ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

A pre-trained dictionary learning framework combined with projections onto convex sets for seismic data reconstruction and denoising

Jianlei Zhang[1,2,3], Bo Yang[4]*, Min Bai[4], Xilin Qin[4], Baobin Wang[2,3], Zhen Zou[2,3], and Boyuan Lv[4]

[1]College of Geophysics, China University of Petroleum (Beijing), Beijing, China

[2]National Engineering Research Center for Computer Software for Oil & Gas Exploration, Research Institute of Petroleum Exploration and Development, Zhuozhou, Hebei, China

[3]BGPInc., China National Petroleum Corporation, Zhuozhou, Hebei, China

[4]Key Laboratory of Exploration Technologies for Oil and Gas Resources of Ministry of Education, Yangtze University, Wuhan, Hubei, China

Abstract

The dictionary learning approach has proven effective in seismic data denoising and interpolation. Its core advantage lies in the ability to continuously update the initial dictionary, thereby adapting to the complex structural characteristics of seismic data. However, many existing implementations rely on predefined transforms (e.g., discrete cosine transform) for dictionary initialization. These fixed, data-agnostic bases often fail to fully capture the unique features of seismic signals, which may compromise the sparsity and fidelity of signal representation. Such a limitation can significantly degrade the performance of tasks requiring high-precision reconstruction or noise attenuation. To address this issue, we propose an innovative dictionary learning framework based on a variational sparse representation model. Specifically, this framework first extracts small data patches from arbitrary locations in seismic data, and then constructs a pre-training dataset using a windowing algorithm to preserve fine-grained data features. This process yields an initial dictionary that inherently encodes the intrinsic characteristics of the input seismic data. Subsequently, the initial dictionary is separately refined and updated through the K-singular value decomposition (K-SVD) and sequential generalization of K-means (SGK) algorithms, resulting in an optimized dictionary with more accurate and data-adaptive features. In addition, we integrate a multi-iteration projections onto convex sets algorithm to compensate for missing data features, ultimately achieving high-precision seismic data interpolation and noise attenuation. Numerical experiments demonstrate that the proposed methods (variational SGK and variational K-SVD) outperform the conventional K-SVD and SGK algorithms in both interpolation accuracy and denoising performance.

Keywords: K-singular value decomposition; Sequential generalization of K-means; Variational sparse representation; Projections onto convex sets; Seismic data reconstruction and denoising

***Corresponding author:**
Bo Yang
(ycl2980738443@163.com)

Citation: Zhang J, Yang B, Bai M, *et al*. A pre-trained dictionary learning framework combined with projections onto convex sets for seismic data reconstruction and denoising. *J Seismic Explor*. 2026;35(1):282-299.
doi: 10.36922/JSE025440100

Received: October 31, 2025

Revised: January 5, 2026

Accepted: January 15, 2026

Published online: February 25, 2026

1. Introduction

During the field data acquisition process, the complexity of geological and construction conditions can lead to incomplete data collection. To minimize costs, the acquisition system is often not dense enough, resulting in overly sparse datasets. In addition, the presence of restricted areas and obstacles can limit the deployment of receivers in certain locations, making it impossible to fully capture reflection information from these regions. Beyond the aforementioned issues, the raw data also contains numerous anomalous signals caused by environmental influences and unstable noise. Simply removing these anomalies may lead to significant losses in seismic data. Therefore, the raw data collected in the field must be processed to attenuate noise interference, reconstruct missing data, and restore complete seismic records. This will significantly improve the signal-to-noise ratio (SNR), resolution, and fidelity, thereby enabling more reliable analysis and evaluation of hydrocarbon reservoir characteristics in specified regions.

After years of research and development by numerous scholars in this field, the reconstruction methods can generally be categorized into several types. Wave equation-based seismic data reconstruction methods reconstruct data based on the characteristics of seismic wave propagation underground and often require substantial computational resources.[1,2] The method based on prediction filtering applies convolution operations between target data and prediction filters to estimate missing trace data corresponding to specific frequencies.[3,4] For regular or uniform sampling of seismic data, insufficient spatial sampling rates can often lead to spatial aliasing phenomena. Methods based on rank reduction optimization accomplish signal recovery through either Hankel matrix rank reduction or data dimension reduction techniques.[5,6] Nevertheless, the computational burden is significant for complex data. Seismic data reconstruction methods based on mathematical transforms can generally be categorized into fixed-basis methods and dictionary learning methods. Fixed-basis approaches (e.g., Fourier transform,[7] Radon transform,[8] Curvelet transform,[9] and Seislet transform[10,11]) rely on predefined transforms for sparse representation, and the reconstruction performance is inherently influenced by the chosen mathematical basis. Although these methods are computationally efficient and straightforward to implement, they cannot adaptively model diverse seismic structural features.[12] Representative methods, such as Curvelet-based reconstruction, enforce sparsity via thresholding in the Curvelet domain while applying projections onto convex sets (POCS)-based data consistency constraints, thereby achieving denoising and interpolation simultaneously and enhancing reflector continuity.[13] In contrast, dictionary learning methods suppress noise through sparse coding and iteratively learn a data-adaptive dictionary that better captures intrinsic structural characteristics, enabling unified interpolation and denoising, particularly with advantages in scenarios involving weak signals or high levels of missing data.[14]

The emergence of compressed sensing technology has driven significant progress in dictionary learning methods across multiple application fields. A novel statistical-based masking strategy has been proposed and applied to the compressed sensing framework for outlier denoising, offering new insights to enhance the applicability of compressed sensing in processing data contaminated with abnormal noise.[15] Building on this foundation, a new dictionary learning framework based on fully adaptive boundary empirical mode decomposition has been further developed and applied to seismic data denoising. The integration of this framework with dictionary learning improved the adaptability and noise removal efficacy for complex seismic data.[16] These approaches typically involve two key phases: performing sparse coding and then updating the dictionary, where iterative optimization between these stages yields an ideal sparse representation model.[17,18] Compared with signal transform methods, the advantage of dictionary learning methods lies in their ability to represent signals through a few sparse coefficients, which reduces computational costs and separates the desired components from other components. To address the problem of structured noise removal in seismic data, relevant studies have explored the combination of dictionary learning and atomic filtering, focusing on investigating the impact of atomic filtering operations on structured noise removal performance. These studies provide a reference for parameter optimization and strategy design of dictionary learning in handling specific types of noise.[19] The key difference between dictionary transforms and traditional transforms is that dictionary learning methods use a variable dictionary, whereas traditional transforms (such as wavelet transform) maintain a fixed dictionary.[20-22] Currently, such learned dictionaries with excellent representation performance are widely employed in interpolation and noise reduction applications.[23,24] Numerous overlapping small patches are segmented from the data and utilized to train the dictionary.[25] After a sufficient number of iterations, the resulting dictionary effectively captured reflection signals while attenuating noise. A dictionary with coherence constraints was introduced to facilitate learning without requiring any pre-existing information,[26] alleviating the issue of unstable noise prior information at different locations. An optimized approach for matrix rank reduction was utilized to simultaneously accomplish both

noise suppression and data interpolation.[27] An effective approach was adopted to extract leaked signals from noise; algebraic means were used for dictionary update instead of singular value decomposition (SVD),[28] addressing the high computational burden of training dictionaries in high-dimensional seismic data.

The K-singular value decomposition (K-SVD) algorithm ranks among the most traditional methods for dictionary learning.[29,30] This algorithm begins with a discrete cosine transform (DCT) dictionary as the initial dictionary and iteratively refines it to extract features. However, DCT dictionary atoms do not learn features, resulting in limited extraction of features, which in turn affects the accuracy of reconstruction and denoising.[31-34]

In this study, we introduce a novel dictionary learning framework derived from a variational sparse representation model, offering enhanced signal representation capabilities. First, we extract the features from seismic data using this model to obtain a preliminary dictionary. Subsequently, the traditional dictionary learning methods, such as K-SVD or sequential generalization of K-means (SGK), are applied to update the initial dictionary. We utilize the dictionaries updated by both algorithms for reconstructing and denoising. The POCS algorithm can effectively remove noise generated by missing traces through multiple iterations, thereby restoring the missing seismic traces. This approach effectively supplements the missing features in incomplete seismic datasets, enabling both reconstruction and denoising.

2. Theory

This section first introduces the variational sparse representation model, followed by an overview of the basic theories of the K-SVD and SGK methods. Finally, we introduce the reconstruction theory based on the POCS algorithm.

2.1. The variational sparse representation model

The general formula of this model is expressed in **Equation 1:**

$$\mathbf{b} = \sum_{\mathbf{i=1}}^{\mathbf{m}} \mathbf{z_i q_i} = \mathbf{Qz} \tag{1}$$

where the expression $\mathbf{z}^T = [\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_m]$ represents a random vector, in which the components $\mathbf{z}_i$ correspond to their respective coefficients. Here, the vector $\mathbf{z}$ represents the orthogonal projection of $\mathbf{b}$ in the direction of $\boldsymbol{q}_i$. $\mathbf{Q}$ represents the extraction of data-driven basis functions and patches from the input data $\mathbf{Y}$.

The model extracts patches $\mathbf{Q}$ from the input data $\mathbf{Y}$ at arbitrary locations. Next, we constructed a series resembling **Equation 1** by the windowing algorithm to identify data features and establish a preliminary dictionary $\mathbf{D}_0$. Let $\mathbf{D}_0 = [\mathbf{d}_1,\mathbf{d}_2,\ldots,\mathbf{d}_k]$. We utilized the updated dictionary $\mathbf{D}$ derived from the dictionary learning approaches to reduce noise.

The pth window in the model is denoted as $\mathbf{I}_p$. Let $\mathbf{l}_p = \sum_{i=1}^{m} \mathbf{z}_{li}\mathbf{Q}_i$ denote the representation of $\mathbf{I}_p$ as an expansion using the patches $\mathbf{Q}_i$. This equation can be converted into **Equation 1**, where the window $\mathbf{I}_p$ corresponds to the pth implementation of $\mathbf{b}$ by mapping it according to lexicographical order. To implement this approach, k distinct windows must be extracted from the source dataset, with each window producing corresponding realizations of $\mathbf{b}$ and $\mathbf{z}$ that systematically map the patches $\mathbf{Q}_i$ to vectors $\mathbf{q}_i$ according to the lexicographic ordering specified in **Equation 1**.

2.2. The K-SVD algorithm

The K-SVD model can be described as an optimization problem as follows in **Equation 2**:

$$\min_{\mathbf{D},\mathbf{X}} \left\|\mathbf{Y} - \mathbf{DX}\right\|_F^2 \quad \text{s.t.}\forall i, \left\|\mathbf{x}_i\right\|_0 \le T \tag{2}$$

where the observed signal $\mathbf{Y}$ is expressed as the product of the dictionary matrix $\mathbf{D}$ and sparse coefficients $\mathbf{X}$, with T indicating the desired sparsity.

When determining sparse coefficients, we employed the Orthogonal Matching Pursuit algorithm to find the solution. Furthermore, there are various other similar methods capable of achieving solutions, including Basis Pursuit and Bregman algorithms, among others. The approach updates the dictionary by updating it one column individually at each step. As an example, during the refinement process for the dictionary t_{th} column, the corresponding vector is denoted as $\mathbf{d}_t$, referred to as an atom, while the corresponding t_{th} row vector of $\mathbf{X}$ is recorded as $\mathbf{x}_t$. The mathematical derivation is presented through the subsequent formulation in **Equation 3**:

$$\begin{aligned}\left\|\mathbf{Y} - \mathbf{DX}\right\|_F^2 &= \left\|\mathbf{Y} - \sum_{j=1}^{t} \mathbf{d}_j\mathbf{x}_j\right\|_F^2 \\ &= \left\|\left(\mathbf{Y} - \sum_{j=1, j\neq t} \mathbf{d}_j\mathbf{x}_j\right) - \mathbf{d}_t\mathbf{x}_t\right\|_F^2 = \left\|\mathbf{E}_t - \mathbf{d}_t\mathbf{x}_t\right\|_F^2\end{aligned} \tag{3}$$

according to **Equation 3**, we can obtain **Equation 4**:

$$\mathbf{E}_t = \mathbf{Y} - \sum_{j=1, j\neq t} \mathbf{d}_j\mathbf{x}_j \tag{4}$$

where $\mathbf{E}_t$ represents the error matrix. Meanwhile, the optimization problem can be converted to **Equation 5**:

$$\min_{\mathbf{d}_t, \mathbf{x}_t} \left\| \mathbf{E}_t - \mathbf{d}_t \mathbf{x}_t \right\|_F^2 \tag{5}$$

To make the obtained sparse code as sparse as possible, we need to extract the position in $\mathbf{E}_t$ that is not zero for $\boldsymbol{x}_j$ to get the new $\mathbf{E}_t^{'}$. Here, we used the SVD method to solve this formula using **Equation 6**:

$$\mathbf{E}_t^{'} = \mathbf{U} \pounds \mathbf{V}^{\mathrm{T}} \tag{6}$$

The dictionary's t_{th} column undergoes updating through the initial column of **U**, whereas $\mathbf{x}_t^{'}$ derives from multiplying the initial column of **V** with the initial diagonal element in **Σ**. This update enables the direct substitution of the resultin $\mathbf{x}_t^{'}$ into $\mathbf{x}_t$, effectively minimizing the value of **Equation 3**.

2.3. The sequential generalization of K-means algorithm

We introduce a dictionary learning model, as shown in **Equation 7**:

$$\forall_j \mathbf{x}_j = arg \min_{\mathbf{x}_j} \left\| \mathbf{Y} - \mathbf{DX} \right\|_F^2, s.t. \forall_j \mathbf{x}_j = \mathbf{e}_p \tag{7}$$

where $\mathbf{e}_p$ represents a unit vector that has a value of one for the t_{th} element and is zero for all remaining elements. During the sparse coding process, we can consider the constraint on $\mathbf{x}_j$, specifically $\mathbf{x}_j = \boldsymbol{e}_p$. In the dictionary updating process, **D** needs to be updated in accordance with the K-SVD algorithm. While using K-SVD to address the same dictionary update problem is a common approach, we transitioned to an optimized approach by enforcing unit vector constraints on coefficients. The following optimization functions are defined as **Equation 8**:

$$J = \left\| \mathbf{E}_t^{'} - \mathbf{d}_t \mathbf{x}_t^{'} \right\|_F^2 \tag{8}$$

through computing the partial derivative of J concerning the vector $\mathbf{d}_t$ and equating it to zero, we obtain **Equation (9)**:

$$\mathbf{d}_t = \mathbf{E}_t^{'} \left(\mathbf{x}_t^{'} \right)^T \left(\mathbf{x}_t^{'} \left(\mathbf{x}_t^{'} \right)^T \right)^{-1} \tag{9}$$

assuming that the row vector $\mathbf{x}_t^{'}$ has N^t non-zero elements, then we obtain **Equation 10**:

$$\mathbf{x}_t^{'} \left(\mathbf{x}_t^{'} \right)^T = N^t \tag{10}$$

according to Wang *et al.*[35] we can get **Equation 11:**

$$\mathbf{d}_t = \frac{1}{N^t} \sum_{i=1}^{N^t} \mathbf{s}_i \tag{11}$$

Equation 11 represents the dictionary optimization procedure within the SGK methodology. Notably, this approach replaces the conventional SVD-based iterative approximation of the dictionary element with a computationally efficient summation operation.

2.4. Reconstruction based on the POCS algorithm

The reconstruction problem may be mathematically formulated as **Equation 12**:

$$\mathbf{y} = \mathbf{Mp} \tag{12}$$

where **y** denotes the acquired incomplete data containing noise, **p** represents the target complete seismic data requiring reconstruction, while **M** stands for the stochastic sampling matrix.

The POCS algorithm is mainly divided into the following steps. First, we selected an appropriate threshold parameter λ_i (i = 1, 2,…, N), N representing the iteration number, and the initial input data were the noise data with random missing before reconstruction, which is defined as $\mathbf{y}_0 = \mathbf{y}$.

Then, the initial input data was subjected to curvelet transform, denoted as $\boldsymbol{\delta}_{i-1} = \boldsymbol{f}\boldsymbol{y}_{i-1}$, where $\boldsymbol{\delta}_{i-1}$ corresponds to the sparse representation of signal $\boldsymbol{y}_{i-1}$ within the transformation space **f**. Using a threshold parameter as a constraint, values less than the threshold λ_i were removed and filled with 0, denoted as $\boldsymbol{\delta}_{i-1} = \mathbf{T}_{\boldsymbol{\lambda i}} \mathbf{f} \mathbf{y}_{i-1}$, where $\mathbf{T}_{\boldsymbol{\lambda i}}$ represents the threshold operator.

Next, we took the inverse of $\boldsymbol{\delta}_i$, which is shown in **Equation 13**:

$$\mathbf{y}_i = \mathbf{f}^{-1} \delta_i \tag{13}$$

After filling the non-missing data from the noisy data **y** with random missing into the $\mathbf{y}_i$, we get **Equation 14**:

$$\mathbf{y}_{i+1} = \left(\mathbf{I} - \mathbf{M} \right) \mathbf{y}_i + \mathbf{My} \tag{14}$$

After the updating $\mathbf{y}_i$, we continued the threshold iteration until the accuracy requirements were met. We performed the inverse transformation on the results $\boldsymbol{\delta}_N$ after the last iteration to acquire the reconstructed data. The selection of the threshold parameter had a significant impact on this process. An appropriately chosen threshold value reduces iteration counts and computational overhead without compromising solution precision.

We first used a variational sparse representation model to obtain an initial dictionary with partial features

learned from the seismic data. Then, the K-SVD and SGK algorithms were employed to update the dictionary. Using the updated dictionary and corresponding sparse coefficients, the data was reconstructed to obtain the initial results. Finally, the missing seismic data features were reconstructed using the POCS algorithm, completing the reconstruction and denoising.

3. Methodology

Figure 1 illustrates the overall workflow of the proposed seismic data denoising and reconstruction method based on POCS preprocessing and data-driven dictionary learning. First, a POCS preprocessing step was applied to the noisy and incomplete input seismic data. Through iterative thresholding and data-consistency projection, this step provided initial noise suppression and basic interpolation in missing regions, efficiently producing a structurally more coherent preliminary estimate. Subsequently, using this preprocessed result as input, dictionary initialization was performed through a variational sparse representation scheme, replacing the fixed DCT dictionary to better match the local structural characteristics of the data. Local image patches were then extracted using a sliding-window strategy, and sparse coding and dictionary updating were conducted using the SGK and K-SVD algorithms, respectively. The resulting sparse coefficients were further processed using a percentage-based complex thresholding scheme to suppress residual noise while retaining the dominant signal components. Finally, patch reconstruction and aggregation yielded the output seismic data with simultaneous denoising and reconstruction.

4. Results and discussion

We conducted validation tests using numerical experiments to evaluate the variational sparse representation SGK (VSGK) and variational sparse representation K-SVD (VK-SVD) algorithms' denoising and reconstruction performance. The assessment employed the SNR metric, mathematically expressed in **Equation 15**:

$$\mathrm{SNR} = 10\log_{10}\frac{\mathbf{Y}_{\mathrm{signal}}}{\mathbf{Y}_{\mathrm{noise}}} \tag{15}$$

where $\mathbf{Y}_{\mathbf{signal}}$ represents the power of the effective signal, and $\mathbf{Y}_{\mathbf{noise}}$ represents the power of the noise.

The assessment employed the structural similarity index measure (SSIM) metric, mathematically expressed in **Equation 16**:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{16}$$

where μ_x represents the local mean of the true data, μ_y represents the local mean of the reconstructed data, σ_x represents the local standard deviation of the true data, σ_y represents the local standard deviation of the reconstructed data, σ_{xy} represents the local covariance between the true

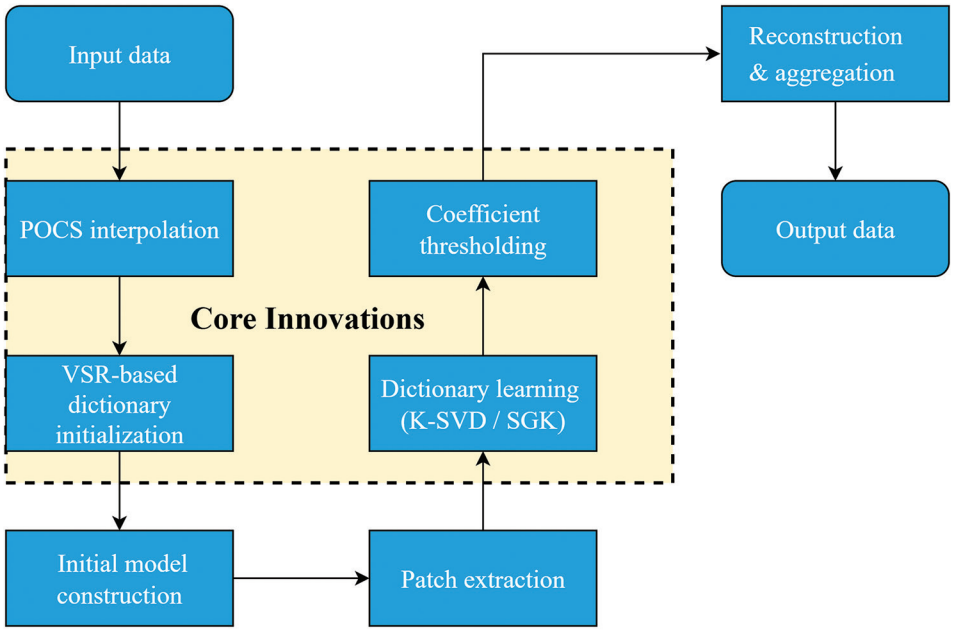

Figure 1. Flow chart of seismic data processing framework based on variational sparse representation
Abbreviations: K-SVD: K-singular value decomposition; POCS: Projections onto convex sets; SGK: Sequential generalization of K-means; VSR: Variational sparse representation.

and reconstructed data, C_1 represents a constant, and C_2 represents another constant.

The assessment employed the peak SNR (PSNR) metric, mathematically expressed as in **Equation 17**:

$$PSNR = 10\log_{10}\left(\frac{X_{max}^2}{MSE}\right) \quad (17)$$

When X_{max} represents the peak amplitude of the signal, and MSE represents the point-wise mean squared error between the reconstructed data and the true data.

4.1. Synthetic data

The synthetic data are shown in Figure 2, and the results of denoising and reconstruction using different methods are presented in Figure 3. SGK and K-SVD effectively removed random noise while partially recovering the seismic data at the missing locations, but some events are not very clear. Significant leakage of useful signals was observed in the areas indicated by the green arrows (Figure 3E and F). In Figure 3C, VSGK shows almost no residual background noise and exhibits better continuity of the events, with only a few useful signal losses in the region marked by the green arrow (Figure 3G). By comparison, the VK-SVD method exhibited less useful signal leakage.

Next, we examined the impact of the variational sparse representation model on overall performance using the data shown in Figure 2B. We replaced the initial dictionary in VSGK and VK-SVD with the DCT dictionary and compared it with the method that uses the dictionary trained by the variational sparse representation model as the initial dictionary. The initial dictionaries and the updated dictionaries for different methods are displayed in Figure 4. For the initial dictionaries, it can be seen that the atoms of the DCT dictionary are unrelated to the

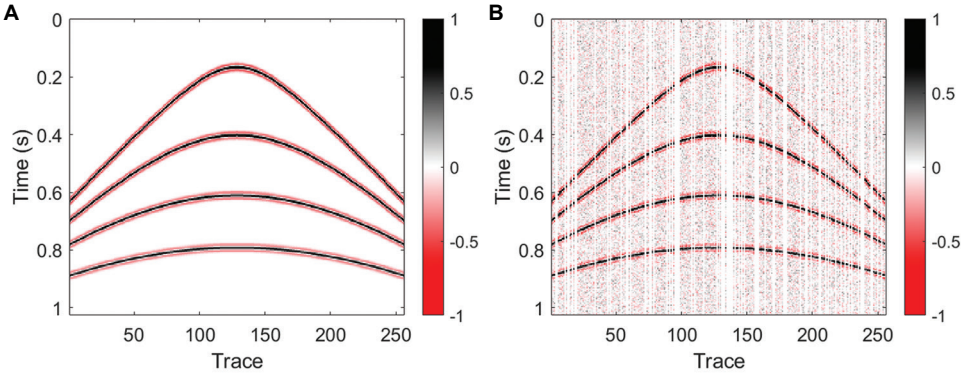

Figure 2. Original data and noisy data with random missing 0.3 data. (A) Original data, (B) noisy data with random missing 0.3 data.

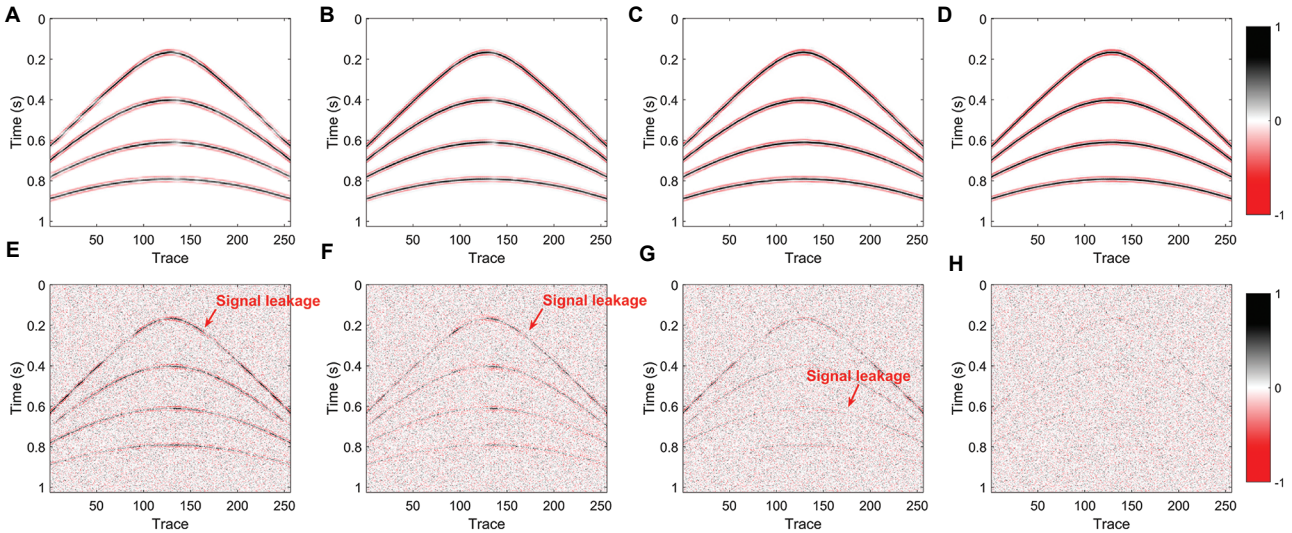

Figure 3. Comparison of results processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The noise that is removed by (E) SGK, (F) K-SVD, (G) VSGK, and (H) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

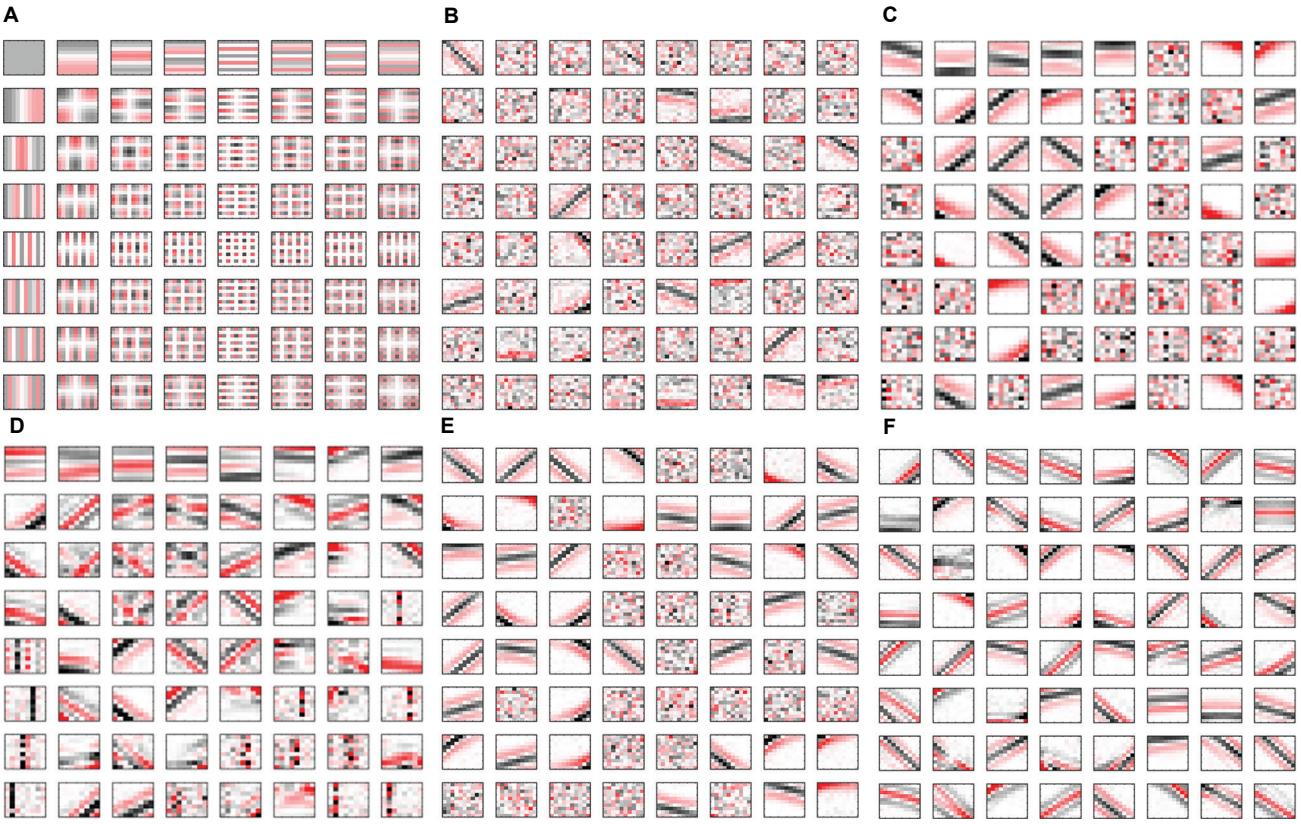

Figure 4. Dictionary comparison. (A) The DCT dictionary, (B) the dictionary obtained from the variational sparse representation (V) model, (C) the dictionary obtained after updating the DCT dictionary using the VSGK method, (D) the dictionary obtained after updating the DCT dictionary using the VK-SVD method, (E) the dictionary obtained after updating (B) using the VSGK method, and (F) the dictionary obtained after updating (B) using the VK-SVD method.
Abbreviations: D: Discrete cosine transform; K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means.

seismic data and have a fixed morphology, while the initial dictionary trained by the variational sparse representation model already captured some characteristics of the seismic data. Figure 4C and D shows the updated dictionaries after updating the DCT dictionary, whereas Figure 4E and F shows the updated dictionaries from the proposed method. The dictionary trained by the variational sparse representation model exhibited a richer morphology after updating and extracted more features from the seismic data. Figure 3 presents the results of reconstruction and denoising using different methods. We can observe that both VSGK and VK-SVD exhibited less effective signal loss compared to SGK and K-SVD, which used the DCT dictionary as the initial dictionary. These methods improved the SNR from 2.07 dB to 10.72 dB (SGK), 11.63 dB (K-SVD), 11.75 dB (VSGK), and 14.43 dB (VK-SVD).

To visualize the details, we conducted a spectrum analysis of the results processed by four methods (Figure 5). Figure 5A presents the spectrum corresponding to the clean data, while Figure 5B displays the spectrum of the noisy data after 0.3 random traces are missing. Figure 5C-F displays the spectrum corresponding to the denoising results from SGK, K-SVD, VSGK, and VK-SVD, respectively. From the figures, it can be seen that both SGK and K-SVD removed some useful signals, but K-SVD experienced slightly lower loss of useful signals. The spectrum obtained through the proposed method closely matches that of the original data. Figure 5G-J shows the difference in spectra between the original data and the processed results. It is evident that VSGK and VK-SVD exhibited fewer useful signals in their denoising residuals, with VK-SVD showing this more prominently.

To further evaluate the four methods, we extracted the 50th trace from the results of each method for single-trace analysis, as shown in Figure 6. In the area enclosed by the red box, it is clear that the curves of the useful signals from VSGK and VK-SVD closely match the clean data, further demonstrating that the proposed method performs better in recovering effective signals.

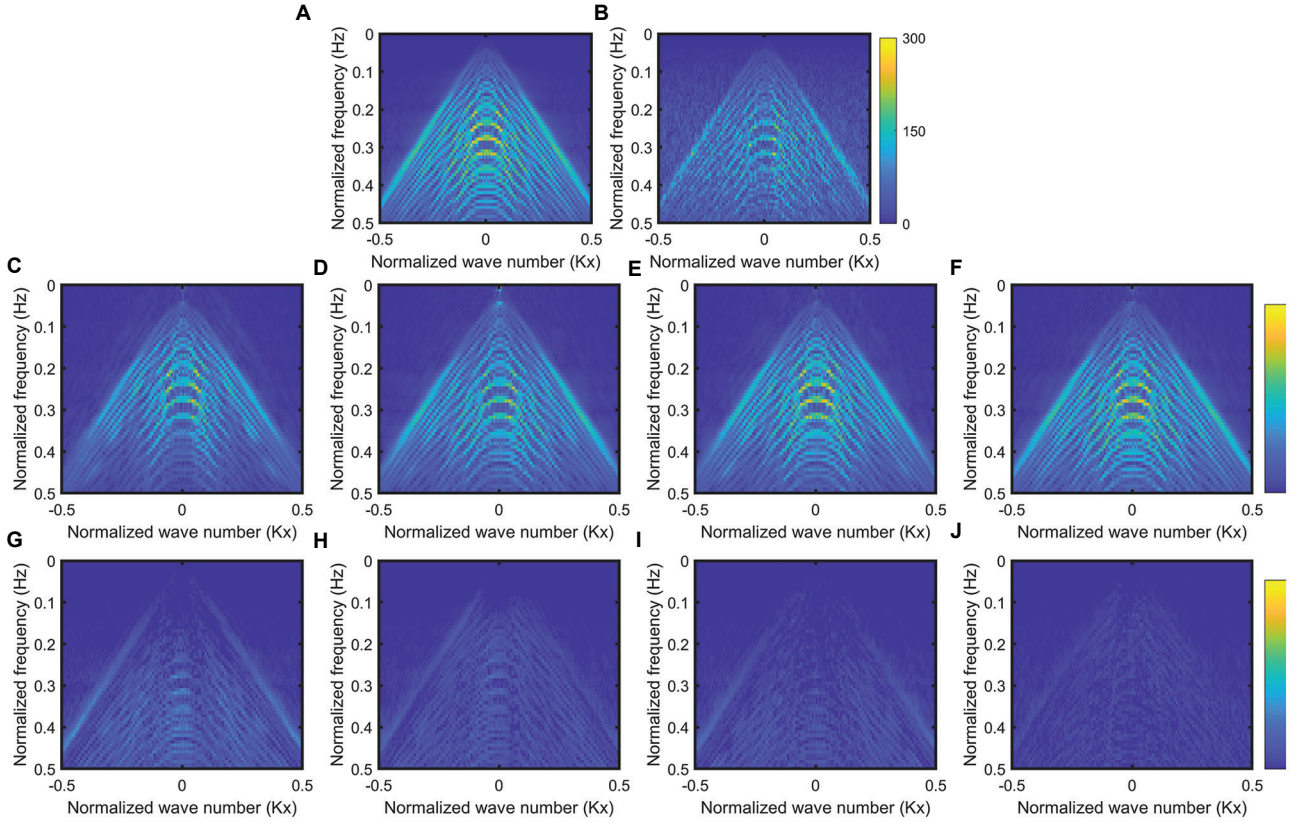

Figure 5. Frequency spectra comparison of various approaches. (A) Original data, (B) noisy data, (C) SGK, (D) K-SVD, (E) VSGK, and (F) VK-SVD. (G-J) The difference between processed data and original data.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

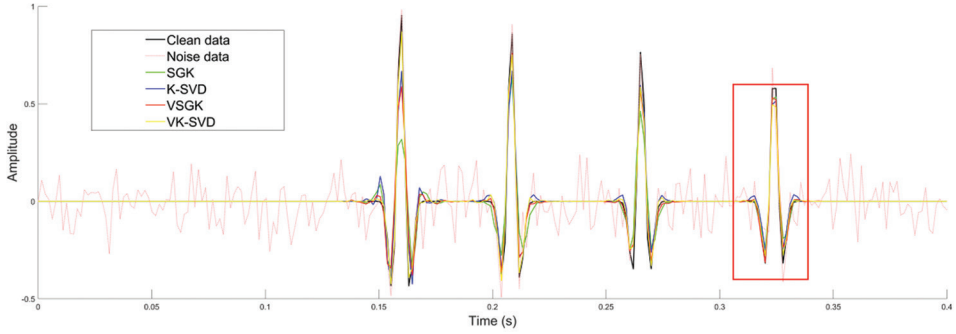

Figure 6. Single-channel seismic record analysis.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

4.2. Field data

This study evaluated the performance of the proposed method in seismic data reconstruction and noise suppression under practical scenarios, utilizing both pre-stack and post-stack field seismic datasets for testing and validation. Among these datasets, Figure 7 shows a pre-stack shot gather, which was acquired using the observation configuration of central-source excitation with bilateral geophone deployment. The propagation characteristics of seismic events and the spatial distribution pattern of noise can be clearly identified from the gather, which was thus employed as the baseline model for real seismic data testing. Figure 7 presents the pre-stack data used in the experiment, which not only exhibited 0.3 random trace missing but also contained significant background noise. To compare the processing effectiveness, four methods—SGK, K-SVD, VSGK, and

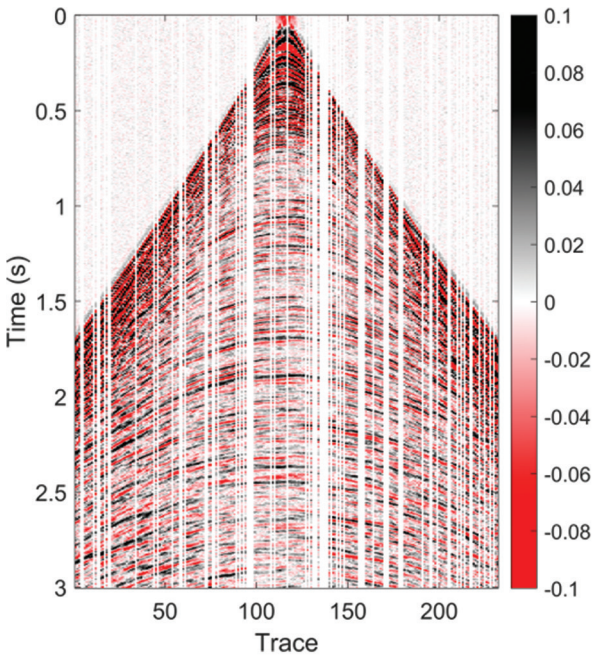

Figure 7. The first field data.

VK-SVD—were applied to this pre-stack data, with the results shown in Figure 8. The experimental results demonstrate that the SGK and K-SVD methods yielded unsatisfactory recovery of the missing traces, leading to noticeable degradation in the spatial continuity of seismic events. In contrast, the proposed VSGK and VK-SVD methods effectively reconstructed the missing seismic traces, significantly enhancing data integrity. However, during the noise suppression process, both methods exhibited a slight leakage of useful signals.

Figure 9 presents a post-stack seismic section in a tilted structural setting. The region with the most complex geological structures within this section was selected as the test target; the distinct tilted structural features of the section are suitable for verifying the applicability of the proposed method in scenarios with complex geological structures. The original data contains different types of events and has 0.3 randomly missing traces, along with a significant amount of noise near the events. Figure 10 displays the results processed by four methods. The results of SGK and K-SVD showed many missing seismic

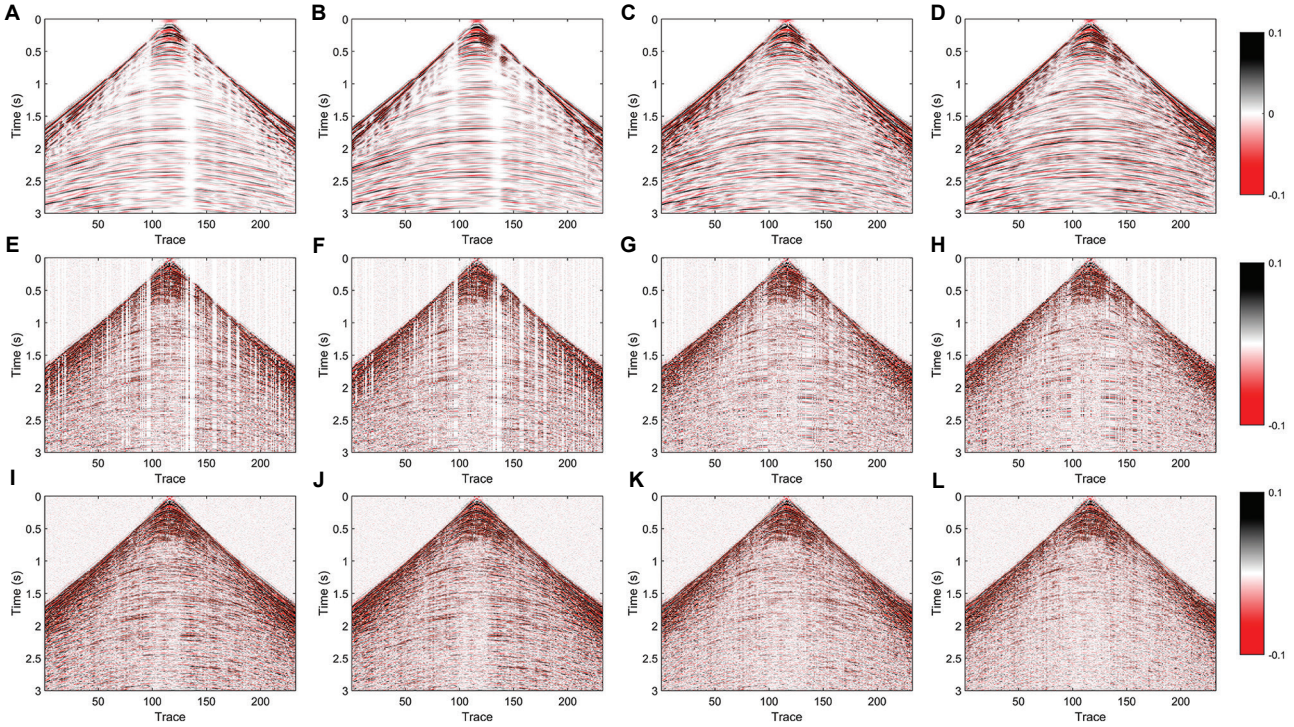

Figure 8. Comparison of the first field data results processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The residual between the (E) SGK-reconstructed data and the original data, (F) K-SVD-reconstructed data and the original data, (G) VSGK-reconstructed data and the original data, and (H) VK-SVD-reconstructed data and the original data. The noise that is removed by (I) SGK, (J) K-SVD, (K) VSGK, and (L) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

data structures (Figure 10A and B), and the recovery of the missing trace was not ideal, compromising the continuity of the events. Figure 10C and G shows the results of reconstruction and denoising using VSGK and the removed noise profile. The experimental results clearly indicate that the method successfully retrieved the missing seismic channels, although there is still some leakage of useful signals. Specifically, the comparative results in Figure 10E-H intuitively demonstrate that the reconstruction performance of the method proposed in this study is significantly superior to that of other comparative methods. After processing with VK-SVD, the seismic events continuity remained effectively maintained, with negligible useful signals detectable in the removed noise profile.

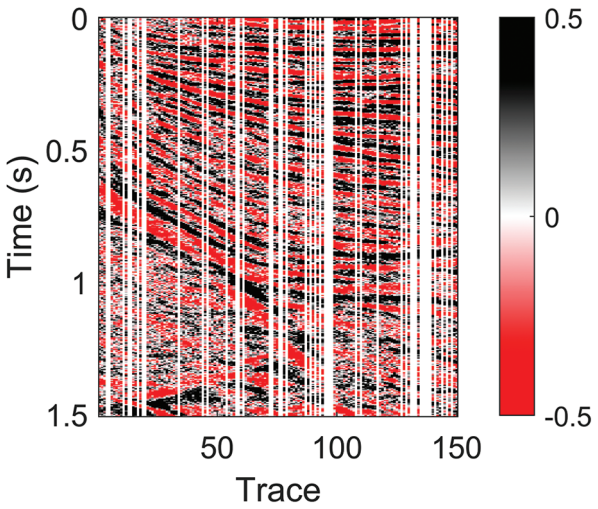

Figure 9. The second field data.

Next, we adopted the data shown in Figure 11 for numerical experiments. Figure 11 displays a post-stack seismic section with layered structures. Extracted from the post-stack seismic dataset of a real working area, this section was used as a test model to further validate the stability and generalization capability of the proposed method in layered geological settings. The results of reconstruction and denoising are presented in Figure 12. It is observed that SGK and K-SVD did not effectively

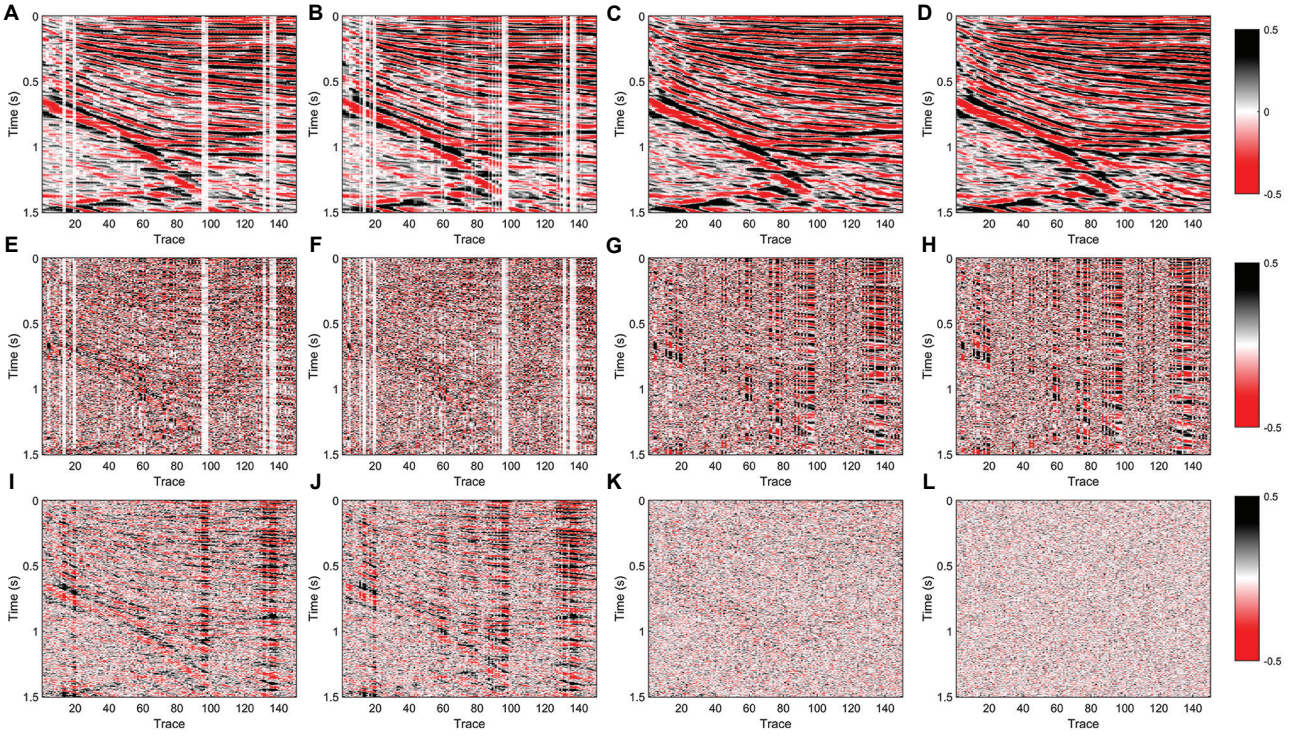

Figure 10. Comparison of the second field data results processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The residual between the (E) SGK-reconstructed data and the original data, (F) K-SVD-reconstructed data and the original data, (G) VSGK-reconstructed data and the original data, and (H) VK-SVD-reconstructed data and the original data. The noise that is removed by (I) SGK, (J) K-SVD, (K) VSGK, and (L) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

recover the partially missing seismic data after processing (Figure 12A and B), and caused damage to a significant amount of useful signals while suppressing noise. In contrast, the results processed by VSGK and VK-SVD (Figure 12C and D) showed clearer events and retained more details of the seismic data structure compared to the first two methods. Further analysis revealed that the proposed method exhibited excellent data reconstruction performance, whereas the SGK and K-SVD methods showed distinctly unsatisfactory reconstruction results (Figure 12E-H).

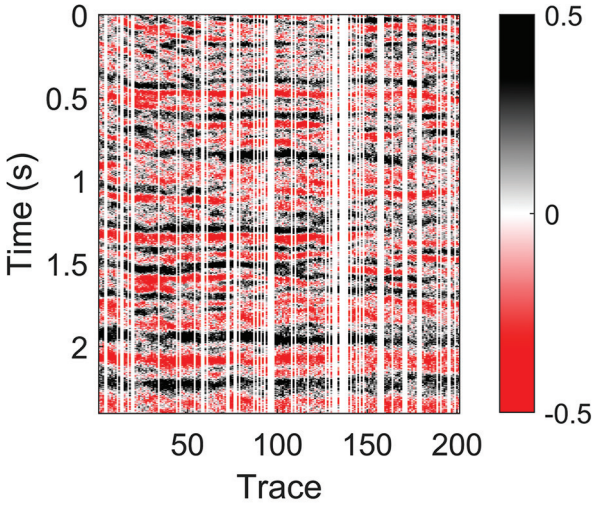

Figure 11. The third field data.

4.3. Robustness analysis in seismic denoising and reconstruction

To verify the applicability of the method proposed in this study in complex data scenarios, the original data were subjected to a 0.5 random missing treatment (Figure 13). From the denoising and reconstruction results (Figure 14), it can be seen that both VSGK and VK-SVD repaired the missing regions in the data and effectively suppressed random noise, fully demonstrating the advantages of the two methods in denoising and reconstruction.

To further investigate the robustness of the methods, the noise intensity was additionally increased based on 0.5 random missing values of the original

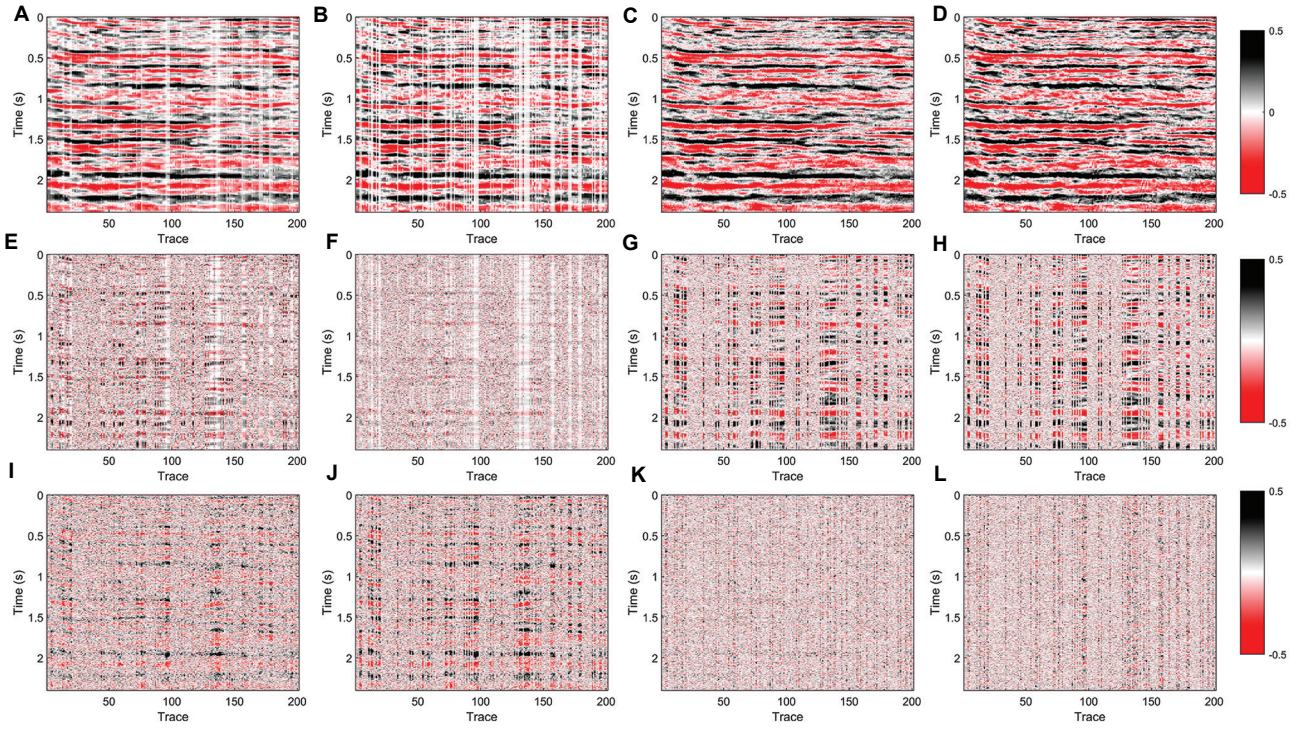

Figure 12. Comparison of the third field data results processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The residual between the (E) SGK-reconstructed data and the original data, (F) K-SVD-reconstructed data and the original data, (G) VSGK-reconstructed data and the original data, and (H) VK-SVD-reconstructed data and the original data. The noise that is removed by (I) SGK, (J) K-SVD, (K) VSGK, and (L) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

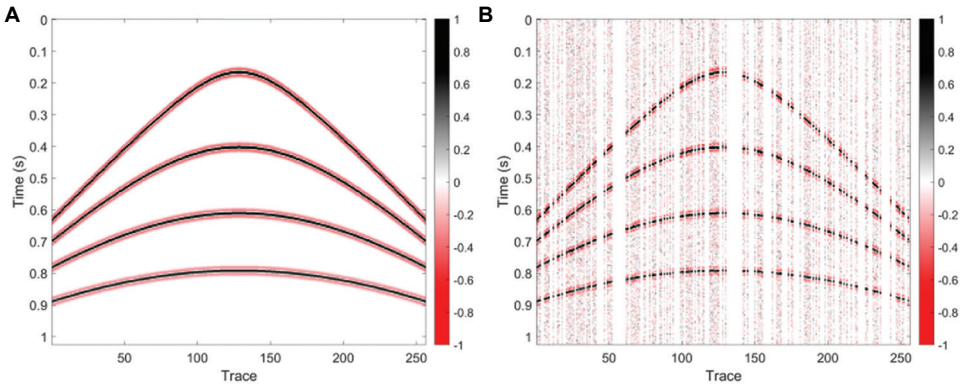

Figure 13. Original data and noisy data with random missing 0.5 data. (A) Original data, (B) noisy data with random missing 0.5 data.

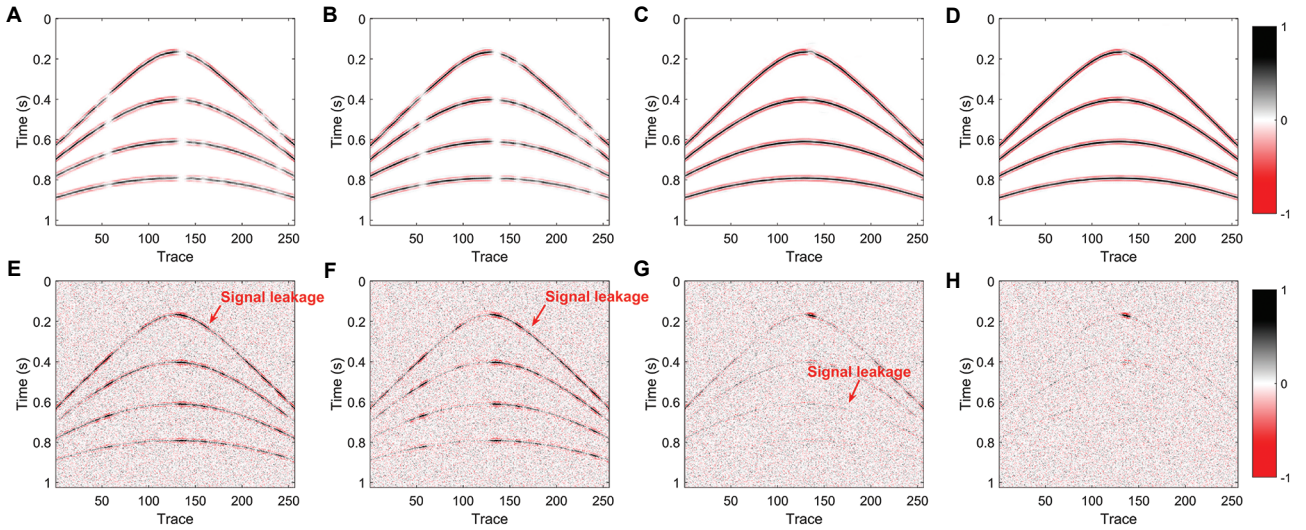

Figure 14. Comparison of the results of data with 0.5 missing treatment processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The noise that is removed by (E) SGK, (F) K-SVD, (G) VSGK, and (H) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

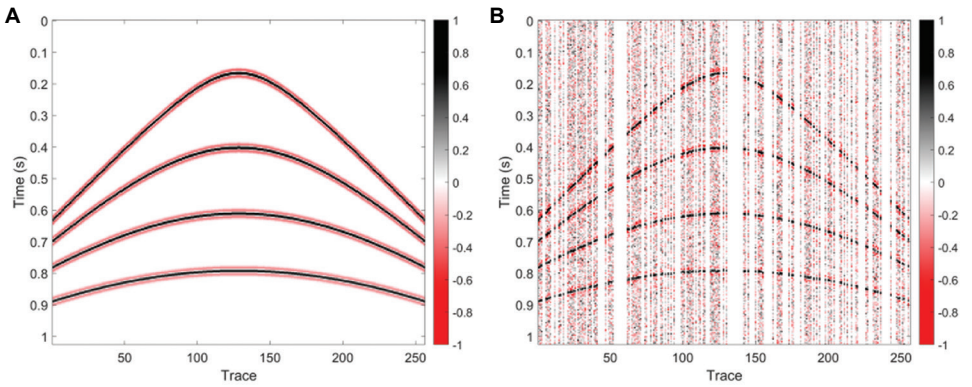

Figure 15. Original data and noisy data with random missing 0.5 data (with enhanced noise intensity). (A) Original data, (B) noisy data with random missing 0.5 data.

data (Figure 15). The comparative experimental results (Figure 16) indicate that the denoising and reconstruction effects of VSGK and VK-SVD were significantly superior to those of the traditional SGK

and K-SVD methods, with obvious advantages in terms of data integrity and SNR improvement. However, due to the dual interference of high missing rate and strong noise, there is still a small amount of noise residue in the reconstruction results of the two methods, which have not yet reached the ideal state of complete noise-free. This also provides a direction for the subsequent optimization of the methods.

To verify the comprehensive performance of the method proposed in this study, two groups of comparative experiments were designed on synthetic seismic data. The first group was subjected to a block-missing combined with noise test experiment. Based on the original data in Figure 17, a noisy dataset with a 0.2 block missing rate was constructed. The denoised and reconstructed results obtained by the proposed method are shown in Figure 18, which indicates that the method still maintains excellent performance under the interference of complex missing and noise. The second group was subjected to a horizontal comparison experiment of algorithms. Taking the original data in Figure 2 as the benchmark, the dataset was processed using three methods, namely VSGK, VK-SVD, and damped rank-reduction (DRR). The experimental results are presented in Figure 19. VSGK and VK-SVD can accurately reconstruct the missing seismic events while denoising, and their performance was significantly superior to that of DRR, whereas DRR exhibited an advantage in computational efficiency.

Seismic data denoising and reconstruction are key steps for subsequent geological structure interpretation. In practical production applications, it is necessary to balance the dual requirements of processing accuracy and computational efficiency.

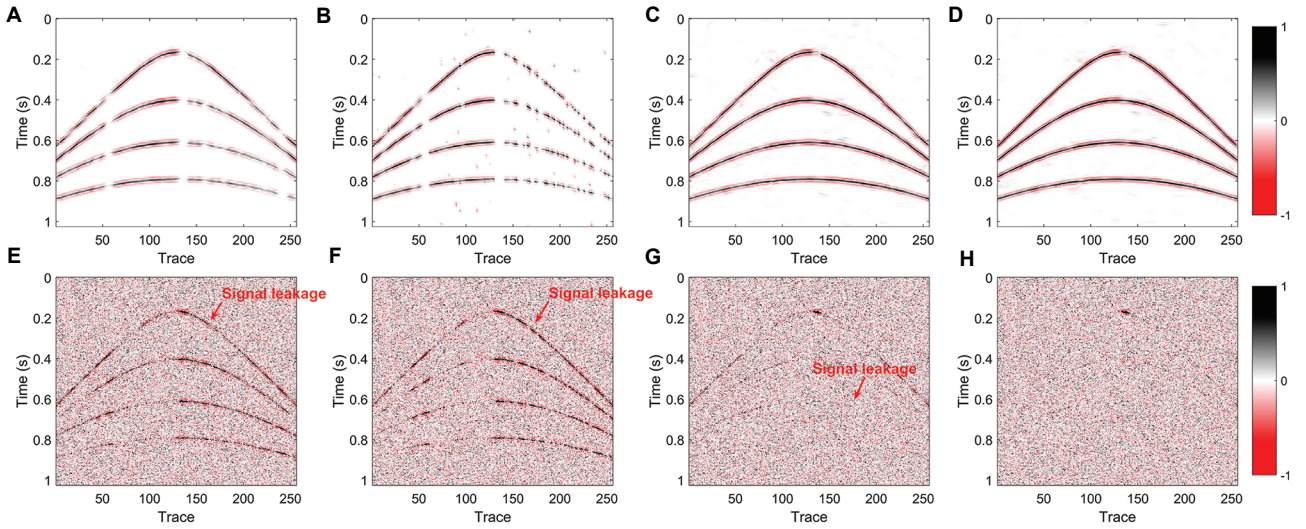

Figure 16. Comparison of the results of data with random missing 0.5 data (with enhanced noise intensity) processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The noise that is removed by (E) SGK, (F) K-SVD, (G) VSGK, and (H) VK-SVD. Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

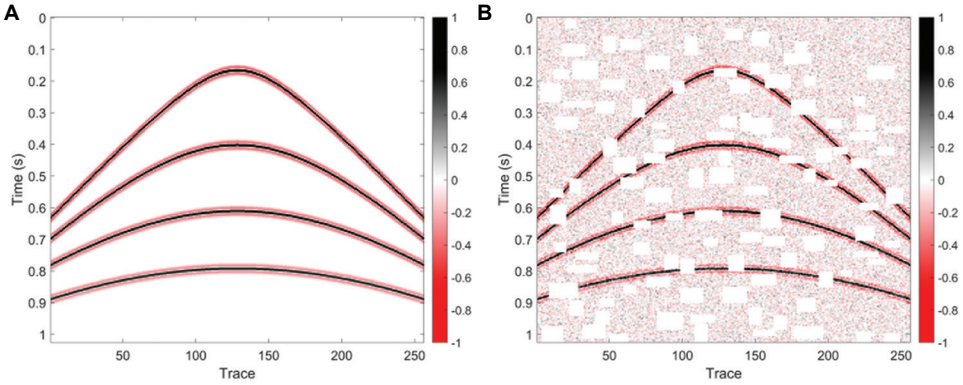

Figure 17. Original data and noisy data with a block missing 0.2 data. (A) Original data, (B) noisy data with block missing 0.2 data.

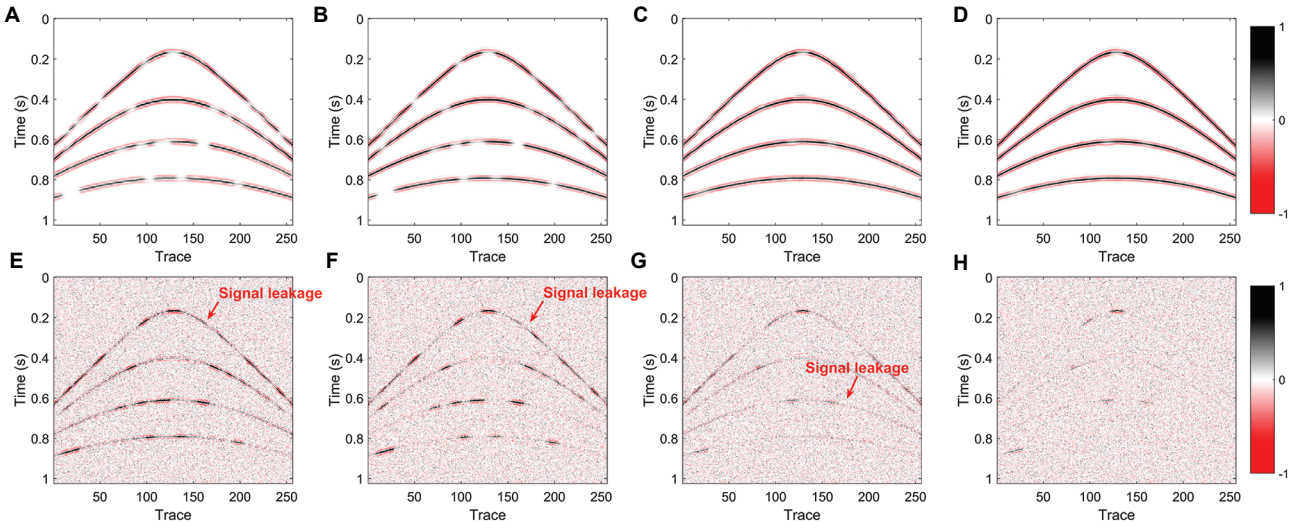

Figure 18. Comparison of the results of the block missing 0.2 data processed using various approaches. The result processed by (A) SGK, (B) K-SVD, (C) VSGK, and (D) VK-SVD. The noise that is removed by (E) SGK, (F) K-SVD, (G) VSGK, and (H) VK-SVD.
Abbreviations: K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

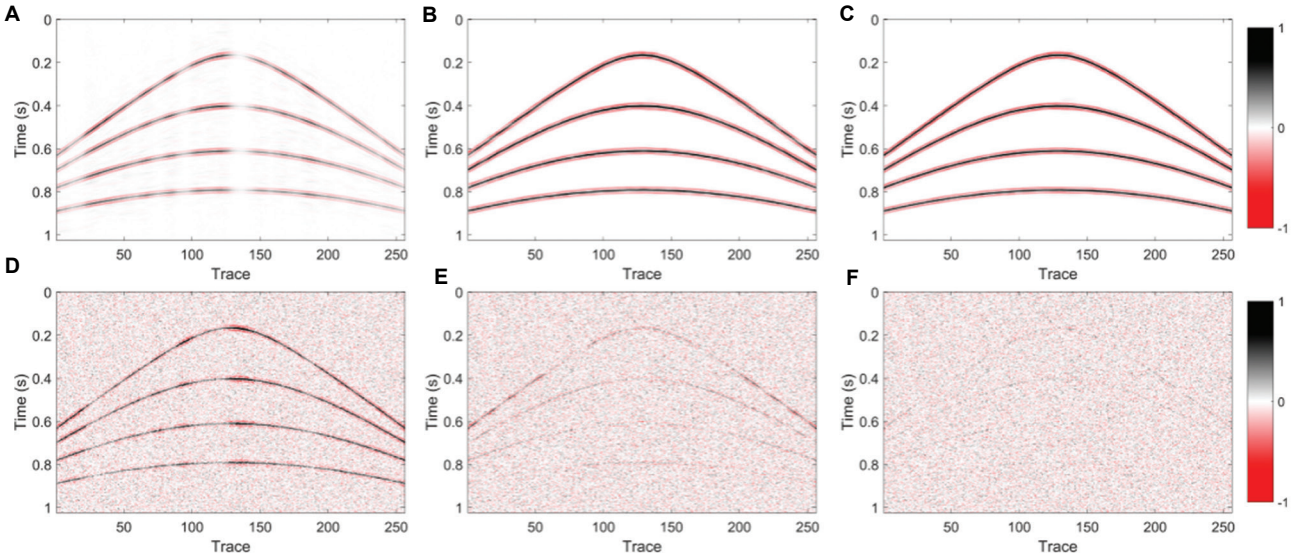

Figure 19. Comparison of the results of the block missing 0.2 data processed using various approaches. The result processed by (A) DRR, (B) VSGK, and (C) VK-SVD. The noise that is removed by (D) DRR, (E) VSGK, and (F) VK-SVD.
Abbreviations: DRR: Damped rank-reduction; K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; V: Variational sparse representation.

Table 1. Comparison of different algorithms

Algorithm	Time (s)	SNR (dB)	Peak SNR (dB)	SSIM
SGK	102.82	6.95	23.51	0.894
K-SVD	1,538.46	10.09	26.65	0.930
VSGK	36.41	11.64	27.77	0.953
VK-SVD	66.11	15.43	32.00	0.964

Abbreviations: DRR: Damped rank-reduction; K-SVD: K-singular value decomposition; SGK: Sequential generalization of K-means; SNR: Signal-to-noise ratio; SSIM: Structural similarity index measure; V: Variational sparse representation.

Table 1 presents the comparison of the four methods when processing 256 × 256-sized data. The running time of VSGK was significantly shorter than that of SGK, and the running time of VK-SVD was lower than that of KSVD. Both improved methods greatly enhanced the efficiency of data processing, with further comparison showing that VSGK and VK-SVD improved the accuracy of reconstruction and denoising. Therefore, the proposed methods of variational sparse representation for pre-trained dictionary learning (VK-SVD and VSGK) not

only improved the accuracy of reconstruction and denoising but also greatly reduced computing time. These improvements could be attributed to the closer resemblance of the variational sparse representation-based pre-trained atoms to the characteristics of seismic data; hence, they converge faster during the iterative process, thus achieving high accuracy and saving computing time.

4.4. Impact of POCS initialization on reconstruction

To quantitatively assess the actual contribution of POCS as an initialization step in the dictionary-driven reconstruction process, this study compared the reconstruction performance of the VSGK and VK-SVD methods with and without POCS initialization on the same test dataset, and systematically contrasted the results with those of conventional SGK, K-SVD, and standalone POCS.

Algorithm 1. Adaptive-threshold projections onto convex sets for seismic data reconstruction

Input: Noisy and incomplete data $\mathbf{d}_0$,
mask $\mathbf{M}$, Maximum iterations N,
threshold decay ε,
Curvelet parameters: n_{scales}, n_{angles}
Output: Reconstructed data $\mathbf{D}_{rec}$
Initialization: $\mathbf{X}^{(0)} \leftarrow \boldsymbol{d}_0$, $\mathbf{1} \leftarrow ones\ (n_1, n_2)$;
for i=1 **to** N $\mathbf{d}_0$

Forward Curvelet transform: $\mathbf{C} \leftarrow$ CURVELET $(\mathbf{X}^{(i-1)}, n_{scales}, n_{angles})$;
Compute maximum coefficient: $c_{max}=\max_{j,k}|\mathbf{C}[j, k]|$;
Adaptive exponential threshold: $\tau_i = c_{max} \cdot \exp\left(\frac{(i-1)\cdot(\ln \varepsilon - \ln c_{max})}{N-1}\right)$;
Hard thresholding: $\mathbf{C}[j, k] \leftarrow \mathbf{C}[j, k] \cdot \mathbf{1}_{|\mathbf{C}[j,k]|>\tau_i}$, $\forall j, k$;
Inverse Curvelet transform: $\mathbf{X}_{temp} \leftarrow$ INVERSE CURVELET $(\mathbf{C})$;
Data consistency (POCS): $\mathbf{X}^{(i)} \leftarrow (\mathbf{1} - \mathbf{M}) \odot \mathbf{X}_{temp} + \mathbf{d}_0$;

Final reconstruction:
$\mathbf{C}_{final} \leftarrow$ CURVELET $(\mathbf{X}^{(N)}, n_{scales}, n_{angles})$
$\mathbf{D}_{rec} \leftarrow$ Re (INVERSE CURVELET $(\mathbf{C}_{final})$)
return $\mathbf{D}_{rec}$

The POCS algorithm employed in this paper is presented in Algorithm 1.

Figure 20 compares the reconstruction results of different algorithms on noisy seismic data. Figure 20A-E shows the reconstructed profiles. POCS exhibited significant residual noise, while VSGK and VK-SVD recovered the seismic events and suppressed random noise reasonably well, though their performance decreased in highly curved regions. Combining POCS with VSGK or VK-SVD (VSGK+POCS and VK-SVD+POCS) significantly improved the reconstruction, enhancing the continuity of seismic events. Figure 20F-J displays the removed noise for each method. POCS removed limited energy, appearing as narrow-band streaks; VSGK and VK-SVD removed more random noise but partially attenuated the useful signals. When combined with POCS, noise was effectively suppressed while preserving the signals. Figure 20K-O presents the residuals between the reconstructed results and the original data, showing that the seismic event energies for VSGK+POCS and VK-SVD+POCS were significantly higher than those of other methods, further demonstrating their superiority in reconstruction accuracy and structural preservation.

Table 2 compares the performance of seven reconstruction algorithms on noisy seismic data in terms of run time, SNR, PSNR, SSIM, SNR gain, and SNR gain rate. Among the methods, VK-SVD + POCS achieved the highest reconstruction quality, with an SNR of 15.43 dB, PSNR of 32.00 dB, and SSIM of 0.964, indicating excellent noise suppression and structural preservation. Although K-SVD and SGK can also reconstruct the data, they require higher computational costs and deliver inferior results compared to methods combined with POCS. POCS alone has limited denoising capability, but when combined with VSGK or VK-SVD, it enhances SNR gain and seismic event continuity, yielding better reconstruction. Overall, POCS-

Table 2. Performance comparison of reconstruction algorithms with and without POCS initialization

Algorithm	Time (s)	SNR (dB)	Peak SNR (dB)	SSIM	SNR gain (dB)	SNR rate (dB/s)
POCS	28.42	3.51	20.07	0.201	1.44	0.0507
VSGK	9.87	9.41	25.97	0.953	7.34	0.7437
VK-SVD	39.78	11.13	27.70	0.947	9.06	0.2277
VSGK+POCS	36.41	11.64	27.77	0.953	9.57	0.2628
VK-SVD+POCS	66.11	15.43	32.00	0.964	13.36	0.2021
SGK	102.82	6.95	23.51	0.894	4.88	0.0475
K-SVD	1,538.46	10.09	26.65	0.930	8.02	0.0052

Abbreviations: K-SVD: K-singular value decomposition; POCS: Projections onto convex sets; SGK: Sequential generalization of K-means; SNR: Signal-to-noise ratio; SSIM: Structural similarity index measure; V: Variational sparse representation.

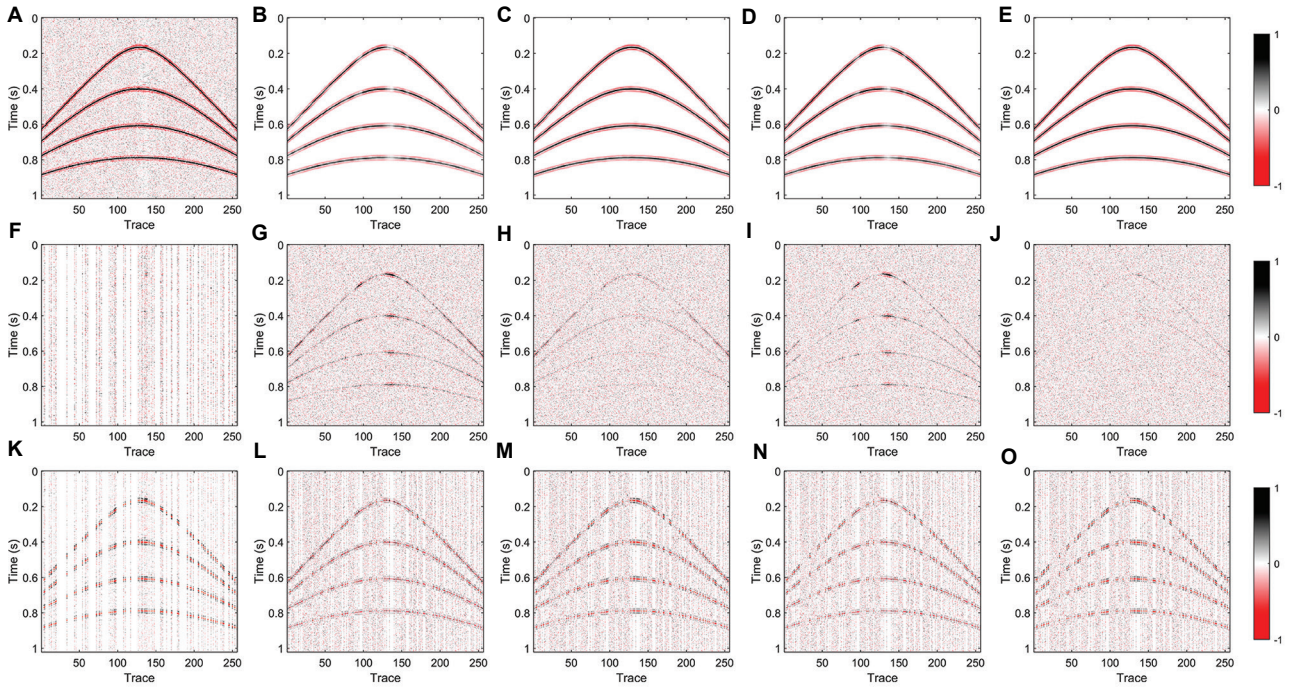

Figure 20. Comparison of the results of noisy seismic data processed using various approaches. The result processed by (A) POCS, (B) VSGK, (C) VSGK + POCS, (D) VK-SVD, and (E) VK-SVD + POCS. The noise that is removed by (F) POCS, (G) VSGK, (H) VSGK + POCS, (I) VK-SVD, and (J) VK-SVD + POCS. The residual between the (K) POCS-reconstructed and the original data, (L) VSGK-reconstructed and the original data, (M) VSGK + POCS-reconstructed and the original data, (N) VK-SVD-reconstructed and the original data, and (O) VK-SVD + POCS-reconstructed and the original data
Abbreviations: K-SVD: K-singular value decomposition; POCS: Projections onto convex sets; SGK: Sequential generalization of K-means; V: Variational sparse representation.

based initialization improves reconstruction performance, with VSGK + POCS and VK-SVD+POCS achieving the best results.

5. Conclusion

We developed a new dictionary learning method based on a variational sparse representation model. This method performed a pre-training through the variational sparse representation model to conduct feature extraction from seismic signals, generating an initial dictionary. The dictionary was then updated by applying the SGK and K-SVD algorithms to further extract deep features from the data. In addition, multiple iterations of the POCS algorithm were employed to supplement some missing data features. The numerical experiments showed that VSGK and VK-SVD can effectively achieve interpolation and denoising. Compared to traditional SGK and K-SVD methods, the proposed methods learnt deeper features from the data, resulting in significant improvements in the performance of denoising and reconstruction, particularly in maintaining lower signal leakage.

Acknowledgments

None.

Funding

This research was supported by the Open Fund of the National Engineering Research Center for Computer Software for Oil & Gas Exploration (grant no. DFWT-ZYRJ-2024-JS-83).

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

Conceptualization: Jianlei Zhang, Min Bai, Xilin Qin, Baobin Wang, Zhen Zou
Formal analysis: Min Bai, Xilin Qin
Investigation: Jianlei Zhang, Bo Yang
Methodology: Jianlei Zhang, Bo Yang
Writing–original draft: Bo Yang
Writing–review & editing: Bo Yang, Boyuan Lv

Availability of data

The data generated or analyzed during this study are available from the corresponding author upon reasonable request.

References

1. Canning A, Gardner GHF. Regularizing 3-D data sets with DMO. *Geophysics*. 1996;61:1103-1114.

   doi: 10.1190/1.1444031

2. Chemingui N, Biondi B. Handling the irregular geometry in wide-azimuth surveys. *SEG Int Expo Annu Meet*. 1996;10: 32-35.

   doi: 10.1190/1.1826634

3. Spitz S. Seismic trace interpolation in the F-X domain. *Geophysics*. 1991;56:785-794.

   doi: 10.1190/1.1443096

4. Li C, Liu G, Hao Z, Zu S, Mi F, Chen X. Multidimensional seismic data reconstruction using frequency-domain adaptive prediction-error filter. *IEEE Trans Geosci Remote Sens*. 2018;56:2328-2336.

   doi: 10.1109/tgrs.2017.2778196

5. Gao J, Sacchi MD, Chen X. A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions. *Geophysics*. 2012;78:V21-V30.

   doi: 10.1190/geo2012-0038.1

6. Zhang D, Zhou Y, Chen H, Chen W, Zu S, Chen Y. Hybrid rank-sparsity constraint model for simultaneous reconstruction and denoising of 3D seismic data. *Geophysics*. 2017;82:V351-V367.

   doi: 10.1190/geo2016-0557.1

7. Gao J, Stanton A, Naghizadeh M, Sacchi MD, Chen X. Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction. *Geophys Prospect*. 2013;61:138-151.

   doi: 10.1111/j.1365-2478.2012.01103.x

8. Trad DO, Ulrych TJ, Sacchi MD. Accurate interpolation with high-resolution time-variant radon transforms. *Geophysics*. 2002;67:644-656.

   doi: 10.1190/1.1468626

9. Hennenfent G, Fenelon L, Herrmann FJ. Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach. *Geophysics*. 2010;75:WB203-WB210.

   doi: 10.1190/1.3494032

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

    doi: 10.1190/1.3380591

11. Gan S, Wang S, Chen Y, Zhang Y, Jin Z. Dealiased seismic data interpolation using seislet transform with low-frequency constraint. *IEEE Geosci Remote Sens Lett*. 2015;12: 2150-2154.

    doi: 10.1109/lgrs.2015.2453119

12. Li H, Han L, Zhang L, Jia S. Seismic data reconstruction using dictionary learning and the alternating direction method of multipliers. *Geophys Prospect Pet*. 2019;58(3):419-426.

    doi: 10.3969/j.issn.1000-1441.2019.03.011

13. Wang B, Lu W, Chen X, Wang Z. Efficient seismic data interpolation using three-dimensional curvelet transform in the frequency domain. *Geophys Prospect Pet*. 2018;57(1):65-71.

    doi: 10.3969/j.issn.1000-1441.2018.01.009

14. Zhu L, Liu E, McClellan JH. Joint seismic data denoising and interpolation with double-sparsity dictionary learning. *J Geophys Eng*. 2017;14(4):802-810.

    doi: 10.1088/1742-2140/aa6491

15. Gelis LE, Gómez JL, Velis DR. Structured Noise Removal with Dictionary Learning in Seismic Data: The Effect of Atom Filtering. In: *Proceedings of the IEEE International Workshop on Remote Sensing with Intelligent Processing (RPIC)*. Salvador, Brazil: IEEE; 2019. p. 204-208.

    doi: 10.1109/rpic.2019.8882163

16. Wang W, Yang J, Huang J, Li Z, Sun M. Outlier denoising using a novel statistics-based mask strategy for compressive sensing. *Remote Sens*. 2023;15(2):447.

    doi: 10.3390/rs15020447

17. Cai J, Ji H, Shen Z, Ye G. Data-driven tight frame construction and image denoising. *Appl Comput Harmon Anal*. 2014;37:89-105.

    doi: 10.1016/j.acha.2013.10.001

18. Zhou Z, Bai M, Wu J, Cui Y. Coherent noise attenuation by kurtosis-guided adaptive dictionary learning based on variational sparse representation. *IEEE Trans Geosci Remote Sens*. 2023;61:1-10.

    doi: 10.1109/tgrs.2023.3286791

19. Wang W, Yang J, Li Z, Huang J, Zhao C. Seismic data denoising using a new framework of FABEMD-based dictionary learning. *IEEE Trans Geosci Remote Sens*. 2024;62:1-9.

    doi: 10.1109/tgrs.2024.3396459

20. Oropeza V, Sacchi M. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis (MSSA). *Geophysics*. 2011;76:V25-V32.

    doi: 10.1190/1.3552706

21. Naghizadeh M. Seismic data interpolation and denoising in the frequency-wavenumber domain. *Geophysics*. 2012;77:71-80.

    doi: 10.1190/geo2011-0172.1

22. Siahsar MA, Gholtashi S, Abolghasemi V, Chen Y. Simultaneous denoising and interpolation of 2D seismic data using data-driven non-negative dictionary learning. *Signal Process*. 2017;141:309-321.

doi: 10.1016/j.sigpro.2017.06.017

23. Yu S, Ma J, Zhang X, Sacchi MD. Interpolation and denoising of high-dimensional seismic data by learning a tight frame. *Geophysics*. 2015;80:V119-V132.

doi: 10.1190/geo2014-0396.1

24. Turquais P, Asgedom EG, Söllner W, Gelius L. Parabolic dictionary learning for seismic wavefield reconstruction across the streamers. *Geophysics*. 2018;83:V263-V282.

doi: 10.1190/geo2017-0694.1

25. Beckouche S, Ma J. Simultaneous dictionary learning and denoising for seismic data. *Geophysics*. 2014;79(3):A27-A31.

doi: 10.1190/geo2013-0382.1

26. Turquais P, Asgedom EG, Söllner W. A method of combining coherence-constrained sparse coding and dictionary learning for denoising. *Geophysics*. 2017;82(3):V137-V148.

doi: 10.1190/geo2016-0164.1

27. Siahsar MA, Gholtashi S, Torshizi EO, Chen W, Chen Y. Simultaneous denoising and interpolation of 3-D seismic data via damped data-driven optimal singular value shrinkage. *IEEE Geosci Remote Sens Lett*. 2017;14: 1086-1090.

doi: 10.1109/lgrs.2017.2697942

28. Chen W, Saad OM, Oboué YASI, Yang L, Chen Y. Retrieving the leaked signals from noise using a fast dictionary-learning method. *Geophysics*. 2021;87(1):V39-V49.

doi: 10.1190/geo2021-0243.1

29. Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*. 2006;54(11):4311-4322.

doi: 10.1109/tsp.2006.881199

30. Rubinstein R, Zibulevsky M, Elad M. Double sparsity: Learning sparse dictionaries for sparse signal approximation. *IEEE Trans Signal Process*. 2009;58(3):1553-1564.

doi: 10.1109/tsp.2009.2036477

31. Wang B, Wu R, Chen X, Li J. Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform. *Geophys J Int*. 2015;201(2):1182-1194.

doi: 10.1093/gji/ggv072

32. Almadani M, Waheed UB, Masood M, Chen Y. Dictionary learning with convolutional structure for seismic data denoising and interpolation. *Geophysics*. 2021;86(5):V361-V374.

doi: 10.1190/geo2019-0689.1

33. Li C, Wen X, Liu X, Zu S. Simultaneous seismic data interpolation and denoising based on nonsubsampled contourlet transform integrating with two-step iterative log thresholding algorithm. *IEEE Trans Geosci Remote Sens*. 2022;60:1-10.

doi: 10.1109/tgrs.2022.3192531

34. Kaplan ST, Sacchi MD, Ulrych TJ. Sparse coding for data-driven coherent and incoherent noise attenuation. In: *SEG Technical Program Expanded Abstracts*; 2009. p. 3327-3331.

doi: 10.1190/1.3255551

35. Wang H, Chen W, Zhang Q, Liu X, Zu S, Chen Y. Fast dictionary learning for high-dimensional seismic reconstruction. *IEEE Trans Geosci Remote Sens*. 2021;59(8):7098-7108.

doi: 10.1109/tgrs.2020.3030740