Acoustic impedance estimation using a gradient-based algorithm with total variation semi-norm regularization

Daniel O. Pérez (FCAG-UNLP; CONICET; Y-TEC S.A.)* and Danilo R. Velis (FCAG-UNLP; CONICET), La Plata, Argentina.

Abstract

We present an algorithm to estimate blocky images of the subsurface acoustic impedance (AI) from seismic reflection data. We use the total variation semi-norm (TV) to regularize the inversion and promote blocky solutions which are, by virtue of the capability of TV to handle edges properly, adequate to model layered earth models with sharp contrasts. In addition, the use of the TV leads to a convex objective function that can be minimized using a gradient-based algorithm that only requires matrix-vector multiplications and no direct matrix inversion. The latter makes the algorithm numerically stable, easy to apply, and economic in terms of computational cost. Besides, given appropriate *a priori* information, the algorithm allows to easily incorporate into the inversion scheme the low frequency trend that is missing from the data. Numerical tests on noisy 2D synthetic and field data show that the proposed method is capable of providing consistent and blocky AI images that preserve edges and the subsurface layered structure.

Introduction

The inversion of poststack seismic data for AI is a common technique used to obtain information about the structure of the subsurface. It allows to establish relations between the recorded seismic data and the geology (Oldenburg et al., 1983). Inverse problems are usually solved by minimizing a cost function that measures the differences between the observed and the modeled data (Tarantola, 2005). Unfortunately, the solutions of most geophysical inverse problems are inherently non-unique, for there exists several solutions that honor the data equally well. Furthermore, seismic inversion is an ill-posed problem, meaning that little amounts of noise in the observed data lead to large errors in the estimated solutions. For the sake of stabilization, and to avoid meaningless solutions, an appropriate regularization must be used during the inversion process. A well-chosen regularization can also impart desirable characteristics to the estimated solution (Ulrych & Sacchi, 2005). In addition, as consequence of the band-limited nature of the seismic data, there is a lack of low frequency information that must be incorporated into the inversion process to properly constrain the estimated AI solutions. The inversion strategy must overcome the aforementioned drawbacks in a com-putationally efficient way, as the amount of seismic data to process is usually large.

The AI inversion can be separated into reflectivity domain inversion and data domain inversion (Gholami, 2016). In the former the reflectivity is first estimated from the seismic data via deconvolution, and then the AI is derived from the reflectivity by recursive integration. In the later the AI is directly estimated from the seismic data. Also, depending on if the source wavelet is known *a priori* or it is estimated during the inversion process, the inversion can be separated into non-blind and blind. In this sense, several authors have developed AI inversion techniques with very interesting results (Oldenburg et al., 1983; Cooke & Schneider, 1983, e.g.). In the context of blocky AI inversion, Velis (2008) proposed to use a global optimization algorithm known as very fast simulated annealing to perform the non-blind AI inversion in the reflectivity domain. Also in this domain, Gholami & Sacchi (2013) derived an AI inversion algorithm based on the split Bregman iteration method (Goldstein & Osher, 2009). Based on the same strategy, Gholami (2015, 2016) developed various techniques to perform non-blind and blind AI inversions in the data domain.

In this work we propose an alternative algorithm for the non-blind multichannel AI linear inversion in the data domain. The proposed algorithm uses the TV semi-norm (Rudin et al., 1992; Chambolle, 2004) as regularization term to promote blocky solutions of the AI. The use of the TV regularization leads to a convex cost function that can be minimized using an efficient iterative gradient-based algorithm that requires no matrix inversion (Beck & Teboulle, 2009a). Blocky solutions are desirable because they lead to sharply resolved discontinuities that overcome the band-limitation of classical l_2-norm solutions. Based on impedance well logs information, one may argue that the real structure of the subsurface is continuous rather than layered (Cooke & Schneider, 1983). Nevertheless, it is known from well log data studies that the amplitudes of the reflection coefficients associated with the interfaces follow a non-Gaussian distribution (Walden & Hosken, 1986; Velis, 2003). This fact indicates that the main lithological units can be represented by layers with certain properties.

This work is organized as follows. First, we explain the proposed method, we set up the necessary hypothesis and define all relevant equations. In addition, we provide a step-by-step description of the inversion algorithm. Next, we test the method on 2D synthetic poststack seismic data from the Marmousi model and analyze the corresponding results. Then, we test the algorithm using a 2D field data set. Finally, the obtained results are summarized in the conclusions.

Method

Given a layered AI model $\mathbf{Z}$, a source wavelet $\mathbf{w}$, and assuming the validity of the convolutional model, the noisy seismic section can be expressed as

$$\mathbf{S} = \mathbf{WDX} + \mathbf{N}, \qquad (1)$$

with

$$\mathbf{X} = \frac{1}{2} \log(\mathbf{Z}), \qquad (2)$$

where $\mathbf{D}$ is the first order difference operator, $\mathbf{W}$ is the Toeplitz matrix associated with the source wavelet and $\mathbf{N}$ is the additive noise term. Equation (2) relies in the hypothesis that the contrasts of the AI across the boundaries of the layers are small compared to their absolute values. This assumption allow us to estimate the AI through a linear inverse problem (Cooke & Schneider, 1983).

Assuming that $\mathbf{N}$ contains uncorrelated Gaussian noise, the inversion is carried out minimizing the following cost function:

$$J(\mathbf{X}) = \frac{1}{\sigma^2} \|\mathbf{AX} - \mathbf{S}\|^2 + \mu TV(\mathbf{X}), \qquad (3)$$

where σ^2 is the noise variance, μ is a trade-off parameter, $\mathbf{A} = \mathbf{WD}$ and

$$TV(\mathbf{X}) = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \sqrt{(x_{i,j} - x_{i+1,j})^2 + (x_{i,j} - x_{i,j+1})^2} \qquad (4)$$

is the isotropic TV semi-norm (Chambolle, 2004). In equation (4), m and n denote the dimension of the matrix $\mathbf{X}$.

The cost function $J(\mathbf{X})$ is composed of two terms, each one imposing a different constraint on $\mathbf{X}$. The first term represents the misfit between the modeled and the observed data. Its minimization will ensure that the estimated solution honors the observed data, a constraint that must be satisfied. On the other hand, the second term represents the regularization term. Its minimization will impart desirable characteristics to the estimated solutions. The overall impact of the regularization is controlled by the trade-off parameter μ. The TV is a regularization approach capable of handling edges properly. As the AI model is related to the lithology, for layered earth models with sharp contrasts between adjacent layers the TV seems to be an adequate choice. This norm will impose sharpness and appropriate amplitude constraints for the estimated image $\mathbf{X}$ (and thus for $\mathbf{Z}$), as we will show in the Examples section.

The described inversion problem is viewed as the combination of a deconvolution problem and a TV denoising problem. The inherent large scales of the problems to be solved require the use of fast and simple numerical methods. For this reason, we minimize equation (3) by means of the iterative gradient-based algorithm proposed by Beck & Teboulle (2009a). The algorithm is composed of two nested iterative processes, as detailed in Algorithm 1. The outer process performs the deconvolution of the traces and is based on the Fast Iterative Shrinkage-thresholding Algorithm (FISTA) developed by Beck & Teboulle (2009b), an algorithm that requires no matrix inversions. The inner iterative process performs the TV denoising of the deconvolved images (line 4 of Algorithm 1). The denoising problem does not lead to a closed-form expression, then it is necessary to rely on an iterative algorithm. To this end we use the dual TV denoising iterative algorithm developed by Chambolle (2004). Like FISTA, this globally convergent TV denoising algorithm does not require any matrix inversion neither. Readers are referred to the cited articles for a detailed description of the algorithms. Although the dual TV denoising algorithm seems to be appropriate in the current context, other methods can be used without loss of generality. We stress the fact that the resulting algorithm that we propose to minimize the cost function $J(\mathbf{X})$ only requires matrix-vector multiplications, making it numerically stable, easy to apply and economic in computational terms.

Due to the band-limited nature of the seismic data, it is not possible to recover the low-frequency trend $\mathbf{T}$ of $\mathbf{X}$ (and thus of $\mathbf{Z}$) during the inversion process. The low-frequency trend information must be obtained *a priori* from other sources (e.g. stacking velocities or well logs), and then incorporated into the inversion process by means of appropriate constraints. In this case, as indicated in the line 2 of Algorithm 1, we incorporate this information using $\mathbf{T}$ as the initial solution of the iterative process.

Algorithm 1

input: $\mathbf{A}, \mathbf{S}, \mathbf{T}, \mu$
output: estimated AI image $\mathbf{Z}$

1: Set L as the Lipschitz constant of $\nabla \|\mathbf{AX} - \mathbf{S}\|^2$
2: set $\mathbf{X}_0 = \mathbf{Y}_1 = \mathbf{T}$ and $k = t_1 = 1$
3: **loop**
4:      $\mathbf{W}_k = $ TV-denoising $\left(\mathbf{Y}_k - \frac{2}{L} \mathbf{A}^T (\mathbf{AY}_k - \mathbf{S}), L, \mu \right)$
5:      $\mathbf{X}_k = \text{argmin}(J(\mathbf{U}) : \mathbf{U} = \mathbf{W_k}, \mathbf{X_{k-1}})$
6:      $t_{k+1} = \dfrac{1 + \sqrt{1 + 4t_k^2}}{2}$
7:      $\mathbf{Y}_{k+1} = \mathbf{X}_k + \left(\dfrac{t_k}{t_{k+1}} \right) (\mathbf{W}_k - \mathbf{X}_k) + \left(\dfrac{t_k - 1}{t_{k+1}} \right) (\mathbf{X}_k - \mathbf{X}_{k-1})$
8:      Check convergence or stopping condition.
9:      k=k+1
10: **end loop**
11: $\mathbf{Z} = \exp(2\mathbf{X}_k)$

The Lipschitz constant of $\nabla \|\mathbf{AX} - \mathbf{S}\|^2$ required in the line 1 of Algorithm 1, is given by $L = 2\lambda_m(\mathbf{A}^T\mathbf{A})$ (Palomar & Eldar, 2010; Gramfort et al., 2013), where $\lambda_m(\cdot)$ denotes the maximum eigenvalue of its argument. In practice, we obtain the maximum eigenvalue using the power iteration method (Larson & Edwards, 1999, e.g.).

Another important practical issue to take into account is the stopping condition of the iterative process (line 8 of Algorithm 1). Ideally, the process should stop when the misfit is equal to the noise level. Unfortunately, depending on the selected trade-off parameter μ, the misfit often shows an asymptotic behavior towards the noise level and convergence is not achieved in a acceptable number of iterations. To overcome this problem we decided to stop the iterative process whenever the cost function does not show a significant variation for a given number of iterations.
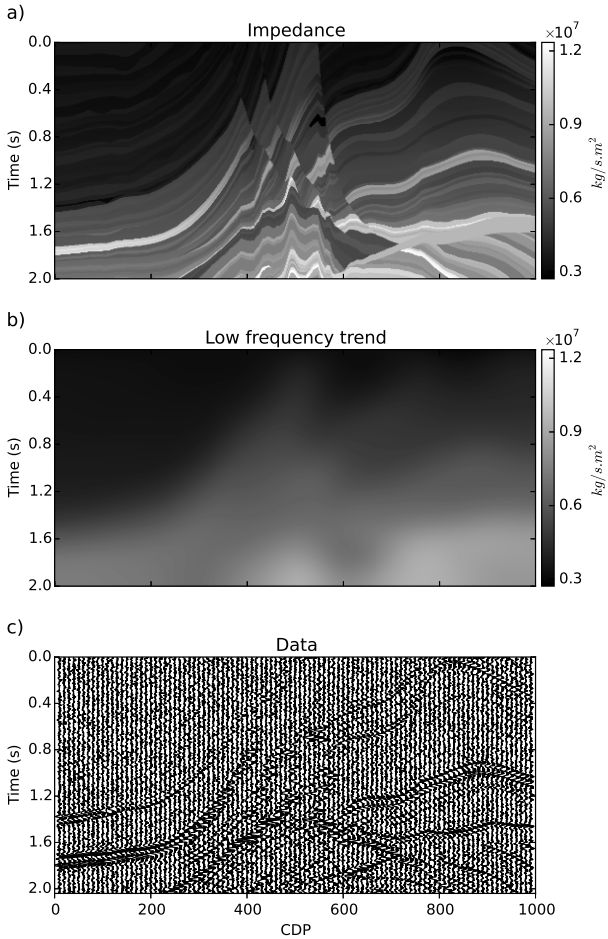
a)

b)

c)

Figure 1 – a) Actual AI model generated from the Marmousi2 elastic model, b) low-frequency trend obtained from the AI model using a low-pass filter and c) noisy seismic section with S/N=10 obtained from the AI model using a Ricker wavelet with $f_0 = 30$ Hz.

Examples

Synthetic data example

In this example we test the proposed algorithm on 2D synthetic data generated from the Marmousi2 elastic model (Martin G. S. & Marfurt, 2006). Figure 1 shows the AI of the model, the low-frequency trend obtained from smoothing the AI using a low-pass filter, and the corresponding seismic section. For the sake of clarity, Figure 1c shows one of every ten traces only. The data was generated using equations (1) and (2) with a sampling interval of 4 ms and a Ricker wavelet of central frequency $f_0 = 30$ Hz. We added uncorrelated band-limited Gaussian noise with $\sigma = \max(|\mathbf{S}|)/10$ (i.e. S/N=10).

To apply the proposed inversion algorithm to the dataset we first need to estimate the trade-off parameter μ. In general, the selection of this parameter depends on the noise level of the data at hand. If μ is too large the regularization term in equation (3) will be predominant over the misfit term, leading to solutions that might not honor the observed data.
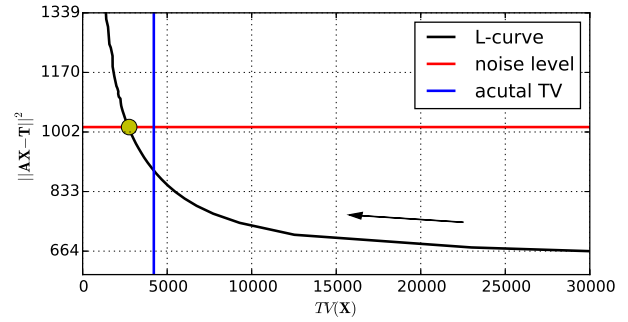

Figure 2 – Pareto curve after the inversion of the noisy data shown in Figure 1c using Algorithm 1, for various trial values of μ (the arrow show the direction of increasing μ). The horizontal red line indicate the noise level, the vertical blue line indicate the TV of the actual AI and the yellow dot denote the selected optimum μ according to the discrepancy principle.

Contrarily, if μ is too small the noisy data might be overfit and the solutions might be too smooth. It is important to note that the trade-off parameter is not unique, for there might exist a range of μ values for which the corresponding solutions honor the observed seismic data equally well. Often, the selection of a particular value is based on the analysis of the solutions and personal judgment, especially when the data noise is unknown (Farquharson & Oldenburg, 2004). Even so, there are various methods reported in the literature that can be used as guide to choose a value of μ. For instance, one can adopt the so-called L-curve criterion, the discrepancy principle, the generalized cross-validation criterion or the empirical Bayes method (Farquharson & Oldenburg, 2004; Malinverno & Briggs, 2004; van den Berg & Friedlander, 2008; Hennenfent et al., 2008).

In this numerical example we estimate μ using the discrepancy principle and the L-curve. This formulation is preferred when an estimation of the noise level is available. To this end, we construct the L-curve of the data, also known as Pareto curve (van den Berg & Friedlander, 2008; Hennenfent et al., 2008), by carrying out the inversion using various trial μ values and plotting the resulting $TV(\mathbf{X})$ versus misfit. Then, we choose the optimum μ as the one that minimizes the $TV(\mathbf{X})$ while the misfit remains less or equal than the noise level. Figure 2 shows the Pareto curves corresponding to the noisy seismic section shown in Figure 1c. The blue vertical line indicates the $TV(\mathbf{X})$ associated with the actual AI image shown in Figure 1a, the red horizontal line the noise level, and the black arrow the direction of increasing μ. As expected, the larger the μ, the smaller the $TV(\mathbf{X})$ and the larger the misfit, and *vice versa*. Following the discrepancy principle, we estimated the optimum values $\mu = 0.05$, denoted by the yellow circle.

It is worth noting that the L-curve does not cross the intersection of the noise level and the actual TV of the AI, which would be the optimal solution of the problem. This behavior is expected because, as it is well-known, the TV regularization tends to underestimate the amplitudes of the estimated solutions (Paragios et al., 2005). Several authors

have developed strategies to improve the amplitudes, such as the the L^1 fitting (Nikolova, 2002) or the iterative refinement (Osher et al., 2005). Also, the use of a debiasing step to adjust the amplitudes after the iterative process was proposed in the context of other regularized inversion problems (Figueiredo et al., 2007; Pérez et al., 2013). In this work, as the underestimation of the amplitudes resulted to be small, we choose not to apply any correction in benefit of the efficiency of the method.

Figure 3 shows the results of the inversion. We observe that the proposed method is capable of estimating a consistent and sharp AI image. The solution clearly resembles the actual AI showed in Figure 1a. The amplitudes were correctly recovered and follow the low frequency trend. The image shows good lateral continuity between CDPs, allowing the identification of the major geological structures. Also, considerable detail was recovered. Thin layers, faults and other discontinuities are easily identifiable. As shortcoming, some vertical discontinuities as the one located near the CDP 500 in the upper part of the Figure 1a were not correctly recovered. Nonetheless, this is expected as this structure is barely visible in the noisy seismic data shown in Figure 1c. The estimated seismic section is shown in Figure 3b, while Figure 3c shows the residual. Note that the estimated data honors the observed data quite well, being the residual only Gaussian noise.

Field data example

In this section we test the method using field data. Field data complicates the application of our method because not only the noise level is unknown, but also the source wavelet is often not available. We assume that the data has been properly processed to preserve amplitudes. The wavelet used in the inversion was estimated from the seismic section by assuming zero phase (Robinson & Treitel, 2002). The low-frequency trend was estimated using well log data near to the area where the data was acquired. In this example we estimate the trade-off parameter μ by trial and error. The field dataset consists of 400 traces with a sampling interval of 2 ms. The inversion was carried out in a time window of 0.5 s.

Figure 4 shows observed field data (panel a), the estimated AI (panel b) and the reconstructed seismic data after the inversion (panel c). We observe that the reconstructed data is quite similar to the actual data, showing that the estimated AI honors the observations very well.

Conclusions

In this work we presented a multichannel inversion algorithm to estimate sharp images of the acoustic impedance from poststack seismic data. When dealing with noisy and band-limited data, and due to the ill-posed nature of inverse problem, an appropriate regularization is needed to obtain meaningful solutions. To this end, we advocate the use of the total variation semi-norm, and adequate choice because it allows to handle edges properly. The cost function associated with the inversion resulted to be a convex function, because both the misfit between the calculated and
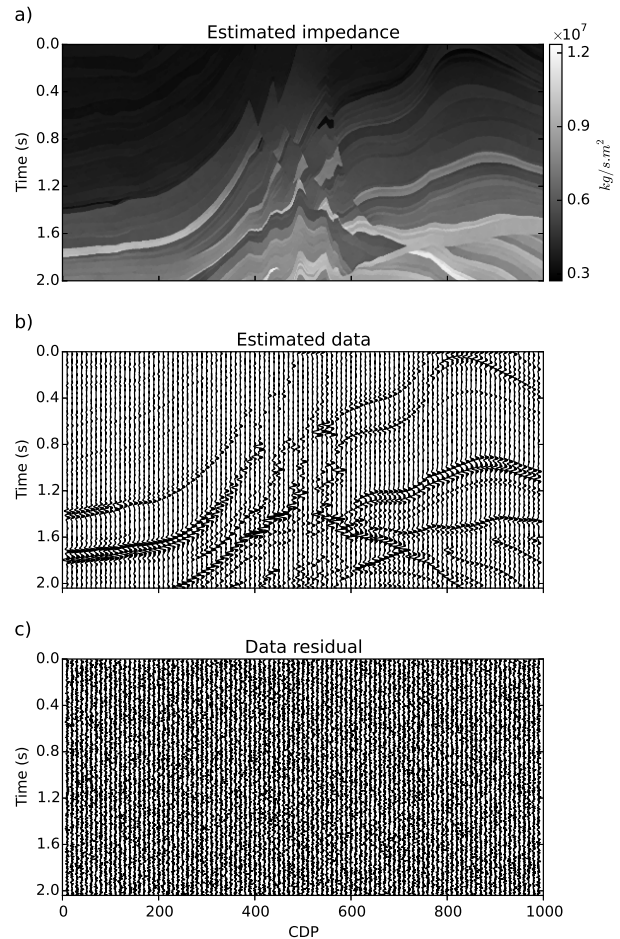

Figure 3 – *a) Estimated AI from the seismic data shown in Figure 1c using Algorithm 1, b) reconstructed seismic section from the estimated AI and c) residual between the actual and the reconstructed seismic data.*

observed data and the total variation are convex functions. This cost function was minimized very efficiently using a gradient-based algorithm. Since this algorithm relies only on matrix-vector multiplications the inversion method was economic in terms of computational cost and numerically stable. The strategy was very versatile and easy to apply, allowing us to easily incorporate the low frequency trends required to recover the corresponding acoustic impedance. We tested the algorithm on 2D synthetic and field data. In the synthetic data example, before applying the algorithm, we used the discrepancy principle and the Pareto curve to estimate the optimum trade-off parameter μ. This strategy proved to be useful since we could also see that the solutions are affected by an underestimation of the amplitudes, an issue that is common when using the total variation as regularization term. Both examples showed that the inversion algorithm is capable of obtaining sharp AI images that honor the observed data, showing good lateral continuity and properly estimated magnitudes while preserving the low frequency trend.
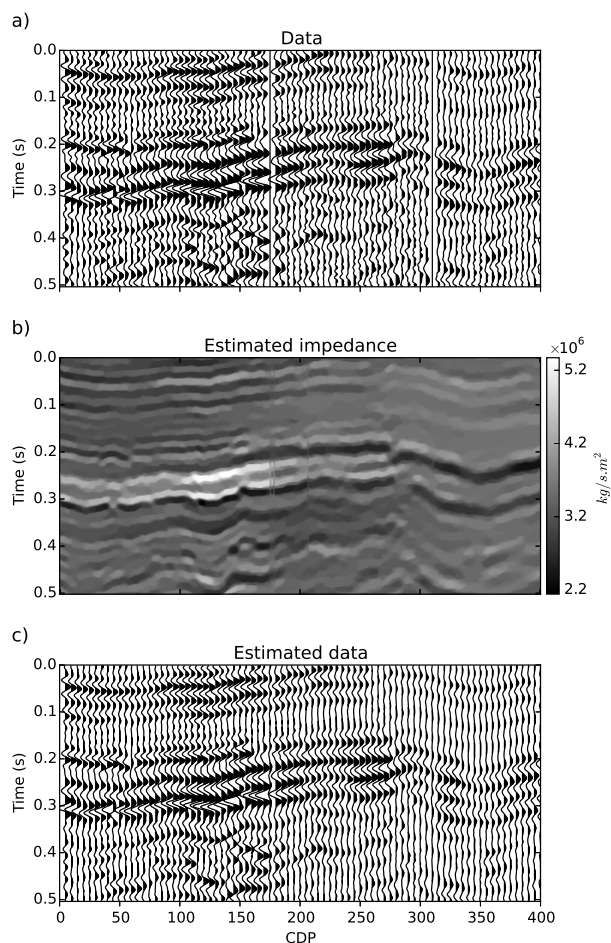
Figure 4 – a) Actual field data, b) estimated impedance and c) reconstructed seismic data.

Acknowledgments

References

BECK A. & M. TEBOULLE, 2009a, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems: IEEE Transactions on Image Processing, 18, 2419–2434.

———, 2009b, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM Journal on Imaging Sciences, 2, 183–202.

CHAMBOLLE A., 2004, An algorithm for total variation minimization and applications: Journal of Mathematical Imaging and Vision, 20, 89–97.

COOKE D. & W. SCHNEIDER, 1983, Generalized linear inversion of reflection seismic data: Geopphysics, 46, 665–676.

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

FIGUEIREDO M. A. T., R. D. NOWAK & S. J. WRIGHT, 2007, Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems: IEEE journal of selected topics in signal processing, 1, 586–597.

GHOLAMI A., 2015, Nonlinear multichannel impedance inversion by total-variation regularization: Geophysics, 80, R217–R224.

———, 2016, A fast automatic multichannel blind seismic inversion for high-resolution impedance recovery: Geophysics, 81, V357–V364.

GHOLAMI A. & M. D. SACCHI, 2013, Fast 3d blind seismic deconvolution via constrained total variation and gcv: SIAM Journal on Imaging Sciences, 6, 2350–2369.

GOLDSTEIN T. & S. OSHER, 2009, The split bregman method for l1-regularized problems: SIAM Journal on Imaging Sciences, 2, 323–343.

GRAMFORT A., D. STROHMEIER, J. HAUEISEN, M. S. HÄMÄLÄINEN & KOWALSKI A., 2013, Time-frequency mixed-norm estimates: Sparse m/eeg imaging with non-stationary source activations: NeuroImage, 70, 410 – 422.

HENNENFENT G., E. VAN DEN BERG, M. P. FRIEDLANDER & F. J. HERMANN, 2008, New insights into one-norm solvers from the Pareto curve: Geophysics, 73, 23–26.

LARSON R. & B. H. EDWARDS, 1999, Elementary linear algebra, 4th ed.: Houghton Mifflin Company.

MALINVERNO A. & V. A. BRIGGS, 2004, Expanded uncertainty quantification in inverse problems: Hierarchical bayes and empirical bayes: Geophysics, 69, 1005–1016.

MARTIN G. S., R. WILEY & K. J. MARFURT, 2006, Marmousi2: an elastic upgrade for Marmousi: The Leading Edge, 25, 156–166.

NIKOLOVA A., 2002, Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers: SIAM Journal on Numerical Analysis, 40, 965–994.

OLDENBURG D. W., T. SCHEUER & S. LEVY, 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, 48, 1318–1337.

OSHER S., M. BURGER D. GOLDFARB, J. XU & W. YIN, 2005, An iterative regularization method for total variation-based image restoration: Multiscale Modeling & Simulation, 4, 460–489.

PALOMAR, D. P. & Y. C. ELDAR, eds., 2010, Convex optimization in signal processing and communications: Cambridge University Press.

PARAGIOS, N., CHEN, Y. & O, D. FAUGERAS, 2005, Handbook of mathematical models in computer vision: Springer US.

PÉREZ D. O., D. R. VELIS & M. D. SACCHI, 2013, High-resolution prestack seismic inversion using a hybrid FISTA least-squares strategy: Geophysics, 78, R185–R195.

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

ROBINSON, E. A., AND S. TREITEL, 2002, Geophysical signal analysis: SEG.

RUDIN I. L., S. OSHER & E. FATEMI, 1992, Nonlinear total variation based noise removal algorithms: Physica D: Nonlinear Phenomena, 60, 259 – 268.

TARANTOLA A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM. (Available at the Web).

ULRYCH T. J. & M. D. SACCHI, 2005, Information-based inversion and processing with applications: Elsevier.

VAN DEN BERG E. & M. FRIEDLANDER, 2008, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal of Scientific Computing, 31, 890–912.

VELIS D. R., 2003, Estimating the distribution of primary reflection coefficients: Geophysics, 68, 1417–1422.

———, 2008, Stochastic sparse-spike deconvolution: Geophysics, 73, R1–R9.

WALDEN A. T. & J. W. J. HOSKEN, 1986, The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution: Geophysical Prospecting, 34, 1038–1066.