Asociación Argentina de Mecánica Computacional

Mecánica Computacional Vol XXXII, págs. 1701-1713 (artículo completo)
Carlos G. García Garino, Aníbal E. Mirasso, Mario A. Storti, Miguel E. Tornello (Eds.)
Mendoza, Argentina, 19-22 Noviembre 2013

A HYBRID STRATEGY BASED ON FAST ITERATIVE SHRINKAGE-THRESHOLDING ALGORITHM AND VERY FAST SIMULATED ANNEALING: APPLICATION TO THE PRESTACK SEISMIC INVERSE PROBLEM

Daniel O. Pérez[a], Danilo R. Velis[a] and Mauricio D. Sacchi[b]

[a]*Facultad de Ciencias Astronómicas y Geofísicas, UNLP, Paseo del Bosque s/n, 1900 La Plata; and CONICET, Argentina, dperez@fcaglp.unlp.edu.ar, velis@fcaglp.unlp.edu.ar*

[b]*Department of Physics, University of Alberta, Edmonton, Canada, msacchi@ualberta.ca*

Keywords: VFSA, FISTA, inversion, l_1-norm, prestack data

Abstract. With the purpose of characterizing the Earth subsurface, one of the objectives of the inversion of prestack seismic data is to determine contrasts between rock properties from the information contained in the variation of the amplitudes of the reflected compressional waves with the angle of incidence. This amplitude-versus-angle (AVA) variation can be described by various approximations to the so-called Zoeppritz equations, a set of non-linear equations that depend on the physical characteristics of the medium at each side of the interface where the compressional wave strikes. The coefficients of such approximations constitute AVA attributes that may provide important information about fluid content, a key issue for the characterization of hydrocarbon reservoirs. In this work we present a new inversion strategy to estimate efficiently and accurately high-resolution AVA attributes from prestack data. The proposed technique promotes sparse-spike reflectivities that, when convolved with the source wavelet, fit the observed data. Sparse solutions are desirable because they can be used to characterize significant and close reflectors more accurately than using conventional solutions. The inversion is carried out using a hybrid two-step strategy than combines Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) and Very Fast Simulated Annealing (VFSA). FISTA provides sparse solutions by minimizing both the misfit between the modeled and the observed data, and the l_1-*norm* of the solution. VFSA is an stochastic computational algorithm to finding near-optimal solutions to hard optimization problems. At the first stage, FISTA sparse-solutions provide an estimate of the location in time of the main reflectors, information that is subsequently used as an initial guess for the second stage, where accurate reflectivity amplitudes are estimated by solving a more stable overdetermined inverse problem. The second stage also involves the use of VFSA for tuning the location in time of the main reflectors and the source wavelet. FISTA does not require the inversion of matrices in explicit form. At each iteration only matrix-vector multiplications are involved, making it easy to apply, economic in computational terms, and adequate for solving large-scale problems. As a result, the FISTA+VFSA strategy represents a simple and cost-effective new procedure to solve the high-resolution AVA inversion problem. Results on synthetic data show that the proposed hybrid method can obtain high-resolution AVA attributes from noisy observations, even when the number of reflectors is not known *a priori* and the utilized wavelet is inaccurate, making it an interesting alternative to conventional methods.

1 INTRODUCTION

The inversion of prestack seismic data has been long recognized as a useful technique for the characterization of hydrocarbon reservoirs. Like any other inversion problem it is based on an appropriate set of observations, and a general knowledge of the desired solution. The objective of the prestack seismic inverse problem is to make inferences about the contrasts between rock properties such as P and S-wave propagation velocities and rock densities. These contrasts are related to the variation of the amplitudes of the reflected compressional waves as a function of the angle of incidence, for a given reflecting interface (or reflector) that separates two different media. This variation, which is known as amplitude versus angle variation (AVA) or amplitude versus offset variation (AVO), can be described by the Zoeppritz equations (Zoeppritz, 1919; Yilmaz, 2001). Zoeppritz equations depend on the angle of incidence and six independent elastic parameters, three on each side of the reflector. The high non-linearity of these equations makes them impractical for their use in inversion problems and data interpretation. Because of this, several linear approximations have been developed by various authors over the last decades (Aki and Richards, 1980; Shuey, 1985; Fatti et al., 1994). These linear approximations depend on the angle of incidence and various coefficients related to the physical properties of the medium. The most common approximations, such as the ones developed by the authors indicated above, only require two or three coefficients to approximate the Zoeppritz equations with an acceptable error for angles of incidence smaller than the critical angle. The coefficients of such approximations, which are the objective of the inversion, constitute AVA attributes that may provide important information about fluid content, a key issue for the characterization of hydrocarbon reservoirs.

The proposed inverse problem is multidimensional and ill posed. In practice, seismic records contain many reflectors and therefore there will be as many sets of AVA coefficients as there are reflectors within the selected record length. Also, uncertainties in the seismic wavelet used in the forward model increase the no-uniqueness of the inverse problem. The main problem is that not only there exist several sets of coefficients that honor the data equally well, but some of them might exhibit a huge norm and thus are meaningless. Also, these solutions are strongly affected by the omnipresent noise in the observed data. To alleviate these difficulties the space of possible solutions must be constrained using an adequate regularization or *a priori* information. Sparseness is a property that can be incorporated as *a priori* constraint via regularization (Taylor et al., 1979; Oldenburg et al., 1983; Sacchi, 1997). Sparse solutions are desirable because they can be used to characterize significant and close reflectors more accurately that using conventional solutions (Debeye and van Riel, 1990).

Many authors have studied the sparse-spike AVA inversion with very interesting results. In the works of Downton and Lines (2003), Misra and Sacchi (2008) and Alemie and Sacchi (2011) sparseness is obtained through a Bayesian approach, using an appropriate long-tailed *a priori* probability distribution. Pérez and Velis (2011) obtained sparse solutions by fixing the number of reflectors and finding the model that best adjusts the observed data using a global optimization algorithm. Further, Pérez et al. (2012) obtained sparse solutions by minimizing, simultaneously, the l_1-*norm* of the solution and the misfit between the modeled and the observed data using an iterative shrinkage-thresholding algorithm.

In this work we propose an hybrid algorithm by combining an iterative shrinkage-thresholding algorithm known as Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) (Beck and Teboulle, 2009) and a global optimization technique known as Very Fast Simulated Annealing (VFSA) (Ingber, 1989). FISTA is used, in a first stage, to find an initial estimation of the number of

reflectors and their positions. Then, in a second stage, VFSA is used to improve the solution given by FISTA, tuning the seismic wavelet and adjusting, if necessary, the position of the reflectors at the same time. In each iteration of the VFSA, the amplitudes of the reflectors are estimated using least-squares (LS). This LS strategy was also used by Velis (2008) in the context of sparse-spike deconvolution of postack data. It is worth mentioning that FISTA only uses matrix-vector multiplications and no matrix inversion. The initial solution provided by FISTA allows one to increase the computational efficiency of a versatile, but very expensive algorithm like VFSA. The use of a global optimization technique such as VFSA is necessary because the inversion problem proposed in this work is highly non-linear in terms of the position of the reflectors and the parameters that define the seismic wavelet. In that sense, the goal of the proposed hybrid method is to get the best out of these two algorithms: the simplicity of FISTA and the versatility of VFSA, with a reasonable computational burden. One advantage of the method is that, since VFSA is an stochastic algorithm, the uncertainty of the solutions can be estimated due to the large number of solutions that are tested during the inversion process.

The proposed method was tested on noisy synthetic normal-move-out (NMO) corrected prestack data. For the approximation to the Zoeppritz equations we used the well-known two-term Shuey's equation (Shuey, 1985). To account for absorption and dispersion effects associated with the wave propagation through the earth, we considered a time-varying Ricker wavelet (Ricker, 1940) as the source signature both to generate the data and during the inversion process. To this end, we allowed for a wavelet central frequency linearly decreasing with time, and a constant phase rotation linearly varying with time. These parameters were estimated during the inversion process, together with the reflectors positions and amplitudes, as mean to take into account the fact that only an approximate source wavelet might be available. The results show that high-resolution AVA attributes can be derived from noisy data accurately, even when the number of reflectors is not known *a priori*, and the initial wavelet is inaccurate.

2 THEORY

The convolutional model is used as the base of the inversion method (Yilmaz, 2001). This model assumes that the medium is composed of a series of flat, parallel, homogeneous and isotropic layers. Then, the trace $\mathbf{s}$ for the i-th angle of incidence θ_i can be expressed as

$$\mathbf{s}(\theta_i) = \sum_{t=1}^{N_r} \mathbf{w}_t * \mathbf{I}_t \mathbf{r}(\theta_i) + \mathbf{n}(\theta_i), \qquad i = 1, \cdots, N \tag{1}$$

where the symbol "$*$" denotes convolution, $\mathbf{w}_t$ is source wavelet of dimension L_w at time t, $\mathbf{r}(\theta_i)$ is the vector of dimension L_r that contains the primary reflection coefficients (reflectivity), and $\mathbf{n}(\theta_i)$ is the random noise. Both $\mathbf{n}(\theta_i)$ and $\mathbf{s}(\theta_i)$ have dimension $L_s = L_w + L_r - 1$. Finally, $\mathbf{I}_t$ is a $L_r \times L_r$ matrix given by

$$\mathbf{I}_t\{l, m\} = \begin{cases} 1 & l, m = t, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Equation (1) can be expressed in matrix form as

$$\mathbf{s}(\theta_i) = \sum_{t=1}^{N_r} \mathbf{W}_t \mathbf{I}_t \mathbf{r}(\theta_i) + \mathbf{n}(\theta_i), \tag{3}$$

where $\mathbf{W}_t$ is the convolution matrix of dimension $L_s \times L_r$ corresponding to the wavelet $\mathbf{w}_t$. Then, setting

$$\mathbf{W} = \sum_{t=1}^{N_r} \mathbf{W}_t \mathbf{I}_t, \tag{4}$$

equation (3) can be expressed as

$$\mathbf{s}(\theta_i) = \mathbf{W}\mathbf{r}(\theta_i) + \mathbf{n}(\theta_i), \qquad i = 1, \cdots, N. \tag{5}$$

For a given angle of incidence θ, the reflection coefficient for a reflector at sample t can be approximated, in a general form, as

$$r_t(\theta) = \sum_{k=1}^{n} x_{tk} g_k(\theta), \tag{6}$$

where x_{tk} are coefficients that depend on physical properties of the rocks at each side of the interface (velocities and densities), and n is the order of the chosen Zoeppritz approximation (usually $n = 2$ or 3). The functions $g_k(\theta)$ depend on the angle of incidence which must be less than the critical angle.

When the reflectivity is sparse-spike, we can use equation (5) to represent the trace $\mathbf{s}(\theta_i)$ by only changing the appropriate indices of the matrix $\mathbf{I}_t$. If there are M non-zero spikes at the samples τ_j, where $j = 1, ..., M$ and $M \ll L_r$, then $t = \tau_1, ..., \tau_M$. Then, matrices $\mathbf{W_t}$, and $\mathbf{W}$, are quite sparse. Thus, combining equations (1) to (6), and omitting the noise term for simplicity, the trace $\mathbf{s}(\theta_i)$ associated with a sparse-spike reflectivity can be rewritten as

$$\mathbf{s}(\theta_i) = \mathbf{A}(\theta_i)\mathbf{x}, \qquad i = 1, \cdots, N \tag{7}$$

where $\mathbf{x} = (x_{11}, \cdots, x_{M1}, \cdots, x_{1n}, \cdots, x_{Mn})^T$, and $\mathbf{A}(\theta_i)$ is a matrix of dimension $L_s \times Mn$ that can be expressed as

$$\mathbf{A}(\theta_i) = \left(\hat{\mathbf{W}} g_1(\theta_i) | ... | \hat{\mathbf{W}} g_n(\theta_i)\right). \tag{8}$$

Here, $\hat{\mathbf{W}}$ is a convolution matrix of dimension $L_s \times M$ defined in the same way that $\mathbf{W}$, but now considering that $t = \tau_1, ..., \tau_M$, only. Since usually $M \ll L_r$, the constraint given by the sparse-spike assumption reduces significantly the size of the matrices required for the inversion process, decreasing at the same time the computational cost of the method.

The N systems of equations (7) can be re-arranged into a unique system of equations in the form

$$\mathbf{s} = \mathbf{A}\mathbf{x}, \tag{9}$$

where $\mathbf{A}$ and $\mathbf{s}$ are the "stacks" of $\mathbf{A}(\theta_i)$ and $\mathbf{s}(\theta_i)$, respectively.

Assuming Gaussian noise, the solution $\hat{\mathbf{x}}$ can be obtained by minimizing the discrepancy E between the observed data $\mathbf{s}$ and the modeled data:

$$E = \| \mathbf{A}\hat{\mathbf{x}} - \mathbf{s} \|_2^2 . \tag{10}$$

In an inverse problem the relation between the model parameters and the measured data is not necessarily linear, then a global optimization technique can be used to find a correct solution

to avoid local minima and other convergence issues (Buland and Omre, 2003). In practice, we minimize E with respect to the positions τ_j and the parameters that define the wavelets $\mathbf{w}_t$ using Very Fast Simulated Annealing (VFSA) (Ingber, 1989) (see Appendix). In each iteration of the VFSA, given the M values of τ_j and the parameters that define the wavelets, equation (9) becomes linear. Hence, the optimum vector $\hat{\mathbf{x}}$ can estimated by solving the sparse least-squares problem (10), which leads to

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T\mathbf{s}\right). \tag{11}$$

VFSA is a powerful and versatile non-linear optimization algorithm that does not require the use of gradients or derivatives to find the global minimum of a function. It is also independent of the initial model. However, the computational cost of the iterative process until convergence can be very high if the number of model parameters is unnecessarily large, and the initial model is too far from the optimum solution. To avoid these undesired consequences, we decided to estimate the number of reflectors M and their initial approximate position τ_j using a shrinkage-thresholding algorithm.

To this end we first obtain an estimated sparse solution of the system of equations (9) using $M = L_r$. This new system can be expressed as

$$\mathbf{s} = \mathbf{B}\mathbf{y}, \tag{12}$$

where $\mathbf{B}$ is a matrix of dimension $L_s \times nL_r$, and $\mathbf{y}$ is a vector of dimension nL_r. To construct the model matrix $\mathbf{B}$ an initial estimate of the wavelet is required. A sparse solution of the system given by equation (12) can be estimated through l_1-*norm* regularization (Oldenburg et al., 1983), which implies to find the vector $\mathbf{y}$ that minimizes

$$\| \mathbf{B}\mathbf{y} - \mathbf{s} \|_2^2 + \lambda \| \mathbf{y} \|_1, \tag{13}$$

where λ is a trade-off parameter used to balance the weight or impact of the two terms. We use the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) (Beck and Teboulle, 2009) to solve this non-linear inverse problem (see Appendix).

Step-by-step, the proposed hybrid inverse strategy is as follow:

1. An initial seismic wavelet is estimated from the available data.
2. Matrix $\mathbf{B}$ (equation (12)) is constructed using the approximate wavelet, and an initial sparse solution $\hat{\mathbf{y}}$ is estimated using FISTA through the minimization of equation (13).
3. Both M and τ_j, $j = 1, \cdots, M$, are determined based on the non-zero components of $\hat{\mathbf{y}}$.
4. VFSA is used to minimize equation (10), which includes the tuning of the initial wavelet and adjusting the position of the reflectors. Amplitudes are obtained using equation (11) at each annealing iteration.

3 NUMERICAL EXAMPLES

3.1 The synthetic data

To test the proposed method we generated a synthetic NMO-corrected gather consisting of 31 traces with $\theta_i \in (0°, 30°)$, and 6 reflectors with times between 0.04 and 0.24 s. The AVA re-

sponse was modeled using the two-terms Shuey approximation (Shuey, 1985). So, equation (6) becomes

$$r_t(\theta) = I_t + G_t \sin^2(\theta). \tag{14}$$

The coefficients I and G, which are the objective of the inversion, are known as *Intercept* and *Gradient*. Data traces were generated by convolving the reflectivities with a phase-rotated Ricker wavelet of central frequency f_0 (Ricker, 1940). In order to account for the expected time-varying attenuation and other propagation effects that change the shape of the source wavelet, in this example the central frequency f_0 and the phase rotation φ_0 are allowed to vary linearly from $f_0^i = 30$ Hz and $\varphi_0^i = 20°$ for the first sample of the data, to $f_0^f = 20$ Hz and $\varphi_0^f = 40°$ for the last sample of the data. Gaussian noise was added with standard deviation $\sigma = \max_{ti} |s_t(\theta_i)|/\text{SNR}$, where SNR is the signal-to-noise ratio. Figures 1a and 1b show the *Intercept* and *Gradient* used to generate the data. Figures 1c and 1c show the noisy data with SNR=20 and 10, respectively. Figure 1e shows the wavelets for those particular times where the reflectors are located. Note the broadening of the pulse due to the f_0 and φ_0 time variations.

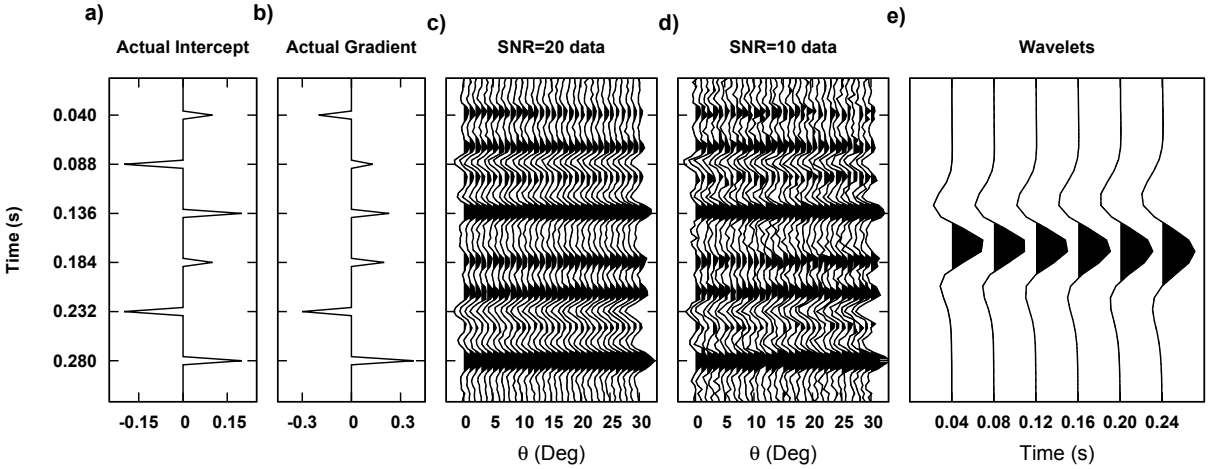

Figure 1: a) Actual Intercept and b) actual Gradient, c) noisy data with SNR=20, d) noisy data with SNR=10, and d) evolution of the wavelet as a function of the propagation time.

3.2 Using FISTA to estimate the initial model

To use FISTA to estimate an initial model we first need to select the appropriate trade-off parameter λ in equation (13). In general, the selection of this parameter depends on the noise level of the data at hand. Clearly, if λ is too large, a very sparse solution would be obtained, but data may not be honored and thus some of the reflectors may be missed. Conversely, if λ is too small, data would be over-fit and there would be too many spurious reflectors, leading to non-sparse solutions. Further, an excessively large number of reflectors implies an unnecessary computational cost in the VFSA stage. Therefore, an appropriate strategy is required to select this key parameter. There are various methods to estimate λ, such as the L-curve criterion, the discrepancy principle, and the generalized cross-validation criterion (Farquharson and Oldenburg, 2004). Since we are interested in sparse solutions that fit the data, we want to keep the number of reflectors to a minimum, provided that data are honored without overfitting. To estimate the correct parameter λ we analyzed the variation of the different magnitudes of interest that are involved in the equation (13), and built a set of curves as a guide (Figure 2). These curves

have been built by applying FISTA on both data sets, and then counting the number of non-zero reflectors while varying λ from 0 to 15. Thick continuous lines in Figure 2a and 2b show the variation of the number of reflectors as a function of λ, while thin dashed lines show the actual number of reflectors, which is equal to 6 in both cases. From these curves we can infer that, to obtain a solution with 6 (or less) reflectors, λ should be no smaller than 14. Since the actual number of reflectors is not known *a priori*, and in order guarantee that data would be honored, we decided to choose a smaller value for λ. This gave us the opportunity to test the behavior of VFSA against a number of reflectors larger than the actual one. Hence, in these numerical examples we chose $\lambda = 12$, which lead to solutions with 8 and 10 reflectors for the data with SNR=20 and 10, respectively.

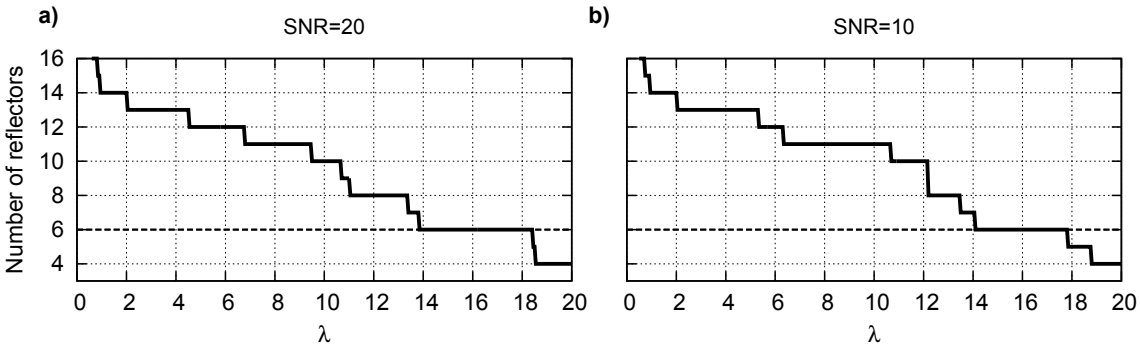

Figure 2: Number of reflectors as a function of the trade-off parameter λ: a) SNR=20, b) SNR=10.

Once estimated an appropriate trade-off parameter λ we analyzed the behavior of the solutions given by FISTA using different wavelets. Figures 3a, 3b, 3f and 3g show the solutions for the *Intercept* and *Gradient* obtained using a wavelet with an incorrect variation of the phase rotation φ_0, while the variation of the central frequency f_0 is the one used to generate the data (i.e. 30 Hz - 20 Hz). On the other hand, Figures 3d, 3e, 3i and 3j show the solutions using a wavelet with the correct variation of the phase rotation (i.e. $20° - 40°$), but with an incorrect variation of the central frequency. For comparison, Figures 3c and 3h show the solution given by FISTA using the exact wavelet. All these solutions where obtained applying FISTA on the data with SNR=10 and using the trade-off parameter $\lambda = 12$ as indicated above. As can be observed, FISTA solutions are very sensitive to the wavelet. The use of the correct wavelet produces a very good solution, as concluded by comparing Figures 3c and 3h with the actual *Intercept* and *Gradient* shown in Figures 1a and 1b. All reflector positions and amplitudes are correctly estimated, with almost no spurious spikes. On the other hand, the use of an incorrect wavelet gives solutions that are sparse, as desired, but unstable, showing a relatively large number of spurious spikes of significant amplitudes.

3.3 The FISTA+VFSA hybrid algorithm

Clearly, FISTA solutions using incorrect wavelets are not acceptable. Even though, they might be good initial models for the VFSA stage, as we will show later. In practice, the estimation of an accurate wavelet from noisy data is a very difficult task, usually requiring complex processes and the use *a priori* information such as sonic profiles from well logs (Henry, 1997), which are rarely available. As an initial estimate to be used by FISTA, we decided to use a unique time invariant zero-phase Ricker wavelet with $f_0 = 25$ Hz. This is a very raw approximation to the actual wavelet contained in the data, but sufficient to provide an initial estimate to

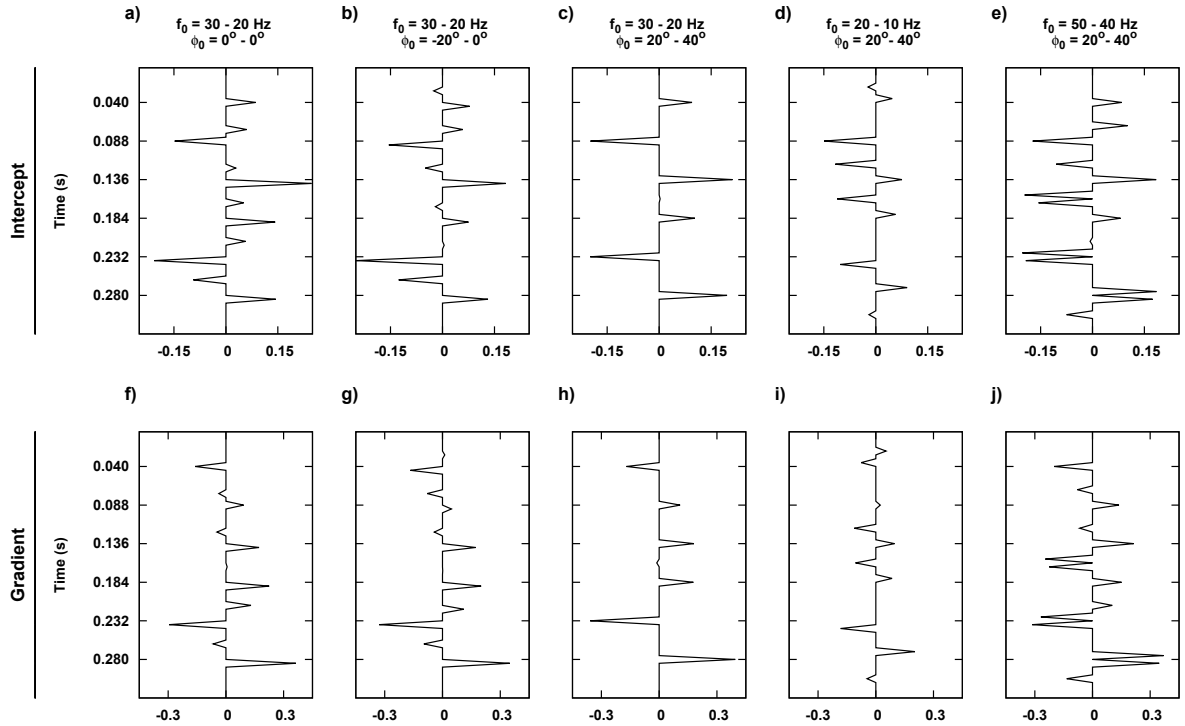

Figure 3: FISTA solutions using different wavelets (data with SNR=10); a), b), f) and g) wavelets with correct central frequency but incorrect phase rotation; d), e), i) and h) wavelets with correct phase rotation but incorrect central frequency; c) and h) solutions with the correct wavelet.

be fed into the VFSA stage. During this inversion stage, we set the search ranges for the time-varying central frequencies and phase rotations $(10, 60)$ Hz and $(-90°, 90°)$, respectively. The reflectors were free to move in time, but as a constraint side-by-side reflectors were forbidden. For statistical purposes we carried out 100 inversions using different seeds for the generation of the pseudo random numbers used in VFSA. As a stop criterion was set a maximum of 2000 iterations (cost function evaluations) or when the misfit was less or equal than the noise level.

Table 1 summarizes the statistical results of the estimated wavelets. As it can be seen, all mean solutions are very close to the actual values for all the estimated magnitudes. The small values of the standard deviations show the capacity of the proposed method to successfully estimate a time-varying Ricker wavelet given a very raw initial estimate.

	SNR=20		SNR=10	
	Mean value	Standard deviation	Mean value	Standard deviation
f_0^i	28.9	0.71	29.5	0.69
f_0^f	20.5	0.43	20.3	0.41
φ_0^i	18.4	3.59	31.3	1.45
φ_0^f	40.5	2.48	39.2	0.77

Table 1: Statistical results of the estimated wavelet parameter after 100 realizations. Actual central frequencies and phase rotation parameters are $(30, 20)$ Hz and $(20°, 40°)$, respectively.

Figure 4 shows the *Intercept* and *Gradient* sparse solutions. We can observe that the method gives good results even in the presence of a significant amount of noise. Figure 4a and 4e show the actual *Intercept* and *Gradient*. Figures 4b and 4f show the sparse solutions obtained by

FISTA using the initial approximate wavelet for the data with SNR=10. Actually, these are the initial models used for the subsequent VFSA stage. Figures 4c, 4d, 4g and 4h show the mean sparse solutions and their corresponding standard deviations σ, for both the SNR=20 and SNR=10 cases, after the VFSA stage. As expected, the estimated solutions from the data with SNR=20 are more accurate than the ones obtained from the data with SNR=10. As shown in the figure, the mean solutions in the two cases are very similar to the actual values. As in the case of the wavelet parameters (Table 1), the standard deviations are very small, showing the consistence of the method to find sparse solutions that honor the observed data.

It is worth noting that the FISTA solution using the approximate initial wavelet is, in fact, not too far from the desired solution. However, as shown in the previous section, it has too many spurious spikes of significant magnitude. In the solution obtained by VFSA those spurious spikes did not disappear at all, but due to the capacity of the algorithm to move the position of the reflectors, most of them where relocated to places where there is only noise. Then, their amplitudes were reduced to negligible values.

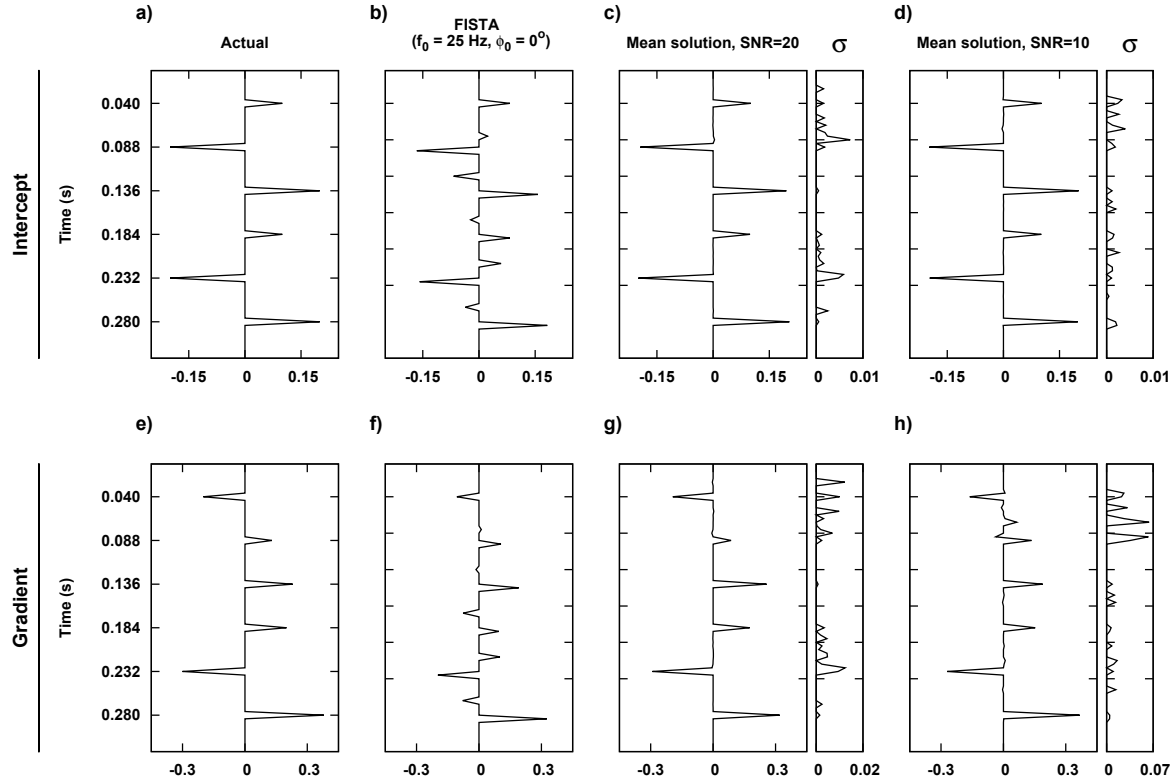

Figure 4: a) and e) Actual *Intercept* and *Gradient*; b) and f) FISTA solutions with the approximate wavelet (SNR=10); c) and g) FISTA+VFSA solutions for the data with SNR=20; d) and h) FISTA+VFSA solutions for the data with SNR=10.

4 CONCLUSIONS

In this work we presented a hybrid algorithm for sparse-spike AVA inversion. The proposed hybrid algorithm is based on the usage of the Fast Iterative Shrinkage-Thresholding Algorithm and Very Fast Simulated Annealing. The objective of the method is to obtain high resolution AVA attributes from prestack data, such as *Intercept* and *Gradient* and, simultaneously, perform an adjustment of the source wavelet. To this end, an initial sparse solution is obtained with the very efficient FISTA using an initial approximate wavelet. Then, in a second stage, VFSA

is used to tune both the the initial wavelet estimate and the reflector positions. Numerical test using synthetic data showed that the proposed method is capable to obtain satisfactory sparse solutions and a good estimation of the source wavelet, even in the presence of noise and a poor initial estimation of the wavelet. The use of a stochastic optimization method, such as VFSA, allowed one to obtain many solutions that fit the data by using different seeds, and thus the solution uncertainties can be estimated. The small values of standard deviations obtained for all the inverted magnitudes tell us about the consistency of the method.

5 ACKNOWLEDGMENTS

This work was partially supported by Agencia Nacional de Promoción Científica y Tecnológica (PICT-2010-2129).

APPENDIX

About the Fast Iterative Shrinkage-Thresholding Algorithm

Given the cost or objective function

$$J = \| \mathbf{Ax} - \mathbf{s} \|_2^2 + \lambda \| \mathbf{x} \|_1, \tag{15}$$

where $\mathbf{A}$ is the model matrix, $\mathbf{s}$ is the observed data, $\mathbf{x}$ is the unknown, and λ is a trade-off parameter, *Fast Iterative Shrinkage-Thresholding Algorithm* (FISTA) is a powerful algorithm to minimize J very efficiently in terms of computational cost (Beck and Teboulle, 2009). FISTA is based on the Iterative Shrinkage Thresholding Algorithm (ISTA) (Daubechies et al., 2004). These algorithms are an extension of the classical gradient algorithm used to solve large-scale linear inverse problems in a simple way. One of the advantages of ISTA and FISTA over other known methods is that at each iteration only matrix-vector multiplications are required, making them simple and cost-effective. ISTA is known to have linear convergence, but FISTA is shown to be faster by several orders of magnitude (Beck and Teboulle, 2009). In this context, FISTA promotes sparse-spike solutions in a simple and effective way.

The process carried out by FISTA to minimize equation (15) is summarized as follows:

1. Set $\alpha \geq \sigma_{max}$, where σ_{max} is the maximum eigenvalue of $\mathbf{A}^T\mathbf{A}$.
2. Set $\mathbf{z}_1 = \mathbf{x}_0$ and $t_1 = 1$, where $\mathbf{x}_0$ is an initial solution.
3. For each iteration $k = 1, 2, 3, \cdots$:

   (a)

   $$\mathbf{x}_k = T_{\lambda/2\alpha}\left\{\mathbf{z}_k - \frac{1}{\alpha}\mathbf{A}^T\left(\mathbf{Az}_k - \mathbf{s}\right)\right\}, \tag{16}$$

   where $T_\beta\{\cdot\}$ is a soft-thresholding function which is applied to each element of its vectorial argument and is defined by

   $$T_\beta\{g\} = \begin{cases} g\left(1 - \beta/|g|\right), & |g| \geq \beta \\ 0. & |g| < \beta \end{cases} \tag{17}$$

   (b)

   $$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}. \tag{18}$$

(c)

$$\mathbf{z}_{k+1} = \mathbf{x}_k + \frac{t_k - 1}{t_{k+1}} \left(\mathbf{x}_k - \mathbf{x}_{k-1} \right). \tag{19}$$

(d) Check convergence or stopping condition.

Step 1 is required to prevent the argument of the soft-thresholding function to become negative. For more details about FISTA, please refer to Beck and Teboulle (2009).

About the Very Fast Simulated Annealing

Simulated Annealing (SA) is a powerful non-linear optimization algorithm that does not require the use of gradients or derivatives to find the global minimum of any given cost or objective function. It is named after the *annealing* process, a metallurgical technique consisting of heating and cooling a metallic material to change its physical properties (e.g. hardness, ductility, etc.). These changes have been studied by Metropolis et al. (1953) using Monte Carlo techniques. Kirkpatrick et al. (1983) generalized this concept and applied it to non-linear optimization problems. The unknown parameters of the model play the role of the particles of the metallic material, and the energy state of the system is represented by a cost function. SA optimization is an iterative process. At each iterative step the model configuration, analogous to the position of the particles in the metal, is selected "randomly" using a predefined probability density function that depends on a control parameter known as *temperature*. This new model configuration is later accepted (or rejected) according to the *Metropolis criterion*. The Metropolis criterion states that if the cost function decreases, the model is accepted with a probability equal to one, but, if the cost function increases, the model is accepted with a non zero probability. This strategy allows the algorithm to accept solutions that increase the cost function, thus avoiding getting trapped in local minima. It is important to notice that the temperature should be slowly decreased during the optimization process, following a preset cooling schedule. When the temperature is high, the space of possible models is explored in an approximately uniform way. On the other hand, at low temperatures the models with smaller values of the cost function are chosen. Should the temperature of the system be reduced too fast, the particles of the material, represented by the unknown parameters of the model, would not reach the state of minimum energy, represented by the global minimum of the cost function. For this reason, the selection of an appropriate cooling schedule is essential for the proper operation of the algorithm. Finally, note that this acceptance probability depends on the cost function and the temperature. The lower the cost function, the higher the probability, and the lower the temperature, the lower the probability. Convergence is achieved when, at low temperatures, there is not significant decrease in the cost function and the system is in the lowest energy state.

Very Fast Simulated Annealing (VFSA) is a variation of the SA algorithm proposed by Ingber (1989). The main differences are the cooling schedule and the probability density function used to select new models at each iteration. VFSA uses a long-tailed Cauchy-like distribution that allows to explore the model space in a more efficient way than using Gaussian or uniform distributions as in other conventional SA algorithms. This allows to select a faster cooling schedule to accelerate convergence without limiting the capacity of the algorithm to escape from local minima. Figure 5 shows the workflow of the VFSA.

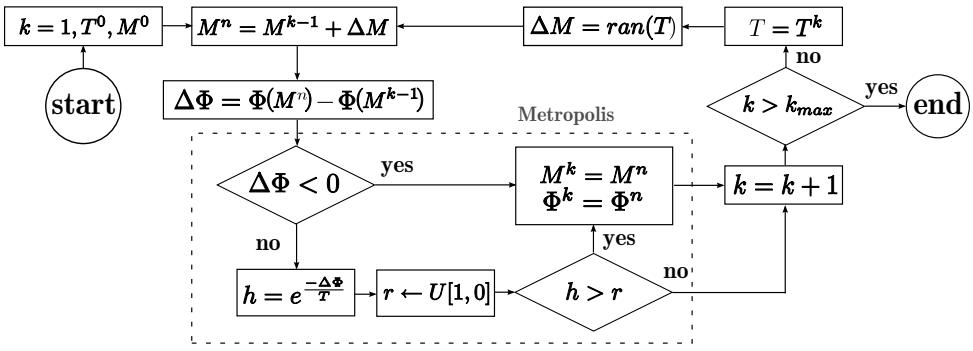

Figure 5: Workflow of the Very Fast Simulated Annealing. Φ is the cost function, M^k is the model at the k-iteration and M^n is the temporary model with acceptance probability h.

REFERENCES

Aki K. and Richards P. *Quantitative seismology: theory and methods*. W.H. Freeman and Co., 1980.

Alemie W.M. and Sacchi M.D. High-resolution three-term AVO inversion by means of a trivariate Cauchy probability distribution. *Geophysics*, 76(3):R43–R55, 2011.

Beck A. and Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal of Imaging Sciences*, 2:183–202, 2009.

Buland A. and Omre H. Bayesian linearized AVO inversion. *Geophysics*, 68(1):185–198, 2003. doi:10.1190/1.1543206.

Daubechies I., Defrise M., and Mol C.D. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.

Debeye H.W.J. and van Riel P. L_p-norm deconvolution. *Geophysical Prospecting*, 38:381–403, 1990.

Downton J.E. and Lines L.R. High-resolution AVO analysis before NMO. In *Expanded Abstracts*, volume 22, pages 219–222. SEG, 2003. ISSN 1. doi:10.1190/1.1817781.

Farquharson C.G. and Oldenburg D.W. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. *Geophysical Journal International*, 156:411–425, 2004.

Fatti J.L., Smith G.C., Vail P.J., Strauss P.J., and Levitt P.R. Detection of gas in sandstone reservoirs using AVO analysis: A 3-D seismic case history using the Geostack technique. *Geophysics*, 59(9):1362–1376, 1994.

Henry S.G. Catch the (seismic) wavelet. *AAPG Explore*, 18(3):36–38, 1997.

Ingber L. Very fast simulated re-annealing. *Journal of Mathematical Computation and Modelling*, 12:967–973, 1989.

Kirkpatrick S., Gellat C.J., and Vecchi M. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., and Teller E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

Misra S. and Sacchi M.D. Global optimization with model-space preconditioning: Application to AVO inversion. *Geophysics*, 73(5):R71–R82, 2008.

Oldenburg D.W., Scheuer T., and Levy S. Recovery of the acoustic impedance from reflection seismograms. *Geophysics*, 48(10):1318–1337, 1983.

Pérez D.O., Velis D., and Sacchi M.D. Inversion of prestack seismic data using FISTA. In

Mecánica Computacional, volume 31, pages 3255–3263. AMCA, 2012. ISSN 1666-6070.

Pérez D.O. and Velis D.R. Sparse-spike AVO/AVA attributes from prestack data. In *Expanded Abstracts*, volume 30, pages 340–344. SEG, 2011. doi:10.1190/1.3627906.

Ricker N. The form and nature of seismic waves and the structure of seismograms. *Geophysics*, 5:348–366, 1940.

Sacchi M.D. Reweighting strategies in seismic deconvolution. *Geophysical Journal International*, 129(3):651–656, 1997. doi:10.1111/j.1365-246X.1997.tb04500.x.

Shuey R. A simplification of the Zoeppritz equations. *Geophysics*, 50(4):609–614, 1985.

Taylor H.L., Banks S.C., and McCoy J.F. Deconvolution with the L_1-norm. *Geophysics*, 44(1):39–52, 1979. doi:10.1190/1.1440921.

Velis D.R. Stochastic sparse-spike deconvolution. *Geophysics*, 73:R1–R9, 2008.

Yilmaz O. *Seismic Data Analysis: processing, inversion, and interpretation of seismic data*. Investigations in Geophysics. SEG, 2001.

Zoeppritz K. Über Reflexion and Durchgang seismischer Wellen durch Unstetigkeits-flächen. *Gott. Nachr. Math. Phys*, K1:66–84, 1919.