

---

---

# A Fault Diagnosis Approach for Rolling Bearing Based on Wavelet Packet Decomposition and GMM-HMM

**Liangpei Huang**

*Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan, 411201, China.*

*School of Electromechanical Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China.*

**Hua Huang**

*School of Electromechanical Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China.*

**Yonghua Liu**

*Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan, 411201, China.*

*School of Electromechanical Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China.*

(Received 1 August 2015; accepted 19 February 2018)

Considering frequency domain energy distribution differences of bearing vibration signal in the different failure modes, a rolling bearing fault pattern recognition method is proposed based on orthogonal wavelet packet decomposition and Gaussian Mixture Model-Hidden Markov Model (GMM-HMM). The orthogonal three-layer wavelet packet decomposition is used to obtain wavelet packet decomposition coefficients from low frequency to high frequency. Rolling bearing raw vibration signals are firstly decomposed into the wavelet signals of different frequency bands, then different frequency band signals are reconstructed respectively to extract energy features, which form feature vectors as the model input of GMM-HMM. A large number of samples are trained to get model parameters for different bearing faults, then several groups of test data are adopted to verify GMM-HMMs so different fault types of rolling bearings are recognized. By calculating the current state appearance probability of monitoring data in GMM-HMMs, different failure patterns are recognized and evaluated from the maximum probability. Similarly, we establish GMM-HMMs for different grade fault samples and evaluated the performance degradation state. Test results show that the proposed fault diagnosis approach can identify accurately the fault pattern of rolling bearings and evaluate performance degradation of bearings.

---

## 1. INTRODUCTION

Safe and reliable operation of the mechanical equipment has aroused wide attention and great interest of researchers in recent years, so we usually carry on condition monitoring, fault diagnosis of early stages, and failure identification and prediction to ensure safe and reliable operation of the mechanical equipment. Rolling bearings are key components of rotating machinery; the fault and failure of rolling bearings can cause abnormal vibration and noise, even direct destruction for mechanical equipment, and affect safe and reliable operation. According to statistics, among the rotating machinery about 30% of mechanical failures are associated with rolling bearing damages.<sup>1</sup> The diagnosis technology on rolling bearing fault has always been a hot spot for many research scholars. Xia Ruihua applied time domain analysis method to feature vector extraction of rolling bearing vibration signal and developed fault diagnosis technology for rolling bearings.<sup>2</sup> Huang Zhonghua proposed a fault diagnosis method of rolling bearings with

Hilbert transformation considering the modulation feature of vibration signal.<sup>3</sup> In order to solve the application problem of traditional resonance demodulation technique, which is seriously impacted by the low signal to noise ratio of rolling bearing original vibration and the choice of filter parameters depending on manipulator's subjective experience, Zhou Zhi presented the method of combining adaptive noise reduction with adaptive resonance demodulation.<sup>4</sup> These studies have made some progress, but they still fail to break through these disadvantages of time domain analysis, which has lower frequency resolution at high frequencies and has poor temporal resolution at low frequencies.<sup>5</sup> Wavelet analysis is a localized method of time frequency analysis, and it can gradually make the signal refined on multi-scale by the telescopic pan operations and ultimately achieve time segments at high frequency and frequency segments at low frequency, which can automatically adapt to the requirements of frequency signal analysis.<sup>6</sup>

With strict data structure and reliable computing capabil-

ity, Hidden Markov Model (HMM) has been widely applied in speech recognition, mouth recognition, face detection, and other fields. In recent years, some researchers have begun to introduce HMM to fault diagnosis and condition monitoring.<sup>7</sup> Xiao Wenbin proposed a method to assess performance degradation of bearings using Wavelet transform and HMM.<sup>8</sup> This method used wavelet transform to analyze the bearing vibration signal, extract the total node energy as a feature, establish a performance degradation assessment model, and make a quantitative assessment about the extent of degradation of the bearing. S.S.H. Zaidi used the time-frequency features extracted from the motor current as machine health indicators and predicted the future state of fault severity using HMMs.<sup>9</sup> Jianbo Yu proposed an adaptive learning method for machine faulty detection and health degradation monitoring, and adaptive HMM is used for online learning the dynamic health changes of machines in their full life.<sup>10</sup> Mitchell Yuwono, et al. proposed an automatic bearing defect diagnosis method based on Swarm Rapid Centroid Estimation and HMM using the defect frequency signatures extracted with Wavelet Kurtogram and Cepstral Liftering.<sup>11</sup> Haitao Zhou, et al. developed a fault diagnosis model for rolling element bearings based on shift-invariant dictionary learning and HMM.<sup>12</sup> Zefang Li developed a likelihood-based Continuous Hidden Markov Model (CHMM) for averaging estimator, and bearing fault diagnosis is effectively performed.<sup>13</sup> Omid Geramifard introduced a semi-nonparametric approach based on HMM for fault detection and diagnosis in synchronous motors.<sup>14</sup> V. Purushotham presented a new method for detecting localized bearing defects based on wavelet transform.<sup>15</sup> Bearing race faults have been detected by using discrete wavelet transform.

However, these studies do not deeply reveal the inherent regularity and particularity of HMM in mechanical fault diagnosis, especially for Gaussian Mixture Model-Hidden Markov Model (GMM-HMM). In this paper, we present a rolling bearing fault pattern recognition method based on wavelet packet decomposition and GMM-HMM. Because the energy features for vibration signals of different faults have differences in different frequency bands, wavelet packet decomposition algorithms are proposed to extract energy feature vectors of rolling bearings, which provide input feature vectors of GMM-HMM. After model train and verification, fault pattern recognition and degradation evaluation of rolling bearings can be effectively realized.

## 2. THEORETICAL FOUNDATION

### 2.1. Wavelet Packet Decomposition Algorithm

The basic idea of wavelet packet analysis is as follows:<sup>16</sup> assuming that the discrete approximation  $A_j f(t)$  has been calculated for the signal  $f(t) \in L^2(R)$  in the resolution of  $2^j$  and the discrete approximation  $A_{j-1} f(t)$  of  $f(t)$  in the resolution of  $2^{j-1}$  can be obtained from  $A_j f(t)$  by discrete low-pass filter  $H$ , according to Mallet algorithm,  $A_j f(t)$  is decomposed

into rough image  $A_{j-1} f(t)$  and detail  $D_{j-1} f(t)$ ,

$$A_j f(t) = A_{j-1} f(t) + D_{j-1} f(t). \quad (1)$$

Once the scaling function  $\phi(t)$  is determined for multi-resolution analysis, the wavelet function  $\psi(t)$  can also be constructed. If the scaling function is orthogonal, this multi-resolution analysis is called orthogonal analysis. Mallet algorithm is based on orthogonal multi-resolution analysis. Discrete approximation  $A_j f(t)$  and detail  $D_j f(t)$  can be expressed as:

$$A_j f(t) = \sum_{k=-\infty}^{+\infty} c_{j,k} \phi_{j,k}(t); \quad (2)$$

$$D_{j-1} f(t) = \sum_{k=-\infty}^{+\infty} d_{j,k} \psi_{j,k}(t); \quad (3)$$

where  $c_{j,k}$  and  $d_{j,k}$  are respectively scale coefficients and wavelet coefficients.

According to the scalability and inclusiveness of multi-resolution analysis, two-scale equation scaling function and wavelet function can be obtained:

$$\phi(t) = \sqrt{2} \sum_{k=-\infty}^{+\infty} h_k \phi(2t - k); \quad (4)$$

$$\psi(t) = \sqrt{2} \sum_{k=-\infty}^{+\infty} g_k \phi(2t - k). \quad (5)$$

In Eq. (4),  $h_k$  is two-scale expansion coefficients of scaling function, also known as the low-pass filter coefficients. In Eq. (5),  $g_k$  is two-scale equation wavelet expansion coefficients, also known as a high-pass filter coefficients. The relationship of  $g_k$  and  $h_k$  can be expressed as:

$$g_k = (-1)^{k-1} h_{1-k}, \quad k \in Z. \quad (6)$$

The method of wavelet packet decomposition divides the frequency band into a multilayer structure, and it allows the high frequency part to make further decomposition, which is not required in normal multi-resolution analysis or wavelet analysis. Moreover, according to the analysis of signal features, it can select adaptively the appropriate frequency band to match the signal spectrum, thus improving the resolution of time-frequency analysis.<sup>17</sup> Wavelet packet decomposition and reconstruction formulas are as follows:

$$c_{j-1,k} = \sum_{m=-\infty}^{+\infty} \hat{h}_{m-2k} c_{j,m}; \quad (7)$$

$$d_{j-1,k} = \sum_{m=-\infty}^{+\infty} \hat{g}_{m-2k} c_{j,m}. \quad (8)$$

In Eq. (7) and (8),  $\hat{h}_k$ ,  $\hat{g}_k$  respectively represent conjugate coefficients of low-pass and high-pass orthogonal filters, and  $c_{j,k}$ ,  $d_{j,k}$  are scale coefficients and wavelet coefficients.

The wavelet packet decomposition algorithm can be expressed as:

$$c_{j,k} = \sum_{m=-\infty}^{+\infty} h_{k-2m}c_{j-1,m} + \sum_{m=-\infty}^{+\infty} g_{k-2m}d_{j-1,m}; \quad (9)$$

where  $h_k, g_k$  are the coefficients of low-pass and high-pass orthogonal filters.

The wavelet packet reconstruction formula can be expressed as:

$$c_{j,k} = \sum_{m=-\infty}^{+\infty} h_{k-2m}c_{j-1,m} + \sum_{m=-\infty}^{+\infty} g_{k-2m}d_{j-1,m}. \quad (10)$$

For the discrete original signals, the wavelet packet decomposition and reconstruction algorithms are used to obtain wavelet signals of some certain frequency or several frequency bands. In this paper, according to the corresponding relation between rolling bearing fault states and feature vectors, original vibration signals are decomposed and reconstructed to obtain wavelet signals of eight frequency bands, then the energy feature vectors are extracted from different frequency band signals.

## 2.2. GMM-HMM

In GMM-HMM, HMM is the main model. The parameters of HMM are defined as follows:<sup>18</sup>

1.  $N$ , the number of states in the model. Set  $S = \{s_1, s_2, \dots, s_N\}$  consists of the individual states, if the state is  $q_t$  at  $t$  time,  $q_t \in \{s_1, s_2, \dots, s_N\}$ ;
2.  $M$ , the number of distinct observation symbols per state. The individual observation symbols are denoted as  $V = \{v_1, v_2, \dots, v_M\}$ . If the observation symbol is  $o_t$  at  $t$  time,  $o_t \in \{v_1, v_2, \dots, v_M\}$ ;
3.  $A$ , the state transition probability matrix.  $A = (a_{ij})_{N \times N}$  where  $a_{ij} = P[q_{t+1} = s_j | q_t = s_i], 1 \leq i, j \leq N$ ;
4.  $B$ , the observation probability matrix  $B = (b_{jk})_{N \times M}$  where  $b_{jk}$  indicates observation symbol probability distribution in state  $s_j$ ,  $b_{jk} = P(o_t = v_k | q_t = s_j), 1 \leq k \leq M$ ;
5.  $\pi$ , the initial state distribution.  $\pi = (\pi_1, \pi_2, \dots, \pi_N)$  where  $\pi_t = P(q_1 = s_t), 1 < i < N$ .

Typically, HMM assumes that the observation symbols in per state are discrete values and uses the method of discrete probability density to build a model for transition probabilities and observation probabilities. However, in the actual process of condition monitoring and fault diagnosis, the observation symbols of a characteristic signal are usually continuous variations, not a few definite values. Although a continuous signal can be processed with a variety of discrete ways, such as vector quantization, the process may cause great loss of effective information. Therefore, in the process of condition monitoring

and fault diagnosis, using HMMs with continuous observation probability distribution will have more advantages than to use HMMs with discrete observation symbols. The actual fault characteristic signals are often multidimensional feature vectors, and the multi-dimensional characteristic signal of continuous variation can be mixed with several probability density functions to characterize. Since the GMM can be infinitely close to arbitrary distribution, GMM is used to fit the probability density function of observation vector in each state, namely,

$$b_{jm}(o_t) = \sum_{m=1}^M w_{jm}G_{jm}(o_t) = \sum_{m=1}^M w_{jm}N(o_t, \mu_{jm}, c_{jm}), \quad 1 \leq m \leq M; \quad (11)$$

where  $M$  is the Gaussian component number of the state  $s_j$ , equivalent to the number of observation symbols corresponding to each state in the discrete HMM.  $w_{jm}$  is the weights of the  $m$ -th Gaussian distribution of the state  $s_j$ ,  $\mu_{jm}$  and  $c_{jm}$  are mean vector, and the covariance matrix of the  $m$ -th Gaussian distribution in the state  $s_j$ ,  $G_{jm}(o_t)$  is the  $m$ -th Gaussian distribution of state  $s_j$ . This distribution is a multi-dimensional normal random variable probability density function, and its dimension  $D$  is the dimension of feature vector. According to the analysis of multi-dimensional feature vector, multi-dimensional normal random variable probability density functions are:<sup>19</sup>

$$G_{jm}(o_t) = N(o_t, \mu_{jm}, C_{jm}) = \frac{1}{\sqrt{(2\pi)^D |C_{jm}|}} \exp\left(-\frac{1}{2}(o_t - \mu_{jm})C_{jm}^{-1}(o_t - \mu_{jm})\right). \quad (12)$$

From the above analysis, the observations matrix  $B$  of GMM-HMM is the use of multi-dimensional Gaussian density function to model, which uses mean vector  $\mu_{jm}$ , covariance matrix  $C_{jm}$ , and mixing coefficients (weights)  $w_{jm}$  to characterize  $b_j(o_t)$ .

## 2.3. GMM-HMM Parameter Reevaluation

In GMM-HMM, forward-backward algorithm, Viterbi algorithm, and Baum-Welch algorithm are basically the same as discrete HMM.<sup>20</sup> However, when calculating and reevaluating the probability distribution matrix  $B$ , the former is more complex than the latter. Suppose the model parameters  $\lambda$  and the sequence of observations  $o$  is given,  $\gamma_t(j, m)$  is defined as the joint probability of the  $m$ -th Gaussian distribution when the state is  $q_t$  at  $t$  time is  $s_j$ :

$$\gamma_t(j, m) = P(q_t = s_j, x_{j,t} = X_{j,m} | O, \lambda); \quad (13)$$

where  $x_{j,t}$  stands for the Gaussian distribution of  $s_j$  at time  $t$  and  $X_{j,m}$  stands for the  $m$ -th Gaussian distribution of  $s_j$ ,  $\gamma_t(j, m)$ , can be obtained by forward variable  $\alpha_t(i)$  and backward variable  $\beta_t(i)$ ,

$$\gamma_t(j, m) = \frac{\alpha_t(j)\beta_t(j)}{\sum_{i=1}^n \alpha_t(i)\beta_t(i)} \times \frac{w_{j,m}N(o_t, \mu_{j,m}, C_{j,m})}{\sum_{j,n} w_{j,n}N(o_t, \mu_{j,n}, C_{j,n})}. \quad (14)$$

Then, according to  $\gamma_t(j, m)$ , weights  $w_{j,m}$ , mean vector  $\mu_{j,m}$ , and covariance matrix  $C_{j,m}$  are revalued as follows:

$$\hat{w}_{j,m} = \frac{\sum_{t=1}^T \gamma_t(j, m)}{\sum_{t=1}^T \sum_{n=1}^M \gamma_t(j, n)}; \tag{15}$$

$$\hat{\mu}_{j,m} = \frac{\sum_{t=1}^T \gamma_t(j, n) o_t}{\sum_{t=1}^T \gamma_t(j, m)}; \tag{16}$$

$$C_{j,m} = \frac{\sum_{t=1}^T \gamma_t(j, m) (o_t - \mu_{j,m})(o_t - \mu_{j,m})^T}{\sum_{t=1}^T \gamma_t(j, m)}. \tag{17}$$

### 3. FAULT DIAGNOSIS BASED ON GMM-HMM

In the classic study case on HMM, these hermits used wet degrees of the seaweed (four observation symbols of soggy, damp, dryish, dry) to predict the states of weather (four hidden states of sunny, cloudy, rain, heavy rain). The HMM prediction model library established they are not the same in different seasons, and the prediction results are different. Fault diagnosis method based on GMM-HMM can use this classic case to illustrate. For example, the vibration signals of rolling bearings are classified into four modes (normal, inner ring fault, rolling body faults, and outer fault), which is similar to the seasons spring, summer, autumn and winter in the weather prediction case; according to feature vectors of four different failure modes, the HMM model library including four modes is established. In the forecast of weather, the Markov chain consists of four weather conditions (sunny, cloudy, rain, and heavy rain), which can transform into each other.

In the process of fault diagnosis and prognosis, Markov chains are generally set the left-right type, which includes four states ( $s_1, s_2, s_3, s_4$ ). The four hidden states do not seem to have clear meaning on the surface, but they often can be characterized by the damage degree of components in some failure mode. For example, the bearing crack failures are divided into four hidden states according to the length of the crack. Four hidden states can be considered as the internal mechanism of bearing which can generate different vibration signals, and these vibration signals are extracted to feature vendors to input HMM as an observation symbol. Usually, the fault diagnosis signals collected are not discrete values but continuous, so continuous HMM would have a better advantage.

Figure 1 is the flow chart of GMM-HMM fault diagnosis and prediction, the prognosis process includes two parts: offline training and online test. The work of offline training is to build the GMM-HMM model library by means of fault samples. The original data signals sampled by the sensor are pretreated, using Wavelet Packet Decomposition (WPD) to extract feature vendors, namely, the n-th layer node energies after WPD. Then node energies as observation symbols are used to

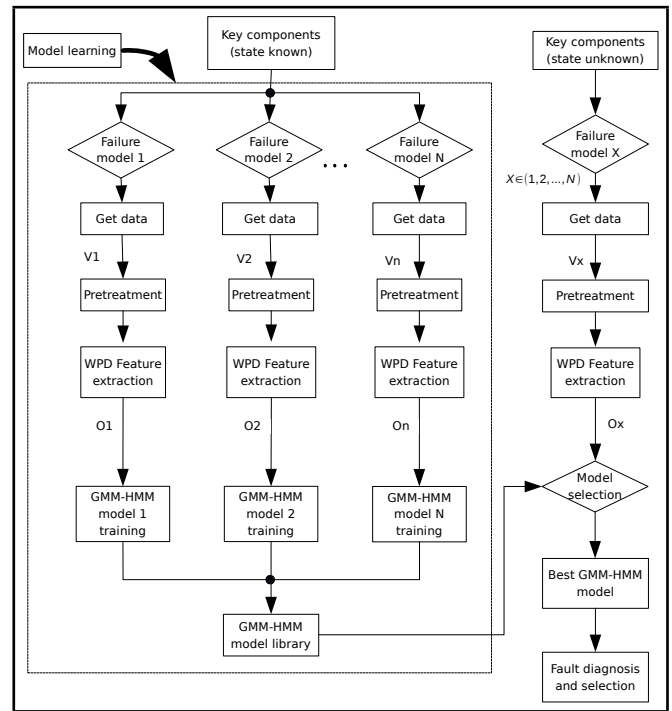


Figure 1. The flow chart of GMM-HMM fault diagnosis and prediction.

train and revalue the model parameters of GMM-HMM. During the course of establishing the model library, the parameters of each GMM-HMM are learned and trained by using the Baum-Welch algorithm.

The online test involves signal process, fault diagnosis and prediction aiming at online test data. Wavelet packet analysis is adopted to extract feature vectors of the current condition monitoring data, which is the observation vector of GMM-HMM. Forward algorithm is used to calculate probability  $P(O|\lambda)$  of the current observation vector in the GMM-HMM model library, and Viterbi algorithm is adopted to determine the observation sequence of optimal path and to calculate maximum probability; then, according to the maximum probability of failure, different failure mode and damage degree of the components can be recognized and evaluated.

## 4. APPLICATION AND RESULTS

### 4.1. Test Systems and Data Acquisition

In order to verify the effectiveness of this fault diagnosis method presented previously, rolling bearing is selected as the object to carry out test research. The raw data of test research come from the rolling bearing experimental platform provided by Case Western Reserve University.<sup>21,22</sup> As shown in Fig. 2 below, the test platform consists of a 2 hp motor, a torque transducer, a dynamometer, and control electronics. The test bearings support the motor shaft. Single point faults were introduced to the test bearings using electro-discharge machining with fault diameters of 0.007 inches, 0.014 inches, 0.021 inches, and 0.028 inches. Normal fault data of fan end are acquired at 12 kHz frequency; bearing failure of drive end is collected at 12 kHz frequency data, but also at 48 kHz frequency to acquire.

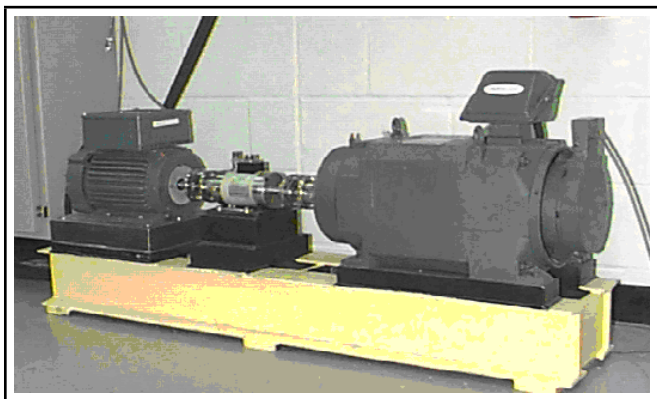


Figure 2. The test platform for rolling bearing faults.

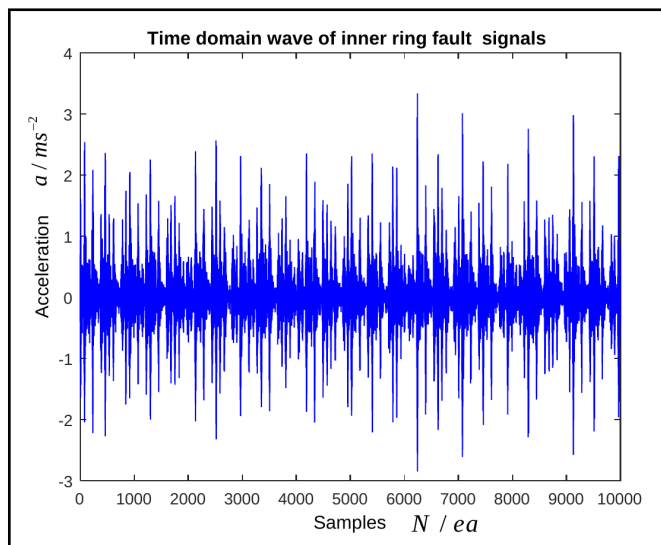


Figure 3. Raw signals for bearing inner ring fault.

In this paper, the experimental data selected is drive end data whose sampling frequency data is 12 kHz, fault diameter is 0.021 inches, motor speed is 1750 rpm, and motor load is 2 horsepower. The types of data are classified by normal, inner ring fault, balls fault, and outer ring fault. Among these data, 20 samples are used to train the GMM-HMM model library and 10 samples as model test and validation. Additionally, in order to verify the feasibility of this approach in evaluation of part degradation, we select four different grades data for the same type of fault bearings, which are inner ring fault bearings with fault diameters of 0.007 inches, 0.014 inches, 0.021 inches, and 0.028 inches.

### 4.2. Feature Extraction of Vibration Signal

The different fault vibration signals of rolling bearings are sampled by the sensor. The raw fault signals of bearing inner ring are shown in Fig. 3. Figure 4 is the spectrum of fault signals for bearing inner ring. Seen from Fig. 4, there are several peaks at some frequencies, but the fault features of inner ring are not very obvious, so that inner ring fault of rolling bearing corresponding to characteristic frequencies cannot be accurately identified.

According to the wavelet package algorithm previously, the

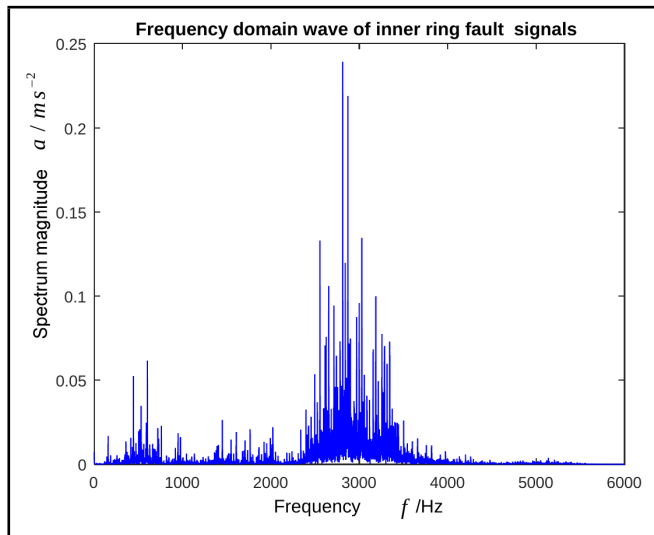


Figure 4. Spectrum of bearing inner ring fault signals.

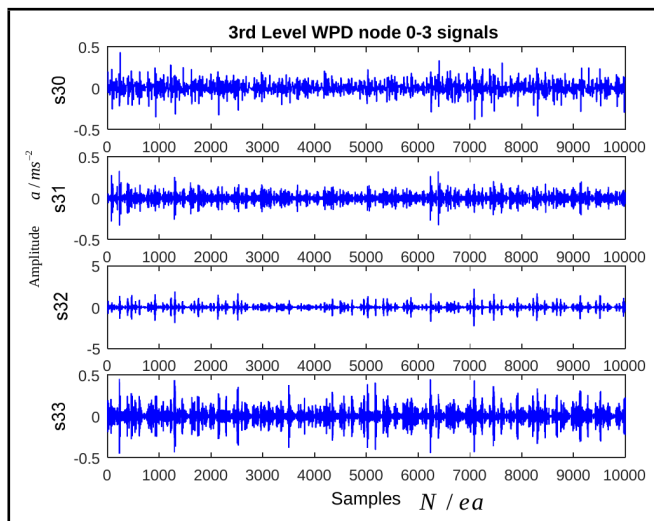


Figure 5. Wavelet reconstruction s30-s33 for inner ring fault.

fault signals are decomposed into a three-layer wavelet packet to effectively extract the fault feature.<sup>23,24</sup> After wavelet decomposition and reconstruction, 8 band signals  $s_{3j}$  are shown in Fig. 5 and Fig. 6.

As shown from Fig. 5 and Fig. 6, compared to the original signals, each node signal was more moderate and clear by means of three layer wavelet decomposition and reconstruction, which achieves the purpose of eliminating noise. Through analysis of different frequency band wavelet signals, we can obtain dynamic behavior characteristics of vibration signals in each frequency band. After reconstructed signal  $s_{3j}(t)$  in each band is further processed, node energy  $E_{3j}$  is given below:

$$E_{3j} = \int |s_{3j}(t)|^2 dt = \sum_{k=1}^n |x_{jk}|^2; \quad (18)$$

where  $x_{jk}(j = 0, 1, \dots, 7, k = 1, 2, \dots, n)$  denotes the magnitude of  $n$  discrete points in the reconstructed signal  $s_{3j}$ , and the third WPD node energy for inner ring fault is shown in Fig. 7.

In order to improve the clustering of feature vectors, fea-

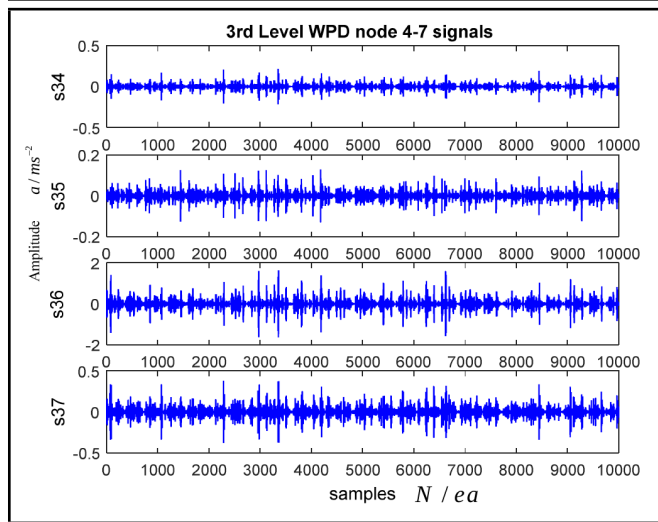


Figure 6. Wavelet reconstruction s34–s37 for inner ring fault.

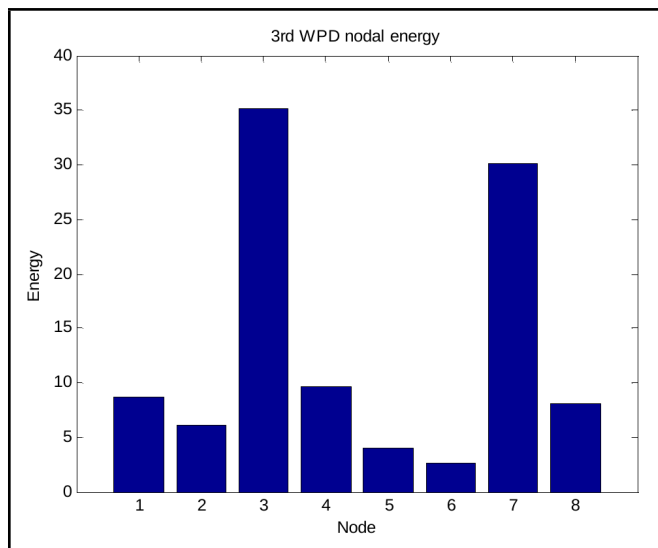


Figure 7. The histogram of node energy for bearing inner fault.

ture vectors need to be normalized to facilitate network input and recognition. Among the normalized feature vector  $S_{3j} = [s_{30}, s_{31}, \dots, s_{37}]$ , the range of each element  $s_{3j}$  is between 0 and 1. The normalization method is expressed as

$$s_{3j} = \frac{e_{3j} - \max(E_{3j})}{\max(E_{3j}) - \min(E_{3j})}. \tag{19}$$

### 4.3. Diagnosis of Different Fault Bearings

#### 4.3.1. GMM-HMM for Different Faults

Aiming at four different fault features of rolling bearings, including normal, inner ring, rolling body fault, and outer ring fault, there was a total of 80 groups of sample data. Twenty groups of each type of fault data were selected for model learning and parameter training of the GMM-HMM library. After sub-frame processing and wavelet packet feature extraction, the original signals are converted into feature vectors to input GMM-HMMs as observation symbols. The K-means clustering algorithm is used to estimate an initial set of model parameters, then Baum-Welch algorithm is adapted to train model

parameters and to obtain fault feature GMM-HMM model library.

The GMM-HMM train processes of different modes are shown in Fig. 8. As can be seen from Fig. 8, corresponding to rolling bearing four different sample data (normal, inner ring fault, rolling element fault, and outer ring fault), we established four GMM-HMMs. The number of training and the convergence degree of each model is not consistent. The training numbers of four GMM-HMMs are respectively 22, 32, 15, and 19 times. Through the above study and training, we can obtain four different types of rolling bearing GMM-HMM library for normal, inner ring fault, rolling body fault, and outer ring fault.

Due to space limitations, this article only gives model parameters of GMM-HMM for the inner fault. Initial probability  $\pi$  of GMM-HMM is given as  $\pi = (1, 0, 0, 0)'$ . After model training and parameters reevaluation, the state transition matrix  $A$  of GMM-HMM for the inner ring fault is shown as follows:

$$A = \begin{pmatrix} 0.9687 & 0.0313 & 0 & 0 \\ 0 & 0.7627 & 0.2373 & 0 \\ 0 & 0 & 0 & 0.9959 & 0.0041 \\ 0 & 0 & 0 & 1.0000 \end{pmatrix} \tag{20}$$

When we use GMM to approximate the probability density function  $b_{jm}(o_t)$  of observation vector in each implied state, the assumptions are as follows:

1. there are 4 hidden states  $s_j, j = 1, 2, 3, 4$  in each fault mode;
2. after these observation vectors are clustered to three datasets, which are respectively subject to three Gaussian probability distribution functions in each hidden state.

By means of parameter reevaluation, we obtained the parameters of Gaussian mixture distribution  $b_{jm}(o_t)$  for the inner ring fault of rolling bearings. Mean vectors  $\mu_{j,m}$  are listed in Table 1, covariance matrices  $C_{j,m}$  are listed in Table 2, and weights  $w_{j,m}$  are listed in Table 3.

#### 4.3.2. Model Verification and Fault Recognition

In order to verify the performance of GMM-HMM applying in fault diagnosis and failure pattern recognition, a total 40 groups of sample data (10 groups for each type, normal, inner ring fault, balls fault, and outer ring fault) are selected to make fault prognosis. Based on the GMM-HMM library, we can calculate the maximum probability of test data that appears in the HMM database by Vertibi algorithm, and we can deduce which GMM-HMM these test data belong to. Furthermore, online diagnosis and prognosis can be realized. The test results of different faults are shown in Fig. 9.

Figure 9 shows that the fault diagnosis method based on GMM-HMM is feasible, and all test results are accurate without mistaken judgment. Moreover, the diagnosis and recognition processes have a fast computing speed and high recognition rates. Meanwhile, this demonstrates that the fault diagnosis method using GMM-HMM has higher recognition accuracy

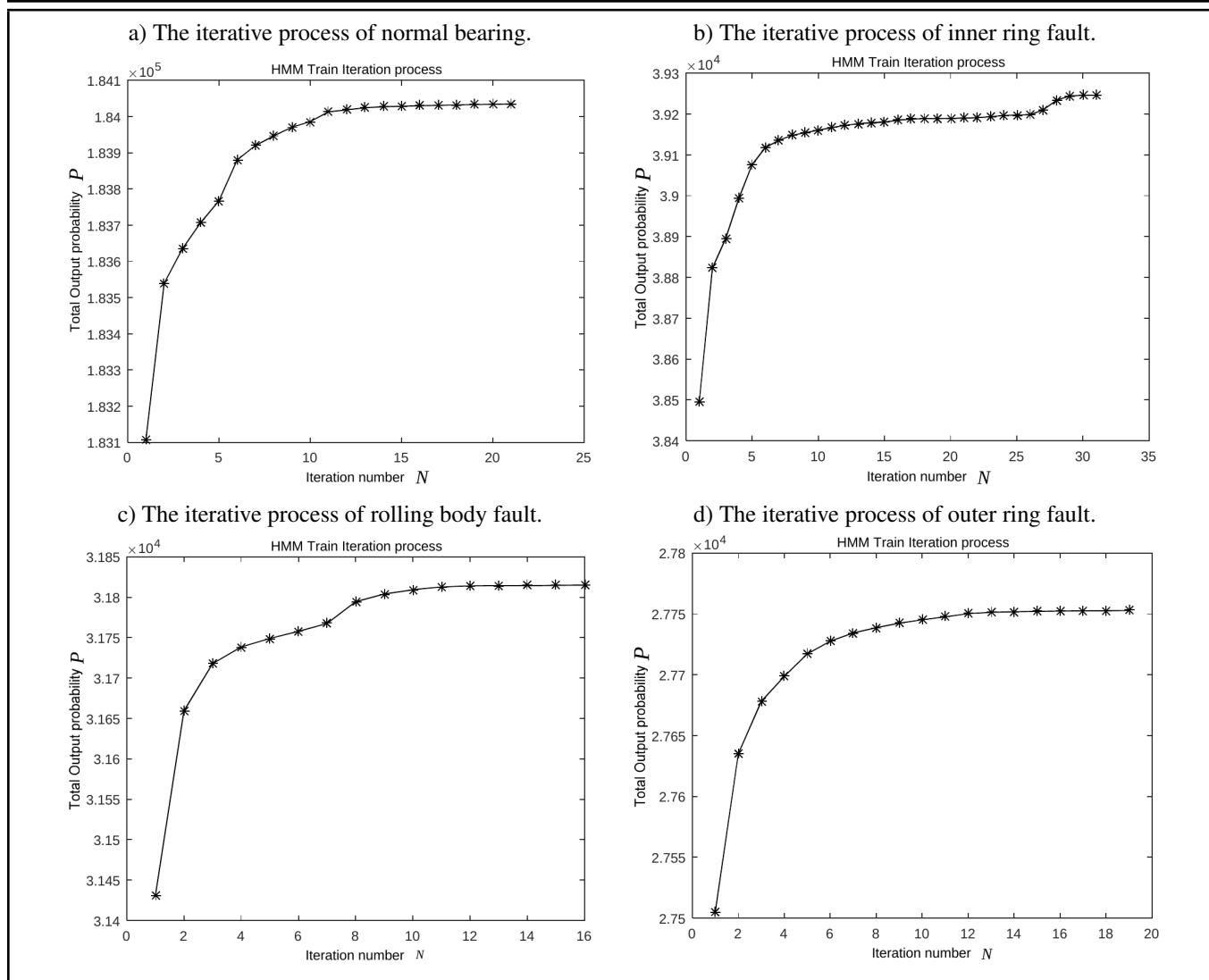


Figure 8. GMM-HMM train processes for different faults.

Table 1. GMM-HMM reevaluation parameter  $\mu_{j,m}$ .

$\mu_{1,m}$	0.1010	0.1112	0.7389	0.1498	0.0423	0.0445	0.6247	0.1117
	0.1442	0.1365	0.5267	0.1388	0.0411	0.0426	0.7997	0.1331
	0.1580	0.1476	0.6299	0.1512	0.0403	0.0508	0.7146	0.1257
$\mu_{2,m}$	0.1604	0.1591	0.5275	0.1488	0.0440	0.0485	0.7894	0.1394
	0.1569	0.1542	0.5943	0.1554	0.0426	0.0488	0.7413	0.1352
	0.1591	0.1621	0.6479	0.1622	0.0424	0.0452	0.6934	0.1243
$\mu_{3,m}$	0.1636	0.1676	0.6640	0.1659	0.0403	0.0484	0.6741	0.1210
	0.1342	0.1178	0.5448	0.1291	0.0418	0.0429	0.7938	0.1372
	0.1214	0.1179	0.6177	0.1384	0.0380	0.0429	0.7419	0.1201
$\mu_{4,m}$	0.1397	0.1348	0.5750	0.1426	0.0405	0.0422	0.7678	0.1297
	0.1705	0.1754	0.5677	0.1578	0.0449	0.0503	0.7522	0.1347
	0.1362	0.1346	0.6445	0.1504	0.0402	0.0425	0.7108	0.1216

compared to the method proposed by literature.<sup>5</sup> The output of the RBF network was 0.945, 0.983, 0.969, and 0.932.

#### 4.4. Performance Degradation State Evaluation for the Same Fault Bearings

##### 4.4.1. GMM-HMM for Different Grades

The failure of the mechanical part is the process of performance degradation, which can appear as different faults, or be-

have as different grades for the same fault. Here we choose the inner ring fault to characterize the deterioration degree of bearing performance, and the inner ring fault bearings are divided into different grades according to the size of the machining fault diameter (0.007 inches, 0.014 inches, 0.021 inches, and 0.028 inches). For the convenience of description, the data of inner ring fault are abbreviated as IR007, IR014, IR021, and IR028. A total 80 groups of test data for four different grades,

Table 2. GMM-HMM revaluation parameter  $C_{j,m}$ .

$C_{1,m}$	0.0002	0.0008	0.0007	0.0003	0.0000	0.00001	0.0005	0.00001
	0.0004	0.0005	0.0013	0.0001	0.0000	0.00001	0.0004	0.0003
	0.0005	0.0007	0.0006	0.0004	0.0000	0.0001	0.0006	0.0001
$C_{2,m}$	0.0003	0.0004	0.0005	0.0001	0.00001	0.00002	0.0002	0.00004
	0.0006	0.0005	0.0004	0.00004	0.00002	0.00003	0.00033	0.00004
	0.00031	0.00032	0.00051	0.00035	0.00002	0.00005	0.00005	0.00011
$C_{3,m}$	0.0005	0.0004	0.0013	0.0003	0.00001	0.00001	0.0011	0.0001
	0.0001	0.0002	0.0012	0.0001	0.0000	0.00001	0.0005	0.0001
	0.0003	0.0004	0.0009	0.0002	0.0000	0.00001	0.0006	0.0001
$C_{4,m}$	0.0002	0.0003	0.0005	0.0001	0.0000	0.00001	0.0003	0.0001
	0.0005	0.0006	0.0019	0.0002	0.0000	0.00001	0.0012	0.0001
	0.0003	0.0003	0.0005	0.0002	0.0000	0.00001	0.0004	0.0001

Table 3. GMM-HMM revaluation parameter weight  $w_{j,m}$ .

$w_{1,m}$	0.1080	$w_{2,m}$	0.2614	$w_{3,m}$	0.3553	$w_{4,m}$	0.3388
	0.3953		0.4773		0.2355		0.4072
	0.4967		0.2614		0.4091		0.2540

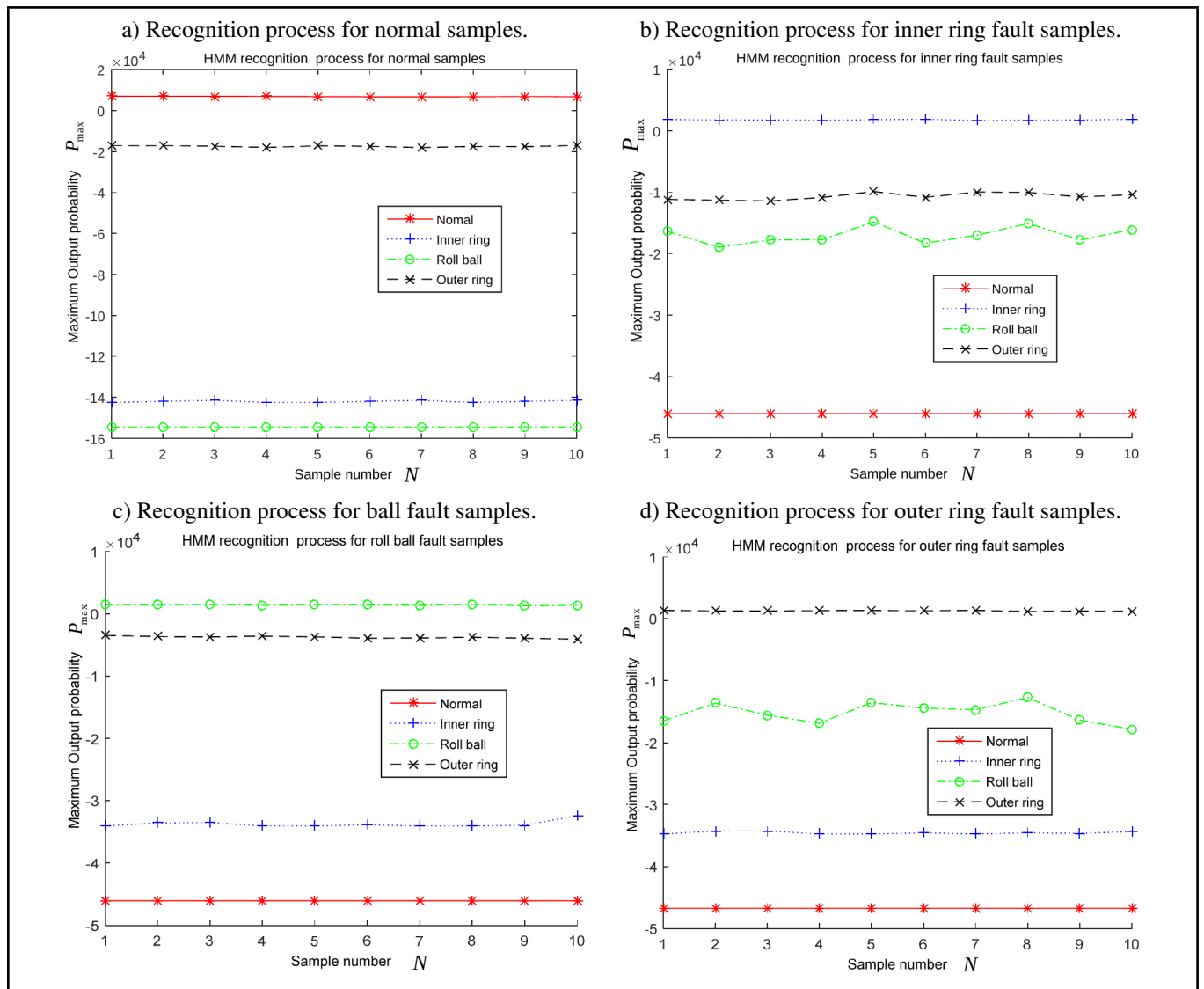


Figure 9. The test results of different faults.



20 groups of each type of fault data, were selected for model learning and parameter training of the GMM-HMM library.

Similarly, these raw signals are converted to feature vector to input GMM-HMM as observation symbols, the K-means clustering algorithm is used to estimate an initial set of model parameters, and Baum-Welch algorithm is adapted to train model parameters and to obtain the GMM-HMM library for different grade fault features. The GMM-HMM train processes are shown in Fig. 10. As can be seen from Fig. 10, corresponding to four different grade sample data of 0.007, 0.014, 0.021, and 0.028 for inner ring fault bearings, we established four GMM-HMMs. Obviously, the number of training and convergence degree of four models is different. The training numbers of four models are respectively 40, 38, 15, and 18 times. As space is limited, all model parameters of GMM-HMMs for different grades are not listed.

#### 4.4.2. Model verification and Degradation Evaluation

In the same way, a total 40 groups of sample data (10 groups for each grade, 0.007, 0.014, 0.021, and 0.028 ) were selected to verify the performance of GMM-HMM. Utilizing the above-trained four-model library, we can calculate the maximum probability of the test data that appears in the GMM-HMM database and determine which model these test data are subject to. Furthermore, the performance degradation state of rolling bearing can be judged and evaluated. The test results of the different grade fault samples are shown in Fig. 11.

As can be seen from Fig. 11, the evaluation method of the degradation state is feasible and effective for different grade fault rolling bearings based on GMM-HMM, which have the same fault pattern with inner ring. Moreover, all test results are accurate without mistaken judgment, and the evaluation and recognition processes have fast computing speed and high recognition rates.

## 5. CONCLUSIONS

In this paper, we put forward fault diagnosis and pattern recognition technology combining wavelet packet decomposition with GMM-HMM. Using WPD to extract energy feature vectors can eliminate noise effectively from raw signals. The method of using GMM to approximate observation probability density function of the feature vector can avoid losing effective information. The experimental results show that fault prognosis and state evaluation based on GMM-HMM have higher identification accuracy and fast speed for fault pattern recognition and degradation state judgment of rolling bearings.

From above studies, whether training sample data or test sample data are from the laboratory and sample data are concentrated and have less interference factors, the recognition rate of test samples are very high. In the follow-up study, according to the different damage degree of components, we will select as many as possible training and testing samples to train and verify GMM-HMM and consider the impact of interfering factors in the field environment. In addition, we will extend our

research based on GMM-HMM from rolling bearings to other mechanical parts and from fault diagnosis to pattern recognition, life prediction and other fields.

## ACKNOWLEDGEMENT

The project is jointly supported by National Natural Science Foundation of China and Civil Aviation Administration of China, as well as jointly supported by Hunan Provincial Natural Science Foundation of China and Xiangtan Municipal Science and Technology Bureau.

## REFERENCES

- Jiang, T. and Yuan, S. Rolling bearing faults diagnosis based on the improved wavelet neural network, *Journal of Huazhong Agricultural University*, **33** (1), 131–136, (2014). <http://10.13300/j.cnki.hnlkxb.2014.01.005>
- Xia, R. *Fault diagnosis for rolling bearing based on time series analysis*, Xinjiang University, Urumqi, 2006.
- Huang, Z. and Xie, Y. Fault diagnosis of roller bearing with inner and external fault based on Hilbert transformation, *Journal of Central South University (Science and Technology)*, **42** (7), 1992–1996, (2011).
- Zhou, Z., Zhu, Y. S., Zhang, Y. Y., Zhu, C. F. and Wang, P. Adaptive fault diagnosis of rolling bearings based on EEMD and demodulated resonance, *Journal of Vibration and Shock*, **32** (2), 76–80 (2013). <http://10.13465/j.cnki.jvs.2013.02.021>
- Wang, G. F. and Wand Z. L. Fault diagnosis of the rolling bearing using Wavelet Packet and RBF neural network, *The Journal of Beijing University of Science and Technology*, **26** (2), 184–187, (2004).
- Wang, G. D. and Zhang, J. Y. Fault pattern recognition of bearing rolling using Wavelet Packet Neural Network, *Measurement and Instrument*, **21** (2), 42–49, (2008). <http://10.19533/j.issn1000-3762.2007.01.011>
- Feng, C. J., Shao, Q. and Kang, J. Faults Diagnosis for Bearing Based on HMMs with Gaussian Mixtures, *Turbine Technology*, **53** (3), 205–208, (2011).
- Xiao, W. B., Chen, J., Zhou, Y., Wang, Z. Y. and Zhao, F. G. Wavelet packet transform and hidden Markov model based bearing performance degradation assessment, *Journal of Vibration and Shock*, **30** (8), 32–35, (2011). <http://10.13465/j.cnki.jvs.2011.08.011>
- Zaidi, S. S. H., Aviyente, S., Salman, M., Shin, K. K. and Strangas, E. G. Prognosis of gear failures in DC starter motors using Hidden Markov Models, *IEEE Transactions on Industrial Electronics*, **58** (5), 1695–1706, (2011). <https://dx.doi.org/10.1109/tie.2010.2052540>

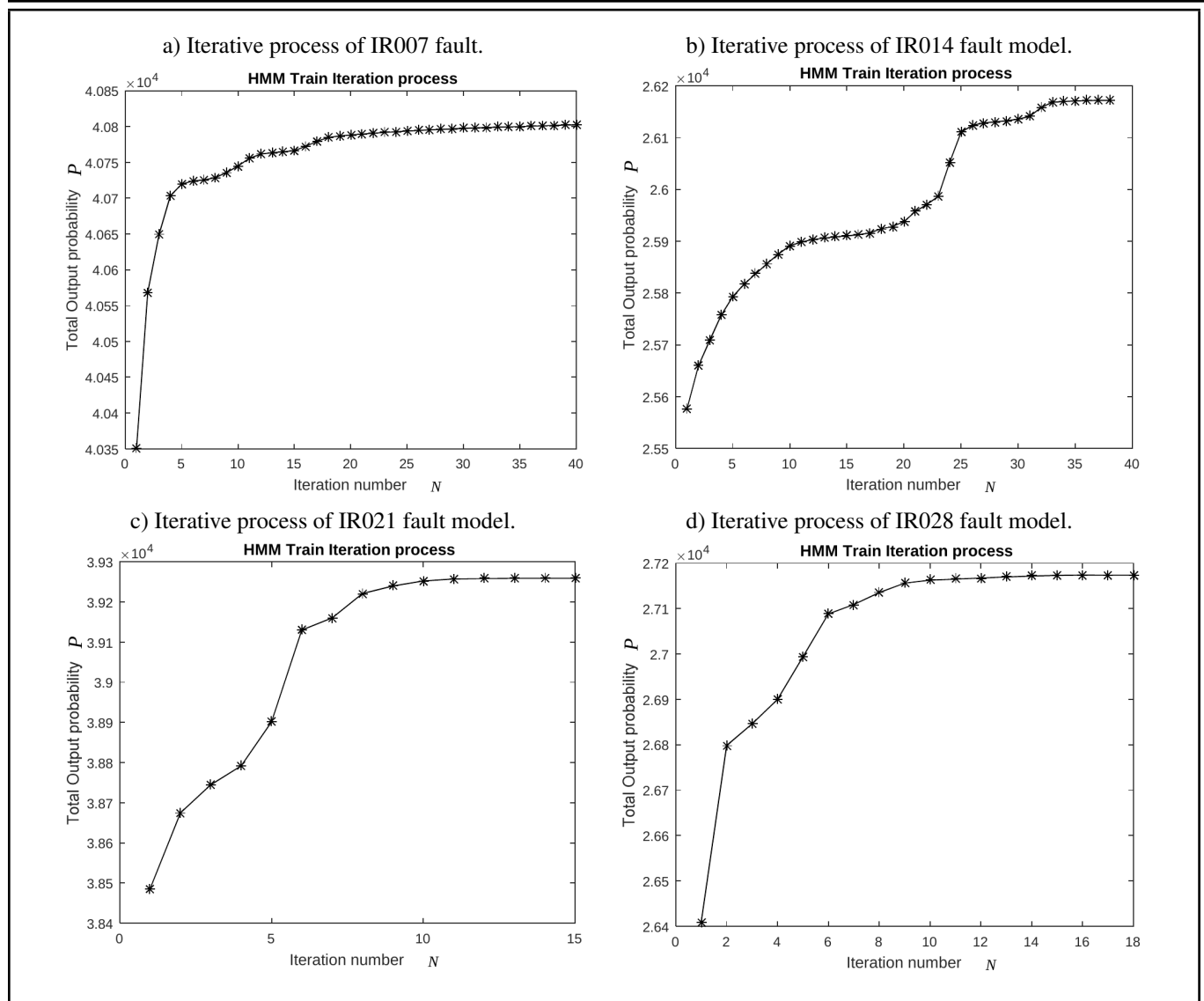


Figure 10. GMM-HMM train processes for different grades.

- <sup>10</sup> Yu, J. Adaptive hidden Markov model-based online learning framework for bearing faulty detection and performance degradation monitoring, *Mechanical Systems and Signal Processing*, **83** (1), 149–162, (2017). <https://dx.doi.org/10.1016/j.ymsp.2016.06.004>
- <sup>11</sup> Yuwono M., Qin Y., Zhou J. Guo, Y., Celler, B. G. and Su, S. W. Automatic bearing fault diagnosis using particle swarm clustering and Hidden Markov Model, *Engineering Applications of Artificial Intelligence*, **47** (1), 88–100, (2016). <https://dx.doi.org/10.1016/j.engappai.2015.03.007>
- <sup>12</sup> Zhou, H., Chen, J., Dong, G. and Wang, R. Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model, *Mechanical Systems and Signal Processing*, **72** (5), 65–79, (2016). <https://dx.doi.org/10.1016/j.ymsp.2015.11.022>
- <sup>13</sup> Li, Z., Fang, H. and Huang, M. Diversified learning for continuous hidden Markov models with application to fault diagnosis, *Expert Systems with Applications*, **42** (23), 9165–9173, (2015). <https://dx.doi.org/10.1016/j.eswa.2015.08.027>
- <sup>14</sup> Geramifard, O., Xu, J.X. and Panda, S.K. Fault detection and diagnosis in synchronous motors using hidden Markov model-based semi-nonparametric approach, *Engineering Applications of Artificial Intelligence*, **26** (8), 1919–1929, (2013). <https://dx.doi.org/10.1016/j.engappai.2013.06.005>
- <sup>15</sup> Purushotham, V., Narayanan, S. and Prasad, S.A. Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, *NDT&E International*, **38** (8), 654–664, (2005). <https://dx.doi.org/10.1016/j.ndteint.2005.04.003>
- <sup>16</sup> Sheng, Y. *Wavelet Transform*, CRC Press LLC, Boca Raton, Florida, United States, (2000).
- <sup>17</sup> Zeng, X.W., Zhao, W.M. and Sheng, J.Q. Corresponding relationships between nodes of decomposition tree of wavelet packet and frequency bands of signal sub-

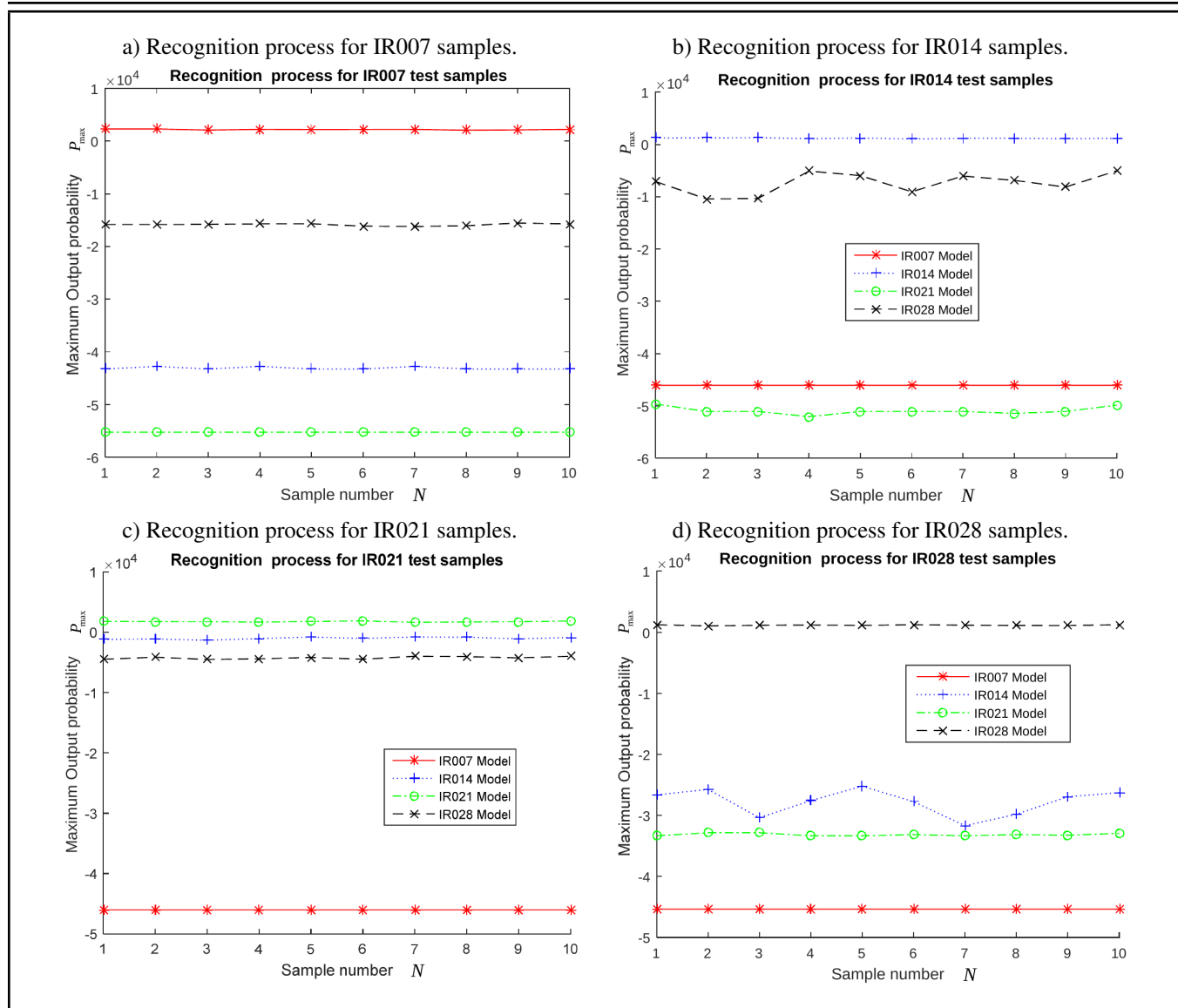


Figure 11. The test results of different grades.

space, *Acta Seismologica Sinica*, **21** (1), 91–97, (2008).  
<https://dx.doi.org/10.1007/s11589-008-0091-x>

- <sup>18</sup> Yu, J. *Early Fault Detection for Gear Shaft and Planetary Gear Based on Wavelet and Hidden Markov Model*, University of Toronto, Toronto, (2011).  
<https://tspace.library.utoronto.ca/handle/1807/31991>
- <sup>19</sup> Boutros, T. and Liang, M. Detection and diagnosis of bearing and cutting tool faults using Hidden Markov models, *Mechanical Systems and Signal Processing*, **25** (6), 2102–2124, (2011).  
<https://dx.doi.org/10.1016/j.ymssp.2011.01.013>
- <sup>20</sup> Miao, Q. and Makis, V. Condition Monitoring and Classification of Rotating Machinery using Wavelets and Hidden Markov Models, *Mechanical System and Signal Process*, **21** (2), 840–855, (2007).  
<https://dx.doi.org/10.1016/j.ymssp.2006.01.009>
- <sup>21</sup> Case Western Reserve University Bearing Data Center, <https://csegroups.case.edu/bearingdatacenter/pages/download-data-file>
- <sup>22</sup> Qiang, L., Taiyong, W., Yonggang, L., Wei, W. and Guofeng, W. Engineering signal processing based on adaptive step-changed stochastic resonance, *Mechanical Systems and Signal Processing*, **21** (5), 2267–2279, (2007).  
<https://dx.doi.org/10.1016/j.ymssp.2006.10.003>
- <sup>23</sup> Zhou, H. Q. H. Z. Z. and Yaguo, Z. Y. L. Intelligent diagnosis for incipient fault based on lifting wavelet package transform and support vector machines ensemble, *Chinese Journal of Mechanical Engineering*, **42** (8), 17–22, (2016).  
<https://dx.doi.org/10.3901/jme.2006.08.016>
- <sup>24</sup> Li, B., Zhang, P., Liu, D., Mi, S. and Ren, G. Feature extraction of roller bearing fault using morphological lifting wavelets, *Journal of Vibration, Measurement & Diagnosis*, **31** (1), 36–40, (2011).  
<https://10.16450/j.cnki.issn.1004-6801.2011.01.022>