IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 17, 2020, accepted January 7, 2021, date of publication January 12, 2021, date of current version January 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051069

3-D Seismic Noise Attenuation via Tensor Sparse Coding With Spatially Adaptive Coherence Constraint

XIN HE[1], FENG QIAN[2], (Member, IEEE), YUQI FAN[2], (Graduate Student Member, IEEE), LINGTIAN FENG[2], (Student Member, IEEE), AND GUANGMIN HU[2]

[1]School of Resources and Environment, Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]School of Information and Communication Engineering, Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding author: Xin He (17844533663@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 41874155 and Grant U1562218.

ABSTRACT Tensor sparse coding (TSC) is a method used to excavate 3D volume structures extended by sparse coding (SC), which is increasingly applied in data noise attenuation. Existing TSC approaches control the intensity of noise attenuation by using a predetermined soft or hard threshold that relies on the noise variance. However, the noise variance in seismic data is unknown and varies with time and space, leading to the conventional TSC method not being able to track this change. To address this issue, we proposed a tensor sparse coding model with spatially adaptive coherent constraint (TSC-SAC) to find an optimal adjustable threshold without the demand for prior knowledge of the noise variance. The threshold is determined by the coherence of the residual with respect to the dictionary. Moreover, a tensor spatial coherence orthogonal matching pursuit algorithm (TSC-OMP) is developed for solving sparse representation. Unlike the existing threshold strategy in traditional TSC methods, TSC-OMP utilizes an ideal spatially adaptive coherence threshold to regulate the sparsity, which can effectively preserve the valuable information in processing for noise suppression. By comparing with four state-of-the-art denoising algorithms, we then demonstrated the superior performance of TSC-SAC on both a synthetic and two field data sets.

INDEX TERMS Tensor sparse coding, spatially adaptive coherence threshold, tensor spatial coherence orthogonal matching pursuit (TSC-OMP).

I. INTRODUCTION

Sparse coding (SC) is an unsupervised method to learn the useful structure or pattern of real data, which is widely used in signal denoising [1], [2], natural image denoising [3], [4], video denoising [5], [6] [7], and medical image denoising [8], [9] [10]. With the increasing demand for high-quality subsurface images, SC is naturally applied in seismic data denoising [11], [12]. However, since the seismic noise varies with time and space, the traditional SC methods cannot track this change due to the applied preset soft or hard threshold strategy [13], [14], thus damaging some useful signals while eliminating the noise [15]. Therefore, the essential challenge for SC is to find an adaptive threshold strategy to achieve the optimal balance between the spatio-temporal noise reduction and the preservation of detailed information.

In the seismic literature, most threshold strategies are generally based on the transform domain to regulate the data sparsity, which can be divided into the fixed-basis-based transform and the SC-based method. The fixed-basis-based transform method utilizes a linear combination of fixed basis, representing the signal sparsely by a predetermined soft or hard threshold in the transform domain, namely, Fourier transform [16], wavelet transform [17], curvelet transform [18], [19], shearlet transform [20], [21], and seislet transform [22], [23]. However, the fixed basis is determined in advance, which can lead to an unsuitable adjustment to the varying strength of seismic noise. The SC-based method encodes several basic elements as linear combinations in the dictionary domain, controlling the sparsity with a set threshold value to attenuate the seismic noise, including the

The associate editor coordinating the review of this manuscript and approving it for publication was Anubha Gupta.

K-singular value decomposition (K-SVD) [24], the data-driven tight frame (DDTF) [25], the double-sparsity dictionary learning [26], the multichannel singular spectrum analysis (MSSA) [27], and the coherence-constrained sparse coding [28]. Nevertheless, the mentioned method has difficulty in eliminating spatio-temporal noise virtually, due to the deficiency of spatial information in seismic data.

Recent research has suggested that the tensor-tensor product (t-product) can be introduced into SC models, which enables the direct data to process on high dimensional structures [29], [30], such as K-tensor singular value decomposition (K-TSVD) [31], and two-dimensional sparse coding (2DSC) [32]. However, the existing approaches rely on the preset fixed threshold strategy, leading to an inability to effectively extract useful features from seismic data [33]–[35]. In the absence of noise variance as a priori knowledge, the application of predetermined threshold strategies will cause part of the valuable information to be removed together with the seismic noise. To address this issue, we introduced a tensor space coding model, namely, tensor space coding with spatially adaptive coherence (TSC-SAC), in which the sparsity of seismic data is restricted by spatial coherence. Since our adaptive coherence threshold is independent of the noise variance, TSC-SAC can efficiently address the seismic noise varying spatio-temporally.

More specifically, the main contributions of this paper lie in four aspects. First, we introduce tensor sparse coding to extract the spatial-time structure features of seismic data. Second, by combining the tensor space coherence constraint with orthogonal matching pursuit (OMP) [36], we establish a tensor spatial coherence orthogonal matching pursuit algorithm (TSC-OMP), which we then use it as an algorithm for solving sparse representation. Third, we deduce an ideal spatial adaptive coherence threshold, which is used as a stop criterion to find the sparsity of each lateral slice. Finally, we apply TSC-SAC to synthetic and field data sets, and compare its noise attenuation effect with four state-of-the-art methods; thereby proving our model's competitive performance.

The rest of this paper is structured as follows. Section II introduces the notations and preliminaries. Section III proposes the problem statement and our model. Section IV introduces solutions to TSC-SAC. Section V presents the experimental results conducted on synthetic data and field data. We summarize this work in section VI.

II. PRELIMINARIES

A. NOTATIONS

Throughout the paper, real third-order tensors are discussed in the space $\mathbb{R}^{N_1\times N_2\times N_3}$. Tensors are represented in boldface calligraphic letters, i.e., $\mathcal{Y}$; matrices are represented in boldface capital letters, i.e., Y; vectors are represented in boldface lowercase letters, i.e., $\mathbf{y}$. The xline, inline, and time dimension of a tensor are indexed by i, j, k, respectively, and $i \in [N_1], j \in [N_2], k \in [N_3]$, where $[N]$ denotes the set $\{1, 2, \ldots, N\}$. For a third-order tensor $\mathcal{Y} \in \mathbb{R}^{N_1\times N_2\times N_3}$, we use $\mathcal{Y}(i, :, :)$, $\mathcal{Y}(:, j, :)$, and $\mathcal{Y}(:, :, k)$ to denote the horizontal, lateral, and frontal slices, respectively. Specifically, we represent $\overrightarrow{\mathcal{Y}}_j = \mathcal{Y}(:, j, :)$ and $\mathcal{Y}^{(k)} = \mathcal{Y}(:, :, k)$. Furthermore, we use $\mathcal{Y}(:, j, k)$, $\mathcal{Y}(i, :, k)$, and $\mathcal{Y}(i, j, :)$ to denote the mode-1, mode-2, and mode-3 tubes, respectively.

Additionally, we introduce the high-dimensional relations of a tensor given by the concept of the t-product, which is a natural generalization of matrix multiplication [37]. Next, we provide its definition and some related definitions.

Definition 1 (t-Product [38]): The t-product of $\mathcal{D} \in \mathbb{R}^{N_1\times r\times N_3}$ and $\mathcal{X} \in \mathbb{R}^{r\times N_2\times N_3}$ is a tensor whose (i, j)-th tube $\mathcal{Y}(i, j, :)$ is given by

$$\mathcal{Y}(i, j, :) = \sum_{q=1}^{r} \mathcal{D}(i, q, :) * \mathcal{X}(q, j, :). \tag{1}$$

Definition 2 (Tensor Transpose [38]): The transpose of a tensor $\mathcal{Y}$ is $\mathcal{Y}^\dagger \in \mathbb{R}^{N_2\times N_1\times N_3}$ obtained by transposing each of the frontal slices and then reversing the order of the transposed frontal slices 2 through N_3.

Definition 3 (Orthogonal Tensor [38]): A tensor $\mathcal{A} \in \mathbb{R}^{N_1\times N_1\times N_3}$ is orthogonal if it satisfies

$$\mathcal{A}^\dagger * \mathcal{A} = \mathcal{A} * \mathcal{A}^\dagger = \mathcal{I}, \tag{2}$$

where the identity tensor $\mathcal{I} \in \mathbb{R}^{N_1\times N_1\times N_3}$ is a tensor whose first frontal slice $\mathcal{I}(:, :, 1)$ is the $N_1 \times N_1$ identity matrix and all other frontal slices $\mathcal{I}^{(k)}$, $(k = 2, \ldots, N_3)$ are zero matrices.

Definition 4 (f-Diagonal Tensor [38]): A tensor is called f-diagonal as each frontal slice of the tensor is a diagonal matrix.

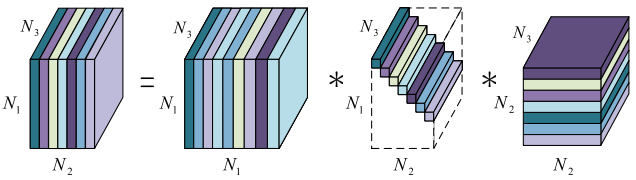

FIGURE 1. The t-SVD of an $N_1 \times N_2 \times N_3$ tensor.

Definition 5 (Tensor Singular Value Decomposition (t-SVD) [38], [39]): The t-SVD of $\mathcal{Y}$ is defined as (see Fig. 1)

$$\mathcal{Y} = \mathcal{U} * \mathcal{S} * \mathcal{V}^\dagger, \tag{3}$$

where $\mathcal{U} \in \mathbb{R}^{N_1\times N_1\times N_3}$ and $\mathcal{V}^\dagger \in \mathbb{R}^{N_2\times N_2\times N_3}$ are orthogonal tensors, and $\mathcal{S} \in \mathbb{R}^{N_1\times N_2\times N_3}$ is a f-diagonal tensor. The $\mathcal{U}$ and $\mathcal{V}$ of t-SVD obey the orthogonal structures, and the $\mathcal{S}$ satisfies the diagonal property. Note that each frontal slice of t-SVD is indeed a matrix SVD in the frequency domain.

B. TENSOR SPARSE CODING (TSC)

For the consideration of being self-contained, we summarily review the fundamental idea of the TSC model. For a third-order tensor $\mathcal{Y}$, the TSC model aims at finding an overcomplete dictionary $\mathcal{D} \in \mathbb{R}^{N_1\times r\times N_3}$ and the tensor sparse coefficient $\mathcal{X} \in \mathbb{R}^{r\times N_2\times N_3}$. The tensor dictionary $\mathcal{D}$ consists

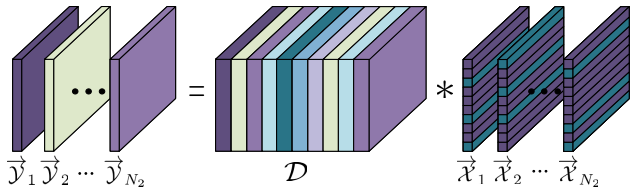

FIGURE 2. Tensor lateral slices are represented by the t-product of tensor dictionary and tubal-sparse coefficient tensors. The blue tubes in the coefficient tensors stand for the non-zero tubes and the purple ones are zero tubes.

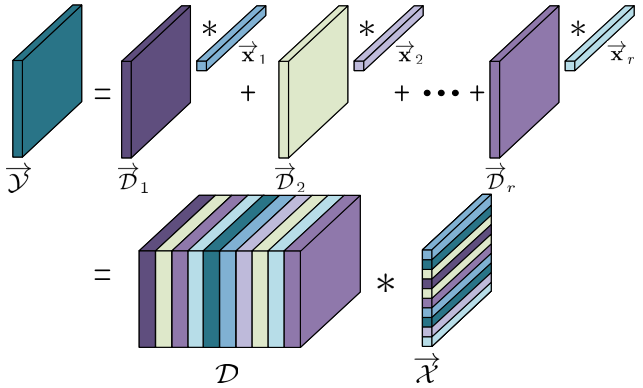

FIGURE 3. A tensor lateral slice is represented by a t-linear combination of r tensor dictionary lateral slices.

of r atoms where each lateral slice $\vec{\mathcal{D}}_j$ represents an atom. Each lateral slices of the tensor sparse coefficient $\mathcal{X}$ also consists of r tubes. After that, we give the following definitions to represent a lateral slice of the tensor $\mathcal{Y}$.

Definition 6 (Tensor-Linear Combination [31]): Given r tensor lateral slices $\vec{\mathcal{D}}_j \in \mathbb{R}^{N_1\times 1\times N_3}$, Fig. 3 illustrates a tensor signal $\vec{\mathcal{Y}} \in \mathbb{R}^{N_1\times 1\times N_3}$ using the t-linear combination of the given tensor dictionaries as follows:

$$\vec{\mathcal{Y}} = \sum_{j=1}^{r} \vec{\mathcal{D}}_j * \vec{\mathbf{x}}_j = \mathcal{D} * \vec{\mathcal{X}}, \tag{4}$$

where $\left\{\vec{\mathbf{x}}_j\right\}_{j=1}^{r}$ are tubes of size $1\times 1\times N_3$ and $\vec{\mathcal{X}} \in \mathbb{R}^{r\times 1\times N_3}$ is obtained by aligning all the $\vec{\mathbf{x}}_j$.

Definition 7 (Tensor Tubal Sparsity [31]): Let the tensor tubal sparsity $\|\cdot\|_{TS}$ denote the number of non-zero tubes of a lateral slice $\vec{\mathcal{X}}$ in the third dimension.

Invoking the above definitions, one can obtain the TSC model based on tensor tubal sparsity constraints displayed in Fig. 2 [31].

$$\begin{aligned} &\min_{\mathcal{D},\vec{\mathcal{X}}_j} \sum_{j=1}^{N_2} \left\| \vec{\mathcal{Y}}_j - \mathcal{D} * \vec{\mathcal{X}}_j \right\|_F^2, \\ &\text{s.t. } \|\vec{\mathcal{X}}_j\|_{TS} \le T, \quad j \in [N_2]. \end{aligned} \tag{5}$$

where T is the threshold dictated by the tensor tubal sparsity of the solution, and $\|\cdot\|_F$ is the Frobenius norm of a tensor, i.e., $\|\mathcal{Y}\|_F = \sqrt{\sum_{ijk} |\mathcal{Y}_{ijk}|^2}$. Note that the j-th tube $\mathcal{X}(j,1,:)$ is zero indicates that $\vec{\mathcal{D}}_j$ is not being used in the representation of $\vec{\mathcal{Y}}_j$.

Algorithm 1 TSC-SAC Algorithm

Input: $\mathcal{Y}, \mathcal{D}, \bar{\mu}, r, N_{it}$.

Initialize: Randomly initialize $\mathcal{D} \in \mathbb{R}^{N_1\times r\times N_3}$.
while $n < N_{it}$ **do**
  // Tensor coefficients Learning
  $\hat{\mathcal{X}} \leftarrow$ TSC-OMP($\mathcal{Y}, \mathcal{D}, \bar{\mu}$).
  // Tensor Dictionary Learning
  for $i = 1, 2, \cdots r$ **do**
    $\boldsymbol{\varepsilon}_i = \mathcal{Y} - \sum_{(j\neq i)} \mathcal{D}(:,j,:) * \hat{\mathcal{X}}(j,:,:), j \in [r]$,
    $w_i = \{j | \mathcal{X}(i,j,:) \neq 0, j = 1, 2, \cdots n\}$,
    $\mathcal{Z}(:,j,:) = \varepsilon(:, w_i(j), :)$,
    $[\mathcal{U}, \mathcal{S}, \mathcal{V}] =$ t-SVD $(\mathcal{Z})$,
    $\vec{\mathcal{D}}_i = \mathcal{U}(:,1,:)$.
  end for
  $\hat{\mathcal{D}} = \left\{\vec{\mathcal{D}}_i\right\}_{i=1}^{r}$.
end while

Output: $\hat{\mathcal{Y}} = \hat{\mathcal{D}} * \hat{\mathcal{X}}$.

III. PROBLEM STATEMENT AND FORMULATION

A. PROBLEM STATEMENT

In the process of seismic data acquisition, the ideal tensor is observed in the presence of random noise [28], thus the mathematical model is given by

$$\mathcal{T} = \mathcal{Y} + \mathcal{N}, \tag{6}$$

where $\mathcal{T} \in \mathbb{R}^{N_1\times N_2\times N_3}$ is the measured data containing the useful information $\mathcal{Y}$ and a noise component $\mathcal{N}$. The basic idea behind most denoising methods is to utilize the concise structure that only exists in $\mathcal{Y}$ to remove $\mathcal{N}$ [40]. Following this key idea, the SC method vectorizes the image, making a sparse approximation to estimate the useful structure where the clean image $\mathcal{Y}$ obeys the criteria [41].

However, since seismic data has high-dimensional properties, the vectorization operation of the traditional SC approach pulls down the original multidimensional structure of the volume [33]. To overcome this defect, we introduce the t-product to the SC method [29], [30], whereby equation (6) can be rewritten as

$$\mathcal{T} = \mathcal{D} * \mathcal{X} + \mathcal{N}. \tag{7}$$

To obtain the optimal tensor sparse coefficient $\mathcal{X}$, the TSC model controls the sparsity by different threshold strategies, seeking best approximations of $\mathcal{Y}$. From the perspective of the threshold strategy, the conventional TSC model can be divided into two categories, namely, the fidelity term constrained model and the tensor tubal sparsity constrained model. The fidelity term constrained model deals with problem 1, and the tensor tubal sparsity constrained model deals with problem 2, where problems 1 and 2 are as follows.

Problem 1 (The Unknown Tensor Tubal Sparsity): Since the tensor tubal sparsity is unknown, the TSC model will preset

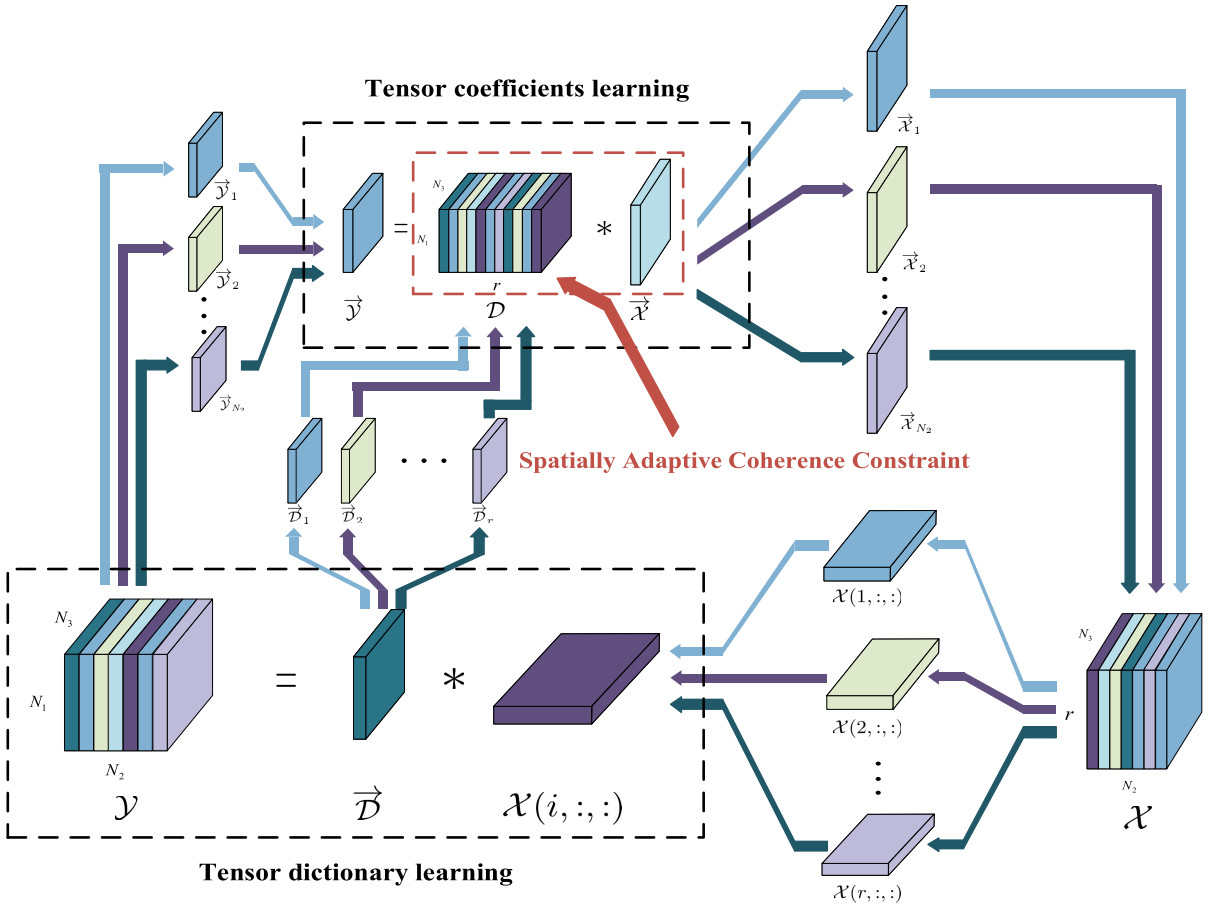

FIGURE 4. Flow diagram of the TSC-SAC model that is applied to 3-D seismic data.

a soft or hard threshold based on experience to control the fidelity term [42]

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \|\mathcal{X}\|_{TS}, \quad \text{s.t. } \|\mathcal{Y} - \mathcal{D} * \mathcal{X}\|_F^2 \leq \epsilon, \tag{8}$$

where ϵ is the error threshold ordained by the noise variance σ^2. Remarkably, the error threshold ϵ underlying the TSC model is to remove the noise while retaining the useful information as much as possible.

Problem 2 (The Unknown Noise Variance): Since the variance σ^2 of the noise is unknown, the TSC model places a hard threshold on the tensor tubal sparsity [43], [44]

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \|\mathcal{Y} - \mathcal{D} * \mathcal{X}\|_F^2, \quad \text{s.t. } \|\mathcal{X}\|_{TS} \leq T. \tag{9}$$

In terms of processing the mentioned problems, the existing TSC approach can address either problem 1 or 2. Nevertheless, problem 1 arising in seismic data noise reduction coexists with problem 2, which results in difficulty for the existing TSC model to solve it.

B. PROBLEM FORMULATION

To overcome the above difficulty, the threshold of sparsity needs to be independent of the noise variance [45]. For the consideration of satisfying this condition, [28] exploited the coherence threshold of residual $\mathcal{R}$ with respect to the dictionary $\mathcal{D}$. Extending this to the tensor version, we propose the TSC-SAC model to solve these two problems, constraining the sparsity by tensor spatially adaptive coherence.

Unlike equations (8) and (9), our threshold strategy is chosen adaptively for each slice of the seismic data, rather than using the same stop criteria for all slices [45]. Fig. 4 provides a diagrammatic representation of our proposed approach. Different from the coherence in 2-D cases [28], we use the tensor-linear combination structure to define the higher dimensional coherence as

$$\mu(\vec{\mathcal{R}}, \mathcal{D}) = \sup_{j\in[r]} \frac{\left|\left\langle \vec{\mathcal{R}}, \vec{\mathcal{D}}_j \right\rangle\right|}{\|\vec{\mathcal{R}}\|_F} = \sup_{j\in[r]} \frac{\left|\sum_{i,k} \mathcal{R}(i,1,k)\mathcal{D}(i,j,k)\right|}{\left(\sum_{i,k} \mathcal{R}(i,1,k)^2\right)^{1/2}}, \tag{10}$$

where $\vec{\mathcal{R}}$ is a slice of residual volume such that $\vec{\mathcal{R}} = \vec{\mathcal{Y}}_i - \mathcal{D} * \vec{\mathcal{X}}_i$, and $\vec{\mathcal{D}}_j$ is one of the lateral slices in dictionary $\mathcal{D}$.

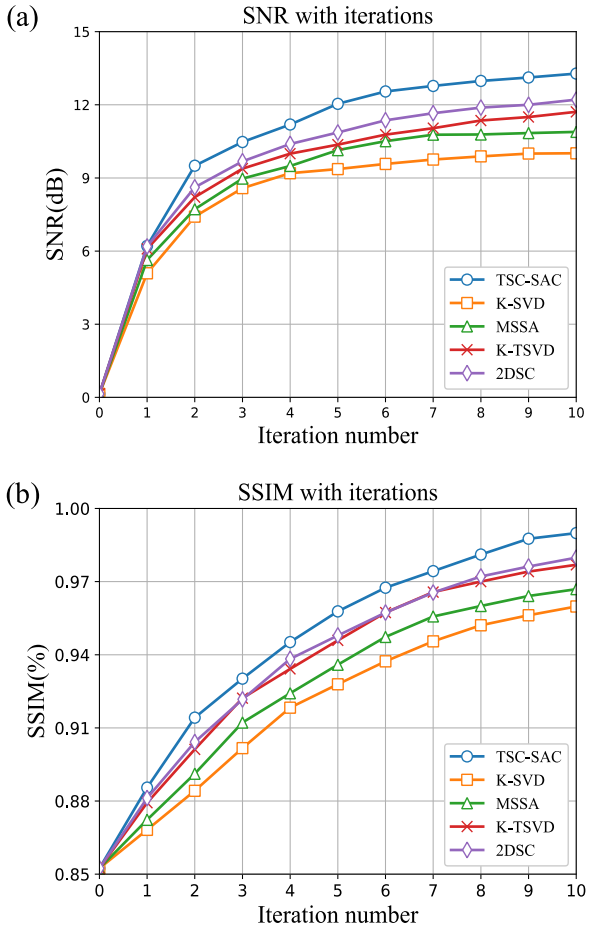

FIGURE 5. The SNR and SSIM comparison with regard to the iteration number of the proposed and existing methods under the 3-D synthetic example with noise of constant variance ("○, □, △, ×, and ◇" correspond to TSC-SAC, K-SVD, MSSA, K-TSVD, and 2DSC, respectively).

Notice that sup(·) is the minimum upper bound of a set, and $|\cdot|$ denotes the absolute value.

To constraint the tensor tubal sparsity and the tensor coherence, the objective function of the TSC-SAC is given by

$$(\hat{\mathcal{D}}, \hat{\mathcal{X}}) = \underset{\mathcal{D},\mathcal{X}}{\operatorname{argmin}} \|\mathcal{Y} - \mathcal{D} * \mathcal{X}\|_F^2, \quad \text{s.t.} \begin{cases} \min\limits_{\mathcal{X}} \|\mathcal{X}\|_{TS}, \\ \mu(\overrightarrow{\mathcal{R}}, \mathcal{D}) \le \bar{\mu}, \end{cases} \tag{11}$$

where $\bar{\mu}$ is the ideal spatially adaptive coherence threshold. In the coherent constraint term, the dictionary $\mathcal{D}$ is found at the dictionary update step of the previous iteration. [28] showed that the threshold $C(2\log(K)/N)^{1/2}$ is ideal for filtering out random noise by coherent denoising with respect to a redundant dictionary $D \in \mathbb{R}^{N\times K}$. Following the same track, we establish the derivation of $\bar{\mu}$ in the Appendix.

IV. SOLUTIONS

Due to the convolution operator and the highly underdetermined objective function, the problem in equation (11) cannot be solved directly [32]. Similiar to [24], [31], we split equation (11) into two iteration optimization stages: *1)* Obtaining the tensor coefficient $\mathcal{X}$, and *2)* Learning the tensor dictionary $\mathcal{D}$.

For the first stage, we fix the tensor dictionary $\mathcal{D}$ to learn the tensor sparse coefficient $\mathcal{X}$, and then equation (11) becomes

$$\hat{\mathcal{X}} = \underset{\mathcal{X}}{\operatorname{argmin}} \|\mathcal{Y} - \mathcal{D} * \mathcal{X}\|_F^2 + \|\mathcal{X}\|_{TS}, \quad \text{s.t. } \mu(\overrightarrow{\mathcal{R}}, \mathcal{D}) \le \bar{\mu}. \tag{12}$$

However, since equation (12) is considered as a non-deterministic polynomial-time (NP) hard problem, which is difficult to contract to a realistic seismic data size, therefore, [17] utilized the coherent matching pursuit algorithm, which uses the iterative greedy scheme of the orthogonal matching pursuit (OMP) [36]. Based on this, we propose a tensor spatial coherence orthogonal matching pursuit algorithm (TSC-OMP) approach to obtain an approximate solution in the TSC frame.

Algorithm 2 TSC-OMP Algorithm

Input: $\mathcal{Y}, \mathcal{D}, \bar{\mu}$.

function TSC-OMP$\mathcal{Y}, \mathcal{D}, \bar{\mu}$
Initialize: $\mathcal{R}^0 = \mathcal{Y}$, $\mathcal{D}^0 = \varnothing$, $\mathcal{X}^0 = 0$, and $t = 0$.
for $j = 1, 2, \cdots N_2$ **do**
  while $\mu\left(\overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_m^t\right) < \bar{\mu}$ **do**
    $\left|\left\langle \overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_m^t \right\rangle\right| = \sup\left|\left\langle \overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_n \right\rangle\right|$;
    $\hat{\mathcal{D}}_m^t = \overrightarrow{\mathcal{D}}_m^t - \sum_{t=0}^{t} \frac{\left\langle \overrightarrow{\mathcal{D}}_m^t, \hat{\mathcal{D}}_m^{t-1} \right\rangle}{\left\|\hat{\mathcal{D}}_m^{t-1}\right\|_F^2} \hat{\mathcal{D}}_m^{t-1}$;
    $\mathcal{D}^t = \mathcal{D}^{t-1} \bigcup \left\{\overrightarrow{\mathcal{D}}_m^t\right\}$;
    $\overrightarrow{\mathcal{X}}_j = \underset{\mathcal{X}_j}{\operatorname{argmin}} \left\|\overrightarrow{\mathcal{Y}}_j - \mathcal{D}^t * \mathcal{X}_j\right\|_F^2$;
    $\overrightarrow{\mathcal{R}}_j^t = \overrightarrow{\mathcal{Y}}_j - \mathcal{D}^t * \overrightarrow{\mathcal{X}}_j$;
    $t = t + 1$.
  end while
end for

Output: $\hat{\mathcal{X}} = \left\{\overrightarrow{\mathcal{X}}_j\right\}_{j=1}^{N_2}$.

To learn the tensor sparse coefficient $\mathcal{X}$, the TSC-OMP algorithm first obtains a matching dictionary set $\mathcal{D}$, and then learns the tensor sparse coefficient $\mathcal{X}$ based on $\mathcal{D}$. For clarity, we show the details of learning $\mathcal{X}$, as shown in Algorithm 2.

Firstly, in the over-complete dictionary set $\mathcal{D}_c$, we find a lateral slice $\overrightarrow{\mathcal{D}}_m$ that has the largest inner product with $\overrightarrow{\mathcal{Y}}_j$, namely

$$\left|\left\langle \overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_m^t \right\rangle\right| = \sup\left|\left\langle \overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_n \right\rangle\right|, \tag{13}$$

where $\overrightarrow{\mathcal{D}}_m^t, \overrightarrow{\mathcal{D}}_n \in \mathcal{D}_c$ and $\overrightarrow{\mathcal{D}}_m^t, \overrightarrow{\mathcal{D}}_n \notin \mathcal{D}^t$. $\overrightarrow{\mathcal{R}}_j^t$ represents the residual of the j-th lateral slice $\overrightarrow{\mathcal{Y}}_j$ after t iterations.

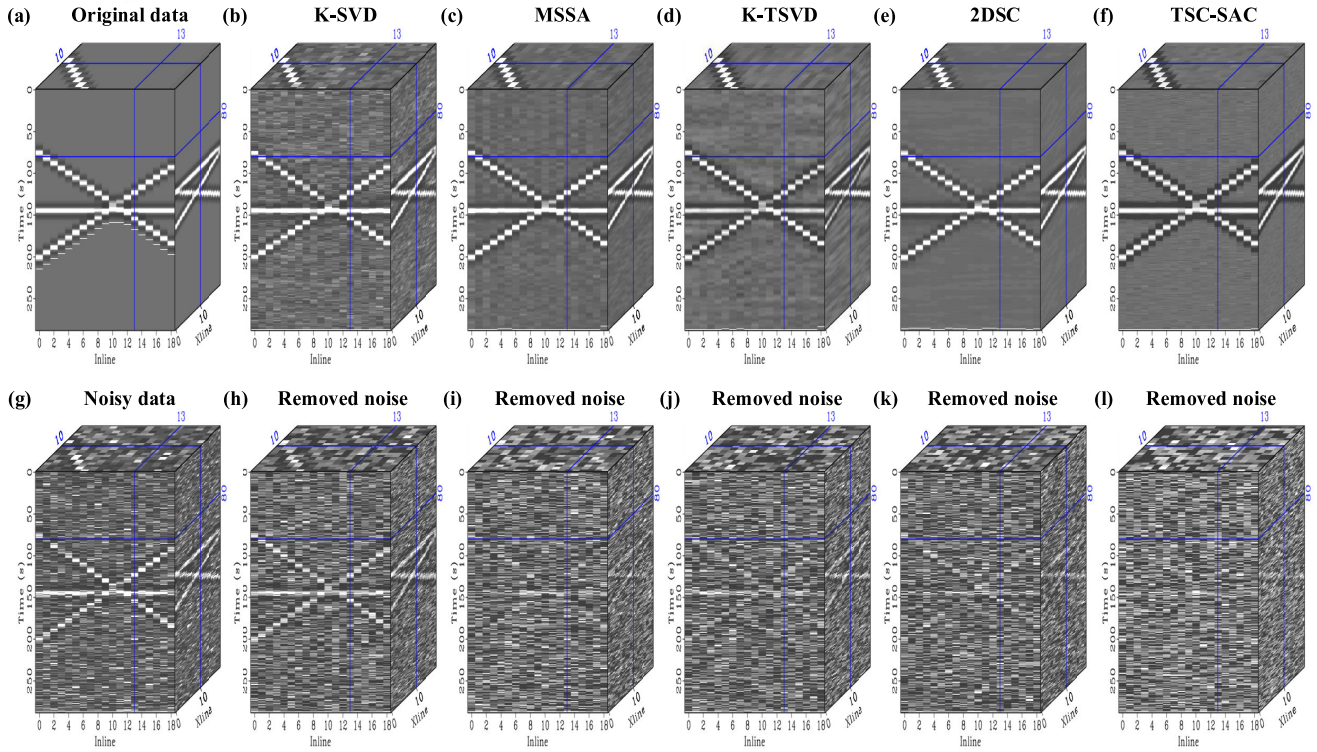

FIGURE 6. Denoising comparison of the 3-D synthetic example that adding random noise with constant variance. (a) and (g) Original noise-free data and noisy data, respectively. (b)-(f) Denoised data using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. (h)-(l) Removed noise using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively.

Secondly, we orthogonalize the selected slice $\overrightarrow{\mathcal{D}}_m^t$ with other $t-1$ selected slices

$$\hat{\mathcal{D}}_m^t = \overrightarrow{\mathcal{D}}_m^t - \sum_{t=0}^{t} \frac{\left\langle \overrightarrow{\mathcal{D}}_m^t, \hat{\mathcal{D}}_m^{t-1} \right\rangle}{\left\| \hat{\mathcal{D}}_m^{t-1} \right\|_F^2} \hat{\mathcal{D}}_m^{t-1}. \tag{14}$$

For each iteration, the selected lateral slice is placed into $\mathcal{D}^t$ as

$$\mathcal{D}^t = \mathcal{D}^{t-1} \bigcup \left\{ \hat{\mathcal{D}}_m^t \right\}, \tag{15}$$

where $\mathcal{D}^0 = \varnothing$. We then find the sparsity coefficient of the lateral slice $\overrightarrow{\mathcal{Y}}_j$ can be written as

$$\overrightarrow{\mathcal{X}}_j = \underset{\mathcal{X}_j}{\operatorname{argmin}} \left\| \overrightarrow{\mathcal{Y}}_j - \mathcal{D}^t * \mathcal{X}_j \right\|_F^2. \tag{16}$$

Naturally, we calculate the approximation error for $\overrightarrow{\mathcal{Y}}_j$:

$$\overrightarrow{\mathcal{R}}_j^t = \overrightarrow{\mathcal{Y}}_j - \mathcal{D}^t * \overrightarrow{\mathcal{X}}_j. \tag{17}$$

When the tensor spatially adaptive coherence $\mu\left(\overrightarrow{\mathcal{R}}_j^t, \overrightarrow{\mathcal{D}}_m^t\right)$ is less than $\bar{\mu}$, we can get the tensor sparse coefficient $\hat{\mathcal{X}}$ as follows:

$$\hat{\mathcal{X}} = \left\{ \overrightarrow{\mathcal{X}}_j \right\}_{j=1}^{N_2} \in \mathbb{R}^{r \times N_2 \times N_3}. \tag{18}$$

For the second stage, we use the updated sparsity coefficient to learn the tensor dictionary $\mathcal{D}$. We note that the spatially adaptive coherence threshold $\bar{\mu}$ just constrains the tensor coefficient [45], and then equation (11) becomes

$$\hat{\mathcal{D}} = \min_{\mathcal{D}} \|\mathcal{Y} - \mathcal{D} * \mathcal{X}\|_F^2. \tag{19}$$

To simplify matters, we use the K-TSVD algorithm to update $\mathcal{D}$ [31]. The optimization problem (19) can be solved as follows:

$$\overrightarrow{\mathcal{D}}_i = \underset{\mathcal{D}}{\operatorname{argmin}} \|\boldsymbol{\varepsilon}_i - \mathcal{D}(:, i, :) * \mathcal{X}(i, :, :)\|_F^2, \tag{20}$$

where $\boldsymbol{\varepsilon}_i$ represents the representation error of the dictionary apart from $\overrightarrow{\mathcal{D}}_i$. Then the problem of (20) becomes

$$\boldsymbol{\varepsilon}_i = \mathcal{Y} - \sum_{(j \neq i)} \mathcal{D}(:, j, :) * \mathcal{X}(j, :, :). \tag{21}$$

Since each horizontal of $\mathcal{X}$ employs part lateral slices in $\mathcal{D}$, we select w_i in $\mathcal{Y}$ and the corresponding tensor column to get the following variable $\mathcal{Z}_i$:

$$\mathcal{Z}(:, j, :) = \boldsymbol{\varepsilon}(:, w_i(j), :), \quad j \in [w_i], \tag{22}$$

where $w_i = \{j \mid \mathcal{X}(i, j, :) \neq 0, j = 1, 2, \cdots N_2\}$ represents the set of indices that data $\mathcal{Y}$ use dictionary slices $\mathcal{D}(:, j, :)$.

In order to minimize the error term, we apply t-SVD on $\mathcal{Z}_i$:

$$\mathcal{Z}_i = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \tag{23}$$

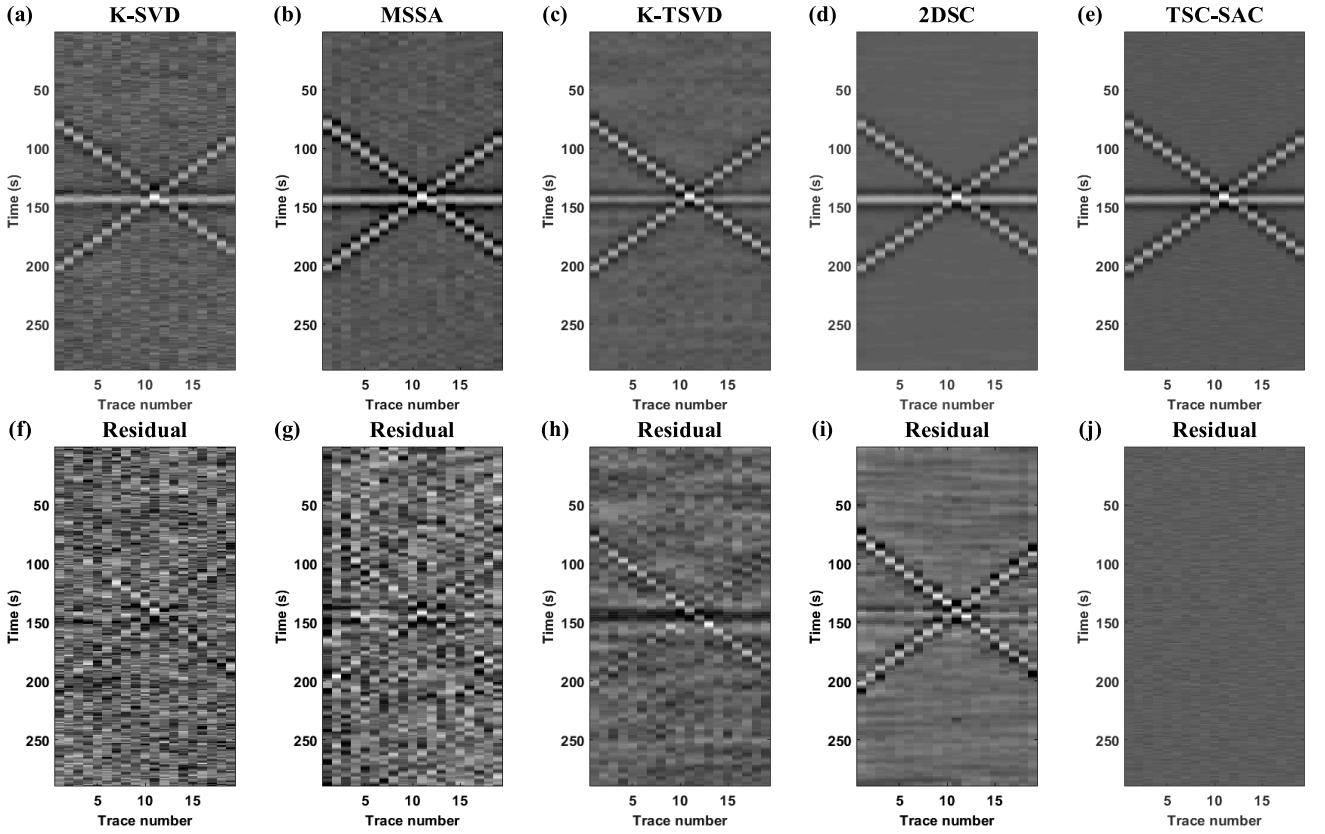

FIGURE 7. The 12th crossline section of the 3-D synthetic example. (a)-(e) Denoised data using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. (f)-(j) The corresponding residuals of K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively.

then we update $\overrightarrow{\mathcal{D}}_i$ as follows:

$$\overrightarrow{\mathcal{D}}_i = \mathcal{U}(:, 1, :), \tag{24}$$

and we can obtain the tensor dictionary $\hat{\mathcal{D}}$ as below

$$\hat{\mathcal{D}} = \left\{ \overrightarrow{\mathcal{D}}_i \right\}_{i=1}^{r} \in \mathbb{R}^{r \times N_2 \times N_3}. \tag{25}$$

Finally, we fix $\hat{\mathcal{D}}$ and optimize $\mathcal{X}$ in the next iteration until our termination conditions is satisfied.

V. PERFORMANCE EVALUATIONS

In this section, we use synthetic and field data sets, comprising post-stack and pre-stack field seismic data, for evaluating the denoising performance of our TSC-SAC model. In section V-A, we first describe the experimental setup for the entire evaluation process. Then, we compare our proposed technique with the K-SVD, MSSA, K-TSVD, and 2DSC methods using both the synthetic (see section V-B) and field examples (see section V-C).

A. EXPERIMENTAL SETUP

1) PERFORMANCE METRICS

For evaluation purposes, some metrics will be determined to measure the quality of the noise attenuation results. Given the original data $\mathcal{Y}$ and the denoised data $\hat{\mathcal{Y}}$, two quantitative assessments criteria are explicitly described as follows: one is the SNR, which is defined as

$$\text{SNR[dB]} = 20 \log_{10} \frac{\|\mathcal{Y}\|_F}{\|\mathcal{Y} - \hat{\mathcal{Y}}\|_F}, \tag{26}$$

and the other is the spatial domain SSIM index [46], in terms of the similarities structure between the reference image and distorted image, which is defined as

$$\text{SSIM}(\mathcal{Y}, \hat{\mathcal{Y}}) = \frac{\left(2\mu_{\mathcal{Y}}\mu_{\hat{\mathcal{Y}}} + c_1\right)\left(2\sigma_{\mathcal{Y}\hat{\mathcal{Y}}} + c_2\right)}{\left(\mu_{\mathcal{Y}}^2 + \mu_{\hat{\mathcal{Y}}}^2 + c_1\right)\left(\sigma_{\mathcal{Y}}^2 + \sigma_{\hat{\mathcal{Y}}}^2 + c_2\right)}, \tag{27}$$

where $\mu_{\mathcal{Y}}, \sigma_{\mathcal{Y}}$, and $\sigma_{\mathcal{Y}\hat{\mathcal{Y}}}$ are the mean, standard deviation, and cross correlation between $\mathcal{Y}$ and $\hat{\mathcal{Y}}$, respectively. Here c_1 and c_2 are used to avoid instability when the means and variances are close to zero. Note that the whole image SSIM index is averaging the local SSIM indices as calculated by a sliding window.

2) COMPARED ALGORITHMS

To reveal the superior performance of our proposed method, we compared two classes of algorithms,

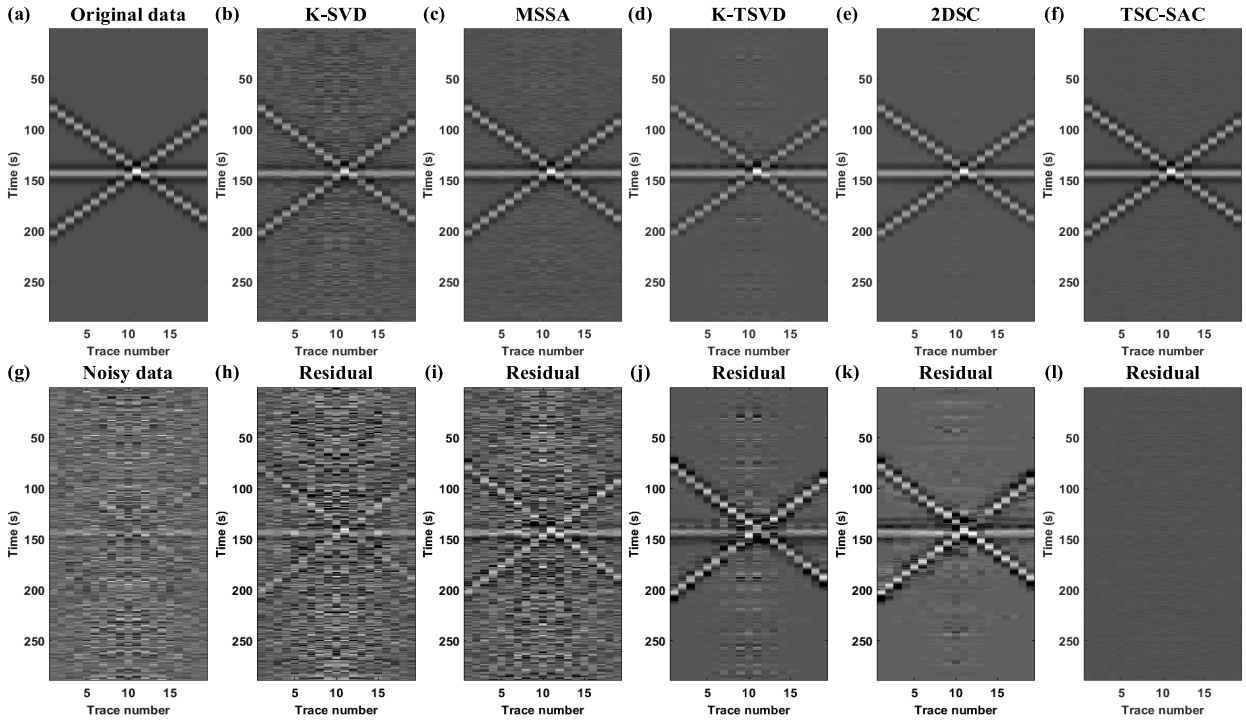

FIGURE 8. Denoising comparison of the 12th crossline section of the 3-D synthetic example that adding random noise with varying variance. (a) and (g) Original noise-free data and noisy data, respectively. (b)-(f) Denoised data using K-SVD, MSSA, K-TSVD, T2DSC, and our TSC-SAC model, respectively. (h)-(l) The corresponding residuals of K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively.

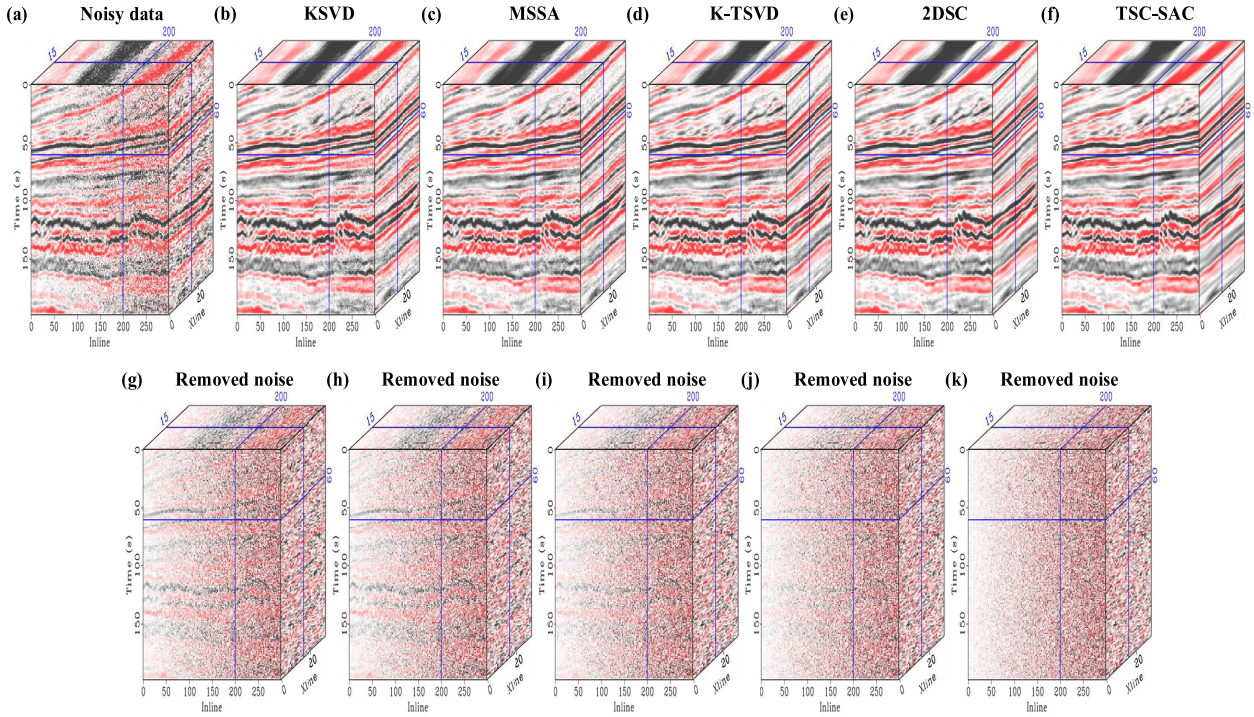

FIGURE 9. Denoising comparison of the 3-D post-stack field data $\mathcal{Y}_1$. (a) Noisy data. (b)-(f) Denoised data using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. (g)-(k) Removed noise using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively.

namely, SC and TSC methods. After considerate evaluation, we selected four state-of-the-art reference approaches that have published confirmed results and open-source codes.

- *K-SVD [24]:* Using t-SVD decomposition for sparse coding of the examples based on the current dictionary and updating the dictionary atoms alternately to fit the data better.

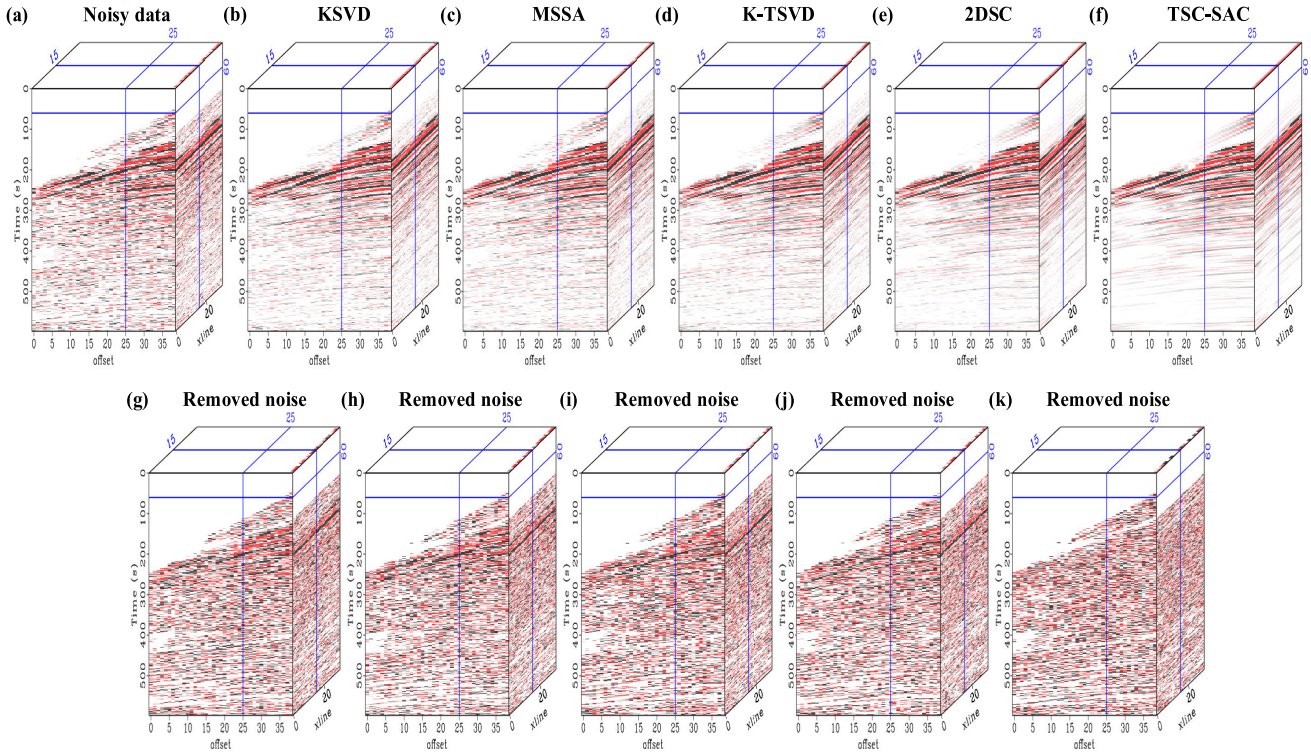

FIGURE 10. Denoising comparison of the 3-D pre-stack field data $\mathcal{Y}_2$. (a) Noisy data. (b)-(f) Denoised data using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. (g)-(k) Removed noise using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively.

- *MSSA [27]:* Decomposing the vector space of the Hankel matrix of the noisy signal into a signal and a noise subspace by truncated t-SVD.
- *K-TSVD [31]:* Using the Fourier transform and group-sparsity over multidimensional coefficient vectors to alternate between estimating a compact representation and dictionary learning.
- *2DSC [32]:* Investigating the input images as tensor-linear combinations under a novel algebraic model. It uses the circular convolution operator since the shifted versions of atoms learned by conventional SC are treated as the same ones.

In the implementation of TSC-SAC and for comparison with the approaches, the parameters for the model and the numerical algorithm need to be set as follows: We use the ideal coherence threshold $\bar{\mu}$ in our TSC-SAC method, and the error threshold ε of K-SVD, MSSA, K-TSVD, and 2DSC are set to $\sqrt{N}\sigma$, where σ is the standard deviation of Gaussian noise. Note that the maximum number of iterations $N_{it} = 10$ during our experiments.

To test our TSC-SAC method, we applied the synthetic and field data sets to verify the effectiveness of our proposed approach. Introductions and details of the synthetic and field data sets are as follows:

- *3-D synthetic data [34]:* The 3-D synthetic example is composed of three crossing linear events in the t-x-y domain, containing $N_x \times N_y$ traces in all. The temporal length of the data is 600 ms, whereas the sampling interval is 2 ms.
- *Liziba data set [47]:* The post-stack field data cover the Liziba region of the Sichuan Basin in southwestern China. The Liziba data set consists of 950 inlines by 550 crosslines for a total of nearly 50,000 CDPs, each sampled at 2 ms.
- *Avo Viking data set [48]:* A pre-stack seismic data set containing marine seismic data and well-log measurements of key elastic parameters was compiled from data acquired in an area containing hydrocarbon reservoirs.

During the experiments, we demonstrated our model under noise with constant variance and varying variance in the synthetic example sized $300 \times 19 \times 19$ separately (see Figs.6, 7, and 8). To illustrate the superior performance of our method in practice, we selected a field example $\mathcal{Y}_1$ with a size of $200 \times 31 \times 300$ from the Liziba data set (see Fig. 9). We further verified the effectiveness of the proposed TSC-SAC approach by applying it to a field example $\mathcal{Y}_2$ sized $600 \times 39 \times 31$ from the pre-stack seismic data set shown in Fig.10.

B. VALIDATION ON SYNTHETIC DATA

For noise with constant variance, the original noise-free data is displayed in Fig. 6(a). We add zero-mean white Gaussian noise to the signal, obtaining noisy data with a signal-to-noise ratio of 0.1403dB, as shown in Fig. 6(g). The denoising

results of K-SVD, MSSA, K-TSVD, 2DSC, and TSC-SAC are revealed in Figs. 6(b)-6(f), respectively. Fig. 6(h)-6(l) provides information about the removed noise corresponding to the above methods. For an even better demonstration, we selected the slice with the crossline section of the 12th to illustrate the effect in Fig. 7. Figs. 7(a)-7(e) compares K-SVD, MSSA, K-TSVD, 2DSC, and TSC-SAC in terms of the denoising results. To observe the performance of noise attenuation more intuitively, we calculated the residuals between the denoised data and the original noise-free data of each method, which are shown in Figs. 7(f)-7(j), respectively. Compared to our model, we can see apparent useful information in the residuals of K-SVD, MSSA, K-TSVD, and 2DSC. Through this comparison, in the absence of prior noise information, our proposed method demonstrate a superior performance on attenuating the noise of constant variance.

To explicate the effect of noise with varying variance, we selected the 12th crossline section of the 3-D synthetic example, as shown in Fig. 8. The original noise-free data is exemplified in Fig. 8(a). We added noise with a large variance in the middle of the synthetic data, as displayed in Fig. 8(g). Fig.8(b)-8(f) compare K-SVD, MSSA, K-TSVD, 2DSC, and TSC-SAC in terms of the denoising performance. Similar to the test of noise with constant variance, we calculated the residuals between the denoised data and the original noise-free data of each method, which are revealed in Fig.8(h)-8(l), respectively. Fig. 8(h)-8(k) illustrate the denoising aftermath of K-SVD, MSSA, K-TSVD, and 2DSC, where obvious noise remained, this is because the Frobenius norm of noise in these places is higher than the error threshold we set. For numerical purposes, the SNR and SSIM of the denoising results are shown in Fig. 5, illustrating that the performance of our proposed method is better than the other four reference approaches.

C. VALIDATION ON FIELD DATA

For the 3-D post-stack field data $\mathcal{Y}_1$, the noisy data is illustrated in Fig. 9(a). The denoised data of K-SVD, MSSA, K-TSVD, 2DSC, and TSC-SAC are shown in Fig. 9(b)-9(f), respectively. Fig. 9(g)-9(k) show the removed noise corresponding to the above methods. To highlight the denoising result, we chose the slice with the crossline section of the 9th to expose the effect in Fig. 11. The black boxes show that the denoising upshot of K-SVD, MSSA, K-TSVD, and 2DSC is that obvious noise remained, as can be seen in Fig. 11(a), Fig.11(c), Fig.11(e), and Fig.11(g). In numerical, we computed the variance of the two noise data blocks with the size of $20 \times 20 \times 20$ as shown in Fig. 11(b) where the variance of σ_2 is 2.6 times larger than the variance of σ_1. This demonstrate that TSC-SAC can attenuate the noise of the varying variance in the actual 3-D post-stack data. The black boxes in Fig. 11(b), Fig.11(d), Fig.11(f), and Fig.11(h) clarify that an apparent part of the useful information is contained in the removed noise of K-SVD, MSSA, K-TSVD, and 2DSC. In the case where the noise variance is varying, this leads us to the conclusion that the proposed approach can achieve better denoising results than the other four methods.

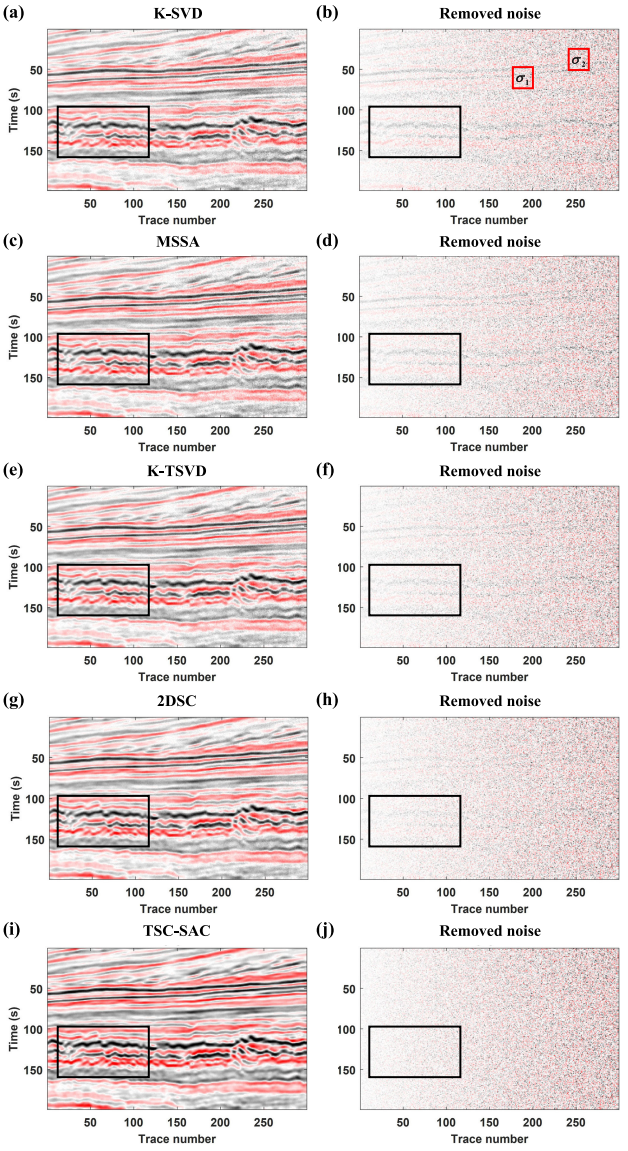

FIGURE 11. The 9th crossline section of the 3-D post-stack field data $\mathcal{Y}_1$. The left column is the denoising performance using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. The right column is the residual corresponding to the left column.

Furthermore, we tested the denoising performances on the 3-D pre-stack field noisy data $\mathcal{Y}_2$ as indicated in Fig. 10(a). Fig. 10(b)-10(f) compare K-SVD, MSSA, K-TSVD, 2DSC, and TSC-SAC in terms of the denoising results. The removed noise corresponding to the above methods are shown in Fig. 10(g)-10(k), respectively. For an even better demonstration, we selected the slice with the crossline section of the 15th to illustrate the effect in Fig. 12. The black boxes show that the denoising upshot of K-SVD, MSSA, K-TSVD, and 2DSC is that noise remained, as can be seen in Fig. 12(a), Fig. 12(c), Fig. 12(e), and Fig. 12(g). The black boxes in Fig. 12(b), Fig. 12(d), Fig. 12(f), and Fig. 12(h)

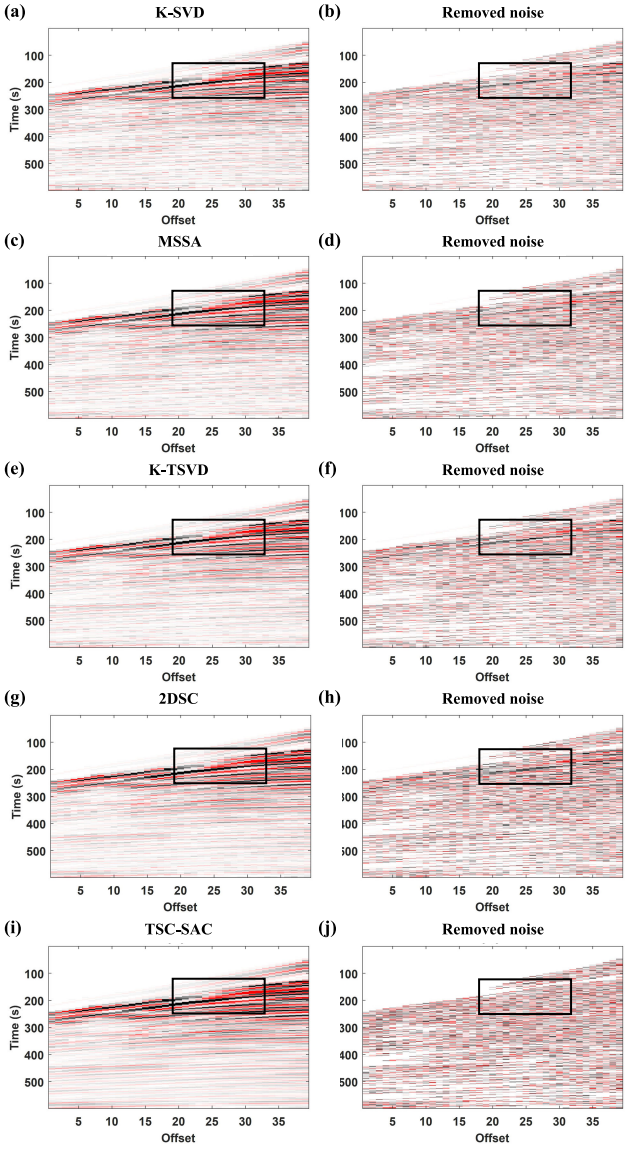

FIGURE 12. The 15th crossline section of the 3-D pre-stack field data $\mathcal{Y}_2$. The left column is the denoising performance using K-SVD, MSSA, K-TSVD, 2DSC, and our TSC-SAC model, respectively. The right column is the residual corresponding to the left column.

exhibit that the removed noise of K-SVD, MSSA, K-TSVD, and 2DSC still contained useful information. This show that the denoising performance of TSC-SAC is superior to the other four methods when the noise variance is varying.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed a TSC-SAC model to remove seismic data noise, whose adaptive threshold controls the sparsity without the requirement for noise variance. With the constraint of spatially adaptive coherence, our model can find an adaptive threshold strategy to reach the optimal balance between the spatio-temporal noise reduction and the preservation of detailed information. Then, we developed a TSC-OMP algorithm to solve the tensor coefficient. Furthermore, we derived an ideal spatially adaptive coherence threshold, which is used as a stop criterion to find the sparsity of each lateral slice. Compared with some well-known noise reduction methods, such as K-SVD, MSSA, K-TSVD, and 2DSC, we used a synthetic example and two field examples to illustrate the superior performance of TSC-SAC in terms of the denoising results.

In future work, we will further investigate the potential of our proposed method in seismic data reconstruction [49], [50], and seismic facies analysis [51]. Meanwhile, some follow-up works can apply the TSC-SAC model to video processing [5]–[7], and medical image processing [8]–[10].

APPENDIX
DERIVATION OF IDEAL SPATIALLY ADAPTIVE COHERENCE THRESHOLD

With respect to the over-complete dictionary $\mathcal{D}$, we derive the ideal upper bound $\bar{\mu}$ for coherence of the residual signal $\mathcal{R}$:

$$\begin{aligned}\mu(\overrightarrow{\mathcal{R}}, \mathcal{D}) &= \sup_{j\in[r]} \frac{\left|\left\langle \overrightarrow{\mathcal{R}}, \overrightarrow{\mathcal{D}}_j \right\rangle\right|}{\|\overrightarrow{\mathcal{R}}\|_F}, \\ &= \sup_{j\in[r]} \frac{\left|\sum_i^{N_1}\sum_k^{N_3} \mathcal{R}(i,1,k)\mathcal{D}(i,j,k)\right|}{\sqrt{\left(\sum_i^{N_1}\sum_k^{N_3} \mathcal{R}(i,1,k)^2\right)}}, \\ &= \sup_{j\in[r]} \left|\sum_i^{N_1}\sum_k^{N_3} \frac{\mathcal{R}(i,1,k)}{\|\mathcal{R}(i,1,k)\|_F}\mathcal{D}(i,j,k)\right|, \end{aligned} \tag{28}$$

where the normalized residual signal $\mathcal{R}(i,1,k)/\|\mathcal{R}(i,1,k)\|_F$ follows the Gaussian distribution:

$$\frac{\mathcal{R}(i,1,k)}{\|\mathcal{R}(i,1,k)\|_F} \sim N\left(0, \frac{1}{N_1 \times N_3}\right), \tag{29}$$

on the basis of the linear combination in [52], the projection of the normalized residual signal on a dictionary lateral slice follows the distribution:

$$\frac{\mathcal{R}(i,1,k)}{\|\mathcal{R}(i,1,k)\|_F}\mathcal{D}(i,j,k) \sim N\left(0, \frac{1}{N_1 \times N_3}\right). \tag{30}$$

The coherence of the residual signal with regard to the dictionary can be reformulated as the maximum of the absolute value of N_2 dependent [53], namely:

$$\sup_{j\in[N_2]} \left|\sum_i^{N_1}\sum_k^{N_3} \frac{\mathcal{R}(i,1,k)}{\|\mathcal{R}(i,1,k)\|_F}\mathcal{D}(i,j,k)\right| \leq \sqrt{\frac{2\log(N_2)}{\sigma^2}}, \tag{31}$$

combining (30) and (31), the coherence between the residual signal $\mathcal{R}$ and the dictionary $\mathcal{D}$ is bounded:

$$\mu(\mathcal{R}, \mathcal{D}) \leq \sqrt{\frac{2\log(N_2)}{N_1 \times N_3}}. \tag{32}$$

REFERENCES

[1] H. Bristow, A. Eriksson, and S. Lucey, "Fast convolutional sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 391–398.

[2] Q. Barthélemy, A. Larue, A. Mayoue, D. Mercier, and J. I. Mars, "Shift & 2D rotation invariant sparse coding for multivariate signals," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1597–1611, Apr. 2012.

[3] M. Zheng, J. Bu, C. Chen, C. Wang, L. Zhang, G. Qiu, and D. Cai, "Graph regularized sparse coding for image representation," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1327–1336, May 2011.

[4] K. Singh, D. K. Viswakarma, G. S. Walia, and R. Kapoor, "Sparse coding based robust image denoising via coupled dictionary," in *Proc. 1st IICIP*, Aug. 2016, pp. 1–4.

[5] B. Mallik, A. Sheikh-Akbari, and A.-L. Kor, "HEVC based mixed-resolution stereo video codec," *IEEE Access*, vol. 6, pp. 52691–52702, Sep. 2018.

[6] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[7] G. Tech, Y. Chen, K. Müller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 35–49, Jan. 2016.

[8] A. S. Panayides, A. Amini, N. D. Filipovic, A. Sharma, S. A. Tsaftaris, A. Young, D. Foran, N. Do, S. Golemati, T. Kurc, K. Huang, K. S. Nikita, B. P. Veasey, M. Zervakis, J. H. Saltz, and C. S. Pattichis, "AI in medical imaging informatics: Current challenges and future directions," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 1837–1857, Jul. 2020.

[9] G. Wang, W. Li, M. A. Zuluaga, R. Pratt, P. A. Patel, M. Aertsen, T. Doel, A. L. David, J. Deprest, S. Ourselin, and T. Vercauteren, "Interactive medical image segmentation using deep learning with image-specific fine tuning," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018.

[10] A. S. Panayides, M. S. Pattichis, M. Pantziaris, A. G. Constantinides, and C. S. Pattichis, "The battle of the video codecs in the healthcare domain—A comparative performance evaluation study leveraging VVC and AV1," *IEEE Access*, vol. 8, pp. 11469–11481, Dec. 2020.

[11] L. Liu and J. Ma, "Structured graph dictionary learning and application on the seismic denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1883–1893, Apr. 2019.

[12] Y. Li, C. Xu, C. Chen, H. Yin, L. Yi, and X. He, "Adaptive denoising approach for high-rate GNSS seismic waveform preservation: Application to the 2010 EI Mayor-Cucapah earthquake and 2012 Brawley seismic swarm," *IEEE Access*, vol. 7, pp. 173166–173184, Nov. 2019.

[13] M. Sun, Z. Li, Z. Li, Q. Li, Y. Liu, and J. Wang, "A noise attenuation method for weak seismic signals based on compressed sensing and CEEMD," *IEEE Access*, vol. 8, pp. 71951–71964, Mar. 2020.

[14] S. Beckouche and J. Ma, "Simultaneous dictionary learning and denoising for seismic data," *Geophysics*, vol. 79, no. 3, pp. A27–A31, May 2014.

[15] L. Jia, S. Song, L. Yao, H. Li, Q. Zhang, Y. Bai, and Z. Gui, "Image denoising via sparse representation over grouped dictionaries with adaptive atom size," *IEEE Access*, vol. 5, pp. 22514–22529, Oct. 2017.

[16] M. D. Sacchi, T. J. Ulrych, and C. J. Walker, "Interpolation and extrapolation using a high-resolution discrete Fourier transform," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 31–38, Jan. 1998.

[17] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.

[18] E. Candes, "Curvelets—A surprisingly effective nonadaptive representation for objects with edges," *Curves Surf. Fitting Saint Malo*, 2000.

[19] E. J. Candès and D. L. Donoho, "Recovering edges in ill-posed inverse problems: Optimality of curvelet frames," *Ann. Statist.*, vol. 30, no. 3, pp. 784–842, Jun. 2002.

[20] H. Shan, J. Ma, and H. Yang, "Comparisons of wavelets, contourlets and curvelets in seismic denoising," *J. Appl. Geophys.*, vol. 69, no. 2, pp. 103–115, Aug. 2009.

[21] S. Liu, M. Liu, P. Li, J. Zhao, Z. Zhu, and X. Wang, "SAR image denoising via sparse representation in shearlet domain based on continuous cycle spinning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2985–2992, May 2017.

[22] S. Fomel, Y. Liu, and J. University, "Seislet transform and seislet frame: Tools for compressive representation of seismic data," in *Proc. 73th Eage Conf. Exhib.*, May 2011.

[23] Y. Liu and S. Fomel, "OC-seislet: Seislet transform construction with differential offset continuation," *Geophysics*, vol. 75, no. 6, pp. WB235–WB245, Nov. 2010.

[24] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[25] J.-F. Cai, H. Ji, Z. Shen, and G.-B. Ye, "Data-driven tight frame construction and image denoising," *Appl. Comput. Harmon. Anal.*, vol. 37, no. 1, pp. 89–105, Jul. 2014.

[26] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, Mar. 2010.

[27] 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.

[28] P. Turquais, E. G. Asgedom, and W. Söllner, "A method of combining coherence-constrained sparse coding and dictionary learning for denoising," *Geophysics*, vol. 82, no. 3, pp. V137–V148, May 2017.

[29] N. Qi, Y. Shi, X. Sun, and B. Yin, "TenSR: Multi-dimensional tensor sparse representation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 5916–5925.

[30] N. Qi, Y. Shi, X. Sun, J. Wang, and B. Yin, "Two dimensional synthesis sparse model," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2013, pp. 1–6.

[31] Z. Zhang and S. Aeron, "Denoising and completion of 3D data via multidimensional dictionary learning," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2016, pp. 2371–2377.

[32] F. Jiang, X.-Y. Liu, H. Lu, and R. Shen, "Efficient two-dimensional sparse coding using tensor-linear combination," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, Mar. 2017, pp. 1–7.

[33] S. Liao, X.-Y. Liu, F. Qian, M. Yin, and G.-M. Hu, "Tensor super-resolution for seismic data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 8598–8602.

[34] D. Zhang, Y. Zhou, H. Chen, W. Chen, S. Zu, and Y. Chen, "Hybrid rank-sparsity constraint model for simultaneous reconstruction and denoising of 3D seismic data," *Geophysics*, vol. 82, no. 5, pp. V351–V367, Sep. 2017.

[35] N. Kreimer and M. D. Sacchi, "A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation," *Geophysics*, vol. 77, no. 3, pp. V113–V122, May 2012.

[36] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput.*, Nov. 1993, pp. 40–44.

[37] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, Jan. 2013.

[38] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[39] F. Qian, C. Zhang, L. Feng, C. Lu, G. Zhang, and G. Hu, "Tubal-sampling: Bridging tensor and matrix completion in 3-D seismic data reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 854–870, Jan. 2021.

[40] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.

[41] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[42] R. Zhao, J. Fu, L. Ren, and Q. Wang, "Atom-refined multiway greedy algorithm for tensor-based compressive sensing," *IEEE Access*, vol. 7, pp. 23038–23054, Feb. 2019.

[43] S. Pei, L. Li, L. Ye, and Y. Dong, "A tensor foreground-background separation algorithm based on dynamic dictionary update and active contour detection," *IEEE Access*, vol. 8, pp. 88259–88272, May 2020.

[44] L. Zhai, Y. Zhang, H. Lv, S. Fu, and H. Yu, "Multiscale tensor dictionary learning approach for multispectral image denoising," *IEEE Access*, vol. 6, pp. 51898–51910, Sep. 2018.

[45] X. Wang and J. Ma, "Adaptive dictionary learning for blind seismic data denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1273–1277, Jul. 2020.

[46] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[47] L. Wei, Z. Kun, and T. Jiali, "The reservoir characteristics and main control factors of Maokou formation in Lujiaochang-Liziba area," *J. Yangtze Univ., Natural Sci. Ed.*, no. 2, p. 2, 2015.

[48] R. G. Keys and D. J. Foster, "A data set for evaluating and comparing seismic inversion methods," in *Comparison of Seismic Inversion Methods on a Single Real Data Set*. 1998, pp. 1–12.
[49] N. Kreimer, A. Stanton, and M. D. Sacchi, "Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction," *Geophysics*, vol. 78, no. 6, pp. V273–V284, Nov. 2013.
[50] Y. Chen, M. Bai, and Y. Chen, "Obtaining free USArray data by multi-dimensional seismic reconstruction," *Nature Commun.*, vol. 10, no. 1, pp. 1–13, Sep. 2019.
[51] K. Li, B. She, Z. Liu, M. Su, and G. Hu, "Supervised seismic facies analysis using discrimination dictionary," in *Proc. 88th Annu. Int. Meeting, SEG Tech. Program Extended Abstr.*, Aug. 2018, pp. 1585–1589.
[52] B. Eisenberg and R. Sullivan, "Why is the sum of independent normal random variables normal?" *Math. Mag.*, vol. 81, no. 5, pp. 362–366, Dec. 2008.
[53] J. A. Hartigan, "Bounding the maximum of dependent random variables," *Electron. J. Statist.*, vol. 8, no. 2, pp. 3126–3140, 2014.

XIN HE received the B.E. degree in communication engineering from Nanchang University, China, in 2017. He is currently pursuing the M.E. degree in electronics and communication with the University of Electronic Science and Technology of China, Chengdu, China.

His research interests include seismic data processing and machine learning.

FENG QIAN (Member, IEEE) received the B.S. and M.S. degrees from Sichuan University, Chengdu, China, in 2001 and 2004, respectively, and the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, in 2008.

Since 2009, he has been with the School of Information and Communication Engineering and also with the Center for Information Geoscience, University of Electronic Science and Technology of China, where he is currently an Associate Professor. From August 2014 to August 2015, he was also a Visiting Scholar with Columbia University, New York, NY, USA. His current research interests include tensor theory, non-convex optimization, deep learning, big data analysis, and their applications in seismic data processing.

YUQI FAN (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from the University of Electronic Science and Technology, Chengdu, China, in 2017, where she is currently pursuing the master's degree in communication engineering.

Her research interests include big data analysis and seismic data processing.

LINGTIAN FENG (Student Member, IEEE) received the B.E. degree in communication engineering from Hainan University, China, in 2018. He is currently pursuing the M.E. degree in electronics and communication from the University of Electronic Science and Technology of China.

His research interests include seismic data processing and transfer learning.

GUANGMIN HU received the B.S. degree from the Department of Computer Science, Nanjing University, Nanjing, China, in 1986, and the M.S. and Ph.D. degrees in geophysical prospecting from the Chengdu University of Technology, Chengdu, China, in 1992 and 2000, respectively.

He is currently a Full Professor and the Dean of the School of Resources and Environment with the University of Electronic Science and Technology of China, Chengdu, where he was a Postdoctoral Researcher with the School of Communication and Information Engineering, from 2000 to 2003. From 2002 to 2003, he was also a Visiting Scholar with Hong Kong Polytechnic University, Hong Kong. As the principle investigator, he has completed over 50 scientific research projects, including the National Natural Science Foundation of China (General and Major Program), the 973 Program, and other crosswise projects. He was supported by the Program for New Century Excellent Talents in University, Ministry of Education, China. He has published about 100 articles indexed by SCI or EI. His current research interests include seismic data processing and signal processing.

Dr. Hu was a recipient of the Excellent Young Academic Leaders and the Excellent Scientists with Outstanding Scientific Contributions in Sichuan Province, the State Science and Technology Progress Award (Third Class once), the Science and Technology Progress Award (First Class once), Ministry of Education, China, and the Science and Technology Award (Second Class twice and Third Class once), Ministry of Land and Resources, China.