Health Indicator Construction Based on MD-CUMSUM with Multi-Domain Features Selection for Rolling Element Bearing Fault Diagnosis

Qiang Li1, Changfeng Yan*1, Wei Wang2, Abdalla Babiker1, Lixiao Wu1
1 School of Mechanical & Electronic Engineering, Lanzhou University of Technology, Lanzhou 730050, China
2 Xi'an CAS Photoelectric Precision Engineering Co., Ltd, Xi'an, Shanxi 710000, China
Corresponding author: Changfeng Yan (e-mail: changf_yan@163.com).

This work is supported by the program of the National Natural Science Foundation of China (grant number. 51765034) and the Hongl iu First-class Disciplines Development Program of Lanzhou University of Technology.

ABSTRACT The initial fault signal of rolling element bearing is extremely weak and could be easily masked by strong background noise. Different features of vibration signal can be different sensitivity to initial fault and performance degradation. Moreover, individual features cannot reflect bearing fault rationally and these features reveal non-monotonic behavior when the bearing condition deteriorates. A Health Indicator (HI) is proposed based on Mahalanobis Distance and Cumulative Sum (MD-CUMSUM). The time-frequency domain features extracted through Singular Value Decomposition based on Variational Mode Decomposition (VMD-SVD) and several optimal time domain features are used to calculate Mahalanobis Distances (MDs). The coarse-to-fine diagnosing strategy is proposed to determine the initial fault of rolling bearing. The obtained HI is utilized to estimate the different performance degradation stages of the bearing depending on the thresholds. This method is verified by utilizing two different experiments. The results demonstrate that the approach has the capability of estimating initial fault and determining degradation stages of bearing.

INDEX TERMS coarse-to-fine diagnosing strategy, health indicator, initial fault, MD-CUMSUM, VMD-SVD

I. INTRODUCTION Rolling bearings play an important role in rotating machinery and are widely used in manufacturing equipment [1]. Accidental faults in running bearings can cause shutdown of the entire machine, which will result in huge economical losses [2]. Therefore the ability to quickly and accurately diagnose the initial fault and assess severity of performance degradation is very important in running bearings [3]-[5]. Recently, the product of health monitoring system has been becoming critical for intelligent maintenance because of its capacity to detect, classify and prognose the impending faults intelligently [6]. Thus, extracting useful fault information from stationary and non-stationary vibration signals is significant for detecting the wear and diagnosing the fault of bearings [7]-[12]. The models of fault diagnosis and prognostic for rolling element bearings are divided into physical models, knowledgeable models, data-driven models and combined models [13]-[15].

Performance analysis (e.g. stages of bearing wear) of rolling element bearings is a valid manner to determine whether the machine is operating normally or not [16]. There are a lot of efforts on extracting fault information from the vibration signals and detecting early fault [17]-[19]. In fact, one feature is only sensitive to specific fault in the particular performance degradation phase, while time domain, frequency domain and time-frequency domain features cannot simultaneously satisfy the sensitivity and stability of rolling bearing performance degradation. Therefore, it is important to establish a composite indicator which is sensitive to incipient fault and rises steadily as the damage increases [20], [21].

In recent years, a variety of data-driven methods (such as artificial neural network, principal component analysis, machine learning and cluster analysis) have been used for condition monitoring of rolling bearing [1]-[3], [6], [8]. The performance degradation process could be quantified through
data-driven methods without creating a specific model that is difficult to obtain. Generally, data-driven prediction for rolling bearing involves three stages containing data acquisition, Health Indicator (HI) construction and residual useful life prediction [22]. The performance degradation process could be identified and quantified by means of the HI [23]. The HI should fuse the information of multi-domain features. [24]. Therefore, it is of great significance to construct HIs to effectively reveal the degradation process of rolling bearings. The Mahalanobis-Taguchi System (MTS) has been used in fault diagnosis and prognosis for rolling bearings [25], [26]. MTS is also a multivariate pattern recognition technique and can be easily applied to measure the degree of abnormality of data and provide accurate results even with smaller sample sizes and without any prior information [27]. MTS uses Mahalanobis Distance (MD) to fuse all relevant information into a single metric. Additionally, the Signal-to-Noise Ratio (SNR) is used to evaluate the effectiveness of MTS. Orthogonal Array (OA) is used to select the variables with the largest SNR [25].

MTS has succeeded in the field of fault diagnosis and prognosis for rolling bearing. Hu et al. [28] utilized MTS-SOM to identify and assess the bearing degradation. Shakya et al. [29], [30] proposed a method by using Mahalanobis-Taguchi-Gram-Schmidt, which can detect different performance degradation stages of bearings online. Wang et al. [31] used the MTS based on EMD-SVD to assess the degradation of rolling bearings. Soylemezoglu et al. [32] presented the MD-based fault clustering method to detect the rolling bearing fault.

Aiming at applying MTS in fault diagnosis, it is necessary to extract appropriate feature parameters from multi-domain features of bearing vibration signals [31]. Lin et al. [33] proposed an approach by utilizing multifractal detrended function analysis as the fault parameter of MTS to extract features. Shen et al. [20] utilized Fuzzy Support Vector Data Description (FSVDD) to construct a monotonic degradation assessment index for bearings. Lei et al. [34] summarized the research and development of EMD for rolling bearing fault diagnosis. Chen et al. [35] utilized Ensemble Empirical Mode Decomposition (EEMD) and adjustment MTS to assess the health states for rolling bearing. Despite EEMD alleviates the mode mixing problem, some apparent limitations still exist [36]. The Variational Mode Decomposition (VMD) method is proposed to avoid these disadvantages. VMD theory was founded by Dragomiretskiy and Zosso. A multi-component signals could be nonrecursively decomposed into multiple Band-Limited Intrinsic Mode Functions (BLIMF) [37]. In particular, the VMD algorithm could more effectively eliminate the noise and decompose signals compared with the EMD-based adaptive decomposition method [37]. Matrix Singular Value Decomposition (SVD) techniques have the ability of noise reduction and signal estimation [38]. The Intrinsic Mode Functions (IMFs) decomposed by VMD are used to establish the original matrix for SVD.

It is difficult for the initial fault diagnosis and performance degradation monitoring to be implemented because of the nonlinearity and uncertainty for the vibration signal [28]. Yu [39] used contribution-analysis-based method and Hidden Markov Model (HMM) to assess the performance degradation of rolling bearing. Theodoros et al. [40] utilized probabilistic Support Vector Regression (SVR) to predict the residual useful life for rolling bearings. Dawn et al. [41] proposed an approach to predict the remaining useful life for rolling bearing by using a degraded characteristic based on the amplitude reduction of a particular frequency. The change of entropy at particular frequencies are used to extract degradation information from vibration signal. All of these methods mentioned above show the capacities of residual service life prediction for rolling bearings. However, these methods would fail to identify the different performance degradation stages of rolling bearing. Reuben et al. [42] presented an approach for estimating bearing failure time by utilizing the features of spectral analysis. The models can be used to identify the third and the fourth performance degradation stages of the rolling bearing, but they failed to estimate the normal and initial fault stages. Sadok et al. [43] put forward an approach named TALAF to track rolling bearing surface degradation using time domain indicator. This method can identify four performance degradation stages for rolling element bearing by using the indicator changes in slope. However, it can only intuititionally identify the different performance degradation stages rather than quantitatively determine the degradation stages.

Therefore, four key issues should be paid more attention to determine the initial fault as early as possible and identify the different performance degradation stages quantitatively, such as (1) a relatively appropriate approach should be chosen to extract features from vibration signals, (2) the incipient fault stage could be detected as accurately and soon as possible before the severe degradation occurs, (3) the most representative and sensitive feature parameters should be optimized from these original characteristics to reveal the performance degradation for rolling element bearing, (4) the HI should be constructed to reflect the initial fault and performance degradation, and quantitatively identify the different performance degradation stages based on the thresholds.

Aiming at solving the four issues mentioned above and effectively realize initial fault detection and condition monitoring, an approach of HI construction is proposed based on MD-CUMSUM with multi-domain features selection. Firstly, vibration signal features are extracted through VMD-SVD. Secondly, the optimal time domain and time-frequency domain characteristics are fused into Mahalanobis Distances (MDs). The coarse-to-fine diagnosing strategy is used to determine the initial fault of rolling bearing. These features associated with the
performance degradation are fused into MD1 and Cumulative Sum (CUMSUM) is utilized to obtain monotonically HI with bearing degradation process. Finally, different four performance degradation stages could be identified depending on the thresholds of HI and the sampling point of initial fault.

The rest of this paper is arranged as follows: In section 2, the methodologies such as VMD, SVD and MTS are introduced briefly. In section 3, the proposed method about coarse-to-fine diagnosing strategy for initial fault and HI construction utilizing MD-CUMSUM are described. In section 4, the proposed method is validated by the experimental results and some discussions are made. In section 5, some conclusions are given.

II. Methodologies

A. VARIATIONAL MODE DECOMPOSITION

An ensemble of band-limited IMFs could be obtained when a multi-component signal is nonrecursively decomposed by VMD [37]. IMF \( u_k(t) \) here is defined as an AM-FM signal [37].

\[
\begin{align*}
\phi_\ell(t) &= A_\ell(t) \cos(\phi_\ell(t)) \\
\phi_\ell(t) &= \text{a non-decreasing function}, \quad A_\ell(t) = \text{an envelope}, \\
\phi_\ell(t) &= \text{non-negative and varying slower than the } \phi_\ell(t).
\end{align*}
\]

VMD is represented by the following equation [37]:

\[
\begin{align*}
\min_{\{u_k\}, \{\omega_k\}} & \left\{ \sum_k \| \delta(t) + \frac{j}{\pi t} u_k(t) \| e^{-j\omega_k t} \right\} \\
\text{s.t.} & \sum_k u_k = f 
\end{align*}
\]

where \( u_k \) is the \( k^{th} \) mode. \( \omega_k \) is the center frequency. 

Eq. (2) is determined by the quadratic penalty and Lagrangian multipliers. The augmented Lagrangian is given in the following equation:

\[
\begin{align*}
L(\{u_k\}, \{\omega_k\}, \lambda) &= \alpha \sum_k \| \delta(t) + \frac{j}{\pi t} u_k(t) \| e^{-j\omega_k t} \| \\
+ & \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle
\end{align*}
\]

where \( \alpha \) is the balancing parameter, and \( \lambda \) denotes Lagrangian multiplier.

Eq. (3) is decided by the alternate direction method of multipliers. The modes obtained from spectral domain are given as:

\[
\hat{u}_k(\omega) = \frac{\hat{f}(\omega) - \sum_{\nu \neq k} \hat{u}_\nu(\omega) + (\hat{\lambda}(\omega)/2)}{1 + 2\alpha(\omega - \omega_k)^2}
\]

VMD mainly includes the following steps.

1) The modes \( \hat{u}_k^{n+1}(\omega) \) are updated as shown in Eq. (5).

2) The center frequencies \( \omega_k^{n+1} \) are updated by Eq. (6).

3) \( \hat{\lambda}^{n+1} \) is updated for all \( \omega \geq 0 \) by Eq. (7) until

\[
\sum_k \left( \| \hat{u}_k^{n+1} - \hat{u}_k^n \|_2^2 / \| \hat{u}_k^n \|_2^2 \right) < \epsilon
\]

\[
\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau \left( \hat{f} - \sum_{k=1}^K \hat{u}_k^{n+1} \right)
\]

VMD method described in detail is shown in Ref. [37].

B. SINGULAR VALUE DECOMPOSITION

The feature extraction of rolling bearing fault diagnosis has been implemented by using SVD [31], [44]. The SVD is defined as follows:

\[
A = U \Sigma V^T
\]

where \( U \) and \( V \) are orthogonal matrices and \( \Sigma \) is diagonal matrix \((\sum_{i \neq j} \neq 0 \text{ if } i = j \text{ and } \sum_{i \neq j} \geq \sum_{j \neq k} \geq \cdots \geq 0)\).

Singular vectors with large singular values represent more information for the original matrix.

C. MAHALANOBIS-TAGUCHI SYSTEM

MTS has been used for pattern recognition, multivariable diagnosis and forecasting [26]. In MTS, MD is calculated to distinguish the abnormal data from the original data [31], [32]. The performance degradation of rolling bearing will be monitored through the tendency of MD. The MTS shown in Fig. 1 consists of three stages:

Stage 1 Construction of the Mahalanobis Space (MS)
In order to distinguish degradation stages of the bearings through the tendency of MD, the Mahalanobis reference space should be constructed in advance. The reference space is determined by the variables of normal condition obtained through the feature parameters of properly lubricated new bearings. The steps to construct MS are as follows [32].

1. Let \( x_i \) be the \( i \)-th feature in the \( j \)-th observation. \( n \) is the number of observations and \( k \) is the total amount of characteristics. Mean \( \bar{x}_{ij} (i=1, 2, 3, \ldots, k) \) for the normal dataset is calculated as:

\[
\bar{x}_i = \frac{\sum_{j=1}^n x_{ij}}{n}
\]
Standard deviation $s_i$ ($i=1, 2, 3, \ldots, k$) for each characteristic is calculated as:

$$s_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2}$$  \hspace{1cm} (10)

3. Normalize each characteristic, and take its transpose $Z_{ij}^T$ from the normalized data matrix $Z_{ij}$:

$$Z_{ij} = \frac{(x_{ij} - \bar{x}_i)}{s_i}$$  \hspace{1cm} (11)

4. $c_{ij}$ of the correlation matrix $C$ are computed as:

$$c_{ij} = \frac{\sum_{m=1}^{n} (Z_{im} Z_{jm})}{n-1}$$  \hspace{1cm} (12)

5. MD of normal data $MD_{normal}$ is calculated as:

$$MD_{normal} = \frac{1}{k} Z_{ij}^T C^{-1} Z_{ij}$$  \hspace{1cm} (13)

where $MD_{normal}$ is the MD of the normal state of the running bearings.

Stage 2 Validation of MS

Aiming at verifying the MS, the correlation matrix, mean and standard deviation of the normal case are utilized to calculate the abnormal case of MD. The MS will be valid if the MD have the capacity to distinguish the abnormal data with the normal data. If the abnormal data cannot be identified by MD, then the combination of the other characteristics is used to detect until the correct features set are found.

Stage 3 Calculation of the MD of sample space

The mean $\bar{x}_i$, correlation matrix $C$ and standard deviation $s_i$ of the valid reference space are utilized to normalize the data of sample space and calculate the MD.

The performance degradation of the running bearing will be monitored according to the tendency of MDs.

$$\hat{Y}_j = \frac{(Y_{ij} - \bar{Y}_j)}{s_j} \quad (i=1, 2, \ldots, k; j=1, 2, \ldots, n)$$  \hspace{1cm} (14)

$$MD_{sample} = \frac{1}{k} \hat{Y}_j C^{-1} \hat{Y}_j \quad (j=1, 2, \ldots, n)$$  \hspace{1cm} (15)

where $\hat{Y}_j$ is the normalized data matrix of sample space, and $MD_{sample}$ is the MD of sample space.

III. PROPOSED METHODS

The proposed method shown in Fig. 2 mainly consists of two stages such as coarse-to-fine diagnosing strategy for initial fault and HI construction by using MD-CUMSUM. In Fig. 2, the features sensitive to initial fault are fused into MD2, which could provide useful information for the coarse-to-fine diagnosing strategy when diagnose the initial fault. MD1 fuses the features related to the performance degradation, which is the basis for the construction of HI. The detailed description of the methods is as follows.

A. COARSE-TO-FINE DIAGNOSING STRATEGY FOR INITIAL FAULT

Initial fault diagnosis plays an important role in rolling bearing condition monitoring. However, the initial fault vibration signal is difficult to be extracted from the highly background noise. In order to overcome this tough problem, the coarse-to-fine diagnosing strategy shown in Fig. 3 is proposed to determine the earliest fault, which mainly contains three parts.

Part 1 3σ criterion of $\Delta MD2$

The optimal time domain features being sensitive to initial fault and singular values of time-frequency domain features are used to construct Mahalanobis reference space and the MD2 is calculated.

The application of 3σ criterion of the increment $\Delta MD2$ is under the condition that $\Delta MD2$ satisfies normal distribution. The probability distribution function of the normal distribution is as:

$$f(X_i, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right)$$  \hspace{1cm} (16)

where $n$ is the total number of random variables, $X_i$ is the $i^{th}$ random variables, $y$ is the probability density of $X_i$, $\mu$ is the mean of $X_i$ and $\sigma$ is the standard deviation of $X_i$.

When the increment $\Delta MD2 \sim N(\mu, \sigma^2)$, the probability of $\Delta MD2$ in $(\mu - 3\sigma, \mu + 3\sigma)$ is 0.9974. It could be regarded as the abnormal state of rolling bearings when the $\Delta MD2$ exceeds the range of $\mu \pm 3\sigma$.

Part 2 Determining fault sample

The vibration signal $S(t)$ of rolling bearing is decomposed through VMD and the IMFs can be obtained. The optimal IMFs are obtained through the maximum value of kurtosis.
and correlation coefficient. Then the envelope spectrum of the optimal IMFs denoised by Multipoit Optimal Minimum Entropy Deconvolution Adjusted (MOMEDA) [45] is obtained through Hilbert transformation. The Autogram [46] is used to optimize the central frequency and bandwidth for the envelope analysis. The fault could be identified by the fault feature frequency from the envelope spectrum.

Part 3 Backtracking algorithm

The sampling point determined by $3\sigma$ criterion of $\Delta$MD2 indicates that the bearing has been in abnormal state, but it may not be the real initial fault. Aiming at this problem, the backtracking algorithm is used to determine the earliest sampling point of fault, which mainly contains the following steps.

**Step 1** Initially, let $a = 0$ and $b = \theta$. Where $\theta$ is the number of rolling bearing sampling point determined by $3\sigma$ criterion of MD2.

**Step 2** Let

$$S_c = S_{[(a+b)/2]}$$  \hspace{1cm} (17)

where $[\cdot]$ is the integral function that could be used to obtain the largest integer that does not exceed the real number. $S_c$ is the $c^{th}$ sampling point of rolling bearing.

**Step 3** If $S_c$ is the fault sample, let $b = c$, else let $a = c$. Then go to Step 2 if $(b-a) > 1$, else go to Step 4.

**Step 4** $S_b$ could be regarded as the initial fault sample if it meets the $\beta$ deviation criterion (the $\beta$ deviation criterion can only be executed once).

The content of the $\beta$ deviation criterion is: Let $\beta = 10$ (the value of $\beta$ is determined by the empirical analysis), $a = c - \beta$ and $b = c$. Then go to Step 2.

The reason for using $\beta$ deviation criterion is to avoid $S_b$ is not the expected initial fault due to the complex operation condition of rolling bearings. For example, the bearing fault feature could be detected at a certain time, however the fault feature could not be detected in the following sampling time due to the uncertainty of the operating state, then the fault feature could be detected again.

B. HI CONSTRUCTION BY USING MD-CUMSUM

The optimal multiple domain features are used to construct Mahalanobis reference space. $n$-IMFs of vibration signal are extracted using VMD. The singular value (time-frequency domain) of each IMF is calculated through SVD. These features in time-frequency domain and time domain consistent with the performance degradation are used to construct the Mahalanobis reference space and fused into MD1. The performance degradation condition could be evaluated by MD1.

Although the MD1 has a good performance in reflecting the similarity among these data, it is difficult to simply reflect the state of each data sample from the MD1 when there is a
continuous small change among these data samples. Therefore, the CUMSUM method is introduced to optimize the MD1 and extract the required HI, which can accurately reflect the monotony of the bearing health status.

CUMSUM is a perfect way to monitor the out-of-control processes [47], [48]. The CUMSUM can provide a monotonous growth curve when the rolling bearing starts to degrade [49]:

\[
0_{1max}(0, 1(0)]_{i} = (\mu_i + e) + CS_{MD}^{+} - \\
0_{1max}(0, (1)]_{i} = (-\mu_i - e) - CS_{MD}^{-}
\]

where the \( CS_{MD}^{+} \) and \( CS_{MD}^{-} \) are the upward and downward CUMSUM. \( \mu_i \) is the target value, which is calculated by the mean of the standard values. \( e \) is the error value, which is typically half of the sample standard deviation.

The HI is given as following:

\[
HI^\prime = CS_{MD}^{+}
\]

(20)

In order to facilitate calculation and analysis, the HI is obtained by normalizing \( HI^\prime \) through Eq. (21).

\[
HI = \frac{HI^\prime - N_{\text{min}}}{N_{\text{max}} - N_{\text{min}}}
\]

(21)

where \( 0 \leq HI \leq 1 \).

A larger HI means a worse bearing health state, and it indicates that the bearing is completely damaged when \( HI=1 \). The different performance degradation stages of rolling element bearing will be determined by the thresholds of HI.

The performance degradation process of rolling bearing should generally involve different four stages such as normal, incipient fault, severe degradation and failure [8], [28], [50].

In order to determine the different performance degradation stage, the standard of the degradation stages takes the advantages of the slope change in the curve indicator and setting the appropriate thresholds to clearly differentiate the different degradation stages [28], [43].

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Two datasets from different experiments [51], [52] are utilized to validate the proposed method. The rolling bearings vibration signals with natural evolution defects are obtained in these two experiments.

A. EXPERIMENT 1

<table>
<thead>
<tr>
<th>TABLE 1. Parameters of the tested bearing.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>Ball number</td>
</tr>
<tr>
<td>Ball diameter (mm)</td>
</tr>
<tr>
<td>Pitch diameter (mm)</td>
</tr>
<tr>
<td>Contact angle (°)</td>
</tr>
<tr>
<td>Rotation speed (rpm)</td>
</tr>
</tbody>
</table>

FIGURE 4. Test rig of PRONOSTIA.

FIGURE 5. RMS and Peak of raw data collected from the test rig.

FIGURE 6. Singular values of IMFs of raw data.

FIGURE 7. Five optimal time domain features of raw data.

FIGURE 8. MD2 of initial fault of rolling element bearing.
The dataset acquired from “IEEE PHM 2012 Prognostic Challenge” with an experimental rig called PRONOSTIA are performed firstly. PRONOSTIA test rig is shown in Fig. 4 [51]. The motor power equals to 250 W and the secondary shaft speed is less than 2000 rpm. The radial force reduces the service life by setting the maximum dynamic load 4000 N. The vibration data were collected every 10s. Besides, the length of data was 2,560 points with the sampling rate 25.6 kHz. The basic parameters of the tested bearing are shown in Table 1 and the BPFI is calculated as 221.66 Hz.

The dataset is processed by using the proposed method mentioned in Section 3. According to the research results of our team, the trends of RMS and Peak shown in Fig. 5 are consistent with the bearing performance degradation through calculating the correlation between the features in time domain and MD [52]. The trends of singular values of 4 IMFs extracted from time-frequency domain features through VMD are shown in Fig. 6, which are similar to RMS and Peak and could reveal the degrading process of rolling element bearing. Therefore the RMS, Peak and singular values of 4 IMFs could be used to construct the Mahalanobis reference space and fused into MD1. The five optimal time domain features shown in Fig. 7 containing Clearance factor, Kurtosis, Crest factor, Impulse factor and Shape factor are sensitive to initial failure of rolling element bearing [52], which have the maximum value in the same time.

Considering the singular values of 4 IMFs all abruptly increase simultaneously and the moment of abrupt increase is similar to that five optimal time domain features. It means that when a sudden increase occurs, the rolling bearing moves from the healthy state to the degraded stage [30]. So these five optimal time domain features and the singular values of 4 IMFs could be used to construct the Mahalanobis reference space. Since the test bearing is in good condition before the experiment, the first 30 datasets of MD2 are used as normal condition data by vibration signal analysis. The MD2 shown in Fig. 8 is obtained.

The coarse-to-fine diagnosing strategy is used to determine the initial fault of bearing. MD2 should be confirmed to
obey the normal distribution before applying the $3\sigma$ criterion. The values of the increment of MD2 are calculated as shown in Fig. 9. It can be calculated that $\Delta MD2 \sim N(0.1170, 6.2411^2)$ as shown in Fig. 10. From Fig. 9 it exceeds the threshold of $3\sigma$ from the 1073$^{\text{th}}$ sampling point.

Then the earliest fault of bearing is determined by the backtracking algorithm. After using the backtracking algorithm, the envelope spectrums of the 1071$^{\text{th}}$ and 1070$^{\text{th}}$ sampling points are shown in Fig. 11. It is shown in Fig. 11(a) that the high impact of the 1071$^{\text{th}}$ sampling point can be seen most conspicuously at 225 Hz and the harmonics around 2X of BPFI. There is a difference of 3.34 Hz between the identified characteristic frequency (225 Hz) and the theoretical value (221.66 Hz), because of the slip and preload of rolling bearing. However, the envelope spectrum of the 1070$^{\text{th}}$ sampling point presented in Fig. 11(b) does not clearly reveal the 2X of inner race fault characteristic frequency, which is masked by the background noise. Therefore the 1071$^{\text{th}}$ data point can be regarded as the sampling point of initial fault, which is the boundary between normal stage and incipient fault of rolling bearing.

To determine the boundary between incipient fault and severe degradation and the boundary between severe degradation and failure of rolling element bearing, HI using MD-CUMSUM needs to be calculated. MD1 shown in Fig. 12 demonstrates severe fluctuation because of unstable damage propagation in rolling element bearing. The first 30 datasets of MD1 are also used as normal condition data of rolling element bearing, which are regarded as standard values, therefore the target value $\mu_p$ is calculated as 0.9401. The error value $e$ is 47.1415, which is half of the sample standard deviation.

According to Eq. (18), Eq. (20) and Eq. (21), HI is calculated as shown in Fig. 13. Compared with Fig. 12, the HI using MD-CUMSUM is more stable and consistent with the monotonic principle of bearing performance degradation state. In Fig. 13, the 1179$^{\text{th}}$ data point exceeds the threshold of 0.1 and the 1304$^{\text{th}}$ data point exceeds the threshold of 0.5. Therefore, the 1179$^{\text{th}}$ data point is regarded as the boundary between incipient failure and severe degradation of rolling element bearing, and the 1304$^{\text{th}}$ data point is the boundary between severe degradation and failure of bearing.

According to the determined sampling point of initial fault and the thresholds in Fig. 13, four degradation stages of rolling element bearing can be determined by HI as Fig. 14, such as (1) Normal stage: from the 1$^{\text{st}}$ sampling point to the 1070$^{\text{th}}$ sampling point, (2) Incipient fault: from the 1071$^{\text{th}}$ sampling point to the 1178$^{\text{th}}$ sampling point, (3) Severe degradation: from the 1179$^{\text{th}}$ sampling point to the 1303$^{\text{th}}$ sampling point, (4) Failure: from the 1304$^{\text{th}}$ sampling point to the last sampling point.

### B. EXPERIMENT 2

Aiming at making a further verification of the presented method, the second tested dataset is used. This data set was provided by the Intelligent Maintenance System (IMS) Center. Four Rexnord ZA-2115 double-row bearings were installed on a shaft of the test rig as shown in Fig. 15 [53]. The shaft was driven by an AC motor at 2000 rpm and connected by a friction belt, and a 6000 lbs radial load was put on the shaft by a spring mechanism. The vibration signal data was collected every 10 minutes with a sampling rate of 20 kHz and the data length was 20,480 points. It took 7 days for the rolling element bearing to fail. The basic parameters for the tested bearing are shown in Table 2 and BPFO is 236.42 Hz.

MD1 and MD2 are calculated by using the same method as experiment 1. The first 30 datasets of MD2 shown in Fig. 16 are used as normal state data for rolling bearing. The values for the increment of MD2 are calculated as shown in Fig. 17. Because $\Delta MD2 \sim N(0.1267, 5.1794^2)$ as shown in Fig. 18, the $3\sigma$ criterion could be utilized. It exceeds the threshold of $3\sigma$ from the 703$^{\text{th}}$ sampling point as shown in Fig. 17. Autogram is a newly developed and valid method to optimize the central frequency and bandwidth for envelope analysis [46]. After the backtracking algorithm, the 533$^{\text{rd}}$ sampling point is shown in Fig. 19. It could be observed in Fig. 19 that the central frequency is 4921.875 Hz and the bandwidth is 156.25 Hz. However, according to Ref. [54], it is clear and important for diagnosis when the selected

![FIGURE 14. HI using MD-CUMSUM.](https://example.com/fig14)

![FIGURE 15. Test rig of bearings run to failure.](https://example.com/fig15)

| TABLE 2. Parameters of tested Rexnord ZA-2115 bearing. |
|------------------------|------------------------|
| **Parameter**         | **Value**              |
| Ball number           | 16                     |
| Ball diameter (mm)    | 8.4                    |
| Pitch diameter (mm)   | 71.5                   |
| Contact angle (°)     | 15.17                  |
| Rotation speed (r/min)| 2000                   |

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
The optimal center frequency shown in Fig. 20 is 4375 Hz and the bandwidth is 1250 Hz, which could be regarded as the optimal filter parameters. The squared envelope spectrum based filtering algorithm of the 533rd sampling point is displayed in Fig. 21(a). It is shown in Fig. 21(a) that the high
impact of the 533th sampling point can be seen most conspicuously at 230.7 Hz and the harmonics around 2X and 3X. These high impacts are close to the theoretical value 236.42 Hz of the outer race fault characteristic frequency. However, it could be observed in Fig. 21(b) that the squared envelope spectrum of the 532th sampling point does not clearly reveal the outer race fault characteristic frequency, which is masked by the background noise. Therefore the 533th sampling point can be regarded as incipient failure, which is the boundary between normal stage and incipient fault of rolling bearing.

Aiming at determining the boundary between incipient fault and severe degradation and the boundary between severe degradation and failure of rolling element bearing, MD1 is calculated and shown in Fig. 22. The first 30 datasets of MD1 are used as normal condition data of rolling element bearing, therefore, the target value $\mu_0$ is calculated as 0.9187 and the error value $\epsilon$ is 13.9740. According to Eq. (18), Eq. (20) and Eq. (21), the calculation of HI MD-CUMSUM is shown in Fig. 23. In Fig. 23, it exceeds the given threshold of 0.1 from the 727th data point and exceeds the threshold of 0.5 from the 921th data point. Therefore, the 727th data point can be considered as the boundary between incipient fault and severe degradation, and the 921th data point is the boundary between severe degradation and failure of bearing.

According to the sampling point of initial fault and thresholds in Fig. 23, four degradation stages of rolling bearing can be identified by HI as shown in Fig. 24, such as (1) Normal stage: from the 1st data point to the 532th data point, (2) Incipient fault: from the 533th data point to the 726th data point, (3) Severe degradation: from the 727th data point to the 920th data point, (4) Failure: from the 921th data point to the last data point.

C. DISCUSSIONS

Aiming at demonstrating the advantages of the proposed method when detecting the initial fault of rolling bearing, an approach by using ICD and tunable Q-factor wavelet transform [19] is used for comparison. In Ref. [19], the first experimental dataset was provided by IMS, which is the same with our proposed method as described in section 4.2. In their method, the result by means of kurtosis indicator is presented in Fig. 25, and the kurtosis indicator increases abruptly at the 643th sampling point. The occurring time of the 643th sampling point could be regarded as the severe fault [55]. Because the value of kurtosis at the 534th sampling point is nearly the same as the normal state (1-533 sampling point), the 534th sampling point is regarded as the incipient fault of rolling bearing. The envelope demodulation spectrum of the 534th data group after using ICD-OTQWT method is presented in Fig. 26, and the outer race fault feature frequency $f_{BPF0} = 236.42$ Hz and its side frequencies can be extracted. However, compared with our result as shown in Fig. 21(a), our method has the capacity to reveal the outer race fault of the rolling bearing from the data group 648 to 533, which is 1 data group (10 minutes) earlier than Fig. 26. Even Fig. 26 fails to reveal the 2X and 3X of outer race fault characteristic frequency, which is immersed in heavy background noise. The reason is that the selecting of frequency band is important for the demodulation, and using the de-noised process to increase the SNR before envelope analysis with demodulation is necessary. The comparison between ICD-OTQWT and the proposed method is shown in Table 3.

Aiming at verifying the presented method when determining the performance degradation stages of rolling element bearing, a case in Ref. [56] is used for comparison. In their method as shown in Fig. 27, the whole life of rolling element bearing is drawn by the RMS of the vibration signal, which uses the same test data provided by IMS as described in section 4.2. According to the change of RMS, the performance degradation of rolling bearing in the entire life is divided into three stages [56], [57]. A test point selected one day before the beginning of stage II is used to accurately detect the incipient fault of rolling bearing. From Fig. 28, the fault characteristic frequency could be revealed through the envelope spectrum of the target mode. The harmonics could be clearly seen after using their proposed method, which indicates that the selected test point is fault and could be regarded as initial fault of the rolling bearing. Compared with their result, the tendency of the condition monitoring curve MD1 shown in Fig. 29 is basically the same as Fig. 27, and the boundary data point between each performance degradation stage depended on the Fig. 24 is similar to theirs. Furthermore, the sampling point of initial fault (boundary between Stage I and Stage II) shown in Fig. 24 is determined
by using coarse-to-fine diagnosing strategy, which can reveal the fault at an early time more reasonably than the test point selected intuitively as shown in Fig. 27. Therefore the correctness and rationality of the HI are verified when dividing the stages of performance degradation for rolling bearing. The comparison between RMS and the proposed method is shown in Table 4.

Because the features correlated with performance degradation of rolling element bearing are fused into a single HI by using MD-CUMSUM, the HI is more robust than the single feature RMS. Through comparison and discussion, it can be drawn that the HI is effective in estimating the performance degradation state and identifying the fault at an early stage of rolling element bearing.

Although the choice of the two experimental data are low speed (around 2000 rpm), the proposed method is still applicable to high speed bearings experiment. When it comes to high speed bearings, the amplitude of HI will be larger than that in low speed bearings, and it will take less time to cause bearing failure, especially in Stage II, Stage III and Stage IV.

Compared with experiment 2, the Autogram has been actually tried for experiment 1, but it fails to reveal the appropriate demodulation band for square envelope analysis.
(1) The coarse-to-fine diagnosing strategy could be used to determine the initial fault of rolling bearing as early and accurately as possible.

(2) The HI using MD-CUMSUM overcomes the shortcoming that MD is inconspicuously to reflect the small variation of sample and the monotonicity of the bearing health state. The four performance degradation stages of bearing can be more rationally reflected depend on the sampling point of initial fault and the thresholds of normalized HI.

(3) The proposed method based on MD-CUMUM with multi-domain features selection has the ability to be employed in real-time production due to the effectiveness and efficiency of the used algorithms.

REFERENCES


Qiang Li is currently an M.S. candidate in School of Mechanical and Electrical Engineering at Lanzhou University of Technology. He received his B.S. from School of Mechanical and Electrical Engineering at Lanzhou University of Technology in 2017. His research interests include prognostics and health management, condition monitoring and fault diagnosis.

Changfeng Yan is currently a professor in School of Mechanical & Electronic Engineering at Lanzhou University of Technology. He received his B.S. from Huazhong University of Science & Technology in 1996, M.S. from Shenyang University of Technology in 2002 and Ph.D from Tongji University in 2010. His research interests include fault diagnosis and signal process.