HIGHER ORDER STATISTICS APPLIED TO WAVELET IDENTIFICATION OF MARINE SEISMIC SIGNALS

Mohammed Boujida & Jean-Marc Boucher

Télécom Bretagne, Département Signal et Communications
BP 832, 29285 BREST Cedex, FRANCE
Tel : 98 00 13 57, Fax: 98 00 10 12, E-mail : JM.Boucher @enst-bretagne.fr

ABSTRACT

The purpose of this paper is to present the use of higher order statistics to solve the blind identification problem of reflection seismic data. We develop and compare some non-parametric and parametric methods based on higher order statistics. To compare these methods, non-minimum phase wavelet and non-gaussian reflectivity function are simulated. They are then applied to real data of high resolution marine seismic reflection.

1 INTRODUCTION

The improvements in seismic data representation and seismic data processing have clearly become more useful in the case of rig site survey assessment, imaging superficial sea-bottom layers as well as their discontinuities, and in spatially extending the findings of geotechnical bore hole logging. The capability of understanding the seabed geology at a very high resolution (which means more than 100 meters of penetration with at least 1 meter of resolution in the interface distinction) has consequently improved rig installations at sea [1].

High resolution marine seismic reflection uses varied sources (Sediment sounder, Sparker, Water-gun) to analyse sea-floor morphology and underlie shallow stratigraphy. Non-minimum phase wavelet produced by these sources is reflected by different layers composing the earth. In seismic reflection it is customary to model the discrete seismic trace x(k) as the convolution of the seismic wavelet w(k) and the reflectivity function r(k) containing the subsurface information with an additive white gaussian noise:

$$x(k) = h(k) * w(k) + n(k) \tag{1}$$

In practical cases, only the trace signal is recorded. In order to deconvolve the reflectivity sequence from observation of the reflected signal, the wavelet has to be estimated . The problem is not new, and several principal classical approaches were proposed, but they are not suited to this type of application. Predictive deconvolution supposes that the wavelet is a minimum phase signal , and calculates the inverse filter. Homomorphic deconvolution [2] requires that the reflectivity sequence is completely separated from the wavelet in the cepstrum domain. It may be very sensitive to observation noise.
We have already propose to associate a non-minimum phase ARMA model for the source signal, and a Bernoulli-Gaussian model for the sequence reflectivity [3]. The parameters of the wavelet and the hyperparameters are unknown, and restoration of reflectivity is a blind deconvolution problem that requires different operations: identification of the wavelet, estimation of the hyperparameters and restoration of the reflectivity [3].

Here we focus our interest on source signal identification because the result of deconvolution is very sensitive to the estimation of the wavelet.

The purpose of this paper is to explore the use of higher order statistics (HOS) in order to obviate the minimum-phase requirement. This is done either in the frequency domain using non-parametric approaches [4,5], or time domain using parametric approaches based on MA, or ARMA models. We tested different methods ,which determine the wavelet from the reflection seismogram . Methods are based on : phase estimation from the Bispectrum [4] , complex cepstrum of higher order cumulants [5], parametric methods with an ARMA model: Giannakis-Mendel Algorithm [7] based on cumulants, and the method based on spectrally equivalent minimum phase completed by higher order statistics [6,9].

These methods are applied to different synthetic seismic signals, and to real signals.

2 CUMULANTS, POLYSPECTRA AND POLYCEPSTRA

Let the trace $\{x(k)\}$ be as zero-mean, random, stationary, real discrete time process which is generated as follows:

$$x(k) = \sum_{i=-\infty}^{+\infty} h(l) w(k-l) \tag{2}$$

where $w(k)$ is zero -mean non-gaussian, white identically independent distributed (i.i.d) noise.
The k-th order cumulant is given by:

$$C_k^x(n_1,\ldots,n_{k-1}) = m_k^x(n_1,\ldots,n_{k-1}) - m_k^g(n_1,\ldots,n_{k-1}) \tag{3}$$

where:

$$m_k^x(n_1,\ldots,n_{k-1}) = E\left[x(k)x(k+n_1)\ldots x(k+n_{k-1})\right]$$

and $\{g(k)\}$ is a gaussian random-process with the same second-order statistics as $\{x(k)\}$

$$C_k^x(n_1,\ldots,n_{k-1}) = \gamma_k^w \sum_i h(i)h(i+n_1)\ldots h(i+n_{k-1}) \tag{4}$$

γ_k^w is the k-th cumulant of the reflectivity sequence:

$$\gamma_k^w = C_k^w(n_1,\ldots,n_{k-1})$$

The Polyspectrum of the output trace is defined by :

$$S_k^x(\nu_1,\ldots,\nu_{k-1}) = \sum_{n_1}\ldots\sum_{n_{k-1}} C_k^x(n_1,\ldots,n_{k-1}) \cdot \exp\left[-j2\pi(n_1\nu_1+\ldots+n_{k-1}\nu_{k-1}\right] \quad (5\text{-}1)$$

$$S_k^x(\nu_1,\ldots,\nu_{k-1}) = \gamma_k^w H(\nu_1)\ldots H(\nu_{k-1})H^*(\nu_1+\ldots+\nu_{k-1}) \quad (5\text{-}2)$$

The Polycepstrum is defined using Z-transform by :

$$F_k^x(m_1,\ldots,m_{k-1}) = DFT^{-1}\left\{Ln\left[S_k^x(\nu_1,\ldots,\nu_{k-1})\right]\right\} \quad (6)$$

3 NON-PARAMETRIC APPROACH

In this paragraph we show how to determine the wavelet from the output trace using a non-parametric approach based on polyspectrum or polyceptrum at third-order.
Some results of computer simulation will be included to illustrate the variance of the phase estimation.

3.1 Wavelet estimation using the Bispectrum

We use a nonrecursive approach [4] with all bispectral values because recursive algorithms are very sensitive to errors. We have to make the estimation of the amplitude $|H(\nu)|$ and the phase $\phi(\nu)$.
The amplitude $|H(\nu)|$ is estimated by any algorithm available for spectral estimation. The wavelet phase is obtained from the Bispectrum phase $\psi(\nu_1,\nu_2)$.

$$S_3^x(\nu_1,\nu_2) = \gamma_3^w H(\nu_1)H(\nu_2)H^*(\nu_1+\nu_2) \quad (7)$$

From (7) we can write:

$$\psi(\nu_1,\nu_2) = \phi(\nu_1)+\phi(\nu_2)-\phi(\nu_1+\nu_2) \quad (8)$$

Using (8) and the symmetry properties of the Bispectrum, we form all the possible equations for $\nu_1 = 1,2,\ldots,N/2$ and $\nu_2 = \nu_1, \nu_1+1,\ldots,N-\nu_1$, then we obtain the following matrix equation :

$$\mathrm{A}\Phi = \Psi \quad (9)$$

The unknown phase vector Φ is determined by the least square solution:

$$\Phi = (A^T A)^{-1} A\Psi \quad (10)$$

We used the (marron,al.[8]) method to unwrap the Bispectrum phase.

3.2 Wavelet identification using the Bicepstrum

This method reconstructs the minimum and maximum phase components of the wavelet separately, and it is flexible enough to accommodate a general ARMA non-gaussian process and it does not require model order selection criteria.
We compute the impulse response of the wavelet $h(k)$ as the convolution : $h(k) = i(k)*o(k)$.
where $o(k)$ is the maximum phase component and $i(k)$ is the minimum one. They are computed recursively by :

$$i(k) = \frac{-1}{k}\sum_{n=2}^{k+1} i_d(n)i(k-n+1) \qquad k \geq 1$$

$$o(k) = \frac{-1}{k}\sum_{n=k+1}^{0} o_d(n)o(k-n+1) \qquad k \leq -1$$

$$i(0) = o(0) = 1 \quad (11)$$

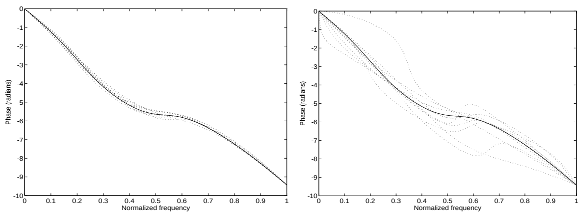
Where $i_d(k)$ and $o_d(k)$ are the differential cepstra of $i(k)$ and $o(k)$, respectively.
The differential cepstrum coefficients of the system impulse are estimated by resolving the linear convolution formula:

$$C_k^x(n_1,n_2)*\left[-mF_k^x(m_1,m_2)\right] = C_k^x(n_1,n_2) \quad (12)$$

Pan & Nikias [6] propose two methods to solve the polycepstral equation (12). The first one is based on Least-Squares estimator , the second is based on Fourier transform.

3.3 Simulation results

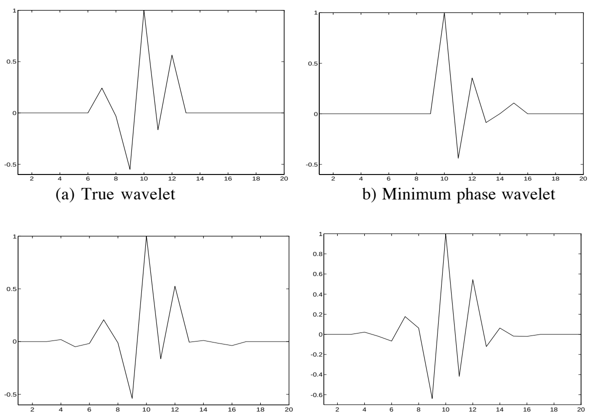
To compare the two previous methods, we illustrate an example in which the wavelet is a non-minimum phase moving average MA(5) with zeros located at : -2,093 ; 1,1856 ± j 0,091 ; -0,073 ± j 0,703 .
The system driving noise $\{w(k)\}$ is zero-mean exponentially distributed, white, with $\gamma_3^w \neq 0$. Different lengths of output data have been used for simulations.
Figure.1 shows the phase response at each frequency for 10 different output data realisations of the same statistical description using the Bicepstrum method (a) and Bispectrum method (b). The values of the Bispectrum and cumulant was computed from 32 records of 128 points.



(a) Bicepstrum method (b) bispectrum method
__ : true phase ... : estimated phase

Fig.1

Figure.2 shows the estimated wavelets.



(a) True wavelet b) Minimum phase wavelet
(c) Estimated wavelet using Bicepstrum (d) Estimated wavelet using Bispectrum

Fig.2

When data length is sufficiently long, the cepstrum method is preferred to the spectrum method. But in case of seismic reflection, both methods are not useful , because of the high variance due to the short data length.

4 PARAMETRIC APPROACH

We present two methods based on cumulants, the Giannakis-Mendel [7] algorithm (GM) and the spectrally equivalent minimum phase (SEMP) method.

4.1 The Giannakis-Mendel method

We suppose that the source wavelet is modelled as a causal invariant time ARMA filter with z-transform:

$$H(z) = \frac{B(z)}{A(z)} = \frac{1+\sum_{i=1}^{q} b_i z^{-i}}{1+\sum_{i=1}^{p} a_i z^{-i}} \tag{13}$$

The information about the non-minimum phase is contained in zero locations of the MA filter. To identify the AR coefficients we can use the Yule-Walker equations based on second or higher order cumulants.

The output trace is filtered by the inverse AR filter :

$$x'(n) = x(n) - \sum_{k=1}^{p} \hat{a}(k)\, x(n-k) \tag{14}$$

Then the MA non-minimum phase identification algorithm[7] is applied to identify the MA parameters .

4.2 The Spectrally equivalent minimum phase method (SEMP)

To solve the problem of the high variance of higher order statistics estimation, especially with "short" data records, we prefer to first estimate the spectrally equivalent minimum phase filter using second-order statistics, then we systematically test all the zero configurations inside and outside the unit circle which gives the same power spectra density . The true filter can be obtained by a minimum entropy deconvolution approach (MED) which is based on the maximisation of the objective function [9] :

$$Q_{\hat{w}} = \left| \frac{C_4^{\hat{w}}}{(\sigma^2)^2} - 3 \right| \tag{15}$$

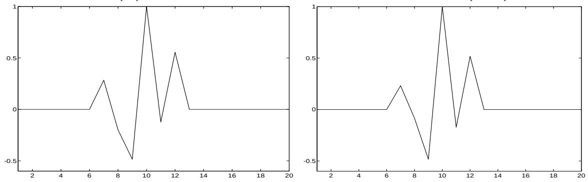
where $C_4^{\hat{w}}$ is the fourth-order cumulant of the deconvolved sequence $\hat{w}(n)$. But this method is limited by the difficulty of the deconvolution operation, especially in case of real seismic signal , where the reflectivity is very difficult to estimate by a simple deconvolution.

In order to obviate this limitation we prefer to estimate the true filter without the deconvolution operation. To do that we compute the true filter by minimising the cumulant difference square sum of the observed data and the ones given by the i-th filter, which is obtained from the i-th zeros location. The objective function to minimise is :

$$J = \sum_{k_1} \cdots \sum_{k_{M-1}} \left(c_M^x(k_1,\ldots,k_{M-1} / \hat{h}_i) - \hat{c}_M^x(k_1,\ldots,k_{M-1}) \right)^2$$

where $\hat{h}_i$ is the i-th filter.

Fig.3 shows the estimated impulse response using GM (a) and SEMP (b) method for the same wavelet as (3.3)



(a) estimated wavelet using GM method (b) estimated wavelet using SEMP method

Fig.3

5 APPLICATION TO SEISMIC SIGNAL

Synthetic signal : The previous methods were evaluated on synthetic signals with different ARMA models, different signal to noise ratios and different data lengths . The reflectivity sequence was simulated by an asymmetric ($\gamma_3^w \neq 0$)and a symmetric ($\gamma_3^w = 0$, $\gamma_4^w \neq 0$) distributed sequence. The signal spectra are chosen to be close to those of real wavelet in seismic marine reflection. The following example shows the power spectra of the simulated non-minimum phase ARMA (6,5) source wavelet (Fig.4).

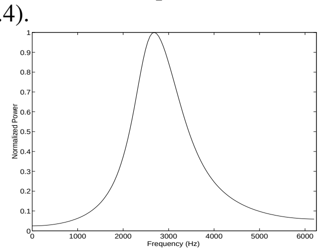
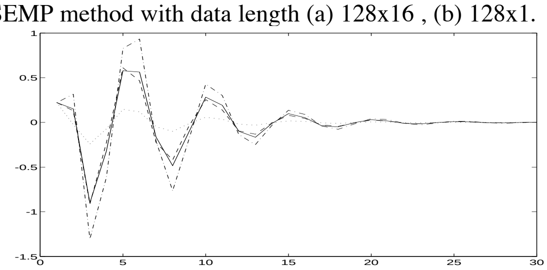
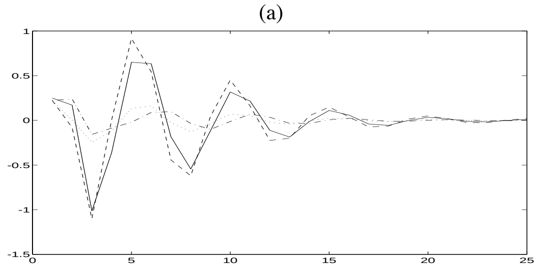


Fig. 4 wavelet Power spectra

The driving input reflectivity was an i.i.d Bernoulli-gaussian (B-G) process, which has been used to model the sparse reflectivity sequence. Hence the cumulant order is M=4 . Fig.5 shows the estimated wavelet using G-M , and SEMP method with data length (a) 128x16 , (b) 128x1.



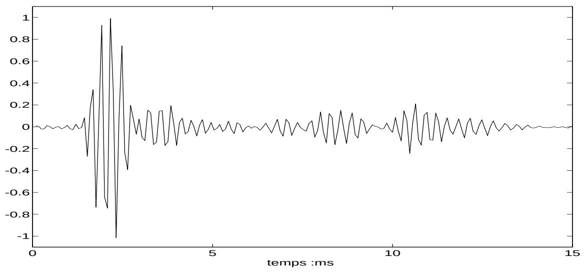
(a)



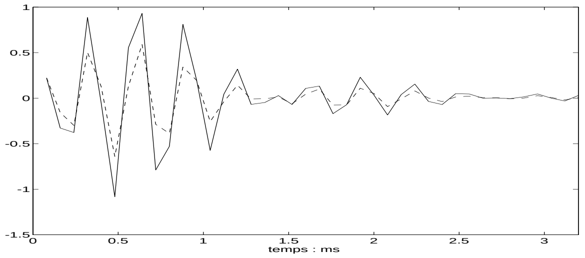
(b)

Fig.5 ---: True wavelet ... : M Phase --.--: GM -- --:SEMP

Real signal : These methods are applied to high resolution marine seismic traces given by IFREMER supplied by "Sediments sounder at 2,5khz" collected during IFREMER ESCOMED survey in the Mediterranean sea, (figure.6) and the ones supplied by " water-gun " collected during PASSISAR Campaign (figure.7).

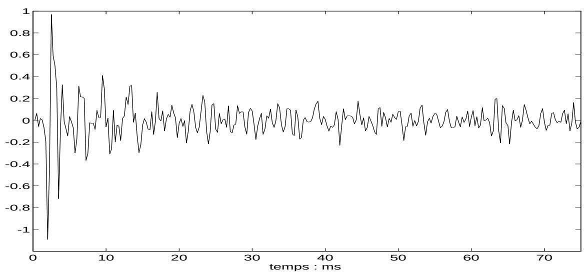


(a) recorded trace

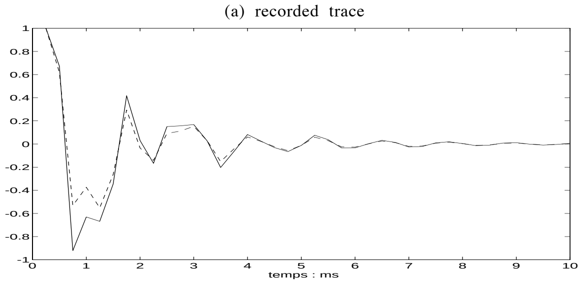


(b) estimated wavelet __ : NM Phase - - : M Phase

Fig.6



(a) recorded trace



(b) estimated wavelet __ : NM Phase - - : M Phase

Fig.7

CONCLUSION

To solve the blind identification problem of reflection seismic data, we use the HOS to obviate the minimum-phase requirement. Non-parametric and parametric approachs are presented. All these methods give a good estimate of the impulse response at long record length and low order ARMA parameters, because of the good estimation of higher order spectra or cumulants. The parametric methods present a lower estimation variance than non-parametric ones, which prove a high sensitivity to data length and to pole and zero locations.

On the other hand, the SEMP method gives better results than the GM method due to good spectra magnitude estimation using second-order statistics.

In the case of real signals, where we dispose of "short" data length, the SEMP methods (which use all the information contained in second-order statistics, and the strict necessary information contained in higher order statistics) are better suited to our application than the others. They give a satisfactory result for the data records given by "Sediment-sounder" source, but for the data records obtained by the "water-gun" the low signal to noise ratio leads to less good results.

Acknowledgments

This work is carried out with IFREMER collaboration (Dpt Geosciences marines (Lericolais,G.) and Dpt Genie Oceanique (Marsset,B.& Meunier,J.). We thank them for allowing the publication of the results obtained on real signals

References :

[1] G.Lericolais, J.P.Allenou, S.Berné, and P.Morvan " A new system for acquisition and processing of very high-resolution seismic reflection data " Geophysics, Vol:55, N°:8,August 1990, p:1036-1046.

[2] R.Hamidi, J-M Boucher, J-L Boelle " Combined use of homomorphic filtring and multipulse modeling for the deconvolution of seismic reflection signals " Signal Processing IV : Theories and applications. B.V EURASIP 1989 , pp:957-978.

[3] M.Boujida, J-M Boucher " Déconvolution en sismique réflexion marine par modèle Bernoulli-gaussien associ à une source à phase non-minimale " GRETSI , 1995, p: 1327-1330.

[4] T.Matsuoka, T.J.Ulrych " Phase Estimation using the Bispectrum " Proceeding of the IEEE, Vol:72, N°: 10, October, 1984, PP:1403-1411.

[5] R.Pan, C.L.Nikias " The complexe cepstrum of higher order cumulants and nonminimum phase system identification " IEEE Transactions on ASSP, Vol:36, N°: 2, February, 1988, pp:186-205.

[6] K.S.Lii, M.Rosenblatt" Deconvolution and estimation of transfer function phase and coefficients for non-gaussian linear processes" Ann.Statist, Vol:10, 1982, pp: 1195-1208.

[7] G.B.Giannakis, J.M.Mendel " Identification of non-minimum phase systems using higher order statistics " IEEE Transactions on ASSP, Vol:37, N°: 3, march, 1989, pp:360-377.

[8] J.C.Marron, P.P.Sanchez, and R.C.Sullivan "Unwrapping algorithm for least-squares phase recovery from the modulo 2π Bispectrum phase" J.Opt.soc.Am.A7(1), anuary, 1990, pp:14-20.

[9] D.Donoho " On minimum entropy deconvolution " in: D.E.Findley, ed., Applied Time Series Analysis II, Academic press, New York, 1981, pp: 565-608.