

Article

# Improved Dynamic Mode Decomposition and Its Application to Fault Diagnosis of Rolling Bearing

Zhang Dang <sup>1,2,3</sup> , Yong Lv <sup>1,2,\*</sup>, Yourong Li <sup>1,2</sup> and Guoqian Wei <sup>1,2</sup>

<sup>1</sup> Key Laboratory of Metallurgical Equipment and Control Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; dangzhang@wust.edu.cn (Z.D.); liyourong@wust.edu.cn (Y.L.); Weiguqian@wust.edu.cn (G.W.)

<sup>2</sup> Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

<sup>3</sup> National Demonstration Center for Experimental Mechanical Education, Wuhan University of Science and Technology, Wuhan 430081, China

\* Correspondence: lvyong@wust.edu.cn; Tel.: +86-027-6886-2857; Fax: +86-027-6886-2212

Received: 17 May 2018; Accepted: 16 June 2018; Published: 19 June 2018



**Abstract:** To solve the intractable problems of optimal rank truncation threshold and dominant modes selection strategy of the standard dynamic mode decomposition (DMD), an improved DMD algorithm is introduced in this paper. Distinct from the conventional methods, a convex optimization framework is introduced by applying a parameterized non-convex penalty function to obtain the optimal rank truncation number. This method is inspired by the performance that it is more perfectible than other rank truncation methods in inhibiting noise disturbance. A hierarchical and multiresolution application similar to the process of wavelet packet decomposition in modes selection is presented so as to improve the algorithm's performance. With the modes selection strategy, the frequency spectrum of the reconstruction signal is more readable and interference-free. The improved DMD algorithm successfully extracts the fault characteristics of rolling bearing fault signals when it is utilized for mechanical signal feature extraction. Results demonstrated that the proposed method has good application prospects in denoising and fault feature extraction for mechanical signals.

**Keywords:** dynamic mode decomposition; optimal rank truncation threshold; dominant modes selection strategy; fault diagnosis

## 1. Introduction

Rolling bearing plays an important role in modern equipment [1], relating to the health condition and the remaining service life. Amplitude-modulated and/or frequency-modulated (AM-FM) multi-component signals would be introduced when local damage such as cracks and exfoliation corrosion occurred during the operation [2]. It becomes intractable to extract the fault feature as the fact of that the measured signals are usually immersed with intensive background noise, especially in the incipient failure. Thus, robust fault feature extraction method should perfectly eliminate the noisy components and acquire the useful features generated from the non-stationary time series, which provides the proof for the maintenance plan before the whole machine gets disabled.

Recent decades have witnessed significant advances in the feature extraction of fault data collected from either experiments or numerical simulations. Short-time Fourier Transform (STFT) [3] adopts the method of piecewise averaging of the unstable signal. Several segments are sliced by a constant sliding window in accordance with the order of the original timeline. Spectrum characteristics are then received by Fourier Transform (FT) on each segment. However, the time-frequency resolution is constant and cannot transform over the change of time and frequency, which is the shortcoming

of STFT. Wavelet Transform (WT) [4], a method of local time and frequency conversion, was proposed to represent the original signal with a mother wavelet by scaling and translation operation. While it can effectively extract information with multiscale refinement analysis, WT has an inherent problem of low frequency resolution of high frequency band. Wavelet Packet Transform (WPT) [5] was an extension of WT. Its basic idea was to make information energy concentration and decompose the signal into low-frequency components and high-frequency components. With further decomposition of the high frequency part, the time-frequency resolution of WPT is higher than that of WT. However, the choice of mother wavelets for WPT is hard to determine. Empirical Mode Decomposition (EMD) [6] was proposed to decompose the original time series to different intrinsic mode functions (IMFs) by means of adaptive low frequency and high frequency components. EMD is a method of adaptive signal processing that can process nonlinear and non-stationary signals. However, the EMD algorithm has many imperfect problems that need to be solved and improved, such as the effect of the endpoint and mode aliasing problems [7]. At the same time, the algorithm lacks physical significance. Singular value decomposition (SVD) [8] can also be used in processing the noise-corrupted signal, but the decomposition effect differs from the various choice of singular values. SVD may not be helpful when the noise energy reaches a certain level [9].

Dynamic mode decomposition (DMD), stemming from the fluid mechanics community [10], has been characterized as a powerful tool for Koopman spectrum analysis technology based on equation-free and data driven methods. It can extract the spatiotemporal coherent characteristics by decomposing the complex flow field into a series of simple expressions based on its inherent space-time. Recently, DMD is successfully used for prediction, state estimation and complex system control in the field of fluid mechanics [11]. Schmid [12] firstly defined the DMD algorithm by demonstrating that the theory was capable of describing the internal mechanism based on the numerical simulation and measured experimental data in a high-dimensional flow field. The core of the DMD algorithm can be regarded as a space dimension reduction technique, similar to the algorithm of proper orthogonal decomposition (POD). While POD decomposes data into a series of multiple orthogonal modal frequency components, and ranks them in terms of energy content, the DMD algorithm can describe the dynamic characteristics in a series of single non-orthogonal frequency modes and superpose them with coefficients according to the time scale [13]. Rowley [14] pointed out that DMD was closely related to the Koopman operator theory, a linear but infinite-dimensional operator that represented the action of a nonlinear dynamic system. Thus, DMD is applicable to deal with nonlinear systems such as roller bearing signal (which we attempt in this paper) based on the consideration that DMD is a numerical approximation to Koopman spectral analysis. The standard DMD algorithm for mechanical time-series signal processing is outlined in Section 2 with the theory provided by literature [12]. Since the appearance of the DMD algorithm, widespread use and great achievements have been made in the field of fluid application [11,13–20]. Nowadays, research regarding DMD mainly focuses on two aspects. Firstly, noise suppression algorithms were presented to correct the decomposition results due to the influence of noise. DMD cannot accurately extract dynamic characteristic information when the snapshots, periods of small data, are immersed by noise. These strategies range from a noise-corrected dynamic mode decomposition with control (ncDMD) [15], forward-backward DMD (fbDMD) [16], total least squares DMD (tlsDMD) [17], compressed DMD (cDMD) [18], and sparsity-promoting DMD (spDMD) [19]. Though most of the variational DMD algorithms can extract fluid dynamic information more accurately, they are not self-adaptive as they need to give a rank truncation order in their processes. Secondly, it aims at computational complexity reduction in pre-processing and/or post-processing when big data are processed. This problem does not apply to the mechanical signals, since the amount of mechanical signals is usually smaller relative to the fluid domain.

Whether the DMD reconstruction matrix best approximates the original system dynamics' characteristics largely depends on two aspects: the optimal rank truncation order of the companion/similarity matrix and the dominant modes selection strategy. In a standard DMD algorithm, a certain rank order  $r$  of the similarity matrix needs to be determined for POD decomposition.

It will increase the calculation time if over-estimating the  $r$  value because of a large amount of data. Simulation results showed that the cross-correlation coefficient between the reconstruction data and the original synthetic signal even get a bit smaller along with the increase of the selected  $r$ . Conversely, the calculation result may not be accurate if the  $r$  value is too small, which would deduct some useful signal characteristics. In the past, most scholars presented the rank truncation criteria based on SVD, including selecting an “elbow” of the singular value with a logarithmic curve and a threshold such as a certain proportion of selected data variance. They mainly use the idea of discarding noises by substituting a smaller (compressed) matrix that captures the essential characteristics of the signal. A recent breakthrough by Gavish [21] provided a hard optimal threshold foundation in theory whether the noise magnitude is known or unknown. The methods of dominant mode selection strategy with a reduced order model are not identical. For a review of criterion to select dominant modes, we refer the reader to [22], which summarized the recent research situation and analyzed their respective advantages and disadvantages. At the same time, Kou [22] developed a method that multiplied each normalized DMD mode by its time coefficient. This method considered the evolution of each mode, but neglected the faint modes that should be treated as noise modes. Dang [23] introduced a tool, multiscale permutation entropy, to calculate the complexity of each DMD mode. Modes whose entropies were smaller than a threshold were chosen to recover the signal. However, for filtering the noise components, an entropy threshold needs to be set by multiple attempts.

In this paper, a DMD framework for mechanical fault signal processing is proposed, solving the intractable problems of optimal rank truncation threshold and dominant modes selection strategy based on the conventional DMD algorithm. The optimal rank truncation  $r$  of the similarity matrix is calculated with parametric non-convex penalty function for low-rank approximation. Inspired by WPT, which splits the signal up into a collection of different signals according to its level of decomposition, DMD modes are divided into slow modes and fast modes by splitting the continuous-time eigenvalues of the similarity matrix. Hierarchical application of the basic DMD algorithm is applied with a given level  $l$ . DMD modes are reconstructed by a number of bottomed modes. Fast Fourier Transform (FFT) is performed on the reconstructed signal for the purpose of extracting the fault features. The improved DMD algorithm is applied to process the rolling bearing’s simulation signal for noise reduction and feature extraction. Amazingly, by comparing the signal-to-noise ratio (SNR), our method has tremendous advantages in noise removal and fault feature extraction, in comparison with traditional mature noise reduction methods such as EMD, SVD and WT. Then, the proposed method is applied to the bearing signal in practice; as a result, it can successfully extract the fault characteristics that are highly consistent with the actual field.

This paper is organized as follows. Section 2 introduces the basic DMD algorithm for mechanical fault signal processing. Section 3 measures the solving of the optimal rank truncation order of the similarity matrix, and the strategy of dominant modes selection are outlined. Analysis results of dynamic simulation signal and practical bearing fault signal are described in Section 4. Conclusions are summarized in Section 5.

## 2. DMD Algorithm for Mechanical Signal Processing

DMD is an equation-free Koopman frequency analysis technique based on the theory of SVD and mode decomposition. Essentially, it is a kind of order reduction method that can represent potential dynamics’ characteristics of a complex high-dimensional system by extracting a series of single frequency. An order reduction algorithm decomposed from DMD in flow fields is divided into two kinds of mathematical expressions. One introduces a companion matrix approximating to the infinite dimensional linear Koopman operator, and another presents a similarity matrix for a linear mapping operation on the singular values of POD. For detailed description of the two methods, we suggest reading [22]. Both methods are capable of implementing the core algorithm, but the latter is more commonly used in numerical calculations for its robustness. We generalize the DMD algorithm framework for mechanical time-series signal processing based on the similarity matrix.

Supposing that a size of  $N$  sample points is acquired by a sensor with equal intervals between two successive sampling points, we define the one-dimensional time series as  $S = [x_1, x_2, \dots, x_i \dots x_N]$ ,  $x_i \in R$ , while  $\Delta t = x_{i+1} - x_i$ .  $S$  can also be made from a digital simulation signal. A  $m \times n$  shift-stack Hankel matrix  $\mathbf{X}_S$  can be constructed by Equation (1):

$$\mathbf{X}_S = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_2 & x_3 & \cdots & x_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_i & x_{i+1} & \cdots & x_{n+i-1} \\ \cdots & \cdots & \cdots & \cdots \\ x_m & x_{m+1} & \cdots & x_{m+n-1} \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ X_1 & X_2 & \cdots & X_n \\ | & | & \cdots & | \end{bmatrix}, \mathbf{X}_S \in R^{m \times n}. \quad (1)$$

In order to achieve the maximum spatial and temporal complexity of the original noisy signal (As in Equation (7), When the matrix is close to the square matrix, the maximum characteristic frequency components can be obtained with SVD), the product of  $m$  and  $n$  should be as large as possible. According to the principle of inequality, the product achieves maximum when  $m$  and  $n$  are equal or close to each other. The dimension  $m$  is defined as follows:

$$m = \begin{cases} N/2 & \text{if } N = 2z; \\ (N+1)/2 & \text{if } N = 2z-1; \end{cases} \quad (2)$$

where  $z \in R$  is positive integer sequence that  $z = 1, 2, 3, \dots$

It is possible to arrange the  $n$  column vectors into two  $m \times (n-1)$  data matrices:

$$\mathbf{X} = \begin{bmatrix} | & | & \cdots & | \\ X_1 & X_2 & \cdots & X_{n-1} \\ | & | & \cdots & | \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} | & | & \cdots & | \\ X_2 & X_3 & \cdots & X_n \\ | & | & \cdots & | \end{bmatrix}, \mathbf{X}, \mathbf{Y} \in R^{m \times (n-1)}. \quad (3)$$

DMD is algorithmically a regression by assuming an optimal local linear approximation for mapping the current data to the subsequent data. The best-fit linear operator  $\mathbf{A}$  is then used in terms of these signal matrices as:

$$\mathbf{Y} = \mathbf{A}\mathbf{X}. \quad (4)$$

Apparently, the evolution of the sequence  $\mathbf{X}$  and  $\mathbf{Y}$  is determined by the eigenvalues of  $\mathbf{A}$  for the linear system. It can also be taken as an operator to approximate the dynamic characteristics when the data are produced in nonlinear systems, such as bearing vibration signal.

This solution minimizes the error:

$$\operatorname{argmin}_{\mathbf{A}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F, \quad (5)$$

where  $\|\bullet\|_F$  is the Frobenius norm, given by

$$\|\mathbf{C}\|_F = \left( \sum_{j=1}^n \sum_{i=1}^m C_{ij}^2 \right)^{0.5}. \quad (6)$$

In what follows, we outline the key steps of the DMD algorithm for mechanical signal processing.

1. Firstly, seek an invertible matrix decomposition by SVD on matrix  $\mathbf{X}$ :

$$\mathbf{X} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*, \quad (7)$$

where  $U$  and  $V$  are orthonormal, called the left and right singular vector, respectively,  $U^*U = I, V^*V = I$ . The symbol  $*$  denotes the complex conjugate transpose.  $\Sigma \in R^{p \times p}$  contains a number of non-zero singular values  $\{\sigma_1, \dots, \sigma_p\}$  by descending sequence in its diagonal.

The matrix  $A$  of Equation (4) may be obtained by the pseudoinverse of  $X$ :

$$A = YV\Sigma^{-1}U^*. \quad (8)$$

2. It is quite normal that the matrix  $A$  contains such a large amount of data that leads to the calculating processing reluctant. Thus, we choose a given number of truncated rank  $r$ , and project it onto POD modes with the order of characteristic vectors. Thus, we have a similarity matrix  $\tilde{A}$  as follows:

$$\tilde{A} = U_r^*AU_r = U_r^*YV_r\Sigma_r^{-1}, \quad (9)$$

where  $\tilde{A} \in R^{r \times r}$ . The eigenvalues and eigenvectors of  $A$  are then represented by those of  $\tilde{A}$  as they process the same dynamical features.

3. Perform eigenvalue decomposition on the similarity matrix  $\tilde{A}$ :

$$\tilde{A} = W\Lambda W^{-1}, \quad (10)$$

where  $W = [\omega_1, \omega_2, \dots, \omega_r] \in R^{r \times r}$  is the eigenvectors of similarity matrix  $\tilde{A}$ , and  $\Lambda = \text{diag}([\lambda_1, \lambda_2, \dots, \lambda_r]) \in R^{r \times r}$  is a diagonal matrix containing the corresponding complex eigenvalue  $\lambda_i$ .

4. Compute the reconstruction matrix of DMD. The evolution of the signal characteristics can be characterized by the similarity matrix. In addition, the  $i$ -th eigenvector of original operator is presented by the relevant feature vector of the similarity matrix:

$$\phi_i = YV_i \sum_i^{-1} W_i. \quad (11)$$

Equation (11), the project DMD approximate solution, is often called standard DMD mode. By firstly rewriting for convenience  $\omega_i = \ln(\lambda_i)/\Delta t$ , the approximate solution of reconstruction matrix of DMD is then given by:

$$X_{DMD} = \sum_{i=1}^r \phi_i \exp(\omega_i \Delta t) b_i = \Phi \exp(\Omega t) b, \quad X_{DMD} \in R^{m \times n}, \quad (12)$$

where  $\Phi$  is a matrix consisting of DMD modes  $\phi_i$ ,  $\Omega = \text{diag}(\lambda_i)$  is a diagonal matrix whose enters are continuous-time eigenvalues of the similarity matrix  $\tilde{A}$ ,  $b$  is a vector containing the initial amplitude of each mode, and  $b = \Phi^\Gamma X$ ,  $\Gamma$  denotes the Moore–Penrose pseudoinverse.

5. Finally, signal reconstruction and feature extraction can proceed based on the reconstruction matrix  $X_{DMD}$ . We take the first column of  $X_{DMD}$  as the recuperative signal that the length is just about half of the original signal. Then, the Fourier transform is applied on the recuperative signal for a spectrogram, determining whether there is a failure frequency.

### 3. Improved DMD Framework

#### 3.1. Rank Estimation Based on Convex Optimization

As in Section 2, non-zero singular values  $\{\sigma_1, \dots, \sigma_p\}$  are arranged in a descending sequence when performing SVD on  $X$ . Singular spectrum analysis treats the ahead  $k$  values representing the useful signals, and the residual  $p - k$  orders for part of the noise. The difficulty lies in the determination of optimized rank  $k$ . It cannot perfectly restore the dynamic information with the method of singular spectrum or hard optimal threshold foundation for their truncated rank orders  $k$  are smaller than the

ideal ones, which is proved in Section 4.1. Here, a convex optimization framework is introduced to estimate the best rank  $r$  for Equation (9).

The noise polluted Hankel matrix  $\mathbf{X}$  can be supposed to be a combination of an expectation matrix and a residual noise matrix:

$$\mathbf{X} = \mathbf{X}_E + \mathbf{X}_N, \quad \mathbf{X}, \mathbf{X}_E, \mathbf{X}_N \in R^{m \times n}, \quad (13)$$

where  $\mathbf{X}_E$  is the desired low rank matrix, and  $\mathbf{X}_N$  is a zero mean Gaussian white noise matrix. Note that for unitary matrices  $\mathbf{U}$  and  $\mathbf{V}$ ,  $\Phi(\mathbf{X}) = \Phi(\mathbf{U}^T \mathbf{X} \mathbf{V})$ . Define the sparseness-inducing regularization  $\|\mathbf{X} - \mathbf{X}_E\|_F$  by adding a penalty term  $\lambda \Phi(\mathbf{X})$ . The typical resulting low-rank approximation problem can be written as follows:

$$\begin{aligned} \hat{\mathbf{X}} &= \underset{\mathbf{X}_E}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{X} - \mathbf{X}_E\|_F^2 + \lambda \Phi(\mathbf{X}_E) \right\} \\ &= \underset{\mathbf{X}_E}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\Sigma - \mathbf{U}^T \mathbf{X}_E \mathbf{V}\|_F^2 + \lambda \Phi(\mathbf{U}^T \mathbf{X}_E \mathbf{V}) \right\}, \\ &= \mathbf{U} \cdot \underset{\mathbf{X}_E}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{X} - \mathbf{X}_E\|_F^2 + \lambda \Phi(\mathbf{X}_E) \right\} \mathbf{V}^T \end{aligned} \quad (14)$$

where  $\lambda$  is a positive regularization parameter balancing the least-square fitting term,  $\Phi(\mathbf{X}_E) = \sum_{i=1}^k \phi(\sigma_i(\mathbf{X}_E); \alpha)$ ,  $\phi$  is a sparseness-inducing function,  $\sigma_i(\mathbf{X}_E)$  is the eigenvalue of  $\mathbf{X}_E$ , and  $0 \leq \alpha \leq 1/\lambda$ . When the function  $\phi(x)$  is represented by  $|x|$ , the convex issue of low-rank approximation problem for rank minimization turns to a typical nuclear norm problem [24].

Define that  $\Theta(\Sigma, \lambda, \alpha) = \underset{\mathbf{X}_E}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{X} - \mathbf{X}_E\|_F^2 + \lambda \Phi(\mathbf{X}_E) \right\}$ .

Equation (14) can be equivalently written as:

$$\hat{\mathbf{X}} = \mathbf{U} \cdot \Theta(\Sigma, \lambda, \alpha) \cdot \mathbf{V}^T. \quad (15)$$

Dang [24] and Selesnick [25] proved that parametric non-convex penalty function method can be more accurate than using nuclear norm in estimating non-zero singular values, when the signal contains a large noise level. Therefore, the non-convex penalty function is used to obtain the optimized estimation of singular value as the following equation:

$$\Theta(\Sigma, \lambda, \alpha) = \min\{|\Sigma|, \max\{(|\Sigma| - \lambda)/(1 - \alpha\lambda), 0\}\} \operatorname{sign}(\Sigma). \quad (16)$$

In this paper, the optimal rank  $r$  of singular value, which should be determined in the second step of the DMD algorithm, is obtained by the convex optimization framework.

### 3.2. Dominant Modes Selection Strategy

Reconstruction matrix of DMD is obtained by a superposition of  $\phi_i \exp(\omega_i \Delta t) b_i$ , and  $\omega_i$  is continuous-time eigenvalues of the similarity matrix  $\tilde{\mathbf{A}}$ , which constitutes a diagonal matrix. Every single  $\omega_i$  corresponds to a DMD mode  $\phi_i$ . In this part, we introduce a method to select optimal dominant modes for DMD with two steps. Firstly, subtract modes relating to  $\|\omega_i\| = \|\ln(\lambda_i)/\Delta t\| \approx 0$  from all the dynamic features, and then define a hierarchical application of basic DMD algorithm.

Step 1. By the construction of DMD methodology,  $b = \Phi^T \mathbf{X}$ , which means that  $\phi b$  renders the columns of  $\mathbf{X}$  with a dimensionality reduction. Thus, the diagonal matrix of frequencies  $\lambda_i$  dictates how the first period of signal  $\mathbf{X}$  gets altered over time to reconstruct the subsequent periods. It becomes apparent that any period of the signal that does not change or changes very slowly in time must have an associated Fourier mode located near the first period with the frequency  $\|\omega_i\| \approx 0$ . We regard the slow oscillations as the background modes (noises) by setting a threshold  $\zeta$ , and produce a representation

of the low-frequency reconstruction matrix  $\mathbf{X}_{DMD}^{\|\omega_i\| \leq \zeta}$  and remaining frequency reconstruction matrix  $\mathbf{X}_{DMD}^{\|\omega_i\| > \zeta}$  of the form:

$$X_{DMD} = X_{DMD}^{|\omega_i| \leq \zeta} + X_{DMD}^{|\omega_i| > \zeta} = \sum_{|\omega_i| \leq \zeta} \phi_i \exp(\omega_i \Delta t) b_i + \sum_{|\omega_i| > \zeta} \phi_i \exp(\omega_i \Delta t) b_i. \tag{17}$$

This observation makes it possible to pick out background modes from all the dynamic features with DMD, thus we only take the remaining frequency reconstruction matrix into account in this first step; thus,

$$X_{DMD} \approx X_{DMD}^{|\omega_i| > \zeta} = \sum_{|\omega_i| > \zeta} \phi_i \exp(\omega_i \Delta t) b_i. \tag{18}$$

Step 2. A hierarchical DMD is introduced to split the reconstructed signal with several levels, inspired by the idea of wavelet packet decomposition. Mathematically, we divide the solution of the DMD approximate solution of Equation (18) into slow modes and fast modes, by splitting the number of  $\omega_i > \zeta$  at its median:

$$\begin{aligned} X_{DMD}^{(1)} &= \sum_{i=1}^{r_1} \phi_i^{(1)} \exp(\omega_i^{(1)} \Delta t) b_i^{(1)} \\ &= \sum_{i=1}^{r_1/2} \phi_i^{(1)} \exp(\omega_i^{(1,1)} \Delta t) b_i^{(1)} + \sum_{i=(r_1/2+1)}^{r_1} \phi_i^{(1)} \exp(\omega_i^{(1,2)} \Delta t) b_i^{(1)}, \tag{19} \\ &= X_{DMD}^{S(1)} + X_{DMD}^{F(1)} \end{aligned}$$

where the symbol (1) at the top right corner of  $X_{DMD}^{(1)}$  represents the first level, and  $r_1$  equals the optimal rank  $r$  minus the number of  $\|\omega_i\| \leq \zeta$ .  $X_{DMD}^{S(1)}$  and  $X_{DMD}^{F(1)}$  are the slow modes matrix and fast modes matrix, respectively. Then, we apply the second segmentation on each of them:

$$\begin{aligned} X_{DMD}^{(2)} &= \sum_{i=1}^{r_1^1/2} \phi_i^{(2,1)} \exp(\omega_i^{(2,1)} \Delta t) b_i^{(2,1)} + \sum_{i=r_1^1/2+1}^{r_1^1} \phi_i^{(2,2)} \exp(\omega_i^{(2,2)} \Delta t) b_i^{(2,2)} + \\ &\quad \sum_{i=1}^{r_1^2/2} \phi_i^{(2,3)} \exp(\omega_i^{(2,3)} \Delta t) b_i^{(2,3)} + \sum_{i=r_1^2/2+1}^{r_1^2} \phi_i^{(2,4)} \exp(\omega_i^{(2,4)} \Delta t) b_i^{(2,4)}, \tag{20} \\ &= X_{DMD}^{SS(2,1)} + X_{DMD}^{FS(2,2)} + X_{DMD}^{SF(2,3)} + X_{DMD}^{FF(2,4)} \end{aligned}$$

where the symbol (SS) at the top right corner represents slow modes in the first level and also the same in the second level, while (SF) represent fast modes in the first level and slow modes in the second level.

The solution (20) can be made more precise by accounting for more numbers of segmentation; here, we say level ( $L$ ). The arbitrary  $L$ -th DMD modes can be presented as follows:

$$\mathbf{X}_{DMD}^{(L)} = \sum_l^L \sum_j^J \sum_{i=1}^{r_l} \phi_i^{(l,j)} \exp(\omega_i^{(l,j)} \Delta t) b_i^{(l,j)}, \tag{21}$$

where  $l = 1, 2, \dots, L$  represents the decomposition levels,  $j = 1, 2, \dots, 2^{l-1}$  represents the number of bins per level, and  $r_l$  represents the truncation rank of  $l$  level.

Figure 1 demonstrates the process in terms of the multiresolution DMD solution. Each  $X_{DMD}^{(l,j)}$  can be obtained in each level  $l$  and each bin  $j$  by linear operator  $\mathbf{A}^{(l,j)}$ :

$$\mathbf{Y}^{(l,j)} = \mathbf{A}^{(l,j)} \mathbf{X}^{(l,j)}. \tag{22}$$

The solution minimizes the error yielding the least-square fit. It can be given similarly by Equation (5).

As in the fifth step of DMD core algorithm in Section 2, Fourier transform can be applied on the recuperative signal for spectrogram with each  $X_{DMD}^{(l,j)}$ , determining whether there is a failure frequency.

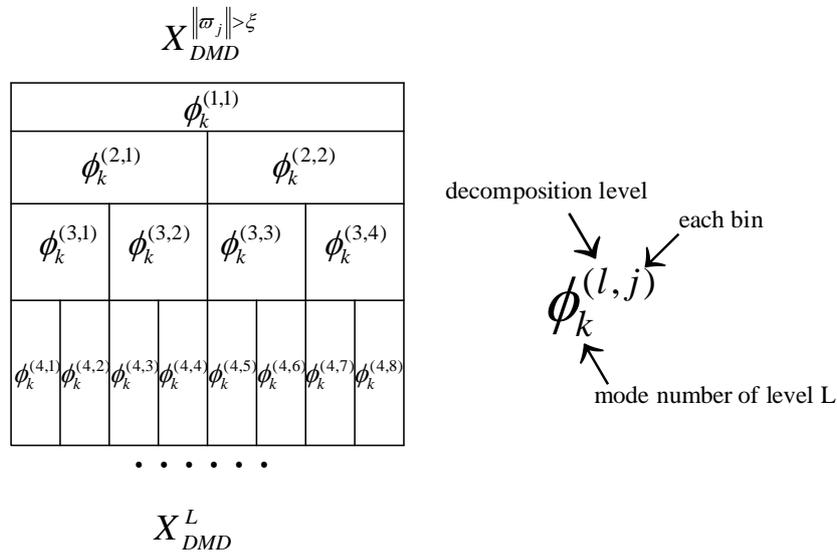


Figure 1. Illustration of the hierarchical DMD solution.

Rewrite the Equation (21) as follows:

$$\begin{aligned}
 X_{DMD}^{(L)} &= \sum_l \sum_j \sum_{i=1}^{r_l} g^{(l,j)} \phi_i^{(l,j)} \exp(\omega_i^{(l,j)} \Delta t) b_i^{(l,j)} \\
 &= g^{(l,j)} \Phi^{(l,j)} \exp(\Omega^{(l,j)} t) b^{(l,j)}
 \end{aligned}
 \tag{23}$$

Define that  $g^{(l,j)} = 1$  when there is a failure frequency for  $X_{DMD}^{(l,j)}$ , accordingly,  $g^{(l,j)} = 0$  when there is no failure frequency.  $g^{(l,j)}$  acts as a sifting operation for each bin  $j$  at level  $l$ .

Ultimately, the flowchart of the improved DMD method is shown in Figure 2.

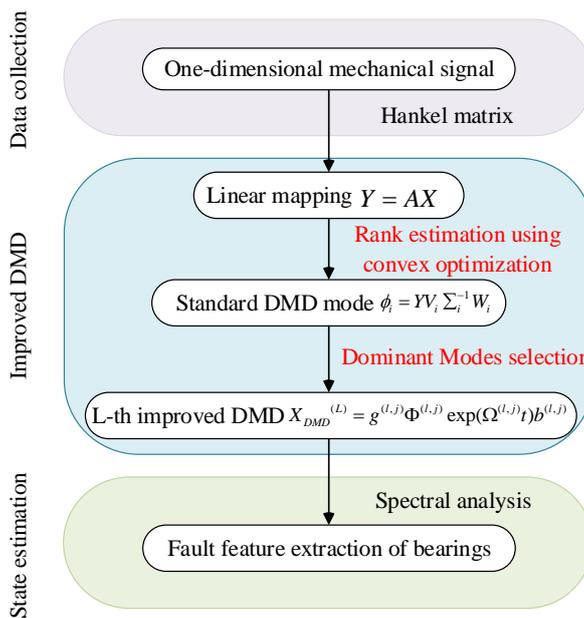


Figure 2. A flow chart of applying improved DMD.

## 4. Bearing Fault Signal Processing with Improved DMD

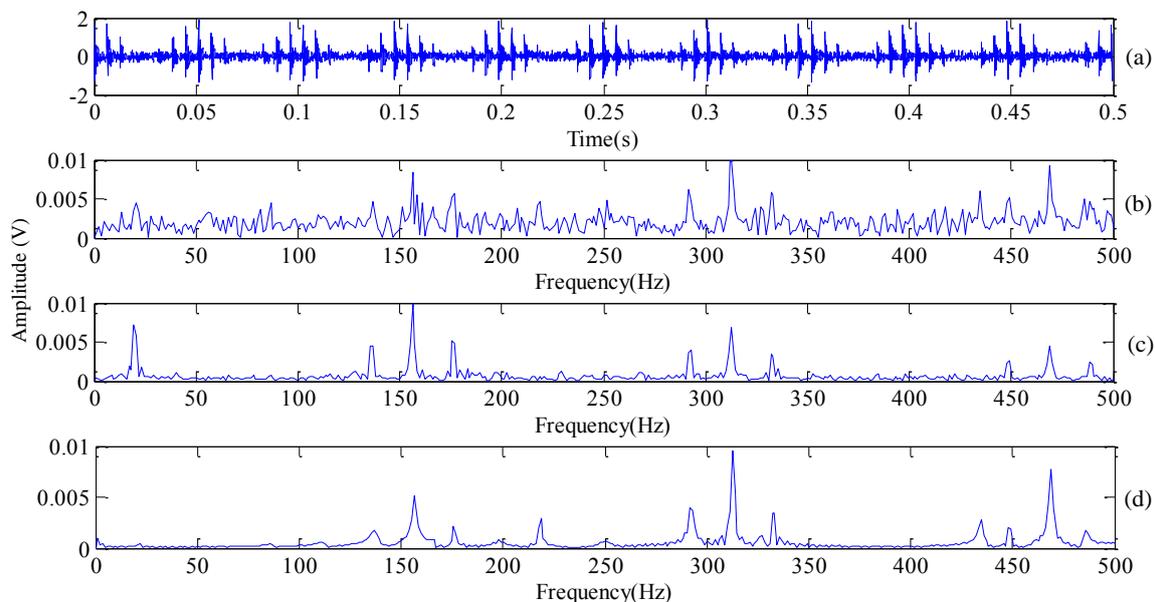
### 4.1. Analysis of Simulated Signal

There are many kinds of dynamic models simulating bearing fault signal. Comparatively speaking, the model first established by McFadden [26] and later refined by Randall [27] can better reflect the characteristics of rolling bearing vibration. According to the structure characteristics, short-term impact signal should be produced when bearing parts revolve through the fault location. The impact motivates the bearing system oscillating by its natural frequency with high frequency attenuation. Define  $t$  as a period of two intervals,  $S(t)$  as the natural frequency oscillation function, and  $A_k$  as the  $i$ -th amplitude of shock response. The interference of additive noise  $n_i$  (zero mean additive background noise) should be taken into consideration along with the simulation signal in view of the poor working environment of rolling bearings. Therefore, the mathematical model of fault rolling bearing can be described as:

$$\begin{cases} f(t) = \sum_{k=1}^M A_k S(t - kT - \tau_k) + n(t) \\ A_k = a_k \cos(2\pi f_m t + \varphi_A) + c_A \\ S(t - kT - \tau_k) = \exp[-B(t - kT - \tau_k)] \cdot \sin[2\pi f_n(t - kT - \tau_k) + \gamma] \end{cases}, \quad (24)$$

where  $a_k$  is the  $k$ -th impact energy;  $\gamma$  and  $\varphi_A$  are the initial phases;  $f_m$  is the modulation frequency;  $B$  is the attenuation coefficient depended on the bearing system;  $\tau_k$  is the time lag from its mean period due to the presence of tiny fluctuation, and  $c_A$  is a random constant.

Most frequently, the fault of rolling bearing appears as the damage in outer race, inner race, and/or the surface of rolling elements. Generally, three common types of faults on outer race, inner race and rolling elements can be modeled as  $f_m = 0$ ,  $f_m = f_r$ ,  $f_m = f_c$ , respectively. Here,  $f_r$  and  $f_c$  are denoted as rotational frequency and cage frequency. The time-domain fault simulation signal of the inner race is plotted in Figure 3a, added by the Gaussian white noise so the SNR ratio is 6. The parameters selection of Equation (24) is listed in Table 1.



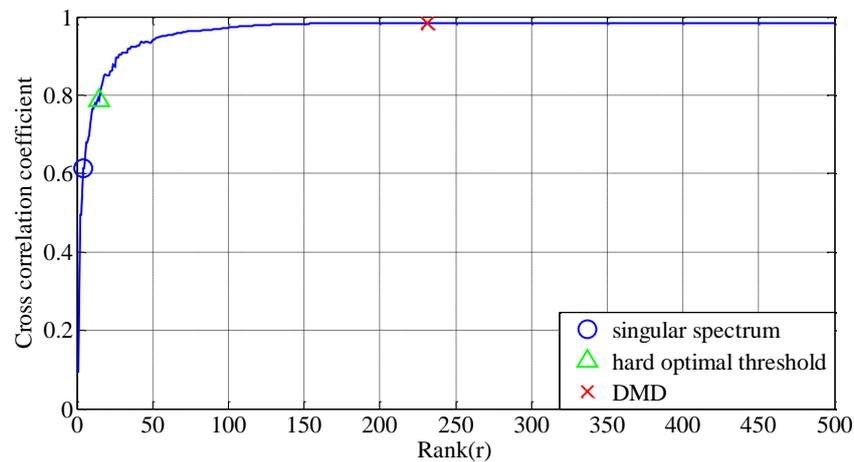
**Figure 3.** Improved DMD analysis with simulation signal of bearing inner race (a) time domain fault simulation signal; (b) FFT spectrum; (c) envelope spectrum; (d) frequency spectrum with improved DMD.

**Table 1.** The parameters selection with simulation signal of bearing inner race.

$a_k$	$\mathbf{fl}$	$'_A$	$f_r(\text{Hz})$	$\mathbf{B}$	$\boldsymbol{\sigma}_k$	$c_A$	<i>Inner Race</i> $f_i(\text{Hz})$
3	0	0	20	800	0.01	1	156

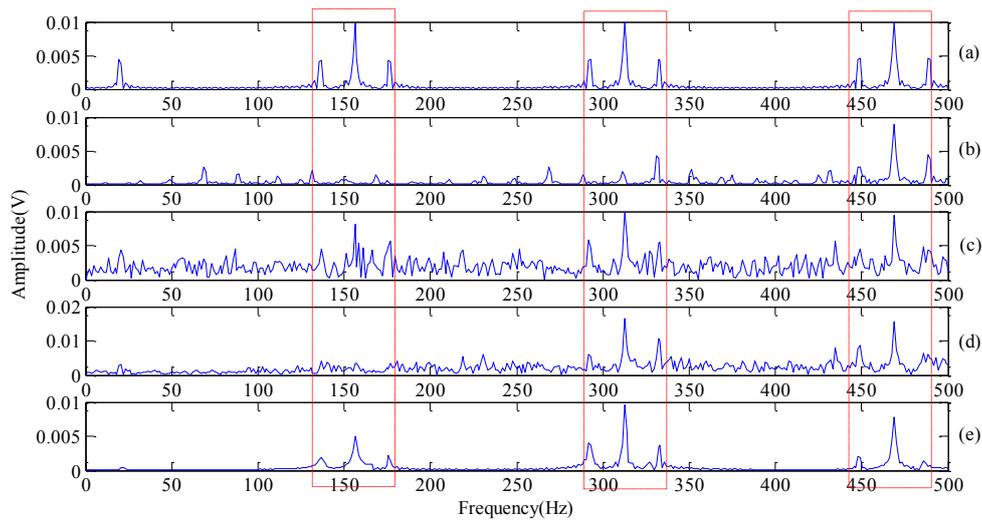
We choose  $\xi = 0.04$  (retain 90% of  $\omega_i$ ) and take the decomposing level  $L$  as 2. Frequency spectrum with improved DMD plotted in Figure 3d, while FFT frequency and envelope spectrum are also plotted for comparison in Figure 3b,c, respectively. Inner race fault frequency can be identified, but the modulation frequency (FM) is not clear connected to the intuition with the FFT frequency spectrum. Inner race frequency and rotational modulation frequency are visibly presented in the envelope spectrum, but there is still a certain amount of noise frequency surrounding the principal frequencies. By adopting our improved DMD, not only can the frequency of inner race fault be clearly visualized, but there is also a little amount of noise outlined in Figure 3d compared with Figure 3c.

Variations of cross correlation coefficient along with the truncated rank order are shown in Figure 4. The best truncation rank number of SVD, hard optimal threshold and convex optimization DMD were also marked on the curve. It can be concluded that the sparse optimization framework outperforms the other methods in the quality of signal recovery.



**Figure 4.** Cross correlation coefficient between DMD reconstruction signal and original simulation signal with different truncation rank  $r$ . Cross correlation coefficient is first ascending and then slightly declining. The rank ( $r = 4$ ) in blue circle point obtained by singular spectrum; the rank ( $r = 13$ ) with the green triangle point obtained by the hard optimal threshold; the rank ( $r = 231$ ) with a red cross mark obtained by our method.

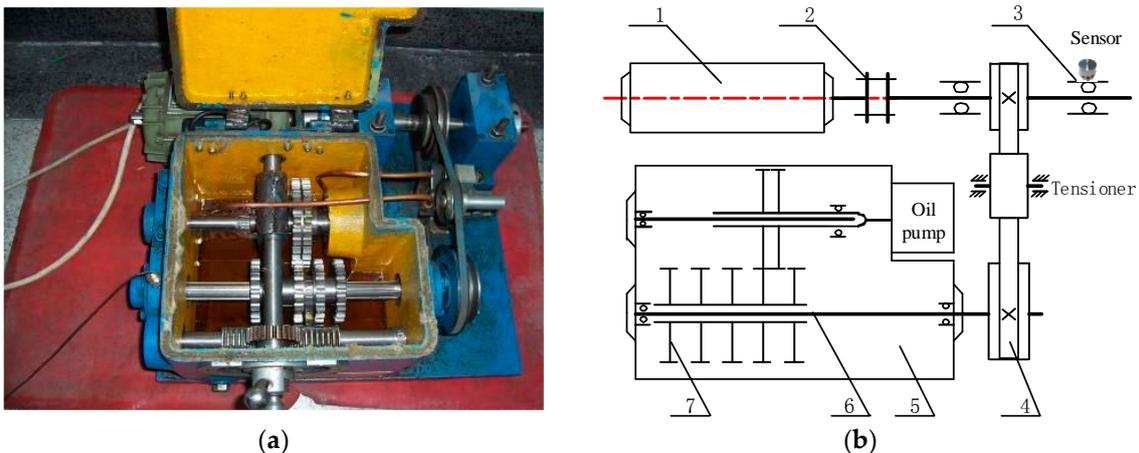
Comparisons between improved DMD with other conventional spectral analysis methods are characterized in Figure 5. Figure 5a is the frequency spectrum of the original situation signal without noise taking by FFT. Figure 5b is the result of SVD, and we select the first four singular values for signal reconstruction. Figure 5c is the multi-scale WT signal spectrum; here, we use wavelet function db4 with three scale wavelet decomposition [5], and choose the high frequency component of the second layer for signal reconstruction. Figure 5d is the frequency spectrum by applying FFT on superposed IMFs of EMD, and we choose IMFs of 1, 4, 5, 6 for superposition. Figure 5e was obtained by our improved DMD method, while  $\zeta = 0.04$ ,  $L = 3$ . SNR of the reconstruction signal is calculated as 7.8, 9.5, 6.6, 14.2 from Figure 5b–e successively. Our method has a tremendous advantage in both denoising and fault feature extraction.



**Figure 5.** Frequency spectrum with different methods (a) FFT of noise-free signal; (b) SVD; (c) WT; (d) EMD; (e) improved DMD.

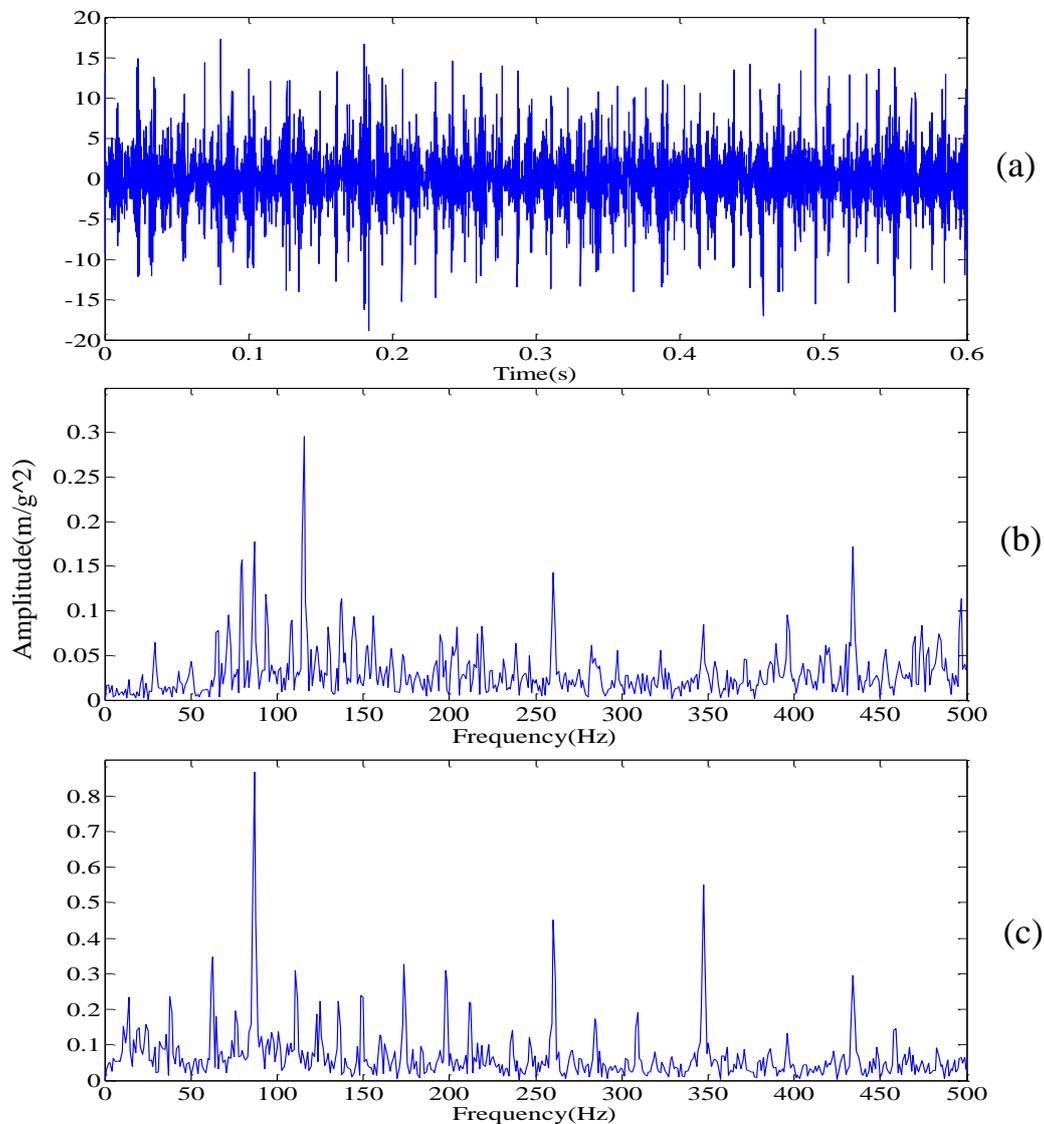
#### 4.2. Analysis of Measured Signal

Vibration signals of the bearing system in an industrial spot are more complex than the dynamic simulation signal [28]. To verify the effect of fault features extraction, the improved DMD algorithm is applied on real environment fault signals. An integrated fault apparatus of gear-bearing transmission is shown in Figure 6. The high-speed shaft is driven by an AC motor (550 w, 220 v/50 Hz) with a belt, and the speed of the motor is 1450 r/min. A corrosion pit with the depth of 1.0 mm was fabricated by electric spark processing in the out race of the testing bearing (deep groove ball bearing, bearing number 6207). An acceleration sensor (PCB-352C33, Depew, NY, USA) is installed on the bearing pedestal along the vertical direction. The vibration signal is then obtained by a machinery health analyzer CSI2130 (Emerson, St. Louis, MO, USA) with sampling frequency of 16,384 Hz. The number of rolling elements is  $Z = 9$ , and ball diameter is  $d = 11.1$  mm, pitch diameter of the testing bearing is  $D = 53.5$  mm. Thus, rotation frequency and outer ring fault frequency can be calculated as 24.17 Hz and 87.01 Hz, respectively [29].



**Figure 6.** Gear-bearing comprehensive failure device: (a) test rig general view; (b) structure diagram of test device: 1—motor, 2—coupling, 3—test rolling bearing from 6207 series, 4—belt pulley, 5—housing, 6—transmission shaft, 7—gear.

A total number of 10,000 samples, shown in Figure 7a in the time domain, are analyzed with improved DMD. The FFT spectrum is shown in Figure 7b. Pang [30] proved the superiority of the envelope spectrum over the FFT in the incipient rolling bearing fault diagnosis. For comparison, the envelope spectrum of the measured signal is shown in Figure 7c.



**Figure 7.** The measured signal in the time domain and frequency spectrum: (a) time domain; (b) FFT spectrum; (c) envelope spectrum.

Obviously, we can't judge whether there is a bearing fault from Figure 7b for a large amount of interference frequencies presented in the frequency spectrum. Figure 7c is slightly better than Figure 7b, but there are still some noise frequencies presented in the frequency spectrum. Meanwhile, in order to illustrate the effectiveness of our method, the frequency spectrum of WPT is applied and the results are shown in Figure 8. Here, we use wavelet function db15 with eleven layers of decomposition. Though it is obvious that the fault frequency  $f_o$  and its harmonics ( $2f_o$ ,  $3f_o$ ,  $4f_o$ ,  $5f_o$ ) can be identified in Figure 8, where  $f_o$  is the frequency of the outer race, the noise is still present in the acquired signal, as we can see in the red elliptical mark.

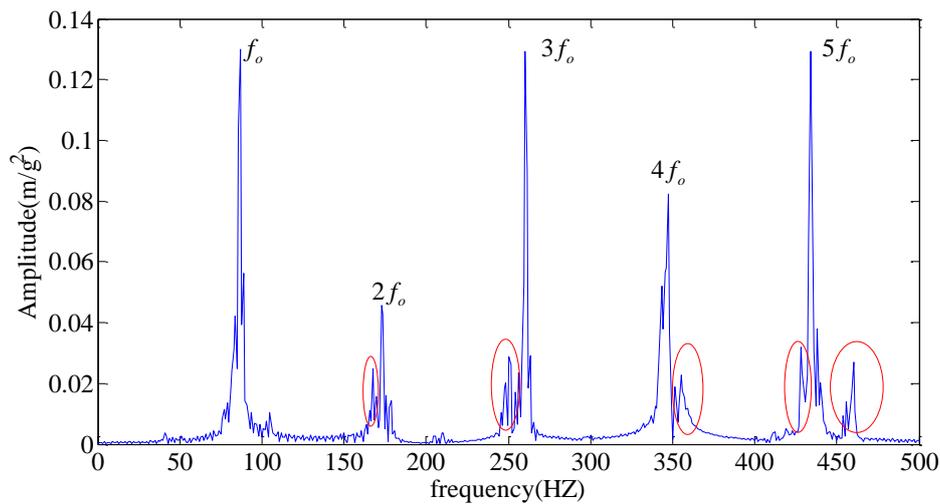


Figure 8. Frequency spectrum of WPT.

Subsequently, we apply our improved DMD method on the measured signal with chosen parameters  $m = 5000$ ,  $\xi = 0.05$ ,  $L = 7$ . The non-fault signal modes are filtered out and the useful signals with characteristic frequencies are retained by examining each bin (total sixty-four pins) of the seventh level. Then, the reconstruction matrix can be obtained by Equation (23). Figure 9 shows the spectrum of our improved DMD method. Obviously, not only are the fault frequency  $f_o$  and its harmonics ( $2f_o$ ,  $3f_o$ ,  $4f_o$ ,  $5f_o$ ) easy to be identified, but there is also little interference frequency close to them. That is to say, we are more confident to determine that the outer race of the testing rolling bearing has failed, which is highly consistent with the actual situation.

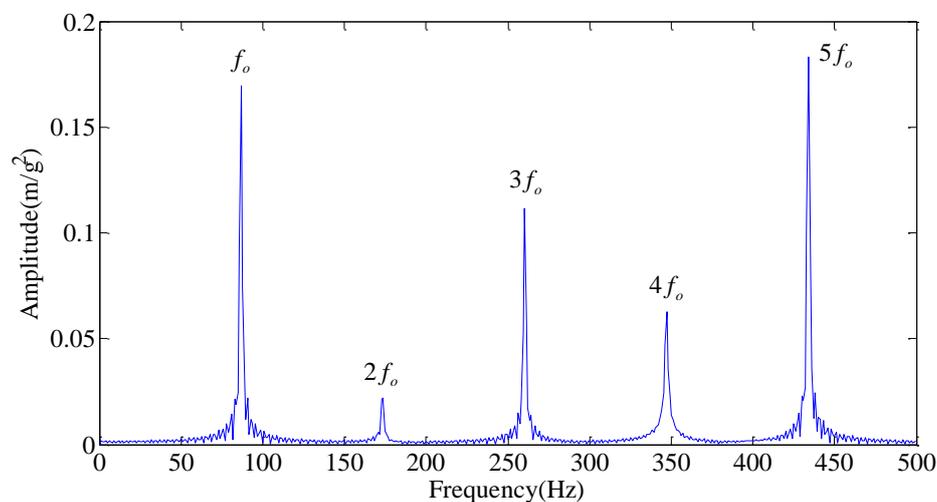


Figure 9. The result provided by improved DMD.

## 5. Conclusions

This paper puts forward an improved DMD decomposition algorithm, which successfully applied on the bearing fault signal. Main solution of our work includes the following two aspects: (1) optimal rank truncation number of the similarity matrix is obtained by using convex optimization through introducing a parameterized non-convex penalty function. Our rank determinative method is more perfectible than others in feature extraction with standard DMD algorithms; (2) in order to archive the dominant modes selection strategy, we firstly subtract slow oscillations modes by setting a threshold, and then define a hierarchical application of the basic DMD algorithm. We apply our improved DMD

framework both on simulated and measured signals for roller bearing signal fault diagnosis. SNR is better than other classical noise reduction methods after the bearing simulation signal is processed by our improved DMD algorithm. In addition, the frequency spectrum of the reconstruction signal is more readable and interference-free. Results show that our method has a good application prospect in denoising and feature extraction with a mechanical signal.

**Author Contributions:** Z.D. was in charge of contributing the analysis method and the program codes, analyzing the data, and writing the paper. Y.L. and Y.L. were in charge of the supervision of this research and helping with revision. C.Y. was responsible for the experimental design and data analysis. All authors have read and approved the final manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China under Grants No. 51475339 and No. 51575408, and the Natural Science Foundation of Hubei province under Grant No. 2016CFA042. The Applied Basic Research Programs of the Wuhan Science and Technology Bureau under Grant No. 2017010201010115.

**Conflicts of Interest:** The authors declare that there is no conflict of interests regarding the publication of this article.

## References

1. Lv, Y.; Zhu, Q.; Yuan, R. Fault diagnosis of rolling bearing based on fast nonlocal means and envelop spectrum. *Sensors* **2015**, *15*, 1182–1198. [[CrossRef](#)] [[PubMed](#)]
2. Golafshan, R.; Sanliturk, K.Y. SVD and Hankel matrix based de-noising approach for ball bearing fault detection and its assessment using artificial faults. *Mech. Syst. Signal Process.* **2016**, *70–71*, 36–50. [[CrossRef](#)]
3. Griffin, D.; Lim, L. Signal estimation from modified short-time Fourier transform. *IEEE Trans. Acoust. Speech Signal Process.* **1984**, *32*, 236–243. [[CrossRef](#)]
4. Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory* **1990**, *36*, 961–1005. [[CrossRef](#)]
5. Hemmati, F.; Wasim, O.; Mohamed, S.G. Roller bearing acoustic signature extraction by wavelet packet transform, applications in fault detection and size estimation. *Appl. Acoust.* **2016**, *104*, 101–118. [[CrossRef](#)]
6. Huang, N.; Shen, Z.; Long, S.; Wu, M.; Hsing, H.; Zheng, Q.; Yen, N.; Chi, C.; Liu, H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **1998**, *451*, 903–995. [[CrossRef](#)]
7. Lv, Y.; Yuan, R.; Song, G. Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing. *Mech. Syst. Signal Process.* **2016**, *81*, 219–234. [[CrossRef](#)]
8. Vautard, R.; Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Phys. D* **1989**, *35*, 395–424. [[CrossRef](#)]
9. Yuan, R.; Lv, Y.; Song, G. Multi-Fault Diagnosis of Rolling Bearings via Adaptive Projection Intrinsically Transformed Multivariate Empirical Mode Decomposition and High Order Singular Value Decomposition. *Sensors* **2018**, *4*, 1210. [[CrossRef](#)] [[PubMed](#)]
10. Jonathan, H.T.; Rowley, C.W.; Luchtenburg, D.M.; Brunton, S.L.; Kutz, J.N. On dynamic mode decomposition: Theory and applications. *J. Comput. Dyn.* **2013**, *1*, 391–421. [[CrossRef](#)]
11. Schmid, P.J.; Li, L.; Juniper, M.P.; Pust, O. Applications of the dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **2011**, *25*, 249–259. [[CrossRef](#)]
12. Schmid, P.J. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **2010**, *656*, 5–28. [[CrossRef](#)]
13. Hemati, M.S.; Williams, M.O.; Rowley, C.W. Dynamic mode decomposition for large and streaming datasets. *Phys. Fluids* **2014**, *26*, 5–28. [[CrossRef](#)]
14. Rowley, C.W.; Mezic, I.; Bagheri, S.; Henningson, D.S. Spectral analysis of nonlinear flows. *J. Fluid Mech.* **2009**, *641*, 115–127. [[CrossRef](#)]
15. Dawson, S.; Hemati, M.; Williams, M.; Rowley, C.W. Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Exp. Fluids* **2016**, *57*, 1–19. [[CrossRef](#)]
16. Kutz, J.N.; Fu, X.; Brunton, S.L.; Erichson, N.B. Multi-resolution dynamic mode decomposition for foreground/background separation and object tracking. In Proceedings of the IEEE International Conference on Computer Vision Workshop, Santiago, Chile, 7–13 December 2015; pp. 921–929.

17. Hemati, M.S.; Rowley, C.W.; Deem, E.A.; Cattafesta, L.N. De-biasing the dynamic mode decomposition for applied koopman spectral analysis. *J. Nonlinear Sci.* **2015**, *25*, 1–40. [[CrossRef](#)]
18. Erichson, N.B.; Brunton, S.L.; Kutz, J.N. Compressed dynamic mode decomposition for background modeling. *J. Real-Time Image Process.* **2016**, 1–14. [[CrossRef](#)]
19. Jovanovic, M.; Schmid, P. Sparsity-promoting dynamic mode decomposition. *Phys. Fluids* **2014**, *26*, 561–571. [[CrossRef](#)]
20. Erichson, N.B.; Donovan, C. Randomized low-rank Dynamic Mode Decomposition for motion detection. *Comput. Vis. Image Underst.* **2016**, *146*, 40–50. [[CrossRef](#)]
21. Gavish, M.; Donoho, D.L. The optimal hard threshold for singular values is  $4/\sqrt{3}$ . *IEEE Trans. Inf. Theory* **2014**, *60*, 5040–5053. [[CrossRef](#)]
22. Kou, J.; Zhang, W. An improved criterion to select dominant modes from dynamic mode decomposition. *Eur. J. Mech.* **2017**, *62*, 109–129. [[CrossRef](#)]
23. Dang, Z.; Lv, Y.; Li, Y.; Yi, C. Optimized Dynamic Mode Decomposition via Non-Convex Regularization and Multiscale Permutation Entropy. *Entropy* **2018**, *20*, 152. [[CrossRef](#)]
24. Yi, C.; Lv, Y.; Dang, Z.; Xiao, H.; Yu, X. Quaternion singular spectrum analysis using convex optimization and its application to fault diagnosis of rolling bearing. *Measurement* **2017**, *103*, 321–332. [[CrossRef](#)]
25. Selesnick, I.W.; Bayram, İ. Enhanced sparsity by non-separable regularization. *IEEE Trans. Signal Process.* **2015**, *64*, 2298–2313. [[CrossRef](#)]
26. Mcfadden, P.D.; Smith, J.D. Model for the vibration produced by a single point defect in a rolling element bearing. *J. Sound Vib.* **1984**, *96*, 69–82. [[CrossRef](#)]
27. Randall, R.B.; Antoni, J.; Chobsaard, S. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mech. Syst. Signal Process.* **2001**, *15*, 945–962. [[CrossRef](#)]
28. Hao, Y.; Song, L.; Ke, Y.; Wang, H.; Chen, P. Diagnosis of Compound Fault Using Sparsity Promoted-Based Sparse Component Analysis. *Sensors* **2017**, *17*, 1307. [[CrossRef](#)] [[PubMed](#)]
29. Han, T.; Liu, X.; Tan, A. Fault diagnosis of rolling element bearings based on Multiscale Dynamic Time Warping. *Measurement* **2017**, *95*, 355–366. [[CrossRef](#)]
30. Bin, P.; Guiji, T.; Tian, T.; Chong, Z. Rolling Bearing Fault Diagnosis Based on an Improved HTT Transform. *Sensors* **2018**, *18*, 1203. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).