



Noise attenuation by 3D greedy Radon transform

Juefu Wang,

Primary GeoServices Ltd., wangjuefu@yahoo.com

Summary

I present a flexible noise attenuation using a frequency domain 3D greedy Radon transform that simulates local seismic data with limited number of dips in two directions. For 2D case, the two directions correspond to CDP and offset, and for 3D case, they correspond to inline and crossline of cross-spread gathers. The solution is computed by a greedy approach. Particularly events with larger amplitude are solved earlier than those with smaller amplitude. Both synthetic and real data examples show that the algorithm can effectively suppress random noise and preserve signal. It can also reduce high-dip linear noise because only a limited number of moveout parameters are used in the inversion. Due to the flexibility of Radon transform in handling survey geometry, the algorithm can process data with irregular spatial sampling, and it does not lose seismic traces due to binning as required by many conventional methods.

Theory

Seismic data contains both random noise and coherent noise. Usually they are treated in different ways in seismic data processing. In this paper I mainly focus on removing random noise, although the proposed method can also reduce aliased linear noise as I show later in this paper. Because random noise is not predictable, it cannot be modelled by linear systems. We can reduce random noise by simulating seismic data with linear equations. Naturally the residual is the random noise. We can have different random noise attenuation methods based on different linear systems. For example, the classic prediction filtering method (Canales, 1984) assumes that a seismic trace can be predicted by summing nearby traces with proper weights in the FX (frequency and space) domain. The task is to optimize the weights with redundant linear equations, then predict clean data with the weighted sum of nearby traces using the optimized weights. Another example is the polynomial approach (Wang, 2004). The method tries to fit data with polynomials based on observed data and its spatial coordinates by solving for coefficients of polynomials. The clean data is then modelled with the optimal coefficients and the coordinates. A recent popular method called Cadzow filtering (Trickett, 2002; Trickett, 2015) assumes that coherent signal is contained in the first few strong eigenvectors of a Henkel matrix. Note that both the prediction filtering method and Cadzow filtering method need to have data on a regular grid. In reality, this requirement is almost always compromised due to limited financial budget and complex surface condition. A practical solution is to bin the data on a predefined grid. The binning process introduces errors in the data prediction, and if there are multiple traces located in a bin, only one trace will be processed. The polynomial method can handle irregular geometry, but it struggles to handle structured data.

To overcome the shortcomings of conventional methods, I use a 3D Radon transform to setup a linear system to predict seismic data. Radon operators can honor true locations of seismic traces, and they can handle spatially aliased data with sparseness constraints (Thorson and Claerbout, 1985; Sacchi and Ulrych, 1995; Herrmann et al., 2000; Trad et al, 2003; Hugonnet

and Boelle, 2007). In the past, Radon transform has mainly been used as a tool to remove coherent noise (e.g., linear noise and multiples). The idea is to focus seismic events in the Radon domain, and then the model of noise is extracted and transferred back to the space and time domain. Finally the predicted coherent noise is subtracted from the input. Here for random noise attenuation, I simply subtract the data misfit of inversion from the input.

The 3D Radon transform is much more expensive than the conventional 2D Radon transform because the model space grows by one dimension. To make the algorithm practical, we have to limit the model space by amplitude thresholding. One example is to conduct a full-rank inversion over filtered data and then solved most dominant events for full-bandwidth inversion (Wang and Nimsaila, 2014). Here I show that we can accomplish the goal by prioritizing the adjoint solution for each frequency.

In the space and temporal frequency (FX) domain, noise-free seismic data can be modelled using the following equation:

$$d(x, y, \omega) = \sum_{p_x, p_y} m(p_x, p_y, \omega) e^{i\omega(p_x x + p_y y)}, \quad (1)$$

where x and y are two spatial coordinates, p_x and p_y are slopes (normalized moveout parameters) in two spatial directions, $d(x, y, \omega)$ is seismic data in the FX domain, $m(p_x, p_y, \omega)$ is the Radon model in the p and temporal frequency domain, The adjoint of Equation (1) can be expressed as

$$m(p_x, p_y, \omega) = \sum_{x, y} d(x, y, \omega) e^{-i\omega(p_x x + p_y y)}. \quad (2)$$

Equation (2) describes a slant-stacking process in the FX domain to transform seismic data to a Radon model.

Considering noise we can modify Equation (1) in a matrix form as below

$$d = Lm + n, \quad (3)$$

where d is the seismic data as a vector, L is the forward operator as a matrix, m is the Radon model as a vector, and n is the random noise as a vector. To attenuate the noise is equivalent to minimize the following cost function:

$$J(m) = \|d - Lm\|^2. \quad (4)$$

To avoid overfitting the noise, I regularize the model using the following cost function instead

$$J(m) = \|d - Lm\|^2 + \mu R(m), \quad (5)$$

where R is a regularizing function based on a *priori* information of the model, and μ is a trade-off parameter to determine the amount of regularization. A lot of work has been devoted to imposing sparseness to the Radon model. For example, $L1$ norm or Cauchy norm can be used to enhance model resolution (Sacchi & Ulrych, 1995). Alternatively, a greedy approach (Ng and Perz, 2004) can be adopted to obtain a model with even higher resolution. There is no explicit regularizing function for the greedy method. The sparse feature is achieved by a prioritized inversion based on the amplitude of the Radon model. Below is a greedy implementation for the frequency domain 3D Radon transform:

1. Convert data from $d(x, y, t)$ to $d(x, y, \omega)$ (i.e. from TXY to FXY domain) by 1D Fourier transform along the time direction.
2. Loop over frequencies to find an optimal Radon model for each frequency using the following algorithm in pseudo code.

```

 $d_{resi} = d, \quad m = 0;$ 
for iter = 1, nitr
     $m_e = L' d_{resi};$ 
    //Note:  $L'$  is the adjoint operator described by Equation (2).

    sort  $m_e$  in amplitude descending order and keep the sorting indices of first  $N$ 
    samples as  $I_s^i, i = 1, N;$ 
    for  $i = 1, N$ 
         $m_e^{I_s^i} = L^{i'} d_{resi};$ 
        //Note:  $m_e^{I_s^i}$  is the adjoint model for one single pair of  $p_x$  and  $p_y$ , and  $L^{i'}$  is simply
        //to apply the adjoint operator for a single pair of  $p_x$  and  $p_y$ .

         $d_e = L^i m_e^{I_s^i};$ 
        //Note:  $L^i$  is to apply the forward operator describe by Equation (1) for a single pair of
        //  $p_x$  and  $p_y$ ,

         $\alpha = \frac{\langle m_e^{I_s^i}, m_e^{I_s^i} \rangle}{\langle d_e, d_e \rangle};$ 
        //Note:  $\langle \rangle$  denotes inner product of two vectors. In this step, a steepest descent
        //algorithm is used to find the optimal scalar.

         $m = m + \alpha m_e^{I_s^i};$ 
         $d_{resi} = d_{resi} - \alpha d_e;$ 
    end;
end;
 $d_{clean} = d - d_{resi};$ 

```

3. Convert clean data from $d_{clean}(x, y, \omega)$ to $d_{clean}(x, y, t)$ by 1D inverse Fourier transform.

The algorithm is similar to the anti-leakage Fourier transform (Xu et al, 2005) except that it uses a different operator and solver. First of all, Radon transform instead of Fourier transform is used so that it is more straightforward to control the dip range of seismic data. Second of all, in this algorithm there is no need to recalculate the full adjoint model after optimizing a single pair of moveout parameter (p_x and p_y). A series of similar moveout parameters are solved after each amplitude prioritization, which can improve convergence and significantly reduce the computational cost without sacrificing the resolution. Typically 6-8 iterations of the outer loops are enough to obtain a satisfactory result. The number of dips (N as described in the pseudo code) to solve in the inner loop is usually set to 20-40, a small fraction of the total number of the Radon model elements. Therefore the main cost of the algorithm is the step of computing the full adjoint model in the outer loop, which is used to find dips with maximum amplitude.

It is straightforward to see that the algorithm can be utilized to attenuate random noise for 3D data in the cross-spread domain. The spatial coordinates are corresponding to those of inline and crossline CDP distance within a cross-spread gather. For 2D data, we can apply the method to CDP domain data by modifying the operators as below

$$d(x, h, \omega) = \sum_{p_x, p_h} m(p_x, p_h, \omega) e^{i\omega(p_x x + p_h h^2)}, \quad (6)$$

where x is the CDP distance, and h is the offset. Note that a parabolic transform is used in the offset direction to accommodate data after NMO correction.

Accordingly, the adjoint of Equation (6) is

$$m(p_x, p_h, \omega) = \sum_{x, h} d(x, h, \omega) e^{-i\omega(p_x x + p_h h^2)}. \quad (7)$$

In practice, we often limit the ranges of moveout parameters to be big enough to model the signal but not linear noise with large dips. Consequently the linear noise is not modelled, and it remains in the residual. Therefore the algorithm can naturally remove linear noise with large dips (often aliased).

Synthetic example

To validate the new algorithm, I prepared a set of 2D CDP gathers with parabolic moveout in the offset direction and linear moveout in the CDP direction. To demonstrate the advantage of Radon transform in handling irregular geometry, the offsets of CDP gathers were set to irregular values as shown in Figure 1.

Figure 2 compares the results of FXY prediction filtering (Canales, 1984) and the proposed method. It can be seen that the new method provides a slightly cleaner result than the classic FXY prediction filtering method. More importantly, it can preserve the signal better when the offset binning error cannot be ignored. The FXY method tends to smooth out events when there is jittering related to irregular offsets. On the other hand, the new method is not affected by this

problem because true offsets are honored in the forward and adjoint operators during inversion. In addition, there are fewer artifacts in the greedy Radon result than those in the FXY result. FXY filtering seems to organize random noise and generate unwanted linear events.

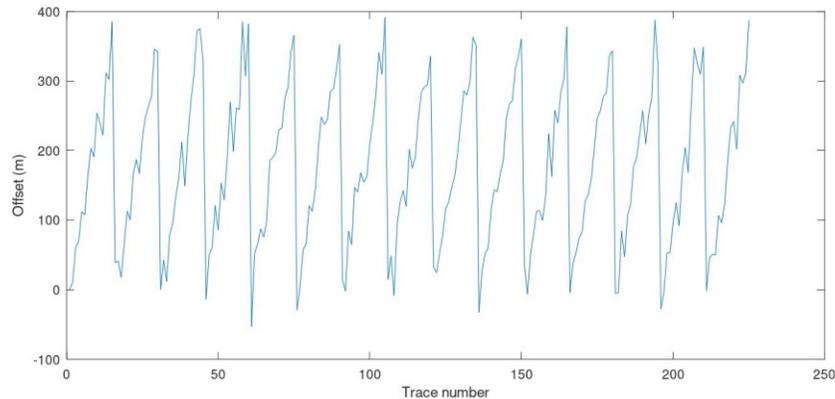


Figure 1: Offset distribution of the synthetic data used to compare two different denoise algorithms. The data contains 15 CDPs, and each CDP contains 15 irregular offsets.

Field data examples

I also tested the algorithm with a real 2D dataset and a 3D dataset. Figure 3 compares the denoising results of two-pass FX prediction filtering and the greedy Radon transform. For simplicity, I ran FX prediction filtering in the common shot domain and then sorted data to the common receiver domain. The second pass FX prediction filtering was applied in the common receiver domain. Finally the data was sorted back to the shot domain for quality control. The 3D greedy Radon was applied in the CDP domain directly. The result was then sorted to the shot domain for comparison. It is clear that the greedy Radon transform more effectively suppresses the random noise at the near offset. Erratic noise is also suppressed. Careful comparison of the data residuals (Figure 3d and Figure 3e) shows that the new method can preserve the signal better than the FX prediction method. Note that the minor signal leakage can be easily recovered by matched filtering after we obtain a clean dataset, but the topic is beyond the scope of this abstract. Of course the result of FX prediction filtering can also be improved by the matched filtering but with less success.

Interestingly, the linear noise with large dips is also suppressed by the greedy Radon transform. As explained in the theory, this is not a surprise because the dips of the linear noise are beyond the predefined moveout range for the inversion. By contrast, the linear noise can be predicted by the FX prediction method. Therefore and it is not removed. Note that most of the linear noise has been removed in the pre-processing stage using shot domain algorithms. Here the new method only deals with residual linear noise that is difficult to remove with conventional methods. Theoretically we can also remove linear noise by modeling and subtraction, but it will be more expensive because we need to increase the range of moveout parameters.

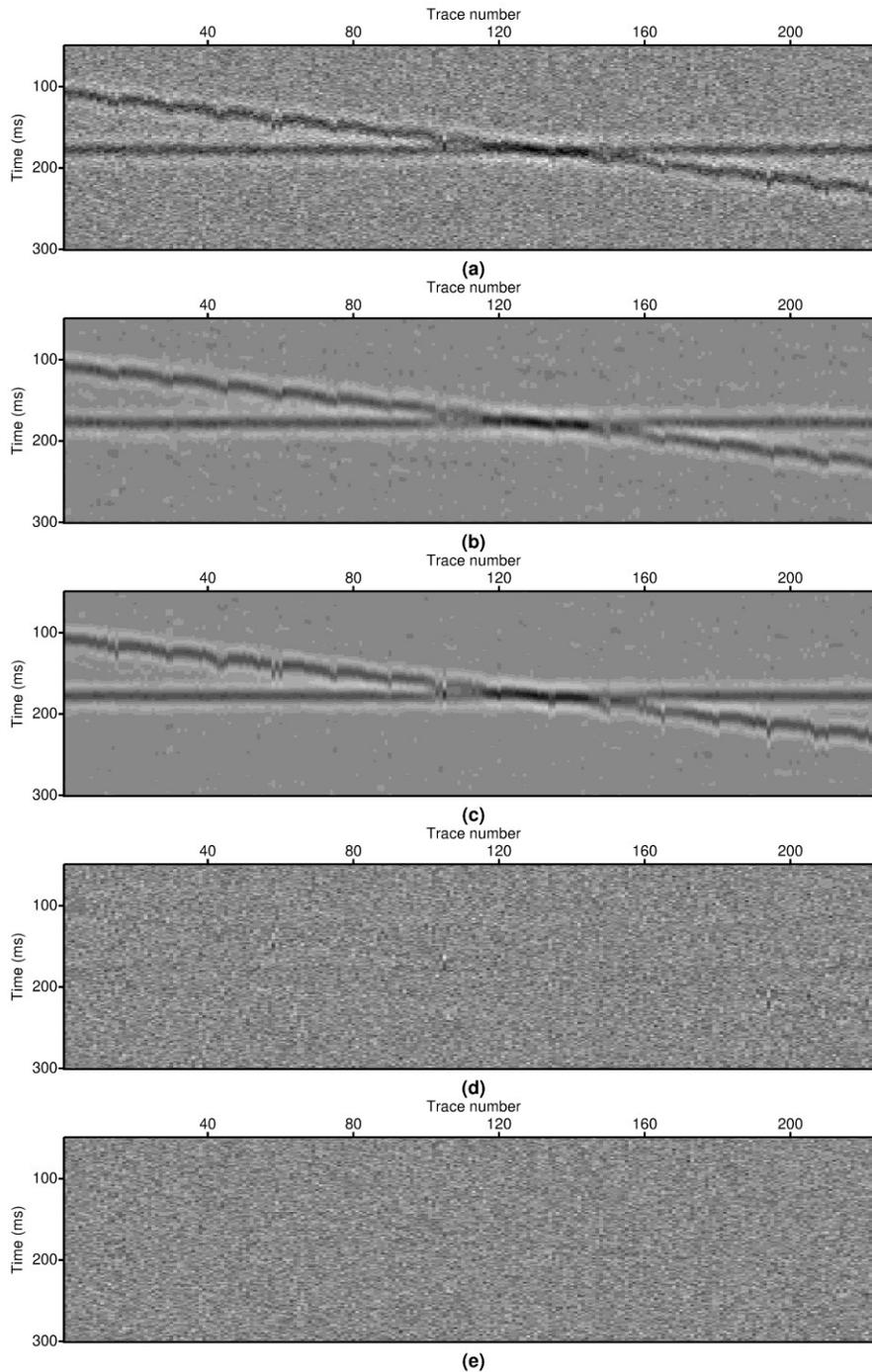


Figure 2: A comparison of denoising with the FXY prediction filtering and the greedy Radon transform. (a) Input synthetic CDP gathers. (b) Result of the FXY prediction filtering. (c) Result of the greedy Radon transform. (d) Residuals of the FXY prediction filtering. (e) Residuals of the greedy Radon transform.

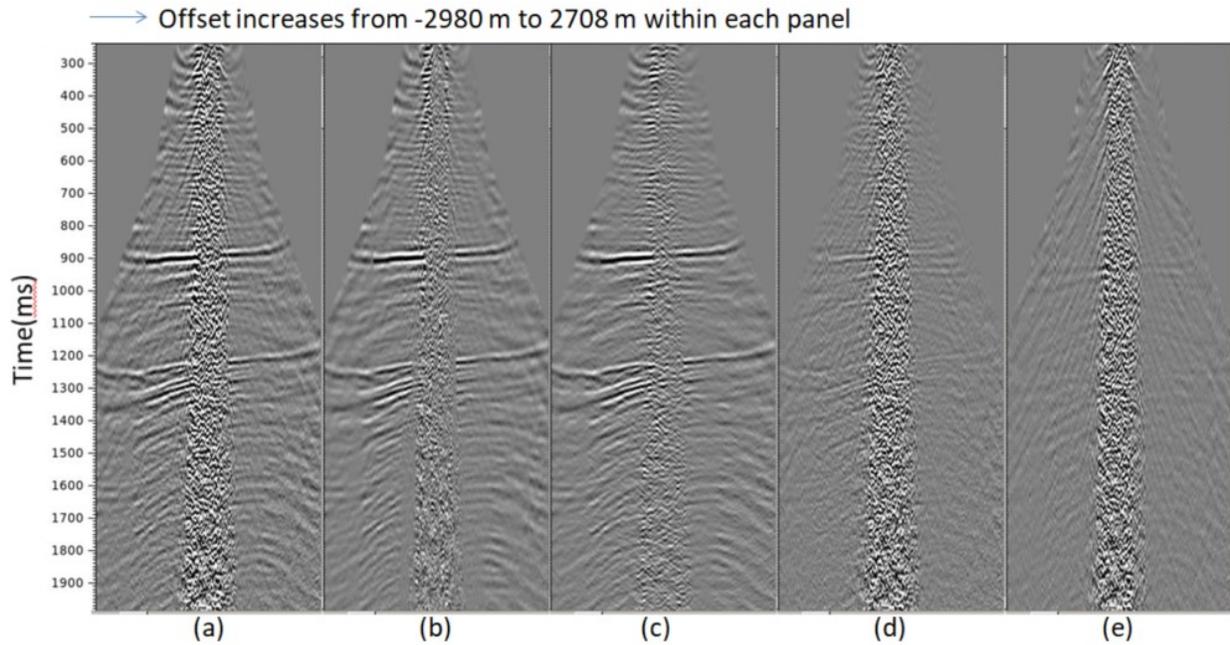


Figure 3: A real 2D example to compare denoising with the two-pass FX prediction filtering and the greedy Radon transform. (a) Input shot gather. (b) Result of the two-pass FX prediction filtering. (c) Result of the greedy Radon transform. (d) Residuals of the two-pass FX prediction filtering. (e) Residuals of the greedy Radon transform.

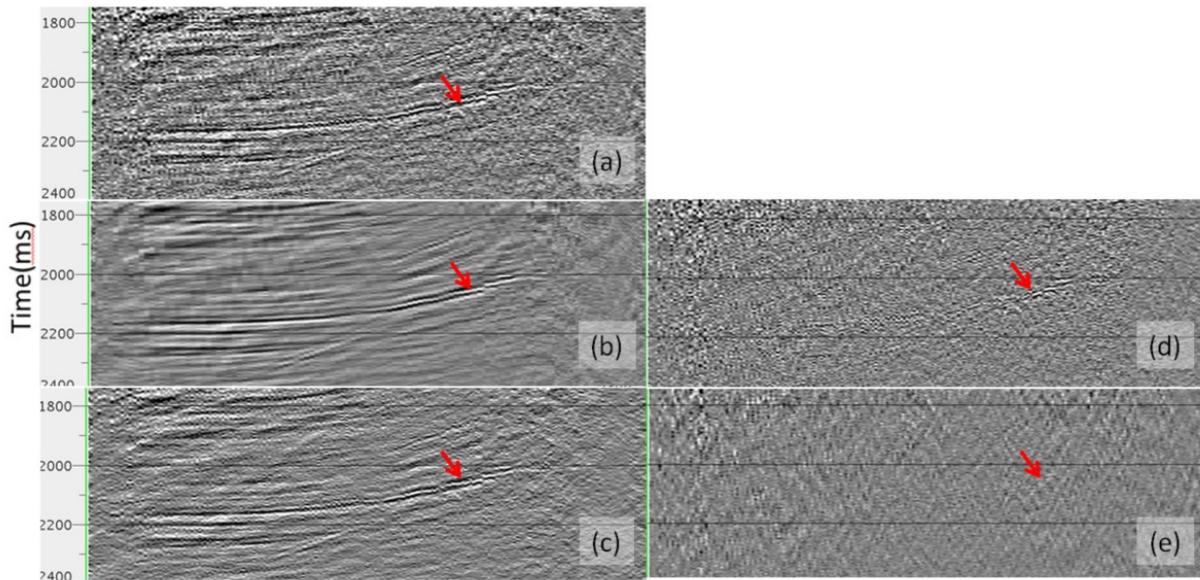


Figure 4: A real 3D example to compare denoising with FXY filtering and the greedy Radon transform. (a) Input cross-spread gather (one receiver line). (b) Result of FXY filtering. (c) Result of the greedy Radon transform. (d) Difference of FXY filtering. (e) Difference of the greedy Radon transform.

Figure 4 compares the denoising results of a 3D dataset using FXY prediction filtering and the greedy Radon transform. FXY filtering is very efficient in smoothing events in this case, but it tends to lose short wavelength details of seismic events (see the difference plot in Figure 4d). The red arrows in Figure 4 highlight the event with subtle structure variations that can be used to evaluate the performance of different noise attenuation algorithms. Figure 4c shows that the 3D greedy Radon transform can preserve geological details, and the data residuals (see Figure 4e) confirm this observation. One advantage of Radon transform is that it has more freedom to control data fitting by adjusting range of moveout parameters and number of dips to solve in each greedy search as described in the pseudo code.

Conclusions and discussions

I have implemented an efficient noise attenuation algorithm using a frequency domain 3D greedy Radon transform. The method can effectively reduce random noise and residual linear noise and preserve signal even when data is highly structured. The algorithm can be used for both 2D and 3D data in different domains.

The idea of optimizing multiple dips in the inner loop after sorting all events based on amplitude enables us to reduce the computational cost of greedy algorithms for higher dimension data. A straightforward application is to develop 4D/5D interpolation algorithms to improve spatial sampling of seismic data.

Acknowledgements

I would like to thank Primary GeoServices and Absolute Imaging for allowing me to publish this paper. I would also like to acknowledge the processing team at Absolute Imaging for their effort in processing the real data and reviewing the denoising results. Special thanks to Professor Mauricio Sacchi for the discussion about noise attenuation.

References

- Canales, L. L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstracts, 525-527.
- Herrmann, P., T. Mojeskey, M. Magesan and P. Hugonnet, 2000, De-aliased, high-resolution Radon transforms: 70th Annual International Meeting, SEG, Expanded Abstracts, 1953-1956.
- Hugonnet, P. and J. L. Boelle, 2007, Beyond aliasing regularization by plane wave extraction: 69th Conference & Exhibition, EAGE, Extended Abstracts, 114-117.
- Ng, M. and M. Perz, 2004, High resolution Radon transform in the t-x domain using “intelligent” prioritization of the Gauss-Seidel estimation sequence: 84th Annual International Meeting, SEG, Expanded Abstracts, 2160-2163.
- Sacchi, M. D. and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: *Geophysics*, **60**, 1169-1177.
- Thorson, J. R. and J. F. Claerbout, 1985, Velocity stack and slant stack stochastic inversion: *Geophysics*, **50**, 2727-2741.

Trad, D., T. Ulrych and M. D. Sacchi, 2003, Latest views of the sparse Radon transform: *Geophysics*, **68**, 386-399.

Trickett, S., 2002, F-xy eigenimage noise suppression: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2166-2169.

Trickett, S. 2015, Preserving signal: Automatic rank determination for noise suppression: 85th Annual International Meeting, SEG, Expanded Abstracts, 4703-4707.

Wang, P. and N. Kawin, 2014, Fast progressive sparse Tau-P transform for regularization of spatially aliased seismic data: 84th Annual International Meeting, SEG, Expanded Abstracts, 3589-3593.

Wang, X., 2004, 3D prestack seismic trace interpolation in time domain with input optimization: CSEG National Convention.

Xu, S., Y. Zhang, D. Pham and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, v87-v95.