ARCHIVE OF MECHANICAL ENGINEERING

DOI: 10.24425/ame.2025.153736 2025, Vol. 72, No. 1, pp. 105–130

Yahia BOUSSELOUB [1], Farida MEDJANI [2], Ahmed BENMASSOUD [3], Nadir BENAMIRA [1], Ali BELHAMRA [1]

An optimized EMD method based on cross correlation and root mean square statistical analysis and its application to bearing faults detection

Received 10 June 2024, **Revised** 7 January 2025, **Accepted** 21 January 2025, **Published online** 25 February 2025

Keywords: vibration signal, signal processing, bearing fault detection, empirical mode decomposition (EMD), root mean square (RMS)

The early detection of bearing faults is critical for ensuring the reliability and performance of electromechanical systems. Vibration signals provide valuable insights into fault characteristics. However, effectively extracting fault-related features remains challenging due to issues like over-decomposition and mode mixing in traditional signal processing methods such as Empirical Mode Decomposition (EMD). To address these challenges, this paper proposes an optimized Empirical Mode Decomposition (OEMD) technique enhanced by cross-correlation (CC) and root mean square (RMS) statistical analysis. The proposed method introduces three novel correlation-based stopping criteria to ensure the independence of Intrinsic Mode Functions (IMFs) in the decomposition result. Furthermore, an RMS-based selection strategy is implemented to identify optimal IMFs that retain fault-related information. The proposed approach is validated using real-world vibration signals from two datasets: an experimental bearing vibration dataset and a public dataset from the Case Western Reserve University (CWRU). The results highlight the feasibility and effectiveness of the method in accurately and automatically selecting the optimal Intrinsic Mode Functions (IMFs) containing high-amplitude peaks at defect characteristic frequencies. These findings demonstrate the robustness of the proposed method in both fault detection and identification.

✉ Yahia BOUSSELOUB, e-mail: yahia.bousseloub@univ-annaba.org

[1]Electromechanical Systems Laboratory, Badji Mokhtar – Annaba University 12, P.O.Box, 23000 Annaba, Algeria

[2]Mathematics and their interactions Laboratory, Abdelhafid Boussouf University Center of Mila, Innovation Academy of Mila, Algeria

[3]University of Science and Technology Houari-Boumédiène, Innovation Academy of Mila, Algeria

1. Introduction

Induction machines (IM) are among the most used machines in the industry, with over 60% of the industrial electricity was consumed by IMs. They play a significant role in power generation, manufacturing, transportation, HVAC systems and electric vehicles. Bearings are one of the critical components in IMs, with mechanical failures related to bearing faults accounting for 42% of total mechanical faults [1]. Hence, the early detection and analysis of faults are very important to avoid the total breakdown of IM components. Various condition monitoring techniques have been used to analyze and detect IMs faults, including chemical analysis, thermal monitoring, acoustic noise measurement and vibration monitoring. Due to the effectiveness and high accuracy of vibration analysis techniques and their ability of detecting and distinguishing the most electrical and mechanical faults, it is the most used in the industry [1].

The vibration signals acquired from IMs are non-stationary and nonlinear. Therefore, to extract the useful information from these signals, several methods have been used, such as Empirical Mode Decomposition (EMD) [2], Ensemble Empirical Mode Decomposition (EEMD) [3], and Local Mean Decomposition (LMD) [4]. Qin and Zhong [5], presented a mathematical model of Fourier transform (FT), short time Fourier transform (STFT) and wavelet transform (WT) giving theorical insights and also its application. Nevertheless, the mode mixing is one of the significant limitations of these methods [6]. Albert and Nii [7], used the correlation-coefficient between the original signal and the resulted IMFs for modes selection. However, their proposed approach faced challenges in certain application, especially in the presence of noise. Su et al. [8], introduced new method by combining EMD with correlation-coefficient and spectral kurtosis for bearing fault detection. A method that combines Empirical Mode Decomposition (EMD) with the Hausdorff distance between the probability density functions (PDFs) of the noisy signal and its modes is employed to select the most relevant modes [9]. EMD interval thresholding denoising based on similarity measure to select relevant modes is presented in [10]. Dragomiretskiy and Zosso [11], developed a new method called Variational Mode Decomposition (VMD). Its primary challenge is to predetermine the optimal number of modes (K). Therefore, many researches have been introduced a solution for VMD limitation to enhance its efficacity and accuracy. Yang, Liu and Zhang [12], developed a method which addressed the over- under decomposition problem in VMD decomposition results. Ni et al. [13], introduced a fault information-guided variational mode decomposition (FIVMD), for bearing fault detection. Yahia et al. [14], developed a new method for bearing fault diagnosis (OVMD) based on cross-correlation and root mean square (RMS) algorithms to allows the automatically estimation of the optimal mode number (K) of VMD.

Currently, artificial intelligence (AI) methods are being used for mechanical and electrical faults monitoring (AI- based fault diagnosis), these methods are being integrated with traditional monitoring techniques. Li et al. [15], pro-

posed a novel fault diagnosis method, named VGAIC-FDM, which integrates a variational autoencoder generative adversarial network with an improved convolutional neural network. The results demonstrate that this method enhances fault diagnosis performance, particularly when addressing unbalanced datasets, thereby achieving higher accuracy. Zhao Huimin, Gao Yongshun, and Deng Wu [16], developed a lightweight defect detection model for turbine blades, utilizing Single Shot Multibox Detection (SSD) based on ShuffleNetv2 and Coordinate Attention (SN-CA-SSD) within the context of IoT. The results demonstrate that the algorithm significantly improves accuracy, detection efficiency, and the capability to detect small targets, achieving a balance between precision and efficiency. Furthermore, by integrating the model with an interpretable algorithm, the decision-making process is analyzed through representation visualization, thereby enhancing the interpretability of the algorithm. Huimin et al. [17], proposed a novel parametric time redistribution multisynchronous compression transform method, referred to as GTMSST. The results indicate that GTMSST achieves high time-frequency energy concentration and demonstrates superior performance under varying noise levels and operating conditions. Furthermore, it proves to be effective in diagnosing bearing faults. Ran et al. [18], introduced an innovative automatic K-means clustering algorithm, named HGA-FACO. They validated the superiority of HGA-FACO over conventional K-means clustering (KMeans) and other intelligent clustering approaches, including ACO-KMeans, GA-KMeans (GAK), Particle Swarm Optimization KMeans (PSOK), and ACO-GAK. Huang et al. [19], presented a novel CSO with dynamic multi-competitions and convergence accelerator, namely DM-CACSO. Song et al. [20], proposed a dual-time dual-population multi-objective evolutionary algorithm (DTDP-EAMO) to address the challenges encountered by multi-objective differential evolution algorithms in solving complex optimization problems. Shah et al. [21], introduced a combination of EMD and an artificial neural network (ANN) for bearing fault diagnosis. Shady et al. [22], developed a new approach for three phase induction motor based on EMD and ANN to decompose stator signal and pattern classification, the results show that it can effectively detect outer race fault of bearing. Ali et al. [23], presented an application of EMD and ANN for automatic bearing fault diagnosis without human intervention. Amarouayache et al. [24], developed an intelligent fault diagnosis approach based on EEMD and a convolutional neural network (CNN), the results demonstrate the effectiveness of the deep learning model that identifies and can successfully classify and locate the faulty bearing vibration signals. Lee, Ahn and Koh [25], proposed a fault monitoring method based on EEMD, particle swarm optimization (PSO), principal component analysis (PCA) and Isomap for bearing fault classification. Nishat Toma, Cheol-Hong and Jong-Myon [26], introduced a method for multiple faults classification based on EEMD and CNN.

To overcome the limitations of the Empirical Mode Decomposition (EMD) method, this paper introduces an optimized EMD (OEMD) approach for bearing fault detection. The proposed method integrates two complementary algorithms.

The first algorithm utilizes cross-correlation statistical analysis to enhance the EMD process by incorporating three novel conditions. These conditions ensure the independence of all Intrinsic Mode Function (IMF) components within their respective center frequencies, effectively mitigating the issue of mode mixing.

The second algorithm employs root mean square (RMS) histogram analysis to identify IMFs containing relevant information for vibration signal analysis, particularly those associated with fault characteristic frequencies. By estimating the RMS of each mode and determining a corresponding threshold, the algorithm automatically selects IMFs with RMS values above the threshold.

Experimental studies validate the effectiveness of the proposed approach for bearing vibration analysis. The method was applied to analyze bearing defects using two datasets: one collected experimentally and the other sourced from the Case Western Reserve University (CWRU) Bearing Data Center [27]. In the first dataset, the vibration signal clearly exhibits pronounced peaks at the characteristic frequency of the inner race fault (f_i = 134.27 Hz) and its harmonics ($2f_i, 4f_i$). Similarly, in the second dataset, the inner race fault is effectively identified at (f_i = 159.48 Hz), along with its harmonics ($2f_i$, $4f_i$, $6f_i$, $8f_i$). The results demonstrate that the proposed approach successfully eliminates low-energy IMFs and addresses the limitations of the EMD method, enabling accurate detection and identification of bearing faults.

This study distinguishes itself from OVMD [14] in several significant ways. Specifically, it implements cross-correlation and root mean square (RMS) algorithms within the theoretical frameworks of two distinct methodologies: Empirical Mode Decomposition (EMD) and Variational Mode Decomposition (VMD). Moreover, in real-time analysis, the Optimized Empirical Mode Decomposition (OEMD) method demonstrates significantly greater efficiency, achieving a 9-to-10-fold improvement in optimal mode extraction speed compared to OVMD. Additionally, OEMD requires less memory than OVMD, further enhancing its practicality for real-world applications.

This article is organized as follows. In Section 2, the main motivation and contribution is presented. The proposed methodology is presented Section 3. Then its experimental validation is presented in Section 4. In Section 5, comparative analysis with other decomposition methods is presented. Finally, the conclusion is presented in Section 6.

2. Motivation and contribution

Recent studies have increasingly focused on diagnosing bearing faults using signal decomposition techniques [14–20, 25, 26]. However, the precise selection of optimal modes within the decomposition process provides deeper insights into the nature of the defects. In this work, a novel optimization of the EMD method, based on cross-correlation (CC) and root mean square (RMS) algorithms, is proposed for enhanced bearing fault identification. The effectiveness of the method is validated using an experimental setup for data acquisition through vibratory signals.

3. Methods

3.1. Empirical mode decomposition method

The EMD technique is based on the premise that every time series is composed of distinct and elementary sub-modes. The fundamental principle of the approach is to identify these IMFs in the data, by empirical observation of their distinctive time intervals, and then decompose the data appropriately. By using the technique of sifting, it is possible to remove the majority of riding waves, which are oscillations that do not have any zero crossings between their highest and lowest points (monotonic signals). The EMD algorithm examines the oscillations of the signals at a highly localized level and decomposes the input signal into sun-components that do not overlap with each other in a local context. Fig. 1 present the flowchart of EMD.

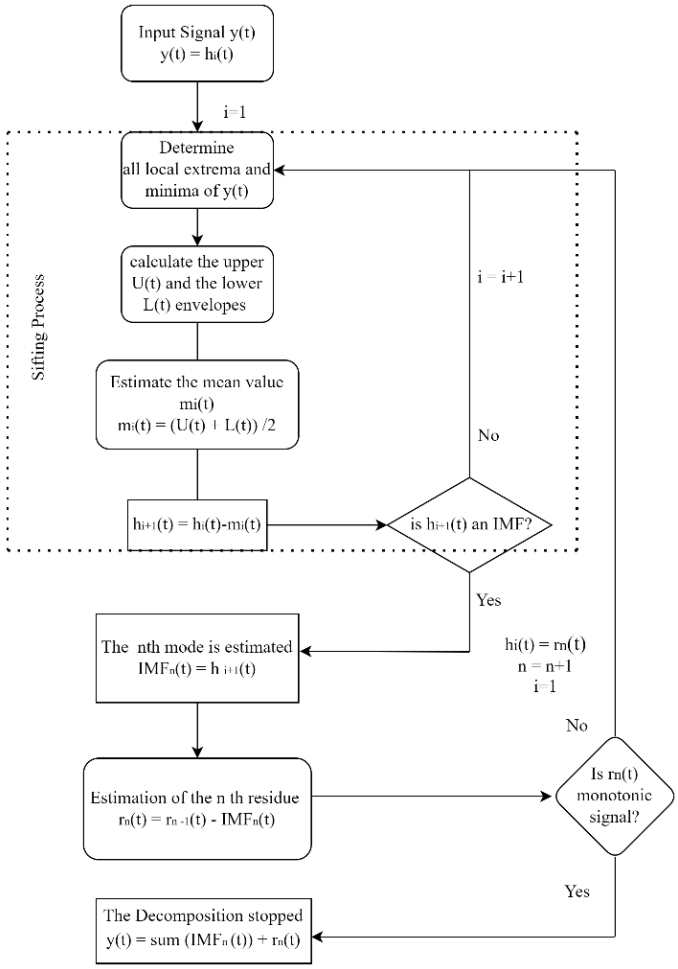

Fig. 1. EMD diagram

The signal $y(t)$ is decomposed into its constituent Intrinsic Mode Functions (IMFs) that adhere to two certain conditions [2]:

- The IMF exhibits a single extremum between two consecutive zero crossings, meaning that the difference between the number of local minima and maxima is at most one.
- The IMF has a mean value equal zero, which is defined by local minima and maxima.

For a given signal $y(t)$, the algorithm of EMD is presented as follows [3]:

Step 0: Determine all local extrema and minima of $y(t)$.

Step 1: Apply cubic spline interpolation to connect all local maxima and minima, thereby obtaining the upper envelope $U(t)$ and the lower envelope $L(t)$, respectively.

Step 2: Calculate the mean of the envelopes:

$$m_i(t) = (U(t) + L(t))/2. \tag{1}$$

Step 3: Subtract $m_i(t)$ from the $y(t)$:

$$h_i(t) = y(t) - m_i(t). \tag{2}$$

Step 4: If: $h_i(t)$ verify the two certain conditions mentioned above, $h_i(t) = \mathrm{IMF}_1(t)$ considered as the first IMF.
Else: $y(t) = h_i(t)$, repeat the steps 0–4.

Step 5: Separating $\mathrm{IMF}_1(t)$ from $y(t)$, we get the following:

$$r_1(t) = y(t) - \mathrm{IMF}_1(t). \tag{3}$$

From this step, $r_1(t)$ will be the new input signal and repeating the sifting process n times and the result is given as follows:

$$r_n(t) = r_{n-1}(t) - \mathrm{IMF}_n(t). \tag{4}$$

Step 6: When $r_n(t)$ is a monotonic function, the decomposition is stopped.
The original signal $y(t)$ is the sum of the decomposed modes and the residue $r_n(t)$.

$$y(t) = \sum_{i=1}^{n} \mathrm{IMF}_i + r_n(t). \tag{5}$$

3.1.1. EMD application on simulated signal

The over-decomposition in signal processing techniques refers to resulting IMFs contain unnecessary or redundant information, increasing noise and computational complexity.

In this section, EMD method has been used to decompose a simulated signal $y(t)$ which is presented in Fig. 2, it is constructed with 3 harmonics $f_1 = 2$ Hz, $f_2 = 24$ Hz and $f_3 = 288$ Hz.

$$y(t) = \frac{1}{4}\cos(2\pi f_1 t) + \frac{1}{3}\cos(2\pi f_2 t) + \frac{1}{2}\cos(2\pi f_3 t). \tag{6}$$

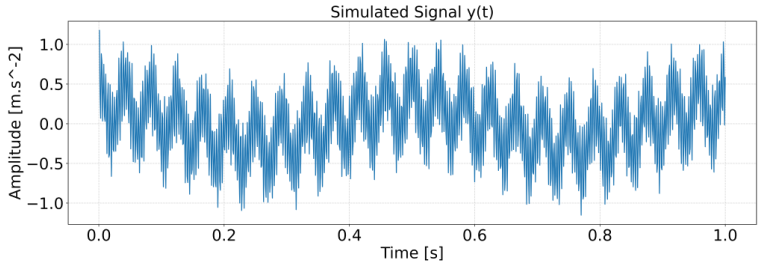

Fig. 2. The simulated signal $y(t)$

Fig. 3 and Fig. 4, present the decomposition results of the simulated signal $y(t)$ using EMD. It can be observed that IMF_1, IMF_3 and IMF_4 have the same center frequencies as $y(t)$, with $f_1 = 288$ Hz, $f_3 = 24$ Hz and $f_4 = 2$ Hz, respectively. In contrast, IMF_2 and IMF_5 can be considered as noise, as they have smaller amplitudes compared to the other IMFs and exhibit new center frequencies.

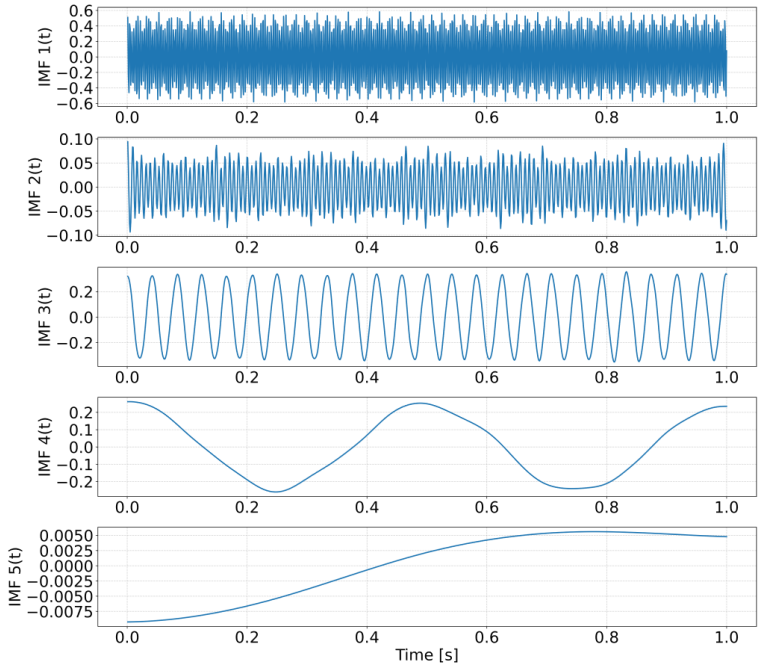

Fig. 3. EMD decomposition result for $y(t)$ in time domain

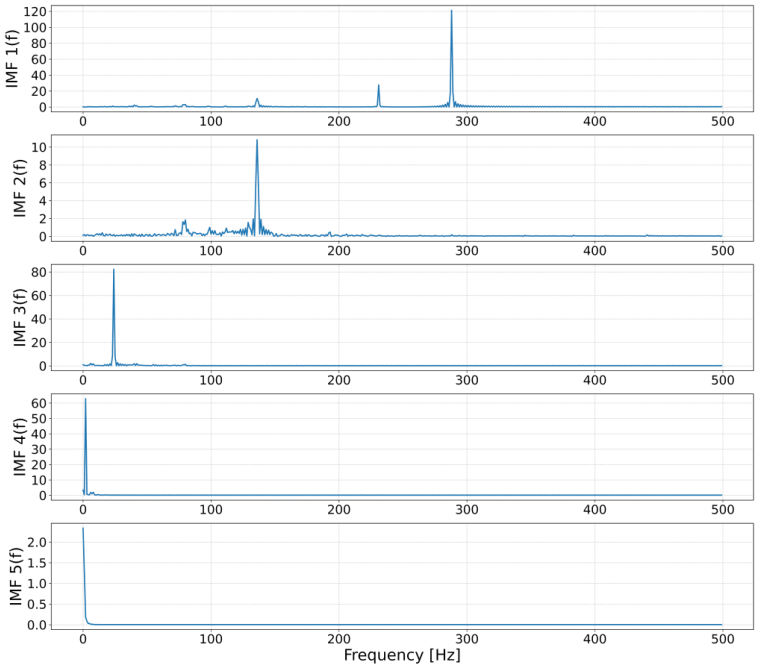

Fig. 4. EMD decomposition result for $y(t)$ in frequency domain

Therefore, to extract the exact component of $y(t)$, IMF_2 and IMF_5 can be eliminated. The over decomposition problem in the decomposition results occurs due to the inappropriate stopping criteria in EMD algorithm and its sensitivity to noise.

In certain applications, where the signal is no-stationary and no-linear in nature and in the presence of noise, EMD face a significant challenge, especial in real-time vibration signal analysis. Therefore, it is important to optimize the EMD algorithm to enable automatic decomposition and optimal selection of the useful modes for any input signal. To this end, we have developed a novel approach called the optimized empirical mode decomposition (OEMD) to address this issue.

3.2. The proposed method

Enhancing the Empirical Mode Decomposition (EMD) algorithm through the integration of novel constraints is essential. The proposed approach comprises two distinct algorithms. The first algorithm is grounded in cross-correlation conditions, which serve as new stopping criteria for the EMD algorithm to ensure that the Intrinsic Mode Functions (IMFs) remain independent within their respective center frequencies. To automatically select the optimal IMFs for any given input signal and address the limitations of EMD in the decomposition results, three cross-correlation-based conditions are introduced.

The first condition prevents over-decomposition of the residual signal and stops the decomposition when the residue no longer contains significant information. The second and third conditions are specifically designed to address the issue of mode mixing.

These conditions are applied as follows [14]:

- The cross-correlation between the original signal and the residual signal must be less than 0.09, as determined through experimental evaluation.
- The cross-correlation between IMF_i and its subsequent mode, IMF_{i+1}, must be below a threshold value of 0.1, as evaluated experimentally.
- Finally, the correlation of each IMF_i with the original signal must be greater than the correlation of IMF_i with its subsequent IMF_{i+1}.

The implementation steps of this algorithm follow next.

Algorithm 1: EMD optimization based on CC

(1) Initialize k = 2;
(2) Do: K = k + 1;
(3) Initialize $U(t)$, $L(t)$;
(4) **for** k = 1: K,
if $CC_1 > 0.1$ stop decomposition;
elif $CC_1 > CC_2$ stop decomposition;
Update $m_i(t)$ and $h_i(t)$:

$$m_i(t) = (U(t) + L(t))/2,$$
$$h_i(t) = y(t) - m_i(t),$$
$$r_i(t) = r_{i-1}(t) - \mathrm{IMF}_i(t),$$
$$CC_1 = \textit{cross correlation}\,(\mathrm{IMF}_i,\ \mathrm{IMF}_{i+1})\,,$$
$$CC_2 = \textit{cross correlation}\,(\mathrm{IMF}_i,\ y)\,.$$

(5) **repeat** steps (2)–(8), until

$$CC_3 = \textit{cross correlation}\,(r_i(t),\ y(t))\,,$$
$$CC_3 < \sigma, \text{ with } \sigma = 0.09.$$

The objective of the second algorithm is to automate the identification of Intrinsic Mode Functions (IMFs) that contain valuable information for vibration signal analysis. This process selects the optimal IMFs without human intervention. By estimating the Root Mean Square (RMS) of each mode in the frequency domain and calculating the corresponding threshold, the algorithm enables the selection of IMFs with RMS values exceeding the threshold. Consequently, the second algorithm serves as a useful tool for comparing the energy content of different IMFs, facilitating the elimination of those that do not contribute useful information for bearing fault detection. The process is summarized as follows [14]:

1. The first algorithm is used to decompose the input signal it to a series of $\mathrm{IMF}s(t)$.

2. Fast Fourier Transform (FFT) is applied to convert the modes from time domain to frequency domain.

$$\mathrm{IMF}(f) = \sum_{n=0}^{N-1} \mathrm{IMF}(n) e^{-j2\pi f n/N}. \tag{7}$$

Here, $\mathrm{IMF}(f)$ represents the mode in frequency domain, f is the frequency index, n is the time index and j is the imaginary unit.

3. The RMS is calculated for each mode ($\mathrm{IMF}(f)$).

$$\mathrm{IMF}(f)_{\mathrm{RMS}} = \sqrt{\frac{1}{n}\sum_{i=1}^{i=m} [\mathrm{IMF}(i)]^2}, \tag{8}$$

where $\mathrm{IMF}(i)$ is the amount of vibration signals at the sampling frequency i, and m is the total number of sampling frequency used.

4. The expected RMS percentage for each mode is estimated .
5. The threshold (ε) is determined.
6. Modes with RMS values above the threshold are selected.
7. The signal $y(f)$ is reconstructed by summing the selected modes($\mathrm{IMF}(f)$).

$$y(f) = \sum_{i=1}^{i=K} \mathrm{IMF}_i(f). \tag{9}$$

The reconstructed spectrum $y(f)$ allows for the identification of high-amplitude peaks at characteristic fault frequencies, which correspond to defect types. The peak corresponds to the maximum amplitude in the signal and can be expressed as:

$$\mathrm{Peak} = \max |y(f)|. \tag{10}$$

The algorithm can be presented as follows:

Algorithm 2: Mode's selection based on RMS

---

n: the length of modes.
K: The number of modes given by the first algorithm
(1) **for** i = 1: K,

$$\mathrm{IMF}_{\mathrm{rms}_i} = \sqrt{\frac{\mathrm{IMF}_i[0]^2 + \mathrm{IMF}_i[1]^2 + \mathrm{IMF}_i[2]^2 + \ldots + \mathrm{IMF}_i[n]^2}{n}}. \tag{11}$$

(2) Find the largest value among *rms* values;

$$\mathrm{rms}_{\max} = \max \left|\mathrm{IMF}_{\mathrm{rms}_i}\right|. \tag{12}$$

(3) Calculate the expected percentage of IMFs;

$$\mathrm{IMF}_{\mathrm{rms}_i}(\%) = \frac{\mathrm{IMF}_{\mathrm{rms}_i} \times 100}{\mathrm{rms}_{\max}}. \tag{13}$$

(4) Calculate the average;

$$\alpha = \sum_{i=1}^{K} \frac{\mathrm{IMF}_{\mathrm{rms}_i}(\%)}{K}. \tag{14}$$

(5) Calculate the threshold;

$$\varepsilon = \alpha \times 0.8. \tag{15}$$

(6) **if** $\mathrm{IMF}_{\mathrm{rms}_i}(\%) > \varepsilon$, IMF_i is a selected mode;
else, drop off all the remained IMFs.

The flowchart of the proposed approach is presented in Fig. 5.

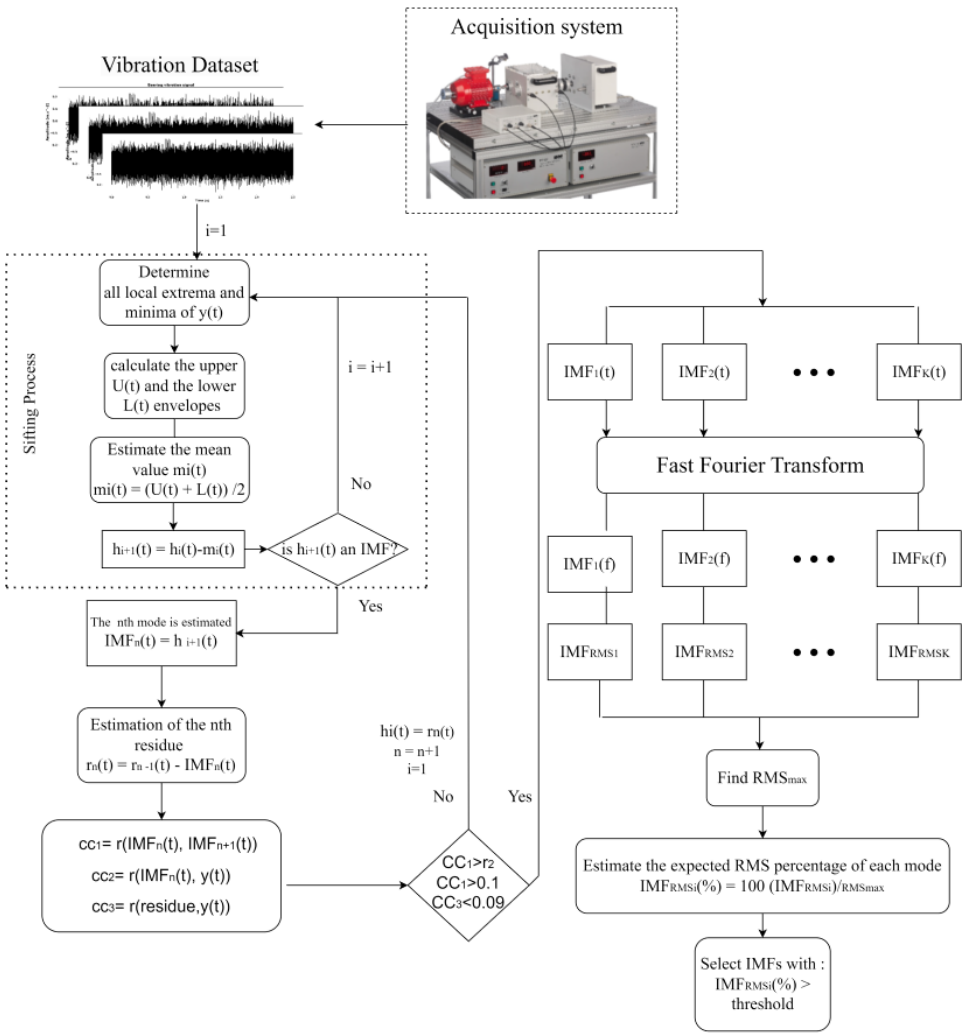

Fig. 5. OEMD flowchart

3.2.1. OEMD application on simulated signal

To verify the validity of the proposed approach, we have analyzed the same simulated signal $y(t)$ with OEMD and compared with EMD decomposition results. As presented in Eq. (6), the signal contains three sub-components (3 harmonics). Table 1 present the calculated RMS of each IMFs using the proposed method.

Table 1. The estimated RMS for each mode using OEMD algorithm

Modes	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5
RMS [m/s^2]	1.119	0.403	0.550	0.584	0.129
RMS (%)	100	36.056	49.195	52.202	11.595

Based on the RMS histogram in Fig. 6, the energies of IMF_3, IMF_4 and IMF_5 are above the energy threshold level of 39.84%, indicating that the modes number 2 and 5 should be eliminated. Indeed, form Fig. 7, it can be observed that OEMD extracts the exact harmonics in the simulated signal ($f_1 = 2$ Hz, $f_2 = 24$ Hz, $f_3 = 288$ Hz).

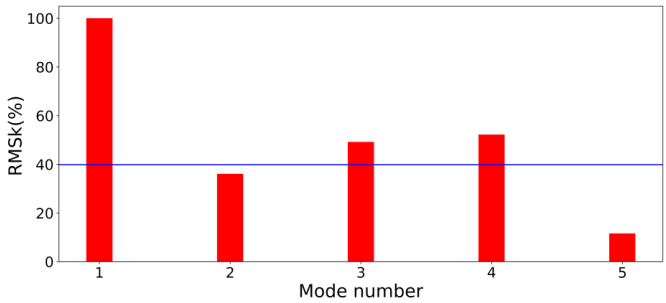

Fig. 6. The RMS histogram of each mode with threshold $\varepsilon = 39.84\%$

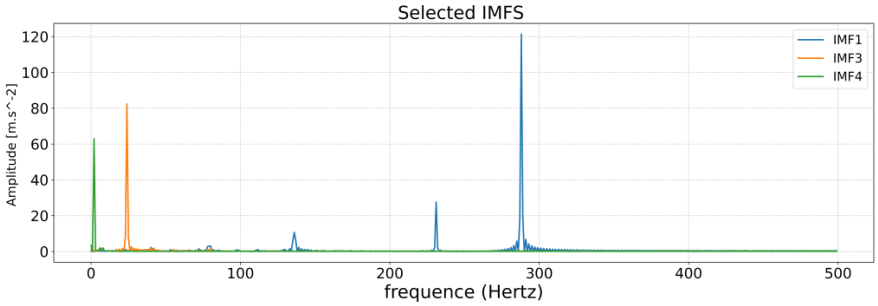

Fig. 7. The selected modes using OEMD for the simulated signal $y(t)$

4. Experimental study

The proposed approach was applied to analyze vibration signal defects from two datasets. The first dataset, obtained experimentally from the Electromechanical System Laboratory (LSEM), involved measurements under varying conditions using the setup shown in Fig. 8 for bearing vibration signal acquisition. The second dataset is sourced from the Case Western Reserve University (CWRU) Bearing Data Center [27], which serves as a widely recognized reference for testing new diagnostic algorithms.

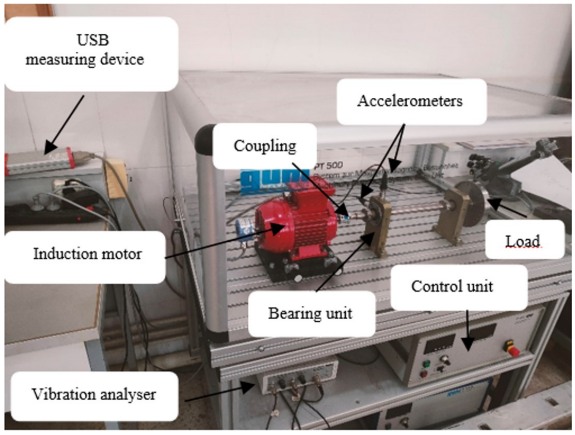

Fig. 8. The experimental test strip

4.1. Bearing vibration signal acquisition

The test rig for vibration signal acquisition is presented in Fig. 8. The vibration signals are generated by a three-phase induction motor (0.37 kW) through an elastic clow coupling, a PC, an accelerometer, a balanced flywheel (load), a bearing unit contain bearing type 6004-2RSH SKF, USB measuring device, and a control unit. The faulty vibration signal (inner race fault) is measured at 8 kHz, for two rotational speeds of 25 and 33 rps.

The characteristic frequencies of bearing faults can be represented as:

- Inner race fault frequency

$$f_{ir} = \frac{n \times f_r}{2}\left(1 + \frac{d}{D}\cos\varphi\right). \tag{16}$$

- Outer race fault frequency

$$f_{or} = \frac{n \times f_r}{2}\left(1 - \frac{d}{D}\cos\varphi\right). \tag{17}$$

- Roling element fault frequency

$$f_{re} = \frac{D \times f_r}{2d}\left(1 - \left[\frac{d}{D}\cos\varphi\right]^2\right), \tag{18}$$

where f_r is the shaft speed, n is the number of rolling elements, φ is the angle of the load from the radial plane, D is bearing pitch diameter and d is the rolling element diameter. The characteristics bearing fault frequencies values are established in Table 2.

Table 2. Characteristics bearing fault frequencies

Rotational speed	25 rps	33 rps
f_{ir}	134.27 Hz	179.01 Hz
f_{or}	90.73 Hz	120.96 Hz
f_{re}	62.16 Hz	82.88 Hz

4.1.1. Results and discussions

The measured vibration signal of inner race fault is presented in Fig. 9, and its decomposition using algorithm 1 (EMD optimization based on CC) is presented in Figs. 10 and 11.

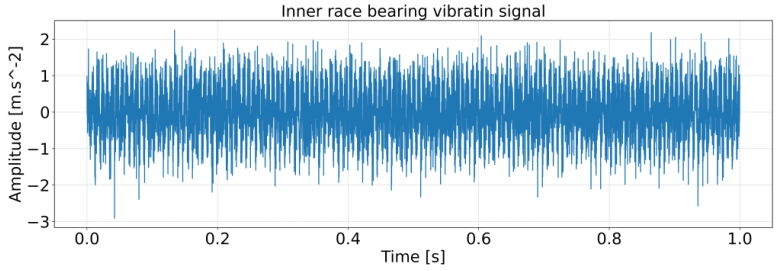

Fig. 9. The measured inner race vibration signal

From Fig. 10 and Fig. 11 the occurrence of over-decomposition is clearly observable. Notably, IMF_7, IMF_8, and IMF_9 display substantially lower amplitude levels in comparison to the other modes. Since these modes lack meaningful information, they are identified as noise and should be discarded. As a result, it is concluded that only the first six modes are relevant for analysis. This conclusion is further validated by the second component of the proposed approach (the RMS-based algorithm), which automatically selects the useful IMFs, as demonstrated in Fig. 12.

Based on Table 3 and the RMS histogram in Fig. 12, it can be concluded that the last three modes should be eliminated, as their RMS percentages fall below the threshold of $\varepsilon = 35.90\%$.

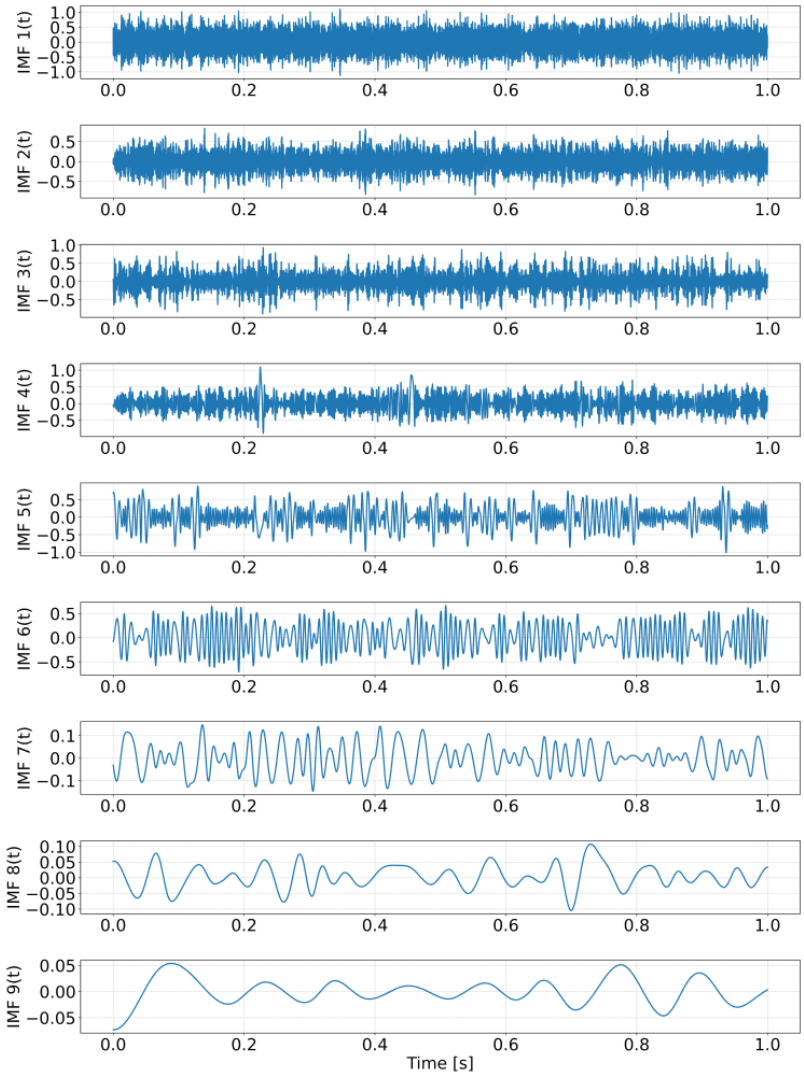

Fig. 10. The decomposition results of the first algorithm in time domain IMF(t)

Table 3. The estimated RMS for each mode using OEMD algorithm

Modes	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7	IMF_8	IMF_9
RMS [m/s^2]	3.808	2.703	2.337	1.965	1.804	1.447	0.625	0.393	0.297
RMS (%)	100	71.00	61.37	51.61	47.37	38.00	16.43	10.33	7.80

The reconstructed spectrum in Fig. 13 reveals a prominent peak at a frequency of 136 Hz, which closely aligns with the characteristic frequency of the inner race fault (f_i = 134.27 Hz). Additionally, its harmonics ($2f_i$, $4f_i$) are clearly

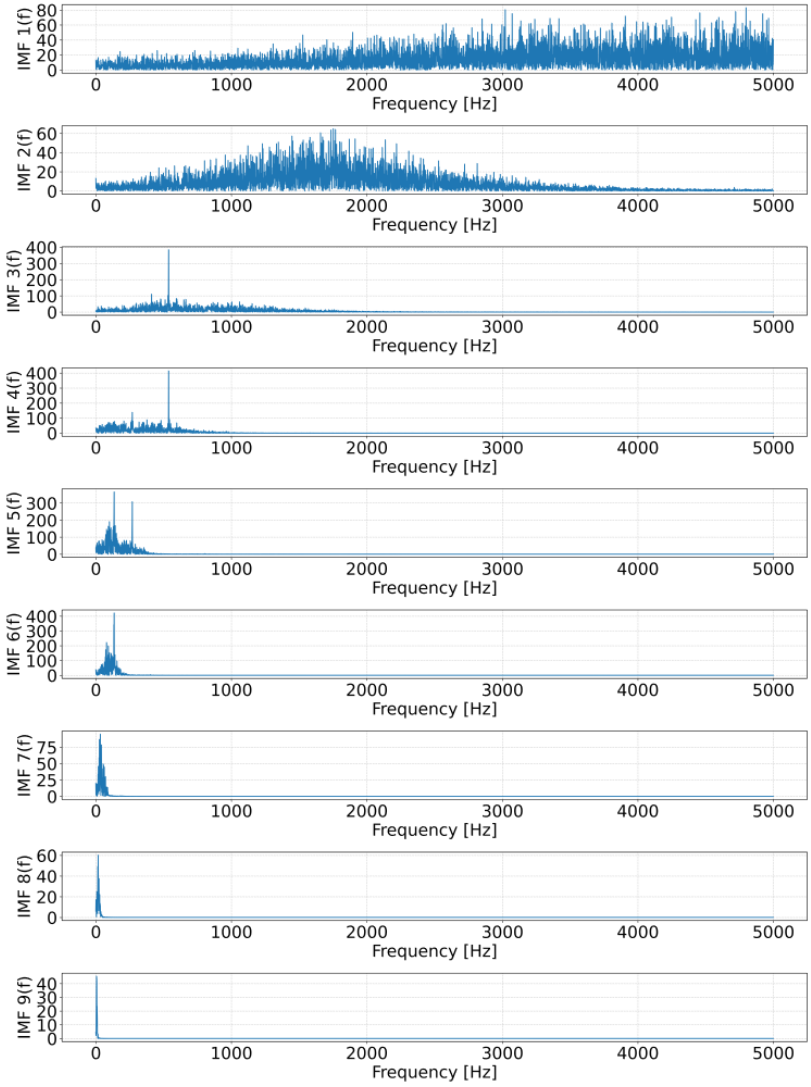

Fig. 11. The decomposition results of the first algorithm in frequency domain IMF(f)

identified. These results confirm the feasibility and effectiveness of the proposed method in accurately identifying inner race faults. The automatic selection of modes containing valuable fault-related information for bearing inner race fault detection is thus demonstrated.

To validate the effectiveness and robustness of the proposed method in selecting useful modes from bearing vibration signals, it is applied to analyze signals from the CWRU database. The results are presented in the following section.

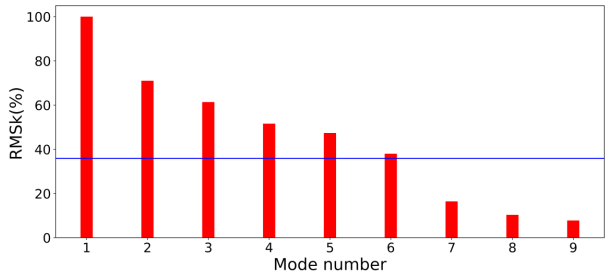

Fig. 12. The RMS representation of each mode with threshold = 35.90%

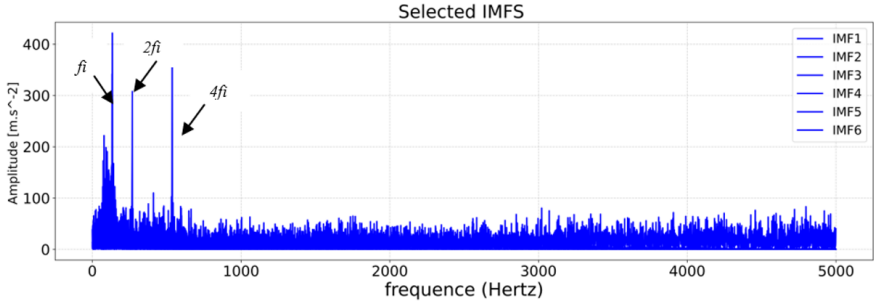

Fig. 13. The selected modes using OEMD for inner race vibration signal

4.2. CWRU database analysis

In this study, we utilized vibration data from the CWRU database. The experimental setup consists of an electric motor and two bearings [27]. The key characteristics of the vibration signals analyzed in both healthy and faulty states for the two cases are summarized in Tables 4 and 5.

Table 4. Vibration signal characteristic

Bearing type	Operating speed	Operating load	Sampling frequency
SKF 6205-2RS JEM	1772 rpm	1491.4 Nm/s	48000 Hz

Table 5. Fault characteristics

Fault types	Fault diameter	Coefficient	f_i
BPFI	0.177 mm	5.415	159.48 Hz

The characteristic frequency of inner race fault can be represented as:

$$f_i = \text{BPFI} \times \textit{Shaft speed}(\text{RPS}) = 159.48 \text{ Hz}. \tag{19}$$

4.2.1. Results and discussions

The healthy bearing vibration signal (CWRU) is presented in Fig. 14, and its decomposition using Algorithm 1 (EMD optimization based on CC) is presented in Figs. 15 and 16.

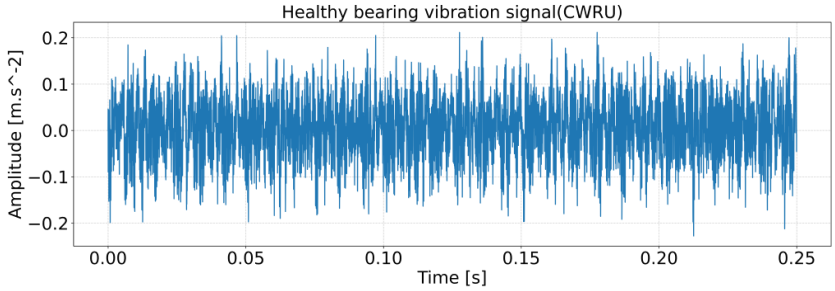

Fig. 14. Healthy state of vibration signal

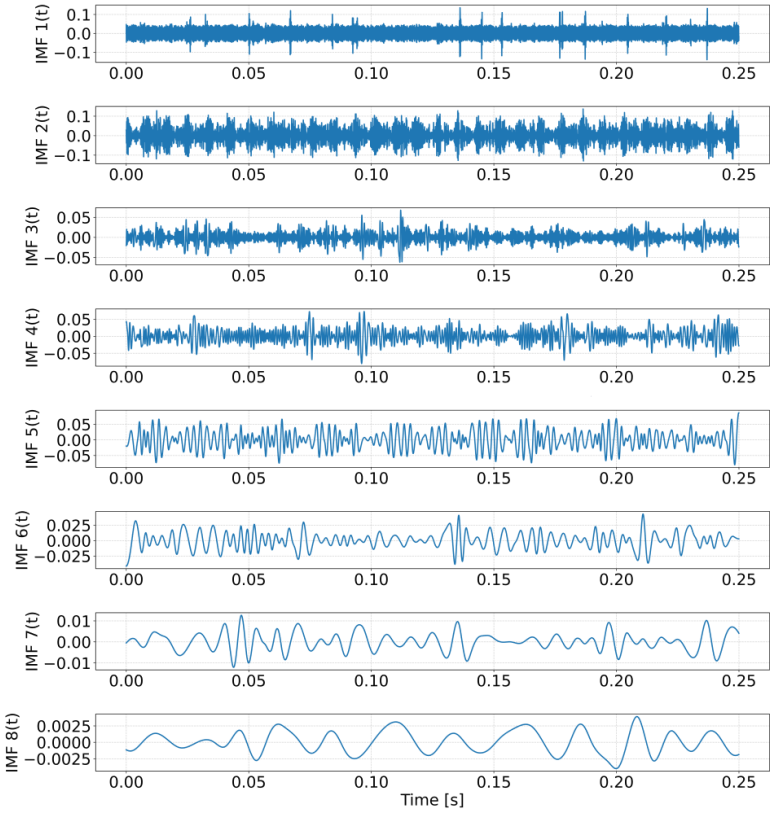

Fig. 15. The decomposition results of the healthy bearing vibration (CWRU) using the proposed approach (Algorithm 1) in time domain

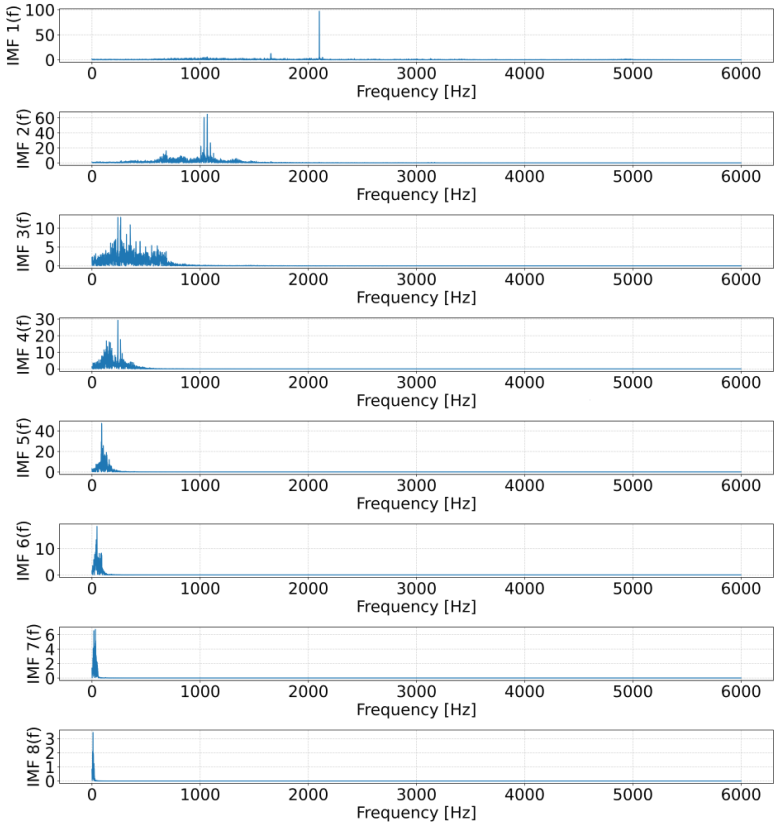

Fig. 16. The decomposition results of the healthy bearing vibration (CWRU) using the proposed approach (Algorithm 1) in frequency domain

From Fig. 15 and Fig. 16 the presence of over-decomposition is evident. Specifically, IMF_6, IMF_7, and IMF_8 exhibit significantly smaller amplitude levels compared to the other modes. As these modes do not contain valuable information, they are classified as noise and should be eliminated. Consequently, it is determined that only the first five modes should be considered. This conclusion is further validated by the second component of the proposed approach (the RMS-based algorithm), which automatically selects the useful IMFs, as demonstrated in Fig. 17.

Based on the data presented in Table 6 and the RMS histogram in Fig. 17, it can be concluded that the last three modes should be excluded, as their RMS

Table 6. The estimated RMS for each mode using OEMD algorithm

Modes	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7	IMF_8
RMS [m/s^2]	0.652	0.745	0.490	0.477	0.446	0.300	0.141	0.071
RMS (%)	87.51	100	65.74	63.97	59.83	40.36	18.99	9.58

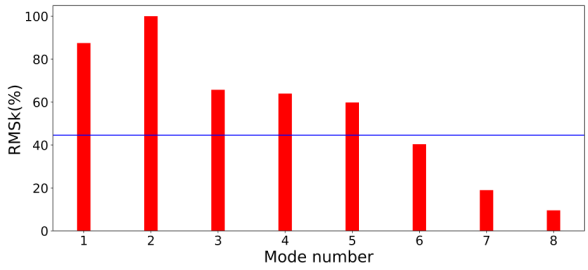

Fig. 17. The RMS histogram for inner race vibration signal with threshold $\varepsilon = 44.60\%$

percentages fall below the threshold of $\varepsilon = 44.60\%$. Furthermore, Fig. 18 illustrates the effectiveness of the proposed method in automatically selecting modes that contain valuable information.

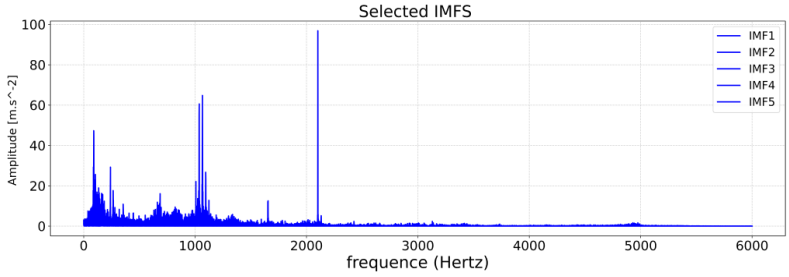

Fig. 18. The selected modes using OEMD for healthy state vibration signal (CWRU)

In the faulty state of the bearing, the vibration signal for the inner race fault is illustrated in Fig. 19. The decomposition results using the proposed approach (the first algorithm) are shown in Figs. 20 and 21.

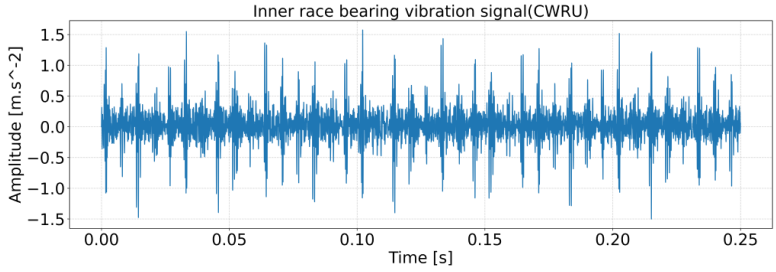

Fig. 19. Faulty state of vibration signal

In the decomposition of the faulty state, shown in Figs. 20 and 21, over-decomposition occurs as the three last modes exhibit significantly lower amplitude levels compared to IMF_1, IMF_2, IMF_3, and IMF_4. These modes are thus regarded as noise and should be excluded, as they do not provide meaningful information.

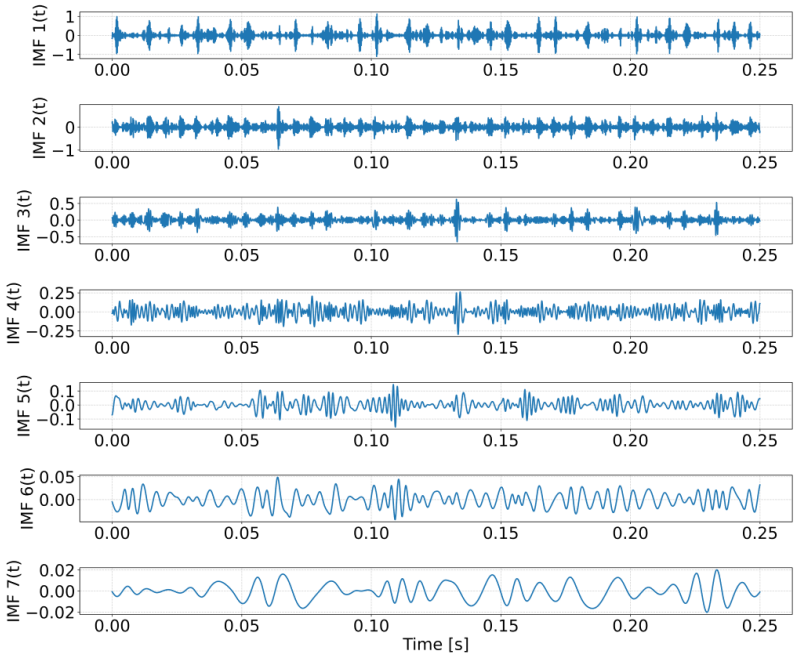

Fig. 20. The decomposition results of the healthy bearing vibration (CWRU) using the proposed approach (Algorithm 1) in time domain

Therefore, it is concluded that only the first four modes should be considered in this case. This conclusion is corroborated by the second component of the proposed approach (the RMS algorithm), which facilitates the automatic selection of useful Intrinsic Mode Functions (IMFs), as detailed in Table 7 and Fig. 22.

Table 7. The estimated RMS for each mode using OEMD algorithm

Modes	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7
RMS [m/s^2]	1.883	1.616	1.125	0.821	0.565	0.333	0.171
RMS (%)	100	85.82	59.74	43.63	30.01	17.69	9.10

Fig. 23 demonstrates the automatic selection of modes containing valuable information. Additionally, the characteristic frequency of the inner race fault, $f(i) = 159.48$ Hz along with its harmonics ($2f_i$, $4f_i$, $6f_i$, $8f_i$), has been accurately identified, thereby validating the feasibility and effectiveness of the proposed method.

The results show that the proposed method effectively and autonomously selects the optimal modes containing relevant information, successfully reducing the impact of noise. Moreover, it accurately identifies the characteristic frequency and its harmonics associated with the faulty condition, specifically the inner race defect.

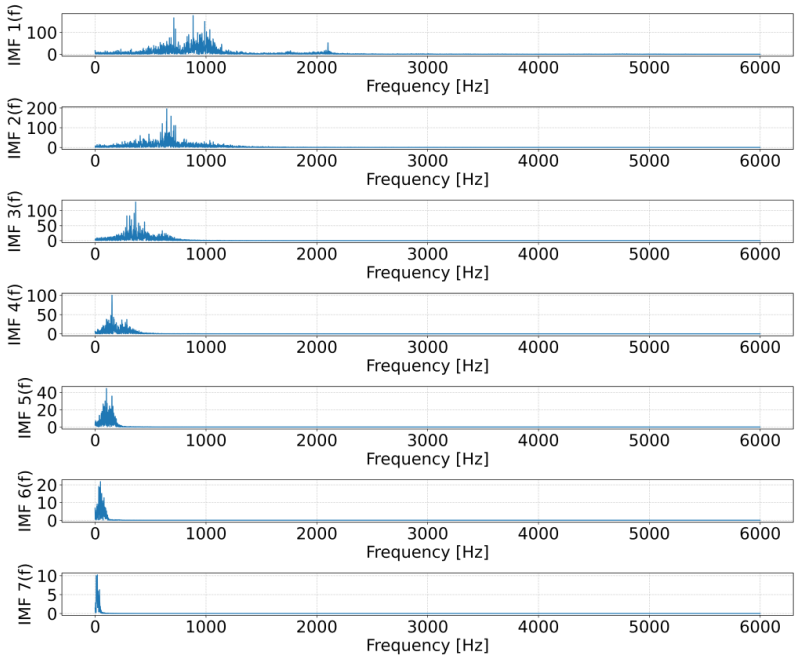

Fig. 21. The decomposition results of the faulty bearing vibration (CWRU) using the proposed approach (Algorithm 1) in frequency domain

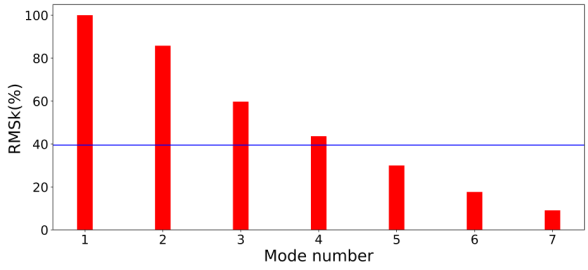

Fig. 22. The RMS histogram for inner race vibration signal with threshold $\varepsilon = 39.54\%$

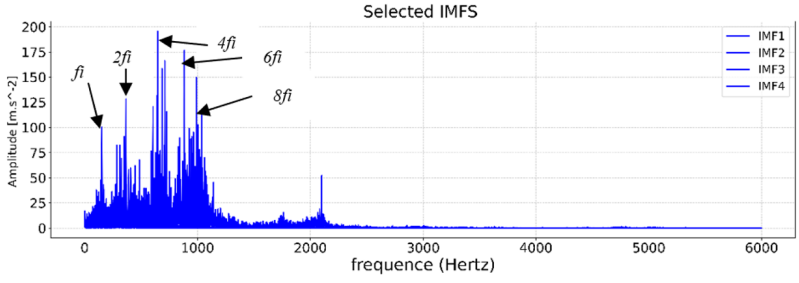

Fig. 23. The selected modes using OEMD for inner race vibration signal (CWRU)

5. Comparative analysis with other decomposition methods

To assess the efficacy of the proposed method for analyzing bearing vibration signals, this section presents a comparison of both the maximum center frequency observation (MCFO) and the center frequency statistical analysis (CFSA) methods.

The MCFO methodology is based on analyzing the distribution of the highest center frequencies associated with the modes. Within the framework of the variational mode decomposition (VMD) technique, it is observed that the center frequency of each intrinsic mode function (IMF) increases progressively with the mode number. The threshold at which the maximum center frequencies stabilize serves as an indicator for determining the optimal mode count [29].

The primary concept underlying the CFSA method involves calculating the frequencies of modes that exceed the mean frequency value. This calculated figure is then regarded as the optimal mode count. The procedural steps of the algorithm are outlined as follows [29]:

Algorithm 3: CFSA algorithm

1. Initialize the VMD parameters (k, α).
2. Extract the intrinsic mode functions (IMFs) from the input signal using VMD.
3. Estimate the center frequency for each IMF.
4. Create a histogram of the center frequencies.
5. Calculate the mean of the center frequencies and count the number (N) of frequencies that exceed the average.
6. Repeat steps 1–5 with ($k = k + 1$) until the count of exceeding frequencies no longer increases; this will indicate the cessation of decomposition, at which point the optimal IMF value is N.

The CFSA and MCFA methods are applied to both states of vibration signals (healthy and faulty states) and the results are presented in Fig. 24 and Fig. 25, respectively.

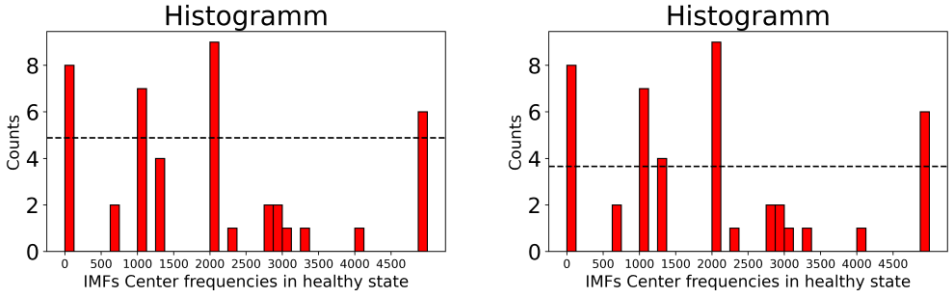

Fig. 24. Center frequency histogram for healthy and faulty states

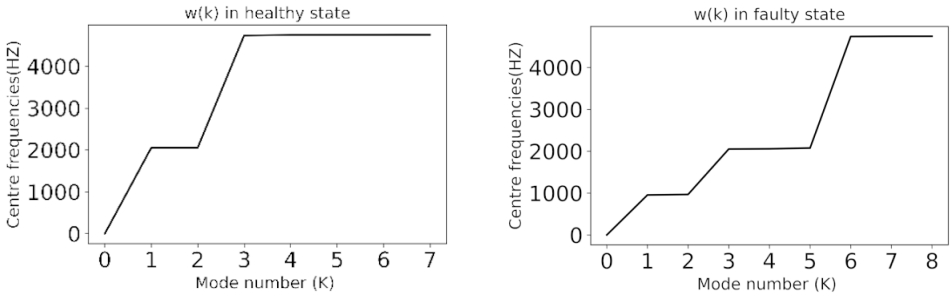

Fig. 25. MCFA trend of bearing vibration signal in healthy and faulty states

Fig. 24 illustrates that there are five dominant center frequencies exceeding the average count ($n = 3.66$) in the healthy state, while four dominant center frequencies surpass the average count ($n = 4.88$) in the faulty state. Consequently, based on the CFSA method, the optimal mode count is determined to be five in the healthy state and four in the faulty state.

The outcomes from the implementation of the MCFO method are depicted in Fig. 25. In the healthy state, it is noted that the center frequencies of the modes stabilize at $K = 3$, suggesting that the optimal number of intrinsic mode functions (IMFs) is 3. Additionally, in the faulty state the center frequencies of the modes stabilize at $K = 6$, indicating that the optimal number of IMFs is 6.

Table 8 summarizes the results of determining the optimal mode count in both healthy and faulty states of the vibration signal.

Table 8. The optimal mode number for each method

The method	Optimal number in healthy state	Optimal number in faulty state
MCFO	3	6
CFSA	5	4
OEMD	5	4

6. Conclusions

In this paper, we introduce a novel method for bearing fault detection, based on cross-correlation (CC) and root mean square (RMS) algorithms, developed to overcome EMD limitations in the decomposition results. The proposed approach is applied to real-world vibration signals, obtained from an experimental study of bearing vibration dataset and a public online dataset from the Case Western Reverse University (CWRU) to assess its feasibility and the effectiveness. The results demonstrate that the proposed method can automatically and accurately select the optimal modes containing high-amplitude peaks at the defects characteristic frequencies, thereby enabling both the detection and the identification of faults. In the first case, the vibration signal exhibits very high amplitude peaks at the

characteristic frequency of the inner race fault (f_i = 134.27 Hz) and its harmonics ($2f_i$, $4f_i$). In the second case, the inner race fault at f_i = 159.48 Hz along with its harmonics ($2f_i$, $4f_i$, $6f_i$, $8f_i$), is also clearly determined.

However, the online application of the proposed approach for bearing faults diagnosis is limited by the computer's available memory, preventing continuous analysis of long time series. Nevertheless, future work will demonstrate the potential of analyze and interpret large, complex time series and extend the application of this method to diagnosing various types of faults.

References

[1] P. Gangsar and R. Tiwari. Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review. *Mechanical Systems and Signal Processing*, 144:106908, 2020. doi: 10.1016/j.ymssp.2020.106908.

[2] A. Zeiler, R. Faltermeier, I.R. Keck, A.M. Tomé, C.G. Puntonet and E.W. Lang. Empirical mode decomposition-an introduction. In: *The 2010International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, Barcelona, Spain, 2010. doi: 10.1109/IJCNN.2010.5596829.

[3] J.R. Yeh, J.S. Shieh, and N.E. Haung. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, 2(2):135–156, 2010. doi: 10.1142/s1793536910000422.

[4] Y. Li, S. SI, Z Liu, and X. Liang. Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *Journal of Systems Engineering and Electronics*, 30(4):799–814, 2019. doi: 10.21629/jsee.2019.04.17.

[5] S.R. Qin and Y.M. Zhong. Research on the unified mathematical model for FT, STFT and WT and its applications. *Mechanical Systems and Signal Processing*, 18(6):1335–1347, 2004. doi: 10.1016/j.ymssp.2003.12.002.

[6] R. Hao and F. Li., A new method to suppress the EMD end point effect. *Zhendong Ceshi Yu Zhenduan/Journal of Vibration, Measurement and Diagnosis*, 38(2):341–345, 2018.

[7] A. Ayenu-Prah and N. Attoh-Okine. A criterion for selecting relevant intrinsic mode functions in empirical mode decomposition. *Advances in Adaptive Data Analysis*, 2(1):1–24, 2010. doi: 10.1142/S1793536910000367.

[8] W. Su, F. Wang, Z. Zhang, Z. Guo, and H. Li. Application of EMD denoising and spectral kurtosis in early fault diagnosis of rolling element bearings. *Journal of Vibration and Shock*, 29(3):18–21, 2010 (in Chinese). doi: 10.13465/j.cnki.jvs.2010.03.046.

[9] A. Komaty, A. Boudraa, and D. Dare. EMD-based filtering using the Hausdorff distance. In: *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 000292-000297, Vietnam, 2012. doi: 10.1109/ISSPIT.2012.6621303.

[10] G. Yang, Y. Liu, Y. Wang, and Z. Zhu. EMD interval thresholding denoising based on similarity measure to select relevant modes. *Signal Processing*, 109:95–109, 2015. doi: 10.1016/j.sigpro.2014.10.038.

[11] K. Dragomiretskiy and D. Zosso. Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, 62(3):531–544, 2014. doi: 10.1109/tsp.2013.2288675.

[12] H. Yang, S. Liu, and H. Zhang. Adaptive estimation of VMD modes number based on cross correlation coefficient. *Journal of Vibroengineering*, 19(2):1185–1196, 2017. doi: 10.21595/jve.2016.17236.

[13] Q. Ni, J.C. Ji, K. Feng, and B. Halkon. A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis. *Mechanical Systems and Signal Processing*, 164:108216, 2022. doi: 10.1016/j.ymssp.2021.108216.

[14] Y. Bousseloub, F. Medjani, A. Benmassoud, T. Kezai, A. Belharma, and I. Attoui. New method for bearing fault diagnosis based on variational mode decomposition technique. *Diagnostyka*, 25(2):2024208, 2024. doi: 10.29354/diag/186751.

[15] W. Li, D. Liu, W. Deng at al. Fault diagnosis using variational autoencoder GAN and focal loss CNN under unbalanced data. *Structural Health Monitoring*, 2024. doi: 10.1177/14759217241254121.

[16] H. Zhao, Y. Gao, and W. Deng. Defect detection using shuffle Net-CA-SSD lightweight network for turbine blades in IoT. *IEEE Internet of Things Journal*, 11(20):32804–32812, 2024. doi: 10.1109/JIOT.2024.3409823.

[17] H. Zhao, L. Wang, Z. Zhao, and W. Deng. A new fault diagnosis approach using parameterized time-reassigned multisynchrosqueezing transform for rolling bearings. *IEEE Transactions on Reliability*, 2024. doi: 10.1109/TR.2024.3371520.

[18] X. Ran, N. Suyaroj, W. Tepsan, J. Ma, X. Zhou, and W. Deng. A hybrid genetic-fuzzy ant colony optimization algorithm for automatic K-means clustering in urban global positioning system. *Engineering Applications of Artificial Intelligence*, 137(Part B):109237, 2024, doi: 10.1016/j.engappai.2024.109237.

[19] C. Huang, D. Wu, X. Zhou, Y. Song, H. Chen, and W. Deng. Competitive swarm optimizer with dynamic multi-competitions and convergence accelerator for large-scale optimization problems. *Applied Soft Computing*, 167(Part A):112252, 2024. doi: 10.1016/j.asoc.2024.112252.

[20] Y. Song, L. Han, B. Zhang, and W. Deng. A dual-time dual-population multi-objective evolutionary algorithm with application to the portfolio optimization problem. *Engineering Applications of Artificial Intelligence*, 133(Part F):108638, 2024. doi: 10.1016/j.engappai.2024.108638.

[21] A.K. Shah, A. Yadav, and H. Malik. EMD and ANN based intelligent model for bearing fault diagnosis. *Journal of Intelligent & Fuzzy Systems*, 35(5):5391–5402, 2018. doi: 10.3233/JIFS-169821.

[22] S.S. Refaat, H. Abu-Rub, M.S. Saad, E.M. Aboul-Zahab, and A. Iqbal. ANN-based for detection, diagnosis the bearing fault for three phase induction motors using current signal. In: *IEEE International Conference on Industrial Technology (ICIT)*, pages 253–258, Cape Town, South Africa, 2013. doi: 10.1109/ICIT.2013.6505681.

[23] J. Ben Ali, N. Fnaiech, L. Saidi B. Chebel-Morello, and F. Fnaiech. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics*, 89:16–27, 2015. doi: 10.1016/j.apacoust.2014.08.016.

[24] I.I.E. Amarouayache, M.N. Saadi, N. Guersi, and N. Boutasseta. Bearing fault diagnostics using EEMD processing and convolutional neural network methods. *The International Journal of Advanced Manufacturing Technology*, 107:4077–4095, 2020. doi: 10.1007/s00170-020-05315-9.

[25] D.H. Lee, J.H. Ahn, and B.H. Koh. Fault detection of bearing systems through EEMD and optimization algorithm. *Sensors*, 17(11):2477, 2017. doi: 10.3390/s17112477.

[26] R.N. Toma, C.H. Kim, and J.M. Kim. Bearing fault classification using ensemble empirical mode decomposition and convolutional neural network. *Electronics*, 10(11):1248, 2021 doi: 10.3390/electronics10111248.

[27] Case Western Reserve University Bearing Data Center. Website: 48k Drive End Bearking Fault Data | Case School of Engineering | Case Western Reserve University.

[28] W.A. Smith and R.B. Randall. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100–131, 2015. doi: 10.1016/j.ymssp.2015.04.021.

[29] S. Wu, F. Feng, J. Zhu, C. Wu, and G. Zhang. A method for determining intrinsic mode function number in variational mode decomposition and its application to bearing vibration signal processing. *Shock and Vibration*, 2020:8304903, 2020. doi: 10.1155/2020/8304903.