A Robust Performance Degradation Modeling Approach Based on Student’s t-HMM and Nuisance Attribute Projection

HUIMING JIANG1, JING YUAN1, QIAN ZHAO1, HAN YAN2, SEN WANG3, AND YUNFEI SHAO1

1School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
2School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
3Shanghai Aerospace Control Technology Institute, Shanghai 201109, China

Corresponding author: Jing Yuan (yuanjing_802@163.com)

This work was supported in part by the Shanghai Sailing Program under Grant 18YF1417800, in part by the National Natural Science Foundation of China under Grant 51975377, in part by the Shanghai Special Funds for Industrial Transformation, Upgrading and Development, under Grant GYQJ-2019-1-03, in part by the Shanghai Aerospace Science and Technology Innovation Fund Project under Grant SAST2019-100, and in part by the Key Laboratory of Vibration and Control of Aero-Propulsion System Ministry of Education, Northeastern University, under Grant VCAME201907.

ABSTRACT
Performance degradation assessment (PDA) is of great significance to ensure safety and availability of mechanical equipment. As an important issue of PDA, the robustness of the trained model directly affects the assessment efficiency and restricts its application in practice. This paper proposes a robust modeling approach based on Student’s t-hidden Markov model (Student’s t-HMM) and nuisance attribute projection (NAP). NAP can remove nuisance attributes caused by individual differences from the feature space. Student’s t-HMM utilizes the finite Student’s t-mixture models (SMMs) to describe the observation emission densities associated with each hidden state, which can be more tolerant towards outliers than conventional HMMs. Based on these two techniques, the proposed method is supposed to be more robust and can assess the performance degradation process of new objects based on data of tested objects. The superiority of the proposed approach is evaluated using the vibration data from the accelerated life tests of bearings and the public XJTU-SY Bearing Datasets. The results demonstrate the robustness and effectiveness of the proposed approach for PDA modeling of rolling element bearings.

INDEX TERMS
Robustness, performance degradation assessment, student’s t-HMM, nuisance attribute projection, bearings.

I. INTRODUCTION
Prognostics and health management (PHM) covers a series of research topics trying to prevent failure of an engineered system in advance by the state assessment, fault diagnosis, remaining life prediction and so on [1]. In fact, equipment fault is not only a state, but also a developing process from normal stage to fault initiation, then to fault deterioration and then to final failure. Performance degradation assessment (PDA) pays much attention to analyzing the current state and the trend of the whole life process [2], [3]. While realizing fault diagnosis, it is more conducive to the later life prediction. Therefore, PDA is of great significance for ensuring the reliability and safety of operation and reducing maintenance cost of equipment [4].

Many research works have been carried out on PDA of equipment covering bearings, gears, pumps and mechanical systems. Some works research the degradation process utilizing mathematical or physical models derived from the first principles and the knowledge of failure mechanisms, which are the so-called model-based approaches [5]. Liao et al. utilized the Pairs-Erdogan model to describe the fault growth of bearings [6]. Li et al. used the exponential model and particle filtering (PF) to model and update parameters of the degradation process of bearings [7]. Advantages of the model-based approach are obvious when the degradation process can be described accurately [8]. However, the failure mechanisms are complicated and not straightforward in most
cases. Accurate degradation expressions are unavailable. As a result, applications of the model based methods in industry are limited. Furthermore, large amounts of monitoring signal containing abundant degradation information can be available, which gives a great development space to the data-driven approach.

The data-driven approach trains an assessment model using monitoring data with machine learning (ML) methods, and then utilizes the trained model to assess the degradation trend with current data. Cheng et al. proposed a health degradation monitoring method based on self-organizing mapping (GSOM) and clustered support vector machine (CSVM) for rolling element bearings [9]. Features extracted from the online monitoring data were compared with the BMU in the trained GSOM to calculate the minimum quantization error (MQE), which was supposed to characterize the degradation state. The proposed method showed better performance compared with several other methods. Rai et al. proposed a PDA approach based on an amalgamation of empirical mode decomposition (EMD) and k-medoids clustering [10]. The cluster center of training data from normal state was obtained based on the k-medoids algorithm. Then the confidence value (CV) based on dissimilarity of the test data object to the normal state was calculated and treated as the degradation indicator for assessing the health of bearings. Ma et al. proposed a bearing PDA approach based on subspace-based minimum volume ellipsoid (SMVE), which was applied for the life-time test of aero-engine bearings [11]. Many other ML techniques have been introduced into the PDA of equipment, e.g., artificial neural networks (ANNs) [12], relevance vector machine (RVM) [13], recurrent neural network (RNN) [14] and hidden Markov model (HMM). Compared with other ML algorithms, HMM has an outstanding advantage. HMM provides a convenient way of modeling observations appearing in a sequential manner, and infers invisible state changes by observations, which is consistent with the idea of PDA [15], [16]. Recently, many extensions of the HMM based methods have been applied to machinery systems. Ocak et al. pioneered the application of HMM into the condition monitoring and fault diagnosis of bearings [17]. Li et al. studied the method of modeling and analysis for multi-state systems based on HMM [18]. Fort et al. proposed an extension of HMM to redefine the system availability performance [19]. Vrignat et al. verified the superiority of HMM-based degradation process modeling in fault prediction by comparative analysis [20]. Alaswad et al. pointed out that hidden semi-Markov model (HSSM) could effectively model and evaluate multi-state degenerate systems [21]. Chen et al. used non-homogeneous HSMM (NHSSM) based on hyper parametric optimization to model the non-linear multi-state degradation process [22]. The validity of this method in performance degradation assessment and remaining useful life (RUL) prediction has been verified by experimental data. Li et al. updated the state transition probability based on the dimensionless performance index degradation function and realized the PDA of wind turbine bearings under small samples based on the degradation-HMM [23]. The above researches have promoted the applications of HMMs in the field of PDA.

As an important subtopic of PDA, the robustness of the trained model has a direct impact on the final evaluation results. In most cases, data of testing objects encountered in real cases and data of objects used in the PDA model training stage have quite different distributions even under the same fault state, which can be caused by differences of individuals, or sampling channels or environmental issues. In fact, even in the laboratory, data from different individuals of the same class objects under the same operation conditions can have different distributions, too. So it often fails to obtain satisfying results when evaluating other same objects based on models trained by several testing objects. As a result, the application of PDA encounters obstacles in practice, which means the PDA model has low robustness. In recent years, transfer learning has emerged as a new learning framework, which tries to learn the target attributes of data from different domain [24], [25]. Lei et al. proposed a feature-based transfer neural network (FTNN) to realize the identification of the health states of bearings used in real-case machines (BRMs) based on the diagnosis knowledge from bearings used in laboratory machines (BLMs) [26]. Transferable features extracted by the FTNN model could reduce the discrepancy between datasets from BLMs and from BRMs efficiently, which proposed a very meaningful idea for the robustness improvement of PDA. In this paper, from a different view, a robust performance degradation modeling approach based on Student’s t-HMM and nuisance attribute projection (NAP) is proposed. This method is developed based on the superiority of HMM on modeling time series data and the advantage of NAP on removing the nuisance attributes from the feature space [27], which is quite helpful for PDA. On one hand, Student’s t-HMM shows more tolerance to outliers than conventional HMMs [16]. On the other hand, NAP is utilized to remove the nuisance attributes caused by individual differences through projection. These two techniques improve the robustness of the PDA modeling approach effectively which can be verified through practical applications.

The rest of this paper is organized as follows. Section 2 briefly introduces the theoretical background of Student’s t-HMM and NAP. Section 3 details the proposed PDA modeling approach. In Section 4, the vibration data from the accelerated life tests of bearings and the public XJTU-SY Bearing Datasets are used to verify the effectiveness and robustness of the proposed method respectively. Conclusions are carried out in Section 5 finally.

II. THEORETICAL BACKGROUND
A. STUDENT’S t-HMM
As one of the most successful applications of dynamic Bayesian networks (DBNs), HMM utilizes two sequences to describe probability events [28], [29]. The observation
emission densities associated with each hidden state are usually described by finite Gaussian mixture models (GMMs), yielding the so-called continuous HMMs (CHMMs). GMMs are successful to approximate unknown random distributions, which promotes the vast popularity of CHMMs. Nevertheless, GMMs are adversely affected by outliers of data in the model fitting stage, which retains the effectiveness and robustness of the trained HMM finally. Unfortunately, outliers often occur in real-world applications. To solve this drawback, the finite Student’s t-mixture models (SMMs) have been proposed as an alternative to GMMs, which are highly tolerant to outliers [30]. SMMs provide a much more robust approach to untypical training data and the significant tolerance has been experimental depicted in many articles.

A Student’s t-distribution has three parameters, which are mean vector \( \mu \), covariance parameter \( \Sigma \) and degrees of freedom \( \nu (\nu > 0) \). Its probability density function (pdf) can be described as follows

\[
t(\mathbf{o}_t; \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)|\Sigma|^{-1/2}(\pi)^{-p/2}}{\Gamma(\nu/2)[1 + d(\mathbf{o}_t, \mu; \Sigma)/\nu]^{(\nu+p)/2}}
\]

where \( p \) is the dimensionality of the observations \( o_t \), \( d(\mathbf{o}_t, \mu; \Sigma) \) is the squared Mahalanobis distance between \( o_t \) and \( \mu \) with covariance matrix \( \Sigma \), and \( \Gamma(s) \) is the Gamma function.

In essence, the Student’s t-distribution corresponds to a scaled model of Gaussian [31]. And the precision scalar is closely related to the degrees of freedom of the Student’s t-density, which follows a Gamma-distributed latent variable. For a Student’s t-distribution, it is described as

\[
o_t \sim t(\mu, \Sigma, \nu)
\]

which is equivalent to

\[
o_t \mid \mu_t \sim N(\mu, \Sigma/\nu_t)
\]

\( \nu_t \) is the precision scalar which follows

\[
u_t \sim \zeta(\nu, \nu/2)
\]

\( N(\mu, \Sigma) \) is a normal Gaussian distribution with mean \( \mu \) and covariance matrix \( \Sigma \), and \( \zeta(\alpha, \beta) \) is the Gamma distribution. Then the density of SMMs can be given as

\[
p(o_t; \Theta) = \sum_{j=1}^{J} w_j f(o_t; \mu_j, \Sigma_j, \nu_j)
\]

where \( w_1, \ldots, w_J \) are mixing weights of components and \( \Theta = \{w_j, \mu_j, \Sigma_j, \nu_j\}_{j=1}^{J} \). \( p(\cdot) \) is defined as a generic notation for a probability function.

The Student’s t-HMM is defined as a finite state-space hidden Markov model whose observation emission distributions of each hidden state are modeled by SMMs. Suppose that one Student’s t-HMM has \( I \) hidden states, and the number of the components of SMMs is \( J \). Then one Student’s t-HMM can be expressed by the following parameters:

1. \( \pi \): the initial state probability distribution, \( \pi = \{\pi_i\} \), and \( \pi_i = P(q_1 = S_i) \) for \( 1 \leq i \leq I \).
2. \( A \): the state transition probability matrix, \( A = \{a_{ij}\} \), and \( a_{ij} = P(q_{t+1} = j | q_t = i) \) for \( 1 \leq i, j \leq I \).
3. \( \Theta \): the observation probability distribution parameter sets based on the Student’s t-HMM, \( \Theta = \{\Theta_j\} \), and \( \Theta_j = \{w_{ij}, \mu_{ij}, \Sigma_{ij}, \nu_{ij}\}_{j=1}^{J} \) for \( 1 \leq i \leq I \). Then the probability density of the observation \( o_t \) emitted from the \( i \)-th hidden state can be calculated as

\[
p(o_t; \Theta_j) = \sum_{j=1}^{J} w_{ij} f(o_t; \mu_{ij}, \Sigma_{ij}, \nu_{ij})
\]

Then one Student’s t-HMM can be described by \( \pi, A \) and \( \Theta \). For convenience, the notation \( \lambda = (\pi, A, \Theta) \) is used to indicate the complete parameter set of one Student’s t-HMM.

A graphical illustration of Student’s t-HMM can be displayed in Fig. 1.

![FIGURE 1. A graphical illustration of Student’s t-HMM.](image-url)

In order to further explain the reason why the proposed method adopts Student’s t-HMM, the advantages of Student’s t-HMM over the conventional HMMs are further summarized and explained. The difference between these two models is that the probability distribution function of observations in the hidden state is different. The conventional HMMs adopt GMMs. GMMs are well known to be highly intolerant to the presence of untypical data within the training data sets used for their estimation [16]. Outliers are quite common in applications. To retain their modeling effectiveness, bigger model size and more training data are needed for the conventional HMMs. SMMs have recently emerged as a heavier-tailed, robust alternative to GMMs, overcoming these hurdles. Student’s t-HMM adopts SMMs to exploit these merits for time series modeling. The advantages of Student’s HMM are supposed to improve the robustness and effectiveness of the proposed PDA modeling approach.

B. NUISANCE ATTRIBUTE PROJECTION

Nuisance attribute projection was initially proposed for removing the effect of the channel variability, which has got successful applications in speaker identification, face or image recognition and so on [32]. Gopinath et al. firstly introduced NAP into the engineering area. NAP was utilized to identify and remove the system dependent features to model a system-independent feature space, which improved the robustness of features across the two different capacity
synchronous generators [33]–[36]. Jiang et al. studied the NAP based feature transformation technique and combined NAP with HMM to get better diagnosis and performance degradation assessment results for bearings [3], [27]. Therefore, this paper takes full advantages of NAP to improve the robustness of the modeling approach based on the Student’s t-HMM.

The working of the NAP algorithm is pictorially explained in Fig. 2. \( x \) represents the feature vectors with the system information containing the target attribute which is the degradation state of bearings in this paper and other nuisance attributes. These nuisance attributes can be caused by sample channels, or individual differences of objects and so on, which have nothing to do with the target attribute. \( x' \) is the system independent feature vector containing the target attribute. The Sub-Space representing nuisance information contains little information about the target attribute and we try to build a projection that zeros that out.

![Sub-Space representing nuisance information](image)

**FIGURE 2.** Nuisance attribute projection.

Consider a \( N \) – dim feature space, feature samples can be described as a matrix \( X = [x_1, x_2, \ldots, x_n] \), where \( x_i \in \mathbb{R}^{N \times 1} \) and \( n \) is the sample number. Through minimizing the figure of merit (FOM) which can be expressed as

\[
\delta = \sum_{ij} w_{ij} \| P x_i - P x_j \|^2
\]

a projection matrix \( P \) can be obtained as following:

\[
P = I - \sum_{i=1}^{d} v_i v_i^T
\]

where \( \{ v_i \} \) are eigenvectors obtained from the following eigenvalue problem which is transferred from the objective function in (7). The eigenvalue problem equation can be described as following:

\[
XZ(W)X^T v_i = \lambda_i v_i \quad (i = 1, \ldots, d)
\]

where the matrix \( Z(W) = diag(W \cdot I) - W \). And \( W = \{ w_{ij} \} (1 \leq i, j \leq n) \) is the weight matrix and directly reflects nuisance attributes. The differences caused by individuals, sampling channels, or environmental issues have nothing to do with the fault stage or degradation state. However, these differences unavoidably disturb the final diagnosis and assessment results. Therefore, issues like these are called nuisance attributes. NAP works on removing these attributes from the feature space. So the weight matrix in our application is

\[
w_{ij} = \begin{cases} 1 & \text{Case 1;} \\ 0 & \text{Otherwise.} \end{cases}
\]

where Case 1 means that if the differences between \( x_i \) and \( x_j \) are mainly caused by nuisance attributes. The nuisance attributes can be removed from the feature space with the following equation

\[
x' = P x
\]

### III. METHODOLOGY

Poor robustness is a common problem for the current PDA methods. In most cases, data of testing objects encountered in real cases and data of objects used in the PDA model training stage have different distributions even under the same fault state, which can be caused by differences of individuals, sampling channels or environmental issues. In fact, even in the laboratory, data from different individuals of the same class objects under the same operation conditions can have different distributions, too. So it often fails to obtain satisfying results when evaluating other same objects based on models trained by several testing objects. To deal with this problem, a robust performance degradation modeling approach based on Student’s t-HMM and NAP is proposed in this paper. On one hand, NAP is applied to remove the nuisance attributes from the feature space to weaken the interference caused by individual differences to obtain a more robust feature space. On the other hand, Student’s t-HMM which is more tolerant to outliers is utilized to model time series data to improve the robustness of the modeling approach. These two techniques, Student’s t-HMM and NAP, are combined to improve the robustness of the PDA modeling approach. The whole frame of PDA based on the proposed method is shown in the Fig. 3. The details are described as follows:

**FIGURE 3.** The whole frame of the proposed method.

### A. STEP 1: SIGNAL PREPROCESSING

First of all, common feature extraction methods are utilized and vibration data are processed to obtain multidimensional feature vectors. In the training procedure, training features from different sample channels or different experiment objects are put into NAP processing to get a projection matrix \( P \). Through NAP, interference caused by nuisance...
attributes is projected out from the feature space. A more robust feature space is obtained. In the testing procedure, testing features are directly transferred with a projection of \( P \). Through STEP 1, more robust feature vectors are prepared.

### B. STEP 2: STUDENT’S T-HMM MODELING

Training feature vectors under normal operation condition obtained after STEP 1 are used as inputs of Student’s t-HMM to train a normal PDA model. Training of Student’s t-HMM can be easily conducted through EM algorithm which is the same with ordinary HMMs. The properties of Student’s t-distribution as described in Equation (3-4) are utilized. Then all the algorithms of ordinary HMMs can be applied. Detailed algorithms have been provided in reference [16]. Then a normal Student’s t-HMM can be obtained.

### C. STEP 3: PERFORMANCE DEGRADATION ASSESSMENT

Feature vectors of testing data which are obtained through feature extraction and transformation with a projection of \( P \) are put into the normal Student’s t-HMM obtained in STEP 2. Similar to ordinary HMMs, Forward-backward algorithm is utilized to calculate the likelihood probability. NAP has projected nuisance attributes out from the feature space, and Student’s t-HMM shows more tolerance to outliers. The likelihood probability is much more sensitive to changes of the degradation process. If the testing data is sampled from objects operating under normal condition, it matches the probability distribution of the trained normal Student’s t-HMM well. As a result, the calculated likelihood probability tends to be around the usual values of trained data. However, if the state of the testing object is away from the normal condition, the likelihood probability tends to be smaller than usual values. And it decreases by degrees with the degradation process. Therefore, the likelihood probability is commonly chosen as the performance index (PI).

Practical issues and details of utilization of the proposed method will be discussed during the following applications.

### IV. RESULTS AND DISCUSSIONS

In practical applications, although all objects belong to a particular type, the PDA model trained by experimental data commonly cannot obtain satisfying results when applied for other objects. The proposed method finds a solution through feature transformation and the modified modeling method. In order to verify the effectiveness of the proposed method, data from two different experiments are utilized. Analysis, results and discussions are provided in the following contexts.

#### A. CASE STUDY 1: ACCELERATED LIFE TESTS OF ROLLING ELEMENT BEARINGS

1) EXPERIMENT INTRODUCTION

In the experiment, accelerated life tests of rolling element bearings are carried out with the support of Hangzhou Bearing Test & Research Center (HBRC). The testbed is shown in Fig.4. The platform consists of an AC motor, a transmission system, a lubrication system, a loading system and a data acquisition system. The appended load \( P \) is 12.744kN to accelerate the degradation of bearing, and three acceleration sensors, type PCB348A, are placed on the housing of the testing bearings, as shown in Fig.5. Four bearings, type 6307, are mounted on the testbed at the same time. Their corresponding parameters and operation conditions are displayed in TABLE 1. Five characteristic frequencies are displayed in TABLE 2. The sampling frequency is 25.6 kHz, and 20480 samples (i.e., 0.8s) are recorded per minute. During the whole test, Bearing B1 at the location Bearing 1 and bearing B2 at the location Bearing 3 run to failure in turn. So data sampled from B1 and B2 are utilized in this section.

![FIGURE 4. Accelerated bearing life tester (ABLT-1A).](image)

![FIGURE 5. Load and locations of testing bearings: (a) Sketch of load and bearing arrangement; (b) Location of sensors.](image)

| TABLE 1. Rolling element bearing corresponding parameters and operation conditions. |
|--------------------------------|---|---|---|---|---|
| Type | Ball number | Ball diameter (mm) | Pitch diameter (mm) | Contact angle (°) | Motor speed (rpm) | Load (kN) |
| 6307 | 8 | 13.494 | 58.5 | 0 | 3000 | 12.744 |

| TABLE 2. Five characteristic frequencies of rolling element bearing 6307(Hz). |
|----------------|---|---|---|---|
| \( f_r \) | \( f_c \) | \( f_b \) | \( f_i \) | \( f_o \) |
| 50 | 19 | 102 | 246 | 153 |
2) DATA ANALYSIS

The life time of B1 and B2 are 2469 minutes and 4431 minutes respectively. When B1 ran to failure at 2469min, a new bearing was reinstalled at the location Bearing 1. Then B2 ran to failure 1962 minutes later. B2 was found to be in normal condition when B1 was replaced. So data of these 1962 minutes after the replacement of B1 is analyzed in the following. Root Mean Square (RMS) and Peak-to-Peak value (PP) of the whole life are displayed in Fig. 6 and Fig. 7. The robustness of the modeling method is the key issue discussed in this section. So channel 3 which is more sensitive to degradation process is selected for both bearings. From the whole life time plot, it seems that bearings were under normal condition for most time of the period. During around the beginning 2000 minutes for B1 and 1800 minutes for B2, RMSs and PPs fluctuate within a certain range without any obvious stage change. Then for the last 300 to 200 minutes, RMSs and PPs both went through a process of mutations until the downtime. So there are basically two stages during the whole process of these two bearings from RMS and PP plots. However, the detailed degradation processes are not very clear.

3) FEATURE EXTRACTION AND TRANSFORMATION BASED ON NAP

Degradation information is not only contained at every time point, but also among time series. 20,480 samples are separated into 10 segments. One feature vector is extracted for each segment, then a 10 points time series is obtained for each minute. Many works have been published on feature extraction methods for fault diagnosis and performance degradation assessment [37]–[43]. In the practical engineering, time-domain statistical features are commonly used indexes for the condition monitoring of rolling element bearings. The commonly used 11 time-domain features contain peak value, peak-to-peak value, average amplitude, root amplitude, root mean square, shape factor, impulse factor, crest factor, clearance factor, skewness and kurtosis. The components which often fail in a rolling element bearing are the outer race, the inner race, the rolling body and the cage. Such failures generate a series of impact vibrations in short time intervals, which occur at bearing characteristic frequencies (BCF). Envelope analysis is always used with FFT to identify faults occurring at the BCF. Amplitude spectrum entropy and envelope spectrum entropy are two commonly used frequency-domain features. As one time–frequency analysis method, wavelets have been established as the most widespread tool in many areas of signal processing, due to their flexibility and their efficient computational implementation. The wavelet packet transform (WPT) is a generalization of the wavelet transform and has been used in several applications of signal processing. The distribution of the wavelet packet decomposition node energies relates to the state of the bearing. The wavelet packet decomposition node energies and energy entropies are usually extracted as features. Finally, 29 commonly used features are extracted, which consist of eleven-time domain features, eight wavelet packet decomposition node energies and eight energy entropies whose level is 3, amplitude spectrum entropy and envelope spectrum entropy. Then the 29-dimension feature vector series are obtained.

To get better robustness, feature vector series of two testing bearings during the beginning 20 minutes, which are supposed to be sampled from bearings under normal stage, but from two different objects, are put into the NAP processor. These features have the same target attribute, that they are from normal stage. But obvious differences exist between these features, which are caused by vibration random, individual differences of objects and so on. To reduce the influence of individual differences on features, the weight matrix in NAP is defined as follows:

\[
w_{ij} = \begin{cases} 
1 & \text{if } x_i \text{ and } x_j \text{ are from different bearings;} \\
0 & \text{otherwise.}
\end{cases}
\]
Then the projection matrix $P$ can be obtained through NAP calculation, which will be used in the feature transformation procedure for both training and testing data.

4) MODELING AND ASSESSMENT
In this experiment, there are only two bearings running to failure. So normal data of B1 is used to train normal Student’s t-HMM. Then the normal model is utilized to assess the whole life of these two bearings.

Although the accurate degradation process cannot be obtained from the RMS and PP plots in Fig. 6-7, the beginning 100 minutes are supposed to be under the normal stage for B1 based on our experience. So data from the beginning 100 minutes of B1 are processed through feature extraction and transformation by the projection matrix $P$ calculated in the Part 3 of this section. These feature vector series are treated as the training data for PDA modeling. The initialization of student’s t-HMM parameters is processed as follows. The number of hidden states is set to 4. The number of the components of SMMs is set to 2. The convergence threshold is set to $10^{-4}$, and the max times of EM iterations is set to 10. Based on these training data, a normal stage Student’s t-HMM can be obtained. Then whole life feature vector series from the B1 and B2 are put into the trained model. The results are shown in Fig.8.

FIGURE 8. PDA based on NAP and Student’s t-HMM.

5) RESULTS AND DISCUSSIONS
From results shown in Fig.8, the whole life of B1 can be separated into three obvious stages basically, which are Stage 1, Stage 2 and Stage3. During Stage 1, PIs keep at a high and stable level with small fluctuations for quite a long time. Then during the period I as marked in the figure, PI curve experiences a step-down and keep at this lower level with little bigger fluctuations, which marks that the degradation process enters Stage 2. During Stage 2, PI curve experiences a slightly higher level in the period I as marked in the figure. Then PIs keep stable at this level for a long time until the changes happen in the period III. After the abnormal and short time increase, decreases like a waterfall start abruptly and last for a period, which marks the coming of Stage 3. During Stage 3, although the fluctuation ranges of PIs are very big, the overall trend is sharply downward.

Based on these characteristics of B1’s PI curve, it can be inferred that the testing bearing B1 was under normal condition in Stage1, and ran to early degradation state in Stage 2. And obvious fault occurred and developed to failure in Stage 3. Beyond those, there are some changes of PI plot that need to be explained. At the beginning of Stage 2, the occurrence of early degradation lead to the decreases of PIs. After a short time of running, the condition of bearing tended to be stable. So PIs increase slightly and keep on the level for a long time. At the end of Stage 2, changes happened in the vibration feature distributions. PI curve seems to increase a little. But the subsequent sharp decline is more convincing, which marks the bearing occurred fault and ran to failure. To verify these inferences, the vibration signal and the envelope spectra of five representative moments as marked in the figure are plotted in Fig.9-10.

FIGURE 9. The vibration plots of Time 1-5.

Time 1, Time 2 and Time 3 seem to be similar from the time plots in Fig.9. Time 1 is at the end of Stage 1 and Time 2 is at the beginning of Stage 2. Compared with Time 1, the amplitude at the rotating frequency is much bigger and the ball pass inner race frequency (BPIF) component can be observed at Time 2. These characteristics indicate the early stage of fault. With the development of degradation process, the system gradually self-balanced. The amplitude at the rotating frequency decreases, but the amplitude at BPIF is basically the same at Time 3. These indicate that the early fault happened and kept stable in Stage 2. Time 4 is at the end of Stage 2. Time 5 is at the beginning of Stage 3. From the
time plots in Fig.9, the fault characteristic is quite obvious at Time 5. From the spectra in Fig.10, the amplitude at the rotating frequency and the BPIF component accompanied by the rotating frequency are quite obvious and much bigger than those at Time 4. These characteristics indicate that B1 covered severe fault in Stage 3 and the fault developed quickly. B1 ran to failure at the end of Stage 3. Fig. 11 gives the picture of the failure bearing after experiment. B1 ran to inner race fault, which verifies the assessment results provided above. B1 ran under normal state in Stage 1. Early degradation happened in the period I. B1 ran with early degradation in Stage 2. Fault deterioration happened in the period II. B1 ran to failure quickly in Stage 3. So the PI plot obtained based on the proposed method describes the whole life of B1 faithfully, and it is quite effective for the performance degradation assessment of B1.

The proposed method is supposed to be more robust for different testing bearings. From the PDA results of B2 in Fig.8, the whole life of B2 can be roughly divided into two obvious stages. In Stage 1, PIs keep at a high level with small fluctuations for quite a long time. Specially, the PI curve decreases slightly to a little lower level during the period I. Then it goes back to the normal level. Sharp drops happen in the PI curve in the period II. Then the decline continues and large fluctuations happen during Stage 2. From these characteristics of the PI curve, it can be inferred that the testing bearing B2 was under normal condition at the beginning of Stage 1, and ran to early degradation in the period I. And during Stage 1, no obvious fault occurred. Severe fault happened in the period I, and the fault developed quickly and B2 ran to failure at the end of Stage 2. To verify these inferences, the vibration signal and the envelope spectra of five representative moments as marked in the Fig.8 are plotted in Fig. 12-13.

From the time plots in Fig.12 and the envelop spectra in Fig.13, Time 1, Time 2, Time 3 and Time 4 are at in Stage 1. The time plots and spectra of these four time points keep similar. The differences among these time points cannot show the fault occurrences or developments. So the degradation state keeps unchanged in the period I. Time 5 is at the end of Stage 1. Time 5 is at the beginning of Stage 2. Comparing these two time points, obvious fault characteristics can be observed in the vibration signal and the spectrum of Time 5. The amplitude at the rotating frequency and the BPIF component accompanied by the rotating frequency are quite obvious and much bigger than those at Time 4. These characteristics show obvious fault happened at Time 5. B2 run to fault stage and the fault developed quickly in Stage 2. B2 covered serious fault in inner race, and ran to failure finally at the end of Stage 2. The picture of the failure bearing after experiment is given in Fig.14 verifying these assessment results. So PDA model trained based on normal data of B1 also get excellent results when it is applied to B2.

---

**FIGURE 10.** The envelop spectra of Time 1-5.

**FIGURE 11.** The picture of the failure bearing B1: the fault part is marked in red circle.

**FIGURE 12.** The vibration plots of Time 1-5.
FIGURE 13. The envelop spectra of Time 1-5.

FIGURE 14. The picture of the failure bearing B2: the fault part is marked in red circle.

6) COMPARISONS

To show the advantages of the proposed method, results based on three other methods are also provided. Method 1 is based on HMM. Method 2 is based on Student’s t-HMM. Method 3 is based on NAP and HMM. Except different choices of methods, all the feature extraction methods and detailed parameter choices all remain the same. The results are shown in Fig. 15-17. From these results, it can be seen that all these three methods can give assessment results for both B1 and B2 based on PDA model trained by normal data of B1. But some differences can be found. The details are described as follows:

1) Compared PAD results based on HMM in Fig.15 and those based on Student’s t-HMM in Fig.16, the method based on Student’s t-HMM shows obvious better performance. For B1, the method based HMM is less sensitive to the early degradation happened in the period I and the changes in period II. For B2, the method based on Student’s t-HMM performs better than that based on HMM as well. The method based on Student’s t-HMM is more sensitive to the degradation occurrences in period I.

2) Compared the results in Fig.16 with those in Fig.8, PIs obtained by the method based on Student’s t-HMM have bigger fluctuations. NAP can effectively remove interference information caused by random issues and individual differences from the feature space. So the method based on NAP and Student’s t-HMM is more sensitive to the degradation process instead of the nuisance attributes. Furthermore, by comparing PAD results based on HMM in Fig.14 and those based on NAP and HMM in Fig.17, the advantages of NAP are more obvious. With the help of NAP, PI curve shows more sensitive to the degradation process, and the fluctuations of PIs are much smaller. So NAP is actually helpful to remove the nuisance attributes from the feature space, and can improve the robustness of the PDA modeling.
3) Compare all results in Fig.15-17, the method based on Student’s t-HMM stands out. The model based on this method shows more sensitive to the degradation process for both B1 and B2. However, compared with the results in Fig.8 based on the proposed method, the fluctuations of PI values during each stage are relatively bigger, which makes it easy to mask important changes in the degradation process.

Furthermore, there are two obvious outliers in the PI plot of B2 in Fig.17. These two outliers happened at 867min and 1355min. In the whole life time PP plot of B2, outliers at 867min and 1335min were observed in Fig.7, too. The time-plots of 867min and 1335min are displayed in Fig.18. So it can be found that the outliers in PIs of Fig.17 are caused by the abnormal value of the original data, independent of the degradation state. The PI plot of B2 obtained by the proposed method in Fig.8 also contains two outliers at 867min and 1335min. However, the deviation of these two outliers from the surrounding points are far less than those in Fig.17. This results also verify that SMMs have significant tolerance and provide a much more robust approach to untypical training data. From the above analysis and comparisons, the PDA modeling based on NAP and Student’s t-HMM is verified to be more effective. And the advantages of NAP and the superiority of Student’s t-HMM on PAD modeling are obvious. However, only two testing bearings are provided in this experiment, it’s not enough to verify the robustness of the proposed method. So the proposed method is applied in the public XJTU-SY Bearing Datasets in the following part.

B. CASE STUDY 2: THE PUBLIC XJTU-SY BEARING DATASETS

1) DATE DESCRIPTION
XJTU-SY bearing datasets [5] are provided by the Institute of Design Science and Basic Component at Xi’an Jiaotong University (XJTU), Shaanxi, P.R. China [44] and the Changxing Sumyoung Technology Co., Ltd. (SY), Zhejiang, P.R. China [45]. Details of the testbed are shown in Fig. 19. The type of tested bearings is LDK UER204. Three different operating conditions and five bearings under each operating condition were carried out during the experiment as tabulated in TABLE 3. Two PCB 352C33 accelerometers positioned at 90° to each other were placed on the housing of the testing bearings, sampling vibration signal of the vertical and the horizontal direction respectively. The sampling frequency is 25.6kHz, and 32768 samples (i.e., 1.28s) are recorded per minute. The complete run-to-failure datasets of fifteen rolling element bearings are provided. These datasets are utilized to verify the robustness of the proposed method.

2) DATA ANALYSIS
As the load direction is horizontal, the horizontal vibration signal is selected to assess the performance degradation. Fig.20 shows the horizontal vibration signals of five bearings under Condition 1, for example. All these fives testing bearings belong to the same type, and they all operated under the same conditions. However, the whole life time vibration signals are obviously different from each other. And the life times and degradation processes are quite different. This situation is very common in practical applications. And this
is the reason that the PDA model trained based on data of one bearing usually could not assess other bearings exactly.

3) RESULTS AND DISCUSSIONS
There are five whole life datasets for each operation conditions. Take datasets from Condition 1 as an example. When applying the proposed method, the same feature extraction method as utilized in Case 1 is chosen. Although the detailed degradation processes of these five bearings are still unclear, one thing is certain that bearings were running under normal condition during the beginning 20 minutes. The beginning 20 minutes’ feature sets after transformation from any of these four bearings are utilized to train a normal Student’s t-HMM as the PDA model. Then feature sets of all these five bearings are put into the trained model to assess the whole life degradation processes. The results are displayed in Fig. 21. The proposed method has strong robustness. It can effectively assess all these five bearings based on the trained PDA model. The whole life of bearings can be separated into two obvious stages. In Stage 1, no obvious changes can be observed. In Stage 2, PIs decrease along with degradation process. With proper abnormal point detection methods, the PI plot in Stage 2 is quite helpful for the further prediction such as RUL. The PI curves obtained based on the proposed method have the similar trends and stage changes with the whole life vibration plots in Fig.20. But the PI curves have better tendency and consistence for performance degradation processes.

3) RESULTS AND DISCUSSIONS
There are five whole life datasets for each operation conditions. Take datasets from Condition 1 as an example. When applying the proposed method, the same feature extraction method as utilized in Case 1 is chosen. Although the detailed degradation processes of these five bearings are still unclear, one thing is certain that bearings were running under normal condition during the beginning 20 minutes. The beginning 20 minutes’ feature sets after transformation from any of these four bearings are utilized to train a normal Student’s t-HMM as the PDA model. Then feature sets of all these five bearings are put into the trained model to assess the whole life degradation processes. The results are displayed in Fig. 21. The proposed method has strong robustness. It can effectively assess all these five bearings based on the trained PDA model. The whole life of bearings can be separated into two obvious stages. In Stage 1, no obvious changes can be observed. In Stage 2, PIs decrease along with degradation process. With proper abnormal point detection methods, the PI plot in Stage 2 is quite helpful for the further prediction such as RUL. The PI curves obtained based on the proposed method have the similar trends and stage changes with the whole life vibration plots in Fig.20. But the PI curves have better tendency and consistence for performance degradation processes.
The proposed method also successfully calculates PIs of all these bearings.

4) COMPARISON WITH NAP AND HMM

In the proposed method, student’s t-HMM which is more tolerant to outliers is utilized to model time series data to improve the robustness of the modeling approach. To verify the robustness of the proposed method, the method based on NAP and HMM is carried out for PDA of all these bearings. All parameters and training sets keep same for these two methods. The results based on NAP and HMM is given in fig. 26-28. Through comparing PDA results of bearings under condition 1 which are displayed in Fig. 21 and Fig. 26, it is obvious that the PI curves based on the proposed method have much better tendency and consistence for performance degradation processes. Furthermore, the proposed method keeps good performance for all these five bearings, especially for bearing 1_4. Comparing PIS in Fig. 22-23 and Fig. 27-28, the proposed method still has its own advantages in tendency and robustness. So it is obvious that the PDA modeling approach based on NAP and student’s t-HMM is effective and has strong robustness. Student’S t-HMM can improve the robustness of the modeling approach effectively.
5) COMPARISON WITH SOM

Self-organizing mapping (SOM) is one popular method utilized for PDA of bearings. Readers are advised to go through [9], [10] for details on the SOM algorithms and the construction method of health indices (HIs). The same feature extraction method and training data set are selected as utilized for Student’s t-HMMs and HMMs above. In the initial stage of SOM, the neighborhood radius $\sigma_0$ is set to 0.5, and the learning rate $\alpha_0$ is set to 0.1, which are same with Ref. [9]. HIs of bearings under three different cases are calculated and displayed in Fig. 29-31. Comparing PIs obtained through the proposed method in Fig. 21-22,24 and HIs obtained by SOM in Fig. 29-31, it can be found that both methods have good consistency and tendency to the performance degradation of bearings. However, for bearings under Condition 1, the proposed method is more sensitive to the degradation stage especially for Bearing 1_3. For bearings under Condition 3, the proposed method shows better robustness especially for bearing 3_5. In conclusion, the proposed method shows better robustness and more sensitiveness to the degradation stage of bearings, which means a lot for further accurate state identification and RUL (Remaining useful life, RUL) prediction.
Through this procedure, feature vector series are less sensitive to individual differences and more sensitive to degradation process. Second, an improved HMM, Student’s t-HMM is utilized for normal condition modeling. Student’s t-HMM exploits the merit of Student’s t-mixture models to be more tolerant towards outliers than conventional HMMs. So the PDA modelled based on Student’s t-HMM has more stable performance in applications. The method combined by these two techniques is able to be more effective and has better robustness. The results in Case study 1 show the effectiveness and the sensitivity to degradation processes of the proposed method. The results in Case study 2 demonstrate the robustness of the proposed method when applied to different objects. Furthermore, the performance of the proposed is very stable. In conclusion, the proposed method is verified to be robust and effective for performance degradation assessment of rolling element bearings. The results are supposed to be helpful for further prediction such as RUL. Consequently, more comprehensive equipment health management system will be developed in the future research.

V. CONCLUSION

In this paper, a robust performance degradation modeling approach is proposed and applied to rolling element bearings. The proposed method covers two important techniques which contribute a lot for the effectiveness. First, NAP is utilized to remove interferences caused by individual differences of testing objects. Feature sets from different testing objects under normal operation stages are utilized to calculate the projection matrix. Not only features used for PDA modeling, but also features put into PDA model to assess degradation all need to be transferred with this projection matrix. Through this procedure, feature vector series are less sensitive to individual differences and more sensitive to degradation process. Second, an improved HMM, Student’s t-HMM is utilized for normal condition modeling. Student’s t-HMM exploits the merit of Student’s t-mixture models to be more tolerant towards outliers than conventional HMMs. So the PDA modelled based on Student’s t-HMM has more stable performance in applications. The method combined by these two techniques is able to be more effective and has better robustness. The results in Case study 1 show the effectiveness and the sensitivity to degradation processes of the proposed method. The results in Case study 2 demonstrate the robustness of the proposed method when applied to different objects. Furthermore, the performance of the proposed is very stable. In conclusion, the proposed method is verified to be robust and effective for performance degradation assessment of rolling element bearings. The results are supposed to be helpful for further prediction such as RUL. Consequently, more comprehensive equipment health management system will be developed in the future research.

REFERENCES

[1] D. Wang, K.-L. Tsui, and Q. Miao, “Prognostics and health management: A review of vibration based bearing and gear health indicators,” *IEEE Access*, vol. 6, pp. 665–676, 2018.
[2] A. Rai and S. H. Upadhyay, “Intelligent bearing performance degradation assessment and remaining useful life prediction based on self-organising map and support vector regression,” *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, vol. 232, no. 6, pp. 1118–1132, Mar. 2018.
[3] H. Jiang, J. Chen, G. Dong, and R. Wang, “An intelligent performance degradation assessment method for bearings,” *J. Vib. Control*, vol. 23, no. 18, pp. 3023–3040, Oct. 2017.
[4] G. Ni, J. Chen, and H. Wang, “Degradation assessment of rolling bearing towards safety based on random matrix single ring machine learning,” *Saf. Sci.*, vol. 118, pp. 403–408, Oct. 2019.
[5] B. Wang, Y. Lei, N. Li, and N. Li, “A hybrid prognostics approach for estimating remaining useful life of rolling element bearings,” *IEEE Trans. Rel.*, vol. 69, no. 1, pp. 401–412, Mar. 2020.
[6] L. Liao, “Discovering prognostic features using genetic programming in remaining useful life prediction,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2464–2472, May 2014.
[7] N. Li, Y. Lei, J. Lin, and S. X. Ding, “An improved exponential model for predicting remaining useful life of rolling element bearings,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 12, pp. 7762–7773, Dec. 2015.
[8] J. Son, Y. Zhang, C. Sankavaram, and S. Zhou, “RUL prediction for individual units based on condition monitoring signals with a change point,” *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 182–196, Mar. 2015.
[9] Y. Cheng, H. Zhu, K. Hu, J. Wu, X. Shao, and Y. Wang, “Health degradation monitoring of rolling element bearing by growing self-organizing mapping and clustered support vector machine,” *IEEE Access*, vol. 7, pp. 135322–135331, 2019.
[10] A. Rai and S. H. Upadhyay, “Bearing performance degradation assessment based on a combination of empirical mode decomposition and K-medoids clustering,” *Mech. Syst. Signal Process.*, vol. 93, pp. 16–29, Sep. 2017.
[11] M. Ma, C. Sun, C. Zhang, and X. Chen, “Subspace-based MVE for performance degradation assessment of aero-engine bearings with multimodal features,” *Mech. Syst. Signal Process.*, vol. 124, pp. 298–312, Jun. 2019.
[12] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, “A recurrent neural network based health indicator for remaining useful life prediction of bearings,” *Neurocomputing*, vol. 240, pp. 98–109, May 2017.
[13] Y. Zhang and B. Guo, “Online capacity estimation of lithium-ion batteries based on novel feature extraction and adaptive multi-kernel relevance vector machine,” *Energies*, vol. 8, no. 11, pp. 12439–12457, 2015.
[14] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data,” *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
[15] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*. New York, NY, USA: Springer, 2005.
Y. Zhang, S. Tong, F. Cong, and J. Xu, “Research of feature extraction schemes for rolling element bearings,” J. Vib. Acoust., vol. 127, no. 4, pp. 299–306, Aug. 2005.

Y. Li, L. Cui, and C. Lin, “Modeling and analysis for multi-state systems with discrete-time Markov regime-switching,” Rel. Eng. Syst. Saf., vol. 166, pp. 41–49, Oct. 2017.

A. Fort, M. Mugnami, and V. Vignoli, “Hidden Markov models approach for understanding data parameters estimations,” Rel. Eng. Syst. Saf., vol. 136, pp. 85–91, Apr. 2015.

P. Vignat, M. Avila, F. Duculty, and F. Kratz, “Failure event prediction using hidden Markov model approaches,” IEEE Trans. Rel., vol. 64, no. 3, pp. 1038–1048, Sep. 2015.

S. Alasawad and Y. Xiang, “A review on condition-based maintenance optimization models for stochastically deteriorating system,” Rel. Eng. Syst. Saf., vol. 157, pp. 54–63, Jan. 2017.

G. Chen, J. Chen, Y. Zi, and H. Miao, “Hyper-parameter optimization based nonlinear multistate deterioration modeling for deterioration level assessment and remaining useful life prognostics,” Rel. Eng. Syst. Saf., vol. 167, pp. 517–526, Nov. 2017.

J. Li, X. Zhang, X. Zhou, and L. Lu, “Reliability assessment of wind turbine bearing based on the degradation-Hidden-Markov model,” Renew. Energy, vol. 132, pp. 1076–1087, Mar. 2019.

S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE Trans. Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, “Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabelled data,” IEEE Trans. Ind. Electron., vol. 66, no. 9, pp. 7316–7325, Sep. 2019.

B. Yang, Y. Lei, F. Jia, and S. Xing, “An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings,” Mech. Syst. Signal Process., vol. 122, pp. 692–706, May 2019.

H. Jiang, J. Chen, and G. Dong, “Hidden Markov model and nuisance attribute projection based bearing performance degradation assessment,” Mech. Syst. Signal Process., vol. 72–73, pp. 184–205, May 2016.

L. R. Rabine, “A tutorial on hidden Markov models and selected applications in speech recognition,” Proc. IEEE, vol. 77, no. 2, pp. 257–286, Feb. 1989.

H. Jiang, J. Chen, G. Dong, T. Liu, and G. Chen, “Study on Hankel matrix-based SVD and its application in rolling element bearing fault diagnosis,” Mech. Syst. Signal Process., vol. 52–53, pp. 338–359, Feb. 2015.

D. Peel and G. J. McIlachlan, “Robust mixture modelling using the t distribution,” Statist. Comput., vol. 10, pp. 339–348, Oct. 2000.

C. Liu and D. B. Rubin, “ML estimation of the t distribution using EM and its extensions, ECM and ECME,” Stat. Sinica, vol. 50, no. 1, pp. 19–39, 2007.

A. Solomonoff, W. M. Campbell, and C. Quillen, “Nuisance attribute projection,” in Proc. Speech Commun., 2009, p. 1007.

R. Gopinath, C. Santhosh Kumar, K. I. Ramachandran, V. Upendranath, and P. V. R. Sai Kiran, “Intelligent fault diagnosis of synchronous generators,” Expert Syst. Appl., vol. 45, pp. 142–149, Mar. 2016.

R. Gopinath, C. S. Kumar, and K. I. Ramachandran, “Scalable fault models for diagnosis in a synchronous generator using feature mapping and transformation techniques,” Int. J. Prognostics Health Manage., vol. 9, no. 2, p. 11, 2018.

J. Gandhi, R. Gopinath, and C. S. Kumar, “Nuisance attribute projection for system independent fault diagnosis of synchronous generator,” in Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET), Mar. 2017, pp. 544–549.

J. Gandhi, R. Gopinath, and C. S. Kumar, “System independent fault diagnosis for synchronous generator,” Int. J. Prognostics Health Manage., vol. 8, no. 2, p. 11, 2017.

C. Zhang, B. Li, B. Chen, H. Cao, Y. Zi, and Z. He, “Weak fault signature extraction of rotating machinery using flexible analytic wavelet transform,” Mech. Syst. Signal Process., vol. 64–65, pp. 162–187, Dec. 2015.

Y. Zhang, S. Tong, F. Cong, and J. Xu, “Research of feature extraction method based on sparse reconstruction and multiscale dispersion entropy,” Appl. Sci., vol. 8, no. 6, p. 888.

J. Yuan, F. Ji, Y. Gao, J. Zhu, C. Wei, and Y. Zhou, “Integrated ensemble noise-reconstructed empirical mode decomposition for mechanical fault detection,” Mech. Syst. Signal Process., vol. 104, pp. 323–346, May 2018.

J. Yuan, Y. Wang, Y. Peng, and C. Wei, “Weak fault detection and health degradation monitoring using customized standard multiwavelets,” Mech. Syst. Signal Process., vol. 94, pp. 384–399, Sep. 2017.

Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, “Machinery health prognostics: A systematic review from data acquisition to RUL prediction,” Mech. Syst. Signal Process., vol. 104, pp. 799–834, May 2018.

W. He, Y. Zi, B. Chen, F. Wu, and Z. He, “Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform,” Mech. Syst. Signal Process., vol. 54–55, pp. 457–480, Mar. 2015.

Z. Wei, Y. Wang, S. He, and J. Bao, “A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection,” Knowl.-Based Syst., vol. 116, pp. 1–12, Jan. 2017.

P. R. Shamsi, The Institute of Design Science and Basic Component at Xi’an Jiaotong University (XJTU), Xi’an, China. Accessed: Mar. 2020. [Online]. Available: http://gr.xjtu.edu.cn/web/yaguolei

The Changxing Sumyoung Technology. Zhejiang, China. Accessed: Mar. 2020. [Online]. Available: https://www.sumyoutech.com.cn

HUIMING JIANG received the B.S. degree in mechanical engineering from Xi’an Jiaotong University, in 2011, and the Ph.D. degree in mechanical engineering from Shanghai Jiao Tong University, in 2017. She is currently a Lecturer with the School of Mechanical Engineering, University of Shanghai for Science and Technology. Her current research interests include machinery condition monitoring and fault diagnosis, intelligent fault diagnostics, and performance degradation assessment.

JING YUAN was born in Chongqing, China, in 1983. She received the B.S. and Ph.D. degrees in mechanical engineering from Xi’an Jiaotong University, in 2006 and 2011, respectively. From 2009 to 2010, she visited the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA. From 2011 to 2018, she was a Senior Engineer with the Shanghai Academy of Spaceflight Technology. Since 2018, she has been an Associate Professor with the School of Mechanical Engineering, University of Shanghai for Science and Technology. She is the author of one book, more than 30 articles, more than ten inventions, and two industry standards. Up to now, she has undertaken more than ten projects from various national foundations and enterprises. Her current research interests include mechanical condition monitoring, fault diagnosis, signal processing, and feature extraction.

QIAN ZHAO received the Ph.D. degree in mechanical engineering from Northeastern University, in 2016. She is currently a Lecturer with the School of Mechanical Engineering, University of Shanghai for Science and Technology. Her current research interests include rotating machinery dynamics and vibration control.
HAN YAN received the B.S. degree in mechanical engineering from Xi’an Jiao Tong University, Shanxi, China, in 2011, and the M.S. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. He is currently a Postdoctoral Researcher with the State Key Laboratory of Mechanical System and Vibration and the School of Mechanical Engineering, Shanghai Jiao Tong University. His current research interests include microfluidic devices, nonlinear dynamics, and vibration analysis in MEMS/NEMS.

SEN WANG received the B.S. and M.A.Eng. degrees in mechanical engineering from the Huazhong University of Science and Technology, in 2010 and 2012, respectively. He is currently an Engineer with the Shanghai Aerospace Control Technology Institute. He is also mainly involved with the design and development of The Space Inertia Actuator for satellite attitude control system.

YUNFEI SHAO received the B.E. degree from the Merchant Marine College, Shanghai Maritime University, Shanghai, China, in 2018. He is currently pursuing the M.E. degree with the School of Mechanical Engineering, University of Shanghai for Science and Technology. His current research interest includes mechanical fault diagnosis.

***