

Structure-oriented singular value decomposition for random noise attenuation of seismic data^a

^aPublished in Journal of Geophysics and Engineering, 12, 262-272, (2015)

*Shuwei Gan**, *Yangkang Chen†*, *Shaohuan Zu**, *Shan Qu** and *Wei Zhong**

ABSTRACT

Singular value decomposition (SVD) can be used both globally and locally to remove random noise in order to improve the signal-to-noise ratio (SNR) of seismic data. However, they can only be applied to seismic data with simple structure such that there is only one dip component in each processing window. We introduce a novel denoising approach that utilizes a structure-oriented SVD and this approach can enhance seismic reflections with continuous slopes. We create a third dimension for a 2D seismic profile by using the plane-wave prediction operator to predict each trace from its neighbour traces and apply SVD along this dimension. The added dimension is equal to flattening the seismic reflections within a neighbouring window. The third dimension is then averaged to decrease the dimension. We use two synthetic examples with different complexities and one field data example to demonstrate the performance of the proposed structure-oriented SVD. Compared with global and local SVDs, and $f - x$ deconvolution, the structure-oriented SVD can obtain much clearer reflections and preserve more useful energy.

INTRODUCTION

The attenuation of random noise is an important subject in seismic data processing. The enhanced seismic signals with higher signal-to-noise ratio (SNR) can help interpreters to make more accurate decisions. There are generally four different categories of random noise attenuation approaches that exist in the exploration geophysics literatures. The first is based on the predictive property of useful signals in small spatial-temporal windows, such as $f - x$ deconvolution (Canales, 1984; Chen and Ma, 2014), $f - x - y$ prediction filtering (Wang, 1999, 2002), $t - x$ prediction error filtering (Abma and Claerbout, 1995), and $f - x$ non-stationary polynomial fitting (Liu et al., 2011). Yang et al. (2015) proposed a novel trace-by-trace random noise attenuation approach based on the predictable spectral components property of useful seismic reflections using regularized non-stationary autoregression (Fomel, 2013). This approach does not require the spatial coherency assumption and has the potential to be widely used to denoise microseismic signal, the SNR of which is very low. The second is based on the statistical property of seismic profile, such

as the mean filter (Bonar and Sacchi, 2012), median filter (Liu, 2013; Chen, 2014; Chen et al., 2014b). The third is based on extracting the principal components of seismic data, such as multichannel singular spectrum analysis (MSSA) (Oropeza and Sacchi, 2011; Chen et al., 2015) and EMD based approaches (Chen et al., 2014c). This type of approach is also related with those rank-reduction based approaches, eigenimage-based approaches. The fourth is based on a transformed domain thresholding strategy (Neelamani et al., 2008; Fomel and Liu, 2010; Chen et al., 2014a). The transform operator can be a fixed-basis sparsity-promoting transform, and can also be an adaptively-learned dictionary, while the fixed-basis transform enjoys better efficiency and learning-based dictionary enjoys better adaptivity. Apart from those aforementioned one-step noise attenuation, Chen and Fomel (2015) proposed a two-step approach for retrieving the lost useful signals from the noise section, which can be seen as the residual random noise attenuation.

Freire and Ulrych (1988) proposed to carry out rank reduction of seismic images in the $t-x$ domain via singular value decomposition (SVD). They applied SVD to the seismic data matrix and extracted the first singular value in order to remove random noise based on the assumption that the seismic data with only horizontal events have a rank of one. We refer to this method as the global SVD (GSVD), because this SVD method does not require to be implemented in small windows. The GSVD does not require regularly sampled data because no convolutional operator is used. As long as the seismic profiles are composed with horizontal events, GSVD can obtain a good denoising performance. However, the seismic profiles do not necessarily meet the requirement that there are no dipping events. For those complex profiles, GSVD cannot be applied. Freire and Ulrych (1988) also demonstrated that the GSVD is actually equivalent to Karhunen-Loeve (KL) (or principal component transformation) approach (Jones and Levy, 1987). Because of the horizontal-energy selective property of GSVD, Milton et al. (2009) proposed to use GSVD filtering to remove dipping interference, such as ground-rolls.

The $f-x$ domain rank reduction approaches are independent from dip, and therefore, do not require flattening. They require an eigen-decomposition of the spectral matrix of data, which is connected with Cadzow filtering (Cadzow, 1988) and multichannel singular spectrum analysis (MSSA) (Vautard et al., 1992; Oropeza and Sacchi, 2011). Although MSSA and Cadzow filtering are equivalent, they come from different signal analysis subfields. Cadzow filtering was proposed to denoise images, whereas MSSA was proposed to decompose time series arising in the study of dynamical systems (Oropeza and Sacchi, 2010). Cadzow filtering and MSSA are also referred to as the $f-x$ SVD (FXSVD). For FXSVD, the eigen-decomposition of spectral matrix is usually done by applying an SVD to a pre-constructed (block) Hankel matrix. However, this type of methods depends on the low-rank property of seismic data, in other words, the seismic profile should contain few linear events (2D) or few plane-wave components (3D). The low-rank and linear-events assumptions of the FXSVD are not always met, which enforces the FXSVD to be implemented in local processing windows. The division of local windows and decision of rank value, however, are sometimes an empirical and user-unfriendly process.

In this paper, we propose a novel SVD approach for enhancing useful reflections by removing random noise. The novel SVD is implemented obeying the structural information. We first flatten the seismic reflections according to the local slopes by trace prediction. In the flattened domain, we apply a GSVD to horizontal events in order to remove random noise. The flattened traces are then transformed back to the original structural shape by stacking the traces in the flattened domain. We refer to the proposed novel SVD denoising approach as structure-oriented SVD (SOSVD). The SOSVD does not require local processing windows as LSVD and FXSVD do, and it can distinguish signal and noise better than GSVD. Flattened events make the largest singular value components mainly correspond to useful events, which are much easier to be separated. However, GSVD may not be applied to flattened events.

We organize the paper as follows: we first introduce the well-known GSVD, LSVD along with dip steering, and the theoretical aspects of SOSVD. A brief review of the basic theory of the commonly used $f - x$ deconvolution is also provided in the appendix. We then apply three different SVDs and $f - x$ deconvolution onto three different synthetic examples (with increasing complexity) and one field data example that comes from the North Sea, and compare their performances. The comparisons show that the SOSVD can obtain much better performance in removing noise and preserving useful signals.

THEORY

Noise attenuation using SVD

2D seismic data can be expressed as a data matrix \mathbf{D} ($M \times N$), consisting of N traces and M time samples. The SVD of the data matrix \mathbf{D} can be expressed as:

$$\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T. \quad (1)$$

Here, \mathbf{U} is composed of the eigenvectors of $\mathbf{D}\mathbf{D}^T$. \mathbf{V} is composed of the eigenvectors of $\mathbf{D}^T\mathbf{D}$. $\mathbf{\Sigma}$ is a diagonal matrix composed of the decreasing singular values. Let us denote \mathbf{U} , $\mathbf{\Sigma}$, and \mathbf{V} in the following form:

$$\begin{aligned} \mathbf{U} &= [u_1, u_2, \dots, u_r], \\ \mathbf{\Sigma} &= \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), \\ \mathbf{V} &= [v_1, v_2, \dots, v_r], \end{aligned} \quad (2)$$

where r is the rank of \mathbf{D} . The vectors \mathbf{u}_k and \mathbf{v}_k are also called the propagation vectors and the eigen-wavelets, respectively (Vrabie et al., 2004). The singular values σ_k are sorted such that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$. They can be obtained by calculating the positive square roots of the eigenvalues of the data covariance matrix $\mathbf{D}\mathbf{D}^T$.

Equation 1 can be expressed as:

$$\mathbf{D} = \sum_{k=1}^r \lambda_k \mathbf{u}_k \mathbf{v}_k^T, \quad (3)$$

where $\mathbf{u}_k \mathbf{v}^T$ is the rank-one matrix called the k th eigenimage of \mathbf{D} . Thus, from equation 3, the seismic image can be decomposed into r eigenimages, the energy of which corresponds to the value of each element in matrix Σ .

We can remove the random noise in seismic data in order to enhance the seismic reflections by only selecting the first several eigenimages (Freire and Ulrych, 1988):

$$\hat{D}_{svd} = \sum_{k=1}^p \sigma_k \mathbf{u}_k \mathbf{v}_k^T. \quad (4)$$

Random noise attenuation approach by SVD utilizes the property that useful seismic signals are horizontally coherent to separate signal and noise. The decomposition is based on the criterion of horizontal coherency. The energy of coherent signals lays in the first several parts of the decomposition, which is represented by the value of bigger values in Σ . For seismic data, the horizontal coherent part corresponds to useful signal while random noise or coherent dipping events are not horizontally coherent. Thus, by choosing the first several elements in Σ and making the others equal to zero, we can remove the noise and enhance seismic data quality. SVD acts as a data-driven, low-pass filter by rejecting highly uncorrelated traces (Bekara and van der Baan, 2007). Here, we refer to those denoising approaches that simply apply SVD to the original seismic data in order to remove random noise as global SVD (GSVD).

Dip steering and local SVD

Supposing the events in a processing window have one slope, dip steering algorithm aims to flatten the dipping events by shifting each trace such that the flattened events are suitable for applying a GSVD based denoising approach, as shown in equation 4. This flattening strategy is called dip steering (Bekara and van der Baan, 2007). After removing noise, applying an inverse flattening to the data in the processing window can obtain the denoised data. We refer to those denoising approaches that apply a GSVD in the dip-steering flattened domain as local SVD (LSVD). The time shifts for each trace in order to flatten the events are obtained by selecting the time delays that can maximize cross-correlation between each trace and a reference trace.

$$\tau_n = \arg \max_{\tau} \sum_{m=\min(1-\tau,1)}^{\max(M_w, M_w-\tau)} r(m) d(m+\tau, n), \quad (5)$$

where τ_n is the optimal time shift for n th trace in the processing window, M_w denotes the number of time samples in the processing window and $r(m)$ denotes the m th sample of the reference trace and $d(m, n)$ denotes the data value in m th sample and n th trace. $\tau_n > 0$ corresponds to shifting trace downward with respect to the reference trace, and $\tau_n < 0$ corresponds to shifting trace upward.

As can be seen from the equation 5, the selection of reference trace is crucial for the effectiveness of LSVD. It can be chosen as random trace in the processing window if the noise level is not high, or can be chosen as a stacked trace after normal-moveout (NMO) of common-midpoint gathers. Bekara and van der Baan (2007) proposed an iterative strategy for selecting the optimal reference trace stating that the resulting shifted traces are stacked after each cross-correlation pass for updating the new reference traces and the process of cross-correlation, shifting, and stacking is repeated until the process converges. Figure 1 shows a demonstration for dip steering and the processes of LSVD. Figure 1a is the original noisy dipping event. Figure 1b denotes the flattened event after forward dip steering. Figure 1c shows the LSVD denoised result by choosing only one singular value. The final denoised result after inverse dip steering on Figure 1c is shown in Figure 1d. The time shifts after the optimizing equation 5 are shown in Figure 2. The reference trace is simply chosen as the first trace of the original data, as shown in Figure 1a.

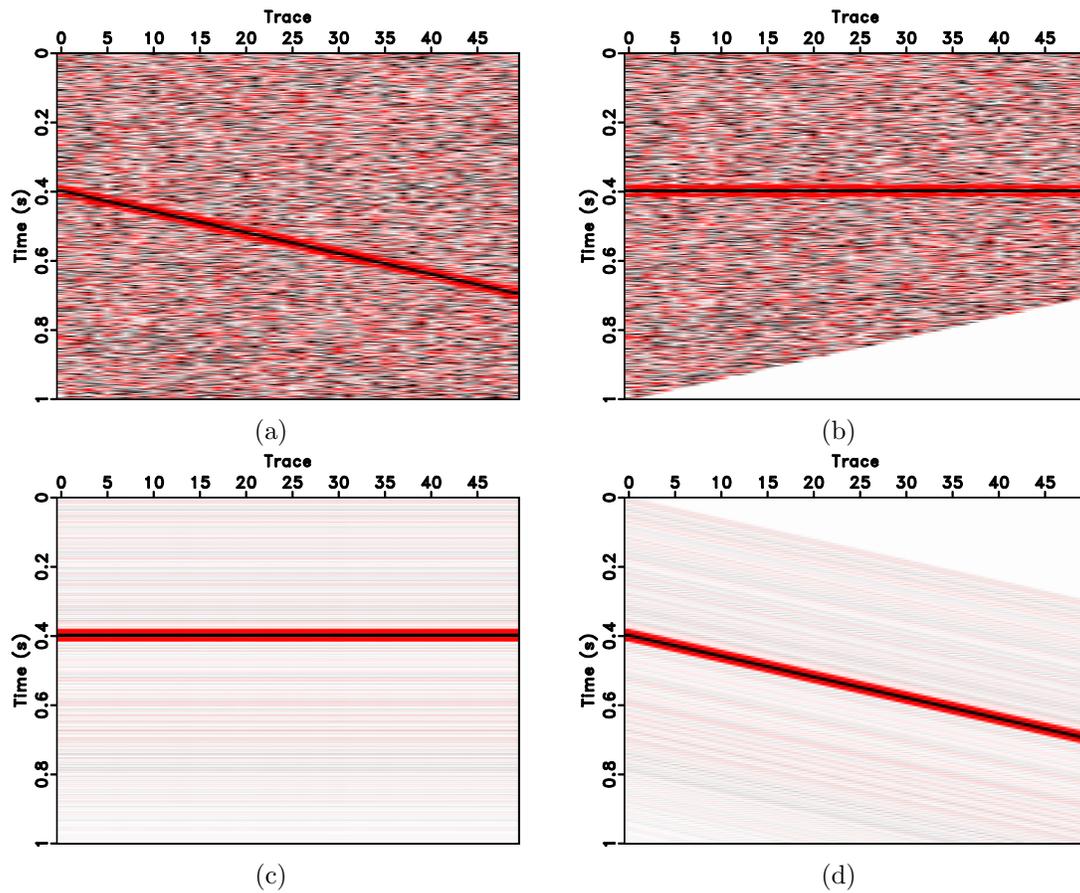


Figure 1: Demonstration for dip steering and LSVD. (a) Noisy dip reflector. (b) Flattened event using dip steering. (c) SVD denoised result by selecting only one eigen-image. (d) Denoised data after inverse dip steering.

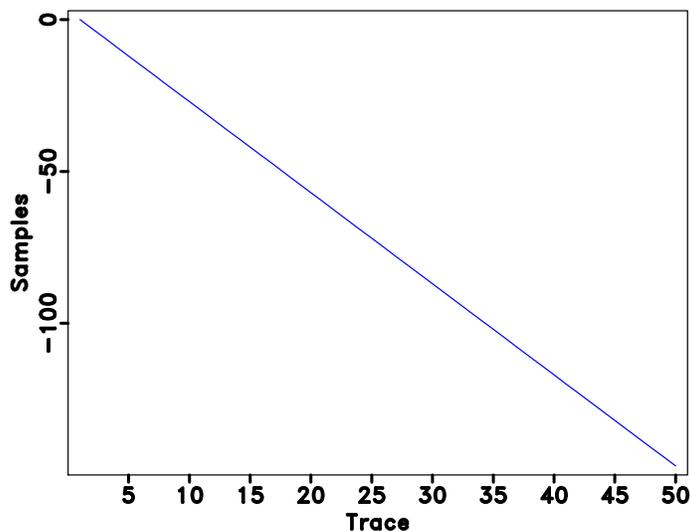


Figure 2: Time shifts for each trace for the forward dip steering as shown in Figure 1 (reference trace selected as the left trace in Figure 1a).

Structure-oriented SVD

The structure-oriented SVD (SOSVD) refers to two processes: flattening along the local structure in a local spatial window and applying GSVD in the flattened local spatial window. The procedures can be summarized as:

$$\mathbf{D} \rightarrow \mathbf{D}_j^R (j = 1, 2, \dots, N) \rightarrow \overline{\mathbf{D}}_j^R \rightarrow \overline{\overline{\mathbf{D}}}_j^R \rightarrow \mathbf{d}_j \rightarrow \hat{\mathbf{D}}_{\text{sosvd}}. \quad (6)$$

Here, \mathbf{D}_j^R denotes the j th spatial window (corresponding to j th trace) with a radius of R , $\overline{\mathbf{D}}_j^R$ denotes the flattened local spatial window, $\overline{\overline{\mathbf{D}}}_j^R$ denotes the SVD denoised local spatial window, \mathbf{d}_j denotes the averaged local spatial window, and $\hat{\mathbf{D}}_{\text{sosvd}}$ denotes the output data using SOSVD.

As we can see from the workflow, the key step that distinguishes SOSVD with other types of SVD approaches is the flattening in the local spatial window. The flattening corresponds to applying a flattening operator to the data (here we use a prediction operator according to local slope) so that the output data have horizontal events:

$$\mathbf{P}_j \mathbf{D}_j^R = \overline{\mathbf{D}}_j^R. \quad (7)$$

where \mathbf{P}_j is the j th flattening operator. Here, the flattening operator is chosen as a plane-wave prediction operator related with the local slope. Equation 7 has the

following detailed form:

$$\begin{aligned}
 & \left[\begin{array}{cccc}
 \mathbf{P}_{(1,j) \rightarrow (1+R,j)}(\sigma_{1,j}) & & & \\
 & \ddots & & \\
 & & \mathbf{P}_{(1+R,j) \rightarrow (1+R,j)}(\sigma_{1+R,j}) & \\
 & & & \ddots \\
 & & & & \mathbf{P}_{(1+2R,j) \rightarrow (1+R,j)}(\sigma_{1+2R,j})
 \end{array} \right] \\
 & [\mathbf{d}_{1,j}, \dots, \mathbf{d}_{1+R,j}, \dots, \mathbf{d}_{1+2R,j}] \\
 & = [\bar{\mathbf{d}}_{1,j}, \dots, \bar{\mathbf{d}}_{1+R,j}, \dots, \bar{\mathbf{d}}_{1+2R,j}].
 \end{aligned} \tag{8}$$

Here, $\mathbf{P}_{(i,j) \rightarrow (k,j)}(\sigma_{i,j})$ denotes the prediction operator from trace i to trace k in j th spatial window, which is connected with the local slope of i th trace. Prediction of a trace consists of shifting the original trace along dominant event slopes (Fomel, 2010). Prediction of a trace from a distant neighbour can be accomplished by simple recursion (Liu et al., 2010), i.e., predicting trace k from trace 1 is simply

$$\mathbf{P}_{(1,j) \rightarrow (k,j)}(\sigma_{1,j}) = \mathbf{P}_{(k-1,j) \rightarrow (k,j)}(\sigma_{k-1,j}) \cdots \mathbf{P}_{(2,j) \rightarrow (3,j)}(\sigma_{2,j}) \mathbf{P}_{(1,j) \rightarrow (2,j)}(\sigma_{1,j}). \tag{9}$$

The prediction operator is a numerical solution of the local plane differential equation

$$\frac{\partial P}{\partial x} + \sigma \frac{\partial P}{\partial t} = 0, \tag{10}$$

for local plane wave propagation in the x direction.

The dominant slopes are estimated by solving the following least-square minimization problem using regularized least-squares optimization:

$$\hat{\sigma} = \arg \min_{\sigma} \| \mathbf{W}(\sigma) \mathbf{D} \|_2^2, \tag{11}$$

where \mathbf{W} is the destruction operator defined as

$$\mathbf{W} = \begin{bmatrix}
 \mathbf{I} & 0 & 0 & \cdots & 0 \\
 -\mathbf{P}_{1 \rightarrow 2} & \mathbf{I} & 0 & \cdots & 0 \\
 0 & -\mathbf{P}_{2 \rightarrow 3} & \mathbf{I} & \cdots & 0 \\
 \cdots & \cdots & \cdots & \cdots & \cdots \\
 0 & 0 & \cdots & -\mathbf{P}_{N-1 \rightarrow N} & \mathbf{I}
 \end{bmatrix},$$

where \mathbf{I} stands for the identity operator, and $\mathbf{P}_{i \rightarrow k}$ describes prediction of trace k from trace i (same as the previous version $\mathbf{P}_{(i,j) \rightarrow (k,j)}(\sigma_{i,j})$ except for not in a specific spatial window). The optimization approach as shown in equation 11 for obtaining local slope estimation is called plane wave destruction (PWD) (Fomel, 2002).

EXAMPLES

In this section, we first use two synthetic examples with different complexities. Then a field data example is presented for better demonstration of the proposed approach. For all the three examples, we compare the denoising performances for GSVD, LSVD, SOSVD, and the commonly used $f - x$ deconvolution. A brief review of the theory of $f - x$ deconvolution is provided in the appendix.

The first synthetic example is a simple seismic profile that has hyperbolic events. It is generated from the SeismicLab library. The clean data and noisy data with simulated Gaussian white noise are shown in Figure 3. There are three hyperbolic events in the synthetic data. Two events have small slopes and the other one has high slope. The denoised results using GSVD, LSVD, and SOSVD are shown in Figures 4a, 4b, and 4d, respectively. As a reference, we also show the denoised result using $f - x$ deconvolution in Figure 4c. From the comparison of denoised results shown in Figure 4, we can initially conclude that the GSVD removes the least noise, there are some damages to the events for LSVD, $f - x$ deconvolution does a little damage to the dipping hyperbolic event, and the denoised result using the proposed SOSVD obtains a nearly perfect result. The removed noise sections are shown in Figure 5. From the noise sections, we can confirm the observation made from the denoised results shown in Figure 4. The GSVD cannot remove too much random noise because we cannot remove too many eigen images, otherwise the damages to useful events are very large. The GSVD causes some damages to the steep event, LSVD causes some damages to areas that have conflicting slopes, $f - x$ deconvolution causes damages to both steep event and more horizontal events. There are nearly no coherent signals lost in the noise section for the proposed SOSVD.

The second synthetic example is a linear-event section. Figure 6 shows the clean and noisy data. There are one horizontal and three dipping events in this section. Two of the dipping events cross with each other. After using the GSVD, LSVD, $f - x$ deconvolution, and the proposed SOSVD, we obtain four denoised sections, as shown in Figure 7. The GSVD still does a bad job because of the dipping events. The LSVD seems to cause more damages to the useful events than the previous example because almost in each processing window there are more than one slopes, which makes the dip steering strategy fail. $f - x$ deconvolution damages both horizontal and dipping events. The proposed SOSVD removes the most noise and preserves the useful energy best. From the noise sections as shown in Figure 8, we conclude that the SOSVD gets an excellent performance except for small damage to the crossing point of the original noisy data. The damage for the crossing point comes from the fact that the PWD algorithm cannot obtain a precise slope estimation in the region and thus causes the flattening along the structure more difficult. In fact, the only disadvantage of the proposed approach is the incapability to handle crossing seismic events. The crossing points will appear in the noise section. The problem can be handled by slope-separated processing using the same approach. However, as a byproduct, the proposed approach may has potential to be used as a diffraction detector, which will be helpful for other important tasks in seismic data processing and imaging. This

topic regarding to the crossing point may be the subject of future investigation.

The third example is a field data example. This example is a part extracted from a 3D North Sea data (Lomask et al., 2006; Fomel, 2010). The original image is shown in Figure 9. As we can see, the seismic data contain a lot of dis-continuous events caused by a low SNR, which makes the interpretation and evaluation of the useful signals impossible. After using GSVD, LSVD, $f - x$ deconvolution and SOSVD, we obtain four denoised sections that are shown in Figure 10. It's obvious that GSVD and LSVD both cause some damages to the useful signals. Although $f - x$ deconvolution preserves the useful reflections well, there seems no improvement for SNR of the data. By using SOSVD, we can obtain a very good denoised image (Figure 10d). The image becomes clean and seismic reflections becomes continuous, which is beneficial for the following interpretation. From the noise sections shown in Figure 11, we can get the similar conclusion that SOSVD preserves the most useful energy while removing the most noise.

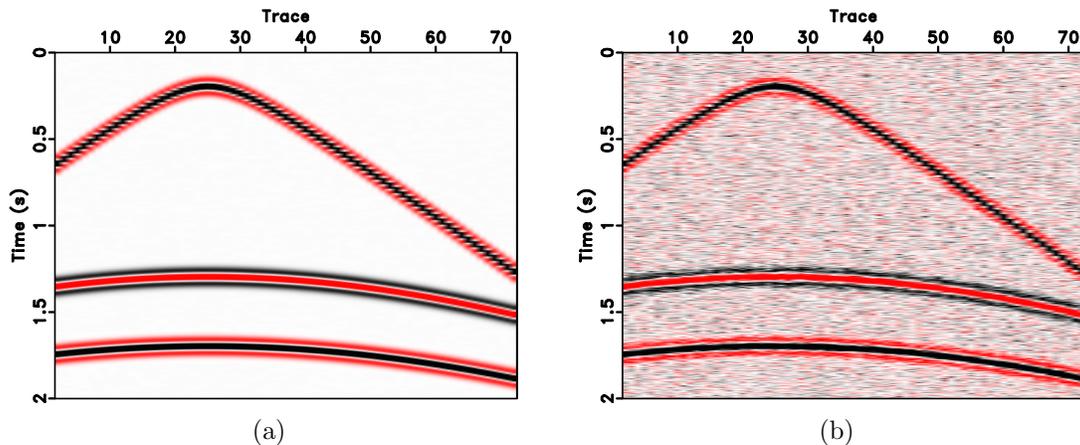


Figure 3: Hyperbolic-events synthetic seismic profile. (a) Clean data. (b) Noisy data.

CONCLUSIONS

We have proposed a novel structure-oriented singular value decomposition (SOSVD) for enhancing useful reflections by removing random noise of seismic data. The SOSVD locally flattens the useful reflections by predicting each trace from its neighbouring traces. The plane-wave destruction (PWD) operator serves as the slope estimator for the prediction operator. In the flattened local processing window, we can apply a traditional SVD to remove incoherent noise. The SVD denoised windows are then stacked to output the denoised trace. Compared with GSVD, the SOSVD can preserve much more dipping events by slope flattening. Compared with LSVD, the SOSVD skips the use of dip steering that depends on a uniform shift for each trace and assumes a uniform-slope processing window, thus is more robust and effective in handling complex structures. Compared with $f - x$ deconvolution, the SOSVD can be

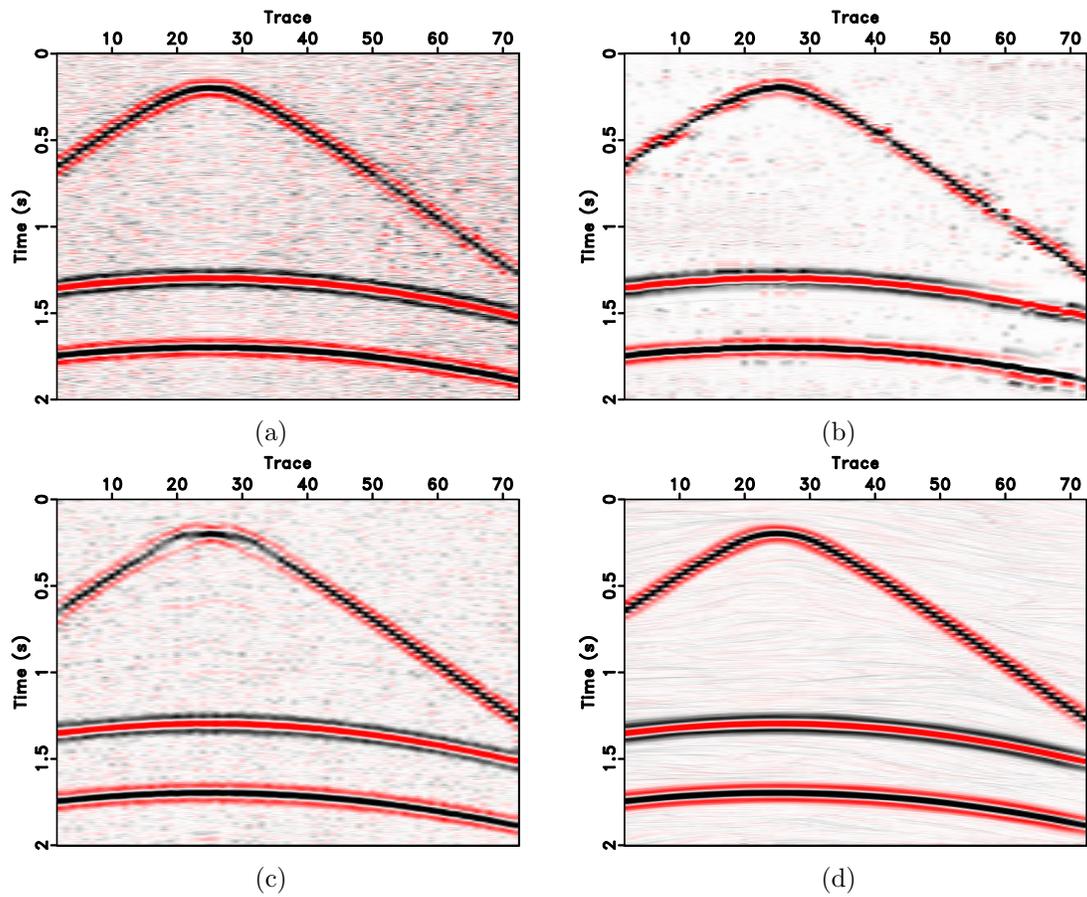


Figure 4: Comparison of denoised data for the hyperbolic-events synthetic example. (a) Using GSVD. (b) Using LSVD. (c) Using $f - x$ deconvolution. (d) Using SOSVD.

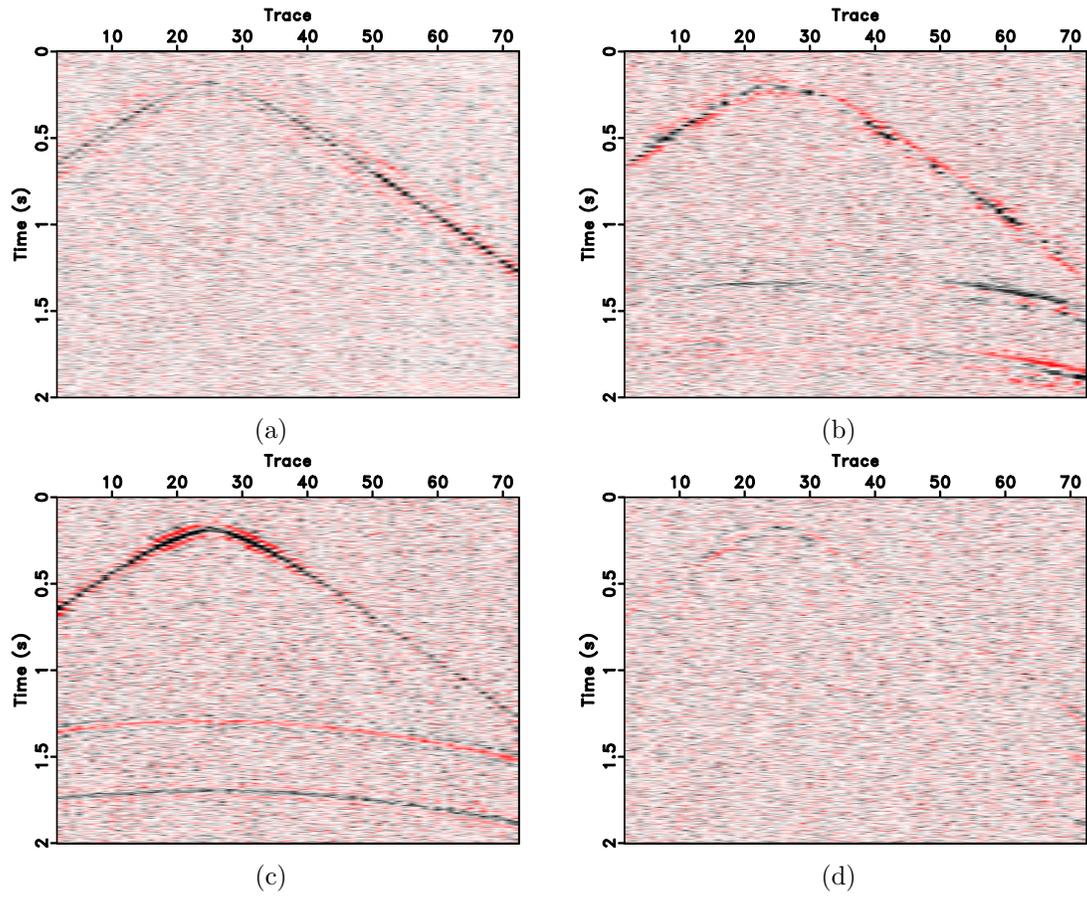


Figure 5: Comparison of removed noise for the hyperbolic-events synthetic example. (a) Using GSVD. (b) Using LSVD. (c) Using $f-x$ deconvolution. (d) Using SOSVD.

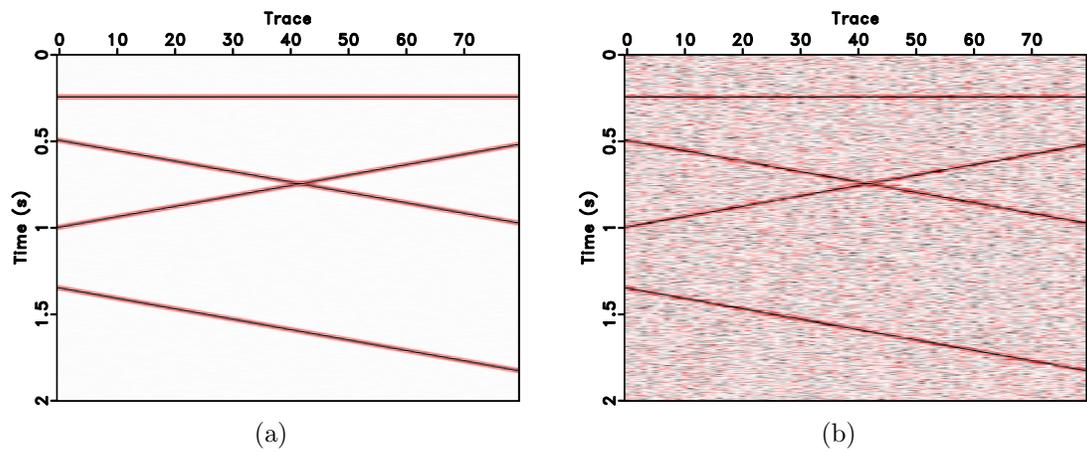


Figure 6: Complex-events synthetic seismic profile. (a) Clean data. (b) Noisy data.

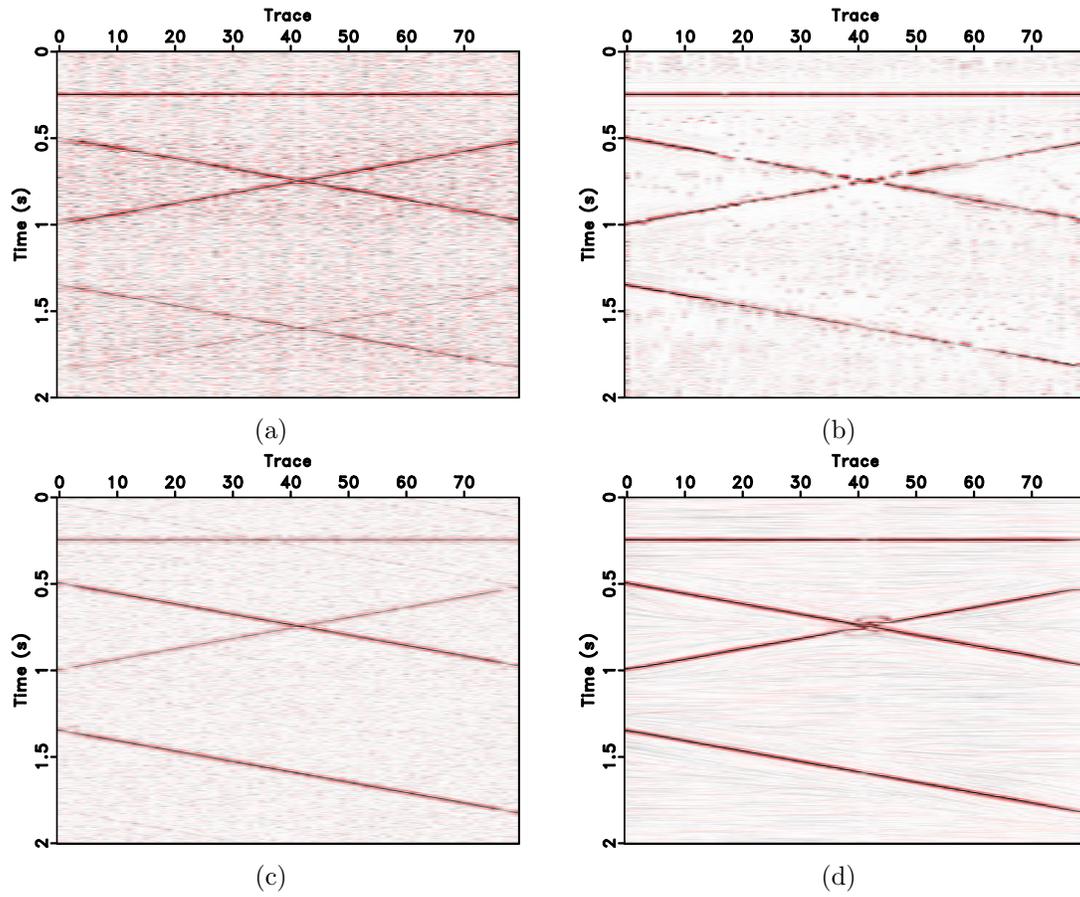


Figure 7: Comparison of denoised data for the complex-events synthetic example. (a) Using GSVD. (b) Using LSVD. (c) Using $f-x$ deconvolution. (d) Using SOSVD.

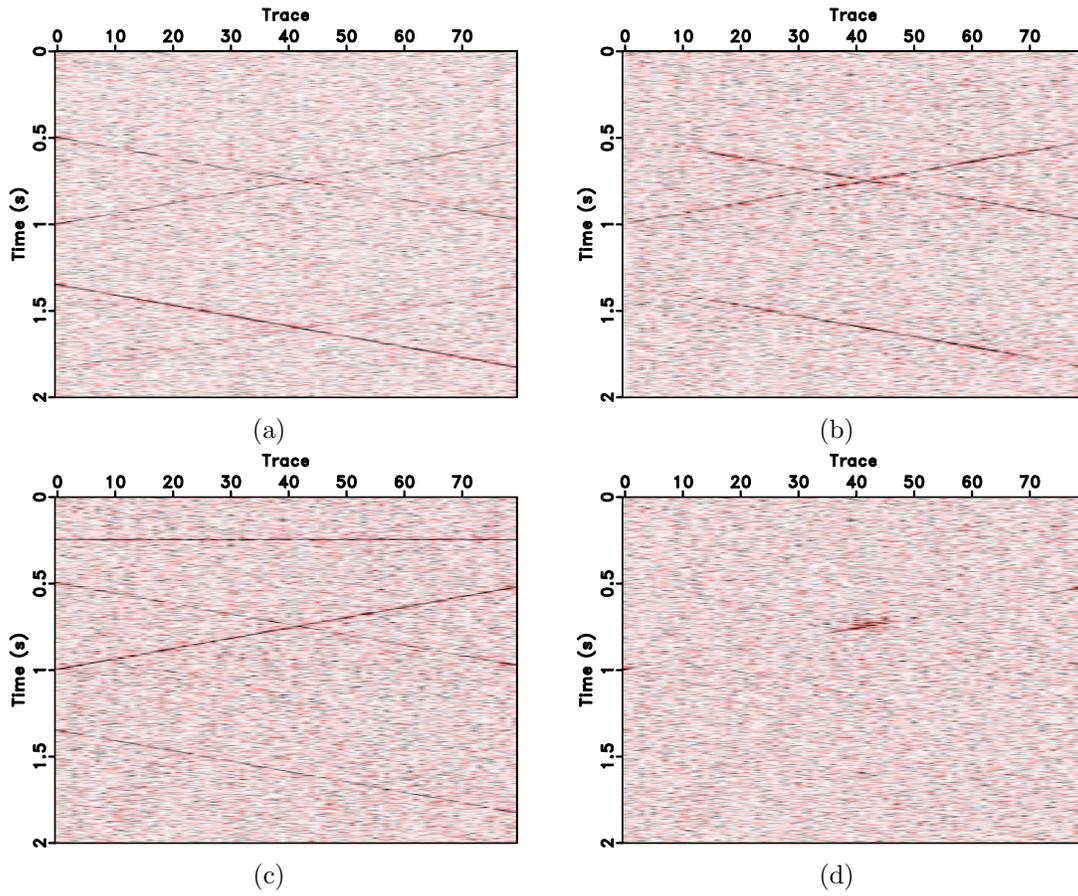


Figure 8: Comparison of removed noise for the complex-events synthetic example. (a) Using GSVD. (b) Using LSVD. (c) Using $f-x$ deconvolution. (d) Using SOSVD.

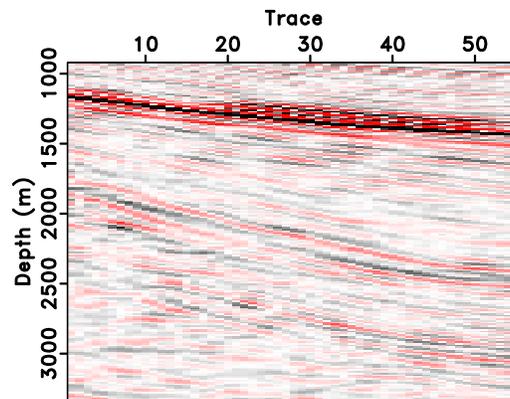


Figure 9: Field data example.

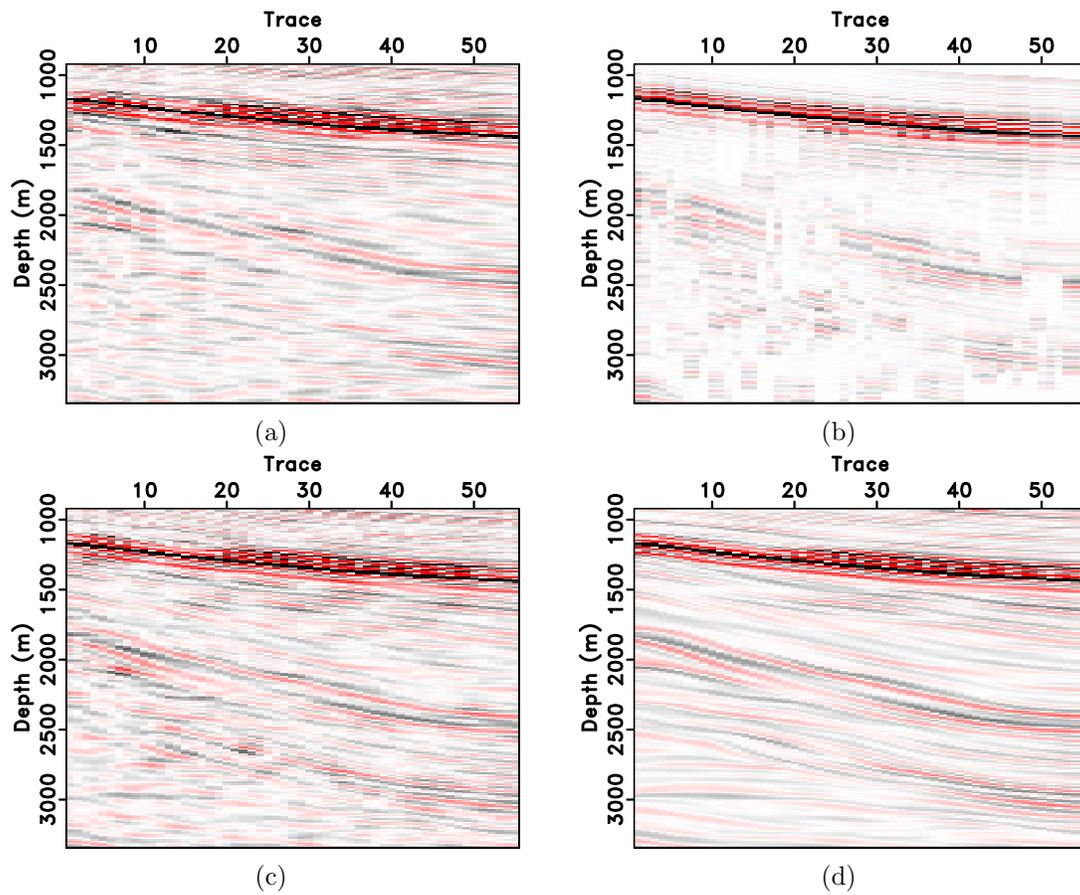


Figure 10: Comparison of denoised data for field data. (a) Using GSVD. (b) Using LSVD. (c) Using $f-x$ deconvolution. (d) Using SOSVD.

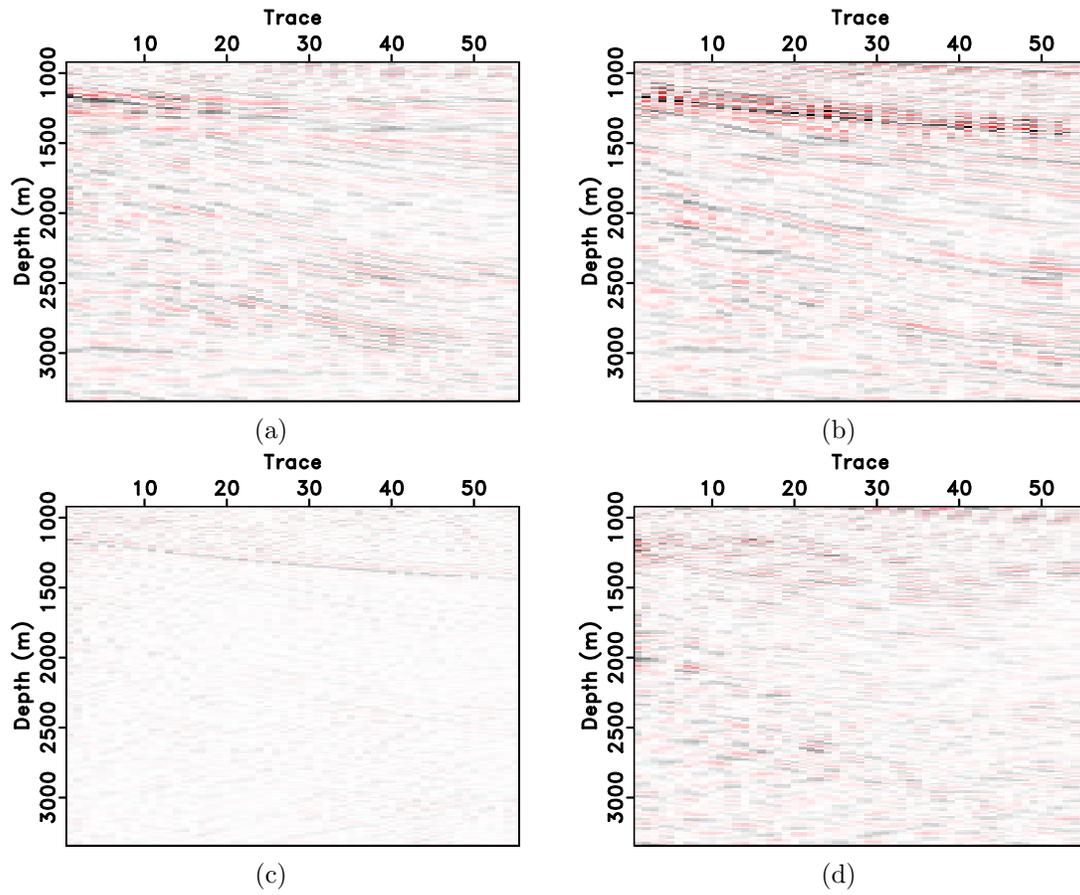


Figure 11: Comparison of removed noise for field data. (a) Using GSVD. (b) Using LSVD. (c) Using $f-x$ deconvolution. (d) Using SOSVD.

significantly more effective both in the removal of random noise and the preservation of useful reflections.

ACKNOWLEDGEMENT

We would like to thank three anonymous reviewers for very constructive suggestions, which help us improve the manuscript greatly.

APPENDIX A: REVIEW OF $F - X$ DECONVOLUTION

$f - x$ deconvolution is one of the most widely used approaches for random noise attenuation. In this appendix, we briefly review the theory of $f - x$ deconvolution. Let $s_t(x)$ be the seismic signal located at trace x and time t . If the slope of a linear event with constant amplitude in a seismic section is ψ , then:

$$s_t(x + 1) = s_{t-x\psi\Delta h}(1), \quad (12)$$

where Δh denotes the spatial interval. Equation 12 can be directly transformed into the frequency domain utilizing the time-shift property of the Fourier transform:

$$S_f(x + 1) = S_f(1)e^{-i2\pi fx\psi\Delta h}. \quad (13)$$

For a specific frequency f_0 , from equation 13 we can obtain a linear recursion, which is given by:

$$S_{f_0}(x + 1) = a_{f_0}(1)S_{f_0}(x), \quad (14)$$

where $a_{f_0}(1) = e^{-i2\pi f_0\psi\Delta h}$. This recursion is also known as an auto-regressive (AR) model of order 1 (Canales, 1984). Similarly, superposition of p linear events in the $t - x$ domain can be represented by an AR model of order p (Tufts and Kumaresan, 1982; Harris and White, 1997) as the following equation:

$$S_{f_0}(x + 1) = a_{f_0}(1)S_{f_0}(x) + a_{f_0}(2)S_{f_0}(x - 1) + \cdots + a_{f_0}(p)S_{f_0}(x + 1 - p), \quad (15)$$

Equation 15 can be formulated as a convolutional form:

$$\mathbf{d} = \mathbf{f} * \mathbf{a}, \quad (16)$$

where \mathbf{d} denotes the vector composed of $S_{f_0}(x + 1)(x = 1, 2, \cdots, X)$, \mathbf{f} denotes the vector composed of $S_{f_0}(x)(x = 1, 2, \cdots, X)$, \mathbf{a} denotes the vector composed of $a_{f_0}(x)(x = 1, 2, \cdots, X)$, and X denotes the number of traces.

Equation 16 can be formulated as a matrix vector form:

$$\mathbf{d} = \mathbf{F}\mathbf{m}, \quad (17)$$

where \mathbf{F} is the convolution matrix composed by \mathbf{f} . Suppose $p = 4$, the detailed form of equation 17 can be expressed as:

$$\begin{bmatrix} f_2 \\ f_3 \\ \vdots \\ f_X \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} f_1 & 0 & 0 & 0 \\ f_2 & f_1 & 0 & 0 \\ \vdots & f_2 & f_1 & 0 \\ f_{X-1} & \vdots & f_2 & f_1 \\ 0 & f_{X-1} & \vdots & f_2 \\ 0 & 0 & f_{X-1} & \vdots \\ 0 & 0 & 0 & f_{X-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}. \quad (18)$$

However, equation 17 is based on clean signal model. In reality, the seismic data is composed of random noise. Thus, we have to solve \mathbf{a} from the noise corrupted observation \mathbf{d} based on some optimization schemes. Based on equation 15, we can formulate an optimization problem based on the minimum prediction error energy assumption. The predictive error filter can be solved by minimizing the following objective function:

$$J = \|\mathbf{F}\mathbf{m} - \mathbf{d}\|_2^2, \quad (19)$$

where $\|\cdot\|_2^2$ denotes the squares of L_2 norm.

Taking derivatives of the cost function 19 with respect to \mathbf{m} , and setting the result to zero, we can obtain the following equation:

$$\mathbf{F}^T \mathbf{d} = \mathbf{F}^T \mathbf{F} \mathbf{m}, \quad (20)$$

where $[\cdot]^T$ denotes transpose. Note that $\mathbf{F}^T \mathbf{F}$ is a Toeplitz form and thus can be efficiently solved using Levinson's recursion. In order to stabilize the recursion for solving $\hat{\mathbf{a}}$, we need to add a small perturbation to the diagonal of the Toeplitz matrix:

$$\hat{\mathbf{a}} = (\mathbf{F}^T \mathbf{F} + \mu \mathbf{I})^{-1} \mathbf{F}^T \mathbf{d}. \quad (21)$$

Finally, the estimated clean data (denoised data) can be expressed as:

$$\hat{\mathbf{d}} = \mathbf{F} \hat{\mathbf{a}}. \quad (22)$$

It is worth to be mentioned that, the $f - x$ deconvolution approach introduced here only applies a forward AR model to estimate the signal. A hybrid forward and backward version of AR model was proposed in Wang (1999).

REFERENCES

- Abma, R., and J. Claerbout, 1995, Lateral prediction for noise attenuation by $t - x$ and $f - x$ techniques: *Geophysics*, **60**, 1887–1896.
- Bekara, M., and M. van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data: *Geophysics*, **72**, V59–V65.

- Bonar, D., and M. Sacchi, 2012, Denoising seismic data using the nonlocal means algorithm: *Geophysics*, **77**, A5–A8.
- Cadzow, J. A., 1988, Signal enhancement - a composite property mapping algorithm: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **1**, 49–62.
- Canales, L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstracts, 525–527.
- Chen, Y., 2014, Deblending using a space-varying median filter: *Exploration Geophysics*, doi:<http://dx.doi.org/10.1071/EG14051>.
- Chen, Y., and S. Fomel, 2015, Random noise attenuation using local signal-and-noise orthogonalization: *Geophysics*, in print.
- Chen, Y., S. Fomel, and J. Hu, 2014a, Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization: *Geophysics*, **79**, V179–V189.
- Chen, Y., S. Gan, T. Liu, J. Yuan, Y. Zhang, and Z. Jin, 2015, Random noise attenuation by a selective hybrid approach using f-x empirical mode decomposition: *Journal of Geophysics and Engineering*, **12**, 12–25.
- Chen, Y., and J. Ma, 2014, Random noise attenuation by *f-x* empirical mode decomposition predictive filtering: *Geophysics*, **79**, V81–V91.
- Chen, Y., J. Yuan, Z. Jin, K. Chen, and L. Zhang, 2014b, Deblending using normal moveout and median filtering in common-midpoint gathers: *Journal of Geophysics and Engineering*, **11**, 045012.
- Chen, Y., C. Zhou, J. Yuan, and Z. Jin, 2014c, Application of empirical mode decomposition to random noise attenuation of seismic data: *Journal of Seismic Exploration*, **23**, 481–495.
- Fomel, S., 2002, Application of plane-wave destruction filters: *Geophysics*, **67**, 1946–1960.
- , 2010, Predictive painting of 3-d seismic volumes: *Geophysics*, **75**, A25–A30.
- , 2013, Seismic data decomposition into spectral components using regularized nonstationary autoregression: *Geophysics*, **78**, O69–O76.
- Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: *Geophysics*, **75**, V25–V38.
- Freire, S. L. M., and T. J. Ulrych, 1988, Application of singular value decomposition to vertical seismic profiling: *Geophysics*, **53**, 778–785.
- Harris, P., and R. White, 1997, Improving the performance of f-x prediction at low signal-to-noise ratios: *Geophysical Prospecting*, **45**, 269–302.
- Jones, I. F., and S. Levy, 1987, Signal-to-noise ratio enhancement in multichannel seismic data via the karhunen-loeve transform: *Geophysical Prospecting*, **35**, 12–32.
- Liu, G., X. Chen, J. Du, and J. Song, 2011, Seismic noise attenuation using nonstationary polynomial fitting: *Applied Geophysics*, **8**, 18–26.
- Liu, Y., 2013, Noise reduction by vector median filtering: *Geophysics*, **78**, V79–V87.
- Liu, Y., S. Fomel, and G. Liu, 2010, Nonlinear structure-enhancing filtering using plane-wave prediction: *Geophysical Prospecting*, **58**, 415–427.
- Lomask, J., A. Guitton, S. Fomel, J. Claerbout, and A. A. Valenciano, 2006, Flattening without picking: *Geophysics*, **71**, P13–P20.

- Milton, P., M. G. Silva, E. M. Melo, and B. Ursin, 2009, Ground-roll attenuation based on svd filtering: 79th Annual International Meeting, SEG, Expanded Abstracts, 3381–3385.
- Neelamani, R., A. Baumstein, D. Gillard, M. Hadidi, and W. Soroka, 2008, Coherent and random noise attenuation using the curvelet transform: *The Leading Edge*, **27**, 240–248.
- Oropeza, V., and M. Sacchi, 2010, A randomized svd for multichannel singular spectrum analysis (mssa) noise attenuation: 80th SEG expanded abstracts:83rd Annual international meeting, 3539–3544.
- , 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, V25–V32.
- Tufts, D., and R. Kumaresan, 1982, Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood: *Proceedings of the IEEE*, **70**, 975–989.
- Vautard, R., P. Yiou, and M. Ghil, 1992, Singular-spectrum analysis: A toolkit for short, noisy chaotic signals: *Physica D: Nonlinear Phenomena*, **1**.
- Vrabie, V. D., J. L. Mars, and T. L. Lacoume, 2004, Modified svd by means of ica: *Signal Processing*, **84**, 645–652.
- Wang, Y., 1999, Random noise attenuation using forward-backward linear prediction: *Journal of Seismic Exploration*, **8**, 133–142.
- , 2002, Seismic trace interpolation in the f-x-y domain: *Geophysics*, **67**, 1232–1239.
- Yang, W., R. Wang, Y. Chen, J. Wu, S. Qu, J. Yuan, and S. Gan, 2015, Application of spectral decomposition using regularized non-stationary autoregression to random noise attenuation: *Journal of Geophysics and Engineering*, **12**, 175–187.