An Adaptive-Rank Singular Spectrum Analysis for Simultaneous-Source Data Separation

Yaru Xue, Libo Niu, Chong Chen, and Xin Xu

***Abstract*— Simultaneous-source exploration improves efficiency and reduces the cost when acquiring seismic data. However, the adjacent shot records interfere with each other, and an efficient deblending way is needed. The traditional truncated singular spectrum analysis (SSA) algorithm is employed in the local window to predict coherent events. After all the local events are predicted, the whole dither noise could be estimated completely. Traditional processing in the time domain complicates deblending. In this letter, a global-frequency SSA is proposed to predict dither noise with a simple iteration scheme. This method will lead to an increase in the rank in the Hankel matrix. Thus, a trigonometric function is introduced to adaptively determine the rank instead of the rank-truncated method. The experiments on actual seismic data show that the proposed method not only improves the deblending performance but also enjoys high efficiency.**

***Index Terms*— Adaptive rank-reduction (RR), simultaneous-source separation, singular spectrum analysis (SSA), trigonometric function.**

I. Introduction

IN SEISMIC exploration, the time interval of source excitation is usually set to long enough to prevent crosstalk from adjacent seismic sources, which results in low acquisition efficiency, especially in marine exploration. The simultaneous-source seismic exploration method permits records from different sources to overlap in the time domain so that the acquisition efficiency can be significantly improved [1]–[5]. However, it is necessary to separate the simultaneous-source record that is blended in the time domain for the subsequently traditional process.

Simultaneous-source separation is generally posed as an inversion problem to estimate the coherent signals and then subtract dither noise from the blended gathers. Because of the ill-posed nature of the blending problem, a regularization term is often introduced in the coherent events' estimation procedure. Sparsity promotion and low-rank promotion are the currently often-used regularization terms. Mahdad *et al.* [6] introduced the $f-k$ filter to regularize the coherent events. Zu *et al.* [7] proposed a coherency-pass shaping operator to separate simultaneous source data, but it may leave residual noise when there is strong blending interference. Chen [8] used the seislet-domain shaping regularization to map the coherent events to the more admissible model. Gan *et al.* [9] used seislet frames with two corresponding local dips to sparsify each signal component. An amplitude-preserving high-order Radon transform was incorporated with a regularization inversion method to achieve AVO-preserving deblending performance [10].

No matter for the seislet transform or other kinds, these sparsity promotions are based on fixed basis functions, and their sparsity is depended on the similarity between signal and basis functions. Based on the linear event predictability, a low-rank property is demonstrated in the singular spectrum. Its basis functions are driven by the data, which is more conducive to data sparsity. Singular spectrum analysis (SSA) has been widely used in denoising and data reconstruction [11], [12]. Cheng and Sacchi [13] introduced SSA to separate the simultaneous-source data in the local window. The events in the small local window could be regarded as linear with low-rank properties. In the Hankel matrix, the rank is equal to the number of events, which is difficult to determine in field data. Cheng and Sacchi [13] calculated the initial rank and step size through many simulations. Similar rank-reduction (RR) strategy deployed in the data reconstruction using SSA [14]. A simple rank increasing (RI) was proposed by Xue *et al.* [15], which sets the initial reconstructed rank to 1 and increases the rank step by step with iterations. This algorithm is simple but converges slowly.

The local scheme is another strategy in the SSA algorithm. In the small local $t-x$ window, not only coherent events but also blending noise from another window are included. Therefore, it is necessary to estimate all the window coherent data before noise prediction, which brings algorithm complexity. In this letter, we propose to divide the window only in the spatial domain to simplify the dither noise prediction-subtraction scheme. A trigonometric function is also introduced to adaptively estimate the Hankel rank.

II. Method

A. Simultaneous-Source Acquisition Model

Here, the blending model is reviewed in brief. Taking two-sources acquisition as example, the two sources are fired alternatively and pseudosynchronously. The observed blending data D^{obs} with two shots D_1 and D_2 are related by the following expression:

$$D^{\text{obs}} = D_1 + D_2\Gamma \tag{1}$$

Manuscript received January 4, 2020; revised March 26, 2020; accepted April 18, 2020. *(Corresponding author: Libo Niu.)*

The authors are with the Beijing State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing, China (e-mail: 15624977344@163.com).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2020.2989750

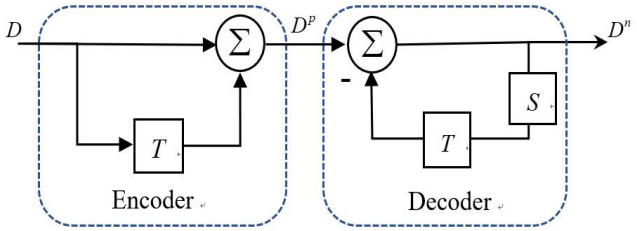

Fig. 1. Deblending encoder and decoder diagram.

where Γ is the relative delay matrix

$$\Gamma = \text{diag}(1, 1, 1 \ldots, 1)_N \times \exp(-i\omega\delta t_n) \tag{2}$$

where ω is the angular frequency, and δt_n is the relative delay time of the nth shot. The coherent D_1 is superposed by the dithered gather D_2.

Introducing the adjoint dithering code Γ^H, we obtained the pseudodeblended data

$$D^p = \begin{bmatrix} D^{\text{obs}} \\ D^{\text{obs}}\Gamma^H \end{bmatrix} = \begin{bmatrix} I & \Gamma \\ \Gamma^H & I \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = D + DT \tag{3}$$

where

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

is the expected deblended data and

$$T = \begin{bmatrix} 0 & \Gamma \\ \Gamma^H & 0 \end{bmatrix}.$$

Most of deblending processing are carried out in the pseudodeblending domain. From formula (3), an easy iterative deblending scheme can be achieved. The shaping operator S are embedded to attenuate the crosstalk noise

$$D^{n+1} = D^p - T[S(D^n)]. \tag{4}$$

To clearly explain the relationship between pseudodeblending and deblending data, the pseudodeblending is defined as an encoder and deblending as a decoder, and their relations are plotted in Fig. 1. The decoder is the inverse of encoder. Without shaping operator, the decoder will converge very slowly because spectrum radium of T is close to 1. The shaping operator is introduced to speed up the deblending.

B. Deblending Scheme: Rank-Reduction in the Frequency Domain

The SSA operator has been used in denoising, data reconstruction. Its detail algorithm refer to [13]. Here, we talk about the deblending scheme.

Generally, SSA is used to estimate the coherent events; then, the blending noises predicted with dither code Γ are subtracted from observed data, and this procedure is called the prediction-subtraction scheme. To benefit from the linear prediction of events and low-rank property, the seismic profile is divided into many local windows in the $t - x$ domain, as shown in Fig. 2(a). The formula (2) tells us that the coherent dither noise

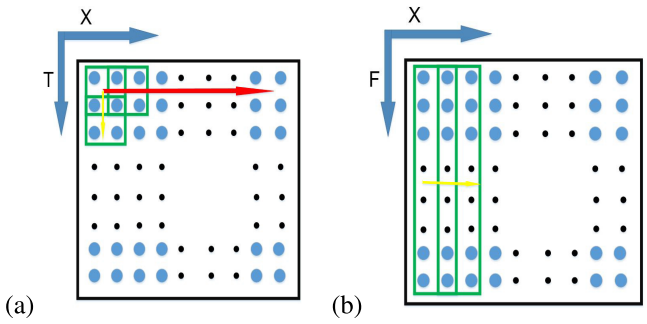

Fig. 2. Local window scheme. (a) Local window strategy for the RR method. (b) Local window strategy for the proposed method.

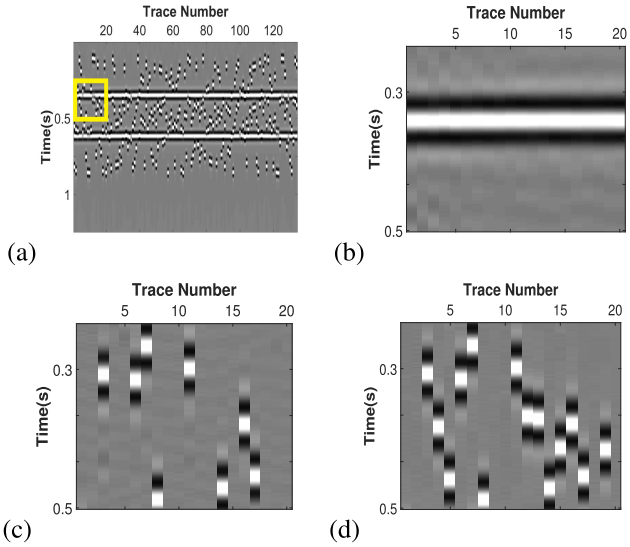

Fig. 3. Disadvantage of local $t - x$ window prediction. (a) Blended data. (b) Data reconstructed from the yellow box. (c) Noise predicted by reconstructing data. (d) Noise in yellow box.

is up or down along the trace, which means that there are two seismic reflection records in each trace. Thus, in a local $t - x$ window, besides the coherent events and their corresponding dithered noise, the noises from the adjoint window are also included.

Fig. 3 explains this phenomenon clearly. Fig. 3(a) simulates a two-layer gather. The coherent event in the green box is estimated using SSA shown in Fig. 3(b) and its corresponding dither noise in Fig. 3(c). Compared with the whole dither noise of green box zoomed in Fig. 3(d), it only includes the first event dither noise. The result of dither noise distributed in a multilocal window brings that the prediction-subtraction only carried out until whole local profiles are processed. The abovementioned deblending pseudocode is listed in Algorithm 1. In this algorithm, coherent even estimations are executed in the frequency domain, and the dither noise subtractions are in the time domain. This transform between the time and frequency domains brings complexity in the iteration.

To overcome the complexity of prediction in all the local $t-x$ windows, a novel local strategy is proposed in the $f - x$ domain, and localizations are carried out only in the spatial domain, as shown in Fig. 2(b). In each window, the dither noise DT and its corresponding coherent data D are all

Algorithm 1

Inputs: The blended data D^{obs}, dither code Γ and error threshold ε

Initialize: The pseudodeblending data :$D^p = \begin{bmatrix} D^{obs} \\ D^{obs}\Gamma^H \end{bmatrix}$

divide the D^p into a set of local window data $\left\{D^i_\omega, i = 1, 2, \ldots, n\right\}$

Prediction and subtraction iteration:

1. For each local window data D^i_ω in the time domain
   - *A* Transform them to frequency domain
   - *B* For each frequency, execute SSA algorithm.
   - *C* Transform it to time domain and get the estimation $\hat{D}^i_\omega$ of D^i_ω

   } Many FFT and its inverse are involved here
2. Patch the $\hat{D}^i_\omega$ into a whole profile and get the current coherent estimation $\hat{D}$
3. Transform $\hat{D}$ into dither noise with operator T and subtract from D^p
4. If $\left\|\hat{D} + \hat{D}T - D^p\right\| \le \varepsilon$, end, otherwise return 1

} Once iteration

included, no other window crosstalk introduced. The coherent prediction and dither noise subtraction can be all accomplished in this window and avoid the transform between the time and frequency domains. The deblending scheme becomes as simple as in Algorithm 2. It is clear that this scheme is more efficient than Algorithm 1.

C. Adaptive Rank Determination Rule

Expanding windows will increase the amount of valid information contained in the window [16]. The abovementioned proposed method will lead to a high rank because all the events are included in the local–global-frequency window. This increases the difficulty of rank determination. A new trigonometric method is proposed here to adaptively solve this rank, and its diagram is shown in Fig. 4. First, the largest and smallest singular value points λ_1 and λ_N are fixed in the graph; another singular value point λ_k is then picked to form a triangle with an angle α. To compute the angle α, the distance of λ_1 to λ_k is computed by

$$d_{1k} = \sqrt{(\lambda_1 - \lambda_k)^2 + (k-1)^2}. \tag{5}$$

Similarly, the distances d_{1N} between λ_1 and λ_N and d_{kN} between λ_k and λ_N can be computed. Then, the angle α is calculated using a trigonometric function easily.

Seismic signals show low rank in the singular spectrum because of events' strong correlation [17]; their singular values weight more on the first few ones and converge quickly like the region a in Fig. 4. However, the noise singular values distribute smoothly like the region b. The demarcation of the region a and b signifies the rank of coherent signals. From the demonstration of Fig. 4, α reaches its minimum value in the demarcation.

In each iteration of this process, the corresponding rank is determined automatically and adaptively changed with a signal-to-noise ratio (SNR), which can greatly speed up the convergence. Applied to all frequencies, a frequency-dependent rank scheme is acquired.

III. Data Experiments

Marine simultaneous-source exploration has broad prospects. We tested the proposed algorithm with real

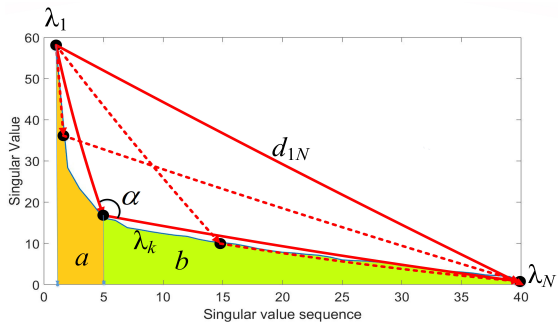

Fig. 4. Trigonometric function estimated rank.

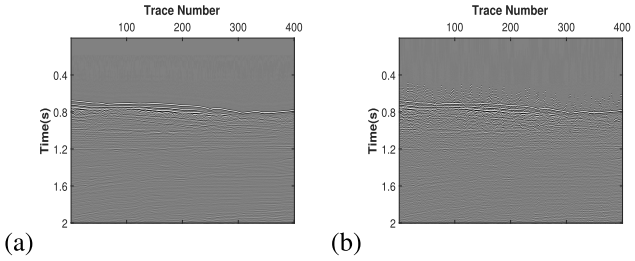

Fig. 5. Synthetic model. (a) Original unblended data. (b) Blended data.

marine data. For towed-streamer exploration, the common offset gathers (COGs) have the same size data, and all the experiments are carried out in the COG.

The original section consists of 400 traces with 1000 samples shown in Fig. 5(a) and the blended data in Fig. 5(b). In the experiment, the local window width is set to 100 traces and the step size to 50 traces, which means that each trace is processed twice and their average is the last results. The RI method is employed for comparison with the time width set to 200 examples.

The deblended results of the proposed method are shown in Fig. 6(a) and the RI method in Fig. 6(b). Their corresponding zoomed-in images are shown in Fig. 6(c) and (d), respectively. It is clear that the dither noise is greatly attenuated in the proposed method, while the noise is still clear in the RI results.

To evaluate the performance of deblending, the SNR is defined as follows:

$$\mathrm{SNR} = 10\log_{10}\frac{\|D\|_2^2}{\|D - \hat{D}\|_2^2} \tag{6}$$

Algorithm 2

Inputs: The blended data D^{obs}, dither code Γ and error threshold ε

Initialize: The pseudodeblending data :$D^p = \begin{bmatrix} D^{obs} \\ D^{obs}\Gamma^H \end{bmatrix}$, transform the D^p into frequency D_f

Divide the D_f into a set of local window data $\left\{D^i_{f\omega}, i = 1, 2, \ldots m\right\}$, here m is much less than window number n in the algorithm 1

Prediction and subtraction iteration:

1. For each local window data $D^i_{f\omega}$

A For each frequency, execute SSA algorithm and get the estimation $\hat{D}^i_{f\omega}$

B Transform $\hat{D}^i_{f\omega}$ into dither noise with operator T, subtract from D^i_f and get the current deblended data $\hat{D}_f$

} Once iteration } Only one time FFT and its inverse

2. if $\left\|\hat{D}_f + \hat{D}_f T - D_f\right\| \le \varepsilon$, go to 3, otherwise return 1

3. Transform $\hat{D}_f$ into time domain and get the deblending data $\hat{D}$

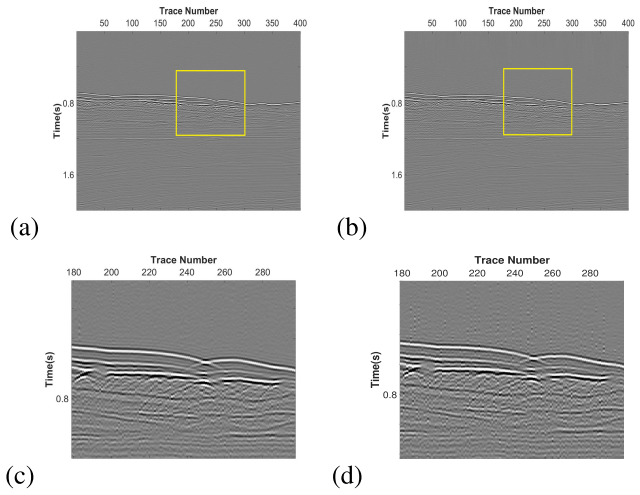

Fig. 6. Deblended comparison for the real data in COG. (a) Deblended data by the proposed method. (b) Deblended data by the RI method. (c) Zoomed-in image in the yellow box of the proposed method. (d) Zoomed-in image in the yellow box of the RI method.

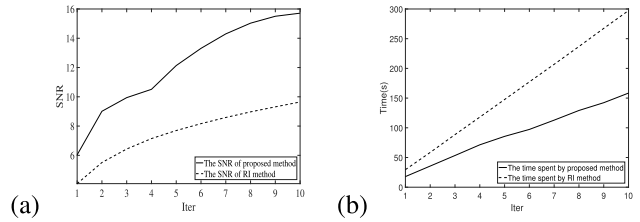

Fig. 7. Algorithm performance evaluation. (a) SNR with the number of iterations. (b) Time required for the number of iterations.

where D stands for the original data and $\hat{D}$ stands for the deblended data. The SNR of proposed and RI method changes with iteration is shown in Fig. 7(a). The proposed method converges after ten iterations. For the RI method, it converges slowly because of the unit rank step size. It is until the almost 35 iterations to reach the same SNR of the proposed method.

The computation time is also compared. The machine CPU for this test is i7-8750H, the main frequency is 2.2 GHz, and the software is MATLAB 2016a. The computation time with iterations is plotted in Fig. 7(b). While the SNR increases, the time required for the proposed method increases slowly. The higher the number of iterations, the more the time is saved. The time saving is mainly from two procedures. First, the local window number of this method is far less than the $t-x$ method. In this experiment, there are eight local windows in the proposed method and 80 in the RI method. The second way to save time is that fewer FFT computations are needed, only once the Fourier transform in the proposed method. Although the demarcation α is introduced, its computation is very easy.

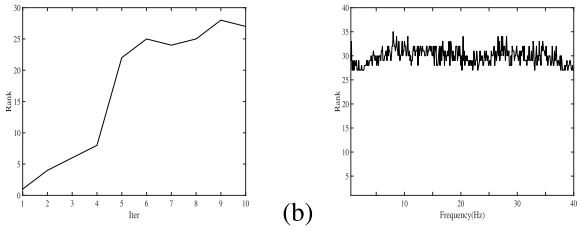

Fig. 8. Analysis of the rank of proposed method. (a) Rank selected by the trace at 20 Hz with different iterations. (b) Rank selected by the trace with different frequencies in the last iteration.

The rank adaptive variation at 20 Hz is also analyzed. In Fig. 8(a), the rank is gradually increasing with iteration, which is consistent with SNR increasing. In early iterations, it is shown that the first few ranks are small and enlarge slowly; because the dither noise is strong and distributed on almost every eigenvalue, a small rank is expected. As the iteration goes, the SNR increases, noise weakens, and rank enlarges quickly. This rank-determination rule accelerates the convergence, and the expected results can be obtained in fewer iterations.

In Fig. 8(b), the rank in the last iteration of each frequency is shown, and the ranks are different for different frequencies, so the proposed adaptive rank determination scheme is frequency-dependent. There is no explicit relationship between rank and frequency because of the random dithering code.

Furthermore, the spectrums of processed data are showed. The proposed method spectrum is plotted in Fig. 9(a) with a red line and the RI method with a blue line. The real data spectrum and blended data spectrum are also plotted in green and black lines, respectively. For a clear comparison, the relative errors are demonstrated in Fig. 9(b). The proposed

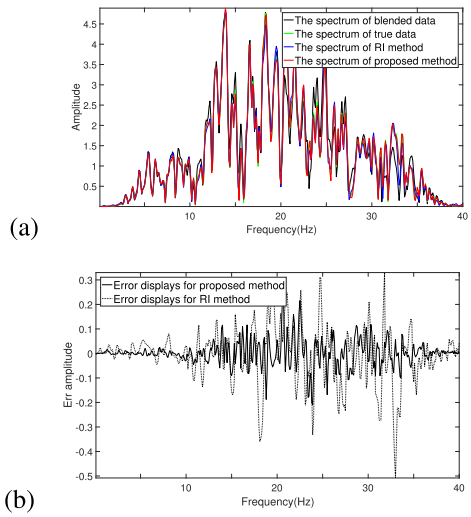

Fig. 9. Evaluating the results of spectrum convergence. (a) Spectrum of data. (b) Errors compared with clean data.

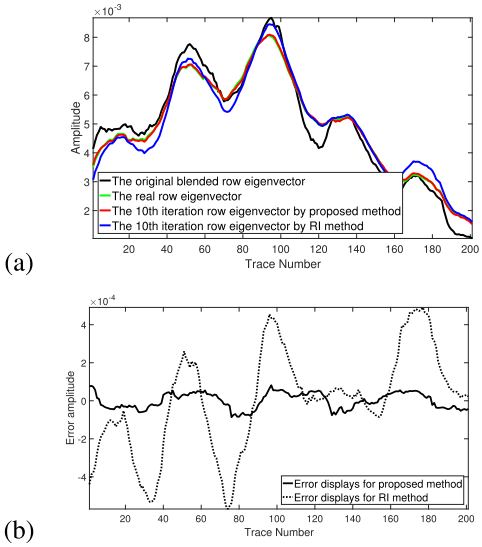

Fig. 10. Evaluating eigenvector convergence of a single frequency. (a) Eigenvector spectrum at 20 Hz. (b) Errors compared with clean data.

method shows a smaller deviation from the real data than the RI method.

Finally, to explain the convergence performance of the proposed method, the eigenvectors at 20 Hz are extracted. The singular spectrum eigenvectors demonstrate the principal components in the seismic data. Fig. 10(a) shows the first eigenvector of different methods for comparison. Fig. 10(b) shows the difference between the real data. The eigenvector of the proposed method converges almost to the real data at the 10th iteration, but there are still obvious deviations in the RI method. This result shows our method with better convergence performance.

IV. Summary

In this letter, we propose a novel iterative method for simultaneous-source separation based on the SSA algorithm. There are two main contributions to this work. First, a simple iterative prediction-subtraction scheme is built, which spares too much local processing in the time domain. Second, an adaptive rank-determination scheme is devised to adaptively predict noise in the frequency domain. In the marine data experiment, the proposed method shows good performance with high SNR results and time-saving.

References

[1] G. Hampson, J. Stefani, and F. Herkenhoff, "Acquisition using simultaneous sources," in *Proc. SEG Tech. Program Expanded Abstr.*, Jan. 2008, pp. 2816–2820.

[2] A. J. Berkhout, "Correction to 'the seismic method in the search for oil and gas: Current techniques and future developments,"' *Proc. IEEE*, vol. 75, no. 3, p. 430, Mar. 1987.

[3] J. Cheng and M. D. Sacchi, "An analysis of the distribution of time delays on simultaneous source separation," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2014, pp. 275–279.

[4] A. J. Berkhout, "Blended acquisition with dispersed source arrays," *Geophysics*, vol. 77, no. 4, pp. A19–A23, Jul. 2012.

[5] R. Baardman and R. van Borselen, "Separating sources in marine simultaneous shooting acquisition—Method & applications," in *Proc. SEG Tech. Program Expanded Abstr.*, Sep. 2012, pp. 1–5.

[6] A. Mahdad, P. Doulgeris, and G. Blacquiere, "Separation of blended data by iterative estimation and subtraction of blending interference noise," *Geophysics*, vol. 76, no. 3, pp. Q9–Q17, May 2011.

[7] S. Zu *et al.*, "Iterative deblending of simultaneous-source data using a coherency-pass shaping operator," *Geophys. J. Int.*, vol. 211, no. 1, pp. 541–557, Oct. 2017.

[8] Y. Chen, "Iterative deblending with multiple constraints based on shaping regularization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2247–2251, Nov. 2015.

[9] S. Gan, S. Wang, Y. Chen, and X. Chen, "Simultaneous-source separation using iterative seislet-frame thresholding," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 197–201, Feb. 2016.

[10] Y. Xue, M. Man, S. Zu, F. Chang, and Y. Chen, "Amplitude-preserving iterative deblending of simultaneous source seismic data using high-order radon transform," *J. Appl. Geophys.*, vol. 139, pp. 79–90, Apr. 2017.

[11] V. Oropeza and M. Sacchi, "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis," *Geophysics*, vol. 76, no. 3, pp. V25–V32, May 2011.

[12] S. Trickett and L. Burroughs, "Prestack rank-reduction-based noise suppression: Theory," *CSEG Recorder*, vol. 34, no. 9, pp. 24–31, Jan. 2009.

[13] J. Cheng and M. D. Sacchi, "Separation and reconstruction of simultaneous source data via iterative rank reduction," *Geophysics*, vol. 80, no. 4, pp. V57–V66, Jul. 2015.

[14] W. Huang, R. Wang, Y. Chen, H. Li, and S. Gan, "Damped multichannel singular spectrum analysis for 3D random noise attenuation," *Geophysics*, vol. 81, no. 4, pp. V261–V270, Jul. 2016.

[15] Y. Xue, F. Chang, D. Zhang, and Y. Chen, "Simultaneous sources separation via an iterative rank-increasing method," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1915–1919, Dec. 2016.

[16] J. Cheng and M. D. Sacchi, "Separation of simultaneous source data via iterative rank reduction," in *Proc. SEG Tech. Program Expanded Abstr.*, Sep. 2013, pp. 88–93.

[17] M. Ma, S. Wang, S. Yuan, J. Gao, and S. Li, "Multichannel block sparse Bayesian learning reflectivity inversion with l-norm criterion-based q estimation," *J. Appl. Geophys.*, vol. 159, pp. 434–445, Dec. 2018.