BOLLETTINO DI GEOFISICA TEORICA ED APPLICATA VOL. 41, N. 2, pp. 111-132; JUNE 2000

# MIDEM: an iterative deconvolution of virtual reflections for single channel seismic signals

E. DE MARINIS[1] and L. MIRABILE[2]

[1]*Free lance scientific consultant*
[2]*Istituto di Oceanlogia, Istituto Universitario Navale of Naples, Italy*

(Received September 9, 1999; accepted July 7, 2000)

**Abstract**. The present work introduces MIDEM (**M**inimum-squares **I**terative **DE**convolution of **M**ulti-reflectors): a simple deconvolution approach, designed for single channel seismic signals and well suited for a straightforward computer coding, with no need of any built-in specialised math function call. The approach is based on the introduction of the concept of virtual reflections and on the a priori estimate of the received signature. MIDEM first reduces the problem statement into a vector transformation, then uses as solutive kernel, a classical self-learning, gradient search approach, finding an optimum estimate of the solution (deconvolved signal) in a least squares sense. Stability and convergence criteria have been deduced from theory and reduced into bonds well suited for practical use. A further generalisation of the algorithm for the constrained minimisation problem has been presented as well. MIDEM has been tested both on synthetic and real sea trials data, always resulting in sharper scenarios and a resolution enhancement.

## 1. Introduction

Among others, one of the main limitations in single channel seismic signal analysis is the not negligible time duration and broadband structure of the source signature, each being the time-frequency domain counterpart of the other. Too many reasons, both theoretical and technological, impose lower limits to the time duration of the emitted signature, regardless the nature of the source (Mc Quillin and Bacon, 1979; Parkes and Hatton, 1986; Hatton et al., 1988); this is responsible for the loss of resolution and introduction of the ambiguity in the signal structure, often resulting in a severe limitation of the interpretation capabilities (Berkhout, 1984). Many well-known signal processing techniques try to remove the negative effects of the finite time-frequency structure of the signature on the received signal; most of them use an autoregeressive

Corresponding author: L. Mirabile; Istituto di Oceanologia, Istituto Universitario Navale, Via Acton 38, 80133 Napoli, Italy; phone: +39 081 5513123; fax: +39 081 5521608; e-mail: loremira@tin.it

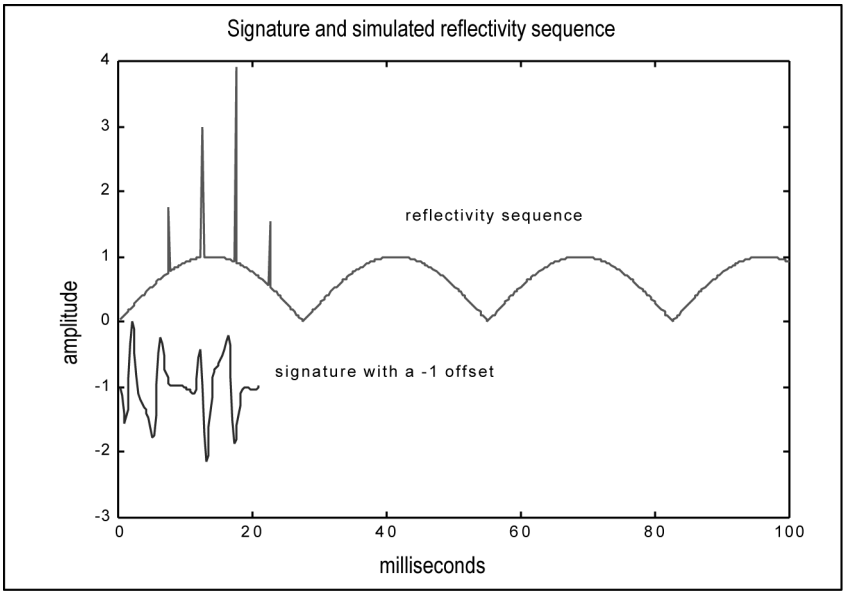


**Fig. 1** - Simulated reflecivity sequence and signature.

and moving average (ARMA) model of the received signal and operate a sort of direct deconvolution of the autocovariance matrix (Robinson, 1983, 1984), based on a minimum phase hypothesis of the signature and a Yule-Walker algorithm (just a faster version of the discrete Wiener-Hopf solution for this class of least squares problems, Haykin, 1986). The performances achievable by this class of direct (spiking or predictive) deconvolution algorithms are often poor (because the minimum phase hypothesis is only very seldom fulfilled) and only an accurate selection of their main driving parameters of these algorithms can result in an acceptable performance. The purpose of the present work is to obtain an analytical form of the ideally received signal well suited for the design of a feasible deconvolution approach, constrained to be robust, non-parametric and with an easy translation into a computer code.

## 2. Problem statement

Say $y_k$ is the $k$-th sample of the received signal, we may write all the data collected from a single shot as the column vector:

$$y^T = \left[ y_1, ..., y_{Nt} \right] \tag{1}$$

and if $f_n$ is the $k$-th sample of the received signature, we define the signature vector as:

$$\mathbf{f}^T = \left\lfloor f_1, ..., f_{N_f} \right\rfloor \tag{2}$$

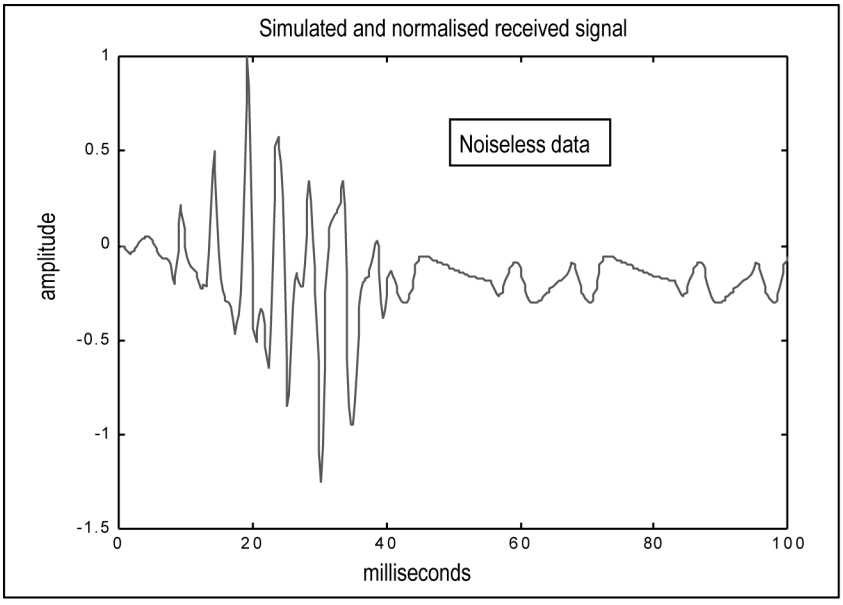


**Fig. 2** - Simulated received signal as convolution between signature and reflectivity sequence. No noise added.

Let's suppose that in correspondence with the time sample *j* an hypothetical reflecting surface is located, with an unknown reflection coefficient $b_j$. The amount of received signal $y_k$ due to the contribution of this single (virtual) reflector is

$$y_k = \begin{cases} 0 & \text{if} \quad k < 1 \\ b_j f_{k-j+1} & \text{if} \quad j \le k \le j + N_f \\ 0 & \text{if} \quad k > j + N_f \end{cases} \tag{3}$$

The next step of this problem statement is to suppose the presence of a virtual reflector with unknown reflection coefficient $b_j$. at each time sample *j*. It doese not matter if these is not (no matter about) any match between a real reflecting structure and this oversized virtual one: if at the *j*-th time sample there is no real reflecting spot, this will ideally result in a zero reflection coefficient of the corresponding virtual reflector. It is worth noting that no bond is imposed, at the moment, on the values of the reflection coefficients, but the inclusion of any constrain will have a deep impact on the design of the deconvolution algorithm. The reflection sequence can therefore be described by an $N_t$ dimensional column vector **b** of all the reflection coefficients (virtual and unknown), each relative at a different time sample.

$$b^T = \left[ b_1, ..., \; b_{Nt} \right]. \tag{4}$$

To achieve useful form of the received signal **y**, we define and introduce the $N_t \times N_t$ matrix of signatures **Φ** as follows:

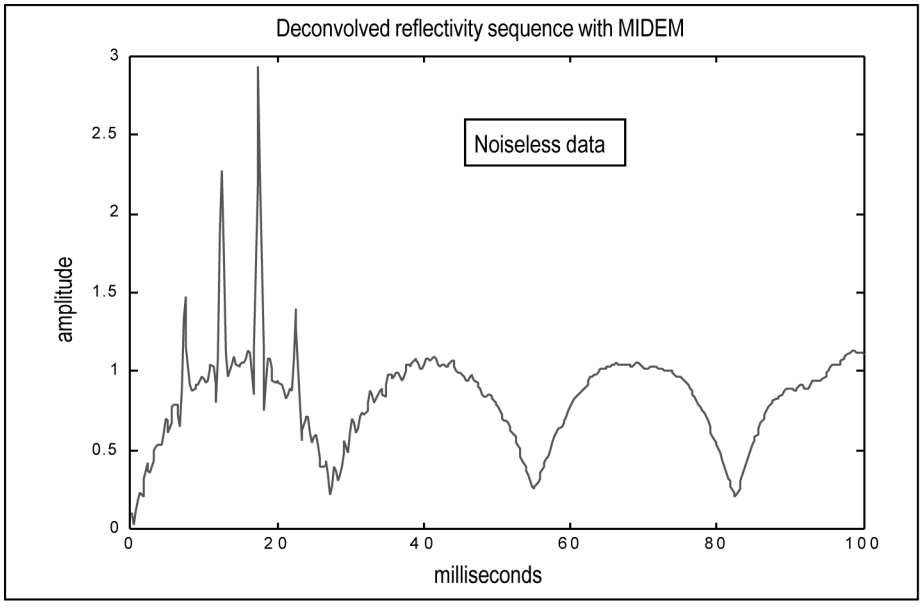


**Fig. 3** - Reflectivity sequence estimated by MIDEM in a noiseless case.

$$\Phi = \begin{bmatrix} \mathbf{f}^T & 0 & \square & 0 \\ 0 & \mathbf{f}^T & \square & 0 \\ \square & \square & \square & \square \\ 0 & \square & 0 & \mathbf{f}^T \\ 0 & \square & 0 & f_1,\ldots,f_{(N_f-1)} \\ 0 & \square & 0 & f_1,\ldots,f_{(N_f-2)} \\ 0 & \square & 0 & \square \\ 0 & \square & 0 & f_1 \end{bmatrix} \tag{5}$$

where **f** is the column vector of the known signature, given by Eq. (2). The physical meaning of the matrix $\mathbf{\Phi}$ is straightforward: the $n$-th row represents the signal that should have been received in presence of a single virtual reflector at the $n$-th time sample with unitary reflection coefficient.

The received signal **y**, at each time sample $j$, is the overlap of all reflections on every virtual reflecting surface, each weighted with the corresponding reflection amplitude $b_j$; that is the overlap of all the contributions given by Eq. (3). We have therefore the following closed form for the received signal vector

$$\mathbf{y} = \Phi^T \mathbf{b} \tag{6}$$

The above relation translates the problem statement into the classical form of a vector

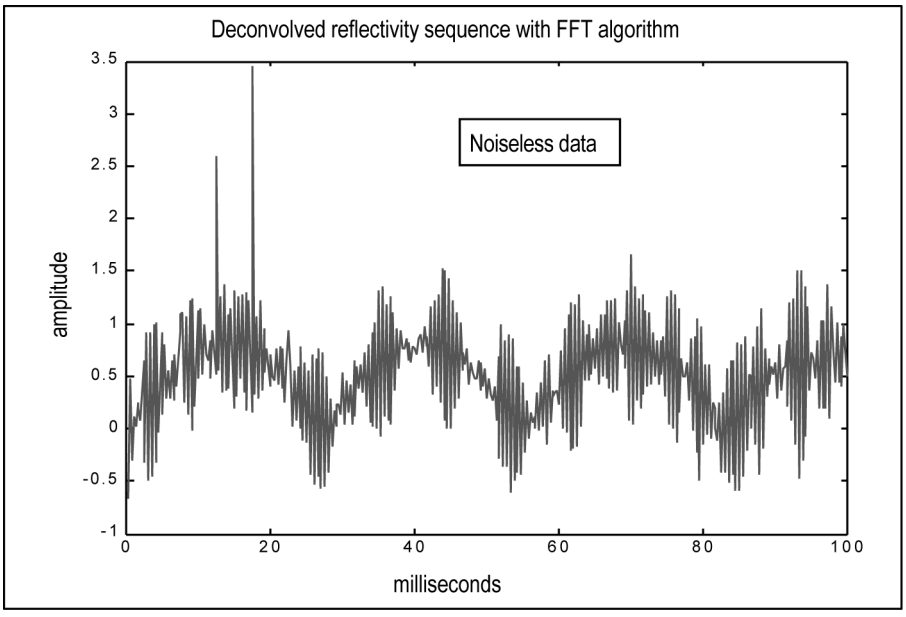


**Fig. 4** - Reflectivity sequence estimated by spectral inversion method of Eq. ( 9 ) with perfect and noiseless data.

transform into an $N_t$ dimensional hyperspace and our objective is to estimate the sequence **b** of the virtual reflection coefficients, once **y** and **f** (and therefore **Φ**) are known. Many reasons (noise, estimation mismatch, roundoff errors, ill-conditioning) make Eq. (6) as having no solution, therefore we have to steer our attention towards finding an estimate $\hat{\mathbf{b}}$ of the unknown vector, minimising some suitable performance function, usually the mean squared error between estimated and received signal

$$\min_{\mathbf{b}\in N_t} |\mathbf{e}|^2 = \min \left| \mathbf{y} - \mathbf{\Phi}^T \mathbf{b} \right|^2 \qquad (7)$$

This kind of problem belongs to the optimisation classes of algorithms, where many different processing techniques are available. The next section will synthetically span some of these optimisation/inversion techniques, trying to identify the best suited for the specific problem we are handling.

## 3. Inversion and optimisation criteria overview

We can roughly split all the families of unconstrained optimisation algorithms into direct open-loop and iterative closed-loop ones. At the state of the art, there is no optimum general approach, but each algorithm may result best suited within some classes of problems. To restrict the research area, we must remember that we have to face a problem with a very high number of

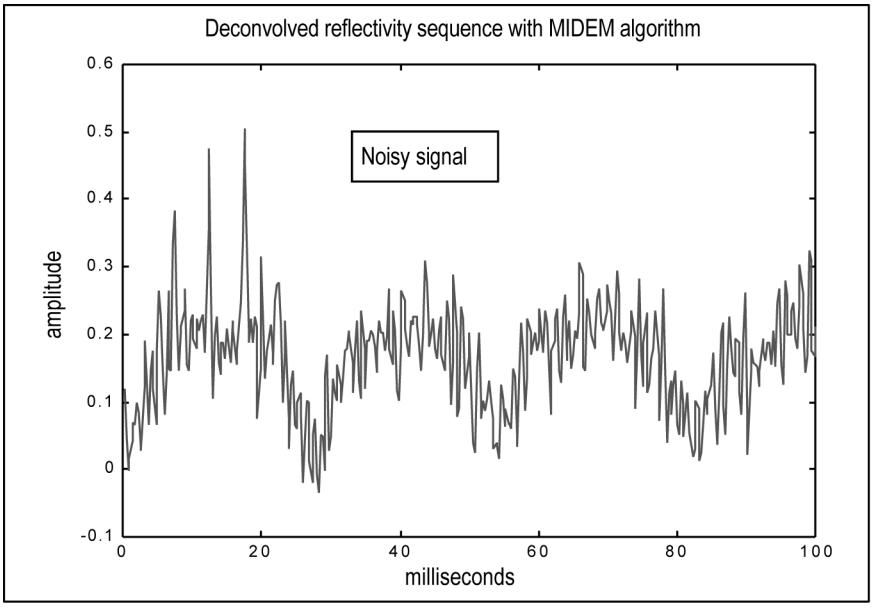


**Fig. 5** - Reflectivity sequence estimated by MIDEM in presence of uniform gaussian white noise with σ=0.1 added to

unknowns (typically several thousands) and with a strong presence of noise. Furthermore, it is also desirable to choose, among all the well behaving algorithms, the one(s) best tailored for the simplest computer coding.

Starting the overview with the direct methods, one classical solution of Eq. (7), and maybe the oldest one, is the deterministic normal equation (Lawson, Hanson, 1974) that, in this specific case gives the least mean squares unique[1] solution

$$\mathbf{b} = (\Phi\Phi^T)^{-1}\Phi\mathbf{y} \tag{8}$$

but from the structure of $\mathbf{\Phi}$ as defined in Eq. (5) and from its huge dimensionality it is easy to guess that the inversion of $\mathbf{\Phi\Phi}^T$ is a very ill-conditioned problem (Strang, 1980). For instance, 1 sec. of signal sampled at 4 kHz should involve the inversion of a 4000 by 4000 matrix. The system theory equivalence of Eq. (8) is that the received signal is the convolution between the emitted signature and the impulse response of the propagation medium, therefore, just formally, an inversion in the frequency domain to get a direct estimate of *b(t)* using an inverse

[1] The solution is unique only if invertibility of $\mathbf{\Phi\Phi}^T$ is ensured; this happens only if nullity of $\mathbf{\Phi}$ is zero; in such a case, recalling Eq. (10)

$$\mathbf{y} = \mathbf{\Phi}^T\mathbf{b}$$

and then, calculating the powwer of the received signal, we have

$$|\mathbf{y}|^2 = (\mathbf{\Phi}^T\mathbf{b})^T\mathbf{\Phi}^T\mathbf{b} = \mathbf{b}^T(\mathbf{\Phi\Phi}^T)\mathbf{b} > 0$$

therefore $\mathbf{\Phi\Phi}^T$ is a positive definite Hermitian form and all its principal minors are positive, therefore $\mathbf{\Phi\Phi}^T$ is non-singular and its inverse exists.

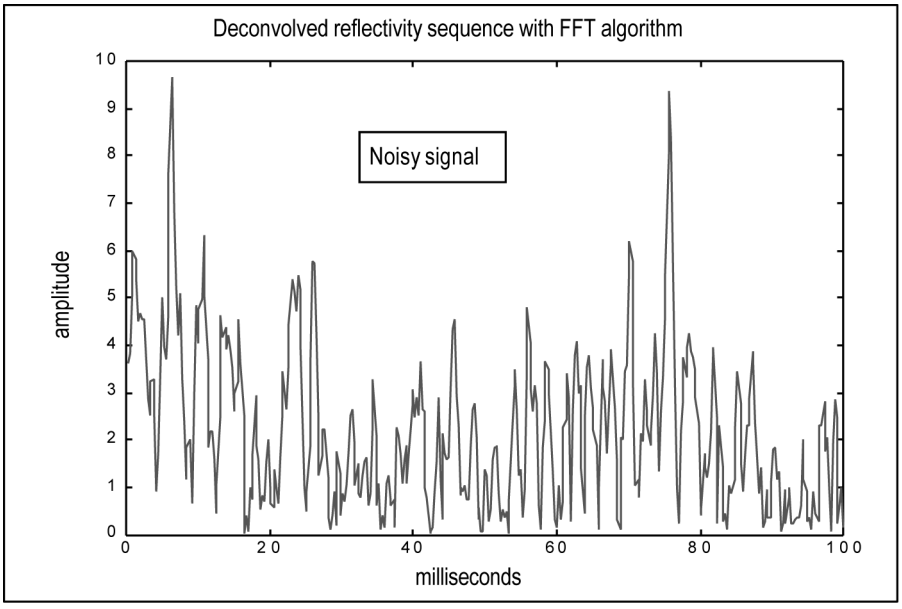


**Fig. 6** - Reflectivity sequence estimated by the spectral inversion method of Eq. ( 9 ) with uniform gaussian white noise with $\sigma$=0.1 added to the signal shown in Fig. 2.

transform is possible

$$y(t) = f(t) * b(t) \Leftrightarrow b(t) = F^{-1}\left[\frac{F[y(t)]}{F[f(t)]}\right] \tag{9}$$

where $F[\ldots]$ is the Fourier transform. In Eq. (9) a strong numerical instability is related to the zeroes of the signature spectrum; furthermore, the above relation also suffers noise hypersensitivity. As these direct approaches are basically equivalent, it can be noticed that the bad condition number of $\mathbf{\Phi\Phi}^T$, the zeroes of the signature spectrum or the poles in a polynomial division (another way to describe the deconvolution), are only formally different, but they merely represent the same analytical problem. Generally speaking, the high dimensionality and the presence of pole-like points, suggests a limited applicability of the direct inversion approach above.

Usually, the iterative optimisation algorithms are to be preferred to direct solutions of problem (7) in presence of high number of unknowns and in presence of noise. A wide range of algorithms are available to get an estimate of the solution of Eq. (7): the simplex search (Nelder and Mead, 1987) that just uses the function values, the steepest descent (Murray, 1972) for which the gradient vector is used, the Newton or quasi-Newton approaches (Mathworks inc, 1998) where the Hessian and Jacobian matrix evaluation is needed, the recursive least squares (RLS) (Cioffi and Kailath, 1984) that is the deterministic counterpart of the state-space Kalman filter theory. It is well known that none of all the iterative approaches can be considered as the best in an absolute sense: case by case, one or two algorithms can be a better compromise among desired

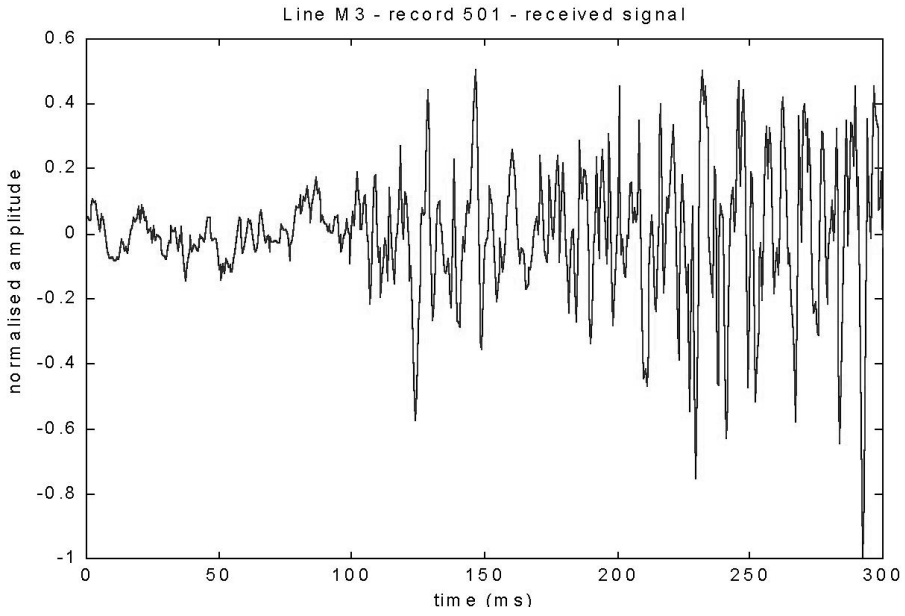


**Fig. 7** - Section of received signal.

accuracy, achievable accuracy (because of noise), effectiveness, computational load, computer coding simplicity and robustness. Coming back to the real nature of the problem we are facing, the measured signal is a single channel seismic record and it is far from being noiseless; each shot record is made of by a couple of (or more) thousands of samples, that is the number of the unknowns to estimate on each shot, meanwhile each seismic line can contain one or two thousand shots. This implies the rejection of the high accuracy family of algorithms (like Newton-like) because their computational charge and complexity have no proportional counterpart in achievable accuracy when compared with simpler gradient search minimisation. Therefore, in order to solve Eq. (7) with the aforesaid signal features, a steepest descent approach seems an acceptable compromise among feasibility, simplicity and effectiveness. As an alternative, an RLS state-space approach could be possible, but a significant feature of the steepest descent is its simplicity: it is slower, but doesn't require any matrix inversion and is designed to find the minimum of an error performance surface without the knowledge of the surface itself. Indeed, it is the simplicity that has made it a standard against which other algorithms are benchmarked.

## 4. Design of the unconstrained MIDEM algorithm

Once focused on the way of getting an estimate of **b** with a self-learning, gradient search, iterative approach, the algorithm must be designed to converge toward an optimum solution in a minimum squares sense, that is a convergence driven by the minimisation of the cost (error)

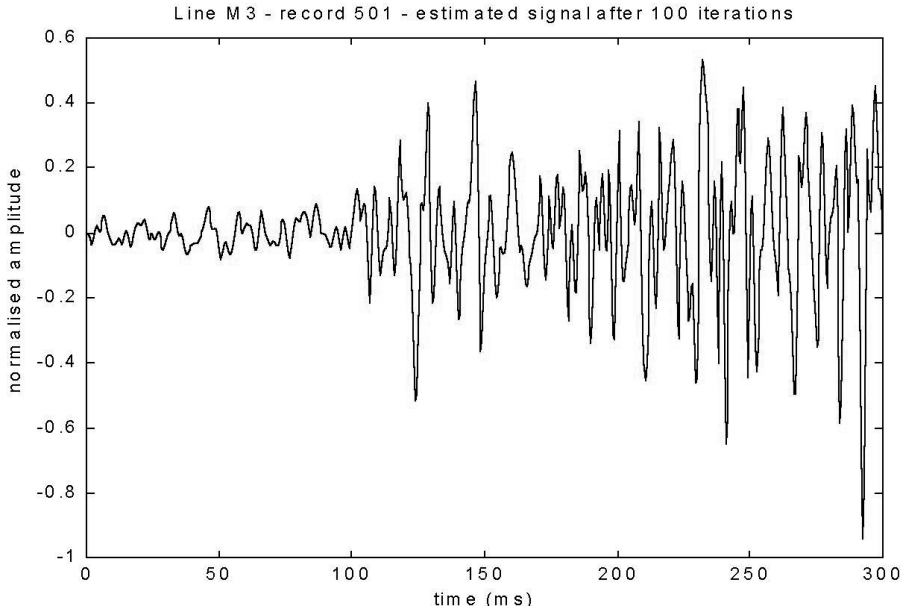


**Fig. 8** - Estimated signal after 100 iterations.

energy at each step of the iterations. Let's suppose that at the *n*-th iteration, an estimate $\hat{\mathbf{b}}$ *(n)* of the virtual reflections vector has been done; because of Eq. (6) there must be a corresponding estimate $\hat{\mathbf{y}}$ *(n)* of the signal that should have been received:

$$\hat{\mathbf{y}}(n) = \Phi^T \hat{\mathbf{b}}(n). \tag{10}$$

This means that at every iteration we have an $N_t$ dimensional error vector given by:

$$\mathbf{e}\,(n) = \hat{\mathbf{y}}(n) - \mathbf{y} = \Phi^T \hat{\mathbf{b}}(n) - \mathbf{y} \tag{11}$$

at the *n*-th iteration, there is an overall amount of error power *Q(n)* given by:

$$\begin{aligned} Q(n) &= \mathbf{e}^T(n)\mathbf{e}(n) = \\ &= \mathbf{y}^T\mathbf{y} + \hat{\mathbf{b}}^T(n)\Phi\Phi^T\hat{\mathbf{b}}(n) - \hat{\mathbf{b}}^T(n)\Phi\mathbf{y} - \mathbf{y}^T\Phi^T\hat{\mathbf{b}}(n) \end{aligned} \tag{12}$$

that can be considered as the error-performance surface of this minimisation problem. Comparing Eq. (12) with Eq. (10) it can be noted that *Q(n)* is a quadratic function of vector $\hat{\mathbf{b}}$ *(n)*, therefore *Q(n)* represents a quadratic error surface in an $N_t$ dimensional hyperspace and has an unique absolute minimum, but there is no unique approach to estimate all the (thousands) components of $\hat{\mathbf{b}}$ *(n)* that jointly minimise *Q(n)*. Following the classical philosophy of the steepest descent (Haykin, 1986), we can try - starting from an arbitrary $\hat{\mathbf{b}}$ (0) - to estimate these components in an iterative way, modifying the estimate of $\hat{\mathbf{b}}$ *(n)* at each iteration with a

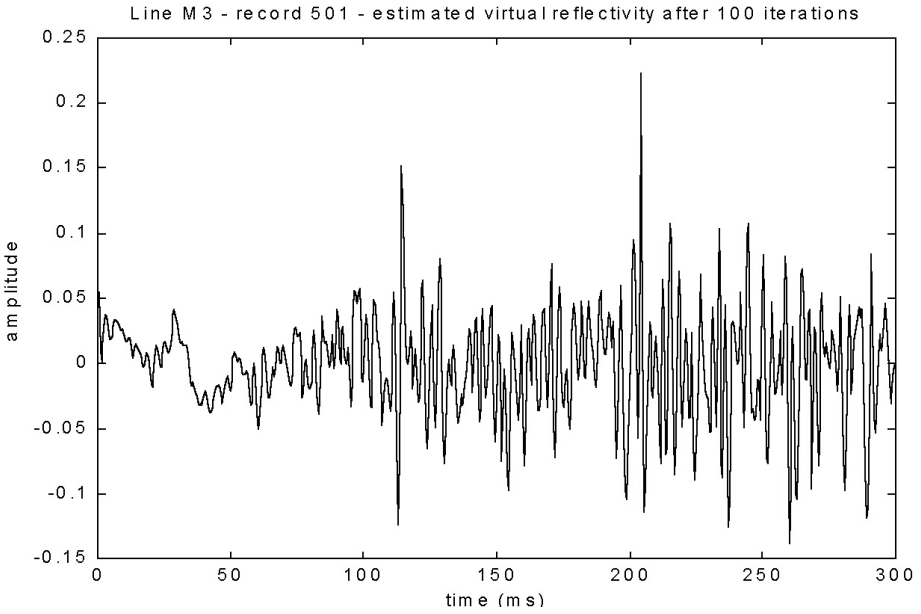


**Fig. 9** - Virtual reflections estimated after 100 iterations.

percentage μ of the local gradient vector to the error surface *Q(n)*, using the following recursive relation

$$\hat{\mathbf{b}}(n+1) = \hat{\mathbf{b}}(n) - \frac{1}{2}\mu\left(\frac{\partial Q(n)}{\partial \hat{\mathbf{b}}(n)}\right) \tag{13}$$

From Eqs. (12) and (11), after some little algebra, it is easy to derive the following closed form for the local gradient vector

$$\frac{\partial Q(n)}{\partial\ \hat{\mathbf{b}}(n)} = 2\Phi\Phi^T\hat{\mathbf{b}}(n) - 2\Phi\mathbf{y} = 2\Phi\mathbf{e}(n) \tag{14}$$

That, at the minimum point, gives the same result reported for the deterministic normal Eq. (8). We can now use the above equality in the recursion formula (13), thus achieving a very simple result for the iterative update of the virtual reflections vector

$$\hat{\mathbf{b}}(n+1) = \hat{\mathbf{b}}(n) - \mu\Phi\mathbf{e}(n). \tag{15}$$

It is worth noting that the estimate at the n+1-th iteration depends only on the previous estimate and on the previous error vector, therefore the recursion (15) doesn't suffer any error propagation through the increasing number of iterations, a typical problem of other step-

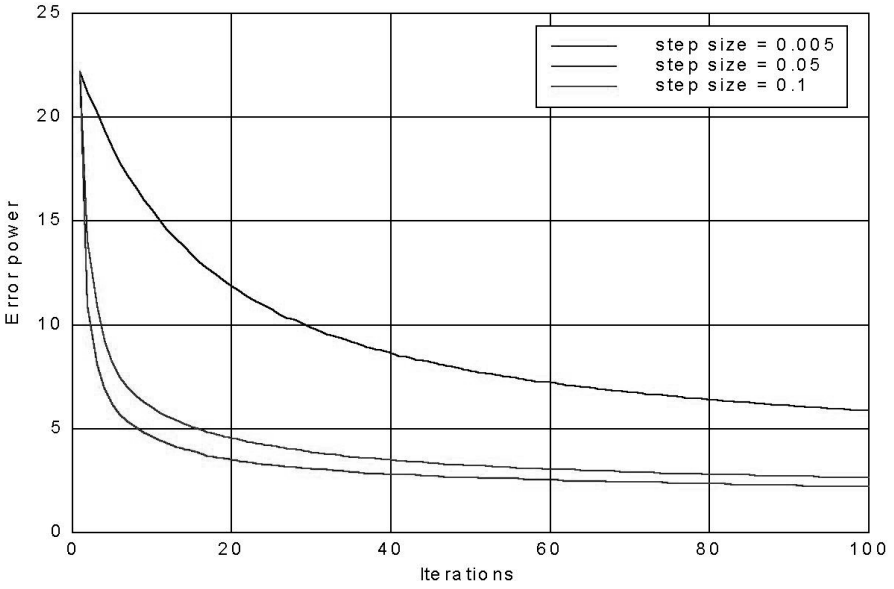


**Fig. 10** - Learning curve (normalised error power) during adaption process.

projection-gradient based minimisation algorithms (Ljung, 1984).

The algorithm design still lacks a discussion about the choice of the step size parameter $\mu$, as its value deeply impacts on convergence and accuracy properties of the algorithm. The bonds on step size coming from the general convergence condition are easy to handle and can be found in open literature (Haykin, 1986); here, the conditions for this specific problem have been briefly adapted and summarised in Appendix A. Once the convergence criteria has been satisfied, the step size parameter $\mu$ can be variable or fixed. A fixed step size will result in a lower computational load within each iteration, but in a slightly higher number of iterations; a variable step size increases both cost and complexity. In running the algorithm we preferred code simplicity, choosing a fixed step size, assigning its value as the one ensuring the faster convergence on a random set of shots.

Once the convergence has been reached, we can give the value of the residual error energy (Ljung, 1984; Haykin, 1986), a measure of the asymptotic behaviour of the algorithm:

$$\lim_{n\to\infty} Q(n) = \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\Phi(\Phi\Phi^T)^{-1}\Phi\mathbf{y}$$

That, looking at Eqs. (12) and (14), equals exactly the minimum value of the error-performance surface. This implies that the approach used is efficient, and no other approach, based on error energy minimisation, can exhibit a better asymptotic performance. This holds, as far as we are concerned, for the asymptotic behaviour, but convergence speed can be increased using an RLS algorithm or conjugate gradient instead of the simpler steepest descent. But once

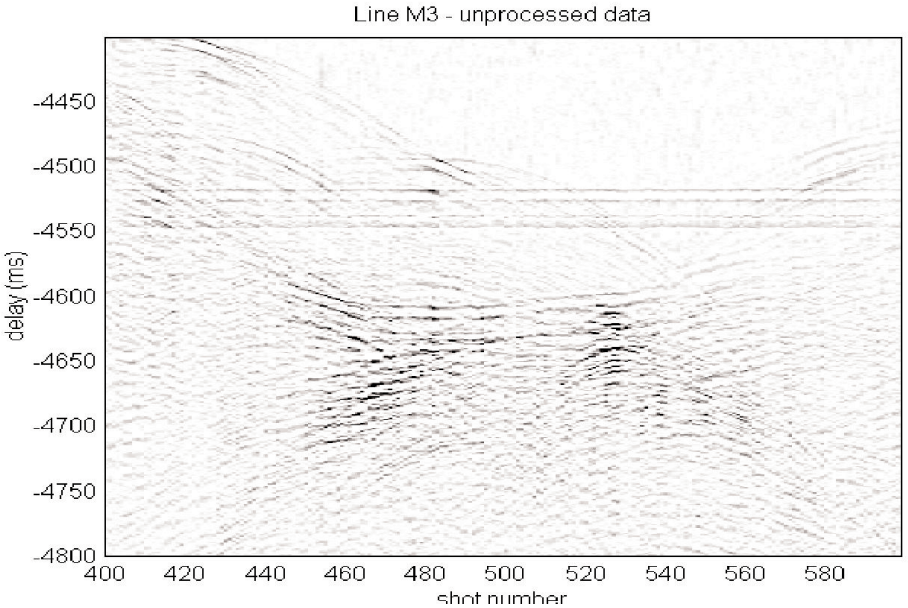


**Fig. 11** - Line M3 - gray scale representation of 200 unprocessed shots.

again the simplicity of the approach followed has been preferred, also because, standing at the aforesaid asymptotic efficiency, we do not expect any dramatic jump in minimisation capabilities. If to this we add that in running the algorithm on real data, the learning curves (error power vs. iterations) showed that convergence is usually attained within the first 100 iterations, it seems that the effort expended in faster implementations would not really be cost-effective.

Lastly, it is worth dropping two lines about computational cost reduction. Eqs. (11) and (15) include matrix multiplication, but, looking at the definition of $\mathbf{\Phi}$ given by Eq. (5), we discover that the signature matrix is mainly composed of null elements and the non zero ones are all gathered around the main diagonal. Furthermore, it can be noticed that there is no need to allocate such an huge matrix as $\mathbf{\Phi}$ in memory, as all its columns are made up of the same vector $\mathbf{f}$, located at different position lags. These considerations may suggest different ways and tricks to save computational cost.

## 5. Introduction of amplitude constraints

As already stated in the problem statement section, no specific hypotheses have been introduced on the possible range along which the numeric values of all the reflection coefficients have to fall; therefore, in principle, every value, from minus to plus infinity, is allowed. In this section we want to remove this degree of freedom, modifying the algorithm to allow only reflection coefficients within the open range ]0, 1[. There is no unique way of achieving this constrained minimisation; many efforts have been steered toward direct methods based on the Moore-Penrose pseudo-inverse (Lawson and Hanson, 1974) or toward open-loop iterative

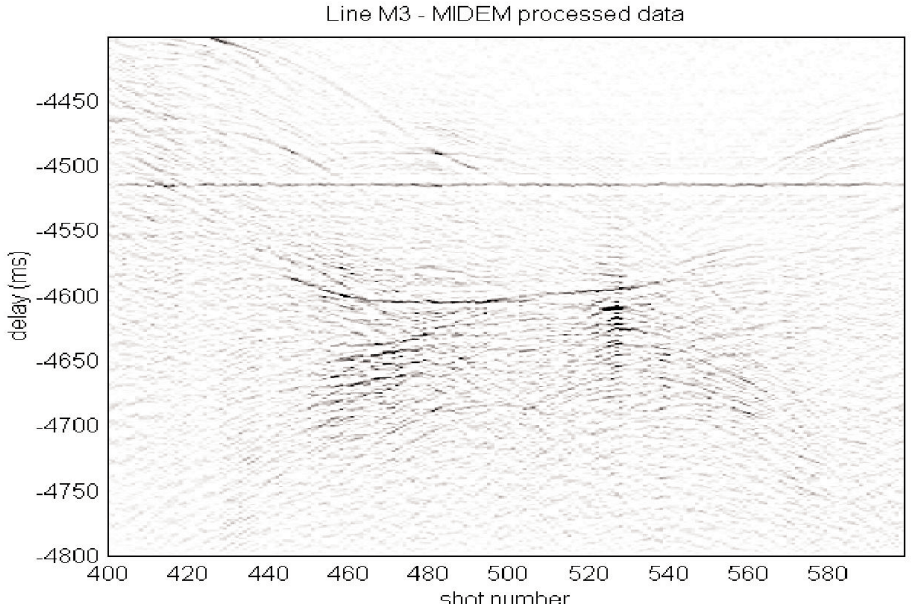


**Fig. 12** - Line M3 - gray scale representation of 200 processed shots.

algorithms (Zimmer, 1987; Berrou and Boullon, 1983); here we choose to search the best mean squares estimate of $\hat{\mathbf{b}}(n)$ that is a sigmoidal function of a dumb vector $\hat{\mathbf{x}}(n)$

$$\hat{\mathbf{b}}(n) = \frac{e^{\hat{\mathbf{x}}(n)}}{1 + e^{\hat{\mathbf{x}}(n)}} \tag{16}$$

in this way all the $\hat{\mathbf{b}}(n)$ will lie between 0 and 1, regardless the value of $\hat{\mathbf{x}}(n)$. Of course the expression (12) for the error energy still holds, but it is now a bowl-shaped function of $\hat{\mathbf{x}}(n)$, therefore the time-update recursion for a gradient-based minimum search has to be

$$\hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n) - \frac{1}{2}\mu\left(\frac{-\partial Q(n)}{\partial \hat{\mathbf{x}}(n)}\right) \tag{17}$$

using Eq. (16) and making the vector derivatives, after some algebra, the recursion for $\hat{\mathbf{x}}(n)$ assumes the form

$$\hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n) - \mu\left(\frac{-\hat{\mathbf{b}}^2(n)}{e^{\hat{\mathbf{x}}(n)}}\right)\Phi\mathbf{e}(n) \tag{18}$$

where the new estimate of $\hat{\mathbf{x}}(n)$ depends only on the previous estimate and on the error vector. Through this recursion, at every iteration, Eq. (16) will give the corresponding bonded values of the virtual reflections.

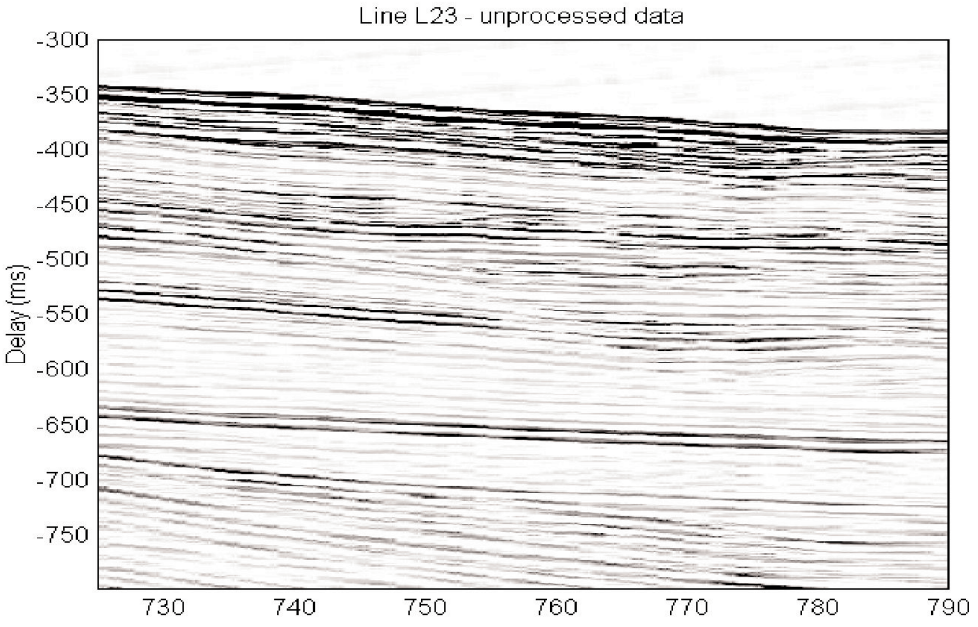


**Fig. 13** - Line L23 - gray scale representation of 66 unprocessed shots.

## 6. Summary of MIDEM algorithm

| | | |
|---|---|---|
| Output | Virtual reflections vector estimate | $\hat{\mathbf{b}}\,(n) = [b_1\,(n), \ldots, b_{N_t}\,(n)]$ |
| Input | Signal vector | $y^T = [y_1, \ldots, y_{N_t}]$ |
| | Signature vector | $\mathrm{f}^T = [f_1, \ldots, f_{N_f}]$ |
| | Initialisation | $\hat{\mathbf{b}}\,(0) \approx \mathbf{0}$ (suggested) |
| *n*-th step recursion | | |
| | Unconstrained | Constrained in ]0, 1[ |
| Constraint | $\hat{\mathbf{b}}\,(n) = \hat{\mathbf{b}}\,(n)$ | $\hat{\mathbf{b}}\,(n) = \dfrac{e^{\hat{\mathbf{x}}(n)}}{1+e^{\hat{\mathbf{x}}(n)}}$ |
| Estimated signal | $\hat{\mathbf{y}}\,(n) = \boldsymbol{\Phi}^T \hat{\mathbf{b}}\,(n)$ | $\hat{\mathbf{y}}\,(n) = \boldsymbol{\Phi}^T \hat{\mathbf{b}}\,(n)$ |
| Estimation error | $\mathbf{e}\,(n) = \hat{\mathbf{y}}\,(n) - \mathbf{y}$ | $\mathbf{e}\,(n) = \hat{\mathbf{y}}\,(n) - \mathbf{y}$ |
| Update | $\hat{\mathbf{b}}\,(n+1) = \hat{\mathbf{b}}\,(n) - \mu\,\boldsymbol{\Phi}\,\mathbf{e}\,(n)$ | $\hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n) - \mu\left(\dfrac{\hat{\mathbf{b}}^2(n)}{e^{\hat{\mathbf{x}}(n)}}\right)\boldsymbol{\Phi}\mathbf{e}(n)$ |

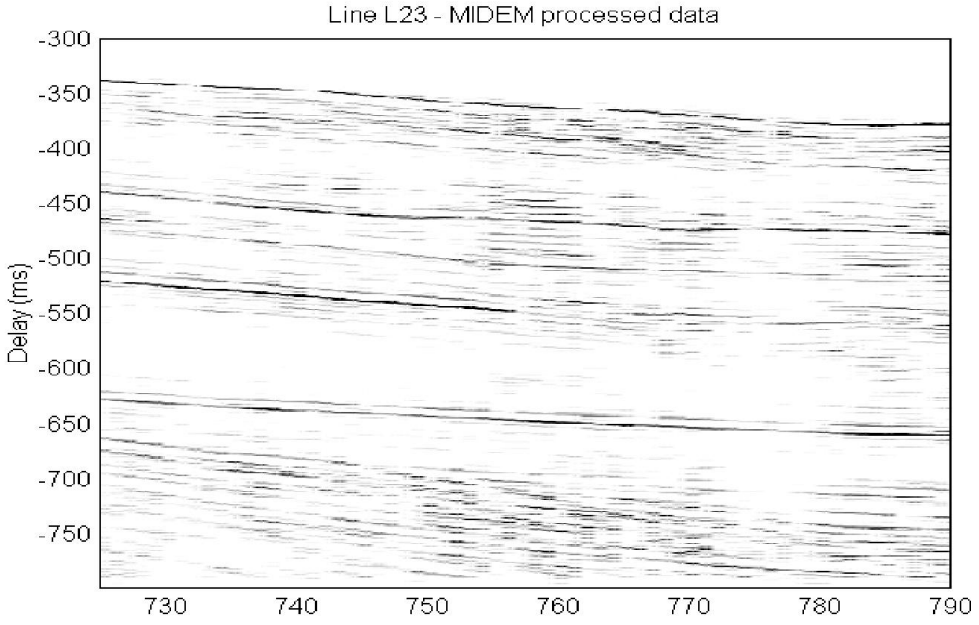


**Fig. 14** - Line L23 - gray scale representation of 66 processed shots.

## 7. MIDEM performances on synthetic data

In a few words, pointless checking the algorithm on simulated data when real survey data are available could seem; nevertheless, it should be pointed out that when handling real data, even if the deconvolution results suggest a well behaved processing, basically, there is a lack of knowledge about the true reflectivity sequence and no comparison can be made between estimated and true results. Therefore, to check the performances of the algorithm, a controlled environment with computer generated signals is first needed. In order to stress the algorithm, the reflectivity test sequence has been chosen as a superposition of four single point reflectors of different amplitudes with 5 ms time separation, a value well below the 21 ms of the signature duration, thus involving an intrinsic sub-resolution nature of this discrete reflecting structure. Furthermore, a continuous structure of reflectivity of variable amplitude has been overlapped, resulting in the synthetic reflectivity sequence shown in Fig. 1. In the same figure the normalised signature used in the processing has also been reported; this signature is not a mathematical function, but the real signature measured during sea trials and it is the same signature that will be used in the next section, on MIDEM performance on real data. Just to avoid a graphical mutual jamming of the lines, the signature has a –1 offset. The received signal is shown with normalised amplitude in Fig. 2 and it is possible to notice both the effect of the continuous, modulated reflectors and the loss of any information about the time-space structure of the four discrete reflectors, packed below the resolution capability of the signature itself. All long the synthetic received signal a strong discrete-continuous interference process takes place and no

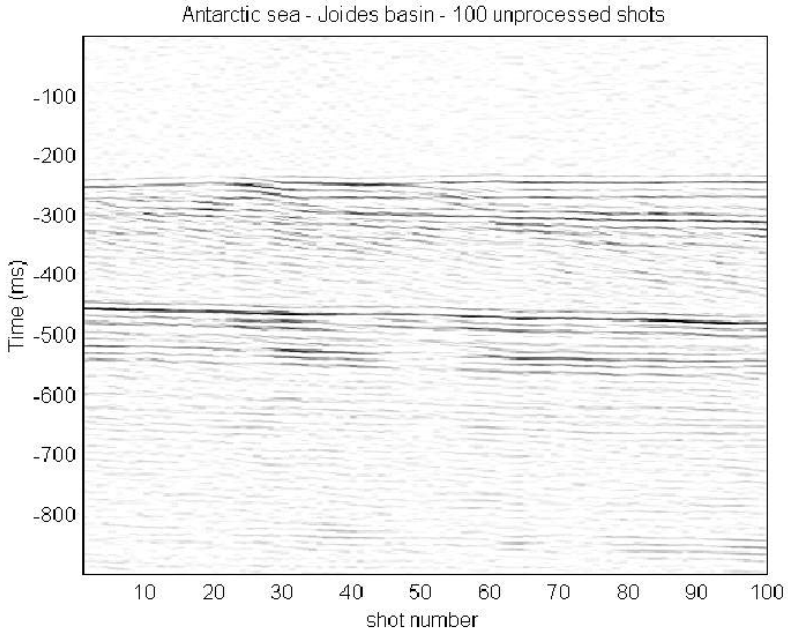


**Fig. 15** - Antarctic sea - gray scale representation of 100 unprocessed shots.

part of this signal resembles the shape of the original reflectivity sequence of Fig. 1. The MIDEM algorithm then has been applied on this signal abd the resulting estimate, is reported in Fig. 3 where both the discrete and the continuous part of the reflectivity sequence have been correctly deconvolved, in spite of the involved signature structure and the non trivial nature of the true reflectivity. To make a comparison, the result of the direct spectral inversion method outlined in Eq. (9) has also been reported in Fig. 4; as can be noted, the direct approach fails to detect two out of the four single discrete reflectors and also results in a somewhat more noisy, but similar deconvolution of that achieved by MIDEM in the continuous part. The noisy nature of the direct approach is not surprising, as it comes from the aforesaid numerical instability of the spectral inversion and even in a limit, perfect, noiseless case, the roundoff error and the finite precision achievable in the sine and cosine evaluation, causes an intrinsic algorithmic noise. The situation is strongly stressed by adding a white Gaussian noise with $\sigma = 0.1$ to the received normalised signal of Fig. 2; the output of MIDEM is reported in Fig. 5 and the result of the direct spectral inversion is shown in Fig. 6. It is easy to notice that the iterative deconvolution still exhibits a good robustness against this additional noise, whereas the direct method shows a total failure in the whole sequence. This preparatory analysis carried out on synthetic data allows us to note that the applicability of the MIDEM algorithm on real data can be really meaningful.

## 8. MIDEM performances on real survey data

The approach presented has been widely tested on real survey data collected during extensive

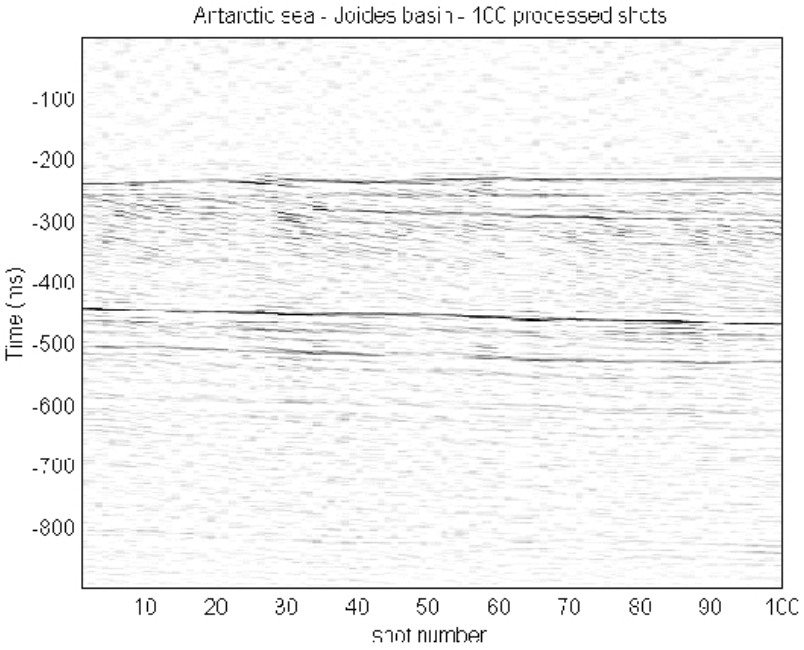


**Fig. 16** - Antarctic sea - gray scale representation of 100 processed shots.

sea trials carried out by the Istituto di Oceanologia of the Istituto Universitario Navale of Naples in the gulf of Naples area, in the Central Mediterranean and in the Antarctic Sea in 1990, 1991 and 1993.

The first data set comes from seismic line M3 of data collected during the MEDRIFF [2] campaign with the Bannok research vessel in 1993. The seismic source involved in this measurement was MEAS (Multiple electrodes Extended Array Sparker), a planar array of 36 equally spaced electrodes powered with a total of 15 kJ. The intrinsic resolution of the source is about 6 meters, but in some situations, the ship's hull reflection degraded this value down to about 25 m. Among the data collected in this campaign, line M3 (3290 shots; start: lat 35°12'54" N - long 21°34'23" E; end: lat 35°17'11" N - long 21°10'00" E) has been selected for our processing, because of the presence in that area of one deep high-salinity lake, a very interesting structure for a deconvolution test, as it consists of a single, perfectly planar, reflecting surface. The received signal has been filtered in a 60-600 Hz band and then digitalised with a 2 kHz sampling frequency.

To give an example of the learning process, the signal received at shot number 501 of the M3 line is presented in Fig. 7. Figure 8 reports the estimated signal after 100 iterations and the resulting sequence of estimated virtual reflections is reported in Fig. 9. As can be seen, the difference between the received and estimated signal is negligible, thus suggesting a good behaviour of the deconvolution process. To give an idea of the adaption process, the learning curve of the algorithm with increasing iterations, plotted for three different values of the step-size parameter $\mu$ has been reported in Fig. 10; it is possible to notice, as shown from theory (Appendix A), that the convergence of the algorithm is the superposition of several exponential decaying modes and only

[2] Contract MAS2CT92-0037. Final report July 1995.

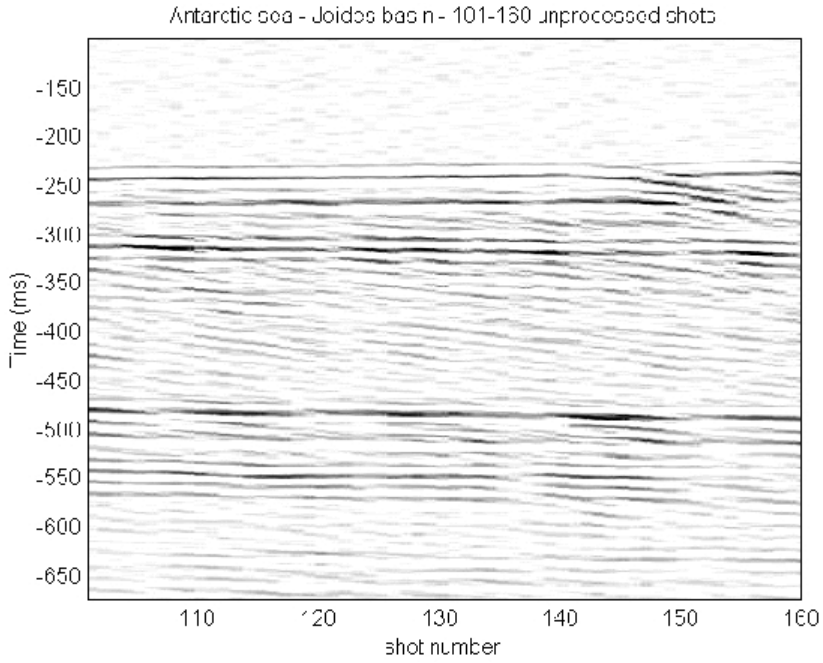


**Fig. 17** - Antarctic sea - gray scale representation of a detail of 60 unprocessed shots.

those related to the higher eigenvalues are dominant. It can also be noticed that, working on real data and with a real signature, the overall number of iterations before attaining a near steady-state is small enough to discourage the use of faster convergence algorithms.

An interpolated grey scale representation of 200 unprocessed shots selected on an M3 line in correspondence with this high salinity lake is reported in Fig. 11. From the plot, the broadening of the reflected signal from the lake interface, resulting in location ambiguity and information loss can be pointed out. In the same way the underlying structures are affected. All these data have been processed and the output of the algorithm is reported in Fig. 12, where the surface of the hyper-saline lake has been correctly deconvolved into a single horizontal line at the right location, with all the other structures of the unprocessed data reproduced in a sharper way, reducing the overall ambiguity, as expected from a well behaved deconvolution process.

The MIDEM algorithm has also been tested on another set of survey data collected by the Istituto di Oceanologia in the Gulf of Naples during the campaign “Naples Gulf 90” in 1990. The seismic line is L23 (start: lat 40°34.11' N - long 14°17.11' E; end: lat 40°46.16' N - long 14°04.71' E) for a total of 1846 shots. The received signal has been filtered in a 60-600 Hz band and then digitalised with a 4 kHz sampling frequency. The interpolated grey scale representation of 66 unprocessed shots of the L23 line has been reported in Fig. 13 and, in Fig. 14, the same set of data has been plotted after processing: as can be seen, in the deconvolved data most of the signal ambiguity has been removed, resulting in sharper and better resolved structures.

A third data set comes from a campaign carried out by the Istituto di Oceanologia of the Istituto Universitario Navale in the Antarctic Sea in 1990-91 in the area Joides basin area. The seismic shots location is lat 74°00.50', long 172°00.00'. The received signals have been

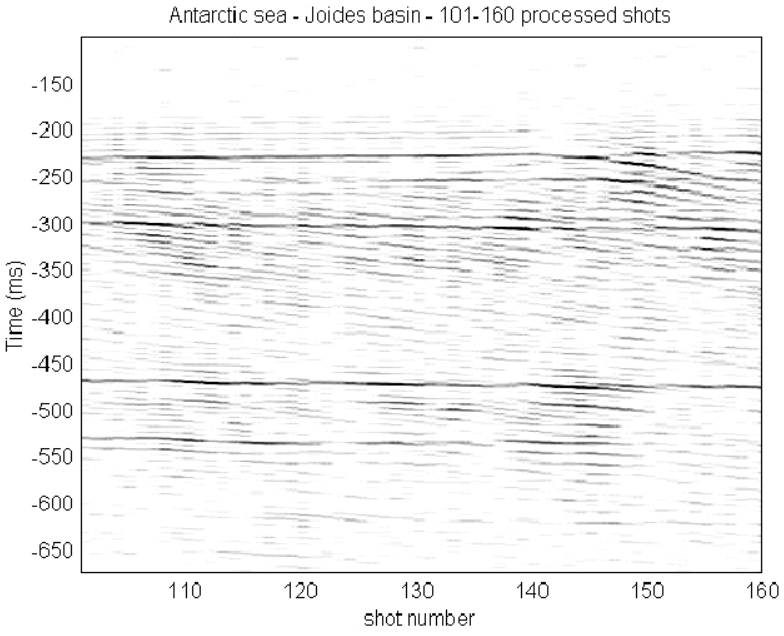


**Fig. 18** - Antarctic sea - gray scale representation of a detail of 60 processed shots.

analogically filtered in a 20-600 Hz band and then sampled at 4 kHz.

A gray-scale representation of the first 100 shots of the seismic line has been reported in Fig. 15, where the origin of the time scale, an offset of 1.1 s, has to be added to have the true depth from the surface. The same lines have been processed and in Fig. 16 the result has been plotted. Once again, a comparison between the two figures gives further encouragement, as the processed signals show a general enhancement in resolution and definition in the areas where the reflecting structures are present. In Fig. 15, most of the signal below –200 ms and beneath –600 ms is just noise, therefore, in order to highlight the signal structures alone, a section of signals belonging to the 60 next shots of those reported in Fig 15 has been presented in Fig. 17.

Fig. 18 reports the gray scale representation of the same signals after processing and the overall benefit in resolution and sharpness is clearly noticeable.

## 9. Conclusions

The aim of the present work has been to design an alternative approach for single-channel seismic signal deconvolution. The theoretical basis has been to use a self-learning algorithm based on a minimum-error iterative gradient search. The algorithm exhibits good numerical properties, it has a very simple structure and is well-suited for computer implementation. It has been extensively tested on both synthetic and real raw seismic data, showing good performances. The MIDEM deconvolution algorithm is not designed to increase the signal to noise ratio: it is tailored to increase the space-time resolution, being able to resolve interfaces or structures lying

within the pulse duration. The next steps in its development will be the massive processing of all the single channel seismic data collected by Ist. Univarsitario Navale of Naples in the Gulfs of Naples and Pozzuoli and in the Antarctic Sea. At the moment some further studies have been planned to improve MIDEM performances using a time-varying (depth-dependant) signature matrix and re-designing it using covariance spectral decomposition.

## References

Berrou J. L. and Boullon P.; 1983: *Une nouvelle methode haute resolution rapide et robuste*; Neuvieme Colloque sur le Traitement du Signal et ses Applications, Nice 16-20 May.

Berkhout A. J.; 1984: *Seismic resolution*. Handbook of Geophysical Exploration, Geophysical Press, **12,** 1-16, 58-75.

Cioffi J. M. and Kailath T.; 1984: *Fast, recursive-least-squares transversal filters for adaptive filtering*. In: IEEE Trans. Acoust., Speech, and Signal Processing, Vol. ASSP-32, pp. 304-337.

Haykin S.; 1986: *Adaptive filter theory*. In: N. J., T. Kailath (ed), Prentice Hall, Englewood Cliffs, ISBN 0-13-004052-5.

Hatton L., Worthington M. H. and Makin J.; 1988: *Seismic data processing: theory and pratice.* Blackwell Scientific Publications, Palo Alto, CA, ISBN 0-632-01374-5.

Lawson C. L. and Hanson R. T.; 1974: *Solving least squares problems.* Prentice Hall, NJ.

Ljung L.; 1984: *Analysis of stochastic gradient algorithms for linear regression problems.* IEEE Trans. Information Theory, vol. IT-30, Special Issue on Linear Adaptive Filtering, pp 151-160.

MathWorks Inc.;1997: MATLAB Optimization toolbox, Ver. 5.

Mc Quillin R., Bacon M. and Barclay W.; 1979: *An introduction to seismic interpretation.* Grahm & Trotman Ltd, London.

Murray W.; 1972: *Numerical methods for unconstrained optimization*. Academic Press, New York.

Nelder J. A. and Mead R.; 1987: *A simplex method for function minimization*. Computer Journal, **7**, pp. 308-313.

Parkes G. and Hatton L.; 1986: *The marine seismic source.* Reidel Publishing.

Robinson E. A.; 1983: *Digital seismic inverse methods.* With S. Treitel, R. A. Wiggins, P. R. Gutowski, Reidel Publishing .

Robinson E. A.; 1984: *Seismic inversion and deconvolution.* Handbook of Geophysical Exploration, Geophysical Press, Vol. 4a, pp. 223-243.

Strang G.; 1980: *Linear algebra and its applications.* Academic Press, NY.

Zimmer W. M. X.; 1987: SACLANT ASW Research Centre Report, SR 110, March.

## Appendix A: Convergence criteria

The recursive relation (15) gives no information about the convergence of MIDEM itself. To investigate this point, we introduce the vector

$$\mathbf{d}(n) = \hat{\mathbf{b}}(n) - \mathbf{b} \tag{A1}$$

that is the distance between the estimate and the true solution at the n-th iteration. The algorithm is said to be convergent *in mean* when

$$\lim_{n\to\infty} \mathbf{d}(n) = 0 \tag{A2}$$

therefore, our target is to find when Eq. (20) holds. Using Eqs. (15) and (11) into the definition (19) we have

$$\mathbf{d}(n+1) = \mathbf{d}(n) - \mu\Phi\left(\Phi^T\hat{\mathbf{b}}(n) - \mathbf{y}\right) \tag{A3}$$

where $\mathbf{I}$ is the identity matrix. As $\mathbf{\Phi\Phi}^{\mathrm{T}}$ is non-singular, its eigenvalue problem has an unique solution and the matrix $\mathbf{\Phi\Phi}^{\mathrm{T}}$ can be rewritten as

$$\Phi\Phi^{\mathrm{T}} = \mathbf{W}\Lambda\mathbf{W}^{\mathrm{T}} \tag{A4}$$

where $\mathbf{\Lambda}$ is the eigenvalue's diagonal matrix and $\mathbf{W}$ is the unitary ($\mathbf{W}^{\mathrm{T}}\mathbf{W}=\mathbf{I}$) transformation matrix, whose columns are the eigenvectors of $\mathbf{\Phi\Phi}^{\mathrm{T}}$. Left multiplying by $\mathbf{W}^{\mathrm{T}}$ and calling $\mathbf{q}(n) = \mathbf{W}^{\mathrm{T}}\mathbf{d}(n)$, we have

$$\mathbf{q}(n+1) - (\mathbf{I} - \mu\Lambda)\mathbf{q}(n) \tag{A5}$$

or equivalently

$$\mathbf{q}(n+1) - (\mathbf{I} - \mu\Lambda)^n\mathbf{q}(0) \tag{A6}$$

that, written for each component $k$, becomes

$$q_k(n) = (1 - \mu\lambda_k)^n q_k(0) \tag{A7}$$

From the above relation, it is obvious that each component of $\mathbf{q}(n)$ converges with a different speed and the sufficient rule to insure the convergence is

$$0 < \mu < \frac{1}{\lambda_k} \quad \text{for all} \quad k \tag{A8}$$