



A dependence-based feature vector and its application on planetary gearbox fault classification



Libin Liu, Xihui Liang, Ming J. Zuo*

Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta T6G 1H9, Canada

ARTICLE INFO

Article history:

Received 6 November 2017

Received in revised form 11 May 2018

Accepted 6 June 2018

Available online 13 June 2018

Handling Editor: K. Shin

Keywords:

Tail dependence
Archimedean copula
Feature vector
Condition indicator
Planetary gearbox
Fault classification

ABSTRACT

To achieve planetary gearbox fault classification, vibration signal analysis has been widely employed with rich information about the health status and easy measurement. It is critical to extract features with enough health status information for fault classification. The self-adaptation of ensemble empirical mode decomposition (EEMD) indicates the dependence between the raw vibration signal and EEMD-decomposed intrinsic mode functions (IMFs). In this study, we develop a novel fault feature vector based on the dependence. To develop the dependence-based feature vector, simulated vibration signals with different sun gear tooth crack levels are analyzed. The dependence between the raw signal and each IMF is investigated by Archimedean copulas. With the goodness-of-fit test, the copula estimation closest to the perfect fit is selected for dependence representation. The parameter of the selected copula is applied to develop the dependence-based feature vector. To test the ability of the dependence-based feature vector in fault classification for a real planetary gearbox, experimental vibration signals with different gear fault levels at different gears are classified by a multi-class support vector machine. The classification accuracy of the developed feature vector is compared with that of a reported indicator. Results show the dependence-based feature vector provides higher classification accuracy than the reported, indicating the developed feature vector contains more health status information. The developed feature vector can serve better for planetary gearbox fault classification.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

As a planetary gearbox can provide a high transmission ratio and a high power density within a compact structure, planetary gearboxes are widely used in heavy industrial applications such as helicopters, wind turbines and construction machinery [1]. However, despite the advantages, a planetary gearbox is vulnerable to gear tooth damages such as tooth crack, tooth pitting and tooth breakage due to the tough working environment and the heavy load [2]. Such gear tooth malfunctions would lead to reduction of transmission efficiency and breakdown of the system, resulting in immense economic losses and even human casualties. Consequently, planetary gearbox fault detection and fault diagnosis have been attracting increasing interest from researchers as an important research topic to guarantee the reliability of the system.

* Corresponding author.

E-mail address: ming.zuo@ualberta.ca (M.J. Zuo).

For machinery fault detection and fault diagnosis, vibration signal analysis has been widely used [3–6] thanks to the easy acquisition of the vibration signal and rich information in the vibration signal about the machinery health status. Gilles [7] reported the empirical Wavelet transform (EWT) to analyze a signal with a wavelet filter bank which is built based on the information in the signal. Chen et al. [8] applied EWT to bearing fault diagnosis for wind turbine generator. With gear faults, relevant fault signatures are introduced into vibration signals [9]. By effective signal processing methods, the features can be extracted to indicate the possible faults [10–12], serving as the condition indicators (CIs) for machinery fault detection and fault diagnosis.

Up to date, various features have been reported as CIs by different statistical extraction approaches in the literature. The conventional features can be calculated from the waveform in the time domain and the spectrum in the frequency domain, such as skewness, shape factor, kurtosis, crest factor, frequency center, standard deviation frequency, and so on [13]. As these conventional features are well established, their descriptions are omitted here. One who are interested can refer to [13] for detailed information. However, due to the complexity of a planetary gearbox vibration signal [14], it is difficult to classify the location or the severity of a fault in the planetary gearbox through the conventional features as CIs [15]. Subsequently, more sophisticated algorithms for advanced CIs have received intensive investigation in recent years, such as algorithm combining fast dynamic time warping (fast DTW) and correlated kurtosis (CK) techniques reported in Ref. [11], the windowing and mapping strategy for gear tooth fault detection reported in Ref. [16], and the accumulative amplitudes of carrier orders (AACOs) reported in Ref. [17] designed specifically for planetary gearboxes. However, for the algorithm reported in Ref. [11], the gear fault location identification is based on the analysis of the residual signal, the absolute difference between the warped signals of the measured signal and the estimated reference signal. The performance of the algorithm highly depends on the estimated reference signal. Only when the accuracy and appropriate usage of the estimated reference signal are guaranteed, can the superiority of the method be highlighted [18]. As for the windowing and mapping strategy reported in Ref. [16], besides the complicated algorithm for the windowing and the mapping, its accuracy depends on the window function selection with a trade-off at the computational cost [19]. The AACO in Ref. [17] is based on the fault mechanism investigation and the observation that the characteristic frequencies of gears in a planetary gearbox are integer multiples of the carrier rotating frequency. For a planetary gearbox with the ring gear fixed, the characteristic frequencies of the planetary gearbox are given as follows in terms of the carrier rotating frequency [1]: $f_p = (N_p - N_r)f_c/N_p$, $f_s = (N_r + N_s)f_c/N_s$, $f_r = 0$, $f_m = N_r f_c$, and $f_{p-p} = M_p f_c$, where N_p , N_r , and N_s are the numbers of teeth of the planet gear, the ring gear, and the sun gear, respectively; M_p is the number of planet gears; f_p , f_s , f_r , and f_c are the rotating frequencies of the planet gear, the sun gear, the ring gear, and the carrier, respectively; f_m is the meshing frequency of the planetary gearbox; f_{p-p} is the passing frequency of the planet gears. By the above equations, it can be noticed that f_r , f_m , and f_{p-p} are integer multiples of f_c while f_p and f_s are not integer multiples of f_c if N_r/N_p and N_r/N_s are not integers. Thus, the AACO may not work as well as presented in Ref. [17] for the fault detection and the fault diagnosis of a planetary gearbox whose N_r/N_p and N_r/N_s are not integers. Therefore, it is desirable to develop new features that are able to better extract the health status information from a vibration signal. In this paper, this concern will be explored and investigated with a novel signal processing method, aiming at developing a feature vector with more health status information to better serve the planetary gearbox fault detection and fault diagnosis. The developed feature vector is expected to detect the existence of a fault and diagnose the fault position and the level.

Specifically, the targeted feature vector is on the strength of the ensemble empirical mode decomposition (EEMD) and the tail dependence between the raw signal and EEMD-decomposed intrinsic mode functions (IMFs). Different from the orthogonal decomposition methods like Fourier transform and Wavelet transform, the EEMD method is an iterative data driven method that does not imply orthogonality amongst the decomposed IMFs and the raw signal [20], which results in the possible dependence between the raw signal and the IMFs [21]. In probability theory, the tail dependence describes the co-movement of variables in the distribution tails [22]. For example, if we have two variables U and V , the upper (lower) tail dependence means that with large (small) values of U , large (small) values of V are expected. By intuitive understanding, when a fault-induced transient impulse is introduced in a vibration signal, the transient impulse will distribute in IMFs decomposed by EEMD, i.e. more extreme values in the raw signal means more chance to observe extreme values in IMFs. Thus, there is possible tail dependence between the raw signal and the IMFs. Besides, different faults introduce different transient impulses which distribute differently in IMFs, meaning that different faults correspond to different tail dependence levels. Consequently, if we could capture and describe the different tail dependences properly, novel features can be developed with potential to achieve machinery fault detection and fault diagnosis.

To describe the tail dependence, copulas are used. A copula is an alternative to correlation for dependence description [23]. More powerful, copulas contain information about the joint behavior of the variables in the distribution tails, i.e. tail dependence, which correlation cannot describe [24]. For a parametric copula, the copula parameter works as the coefficient to describe the tail dependence level [22]. As different faults correspond to different tail dependence levels, copula parameters are different accordingly. Following the above logic and hypothesis, a novel feature vector is developed with the parameter of a parametric copula.

Once the dependence-based feature vector is developed, the next concern is the objective measure to its performance on machinery fault detection and fault diagnosis. To address this concern, its application on the planetary gearbox fault classification as a pattern recognition problem is employed. For a pattern recognition problem, various classification methods have been reported, such as artificial neural network (ANN) [25], decision tree [26], and support vector machine (SVM) [27]. Given a specific classification method, the classification accuracy is affected by the inputs which are the employed features

[28]. To compare the classification accuracies with different inputs, one classification method should be specified. In this study, we employ the SVM method. The more useful information about the health status the input has the more accurate classification the SVM can obtain [25,28]. To demonstrate the advantage of the proposed dependence-based feature vector, it serves as the input of a SVM model to check the classification accuracy. For comparison purpose, the reported AACO designed specifically for planetary gearbox fault diagnosis in Ref. [17] and an AACO-based feature vector are input into the SVM model. Classification accuracies by the SVM model with different inputs are compared as the objective performance measure.

The remainder of this paper is organized as follows. In Section 2, fundamentals and theories on EEMD, copula theory, and SVM are reviewed. The development of dependence-based feature vector is given in Section 3 with the simulated planetary gear set vibration signal analysis. After the development of the proposed feature vector, its application on experimental planetary gearbox vibration signal classification is studied with a multi-class SVM. The performance of alternatives which are the AACO reported in Ref. [17] and the AACO-based feature vector is studied as well. The accuracies are then compared to highlight the advantage of the proposed feature vector. At last, discussions and conclusions are given in Section 5 and Section 6, respectively.

2. Fundamentals and theories

In this section, we provide fundamentals of the EEMD theory and the copula theory as they are needed for the development of the proposed feature vector in Section 3. The EEMD is used to decompose a raw vibration signal into IMFs and the copula theory is used to describe the tail dependence between the raw signal and each IMF. The SVM method is presented since it is used to calculate the classification accuracy subject to a specific input in Section 4, serving as the objective measure to indicate the health status information in the input.

2.1. Ensemble empirical mode decomposition

The empirical mode decomposition (EMD), developed by Huang et al. [29], is first reviewed as it is the fundamental of the EEMD. Essentially, the EMD is an adaptive signal decomposition method to decompose a signal into IMFs [30]. IMFs are oscillatory functions with varying amplitude and frequency, satisfying two conditions [20]: 1) Throughout the whole length of a single IMF, the number of extrema and the number of zero-crossings must either be equal or differ at most by one; 2) At any data location, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. By EMD, one time series $x(t)$ can be decomposed into

$$x(t) = \sum_{i=0}^n c_i + r_n \quad (1)$$

where c_i is the i th IMF and r_n is the residue of $x(t)$ after n IMFs are extracted. The EMD is implemented by the following sifting process using local extrema [30]:

- 1) Initialize: let $r_0(t) = x(t)$ and $i = 1$;
- 2) Extract the i th IMF
 - 2.1) Initialize: let $j = 0$ and $h_{ij}(t) = r_{i-1}(t)$;
 - 2.2) Find the local minima and the local maxima of $h_{ij}(t)$;
 - 2.3) Interpolate the local minima and the local maxima by cubic spline to construct the lower envelope and the upper envelope of $h_{ij}(t)$;
 - 2.4) Calculate the local mean $m_{ij}(t)$ of the lower envelope and the upper envelope;
 - 2.5) Update: $h_{ij}(t) = h_{ij}(t) - m_{ij}(t)$;
 - 2.6) Repeat step 2.2) to step 2.5) until the envelopes are symmetric with respect to zero mean. The final $h_{ij}(t)$ is designated as the i th IMF $c_i(t)$, i.e. $c_i(t) = h_{ij}(t)$;
- 3) Let $r_i(t) = r_{i-1}(t) - c_i(t)$;
- 4) Let $i = i + 1$ and return to step 2) until the residual $r_i(t)$ becomes a monotonic function from which no more IMFs can be extracted.

The EMD works as an effective self-adaptive dyadic filter bank for a white noise series [20]. However, when the data series is of intermittency as a mixture of intermittent high-frequency oscillations riding on a continuous lower-frequency signal, the dyadic property of EMD is compromised, leading to the mode mixing [20]. To overcome the mode mixing, the EEMD was developed in Ref. [20].

The EEMD is a noise-assisted data analysis method taking advantage of the statistical properties of white noise [20]. Adding white noise with finite amplitude could provide a uniformly distributed reference scale. When a signal is added to this uniformly distributed white noise background, the component in different scales of the signal are projected onto proper scales of the white noise, which collates the component of the signal with comparable scale into one IMF [20]. In this way, the drawback of mode mixing of EMD is overcome. Different white noises with finite amplitude are added to the signal for

different trials. By the ensemble mean of enough trials, the noise can be averaged out. The detailed procedure of EEMD is as follows [31]:

- 1) Set the number of trials and the amplitude of the zero-mean white noise;
- 2) Generate the white noise series and add it to the signal;
- 3) Decompose the composite signal with the white noise into IMFs by EMD;
- 4) Return to Step 2) and redo Step 3) for the predefined number of trials. Each trial is with different white noise of the same amplitude;
- 5) Calculate the ensemble mean of corresponding IMFs obtained above as the final result.

The final standard deviation of error ϵ_n introduced by the added white noise follows the statistical rule [20]

$$\epsilon_n = \epsilon / \sqrt{N} \quad (2)$$

where N is the number of trials and ϵ is the amplitude of the added noise. To make EEMD effective in the extrema change with negligible standard deviation of error, the amplitude of the added noise can be 0.2 times the standard deviation of the signal and the number of trials can be a few hundred, as suggested in Ref. [20].

Another concern about EEMD is that the EEMD-decomposed components are not necessarily IMFs as EEMD involves summation of numerous IMFs [20]. To address this concern, a post-processing method with another round of EMD is reported in Ref. [20]. Specifically, EMD is applied to the combination of the first two components obtained from EEMD. The first IMF from EMD is treated as the desired IMF. The remainder is calculated as the difference between the combination and the desired IMF. Then the summation of the remainder and the next component from EEMD is subject to EMD again. This process is carried out consecutively until all the EEMD-decomposed components are traversed through.

2.2. Copula theory

2.2.1. Copulas

Copula, first proposed by Sklar [32], is a mathematical theory to describe the dependence between random variables [22]. A copula \mathbf{C} is a function from $[0, 1]^2$ to $[0, 1]$ with the following properties [33]:

- 1) $\mathbf{C}(u, 0) = \mathbf{C}(0, v) = 0$ for all $(u, v) \in [0, 1]^2$;
- 2) $\mathbf{C}(u, 1) = u$ and $\mathbf{C}(1, v) = v$ for all $(u, v) \in [0, 1]^2$;
- 3) For all $(u_1, u_2, v_1, v_2) \in [0, 1]^4$ with $u_1 \leq u_2$ and $v_1 \leq v_2$, it has $\mathbf{C}(u_2, v_2) - \mathbf{C}(u_1, v_2) - \mathbf{C}(u_2, v_1) + \mathbf{C}(u_1, v_1) \geq 0$;
- 4) For all $(u, v) \in [0, 1]^2$, $\max(u + v - 1, 0) \leq \mathbf{C}(u, v) \leq \min(u, v)$.

To describe the tail dependence between the EEMD-decomposed IMFs and the raw signal, Archimedean copulas are adopted in this study as they have distinct upper and lower tail dependence coefficients [34] with simple mathematical forms. The function of an Archimedean copula is in the form of [22]

$$\mathbf{C}(u, v) = \phi^{-1}(\phi(u) + \phi(v)) \quad (3)$$

where $\phi : (0, 1) \rightarrow [0, +\infty)$ is a decreasing convex function with $\phi(1) = 0$ and $\phi(0) = +\infty$. The function ϕ is called the generator. Archimedean copulas used in this study are Frank copula, Clayton copula, and Gumbel-Hougaard (GH) copula, which can qualitatively show different tail dependences [22]. Specifically, Frank copula has no tail dependence while Clayton

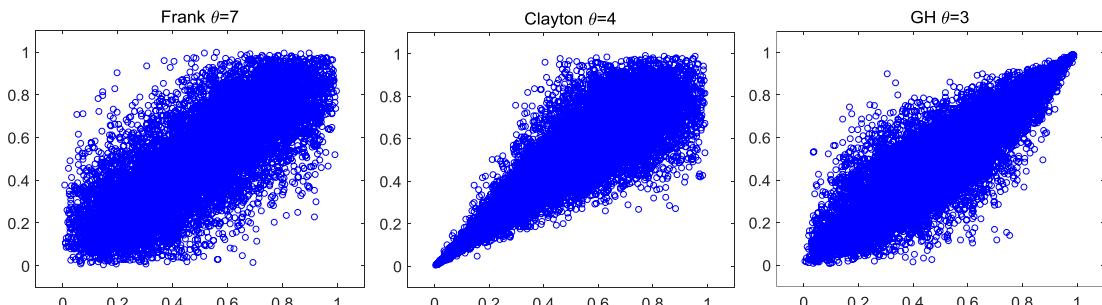


Fig. 1. Tail dependences represented by the three Archimedean copulas.

copula and GH copula have lower tail dependence and upper tail dependence, respectively, as shown in Fig. 1. Table 1 gives the function $\mathbf{C}_\theta(u, v)$, generator $\phi_\theta(t)$, and the range of parameter θ for each of the three Archimedean copulas.

2.2.2. Copula parameter estimation

Suppose we have data samples (X, Y) , for the dependence analysis by copulas, it is recommended to use rank pairs (R, S) associated with data samples (X, Y) [22]. The rank pairs can guarantee a unique copula which is invariant with the monotone transformation of the data samples [24]. To obtain the rank pairs (R, S) , following equations are used [22]

$$R(i) = X(i)/(n + 1) \quad (4)$$

$$S(i) = Y(i)/(n + 1) \quad (5)$$

where n is the number of data samples in X or Y . Note that the number of data samples in X is the same as the number of data samples in Y . With rank pairs (R, S) , we use the Matlab command of “copulafit” to estimate the Archimedean copula coefficients by the maximum likelihood estimation (MLE) method.

2.2.3. Goodness-of-fit test

Statistically, the goodness-of-fit describes how well an estimated model fits a set of data samples. In this study, we use QQ plot [24], short for quantile-quantile plot, to conduct the goodness-of-fit test for the copula estimations. A quantile in the QQ plot means the percentage of points below the given value, for example the 0.3 quantile is the point where 30% of the data fall below and 70% fall above. Specifically, the QQ plot is made by plotting the pairs $(W_{i:n}, W_{(i)})$ [22] where $W_{(1)} \leq W_{(2)} \leq \dots \leq W_{(n)}$ is the order statistics associated with $W_i = \frac{1}{n} \#\{j : X(j) \leq X(i), Y(j) \leq Y(i)\}$ and $W_{i:n}$ is given by

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w \{K_\theta(w)\}^{i-1} (1 - K_\theta(w))^{n-i} dK_\theta(w) \quad (6)$$

In statistics, the i th order statistic is equal to the i th-smallest value in the sample. For Archimedean copulas, $K_\theta(w)$ is formulated as [22]

$$K_\theta(w) = w - \frac{\phi(w)}{\phi'(w)} \quad (7)$$

where $\phi(w)$ is the Archimedean copula generator and $w \in (0, 1)$.

In the QQ plot, a reference diagonal line ($W_{i:n} = W_{(i)}$) is plotted to represent the perfect fit [22]. As a measure to the fitness of the estimated copula on the dependence between the data samples, average distance of the curve regarding the copula estimation to the reference diagonal line is calculated. The average distance is defined as the mean of the absolute difference of the point coordinates on the curve. The shorter the average distance is, the better the fitness that the estimated copula can provide [22].

2.3. Support vector machine

SVM is a machine learning algorithm for pattern recognition to categorize data into different classes based on optimization theory [35]. The basic SVMs are originally developed to tackle the binary classification problem. Consider a binary classification problem with training data as $\{(\mathbf{z}_1, y_1), (\mathbf{z}_2, y_2), \dots, (\mathbf{z}_n, y_n)\}$ where $\mathbf{z}_i \in \mathbb{R}^m$ is the i th input data and $y_i \in \{1, -1\}$ is the class label associated with \mathbf{z}_i . If the training data are linearly separable, a separating plane can be found in the input space, expressed as [13]

Table 1

Functions, generators, and parameter ranges of the three Archimedean copulas.

Copula	$\mathbf{C}_\theta(u, v)$	$\phi_\theta(t)$	Range of θ
Frank	$-\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right)$	$-\ln \frac{e^{\theta t} - 1}{e^\theta - 1}$	$(-\infty, \infty) \setminus \{0\}$
Clayton	$(\max(u^{-\theta} + v^{-\theta} - 1, 0))^{-1/\theta}$	$(t^{-\theta} - 1)/\theta$	$[-1, \infty) \setminus \{0\}$
GH	$\exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta})$	$(-\ln t)^\theta$	$[1, \infty)$

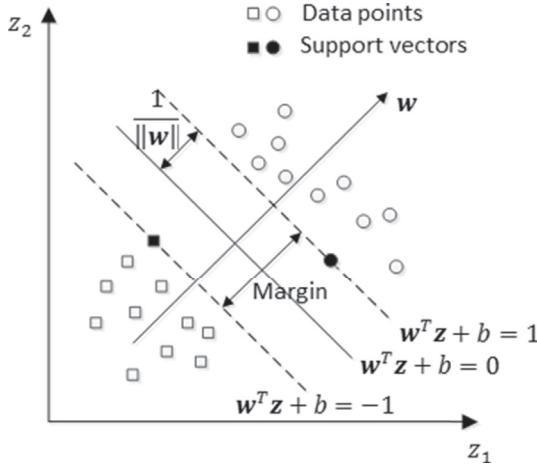


Fig. 2. A linearly separable classification problem with SVM.

$$f(\mathbf{z}) = \mathbf{w}^T \mathbf{z} + b = \sum_{j=1}^m w_j z_j + b = 0 \quad (8)$$

where $\mathbf{w} \in \mathbb{R}^m$ is a weight vector; b is a scalar; and T means the transpose operator.

Fig. 2 demonstrates a linearly separable classification problem in a two-dimensional space. The solid squares and the solid circles on the boundaries $\mathbf{w}^T \mathbf{z} + b = \pm 1$ are called support vectors [13]. All training data are constrained by the following inequality for the ideal without misclassification [13]

$$y_i f(\mathbf{z}_i) = y_i (\mathbf{w}^T \mathbf{z}_i + b) \geq 1 \text{ for } i = 1, 2, \dots, m \quad (9)$$

The distance between the boundaries is called margin [13]. Given above boundaries, the margin can be calculated quantitatively by

$$d = \frac{|1 - (-1)|}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|} \quad (10)$$

where $\|\mathbf{w}\|$ is the 2-norm of \mathbf{w} , defined as $\|\mathbf{w}\| = \sqrt{(w_1)^2 + (w_2)^2 + \dots + (w_m)^2}$.

To obtain the optimal separating plane, SVM employs an optimization process that maximizes margin and minimizes noise with the slack variables [13]. The slack variables are used to relax the constraints by considering points that fail the margin requirement defined by Eq. (9). It allows a soft margin classification which ignores a few noise data to reduce the boundary complexity [13]. The optimization problem is defined as [13]

$$\text{Minimize } \frac{1}{2} \|\mathbf{w}\|^2 + d \sum_{i=1}^m \xi_i \quad (11)$$

Subject to $y_i (\mathbf{w}^T \mathbf{z}_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, m$

$$\xi_i \geq 0 \quad (12)$$

where d is a positive constant serving as the penalty parameter to define the trade-off between the misclassification and boundary complexity; ξ_i is the slack variable representing the distance of a data point of misclassification to the boundary of its true class. This optimization problem can be solved by the Lagrangian method by introducing the Lagrange multipliers α_i and β_i . Its solution process can be found in Ref. [13]. After obtaining solutions to α_i , \mathbf{w} , and b , the decision function is given by Ref. [13]

$$l_f = \text{sign} \left(\sum_{i=1}^m \alpha_i y_i (\mathbf{z}_i^T \mathbf{z}) + b \right) \quad (13)$$

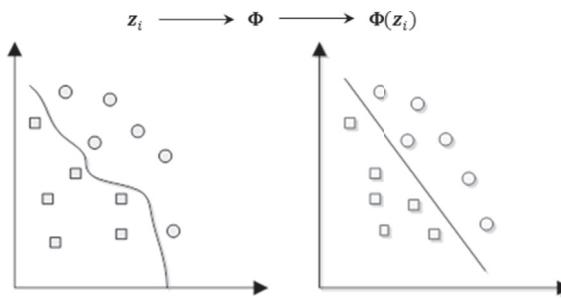


Fig. 3. Feature mapping enabling linear data separation with SVM.

where (\mathbf{z}_i, y_i) is the i th training data pair; \mathbf{z} is a new input data; l_f is the label assigned by the SVM to the new input data \mathbf{z} ; and $\text{sign}(A)$ is the sign function of A which is defined as

$$\text{sign}(A) = \begin{cases} -1 & \text{if } A < 0 \\ 1 & \text{if } A > 0 \end{cases} \quad (14)$$

When the given data are not linearly separable, Eq. (13) is no longer appropriate. To handle the non-linearly separable data, a mapping strategy is introduced to project the original input data to a feature space where the features can be linearly separated. Fig. 3 shows an example of feature mapping from a two-dimensional input space to a two-dimensional feature space using a mapping function $\Phi(\cdot)$.

Accordingly, the non-linear decision function is given with a similar form with Eq. (13) by Ref. [13]:

$$l_f = \text{sign}\left(\sum_{i=1}^m \alpha_i y_i (\Phi^T(\mathbf{z}_i) \Phi(\mathbf{z})) + b\right) \quad (15)$$

By applying a kernel function $K(\mathbf{z}_i, \mathbf{z}) = \Phi^T(\mathbf{z}_i) \Phi(\mathbf{z})$ which satisfies Mercer's theorem [36], the explicit form of $\Phi(\cdot)$ can be avoided since only the inner product of $\Phi^T(\mathbf{z}_i) \Phi(\mathbf{z})$ is needed in Eq. (15) [13]. Several kernel functions are available in the literature such as linear kernel, polynomial kernel, and Gaussian kernel [13]. With a kernel function, the non-linear decision function is then given by Ref. [13]

$$l_f = \text{sign}\left(\sum_{i=1}^m \alpha_i y_i K(\mathbf{z}_i, \mathbf{z}) + b\right) \quad (16)$$

To solve a multi-class pattern recognition problem, a multi-class SVM model disintegrates the multi-class problem into several binary problems [35]. Methods like one-versus-one, one-versus-all, and direct acyclic graph have been reported, among which the one-versus-one method is the most effective one with good generalization ability and less training period [37]. Consequently, we adopt one-versus-one method to solve the multi-class SVM classification problem in this study.

For a problem with N classes, $N(N - 1)/2$ binary SVMs are constructed by one-versus-one method, where each SVM is trained based on the data from two classes [35]. For the new input data \mathbf{z} to be classified, a max wins voting strategy is applied [35], in which if $(\text{SVM})_{ij}$ decides \mathbf{z} to be in the i th class, the vote for the i th class is added by one; otherwise, the vote for the j th class is added by one. This process is being conducted for all the $N(N - 1)/2$ binary SVMs. Eventually, the new input data \mathbf{z} is predicted to be in the class with the maximum votes.

Table 2

Physical parameters of the simulated planetary gear set [14].

Parameters	Sun gear	Planet gear	Ring gear
Number of teeth	19	31(4)	81
Module (mm)	3.2	3.2	3.2
Pressure angle	20°	20°	20°
Mass (kg)	0.7	1.822	5.982
Base circle radius (mm)	28.3	46.2	120.8
Reference circle radius (mm)	30.4	49.6	129.6

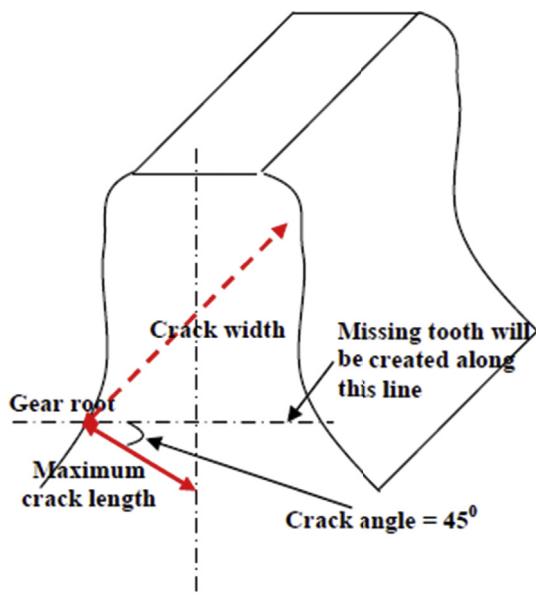


Fig. 4. Tooth crack model [3].

3. Development of the proposed feature vector by simulated vibration signal analysis

In this section, to develop the proposed feature vector, simulated planetary gear set vibrations with different sun gear tooth crack levels are analyzed. With vibration sources generated by the dynamic model in Ref. [3], the simulated vibrations are constructed by the modeling method reported in Ref. [14]. The raw simulated vibration signals are decomposed into IMFs by EEMD. Then the tail dependence between the raw vibration signal and the IMFs is analyzed by the Archimedean copulas. Based on the dependence analysis, the dependence-based feature vector is developed for planetary gearbox fault classification.

3.1. Simulated planetary gear set vibration signal

The simulated planetary gear set has the sun gear and the carrier as the power input and the power output, respectively, with the ring gear fixed [3]. Its physical parameters are shown in Table 2. The number 4 in the parentheses is the number of planet gears. The input rotational speed, i.e. the sun gear rotating speed, is constant at 46.667r/min and the load torque applied on the carrier is 2367Nm [3]. The sun gear tooth crack starts from the gear root circle with a crack angle of 45° along the whole tooth width. When the crack line reaches the tooth central line as demonstrated in Fig. 4, we call it 50% crack. Three sun gear tooth crack levels are considered in this study, namely perfect, 10% crack and 50% crack with crack lengths of 0 mm,

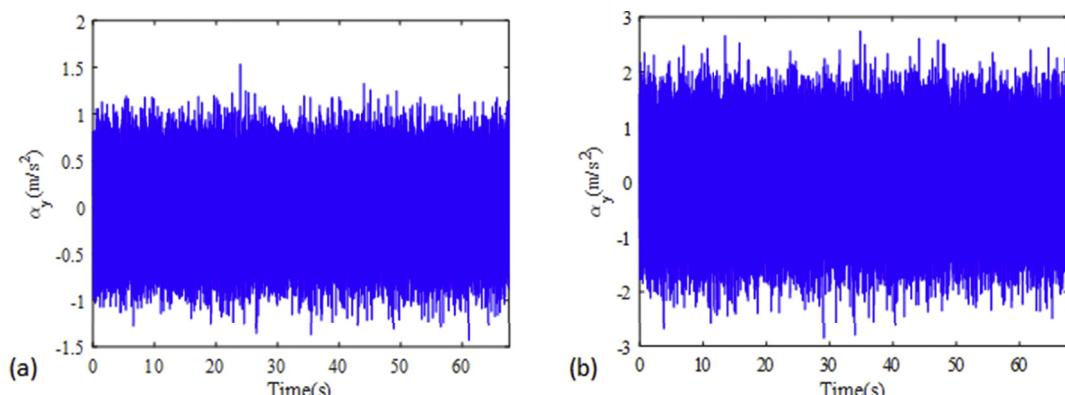


Fig. 5. Waveform of the simulated signal with 50% sun gear tooth crack. (a) SNR10 case; (b) SNR05 case.

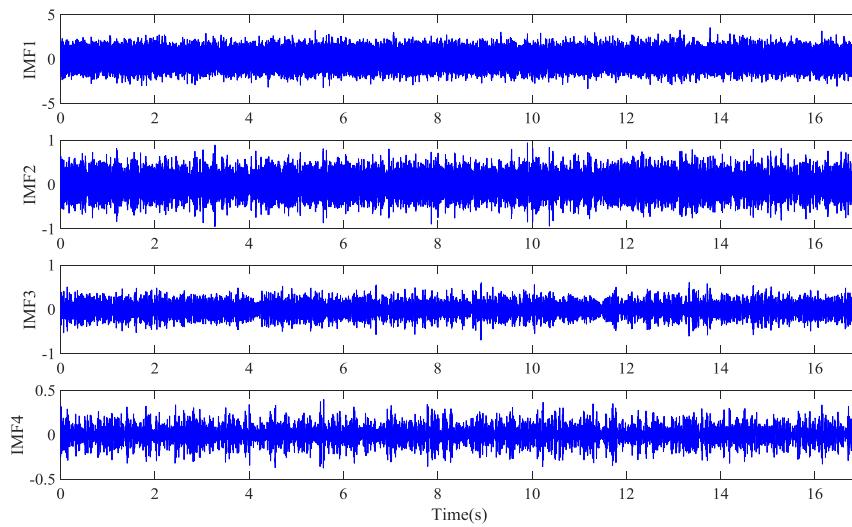


Fig. 6. IMF plots of 50% sun gear tooth crack with SNR05.

0.78 mm and 3.90 mm, respectively [38]. To mimic the background noise, white Gaussian noise is added to the simulated signals. For the robustness analysis of the developed feature vector to the noise interference, two noisy cases are considered with signal-to-noise ratios (SNRs) being 10 dB and 5 dB.

Fig. 5 shows the simulated vibration signals with SNR10 and SNR05 for the 50% sun gear tooth crack case where we omit the other simulated signals to save the article length. As an example to show the EEMD-decomposed IMFs, **Fig. 6** shows the first 4 IMFs of 50% sun gear tooth crack with SNR05. With the raw vibration signal and the IMFs ready, the tail dependence between the raw vibration signal and each IMF is analyzed to develop the dependence-based feature vector.

It is noteworthy that the tail dependence regards the vibration strength but is regardless of the vibration direction. For the waveform of a vibration signal, the positive and negative signs indicate the vibration directions and the amplitudes suggest the vibration strength. Thus, in this study, we apply the signal squared to get the instantaneous vibration energy as the vibration strength for the tail dependence analysis. Alternatively, one also can apply the absolute values for the tail dependence analysis. The rational is that the signal squared and the absolute values are monotone transformation to each other while the tail dependence is invariant with the monotone transformation [39].

3.2. Tail dependence analysis

3.2.1. Copula model selection

With copula candidates as the three Archimedean copulas, namely Frank copula, Clayton copula, and GH copula, we first estimate the copula coefficients for the raw vibration signal and each IMF by Matlab command of “copulafit” with rank pairs associating with the data samples. Then, to exam the fitness of estimated copulas to the tail dependence between the data samples, goodness-of-fit test by QQ plot is conducted and the average distance of the curve regarding the estimated copula to the reference diagonal line is calculated. The copula with the shortest distance among the three copula candidates is selected to describe the tail dependence. Its coefficient is used to develop the targeted feature vector. As an example, **Fig. 7** shows the QQ plots regarding the first 4 IMFs of the raw vibration of 50% sun gear tooth crack case with SNR05. The average distance to the reference diagonal line is presented as well in **Fig. 7** for each copula model. The result shows the GH copula provides the best fitness with the shortest average distance for the crack50 case with SNR05. More comprehensively, **Table 3** gives the average distances of the curves away from the reference diagonal line in the QQ plots for all simulated signals considered in this study.

3.2.2. GH copula coefficient analysis

Table 4 shows the GH copula coefficients for all cases considered in this study. From **Table 4**, following observations can be noted:

- 1) With more severe crack level, it is not necessary to have greater GH copula coefficient for each IMF, neither in SNR10 case nor in SNR05 case.
- 2) With the increase of the crack level, the GH copula coefficients do not have a monotone trend for each IMF, neither in SNR10 case nor in SNR05 case.

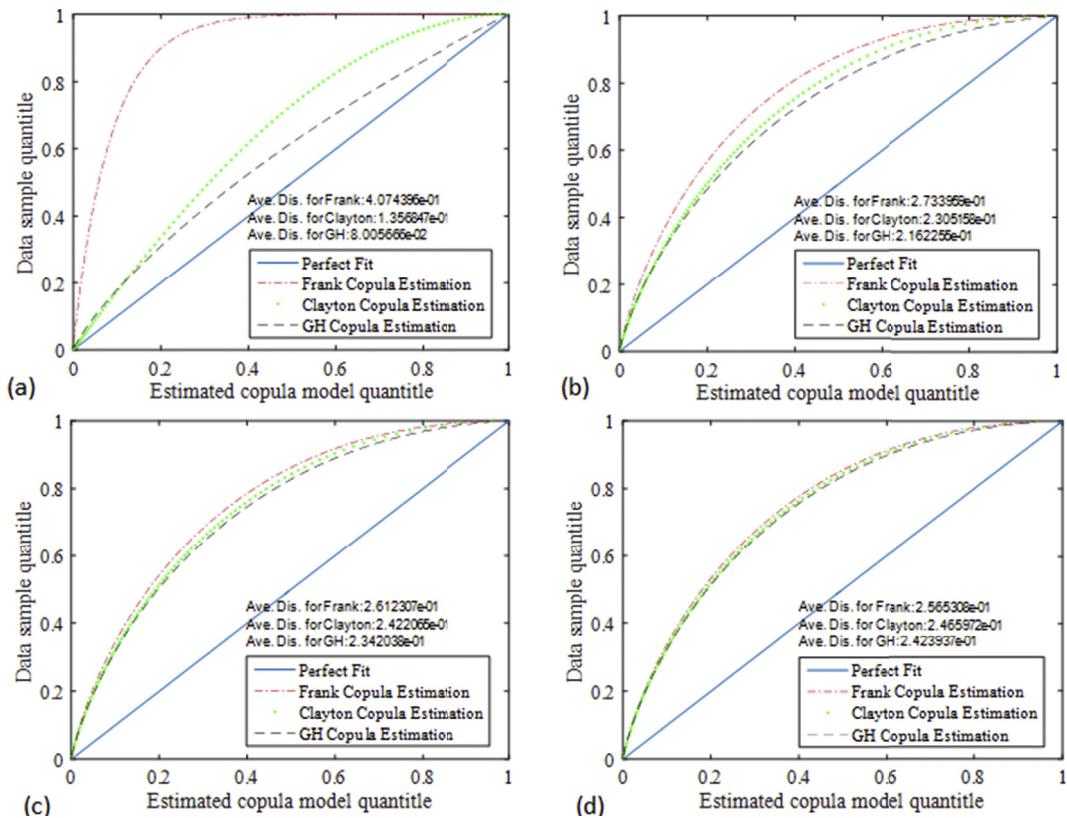


Fig. 7. QQ plot for the estimated Archimedean copula models of Crack50 case with SNR05. (a) QQ plot regarding IMF1; (b) QQ plot regarding IMF2; (c) QQ plot regarding IMF3; (d) QQ plot regarding IMF4.

Table 3

Average distance of estimated copula models to the perfect fit.

Simulated Signal Case	Copula Model	IMF1	IMF2	IMF3	IMF4
Crack00 with SNR05	GH	0.0803	0.2159	0.2350	0.2417
	Frank	0.4074	0.2758	0.2634	0.2581
	Clayton	0.1357	0.2295	0.2416	0.2462
Crack00 with SNR10	GH	0.0798	0.2174	0.2348	0.2418
	Frank	0.4089	0.2746	0.2605	0.2565
	Clayton	0.1353	0.2303	0.2420	0.2464
Crack10 with SNR05	GH	0.0799	0.2172	0.2349	0.2425
	Frank	0.4084	0.2742	0.2610	0.2579
	Clayton	0.1354	0.2300	0.2422	0.2454
Crack10 with SNR10	GH	0.0791	0.2172	0.2358	0.2419
	Frank	0.4079	0.2742	0.2593	0.2553
	Clayton	0.1353	0.2305	0.2436	0.2465
Crack50 with SNR05	GH	0.0801	0.2162	0.2342	0.2424
	Frank	0.4074	0.2734	0.2612	0.2565
	Clayton	0.1357	0.2305	0.2422	0.2466
Crack50 with SNR10	GH	0.0792	0.2191	0.2349	0.2421
	Frank	0.4073	0.2733	0.2628	0.2566
	Clayton	0.1356	0.2313	0.2416	0.2465

3) One earlier IMF always has a greater GH copula coefficient than the later ones for each crack level in both SNR10 case and SNR05 case.

Based on above observations, it can be concluded that by using single GH copula coefficient one cannot achieve the fault detection or the fault degradation level diagnosis.

On the other hand, the combination of two features has been proved to be more effective in fault classification [17]. Inspired by this idea, in this study, we use the combination of the GH copula coefficients to achieve planetary gearbox fault

Table 4

GH copula coefficient for each IMF.

	Crack Level	IMF1	IMF2	IMF3	IMF4
SNR10 Case	Crack00	2.9332	1.1273	1.0540	1.0301
	Crack10	2.9492	1.1273	1.0515	1.0285
	Crack50	2.9464	1.1188	1.0581	1.0297
SNR05 Case	Crack00	2.9134	1.1341	1.0589	1.0326
	Crack10	2.9371	1.1279	1.0572	1.0307
	Crack50	2.9202	1.1309	1.0586	1.0296

classification. Moreover, as shown in Fig. 7 and Table 3, from IMF1 to IMF4, the fitness of the estimated GH copulas to the data samples decreases. Following the principle of choosing copula estimations with fitness as good as possible, the estimated GH copulas regarding the first two IMFs are selected for coefficient combination. Thus, in the following, we pair the GH copula coefficients regarding the first two IMFs and its performance in fault classification is checked, aiming at the development of the dependence-based feature vector.

Moreover, to demonstrate the influence of different copula models with different goodness-of-fit as shown in Fig. 7 and Table 3, the performance of combinations of Clayton copula coefficients and Frank copula coefficients regarding the first two IMFs are presented as well. Correspondingly, results of SNR10 case are shown in Fig. 8 and results of SNR05 case are shown in Fig. 9.

From Figs. 8 and 9, the following observations regarding the categorization by different Archimedean copulas can be noticed for different gear faults:

- 1) Pairs of Frank copula coefficients cannot provide proper categorization for each crack level in neither SNR10 case nor SNR05 case as shown in Fig. 8(a) and Fig. 9(a).
- 2) Pairs of Clayton copula coefficients can separate each crack level well in SNR10 case as shown in Fig. 8(b). However, when the noise level increases to SNR05, Clayton copula coefficient pairs cannot separate crack10 and crack50 properly as shown in Fig. 9(b).

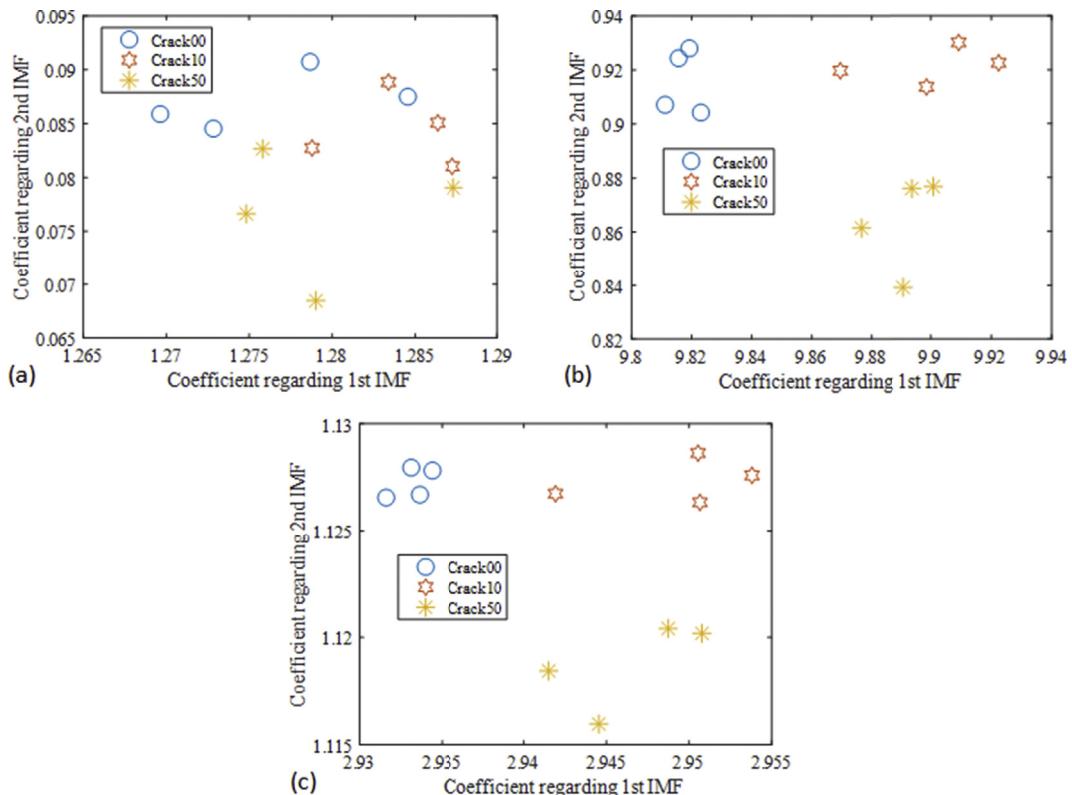


Fig. 8. Coefficient pair by different Archimedean copulas for SNR10 case. (a) Coefficient pair by Frank copula; (b) coefficient pair by Clayton copula; (c) coefficient pair by GH copula.

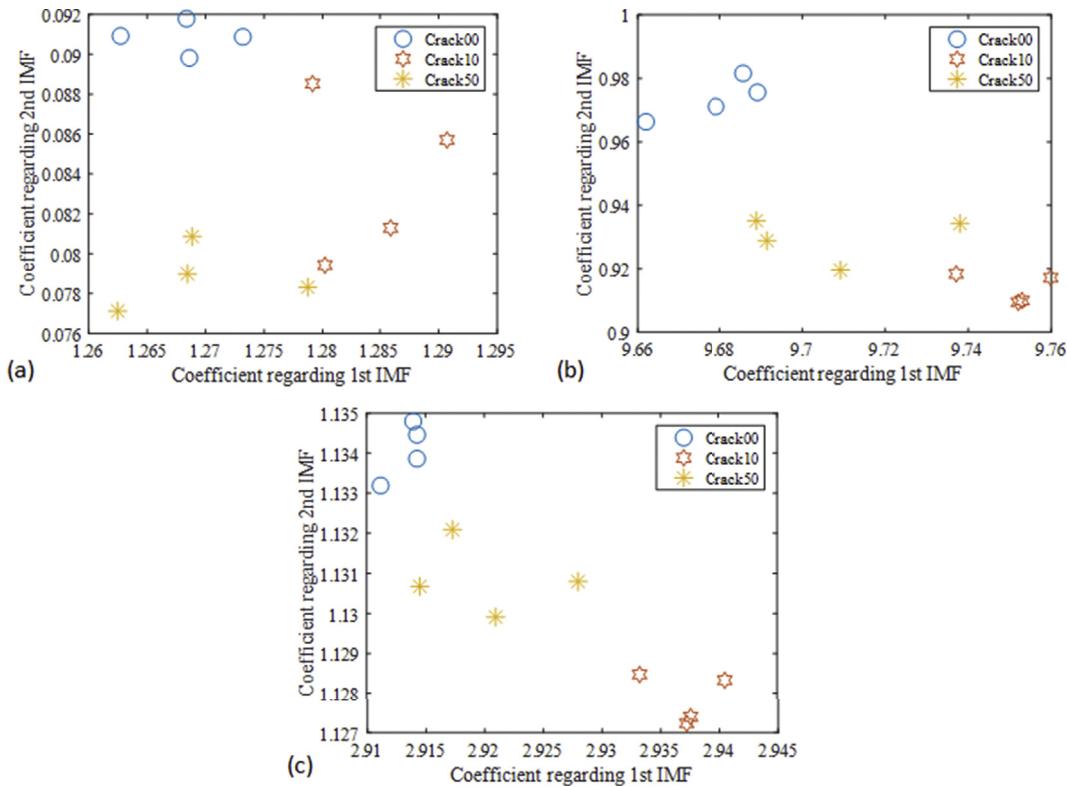


Fig. 9. Coefficient pair by different Archimedean copulas for SNR05 case. (a) Coefficient pair by Frank copula; (b) coefficient pair by Clayton copula; (c) coefficient pair by GH copula.

- 3) Pairs of GH copula coefficients can separate each crack level with proper categorization for both SNR10 case and SNR05 case as shown in Fig. 8(c) and Fig. 9(c).

Based on above observations, the conclusion can be drawn that with the best fitness to tail dependence among the three different Archimedean copulas, GH copula can provide coefficient pairs for fault classification in both SNR10 case and SNR05 case with great robustness to noise.

3.3. Definition of the proposed feature vector

Through above analysis on the tail dependence between the raw vibration signal and the EEMD-decomposed IMFs, it is found that there is an upper tail dependence which can be described by GH copula. Then how the GH copula coefficient can be used in fault detection and fault diagnosis through categorization is analyzed and investigated. It is found that the GH copula coefficient regarding a single IMF cannot achieve fault detection or fault diagnosis because there is no a monotone trend with the increase of the sun gear tooth crack level. On the other hand, when pairs of the GH copula coefficients regarding the first two IMFs are applied, they can be categorized properly with great robustness to noise interference for simulated vibration signals with different sun gear tooth crack levels, i.e. no crack, 10% crack, and 50% crack, indicating its great potential in fault classification for a real planetary gearbox by measured vibration signals. Consequently, we define our proposed dependence-based feature vector as the pair of the GH copula coefficients regarding the first two IMFs. The dependence-based feature vector FV is obtained and expressed as

$$\left\{ \begin{array}{l} FV = (\theta_{GH1}(R_x, S_{IMF_1}), \theta_{GH2}(R_x, S_{IMF_2})) \\ R_x(i) = \frac{x(i)}{n+1} \\ S_{IMF}(i) = \frac{IMF(i)}{n+1} \end{array} \right. \quad (17)$$

where θ_{GH1} and θ_{GH2} are the GH copula coefficients regarding the first two IMFs; R_x and S_{IMF} are the ranks associated with the raw vibration signal x and the EEMD-decomposed IMF, respectively.

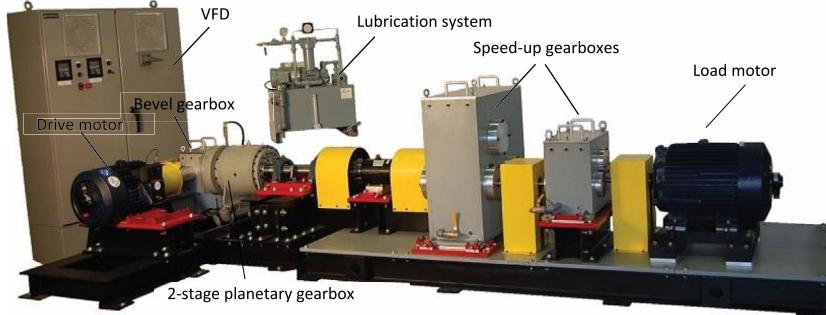


Fig. 10. Configuration of the planetary gearbox test rig [14].

4. Application to experimental planetary gearbox vibration signals

In this section, the developed dependence-based feature vector is applied to the experimental data with gear tooth failures of different levels at different gears to check its ability in fault classification for a real planetary gearbox. Meanwhile, the AACO reported in Ref. [17] and the AACO-based feature vector are applied to the same experimental data for comparison purpose. The accuracy of classifying the gear faults by a multiclass SVM model is compared with different inputs to outstand the advantage of the proposed dependence-based feature vector.

4.1. Experimental setup

The planetary gearbox test rig in Reliability Research Lab (RRL) at the University of Alberta has a configuration as shown in Fig. 10. The main components of the planetary gearbox test rig include a 20HP drive motor, one stage of bevel gearbox, two stages of planetary gearbox, two speed-up gearboxes, and a 40HP load motor. There are three foundations: the drive motor on the first foundation; the bevel gearbox and the planetary gearbox on the second foundation; and the two speed-up gearboxes and the load motor on the third foundation. With this layout, the study target, vibration from the planetary gearbox, can be isolated from interference of vibrations from other components on different foundations.

For the planetary gearbox, all gears are spur gears without tooth profile modification. An accelerometer was installed on the casing of the second stage planetary gearbox vertically to acquire the vibration signal. The second stage planetary gearbox has the same structure configuration and the same gear parameters with the simulated planetary gear set. The input rotational speed to the second stage planetary gearbox and the torque on its carrier are the same with those values in the simulated case, i.e. 46.667r/min and 2367Nm, respectively [14].

In this study, vibration signals with perfect gears and different faulty gears in the second stage planetary gearbox are acquired and analyzed by the developed feature vector for planetary gearbox fault detection and fault diagnosis. The fault diagnosis focuses on distinguishing the fault position and the fault level. The gear faults include the tooth damage on single tooth (namely planet gear tooth crack on ring gear meshing side (PR), planet gear tooth crack on sun gear meshing side (PS), the ring gear tooth crack (RC), and the sun gear tooth crack (SC)), the tooth damage on multiple teeth (namely slight planet gear tooth pitting (SP), moderate planet gear tooth pitting (MP), and critical planet gear tooth pitting (CP)), and the tooth breakage (namely planet gear tooth breakage (PB), ring gear tooth breakage (RB), and sun gear tooth breakage (SB)). Detailed description to these gear damages can be found in Refs. [40] and [41]. The experimental vibration signals are with a time length of 300 s. As waveform examples of experimental vibrations, Fig. 11 shows the waveforms of the experimental data with perfect gears, RC, SC, PB, MP, and CP.

4.2. Result of the dependence-based feature vector and the reported AACO

As illustrated in above discussion with the simulated vibration signal analysis, the performance of the proposed dependence-based feature vector in fault diagnosis is intuitively reflected by the categorization of feature vectors. To achieve the categorization, multiple vibration signals are needed as one vibration signal can only get one feature vector. For this reason, each experimental vibration signal is divided into ten segments with equal time length of 30 s. Note that the lowest frequency component of interest in the 2nd stage planetary gearbox is 0.1478 Hz, the carrier rotating frequency [14], corresponding to a time period of 6.76 s. With the segment time length as 30 s, it is guaranteed that the segment covers multiple (4.47) periods of the lowest frequency component. After the segmentation, the EEMD is then applied to decompose each segment into IMFs. As one example, Fig. 12 shows IMF waveforms of the experimental vibration with planet gear tooth breakage. GH copula coefficient is subsequently estimated to represent the upper tail dependence between the raw vibration signal and each IMF. Pairs of GH copula coefficients regarding the first two IMFs are plotted as the feature vector for classification of experimental planetary gearbox vibration signals. The scatter plot is given in Fig. 13.

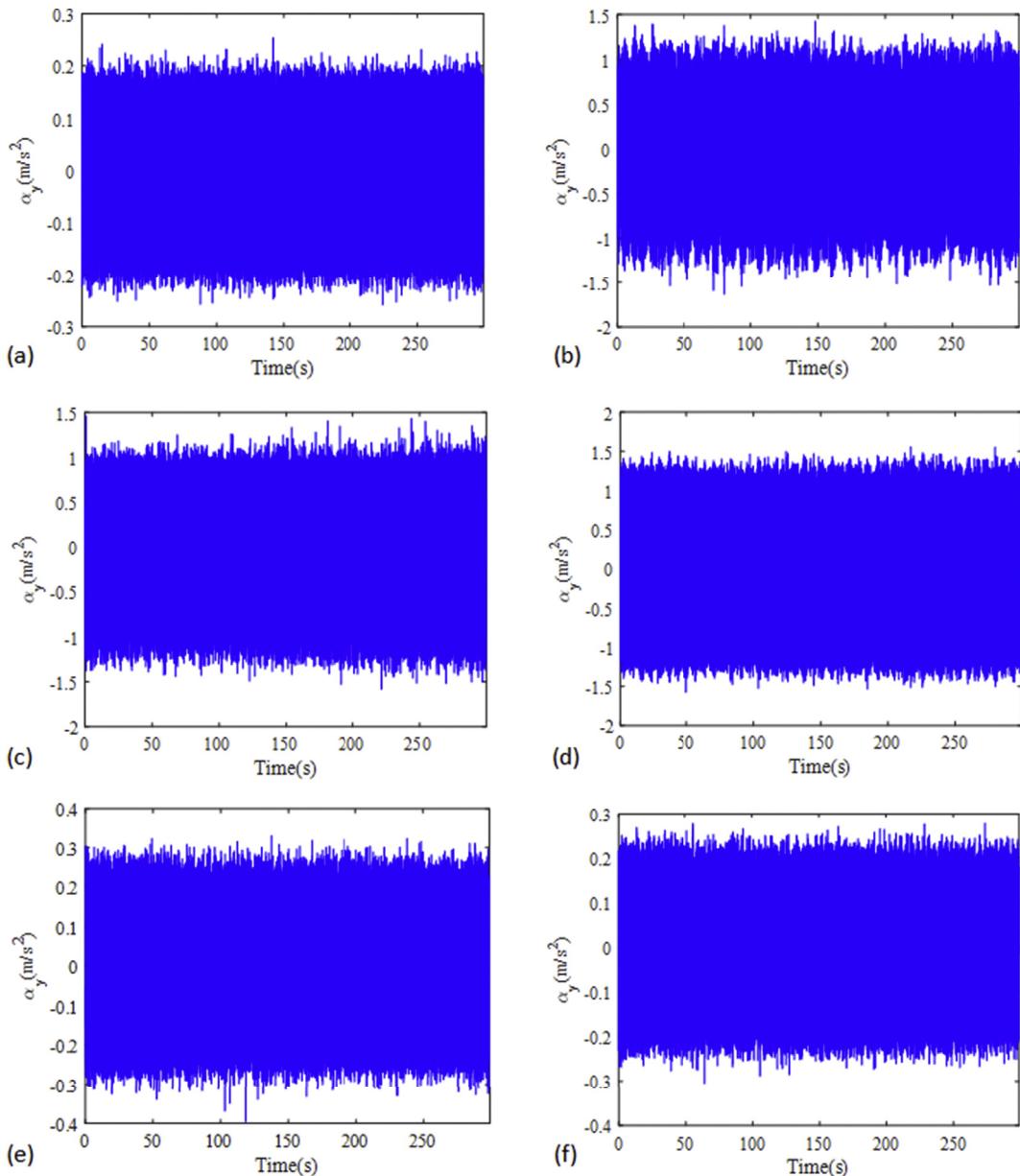


Fig. 11. Experimental vibration waveforms with perfect gears and different gear tooth damages. (a) Waveform with perfect gears; (b) waveform with RC; (c) waveform with SC; (d) waveform with PB; (e) waveform with MP; (f) waveform with CP.

For comparison purpose, we adopt the reported AACO in Ref. [17] as the alternative feature for experimental planetary gearbox fault classification. The AACO is designed especially for the planetary gearbox fault diagnosis based on the order spectrum [17]. The AACO is defined as the summation of the maximum amplitudes around different orders of the carrier rotating frequency in the order spectrum [17]. Different gear faults may result in different amplitude changes at the orders of the carrier rotating frequency. By the AACO, these amplitude changes can be captured to achieve planetary gearbox fault diagnosis. In addition, as we take the combination of GH copula coefficients regarding the first two IMFs as the proposed feature vector, even though Ref. [17] only gives out the individual AACO plot, we also give out the scatter plot of AACO pairs regarding the first two IMFs to be consistent in data processing. The AACO pair regarding the first two IMFs is the AACO-based feature vector. Fig. 14 depicts the plots by AACO including the individual AACO plot of the raw data as given in Fig. 14(a) and the AACO pair plot of the first two IMFs as given in Fig. 14(b).

From Fig. 13, one can notice that the proposed dependence-based feature vector can distinguish the perfect condition with the faulty conditions in the sense of fault detection. In the sense of fault diagnosis, it is observed that the proposed feature vector can categorize the cases of planet gear tooth crack on the ring meshing side, planet gear tooth crack on the sun meshing

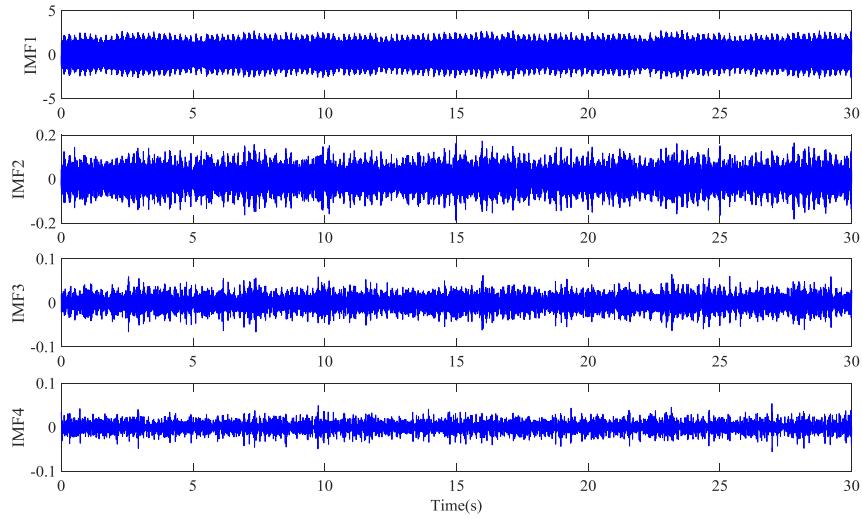


Fig. 12. IMF plots of the experimental vibration with planet gear tooth breakage.

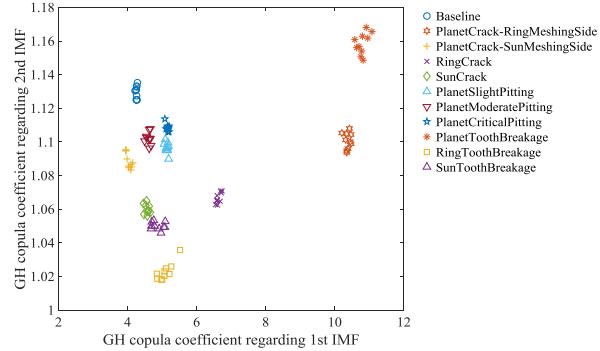


Fig. 13. Scatter plot of dependence-based feature vector for experimental vibration signals.

side, planet gear tooth breakage, ring gear tooth crack, ring gear tooth breakage, planet gear tooth pitting, and the faults (tooth crack and breakage) on the sun gear. Although the coefficient pairs regarding different planet gear tooth pitting levels are very close to each other, there is no overlap for coefficient pairs with different pitting levels. The same observation can be obtained for the coefficient pairs with sun gear tooth crack and sun gear tooth breakage. Overall, by Fig. 13, it can be conclude that the proposed dependence-based feature vector could properly reflect the health status of the system by categorization without overlap.

From Fig. 14, it can be noticed that both the AACO plot and the AACO pair plot can distinguish well the perfect condition with the faulty conditions to achieve fault detection. In the sense of fault diagnosis, by the AACO plot in Fig. 14(a), ring gear tooth crack, ring gear tooth breakage, planet gear tooth crack on ring meshing side, planet gear tooth breakage can be distinguished clearly. However, the lines for sun gear tooth crack, sun gear tooth breakage and planet gear tooth crack on sun meshing side are mixed together; the lines subject to slight planet gear tooth pitting and critical planet gear tooth pitting are mixed together. On the other hand, as shown in Fig. 14(b), the AACO pairs can separate the baseline, planet crack on sun meshing side, ring crack, sun crack, ring tooth breakage and sun tooth breakage. However, the planet gear tooth breakage is mixed with planet gear tooth crack on the ring meshing side, and the slight planet gear tooth pitting, the moderate planet gear tooth pitting, and the critical planet gear tooth pitting are mixed together. Thus, by Fig. 14, it can be found that neither AACO plot nor AACO pair plot is satisfactory for planetary gearbox fault categorization. One reason is that the design of the AACO in Ref. [17] is based on the observation that the rotating frequencies of gears in a planetary gearbox are integer multiples of the carrier rotating frequency, which is not true for the 2nd stage planetary gearbox of the experimental test rig. Given numbers of teeth of the sun gear, the ring gear and the planet gear being 19, 81, and 31, respectively, with 4 planet gears in the gearbox as shown in Table 2, the characteristic frequencies of the 2nd stage planetary gearbox are calculated by the formulas presented in the Introduction and the results are listed in Table 5 where f_s , f_p , f_c , f_{p-p} , and f_m are the sun gear rotating frequency, the planet gear rotating frequency, the carrier rotating frequency, the passing frequency of the planet gears, and the

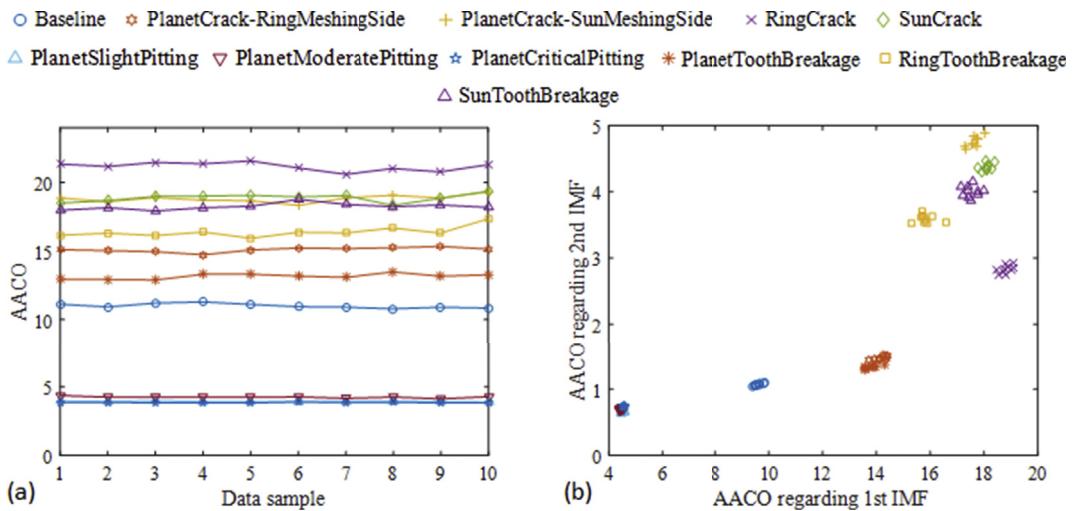


Fig. 14. AACO-based plot for experimental vibration signals. (a) AACO plot; (b) AACO pair plot.

Table 5

Characteristic frequencies of the 2nd stage planetary gearbox.

	f_s	f_p	f_c	f_{p-p}	f_m
Rotating frequency/Hz	0.7778	0.23836	0.1478	0.5913	11.97
Ratio with f_c	5.26	1.61	1	4	81

meshing frequency, respectively. Table 5 shows that f_s and f_p are not the integer multiples of f_c , which results the AACO-based features cannot work well in fault categorization for the 2nd stage planetary gearbox.

4.3. Performance comparison

In Subsection 4.2, the features shown in Figs. 13 and 14 are analyzed subjectively with the observation on the categorization for different gear faults. In this subsection, the accuracy of classifying the gear faults with features in Figs. 13 and 14 is compared objectively which is achieved by a multi-class SVM model. For a SVM-based classification problem, its accuracy is affected by the input which is the employed features. The more useful information the input has the higher classification accuracy the SVM can obtain. Therefore, to check the performance of the proposed dependence-based feature vector in fault classification, the developed feature vector and the reported AACO as well as the AACO-based feature vector are input into a multiclass SVM with same properties. Then classification accuracies by the SVM with different inputs are compared.

As an SVM classifier is a supervised machine learning method, conditions with different gear faults need to be labeled. In this study, we label the cases of baseline, planet gear tooth crack on ring meshing side, planet gear tooth crack on sun gear meshing side, ring gear tooth crack, sun gear tooth crack, slight planet gear tooth pitting, moderate planet gear tooth pitting, critical planet gear tooth pitting, planet gear tooth breakage, ring gear tooth breakage, and sun gear tooth breakage as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, respectively.

With the features as the input and the labels as the output, we use Matlab function of "fitcecoc" to fit the multiclass SVM model by setting the learner of "fitcecoc" as SVM. We then use the cross-validation method by Matlab function of "crossval" to determine how well the estimated SVM model generalizes. Five-fold cross-validation is employed. By the five-fold cross-validation, five models are obtained by the training of in-fold observations which contain 4/5 of data, i.e. 8 segments for each case in this study. The rest 1/5 excluded data, i.e. 2 segments for each case in this study, works as the test fold for prediction. Specifically, the first model is trained with the first 1/5 of data excluded; the second model is trained with the second 1/5 of data excluded, and so on. For the prediction, responses to the excluded data are computed by the model trained with that data excluded, i.e. the first model computes predictions for the first 1/5 of data; the second model computes the prediction for the second 1/5 of data, and so on.

With above setup, confusion matrices of the multiclass SVM with different inputs can be obtained and are given in Fig. 15. The confusion matrices show predictions of the 1/5 excluded data by the corresponding trained model. In the confusion matrix, each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. One benefit of confusion matrix is that it is easy to see if the system is confusing two classes, i.e. mislabeling one as another. In the last row and the last column, the rates of correct predictions and incorrect predictions regarding each column and row are given to show the prediction accuracy and the classification accuracy for each labeled case. Note that the classification accuracy is defined as the rate of correct predictions in an actual class and the prediction accuracy is defined as

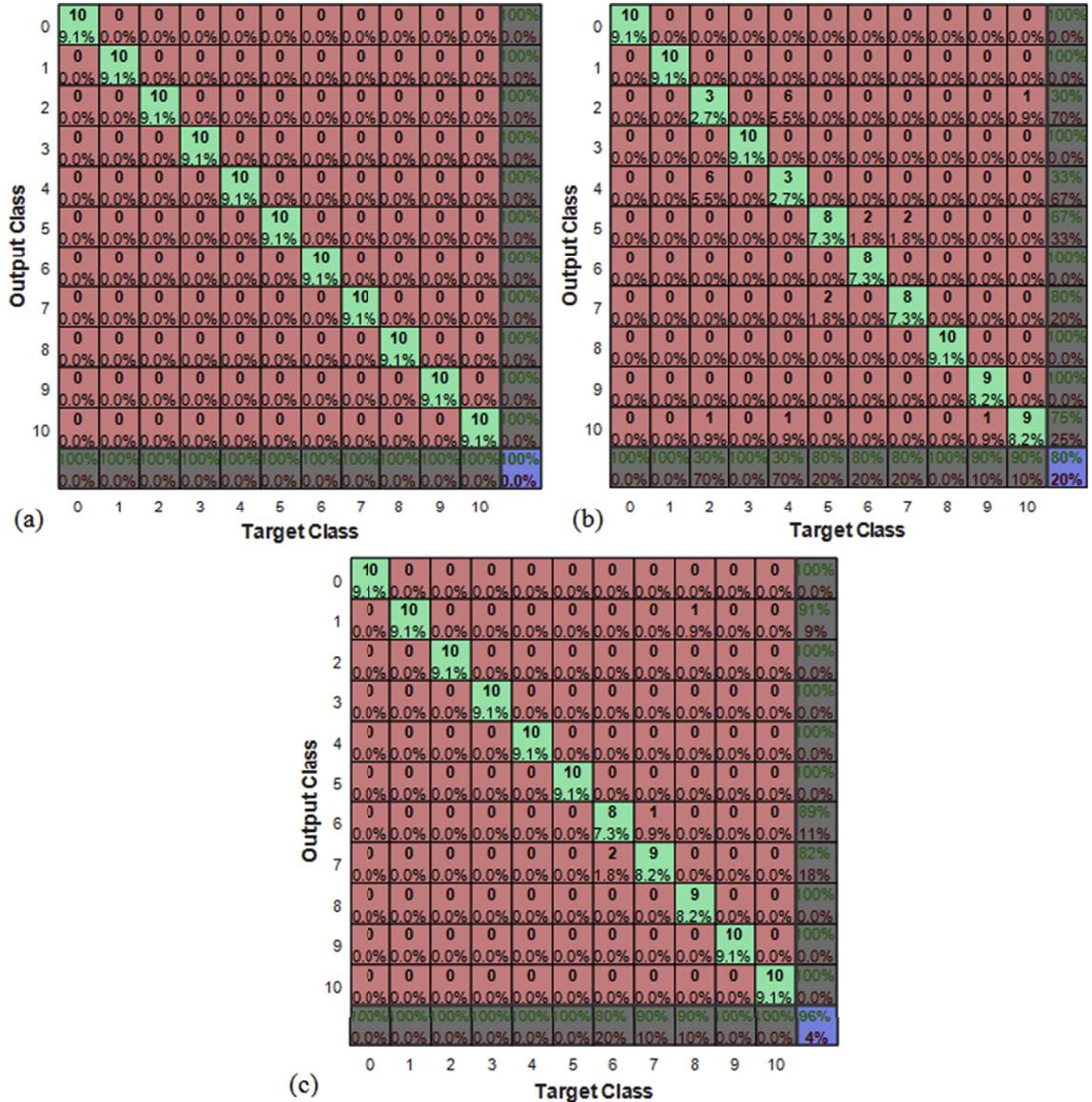


Fig. 15. Confusion matrices with different inputs. (a) Confusion matrix with proposed feature vector; (b) confusion matrix with AACO; (c) confusion matrix with AACO pairs.

the rate of correct predictions in a predicted class. The overall accuracy regarding all cases is given in the last cell at the bottom right corner of the matrix.

By the confusion matrices shown in Fig. 15, the proposed dependence-based feature vector can provide high classification accuracy and prediction accuracy as 100% for each gear fault case; the AACO reported in Ref. [17] provides lower classification accuracies as 30% for labeled 2, 33% for labeled 4, 67% for labeled 5, 80% for labeled 7, and 75% for labeled 10 and lower prediction accuracy as 30% for labeled 2 and labeled 4, 80% for labeled 5, labeled 6, and labeled 7, and 90% for labeled 9 and labeled 10; the AACO pairs provides lower classification accuracy as 91%, 89%, and 82% for labeled 1, labeled 6, and labeled 7, respectively, and lower prediction accuracy as 80% for labeled 6, and 90% for labeled 8 and labeled 9. The overall accuracies of the SVM models with input as the proposed dependence-based feature vector, AACO values and AACO pairs are 100%, 80%, and 96%, respectively, as given in Fig. 15.

Consequently, we can draw the conclusion that when the input is the proposed dependence-based feature vector, it can lead to better results in terms of the classification accuracy, the prediction accuracy and the overall accuracy than the other two AACO-based features. This conclusion suggests that the proposed dependence-based feature vector, i.e. the pair of GH copula coefficients regarding the first two IMFs, can better extract and reflect the health status information about the gearbox from the vibration signal, serving better as a condition indicator for planetary gearbox fault classification.

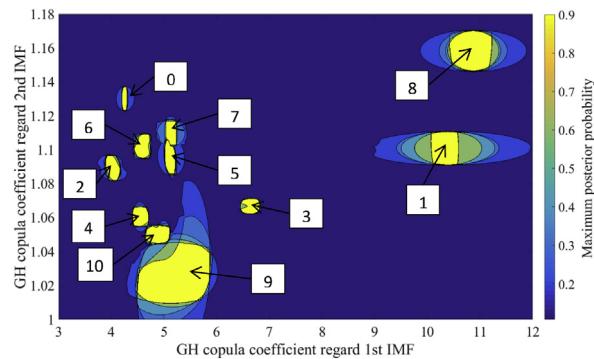


Fig. 16. Maximum posterior probability plot.

5. Discussion

To check the performance of the proposed feature vector in the fault classification by a multi-class SVM, besides the confusion matrix shown in Fig. 15, another way is the maximum posterior probability plot. By setting the “FitPosterior” in Matlab command of “fitcecoc” being true, the binary-leaner classification score can be transformed to posterior probability which can then be extracted by Matlab command of “resubPredict”. Note that the SVM score for classifying the observation is the distance from the observation to the decision boundary. The posterior probability in a SVM model can be calculated based on the Bayesian theorem where the involved prior probabilities are computed from the training data [42]. Defining a grid of values in the predictor space, the posterior probability regarding each class can be calculated for each coordinate on the grid. The maximum posterior probability is plotted in Fig. 16 for the experimental vibration signals with a grid size of 2500×2500 . The numbers on Fig. 16 are the labels of the gear tooth faults in the experimental planetary gearbox as defined in Subsection 4.3. From Fig. 16, it can be found that the decision boundaries for the gear tooth faults are separated clearly. The well-separated decision boundaries in Fig. 16 validate the conclusion drawn based on the subjective categorization observation in Subsection 4.2.

It should be noted that the foundation of the proposed dependence-based feature vector is the correlation/dependence between the extreme values in the EEMD-decomposed IMFs and the extreme values in the raw vibration signal. It is noteworthy that the extreme values in the IMFs can be introduced by both the fault-induced impulses and the noise in the raw vibration signal. In this study, we constrain the raw vibration signal with the same noise level of the same SNR. In this way, we set the extreme values in the IMF introduced by the noise consistent. Thus, the change of the extreme values in the IMF is mainly caused by the occurrence of the fault-induced impulses. Consequently, for the possible real application, the proposed dependence-based feature vector aims at the situation with the operation condition of the same noise level. In the cases the system noise level is changing with the structural dynamic noise and the environmental noise, the acquired vibration signals would be with different noise levels. For these cases, the pre-process of de-noising is needed. The de-noising is to guarantee the noise levels of the vibration signals are similar before the extraction of the dependence-based feature vector for its effectiveness.

6. Conclusion

In this paper, a dependence-based feature vector is developed to detect and diagnose the faults in a planetary gearbox by fault classification. The dependence-based feature vector is developed on the strength of the tail dependence between the raw vibration signal and the EEMD-decomposed IMFs. By the goodness-of-fit test with QQ plot, it is found that the tail dependence can be best described by Gumbel-Hougaard (GH) copula with an upper tail dependence amongst the three Archimedean copula candidates, namely Frank copula, Clayton copula, and GH copula. Accordingly, the GH copula coefficient, representing the upper tail dependence level, is adopted for the development of the targeted dependence-based feature vector. Eventually, the dependence-based feature vector is defined as the pair of the GH copula coefficients regarding the first two IMFs.

To evaluate the performance of the developed dependence-based feature vector, test on experimental planetary gearbox vibration signals with different gear tooth faults are conducted. The gear tooth faults are of different levels at different gears. The developed dependence-based feature vector, reported AACO, and AACO-based feature vector are input into a multi-class SVM with same properties. Then the classification accuracies by the SVM models with different inputs are compared. The results show that the developed dependence-based feature vector leads to a higher classification accuracy, suggesting that the dependence-based feature vector can serve better for planetary gearbox fault classification with more health status information about the planetary gearbox.

For this study, the planetary gearbox of interest is with a single gear tooth fault. The dependence-based feature vector is developed with stationary operation of constant load and constant speed under the same noise level. The further analysis on the non-stationary operation condition, under different noise levels, and with multiple gear tooth faults will be studied in the future work. In those complicated cases, copula coefficients regarding more IMFs may be involved to extend the two-element feature vector to a higher dimensional feature vector. Besides, reported features such as AACO and conventional indicators as kurtosis, skewness, crest indicator, etc. can be added as complementary to form a composite feature vector.

Acknowledgement

This research is supported by the Natural Sciences and Engineering Research Council of Canada, Canada [Grant No. RGPIN-2015-04897], China Scholarship Council, China [Grant No. 201306050030], Alberta Innovates - Technology Futures (AITF) and Alberta Innovation & Advanced Education, Canada. Comments and suggestions from reviewers and the Editor are very much appreciated.

References

- [1] Y. Lei, J. Lin, M.J. Zuo, Z. He, Condition monitoring and fault diagnosis of planetary gearboxes: a review, *Measurement* 48 (2014) 292–305, <https://doi.org/10.1016/j.measurement.2013.11.012>.
- [2] Z. Feng, M.J. Zuo, Vibration signal models for fault diagnosis of planetary gearboxes, *J. Sound Vib.* 331 (2012) 4919–4939, <https://doi.org/10.1016/j.jsv.2012.05.039>.
- [3] X. Liang, M.J. Zuo, M.R. Hoseini, Vibration signal modeling of a planetary gear set for tooth crack detection, *Eng. Fail. Anal.* 48 (2015) 185–200, <https://doi.org/10.1016/j.engfailanal.2014.11.015>.
- [4] W.D. Mark, H. Lee, R. Patrick, J.D. Coker, A simple frequency-domain algorithm for early detection of damaged gear teeth, *Mech. Syst. Signal Process.* 24 (2010) 2807–2823, <https://doi.org/10.1016/j.ymssp.2010.04.004>.
- [5] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, J. Wang, Machine health monitoring using local feature-based gated recurrent unit networks, *IEEE Trans. Ind. Electron.* 65 (2018) 1539–1548, <https://doi.org/10.1109/TIE.2017.2733438>.
- [6] A. Rai, S.H. Upadhyay, Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering, *Mech. Syst. Signal Process.* 93 (2017) 16–29, <https://doi.org/10.1016/j.ymssp.2017.02.003>.
- [7] J. Giles, Empirical wavelet transform, *IEEE Trans. Signal Process.* 61 (2013) 3999–4010, <https://doi.org/10.1109/TSP.2013.2265222>.
- [8] J. Chen, J. Pan, Z. Li, Y. Zi, X. Chen, Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals, *Renew. Energy* 89 (2016) 80–92, <https://doi.org/10.1016/j.renene.2015.12.010>.
- [9] C. Hu, W.A. Smith, R.B. Randall, Z. Peng, Development of a gear vibration indicator and its application in gear wear monitoring, *Mech. Syst. Signal Process.* 76–77 (2016) 319–336, <https://doi.org/10.1016/j.ymssp.2016.01.018>.
- [10] Y. Zhang, B. Tang, Y. Han, L. Deng, Bearing performance degradation assessment based on time-frequency code features and SOM network, *Meas. Sci. Technol.* 28 (2017) 045601, <https://doi.org/10.1088/1361-6501/aa56c9>.
- [11] L. Hong, J.S. Dhupia, A time domain approach to diagnose gearbox fault based on measured vibration signals, *J. Sound Vib.* 333 (2014) 2164–2180, <https://doi.org/10.1016/j.jsv.2013.11.033>.
- [12] Z. Li, J. Chen, Y. Zi, J. Pan, Independence-oriented VMD to identify fault feature for wheel set bearing fault diagnosis of high speed locomotive, *Mech. Syst. Signal Process.* 85 (2017) 512–529, <https://doi.org/10.1016/j.ymssp.2016.08.042>.
- [13] J. Qu, Support-vector-machine-based Diagnostics and Prognostics for Rotating Systems, University of Alberta, 2011, <https://doi.org/10.7939/R3QQ10>.
- [14] L. Liu, X. Liang, M.J. Zuo, Vibration signal modeling of a planetary gear set with transmission path effect analysis, *Measurement* 85 (2016) 20–31, <https://doi.org/10.1016/j.measurement.2016.02.006>.
- [15] Y. Zhang, H. Zuo, F. Bai, Classification of fault location and performance degradation of a roller bearing, *Measurement* 46 (2013) 1178–1189, <https://doi.org/10.1016/j.measurement.2012.11.025>.
- [16] X. Liang, M.J. Zuo, L. Liu, A windowing and mapping strategy for gear tooth fault detection of a planetary gearbox, *Mech. Syst. Signal Process.* 80 (2016) 445–459, <https://doi.org/10.1016/j.ymssp.2016.04.034>.
- [17] Y. Lei, N. Li, J. Lin, Z. He, Two new features for condition monitoring and fault diagnosis of planetary gearboxes, *J. Vib. Contr.* 21 (2015) 755–764, <https://doi.org/10.1177/1077546313486284>.
- [18] Y. Miao, M. Zhao, J. Lin, Y. Lei, Application of an improved maximum correlated kurtosis deconvolution method for fault diagnosis of rolling element bearings, *Mech. Syst. Signal Process.* 92 (2017) 173–195, <https://doi.org/10.1016/j.ymssp.2017.01.033>.
- [19] P.D. McFadden, Window functions for the calculation of the time domain averages of the vibration of the individual planet gears and sun gear in an epicyclic gearbox, *J. Vib. Acoust.* 116 (1994) 179, <https://doi.org/10.1115/1.2930410>.
- [20] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, *Adv. Adapt. Data Anal.* 01 (2009) 1–41, <https://doi.org/10.1142/S1793536909000047>.
- [21] L.G.N. Martins, S.D. Miller, O.C. Acevedo, Using empirical mode decomposition to filter out non-turbulent contributions to air–sea fluxes, *Boundary-Layer Meteorol.* 163 (2017) 123–141, <https://doi.org/10.1007/s10546-016-0215-0>.
- [22] M. Zagórski, M. Orkisz, Dependence structure analysis of the inverter operational data by means of copula theory, *Mech. Syst. Signal Process.* 40 (2013) 439–451, <https://doi.org/10.1016/j.ymssp.2013.06.016>.
- [23] J.C. Rodriguez, Measuring financial contagion: a Copula approach, *J. Empir. Finance* 14 (2007) 401–423, <https://doi.org/10.1016/j.jempfin.2006.07.002>.
- [24] C. Genest, A.-C. Favre, Everything you always wanted to know about copula modeling but were afraid to ask, *J. Hydrol. Eng.* 12 (2007) 347–368, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:4\(347\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)).
- [25] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, F. Fnaiech, Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, *Appl. Acoust.* 89 (2015) 16–27, <https://doi.org/10.1016/j.apacoust.2014.08.016>.
- [26] M. Boumhadi, J.-P. Dron, S. Rechak, O. Cousinard, On the extraction of rules in the identification of bearing defects in rotating machinery using decision tree, *Expert Syst. Appl.* 37 (2010) 5887–5894, <https://doi.org/10.1016/j.eswa.2010.02.017>.
- [27] M. Cerrada, R.-V. Sánchez, C. Li, F. Pacheco, D. Cabrera, J. Valente de Oliveira, R.E. Vásquez, A review on data-driven fault severity assessment in rolling bearings, *Mech. Syst. Signal Process.* 99 (2018) 169–196, <https://doi.org/10.1016/j.ymssp.2017.06.012>.
- [28] Y. Lei, Z. He, Y. Zi, Application of an intelligent classification method to mechanical fault diagnosis, *Expert Syst. Appl.* 36 (2009) 9941–9948, <https://doi.org/10.1016/j.eswa.2009.01.065>.
- [29] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. A Math. Phys. Eng. Sci.* 454 (1998) 903–995, <https://doi.org/10.1098/rspa.1998.0193>.
- [30] Z. Feng, M. Liang, Y. Zhang, S. Hou, Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation, *Renew. Energy* 47 (2012) 112–126, <https://doi.org/10.1016/j.renene.2012.04.019>.

- [31] Y. Lei, J. Lin, Z. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, *Mech. Syst. Signal Process.* 35 (2013) 108–126, <https://doi.org/10.1016/j.ymssp.2012.09.015>.
- [32] M. Sklar, Fonctions de repartition an dimensions et leurs marges, *Publ. Inst. Stat. Univ. Paris* 8 (1959) 229–231. <https://ci.nii.ac.jp/naid/10011938360/> (accessed April 30, 2018).
- [33] H. Yoshida, S. Kikkawa, H. Nakajima, N. Kira, Positive time-frequency distributions: a least square Approach and A Copula-based approach, *Mem. Sch. B.O.S.T. Kinki Univ.* 13 (2004) 57–71.
- [34] J. De Kort, *Modeling Tail Dependence Using Copulas-literature Review*, 2007.
- [35] P. Gangsar, R. Tiwari, Comparative investigation of vibration and current monitoring for prediction of mechanical and electrical faults in induction motor based on multiclass-support vector machine algorithms, *Mech. Syst. Signal Process.* 94 (2017) 464–481, <https://doi.org/10.1016/j.ymssp.2017.03.016>.
- [36] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge university press, 2000.
- [37] Chih-Wei Hsu, Chih-Jen Lin, A comparison of methods for multiclass support vector machines, *IEEE Trans. Neural Network.* 13 (2002) 415–425, <https://doi.org/10.1109/72.991427>.
- [38] X. Liang, M.J. Zuo, M. Pandey, Analytically evaluating the influence of crack on the mesh stiffness of a planetary gear set, *Mech. Mach. Theor.* 76 (2014) 20–38, <https://doi.org/10.1016/j.mechmachtheory.2014.02.001>.
- [39] L. Liu, M.J. Zuo, X. Liang, Dependence analysis of planetary gearbox vibration marginals, in: 2016 Progn. Syst. Heal. Manag. Conf., IEEE, 2016, pp. 1–6, <https://doi.org/10.1109/PHM.2016.7819881>.
- [40] M. Pandey, T. Patel, X. Liang, T. Tian, M.J. Zuo, Descriptions of Pitting Experiments, Run-to-failure Experiments, Various Load and Speed Experiments, and Crack Experiments Carried Out on the Planetary Gearbox Test Rig, Edmonton, Alberta, 2011.
- [41] M.R. Hoseini, Y. Lei, D. Tuan, T. Patel, M.J. Zuo, Experiment Design of Four Types of Experiments: Pitting Experiments, Run-to- Failure Experiments, Various Load and Speed Experiments, and Crack Experiments, Edmonton, Alberta, 2011.
- [42] J. Platt, Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, *Adv. Large Margin Classif.* 10 (1999) 61–74 doi:[10.1.1.41.1639](https://doi.org/10.1.1.41.1639).