



Sparse Multichannel Blind Deconvolution with Regularization by Denoising

Thonia Cardoso Senna, Kenji Nose-Filho, Renato da Rocha Lopes

Copyright 2021, SBGf - Sociedade Brasileira de Geofísica.

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society, held Online, in Brazil, November 8-11, 2021.

Contents of this paper were reviewed by the Technical Committee of the 17th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of The Brazilian Geophysical Society is prohibited.

Abstract

In this paper, we propose the regularization of the multichannel deconvolution technique proposed by Xu et al. (1995) applying a scheme that relies on the Regularization by Denoising Engine proposed by Romano et al. (2016). We compare our approach with the widely known SMBD method of Kazemi and Sacchi (2014), which is robust to noise and provides a sparse solution. However, it also affects the relationship of the reflectors amplitudes, in such a way that many times smaller amplitude reflectors may disappear. Our technique was tested with synthetic and real data examples. The results show that, in addition to the method being robust to noise, its application provide a better image, in terms of preservation of events with lower amplitudes as well as in terms of numerical stability when compared to SMBD.

Introduction

When seismic data is generated, sources that have a finite length and frequency content are used, which, along with the addition of noise, reduces our ability to image layers. Theoretically, the reflectivity series can be calculated by removing the effect of wavelet from the observed seismic trace (Robinson and Treitel, 1980). In geophysical processing, however, direct wavelet measurements are not always available and in its absence, we rely exclusively on the measured data and on a minimal amount of hypotheses concerning the signal of interest (Romano et al., 2010). In these situations, we resort to unsupervised or blind deconvolution techniques, which appears to be a powerful tool for dealing with practical situations (Zhang and Gao, 2011).

Blindly estimating the signal of interest is an ill-posed problem, as, for instance, in the seismic case, infinite combinations of wavelets and reflectivity series may result in the same trace. Thus, one of the main challenges in this case is to establish *a priori* hypotheses about the desired signal in order to achieve stable results of practical value. Several methodologies have been proposed for this purpose, many of which are based on severely limiting assumptions and hypotheses. Normally, these assumptions only work in an ideal scenario (e.g. considering the seismic wavelet a minimum phase sequence, or the absence of noise) and are not consistent with the practical reality of geophysical acquisitions, greatly compromising the results

(Dondurur, 2010). Several methods emerged later in order to circumvent these requirements, thus producing more realistic results. Among them, we can highlight the Minimum Entropy Deconvolution (MED), proposed by Wiggins (1978), which, by assuming that the reflectivity is a sparse signal, is able to estimate a nonminimum phase filter by maximizing a measure of sparsity of the seismic trace (Donoho, 1981).

If, on the one hand, classical approaches are implemented on a trace-by-trace basis, in this work we suggest the application of a multichannel framework, which consider that, e.g., in a common shot, the wavelet present in each trace is the same. This is a problem that, in signal processing, can be modeled as a SIMO (single input, multiple outputs) system, or equivalently, several outputs (the seismic traces) corresponding to a single input (source wavelet) convolved with each channel (reflectivity series), as illustrated in Figure 1. The study of the applications of SIMO systems to the deconvolution of seismic data was introduced in Rietsch (1997a) and Rietsch (1997b).

Besides being a reasonable assumption, supposing that the wavelets are the same has some very interesting consequences. For instance, it is possible to show that, under some further conditions, we may be able to find an algebraic formula that perfectly removes the wavelet, unlike the SISO (single input, single output) methods, which even under ideal conditions, may not guarantee exact recovery of desired signal (Romano et al., 2010). However, as before, these conditions are hard to verify in the practice of seismic signal processing. Thus, the use of multichannel methods for seismic signals requires some careful thoughts and solutions (Nose-Filho et al., 2018).

In their work, Nose-Filho et al. (2018) discussed how the large similarity between neighboring reflections may lead to ill-posedness in multichannel problems and detailed how some works in the literature use regularization methods to find good solutions. An alternative method to tackle this issue is proposed in this work and leverages the results achieved by Romano et al. (2016), where they applied regularization by denoising (RED) to effectively treat general inverse problems. Their technique is guaranteed to converge to the globally optimal result under some assumptions. They demonstrate state-of-the-art results in image deblurring and super-resolution problems.

Besides the important properties mentioned above, SIMO channel frameworks have been studied by many authors due to its remarkable ability to recover blurred signals: in a SIMO model, if one of the channels happen to severely degrade the signal, other received versions may still provide the necessary information to correctly estimate the original signal (Romano et al., 2010). In this paper, we will combine these two attractive strategies, using regularization based on denoising engines for the multichannel blind deconvolution problem. We show, via synthetic and real

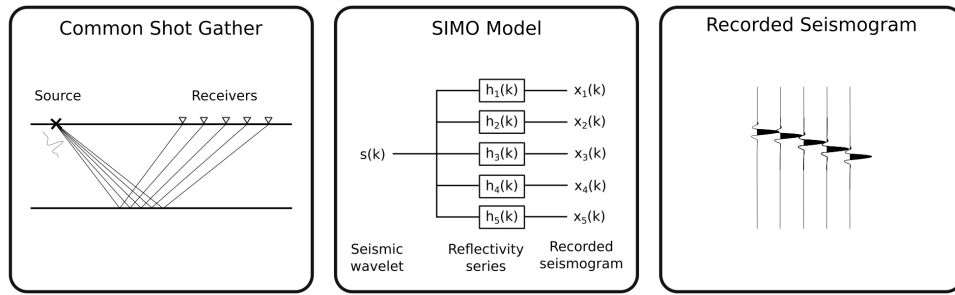


Figure 1: In a SIMO system, the same information (seismic wavelet) is transmitted through different subchannels, since they are all recordings of the same shot. Notice that all received sequences will be distinctly distorted versions of the same message, which accounts for a certain signal diversity. Therefore, it is reasonable to assume that more information about the transmitted signal will be available at the receiver end. It is like looking at the same landscape through different blurry windows: each sight reveals you a different detail that perhaps is not so clear in the others (Romano et al., 2010).

land seismic data, that this solution is indeed effective for blind seismic deconvolution, preserving small events, while also being numerically stable and robust to noise.

In the first section, we describe how exactly SIMO models fit in the framework of the seismic deconvolution of a common shot. In the second section we introduce RED general definitions and present some of its main ideas. In the third section we describe a way to obtain a solution to our problem by performing multichannel blind deconvolution using RED. In the fourth section, we use the theoretical results from previous sections to estimate reflectivity series from both synthetic and real land seismic data. Finally, we concludes this study with our achievements and general conclusions.

Multichannel Blind Deconvolution

We begin by presenting the noiseless SIMO model shown in Figure 1. We model each trace $x_p(k)$ recorded at the p th receiver as the convolution of the source wavelet $s(k)$ with the reflectivity series $h_p(k)$ seen by the signal that reaches this receiver, i.e., $x_p(k) = h_p(k) * s(k)$, $p = 1 \dots J$. In terms of z-transforms, we write

$$X_p(z) = H_p(z)S(z), \quad p = 1 \dots J \quad (1)$$

Now, consider a pair of traces $x_p(k)$ and $x_q(k)$. According to Equation 1, we can write, for each of the traces, $X_p(z) = H_p(z)S(z)$ and $X_q(z) = H_q(z)S(z)$. Then, by isolating the source wavelet, which is a common factor to both of the traces, we have

$$S(z) = \frac{X_p(z)}{H_p(z)} = \frac{X_q(z)}{H_q(z)}. \quad (2)$$

Consequently

$$X_p(z)H_q(z) - X_q(z)H_p(z) = 0. \quad (3)$$

In the time domain, each term in Equation 3 can be seen as the convolution of a trace with a reflectivity function, which $\forall p, q$, can be rewritten in matrix notation as follows:

$$\mathbf{X}_p \mathbf{h}_q - \mathbf{X}_q \mathbf{h}_p = \mathbf{0} \quad (4)$$

where \mathbf{X}_p is the convolution matrix of the p th trace:

$$\mathbf{X}_p = \begin{pmatrix} x_p(0) & 0 & \dots & 0 & 0 \\ x_p(1) & x_p(0) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & x_p(M-1) & x_p(M-2) \\ 0 & 0 & 0 & 0 & x_p(M-1) \end{pmatrix} \quad (5)$$

and $\mathbf{h}_q = [h_q(0), \dots, h_q(K-1)]^T$ is the vector formed with the samples of the q th reflectivity function. The combination of all possible equations leads to the following homogeneous system of equations:

$$\mathbf{A} \mathbf{h} = \mathbf{0} \quad (6)$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{X}_2 & -\mathbf{X}_1 & & & & & & & \\ \mathbf{X}_3 & & -\mathbf{X}_1 & & & & & & \\ \mathbf{X}_4 & & & -\mathbf{X}_1 & & & & & \\ \vdots & & & & & & \ddots & & \\ & \mathbf{X}_3 & -\mathbf{X}_2 & & & & & & \\ & \mathbf{X}_4 & & -\mathbf{X}_2 & & & & & \\ & \mathbf{X}_5 & & & -\mathbf{X}_2 & & & & \\ \vdots & & & & & & & & \ddots \end{pmatrix} \quad (7)$$

and $\mathbf{h} = [\mathbf{h}_1^T, \dots, \mathbf{h}_J^T]^T$. The classical formulation of multichannel deconvolution applied to seismic data, proposed by Rietsch (1997a) estimates the reflectivity by estimating the eigenvector associated to the minimum nonzero eigenvalue of $\mathbf{A}^T \mathbf{A}$. A small amount of noise in the data makes the solution impractical for real data applications Rietsch (1997b). Furthermore, if the signals of interest have zeros in common, the method can lead to a solution that does not correspond to the true solution, as pointed out in Nose-Filho et al. (2018).

Regularization by Denoising

In the present work, we leverage the impressive results of Romano et al. (2016), where they introduced the image denoising engine. They were inspired by the fact that the existing denoising methods are so sophisticated and effective nowadays, and that their great performance could be exploited to treat other tasks in image processing. In fact, their proposal is an alternative to Chan et al. (2016), which

had already proposed the Plug-and-Play prior (PPP), which solves any inverse problem by sequentially applying image denoising steps. However, PPP relies heavily on ADMM optimization technique, whose parameters may be difficult to tune. On the other hand, the use of Regularization by Denoising (RED) from Romano et al. (2016) yields a more powerful and more flexible framework for achieving the same goal, centered in the idea that a denoiser can serve as a regularizer.

They propose an explicit image-adaptive regularization functional. This is in contrast to PPP, which uses an implicit regularization functional. With their method, we are free to choose the iterative optimization procedure for minimizing the mentioned functional, not being tightly coupled to one specific technique, as in the case of the Plug-and-Play Prior. RED makes it possible incorporating any image denoising algorithm, treat general inverse problems very effectively, and is guaranteed to converge to the globally optimal result under some assumptions on the denoiser. The remarkable results that this technique has achieved in image deblurring and super-resolution problems led us to develop an adaptation to geophysical scenario as well.

More specifically, the regularization term they propose is of the following form:

$$\mathcal{R}(\mathbf{h}) = \frac{1}{2} \mathbf{h}^T [\mathbf{h} - f(\mathbf{h})] \quad (8)$$

in which $f(\mathbf{h})$ is the denoised version of the candidate solution \mathbf{h} . The resulting penalty is proportional to the inner-product between this candidate solution and its denoising residual. Surprisingly, under mild assumptions on $f(\cdot)$ ¹, it was shown that the gradient of the regularization term is simple, given as the denoising residual, $\mathbf{h} - f(\mathbf{h})$. Therefore, armed with this regularization expression, they show that any inverse problem can be handled with a simple gradient descent method while calling the denoising engine iteratively (Romano et al., 2016).

RED Multichannel Blind Deconvolution

In order to make multichannel deconvolution a feasible strategy for real data, which have noise and yield ill-conditioned problems, we propose an adaptation for this process in which we optimize a data-fidelity term based on the acquisition model, as is done classically in the area. With the addition of noise in our model, the expressions (1), (3), (4) and (6) become:

$$X_p(z) = H_p(z)S(z) + N_p(z) \quad p = 1, \dots, J \quad (9)$$

analogously, for a pair of traces p and q :

$$X_p(z)H_q(z) - X_q(z)H_p(z) = N_p(z)H_q(z) - N_q(z)H_p(z) \quad (10)$$

or, in matrix notation:

$$\mathbf{X}_p \mathbf{h}_q - \mathbf{X}_q \mathbf{h}_p = \mathbf{N}_p \mathbf{h}_q - \mathbf{N}_q \mathbf{h}_p \quad (11)$$

which can be expressed in a condensed form as:

$$\mathbf{A} \mathbf{h} = \mathbf{e} \quad (12)$$

¹In the paper, the authors pose two necessary conditions on $f(\cdot)$, namely (Local) Homogeneity and Strong Passivity

We propose to find a solution \mathbf{h} that minimizes the ℓ_2 norm of \mathbf{e} combined with the regularization by denoising defined in Romano et al. (2016), where they use denoising engines in defining the regularization of the inverse problem. To summarize, we want to minimize the following cost function with a constraint to avoid the trivial solution (i.e., $\mathbf{h}^T \mathbf{h} = 1$):

$$\begin{aligned} \hat{\mathbf{h}} = \underset{\mathbf{h}}{\operatorname{argmin}} \quad & \frac{1}{2} \|\mathbf{A} \mathbf{h}\|^2 + \lambda \mathbf{h}^T (\mathbf{h} - f(\mathbf{h})) + \beta \mathbf{h}^T (\mathbf{h} - g(\mathbf{h})), \\ \text{subject to} \quad & \mathbf{h}^T \mathbf{h} = 1 \end{aligned} \quad (13)$$

In the above formulation we are going to make use of two different denoising engines $f(\cdot)$ and $g(\cdot)$. $f(\mathbf{h})$ is the output of a simple prediction error filter, with adjustable prediction lag $D + 1$, making it possible to adjust the degree of whitening of our desired reflectivity series. The third term, on the other hand, has a built-in sparsity-promoting denoising function $g(\mathbf{h})$, given by the hard threshold function, specified by a threshold $\varepsilon = \mu \cdot \max(\operatorname{abs}(h))$:

$$g(x) = \begin{cases} 0, & |x| < \varepsilon \\ x, & |x| \geq \varepsilon \end{cases} \quad (14)$$

The trade-off parameters λ and β balance the importance of the term related to whitening and the term related to sparsity.

Both functions satisfy the two necessary conditions described in Romano et al. (2016). For the optimization, we adopted the same steepest descent algorithm used in Kazemi and Sacchi (2014).

Results

We firstly present some simulations with synthetic data as it is essential that we know the value of the real reflectivity in order to evaluate the method's performance during its development phase. The quality of the results will be measured with the mean Pearson correlation coefficient (PCC), which is also a measure of orthogonality, given by:

$$\rho = \frac{1}{J} \sum_j \frac{\mathbf{h}_j^T \hat{\mathbf{h}}_j}{\|\mathbf{h}_j\| \|\hat{\mathbf{h}}_j\|} \quad (15)$$

Another metric employed here to evaluate the preservation of the amplitude of the events is the mean absolute percentage error (MAPE) of the original events and the estimated ones, only in their original positions considering that both signals are normalized with their maximum absolute value. So, the MAPE is calculated as:

$$\operatorname{MAPE} = 100 \frac{1}{\sum_j \sum_k r_j(k)} \sum_j \sum_k \frac{|h_j(k) - \hat{h}_j(k) r_j(k)|}{|h_j(k)|} \quad (16)$$

In this case, we consider that each signal has been previously normalized with their maximum absolute value and the signal $r_j(k)$ can be seen as a mask function, with ones in the positions containing a reflector and zeros otherwise. This could be obtained by simply:

$$r_j(k) = \begin{cases} 0, & h_j(k) = 0 \\ 1, & h_j(k) \neq 0 \end{cases} \quad (17)$$

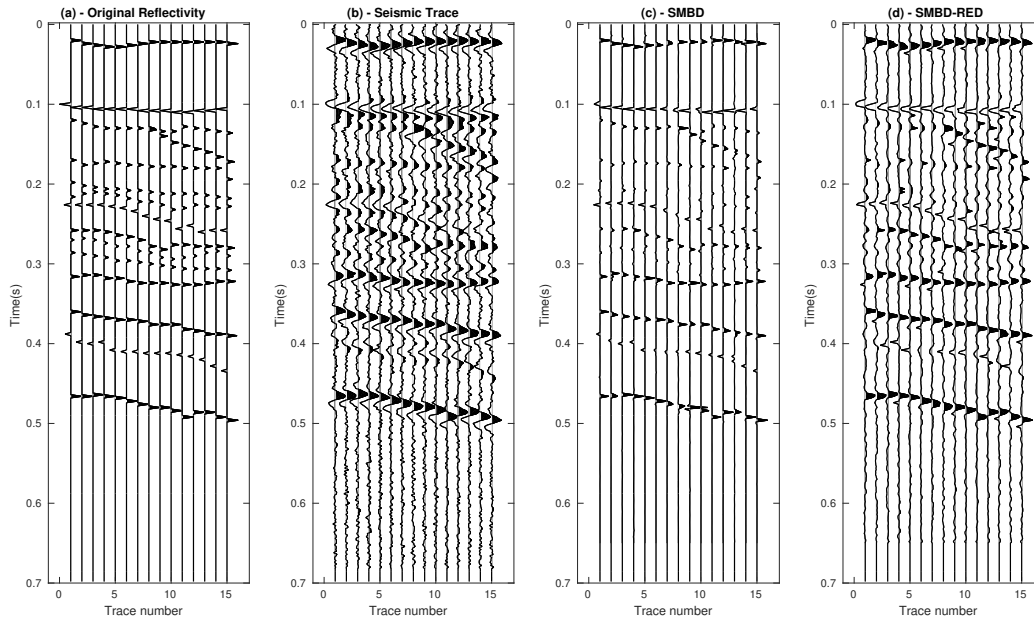


Figure 2: Deconvolution results for SNR= 9dB. (a) Original reflectivity. (b) Generated seismic trace. Deconvolution results for (c) SMBD- 200 iterations and (d) SMBD-RED- 200 iterations.

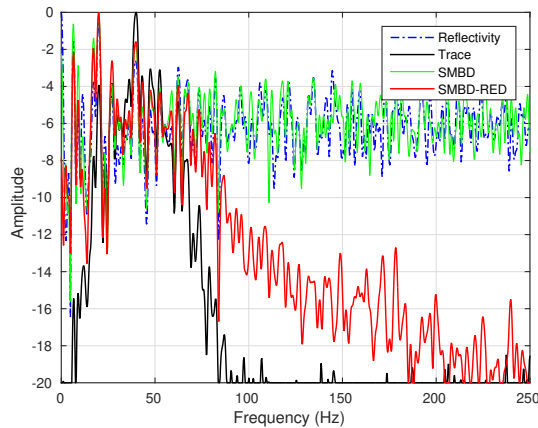


Figure 3: Frequency spectrum for the data in Figure 2 (a), Figure 2 (b), Figure 2 (c) and Figure 2 (d).

To test our method, we first run a synthetic example where 15 traces were generated with a sampling frequency of 500 Hz using the reflectivity shown in Figure 2 (a). The trace shown in Figure 2 (b) is the convolution of the reflectivity with a Ricker wavelet of central frequency 40Hz with 50 degrees of phase shift plus additional white gaussian noise (AWGN) with 9dB of SNR. We tried to recover our reflectivities directly from the traces, using SMBD and SMBD-RED. In this context, several simulations were carried out, in which various combinations of parameters were tested. Here we present the best parameters found by our simulations (that is, those with the best balance between PCC and MAPE), for both SMBD and SMBD-RED. For the SMBD algorithm, the recovered reflectivity is shown in Figure 2 (c).

Its regularization parameter was set to 10, the smoothed norm parameter was set to 0.0005 and 200 iterations were performed. A PCC of 0.771 and a MAPE of 72.704 were obtained. For the SMBD-RED technique, the recovered reflectivity is shown in Figure 2 (d). The regularization parameter λ was set to 50, β was set to 1000, μ was set to 0.4 and 200 iterations were performed. If, on the one hand, the PCC value of 0.608 obtained by SMBD-RED (smaller than the SMBD) could make us believe that this method has no advantages over the first, the MAPE for SMBD-RED method (41.203) could show us the opposite: in fact, we can see in Figure 2 (d) that smaller and close events are being preserved, while SMBD (Figure 2 (c)) makes them disappear, affecting the interpretation of geological events.

In Figure 3 we illustrate the smooth frequency spectrum of the original reflectivity series, the convolved traces and the estimations with SMBD and SMBD-RED. As we can see, despite of the fact that SMBD produces a more flat frequency spectrum, the SMBD-RED spectrum better preserves the shape of the trace and reflectivity spectrum in the narrowband frequencies close to the wavelet central frequency.

To further illustrate the performance of our technique, we applied the method in the seismic data available in Schleicher (2012), corresponding to a 2D terrestrial line from National Petroleum Reserve (NPRA), acquired in the northern region of Alaska. More details about the data can be checked in Schleicher (2012). This data was pre-processed based on scripts available in Schleicher (2012), in order to remove poor quality traces and parts of the data that could affect the processing. The data has a 4 ms sample rate and we run both algorithms on 525 traces and 800 samples, starting at 1s. We performed both SMBD and SMBD-RED deconvolution by separating the data into

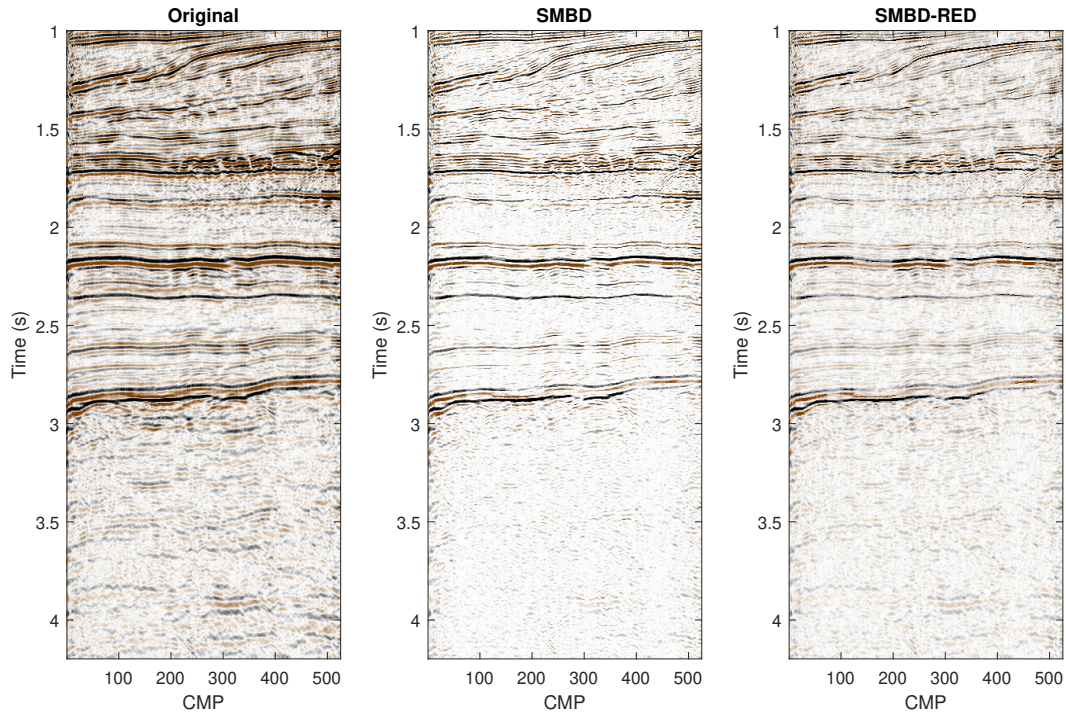


Figure 4: Deconvolution results for the 2D terrestrial line from National Petroleum Reserve (NPRA), available in Schleicher (2012). (a) Original data, (b) SMBD- 150 iterations and (c) SMBD-RED- 150 iterations.

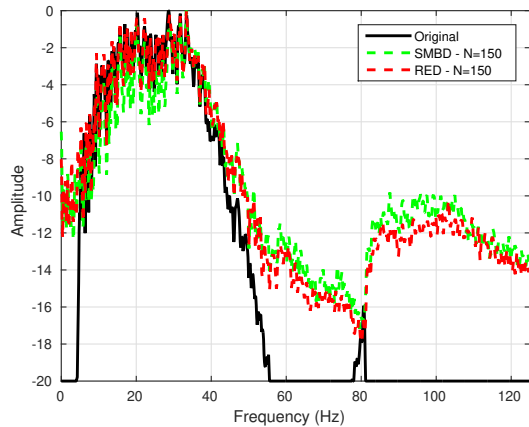


Figure 5: Frequency spectrum for the data in Figure 4 (a), Figure 4 (b) and Figure 4 (c).

windows of 25 traces and 800 samples, totaling 21 windows. Figure 4 shows the input data and the deconvolution results for SMBD and SMBD-RED. It is clear from these results that the proposed algorithm better preserves weaker reflections. The consistency and the quality of the estimates is also easy to observe. Several combinations of parameters were tested, among which we selected one that obtained a good balance between PCC and MAPE:

- SMBD - regularization parameter was set to 10, the smoothed norm parameter was set to 0.001 and 150

iterations were performed.

- SMBD-RED - λ was set to 200, β was set to 1000 and μ was set to 0.4. 150 iterations were performed

In Figure 5 we illustrate the smooth frequency spectrum of the seismic trace and the estimations with SMBD and SMBD-RED. As already observed in the synthetic simulations, the SMBD-RED spectrum better preserved the shape of the trace spectrum in the lower frequencies.

Conclusions

Different authors have exploited the algebraic structure of SIMO model to eliminate a blurring kernel and solve a large set of linear equations directly for the signal of interest, and so, perform deconvolution in a variety of applications, such as biomedical imaging, geophysics, remote sensing, microscopical and astronomical imaging, channel equalization, speech dereverberation, etc.. Recently, Kazemi and Sacchi (2014) have shown that, for seismic signals, pursuing sparse solutions of the wavelet-free linear equations would be feasible and produce results that are more robust to noise. This work gave rise to other variants such as (Nose-Filho et al., 2016), (Kazemi et al., 2016), and (Iqbal et al., 2019). Our main motivation here was proposing a method that is simultaneously robust to noise and that could preserve the relationship between reflectors amplitudes, without making smaller events disappear in the results. By applying RED, an approach originally proposed for image processing applications, we could get high quality seismic deconvolution results, achieving our goals in terms of noise

and preservation of smaller events. Another important property of the proposed method is that it does not require careful parameter tuning. In fact we observed that small changes in parameters in our simulations did not significantly alter the results.

References

- Chan, S. H., Wang, X., and Elgendy, O. A., 2016, Plug-and-play admm for image restoration: Fixed-point convergence and applications: *IEEE Transactions on Computational Imaging*, **3**, no. 1, 84–98.
- Dondurur, D., feb 2010, An approximation to sparse-spike reflectivity using the gold deconvolution method: *Pure and Applied Geophysics*, **167**, no. 10, 1233–1245.
- Donoho, D. L., 1981, *On minimum entropy deconvolution*: Academic Press, pages 656–608.
- Iqbal, N., Liu, E., McClellan, J. H., and Al-Shuhail, A. A., 2019, Sparse multichannel blind deconvolution of seismic data via spectral projected-gradient: *IEEE Access*, **7**, 23740–23751.
- Kazemi, N., and Sacchi, M. D., sep 2014, Sparse multichannel blind deconvolution: *GEOPHYSICS*, **79**, no. 5, V143–V152.
- Kazemi, N., Gholami, A., and Sacchi, M., may 2016, Modified sparse multichannel blind deconvolution: Modified sparse multichannel blind deconvolution:, EAGE Publications BV, 78th EAGE Conference and Exhibition 2016.
- Nose-Filho, K., Takahata, A., Lopes, R., and Romano, J., 2016, A fast algorithm for sparse multichannel blind deconvolution: *Geophysics*.
- Nose-Filho, K., Takahata, A. K., Lopes, R., and Romano, J. M. T., mar 2018, Improving sparse multichannel blind deconvolution with correlated seismic data: Foundations and further results: *IEEE Signal Processing Magazine*, **35**, no. 2, 41–50.
- Rietsch, E., 1997a, Euclid and the art of wavelet estimation, part I: basic algorithm for noise-free data: *Geophysics*, **62**, 1931–1938.
- 1997b, Euclid and the art of wavelet estimation, part II: robust algorithm and field-data examples: *Geophysics*, **62**, 1939–1946.
- Robinson, E., and Treitel, S., 1980, *Geophysical signal analysis*: New York.
- Romano, J. M. T., Attux, R., Cavalcante, C. C., and Suyama, R., 2010, *Unsupervised signal processing: channel equalization and source separation*: CRC Press.
- Romano, Y., Elad, M., and Milanfar, P., 2016, The little engine that could: Regularization by denoising (RED): *CoRR*, **abs/1611.02862**.
- Schleicher, K., 2012, Open data/open source: Seismic unix scripts to process a 2d land line:.
- Wiggins, R. A., 1978, Minimum entropy deconvolution: *Geoexploration*, **16**, no. 1-2, 21–35.
- Xu, G., Liu, H., Tong, L., and Kailath, T., 1995, A least-squares approach to blind channel identification: *IEEE Transactions on Signal Processing*, **43**, 2982–2993.
- Zhang, B., and Gao, J.-H., 2011, Blind deconvolution of seismic data using f-divergences: *Entropy*, **13**, no. 9, 1730–1745.

Acknowledgments

Thonia Cardoso Senna would like to thank the National Council for Scientific and Technological Development (CNPq) for funding this research.