

PAPER • OPEN ACCESS

Synchronous fault feature extraction for rolling bearings in a generalized demodulation framework

To cite this article: Kangning Liu *et al* 2023 *Meas. Sci. Technol.* **34** 095009

View the [article online](#) for updates and enhancements.

You may also like

- [Demodulated synchrosqueezing S-transform and its application to machine-fault diagnosis](#)
Wei Liu, Yang Liu, Shuangxi Li et al.
- [Study on demodulated signal distribution and acoustic pressure phase sensitivity of a self-interfered distributed acoustic sensing system](#)
Ying Shang, Yuan-Hong Yang, Chen Wang et al.
- [A high-speed electrical impedance measurement circuit based on information-filtering demodulation](#)
Shijie Sun, Lijun Xu, Zhang Cao et al.

Synchronous fault feature extraction for rolling bearings in a generalized demodulation framework

Kangning Liu², Juanjuan Shi^{1,*} , Changqing Shen² , Weiguo Huang² 
and Zhongkui Zhu² 

¹ State Key Laboratory of Performance Monitoring and Protecting of Rail Transit Infrastructure, Nanchang 330013, People's Republic of China

² School of Rail Transportation, Soochow University, Suzhou 213151, People's Republic of China

E-mail: jshi091@suda.edu.cn

Received 31 January 2023, revised 17 April 2023

Accepted for publication 5 May 2023

Published 1 June 2023



CrossMark

Abstract

Generalized demodulation (GD) has the potential to process non-stationary vibration signals since it can demodulate a signal with a curved time–frequency (TF) ridge into a signal, with a TF ridge parallel to a time axis with an improved time-frequency representation (TFR) energy concentration level. However, current GD methods require iteration operations and cannot simultaneously deal with vibrations from multiple components of rolling bearings. This paper proposes a method based on the GD framework, which can simultaneously demodulate multiple components of interest using the Hadamard product between matrices. A synchronous extractor is also constructed to post-process TFRs of generalized demodulated signals to further improve the TF aggregation. Unlike the conventional synchronous extraction transform, the synchronous extractor in this paper can be directly applied to TF ridges parallel to the time axis without the estimation of instantaneous frequencies (IFs). Then, the post-processed TF ridges are backward demodulated to restore the actual IF. The proposed synchronous fault feature extraction method in the GD framework also allows for the signal reconstruction. Both simulated and experimental signals are applied to validate the effectiveness of the proposed method for rolling bearing fault diagnosis.

Keywords: bearing fault diagnosis, time-frequency analysis, time-varying speed condition, fault feature extraction, generalized demodulation

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Bearings are vital components of rotating machinery systems and their health condition plays an important role in ensuring the safe operation of rotating machinery. In reality, most rotating machinery often works under time-varying speed conditions. Therefore, conventional fault diagnosis methods, including spectrum analysis and envelope spectrum analysis, are no longer effective due to the time-varying features of vibration signals. Thus, the fault detection and diagnosis of bearings under variable speed conditions has become a hot topic of research [1, 2].

Time–frequency analysis (TFA) is a method that can be used to show the time-varying characteristics of non-stationary signals [3, 4]. However, the smearing and cross-term problems of traditional TFA methods, including short-time Fourier transform (STFT), Wigner–Ville distribution (WVD), wavelet transform (WT) and empirical mode decomposition (EMD), greatly confine their applications for bearing fault diagnosis. Specifically, WVD is a serious cross-term interference problem. Although plenty of research work has been carried out by both national and international scholars on the suppression of cross terms, these methods are at the expense of WVD TF aggregation [5–8]. EMD does not yet have a solid theoretical foundation, and it suffers from problems, such as under-envelope, over-envelope modal confusion, endpoint effects, etc. In addition, the binary filtering characteristics of EMD sieving also make it unable to process signals with a wide range of continuous changes in instantaneous frequencies (IFs) [9–12]. WT methods have excellent TF localization. However, its ‘box’ TF atoms (such as Gabor basis functions, Morlet basis functions, etc) are not suitable for vibration signals of time-varying speeds, especially for start-stop signals of equipment [13]. Li *et al* proposed a method for bearing fault diagnosis based on improved multiscale permutation entropy and least squares support vector machine (SVM) [14].

In order to solve the TF ambiguity problem caused by the influence of TF resolution in the traditional TFA method, Yu *et al* proposed a synchroextracting transform (SET), which can enhance the TF aggregation and sharpen the TF ridges of nonstationary vibration signals [15, 16]. Shi *et al* constructed an adaptive demodulation factor to achieve the generalized demodulation (GD) synchronously by angle matching and found the most suitable window length and angle by the relationship between different window lengths, angles and spectral kurtosis values [17]. Li *et al* used the empirical WT (EWT) to decompose the bearing fault signal to enhance the TF characteristics of the signal [18]. Wang *et al* proposed the iterative use of the matching demodulation transform to improve the accuracy of the extracted IFs [19]. Shi *et al* proposed a fractal dimension and GD-based fault diagnosis method of bearing under variable speed conditions, which avoids the resampling process when processing nonstationary signals [20].

In addition to the TFA methods reviewed above, GD-based TFA methods provide another way of processing bearing vibration signal under time-varying speed conditions since the GD can demodulate the time-varying frequency of the signal to a constant frequency parallel to the time axis, showing

that GD has the potential to process non-stationary vibration signals [21, 22]. Shi *et al* have proposed the generalized step-wise demodulation transform by uniquely utilizing the properties that can effectively enhance the TF readability [23]. Zhao *et al* used the Vold–Kalman filtering method to extract the time-varying non-stationary bearing and gear fault feature components from the envelope spectrum of the original signal [24]. Liu *et al* have redefined the energy factor of GD to effectively improve the capability of noise against when processing vibration signals with time-varying frequencies [25]. Li *et al* proposed a method to construct GD phase functions based on peak engagement multiplier trend lines to achieve iterative GD for diagnosing the rolling bearing faults under variable speeds without tachometers [26]. Liu *et al* proposed a hybrid method that integrates GD and artificial neural networks, and the method was shown to improve the accuracy of fault determination [27]. Pi *et al* obtained the phase function of the multi-component signal components by combining the multiscale line frequency modulation with the sparse signal decomposition method. GD and spectral analysis of the signal were then performed to obtain the gear fault characteristic frequency (FCF) [28]. Liu *et al* proposed a novel flexible iterative GD filtering method to reveal the fault-related frequencies, where the demodulated frequency values are not subject to the speed fluctuation profiles [29].

However, most aforementioned GD-based methods rely on prior knowledge of each signal component of interest and require an iterative process to process multi-component signals, which may make the algorithm largely dependent on each IF estimation and also complicate the algorithms. For post-processing methods, such as SET, the realization requires the estimation of the IF at each time instant, thereby leading to a large computational cost. Little work has been performed on synchronous GD of multi-component signals. Therefore, this paper proposes to synchronously demodulate vibration signals with multi-component components of interest by constructing a demodulation factor matrix without the iterative operation being involved. The synchronous extractor is then constructed for the generalized demodulated signal components, and the TF aggregation of the vibration signal analysis is further enhanced.

The rest of this paper is organized as follows. Section 2 describes the proposed method in detail, including the matrix construction of demodulation factors, pre-determination of fault characteristic coefficients, construction of the extraction matrix and signal reconstruction of the proposed method. Section 3 illustrates the effectiveness of the method using simulated signals. Section 4 verifies the effectiveness of the proposed method through experimental signals, and Section 5 gives the conclusion.

2. Proposed synchronous fault feature extraction method in a GD framework

The proposed method is composed of the following modules: the matrix of demodulation factor, pre-determination of

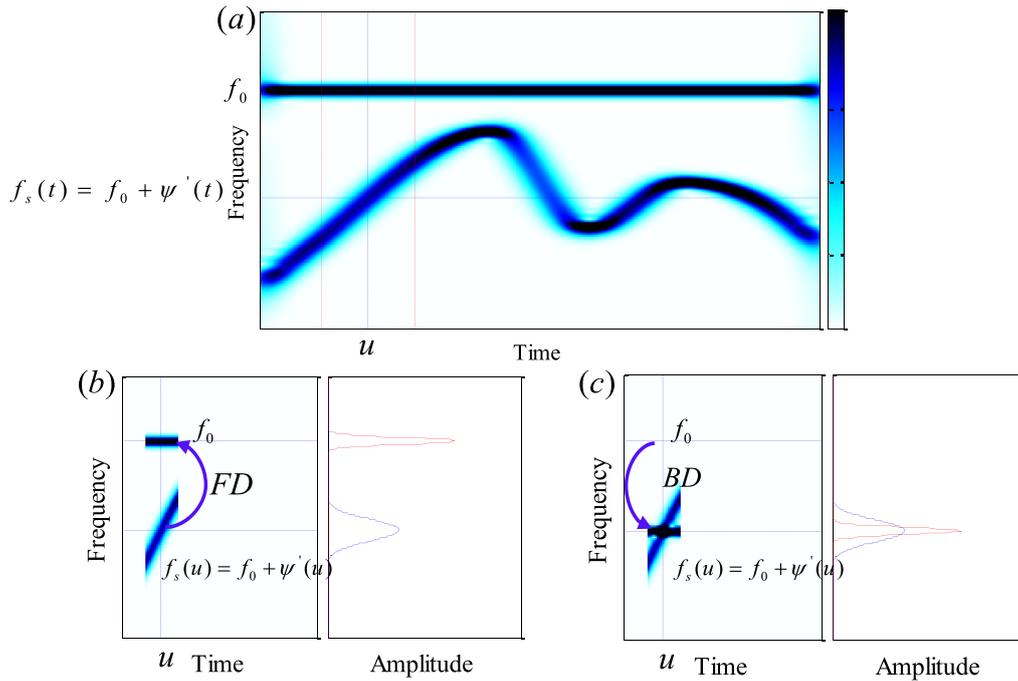


Figure 1. (a) TFR of the signal and demodulated signal, (b) TFR of the original signal segment and forward demodulated signal segment, and their spectra and (c) TFR of the original signal segment and the backward demodulated signal segment and their spectra (FD: forward demodulation, BD: backward demodulation), (the level of darkness reflects the energy level).

fault characteristic coefficients, construction of the extraction matrix. For GD, how to demodulate multiple components directly is a problem. In this paper, the theory of multi-component synchronous GD by constructing a matrix of demodulation factors is proposed, which can obtain the frequency–amplitude (F–A) diagram from the matrix that is obtained by solving the Hadamard product of the signal and the demodulation factor matrix. The extraction matrix is used to extract the components of interest after demodulation. The three modules are illustrated in the following section.

2.1. The matrix of demodulation factor

GD is a useful tool that can help to achieve separation of non-stationary signals with multi-components. GD is based on the generalized Fourier transform (GFT), and a crucial part of GD is the construction of the demodulation factor. Considering an analytical signal $x(t)$, the GFT is expressed as,

$$X_G(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2\pi i(ft+x_0(t))} dt, \quad (1)$$

where $x_0(t)$ is a real-valued demodulator that depends on time only. It can also be considered as the standard Fourier transform (FT) of $x(t) \cdot e^{-2\pi i x_0(t)}$. Observing equation (1), when $X_G(f) \equiv \delta(f - f_0)$, it can be written as,

$$x_d(t) = e^{2\pi i[f_0 t + x_0(t)]}. \quad (2)$$

Observing equation (2), a particular time-varying non-stationary component of the analytic signal can be converted into a linear time-parallel stationary component if its conversion factor satisfies the following equation:

$$f(t) = f_0 + \frac{dx_0(t)}{dt}, \quad (3)$$

where $f(t)$ denotes the IF of a signal component. The TF ridges can be gained by the traditional peak search method based on time frequency representation (TFR). The property of GD of mapping signals with time-varying trajectories to a linear path of arbitrary constant frequency parallel to the time axis can help to avoid frequency overlap, hence facilitating filtering operations and the separation of signal components. In summary, one of the important characteristics of GD is that, after the demodulation, the TF ridge of the demodulated signal is located at a constant frequency f_0 and the energy of the signal $x(t)$ would be concentrated on the constant frequency f_0 on the TFR, as can be seen in figure 1.

Once a basic demodulation factor has been determined, the demodulation factor matrix can be constructed as follows:

$$D = \begin{bmatrix} d_0(t) \\ n_0 \cdot d_0(t) \\ n_1 \cdot d_0(t) \\ n_2 \cdot d_0(t) \\ \vdots \\ n_N \cdot d_0(t) \end{bmatrix}, \quad (4)$$

where $d_0(t)$ is the basic demodulation factor determined by the IF of the signal component of interest, and $n_n (n = 1, 2, 3, \dots, N)$ represents the coefficient of the basic demodulation factor. Hence, expanding the original signal $x(t)$ into a matrix of $n_n \times l$, in which l is the length of the signal analyzed.

Here, we present an algorithm, called the Hadamard product, for two matrices of the same order $A = (a_{i,j})$, $B = (b_{i,j})$, if $c_{i,j} = a_{i,j} \times b_{i,j}$, the matrix $C = (c_{i,j})$ is considered to be the Hadamard product of A and B . It can be defined as,

$$C = A \odot B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{bmatrix}, \quad (5)$$

which can help to achieve synchronous generalization of the two matrices,

$$S = X \odot D = \begin{bmatrix} x(t) \\ x(t) \\ x(t) \\ x(t) \\ \vdots \\ x(t) \end{bmatrix} \odot \begin{bmatrix} d_0(t) \\ n_0 \cdot d_0(t) \\ n_1 \cdot d_0(t) \\ n_2 \cdot d_0(t) \\ \vdots \\ n_N \cdot d_0(t) \end{bmatrix}, \quad (6)$$

where \odot is the Hadamard product and S is the matrix of the demodulated signals. Each row in matrix S represents the corresponding demodulation results obtained with different demodulation coefficients.

2.2. Pre-determination of demodulation coefficients

In order to use equation (6) to synchronously demodulate the signal components of interest, the demodulation coefficients n_n should be pre-determined first. Denote the STFT-resulting TFR of all the demodulated signals stored in matrix S in equation (6) as $\mathbf{SS}^{M \times N}$ (M and N represent the size of the matrix). Since the GD can improve the energy concentration of the signal when the frequency of the demodulation factor best matches the actual one of the signal components, the mean of the TF amplitudes of the frequency where this demodulated signal component is located should be theoretically much larger than the ones of the signal component demodulated by other demodulation factors. Taking the mean of TF amplitudes of all frequencies of the matrix \mathbf{SS} , a new matrix A containing the mean of TF amplitudes of each row in \mathbf{SS} can be expressed as,

$$A = [a_1 \quad a_2 \quad a_3 \quad \cdots \quad a_n], \quad (7)$$

where $a_n (n = 1, 2, 3, \dots)$ means the mean of TF amplitude of the row in \mathbf{SS} . Figure 1 shows the F–A graph of an experimental signal.

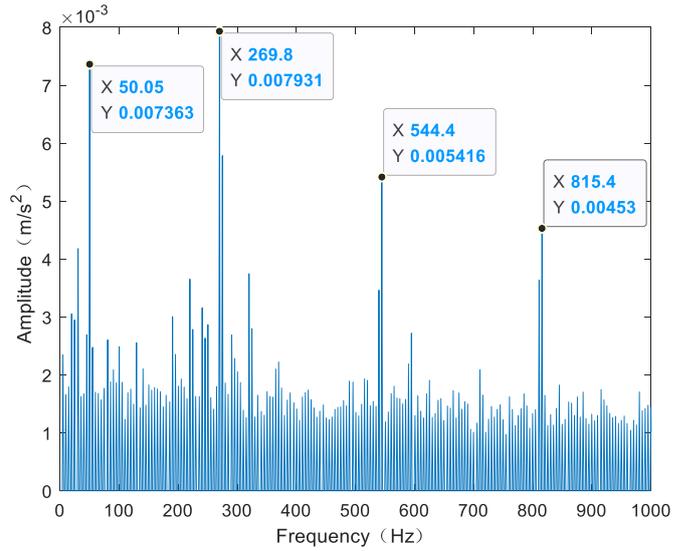


Figure 2. F–A diagram of the signals.

In figure 2, the frequencies of the peaks are denoted by $\omega_k (k = 1, 2, 3, \dots)$. The demodulation coefficients can be predetermined by the ratios of the higher frequencies (ω_2 and ω_3) to the frequency (ω_1) of the first peak. The ratios of frequencies corresponding to these peaks are 5.40 ($=269.8/50.05$), 10.87 ($=544.4/50.05$) and 16.29 ($=815.4/50.05$). It can be seen that 10.87 and 16.29 are double and triple 5.40, respectively. Therefore, it can be predetermined that its basic demodulation coefficient is 5.40, and the other two are two and three times the basic demodulation coefficient, respectively. With the pre-determined demodulation coefficients, the demodulation matrix that is used to extract the TF ridges of interest can be obtained and also paves the way for the extraction matrix construction.

2.3. Construction of the extraction matrix

GD can demodulate a signal component with curved IF into a signal with a horizontal frequency line. On the one hand, the TF ridges of generalized demodulated signals can be further sharpened. On the other hand, the overlap problem may be generated when processing multi-component signals. Hence, a clearer TFR of the demodulated signals is desirable, which can be obtained by extracting the TF ridges at the targeted frequencies of the TFR.

The extraction operator construction can start from STFT and its definition is as follows:

$$G(t, \omega) = \int_{-\infty}^{\infty} g(u-t) s(u) e^{-j\omega u} du, \quad (8)$$

where $s(u)$ is an analyzed signal and $g(u-t)$ denotes the moving window. Let $g_\omega(u) = g(u-t) \cdot e^{j\omega u}$, according to Parseval's theorem, equation (8) can be written as,

$$\begin{aligned}
 G(t, \omega) &= \int_{-\infty}^{\infty} s(u)(g_{\omega}(u))^* du \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\zeta) \cdot [e^{j(\omega-\zeta)t} \cdot G_{\omega}(\omega-\zeta)]^* d\zeta \\
 &= e^{-j\omega t} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\zeta) \cdot G_{\omega}(\omega-\zeta) \cdot e^{-j\zeta t} d\zeta \quad (9)
 \end{aligned}$$

where $S(\zeta)$ and $(g_{\omega}(u))^*$ are the FTs of $s(u)$ and $g_{\omega}(u)$, separately. The superscript * indicates that the complex conjugate $G_{\omega}(\omega-\zeta)$ is the FT of the window function. Since the observed TF spectrum is of amplitude spectrum, adding a phase move $e^{-j\omega t}$ to the STFT results $G(t, \omega)$ does not affect the amplitude, let,

$$\begin{aligned}
 G_e(t, \omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\zeta) \cdot G_h(\omega-\zeta) \cdot e^{-j\zeta t} d\zeta \\
 &= \int_{-\infty}^{\infty} g(u-t) \cdot s(u) \cdot e^{-j\omega(u-t)} du. \quad (10)
 \end{aligned}$$

Constructing the single-component signal $s_h(t) = A \cdot e^{-j\omega t}$, its FT is $S_h(\zeta) = 2\pi A \cdot \delta(\omega - \omega_0)$, where A denotes its amplitude and ω_0 denotes the frequency of $s_h(t)$, by substituting the signal $s_h(t)$ into equation (10), we can get its STFT expression:

$$G_e(t, \omega) = A \cdot G_h(\omega - \omega_0) \cdot e^{-j\omega_0 t}. \quad (11)$$

In equation (11), according to Heisenberg's inaccuracy principle, the window function is chosen as a Gaussian window in order to get the best time and frequency resolution and $G_h(\omega)$ is compact [12]. Due to $|e^{-j\omega_0 t}| = 1$, $|G_e(t, \omega)|$ then achieves its maximum value at $\omega = \omega_0$, where it has the highest energy. Hence, the TF component that demodulates the signal to $\omega = \omega_0$ can be considered to be extracted from the TF domain.

The extraction operator is denoted as $\int_{-\infty}^{\infty} \delta(\omega - \omega_0(t, \omega)) d\omega$ and the extraction transformation can be expressed as,

$$T_{G0}(\omega) = \int_{-\infty}^{\infty} X_G(\omega) \cdot \delta(\omega - \omega_0(t, \omega)) d\omega. \quad (12)$$

When $\omega = \omega_0(t, \omega)$,

$$T_G(\omega_0) = X_G(\omega) \cdot \delta(\omega - \omega_0(t, \omega)). \quad (13)$$

Hence, this extraction operator can be defined as,

$$\delta(\omega - \omega_0(t, \omega)) = \begin{cases} 1 & \omega = \omega_0 \\ 0 & \omega \neq \omega_0 \end{cases}. \quad (14)$$

From equations (13) and (14) we know that in the TF domain, it can be expressed as,

$$T_G(\omega) = \begin{cases} X_G(\omega) & \omega = \omega_0 \\ 0 & \omega \neq \omega_0 \end{cases}. \quad (15)$$

This extraction operator can effectively extract only the TF component $X_G(\omega_0)$ in TFR.

From the above derivation, it is known that the extraction operator can extract the signal component at a constant frequency, and the demodulated frequency positions of all signal components associated with the FCF and its harmonics can be determined from the F-A diagram. Hence, it is possible to extract the signal components in the matrix **SS** corresponding to the target frequencies determined by the F-A diagram by constructing the extraction operator matrix:

$$O = \begin{bmatrix} a & a & a \\ a & a & a \\ a & a & a \end{bmatrix}_{M \times N}, \begin{cases} a = 1 & \omega_m = \omega_k(t, \omega) \\ a = 0 & \omega_m \neq \omega_k(t, \omega) \end{cases}, \quad (16)$$

where **O** is the extraction matrix, m is the m th row of the matrix and ω_k ($k = 1, 2, 3, \dots$) represents frequencies corresponding to the peaks in the F-A diagram.

By calculating the Hadamard product of the matrix **SS** and **O**, the demodulated signal components of interest can be successfully extracted:

$$\begin{aligned}
 S_0 &= \mathbf{SS} \odot \mathbf{O} = \begin{bmatrix} SS_{11} & SS_{12} & \cdots & SS_{1N} \\ SS_{21} & SS_{22} & \cdots & SS_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ SS_{M1} & SS_{M2} & \cdots & SS_{MN} \end{bmatrix}_{M \times N} \\
 &\odot \begin{bmatrix} a & a & \cdots & a \\ a & a & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & a \end{bmatrix}_{M \times N} \\
 &\times \begin{cases} a = 1 & \omega_m = \omega_k(t, \omega) \\ a = 0 & \omega_m \neq \omega_k(t, \omega) \end{cases}, \quad (17)
 \end{aligned}$$

where $SS_{m,n}$ represents the element of the matrix **SS** and ω_k is the frequencies of the peaks determined by the F-A graph. Equation (17) refers to the signal matrix that only contains each component of interest after demodulation.

Hence, the inverse GD can be performed on the basis of the matrix **S**₀ to obtain the TF components with the original frequency distribution trend and enhanced aggregation. Consequently, the accuracy of the fault characterization factor can be further improved.

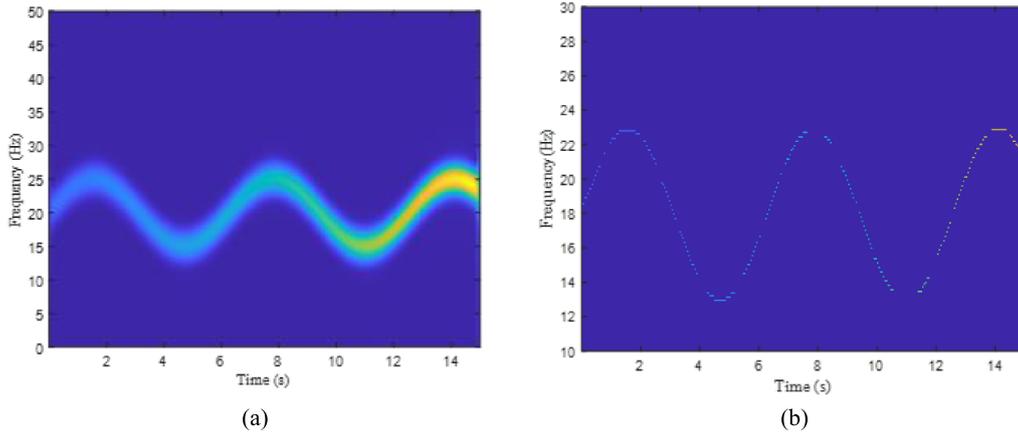


Figure 3. TFR of the simulated signal: (a) TFR of the original signal; (b) TFR of the processed signal.

2.4. Time domain reconstruction

Since the TF energy is redistributed only in the frequency path, the proposed method can also satisfy the time domain of the signal reconstruction, which is proved as follows:

$$\begin{aligned}
 x(u) &= \frac{1}{g(0)} \int_{-\infty}^{\infty} x(t)g(t-u)\delta(t-u)dt \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t)g(t-u)e^{-i\omega(t-u)}dtd\omega \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} S_x^g(u,\omega)d\omega \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} S_x^g(u,\omega)d\omega \int_{-\infty}^{\infty} \delta(\eta - \varpi_x(u,\omega))d\eta \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_x^g(u,\omega)\delta(\eta - \varpi_x(u,\omega))d\eta d\omega \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} T_x^{M[1]}(u,\eta)d\eta \\
 &\vdots \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} T_x^{M[N-1]}(u,\eta)d\eta \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} T_x^{M[N-1]}(u,\eta) \int_{-\infty}^{\infty} \delta(\eta - \varpi_x(u,\eta))d\omega d\eta \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T_x^{M[N-1]}(u,\eta)\delta(\eta - \varpi_x(u,\eta))d\eta d\omega \\
 &= \frac{1}{2\pi g(0)} \int_{-\infty}^{\infty} T_x^{M[N]}(u,\omega)d\omega, \tag{18}
 \end{aligned}$$

where $g(0)$ is the constant determined after the selection of the window function. For multi-component signals where the frequency components can be completely separated, the reconstruction in time domain for each component of the

multi-component signal can be done by the TF ridge reconstruction alone, written as,

$$x_k(u) = (2\pi g(0))^{-1} \int_{|\omega - \varphi'_k(u)| < ds} T_x^{M[N]}(u,\omega)d\omega, \tag{19}$$

where ds is the reconstruction bandwidth, which can be set to the frequency resolution $\Delta\omega$ and its multiplier, indicating the number of TF points used to reconstruct the signal during reconstruction. φ'_k is the IF of the k th component, which can be obtained using the ridge search algorithm.

To verify the effectiveness of the above method, the simulation signal without noise is defined as follows:

$$x(t) = \sin\left(2\pi \int_0^t f(t)dt\right), \tag{20}$$

where f denotes the IF, defined as,

$$f = 10\sin(0.5\pi t) + 22. \tag{21}$$

The sampling frequency is 600 Hz and the signal lasts 15 s. Figure 3(a) shows the TFR of the signal and figure 3(b) shows the TFR after being processed by the method. The TF energy of the signal is concentrated in a smaller frequency range, and the TF aggregation is further improved. The obtained TFR is estimated for the IF, and then the time-domain signal reconstruction is performed based on the estimated component of the target frequency, and the results are shown in figure 4. The effectiveness of the algorithm can be verified by the mean relative error (MRE), which is defined as,

$$\text{MRE} = \frac{1}{L_x} \int \frac{x(u) - \chi(u)}{x(u)} du, \tag{22}$$

where L_x denotes the length of the simulation signal and $\chi(u)$ is the reconstruction signal. Equation (22) can be used to quantify the accuracy of the signal reconstruction. A smaller MRE means a higher accuracy of the reconstructed signals.

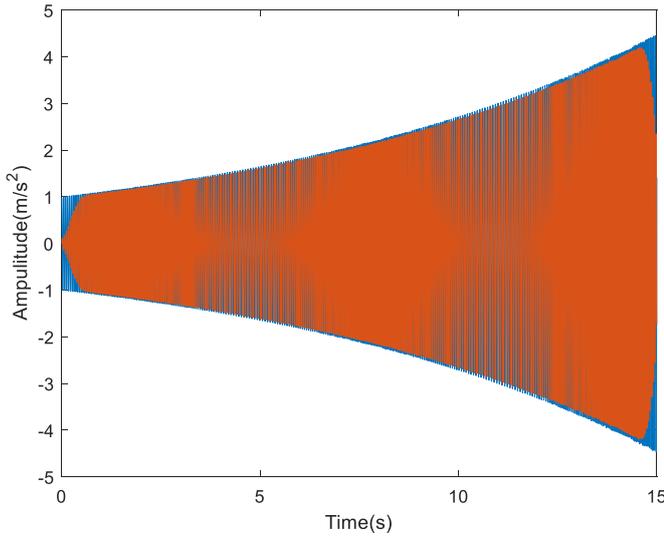


Figure 4. Time domain reconstruction.

The MRE of the reconstructed signal and the original signal was calculated to be 4.2%, which indicates that the proposed method can achieve the signal reconstruction.

3. Simulation study

In this section, the feasibility of the algorithm is verified by mono-component and multi-component simulation signals, respectively.

In order to verify the effectiveness of the proposed method, a simulation signal is constructed. A simulation signal model of the fault bearing vibration is constructed as follows:

$$x(t) = x_1(t) + x_2(t) + \eta(t), \quad (23)$$

where $x_1(t)$ denotes the component associated with the fault feature. It can be defined as,

$$x_1(t) = [1 + \alpha \cos(2\pi x_2(t))] \sum_{m=1}^M A_m e^{-\beta(t-t_m)} \sin(\omega_r(t-t_m)) \times u(t-t_m), \quad (24)$$

where $[1 + \alpha \cos(2\pi x_2(t))]$ indicates the resonance phenomenon demodulated by the speed signal $x_2(t)$, $\alpha = 0.11$ is the modulation amplitude, M is the signal length and A_m denotes the amplitude of the m th impulse response. $\beta = 800$ denotes the coefficient related to damping, ω_r indicates the resonant frequency being stimulated, $u(t)$ is a unit step function and t_m denotes the time of occurrence of the m th pulse, defined as,

$$\begin{cases} t_1 = (1 + \mu_1) [1/x_2(0)] \\ t_m = (1 + \mu_m) [1/x_2(0) + 1/x_2(t_1) + \dots + 1/x_2(t_{m-1})], \\ m = 2, 3, \dots, M, \end{cases} \quad (25)$$

where μ_m indicates the slip rate, which varies between 0.01 and 0.02 [28], $x_2(t)$ indicates the component related to the rotational speed of the rotary axis, $\eta(t)$ indicates Gaussian white noise and $x_2(t)$ is defined as,

$$x_2(t) = A_1 \cos(2\pi\varphi) + A_2 \cos(4\pi\varphi) + A_3 \cos(6\pi\varphi), \quad (26)$$

where $A_{i=1,2,3}$ denotes the amplitude and takes the values of 1.6, 1.2 and 1, accordingly. The instantaneous rotational frequency of the simulated signal is defined as $\dot{\varphi} = 2.49t^2 + 5t + 35$. The FCF is set to 3.7. The sampling frequency is 20 kHz and the sampling time is 5.1 s. For the simulation signal $x(t)$, $X(\tau, f)$ is the STFT—resulting TFR of $x(t)$, which is expressed as follows:

$$X(\tau, f) = \int_{-\infty}^{\infty} x(t)h(t-\tau)e^{-j2\pi ft} dt, \quad (27)$$

where $h(t)$ is the window function. Let $SP(\tau, f) = |X(\tau, f)|^2$, which can be considered as the energy of the signal at time t , frequency f . The peak search method used in this paper is defined as follows:

$$f(m, k) = \arg \max SP(m, k), k = 1, 2, \dots, \frac{N}{2} + 1, \quad (28)$$

$$f(m, k) > \alpha f(m-1, k), \quad (29)$$

$$f(m, k) < \beta f(m-1, k), \quad (30)$$

where α and β are obtained from traditional experience. If equations (29) and (30) are not satisfied, set the maximum energy at that point to zero and repeat equations (29) and (30).

The simulated signal is shown in figure 5(a), and its TFR is shown in figure 5(b).

From figure 5(b), some signal components are overwhelmed by noise, which makes it difficult to accurately extract them. The TF ridges extracted directly from figure 5(b) by the peak search method are shown in figure 6.

In figure 6(a), the second and third harmonics of the IF of the simulation signal are adversely influenced by noise, which makes it difficult to accurately extract all the IF curves of interest from the TFR. Then, the most accurate IF curve extracted (low-frequency one in this case) is used as the basic demodulator to construct the demodulator factor matrix defined by equation (4) to demodulate the signal to obtain the F–A diagram, as shown in figure 6(b). The most accurate IF curve can be determined by the sum of the absolute values of frequencies of adjacent time instants, which corresponds to the one with the minimal summation since mechanical systems generally do not experience sudden changes [30].

According to figure 6(b), the ratio of the two peaks with higher frequency to the frequency corresponding to the first peak can be determined as 2 (=99.61/49.8) and 3.12 (=155.3/49.8). The demodulation coefficients are then pre-determined for the next synchronous demodulation and extraction.

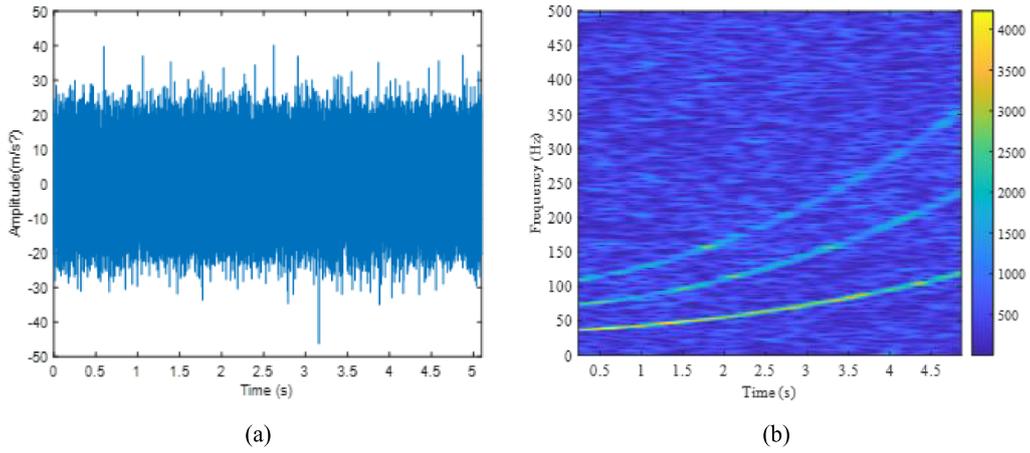


Figure 5. Simulated signal and its TFR: (a) time domain representation of the simulated signal; (b) TFR of the simulated signal.

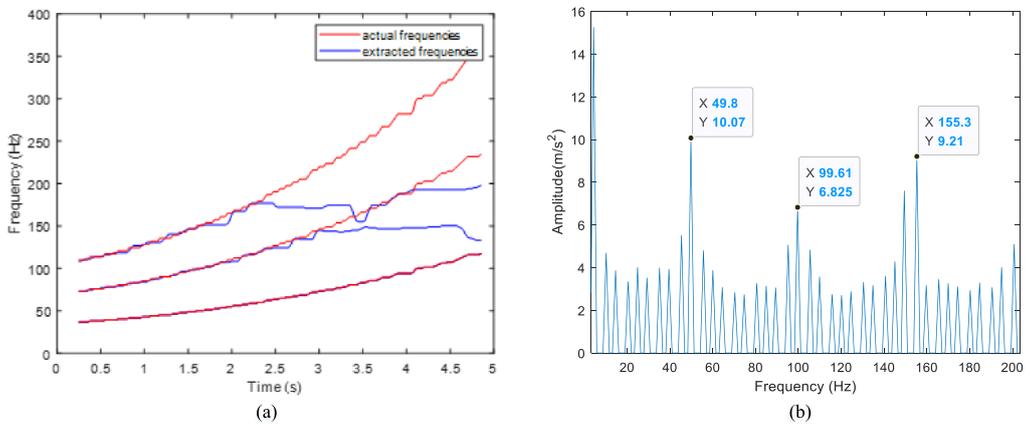


Figure 6. TF ridges of the simulated signal.

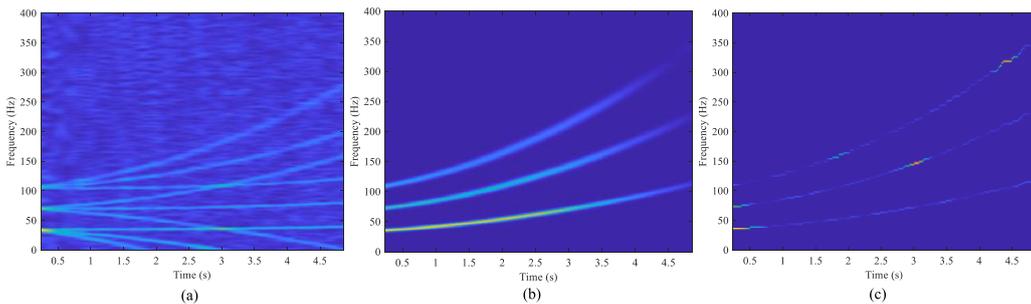


Figure 7. (a) TFR obtained by the traditional GD method without filtering, (b) TFR obtained by the traditional GD method with filtering and (c) TFR obtained by the proposed method.

The simulation signal can be processed by the method proposed in this paper. With a pre-determined demodulation coefficient, only the IF and its second and third harmonics need to be extracted. Figure 7(a) shows the TFR of the simulation signal obtained by the traditional GD method with iteration operations, in which the IFs of signal components are demodulated to constant frequencies. However, the spectrum overlap problem appears. Figure 7(b) shows the TFR obtained by the traditional GD method with filtering. As observed, even though

the overlapping and interference phenomena are relieved, its TF aggregation still has the space to be enhanced. Figure 7(c) is the TFR of the proposed method, and compared with the figures 7(a) and (b), the proposed method can directly extract the TF ridges of signal components of interest and meanwhile enhance its TF aggregation. Figure 8 shows the IF and its harmonic curves extracted from the TFR obtained by the proposed method, which is much more accurate compared with the one in figure 6(a). The comparisons show that the proposed method

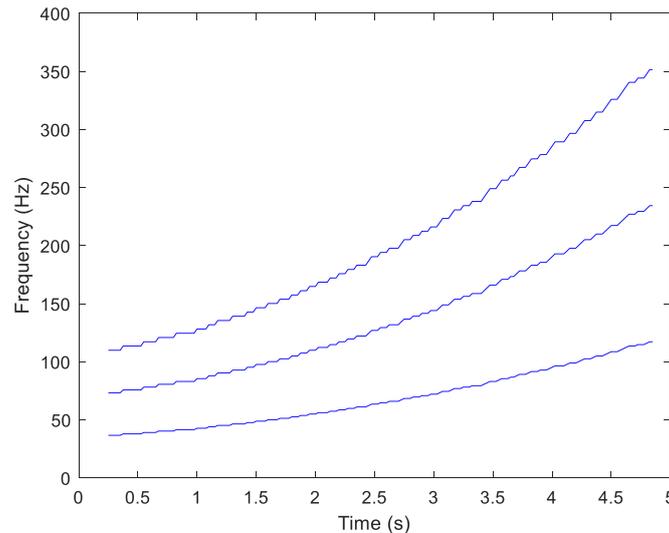


Figure 8. IF and its harmonic curves of the TFR obtained by the proposed synchronous GD.

can enhance the TF ridges of interest and the IFs can then be accurately extracted from the enhanced TFR.

Numerically, comparing the MREs of the extracted IFs in figures 6(a) and 8, it is improved from 6.7% to 0.94 %, verifying that the proposed method can effectively enhance the TF ridges and accurately extract the IF and its harmonics.

From the above discussion, it can be seen that compared with the traditional GD method, the accuracy of the extracted TF ridges is significantly improved (MRE is improved from 6.7% to 0.94%), which in turn reflects the improvement in the TF aggregation achieved by the proposed method.

4. Experimental validation

In order to further verify the effectiveness of the proposed method in real applications, it is applied to the IF extraction of experimental signals with variable speed and fault diagnosis.

4.1. Outer race fault detection

The experiment was conducted on the MKF-PK5M fault simulator in the lab at the University of Ottawa, with the setup shown in figure 9(a). Two ER10K bearings are installed to support the shaft and the load is 5.03 kg. The left bearing has an outer race fault, and the rest of the detailed parameters are shown in table 1. The sampling frequency is 24 kHz and the test lasts for 7 s, in which the shaft speed increases linearly from 18 to 39 Hz.

Similar to the simulated signal, figure 10(a) shows the signal of the experimental platform and figure 10(b) shows the TFR obtained by the STFT. Figure 11 shows the FCF and its harmonics extracted from figure 10(b) by the peak search method. As can be observed in figure 11, the harmonics of FCF cannot be accurately extracted from the TFR in figure 10(b) due to noise interference.

The signal is then processed by the method proposed in this paper. The signal is first processed to find the demodulation

coefficients, as shown in figure 12. The TFR of the synchronous GD of the signal components corresponding to the peaks is shown in figure 13. The TF ridges in figure 13 are then backward demodulated to obtain the final TFR in figure 14(b). Figure 14(a) shows the TFR of the signal generated by the traditional GD method with iteration and filtering operations. By comparing figures 14(a) and (b), it can be seen that the proposed method can enhance and sharpen the TF ridges.

Figure 15 shows the IF and its harmonic curves of the TFR obtained by the proposed method. As can be seen in figure 15, there are few deviations between the extracted IF curves and the true frequency curves.

The MRE of the extracted IFs is calculated to be 1.1%. The fault characteristic coefficient can finally be determined by the ratio of the second lowest IF to the lowest IF, which is 3.08, almost equal to the actual fault characteristic coefficient 3.05. It can then be concluded that there is an outer race fault on the bearing, indicating that the proposed method can be used to extract the fault features of the bearing.

Comparing the MREs of the extracted IFs in figures 11 and 15, it is improved from 9.7% to 1.1%, verifying that the proposed method can effectively enhance the TF ridges and then facilitate the accurate extraction of IF ridges.

4.2. Inner race fault detection

For bearing inner race fault diagnosis, data of two different cases downloaded from [31] are analysed. Two experiments with different shaft rotating frequency changing patterns have been conducted on the experimental setup shown in figure 9(b). Two ER16K bearings are installed to support the shaft and the load. The bearing with the inner raceway failure is located on the left side. The gearbox is connected by a belt.

A load mass of 5.03 kg was installed on a 25.4 mm steel shaft. The data was fed to an NI data acquisition module (NI USB-6212 BNC) and recorded by a computer using LabVIEW software [6]. The specific bearing parameters are shown in table 2.

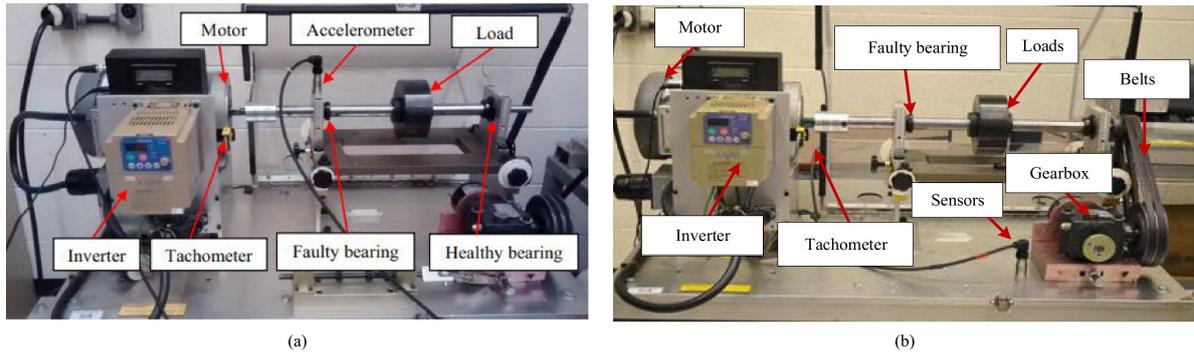


Figure 9. Experimental setup of (a) bearing outer race fault and (b) bearing inner race fault.

Table 1. Parameters of the bearings for the experiment.

| Parameter | Bearing type | Fault type | Pitch diameter (mm) | Ball diameter (mm) | Number of balls | FCF |
|-----------|--------------|------------------|---------------------|--------------------|-----------------|----------------|
| Value | ER10K | Outer race fault | 33.50 | 7.94 | 8 | 3.052 <i>f</i> |

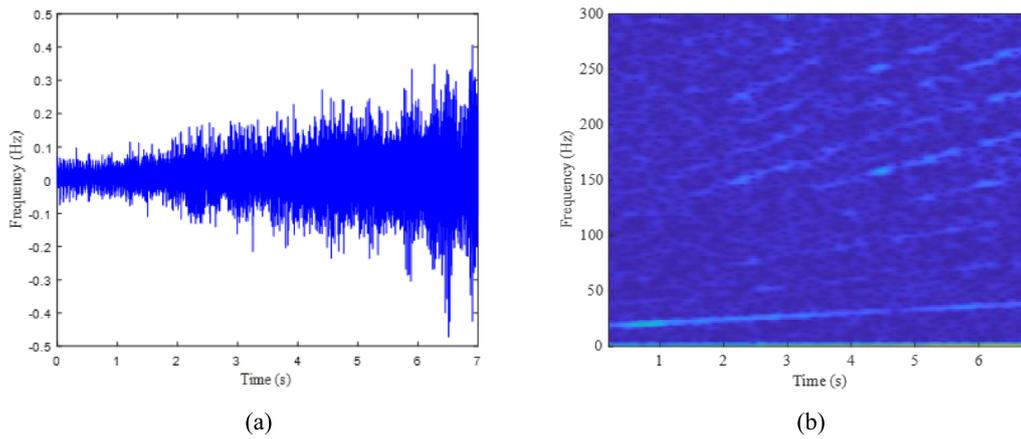


Figure 10. Outer experimental signal and its TFR: (a) time domain representation of the experimental signal; (b) TFR of the experimental signal.

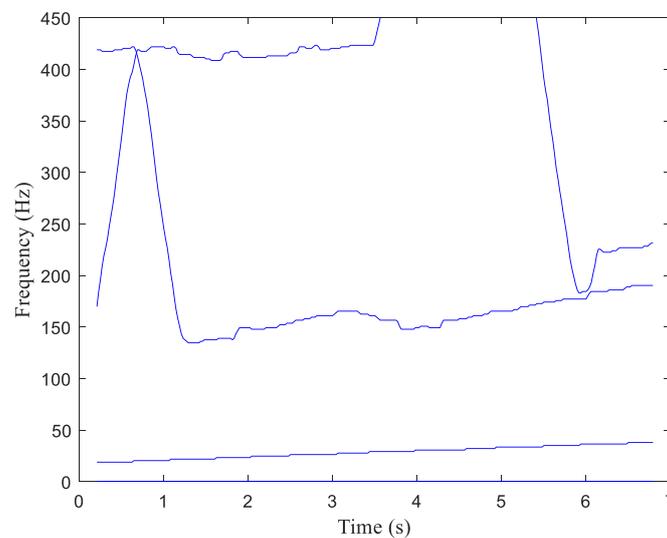


Figure 11. TF ridges of the experimental signal.

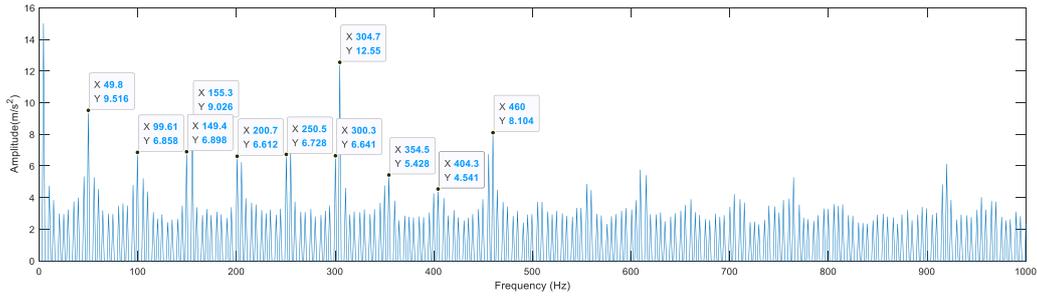


Figure 12. F–A diagram of the outer experimental signal.

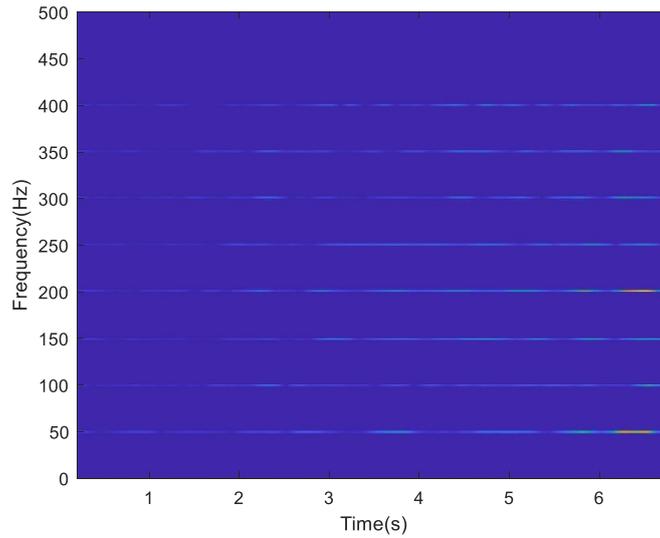


Figure 13. TFR after synchronous GD.

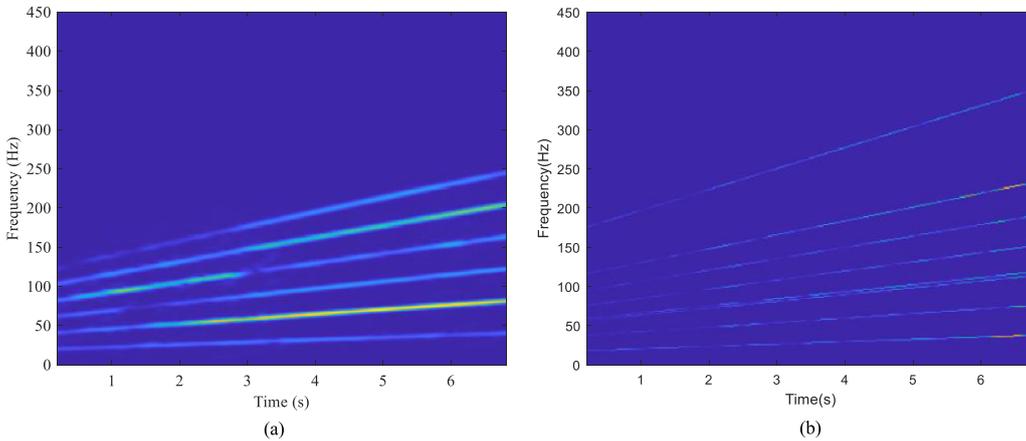


Figure 14. (a) TFR obtained by the traditional GD method with filtering; (b) TFR obtained by the proposed method.

Case 1

For the first case, the sampling frequency is 200 kHz and the test lasts for 4 s, in which the shaft rotational speed frequency is increased from 15.1 to 24.3 Hz and then decreased again to 18.7 Hz.

Similar to the simulated signal, figure 16(a) shows the signal of the experimental platform and figure 16(b) shows the

TFR obtained by STFT. Figure 17 shows the ridges extracted from figure 16(b) by the peak search method. As can be observed in figure 17, the FCF cannot be accurately extracted from the TFR in figure 16(b) due to the noise interference.

Therefore, the signal is processed by the method proposed in this paper. Similar to the simulated signal, from figure 18, the ratio of the peaks with higher frequency to the

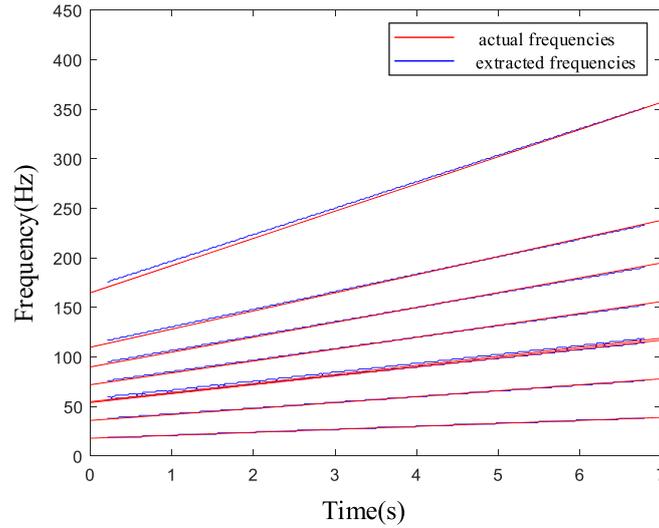


Figure 15. IF and its harmonic curves of the TFR obtained by the proposed synchronous GD (red solid lines denote the actual IFs and blue solid lines denote the extracted IFs).

Table 2. Parameters of the bearings for the experiment.

| Parameter | Bearing type | Fault type | Pitch diameter(mm) | Ball diameter(mm) | Number of balls | FCF |
|-----------|--------------|------------------|--------------------|-------------------|-----------------|----------|
| Value | ER16K | Inner race fault | 38.52 | 7.94 | 9 | 5.43 f |

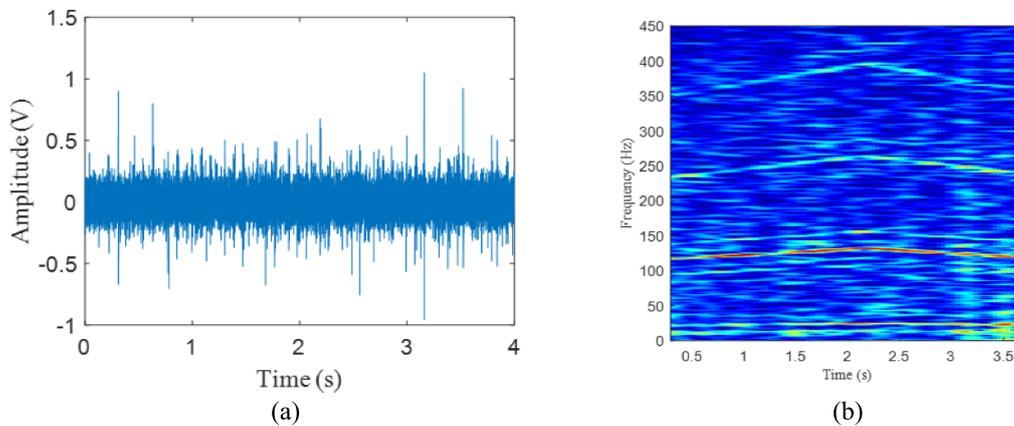


Figure 16. Experimental signal of bearing inner race fault in case 1 and its TFR: (a) time domain representation of the experimental signal; (b) TFR of the experimental signal.

frequency corresponding to the first peak can be determined. The demodulation coefficients are then pre-determined for the next synchronous demodulation and extraction. Figure 19 shows the TFR of the synchronous GD of the signal components corresponding to the peak in figure 18.

Figure 20(a) shows the TFR obtained by the traditional GD method with filtering. However, TF ridges cannot accurately

show the changes of IFs, particularly for those of higher frequencies. Figure 20(b) is the TFR of the signal after processing by the proposed method and compared with figure 20(a), the proposed method can directly extract the TF component of interest and enhance its TF aggregation. Figure 21 shows the IF and its harmonic curves extracted from the TFR obtained by the proposed method, which matches the actual IFs well.

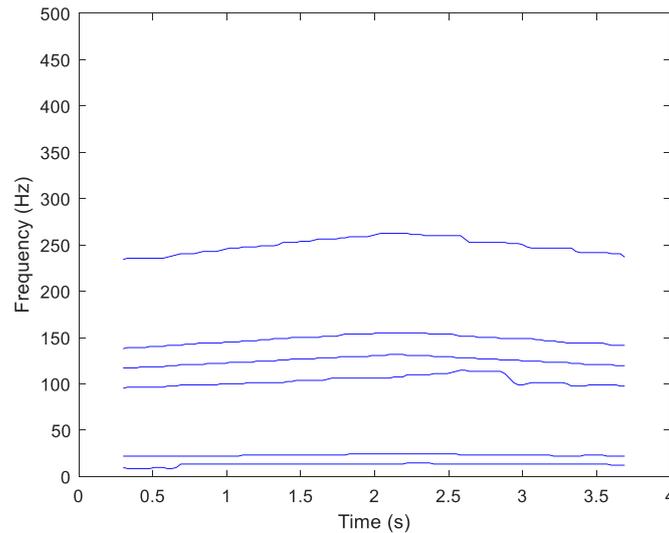


Figure 17. TF ridges of the experimental signal in case 1.

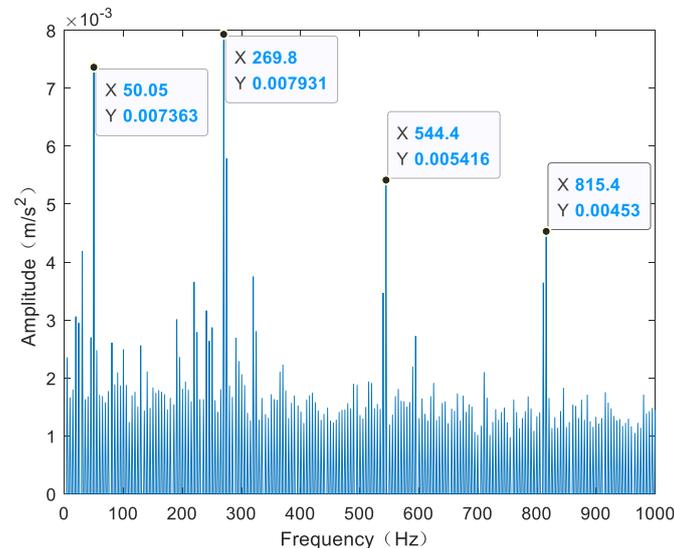


Figure 18. F-A diagram of the inner experimental signal in case 1.

To quantitatively evaluate the accuracy of extracted IFs from the TFR in figure 20(b), MRE is calculated to be 3.0%. Similar to the outer race fault case, the fault characteristic coefficient can also be calculated, which is 5.41, almost identical to the actual value of 5.43. Thus, it can be concluded that the proposed method can be used to extract the fault features of a bearing with an inner race fault and then perform the bearing fault diagnosis task.

Comparing the MREs of the extracted IFs in figures 17 and 21, the results are improved from 4.2% to 3.0%, verifying that the proposed method can effectively enhance the TF ridges and accurately extract shaft IF, IFCF and their harmonics.

Case 2

Using the same experimental setup for case 1, the vibration signals under different speed-varying conditions are measured. The shaft rotational frequency in this case decreases from 25.3 to 14.8 Hz and then increases again to 19.4 Hz.

Figure 22(a) shows the signal of the experimental platform and figure 22(b) shows the TFR obtained by STFT. Figure 23 shows the ridges extracted from figure 22(b) by the peak search method. As can be observed in figure 23, the FCF cannot be accurately extracted from the TFR in figure 22(b) due to the noise interference.

Therefore, the signal is processed by the method proposed in this paper. Similar to the simulated signal, the F-A diagram

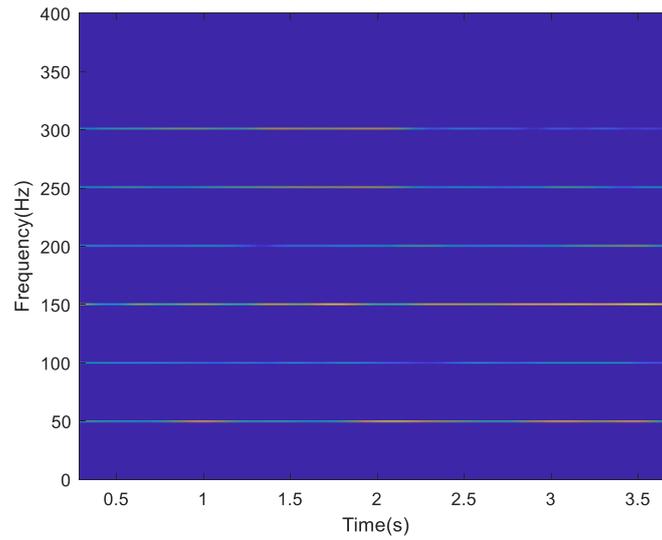


Figure 19. TF diagram after synchronous GD in case 1.

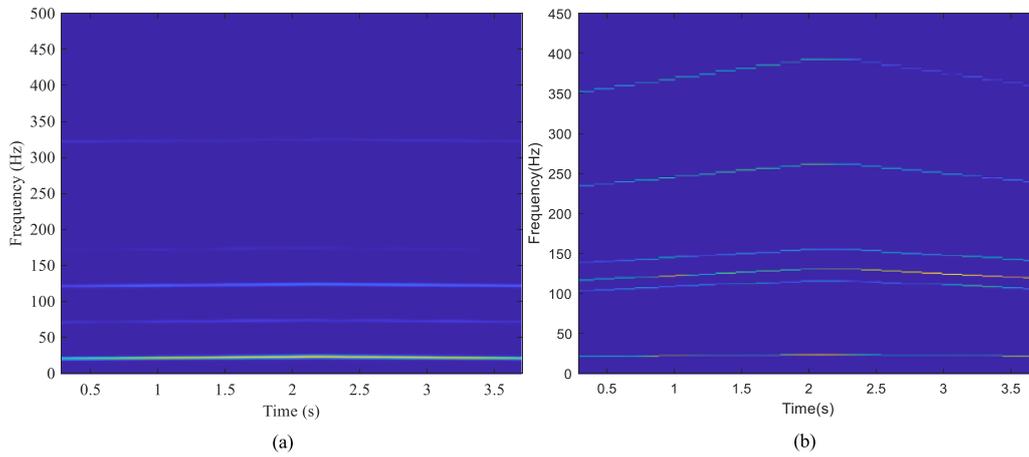


Figure 20. (a) TFR obtained by the traditional GD method with filtering in case 1; (b) TFR obtained by the proposed method in case 1.

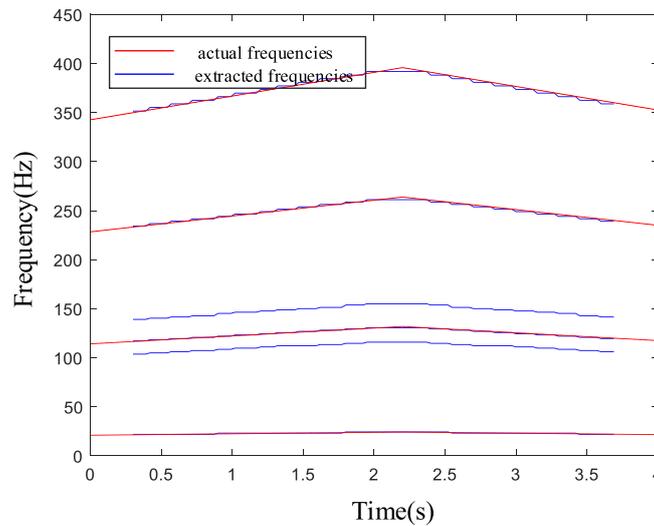


Figure 21. IF and its harmonic curves of the TFR obtained by the proposed synchronous GD in case 1 (red solid lines denote the actual IFs and blue solid lines denote the extracted IFs).

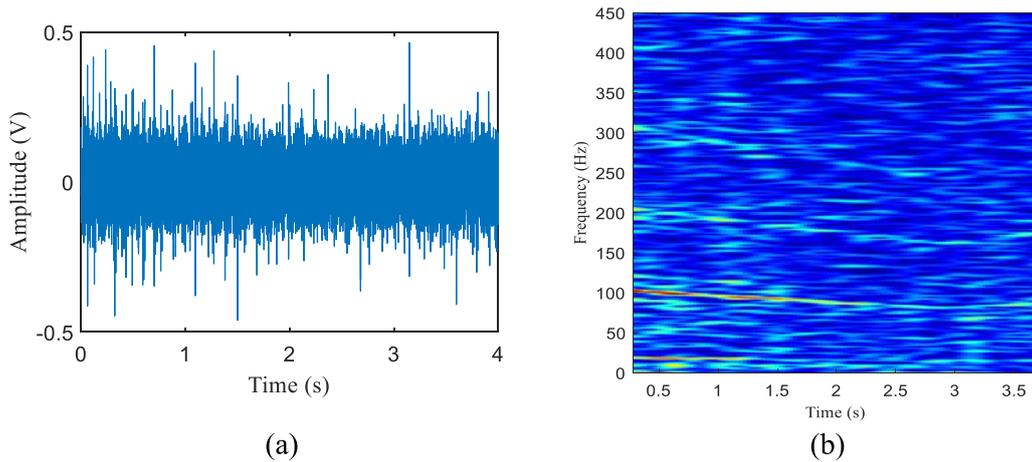


Figure 22. Experimental signal of bearing inner race fault in case 2 and its TFR: (a) time domain representation of the experimental signal; (b) TFR of the experimental signal.

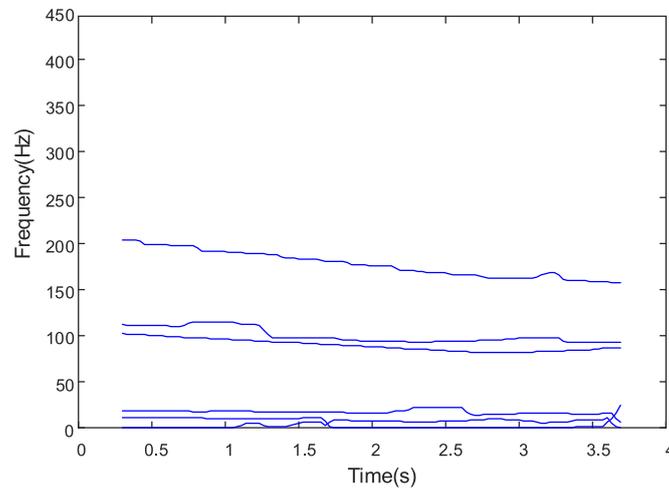


Figure 23. TF ridges of the experimental signal.

is obtained, as displayed in figure 24. The demodulation coefficients are then pre-determined for the next synchronous demodulation and extraction. Figure 25 shows the TFR of the synchronous GD of the signal components corresponding to the peaks labeled in figure 24.

Figure 26(a) shows the TFR obtained by the traditional GD method with filtering. However, TF ridges cannot accurately show the changing pattern of IF ridges. Figure 26(b) is the TFR of the signal processed by the proposed method. Compared with figure 26(a), the proposed method can accurately extract the TF components of interest and enhance their TF aggregation. Figure 27 shows the IFs extracted from the TFR obtained by the proposed method, which matches the actual IFs well.

To quantitatively evaluate the accuracy of extracted IFs from the TFR in figure 26(b), MRE is calculated to be 2.3%. Similar to the outer race fault case and the inner fault detection case 1, the fault characteristic coefficient can also be calculated, which is 5.45, almost identical to the actual value of 5.43. Thus, it can be concluded that the proposed method can be used to extract the fault features of a bearing with an inner race fault and then perform the bearing fault diagnosis task.

In conclusion, the MRE of all extracted IFs shown in figure 27 is 2.3%, which is greatly improved (the MRE of all extracted IFs in figure 23 is 7.2%), indicating that the proposed method can enhance the TF ridges and accurately extract IF ridges.

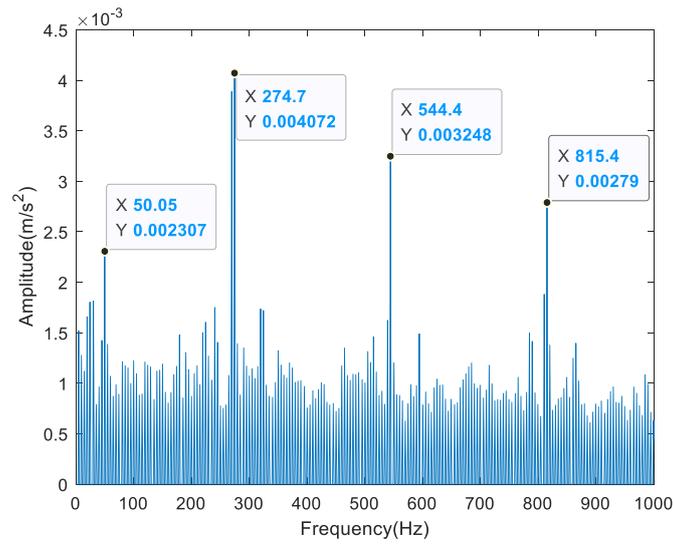


Figure 24. F–A diagram of the inner experimental signal in case 2.

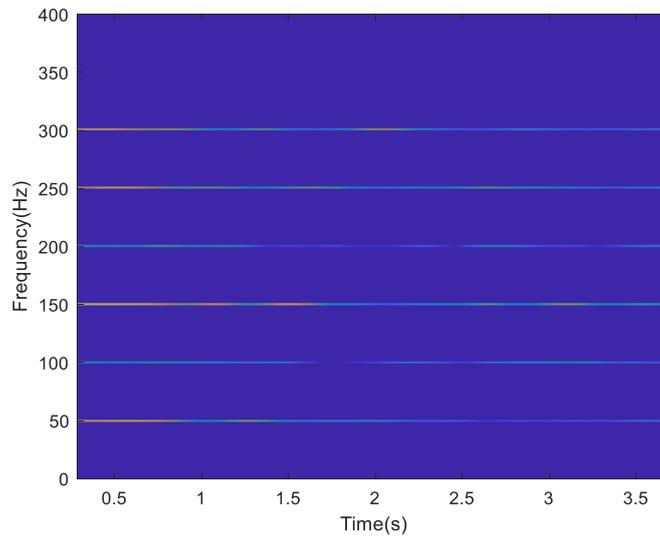


Figure 25. TF diagram after synchronous GD in case 2.

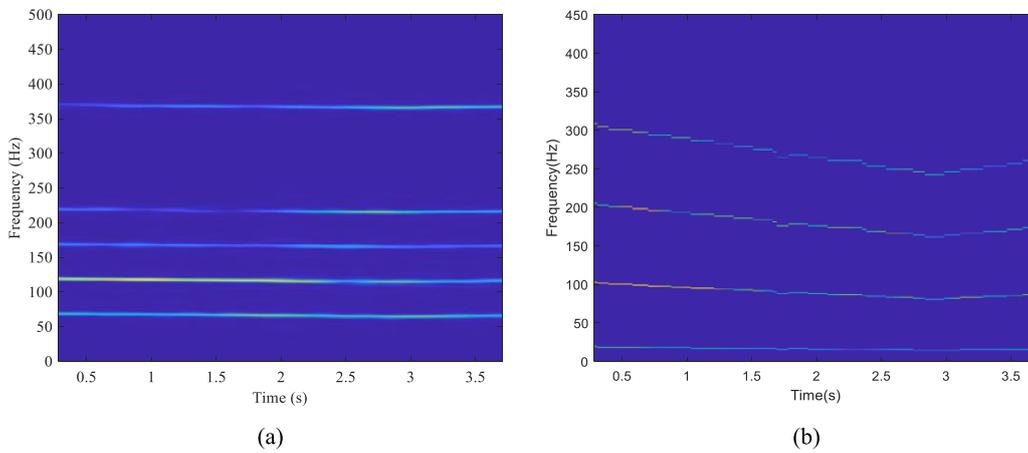


Figure 26. (a) TFR obtained by the traditional GD method with filtering in case 2; (b) TFR obtained by the proposed method in case 2.

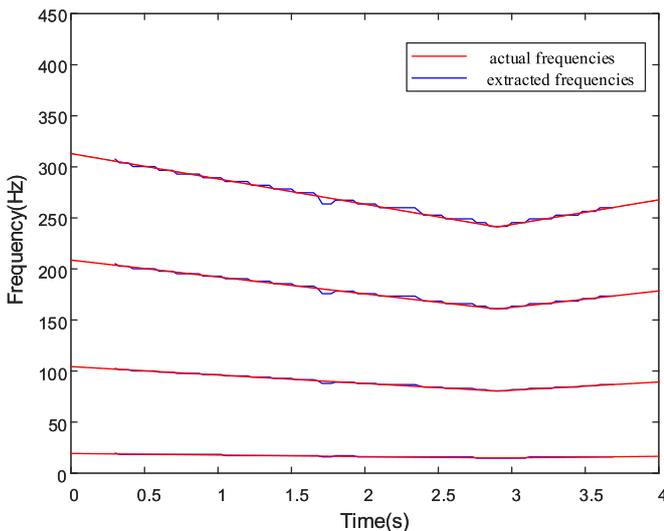


Figure 27. IF and its harmonic curves of the TFR obtained by the proposed synchronous GD in case 2 (red solid lines denote the actual IFs and blue solid lines denote the extracted IFs).

5. Conclusion

This paper proposes a method for feature enhancement that can be applied to bearing fault diagnosis under time-varying speeds without the use of a tachometer, pre-filtering and iteration operations. The proposed method first develops a strategy to realize the synchronous GD for multi-component signals using the Hadamard product between matrices. To further improve the TFR readability, a synchronous extractor independent of the estimation of IFs is proposed to further sharpen the TF ridges to obtain the post-processed TF ridges. The improved TFR with actual IFs is then achieved by backward demodulating the post-processed TF ridges.

The main advantages of this work are summarized as follows: (a) synchronous GD of multi-component vibration signals via constructing a demodulator matrix; (b) construction of the synchronous extractor independent of the IF estimation for TFR readability improvement; (c) accuracy improvement of fault characteristic coefficients led by the improved and sharpened TFR. The effectiveness of the proposed method has been tested using both simulated and experimental data from rolling bearings.

Data availability statement

Data of bearing inner race fault are downloaded from this link: <https://www.sciencedirect.com/science/article/pii/S2352340918314124>. Data of bearing outer race fault are collected in the corresponding author's previous lab and available upon reasonable request from authors.

Acknowledgments

This work is supported by the open project of State Key Laboratory of Performance Monitoring and Protecting of

Rail Transit Infrastructure, East China Jiaotong University (Grant No. HJGZ2021112) and the National Nature Science Foundation of China (Grant No. 52175056).

ORCID iDs

Juanjuan Shi  <https://orcid.org/0000-0001-8634-9083>
 Changqing Shen  <https://orcid.org/0000-0002-5143-8366>
 Weiguo Huang  <https://orcid.org/0000-0002-6734-2019>
 Zhongkui Zhu  <https://orcid.org/0000-0001-9827-4154>

References

- [1] Zhao D, Wang T and Chu F 2020 Adaptive generalized demodulation transform based rolling bearing time-varying nonstationary fault feature extraction *Chin. J. Mech. Eng.* **56** 80–87
- [2] Borghesani P, Ricci R, Chatterton S and Pennacchi P 2013 A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions *Mech. Syst. Signal Process.* **38** 23–35
- [3] Pons-Llinares J, Antonino-Daviu J A, Riera-Guasp M, Bin Lee S, Kang T-J and Yang C 2015 Advanced induction motor rotor fault diagnosis via continuous and discrete time-frequency tools *IEEE Trans. Ind. Electron.* **62** 1791–802
- [4] Vedreno-Santos F, Riera-Guasp M, Henao H, Pineda-Sanchez M and Puche-Panadero R 2014 Diagnosis of rotor and stator asymmetries in wound-rotor induction machines under nonstationary operation through the instantaneous frequency *IEEE Trans. Ind. Electron.* **61** 4947–59
- [5] Choi H-I and Williams W J 1989 Improved time-frequency representation of multicomponent signals using exponential kernels *IEEE Trans. Signal Process.* **37** 862–71
- [6] Zhao Y, Atlas L E and Marks R J 1990 The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals *IEEE Trans. Signal Process.* **38** 1084–91
- [7] Jeong J and Williams W J 1992 Kernel design for reduced interference distributions *IEEE Trans. Signal Process.* **40** 402–12
- [8] Stankovic L 1996 Auto-term representation by the reduced interference distributions: a procedure for kernel design *IEEE Trans. Signal Process.* **44** 1557–63
- [9] Huang N, Wu M, Long S, Shen S S P, Qu W, Gloersen P and Fan K L 2003 A confidence limit for the empirical mode decomposition and Hilbert spectral analysis *Proc. Math. Phys. Eng. Sci.* **459** 2317–45
- [10] Rilling G and Flandrin P 2008 One or two frequencies? The empirical mode decomposition answers *IEEE Trans. Signal Process.* **56** 85–95
- [11] Deng Y, Wang W, Qian C, Wang Z and Dai D 2001 Boundary-processing-technique in EMD method and Hilbert transform *Chin. Sci. Bull.* **46** 954–60
- [12] Flandrin P, Rilling G and Goncalves P 2004 Empirical mode decomposition as a filter bank *IEEE Signal Process. Lett.* **11** 112–4
- [13] Feng Z, Chu F and Zuo M J 2011 Time-frequency analysis of time-varying modulated signals based on improved energy separation by iterative generalized demodulation *J. Sound Vib.* **330** 1225–43
- [14] Li Y, Zhang W, Xiong Q, Luo D, Mei G and Zhang T 2017 A rolling bearing fault diagnosis strategy based on improved multiscale permutation entropy and least squares SVM *J. Mech. Sci. Technol.* **31** 2711–22

- [15] Yu G, Yu X and Xu C 2017 Synchroextracting transform *IEEE Trans. Ind. Electron.* **64** 8042–54
- [16] Daubechies I, Lu T and Wu H 2011 Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool *Appl. Comput. Harmon. Anal.* **30** 243–61
- [17] Shi J, Hua Z and Shen C 2021 A generalized instantaneous-frequency-estimation-free stepwise demodulation transform and its application in vibration signal analysis of rotating machinery *J. Sound Vib.* **40** 1–11, 21
- [18] Li Z, Liu Y, Hu X and Wen C 2021 Empirical wavelet transform-synchroextracting transform and its applications in fault diagnosis of rolling bearing *J. Vib. Eng.* **34** 1284–92
- [19] Wang S, Chen X, Cai G, Chen B, Li X and He Z 2014 Matching demodulation transform and synchrosqueezing in time-frequency analysis *IEEE Trans. Signal Process.* **62** 69–84
- [20] Shi J, Liang M and Guan Y 2016 Bearing fault diagnosis under variable rotational speed via the joint application of windowed fractal dimension transform and generalized demodulation: a method free from prefiltering and resampling *Mech. Syst. Signal Process.* **68** 15–33
- [21] Feng Z and Chen X 2015 Iterative generalized synchrosqueezing transform for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions *Mech. Syst. Signal Process.* **52/53** 360–75
- [22] Olhede S and Walden A 2005 A generalized demodulation approach to time-frequency projections for multicomponent signals *Proc. R. Soc. A* **461** 2159–79
- [23] Shi J, Liang M, Neculescu D and Guan Y 2016 Generalized stepwise demodulation transform and synchrosqueezing for time–frequency analysis and bearing fault diagnosis *J. Sound Vib.* **368** 202–22
- [24] Zhao D, Li J and Cheng W 2019 Bearing and gear compound faults diagnosis based on the Vold-Kalman generalized demodulation under time-varying speeds *J. Shock. Vib.* **38** 172–8
- [25] Liu D, Cheng W and Wen W 2018 Generalized demodulation with tunable E-factor for rolling bearing diagnosis under time-varying rotational speed *J. Sound Vib.* **430** 59–74
- [26] Li H, Tang B, Han Y and Deng L 2018 Fault diagnosis of rolling bearings under a variable rotating speed based on iterative generalized demodulation gear signal separation *J. Shock. Vib.* **37** 38–44
- [27] Liu D, Cheng W, Zhang J, Gao R X and Wen W 2021 Integrated method of generalized demodulation and artificial neural network for robust bearing fault recognition *Proc. Manuf.* **53** 628–35
- [28] Pi W, Yu D and Peng F 2010 Generalized demodulation method based on multi-scale chirplet and sparse signal decomposition and its application to gear fault diagnosis *J. Mech. Eng.* **46** 59
- [29] Liu D, Cui L and Cheng W 2023 Flexible iterative generalized demodulation filtering for the fault diagnosis of rotating machinery under nonstationary conditions *Struct. Health Monit.* **22** 1421–36
- [30] Liu K, Shi J, Shen C and Zhu Z 2022 Fault feature enhancement and extraction using generalized demodulation guided by ridge scale factor probability distribution *IEEE Int. Conf. on Sensing Diagnostics Prognostics, and Control* pp 43–47
- [31] Huang H and Baddour N 2018 Bearing vibration data collected under time-varying rotational speed conditions *Data Br.* **21** 1745–9