Seismic regularization using Matching Pursuit and ALFT: A comparative study

Yuri Shalom de Freitas Bezerra (UFRN), German Garabito (UFRN) and Liacir dos Santos Lucena (UFRN).

Abstract

The application of seismic data regularization and interpolation methods is critical to many processing steps, such as noise suppression, wave equation based migration and modern inversion methods. The Anti-Leakage Fourier Transform (ALFT) and Matching Pursuit (MP) regularization methods use an iterative procedure to calculate the spectrum of irregularly sampled data. The estimated spectrum is used to synthetize data at missing and new spatial positions. In this work these two methods are compared. Specifically, the implementation made by Jahanjooy *et al.* (2016) of the ALFT is compared to a simpler version of the MP method presented by Hollander *et al.* (2012). The 1D and 2D examples demonstrate the similarity of both methods in terms of results. An example of a 2D regularization of seismic data applied to the Marmousi data set clearly demonstrates that the MP algorithm is faster than the ALFT.

Introduction

In a real seismic survey the acquired data are often irregularly sampled in space due to many causes such as presence of valleys and rivers, equipment failures and acquisition costs. Interpolation and reconstruction method can be used to regularize irregular data. For this purpose one can adopt methods based on prediction error filtering (Naghizadeh and Sacchi, 2010), methods based on Fourier synthesis operators (Liu and Sacchi, 2004) or methods based on matrix completion (Naghizadeh and Sacchi, 2013).

This paper concentrates on the comparison of two Fourier synthesis methods. More specifically, we will compare two greedy algorithms: AFLT (Xu *et al.*, 2005) and MP (Mallat and Zhang, 1993). Both methods can interpolate data with irregular spatial attributes. They can also cope with aliasing by establishing strategies for wave-number selection (Schonewille *et al.*, 2013).

The Anti-Leakage Fourier Transform (ALFT) was developed by Xu *el al.* (2005). ALFT is an iterative algorithm and is slower compared to methods that use the Fast Fourier Transform (FFT) algorithm, since it calculates the Discrete Fourier Transform (DFT) in each iteration. However, a faster version that only calculates the DFT only once (Jahanjooy *et al.*, 2016) could also be implemented. The ALFT method has been expanded to high dimensions (Xu *el al.*, 2010).

Another technique very similar to the ALFT is the Matching Pursuit (MP) interpolation, also called Matching Pursuit Fourier Interpolation (MPFI). These methods possess similar advantages and disadvantages.

Different versions of MP interpolation have been proposed. For instance, Hollander *et al.* (2012) proposed an orthogonal version of the MP (OMP), which uses the least squares to fit all coefficients that were found at a given iteration. These authors also compare OMP to the ALFT. Schonewille *et al.* (2013) investigate the problem of interpolating aliased data via MP. They solved the problem of choosing the maximum energy Fourier coefficients from aliased events by using weights (or priors) derived from low frequency un-aliased data.

The objective of this work is to perform a comparative study of ALFT and MP methods, through the development of their algorithms and applications in synthetic data and real land data of Brazilian basins. Here are presented the preliminary results results comparing ALFT and MP interpolation with 1D and 2D harmonic examples. We also provide an example portraying reconstruction of 2D data using the Marmousi data set.

Theory and Method

Anti-Leakage Fourier Transform (ALFT)

The ALFT objective is to estimate the Fourier coefficients of an irregularly sampled data (signal). Suppose that the recorded time samples are on a regular grid, the Fast Fourier Transform (FFT) algorithm is recommended to transform the data from the time domain to the frequency domain. Since the grid is irregular in space, the non-uniform DFT (Xu *et al.,* 2005) is used to map data from space to wavenumber.

Consider a 1D signal data with N_p recorded samples in space $x_l = [x_1, x_2, x_3, \dots, x_{N_p}]$. The ALFT technique retrieves the amplitudes at wavenumbers $k = [k_1, k_2, k_3, \dots, k_{N_k}]$ via

1) Compute the weights $w(x_l)$ and initialize all the Fourier components to zero.
2) Compute the non-uniform Fourier components of the data using equation (1).

$$\hat{f}^s(k) = \sum_{l=1}^{N_p} w(x_l) f^s(x_l) e^{-2\pi i k x_l} \qquad (1)$$

3) Select the Fourier coefficient with the maximum energy and update data by subtracting the contribution of selected coefficient, equation (2), from the input data, equation (3).

$$f^k(x_l) = \hat{f}_{k_p}(k) e^{2\pi i k x_l} \qquad (2)$$

$$f^{s+1}(x_l) = f^s(x_l) - f^k(x_l) \qquad (3)$$

4) Repeat steps 2 and 3 until the updated residual input data to be less than a tolerance.
5) Finally, use all the coefficients selected in the optimal Fourier spectrum to reconstruct the signal at any desired sampling points.

For a more accurate calculation of the Fourier coefficients it is necessary to adopt weights. The regular and irregularly sampled data in the DFT summation need normalization by the weighting function $w(x_l)$ where this is in function of the spatial distribution $x_l - x_{l-1}$.

Matching Pursuit (MP)

The implementation of the MP and ALFT algorithms are very similar. The MP algorithm can be described as follows:

Initialize all the Fourier components to zero.

1) Compute the Fourier components of the data using equation (4).

$$\hat{f}^s(\boldsymbol{k}) = \sum_{l=1}^{N_p} f^s(\boldsymbol{x}_l) e^{-2\pi i \boldsymbol{k} \quad l} \qquad (4)$$

2) Select the Fourier coefficient with the maximum energy and update data by subtracting the contribution of selected coefficient, equation (5), from the input data, solving equation (6).

$$f^k(\boldsymbol{x}_l) = c(\boldsymbol{k}) \hat{f}_{k_p}(\boldsymbol{k}) e^{2\pi i \boldsymbol{k} \boldsymbol{x}_l} \qquad (5)$$

$$f^{s+1}(\boldsymbol{x}_l) = \min_{\boldsymbol{x}_l} \left\| f^s(\boldsymbol{x}_l) - f^k(\boldsymbol{x}_l) \right\|_2^2 \qquad (6)$$

3) Repeat steps 2 and 3 until the updated residual input data to be less than a tolerance.
4) Finally, use all the coefficients selected in the optimal Fourier spectrum to reconstruct the signal at any desired sampling points.

These weights $c(\boldsymbol{k})$ are the least squares solution between a synthetized harmonic and the residual

$$c(\boldsymbol{k}) = \frac{f^s(\boldsymbol{x}_l) \cdot f^k(\boldsymbol{x}_l)}{f^k(\boldsymbol{x}_l) \cdot f^k(\boldsymbol{x}_l)} \qquad (7)$$

As you can see in the algorithm above $c(\boldsymbol{k})$ is calculated on the flight.

Numerical examples

Figure 1 shows the MP and ALFT reconstructions of a randomly sampled 1D signal with two-harmonics: the function $f(x) = 3\cos\left(\frac{\pi x}{16}\right) + 3\sin\left(\frac{\pi x}{8} + 1\right)$. This function generates a signal for 128 samples as seen in Figure 1(a). If only thirty (30) samples are available (as shown in Figure 1(b)), these can be used to calculate the 98 missing positions of the desired signal. Figures 1(c) and (d) show the MP and ALFT reconstruction, respectively. Figures 1(e) and (f) show the reconstruction error by MP and ALFT in relation to the desired signal, respectively.

The result is practically the same, but the number of iterations is different. The MP method has less iterations and is faster than ALFT. In the tests made it possible to conclude that the less samples, the greater the difference between the methods. An important feature of the MP method observed in this example is that it remains stable, maintaining the number of iterations even when the input data is very sparce.

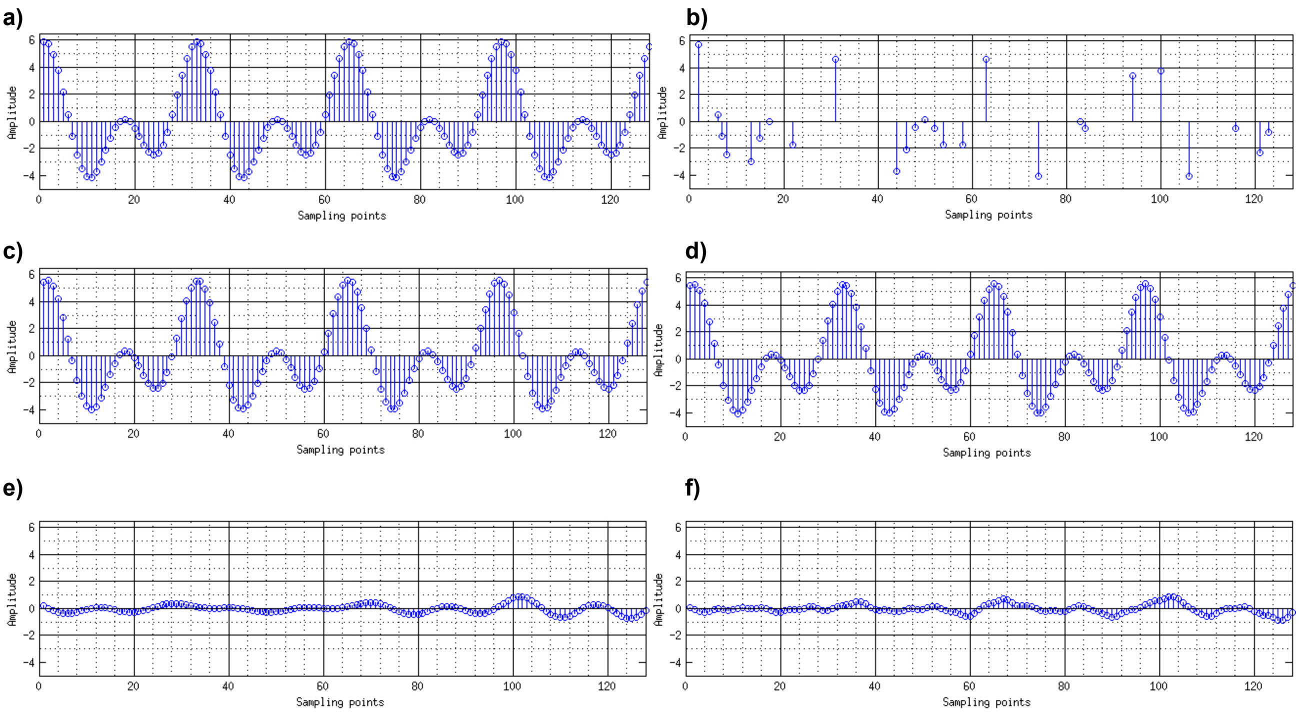

***Figure 1** – 1D regularization example: (a) Desired data. (b) Decimated data. (c) MP data reconstruction. (d) ALFT data reconstruction. (e) MP absolute error. (f) ALFT absolute error.*

Another example for a better understanding of the methods in high dimensions is through a two-dimensional function.

$$f(x,y) = \frac{1}{2}\begin{bmatrix}\cos 2\pi\,[(3x+2y)+(2x+3y)] \\ +\sin 4\pi(x+y)\end{bmatrix} \qquad (8)$$

Equation (8) generates a function distributed in a spacial grid with (128x64) positions in x and y, respectively (Figure 2(a)). As in case 1D, the input data were randomly decimated, leaving a data matrix of (54x36) positions in x and y, Figure 2(b). In other words, it will be necessary to apply the reconstruction in both directions (x and y) to obtain the (128x64) positions in the desired data. In Figures 2(c) and (d) are presented the MP and ALFT data reconstructions, respectively. In Figures 2(e) and (f) are presented the MP and ALFT difference to the desired data or absolute error, respectively.

It is very difficult to perceive difference in the results, Figures 2(c) and (d), even comparing the errors, Figures 2(e) and (f). Again, the number of iterations are different and in most runs, the MP method has fewer iterations and is faster than ALFT.

Application in synthetic seismic data

In order to demonstrate the ability of the MP and ALFT to reconstruct in two spatial dimensions, we applied both

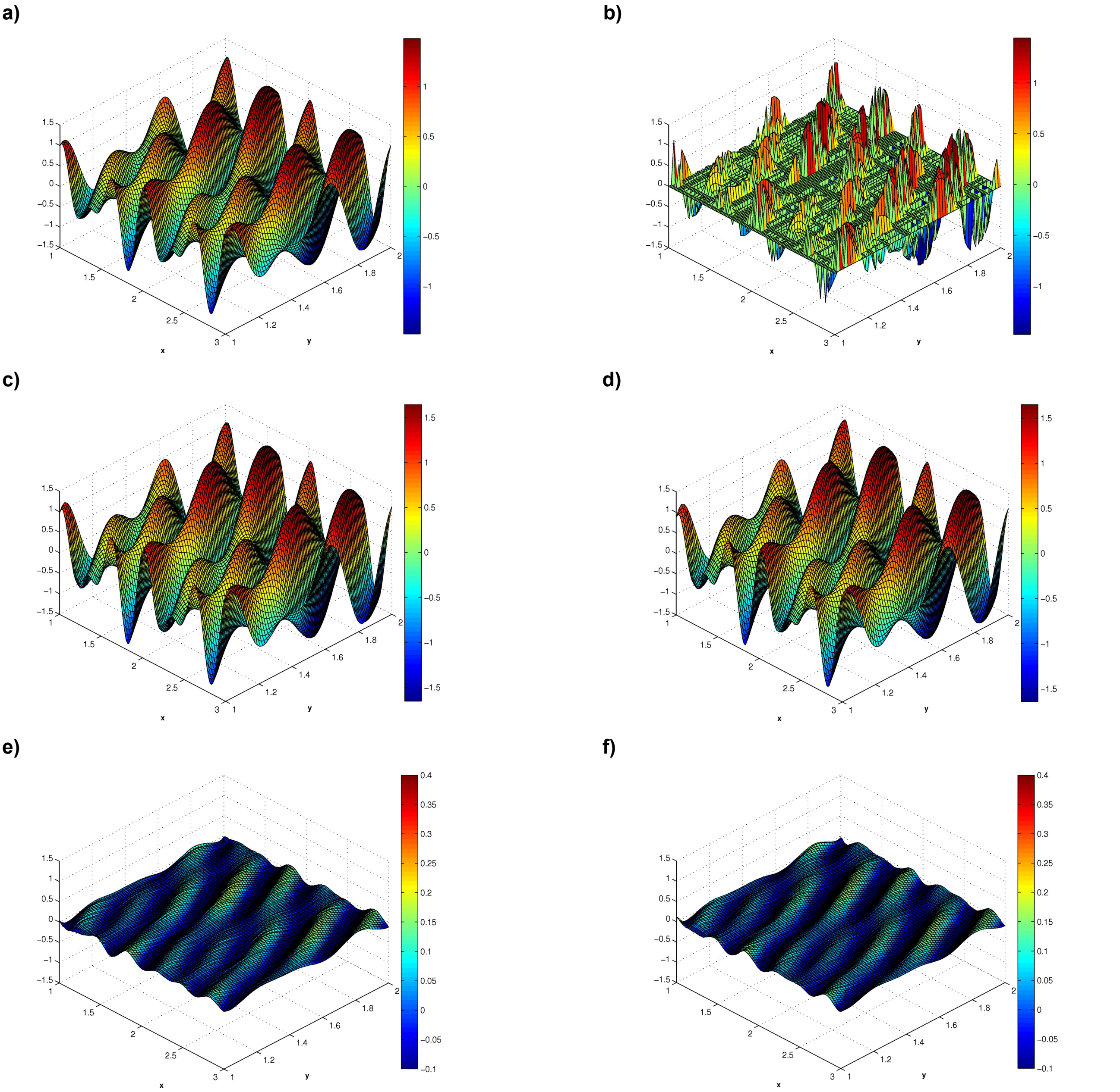

Figure 2 *– 2D regularization example: (a) Desired data. (b) Decimated data. (c) MP data reconstruction. (d) ALFT data reconstruction. (e) MP absolute error. (f) ALFT absolute error*.

methods to a subset of the Marmousi synthetic data. The regularization was done using spatial coordinates shot-receiver, but could be applied also midpoint-offset. The input data subset has 10 shots, interval 100 to 109 in full data, with 96 traces per shot, so it has a total of 960 traces (Figure 3(a) present shots 103 to 105). The offsets in each shot vary from 200 to 2575 with an interval of 25m. This data subset was decimated randomly, removing some shots on the data and some traces in each shot, resulting in a data with 8 shots and a total of 580 traces, as shown Figure 3(b) the same shots in the Figure 3(a).

In the Figures 3(c) and (d) are presented the results of the MP and ALFT reconstructions, for the same shots as in the Figure 3(a), respectively. For a better reconstruction in the f-k_1-k_2 domain it was applied an NMO correction in the input data and then the inverse NMO in output data. This is necessary to improve the ability to reconstruct hyperbolic events.

Both algorithms are still being optimized; it is not possible to make a fair comparison between computational times. But, as has been said in previous sessions, the MP method has its weights calculated on the flight and offers a better estimate of the Fourier coefficients, which entails in fewer iterations and a better performance than the ALFT.

Conclusions

As shown in this work and in the referred papers, the Matching Pursuit (MP) and Anti-Leakage Fourier Transform (ALFT) methods are efficient for seismic data reconstruction, and handles very steep dip events and non-uniform datasets. However, they have a high computational cost, especially the ALFT algorithm.

Were compared the MP and ALFT using simple functions in 1D and 2D. As expected, both methods present fairly similar results. The application in synthetic seismic data also shows that both methods reconstructed very well the decimated data, and no significant differences were observed in the results.

The results presented here are preliminary; we will optimize the MP and ALFT algorithms aiming to reduce the computational time. We will evaluate the performance of both methods by applying to real 2D land dataset of the Parnaiba basin, Brazil.

Acknowledgments

This work was supported by the "Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, Brasil," and by "Parnaíba Gás Natural", through the Investment Clause in Research, Development and Innovation, included in the contracts for Exploration, Development and Production of Oil and Natural Gas.

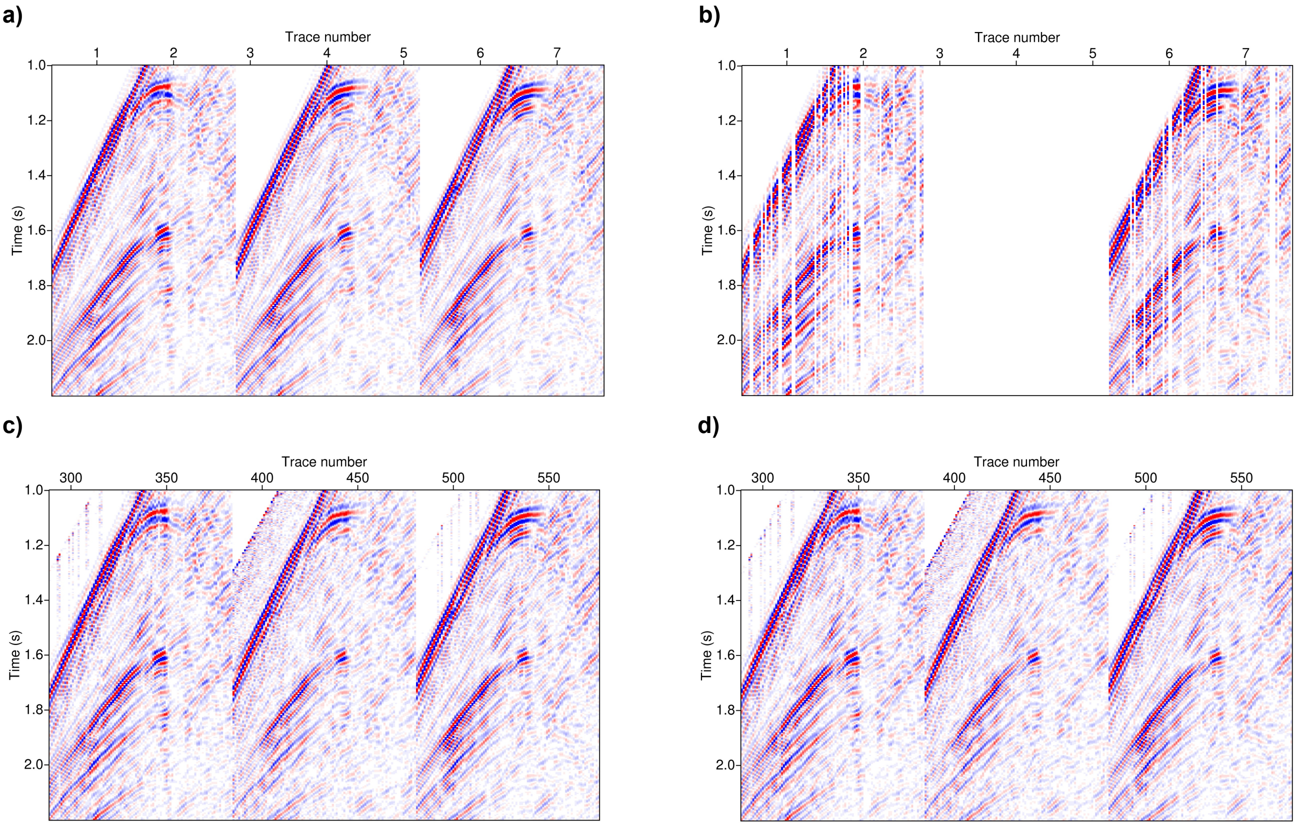

Figure 3 *– 2D regularization result: (a) Desired data. (b) Decimated data. (c) MP data reconstruction. (d) ALFT data reconstruction.*

References

HOLLANDER, Y.; KOSLOFF, D.; KOREN, Z.; BARTANA, A. 2102. Seismic Data Interpolation by Orthogonal Matching Pursuit. 74th EAGE Conference and Exhibition. Expanded Abstracts.

JAHANJOOY, S.; NIKROUZ, R.; MOHAMMED N. 2016. A faster method to reconstruct seismic data using anti-leakage Fourier transform. Journal of Geophysics and Engineering 13, 86-95.

LIU, B. & SACCHI, M.D. 2004. Minimum-weighted norm interpolation of seismic records. Geophysics 69, 1560-1568.

MALLAT, S. G. & ZHANG, Z. 1993. Matching pursuits with time-frequency dictionaries. IEEE Transaction in Signal Processing 41, 3397–3415.

NAGHIZADEH, M. & SACCHI, M.D. 2010. Seismic data reconstruction using multidimensional prediction filters. Geophys. Prospect. 58, 157-173.

NAGHIZADEH, M. & SACCHI, M.D. 2013. Multidimensional de-aliased cadzow reconstruction of seismic records. Geophysics 78, A1-A5.

NGUYEN, T. & WINNETT. R. 2011. Seismic interpolation by optimally matched Fourier components. 81th Annual International Meeting, SEG, Expanded Abstracts, 3085-3089.

POOLE, G. 2007. Multi-dimensional data regularization for modern acquisition geometries. 79th Annual International Meeting, SEG, Expanded Abstracts.2585-2589.

POOLE, G. 2010. 5D Data reconstruction using the anti-leakage Fourier transform. 72th EAGE Conf. & Exhibition, Expanded Abstracts.

SCHONEWILLE, M.; YAN, Z.; BAYLY, M.; BISLEY, R. 2013. Matching pursuit Fourier interpolation using priors derived from a second data set. 82th Annual International Meeting, SEG, Expanded Abstracts, 3651-3655.

XU, S.; ZHANG, Y.; PHAM, D.; LAMBARÉ, G. 2005. Antileakage Fourier transform for seismic data regularization. Geophysics 70, V78-V95.

XU, S.; ZHANG, Y.; PHAM, D.; LAMBARÉ, G. 2010. Antileakage Fourier transform for seismic data regularization in higher dimension. Geophysics 75, WB113-WB120.