Seislet-based morphological component analysis using scale-dependent exponential shrinkage[a]

[a]Published in Journal of Applied Geophysics, doi: 10.1016/j.jappgeo.2015.04.003 (2015)

Pengliang Yang and Sergey Fomel†*

ABSTRACT

Morphological component analysis (MCA) is a powerful tool used in image processing to separate different geometrical components (cartoons and textures, curves and points etc). MCA is based on the observation that many complex signals may not be sparsely represented using only one dictionary/transform, however can have sparse representation by combining several over-complete dictionaries/transforms. In this paper we propose seislet-based MCA for seismic data processing. MCA algorithm is reformulated in the shaping-regularization framework. Successful seislet-based MCA depends on reliable slope estimation of seismic events, which is done by plane-wave destruction (PWD) filters. An exponential shrinkage operator unifies many existing thresholding operators and is adopted in scale-dependent shaping regularization to promote sparsity. Numerical examples demonstrate a superior performance of the proposed exponential shrinkage operator and the potential of seislet-based MCA in application to trace interpolation and multiple removal.

INTRODUCTION

A wide range of applications have been carried out by solving a series of linear inverse problems and using the fact that numerous varieties of signals can be sparsely represented with an appropriate dictionary, namely a certain kind of transform bases. Under the dictionary, the signals have fewer non-zeros of the representation coefficients. However, many complex signals are usually linear superposition of several elementary signals, and cannot be efficiently represented using only one dictionary. The concept of morphological diversity was therefore proposed by Starck et al. (2004, 2005) to combine several dictionaries for sparse representations of signals and images. Then, the signal is considered as a superposition of several morphological components. One has to choose a dictionary whose atoms match the shape of the geometrical structures to sparsify, while leading to a non-sparse (or at least not as sparse) representation of the other signal content. That is the essence of so-called morphological component analysis (MCA) (Starck et al., 2004, 2007; Woiselle et al., 2011).

Seislet transform and seislet frame are useful tools for seismic data compression and sparse representation (Fomel and Liu, 2010). Seislets are constructed by apply-

ing the wavelet lifting scheme (Sweldens, 1998) along the spatial direction, taking advantage of the prediction and update steps to characterize local structure of the seismic events. In the seislet transform, the locally dominant event slopes are found by plane-wave destruction (PWD), which is implemented using finite difference stencils to characterize seismic images by a superposition of local plane waves (Claerbout, 1992). By increasing the accuracy and dip bandwidth of PWD, Fomel (2002) demonstrated its competitive performance compared with prediction error filter (PEF) in the applications to fault detection, data interpolation, and noise attenuation. PWD keeps the number of adjustable parameters to a minimum, endows the estimated quantity with a clear physical meaning of the local plane-wave slope, and gets rid of the requirement of local windows in PEF. Recently, Chen et al. (2013a,b) accelerated the computation of PWD using an analytical estimator and improved its accuracy.

In this paper, we propose seislet-based MCA for seismic data processing. We reformulate MCA algorithm in the shaping-regularization framework (Fomel, 2007, 2008). Successful seislet-based MCA depends on reliable slope estimation of seismic events, which can be done by plane-wave destruction (PWD) filtering. Due to the special importance of an effective shrinkage or thresholding function in sparsity-promoting shaping optimization, we propose a scale-dependent exponential shrinkage operator, which can flexibly approximate many well-known existing thresholding functions. Synthetic and field data examples demonstrate the potential of seislet-based MCA in the application to trace interpolation and multiple removal.

MCA WITH SCALE-DEPENDENT SHAPING REGULARIZATION

Analysis-based iterative thresholding

A general inverse problem combined with a priori constraint $R(x)$ can be written as an optimization problem

$$\min_x \frac{1}{2}\|d_{obs} - Fx\|_2^2 + \lambda R(x), \tag{1}$$

where x is the model to be inverted, and d_{obs} is the observations. To solve the problem with sparsity constraint $R(x) = \|x\|_1$, the iterative shrinkage-thresholding (IST) algorithm has been proposed (Daubechies et al., 2004), which can be generally formulated as

$$x^{k+1} = T_\lambda(x^k + F^*(d_{obs} - Fx^k)), \tag{2}$$

where k denotes the iteration number; and F^* indicates the adjoint of F. $T_\lambda(x)$ is an element-wise shrinkage operator with threshold λ:

$$T_\lambda(x) = (t_\lambda(x_1), t_\lambda(x_2), \ldots, t_\lambda(x_m))^T, \tag{3}$$

in which the soft thresholding function (Donoho, 1995) is

$$t_\lambda(u) = \text{Soft}_\lambda(u) = \begin{cases} u - \lambda \frac{u}{|u|}, & |u| > \lambda, \\ 0, & |u| \leq \lambda. \end{cases} = u. * \max(1 - \frac{\lambda}{|u|}, 0). \tag{4}$$

Allowing for the missing elements in the data, the observations are connected to the complete data via the relation

$$d_{obs} = Md = M\Phi x = Fx, F = M\Phi. \tag{5}$$

where M is an acquisition mask indicating the observed and missing values. Assume Φ is a tight frame such that $\Phi^*\Phi = \text{Id}$, $x = \Phi^*\Phi x = \Phi^* d$. It leads to

$$\begin{aligned} d^{k+1} &= \Phi x^{k+1} \\ &= \Phi T_\lambda(\Phi^* d^k + (M\Phi)^*(d_{obs} - Md^k)) \\ &= \Phi T_\lambda(\Phi^* d^k + \Phi^*(M^* d_{obs} - M^* M d^k)) \\ &= \Phi T_\lambda(\Phi^*(d^k + (d_{obs} - Md^k))), \end{aligned} \tag{6}$$

in which we use $M^* = M = (M)^2$ and $Md_{obs} = M^2 d = Md$. Now we define a residual term as $r^k = d_{obs} - Md^k$, thus Eq. (6) results in

$$\begin{cases} r^k \leftarrow d_{obs} - Md^k \\ d^{k+1} \leftarrow \Phi T_\lambda(\Phi^*(d^k + r^k)), \end{cases} \tag{7}$$

which is equivalent to solving

$$\min_d \frac{1}{2}\|d_{obs} - Md\|_2^2 + \lambda R(\Phi^* d). \tag{8}$$

Note that Eq. (8) analyzes the target unknown d directly, without resort to x and $d = \Phi x$. Eq. (6) is referred to as the analysis formula (Elad et al., 2007). In this paper, we used the analysis formula because it directly addresses the problem in the data domain for the convenience of interpolation and signal separation.

Understanding iterative thresholding as shaping regularization

Note that at each iteration soft thresholding is the only nonlinear operation corresponding to the ℓ_1 constraint for the model x, i.e., $R(x) = \|x\|_1$. Shaping regularization (Fomel, 2007, 2008) provides a general and flexible framework for inversion without the need for a specific penalty function $R(x)$ when a particular kind of shaping operator is used. The iterative shaping process can be expressed as

$$x^{k+1} = S(x^k + B(d_{obs} - Fx^k)), \tag{9}$$

where the shaping operator S can be a smoothing operator (Fomel, 2007), or a more general operator even a nonlinear sparsity-promoting shrinkage/thresholding operator (Fomel, 2008). It can be thought of a type of Landweber iteration followed by

projection, which is conducted via the shaping operator S. Instead of finding the formula of gradient with a known regularization penalty, we have to focus on the design of shaping operator in shaping regularization. In gradient-based Landweber iteration the backward operator B is required to be the adjoint of the forward mapping F, i.e., $B = F^*$; in shaping regularization however, it is not necessarily required. Shaping regularization gives us more freedom to choose a form of B to approximate the inverse of F so that shaping regularization enjoys faster convergence rate in practice. In the language of shaping regularization, the updating rule in Eq. (7) becomes

$$\left\{ \begin{array}{l} r^k \leftarrow d_{obs} - Md^k, \\ d^{k+1} \leftarrow \Phi S(\Phi^*(d^k + r^k)), \end{array} \right. \tag{10}$$

where the backward operator is chosen to be the inverse of the forward mapping.

MCA using sparsity-promoting shaping

MCA considers the complete data d to be the superposition of several morphologically distinct components: $d = \sum_{i=1}^N d_i$. For each component d_i, MCA assumes there exists a transform Φ_i which can sparsely represent component d_i by its coefficients α_i ($\alpha_i = \Phi_i^* d_i$ should be sparse), and can not do so for the others. Mathematically,

$$\min_{\{d_i\}} \sum_{i=1}^N R(\Phi_i^* d_i), \text{subject to } d_{obs} = M \sum_{i=1}^N d_i. \tag{11}$$

The above problem can be rewritten as

$$\min_{\{d_i\}} \frac{1}{2} \left\| d_{obs} - M \sum_{i=1}^N d_i \right\|_2^2 + \lambda \sum_{i=1}^N R(\Phi_i^* d_i). \tag{12}$$

We prefer to rewrite Eq. (12) as

$$\begin{aligned} \min_{\{d_i\}} \frac{1}{2} \left\| \left(d_{obs} - M \sum_{j \neq i} d_j \right) - M d_i \right\|_2^2 \\ + \lambda R(\Phi_i^* d_i) + \lambda \sum_{j \neq i} R(\Phi_j^* d_j). \end{aligned} \tag{13}$$

Thus, optimizing with respect to d_i leads to the analysis IST shaping as Eq. (9). At the kth iteration, optimization is performed alternatively for many components using the block coordinate relaxation (BCR) technique (Bruce et al., 1998): for the ith component d_i^k, $i = 1, \ldots, N$: $\Phi \leftarrow \Phi_i$, $d^k \leftarrow d_i^k$, $d^{k+1} \leftarrow d_i^{k+1}$, $d_{obs} \leftarrow d_{obs} - M \sum_{j \neq i} d_j^k$, yields the residual term $r^k = d_{obs} - M \sum_{i=1}^N d_i^k$ and the updating rule

$$\left\{ \begin{array}{l} r^k \leftarrow d_{obs} - M \sum_{i=1}^N d_i^k \\ d_i^{k+1} \leftarrow \Phi_i S(\Phi_i^*(d_i^k + r^k)). \end{array} \right. \tag{14}$$

The final output of the above algorithm are the morphological components $\hat{d}_i, i = 1, \ldots, N$. The complete data can then be reconstructed via $\hat{d} = \sum_{i=1}^N \hat{d}_i$. This is the main principle of the so-called MCA-based inpainting algorithm (Elad et al., 2005).

SEISLET-BASED MCA SPARSIFIED WITH SCALE-DEPENDENT EXPONENTIAL SHRINKAGE

Seislet transform and local slope estimation

Seislet transform and seislet frame were proposed by Fomel and Liu (2010) for seismic data compression and sparse representation. Seislets are constructed by applying the wavelet lifting scheme (Sweldens, 1998) along the local slope direction. For each level of lifting decomposition, seismic data is split into even and odd parts (**e** and **o**). Then the prediction and update step follows to obtain the detail difference/residual **d** and smooth information **s**:

$$\mathbf{d} = \mathbf{e} - P[\mathbf{o}], \mathbf{s} = \mathbf{e} + U[\mathbf{d}]. \tag{15}$$

Recognizing that seismic data can be organized as collections of traces or records, Fomel and Liu (2010) suggest prediction of one seismic trace or record from its neighbors and update of records on the next scale to follow structural features in seismic data. In the Z-transform notation, the simplest Haar prediction filter can be written as

$$P(Z) = Z, \tag{16}$$

and the linear interpolation prediction filter is

$$P(Z) = (Z + 1/Z)/2. \tag{17}$$

Successful prediction and update play a key role for local slope estimation. By modifying the biorthogonal wavelet construction, the prediction and update operators for a simple seislet transform are defined as

$$\begin{aligned} P[\mathbf{e}]_k &= (S_k^+[\mathbf{e}_{k-1}] + S_k^-[\mathbf{e}_k])/2, \\ U[\mathbf{r}]_k &= (S_k^+[\mathbf{r}_{k-1}] + S_k^-[\mathbf{r}_k])/4, \end{aligned} \tag{18}$$

where S_k^+ and S_k^- are the operators that predict a trace from its left and right neighbors, corresponding to shifting seismic events in terms of their local slopes. The job of local slope estimation can be done by PWD filters. Particularly, it is possible to obtain two or more dips with the help of PWD filters to capture different geometrical components of seismic data. The estimation of slopes involves a least-square optimization problem to be solved (Fomel, 2002), leading to extra computation. It is important to point out that besides PWD, there are other approaches to estimating dips of seismic data, i.e., local slant stack (Ottolini, 1983) and volumetric scan (Marfurt, 2006). However, PWD implements slope estimation through prediction and therefore is appropriate for use with the seislet transform.

Sparsifying MCA with exponential shrinkage shaping

The IST algorithm used by MCA requires soft thresholding function to filter out the unwanted small values. Besides soft thresholding (Donoho, 1995), many other

shrinkage functions can also be applied to obtain possibly better sparseness. One particular choice is hard thresholding:

$$t_\lambda(u) = \text{Hard}_\lambda(u) = u. * (|u| > \lambda ? 1 : 0). \tag{19}$$

where $(\cdot)?A : B$ frequency used hereafter is an if-else judgment in C-code style: The expression equals A if the statement $(\cdot)$ is true, and B otherwise.

Another choice is Stein thresholding (Peyre, 2010; Mallat, 2009):

$$t_\lambda(u) = \text{Stein}_\lambda(u) = u. * \max\left(1 - (\frac{\lambda}{|u|})^2, 0\right). \tag{20}$$

Stein thresholding does not suffer from the bias of soft thresholding, that is,

$$|\text{Stein}_\lambda(u) - u| \to 0, |\text{Soft}_\lambda(u) - u| \to \lambda, \text{if } u \to \infty. \tag{21}$$

Recent advances in nonconvex optimization (Chartrand, 2012; Voronin and Chartrand, 2013; Chartrand and Wohlberg, 2013) show that the shrinkage operator in IST algorithm (Eq. (2)) can be generalized to a p-quasinorm ($0 < p \leq 1$) thresholding operator T_λ, in which

$$t_\lambda(u) = \text{pThresh}_{\lambda,p}(u) = u. * \max\left(1 - (\frac{\lambda}{|u|})^{2-p}, 0\right). \tag{22}$$

A special case is that of $p = 1$, which corresponds to the soft thresholding operator exactly.

Most of these shrinkage functions interpolate between the hard and soft thresholders. It is tempting for us to design a more general shrinkage function to sparsify the transform domain coefficients in shaping regularized MCA. One possibility is multiplying an exponential factor on the elements of original data:

$$t_\lambda(u) = u. * \exp(-(\frac{\lambda}{|u|})^{2-p}). \tag{23}$$

Based on Taylor series, this operator in Eq. (23) enjoys some useful properties:

- It is valuable to point out that the exponential shrinkage can be considered as a smooth ℓ_0 constraint (Mohimani et al., 2009; Gholami and Hosseini, 2011). For $|u| >> \lambda$, it is a good approximation of the p-thresholding operator in Eq. (22), and does not suffer the bias when $p \neq 1$. It reduces to Stein thresholding operator for $p = 0$ and soft thresholding for $p = 1$.

$$\begin{aligned} t_{\lambda,0}(u) &= u. * \exp(-(\frac{\lambda}{|u|})^2) \approx u. * (1 - (\frac{\lambda}{|u|})^2), \\ t_{\lambda,1}(u) &= u. * \exp(-(\frac{\lambda}{|u|})) \approx u. * (1 - (\frac{\lambda}{|u|})). \end{aligned} \tag{24}$$

- It is free of non-differentiable singularity at the thresholding point λ. The transition between the small values and the large values is smoothly stretched. Due to the exponential factor less than 1 ($\exp(-(\frac{\lambda}{|u|})^{2-p}) < 1$), this operator will slightly decrease the data amplitude, even for $|u| < \lambda$.
- Besides the threshold λ, we have another independent parameter p which can be flexibly chosen to achieve better performance.

In the language of shaping regularization, shrinkage-based shaping operator S is equivalent to multiplying the coefficient vector x by a diagonal weighting matrix W to in the sense that

$$S(x) = Wx, \tag{25}$$

where

$$\mathrm{diag}(W_{ii}) = \begin{cases} 1 - \frac{\lambda}{|x_i|} > 0 ? 1 : 0, & \text{Hard} \\ \max(1 - \frac{\lambda}{|x_i|}, 0), & \text{Soft} \\ \max(1 - \left(\frac{\lambda}{|x_i|}\right)^2, 0), & \text{Stein} \\ \max(1 - \left(\frac{\lambda}{|x_i|}\right)^{2-p}, 0), & \text{pThresh} \\ \exp(-(\frac{\lambda}{|x_i|})^{2-p}), & \text{Exponential} \end{cases} \tag{26}$$

For the convenience of comparison, we plot these thresholding operators in Figure 1.

Note that we are using seislet transform which has different scales for signal representation. Usually, large scales of seislet coefficients corresponds to unpredictable noise, while most of the important information gets transformed into smaller scales. We design a scale-dependent diagonal weighting operator:

$$\mathrm{diag}(W_{ii}) = s(x_i) < s_0 ? 1 : 0, \tag{27}$$

where s_0 is user-defined scale, while $s(x_i)$ is the scale that the coefficient x_i correspond to. Putting all things together, in the MCA shaping regularization we are using a scale-dependent exponential shrinkage operator which is a composite operator cascaded with a scale-muting operator W_s and an exponential weighting operator W_{exp}:

$$S(x) = W_{exp} W_s x. \tag{28}$$

The use of scale-dependent exponential shrinkage offers easy control on the separation of the signal components we would like to capture. It is interesting to mention that under the Fourier basis, the scale-muting operator W_s becomes a frequency mask (it behaves like a selective hard thresholding), which can be employed to remove the groundroll in application to seismic interpolation (Gholami, 2014).

By incorporating PWD dip estimation and scale-dependent exponential shrinkage shaping, we summarize the proposed seislet-based MCA algorithm as Algorithm MCA. Seislet transforms associated with different dips form a combined seislet frame (Fomel and Liu, 2010). The threshold in each iteration can be determined with a

predefined percentile according to Hoare's algorithm. Shrinkage operator plays the role of crosstalk removal in MCA algorithm, as explained in more detail in Appendix A.

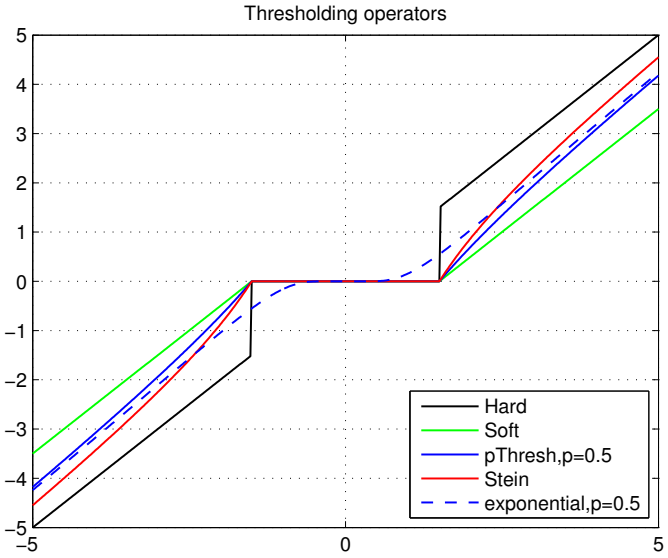

Figure 1: A schematic plot of the shrinkage operators,$\lambda = 1$

Algorithm 1 Seislet-based MCA algorithm

Input: Observed seismic image d_{obs}; sampling mask M; iteration number $niter$; shrinkage shaping parameter p, seislet transform Φ_i associated with the ith estimated slope.

Output: Separated seismic component d_i.

Initialize: $d_i^{(k)} \leftarrow 0$, $i = 1 \dots N$;
for $k = 1 \dots niter$ **do**
  $r^{(k)} \leftarrow d_{obs} - M \sum_{i=1}^{N} d_i^{(k)}$;
  for $i = 1 \dots N$ **do**
    $x_i^{(k)} \leftarrow \Phi_i^*(d_i^{(k)} + r_i^{(k)})$;
    Estimate shaping parameters λ and s_0;
    $d_i^{(k)} \leftarrow \Phi_i S(x_i^k)$;
  end for
end for

NUMERICAL EXAMPLES

Trace interpolation

Interpolation of random missing traces is an important task in seismic data processing. Unlike most studies using only one transform, we consider the seismic data having

two different components, which can be characterized by a seislet frame composed of seislet transforms associated with two different slopes. PWD filter is utilized to estimate the two dip fields (Fomel, 2002). As shown in Figure 2, the complete data is decimated with a random eliminating rate 25%. 10 iterations are carried out to separate these components. The estimated dips (Figure 3) by PWD indicate that the two separated components exhibit different modes: component 2 has positive and negative dips, corresponding to seismic diffractions, while the events of component 1 are more consistent (most values of the dip are positive). The summation of the two components gives a reasonable interpolation result (right panel of Figure 4.) For comparison, we define the signal-to-noise ratio as $\mathrm{SNR} = 10\log_{10}(\frac{|s|^2}{|s-\hat{s}|^2})$ to quantify the reconstruction performance. The resulting SNRs using exponential shrinkage, soft, hard, and generalized p-thresholding are 11.98 dB, 11.29 dB, 5.37 dB and 8.94 dB, respectively. It shows that the proposed exponential shrinkage outperforms the existing methods in MCA interpolation. It is important to point out that approximating the ℓ_0 and ℓ_1 minimization in a smooth constraint has already been validated in seismic interpolation applications by Cao et al. (2011). The use of exponential shrinkage enriches the sparsity-promoting shaping operator and extends the smooth constraint to MCA approach.

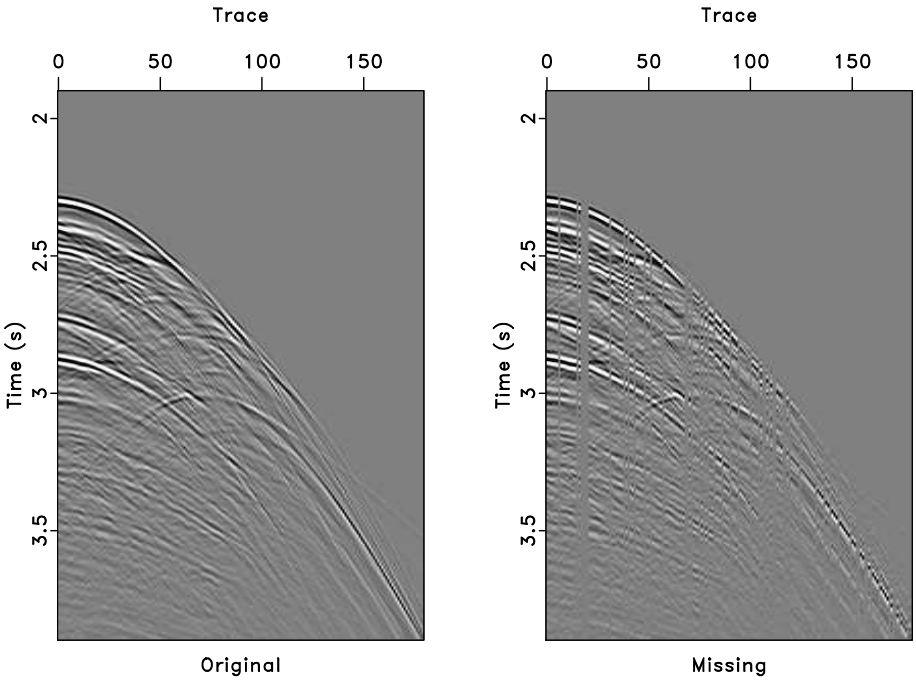

Figure 2: The observed seismic data (right) to be interpolated is obtained by 25% random eliminating the complete data (left).

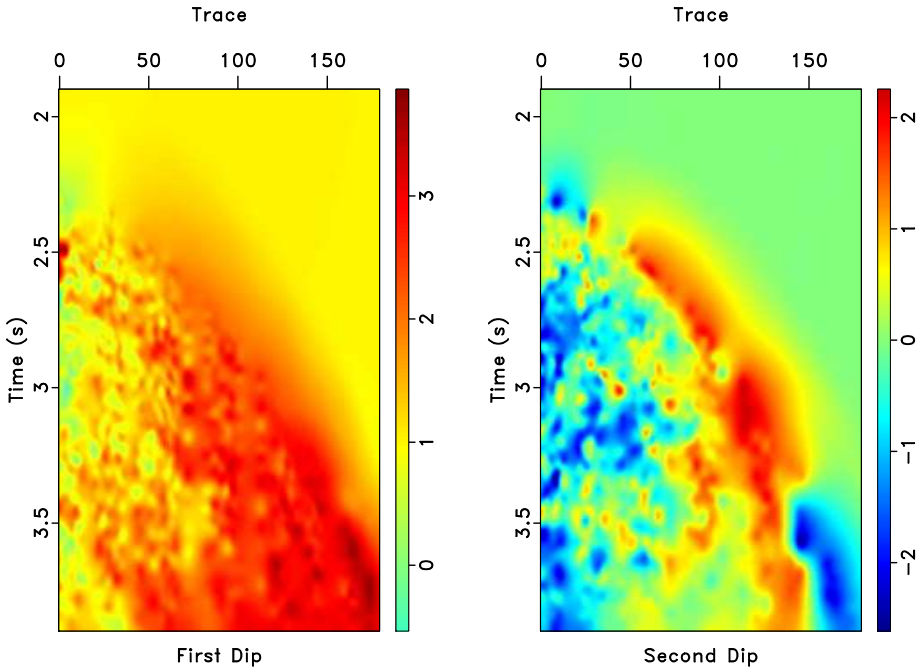

Figure 3: PWD estimated dips: dip estimation for component 1 (Left) are positive-valued, while dip for component 2 includes negative and positive values.

Multiple removal

Our second example is the separation of primaries and multiples for the field CMP gather shown in Figure 5 (Fomel and Guitton, 2006). In the case of signal separation, the mask operator M becomes an identity. The multiples are predicted using surface-related multiple elimination (SRME). Even though SRME fails to predict the correct amplitudes, however, the resulting prediction helps PWD to extract the dominant slopes of multiple events. Before applying the seislet-based MCA method, it is important to point out that the iteratively reweighted least-squares (IRLS) method using the model precondition is another way for sparsity-enforced separation (Daubechies et al., 2010). Thus, we compared multiple removal by two different methods: seislet-based IRLS method and the proposed seislet-based MCA method. For comparison, the separated primaries and multiples using different methods are plotted in Figures 6 and 7. Visually, the seislet-based IRLS method and the seislet-based MCA method output similar primaries, which can also be seen clearly from the velocity scans of the primaries shown in Figure 8. To further confirm our conclusion, we draw the velocity scans of predicted multiples in Figure 9. The panels of velocity scan in Figures 8 and 9 demonstrate that using MCA, the primaries correctly correspond to high velocity part while the multiples are associated with low velocity part. Figure 9 shows that seislet-based MCA outperforms the seislet-based IRLS method in the locations A and B in the velocity scan panel due to the nice match of the corresponding semblance scan of the original data; at locations C and D, the energy of multiples obtained by

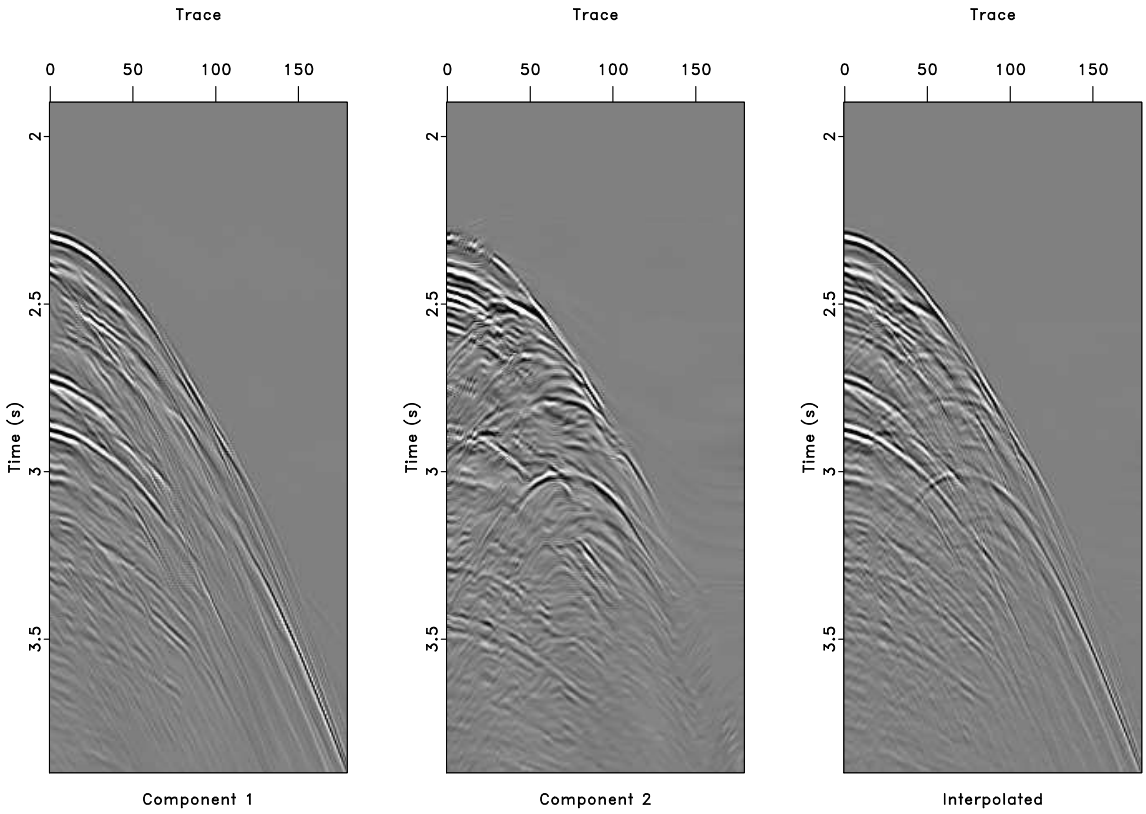

Figure 4: From left to right: MCA reconstructed component 1, component 2 and the final interpolated data.

seislet-MCA has less leakage, compared to the seislet-IRLS method. Note that the seislet-based MCA algorithm only uses 15 iterations to obtain the best separation effect, while the number of iterations for seislet-IRLS method is 1000. Therefore, seislet-based MCA is very efficient to demultiple.

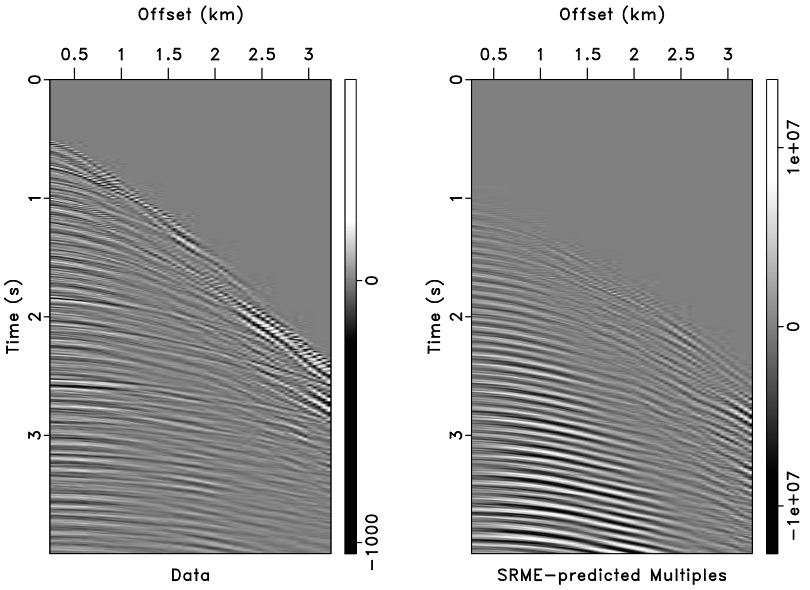

Figure 5: The field CMP data (left) and SRME predicted multiples (right). The amplitudes of SRME prediction needs to be corrected.

CONCLUSION AND DISCUSSION

We have developed a seislet-based MCA method for seismic data processing. PWD filter can be utilized to estimate the slopes of seismic data. An exponential shrinkage function is introduced to diversify the capability of sparsity-promoting shaping operator. The proposed seislet-based MCA using scaled-dependent shaping regularization is promising in the application to seismic trace interpolation, and multiple removal. The numerical results reveal that the exponential shrinkage operator in sparsity-promoting shaping regularization plays an extremely important role in successful seislet-based MCA separation, superior to many existing thresholding operators. The additional parameter p provides us more flexibility for approximating many existing shrinkage operators to achieve better separation performance. Meanwhile, it is free of non-differential singularity and unifies many existing shrinkage operators.

Seislet-MCA using PWD-based dip estimation is of special physical meaning for geophysical data in seismic processing, while the sparse dictionaries reported in Starck et al. (2004) are useful in image processing but lacking seismic attributes. However, a computational expensive optimization problem using least-squares minimization, which is not involved in the method of Starck et al. (2004), has to be solved to estimate the slope fields before applying our seislet-based method. Besides the computational

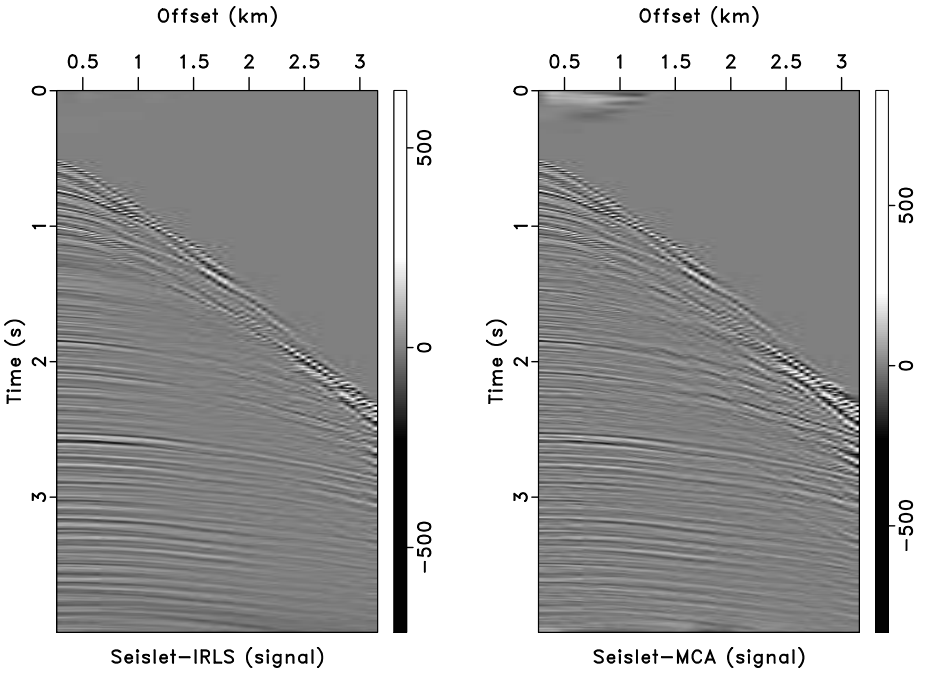

Figure 6: Separated primaries using seislet-IRLS (1000 iterations) and seislet-based MCA (15 iterations).

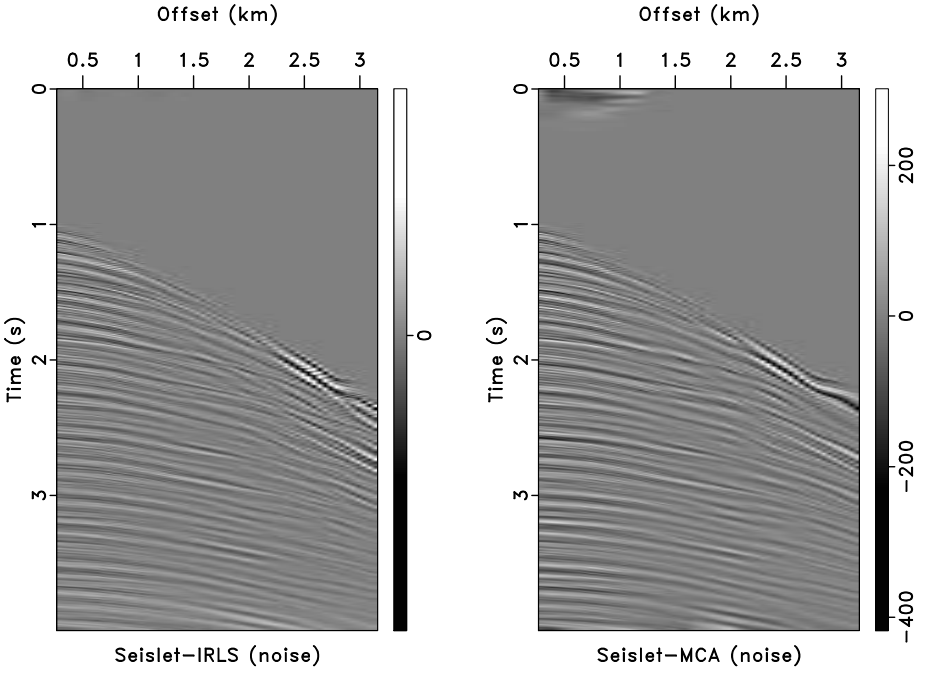

Figure 7: Separated multiples using seislet-IRLS (1000 iterations) and seislet-based MCA (15 iterations).

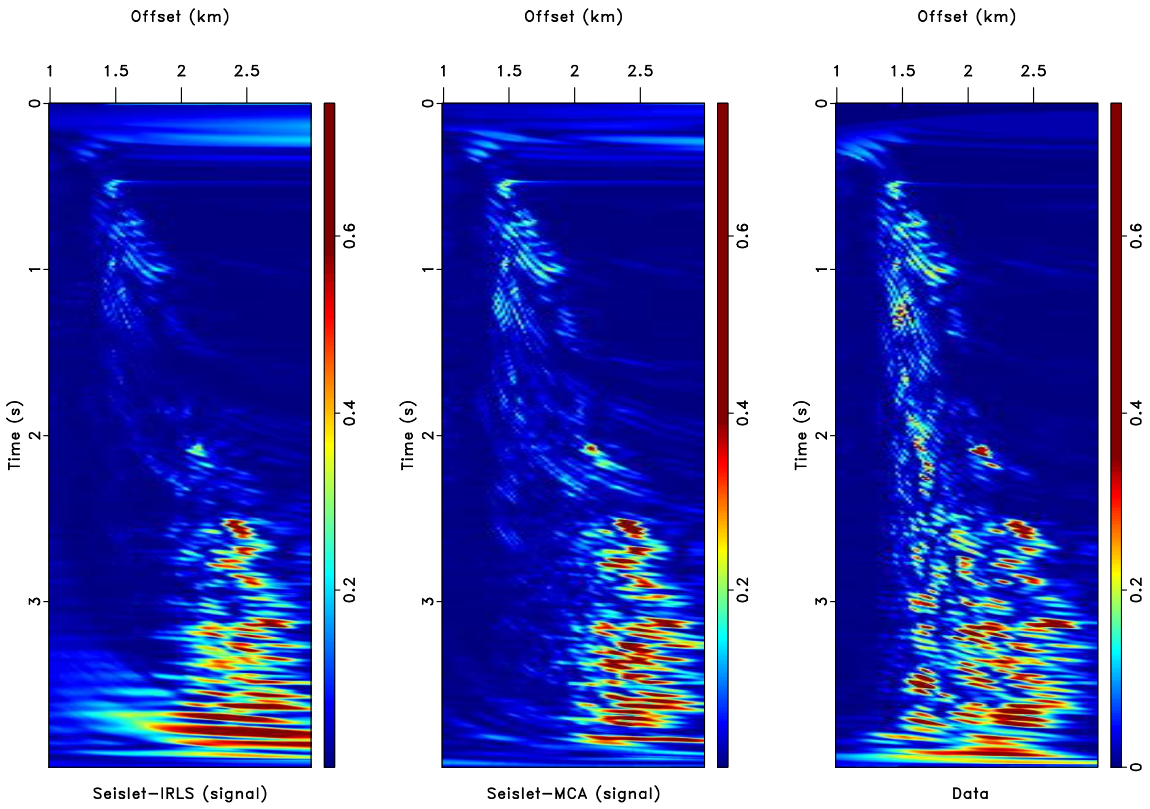

Figure 8: Velocity scan for primaries obtained by seislet-IRLS and seislet-based MCA (first two panels, from left to right). The velocity scan of original data is shown in the last panel for comparison.

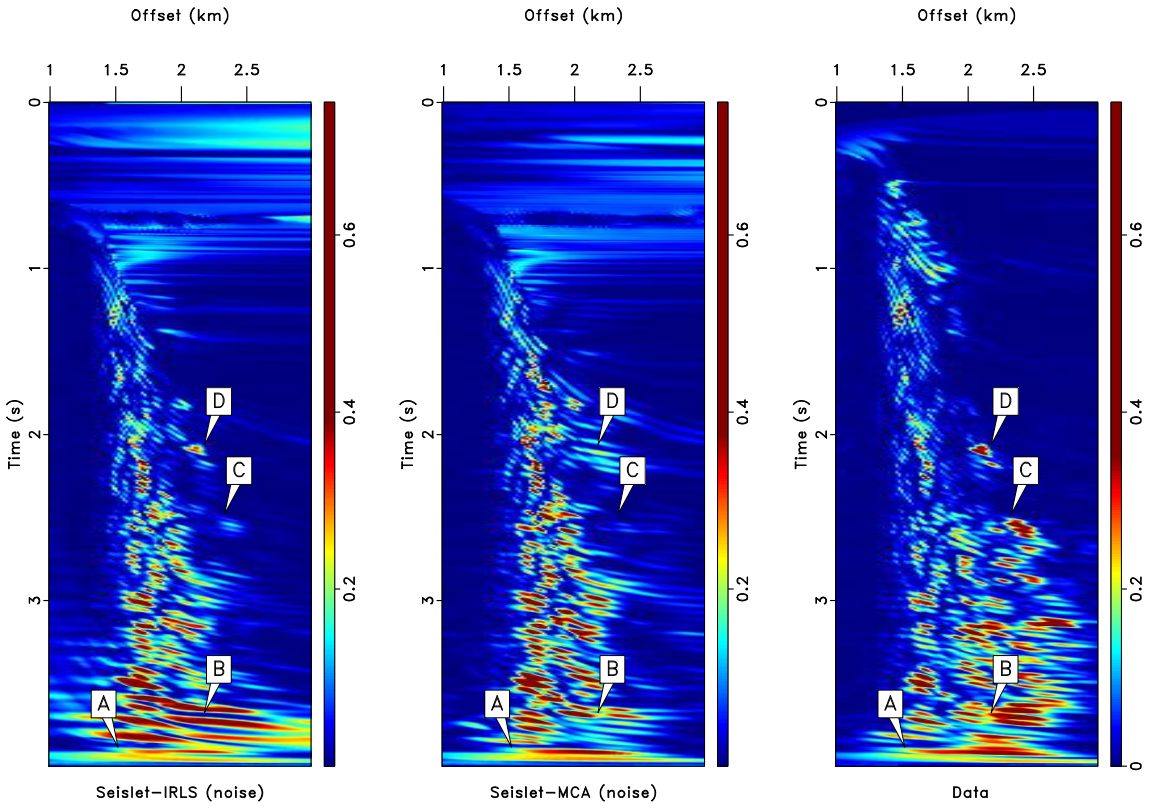

Figure 9: Velocity scan for multiples obtained by seislet-IRLS and seislet-based MCA (first two panels, from left to right). The velocity scan of original data is shown in the last panel for comparison. Seislet-based MCA outperforms seislet-based IRLS method in the locations A and B in the velocity scan panel due to the nice match of the corresponding semblance scan of the original data; at locations C and D, the energy of multiples obtained by seislet-MCA has less leakage, compared to seislet-IRLS method.

expensive slope estimation, the proposed method is very efficient for interpolation and separation. The method fails to interpolate the missing traces when the random decimating rate is larger than 70% for 2D seismic data, which honors the necessity of high dimensional data reconstruction using 3D seislet transform. Although numerically working well, up to now we have no theoretical convergence proof of the nonlinear shaping algorithm, and it remains an open problem for future works.

ACKNOWLEDGMENTS

The work of the first author was supported by the China Scholarship Council (No. 201306280082) during his visit to the Bureau of Economic Geology and the University of Texas at Austin. We would like to thank Wenchao Chen and Jianwei Ma for valuable suggestions. Reproducible examples are created using the Madagascar software package (Fomel et al., 2013).

CONNECTIONS BETWEEN SEISLET FRAME AND SEISLET-MCA ALGORITHM

The complete data d is regarded to be superposition of several different geometrical components, and each component can be sparely represented using a seislet dictionary Φ_i, i.e.,

$$\begin{aligned} d &= \sum_{i=1}^{N} d_i = \sum_{i=1}^{N} \Phi_i \alpha_i \\ &= [\Phi_1, \Phi_2, \cdots, \Phi_N] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \end{aligned} \tag{A-1}$$

where $F = [\Phi_1, \Phi_2, \cdots, \Phi_N]$ is a combined seislet dictionary (i.e. seislet frame), and the backward operator is chosen to be

$$B = \frac{1}{N} \begin{bmatrix} \Phi_1^* \\ \Phi_2^* \\ \vdots \\ \Phi_N^* \end{bmatrix} \tag{A-2}$$

in the sense that

$$FB = \frac{1}{N} \sum_{i=1}^{N} \Phi_i \Phi_i^* = \mathrm{Id}. \tag{A-3}$$

The difference between seislet-MCA algorithm and seislet frame minimization is the use of BCR technique (Bruce et al., 1998): We sparsify one component while keeping

all others fixed. At the $k+1$-th iteration applying the backward operator on the i-th component leads to

$$\tilde{\alpha}_i^{k+1} = \alpha_i^k + \sum_{i=1}^{N} \Phi_i^* r^k = \alpha_i^k + r_i^j + \sum_{j \neq i} \Phi_i^* r_j^k \tag{A-4}$$

where the terms $\sum_{j \neq i} \Phi_i^* r_j^k$ are the crosstalk between the i-th component and the others. An intuitive approach to filter out the undesired crosstalk is shrinkage/thresholding. The proposed exponential shrinkage provides us a flexible control on the performance of the shrinkage/thresholding operator.

REFERENCES

Bruce, A., S. Sardy, and P. Tseng, 1998, Block coordinate relaxation methods for nonparamatric signal denoising: Proceedings of SPIE, **3391**, 75–86.

Cao, J., Y. Wang, J. Zhao, and C. Yang, 2011, A review on restoration of seismic wavefields based on regularization and compressive sensing: Inverse Problems in Science and Engineering, **19**, 679–704.

Chartrand, R., 2012, Nonconvex splitting for regularized low-rank + sparse decomposition: IEEE Transactions on Signal Processing, **60**, 5810–5819.

Chartrand, R., and B. Wohlberg, 2013, A nonconvex admm algorithm for group sparsity with sparse groups: Presented at the Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

Chen, Z., S. Fomel, and W. Lu, 2013a, Accelerated plane-wave destruction: Geophysics, **78**, V1–V9.

———, 2013b, Omnidirectional plane-wave destruction: Geophysics, **78**, V171–V179.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications Cambridge, Massachusetts, USA, **6**.

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Commun. Pure Appl. Math., **57**, 1413–1457.

Daubechies, I., R. DeVore, M. Fornasier, and S. Gunturk, 2010, Iteratively reweighted least squares minimization for sparse recovery: Commun. Pure Appl. Math., **63**.

Donoho, D., 1995, De-noising by soft-thresholding: IEEE Transactions on Information Theory, **41**, 613–627.

Elad, M., P. Milanfar, and R. Rubinstein, 2007, Analysis versus synthesis in signal priors: Inverse Probl., **23**, 947–968.

Elad, M., J. Starck, P. Querre, and D. L. Donoho, 2005, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA): Applied and Computational Harmonic Analysis, **19**, 340 – 358.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

———, 2007, Shaping regularization in geophysical-estimation problems: Geophysics, **72**, R29–R36.

———, 2008, Nonlinear shaping regularization in geophysical inverse problems: SEG Annual Meeting, 2046–2051.

Fomel, S., and A. Guitton, 2006, Regularizing seismic inverse problems by model reparameterization using plane-wave construction: Geophysics, **71**, A43–A47.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: Geophysics, **75**, V25–V38.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

Gholami, A., 2014, Non-convex compressed sensing with frequency mask for seismic data reconstruction and denoising: Geophysical Prospecting, **62**, 1389–1405.

Gholami, A., and S. Hosseini, 2011, A general framework for sparsity-based denoising and inversion: IEEE Transactions on Signal Processing, **59**, 5202–5211.

Mallat, S., 2009, A wavelet tour of signal processing, 3rd ed.: Academic Press.

Marfurt, K. J., 2006, Robust estimates of 3D reflector dip and azimuth: Geophysics, **71**, P29–P40.

Mohimani, G. H., M. Babaie-Zadeh, and C. Jutten, 2009, A fast approach for overcomplete sparse decomposition based on smoothed l-0 norm: IEEE Transactions on Signal Processing, **57**, 289–301.

Ottolini, R., 1983, Signal/noise separation in dip space: Stanford Exploration Project: SEP report, **3**, 143–149.

Peyre, G., 2010, Advanced image, signal and surface processing.

Starck, J., M. Elad, and D. Donoho, 2004, Redundant multiscale transforms and their application for morphological component separation: Advances in Imaging and Electron Physics, **132**, 287–348.

———, 2005, Image decomposition via the combination of sparse representations and a variational approach: IEEE Trans. Image Process., **14**, 1570–1582.

Starck, J., J. Fadili, and F. Murtagh, 2007, The undecimated wavelet decomposition and its reconstruction: IEEE Trans. Image Process., **16**, 297–309.

Sweldens, W., 1998, The lifting scheme: A construction of second generation wavelets: SIAM Journal on Mathematical Analysis, **29**, 511–546.

Voronin, S., and R. Chartrand, 2013, A new generalized thresholding algorithm for inverse problems with sparsity constraints: 38th International Conference on Acoustics, Speech, and Signal Processing, IEEE, 1636–1640.

Woiselle, A., J. Starck, and J. Fadili, 2011, 3-D data denoising and inpainting with the low-redundancy fast curvelet transform: J. Math. Imaging Vis., **39**, 121–139.