

Article

# A Rail Fault Diagnosis Method Based on Quartic $C^2$ Hermite Improved Empirical Mode Decomposition Algorithm

Hanzhong Liu <sup>1,2</sup>, Chaoxuan Qin <sup>3</sup> and Ming Liu <sup>2,\*</sup>

<sup>1</sup> School of Automation, Nanjing Institute of Technology, Nanjing 211167, China

<sup>2</sup> Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China

<sup>3</sup> School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

\* Correspondence: eelium@ust.hk

Received: 1 June 2019; Accepted: 23 July 2019; Published: 26 July 2019



**Abstract:** For compound fault detection of high-speed rail vibration signals, which presents a difficult problem, an early fault diagnosis method of an improved empirical mode decomposition (EMD) algorithm based on quartic  $C^2$  Hermite interpolation is presented. First, the quartic  $C^2$  Hermite interpolation improved EMD algorithm is used to decompose the original signal, and the intrinsic mode function (IMF) components are obtained. Second, singular value decomposition for the IMF components is performed to determine the principal components of the signal. Then, the signal is reconstructed and the kurtosis and approximate entropy values are calculated as the eigenvalues of fault diagnosis. Finally, fault diagnosis is executed based on the support vector machine (SVM). This method is applied for the fault diagnosis of high-speed rails, and experimental results show that the method presented in this paper is superior to the traditional EMD algorithm and greatly improves the accuracy of fault diagnosis.

**Keywords:** empirical mode decomposition (EMD); fault diagnosis; Hermite interpolation; kurtosis; approximate entropy

## 1. Introduction

High-speed trains have developed rapidly in the recent years. Compared to the traditional trains, high-speed trains have advantages such as fast speed, strong carrying capacity, green environmental protection, high safety, and good economic benefits. However, when the high-speed train accelerates and brakes, the rails are prone to severe friction, crushing, and impact. Under harsh working conditions, the rails easily fail. If these failures are not discovered in time, the defects in the rails can expand rapidly causing the rails to break, which can result in serious safety concerns such as accidents. Hence great importance is placed on early fault diagnosis studies of the high-speed rails.

Domestic and foreign scholars have conducted a lot of research on the safety of high-speed rails. Yin Kai used image detection and other related technologies to detect faults in the rails [1]. Li Haiqing analyzed high-speed rail vibration signals and used an empirical mode decomposition (EMD) algorithm to extract the rail damage information for fault detection [2]. Xu Peng et al. used the differential eddy current detection system to detect cracks in high-speed rails and obtained the amplitude and phase changes from the crack detection signals; then they deduced the relation between the detection signal and the crack [3]. Xiongxiang Liu used deep convolutional neural networks to identify the surface defects in the two-dimensional images of the rails [4]. However, the deep learning model has high computational complexity and limited recognition accuracy, which may lead to misjudgment. There are other nondestructive testing methods such as magnetic particle testing [5],

ray detection [6], and ultrasonic testing [7]. Ultrasonic testing is widely used among them, as its accuracy is high and it is easily operated. However, for surface micro-cracks, the detection effect is often affected by the rail surface roughness. At present, the commonly used flaw detection vehicles can only detect the defects on the rail surface, it is difficult to detect damage inside the rail, as internal damage is usually extremely serious.

Common faults in rails include damage at the inner rail head, rail surface wear, damage at the rail waist or bottom [2]. Because of the overload and strong impact during the train running, the fault mode includes various damage, which may be a single fault mode or a combination of fault modes. When the rail is in a fault state, its vibration signal will contain abnormal components such as impact and friction. By analyzing the vibration signal, the fault features of the rail can be obtained at an early stage.

Many research about the fault features extraction from vibration signals and diagnosis have been carried out over the years. Algorithms based on the Fourier transform, such as short-time Fourier transforms, Winger–Ville, and wavelet transforms [8–11], have played important roles, but the performances of these algorithms is limited due to their lack of adaptability when dealing with nonlinear nonstationary signals. Huang presented the EMD algorithm, which is a signal processing method based on the instantaneous frequency [12]. It has strong adaptability to non-stationary signals and has been widely used. However, the traditional EMD algorithm has problems such as curves fitting with the overshoot, undershoot, endpoint swing, and mode mixing. Therefore, the research of EMD focuses on the suppression of the mode mixing and optimization of the sifting procedure. For mode mixing, scholars usually blame the existence of the intermittent signal, as the frequency components of the signal are too close and the amplitude of the low-frequency signal is too large [13,14]. The noise-assistant methods are widely used to settle the mode mixing problem. Wu and Huang [15] added white Gaussian noises to the original signal to strengthen the high frequency component, which is called Ensemble Empirical Mode Decomposition (EEMD) method. Based on the EEMD, the complementary ensemble empirical mode decomposition method (CEEMD) [16] and complete EEMD with adaptive noise method (CEEMDAN) [17] are also the prominent methods to solve the mode mixing problem. In recent years, extensive efforts have been carried out for mode mixing prevention [18–21]. For optimization of sifting procedure, Yang used the B-spline interpolation method to improve the EMD algorithm; though, the signal length was assumed to be infinite and the endpoint effect was introduced [22]. Zhu W used the piecewise cubic Hermite interpolation method to improve the envelope process of the EMD algorithm and obtained good results [23]; however, the curve was bent artificially so that its smoothness was affected. Di Z et al. used the cubic trigonometric cardinal interpolation method with adjustable shape to carry out EMD algorithm, accompanied with great complexity [24].

In this paper, the rail vibration signal is decomposed with the quartic  $C^2$  Hermite interpolation method improved EMD algorithm, and the intrinsic mode function (IMF) components are obtained. Additionally, singular value decomposition [25] (SVD) for the IMF components is performed to determine the principal components of the signal. Then, through the cross correlation coefficient calculation between the IMF components and the original signal, the IMF components with larger correlation coefficients are reconstructed. Then, kurtosis and approximate entropy values of the reconstructed signal are calculated as fault eigenvalues, and finally, early faults of rails are detected and identified by support vector machine [26] (SVM) classifier based on eigenvalues.

The contributions of the paper are as follows:

- (1) In order to solve the undershoot problem caused by traditional EMD algorithm, a quartic  $C^2$  Hermite interpolation algorithm is proposed to make the curve more smooth and flexible.
- (2) Adopt a low-cost and simple method of rail fault diagnosis, use the acceleration sensor to obtain the rail vibration signal, and analyze the signal so as to carry out effective fault diagnosis.

- (3) Use quartic  $C^2$  Hermite improved-EMD algorithm to analyze the rail vibration signal and extract the kurtosis and approximate entropy for fault diagnosis. Experiments show that this method can diagnose the fault effectively.

The rest of the paper is arranged as follows. Section 2 introduced the traditional EMD algorithm and its undershoot problem. Aiming at addressing this problem, an improved EMD algorithm with quartic  $C^2$  Hermite [27] interpolation is presented. Section 3 introduced the singular value decomposition algorithm and the signal reconstruction method used in this paper. Section 4 introduced the fault features used in this paper, which are kurtosis and approximate entropy, and the whole rail fault diagnosis process. Section 5 introduced relevant experiments to verify the effectiveness of the algorithm.

## 2. Empirical Mode Decomposition

### 2.1. EMD Principle

The EMD method was proposed by Huang et al. in 1988. This method exhibits outstanding performance in dealing with non-linear and non-stationary random signals. Therefore, it is widely used in the fields of signal denoising, signal features extraction, and so on. The EMD method decomposes the original signal into multiple IMF components based on the local feature time scale [12]. The conditions that must be met for IMF are as follows:

- (1) The numbers of extreme points and zero points must be equal or at most different in one over the length of the data;
- (2) At any data point, the average of the envelope of the local maximum and that of the local minimum must be zero.

The signal that can be decomposed with the EMD algorithm needs to satisfy the following conditions:

- (1) Have at least two extreme points, including a maximum point and a minimum point;
- (2) A signal feature time scale is determined by the time interval between extreme points;
- (3) Have no extreme points; the first-order or higher-order derivatives of inflection points can be taken as extreme points.

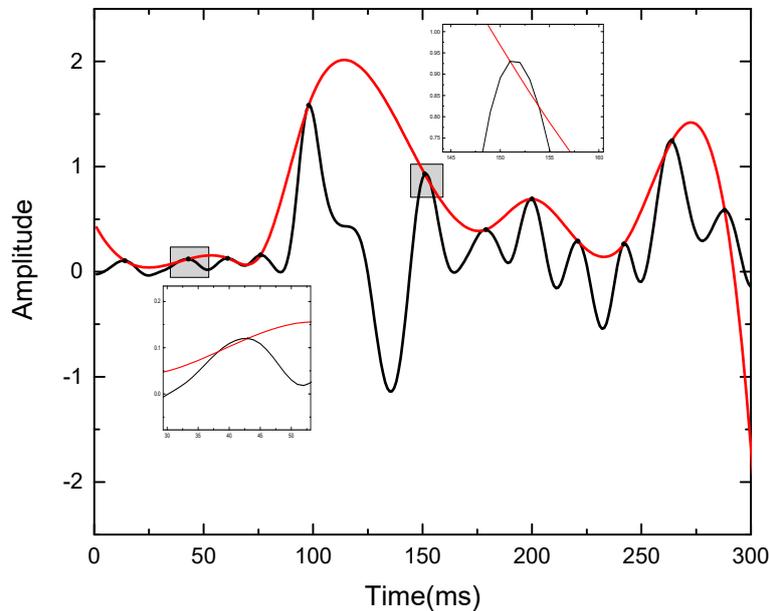
For a signal named  $X$ , the EMD steps are as follows:

- (1) Determine all the local extreme points of the signal, use the cubic spline interpolation to fit all local maximum points to form the upper envelope curve, and fit all local minimum points to form the lower envelope curve.
- (2) Calculate  $h_0(t) = x(t) - \frac{x_0^u(t) + x_0^d(t)}{2}$ .
- (3) Determine whether  $h_0(t)$  satisfies the conditions of the IMF. If they are satisfied, then  $h_0(t)$  is the first IMF component; otherwise, repeat steps 1 and 2 for  $h_0(t)$  until the IMF conditions are satisfied. The IMF component after the first filtration is recorded as  $s_1(t)$ .
- (4) Calculate the remaining signal  $r_1(t) = x(t) - s_1(t)$ , repeat step 1 to 3 for  $r_1(t)$  to get the second IMF component, and in this way, repeat  $n$  times loop to get  $n$  IMF components. Terminate the step when the remainder is a monotonic signal or meets a given condition.

### 2.2. Undershoot

The signal filtration process is the most important part of the EMD algorithm and directly affects the decomposition results. Traditionally, when the cubic spline curve is used to fit extreme points to form the envelope curve, the undershoot phenomenon easily occurs. As shown in Figure 1, the upper envelope curve is fitted by cubic spline interpolation. The ideal upper envelope curve should satisfy  $x^c(i) \geq x(i), \forall i = 1, 2, \dots, N$  anywhere. However, there are two places where the

values of the envelope curve are less than those of the original signal, and this phenomenon is called “undershoot”. The envelope curve cannot surround the original signal curve completely, resulting in meaningless IMF components during the decomposition process, or signal distortion during the signal reconstruction process.



**Figure 1.** Undershoot phenomenon.

The reason for the undershoot phenomenon is that the traditional fitting algorithm with cubic spline interpolation only satisfies  $C^1$  continuous, such that the shape is fixed under the given conditions and lacks flexibility. So the fitting performance is poor when the non-linear and non-stationary signal changes greatly.

### 2.3. Improved EMD Algorithm

In order to solve the drawbacks of the method with cubic spline interpolation, the trigonometric interpolation method was used to fit the envelope curve [24]. Yang proposed a method with piecewise cubic spline interpolation; however, the method has poor continuity and only satisfies  $C^1$  continuous [22]. Oberlin T proposed a method with which the envelope curve can be constructed directly to replace the method with cubic spline interpolation [28]; Zhu Weifang used a method with piecewise cubic Hermite interpolation to fit the envelope curve [23], which mainly used Lagrangian method to get the minimum value and optimized the derivative value at extreme points. However, this method also does not satisfy  $C^2$  continuous. Some common interpolation methods and their pros and cons are shown in Table 1.

**Table 1.** Common interpolation methods and their advantages and disadvantages.

Interpolation Methods	Advantages	Disadvantages
Cubic Spline (CS) CTCP [24]	Common algorithm, good flexibility	Overshoot and undershoot, high computational complexity
B-Spline (BS) [22]	Improve the local characteristics of EMD algorithm	The endpoint extension is assumed to be infinite, and the endpoint effect is not considered.
OS [28]	Adaptive and better than cubic spline	High computational complexity and poor smoothness.

Table 1. Cont.

Interpolation Methods	Advantages	Disadvantages
OPCH [23]	Taking the difference between extremes as the cost function, the results are more accurate and reasonable than cubic spline	Slow computation speed and do not satisfy $C^2$ continuous especially when there are many extreme points and cause undershooting.
MPCI [29]	Through strict theoretical analysis, it is proved that it can solve the overshoot problem caused by the cubic spline interpolation	Only satisfy $C^1$ continuous, the curve smoothness is poor and undershoot.
Piecewise Power Function Algorithm (PPFA) [30]	It is more flexible than cubic spline interpolation, and the curve is smooth.	There is a contradiction between smoothness and flexibility and undershooting occurs.

In this paper, the EMD algorithm is improved by the quartic  $C^2$  Hermite interpolation [27]. Suppose there are  $n + 1$  nodes in the interval  $[a, b]$ , which satisfy  $x_0 < x_1 < \dots < x_n$ , then the traditional standard cubic Hermite basis function is

$$\begin{cases} a_i(t) = 1 - 3t^2 + 2t^3 \\ a_{i+1}(t) = 3t^2 - 2t^3 \\ b_i(t) = t - 2t^2 + t^3 \\ b_{i+1}(t) = -t^2 + t^3 \end{cases} \quad (1)$$

where  $t = \frac{x-x_i}{x_{i+1}-x_i}$ .

This is convenient for the calculation with the standard cubic Hermite interpolation method, which is a common method used in engineering. However, the shape of the fitting curve obtained from this method is fixed, which lacks adaptability and only satisfies  $C^1$  continuous. In this paper, a quartic  $C^2$  Hermite interpolation method is presented, which adds a control factor to adjust the shape of the curve. The basis function is as follows:

$$\begin{cases} F_i(t) = \lambda_i t^4 - 2(\lambda_i - 1)t^3 + (\lambda_i - 3)t^2 + 1 \\ F_{i+1}(t) = -\lambda_i t^4 + 2(\lambda_i - 1)t^3 - (\lambda_i - 3)t^2 \\ G_i(t) = \lambda_i t^4 - (2\lambda_i - 1)t^3 + (\lambda_i - 3)t^2 + t \\ G_{i+1}(t) = -\lambda_i t^4 + (2\lambda_i - 1)t^3 - (\lambda_i - 3)t^2 \end{cases} \quad (2)$$

where  $\lambda$  is a control factor to adjust the shape of the curve.

The basis function satisfies such conditions as follows:

$$\begin{cases} F_i(0) = F_{i+1}(1) = 1 \\ F_i(1) = F_{i+1}(0) = 0; \\ F'_i(0) = F'_i(1) = F'_{i+1}(0) = F'_{i+1}(1) = 0 \\ G_i(1) = G_i(1) = G_{i+1}(0) = G_{i+1}(1) = 0 \\ G'_i(0) = G'_{i+1}(1) = 1 \\ G'_i(1) = G'_{i+1}(0) = 0 \\ F_i(t) + F_{i+1}(t) = 1 \end{cases} \quad (3)$$

The method with quartic Hermite interpolation has the same properties as the one with traditional cubic Hermite interpolation, but the shape of the curve is more flexible due to the parameter  $\lambda$ . When the parameter  $\lambda$  equals 0, the curve is standard cubic Hermite spline curve, which satisfies  $C^1$  continuous. The quartic Hermite spline curve at the interval of  $[a, b]$  is

$$C_i(x) = F_i(t)y_i + F_{i+1}(t)y_{i+1} + \dot{y}_i G_i(t)\Delta x_i + \dot{y}_{i+1} G_{i+1}(t)\Delta x_i, i = 0, 1, \dots, n - 1 \quad (4)$$

where  $y_i$  is the function value of curve,  $\dot{y}_i$  is its first derivative value, and  $\Delta x_i = x_{i+1} - x_i$ . If the curve satisfies  $C^2$  continuous, the second derivative must satisfy the condition as follows:

$$C''(x-0) = C''(x+0) \quad (5)$$

And the continuous equation by calculation is as follows

$$\begin{aligned} & \lambda_i(\Delta y_i + \Delta x_i \Delta \dot{y}_i) \Delta x_{i+1}^2 - \lambda_{i+1}(\Delta y_{i+1} + \Delta x_{i+1} \Delta \dot{y}_{i+1}) \Delta x_i^2 \\ & = -\Delta x_{i+1}^2(3\Delta y_i + \Delta x_i(\dot{y}_i + 2\dot{y}_{i+1})) - \Delta x_i^2[3\Delta y_{i+1} + \Delta x_{i+1}(\dot{y}_{i+2} + 2\dot{y}_{i+1})] \end{aligned} \quad (6)$$

When  $\lambda$  satisfies the constraint condition as in Equation (6), the plotted curve will satisfy  $C^2$  continuous and it will be smoother.

#### 2.4. The Shape of the Envelope Curve Determination

It is known from Section 2.3 that the shape of the fitting curve with quartic  $C^2$  Hermite interpolation is controlled by the parameter  $\lambda$ , and the curve cluster can be obtained with a different parameter  $\lambda$ . This paper draws the idea from reference [23] to find the curve whose length is the shortest in the curve cluster formed by different  $\lambda$ . Generally speaking, excessive length of curve will cause fitting overshoot. The curve length can be expressed as

$$L = \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} \sqrt{1 + \dot{C}(t)^2} dt, i = 1, 2, \dots, n-1 \quad (7)$$

The conditions for searching the optimal curve are as follows:

$$\begin{aligned} & \min L \\ & \text{st.} \begin{cases} 0 < \lambda < 2 \\ L = \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} \sqrt{1 + \dot{C}(t)^2} dt, i = 1, 2, \dots, n-1 \\ C_i(x) = F_i(t)y_i + F_{i+1}(t)y_{i+1} + \dot{y}_i G_i(t)\Delta x_i + \dot{y}_{i+1} G_{i+1}(t)\Delta x_i, i = 0, 1, \dots, n-1 \\ \lambda_i(\Delta y_i + \Delta x_i \Delta \dot{y}_i) \Delta x_{i+1}^2 - \lambda_{i+1}(\Delta y_{i+1} + \Delta x_{i+1} \Delta \dot{y}_{i+1}) \Delta x_i^2 \\ = -\Delta x_{i+1}^2(3\Delta y_i + \Delta x_i(\dot{y}_i + 2\dot{y}_{i+1})) - \Delta x_i^2[3\Delta y_{i+1} + \Delta x_{i+1}(\dot{y}_{i+2} + 2\dot{y}_{i+1})] \end{cases} \end{aligned} \quad (8)$$

After many experiments, it is found that the optimal curves will appear only when  $\lambda < 2$ . If  $\lambda$  is too large, the smoothness of the curve will be affected. The shape of the curve will be determined by using the genetic algorithm to find the optimal value of  $\lambda$ . The detailed steps are as follows.

- (1) Code the variables  $\lambda \in (0, 2]$ , use binary coding, and divide the interval into 1024 subintervals.
- (2) Calculate the fitness value of each individual according to Equation (8).
- (3) Selection operator, elite individual preservation strategy and rotation method selection operator are used to calculate the probability of each selected individual and the cumulative probability in the whole population fitness  $P_i = \frac{\text{fitness}(i)}{\sum_{\text{population}} \text{fitness}(i)}$ ,  $Q_i = \sum_{j=1}^i P_j$ .
- (4) Set the crossover probability  $p_c$ , and generates new individuals according to the single-point crossover rule.
- (5) Set the mutation probability  $p_m$ . The mutation individual is determined randomly and the reverse mutation operator is used for mutation operation.
- (6) Repeat steps (2)~(5) until the termination condition is reached.

### 3. Singular Value Decomposition and Signal Reconstruction

The matrix singular value is usually used for signal feature extraction because of its good stability [20]. The signal matrix named  $X_{m \times n}$  is decomposed as follows:

$$\begin{aligned} X &= U \Sigma V^T \\ &= [u_1, u_2, \dots, u_n] \begin{bmatrix} \delta_1 & & & \\ & \delta_2 & & \\ & & \dots & \\ & & & \delta_n \end{bmatrix} [v_1, v_2, \dots, v_n] \end{aligned} \quad (9)$$

where  $U$  is matrix of left singular vectors and  $V$  is matrix of right singular vectors.  $\Sigma = \text{diag}(\delta_1, \delta_2, \dots, \delta_n)$  is a singular value matrix and satisfies  $\delta_1 \geq \delta_2 \geq \dots \geq \delta_n$ .

The percentage of signal components is calculated as

$$\lambda_k = \frac{\sum_{i=1}^k \delta_i}{\sum_{j=1}^N \delta_j}, k = 1, 2, \dots, N \quad (10)$$

where  $\lambda_k$  represents the percentage of the first  $k$  signal components, and  $N$  is the number of IMF components obtained by EMD. The threshold value of percentage is set to 0.9, which means 90% of the original signal can be obtained. The IMF components are reordered according to the correlation coefficient calculated from each IMF and the original signal  $X$ . When  $\lambda_k > 0.9$ , the first  $k$  IMF components are used for signal reconstruction, and the remaining  $N-k$  IMF components are abandoned. The cross-correlation coefficient can be expressed as follows:

$$\gamma(X(t), IMF_i) = \frac{|\text{cov}(X(t), IMF_i)|}{\sqrt{\text{var}(X(t), X(t)) \cdot \text{var}(IMF_j, IMF_i)}} \quad (11)$$

where  $\text{cov}()$  represents the covariance function,  $\text{var}()$  represents the variance function,  $X(t)$  is the original signal, and  $IMF_i$  represents the  $i$ th IMF component obtained by EMD. The range of the cross-correlation coefficient is  $[0, 1]$ , which reflects the matching degree between these two signals. The closer the value is to 1, the closer the component is to the original signal.

### 4. Fault Diagnosis Method

#### 4.1. Fault Feature

Kurtosis is a dimensionless parameter that reflects the characteristics of signal distribution

$$K = \frac{E(x - \mu)^4}{\delta^4} \quad (12)$$

where  $x$  is the signal to be analyzed,  $\mu$  is the signal mean value, and  $\delta$  is the signal standard deviation. The fault rail vibration signals are usually accompanied by impact components to which the kurtosis is sensitive [31], so the larger the kurtosis value is, the higher the proportion of impact components in the signal, therefore the kurtosis can reflect the rail fault.

Approximate entropy can measure the complexity of the signal [32]. The more complex the signal mode is, the more irregular the signal, and the larger the approximate entropy. The approximate entropy calculation method is as follows.

- (1) The original signal is called  $X$ , which is an  $m$ -dimensional vector composed of continuous points on a time scale

$$X_i = [x(i), x(i+1), \dots, x(i+m-1)], i = 1, \dots, N-m+1 \quad (13)$$

- (2) Calculate vector distance

$$d[X(i), X(j)] = \max_{k=0, \dots, m-1} (|x(i+k) - x(j+k)|) \quad (14)$$

- (3) Given a threshold of similar tolerance represented by variable  $r$ , calculate each vector distance and get the number of less than  $r$ , which is expressed as  $C_i^m(r)$ ,

$$C_i^m(r) = \frac{1}{N-m} \text{Cont}(d[X(i), X(j)] < r) \quad (15)$$

$$i, j = 1, \dots, N-m+1, i \neq j$$

- (4) Calculate the log values of  $C_i^m(r)$  and get their mean value expressed as  $\phi^m(r)$ ,

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \quad (16)$$

- (5) Calculate the approximate entropy,

$$ApEn(m, r) = \lim_{N \rightarrow \infty} (\phi^m(r) - \phi^{m+1}(r)) \quad (17)$$

#### 4.2. Fault Diagnosis Steps

The fault diagnosis steps are as follows and the flow chart is shown in Figure 2.

- (1) Use an acceleration sensor to obtain the original vibration signal;
- (2) Process the vibration signal with the EMD algorithm based on quartic  $C^2$  Hermite interpolation described in Sections 2.3 and 2.4 to obtain the IMF components;
- (3) Process the acquired IMF components by using the SVD algorithm to obtain a singular value matrix;
- (4) Determine the number of primary components of the signal and reconstruct the signal;
- (5) Calculate the kurtosis and approximate entropy of the reconstructed signal as the eigenvalues of fault diagnosis;
- (6) Train the SVM classifier with kurtosis and approximate entropy values of training samples;
- (7) Classify by the SVM classifier based on the different eigenvalues.

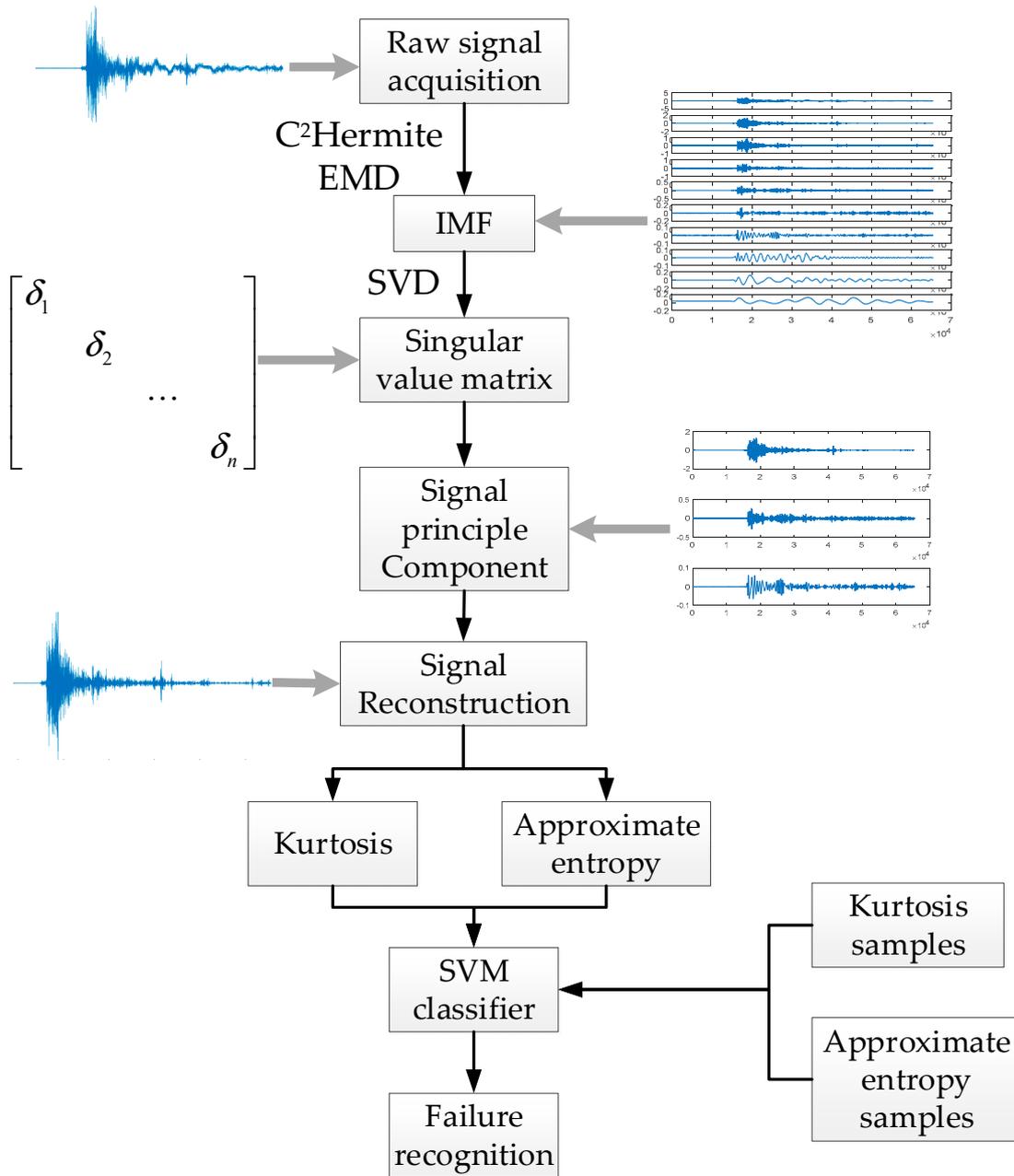
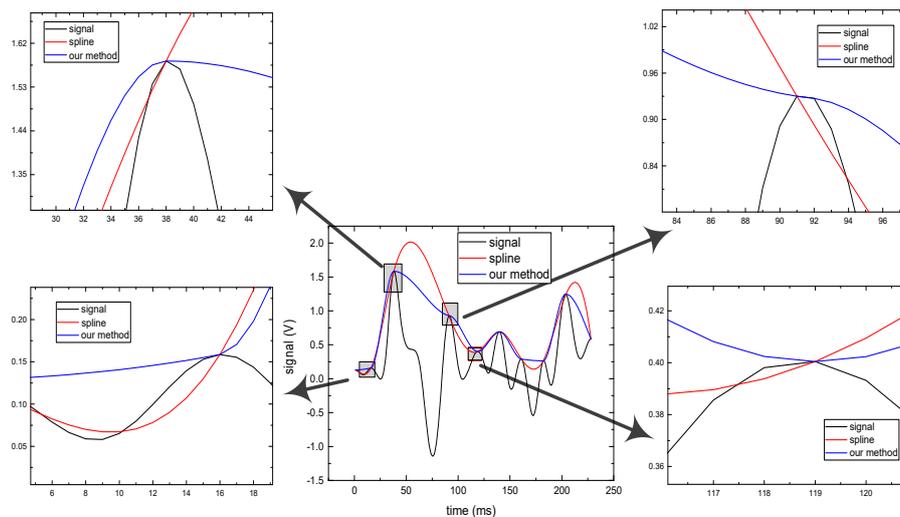


Figure 2. Flow chart of fault diagnosis.

## 5. Experiment

### 5.1. EMD Envelope Experiment

The algorithms with quartic  $C^2$  Hermite interpolation and cubic Hermite interpolation are applied to the vibration signal processing. Figure 3 shows a part of the vibration signal, where the black curve is the original signal. The red is the upper envelope curve fitted with cubic Hermite interpolation algorithm and the blue is the upper envelope curve with the quartic  $C^2$  Hermite interpolation algorithm. It can be seen from Figure 3 that obvious overshoot will appear with the cubic Hermite interpolation algorithm during non-stationary signal processing, and undershoot phenomenon appear in many positions (shown in the zoom Figure 3). But the quartic  $C^2$  Hermite interpolation algorithm presented in this paper effectively solves the overshoot and undershoot problem.



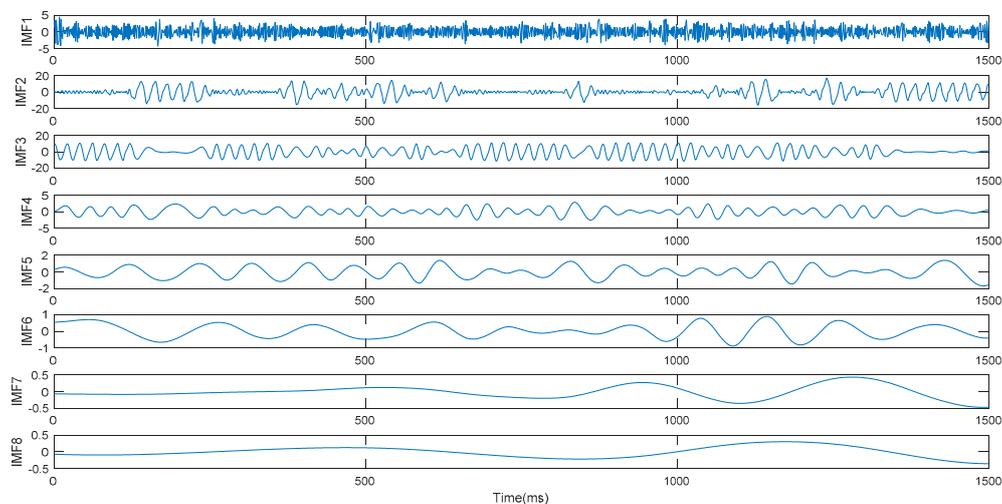
**Figure 3.** Envelope curve fitting of vibration signal.

### 5.2. EMD Simulation Experiment

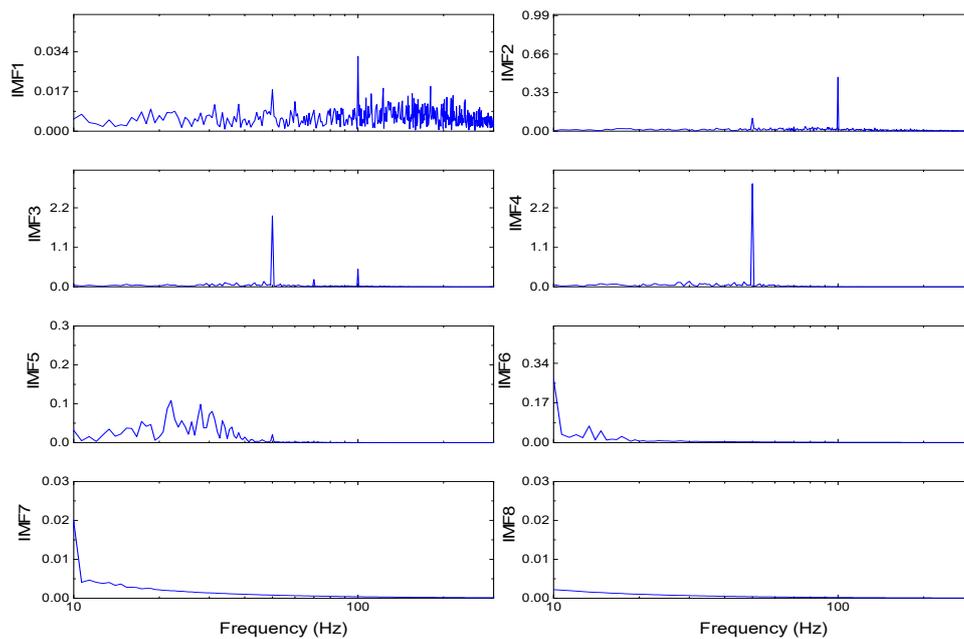
The signal in the simulation experiment is as follows:

$$s = 0.7 \sin(2\pi \times 10t) + 2 \sin(2\pi \times 100t) + (10 + \sin(2\pi \times 20t)) \cos(2\pi \times 50t) + noise \quad (18)$$

The simulation signal is composed of AM-FM and Gaussian white noise. The simulation platform is MATLAB2017a and the sampling frequency is set to 1000 Hz. In the simulation, the signal is processed separately with the traditional cubic Hermite interpolation algorithm, piecewise cubic Hermite interpolation algorithm, and the quartic  $C^2$  Hermite interpolation algorithm, which are presented in this paper. The IMF components obtained by EMD based on the quartic  $C^2$  Hermite interpolation algorithm and the corresponding spectrograms are shown in Figures 4 and 5, respectively. As can be seen from the figure, the frequency of the first IMF component (IMF1) is distributed in each frequency band, and the extracted IMF1 component approximates Gaussian white noise. The spectrums of IMF2 and IMF3 are concentrated at 100 Hz and 50 Hz, while the spectrum of IMF5 is concentrated at 10 Hz. Therefore, every mode of the signal can be effectively extracted with the proposed algorithm.



**Figure 4.** Intrinsic mode function (IMF) components obtained by the empirical mode decomposition (EMD) based on the  $C^2$  Hermite interpolation algorithm.



**Figure 5.** Components spectrograms obtained by EMD based on the  $C^2$  Hermite interpolation algorithm.

Mode mixing may appear in the decomposition [12]. Mode mixing means that an IMF component contains extremely different feature time scales or similar feature time scales distributed in different IMF components, which is caused by uneven distribution of signal extremums, signal interruption, or pulse, etc. The common method to solve mode mixing is to add white noise to the signal to suppress abnormal signal [15,17].

Index of orthogonality (IO) and index of energy conservation (IEC) are important indexes for evaluating the EMD results. The expressions are presented in Equations (19) and (20). The IO shows the orthogonality of the IMF component obtained with the EMD algorithm. The closer the value is to 0, the better the orthogonality between the IMF components. The IEC shows the energy conservation after signal decomposition. The closer the value is to 1, the higher the energy conservation by EMD.

$$\text{IO} = \sum_{t=0}^T \left( \frac{\sum_{i=1}^N \sum_{j=1}^N f_i(t) f_j(t)}{X^2(t)} \right) \quad (19)$$

$$\text{IEC} = \sum_{t=0}^n \frac{\sum_{i=1}^n |f_i(t)|^2}{|X(t) - r(t)|^2} \quad (20)$$

where  $X(t)$  represents the original signal,  $f_i(t)$  and  $f_j(t)$  represent the  $i$ -th and  $j$ -th IMF components by EMD respectively, and  $r(t)$  is a trend item.

Objective evaluation indicators are used to validate the algorithm presented in this paper, the result is shown in Table 2. The IO and IEC obtained with the algorithm presented in this paper are better than those obtained with the traditional cubic Hermite interpolation algorithm and piecewise cubic Hermite interpolation algorithm.

Table 2. Calculation results.

Indicators	Cubic Hermite Interpolation Algorithm	Piecewise Cubic Hermite Interpolation Algorithm	Proposed Algorithm
IO	0.08899	0.05211	0.04573
IEC	0.3129	0.6534	0.9123

### 5.3. The Effect of Noise on the Quartic $C^2$ Hermite Improved Empirical Mode Decomposition Algorithm

The influence of noise on the performance of EMD algorithm is analyzed through simulation experiments, and the simulation signal is:

$$s = \sin(2\pi \times 20t) + \sin(2\pi \times 100t) + noise \quad (21)$$

The signal is composed of sinusoidal signal with frequency of 20 Hz and 100 Hz superimposed with noise signal. The sampling frequency is set as 1000 Hz, the signal length is 1 s, and the noise variance is 0–1. Under different noises, the algorithm proposed in this paper is used to process the signal respectively. Figure 6 shows the IMF component with noise variance of 0.01, IMF2 is approximately a sinusoidal signal of 100 Hz, and IMF3 is approximately a sinusoidal signal of 20 Hz. The blue lines in the figure are IMF decomposed by EMD algorithm, and the gray lines are sinusoidal signals of 100 Hz and 20 Hz respectively. It can be seen from the figure that the algorithm in this paper can effectively decompose the signal and extract typical components of the signal.

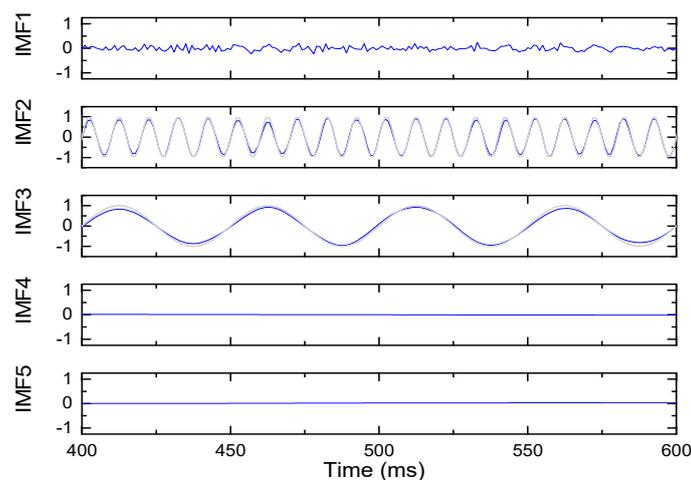
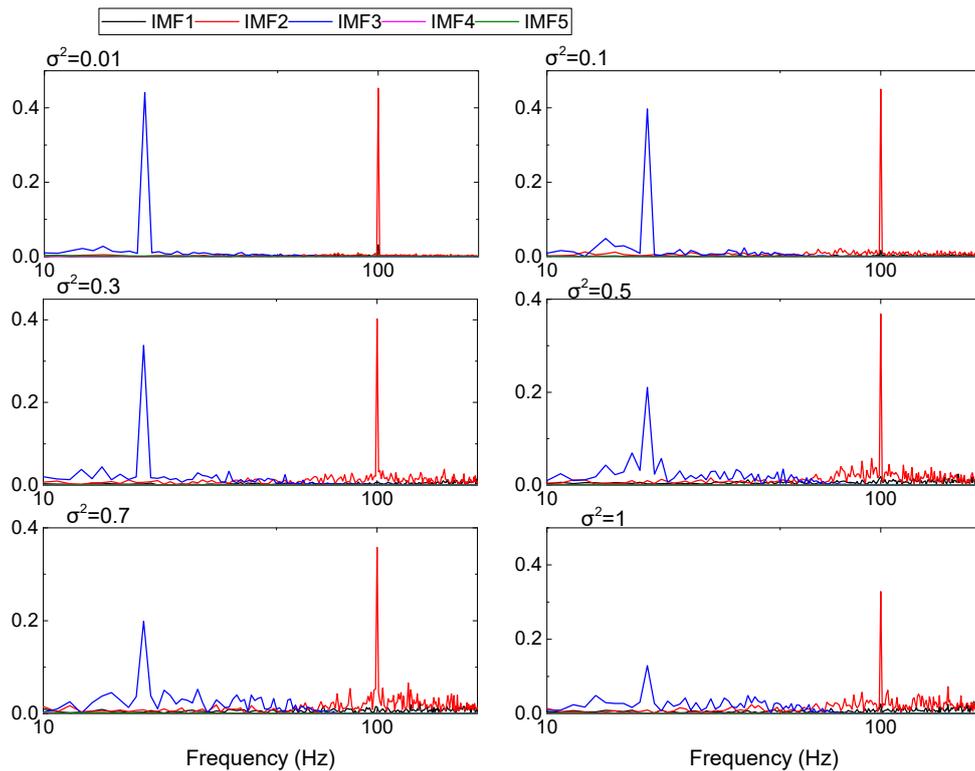


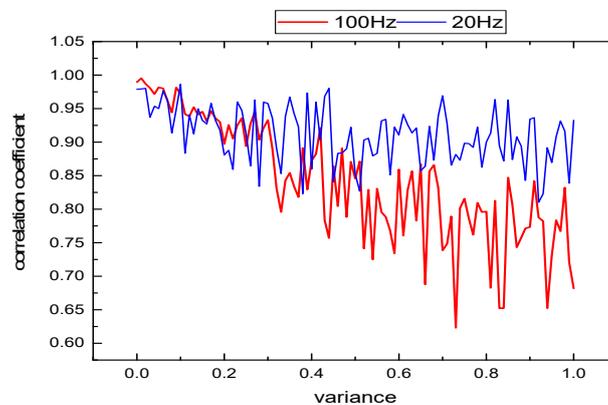
Figure 6. Components obtained by EMD based on the  $C^2$  Hermite interpolation algorithm.

Figure 7 shows the spectrum diagrams of each IMF component decomposed when the variances are 0.01, 0.1, 0.3, 0.5, 0.7, and 1 respectively, and the abscissa is represented by  $\log_{10}$ . As can be seen from the figure, the increase of noise has a certain impact on the EMD decomposition and has a more significant impact on the high-frequency region. However, signal frequency characteristics are not affected, even when the noise variance is large.

The noise variance is increased from 0 to 1 with an interval of 0.01, and the EMD decomposition is carried out and the correlation coefficient is calculated. As shown in Figure 8, the red line is the correlation coefficient of IMF2 and 100 Hz sinusoidal signal, and the blue line is the correlation coefficient of IMF3 and 20 Hz sinusoidal signal. It can be seen that, with the increase of noise variance, the correlation coefficients of both components show a decreasing trend, and the reduction range of 20 Hz signal is significantly smaller than that of the 100 Hz signal. Since EMD algorithm has this characteristic, EMD algorithm can be used for signal noise reduction. In particular, when the target signal is distributed in the low-frequency component, the denoising effect is good. If the target signal frequency is high, the denoising effect is generally poor.



**Figure 7.** Signal decomposition result by EMD based on the  $C^2$  Hermite interpolation algorithm under different noise variances.



**Figure 8.** Correlation coefficient of two components under different noise variances.

#### 5.4. Fault Eigenvalues Extracted from Rail Vibration Signal

The proposed algorithm is applied to vibration signal analysis to validate its effectiveness. The composition of the test system is shown in Figure 9, the rail is made with 60 kg/m standard steel. The sensors are mounted on the bottom of the rail so that the effect on train operation and track mechanical strength is minimal. The sensor used in the experiment is CYD103 piezoelectric acceleration sensor, its sensitivity is 20 pC/g. The sensor is connected to the YE5853 charge amplifier, which is used to modulate and transform the signals collected by the acceleration sensor for display on the computer and subsequent processing. The signal sampling frequency is 50 kHz. The IMF components of the signal obtained by EMD is as shown in Figure 10.

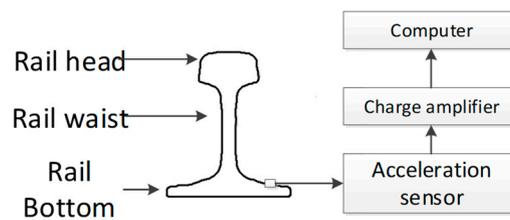


Figure 9. Composition of test System.

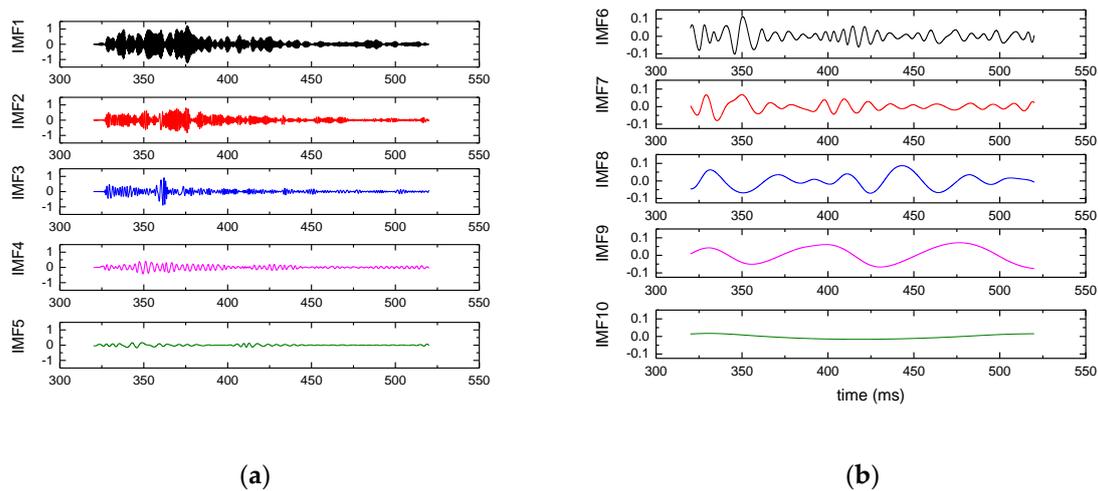


Figure 10. IMF components obtained by EMD. (a) IMF 1 to IMF 5, (b) IMF 6 to IMF 10.

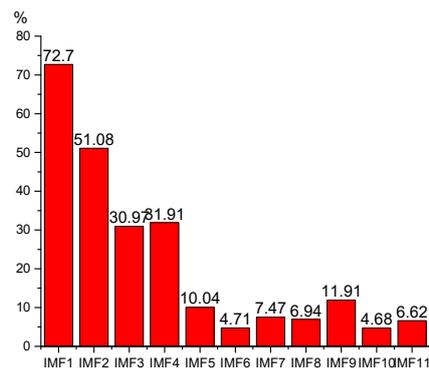
Singular value decomposition is performed for the IMF components obtained; the results are shown in Table 3. Suppose the accumulated percentage of principal components is 90%, then it can be seen that the percentage of the first eight principal component reaches 92.24%. So the signal can be reconstructed by using the first eight IMF components.

Table 3. Principal component of signal.

No.	Eigenvalues	Percentage of Current Eigenvalue in Total	Accumulated Percentage of the First K Eigenvalues in Total
1	29.2743	0.294605	0.294605
2	19.65641	0.197814	0.492419
3	12.7543	0.128354	0.620773
4	10.71655	0.107847	0.72862
5	5.604914	0.056406	0.785026
6	4.869944	0.049009	0.834035
7	4.58757	0.046167	0.880203
8	4.177541	0.042041	0.922244
9	3.021806	0.03041	0.952654
10	2.746282	0.027637	0.980292
11	1.95839	0.019708	1

As is shown in Figure 11, the cross correlation coefficient calculations between the IMF components and the original signal are performed. The first eight IMF components with high cross correlation coefficient are used for reconstruction.

According to a large number of investigations, the main failure forms of high-speed rails in China are rail surface peeling and corrugation, which is shown in Figure 12.



**Figure 11.** Cross correlation for the IMF components and the original signal.

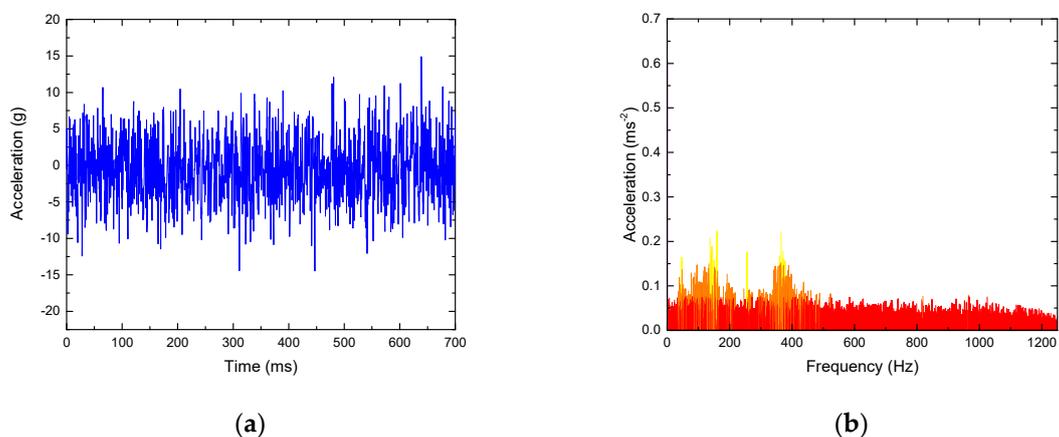


**Figure 12.** Rail surface fault.

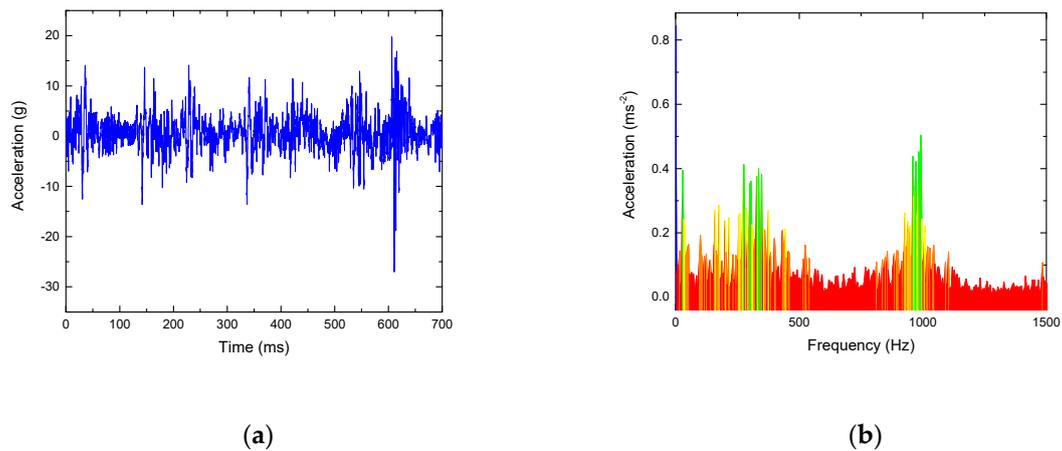
- Rail surface peeling damage

During train operation, due to the large braking, sliding friction is produced by the locked wheels on the rail surface, causing the local temperature to rise rapidly. When the rail cools, the raw surface of the hard and brittle metal gets damaged, resulting in the rail surface peeling phenomenon due to the large cyclical impact between the wheel and rail.

As shown in Figure 13, the signal waveform and frequency spectrum of a typical normal rail surface are mainly distributed in the low frequency part, usually less than 500 Hz. However, for the rail surface with peeling damage, the signal has an obvious distribution in the high frequency band, as shown in Figure 14, and the main frequency increases significantly at 1000 Hz. By calculation, its kurtosis value is 42.3 and approximate entropy is 0.332, which can effectively identify the rail faults.



**Figure 13.** Analysis of a normal rail vibration signal. (a) Waveform; (b) frequency spectrum.

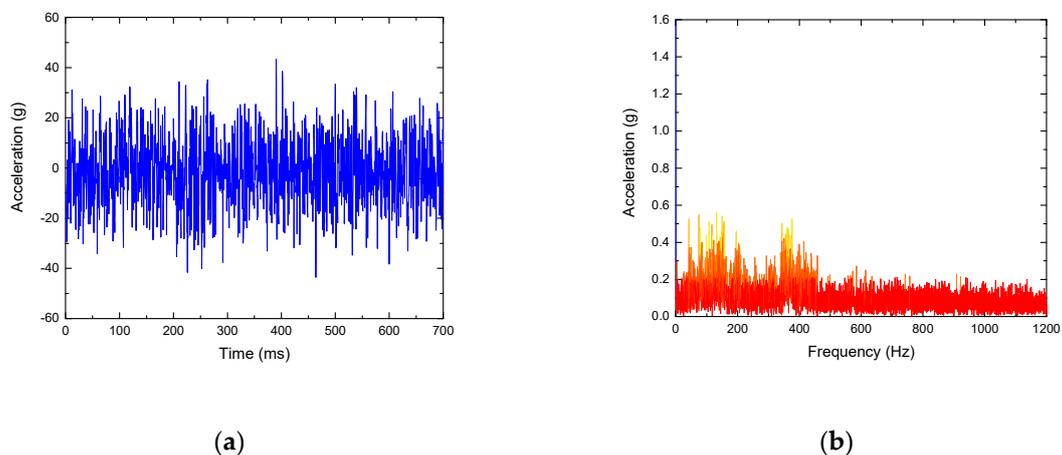


**Figure 14.** Spectrum analysis of a rail vibration signal with peeling damage. (a) Waveform; (b) frequency spectrum.

- Corrugation damage

Undulating wear is the roughness of the rail surface along the longitudinal direction, which appears most often in mountain areas. The causes are complex, involving rolling contact mechanics, material friction, vehicle coupling dynamics, and vibration load, vibration frequency, material characteristics, foundation, etc.

As shown in Figure 15, the signal under undulating wear usually generates abnormal vibration in the frequency region corresponding to the middle frequency band and the undulating frequency band. Since the excitation frequency is the same as the natural frequency of the rail, resonance happens, so the signal has a large amount of energy. Its kurtosis value is 32.8 and approximate entropy is 0.735.



**Figure 15.** Analysis of a rail vibration signal with corrugation damage. (a) Waveform; (b) frequency spectrum.

- Fault identification

A large number of experiments show that the fault signals' kurtosis and approximate entropy calculated from corresponding reconstructed signals are both too large. The approximate entropy of fault signals is usually greater than 0.4 and the kurtosis is usually greater than 20. In this experiment, 100 sets of sample signals are used for identification and classification. The classification results based on the SVM and the fault identification accuracy of the proposed method in the paper are shown in Figures 16 and 17, respectively. It can be seen that all signal types can be identified effectively

except one, which is taking the corrugation damage fault as a normal one. The reason may be the influence of the rail surface roughness. When the kurtosis and approximate entropy calculated from original signals are directly used for identification and classification, there appear eight classification errors totally; the reason may be the influence of surface roughness. Therefore this kind of direct classification accuracy is very low. Besides, three classification errors appear with EMD algorithm based on the traditional cubic Hermite interpolation and the piecewise cubic Hermite interpolation. So it can be concluded that the method presented in this paper improves classification accuracy obviously. Figure 18 shows the accuracy of feature extraction and SVM recognition proposed in this paper by the eight methods mentioned in Table 1, which are CS (Cubic Spline), CTCP (Cubic Trigonometric Cardinal Precise spline [24]), BS (B-Spline [22]), OS (Direct constrained optimization [28]), OPCH (Optimized Piecewise Cubic Hermite [23]), MPCI (Monotone Piecewise Cubic Interpolation [29]), PPFA (Piecewise Power Function Algorithm [30]). By comparison, it is found that the comprehensive accuracy of this algorithm is superior to other algorithms.

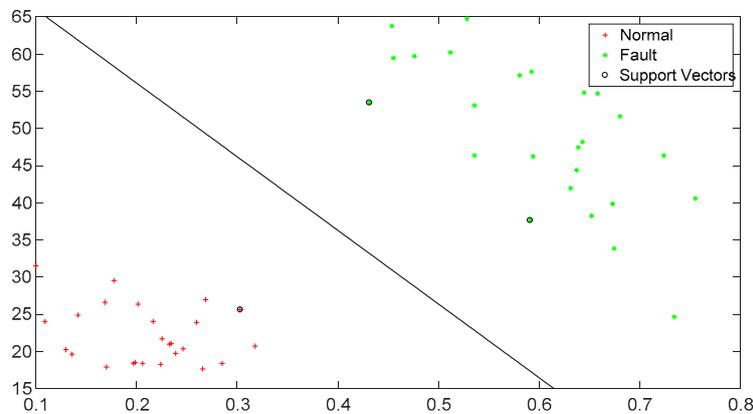


Figure 16. Classification result based on support vector machine (SVM) of the proposed method.

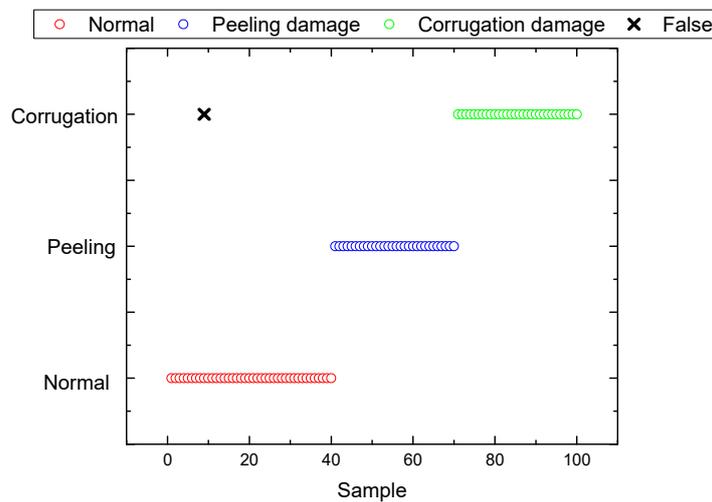


Figure 17. Identification accuracy of the proposed method.

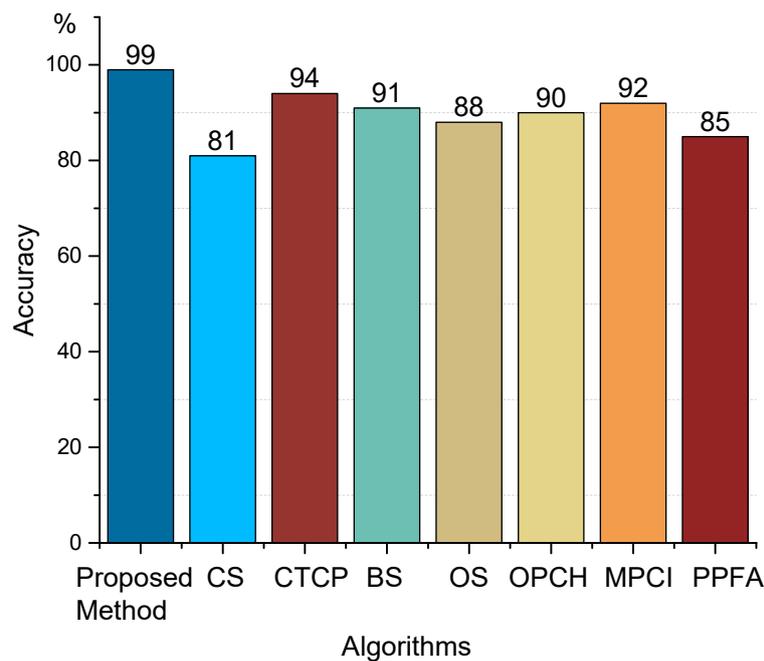


Figure 18. Identification accuracy of different algorithms.

## 6. Conclusions

Addressing the difficulties in rail compound fault diagnosis and the undershoot problem in vibration signal analysis based on the traditional EMD algorithms, an EMD algorithm based on the quartic  $C^2$  Hermite interpolation is presented in this paper. It replaces the traditional cubic Hermite interpolation algorithm and improves the performance of the EMD algorithm, and the undershoot problem in signal filtration based on the EMD algorithm is reduced effectively. As described, the principle components of the signal can be determined and reconstructed after SVD and cross correlation coefficient calculation. Then kurtosis and approximate entropy of the reconstructed signal are calculated as fault eigenvalues, which are used to classify the fault based on the SVM. Experimental results show that the EMD algorithm based on quartic  $C^2$  Hermite interpolation presented in this paper improves classification accuracy greatly.

**Author Contributions:** Conceptualization, H.L.; Methodology, H.L.; Software, C.Q.; Validation, C.Q., H.L.; Data Curation, C.Q.; Writing—Original Draft Preparation, C.Q.; Writing—Review & Editing, H.L.; Supervision, M.L.; Project Administration, H.L.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 51675259 and the Research Foundation of Nanjing Institute of Technology, grant number CKJA201804.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yin, K. Research on Automatic Railway Fastener Defect Detection System Based on Image Processing. Master's Thesis, Nanjing University of Posts and Telecommunications, Nanjing, China, 2018.
2. Li, H. *Research on Nondestructive Monitoring Methods of High Speed Railway Based on Vibration Signal*; Nanjing University of Aeronautics and Astronautics: Nanjing, China, 2013.
3. Xu, P.; Zhu, C.; Xu, Z.; Wang, P. Rail crack identification method based on differential Eddy current testing. *Nondestruct. Test.* **2018**, *40*, 7.
4. Liu, X. Research on Rail Surface Defect Recognition Based on Convolutional Neural Network. Master's Thesis, Southwest University of Science and Technology, Mianyang, China, 2018.
5. Kasai, N.; Takada, A.; Fukuoka, K.; Aiyama, H.; Hashimoto, M. Quantitative investigation of a standard test shim for magnetic particle testing. *Ndt E Int.* **2011**, *44*, 421–426. [[CrossRef](#)]

6. Kechagias-Stamatis, O.; Aouf, N.; Nam, D.; Belloni, C. Automatic x-ray image segmentation and clustering for threat detection. In Proceedings of the Target and Background Signatures III, Warsaw, Poland, 11–14 September 2017; p. 1043200.
7. Li, X.-B.; Hu, H.-W.; Yang, Y.; Ni, P.-J.; Yang, C. Automatic ultrasonic inspection of flaws in a propeller-blade. *Nondestruct. Test. Eval.* **2010**, *25*, 341–351. [[CrossRef](#)]
8. Cai, C.L.; Xiong, H.L.; Yang, L. Blind processing method for frequency hopping signals based on Winger-Ville rearrangement and stationary wavelet. *J. Dalian Marit. Univ.* **2008**, *34*, 115–118.
9. Zhang, Z.; Yi, W.; Wang, K. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *J. Intell. Manuf.* **2013**, *24*, 1213–1227. [[CrossRef](#)]
10. Staszewski, W.J.; Worden, K.; Tomlinson, G.R. Time–frequency analysis in gearbox fault detection using the wigner–ville distribution and pattern recognition. *Mech. Syst. Signal Process.* **1997**, *11*, 673–692. [[CrossRef](#)]
11. Choy, F.K.; Huang, S.; Zakrajsek, J.J.; Handschuh, R.F.; Townsend, D.P. Vibration Signature Analysis of a Faulted Gear Transmission System. *J. Propuls. Power* **2015**, *12*, 2657–2666. [[CrossRef](#)]
12. Huang, N.E.; Zheng, S.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Chi, C.T.; Liu, H.H. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proc. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
13. Xu, B.; Sheng, Y.; Li, P.; Cheng, Q.; Wu, J. Causes and classification of EMD mode mixing. In Proceedings of the Vibroengineering PROCEDIA, 36th International Conference on VIBROENGINEERING, Dubai, UAE, 15–17 March 2019; pp. 158–164. [[CrossRef](#)]
14. Hu, X.Y.; Peng, S.L.; Hwang, W.L. EMD Revisited: A New Understanding of the Envelope and Resolving the Mode-Mixing Problem in AM-FM Signals. *IEEE Trans. Signal Process.* **2012**, *60*, 1075–1086. [[CrossRef](#)]
15. Wu, Z.; Huang, N.E. Ensemble Empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2011**, *1*, 1–41. [[CrossRef](#)]
16. Yeh, J.R.; Shieh, J.S.; Huang, N.E. Complementary Ensemble Empirical Mode Decomposition: A Novel Noise Enhanced Data Analysis Method. *Adv. Adapt. Data Anal.* **2010**, *2*, 135–156. [[CrossRef](#)]
17. Torres, M.I.E.; Colominas, M.A.; Schlotthauer, G.N.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the IEEE International Conference on Acoustics, Prague, Czech Republic, 22–27 May 2011.
18. Alexander, P.; Manjurul, I.; Jaeyoung, K.; Jong-Myon, K. Rub-Impact Fault Diagnosis Using an Effective IMF Selection Technique in Ensemble Empirical Mode Decomposition and Hybrid Feature Models. *Sensors* **2018**, *18*, 2040. [[CrossRef](#)]
19. Chen, D.; Lin, J.; Li, Y. Modified complementary ensemble empirical mode decomposition and intrinsic mode functions evaluation index for high-speed train gearbox fault diagnosis. *J. Sound Vib.* **2018**, *424*, 192–207. [[CrossRef](#)]
20. Wu, W.-H.; Chen, C.-C.; Jhou, J.-W.; Lai, G. A Rapidly Convergent Empirical Mode Decomposition Method for Analyzing the Environmental Temperature Effects on Stay Cable Force. *Comput. Aided Civil Infrastruct. Eng.* **2018**, *33*, 672–690. [[CrossRef](#)]
21. Yang, J.; Li, P.; Yang, Y.; Xu, D. An improved EMD method for modal identification and a combined static-dynamic method for damage detection. *J. Sound Vib.* **2018**, *420*, 242–260. [[CrossRef](#)]
22. Yang, Y.; Miao, C.; Deng, J. An Analytical Expression for Empirical Mode Decomposition Based on B-Spline Interpolation. *Circuits Syst. Signal Process.* **2013**, *32*, 2899–2914. [[CrossRef](#)]
23. Zhu, W.; Zhao, H.; Chen, X. Improving empirical mode decomposition with an optimized piecewise cubic Hermite interpolation method. In Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI2012), Yantai, China, 19–20 May 2012; pp. 1698–1701.
24. Di, Z.; Huang, Z.; Li, H. An improved EEMD method based on the adjustable cubic trigonometric cardinal spline interpolation. *Digit. Signal Process.* **2017**, *64*, 41–48.
25. Wall, M.E.; Rechtsteiner, A.; Rocha, L.M. *Singular Value Decomposition and Principal Component Analysis*; Springer: Boston, MA, USA, 2002.
26. Suykens, J.A.K.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [[CrossRef](#)]
27. Li, J.; Liu, C.; Yang, L. Quartic Hermite interpolating splines with parameters. *J. Comput. Appl.* **2012**, *32*, 1868–1870.

28. Oberlin, T.; Meignen, S.; Perrier, V. An Alternative Formulation for the Empirical Mode Decomposition. *IEEE Trans. Signal Process.* **2012**, *60*, 2236–2246. [[CrossRef](#)]
29. Yang, L.; Yang, Z.; Yang, L. The theoretical analysis for an iterative envelope algorithm. *Digit. Signal Process.* **2015**, *38*, 32–42. [[CrossRef](#)]
30. Qin, S.R.; Zhong, Y.M. A new envelope algorithm of Hilbert–Huang Transform. *Mech. Syst. Signal Process.* **2006**, *20*, 1941–1952. [[CrossRef](#)]
31. Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process.* **2006**, *20*, 282–307. [[CrossRef](#)]
32. Núñez, J.A.; Cincotta, P.M.; Wachlin, F.C. Information entropy. *Celest. Mech. Dyn. Astron.* **1996**, *64*, 43–53. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).