Research on Multichannel Signals Fault Diagnosis for Bearing via Generalized Non-Convex Tensor Robust Principal Component Analysis and Tensor Singular Value Kurtosis

MAO GE\(^1\,^2\), YONG LV\(^1\,^2\) AND YUBO MA\(^1\,^2\)

\(^1\)Key Laboratory of Metallurgical Equipment and Control Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China
\(^2\)Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

Corresponding author: Yong Lv (lvyong@wust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51875416, and in part by the Natural Science Foundation Innovation Group Program of Hubei Province under Grant 2020CFA033.

ABSTRACT Tensor decomposition technique has been widely used in multichannel signals processing for its distinct superiority. The early fault feature signal of bearing is weak and easily inundated by ambient noise. In addition, the interference signals generated by other mechanical components also have a serious impact on the result of fault diagnosis. Aiming at the above issues, and built on the methods of generalized non-convex tensor robust principal component analysis (GNCTRPCA) and tensor singular value kurtosis (TSVK), this paper introduces a new fault diagnosis technique for multichannel bearing signals. First, the attractor tensor is formed by reconstructing the acquired multichannel signals in phase space. Second, based on tensor singular value decomposition (TSVD), the tensor robust principal component analysis (TRPCA) can provide a favorable noise reduction performance. However, TRPCA usually lowers the amplitude of useful singular value tubes (SVTs). To tackle this problem, the GNCTRPCA method is proposed to avoid the amplitude reduction. Third, a new TSVK method is introduced to determine the reconstructed order of SVTs, so as to extract the multichannel fault feature signals. Finally, the bearing fault type can be identified by comparing the peak frequencies of the extracted signals with the theoretical fault-related frequencies. Simulation analysis and experiment studies verify the principle and effectiveness of the proposed technique.

INDEX TERMS Multichannel bearing signals fault diagnosis, tensor singular value decomposition, generalized non-convex tensor robust principal component analysis, tensor singular value kurtosis.

I. INTRODUCTION

As the frequently used components in mechanical equipment, bearings are vulnerable to potential faults under harsh working environments, including high temperature, heavy loads and long online service and so on [1], [2]. Bearing fault will affect the production efficiency, resulting in heavy production loss [3]. Hence, it is of great significance to make accurate detection and diagnosis in the early stage of bearing fault [4].

The faulty bearing can produce periodic impact during operation and the corresponding vibration signal shows strong nonlinear and non-stationary characteristics [5]. In the real industrial environment, the early fault feature signal is extremely weak due to the weak modulation source and easily inundated by the strong ambient noise [6]. Besides, the interference signals generated by other mechanical components may also seriously affect the result of fault diagnosis. Therefore, how to effectively extract the weak fault feature signal from complex mixed signals is the core issue of bearing fault diagnosis research.

In the field of bearing fault diagnosis, a variety of vibration signal analysis methods have been proposed for their capabilities of dynamic information analysis. e.g., wavelet transform (WT) [7], empirical mode decomposition (EMD) [8], synchrosqueezed wavelet transform (SWT) [9], envelope analysis [10], manifold learning [11] and so on.
Envelope analysis is a commonly used amplitude demodulation method in engineering signal analysis, which can separate the low-frequency fault characteristic components from the signal. Borghesani et al. [12] used envelope analysis to realize the bearing fault diagnosis of rolling element under variable operating conditions. However, this method is generally sensitive to strong background noise. Manifold learning is a classical nonlinear dimension reduction analysis method for the feature extraction of vibration signals, which can extract the low-dimensional submanifold representing the feature component of the signal from high-dimensional space. The typical dimensionality reduction algorithms include isometric feature mapping (ISOMAP), locally linear embedding (LLE) and local tangent space alignment algorithm (LTSA) [13]. Su et al. [14] proposed a multifault diagnosis method based on orthogonal supervised LTSA and least square support vector machine for the bearing. However, the dimension of low-dimensional submanifold and the neighborhood of the sample points are hard to be defined. In addition, the anti-noise performance of manifold learning is also insufficient. These methods are generally only suitable for single channel signal processing. Compared with the single channel signal, the multichannel signals acquired from vibration sensors in different directions and positions can provide more abundant information about the operating status of critical equipment. And the multichannel signal processing technology can effectively improve the reliability and accuracy of diagnosis, which has become more and more important and hot in signal processing. Aminghafari et al. [15] proposed a multivariate wavelet denoising method (MWD) for the multichannel signal denoising, but the noise reduction capacity is limited. The independent component analysis (ICA) [16], as a method based on high-order statistical properties, can effectively separate the multicomponent source features from the multichannel mixed signals. Guo et al. [17] used ICA to realize the fault diagnosis of rolling element bearing. However, ICA has two defects, which are noise sensitivity and the source signal’s statistical independence. A novel multivariable EMD method (MEMD) is introduced in [18] for early weak fault feature of bearing. However, MEMD always suffers from the mode aliasing and endpoint effect. Since the instantaneous frequency ridge extracted from time–frequency plane in SWT can reflect the signal oscillation feature information, a multisensor matching SWT method (MMSWT) for mechanical fault condition assessment is proposed in [19]. Nevertheless, the noise can greatly reduce the resolution of time–frequency plane, resulting in the decline of the accuracy of ridge extraction. These contributions enrich the research of multichannel signals fault diagnosis.

As the most natural expression of the whole structure of multidimensional data, high dimensional tensors can keep the data’s inherent structural characteristics to the greatest extent [20]. With the application of multidimensional linear algebra and its powerful function, the data processing technology via tensor decomposition [21], [22] is widely used in multichannel or multisensor signal data’s processing.

As a nonlinear filtering method, the canonical polyadic decomposition (CPD) [23] can effectively extract the useful feature information of tensor data, which is suitable for signal separation and feature extraction. Cheng et al. [24] used CPD to obtain the useful feature from the tensor composed of multiple statistical index values of multichannel signals for the classification of gear faults. In [25], the three dimensional tensor composed from the time-frequency features of the multichannel signals is decomposed by CPD for the identification of bearing compound faults. However, there are two main problems need to be discussed. The first is the applicability of the constructed high dimensional tensor. The tensors in above methods are generally constructed by extracting the relevant information of multichannel signals in the original phase space. In fact, the acquired signal time series is generally considered as the distorted projection of system chaotic motion of high-dimensional phase space in low-dimensional space-time [26]. Thus, it is difficult to show the dynamic characteristics of nonlinear and non-stationary systems in the original phase space. Making a time delay reconstruction of signal phase space to view its dynamics is an effective way. So, the phase space reconstruction technique is employed into this paper for the construction of tensor [27]. Another noteworthy issue is that the improper selection of CPD rank may lead to inaccurate extraction of fault feature, which is actually a nondeterministic polynomial (NP) hard problem in mathematics [28]. In view of this issue, reference [29] introduces a new tensor decomposition model named tensor singular value decomposition (TSVD) and [30] defines a new tensor tubal rank. Similar to matrix SVD [31], the singular value tubes (SVTs) decomposed by TSVD can represent singular subspace of different feature component or noise components in the raw signals. Hence, by selecting an appropriate reconstructed order of SVTs, the relatively clean fault feature signals can be extracted. But there are still two issues need to be solved in the application of TSVD.

The first issue is the influence of noise on the result of fault feature extraction. In fact, the random noise is widely distributed in all orders of SVTs [6]. As such, there is an inevitable noise component in the extracted fault feature signals, which is not conducive for the diagnosis of early weak fault. Thus, eliminating the noise weight in SVTs is an urgent and necessary work. Based on TSVD, the proposed tensor robust principal component analysis (TRPCA) [32] method can provide a favorable noise reduction performance for tensor. According to the manifold learning theory [13], TRPCA believes that the feature component in the attractor tensor is distributed in a low-dimensional submanifold region of the high-dimensional phase space, and this region usually has a low-tubal rank structure. Besides, the noise component in the attractor tensor is generally considered to be sparse. Both of these components can be separated by solving a typical convex optimization procedure about a joint minimization of tensor nuclear norm and penalty function regularization. However, since TRPCA adopts soft threshold [33] to evenly
lower the amplitudes of all SVTs, the energy of the extracted fault feature signals is drastically reduced, which is disastrous for the diagnosis of early weak fault. Recently, Lu et al. [34] proposed an interesting generalized non-convex low-rank minimization method for matrix low-rank restoration in image processing. The problem of amplitude reduction can be effectively avoided by performing a non-convex relaxation on the matrix nuclear norm constraint. Therefore, this paper extends this method into the denoising of tensor.

The other issue is the selection of reconstructed order, which may influence the feature extraction result. It is generally accepted that the first few orders of SVTs dominate the fault feature information [32]. Currently, the selection criterion mainly depends on experience [31] or finding the turning point of SVTs, such as difference spectra [35]. However, the interference features can affect the determination of the turning point. Kurtosis is a numerical statistic reflecting the distribution characteristics of random variables and impressible to impact signals, which can be used to identify early bearing fault [36]. Therefore, employing kurtosis to determine the optimal reconstructed order is a feasible way.

In view of the above analysis, this paper introduces a new fault diagnosis technique for multichannel bearing signals. First, the attractor tensor is constructed via time-delay of the acquired multichannel signals in phase space. Then, the tensor is decomposed to obtain all the SVTs by TSVK.

The rest of the paper is organized as follows: Section II introduces some background theory. Section III describes the proposed fault diagnosis technique for multichannel bearing signals. The simulation analysis in Section IV and experiment analysis in Section V verify the effectiveness of the proposed technique. Section VI draws the conclusions.

II. BACKGROUND THEORY

A. NOTATION

Throughout this paper, vector and matrix are denoted as \( \mathbf{x} \in \mathbb{R}^{n_1} \) and \( \mathbf{X} \in \mathbb{R}^{n_1 \times n_2} \), respectively. Fig. 1 (a) shows a three dimensional tensor with the notation of \( \mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3} \) and set \( n = \min(n_1, n_2) \). We use Matlab notation \( \mathbf{X}(i: : ,) \), \( \mathbf{X}(; i: ,) \), \( \mathbf{X}(;: i, :) \) to represent the i-th horizontal, lateral and frontal slices, and \( \mathbf{X}(i; j: :) \) to represent the tensor tube shown as Fig.1 (b). Particularly, we use \( \mathbf{X}^{(i)} \) to represent \( \mathbf{X}(:, :, i) \) shown as Fig.1(c). The fast Fourier transform (FFT) of \( \mathbf{X} \) along the 3-th dimensional is represented as \( \hat{\mathbf{X}} = \text{fft}(\mathbf{X}(:, :, 3)) \in \mathbb{R}^{n_1 \times n_2 \times n_3} \). \( \mathbf{X} \) can be also computed from \( \hat{\mathbf{X}} \) through inverse FFT as \( \mathbf{X} = \text{ifft}(\hat{\mathbf{X}}(:, :, 3)) \).

The \( l_1 \)-norm and Frobenius norm of \( \mathbf{X} \) and \( \hat{\mathbf{X}} \) are expressed as \( \| \mathbf{X} \|_1 = \sum_{ijk} |\mathbf{X}(i, j, k)| \), \( \| \mathbf{X} \|_F = \sqrt{\sum_{ijk} \mathbf{X}(i, j, k)^2} \), and \( \| \hat{\mathbf{X}} \|_1 = \sum_{ijk} |\hat{\mathbf{X}}(i, j, k)| \), \( \| \hat{\mathbf{X}} \|_F = \sqrt{\sum_{ijk} \hat{\mathbf{X}}(i, j, k)^2} \), respectively.

B. FEATURE EXTRACTION OF THE MULTICHANNEL SIGNALS

1) TENSORIZATION OF THE MULTICHANNEL SIGNALS VIA PHASE SPACE RECONSTRUCTION

We use the phase space reconstruction technique to construct the attractor tensor and the construction process is illustrated in Fig. 2, which can be seen as the tensorization of the multichannel signals. Frist, suppose there are \( n_3 \) channel raw signals \( \mathbf{Z} = [\mathbf{z}_1^T(t), \ldots, \mathbf{z}_{n_3}^T(t)] \in \mathbb{R}^{P \times n_3} \) acquired from vibration sensors in different directions and positions, and each signal contains \( P \) sampling points. Then, each channel signal \( \mathbf{z}_i(t) = [\mathbf{z}_i(1), \ldots, \mathbf{z}_i(P)], i = 1, \ldots, n_3 \) is reconstructed into an attractor matrix \( \mathbf{X}_i \in \mathbb{R}^{n_1 \times n_2} \) via the following embedding.
In this section, we will briefly introduce the decomposition procedure:

\[
X_i = \begin{bmatrix}
  z_i(1) & z_i(2) & \cdots & z_i(n_2) \\
  z_i(1+\tau) & z_i(2+\tau) & \cdots & z_i(n_2+\tau) \\
  \vdots & \vdots & \ddots & \vdots \\
  z_i(1+(n_1-1)\tau) & z_i(2+(n_1-1)\tau) & \cdots & z_i(P) \\
\end{bmatrix}
\]

where embedding dimension \(n_1\) and delay time \(\tau\) are key parameters in phase space reconstruction; \(n_2\) is the window length. The relationship between these parameters can be written as: \(n_2 + (n_1 - 1)\tau = P\).

Third, set these matrices as the frontal slices of tensor \(X\) via arranging them along the third dimension, that is, \(X^{(i)} = X_i\). Finally, the attractor tensor \(\hat{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}\) can be formed by stacking these frontal slices. It can be observed that the correlation feature information of each channel signal is implicit in the corresponding frontal slices. In addition, the third dimension of the tensor concatenates the channel, which is helpful to dig the correlation among the channels.

2) TENSOR DECOMPOSITION AND FAULT FEATURE EXTRACTION VIA TSVD

In this section, we will briefly introduce the decomposition of the attractor tensor via TSVD model and extract the fault feature signal based on this model.

Refer to the definition 2.7 in [29], the attractor tensor \(X\) can be decomposed via the TSVD model as:

\[
\hat{X} = U \circ S \circ \hat{V}^T,
\]

where \(U \in \mathbb{R}^{n_1 \times n_1 \times n_3}\) and \(\hat{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}\) are orthogonal tensors. \(S \in \mathbb{R}^{n_1 \times n_2 \times n_3}\) is a rectangular f-diagonal tensor. It has been proven in [29, 32] that (2) can be efficiently calculated by multiple matrix SVDs in Fourier domain as:

\[
\hat{X}^{(i)} = \hat{U}^{(i)} \hat{S}^{(i)} \hat{V}^{(i)T}, \quad i = 1, \ldots, n_3,
\]

and Algorithm 1 gives the detailed calculation process. This is the process of decomposition of \(X\) via TSVD model in the original domain and the Fourier domain.

Algorithm 1 TSVD

**Input:** attractor tensor \(X \in \mathbb{R}^{n_1 \times n_2 \times n_3}\)

\[
\hat{X} = \text{ifft}(X, [1, 3])
\]

for \(i = 1 : n_3\) do

\[
[U, S, \hat{V}] = \text{svd}(\hat{X}(\cdot, \cdot, i))
\]

\[
\hat{U}(\cdot, \cdot, i) = U, \hat{S}(\cdot, i) = S, \hat{V}(i, \cdot) = V
\]

end

\[
U = \text{ifft}(U, [1, 3]), S = \text{ifft}(S, [1, 3]), \hat{V} = \text{ifft}(\hat{V}, [1, 3])
\]

**Output:** \(U, S, \hat{V}\)

Next, we introduce how to extract the fault feature signal from \(X\) based on this model. Similar to the theory of matrix SVD, we name the orthogonal tensors \(\hat{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}\) and \(\hat{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}\) the left and right singular tensors of \(X\). Moreover, the rectangular f-diagonal tensor \(\hat{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}\) can be seen as the singular value tensor of \(\hat{X}\). For simplicity of notation, we mark the all singular values set in \(\hat{X}\) and \(\hat{X}^{(i)}\) as \(\sigma(\hat{X})\) and \(\sigma(\hat{X}^{(i)})\). Besides, denote the \(j\)-th singular value of \(\hat{X}^{(i)}\) as \(\sigma_j^{(i)}(\hat{X})\). The frontal slice \(\hat{S}^{(i)}\) can be seen as the singular value matrix of \(\hat{X}^{(i)}\) and its singular values set is denoted as \(\sigma(\hat{X}^{(i)}) = \text{diag}(\hat{S}^{(i)}) = \{|\sigma_j^{(i)}(\hat{X})|\}_{1 \leq j \leq n_3}\), which is sorted in descending order. Furthermore, we defined \(j\)-th diagonal tube \(\hat{S}(j, \cdot, \cdot)\) of \(\hat{S}\) as the \(j\)-th order SVTs, which can be viewed as a vector consisting of the \(j\)-th largest singular values of each frontal slice of \(\hat{X}\). Therefore, SVTs can be also regarded as descending order. In addition, there are two new definitions about SVTs as follows, which are tensor tubal rank and tensor nuclear norm.

Definition 2.1. (Tensor Tubal Rank and Tensor Nuclear Norm) [30]: The tensor tubal rank \(r_t\) of \(X\) is defined as the number of non-zero SVTs. The tensor nuclear norm \(\|X\|_{\text{NN}}\) is defined as the sum of singular values in all SVTs:

\[
\|X\|_{\text{NN}} = \sum_{j=1}^{n_3} \sum_{p=1}^{n_2} |\sigma_j^{(p)}(X)|.
\]

According to the mechanism of SVD filtering, we suppose each SVTs can represent the singular subspace of different feature information or noise information in the raw signals. It has been widely accepted that the first few orders of high-amplitude SVTs dominate the main fault feature information of signal [32], which can be seen as the useful SVTs. While the remaining SVTs can be considered to dominate noise or interference information, which are useless SVTs. Thus, we consider \(X\) can be separated into a fault feature tensor \(F_r \in \mathbb{R}^{n_1 \times n_2 \times n_3}\) and a noise or interference tensor \(\mathcal{E}_r \in \mathbb{R}^{n_1 \times n_2 \times n_3}\) via determining an appropriate reconstructed order \(r\) of SVTs (\(r\) can also be called the reconstructed order of the tensor \(X\) or \(X^{(i)}\), and this reconstruction process can be easily computed in the Fourier domain as:

\[
\hat{F}_r^{(i)} = \sum_{j=1}^{r} \sigma_j^{(i)}(\hat{X}) \hat{U}(\cdot, j, i) \hat{V}(j, \cdot, i),
\]

\[
\hat{E}_r^{(i)} = \sum_{j=r+1}^{n_3} \sigma_j^{(i)}(\hat{X}) \hat{U}(\cdot, j, i) \hat{V}(j, \cdot, i).
\]

\(\mathcal{F}_r\) can be got via the inverse FFT of \(\hat{F}_r\) and then the first \(r\)-order extracted signals \(F_r = [F_r^{(1)}, F_r^{(2)}, \ldots, F_r^{(n_3)}] \in \mathbb{R}^{P \times n_3}\) can be obtained via the inverse phase space reconstruction of \(\mathcal{F}_r\). When the value of \(r\) is optimal selected, \(F_r\) is the extracted fault feature signals. The above is the whole idea of using the TSVD model to extract fault feature signals in this paper and a flowchart is shown in Fig.3.

However, the contribution of random noise is widely distributed in all orders [6], which can cause the values in useful SVTs deviate from the true value [2]. Consequently, the extracted fault feature signals inevitably contain noise, resulting in the difficulty in the diagnosis of early weak fault. So, how to effectively eliminate the noise weight in SVTs and the corresponding singular subspace is a key topic of this paper. In addition, the interference components in attractor tensor increases the difficulty of reconstructed order’s determining, which directly determine the quality of feature extraction. Hence, the determination of optimal reconstructed order is another key problem to be solved in this paper.
C. NOISE REDUCTION OF SIGNAL ATTRACTION TENSOR VIA TRPCA

On the basis of TSVD model, the TRPCA can effectively eliminate the noise weights in SVTs via soft threshold function, which has been proved to be a powerful method for the noise reduction of tensor data. And the soft threshold function is defined as follows.

**Definition 2.2. (Soft Threshold Function) [33]:** Fig.4 illustrates the soft threshold function: $\mathbb{R} \rightarrow \mathbb{R}^+$ which can be defined as:

$$\text{soft}(\sigma, \gamma) := \begin{cases} 0, & \sigma < \gamma \\ (\sigma - \gamma), & \sigma \geq \gamma. \end{cases}$$

The principle of TRPCA is shown in Fig.5, which supposes $\mathcal{X}$ can be robustly separated into a sparse tensor $\mathcal{Q} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ representing the noise information and a low-tubal rank tensor $\mathcal{L} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ representing the feature information which include fault feature and other interference features. Both of these components can be extracted by solving the following joint minimization problem of tensor nuclear norm and penalty function regularization:

$$\arg \min_{\mathcal{L}, \mathcal{Q}} \| \mathcal{L} \|_{\text{TNN}} + \lambda \| \mathcal{Q} \|_1, \text{ s.t. } \mathcal{X} = \mathcal{L} + \mathcal{Q},$$

where $\lambda$ is a regularization parameter. Equation (5) is a typical convex optimization procedure and an approximate optimal solution can be acquired via alternating direction method of multipliers (ADMM) [37]. The core operation of this solution is to adopt the soft threshold function to perform an evenly amplitude reduction on all SVTs in the iterative solution process, and the corresponding singular subspace is also eliminated. As shown in Fig.4, the SVTs with smaller singular values decay rapidly to 0, which are generally considered to contain most of the noise information. In the previous section, these SVTs were considered useless in the TSVD model as well. Meanwhile, the values of the SVTs with higher amplitude are also been greatly reduced, thus basically eliminating the noise weight in them. As such, the problem of useful SVTs being polluted by noise in the previous section is basically solved. This theoretically explains the powerful noise reduction performance of TRPCA. Finally, we can obtain a low-tubal rank feature tensor $\mathcal{L}$. Taking it as the attractor tensor after denoising in Fig.3, and then via selecting an appropriate reconstructed order, the relatively pure fault feature signals can be extracted from $\mathcal{L}$. 

---

**FIGURE 3.** The flowchart of the fault feature extraction based on TSVD model.

**FIGURE 4.** Soft threshold function with a threshold parameter $\gamma > 0$.

**FIGURE 5.** Schematic diagram of TRPCA: the attractor tensor can be robustly separated into a sparse component and a low-tubal rank component.
However, there is a fatal flaw in the practical application of TRPCA, that is, the amplitude of the extracted feature signals is often much smaller than that of the raw clean signals. The main reason for this phenomenon is the evenly amplitude reduction strategy for all SVTs, i.e., the value of all SVTs decrease the size of the threshold parameter in each iteration, resulting in a drastically reduction in the amplitude of useful SVTs. Thus, the energy of the extracted weak fault feature signal decrease obviously, which increases the difficulty of the fault diagnosis.

III. MAIN WORK OF THIS PAPER

In this section, we will present our own solution to the two key problem mentioned in the Section II, that is, the amplitude underestimation problem of useful SVTs in the noise reduction process and the problem of difficulty in determining the optimal reconstructed order.

A. NOISE REDUCTION OF SIGNAL ATTRACTOR TENSOR VIA GNCTRPCA

Aiming at the shortcomings of TRPCA in noise reduction process, a novel GNCTRPCA method is introduced to avoid the amplitude reduction of the useful SVTs, so as to effectively maintain the energy of the extracted fault feature signals. Primarily, the definition of generalized non-convex function is given as follow:

Definition 3.1. (Generalized non-convex function) [34] The generalized non-convex functions $g_{\mu,\gamma}(\sigma): \mathbb{R} \rightarrow \mathbb{R}^+$ are continuous, nonnegative, concave and monotonically increasing on $[0, \infty)$. In addition, their gradients are nonnegative and monotonically decreasing on $[0, \infty)$. Table 1 lists several common used generalized non-convex functions, including minimax concave penalty (MCP) [33], smoothly clipped absolute deviation (SCAD) [39], Geman [40], and Logarithm [41]. And Fig.6 illustrates them.

Then, we apply a generalized non-convex relaxation to the convex constraint of tensor nuclear norm in (5) via changing it to the above functions. As such, we can get the following joint minimization problem of generalized non-convex constraint and penalty function regularization:

$$\arg \min_{\mathcal{L}, \mathcal{Q}} \sum_{i=1}^{n_3} \sum_{j=1}^{n} g_{\mu,\gamma}(\sigma_j^{(i)}(\mathcal{L}))+\lambda \|\mathcal{Q}\|_1, \quad s.t. \mathcal{X} = \mathcal{L} + \mathcal{Q}. \tag{6}$$

The solution of (6) faces more challenge than that of the convex problem shown in (5) due to the non-convexity of $g_{\mu,\gamma}$. As shown in Algorithm 2, the solution of (6) can be transformed into two interrelated non-convex optimization (iterative step 1) and convex optimization problems (iterative step 2) via ADMM. And a global optimal solution to the step 2 can be calculated via the soft-thresholding (shrinkage) operator in [42].

Next, we will explain how the proposed GNCTRPCA method maintains the amplitude of the useful SVTs while eliminating the noise weights in them via the solution of the (11) in [29], this problem can be divided into $n_3$ independent minimization problems in Fourier domain:

$$\hat{\mathcal{L}}_{k+1}^{(i)} = \arg \min_{\mathcal{L}^{(i)}} \sum_{j=1}^{n} g_{\mu,\gamma}(\sigma_j^{(i)}(\mathcal{L}))+\lambda \|\mathcal{Q}^{(i)}\|_1, \quad i = 1, \ldots, n_3. \tag{7}$$

where $\varphi$ is a weight constant. It has been proved in [34] that the convergence of (17) can be guaranteed if $\varphi > L(f)$, where $L(f)$ is the Lipschitz constant [43] of the gradient of squared loss function. For simplicity of notation, we denote $\lambda^{(i)} = \mathcal{X} - \mathcal{Z}^{(i)}$ and its TSVD in Fourier domain as $\tilde{\lambda}^{(i)} = \hat{\mathcal{U}}^{(i)} \hat{\mathcal{S}}^{(i)} \hat{\mathcal{V}}^{(i)}$, $i = 1, \ldots, n_3$. And the singular values set of $\lambda^{(i)}$ is denoted as $\sigma^{(i)}(\hat{\lambda}^{(i)}) = \text{diag}(\hat{\mathcal{S}}^{(i)}) = \{\sigma_j^{(i)}(\hat{\lambda}^{(i)})\}_{1 \leq j \leq n}$. Specially, $\hat{\mathcal{S}}_0 = \tilde{\mathcal{S}}^{(i)} = \text{diag}((\sigma_j^{(i)}(\hat{\lambda}^{(i}))_{1 \leq j \leq n})$. The following inequality relation is easily derived via the concavity of $g_{\mu,\gamma}$:

$$g_{\mu,\gamma}(\sigma_j^{(i)}(\hat{\mathcal{L}}^{(i)})) \leq g_{\mu,\gamma}(\sigma_j^{(i)}(\hat{\mathcal{X}}^{(i)})) + m_{j,k}(\sigma_j^{(i)}(\tilde{\mathcal{L}}^{(i)}) - \sigma_j^{(i)}(\hat{\mathcal{X}}^{(i)})). \tag{8}$$
Algorithm 2 Solve (6) by ADMM

Input: signal attractor tensor $X \in \mathbb{R}^{a_1 \times a_2 \times n}$; regularization parameter: $\lambda$; generalized non-convex function parameter: $\mu, \gamma, \varphi$.

Initialize: $\xi = 1 \cdot 9, \mathcal{L}_0 = \mathcal{Q}_0 = \mathcal{Z}_0 = 0$.

while not converged do

1. Update $\mathcal{L}_{k+1}$ by:

$$L_{k+1} = \arg \min_{\mathcal{L}} \sum_{i=1}^{n} \sum_{j=1}^{n} g_{\mu,\gamma}(\sigma_j^{(i)}(\mathcal{L})) + \frac{\varphi}{2} \| \mathcal{L} + \mathcal{Q}_k - \mathcal{X} + \mathcal{Z}_k \|_F^2.$$ 

2. Update $\mathcal{Q}_{k+1}$ by:

$$Q_{k+1} = \arg \min_{\mathcal{Q}} \| \mathcal{Q} \|_1 + \frac{\varphi}{2} \| \mathcal{L}_{k+1} + \mathcal{Q} - \mathcal{X} + \mathcal{Z}_k \|_F^2.$$ 

3. Update $\mathcal{Z}_{k+1}$ by: $\mathcal{Z}_{k+1} = \mathcal{Z}_k + \mathcal{L}_{k+1} + \mathcal{Q}_{k+1} - \mathcal{X}$. 

4. Check the convergence condition:

$$\| \mathcal{L}_{k+1} - \mathcal{L}_k \|_\infty \leq \xi, \| \mathcal{Q}_{k+1} - \mathcal{Q}_k \|_\infty \leq \xi,$$

$$\| \mathcal{L}_{k+1} + \mathcal{Q}_{k+1} - \mathcal{X} \|_\infty \leq \xi.$$ 

end

Output: $\mathcal{L},\mathcal{Q}$

where $m_{j,k} = \partial g_{\mu,\gamma}(\sigma_j^{(i)}(\hat{X}_k))$. Thus, (7) can be relaxed as:

$$\hat{L}_{k+1}^{(i)} = \arg \min_{\hat{L}} \sum_{j=1}^{n} g_{\mu,\gamma}(\sigma_j^{(i)}(\hat{X}_k)) + m_{j,k}(\sigma_j^{(i)}(\hat{L})) - \sigma_j^{(i)}(\hat{X}_k) + \frac{\varphi}{2} \| \hat{L} - X_k^{(i)} \|_F^2$$

$$= \arg \min_{\hat{L}} \sum_{j=1}^{n} m_{j,k}(\sigma_j^{(i)}(\hat{L})) + \frac{\varphi}{2} \| \hat{L} - X_k^{(i)} \|_F^2,$$

$$i = 1, \ldots, n_3.$$ 

Equation (9) is still non-convex. We introduce a new threshold function to give a close form solution to it.

Definition 3.2 (Non-Convex Threshold Function): By using the derivative $\partial g_{\mu,\gamma}$ that varies with $\sigma$ to replace the fixed threshold $\gamma$ in soft threshold function, we can get the following non-convex threshold function:

$$ntf(\sigma, \mu, \gamma, \varphi) = \begin{cases} 
0, & \sigma \leq \partial g_{\mu,\gamma}(\sigma) / \varphi \\
\sigma - \partial g_{\mu,\gamma}(\sigma) / \varphi, & \sigma > \partial g_{\mu,\gamma}(\sigma) / \varphi.
\end{cases}$$

(10)

Fig.7 illustrates a non-convex threshold function, in which GMC is employed as the generalized non-convex function. Comparing it with Fig.7, it can be seen that the main superiority of this new threshold function over the soft threshold function is that the amplitude reduction effect of this function on the $\sigma$ with higher values is gradually weakened.

From the monotone decreasing property of the derivative function $\partial g_{\mu,\gamma}$ and the singular value sequence $\sigma_1^{(i)}(\hat{X}_k) \geq \sigma_2^{(i)}(\hat{X}_k) \geq \ldots \geq \sigma_n^{(i)}(\hat{X}_k)$ in descending order, we can deduce that the gradient sequence $\mathbf{m} = \{m_{j,k}\}_{1 \leq j \leq n}$ in (9) is ascending. Refer to the theorem 2.3 in [44], a close form solution of (9) can be obtained via the weighted singular value shrinkage operator as:

$$\hat{L}_{k+1}^{(i)} = \hat{U}_k^{(i)} M_{\mu,\gamma,\varphi}(\hat{S}_k^{(i)}) \hat{V}_k^{(i)^T},$$

(11)

where $M_{\mu,\gamma,\varphi}(\hat{S}_k^{(i)}) = \text{diag}(ntf(\sigma_j^{(i)}(\hat{X}_k), \mu, \gamma, \varphi))_{1 \leq j \leq n}$. It can be observed the essence of this solution is adopting the non-convex threshold function to perform amplitude reduction on all SVTs in each iteration. As can be seen in Fig.7, the SVTs with smaller singular values are eliminated quickly and directly, which is the same with the result of TRPCA. Furthermore, different from the evenly amplitude reduction strategy on all SVTs adopted by TRPCA, the amplitude reduction effect in this solution for the SVTs with higher amplitude is gradually weakened. Thus, while effectively eliminating noise weights in these SVTs, the weights representing feature information in them are also maintained as much as possible. Especially for the SVTs with very higher singular values, their amplitudes remain almost unchanged.

The above theoretical analysis indicates the advantage of GNCTRPCA over TRPCA, that is, it can significantly avoid the amplitude reduction of the useful SVTs while ensuring powerful noise reduction performance. Thus, the energy of subsequent extracted fault feature signals can be effectively maintained. At the end of the iteration of Algorithm 2, we can get a feature tensor $\mathcal{L}$ with a low-tubal rank $r$. As shown in Fig.3, we can set $\mathcal{L}$ as the attractor tensor after denoising, and then the pure fault feature signals can be extracted via determining an optimal reconstruction order $r$. Note that, if we set $r$ as $r$, the corresponding extracted signals $\mathbf{L}_r = \{\hat{U}_r(t), \hat{V}_r(t), \ldots, \hat{V}_r(n_3(t))\}$ is actually the extracted feature signals after denoising, which mainly composed of fault feature and other interference features. Next, we will introduce a new method to determine the optimal reconstruction order.
B. DETERMINATION OF OPTIMAL RECONSTRUCTED ORDER VIA THE TSVK METHOD

Kurtosis is a numerical statistic to measure the peak of signal waveform, and the kurtosis variety in different types of signals are greatly, the normal signal is generally less than 3, while the impact signal with nonlinear and non-stationary nature will increase substantially [6]. Therefore, kurtosis is suitable for the identification of early weak fault of bearing. Based on this, a new TSVK method is proposed for the determination of optimal reconstructed order. The kurtosis of a vibration signal $x \in \mathbb{R}^P$ can be expressed as:

$$K(x) = E(x - \mu)^4 / \sigma^4,$$  \hspace{1cm} (12)

where $E(x)$ is the expected function, $\mu$ is the mean of $x$ and $\sigma$ is the standard deviation of $x$.

We define TSVK of the first $i$-order extracted signals $F_i = [\tilde{f}_1^i(t), \tilde{f}_2^i(t), \ldots, \tilde{f}_{n^3}^i(t)]$ as the average kurtosis of its each channel signal:

$$tsvk(i) = \frac{1}{n^3} \sum_{j=1}^{n^3} K(\tilde{f}_j^i(t)), \hspace{1cm} i=1, \ldots, r_i.$$  \hspace{1cm} (13)

Then, we define the relative change rate of TSVK as:

$$rtsvk_i = \frac{tsvk(i) - tsvk(i+1)}{tsvk(i+1)}, \hspace{1cm} i=1, \ldots, r_i.$$  \hspace{1cm} (14)

Calculate the $rtsvk_i$ of all order and find the maximum absolute value $rtsvk_1$ in them. If the value $rtsvk_1$ is a positive value, which means the first $I$-th order extracted signals $F_i$ has the maximum TSVK value, i.e., $F_i$ contains the main fault feature. The later $rtsvk_i$ ($i > I$) is getting smaller indicates that the corresponding extracted signals contain interference feature. Therefore, we can confirm that the optimal reconstructed order of $C$ is equal to $I$. In addition, if $rtsvk_1$ is a negative value, which means the first $I+1$ order extracted signals $F_{i+1}$ has the maximum TSVK value. Therefore, the optimal reconstructed order in this case can be set as $I+1$. As a result, the theoretical optimal reconstructed order has been determined.

C. THE PROCESS OF THE PROPOSED FAULT DIAGNOSIS TECHNIQUE FOR MULTICHANNEL BEARING SIGNALS

Fig.8 shows the flowchart of the proposed fault diagnosis technique for multichannel bearing signals, the detailed procedures can be summarized as follows:

**Step 1:** Reconstruct the acquired multichannel vibration signals into the attractor tensor via phase space construction shown in (1).

**Step 2:** Perform the TSVD shown in (2) on the attractor tensor to obtain all SVTs and their corresponding singular subspace.

**Step 3:** Use the proposed GNCTRPCA method shown in (6) and Algorithm 2 to eliminate noise weights in these SVTs as well as maintaining the amplitude of useful SVTs. Moreover, a low-tubal rank feature tensor is obtained and set it as the attractor tensor after denoising shown in Fig.3.

**Step 4:** The TSVK method is employed to determine the optimal reconstructed order, and extract the pure fault feature signals according this order as shown in Fig.3.

**Step 5:** The identification of fault type via comparing the peak frequencies of the extracted signals with the theoretical fault-related frequencies.

IV. SIMULATION ANALYSIS

A. VIBRATION SIGNAL MODEL OF THE BEARING FAULT AND THE SIMULATED MULTICHANNEL SIGNAL

The common bearing local faults mainly occur in the inner race, rolling element and outer race [2]. Refer to [45]–[47], the vibration signal model of bearing fault feature can be expressed as the superposition of multiple impulse excitations:

$$y(t) = \sum_{i=0}^{L} a_i(t) \cos(2\pi f_i t - b_i(t) + \phi_i) + \sum_{j=1}^{J} B_j \sin(2\pi f_j t + \phi_j).$$  \hspace{1cm} (15)

where $\phi_i$ and $f_i$ are the initial phase of the $i$-th impulse excitation and the resonance frequency respectively. $a_i(t)$ and $b_i(t)$ represent the amplitude modulated component and frequency modulated component. $\zeta$ and $\eta_i$ are the system damping ratio and random slip respectively. $f_i$ is the fault feature frequency, and we denote $f_i$ for the inner race fault, $f_o$ for the rolling element fault, $f_o$ for the outer race fault. $T_c$ represent the time period of fault feature frequency. $\phi_{ij}$ represents the initial phase. $A_i$ and $B_j$ are amplitudes.

The $c_i(t)$ in (15) represents the modulated effect produced by the path of propagation from the fault location to the sensor installation location. For the faults of inner race or rolling element, $c_i(t) = C[1 + \cos(2\pi f_i t)],$ where $f_i$ is the rotational frequency of the shaft and $C$ is an amplitude. For the fault of outer race, $c_i(t) = C.$

Fig.9 illustrates two examples of bearing fault feature signals of inner race and outer race. From the spectrum of these two signals, it can be observed the fault-related frequencies of inner race include: fault feature frequency $f_i$, its harmonic frequencies $nf_i$, rotation frequency $f_r$, and the modulated sidebands consisting of their combinations $nf_i \pm f_r$. The fault-related frequencies of outer race include: fault feature frequency $f_o$ and its harmonic frequencies $nf_o$. A common method for identification of fault type is searching for these fault-related frequency contents from the peak frequencies of extracted fault feature signals.

Based on the above analysis, without loss of generality, taking the bearing inner race fault as an example, a raw simulated multichannel signals can be expressed as follow:

$$Z = AX + N.$$  \hspace{1cm} (16)

$$Y = [y_1^T(t), y_2^T(t), y_3^T(t)]$$ is the simulated three source signals. Where $y_1(t)$ is a bearing inner race fault feature signal simulated by (15) and its detailed model parameters is shown.
FIGURE 8. The flowchart of the proposed fault diagnosis technique for multichannel bearing signals via GNCTRPCA and TSVK.
in Table 2. The signal waveform and spectrum of \( y_1(t) \) are shown in Fig.9. \( y_2(t) \) and \( y_3(t) \) are two harmonic interference feature signals and their expression with different feature frequencies \( f_2 = 20 \) Hz and \( f_3 = 50 \) are shown as (17):

\[
\begin{align*}
y_2(t) &= 0.01 \cos(2\pi f_2 t + 20), \\
y_3(t) &= 0.02 \sin(2\pi f_3 t + 10),
\end{align*}
\]

(17)

\( A \) is a \( 3 \times 3 \) random matrix used to mix the source signals into three channel signals, which is employed to simulate the simultaneous acquisition of fault signals via multiple sensors installed in different positions and directions. And its expression is:

\[
A = \begin{bmatrix}
0.7880 & 0.4209 & 0.3884 \\
0.4896 & 0.1142 & 0.2852 \\
0.8529 & 0.3257 & 0.1481
\end{bmatrix}.
\]

(18)

\( N \) represents the zero-mean white Gaussian noise added in each channel. The sampling points and sampling frequency in each channel are \( P = 10000 \) and \( f_s = 10000 \)Hz respectively.

**B. ANALYSIS OF NOISE REDUCTION PERFORMANCE OF GNCTRPCA**

First, the simulated signals with different noise varying SNR from -5 to 5 are created to test the noise reduction performance of the proposed GNCTRPCA method. And the waveform and spectrum of simulated signals with the SNR of -5 are shown in Fig.10. We can observe that the weak fault feature have been overwhelmed by the noise and interference features. To better demonstrate the superior performance of GNCTRPCA, the MEMD and the TRPCA are chosen as the comparative analysis methods. In addition, in order to prove the advantage of the two tensor methods (TRPCA and GNCTRPCA) in processing tensor structure data composed of multichannel signals, a matrix method named RPCA [48] for single-channel signal processing is also chosen for comparison, which performs noise reduction operation for each channel separately. In fact, the two tensor methods and the RPCA method are all noise reduction methods, which are built on the basis of low-rank approximation, and their essence is to extract a low-rank characteristic component and a sparse noise component from the data. And the two tensor methods can be viewed as generalizations of RPCA from the matrix to the high-dimensional tensor.

Without loss of generality, we select the GMC function as the test object in GNCTRPCA. The parameters of phase space reconstruction play very important roles in the methods of RPCA, TRPCA and GNCTRPCA. The minimum embedded dimension can be effectively calculated via the false nearest neighbor algorithm (FNN) [27]. Hence, according to this algorithm and considering the computational efficiency, we set the parameters as \( n_1 = 50 \) and \( r = 200 \) in this section. Thus, the size of the corresponding attractor matrix and attractor tensor are \( X_1 \in \mathbb{R}^{50 \times 200} \) and \( X \in \mathbb{R}^{50 \times 200 \times 3} \). The regularization parameter \( \lambda \) is set to the recommended value \( \lambda = 1/\sqrt{\max(n_1, n_2)} \) in [48] for RPCA, and
\[ \lambda = 1/\sqrt{n_3 \max(n_1, n_2)} \] in [32] for TRPCA and GNCTR-PCA. The Lipschitz constant \( L(f) \) in (7) usually equal to 1, and we set the weight constant \( \varphi = 1.2 \). For the determination of parameters \( \mu, \gamma \) in \( g_{\mu, \gamma} \), we seek the optimal combination of them from a recommended range, which can provide better noise reduction performance in most cases. The energy difference tracking method is adopted [49] in MEMD to identify useful multivariate IMFs.

In order to evaluate objectively the noise reduction performance of the above four methods, we select the average SNR of three channel signals after denoising and a relative error as the evaluation criteria. And the relative error is defined as \[ \frac{\| X - X^* \|_F}{\| X^* \|_F} \], where \( X \) represents the clean simulated signals without noise, and \( X^* \) represents the extracted feature signals after denoising. The energy difference tracking method is adopted [49] in MEMD to identify useful multivariate IMFs.

From Fig. 11, the following information can be summarized: 1) The performance of MEMD is significantly inferior to the other three methods, which shows the advantages of the low-rank approximation methods in signal noise reduction field under strong background noise environment. 2) The performance of the two tensor methods is better than that of RPCA, which indicates that the tensor method has obvious advantages in processing multichannel signals. The main reason for this result is that the RPCA perform noise reduction on each channel signal separately, which cannot comprehensively use the common information among channels. While the tensor methods can take the advantage of overall structure of multidimensional data composed of multichannel signal, which can improve the performance via mining the potential correlation information among channels. 3) The proposed GNCTR-PCA perform the best in these methods, which proves the obvious advantage of this method in processing simulated signals with strong background noise.

C. ANALYSIS OF RECONSTRUCTED ORDER VIA THE TSVK METHOD

After the simulated signals are de-noised by GNCTR-PCA and TRPCA, we can get the feature tensor \( L \). Then, the TSVK method is employed for the determination of optimal reconstructed order of \( L \) to obtain the extracted fault features. Here, we take the simulated signals shown in Fig. 10 as an example. After denoising, the tubal rank of the obtained feature tensor is 7 for TRPCA and 6 for GNCTR-PCA. The corresponding relative change rates of TSVK of these two tensors are shown in Fig. 12. In Fig. 12(a), we can observe that the maximum absolute value appears in the second order and it comes from a positive value. Hence, the optimal reconstructed order in this case is equal to 2. For Fig. 12(b), the maximum absolute value appears in the first order and it comes from a negative value. As such, its optimal reconstructed order is equal to 2 in addition. In other words, we can observe that the relative change rate of TSVK of the attractor tensor \( \lambda_1 \), and the result is shown as Fig. 12(c). It can be observed that its optimal reconstructed order is also equal to 2. The above analysis result demonstrates that the TSVK method can stably identify the reconstructed order of the useful SVTs, and the performance can still maintain strong robustness under strong noise.

D. IDENTIFICATION OF FAULT TYPE

Finally, we still take the simulated signals shown in Fig. 10 as an example to verify the effectiveness of the diagnostic performance of the proposed technique. Meanwhile, the MEMD method and the fault diagnosis technique based on TRPCA and TSVK are chosen for comparison. In addition, in order to indicate the advantage of the two tensor-based techniques in processing the multichannel signals and the advantage of multichannel signals over the single-channel signal in providing fault feature information, three single-channel signal analysis methods, the envelope analysis, the manifold learning based on LTSA and the RPCA, are also chosen for comparison, which perform the fault diagnosis on each channel independently.
FIGURE 13. The analysis results of the three single-channel signal processing methods. (a) Signal envelope spectrum; (b) The spectrum of extracted fault feature signals obtained by the manifold learning method based on LTSA; (c) The spectrum of the top five multivariate IMFs obtained by RPCA.
The analysis results of the three single-channel signal processing methods are shown in Fig.13. As can be seen from the envelope spectrum in Fig.13 (a) that the weak fault features are still overwhelmed by ambient noise and interference components, which is hard to be identified. Therefore, envelope analysis is not suitable for the diagnosis of the simulated signals. In the manifold learning method, the LTSA dimensionality reduction algorithm is employed to extract the feature submanifold of the attractor matrix of each channel signal and the spectrum of the final extracted fault feature signals is shown as Fig.13 (b). It can be observed that the strong ambient noise is suppressed to a certain extent, so that the peaks of several fault-related frequencies can be found, including the rotation frequency ($f_r$) in Channel#1, the fault feature frequency ($f_i$) and its modulated sidebands ($f_i \pm f_r$) in Channel#1 and Channel#2, the double harmonic frequency ($2f_i$) in Channel#2. However, the peaks of the remaining fault-related frequencies are still submerged by the surrounding ambient noise and high-amplitude interference components, which makes it impossible to directly determine the existence of fault. Hence, this method has insufficient performance for the diagnosis of the simulated signals. The spectrum of the fault feature signals extracted by RPCA is shown in Fig.13 (c). Similar to the results in Fig.13 (b), the peaks of some fault-related frequencies are relatively obvious and the strong ambient noise is well suppressed in Fig.13 (c). However, the peaks of other important fault-related frequencies are still hard to be identified, which are almost completely submerged in the surrounding ambient noise and interference components. Therefore, we cannot directly determine that the bearing has failed. The above results indicate that perform the feature extraction on each channel separately cannot provide satisfactory diagnostic performance, because it completely ignores the information across channels. Moreover, from the results shown in Fig.13 (b) and (c), it can be seen that the fault information extracted from each channel is different, including the number of fault-related frequencies and their corresponding amplitude intensity. Hence, when we focus on the diagnosis result of a single channel, the useful information may be very limited or cannot be found at all. For example, we cannot find any fault-related information from the Channel#3 in Fig. 13(b), and Channel#1 or #3 in Fig. 13(c) contains very limited fault-related information. While it can be seen from the comprehensive diagnosis results of multichannel of these two graphs that multichannel signals can provide richer fault feature information, which is of great significance to improve the reliability of fault diagnosis results. These prove that multichannel signals have more advantages than single-channel signals in the field of fault diagnosis.

Next, the multichannel signal processing methods are employed to diagnose the simulated signals. When MEMD is performed on the simulated signals, we can obtain 15 multivariate intrinsic mode functions (IMFs). The spectrums of top five IMFs of three channels are shown in Fig.14. Only partial fault-related frequency contents can be found in the IMF2 of each channel. However, there were still a lot of strong ambient noise and interference frequencies in each channel, which seriously affects the identification of fault type. Therefore, MEMD is not effective for the diagnosis of the simulated signals.

We then investigate the diagnosis performance of two tensor methods combined with TSVK. According to the optimal reconstructed order determined in Fig.12, the corresponding fault feature signals can be extracted. And their spectrums are shown in Fig.15 (a) and (b) respectively. In addition, the spectrum of the simulated clean signals without noise and interference features is shown in Fig.15(c). From the analysis result of the technique based on TRPCA and TSVK shown in Fig.15 (a), we can see that most of the fault feature frequency and their harmonic frequencies can be identified in the peak frequencies of each channel signal. And the ambient noise component is also suppressed at a lower level, which proves the noise reduction performance and feature...
The analysis result of the technique of the combination of two tensor methods and the TSVK.

(a) The spectrum of extracted fault feature signals obtained by the technique based on TRPCA and TSVK;
(b) The spectrum of extracted fault feature signals obtained by the proposed technique; (c) The spectrum of the simulated clean signals without noise and interference feature.

The extraction performance of this technique are better than the previous comparison method. Based on these results, we can preliminarily judge that the bearing fault occurs in the inner race. However, as the important part of the fault-related...
contents, the amplitude of all modulated sidebands and the rotation frequency of each channel are relatively low, and they are almost all mixed with the surrounding noise, which are difficult to identify directly. Especially, the sidebands \(4f_i \pm f_r\) in Channel#1, \((f_i - f_r, 2f_i - f_r)\) in Channel#2, \((4f_i - f_r)\) in Channel#3 and the rotation frequency \(f_r\) in Channel#3 are completely submerged and unrecognized. These situations adversely affect the reliability of the final diagnosis results. Moreover, compared with the frequencies amplitudes of the simulated clean signals shown as Fig. 15 (c), most of the frequencies amplitudes in Fig. 15 (a) are drastically reduced. And the similar adverse result also exists in the RPCA shown in Fig. 13 (c). These results proves that the soft threshold function adopted by TRPCA and RPCA indeed reduces the amplitude of useful SVTs or useful singular values in the process of noise reduction, causing the energy reduction of the extracted signals, which seriously influences the quality of the final diagnosis results.

From the resulting spectrums of the proposed technique in Fig. 15 (b), the peaks of all fault-related frequency content are clearly identifiable and the ambient noise component basically disappears. Thus, it can be firmly confirmed that the fault occurs in the inner race of the bearing. Compared with the results shown in Fig. 15 (a), the amplitude of all frequencies in Fig. 15 (b) is higher, and the side bands as well as the rotation frequencies are clearly visible, which means the quality and reliability of the fault diagnosis performance provided by the proposed technique are obviously superior to that of the technique based on TRPCA and TSVK. Moreover, compared with the frequencies amplitudes shown as Fig. 15 (c), the frequencies amplitudes obtained by the proposed technique are well maintained, and the reduction for most of them is small. And even some of them are even higher than those of the simulated clean signal, which occurs mainly because the noise frequency energy in the signal is superimposed at the location of these fault-related frequencies, thus increasing their amplitude. These results prove that the proposed GNCTRPCA method can effectively maintain the energy of the extracted weak fault feature signals in the process of noise reduction, which is helpful to improve the performance of the fault diagnosis.

In addition, from the spectrum of the extracted signals shown in Fig. 15 (a) and (b), it can be found that the feature frequencies of the two interference signals in the simulated source signal also disappear, which proves that the first few orders of SVTs indeed represent the singular subspace where the fault feature is located and TSVK can accurately identify the corresponding reconstructed order. Furthermore, comparing the results of these two figures with those of Fig. 13, especially the result of RPCA shown in Fig. 13 (c), it is clear that the diagnostic performance of these two tensor-based techniques is superior to that of the three given single-channel signal processing methods. This results indicate that tensor method can improve the performance via taking the advantage of the tensor structure data composed of multichannel signals. Finally, as a summary, the following conclusion can be obtained:

1) Multichannel signals can provide more abundant fault feature information than single-channel signals. And the diagnostic performance of the two tensor-based techniques is obviously better than that of the three single-channel signal processing methods and MEMD method, which proves the advantage of tensor method in analyzing multichannel signals.

2) The noise reduction performance of the proposed GNCTRPCA method is obviously better than the other given methods, which can effectively remove the strong ambient noise in the simulated signals. Moreover, it can better maintain the energy of the extracted weak fault feature signals than TRPCA and RPCA methods, which is helpful to improve the performance of subsequent fault diagnosis.

3) The proposed TSVK method can reliably and accurately determinate the optimal reconstructed order of SVTs where fault feature is located and its performance is still robust under strong ambient noise.

4) The fault diagnosis performance of the proposed technique based on GNCTRPCA and TSVK is the best
among several compared methods, which can effectively extract the weak fault feature signals from the simulated signals and identify the corresponding fault types accurately. The simulated analysis results indicate this technique is suitable for weak fault diagnosis under strong ambient noise.

V. EXPERIMENTAL ANALYSIS

In contrast to simulated signals, the dynamic characteristics of the real bearing vibration signals are more complex and changeable. For testing the practicability of the proposed technique, we analyze two set of real multichannel vibration signals of bearing. One is the bearing outer race pitting fault signals acquired from a lab bearing-gears fault test rig. The other one is an open bearing outer race fault signals provided by the IMS center at the University of Cincinnati [50], [51].

A. TEST RIG FAULT SIGNALS ANALYSIS

The lab bearing-gears fault test rig is shown in Fig.16, which consists of a variable frequency AC motor controlled by an encoder, two couplings, a gearbox with two pair of meshing gears and a magnetic powder loader [46]. The fault bearing is installed on the input shaft away from the motor side shown as the position 3 in Fig.16 (b), which is a single raw tapered roller bearing with the model number of 32206. And the detailed parameters of the bearing are listed in Table 3. Due to the limitation of laboratory hardware conditions, two PCB vibration acceleration sensors were employed to

| Number of roller elements | Roller diameter (mm) | Medium diameter (mm) | Contact angle |
|--------------------------|----------------------|----------------------|--------------|
| z = 17                   | d = 8                | D = 46               | α = 14.04°   |
FIGURE 18. The analysis results of the three single-channel signal processing methods. (a) Signal envelope spectrum; (b) The spectrum of extracted fault feature signals obtained by the manifold learning method based on LTSA; (c) The spectrum of extracted fault feature signals obtained by the RPCA.
simultaneously collect the fault signals of the bearing in this experiment. One is a triaxial sensor, which can simultaneously collect the vibration signals in the axial, horizontal and vertical directions. And the other is a uniaxial sensor, which can only collect the vibration signals in vertical direction. The installation position and measurement direction of sensor has an important effect on the quality of the acquired signals, and then influence the fault diagnosis results. It is generally considered that the placement of sensors near the fault location is a better choice. Hence, as shown in Fig.16 (a), the triaxial sensor is installed on the location of the fault bearing, and the uniaxial sensor is installed on the transmission shaft near the motor side. The sampling frequency and sampling points are \( f_s = 10000 \text{ Hz} \) and \( P = 20000 \). The rotation speed of the input shaft is 433.5 PRM. And according to the parameter of bearing, the fault feature frequency of the bearing outer race can be calculated as \( f_o = 51.05 \text{ Hz} \). Thus, a four channel bearing fault signals can be acquired. The waveform and spectrum of the acquired bearing fault signals are shown as Fig.17. It can be seen that the amplitudes and spectrum intensities of the first three channel signals acquired by the triaxial sensor are significantly higher than that of the channel#4 signal collected by the uniaxial sensor. This indicates the first three channel signals may contain relatively rich fault information. Hence, these two methods are not suitable for the diagnosis of the acquired signals. In the result of RPCA shown in Fig.18 (c), it can be observed that the strong ambient noise is reduced to a certain extent, and the peaks of several fault-related frequencies can be found in the first three channels. This also proves that the signal acquired by the first three channels is more useful than that from the last channel. However, these frequency peaks are not significant enough and are surrounded by noise as well as other high-amplitude interference frequencies, which are difficult to be directly identified. Moreover, other important fault-related frequency contents could not be found as well. Thus, we cannot judge whether there is a fault in the bearing. The above results show that these three single-channel signal processing methods cannot provide satisfactory diagnostic results for the acquired multi-channel signals.

Fig.19 is the spectrum of the top eight multivariate IMFs obtained by MEMD. Only a few peaks of fault-related frequencies can be found in the multivariate IMF4, IMF6, IMF7 and IMF8. But, there are still many interference frequency peaks and strong ambient noise, which seriously affect the identification of these fault-related frequencies, resulting in the inability to determine the components, which cannot be identified. Hence, these two methods are not suitable for the diagnosis of the acquired signals. In the result of RPCA shown in Fig.18 (c), it can be observed that the strong ambient noise is reduced to a certain extent, and the peaks of several fault-related frequencies can be found in the first three channels. This also proves that the signal acquired by the first three channels is more useful than that from the last channel. However, these frequency peaks are not significant enough and are surrounded by noise as well as other high-amplitude interference frequencies, which are difficult to be directly identified. Moreover, other important fault-related frequency contents could not be found as well. Thus, we cannot judge whether there is a fault in the bearing. The above results show that these three single-channel signal processing methods cannot provide satisfactory diagnostic results for the acquired multi-channel signals.
bearing fault type. Therefore, MEMD cannot provide good fault diagnosis results for the acquired signals.

Then, the technique based on TRPCA and TSVK is used to diagnose the acquired signals. According to the FNN, the parameters of phase space reconstruction is set as $n_1 = 90$ and $\tau = 200$. Thus, the size of the corresponding attractor tensor is $\mathbf{X} \in \mathbb{R}^{90 \times 2200 \times 4}$. First, a feature tensor with a low-tubal rank of 25 can be obtained by noise reduction treatment of TRPCA. Second, the relative change rate of TSVK of this tensor was calculated via the TSVK method, and the result is shown as Fig.20 (a). It can be seen that the maximum absolute value appears in the third order and it comes from a positive value. Therefore, the optimal reconstructed order is equal to 3. Finally, according to this reconstructed order, the spectrum of corresponding extracted fault feature signals is shown as Fig.21. Comparing with the analysis results shown as Fig.19 and Fig.18 (c), we can observe there is more fault-related frequency peaks can be search in each channel and the strong ambient noise as well as interference frequency peaks are significantly suppressed. As such, the diagnostic performance of TRPCA is significantly better than that of MEMD and RPCA. However, it is undeniable that the amplitudes of these fault-related frequency peaks are relatively low, and most of them are surrounded by weaker ambient noise. Thus, we cannot accurately judge that fault was occurred in the outer race of the bearing.

Finally, the proposed technique was utilized to address the acquired signals. Frist, a feature tensor with a low-tubal rank of 23 can be obtained by GNCTRPCA. Then, the TSVK method is performing on this tensor to obtain the corresponding relative change rate of TSVK, which is shown as Fig.20 (b). The maximum absolute value appears in
the second order and it comes from a negative value. So, its optimal reconstructed order is also equal to 3. The analysis result of the optimal reconstructed order between the two techniques indicates the accuracy and stability of TSVK in recognizing the optimal reconstructed order. Finally, according to this reconstructed order, the spectrum of corresponding extracted fault feature signals is shown as Fig.22. We can observe that the peaks of fault feature frequency ($f_o$) and first three harmonic frequencies ($2f_o$, $3f_o$, $4f_o$) are very prominent in each channel. In addition, comparing with the analysis result of Fig.21, the energy of ambient noise keep at a very low level and the interference frequency peaks are almost nonexistent, which means the diagnostic performance of the proposed technique is better than that of the technique based on TRPCA and TSVK. Moreover, the amplitudes of these fault-related frequency peaks are significantly higher than that in Fig.18 (c) and Fig.21, which demonstrates that compared with the RPCA and the technique based on TRPCA and TSVK, the proposed technique not only can provide a better noise reduction performance but also can effectively maintain the energy of the extracted weak fault feature. Thus, we can undoubtedly confirm that the fault occurred in the outer ring of the bearing. These analysis results show that the proposed technique has a great diagnostic performance for the acquired signals.

### B. OPEN BEARING FAULT SIGNALS ANALYSIS

The bearing signals are acquired from a bearing test rig of the IMS center at the University of Cincinnati. As shown in Fig.23, the test rig is driven by an AC motor and its shaft is mounted with four Rexnord ZA-2115 double row bearings.
FIGURE 25. The analysis results of the three single-channel signal processing methods. (a) Signal envelope spectrum; (b) The spectrum of extracted fault feature signals obtained by the manifold learning method based on LISA; (c) The spectrum of extracted fault feature signals obtained by the RPCA.
of the original signals, in order to better demonstrate the performance of the proposed method, a small amount of random noise is added to the original signals. Fig. 24(a) and (b) show the waveform and spectrum of the acquired signals. It can be seen that the weak fault feature cannot be identified. Hence, the signals need to be addressed.

First, the three single-channel signal processing methods are employed to deal with the signals and the analysis results are shown in Fig. 25. From the signal envelope spectrum in Fig. 25 (a), it can be seen that the fault feature still cannot be identified. From the result of manifold learning shown in Fig. 25 (b), we can observe that only the peak of double harmonic frequencies ($2f_o$) can be found in Channel#3. From the result of RPCA shown in Fig. 25 (c), we can observe that the peaks of fault feature frequency ($f_o$) can be found in each channel and the ambient noise has been reduced to a certain extent. However, $f_o$ in Channel#2 and #3 are still surrounded by the ambient noise and a large number of high-amplitude interference feature. Moreover, other fault-related frequencies are still completely submerged. As such, it is impossible to determine whether the bearing has malfunctioned. These results show that the three single-channel signal processing methods are not effective for this acquired signal.

When the MEMD is employed to address the signals, and sixteen multivariate IMFs can be obtained. The spectrums of top eight multivariate IMFs are shown in Fig. 26. Only the peak of $2f_o$ appears in Channel#3 of IMF6. However, there are still a large number of noise and interference components in each channel, which seriously affects the final fault diagnosis. Hence, MEMD cannot provide good fault diagnosis performance for the acquired signals.

Then, the technique based on TRPCA and TSVK and the proposed technique are utilized to address the acquired signals. According to the FNN, the parameters of phase space reconstruction is set as $n_1 = 80$ and $\tau = 200$. Thus, the size of the corresponding attractor tensor is $X \in \mathbb{R}^{80 \times 4200 \times 3}$.
After noise reduction, the tubal rank of the feature tensor obtained by TRPCA and GNCTRPCA is 21 and 19 respectively. The corresponding relative change rates of TSVK of these two tensors are shown in Fig. 27. It can be observed that the maximum values in these two figures appear in the first order with a negative value. Therefore, the optimal reconstruction order of both is equal to 2. According to the reconstructed order, the spectrum of corresponding extracted fault feature signals are shown in Fig. 28 and Fig. 29.

As shown in Fig. 28, the peaks of $f_o$ and the $2f_o$, can be found in the Channel#1 and Channel#3, meanwhile, the ambient noise is also effectively suppressed. However, the amplitude of these frequencies is also low, which is seriously interfered by the surrounding weaker ambient noise and is difficult to be recognized directly. Therefore, we cannot accurately determine whether there is a fault in the bearing.

From the resulting spectrums of the proposed technique in Fig. 29, the peaks of $f_o$ and first two harmonic frequencies ($2f_o, 3f_o$) are very significant in each channel and the ambient noise as well as the interference frequency components are kept at a low level. Moreover, the amplitudes of these fault-related frequency peaks is significantly higher than that in Fig. 25(c) and Fig. 28, which also indicates that the diagnostic performance of the proposed technique is significantly better than that of the RPCA and the technique based on TRPCA and TSVK. In addition, compared with the frequency amplitudes of the acquired signals shown in Fig. 24(b), the frequencies amplitudes shown in Fig. 29 are somewhat reduced, but generally remains at a high level, which also indicates that the proposed technique can effectively maintain the energy of the extracted weak fault feature signals. Based on the above analysis, we can conclude that the fault occurred in the outer race of the bearing and the proposed...
technique can provide a good fault diagnosis performance for this acquired signal.

VI. CONCLUSION
The weak fault feature of the bearing is easily affected by the strong ambient noise and interference components, which makes it difficult to be extracted. To handle this problem, a new multichannel bearing signal fault diagnosis technique based on GNCTRPCA and TSVK is proposed. The attractor tensor of raw signals can be decomposed into many SVTs representing different feature component or noise. The GNCTRPCA can effectively eliminate the noise weight in each SVTs and meanwhile, avoid the significant amplitude reduction of useful SVTs. Thus, the energy of the extracted fault feature signals can be well maintained. The TSVK method is employed to determine the optimal reconstructed order of SVKs, so as to extract fault feature signals for further fault diagnosis. The analysis of simulation signals and two acquired bearing fault signals verify that the proposed technique can provide superior fault diagnosis performance.

The parameters of phase space reconstruction have a direct impact on the fault diagnosis results of the proposed technique. The classical FNN algorithm is employed in this paper to determine these parameters and provides good diagnostic performance. However, how to determine the optimal parameters requires further research. In addition, the simulation analysis and experimental analysis in this paper only focus on the diagnosis of signals with several given number of channels, whether different number of channels would affect the analysis results also needs further studied. Moreover, the location of the sensor has a direct influence on the fault diagnosis results, so the selection of their optimal location should be a key issue in the study of the experiment. In the experiment of V-A, due to the limitation of hardware conditions, we just arrange the position of the sensors based on experience to collect the multichannel signal data, and only two sensors can be used. Fortunately, the analysis results indicate the acquired signal data can prove the performance of the proposed technique. In the future work, we will purchase multiple sensors and design multiple sensor layout experiments, including different installation locations and directions, to find the optimal sensor layout, so as to provide the best collected signals data for fault diagnosis.

ACKNOWLEDGMENT
The authors appreciate the free download of the original bearing failure data and two photo picture provided by the IMS Center at the University of Cincinnati.

REFERENCES
[1] Y. Lv, M. Ge, Y. Zhang, C. Yi, and Y. Ma, “A novel demodulation analysis technique for bearing fault diagnosis via energy separation and local low-rank matrix approximation,” Sensors, vol. 19, no. 17, pp. 1–24, 2019.
[2] W. Deng, S. Zhang, H. Zhao, and X. Yang, “A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing,” IEEE Access, vol. 6, pp. 35042–35056, 2018.
[3] R. Yuan, Y. Lv, and G. Song, “Fault diagnosis of rolling bearing based on a novel adaptive high-order local projection denoising method,” Complexity, vol. 2018, pp. 1–15, Oct. 2018.
[4] Z. Wang, J. Wang, Y. Kou, J. Zhang, S. Ning, and Z. Zhao, “Weak fault diagnosis of wind turbine gearboxes based on MED-LMD,” Entropy, vol. 19, no. 6, pp. 1–15, 2017.
[5] R. Yuan, Y. Lv, H. Li, and G. Song, “Robust fault diagnosis of rolling bearings using multivariate intrinsic multiscale entropy analysis and neural network under varying operating conditions,” IEEE Access, vol. 7, pp. 130804–130819, 2019.
[6] H. Li, T. Liu, X. Wu, and Q. Chen, “Research on bearing fault feature extraction based on singular value decomposition and optimized frequency band entropy,” Mech. Syst. Signal Process., vol. 118, pp. 477–502, Mar. 2019.
[7] A. Bhattacharyya and R. B. Pachori, “A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform,” IEEE Trans. Biomed. Eng., vol. 64, no. 9, pp. 2003–2015, Sep. 2017.
[8] Y. Lv, R. Yuan, T. Wang, H. Li, and G. Song, “Health degradation monitoring and early fault diagnosis of a rolling bearing based on CEEMDAN and improved MMSE,” Materials, vol. 11, no. 6, p. 1009, Jun. 2018.
[9] I. Daubechies, J. Lu, and H. T. Wu, “Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool,” Appl. Comput. Harmon. Anal., vol. 30, no. 2, pp. 243–261, Mar. 2011.
[10] M. Kang, J. Kim, L. M. Wills, and J.-M. Kim, “Time-varying and multiresolution envelope analysis and discriminative feature analysis for bearing fault diagnosis,” IEEE Trans. Ind. Electron., vol. 62, no. 12, pp. 7749–7761, Dec. 2015.
[11] Q. Jiang, M. Jia, J. Hu, and F. Xu, “Machinery fault diagnosis using supervised manifold learning,” Mech. Syst. Signal Process., vol. 23, no. 7, pp. 2301–2311, Oct. 2009.
[12] P. Borghesani, R. Ricci, S. Chatterton, and P. Pennacchi, “A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions,” Mech. Syst. Signal Process., vol. 38, no. 1, pp. 23–35, Jul. 2013.
[13] B. Yao, P. Zhen, L. Wu, and Y. Guan, “Rolling element bearing fault diagnosis using improved manifold learning,” IEEE Access, vol. 5, pp. 6027–6035, 2017.
[14] Z. Su, B. Tang, Z. Liu, and Y. Qin, “Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine,” Neurocomputing, vol. 157, pp. 208–222, Jun. 2015.
[15] M. Aminghafari, N. Cheze, and J.-M. Poggi, “Multivariate denoising using wavelets and principal component analysis,” Comput. Statist. Data Anal., vol. 50, no. 9, pp. 2381–2398, May 2006.
[16] G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, and S. M. Smith, “Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers,” NeuroImage, vol. 90, pp. 449–468, Apr. 2014.
[17] Y. Guo, J. Na, B. Li, and R.-F. Fung, “Envelope extraction based dimension reduction for independent component analysis in fault diagnosis of rolling element bearing,” J. Sound Vib., vol. 333, no. 13, pp. 2983–2994, Jun. 2014.
[18] Y. Lv, R. Yuan, and G. Song, “Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing,” Mech. Syst. Signal Process., vol. 81, pp. 219–234, Dec. 2016.
[19] C. Yi, Y. Lv, H. Xiao, T. Huang, and G. You, “Multisensor signal denoising based on matching synchrosqueezing wavelet transform for mechanical fault condition assessment,” Meas. Sci. Technol., vol. 29, no. 4, pp. 1–15, 2018.
[20] S. Kouchaki, S. Sanei, E. L. Arbon, and D. J. Dijk, “Tensor based singular spectrum analysis for automatic scoring of sleep EEG,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 23, no. 1, pp. 1–9, Jan. 2015.
[21] T. D. Pham and H. Yan, “Tensor decomposition of gait dynamics in Parkinson’s disease,” IEEE Trans. Biomed. Eng., vol. 65, no. 8, pp. 1820–1827, Aug. 2018.
[22] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, “Tensor decompositions for signal processing applications: From two-way to multway component analysis,” IEEE Signal Process. Mag., vol. 32, no. 2, pp. 145–163, Mar. 2015.
[23] M. Ge, Y. Lv, C. Yi, Y. Zhang, and X. Chen, “A joint fault diagnosis scheme based on tensor nuclear norm canonical polyadic decomposition and multiscale permutation entropy for gears,” Entropy, vol. 20, no. 3, pp. 1–22, 2018.
