

Seismic data restoration via data-driven tight frame

Jingwei Liang¹, Jianwei Ma², and Xiaoqun Zhang³

ABSTRACT

Restoration/interpolation of missing traces plays a crucial role in the seismic data processing pipeline. Efficient restoration methods have been proposed based on sparse signal representation in a transform domain such as Fourier, wavelet, curvelet, and shearlet transforms. Most existing methods are based on transforms with a fixed basis. We considered an adaptive sparse transform for restoration of data with complex structures. In particular, we evaluated a data-driven tight-frame-based sparse regularization method for seismic data restoration. The main idea of the data-driven tight frame (TF) is to adaptively learn a set of framelet filters from the currently interpolated data, under which the data can be more sparsely represented; hence, the sparsity-promoting ℓ_1 -norm (SPL1) minimization methods can produce better restoration results by using the learned filters. A split inexact Uzawa algorithm, which can be viewed as a generalization of the alternating direction of multiplier method (ADMM), was applied to solve the presented SPL1 model. Numerical tests were performed on synthetic and real seismic data for restoration of randomly missing traces over a regular data grid. Our experiments showed that our proposed method obtains the state-of-the-art restoration results in comparison with the traditional Fourier-based projection onto convex sets, the tight-frame-based method, and the recent shearlet regularization ADMM method.

INTRODUCTION

A successful seismic data acquisition program requires careful and detailed planning before fieldwork begins. Such planning should include budget acquisition costs. Generally, more funds are used for acquisition than data processing; thus, it is necessary

to use as few as possible receivers to reduce the total cost. In practice, receivers spaced along a line are used to record the signal from a series of source points. To avoid information loss, seismic data should be sampled according to the Shannon-Nyquist criterion. However, seismic data are often sparsely or incompletely sampled along the spatial coordinates, partly caused by surface obstacles, dead trace, no-permit areas, and economic constraints.

The goal of interpolation or restoration of traces refers to determining the values at locations where measurements are not acquired, for which increasing sampling ratios and decreasing acquisition costs are allowed. The quality of restoration will impact the subsequent seismic processing steps, e.g., multiple suppression, migration, imaging, and amplitude variation with offset (AVO). Sampling patterns could be regular or irregular in practice. For regular sampling, the traces are placed equidistantly on uniform grids. Irregular sampling data can be further divided into two subcategories: randomly sampling (missing traces) on a uniform regular grid or on an irregular grid (i.e., purely irregular). In this paper, we consider random sampling for which data are acquired from a uniform regular grid.

Over the past few decades, various methods based on wave equations and signal processing techniques have been proposed for seismic data restoration. Because our proposed method emerged from imaging science, we will forgo the introduction of wave-equation-based methods and focus on the latter ones. Most signal-processing-based methods involve representations in some transform domain, such as the Radon transform (Kabir and Verschuur, 1995; Trad et al., 2002), Fourier transform (Sacchi et al., 1998; Gulunay, 2003; Liu and Sacchi, 2004; Xu et al., 2005; Abma and Kabir, 2006; Trad, 2009), seislet transform (Fomel and Liu, 2010; Liu and Fomel, 2010), and curvelet transform (Hennenfent and Herrmann, 2008; Herrmann and Hennenfent, 2008; Naghizadeh and Sacchi, 2010; Wang et al., 2011; Shahidi et al., 2013). In Hennenfent and Herrmann (2008), Herrmann and Hennenfent (2008), Shahidi et al. (2013), and Wang et al. (2011), the authors apply the sparsity-promoting ℓ_1 -norm (SPL1) minimization of curvelet coefficients,

Manuscript received by the Editor 7 July 2013; revised manuscript received 23 November 2013; published online 28 March 2014.

¹Université de Caen, GREYC, CNRS-ENSICAEN, Caen, France. E-mail: jingwei.liang@ensicaen.fr.

²Harbin Institute of Technology, Department of Mathematics, Harbin, China. E-mail: jma@hit.edu.cn.

³Shanghai Jiao Tong University, MOE-LSC and Institute of Natural Sciences, Department of Mathematics, Shanghai, China. E-mail: xqzhang@sjtu.edu.cn.

© 2014 Society of Exploration Geophysicists. All rights reserved.

which leads to improved results for event-preserving restoration. In a recent work, S. Hauser (personal communication, 2012) propose a new method for seismic data restoration by applying directional weighted shearlet regularization. An important issue in trace interpolation is related to aliasing. A popular technique is prediction filters. For instance, the low-frequency nonaliased components of the observed data are used to build antialiasing prediction-error filters, and then these filters are applied to interpolate high-frequency components or missing traces. The prediction-filter method can be implemented in the time-spatial t - x domain (Crawley et al., 1999; Liu and Fomel, 2011), the frequency-spatial f - x domain (Spitz, 1991; Porsani, 1999; Naghizadeh and Sacchi, 2007), the frequency-wavenumber f - x domain (Gulunay, 2003), and even the curvilinear domain (Naghizadeh and Sacchi, 2010).

Rank-reduction methods are also introduced to seismic data restoration. The motivation of these methods is that the seismic data can be approximated by low-rank structures due to redundancy. Trickett et al. (2010) present a truncated singular value decomposition (SVD)-based matrix-rank reduction of constant-frequency slices for trace interpolation. Oropeza and Sacchi (2011) reorganize the seismic data into a Hankel matrix and then use multichannel singular spectrum analysis to solve the rank-reduction problem. The above methods are named the *Cadizow method*. Recently, the Cadizow method was extended to high-dimensional interpolation (Gao et al., 2013a; Naghizadeh and Sacchi, 2013) and dealiasing interpolation. Kreimer and Sacchi (2012) propose a tensor rank-reduction method by higher-order SVD for prestack seismic data interpolation. Alternately, Yang et al. (2013) apply nuclear-norm-minimization-based matrix completion for seismic data restoration. Further, a 3D seismic data completion is presented in Ma (2013).

The method in this paper belongs to the category of transform-based methods. Instead of using a known set of basis functions to estimate representation coefficients, we propose to learn basis functions from the given data, and then use the learned basis for estimating missing data by a sparsity-promoting model. More precisely, we propose to use the data-driven tight frame (TF), recently developed in Cai et al. (2013), as the sparse transform for the seismic data restoration problem. The key advantage of our work is that, unlike framelets, curvelets, and shearlets, the filters of the data-driven TF are adaptively learned from the data, which in turn gives a sparser representation for the data; hence, sparsity-promoting models can lead to a better restoration result. The method has some similarities to the estimation of basis functions via independent component analysis (ICA) proposed in Kaplan and Sacchi (2009). However, our proposed method is a learning algorithm based on a variational sparse representation model. Furthermore, the method in Kaplan and Sacchi (2009) is proposed for denoising whereas ours is for missing data reconstruction. Overall, our method consists of two steps: (1) adaptively learn a set of TF filters from the preprocessed data and (2) reconstruct missing data by using the sparse regularization model with the TF system formed by the learned filters. This procedure is described in Algorithm 1.

The rest of the paper is organized as follows: In the second section, we will introduce the sparse regularization model for seismic data restoration. The split inexact Uzawa algorithm that is used to solve the sparse regularization model will be described in the third section. We present numerical experiments and compare different methods in the fourth section. A brief introduction of the wavelet TF and data-driven TF will be given in Appendix A.

SPARSE REGULARIZATION MODEL

The task of seismic data restoration is to recover the missing traces from the subsampled data. The observation model of this problem can be described as

$$f = Pu + \epsilon, \quad (1)$$

where u is the underground truth that we want to recover, f is the subsampled data with missing traces, P is the trace subsampling operator, and ϵ usually refers to zero-mean white Gaussian additive noise. In this paper, we consider the case that the underlying grid is regular and the sampling traces are randomly chosen from the uniform grid to produce an irregular distribution of sampling traces.

The usage of sparsity as a prior has been widely adopted in various image processing applications. One popular way to achieve sparsity is solving the related ℓ_1 -norm minimization problem due to its simplicity. There are mainly two types of sparsity prior models: analysis-based approach and synthesis-based approach. The analysis-based approach for equation 1 can be formulated as a constrained minimization problem:

$$\min_u \|Wu\|_1 \text{ s.t. } Pu = f, \quad (2)$$

where W is a sparse transform/decomposition operator and $\|\cdot\|_1$ refers to the vectorial ℓ_1 -norm. For the noised case, the constraint $\|Pu - f\|_2 \leq \delta$ can be considered instead of the equality and δ is a parameter related to the standard variance of the noise. The synthesis-based approach solves the following problem:

$$\min_d \|d\|_1 \text{ s.t. } PW^T d = f, \quad (3)$$

where W^T is the corresponding reconstruction operator and d is the coefficients of u under transform W . Similarly, equality constraint can be replaced with $\|PW^T d - f\|_2 \leq \delta$ for the noised case.

The analysis-based approach emphasizes the sparsity of the canonical transformed coefficients, so it tends to recover data with smooth regions; while the synthesis-based approach finds the sparsest approximation of the given data in the transformed domain. Sparse regularization models mainly involve two aspects: how to choose/design the sparse transform W and how to solve the optimization model efficiently. For the first problem, plenty of techniques have been developed for the aforementioned sparse regularization models. In this paper, we use one of these techniques, which has been successfully applied to various image-processing problems and its new development: the wavelet TF (Daubechies et al., 2003; Shen, 2010) and the data-driven TF (Cai et al., 2013). The filters of the former one are fixed, whereas the filters of the latter one are adaptively learned from the data. A frame is tight if it obeys a generalized Parseval identity and it generalizes the orthogonal basis set with a certain amount of redundancy. A brief introduction of the TF is presented in Appendix A. In short, the TF operator W is a linear operator implemented by convolution with filters and has the property $W^T W = I$, while $W W^T$ is not necessarily equal to identity. In other words, if we apply a forward transform to a signal and then apply its adjoint transform, we can exactly reconstruct this signal. However, if we first apply an adjoint transform to a set of coefficients and then apply its forward transform, the TF may not reconstruct the original coefficients. If W

is orthogonal, we have $W^T W = I$ and $W W^T = I$, then the above-mentioned analysis and synthesis model are equivalent.

We outline the framework of our proposed method in Algorithm 1. The data-driven TF is a dictionary learning (or sparse coding) method that builds an adaptive discrete TF to sparsely represent the given data. The key ideas of the data-driven TF are simple:

- 1) Construct a bank of good local filters. The word *local* implies that the filters are compactly supported. If we assume that the size of the support is $r \times r$, or, the dimension of the filter coefficients is r^2 , then we can construct r^2 filters. These filters are constrained to form a TF (even orthogonal), which is used to represent the given data.
- 2) Apply the TF transform using these learned filters, which involves a group of convolution operation on the data to get TF coefficients. The good fitting of the filters is obtained by minimizing the sparsity of the corresponding coefficients.
- 3) The minimization of the sparse-norm requires two alternating steps: (1) Finding the set of coefficients with a fixed filter. This can be implemented by hard thresholding. (2) Finding the filters with the set of coefficients in the last step. This is solved by an SVD decomposition. Under certain assumptions (Cai et al., 2013), one can implement this step much faster than other existing dictionary-learning algorithms (e.g., K-SVD [Elad and Aharon, 2006; Aharon et al., 2006]).

The detail of the data-driven TF is presented in Appendix A. Once we find a good set of filters, we can reconstruct the missing data with the sparse restoration models mentioned above, and we will provide a fast first-order optimization algorithm in the next section to solve these problems.

Algorithm 1. Seismic data restoration via data-driven TF.

initial interpolated data u^0 ,

for $n = 0 \rightarrow N - 1$ **do**

1) learn a set of framelet filters from u^n by Algorithm 3, then generate the analysis operator W^n ;

2) update u^{n+1} with the sparse restoration model (equation 2) using W^n ;

end for

Output u^N , the restored data.

SPLIT INEXACT UZAWA ALGORITHM

In this section, we present the algorithm to solve the model (equation 2). We shall forgo the detailed derivation of the algorithm; interested readers can consult Zhang et al. (2010, 2011) and the references therein.

We first reformulate the model (equation 2) for the noise-free case by adding the auxiliary variable:

$$\min_{u,d} \|d\|_1 \text{ s.t. } Pu = f, \quad Wu = d. \quad (4)$$

This problem can be easily solved by the ADMM (Boyd et al., 2011), also known as the *split Bregman method* (Goldstein and Osher, 2009), a popular algorithm that is widely used in imaging science. In this paper, we apply another efficient method, the split inexact Uzawa method proposed in Zhang et al. (2011), which can be viewed as a generalization of the ADMM.

The augmented Lagrangian of the above minimization model is given as

$$L(u, d; b, c) = \mu \|d\|_1 - \langle Pu - f, c \rangle - \langle d - Wu, b \rangle + \frac{\delta}{2} \|Pu - f\|^2 + \frac{\lambda}{2} \|Wu - d\|^2, \quad (5)$$

where b and c are dual variables and $\mu, \delta, \lambda > 0$. Let M be a positive definite matrix, and we define the induced norm $\|x\|_M^2 = \langle Mx, x \rangle$. The idea of the split inexact Uzawa method is to add a proximal term $\frac{1}{2} \|u - u^k\|_M^2$ to the augmented formula and apply the alternating minimization to the primal and dual variables, which leads to the following scheme: Let $d^0 = b^0 = 0$ and $c^0 = 0$,

$$\begin{cases} u^{k+1} = \arg \min_u (L(u, d^k; b^k, c^k) + \frac{1}{2} \|u - u^k\|_M^2) \\ d^{k+1} = \arg \min_d L(u^{k+1}, d; b^k, c^k) \\ b^{k+1} = b^k + \lambda (Wu^{k+1} - d^{k+1}) \\ c^{k+1} = c^k + \delta (f - Pu^{k+1}) \end{cases}. \quad (6)$$

The purpose of the proximal term $\frac{1}{2} \|u - u^k\|_M^2$ is to cancel out the term Pu and simplify the minimization step by decoupling the variables coupled by the operator P . In this paper, by choosing $M = \text{Id} - \delta P^T P$ and the change of variables $c^k \leftarrow \frac{1}{\delta} c^k, b^k \leftarrow \frac{1}{\lambda} b^k$, we obtain the following scheme for the analysis-based approach using Algorithm 2.

Algorithm 2. Split inexact Uzawa algorithm for the analysis-based approach.

initialed by u^0 , set $b^0 = c^0 = 0, v^0 = 0, d^0 = Wu^0$,

while $k = 0 \rightarrow \text{maxits}$ **do**

1) $v^{k+1} = u^k - \delta P^T (Pu^k - f - c^k)$

2) $u^{k+1} = \arg \min_u (\frac{\lambda}{2} \|Wu - d^k + b^k\|^2 + \frac{1}{2} \|u - v^{k+1}\|^2)$

3) $d^{k+1} = \arg \min_d (\mu \|d\|_1 + \frac{\lambda}{2} \|d - (Wu^{k+1} + b^k)\|^2)$

4) $b^{k+1} = b^k + (Wu^{k+1} - d^{k+1})$

5) $c^{k+1} = c^k + (f - Pu^{k+1})$

end while

Output the result u .

Because the TF satisfies $W^T W = I$, the second subproblem has a close formula:

$$u^{k+1} = \frac{1}{\lambda + 1} (\lambda W^T (d^k - b^k) + v^{k+1}), \quad (7)$$

and the third one is also solved by the pointwise shrinkage formula:

$$d^{k+1} = \text{shrinkage}\left(v^{k+1}, \frac{\mu}{\lambda}\right) \\ := \text{sign}(v^{k+1}) \max\left(|v^{k+1}| - \frac{\mu}{\lambda}, 0\right). \quad (8)$$

Overall, we can see that each step of the algorithm can be efficiently solved without involving any inverse of the large matrix. In practice, the choice of parameters is quite trivial: δ and λ could just set to 1, and $\mu > 0$ for satisfying the convergence condition.

NUMERICAL RESULTS

In this section, we present numerical results on synthetic and real seismic data (see Figure 1). The methods we compare are the wavelet TF sparse restoration model (Cai et al., 2009), shearlet restoration model (S. Hauser, personal communication, 2012), and projection onto convex sets (POCS) (Gao et al., 2013b); the proposed method is denoted as DDTF.

To assess the performance of the methods, we compare the peak signal-to-noise ratio (PSNR) value and the visual quality of the results; the PSNR (dB) is defined as

$$\text{PSNR} = 10 \log \left(\frac{(\max_u - \min_u)^2}{\frac{1}{MN} \sum_{i,j} (u_{i,j} - \tilde{u}_{i,j})^2} \right), \quad (9)$$

where u is the assumed original data, \tilde{u} is the restored data, and M and N denote the total sampling number in the temporal and spatial directions, respectively.

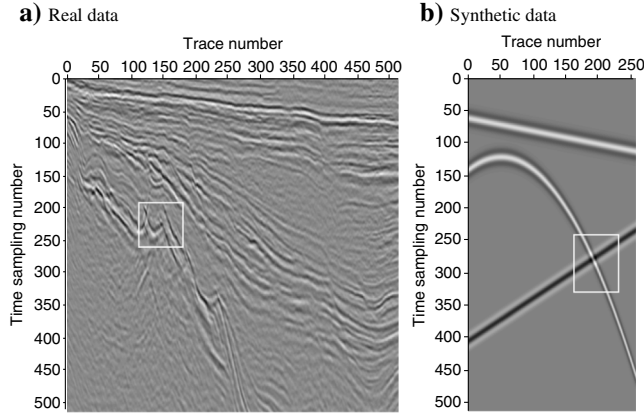


Figure 1. Data used in the experiment; the first dimension denotes the temporal sampling, and the other dimension denotes the spatial trace sampling.

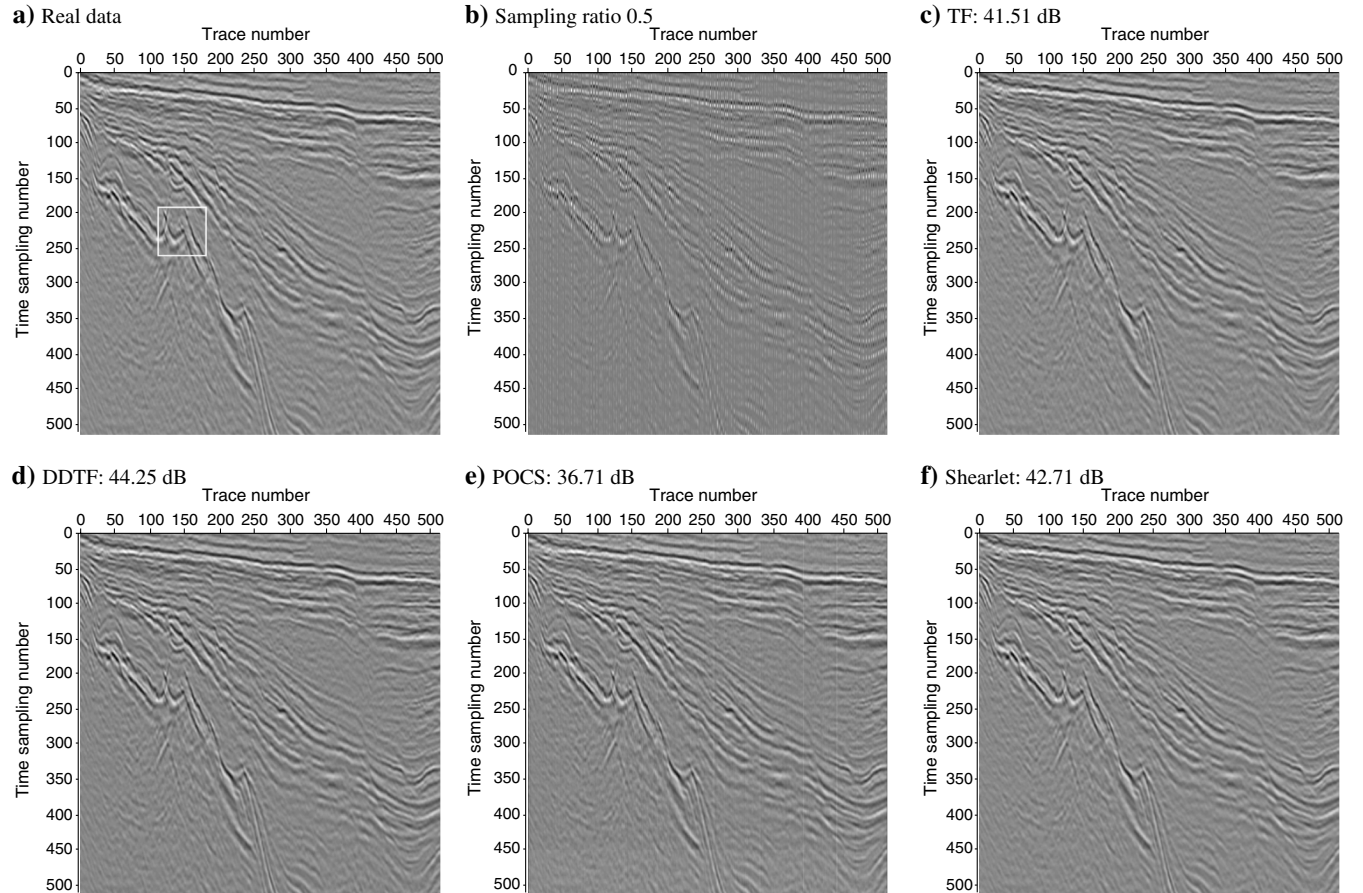


Figure 2. Restored results of four methods on real data with an irregular sampling ratio of 0.5.

For the wavelet TF, a cubic spline wavelet is applied. For the data-driven TF, the filter size is set to be 7×7 . For the POCS method, seismic data are cropped into overlapped patches to suppress the artifacts resulted from the global Fourier transform.

We consider several subsampling ratios vary from 0.2 to 0.7. For the sake of generality, we repeat the test with 10 randomly sampled operators/masks for each subsampling ratio, and then we take the mean value of the 10 PSNRs as the PSNR at this subsampling ratio. Table 1 provides a complete comparison of the PSNR values for all the sampling ratios, and Figures 2–5 show the comparison for one sampling ratio.

Figure 2 shows the results of the four methods under the subsampling ratio 0.5: From the PSNR point of view, DDTF produces the best result. In Figure 3, we compare the wiggle plot of the patch marked in Figure 1a: From a visual quality point of view, DDTF also suffers from the least artifacts. Then, for the synthetic data, Figure 4 shows the results of the four methods: The shearlet method obtains the highest PSNR value. For the patch marked in Figure 1b, a comparison is presented in Figure 5. The shearlet method suffers the least artifacts, and although the PSNR value of the POCS method is lower than that with the TF method, the amplitude of the artifact is smaller than the TF's.

Table 1. PSNR (dB) comparison of four methods, a significant improvement of DDTF over TF.

Sampling ratio		0.2	0.3	0.4	0.5	0.6	0.7
Real data	TF	27.72	30.68	34.12	41.51	41.65	44.35
	DDTF	29.71	33.21	37.21	44.09	44.53	46.78
	POCS	24.39	27.83	31.30	35.34	35.89	37.93
	Shearlet	27.69	32.02	36.45	42.67	42.55	43.14
Synthetic data	TF	32.03	35.05	38.84	48.61	49.91	50.99
	DDTF	32.93	41.69	48.12	58.38	59.27	60.71
	POCS	29.50	34.55	36.29	40.01	40.07	40.92
	Shearlet	37.41	46.54	50.39	60.57	56.14	55.47

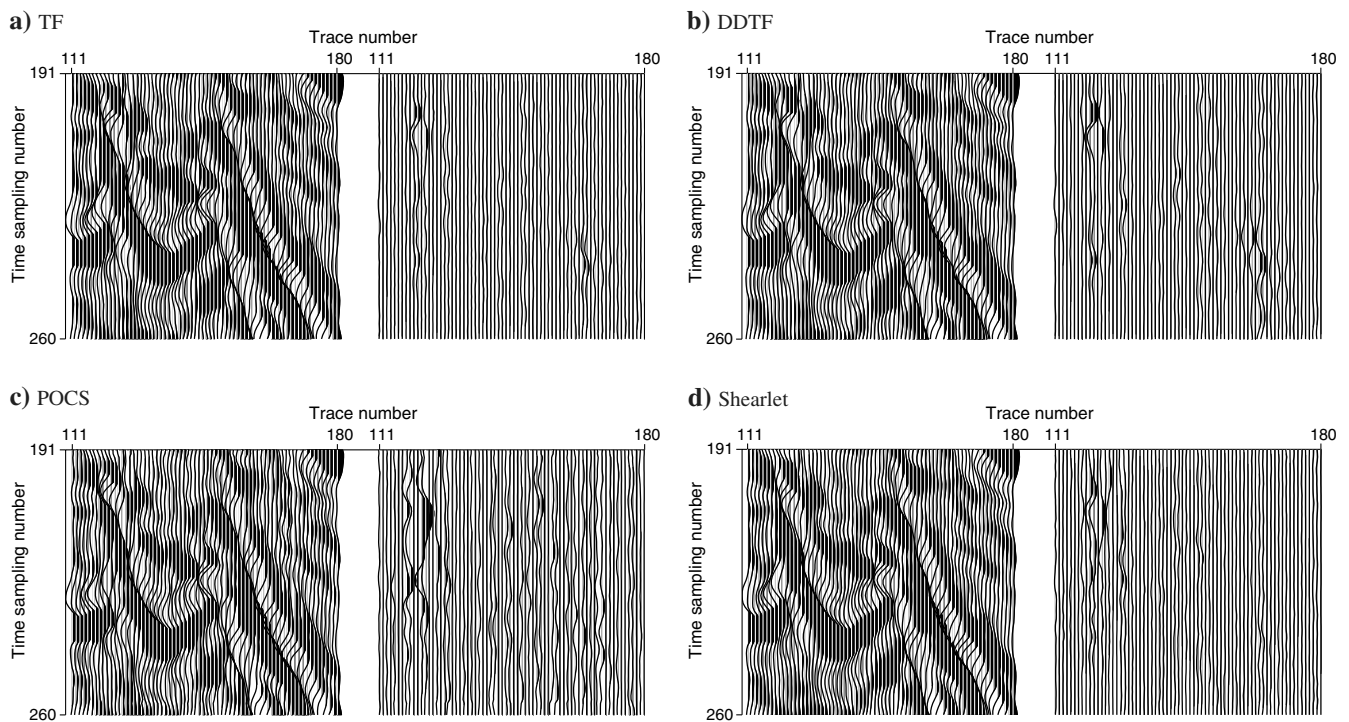


Figure 3. Zoom-in of marked patches and their corresponding differences to the original image patch. The result of DDTF suffers the least artifacts.

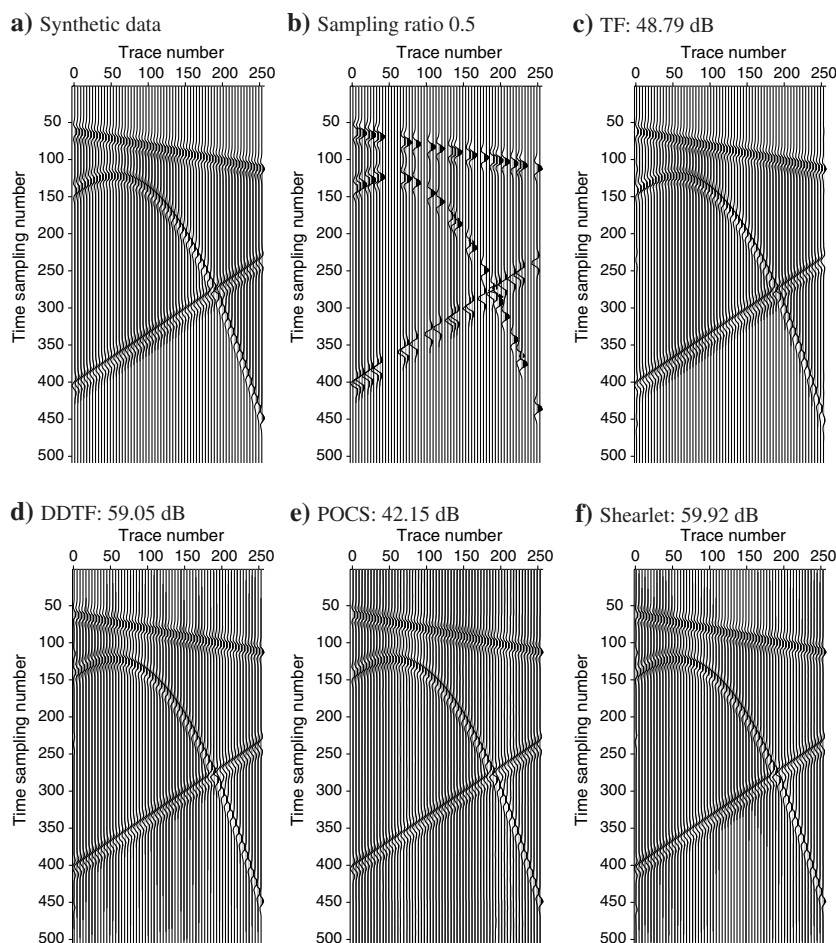
CONCLUSIONS

In this paper, we have presented a data-driven TF approach for seismic data restoration with randomly missing traces. Starting from the initial filters, we can learn a bank of compactly supported filters and reconstruct the unknown data by an SPL1 minimization model. The proposed model can be efficiently solved by a first-order fast method. Our numerical results have shown that the proposed method achieves state-of-art results compared to Fourier-based POCS, the TF method with a fixed basis, and the recently proposed shearlet-based method, especially for real data. Finally, our method can be extended to 3D data and data with noise. Furthermore, the dealiased restoration of the regularly missing trace and the parallelization and acceleration of the proposed method will also be investigated in a future work.

ACKNOWLEDGMENTS

We would like to thank all the reviewers for their useful suggestions for improving the presentation of this paper. The work is supported by National Natural Science Fund of China (grant number: 41374121, 61327013, 91330108, 11101277, 91330102), the Fundamental Research Funds for the Central Universities (grant

Figure 4. Restored results of four methods on synthetic data with irregular sampling ratio 0.5. For the parabolic-like event, the shearlet recovers the best result, due to the fact that shearlet filters are specific for structures such as this.



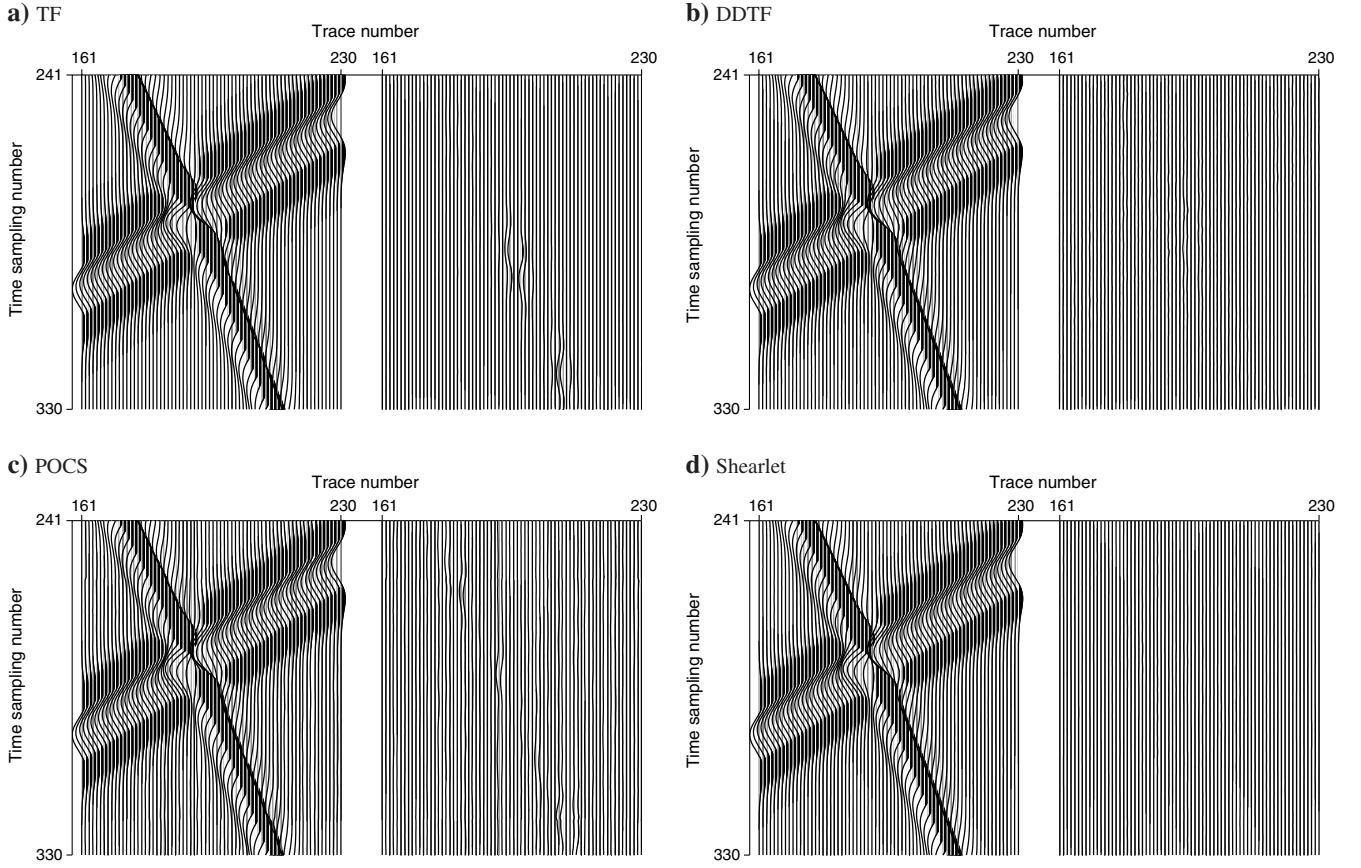


Figure 5. Zoom-in of marked patch of each method and the corresponding differences to the original image patch. The DDTF improves significantly from the TF; however, artifacts are still visible. If we take the filter size of the DDTF to 9×9 or 11×11 , the artifacts can be further suppressed, but this will increase the computational time.

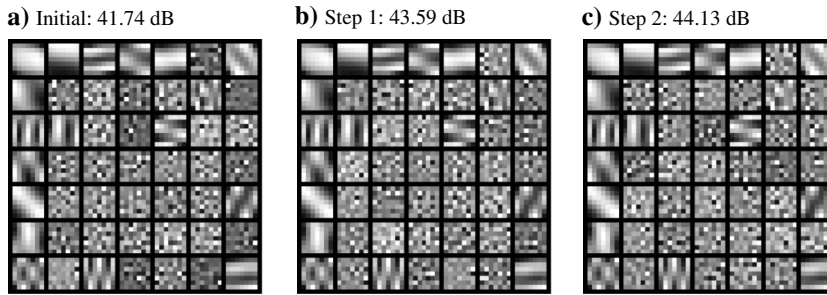


Figure 6. Here, the subsampling ratio is 0.5. (a) Filters learned from the initial data obtained by cubic spline interpolation. (b) Filters learned from the data restored using the filters in (a). (c) Filters learned from the data restored using the filters in (b).

number: HIT.BRETIV.201314), and the Program for New Century Excellent Talents in University (grant number: NCET-11-0804).

APPENDIX A

DATA-DRIVEN TIGHT FRAME

The wavelet TF has been successfully applied to various data processing problems (Chan et al., 2003; Cai et al., 2008, 2009,

2010; Dong et al., 2012). In this section, we first briefly introduce the concept of the wavelet TF (Daubechies et al., 2003; Dong and Shen, 2010; Shen, 2010), and we then introduce the newly developed data-driven TF technique for denoising (Cai et al., 2013).

Wavelet TF

A countable set $\{h_i\}_{i \in \mathcal{I}} \subset \mathcal{H}$ is a frame for \mathcal{H} if there exist two positive constants a and b such that

$$a\|f\|_2^2 = \sum_{i \in \mathcal{I}} |\langle h_i, f \rangle|^2 \leq b\|f\|_2^2, \quad \forall f \in \mathcal{H}, \quad (\text{A-1})$$

where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the inner product and norm of a Hilbert space \mathcal{H} . When $a = b = 1$, frame $\{h_i\}_{i \in \mathcal{I}}$ is called the TF. There are two operators associated with a given frame $\{h_i\}_{i \in \mathcal{I}}$: the analysis operator W defined by

$$W: f \in \mathcal{H} \mapsto \{\langle f, h_i \rangle\} \in \ell^2(\mathbb{N}), \quad (\text{A-2})$$

and its adjoint operator W^T , called the *synthesis operator*, defined by

$$W^T: \{d_i\} \in \ell^2(\mathbb{N}) \mapsto \sum_{i \in \mathcal{I}} a_i h_i \in \mathcal{H}. \quad (\text{A-3})$$

The composition of the two operators forms a so-called *frame operator* $\mathcal{T} = W^T W$ defined by

$$\mathcal{T}: f \in \mathcal{H} \mapsto \sum_{i \in \mathcal{I}} \langle f, h_i \rangle h_i; \quad (\text{A-4})$$

therefore, sequence $\{h_i\}_{i \in \mathcal{I}}$ forms a frame if $a\text{Id} \leq \mathcal{T} \leq b\text{Id}$, and it forms a TF if $\mathcal{T} = \text{Id}$, where Id is the identity operator in \mathcal{H} .

One widely used TF is the wavelet TF, and the corresponding $\{h_i\}_{i \in \mathcal{I}}$ are called *framelets*. To construct the wavelet TF, one can apply the unitary extension principle proposed in Ron and Shen, 1997. For the TF method that we compare in the “Numerical results” section, we adopt the linear B-spline framelet (Chai and Shen, 2007; Cai et al., 2008, 2009), which is as follows:

$$h_1 = \frac{1}{4}[1, 2, 1], \quad h_2 = \frac{\sqrt{2}}{4}[1, 0, -1], \quad h_3 = \frac{1}{4}[-1, 2, -1]. \quad (\text{A-5})$$

The wavelet TF for $L_2(\mathbb{R}^n)$ can be easily constructed by the tensor product of the 1D framelet (Daubechies, 1992; Dong and Shen, 2010).

In a discrete setting, let the wavelet TF system be of m compact supported filters $\{h_i\}_{i=1}^m$, data $u \in \mathbb{R}^N$. For each framelet filter h_i , the convolution operation can be represented by the Toeplitz matrix (Chan et al., 2004) $T_{h_i}: \mathbb{R}^N \rightarrow \mathbb{R}^N$; then, the analysis operator $W \in \mathbb{R}^{mN \times N}$ is defined as

$$W = \begin{bmatrix} T_{h_1} \\ \vdots \\ T_{h_m} \end{bmatrix}_{mN \times N}, \quad (\text{A-6})$$

and the synthesis operator is $W^T \in \mathbb{R}^{N \times mN}$. Note that $W^T W = \text{Id}$; i.e., $u = W^T W u$, while $W W^T$ is not necessarily equal to the identity (Cai et al., 2009).

Data-driven TF

As aforementioned, high-dimensional framelets are generated by the tensor product of 1D framelets. This approach is rather simple to implement and can sparsely represent given data; however, tensor product framelets mainly focus on horizontal and vertical singularities. When the geometry of the data is complex, full of rich textures, for example, then the representation under the tensor product framelets would become less efficient.

Research of the data-driven TF is focused on adopting the machine-learning methodology (Olshausen and Field, 1997; Lewicki and Sejnowski, 2000; Kreutz-Delgado et al., 2003; Aharon et al., 2006; Elad and Aharon, 2006; Mairal et al., 2008, 2009) to construct a set of shift-invariant framelets. In a nutshell and for given data u , a set of real-value filters $\{h_i\}_{i=1}^m$ with compact support can be learned so that the given data could be represented more sparsely. This purpose is realized by solving the following minimization problem:

$$\min_{d, \{h_i\}_{i=1}^m} \lambda \|d\|_0 + \frac{1}{2} \|d - W u\|_2^2, \quad \text{s.t. } W^T W = I, \quad (\text{A-7})$$

where $\|\cdot\|_0$ means to account the number of nonzero elements; d is the coefficient vector, which sparsely approximates the canonical TF coefficient $W u$; W is the analysis operator associated with $\{h_i\}_{i=1}^m$; and W^T is the corresponding synthesis operator.

The minimization problem can be solved alternately by updating coefficient vector d and filters $\{h_i\}_{i=1}^m$. More specifically, let $\{h_i^{(0)}\}_{i=1}^m$ be the initial filters, e.g., the linear B-spline framelet filters, then the algorithm for solving equation A-7 is found below.

Algorithm 3. Data-driven TF.

Initial framelet filters $\{h_i^{(0)}\}_{i=1}^m$,

for $k = 0 \rightarrow K - 1$ **do**

1) Fix framelet filters $\{h_i^{(k)}\}_{i=1}^m$, update the framelet coefficient $d^{(k+1)}$,

$$d^{(k+1)} = \arg \min_d \lambda \|d\|_0 + \frac{1}{2} \|d - W^{(k)} u\|_2^2. \quad (\text{A-8})$$

2) Fix the framelet coefficient $d^{(k+1)}$, update the framelet filters $\{h_i^{(k+1)}\}_{i=1}^m$,

$$\{h_i^{(k+1)}\}_{i=1}^m = \arg \min_{\{h_i\}_{i=1}^m} \|d^{(k+1)} - W u\|_2^2,$$

$$\text{s.t. } W^T W = I, \quad (\text{A-9})$$

end for

Output filters $\{h_i\}_{i=1}^m$.

The solution of step 1 (equation A-8) is given by the standard hard threshold; i.e.,

$$d = T_\lambda(W^{(k)} u), \quad (\text{A-10})$$

$$\text{where } [T_\lambda(W u)](i) = \begin{cases} W u(i), & |W u(i)| \geq \sqrt{2\lambda} \\ 0, & |W u(i)| < \sqrt{2\lambda} \end{cases}.$$

For the subproblem (A-9), in general it is a complex nonconvex problem and hard to solve directly. However, under certain proper assumptions (Cai et al. [2013], subsection 3.3), the optimal solution can be given in a close form, which greatly simplifies the computation.

In Figure A-1, a comparison of the wavelet TF and data-driven TF is demonstrated. Figure A-1a is the initial linear B-spline framelets, and Figure A-1b is the learned data-driven framelet filters via Algorithm 3 for the given seismic data shown in Figure 1a, and Figure A-1c shows the plot of the sorted absolute value of the coefficient of the data under the two framelet systems. The coefficient of the learned filters (denoted by the blue line) decays faster than the linear B-spline filters.

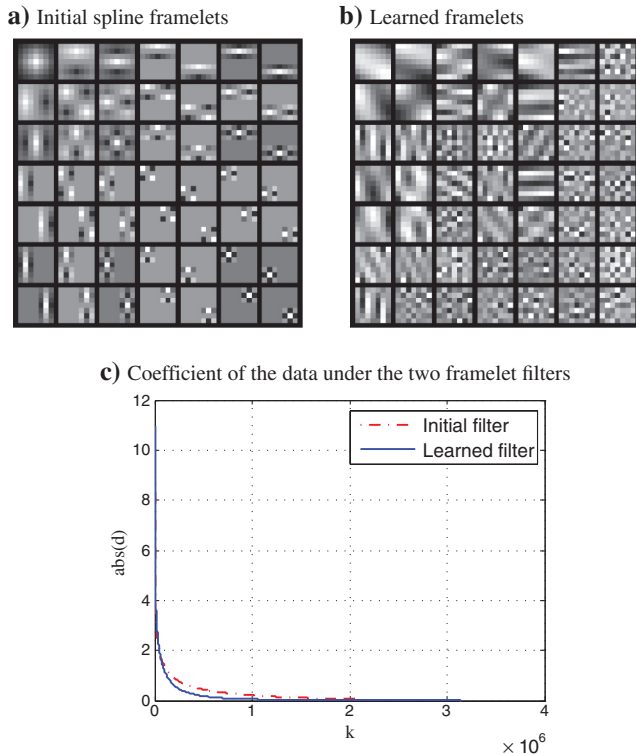


Figure A-1. Comparison of the tensor product framelets and data-driven tight framelets. (a) Initial linear B-spline framelets. (b) Learned framelet filters via Algorithm 3 for the given seismic data shown in Figure 1a. (c) Plot of the sorted absolute value of the coefficient of the data under the two framelet systems.

REFERENCES

- Abma, R., and N. Kabir, 2006, 3D interpolation of irregular data with a POCs algorithm: *Geophysics*, **71**, no. 6, E91–E97, doi: [10.1190/1.2356088](https://doi.org/10.1190/1.2356088).
- Aharon, M., M. Elad, and A. Bruckstein, 2006, An algorithm for designing overcomplete dictionaries for sparse representation: *IEEE Transactions on Signal Processing*, **54**, 4311–4322, doi: [10.1109/TSP.2006.881199](https://doi.org/10.1109/TSP.2006.881199).
- Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein, 2011, Distributed optimization and statistical learning via the alternating direction method of multipliers: *Foundations and Trends in Machine Learning*, **3**, 1–122, doi: [10.1561/22000000016](https://doi.org/10.1561/22000000016).
- Cai, J. F., R. Chan, and Z. Shen, 2008, A framelet-based image inpainting algorithm: *Applied and Computational Harmonic Analysis*, **24**, 131–149, doi: [10.1016/j.acha.2007.10.002](https://doi.org/10.1016/j.acha.2007.10.002).
- Cai, J. F., R. Chan, and Z. Shen, 2010, Simultaneous cartoon and texture inpainting: *Inverse Problems and Imaging*, **4**, 379–395, doi: [10.3934/ipi.2010.4.379](https://doi.org/10.3934/ipi.2010.4.379).
- Cai, J. F., S. Huang, H. Ji, Z. Shen, and G. Ye, 2013, Data-driven tight frame construction and image denoising: *Applied and Computational Harmonic Analysis*, doi: [10.1016/j.acha.2013.10.001](https://doi.org/10.1016/j.acha.2013.10.001).
- Cai, J. F., H. Ji, C. Liu, and Z. Shen, 2009, Blind motion deblurring from a single image using sparse approximation: *Conference on Computer Vision and Pattern Recognition*, IEEE, 104–111.
- Chai, A., and Z. Shen, 2007, Deconvolution: A wavelet frame approach: *Numerische Mathematik*, **106**, 529–587, doi: [10.1007/s00211-007-0075-0](https://doi.org/10.1007/s00211-007-0075-0).
- Chan, R., T. Chan, L. Shen, and Z. Shen, 2003, Wavelet algorithms for high-resolution image reconstruction: *SIAM Journal on Scientific Computing*, **24**, 1408–1432, doi: [10.1137/S1064827500383123](https://doi.org/10.1137/S1064827500383123).
- Chan, R., S. Riemschneider, L. Shen, and Z. Shen, 2004, Tight frame: An efficient way for high-resolution image reconstruction: *Applied and Computational Harmonic Analysis*, **17**, 91–115, doi: [10.1016/j.acha.2004.02.003](https://doi.org/10.1016/j.acha.2004.02.003).
- Crawley, S., R. Clapp, and J. Claerbout, 1999, Interpolation with smoothly nonstationary prediction-error filters: 69th Annual International Meeting, SEG, Expanded Abstracts, 1154–1157.
- Daubechies, I., 1992, Ten lectures on wavelets: CBMS-NSF Lecture Notes, SIAM, nr. 61, 1992.
- Daubechies, I., B. Han, A. Ron, and Z. Shen, 2003, Framelets: MRA-based constructions of wavelet frames: *Applied and Computational Harmonic Analysis*, **14**, 1–46, doi: [10.1016/S1063-5203\(02\)00511-0](https://doi.org/10.1016/S1063-5203(02)00511-0).
- Dong, B., H. Ji, J. Li, Z. Shen, and Y. Xu, 2012, Wavelet frame based blind image inpainting: *Applied and Computational Harmonic Analysis*, **32**, 268–279, doi: [10.1016/j.acha.2011.06.001](https://doi.org/10.1016/j.acha.2011.06.001).
- Dong, B., and Z. Shen, 2010, MRA-based wavelet frames and applications: IAS Lecture Notes Series, Summer Program on The Mathematics of Image Processing: Park City Mathematics Institute.
- Elad, M., and M. Aharon, 2006, Image denoising via sparse and redundant representations over learned dictionaries: *IEEE Transactions on Image Processing*, **15**, 3736–3745, doi: [10.1109/TIP.2006.881969](https://doi.org/10.1109/TIP.2006.881969).
- Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: *Geophysics*, **75**, no. 3, V25–V38, doi: [10.1190/1.3380591](https://doi.org/10.1190/1.3380591).
- Gao, J., M. Sacchi, and X. Chen, 2013a, A fast reduced-rank interpolation method for pre-stack seismic volumes that depend on four spatial dimensions: *Geophysics*, **78**, no. 1, V21–V30, doi: [10.1190/geo2012-0038.1](https://doi.org/10.1190/geo2012-0038.1).
- Gao, J., A. Stanton, M. Naghizadeh, M. Sacchi, and X. Chen, 2013b, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: *Geophysical Prospecting*, **61**, 138–151, doi: [10.1111/j.1365-2478.2012.01103.x](https://doi.org/10.1111/j.1365-2478.2012.01103.x).
- Goldstein, T., and S. Osher, 2009, The split Bregman method for ℓ_1 -regularized problems: *SIAM Journal on Imaging Sciences*, **2**, 323–343, doi: [10.1137/080725891](https://doi.org/10.1137/080725891).
- Gulunay, N., 2003, Seismic trace interpolation in the Fourier transform domain: *Geophysics*, **68**, 355–369, doi: [10.1190/1.1543221](https://doi.org/10.1190/1.1543221).
- Hennenfent, G., and F. Herrmann, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: *Geophysics*, **73**, no. 3, V19–V28, doi: [10.1190/1.2841038](https://doi.org/10.1190/1.2841038).
- Herrmann, F., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: *Geophysical Journal International*, **173**, 233–248, doi: [10.1111/j.1365-246X.2007.03698.x](https://doi.org/10.1111/j.1365-246X.2007.03698.x).
- Kabir, M., and D. J. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: *Geophysical Prospecting*, **43**, 347–368, doi: [10.1111/j.1365-2478.1995.tb00257.x](https://doi.org/10.1111/j.1365-2478.1995.tb00257.x).
- Kaplan, S., M. Sacchi, and T. Ulrych, 2009, Sparse coding for data-driven coherent and incoherent noise attenuation: 79th Annual International Meeting, SEG, Expanded Abstracts, 3327–3331.
- Kreimer, N., and M. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: *Geophysics*, **77**, no. 3, V113–V122, doi: [10.1190/geo2011-0399.1](https://doi.org/10.1190/geo2011-0399.1).
- Kreutz-Delgado, K., J. F. Murray, B. D. Rao, K. Engan, T. W. Lee, and T. J. Sejnowski, 2003, Dictionary learning algorithms for sparse representation: *Neural Computation*, **15**, 349–396, doi: [10.1162/089976603762552951](https://doi.org/10.1162/089976603762552951).
- Lewicki, M., and T. Sejnowski, 2000, Learning overcomplete representations: *Neural Computation*, **12**, 337–365, doi: [10.1162/089976600300015826](https://doi.org/10.1162/089976600300015826).
- Liu, B., and M. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: *Geophysics*, **69**, 1560–1568, doi: [10.1190/1.1836829](https://doi.org/10.1190/1.1836829).
- Liu, Y., and S. Fomel, 2010, OC-seislet: Seislet transform construction with differential offset continuation: *Geophysics*, **75**, no. 6, WB235–WB245, doi: [10.1190/1.3479554](https://doi.org/10.1190/1.3479554).
- Liu, Y., and S. Fomel, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: *Geophysics*, **76**, no. 5, V69–V77, doi: [10.1190/geo2010-0231.1](https://doi.org/10.1190/geo2010-0231.1).
- Ma, J., 2013, Three-dimensional irregular seismic data reconstruction via low-rank matrix completion: *Geophysics*, **78**, no. 5, V181–V192, doi: [10.1190/geo2012-0465.1](https://doi.org/10.1190/geo2012-0465.1).
- Mairal, J., F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, 2009, Non-local sparse models for image restoration: 12th International Conference on Computer Vision, IEEE, 2272–2279.
- Mairal, J., M. Elad, and G. Sapiro, 2008, Sparse representation for color image restoration: *IEEE Transactions on Image Processing*, **17**, 53–69, doi: [10.1109/TIP.2007.911828](https://doi.org/10.1109/TIP.2007.911828).
- Naghizadeh, M., and M. Sacchi, 2007, Multistep autoregressive reconstruction of seismic records: *Geophysics*, **72**, no. 6, V111–V118, doi: [10.1190/1.2771685](https://doi.org/10.1190/1.2771685).
- Naghizadeh, M., and M. Sacchi, 2010, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: *Geophysics*, **75**, no. 6, WB189–WB202, doi: [10.1190/1.3509468](https://doi.org/10.1190/1.3509468).
- Naghizadeh, M., and M. Sacchi, 2013, Multidimensional de-aliased Cadzow reconstruction of seismic records: *Geophysics*, **78**, no. 1, A1–A5, doi: [10.1190/geo2012-0200.1](https://doi.org/10.1190/geo2012-0200.1).
- Olshausen, B., and D. Field, 1997, Sparse coding with an overcomplete basis set: A strategy employed by V1?: *Vision Research*, **37**, 3311–3325, doi: [10.1016/S0042-6989\(97\)00169-7](https://doi.org/10.1016/S0042-6989(97)00169-7).
- Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, no. 3, V25–V32, doi: [10.1190/1.3552706](https://doi.org/10.1190/1.3552706).

- Porsani, M., 1999, Seismic trace interpolation using half-step prediction filters: *Geophysics*, **64**, 1461–1467, doi: [10.1190/1.1444650](https://doi.org/10.1190/1.1444650).
- Ron, A., and Z. Shen, 1997, Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator: *Journal of Functional Analysis*, **148**, 408–447, doi: [10.1006/jfan.1996.3079](https://doi.org/10.1006/jfan.1996.3079).
- Sacchi, M., T. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: *IEEE Transactions on Signal Processing*, **46**, 31–38, doi: [10.1109/78.651165](https://doi.org/10.1109/78.651165).
- Shahidi, R., G. Tang, J. Ma, and F. Herrmann, 2013, Applications of randomized sampling schemes to curvelet-based sparsity-promoting seismic data recovery: *Geophysical Prospecting*, **61**, 973–997, doi: [10.1111/1365-2478.12050](https://doi.org/10.1111/1365-2478.12050).
- Shen, Z., 2010, Wavelet frames and image restorations: in R. Bhatia, P. Arup, G. Rangarajan, M. Srinivas, and M. Vanninathan, eds., *Proceedings of the International Congress of Mathematicians*, ICM, 2834–2863.
- Spitz, S., 1991, Seismic trace interpolation in the f - x domain: *Geophysics*, **56**, 785–794, doi: [10.1190/1.1443096](https://doi.org/10.1190/1.1443096).
- Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: *Geophysics*, **74**, no. 6, V123–V132, doi: [10.1190/1.3245216](https://doi.org/10.1190/1.3245216).
- Trad, D., T. Ulrych, and M. Sacchi, 2002, Accurate interpolation with high-resolution time-variant Radon transforms: *Geophysics*, **67**, 644–656, doi: [10.1190/1.1468626](https://doi.org/10.1190/1.1468626).
- Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation: 80th Annual International Meeting, SEG, Expanded Abstracts, 3829–3833.
- Wang, Y., J. Cao, and C. Yang, 2011, Recovery of seismic wavefields based on compressed sensing by an l_1 -norm constrained trust region method and the piecewise random subsampling: *Geophysical Journal International*, **187**, 199–213, doi: [10.1111/j.1365-246X.2011.05130.x](https://doi.org/10.1111/j.1365-246X.2011.05130.x).
- Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, no. 4, V87–V95, doi: [10.1190/1.1993713](https://doi.org/10.1190/1.1993713).
- Yang, Y., J. Ma, and S. Osher, 2013, Seismic data reconstruction via matrix completion: *Inverse Problems and Imaging*, **7**, 1379–1392, doi: [10.3934/ipi.2013.7.1379](https://doi.org/10.3934/ipi.2013.7.1379).
- Zhang, X., M. Burger, X. Bresson, and S. Osher, 2010, Bregmanized non-local regularization for deconvolution and sparse reconstruction: *SIAM Journal on Imaging Sciences*, **3**, 253–276, doi: [10.1137/090746379](https://doi.org/10.1137/090746379).
- Zhang, X., M. Burger, and S. Osher, 2011, A unified primal-dual algorithm framework based on Bregman iteration: *Journal of Scientific Computing*, **46**, 20–46, doi: [10.1007/s10915-010-9408-8](https://doi.org/10.1007/s10915-010-9408-8).