

Article

Rolling Element Bearing Fault Diagnosis under Impulsive Noise Environment Based on Cyclic Correntropy Spectrum

Xuejun Zhao ^{1,2}, Yong Qin ^{1,2,3,4,*}, Changbo He ⁵, Limin Jia ^{1,2,3,4} and Linlin Kou ^{1,2}

¹ State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China; xuejunzhao0@gmail.com (X.Z.); lmjia@bjtu.edu.cn (L.J.); koull01@163.com (L.K.)

² School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

³ National Engineering Laboratory for System Safety and Operation Assurance of Urban Rail Transit, Guangzhou 510000, China

⁴ Beijing Research Center of Urban Traffic Information Sensing and Service Technologies, Beijing Jiaotong University, Beijing 100044, China

⁵ School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China; strange@mail.dlut.edu.cn

* Correspondence: yqin@bjtu.edu.cn; Tel.: +86-186-1010-5860

Received: 7 December 2018; Accepted: 4 January 2019; Published: 10 January 2019



Abstract: Rolling element bearings are widely used in various industrial machines. Fault diagnosis of rolling element bearings is a necessary tool to prevent any unexpected accidents and improve industrial efficiency. Although proved to be a powerful method in detecting the resonance band excited by faults, the spectral kurtosis (SK) exposes an obvious weakness in the case of impulsive background noise. To well process the bearing fault signal in the presence of impulsive noise, this paper proposes a fault diagnosis method based on the cyclic correntropy (CCE) function and its spectrum. Furthermore, an important parameter of CCE function, namely kernel size, is analyzed to emphasize its critical influence on the fault diagnosis performance. Finally, comparisons with the SK-based Fast Kurtogram are conducted to highlight the superiority of the proposed method. The experimental results show that the proposed method not only largely suppresses the impulsive noise, but also has a robust self-adaptation ability. The application of the proposed method is validated on a simulated signal and real data, including rolling element bearing data of a train axle.

Keywords: fault diagnosis; cyclostationary; kernel method; correntropy; impulsive noise

1. Introduction

Rolling element bearings are one of the most widely used mechanical components in various industrial machines, such as gearboxes, railway axles, and turbines. Their health states tend to degrade due to repeating rotations under harsh working conditions. Fault diagnosis of rolling element bearings is essential to prevent machine breakdown and ensure production efficiency. Therefore, research related to bearing degradation prognostics and fault diagnosis has recently attracted much attention [1–3].

The narrowband-based amplitude demodulation of a vibration signal enables the extraction of more detailed information about the fault signature. The procedure of this method uses a band-pass filter to retain one of the resonant frequency bands. Hilbert transform is then employed to demodulate the signal preprocessed by the band-pass filter and to construct a squared envelope on the filtered signal [4]. At last, the squared envelope spectrum is utilized to identify bearing fault frequencies. Thus, the key of the narrowband-based amplitude demodulation is to find a more informative resonant frequency band for Hilbert demodulation. Much effort has been made to address this issue. Spectral

kurtosis (SK), also known as Kurtogram [5,6], proposed by Antoni, is one of the most pioneering works in the field. Kurtogram uses kurtosis to quantify an analytical signal obtained by a band-pass filter as well as Hilbert transform so as to characterize the impulsiveness of bearing fault signals. Furthermore, an SK-based algorithm named Fast Kurtogram [7] was developed to make it a tool with potential online industrial applications. Fast Kurtogram returns the complex envelopes of the signal in selected frequency bands with an arborescent multi-rate filter-bank structure, thus reducing the computation complexity. This SK-based algorithm has inspired many related works on the fault diagnosis of rotating machines [8–10].

However, vibration signals of bearings contain impulsive noise in some industrial situations, which bring challenges to further applications of this technology. Moreover, the Kurtogram algorithm was found to be sensitive to impulsive noise and this can sometimes lead to incorrect interpretations [11].

A number of improved methods have been proposed in order to analyze and solve this problem, such as the smoothness index-guided approach [12], the sparsity index-guided approach [13], and the Protrugram analysis [14]. Among these works, the Protrugram analysis proposed by Barszcz and Jabłoński has gained the most attention. The Protrugram uses the kurtosis of the envelope spectrum instead of the kurtosis of the filtered time signal to characterize the bearing fault. Despite its unique contribution, the Protrugram has the dual limitations of the spectral Kurtogram and it raises the problem of selecting an optimal mathematical expression in its definition among many possible choices [15].

Motivated by ideas from the field of thermodynamics, a new concept named Infogram was proposed [15], which involves measuring the squared envelope, the squared envelope spectrum, and their average as estimators to select the fault frequency band. It has been proved that the method is able to detect the impulsive and cyclostationary signature in both time and frequency domains. This work stresses the importance of the suppression of impulsive noise; meanwhile, it has the potential to deal with impulsive noise based on cyclostationary modeling.

Cyclostationarity is a property that characterizes stochastic processes whose statistical properties periodically vary with respect to some generic variable [16]. This generality makes cyclostationarity nicely suited to rotating machine signals because of hidden periodicities in its structure. Based on this cyclostationary modeling theory, cyclic spectral analysis, also known as spectral correlation (SC) analysis, has been developed as a reliable tool for fault diagnosis of rolling element bearings [17,18].

As a second-order cyclostationary component extraction indicator, SC analysis has become popular in the field of machine diagnostics. SC analysis [17] can simultaneously describe the carrier and modulation for bearing fault diagnosis by means of a bi-spectral map. Here, the carrier refers to the spectral band of frequencies and the modulation refers to the cyclic frequency. In other words, the cyclic frequency describes the periodicity of the fault impacts, while the spectral band of frequencies corresponds to the resonance characteristic of the mechanical system. Thus, the fault frequency distribution of the signal can be identified over the resonance frequency bands. However, SC may be extremely costly to compute in some situations. To solve this problem, alternative estimators have been proposed, such as averaged cyclic periodogram [17], cyclic modulation spectrum [19], and fast estimator of SC [20]. These estimators greatly relieve SC from computation burdens and contribute to SC analysis for bearing fault diagnosis.

Recently, a new cyclostationary analysis technology named cyclic correntropy (CCE) analysis has emerged to suppress impulsive noise [21–23]. CCE is a kernel-based similarity measure of cyclostationary modeling signals. Related research has shown that it has a good suppression performance when dealing with the binary phase-shift keying signal under an impulsive noise environment [21,23]. However, to the best of our knowledge, only several works related to CCE analysis have been published in the field of communication and its applications need to be developed and reinforced. Furthermore, a CCE analysis of vibration signal processing has never been explored before. Therefore, for the first time, this paper introduces CCE into the analysis of bearing vibration

signal processing. Furthermore, this paper also investigates the influence of kernel size on the CCE performance, which is critical in the CCE analysis. Thus, this is proven to be a promising fault diagnosis method for bearings as well as other rotating machinery in the presence of impulsive noise.

The rest of this paper is outlined as follows. In Section 2, the fundamentals of correntropy function and the cyclic spectral analysis of rotating machine vibration signals are reviewed. In Section 3, cyclostationary analysis based on the correntropy function and its spectrum for bearing fault diagnosis is proposed. In Section 4, simulated and real bearing fault signals including a vibration dataset acquired from industrial railway axle bearings are used to verify the effectiveness of the proposed method. Moreover, the SK-based Fast Kurtogram method is conducted to highlight the superiority of the proposed method. Conclusions are drawn in the final section.

The intended contributions of this study can be summarized as follows:

1. A cyclostationary analysis method based on the correntropy function is introduced into the fault diagnosis of the rolling element bearing under an impulsive noise environment.
2. The kernel size of the correntropy function is investigated to find out its influence on the rolling element bearing fault diagnosis.
3. The diagnosis performance of the proposed method is compared with the Fast Kurtogram method using train axle bearing data to verify its suitability.

2. Fundamentals of Correntropy Function and Cyclic Spectral Analysis

2.1. Correntropy Function

Correntropy function can be regarded a similarity measure based on kernel function. The definition of correntropy function is defined as follows [24]:

Let $\{x_t, t \in T\}$ be a stochastic process with T being an index set and $x_t \in R^d$. The correntropy function $V(t_1, t_2)$ is defined as a function from $T * T$ into R^+ given by Equation (1):

$$V(t_1, t_2) = E[\kappa(x_{t_1} - x_{t_2})] \quad (1)$$

where $E[\cdot]$ denotes the mathematical expectation over the stochastic process x_t . $\kappa(x_{t_1} - x_{t_2})$ corresponds to the positive-definite function which satisfies the Mercer condition [25]. The Gaussian kernel is usually chosen as the Mercer kernel function due to its smoothness and strict positive-definiteness [26]. The kernel function is shown in Equation (2):

$$\kappa(x - y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (2)$$

where σ is the kernel size. The correntropy function is dependent upon the kernel size and it is selected according to certain applications [27].

By applying an extension of the Taylor series to the correntropy function, Equation (1) can be rewritten as follows [24]:

$$V(t_1, t_2) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[\|x_{t_1} - x_{t_2}\|^{2n}] \quad (3)$$

which involves all the even-order moments of the random variable $\|x_{t_1} - x_{t_2}\|$. Specifically, the term corresponding to $n = 1$ in Equation (3) is proportional to:

$$E[\|x_{t_1}\|^2] + E[\|x_{t_2}\|^2] - 2E[\langle x_{t_1}, x_{t_2} \rangle] = \sigma_{x_{t_1}}^2 + \sigma_{x_{t_2}}^2 - 2R_x(t_1, t_2) \quad (4)$$

where $R_x(t_1, t_2)$ is the covariance function of the random process. Thus, the information provided by the traditional correlation function is included within the new function. It can be seen that

the correntropy function is a second-order statistic of the mapped feature space data. Besides, the correntropy function incorporates higher-order moments of the random variable $\|x_{t_1} - x_{t_2}\|$ by adjusting the values of the kernel size.

Thanks to these attractive properties, the correntropy function has been widely applied in machine learning and signal processing. Among those applications, the suppression performance of non-Gaussian noise has been highly researched [28–30], in which the Gaussian kernel function plays an important part. The superiority of the Gaussian kernel function mainly includes two points: transforming values of outliers into zero and extracting higher statistical moments. These two merits enable the correntropy function to handle signals corrupted by non-Gaussian noise, especially impulsive noise.

2.2. Cyclic Spectral Analysis

A cyclostationary process is a stochastic process that exhibits some hidden periodicities in its structure [16]. Denote a bearing fault signal as $x(t_n)$ and its stream as $x[n]$, where $t_n = n/F_s, n = 0, 1, \dots, L - 1$ indicate time instants with sampling frequency F_s . Assume that the bearing fault signal is cyclostationary on the second order, which indicates that the instantaneous autocorrelation function is periodic with period T [20]:

$$R_x(t_n, \tau) = E\{x(t_n)x(t_n - \tau)^*\} = R_x(t_n + T, \tau) \quad (5)$$

where $E\{\cdot\}$ refers to the ensemble average operator; $\tau = m/F_s$; and $*$ is the complex conjugate operator.

One important tool for the cyclic spectral analysis of bearing signals is SC analysis, which is obtained in the form of a bi-spectral map, thus reflecting the whole picture of spectral frequency f and cyclic frequency α in the signal. Based on Equation (5), SC is defined as follows [20]:

$$S_x(\alpha, f) = \lim_{N \rightarrow \infty} \frac{1}{(2N + 1)F_s} \sum_{n=-N}^N \sum_{m=-\infty}^{\infty} R_x(t_n, \tau_m) \exp\left(-j2\pi n \frac{\alpha}{F_s}\right) \exp\left(-j2\pi m \frac{f}{F_s}\right) \quad (6)$$

In the case of the second-order cyclostationary signal, SC is continuous at spectral frequency f and discrete at cyclic frequency α . It can be rewritten as follows:

$$S_x(\alpha, f) = \begin{cases} S_x^k(f), & \alpha = k/T \\ 0, & \alpha \neq k/T \end{cases} \quad (7)$$

where $S_x^k(f), k = 0, \pm 1, \pm 2, \dots$ are cyclic spectra.

The alignment composed of non-zero values at a certain cyclic frequency α demonstrates the existence of a sinusoidal modulation in the signal, which also suggests the presence of fault signature during the fault diagnosis process. Estimators of SC analysis, such as averaged cyclic periodogram, cyclic modulation spectrum, and fast estimator of SC, are proposed and discussed, contributing to the practical applications of SC analysis.

3. Cyclostationary Analysis Based on Correntropy Function

3.1. Fundamental of CCE Function and CCE Spectrum

Similar to formulations of the cyclostationary process of the first and second order, we can obtain the basic expression of the CCE. Let $V_x(t, \tau)$ denote the correntropy function for a stochastic process $x(t)$ that exhibits hidden periodicities in its structure and for which the time shift is τ . Assume the correntropy function is periodic with T_0 , then:

$$V_x(t + T_0, \tau) = V_x(t, \tau). \quad (8)$$

Furthermore, represent $V_x(t, \tau)$ by Fourier series, as shown as Equation (9):

$$V_x(t, \tau) = \sum_{\alpha} V_x^{\alpha}(\tau) e^{j2\pi\alpha t} \quad (9)$$

where $\alpha = n/T$, $n \in Z$ is taken as the cyclic frequency associated with the fault characteristic.

Let us define the CCE function for $x(t)$ as Fourier coefficients $V_x^{\alpha}(\tau)$, computed by:

$$V_x^{\alpha} = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} V_x(t, \tau) e^{-j2\pi\alpha t} dt. \quad (10)$$

According to those formulations, they are functions of time lag τ and are indexed by the cyclic frequency α . Note that for $\alpha = 0$, the CCE function returns to the conventional correntropy function.

Combined with Equation (2), the CCE function can be rewritten as [23]:

$$V_x^{\alpha} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \kappa_{\sigma}(x(t), x(t + \tau)) e^{-j2\pi\alpha t} dt. \quad (11)$$

The CCE function has proved to be very useful in signal processing, although its applications are limited to communication engineering. It is often more effective and natural to transform the structure of a cyclostationary signal in the frequency domain. Note that the CCE is a function with variables t and τ , its frequency domain is a two-dimensional (2D) Fourier transform with two frequency variables α and f , and CCE spectrum (CCES) is defined as a Fourier transform of CCE as follows:

$$S_x^{\alpha}(f) = \int_{-\infty}^{+\infty} V_x^{\alpha}(\tau) e^{-j2\pi f \tau} d\tau. \quad (12)$$

$S_x^{\alpha}(f)$ displays the power distribution of the cyclostationary signal with respect to both the spectral frequency f , which is associated with the system resonance frequency, and the cyclic frequency α , which is also known as the fault frequency.

3.2. Kernel Size Selection of the CCE Function

The Gaussian variance, also known as kernel size, is a free parameter which must be chosen according to certain applications. A well-tuned kernel size can minimize the effect of the noise interruption. Therefore, the diagnosis performance of cyclic correntropy is largely determined by the kernel size. However, in previous studies related to cyclic correntropy, the selection method of kernel size was not addressed in detail [21–23] and kernel size mainly was decided by multiple trials. Thus, it is necessary to find a solution for the kernel size selection.

The correntropy function is an extended work of the information theoretic learning (ITL) theory [31,32]. Thanks to the work of Robert et al. [33], the link between the inner product in a kernel feature and the ITL cost function was discovered. Robert et al. claimed that the ITL cost functions, when estimated by the Parzen method, can be expressed in terms of inner products in a kernel feature space defined by a Mercer kernel. This link offers a criterion for the selection of the Mercer kernel size based on density estimation.

Among those kernel size estimation methods [34–36], Silverman's rule is a classical method because of its low computation cost and robust performance [37]. Therefore, this paper selects the optimal parameter according to Silverman's rule, one of the most widely used kernel estimation methods. Silverman's rule is shown here as Equation (13):

$$\sigma = 0.9AN^{-1/5} \quad (13)$$

where N is the data length. A stands for the minimum of the empirical standard deviation of data and the data interquartile range scaled by 1.34.

The influence of the kernel size on the bearing fault diagnosis result is discussed more in greater depth in Section 4 with detailed examples.

3.3. Estimation of CCES

Based on the introduction of above two sections, the estimation of CCES is shown as follows:

Step 1. Denote the input signal $x[n]$, with signal length n . Then divide the input signal into L blocks, with each block of N samples.

Step 2. Calculate the kernel size σ_l of each block with Silverman's rule, $l = 0, 1, 2, \dots, L - 1$.

Step 3. Calculate the average of the correntropy function for each block $l = 0, 1, 2, \dots, L - 1$:

$$M_l = \frac{1}{N^2} \sum_{\tau_n=0}^{N-1} \sum_{n=0}^{N-1} G_{\sigma_l}(x_l[n], x_l[n + \tau_n]). \quad (14)$$

Step 4. Calculate the following formulation for each block of size N with $\alpha[n] = n/N$, $n = 0, 1, 2, \dots, N - 1$, $l = 0, 1, 2, \dots, L - 1$:

$$V_l^{\alpha_n}[\tau_n] = \sum_{n=0}^{N-1} \left\{ [G_{\sigma_l}(x_l[n], x_l[n + \tau_n]) - M_l] e^{-j2\pi\alpha_n n} \right\}. \quad (15)$$

Step 5. Calculate the mean value $V_l^{\alpha_n}[\tau]$ of all L blocks:

$$V^{\alpha_n}[\tau_n] = \frac{1}{L} \sum_{l=0}^{L-1} V_l^{\alpha_n}[\tau_n]. \quad (16)$$

Step 6. Calculate the discrete Fourier transform for each τ_n :

$$T^{\alpha_n}[k] = \left| \frac{1}{N} \sum_{\tau_n=0}^{N-1} V^{\alpha_n}[\tau_n] e^{-j\frac{2\pi}{N}k\tau_n} \right|. \quad (17)$$

Step 7. To clearly observe the fault frequency, project the CCES into the cyclic frequency domain and obtain the cyclic domain profile.

4. Validation of CCES on Bearing Fault Diagnosis

4.1. Estimation of CCES

The following example aims to illustrate the failure of the Fast Kurtogram in the presence of impulsive noise. Firstly, the simulated bearing fault signal is modeled as a single-degree-of-freedom system according to Reference [38]:

$$y(k) = \sum_r \exp(-\alpha * (k - r * F_s / f_m - \tau_r) / F_s) * \sin(2\pi f * (k - r * F_s / f_m - \tau_r) / F_s) \quad (18)$$

where α is equal to 900; f_m is the fault frequency, which is set to 100 Hz; and F_s is the sampling frequency, which is set to 10,000 Hz. f is the resonant frequency, which is equal to 1000. τ_r is subject to discrete uniform distribution, and is thus used to simulate the randomness caused by roller slippage.

Impulsive noise has a similar behavior to the bearing fault signal. According to the impulsive noise simulation method proposed by Antoni [15], the noise is similarly modeled as a single-degree-of-freedom system with different parameter sets. Typically, α is equal to 300, while f_m is set to 30 Hz. The resonant frequency is equal to 3000. Note that the impulsive noise is not composed of a single transient but rather multiple noisy transients in the time domain. The final synthetic signal of length $L = 10^4$ is displayed in Figure 1. Figure 1a shows that impulsive noise is distributed around sample points 1000, 1334, and so on, whose amplitudes are obviously larger than normal transients

of the simulated fault signal. Figure 1b displays the same signal after the addition of white Gaussian noise with signal to noise ratio (SNR) equalling -6 dB.

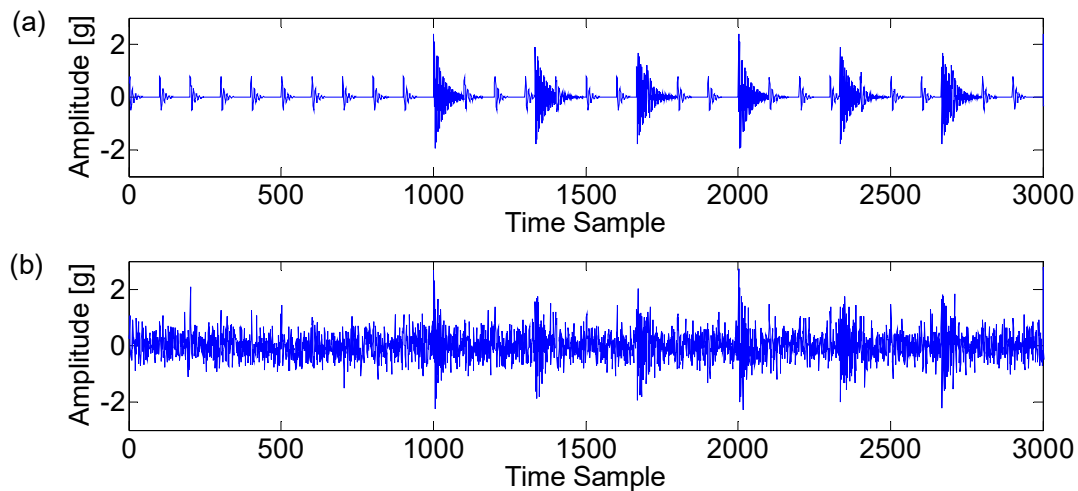


Figure 1. (a) Synthetic signal simulating multiple transients produced by faulty bearing and further interrupted by impulsive noise; (b) the same signal with additive white Gaussian noise (SNR = -6 dB).

The Fast Kurtogram method [7] is applied to select an optimal band for the square envelope analysis and the corresponding Kurtogram is displayed in the Figure 2. It can be seen that the impulsive noise around 3000 Hz dominates the Kurtogram at level 4.5, while the fault signature is masked in the noise. Furthermore, the frequency band centered at 3020 Hz is obtained by the band filter and the corresponding envelope and amplitude spectrum are displayed in Figure 3. Figure 3b shows that the filtered signal band is actually the impulsive noise component, whose frequency is set to be 30 Hz.

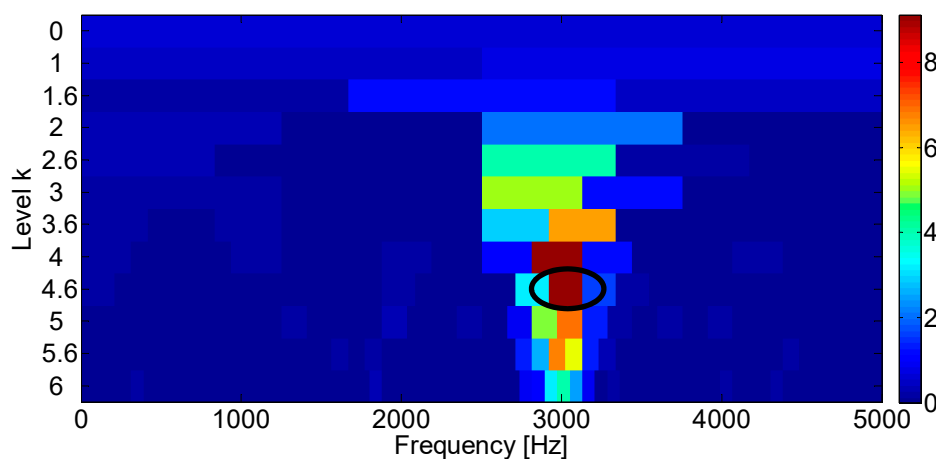


Figure 2. Fast Kurtogram of the simulated bearing fault signal.

Then, the simulated signal is further analyzed with the CCES proposed in Section 3.3. To observe the fault frequency in greater detail, the CCES is projected onto the cyclic domain. According to Silverman's rule, σ is equal to 0.107. The cyclic domain profile when σ is equal to 0.107 is displayed in Figure 4. It can be seen from the figure that the fault frequency is easily identified, marked with red arrows. The spectrum also displays the frequency component of 30 Hz, marked with a red circle. Compared to the fault frequency at 100 Hz and its harmonics, the impulsive noise is largely suppressed. Thus, the analysis results of the simulated signal under an impulsive environment demonstrates that the proposed method can detect the fault frequency effectively.

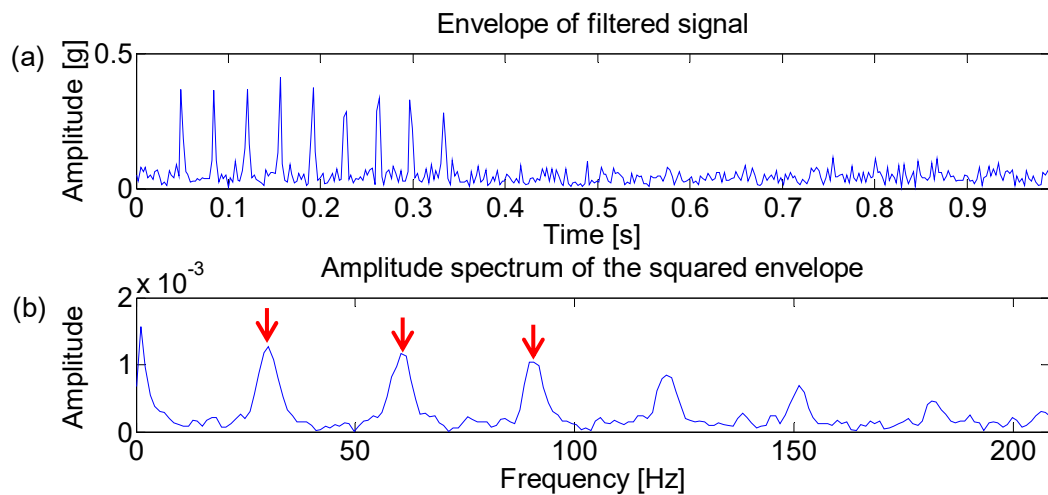


Figure 3. (a) Envelope of the filtered signal which maximizes the Kurtogram; (b) amplitude spectrum of the squared envelope.

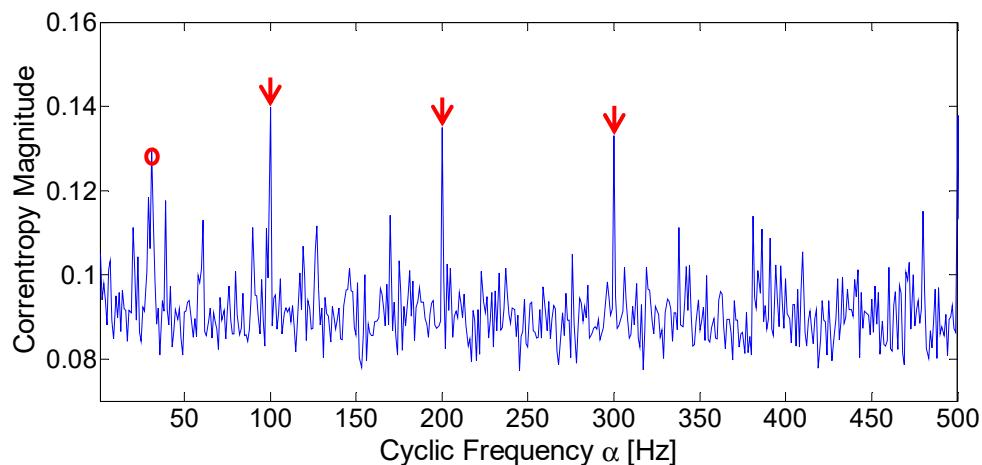


Figure 4. Cyclic domain profile of the CCES for simulated signal.

4.2. Case Study 1

To further testify the performance of the proposed method, comparison experiments are conducted on real bearing data. The first example applies the dataset from the Western Reserve University (WRU) bearing data center [39]. This dataset is free from impulsive noise and can be used to test the performance of the proposed method in the presence of common background noise. The data of inner race faults numbered 169 are chosen for analysis. According to formulas for bearing fault characteristic frequency [40], the inner race fault frequency is 162.185 Hz.

Firstly, the Fast Kurtogram is obtained for the optimal band selection, as shown in Figure 5. It can be seen from the figure that the optimal band is located at level 6 and the frequency center is 2750 Hz, marked with the black circle. This optimal band acquired through the bandpass filter and squared envelope spectrum is displayed in Figure 6. The fault frequency and its harmonics are marked with red arrows. Figure 6b shows that the fault signature is diagnosable but shows some discrete components interrupting the frequency domain.

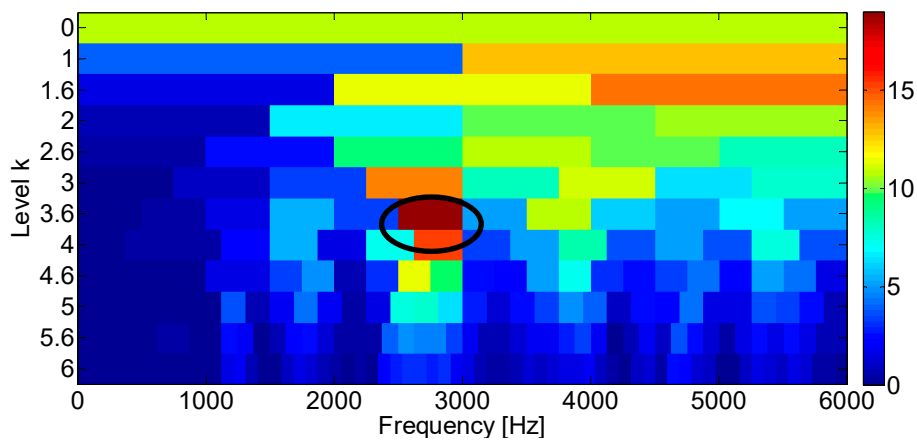


Figure 5. Fast Kurtogram of inner race fault signals numbered 169.

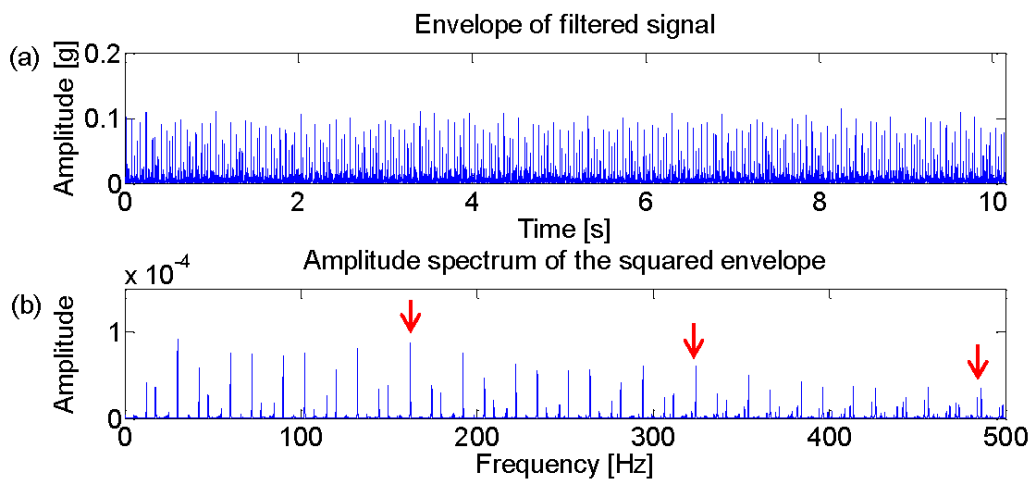


Figure 6. (a) Envelope of the filtered signal which maximizes the Kurtogram; (b) amplitude spectrum of the squared envelope.

Then, with σ set equal to 0.015, the cyclic domain profile of the corresponding signal is displayed in Figure 7. Compared to the squared envelope of the filtered signal based on Fast Kurtogram, the fault frequency and its harmonics are easier to find. The result demonstrates that the proposed method can still obtain a better diagnosis performance compared to the Fast Kurtogram in an environment with common background noise.

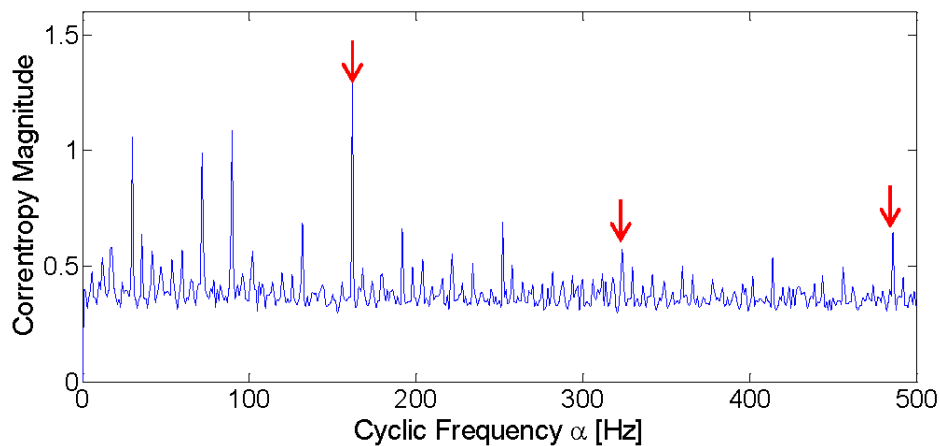


Figure 7. Cyclic domain profile of the CCES for Western Reserve University (WRU) bearing data numbered 169.

4.3. Case Study 2

In this case study, industrial railway axle bearing fault data are used for further comparison experiments. One unique component of the industrial railway axle bearing signal is the impulsive background noise. During the train operation on the rail, impulsive force will be produced when the train passes through curves or small gaps between rail joints. This impulsive force will be conducted to the bearing casing through the train wheel and creates an impulsive noise environment. This phenomenon raises new challenges to the bearing fault diagnosis. Focusing on this issue, railway bearing data under impulsive noise interference are acquired to validate the superiority of the proposed method.

Our experimental platform for collecting railway axle bearing fault data is shown in Figure 8. Through a transmission set, a variable speed DC motor with a speed up to 1480 r/min is used to drive the rotation of an axle at different speeds. Axle bearings are assembled at the ends of the axle. A lateral load set and a vertical load set are installed to impose practical loads during rail vehicle operation. Note that the lateral load is used to simulate multiple impacts when the train passes curves or small gaps between rail joints. Fan motors are installed to simulate the effect of natural wind in the opposite direction of the vehicle's momentum. Sensors are mounted at 12 o'clock (directly in the vertical load zone) and 3 o'clock (orthogonal to the vertical load zone) of the bearing casing to acquire vibration data. Two fault bearings are selected from the railway maintenance center and their degradation conditions are shown in Figure 8b,c respectively. The sampling frequency is set to 12,800 Hz. The simulated speed and vertical load are set to 150 km/h and 272 kN, respectively. The lateral load for generating the impulsive noise is 20 kN. According to the transmission ratio of our experimental platform, the inner race fault frequency and the outer race fault frequency are calculated as 164 Hz and 120 Hz, respectively.

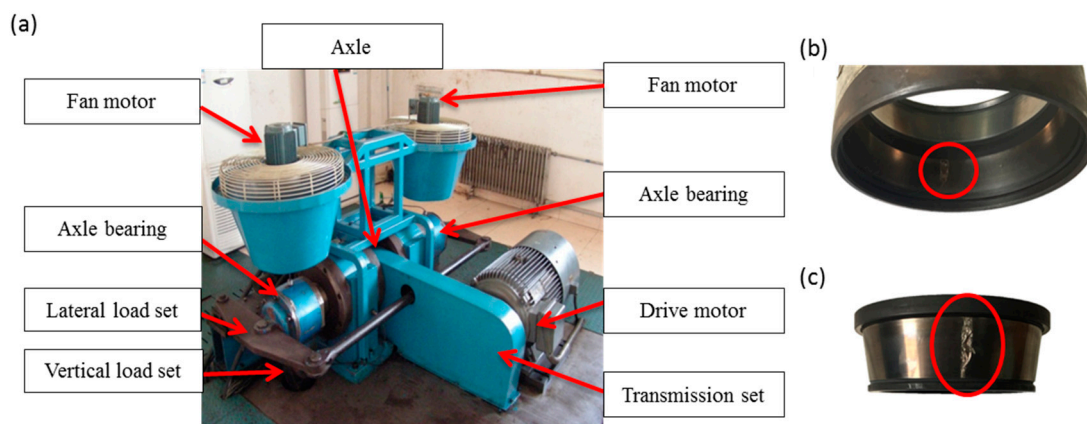


Figure 8. A designed experimental platform and industrial railway axle bearing faults: (a) the designed experimental platform; (b) the outer race fault; (c) the inner race fault.

The input data are firstly analyzed with the Fast Kurtogram method and Figure 9 displays the optimal band marked with the black circle. It can be seen from the figure that the optimal band is obtained at level 3, for which the central frequency is 6000 Hz. The associated squared envelope spectrum is further displayed in Figure 10 based on the optimal band. It is found that the Fast Kurtogram fails to find an informative spectral frequency band relevant to the fault signature for further squared envelope spectrum analysis. In Figure 10b, the frequency components are mainly impulsive noise produced by the lateral load. Therefore, it should be noted that the frequency bands associated with the several largest kurtosis values are not enough to establish a correct spectral frequency range for fault diagnosis of the industrial railway axle bearing.

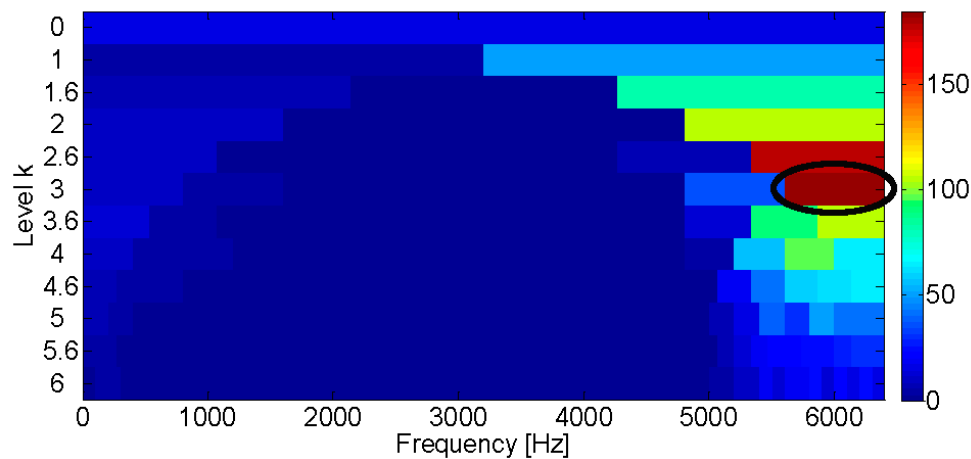


Figure 9. Fast Kurtogram of the axle bearing signal of the inner race fault.

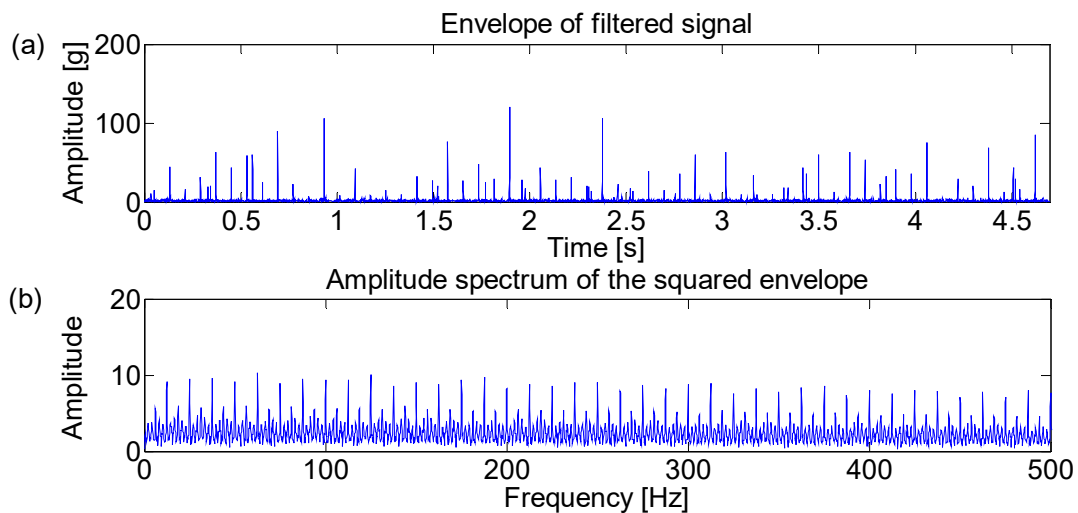


Figure 10. (a) Envelope of the filtered signal which maximizes the Kurtogram; (b) amplitude spectrum of the squared envelope.

Then, the CCES is applied to analyze the input data of the inner race fault and the corresponding cyclic domain profile is displayed in Figure 11. According to Silverman's rule, the σ is set to 4.496 in this experiment. The highlighted frequencies marked with red arrows are 161 Hz, 329 Hz, and 491 Hz. One thing that should be noted is that there is a minor difference between the theoretic fault frequency and the obtained frequency. This is due to random slips between rolling elements and the inner race caused by the large vertical load [41,42]. Furthermore, irregular rotations of the faulty bearing caused by the high speed and vertical load may also have some influence on the fault frequency calculation.

The previous procedure is applied to analyze the axle outer race fault signal and the relevant results are plotted in Figures 12 and 13, respectively. The center frequency of the optimal band is located at level 3 with 6000 Hz. Furthermore, the squared envelope analysis based on the optimal band is displayed in Figure 13. Unfortunately, the fault signature seems to be buried in the impulsive noise again.

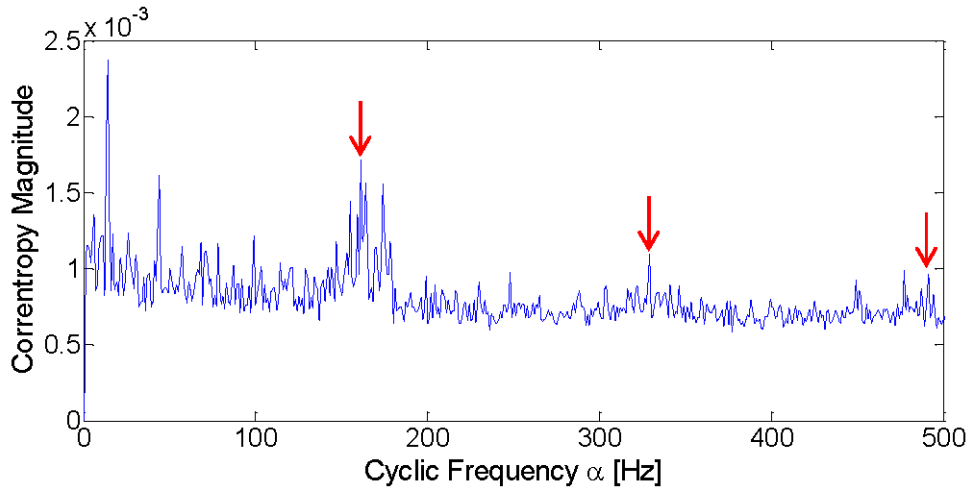


Figure 11. Cyclic domain profile of the CCES for the railway axle bearing signal of the inner race fault.

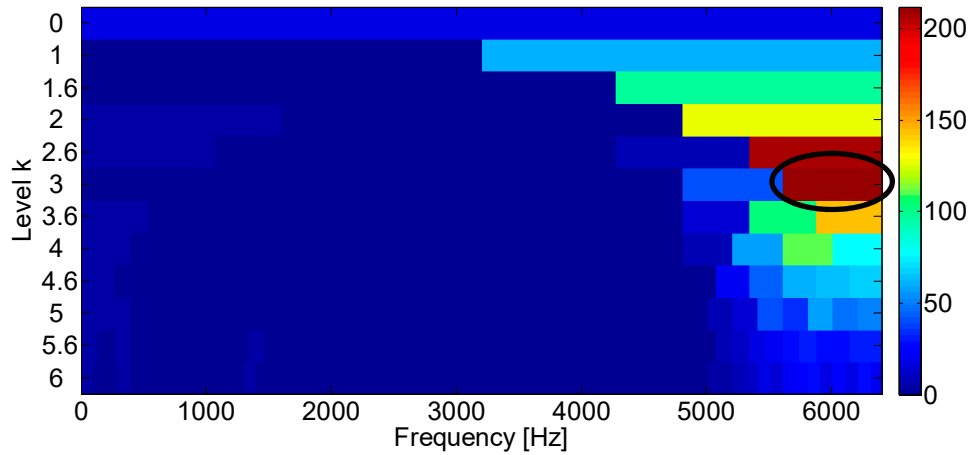


Figure 12. Fast Kurtogram of the axle bearing signal of the outer race fault.

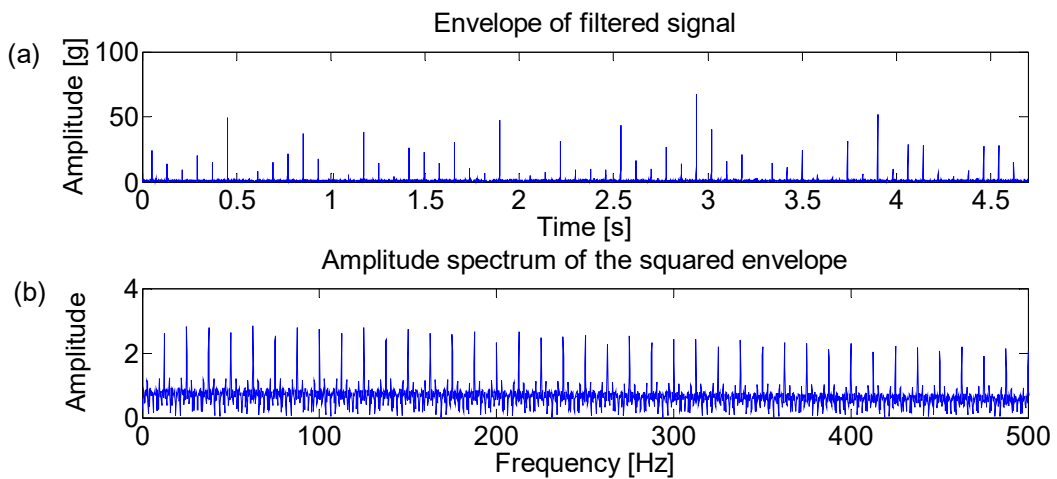


Figure 13. (a) Envelope of the filtered signal which maximizes the Kurtogram; (b) amplitude spectrum of the squared envelope.

The application of the CCES method with σ equal to 3.021 and the cyclic domain profile of outer race fault signal is shown in Figure 14. The highlighted frequencies marked with red arrows are 120 Hz, 241 Hz, and 358 Hz, which are almost the same as the theoretic fault frequency and its harmonics.

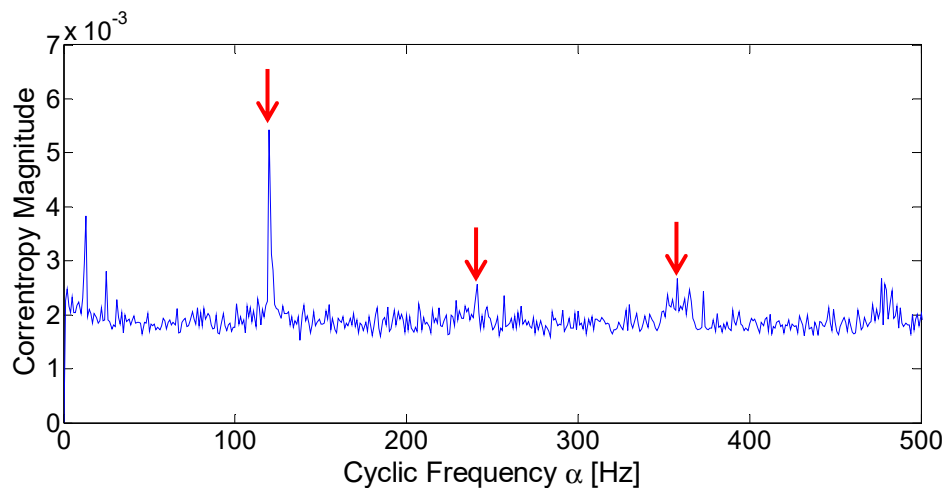


Figure 14. Cyclic domain profile of the CCES for the railway axle bearing signal of the outer race fault.

An interesting point is that the CCES can obtain the cyclic frequency of the fault signature while suppressing the periodic noise, which also is achieved in the simulation signal analysis. Generally speaking, cyclostationary modeling aims to determine all those cyclostationary components in the signal. According to the CCES procedure proposed in Section 3, it is found that this method can easily deal with impulsive noise without any cyclostationarity. When the impulsive noise is cyclostationary, which can also be considered periodic in time domain, the CCES can also detect this component and the cyclic domain profile will show this cyclic frequency in the spectrum plot. However, this frequency component will not dominate the frequency spectrum because this phenomenon only happens when impulsive noise is distributed all along the whole signal and the impulsive amplitude is large enough, which rarely occurs during train operation or in other industries.

4.4. Influence of Kernel Size on the Diagnosis Performance

The Gaussian kernel size is a free parameter selected according to certain applications. The variation of kernel size expands its application range, while improper kernel size may become its potential weakness. Previous studies on the kernel size selection of CCE were mainly based on multiple trials, which does not lead to the optimal parameter for the Gaussian kernel function. In this section, fault diagnosis performance of different kernel size is analyzed to provide an effective mechanism for kernel size selection.

Input data are derived from the above two case studies for comparison, including the WRU data of the inner race fault numbered 169 and the axle bearing data of the outer race fault. The optimal kernel size of each experiment is based on Silverman's rule. In addition, another random kernel size is used for comparisons.

The random kernel size is set to 1 and the optimal kernel size is 0.015 according to Silverman's rule. The analysis results of the WRU data numbered 169 are displayed in Figure 15. Both of the two analysis results can identify the fault frequency and its harmonics. However, Figure 15b with the optimal kernel size displays fewer frequency components than Figure 15a with a random kernel size, especially in the frequency band between 100 Hz and 350 Hz. This is because those frequency components which are uncorrelated to the fault signature have been suppressed.

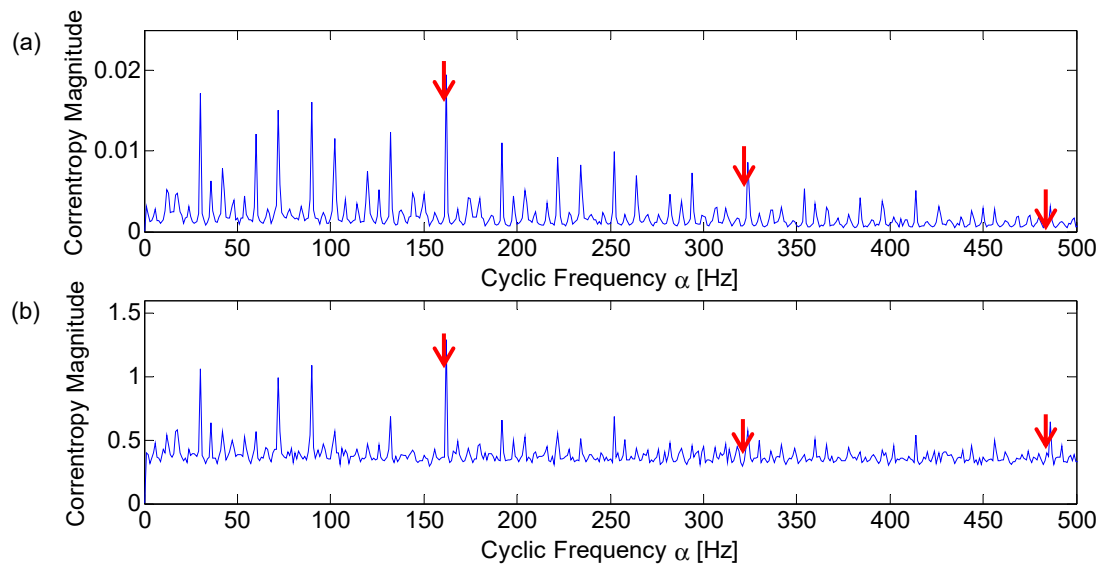


Figure 15. Cyclic domain profile of the CCES for WRU bearing data numbered 169 with different kernel size: (a) $\sigma = 1$; (b) $\sigma = 0.015$.

Similarly, the random kernel size is set to 0.1 and the optimal kernel size is 3.021 according to Silverman's rule. The analysis results of the axle bearing signal of the outer race fault are displayed in Figure 16. It can be seen from the Figure 16a that the result with the randomly chosen kernel size cannot locate frequencies corresponding to the fault characteristic. Compared with Figure 16a, the cyclic domain profile with the optimal kernel size exhibits a much better diagnosis performance, as shown in Figure 16b.

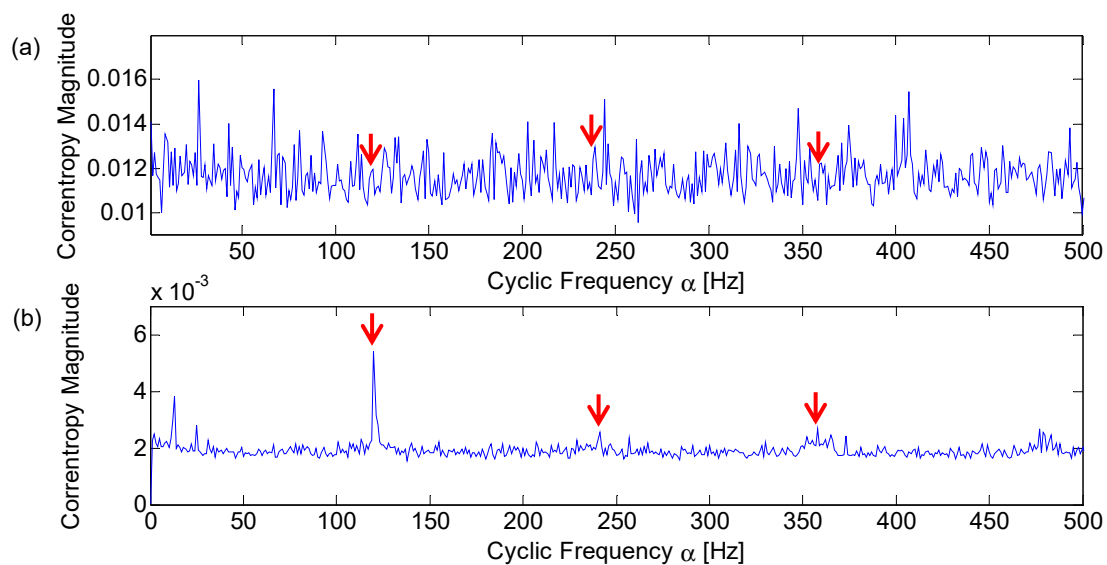


Figure 16. Cyclic domain profile of the CCES for the railway axle bearing signal of the outer race fault with different kernel sizes: (a) $\sigma = 0.1$; (b) $\sigma = 3.021$.

The above two analysis results have shown that the kernel size plays an important role in the fault diagnosis through the CCES method, especially when signal components are complex and affected by impulsive noise. The kernel optimization method based on Silverman's rule is highly efficient and accurate during the fault diagnosis. Thus, CCES can be regarded as a self-adaptation method which may also be applied other objects in mechanical fields.

5. Conclusions

In this paper, we have investigated an impulsive noise suppression method for bearing fault diagnosis based on the CCE function. Silverman's rule was used to obtain an optimal kernel size for CCE input. This adjustable parameter provides an effective mechanism to eliminate the negative effect of impulsive noise. Then, a simulated signal and two real cases were studied to validate the performance of the proposed method. The experiment results show that the proposed method has a good diagnosis performance, especially in the presence of impulsive noise. Furthermore, a powerful frequency band selection method named Fast Kurtogram was applied to analyze the same data for comparison. It was found that the fault diagnosis method based on CCE function outperforms the Fast Kurtogram method.

Further research should mainly focus on the combination of CCE and mode identification methods. Typically, frequency domain features of CCES can be extracted to reflect different fault characteristics. Moreover, considering the computation time of cyclic methodology, a simplified and fast computation method should be developed in subsequent research works. Note that this new method can generate cyclostationary signatures for bearing fault diagnosis. Thus, it should be helpful in the diagnosis of other machinery, such as gears and blades.

Author Contributions: X.Z. and Y.Q. designed the proposed method and prepared the manuscript; X.Z. and C.H. finished writing the code. L.J. and L.K. provided guidance for modifying the manuscript. All authors prepared, revised, and approved the final submission.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities, grant number 2018JBZ006; the National Natural Science Foundation of China, grant number I18A1200011; and the State Key Laboratory of Rail Traffic Control and Safety, contract number RCS2017ZT005.

Acknowledgments: The authors would like to thank Aluisio I.R. Fontes for his insightful suggestions.

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

References

1. Wang, D.; Tsui, K.-L.; Miao, Q. Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access* **2018**, *6*, 665–676. [[CrossRef](#)]
2. Feng, Z.; Liang, M.; Chu, F. Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mech. Syst. Signal Process.* **2013**, *38*, 165–205. [[CrossRef](#)]
3. Chen, X.; Wang, Z.; Zhang, Z.; Jia, L.; Qin, Y. A Semi-Supervised Approach to Bearing Fault Diagnosis under Variable Conditions towards Imbalanced Unlabeled Data. *Sensors* **2018**, *18*, 2097. [[CrossRef](#)] [[PubMed](#)]
4. Ho, D.; Randall, R. Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mech. Syst. Signal Process.* **2000**, *14*, 763–788. [[CrossRef](#)]
5. Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process.* **2006**, *20*, 282–307. [[CrossRef](#)]
6. Antoni, J.; Randall, R. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process.* **2006**, *20*, 308–331. [[CrossRef](#)]
7. Antoni, J. Fast computation of the kurtogram for the detection of transient faults. *Mech. Syst. Signal Process.* **2007**, *21*, 108–124. [[CrossRef](#)]
8. Peter, W.T.; Wang, D. The design of a new sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as “Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement—Parts 1 and 2”. *Mech. Syst. Signal Process.* **2013**, *40*, 499–519. [[CrossRef](#)]
9. He, D.; Wang, X.; Li, S.; Lin, J.; Zhao, M. Identification of multiple faults in rotating machinery based on minimum entropy deconvolution combined with spectral kurtosis. *Mech. Syst. Signal Process.* **2016**, *81*, 235–249. [[CrossRef](#)]
10. Wang, Y.; Xiang, J.; Markert, R.; Liang, M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications. *Mech. Syst. Signal Process.* **2016**, *66*, 679–698. [[CrossRef](#)]
11. Smith, W.A.; Fan, Z.; Peng, Z.; Li, H.; Randall, R.B. Optimised Spectral Kurtosis for bearing diagnostics under electromagnetic interference. *Mech. Syst. Signal Process.* **2016**, *75*, 371–394. [[CrossRef](#)]

12. Bozchalooi, I.S.; Liang, M. A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection. *J. Sound Vib.* **2007**, *308*, 246–267. [[CrossRef](#)]
13. Miao, Y.; Zhao, M.; Lin, J. Improvement of kurtosis-guided-grams via Gini index for bearing fault feature identification. *Meas. Sci. Technol.* **2017**, *28*, 125001. [[CrossRef](#)]
14. Barszcz, T.; Jabłoński, A. A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram. *Mech. Syst. Signal Process.* **2011**, *25*, 431–451. [[CrossRef](#)]
15. Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients. *Mech. Syst. Signal Process.* **2016**, *74*, 73–94. [[CrossRef](#)]
16. Antoni, J.; Bonnardot, F.; Raad, A.; El Badaoui, M. Cyclostationary modelling of rotating machine vibration signals. *Mech. Syst. Signal Process.* **2004**, *18*, 1285–1314. [[CrossRef](#)]
17. Antoni, J. Cyclic spectral analysis in practice. *Mech. Syst. Signal Process.* **2007**, *21*, 597–630. [[CrossRef](#)]
18. Antoni, J. Cyclostationarity by examples. *Mech. Syst. Signal Process.* **2009**, *23*, 987–1036. [[CrossRef](#)]
19. Antoni, J.; Hanson, D. Detection of surface ships from interception of cyclostationary signature with the cyclic modulation coherence. *IEEE J. Ocean. Eng.* **2012**, *37*, 478–493. [[CrossRef](#)]
20. Antoni, J.; Xin, G.; Hamzaoui, N. Fast computation of the spectral correlation. *Mech. Syst. Signal Process.* **2017**, *92*, 248–277. [[CrossRef](#)]
21. Luan, S.; Qiu, T.; Zhu, Y.; Yu, L. Cyclic correntropy and its spectrum in frequency estimation in the presence of impulsive noise. *Signal Process.* **2016**, *120*, 503–508. [[CrossRef](#)]
22. Liu, T.; Qiu, T.; Luan, S. Cyclic Correntropy: Foundations and Theories. *IEEE Access* **2018**, *6*, 34659–34669. [[CrossRef](#)]
23. Fontes, A.I.; Rego, J.B.; Martins, A.d.M.; Silveira, L.F.; Príncipe, J.C. Cyclostationary correntropy: Definition and applications. *Expert Syst. Appl.* **2017**, *69*, 110–117. [[CrossRef](#)]
24. Santamaría, I.; Pokharel, P.P.; Principe, J.C. Generalized correlation function: Definition, properties, and application to blind equalization. *IEEE Trans. Signal Process.* **2006**, *54*, 2187–2197. [[CrossRef](#)]
25. Parzen, E. *Statistical Methods on Time Series by Hilbert Space Methods*; Applied Mathematics and Statistics Laboratory, Stanford University: Stanford, CA, USA, 1959; Volume 23.
26. Chen, B.; Xing, L.; Zhao, H.; Zheng, N.; Pri, J.C. Generalized correntropy for robust adaptive filtering. *IEEE Trans. Signal Process.* **2016**, *64*, 3376–3387. [[CrossRef](#)]
27. Liu, W.; Pokharel, P.P.; Principe, J.C. Correntropy: Properties and applications in non-Gaussian signal processing. *IEEE Trans. Signal Process.* **2007**, *55*, 5286–5298. [[CrossRef](#)]
28. Gunduz, A.; Principe, J.C. Correntropy as a novel measure for nonlinearity tests. *Signal Process.* **2009**, *89*, 14–23. [[CrossRef](#)]
29. Wang, Y.; Tang, Y.Y.; Li, L. Correntropy matching pursuit with application to robust digit and face recognition. *IEEE Trans. Cybern.* **2017**, *47*, 1354–1366. [[CrossRef](#)] [[PubMed](#)]
30. Wu, Z.; Shi, J.; Zhang, X.; Ma, W.; Chen, B.; Senior Member, I. Kernel recursive maximum correntropy. *Signal Process.* **2015**, *117*, 11–16. [[CrossRef](#)]
31. Erdogmus, D.; Principe, J.C. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. *IEEE Trans. Signal Process.* **2002**, *50*, 1780–1786. [[CrossRef](#)]
32. Xu, J.-W.; Paiva, A.R.; Park, I.; Principe, J.C. A Reproducing Kernel Hilbert Space Framework for Information-Theoretic Learning. *IEEE Trans. Signal Process.* **2008**, *56*, 5891–5902.
33. Jensen, R.; Erdogmus, D.; Principe, J.C.; Eltoft, T. Towards a unification of information theoretic learning and kernel methods. In *Machine Learning for Signal Processing, 2004, Proceedings of the 2004 14th IEEE Signal Processing Society Workshop, Sao Luis, Brazil, 29 September–1 October 2004*; IEEE: Piscataway, NJ, USA, 2004; pp. 93–102.
34. Zhao, S.; Chen, B.; Principe, J.C. An adaptive kernel width update for correntropy. In *Proceedings of the 2012 International Joint Conference on Neural Networks (IJCNN), Brisbane, Australia, 10–15 June 2012*; pp. 1–5.
35. Bowman, A.W. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika* **1984**, *71*, 353–360. [[CrossRef](#)]
36. Sheather, S.J. Density estimation. *Stat. Sci.* **2004**, *19*, 588–597. [[CrossRef](#)]
37. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman & Hall: London, UK, 1986; ISBN 0-412-24620-1.
38. Antoni, J.; Randall, R. A stochastic model for simulation and diagnostics of rolling element bearings with localized faults. *J. Vib. Acoust.* **2003**, *125*, 282–289. [[CrossRef](#)]

39. Case Western Reserve University Bearing Data Center Website. Available online: <http://csegroups.case.edu/bearingdatacenter/home> (accessed on 18 January 2018).
40. Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [[CrossRef](#)]
41. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64*, 100–131. [[CrossRef](#)]
42. Zhao, X.; Qin, Y.; Kou, L.; Liu, Z. Understanding real faults of axle box bearings based on vibration data using decision tree. In Proceedings of the 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), Seattle, WA, USA, 11–13 June 2018; pp. 1–5.



© 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/>).