h}}\|. \quad (17)$$

B. Maximum Position Selection

It is well known that the nonminimum phase structure of the wavelet \mathbf{h} makes its estimation complicated. One may think that, due to the stochastic approximation of the expectation in the first step of the SEM algorithm, it should not be sensitive to the initialization parameters, but this is not the case. This problem has already been pointed out in [37], where a simulated annealing version of the SAEM algorithm is proposed to solve the deconvolution problem.

As mentioned previously, the wavelet estimation is not robust with respect to the initialization. If we initialize \mathbf{h} with the vector \mathbf{e}_i , $i = 1, \dots, k_1$, that has all entries equal to zero except the i th one which is equal to 1, we observe that the SEM algorithm yields an estimate of \mathbf{h} with its maximum at entry i . As an example, Fig. 1 shows an estimated wavelet obtained for distinct initializations in the case of Ricker's wavelet.

To understand this phenomenon, let us remark that the knowledge of the data vector \mathbf{y} , given by $\mathbf{y} = \mathbf{h} * \mathbf{r} + \mathbf{w}$, permits us to recover \mathbf{h} up to a time translation and an amplitude shift but not \mathbf{h} , since we also have

$$\mathbf{y} = (a\mathbf{h}_\tau) * (a^{-1}\mathbf{r}_{-\tau}) + \mathbf{w} \quad (18)$$

for any amplitude a and any delay τ , where \mathbf{h}_τ and $\mathbf{r}_{-\tau}$ denote, respectively, the function \mathbf{h} delayed by τ and \mathbf{r} delayed by $-\tau$. This explains why the SEM algorithm does not look for a wavelet with its maximum, apart from the position i of the nonzero component of the initial guess \mathbf{e}_i . This interpretation is impressively confirmed by simulations: For more than 99% of the simulations, the estimated wavelet has its maximum at position i when it is initialized with \mathbf{e}_i .

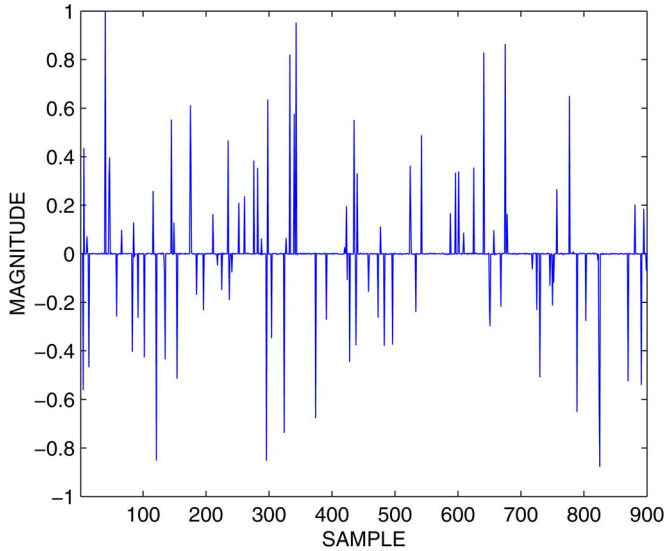


Fig. 2. Gaussian mixture reflectivity sequence.

To end this discussion, let us remark that one should not conclude from relation (18) that any initialization would be satisfactory due to the delay compensation between \mathbf{h}_τ and $\mathbf{r}_{-\tau}$. Indeed, this would be true only if the searched wavelet support were infinite. Here, bad initialization may cause poor estimation of the wavelet, due to the limited wavelet support. We have seen in Section II that limiting the support is necessary to achieve a good parameter estimation performance.

We propose a deterministic procedure for initializing the $MA(L)$ wavelet estimate. First, we have performed the deconvolution for each initialization e_i ($i = 1, \dots, k_1$) of \mathbf{h} . Then, we note $C_i = 1$ if the estimator obtained from initialization e_i has a positive derivative at the origin, and $C_i = -1$ if it is negative. It can be checked on simulations that $\mathbf{C} = (C_i)_{i=1, k_1}$ changes sign for entries i corresponding to local optima of the true wavelet. A justification of this phenomenon is presented in the Appendix. Simulations on several examples show the very good practical behavior of this technique (see Section V). The retained solution for the maximum position is chosen among the entries of \mathbf{C} for which its sign changes, by selecting the one for which the kurtosis of the estimated reflectivity $\hat{\mathbf{r}}$ is maximum.

V. RESULTS

In this section, we evaluate performance of the proposed methods for the wavelet estimation and for the SEM algorithm initialization, through simulations and real data experiments.

A. Simulation Results

We will define the SNR as

$$\text{SNR} = 10 \log \left(\frac{\lambda \sigma_1^2 E_h}{\sigma_w^2} \right) \quad (19)$$

where E_h is the source wavelet energy and L_0 is the wavelet length: $E_h = L_0^{-1} \sum_{i=1}^{L_0} h_i^2$.

1) *Reflectivity and Noise Parameters*: The reflectivity sequence samples are distributed according to a Gaussian mixture

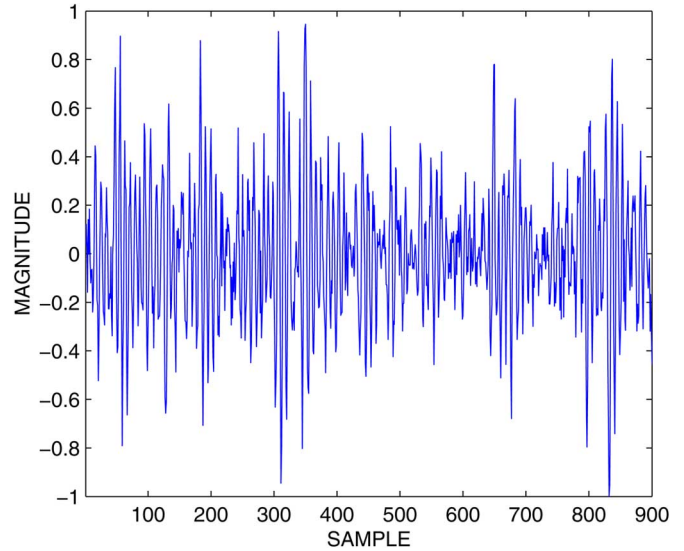


Fig. 3. Noisy seismic data (SNR = 10 dB).

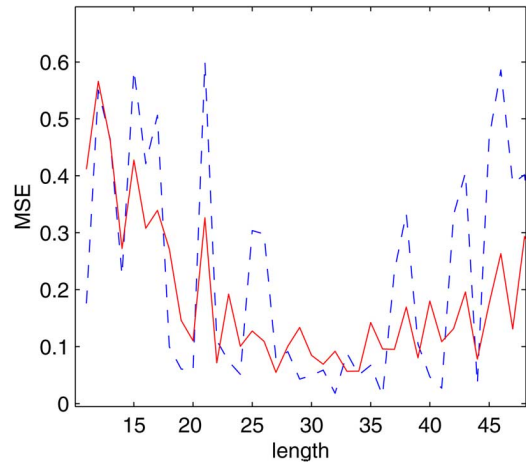
Fig. 4. $\widehat{\text{MSE}}$: “—”; MSE: “- - -”.

TABLE I
SEM DECONVOLUTION: ESTIMATED PARAMETERS FOR SNR = 13 dB

	λ	σ_0^2	σ_1^2	σ_w^2
True	0.1	2.5×10^{-4}	0.1	2.5×10^{-5}
Estimated (long wavelet)	0.23	5×10^{-4}	0.095	2×10^{-4}
Estimated (truncated wavelet)	0.13	1.3×10^{-4}	0.1	3×10^{-4}

with $\lambda = 0.1$, $\sigma_0^2 = 2.5 \times 10^{-4}$, $\sigma_1^2 = 0.1$, and $\sigma_w^2 = 10^{-4}$. A typical reflectivity sequence is presented in Fig. 2. The corresponding seismic trace is presented in Fig. 3. This synthetic trace is generated by using the long convolution wavelet shown in Fig. 5.

To show that it is better to use a truncated wavelet estimator instead of a full length one, we compare the quality of the estimators for parameters $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ in both cases. In addition, we check this for SNR = 8 dB and SNR = 13 dB, which roughly corresponds to the range of SNR values in practical experiments. Results are presented in Tables I and II. It clearly appears that better parameter estimation for $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ is achieved with the truncated wavelet. One main difference be-

TABLE II
SEM DECONVOLUTION: ESTIMATED PARAMETERS FOR SNR = 8 dB

	λ	σ_0^2	σ_1^2	σ_w^2
True	0.1	2.5×10^{-4}	0.1	10^{-4}
Estimated (long wavelet)	0.21	4×10^{-4}	0.08	2.3×10^{-4}
Estimated (truncated wavelet)	0.12	1.2×10^{-4}	0.11	1.5×10^{-4}

TABLE III
KURTOSIS OF $\hat{\mathbf{r}}$ FOR THE SEM PROCEDURE INITIALIZED FROM THE FIRST LOCAL MAXIMA OF $\hat{\mathbf{h}}$ DETECTED FROM \mathbf{C} (WAVELET OF FIG. 5)

Maximum position	4	9	12	16
Kurtosis	0.0079	0.0351	0.0217	0.0070

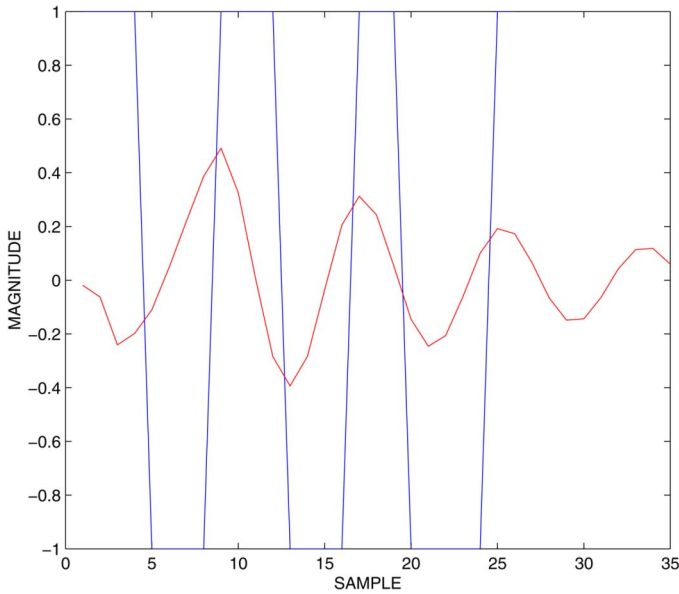


Fig. 5. Estimated maxima positions for a long wavelet.

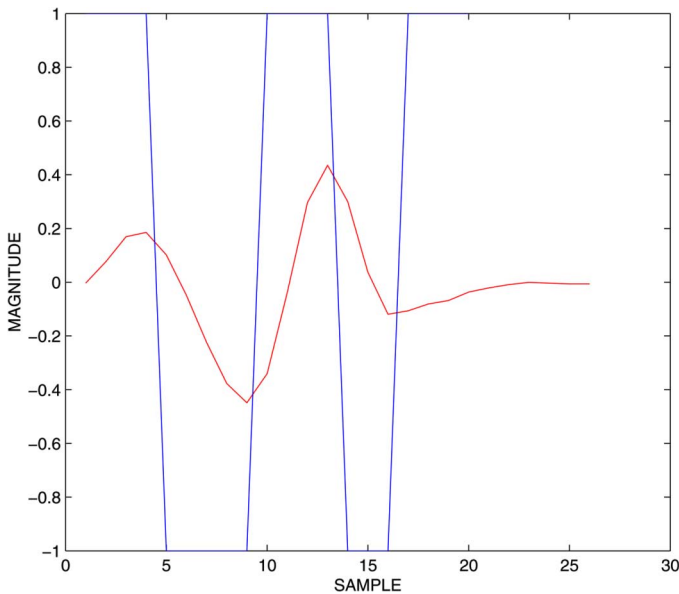


Fig. 6. Estimated maxima positions for Marmousi’s wavelet.

tween using truncated and full size wavelet is that in the second case there are more degrees of freedom; in particular, the truncated wavelet can be seen as a full size one with zeroed last

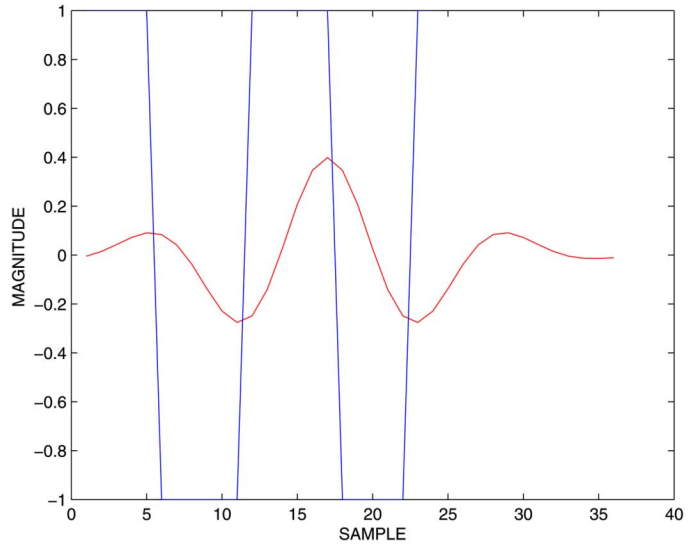


Fig. 7. Estimated maxima positions for Morlet’s wavelet.

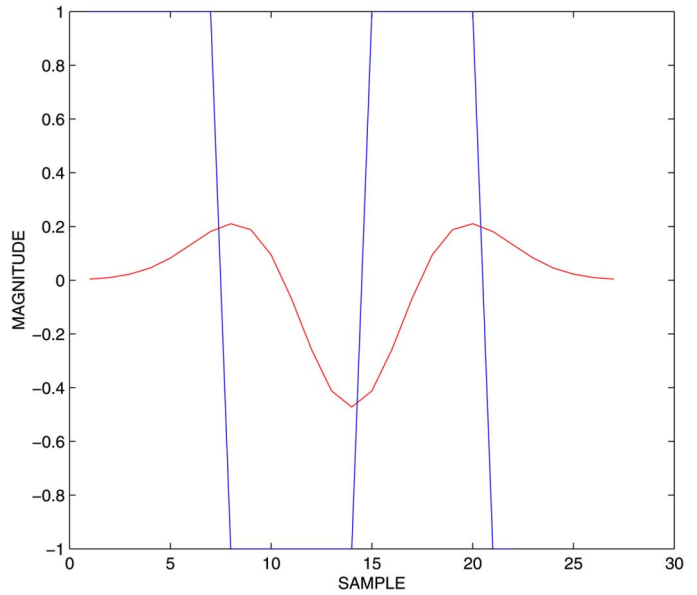


Fig. 8. Estimated maxima positions for Ricker’s wavelet.

coefficients. Thus, the long size wavelet can be adjusted to fit more potential reflectors in the seismic trace. This explains why λ is overestimated when the long wavelet is used.

2) *Wavelet Estimation:* One important issue discussed in Section IV-A was the choice of the length of the truncated wavelet. For data of Fig. 4, we computed the empirical function $\widehat{\text{MSE}}(l \rightarrow \|\hat{\mathbf{h}}_{(l)} - \mathbf{h}\|^2)$ and the true MSE $(l \rightarrow \|\hat{\mathbf{h}}_{(l)} - \hat{\mathbf{h}}\|^2)$. The shape of both criteria is roughly the same, which makes it meaningful choosing the position of the minimum of $\widehat{\text{MSE}}$ to define the length L of the truncated wavelet.

The problem of wavelet initialization led us to define the function \mathbf{C} (see Section IV-B). As mentioned in Section IV-B, initializations with wavelets with only one nonzero entry at a transition of function \mathbf{C} are retained. The kurtosis of the deconvolved sequence with the corresponding estimated wavelets can be seen

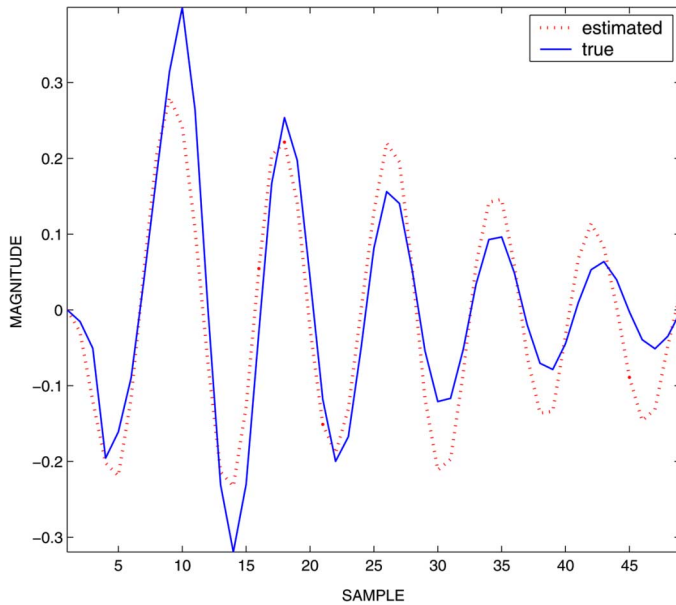


Fig. 9. Estimated long wavelet by a standard SEM procedure.

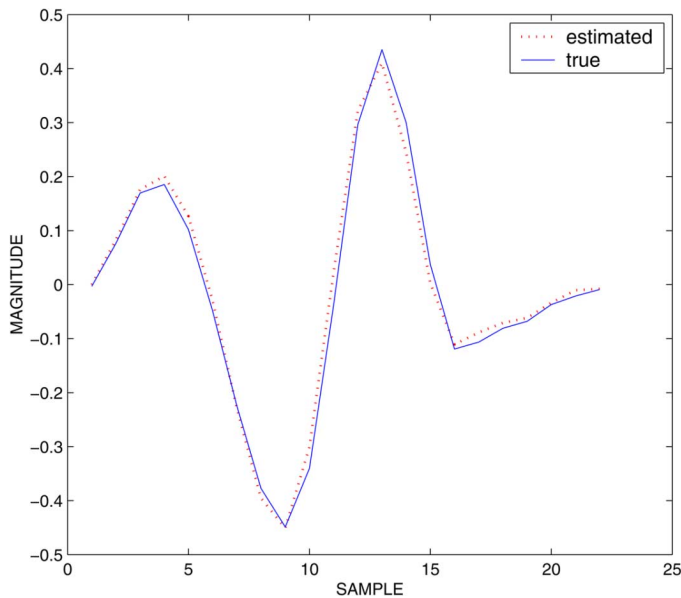


Fig. 10. Estimated Marmousi's wavelet.

in Table III. The maximum of the kurtosis is achieved for initialization at position 9 (e_9) which corresponds to the wavelet maximum position.

We tested the standard SEM procedure [4] with several wavelets. They are represented in Figs. 5–8 together with the corresponding C functions. Vertical lines represent the transitions of C between values -1 and $+1$. We check that in all cases there is a good match between these transitions and the local optima of the wavelet. Figs. 5–8 show the corresponding deconvolution wavelets. We can see that the short wavelets (Figs. 10–12) are well estimated, which justifies using the

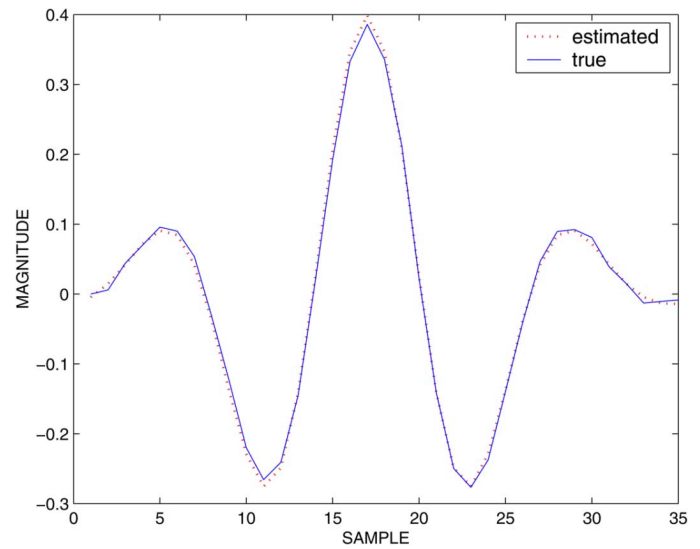


Fig. 11. Estimated Morlet's wavelet.

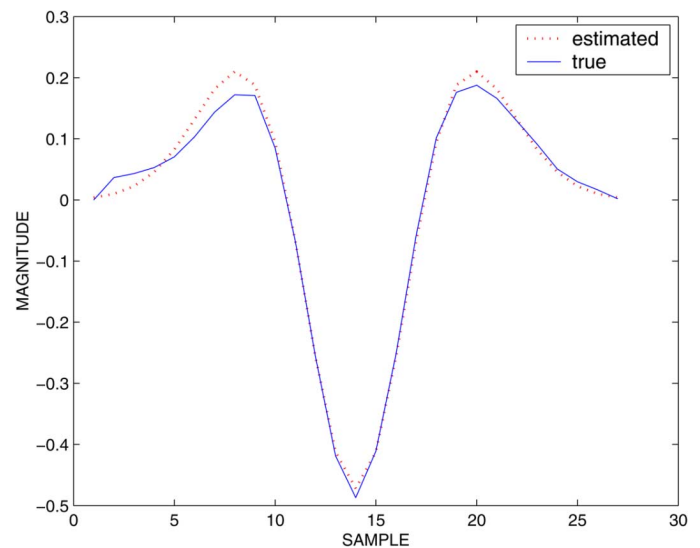


Fig. 12. Estimated Ricker's wavelet.

standard SEM procedure in such cases. On the contrary, the last part of the long wavelet is poorly estimated.

Now, let us show that the proposed SEM + P wavelet estimation procedure works better for long wavelets than the standard SEM. Indeed, we can observe in Fig. 13 that the two-step estimation (truncated wavelet estimation followed by Prony extension) achieves almost perfect long wavelet recovery. Figs. 9 and 13 show the significant gain of the SEM + P against the SEM. Furthermore, to study improvements brought by the deconvolution method, we consider the following performance indices:

$$\text{MSE}_1 = \|\mathbf{h} - \hat{\mathbf{h}}\|^2 \quad \text{and} \quad \text{MSE}_2 = \|\mathbf{r} * \mathbf{h} - \hat{\mathbf{r}} * \hat{\mathbf{h}}\|^2. \quad (20)$$

MSE_1 represents the error energy of the estimated wavelet, while MSE_2 represents the error of noiseless data reconstruc-

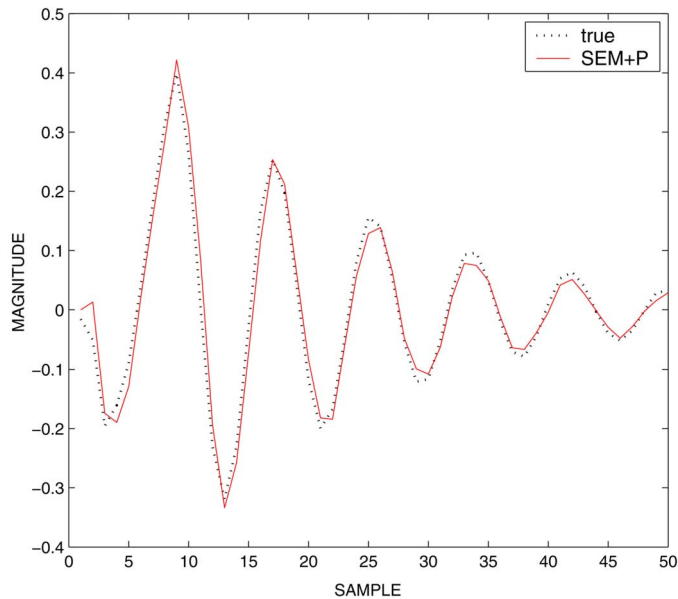


Fig. 13. Estimated wavelet by SEM + P.

TABLE IV
MSE₁ AND MSE₂ PERFORMANCE INDICES FOR SEM AND SEM + P
METHODS (WAVELET OF FIG. 5)

	MSE ₁ (dB)		MSE ₂ (dB)	
SNR (dB)	10	13	10	13
SEM	-0.2	-8.8	17	6
SEM+P	-13.2	-16.6	5	3

tion, that is, the convolution of wavelet and reflectivity. Indeed, direct comparison of \mathbf{r} and $\hat{\mathbf{r}}$ is not suitable since \mathbf{r} is a sparse spike train sequence and thus very small reflector location offsets may lead to high MSE $\|\mathbf{r} - \hat{\mathbf{r}}\|^2$.

Table IV confirms the improvement of the performance indices MSE₁ and MSE₂ [see (20)], with the SEM + P procedure compared to the standard SEM when applied to a long wavelet.

3) *Reflectivity Estimation*: Once the wavelet \mathbf{h} and model distribution parameters $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ are estimated, we are able to recover the reflectivity sequence by means of the MPM algorithm. We are going to see that the quality of the estimator of $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ strongly influences the estimation performance of \mathbf{r} .

The reflectivity sequence estimated by the SEM + MPM and the (SEM + P) + MPM algorithms are given, respectively, in Figs. 14 and 15. Note that after applying the algorithms, it may occur that instead of one reflector two neighboring reflectors are found, either contiguous or separated by only one sample. Then, a postprocessing procedure can be applied to fuse these reflectors at their gravity center with cumulated amplitudes [4]. With the SEM + MPM approach (Fig. 14), there are often many neighboring strong reflectors, which prevents applying this technique, while it has been possible to apply it with the (SEM + P) + MPM (Fig. 15) because distinct reflectors had sufficiently separated contributions, leading to much better results than the standard SEM + MPM procedure (Fig. 14).

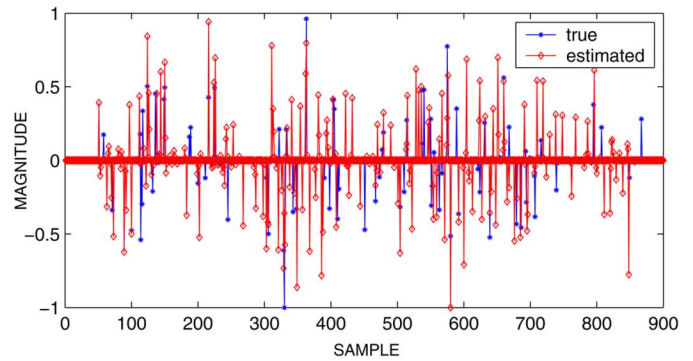


Fig. 14. Estimated reflectivity sequence by SEM + MPM (SNR = 8 dB).

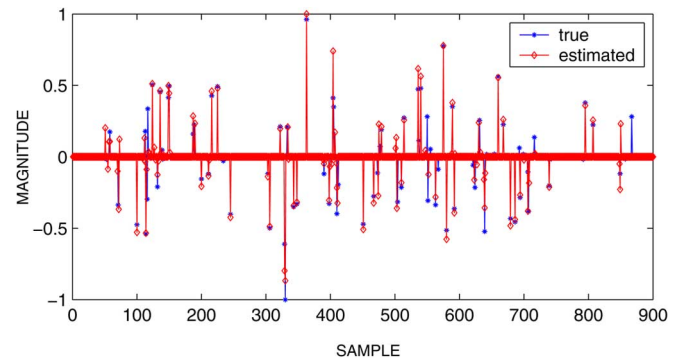


Fig. 15. Estimated reflectivity sequence by (SEM + P) + MPM (SNR = 8 dB).

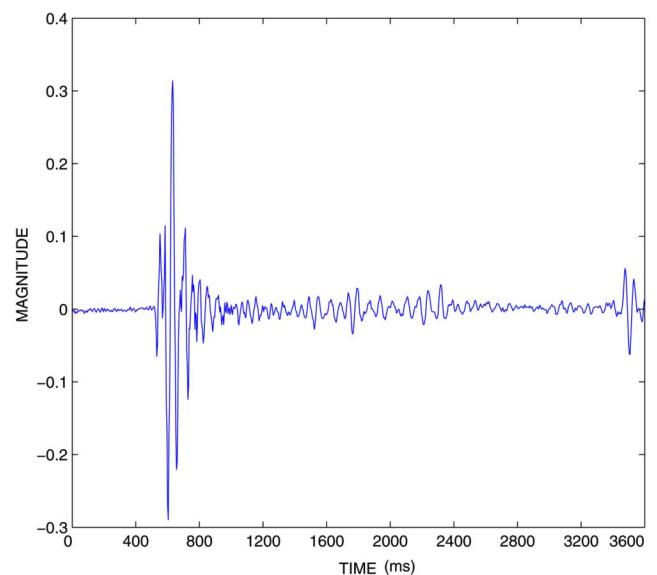


Fig. 16. Data record (trace 602).

B. Real-Data Experiments

Here, we consider real data acquired by IFREMER,¹ consisting of seismic traces obtained from an innovative manner for synchronizing a cluster of air guns to achieve deeper penetration. The fact that the sources are not synchronized in their firing

¹French Research Institute for Exploitation of the Sea—Institut Français de Recherche pour L'Exploitation de la Mer (IFREMER), Bretagne, France.

TABLE V
ESTIMATED PARAMETERS FOR SHOT 1 AND 40, AND MEAN OF 50 SHOTS

Estimated parameter	λ	σ_w^2	σ_1^2	σ_0^2
Shot 1	0.050	3.0×10^{-6}	0.0098	4.5×10^{-5}
Shot 40	0.068	1.3×10^{-6}	0.0089	6.1×10^{-5}
Mean for 50 shots	0.066	1.5×10^{-6}	0.0090	6.0×10^{-5}

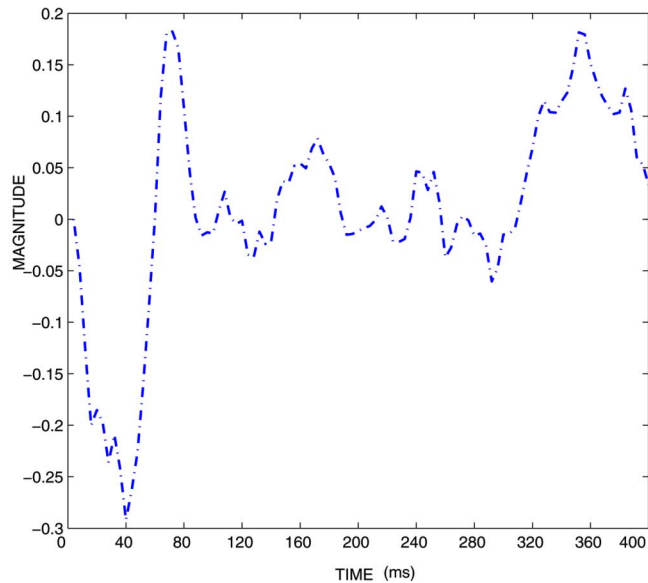


Fig. 17. Full length estimated wavelet.

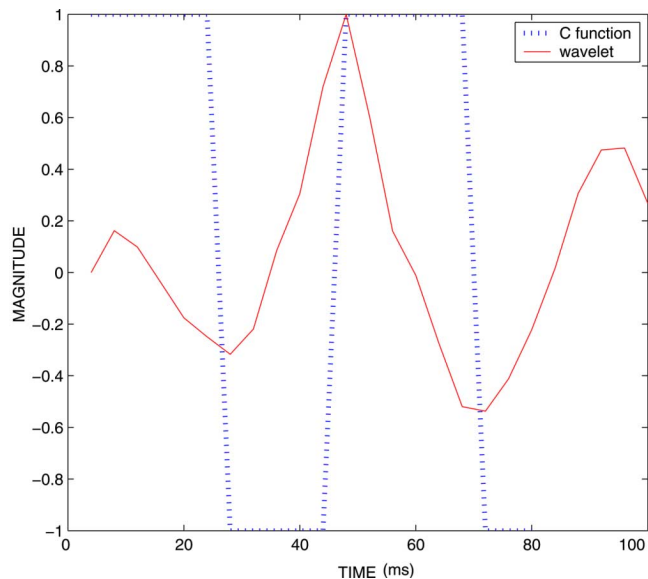


Fig. 18. Truncated estimated wavelet and C function.

order but on their time of maximum energy will decrease the signal bandwidth, and increase significantly the low-frequency content, and thus, the penetration. Each source is composed of 13 air guns, shooting a wavelet every 20 s with spectrum bandwidth 0–128 Hz. The multitrace streamer is composed of 360 clusters of 16 hydrophones. Its role is to improve the SNR by summing the received signals. The length of the streamer is 4.5 km and each cluster is separated by 12.5 m.

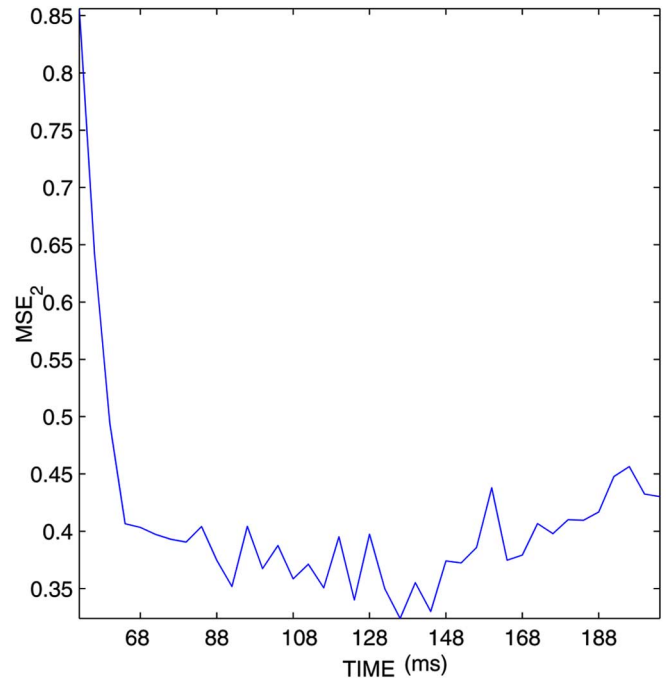


Fig. 19. \widehat{MSE} .

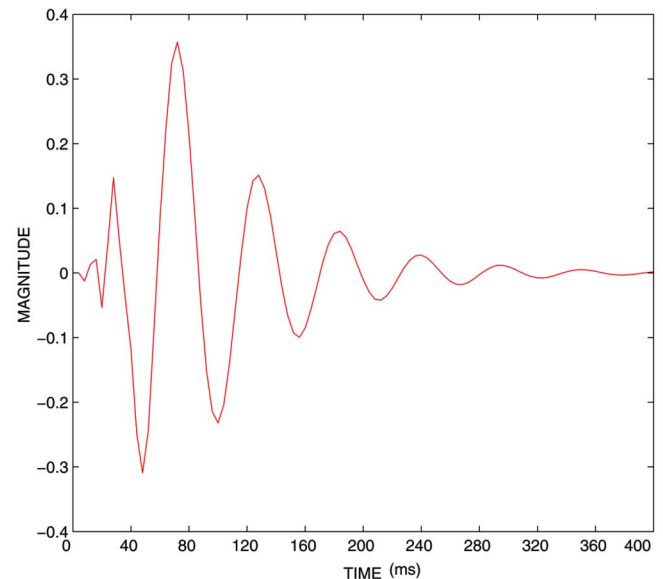


Fig. 20. Estimated wavelet by SEM + P.

A typical recorded trace is presented in Fig. 16. Fig. 17 shows the full length wavelet estimated by SEM, while Fig. 18 represents a truncated wavelet estimated by SEM together with the corresponding C function. The wavelet in Fig. 17 does not correspond to the kind of wavelet generated by air guns.

Fig. 19 shows that a good choice for the truncated wavelet length is 130 ms and Fig. 20 shows the Prony extension of the truncated wavelet estimate (SEM + P estimate). Fig. 21 shows the deconvolution results obtained for several traces. Fig. 22 presents the wavelets estimated by the SEM + P algorithm for different recorded seismic traces. We can observe that these wavelets are somewhat distinct, and thus, wavelet estimation

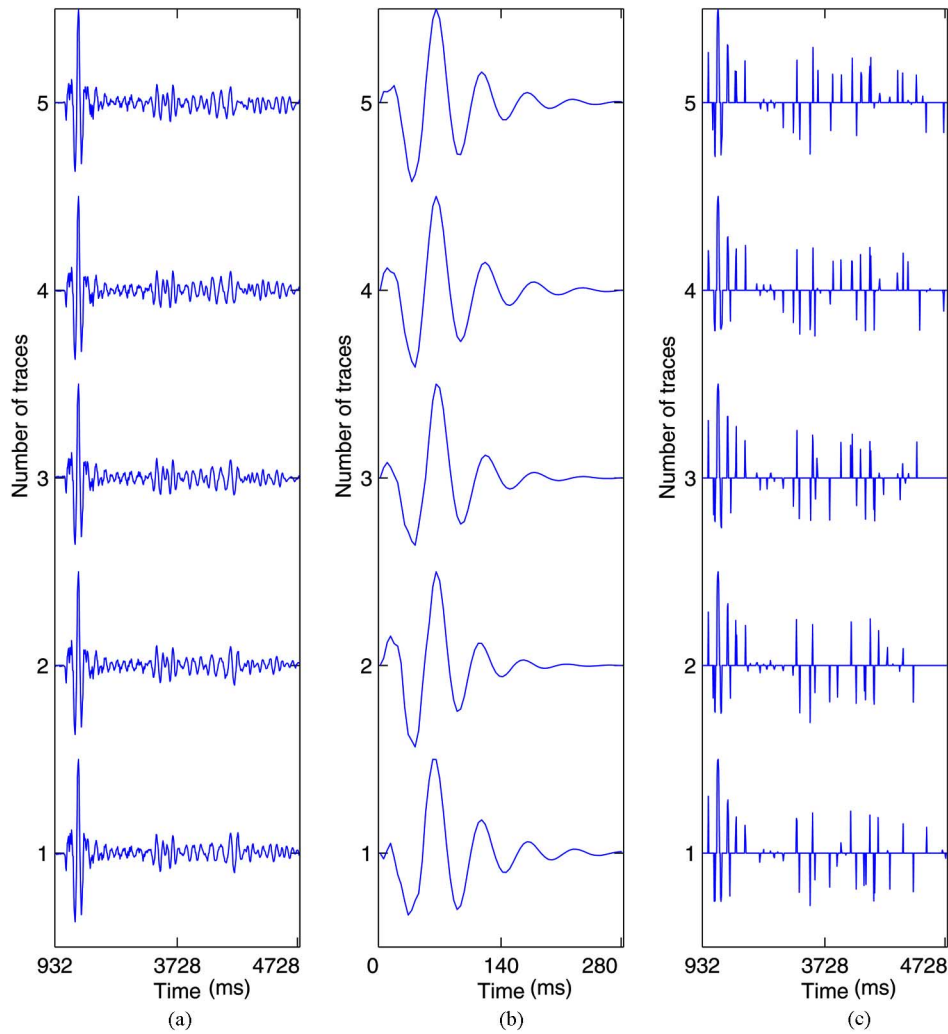


Fig. 21. (a) Real data (traces 402 to 406). (b) Estimated wavelet by SEM + P. (c) Estimated reflectivity by (SEM + P) + MPM.

is necessary on each trace for proper recovery of reflectivity. However, there is a good similarity among them, and thus, the wavelet estimated for one trace could be a good initialization candidate for the deconvolution of the other traces.

The parameters $(\lambda, \sigma_0^2, \sigma_1^2, \sigma_w^2)$ estimated with the SEM algorithm are presented in Table V. The noise variance seems to be the parameter that more varies among traces.

The seismic profile is presented in Figs. 23 and 24. We can see the deconvolution results obtained with our approach. On raw data presented in Fig. 23, the main reflectors are about 200-ms width (see seafloor, for example), and composed of an alternation of two black and three white phases. In Fig. 24, after deconvolution, the reflectors are reduced to a single black phase, about 70 ms wide. Thus, the deconvolution, as expected, improves substantially the resolution of the section. The decrease of the signal width enables us to define more precisely the relationships between the different seismic units: For example, it is now much easier to follow the top of the erosional surface at the top left of the seismic line. It helps discriminating real events from artifacts: A seismic reflector, about 300 ms below the seafloor, with a negative polarity, the polarity and existence of which were not obvious, is now well imaged. Weaker reflectors which were not visible because of their interferences with

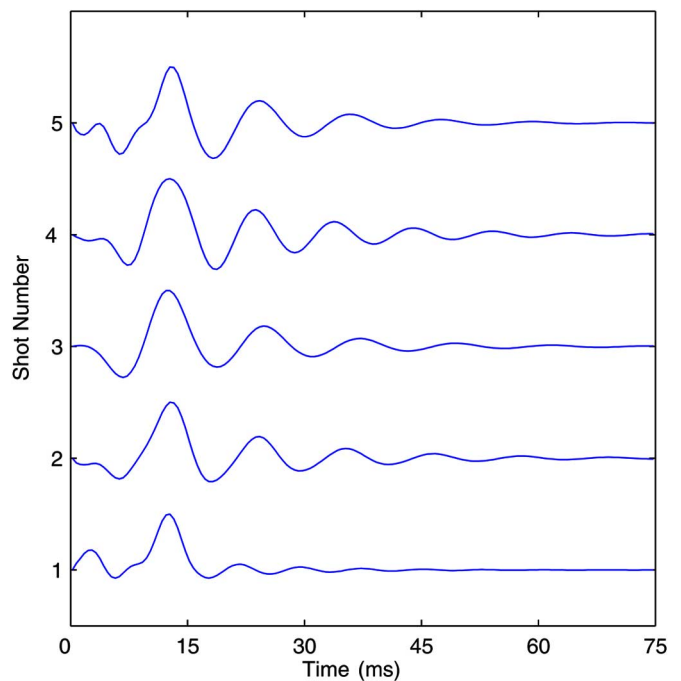


Fig. 22. Estimated wavelets.

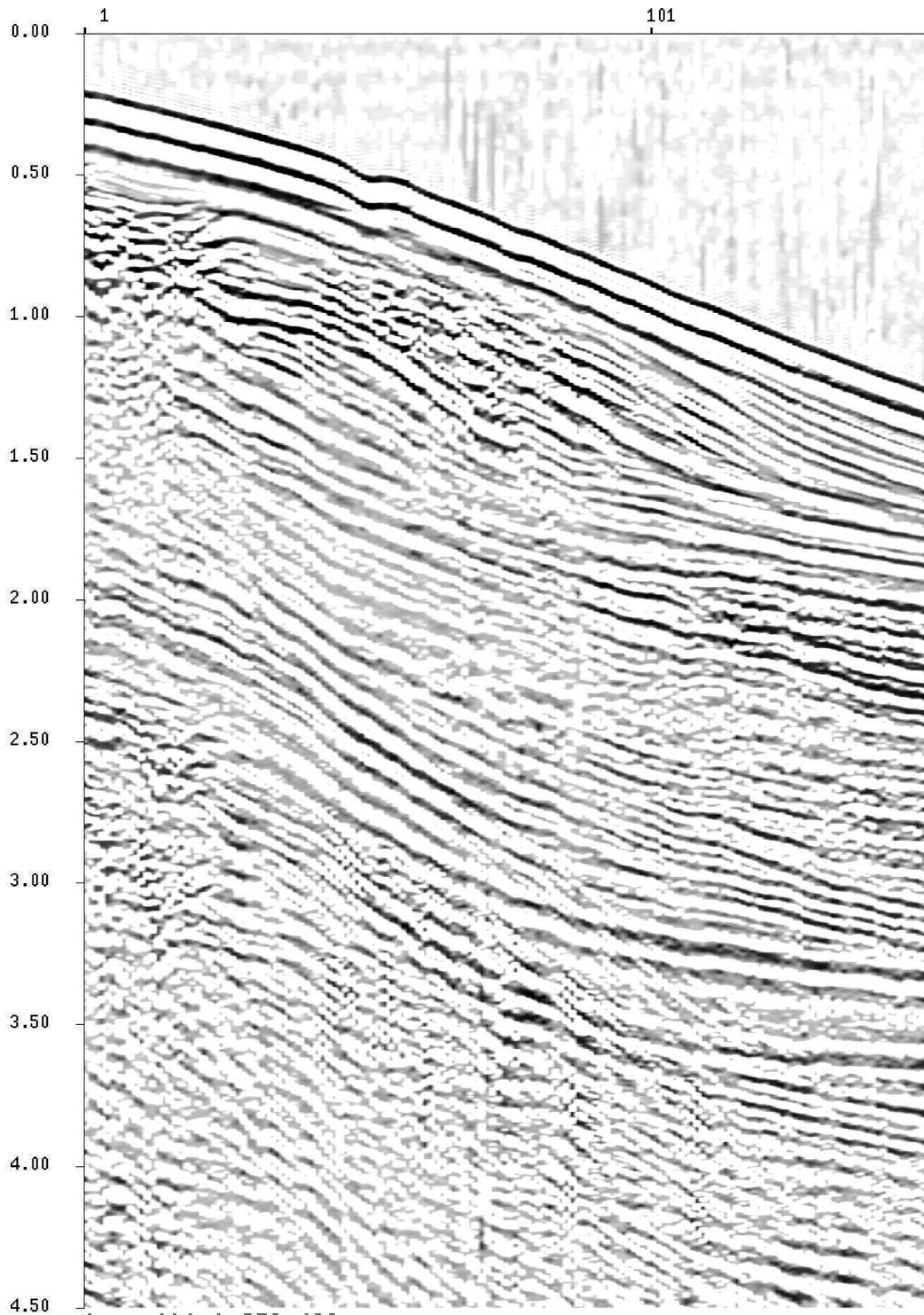


Fig. 23. Seismic profile. Y-label: shot number. X-label: time (s).

bouncing of the major reflectors are now apparent (see, for example, reflectors just below the seafloor).

VI. CONCLUSION

In this paper, we have proposed a new approach for blind deconvolution of seismic data in the presence of a long wavelet.

We have shown that the estimation of a truncated wavelet followed by an ARMA extension (SEM + P) yields improved wavelet estimation, compared to the classical SEM approach. We have also proposed a new method for choosing the truncated wavelet order. Furthermore, an efficient procedure has been proposed for the wavelet initialization in the SEM algorithm. Simu-

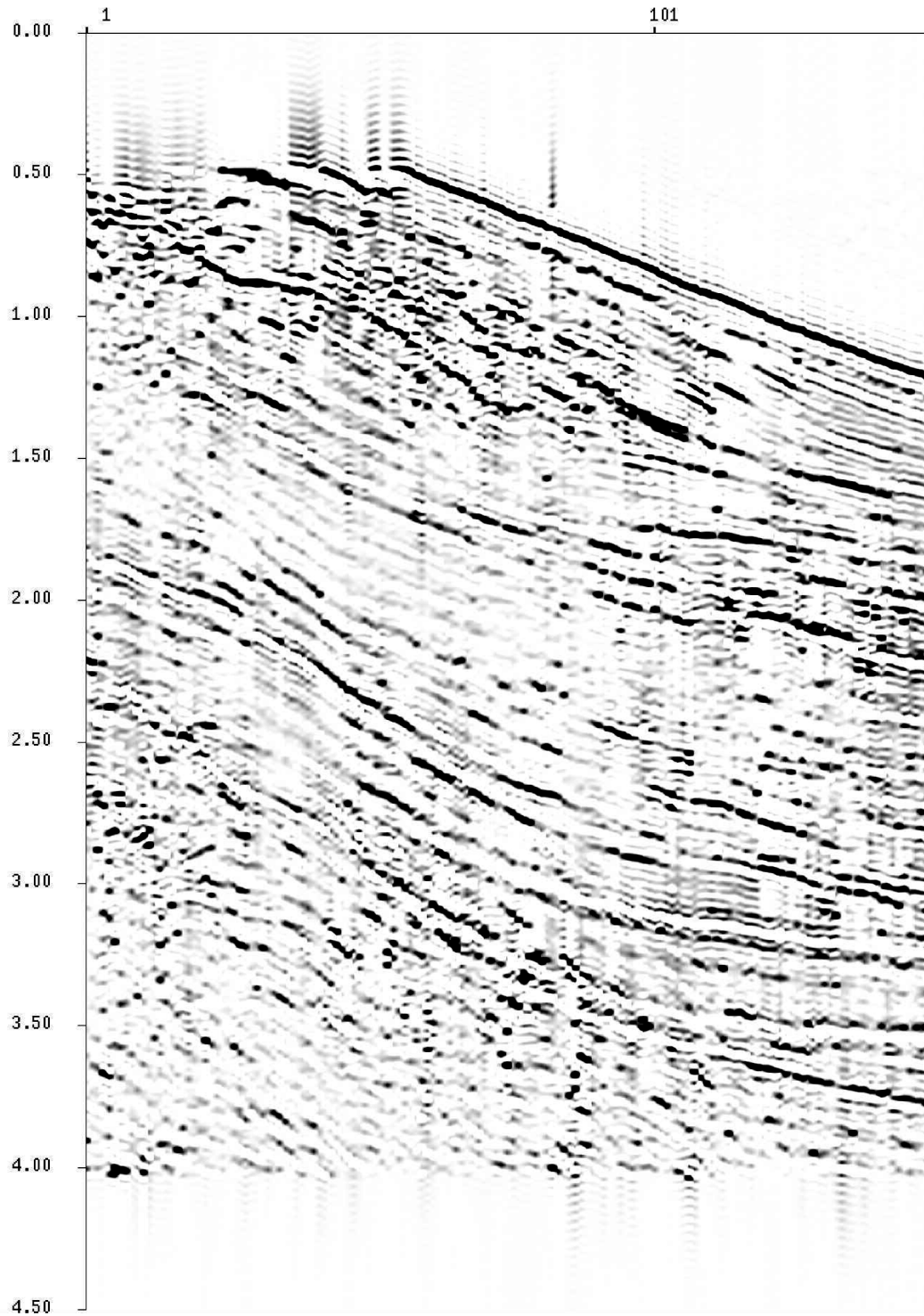


Fig. 24. Deconvolved seismic profile. Y-label: shot number. X-label: time (s).

lations and real-data experiments have shown that our approach achieves significant improvement.

APPENDIX

In this section, we justify why the proposed criterion for maximum position selection works efficiently, as shown in the simulation part. Let us recall, as discussed in Section IV-B, that when

the impulse response \mathbf{h} is initialized with one at the k th entry and zeros at other entries, the SEM algorithm converges to a solution where the estimated \mathbf{h} has its maximum at position k . In other words, we can say that the SEM algorithm looks for a solution that minimizes the norm error $\|\mathbf{y} - \mathbf{R}\mathbf{h}\|$ with a maximum constraint at position k .

Let us rephrase this idea in the continuous time domain. We are led to search for a solution h_t that achieves minimum error

norm under the null derivative constraint $h'_{t_0} = 0$. In mathematical terms, we are considering the following constrained optimization problem:

$$\begin{cases} \min_h \|y_t - (r * h)_t\| \\ h'_{t_0} = 0 \end{cases} \quad (21)$$

Equivalently, problem (21) can be rewritten in the Fourier transform domain

$$\begin{cases} \min_{\hat{h}} \int_B |\hat{y}(f) - \hat{r}(f)\hat{h}(f)|^2 df \\ h'_{t_0} = \int_B (2i\pi f) e^{2i\pi f t_0} \hat{h}_{t_0}(f) df = 0 \end{cases} \quad (22)$$

where $\hat{g}(f)$ denotes the Fourier transform of $g(t)$ and B is the signal bandwidth. Using Lagrange multipliers (see, for instance, [38]) and introducing real and complex variations of \hat{h} yields the following conditions upon the solution of problem (22), denoted by $\hat{h}^{(t_0)}(f)$:

$$\begin{cases} 2\text{Re}\{[\hat{r}(f)\hat{h}^{(t_0)}(f) - \hat{y}(f)]^* \hat{r}(f)\} + \lambda 2i\pi f e^{2i\pi f t_0} = 0 \\ 2i\text{Im}\{[\hat{r}(f)\hat{h}^{(t_0)}(f) - \hat{y}(f)]^* \hat{r}(f)\} + i\lambda 2i\pi f e^{2i\pi f t_0} = 0. \end{cases} \quad (23)$$

Indeed, considering the functional

$$J(\hat{h}, \lambda) = \int_B |\hat{y}(f) - \hat{r}(f)\hat{h}(f)|^2 df + \lambda \int_B (2i\pi f) e^{2i\pi f t_0} \hat{h}_{t_0}(f) df \quad (24)$$

and denoting by $\delta\hat{h}_r(f)$ any real valued small variation of $\hat{h}(f)$, the optimality constraint $J(\hat{h}^{(t_0)} + \delta\hat{h}_r, \lambda) - J(\hat{h}^{(t_0)}, \lambda) = 0$ yields

$$\int_B \left[\hat{r}(f)\hat{h}^{(t_0)}(f) - \hat{y}(f) \right] \hat{r}^*(f) + [\hat{r}(f)\hat{h}^{(t_0)}(f) - \hat{y}(f)]^* \times \hat{r}(f) + \lambda (2i\pi f) e^{2i\pi f t_0} \delta\hat{h}_r(f) df = 0 \quad (25)$$

leading thus to the first equation of (23). The second equation of (23) is derived in a similar way by considering imaginary small variation of $\hat{h}(f)$. Then, summing both equations of (23) yields

$$[\hat{r}(f)\hat{h}^{(t_0)}(f) - \hat{y}(f)] \hat{r}^*(f) = -N(2i\pi f) e^{-2i\pi f t_0} \quad (26)$$

with $N = (1 + i)\lambda/2$. Then, the solution of (22) is of the form

$$\hat{h}^{(t_0)}(f) = \frac{\hat{y}(f)}{\hat{r}(f)} - N \frac{(2i\pi f) e^{-2i\pi f t_0}}{|\hat{r}(f)|^2}. \quad (27)$$

Now, let us denote $\tilde{h}(f) = \hat{y}(f)\hat{r}^{-1}(f)$, which corresponds to the true wavelet in the noise-free case. After we insert solution (27) in the constraint equation $h'_{t_0} = \int_B (2i\pi f) e^{2i\pi f t_0} \hat{h}_{t_0}(f) df = 0$, it follows:

$$N = \tilde{h}'_{t_0} \left(\int_B (2i\pi f)^2 |\hat{r}(f)|^{-2} df \right)^{-1}. \quad (28)$$

Thus, from (27) and (28), we get the time-domain solution

$$h_t^{(t_0)} = \tilde{h}_t - \tilde{h}'_{t_0} \left(\frac{\int_B (2i\pi f) e^{2i\pi f(t-t_0)} |\hat{r}(f)|^{-2} df}{\int_B (2i\pi f)^2 |\hat{r}(f)|^{-2} df} \right). \quad (29)$$

Now, let us remark that

$$h_0^{(t_0)'} = \tilde{h}'_0 - \tilde{h}'_{t_0} \left(\frac{\int_B (2i\pi f)^2 e^{-2i\pi f t_0} |\hat{r}(f)|^{-2} df}{\int_B (2i\pi f)^2 |\hat{r}(f)|^{-2} df} \right). \quad (30)$$

Clearly, if $\tilde{h}'_0 = 0$, $h_0^{(t_0)'}$ changes sign at $t = t_0$ if \tilde{h} has a local optimum at point t_0 . This shows that, providing the true wavelet has an horizontal tangent at point 0, the derivative of the estimated wavelet $h^{(t)}$ changes sign around point $t = t_0$ [note that the term in the parenthesis in (30) is always positive]. We have checked that this result remains true for wavelets \tilde{h} with small tangent slopes at point 0 which corresponds to most practical situations, where wavelets have a smooth shape and thus small slope at the origin.

ACKNOWLEDGMENT

The authors would like to thank B. Marsset and T. Yannick for their data.

REFERENCES

- [1] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1980, pp. 152–161.
- [2] O. Rosec and J. M. Boucher, "Bayesian estimation of non-minimum phase wavelets applied to marine reflection seismic data," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 1999, vol. 5, pp. 2797–2800.
- [3] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*. New York: Academic, 1983.
- [4] O. Rosec, J. M. Boucher, B. Nsiri, and T. Chonavel, "Blind marine seismic deconvolution using statistical MCMC methods," *IEEE J. Ocean. Eng.*, vol. 28, no. 3, pp. 502–512, Jul. 2003.
- [5] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.
- [6] D. L. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis*, D. Findley, Ed. New York: Academic, 1981, pp. 556–608.
- [7] M. Boumahdi, "Blind identification using the Kurtosis with applications to field data," *Signal Process.*, vol. 48, no. 3, pp. 205–216, 1996.
- [8] M. Boujida and J. M. Boucher, "Higher order statistics applied to wavelet identification of marine seismic signal," in *Proc. 8th Eur. Signal Process. Conf.*, 1996, vol. 1, pp. 137–141.
- [9] C. L. Nikias, "ARMA bispectrum approach to nonminimum phase system identification," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 36, no. 4, pp. 513–524, Apr. 1988.
- [10] A. Larue, J. I. Mars, and C. Jutten, "Frequency-domain blind deconvolution based on mutual information rate," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1771–1781, May 2006.
- [11] J. M. Mendel, *Maximum-Likelihood Deconvolution*. New York: Springer-Verlag, 1990, pp. 7–59.
- [12] K. F. Kaarensen, "Deconvolution of sparse spike trains by iterated window maximization," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1173–1183, May 1997.
- [13] C. P. Robert, *The Bayesian Choice*. New York: Springer-Verlag, 1994, pp. 344–355.
- [14] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Ser. B*, no. 39, pp. 1–38, 1977.
- [15] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-36, no. 4, pp. 477–489, Apr. 1988.
- [16] G. Celeux, D. Chauveau, and J. Diebolt, "Stochastic versions of the EM algorithm," *J. Stat. Comput. Simul.*, vol. 55, pp. 287–314, 1996.
- [17] G. Celeux and J. Diebolt, "The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem," *Comput. Stat.*, vol. 2, pp. 73–82, 1985.
- [18] G. Celeux and J. Diebolt, "A stochastic approximation type EM algorithm for the mixture problem," *Stoch. Stoch. Rep.*, vol. 22, pp. 747–761, 1989.
- [19] M. Lavielle, "Bayesian deconvolution of Bernoulli-Gaussian processes," *Signal Process.*, vol. 33, pp. 67–79, 1993.

- [20] A. Doucet and P. Duvaut, "Bayesian estimation of state-space models applied to deconvolution of Bernoulli-Gaussian processes," *Signal Process.*, vol. 57, pp. 147–161, 1997.
- [21] J. Idier and Y. Goussard, "Multichannel seismic deconvolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 5, pp. 961–980, Sep. 1993.
- [22] Q. Cheng, R. Chen, and T. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 377–384, Mar. 1996.
- [23] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Amer. Stat. Assoc.*, vol. 82, pp. 76–89, 1987.
- [24] B. Chalmond, "An iterative Gibbsian technique for reconstruction of m-ary images," *Pattern Recognit.*, vol. 22, pp. 747–761, Jun. 1989.
- [25] S. Geman and D. Geman, "Statistic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [26] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Stat. Soc.*, vol. 48, pp. 259–302, 1986.
- [27] B. Nsiri, J. M. Boucher, and T. Chonavel, "Multichannel blind deconvolution. Application to marine seismic," in *Proc. OCEANS*, Sep. 2003, vol. 5, pp. 22–25.
- [28] P. Farcy, "System acquisition of submarine seismic," IFREMER, Bretagne, France, ESSR4 Campaign Rep., DNIS/ESI/ENS/DTI/99-007, 1999.
- [29] T. Chonavel, *Statistical Signal Processing, Modelling and Estimation*. New York: Springer-Verlag, 2002, pp. 166–182.
- [30] T. W. Parks and C. S. Burrus, *Digital Filter Design*. New York: Wiley, 1987, pp. 226–228.
- [31] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. London, U.K.: Chapman & Hall, 1996, pp. 1–17.
- [32] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *J. Roy. Stat. Soc.*, vol. 85, no. 410, pp. 398–409, 1990.
- [33] O. Rosec, "Déconvolution Aveugle Multicapteur en Sismique Réflexion Marine Très haute Résolution," Ph.D. dissertation, Univ. Bretagne Occidentale, Brest, France, 2000.
- [34] O. Rabaste and T. Chonavel, "Estimation of multipath channels with long impulse response at low SNR via an MCMC method," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 55, no. 4, pp. 1312–1325, Apr. 2007.
- [35] B. Nsiri, T. Chonavel, and J. M. Boucher, "SEM blind identification of ARMA models application to seismic data," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2004, pp. 1099–1102.
- [36] B. Nsiri, T. Chonavel, and J. M. Boucher, "Blind seismic deconvolution with long wavelet impulse response," in *Phys. Signal Image Process.*, Jan. 2003, pp. 5–8.
- [37] M. Lavielle and E. Moulines, "A simulated annealing version of the EM algorithm for non-Gaussian deconvolution," *Stat. Comput.*, vol. 7, pp. 229–236, 1997.
- [38] J. L. Troutman, *Variational Calculus and Optimal Control: Optimization With Elementary Convexity*. New York: Springer-Verlag, 1996, pp. 97–186.



Benayad Nsiri (S'02–A'04) was born in Temara, Morocco, in 1974. He received the D.E.A. degree (French equivalent of M.Sc. degree) in electronics from the Occidental Bretagne University, Brest, France, in 2000 and the Ph.D. and the M.B.I. degrees in computer sciences from Ecole Nationale Supérieure des Télécommunications (ENST) de Bretagne, Brest, France, in 2004 and 2005, respectively.

Currently, he is an Assistant Professor at the Faculty of Sciences Ain Chock, Hassan II University, Casablanca, Morocco, a Member Associate in

GSCM-LRIT group, Rabat Morocco, and a Member Associate at the Signal and Communication Department, ENST. His research interests include but are not restricted to signal processing, adaptive techniques, blind deconvolution, MCMC methods, seismic data, and higher order statistics.



Thierry Chonavel received the Ph.D. degree in signal processing from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1992. His thesis was about trigonometric moment problems and array processing.

He joined the Groupe des Ecoles des Télécommunications (GET) in 1993. Currently, he is a Professor at Ecole Nationale Supérieure des Télécommunications (ENST) de Bretagne, Brest, France. His research interests are in signal processing and communications with works in adaptive processing,

code division multiple access (CDMA), geolocalization, automotive radars, and applications related to underwater acoustics (ocean tomography and seismic).



Jean-Marc Boucher (M'84–SM'05) received the Engineering degree in telecommunications from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1975 and the Habilitation à Diriger des Recherches degree from the University of Rennes 1, Rennes, France, in 1995.

Currently, he is a Professor at the Department of Signal and Communications, ENST, where he is also Education Deputy Director. He is the Deputy Manager of a National Scientific Research Center Laboratory, Traitement Algorithmique et Matériel de

la Communication, de l'Information et de la Connaissance (TAMCIC), UMR 2872, Brest, France. His current research interests are related to statistical signal analysis including estimation theory, Markov models, blind deconvolution, wavelets, and multiscale image analysis. These methods are applied to radar and sonar images, seismic signals, and electrocardiographics. He has published about 100 technical articles in these areas in international journals and conferences.



Hervé Nouzé received the Engineering degree from Ecole Nationale Supérieure des Mines de Nancy in 1992, and the degree from Ecole Nationale du Pétrole et des Moteurs in 1993.

Since 1995, he has been a Geophysicist at the Marine Geosciences Department, Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Bretagne, France. His main research interest is the development of geophysical (mainly seismic) methods to detect and quantify the occurrence of gas and gas hydrates in marine sediments.