



Time-frequency representation based on robust local mean decomposition for multicomponent AM-FM signal analysis

Zhiliang Liu ^{a,*}, Yaqiang Jin ^a, Ming J. Zuo ^{a,b}, Zhipeng Feng ^c

^a School of Mechatronics Engineering, University of Electronic Science and Technology of China, Chengdu 611731, PR China

^b Department of Mechanical Engineering, University of Alberta, Edmonton T6G2G8, Canada

^c School of Mechatronics Engineering, University of Science and Technology Beijing, Beijing 100083, PR China

ARTICLE INFO

Article history:

Received 9 October 2016

Received in revised form 12 March 2017

Accepted 24 March 2017

Keywords:

Time frequency representation

Local mean decomposition

Multicomponent AM-FM signal

Fast kurtogram (spectral kurtosis)

Machinery fault diagnosis

ABSTRACT

Local mean decomposition (LMD) is a promising approach to implement time-frequency representation (TFR) for multicomponent amplitude-modulated (AM) and frequency-modulated (FM) signal analysis; however, its performance usually suffers from end effect and mode mixing problems. To address this issue, this paper proposes a novel comprehensive scheme to improve LMD performance. The novel scheme can automatically determine the fix subset size of the moving average algorithm and the optimal number of sifting iterations in a sifting process. Extensive simulations have been explored for multicomponent AM-FM signal analysis by means of TFR with the improved LMD. Moreover, the improved LMD shows potential application in bearing fault diagnosis in conjunction with the well-known fast kurtogram.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

One of the most important aspects in signal processing is to extract and present useful information from a signal, usually as a function of time, from physical or mechanical systems. Amplitude-modulated (AM) and frequency-modulated (FM) signals are such useful information, and thus, accurate location and representation of AM-FM signals is quite important to interpret systems' health status.

Frequency representation occupies a central role in the vast subject of signal processing. The combination of Hilbert transform (HT) and Fourier transform (FT) is a common approach of frequency representation for AM-FM signals. However, there are two limitations when it is used to analyze multicomponent AM-FM signals [1]. First, HT estimates the instantaneous amplitude and frequency data for monocomponent signals only, and thus, for multicomponent signals, the interference among multiple components leads to a mode mixing problem in HT. Second, the representation of the FM part is more crucial compared to the representation of the AM part due to the fact that the parametric estimation of the time-varying amplitude given the time-varying frequency is always linear. Therefore, locating and representing useful information for multicomponent AM-FM signals is quite challenging in signal processing. The former limitation concerns information location. The most common solution uses signal separation methods, e.g., filters, to extract monocomponent signals from multicomponent signals. The aforementioned HT and FT can further process each monocomponent. The latter limitation concerns information presentation. To alleviate this limitation, time-frequency representation (TFR) is an approach to depict

* Corresponding author.

E-mail address: Zhiliang_Liu@uestc.edu.cn (Z. Liu).

AM-FM signals. Recently, a new trend of combining signal separation and the TFR method as a hybrid solution, e.g., Hilbert-Huang transform (HHT), has emerged to address the two aforementioned limitations for multicomponent AM-FM signals. Inspired by HHT, local mean decomposition (LMD), which Smith [2] pioneered, demodulates multicomponent AM-FM signals into a set of product functions (PFs), each of which is the product of an envelope signal and FM signal. Each PF can be considered a monocomponent AM-FM signal, and TFR can be derived based on instantaneous amplitude and frequency data from all of the PFs. LMD has many promising properties; in particular, it avoids the limitation of doubtful negative instance frequency occurring in HT [3] because LMD can generate instantaneous amplitude and frequency data without using HT, as HHT does. Moreover, LMD is specifically designed for AM-FM signals, whereas this property is not reflected in HHT. From the above evidence, LMD is a suitable and promising technique to process multicomponent AM-FM signals [4,5]. However, end effect and mode mixing problems are two main limitations in LMD. These two limitations can be alleviated by proper parameter selection with regard to boundary condition, envelope estimation, and sifting stopping criterion [6].

Most reported studies that aim to improve the performance of LMD can be grouped into three categories corresponding to the above three parameters. Among them, boundary condition and envelope estimation have been the two main focuses over the past ten years. Regarding boundary condition, most researchers try to extend the left and right side ideally by a variety of methods, e.g., the boundary processing method [7], mirror extension method [8], neural networks [9], self-adaptive waveform matching [10], and integral local waveform matching [11]. Regarding envelope estimation, researchers have put forward new methods in roughly two directions. One uses the moving average algorithm as the smoothing algorithm and tries to find an optimal fixed subset size for better envelope estimation, e.g., a third of the longest local mean [2], and a local mean step-based strategy [7]. The other focuses on alternative smoothing algorithm design, e.g., the rational Hermite interpolation method [12], B-spline interpolation method [13], cubic spline interpolation method [14,15], and cubic Hermite interpolation method [16]. In addition, Cheng et al. [17] proposed an ensemble LMD that uses noise to alleviate the limitation of mode mixing. Smith [2] already defined the principle of sifting stopping criterion in LMD; however, it is so ideal that cannot be directly realized in a numerical computation. Some referable criteria have already been proposed and used in EMD, e.g., the standard deviation criterion [18], three-threshold criterion [6], bandwidth criterion [19], and orthogonality criterion [20]. The sifting process is similar between EMD and LMD; however, the effectiveness of the above criteria still needs to be validated when they are embedded in LMD. Regarding sifting stopping criterion, a few methods that transform the original principle of sifting stopping criterion to a practical expression have been reported. Zhang et al. [21] introduced an orthogonality criterion into LMD. Cheng et al. [5] proposed a two-condition criterion that makes decisions based on evaluation of condition 1: $1 - \delta \leq \text{smoothed local magnitude (i.e., } a_{1p}(n) \text{ in Section 2)} \leq 1 + \delta$ and condition 2: $-1 \leq \text{pure frequency signal (i.e., } s_{1p}(n) \text{ in Section 2)} \leq 1$, where δ is a predefined variation. Li et al. [22] and Jiang et al. [23] have adopted this criterion with the condition 1.

It is notable that boundary condition, envelope estimation, and sifting stopping criterion are all associated with the end effect and the mode mixing problems. In other words, LMD performance cannot be thoroughly enhanced unless all three parameters are properly tuned. Considering the key role of LMD in the proposed method, a comprehensive solution is proposed in this paper to improve the performance of TFR for multicomponent AM-FM signal analysis. Finally, the proposed method will be extensively validated on simulated multicomponent AM-FM signals and tested for its effectiveness in bearing fault diagnosis.

This paper is organized as follows. Section 1 introduces the background, motivation and a brief literature review of the progress of current research. Section 2 introduces the theoretical basis of LMD, which is necessary in the following sections. Section 3 proposes the robust LMD based TFR, and improvements on LMD's three aspects are also detailed. Numerical validations with simulated signals and real-world signals are carried out in Sections 4 and 5, respectively. Finally, Section 6 provide the final conclusion of this paper.

2. LMD fundamentals

LMD is an iterative signal process that aims to extract a set of “best fit” product functions (PFs) of a pure FM signal and an envelope signal. The algorithm of LMD includes a nested loop structure, which is a loop within a loop, i.e., an inner loop within the body of an outer one. The outer loop in LMD aims to extract PFs from an original discrete signal, whereas the inner loop aims to extract pure FM and envelope signals from a PF. In the following, we briefly introduce LMD, which provides basic fundamentals for the proposed scheme. The materials in this section are mostly based on [2].

With a sampling frequency F_s or sampling period T_s ($T_s = 1/F_s$), a continuous signal $x(t)$ is sampled to a discrete signal $x(n)$, where $n = 1, 2, 3, \dots, N_s$, and N_s is the total number of sampling points. The signal $x(n)$ can be decomposed into a set of PFs by LMD with the following steps.

Step 1. Find all local extremes, including maxima and minima, in the signal $x(n)$. The extreme points are denoted as e_k , and the corresponding extreme values are denoted as $x(e_k)$, where $k = 1, 2, 3, \dots$

Step 2. Calculate the smoothed local mean $m(n)$ and the smoothed local magnitude $a(n)$. This is achieved through two substeps. First, the preprocessed local mean $m^0(n)$ and local magnitude $a^0(n)$ are computed following Eqs. (1) and (2), respectively. Obviously, the above two signals are step functions.

$$m^0(n) = \frac{x(e_k) + x(e_{k+1})}{2} \text{ for } e_k \leq n < e_{k+1}. \quad (1)$$

$$a^0(n) = \frac{|x(e_k) - x(e_{k+1})|}{2} \text{ for } e_k \leq n < e_{k+1}. \quad (2)$$

Secondly, $m^0(n)$ and $a^0(n)$ are postprocessed by a smoothing algorithm to generate the smoothed local mean and local magnitude, which are denoted by $m(n)$ and $a(n)$, respectively. Smith [2] suggests the moving average algorithm that can surpass the cubic spline method commonly used in EMD as the smoothing algorithm in LMD, and it can be viewed as a low-pass filter, the cut-off frequency of which is determined by an important parameter, i.e., fixed subset size λ .

Step 3. Calculate the estimated zero-mean signal $h_{11}(n)$ and FM signal $s_{11}(n)$ with the three given signals: $x(n)$, $m(n)$ and $a(n)$. This is achieved through two substeps. First, subtract the signal $m_{11}(n)$ from the original data $x(n)$, and get the estimated zero-mean signal $h_{11}(n)$. Secondly, divide the signal $h_{11}(n)$ by $a_{11}(n)$, and get the estimated FM signal $s_{11}(n)$. The subscripts in $h_{ij}(n)$, $m_{ij}(n)$, $s_{ij}(n)$, and $a_{ij}(n)$ refer to the i th PF and j th sifting process.

The signal $s_{11}(n)$ may not be a pure FM signal after the first round local mean decomposition. If this is the case, we need to use $s_{11}(n)$ as the new signal and repeat the steps 1–3 for p times until the criterion in Eq. (3), which means that the envelope is flat, is satisfied. This is the so-called *sifting process* or the *inner loop* of LMD, as we previously mentioned.

$$\lim_{p \rightarrow \infty} a_{1p}(n) = 1. \quad (3)$$

Step 4. Calculate the pure FM signal $s_1(n)$, envelope signal $a_1(n)$, and product function PF_1 . If the criterion in Eq. (3) is satisfied, the estimated envelope $a_{1p}(n)$ is flat, and the signal $s_{1p}(n)$ is a pure FM signal. Therefore, the extracted pure frequency-modulation signal and corresponding envelope signal can be derived from Eqs. (4) and (5). Multiplying $a_1(n)$ and $s_1(n)$ gives the first product function $PF_1(n)$.

$$s_1(n) = s_{1p}(n). \quad (4)$$

$$a_1(n) = \prod_{j=1}^p a_{1j}(n). \quad (5)$$

Step 5. Extract all of the PFs from the signal $x(n)$. As the natural signal usually contains more than one PFs, the rest of the PFs need to be extracted by the following procedures. The signal $PF_1(n)$ is subtracted from the original signal $x(n)$, and then a residual signal $u_1(n)$ is repeated in the steps 1–5 for q times until $u_q(n)$ is a constant or contains no oscillations. This is the so-called *outer loop* of LMD. Finally, the signal $x(n)$ can be reconstructed according to Eq. (6).

$$x(n) = \sum_{i=1}^q PF_i(n) + u_q(n). \quad (6)$$

3. Proposed robust LMD and time-frequency representation

Based on Section 2, we know how LMD deals with signal processing for multicomponent AM-FM signals. As we also mentioned in Section 1, boundary condition, envelope estimation, and sifting stopping criterion are three crucial aspects of LMD. Therefore, in this section, we propose a comprehensive solution to all three aspects to improve the performance of LMD-based TFR. To distinguish from conventional LMD, we call the proposed method *the robust LMD*. In addition, we also introduce how to implement TFR based on the robust LMD.

3.1. Proposed robust local mean decomposition

3.1.1. Boundary condition

The first step of the robust LMD is to determine local extremes, although the end points may not be the real extrema. If this happens, it induces a wrong trend in the following envelope estimation for the signal end points. This may lead to a divergence phenomenon that propagates to signal components extracted in subsequent sifting processes. As a result, these contaminated components cannot reflect characteristics of the original signal. To deal with the boundary condition, the mirror extension algorithm [6] is employed by the proposed robust LMD. The procedure of this algorithm is briefly summarized as follows.

Step 1. The first step is to determine symmetry points for the left and right ends of the signal. We use only the left end case to illustrate the idea of the mirror extension algorithm, and one can determine the right symmetry point in the same way. The mirror extension algorithm needs to differ four cases for the left end case. If the first local extreme is a maxima, we check whether the function value at the end point is smaller than the first local minimum. If it is true (case A in Fig. 1), we consider the end point as a minima and set it as the symmetry point; otherwise (case B in Fig. 1), we set the first maxima as the symmetry point. If the first local extreme is a minima, we check whether the function value at the end point is larger than

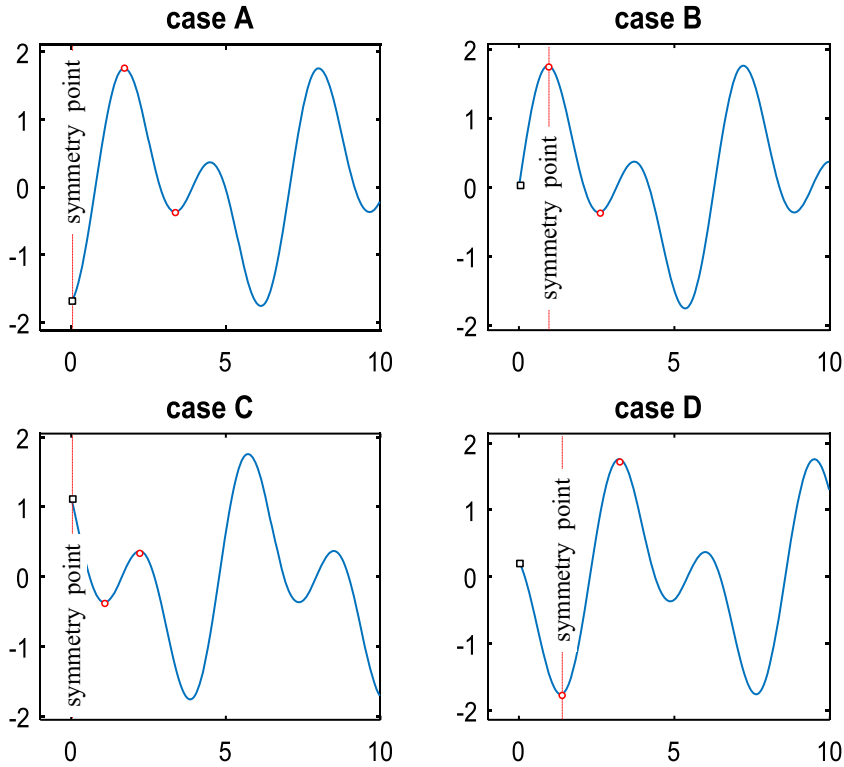


Fig. 1. Four cases to determine the symmetry point.

the first local maximum. If it is true (case C in Fig. 1), we consider the end point as a maxima and set it as the symmetry point; otherwise (case D in Fig. 1), we set the first minima as the symmetry point.

Step 2. Put “mirrors” on the symmetry point determined in the step 1 and extend the signal. The extended ratio controls the total length of the extended signal.

Step 3. Use the extended signal instead of the original signal to perform the moving average algorithm, and then cut the extended outputs back to the same length as the original signal.

3.1.2. Envelope estimation

The moving average algorithm is the smoothing algorithm of envelope estimation for the local mean and local magnitude in Eqs. (1) and (2). The fixed subset size λ is the only parameter that usually needs to be predefined with prior knowledge. It is the number of points used for taking the average. Notably, λ is an important and sensitive parameter that must be properly considered. If the value of λ is small, the envelope signal cannot be totally smoothed. The smoothed envelope signal still includes flat steps, which is exactly the case with a fixed subset size of 10 in Fig. 2. If the value of λ is large, the smoothing process will remove the local characteristics of the envelope signal. Taking the case with a fixed subset size of 1000 in Fig. 2 for example, the global increasing trend is clearly kept, whereas the local fluctuation characteristic disappears after the smoothing process. Both of the two extreme cases deteriorate envelope estimation and then lead to a poor sifting process. Thus, they cannot be adopted in LMD. A proper value of λ , e.g., 50 in Fig. 2, can achieve a relatively accurate envelope estimation. In the following, we propose a statistical method based on the statistics theory to automatically determine a reasonable fixed subset size λ^* for accurate envelope estimation.

In the proposed method, we first obtain all step lengths in the signal of $m^0(n)$ or $a^0(n)$. The step length is numerically equal to $(e_{k+1} - e_k) + 1$. It may be different for different k values. It has no relationship with the moving average algorithm. Second, we apply histogram bin counts to the step length set and obtain the probability in each bin (denoted by $S(k)$) and the bin edges (denoted by $edge(k)$). An automatic binning algorithm developed by MATLAB and embedded in the *hiscounts* function is used to determine the bin edges. Third, we define the center and standard deviation of the step length set by Eqs. (7) and (8), respectively.

$$\mu_s = \sum_{k=1}^{N_b} s(k)S(k), \quad (7)$$

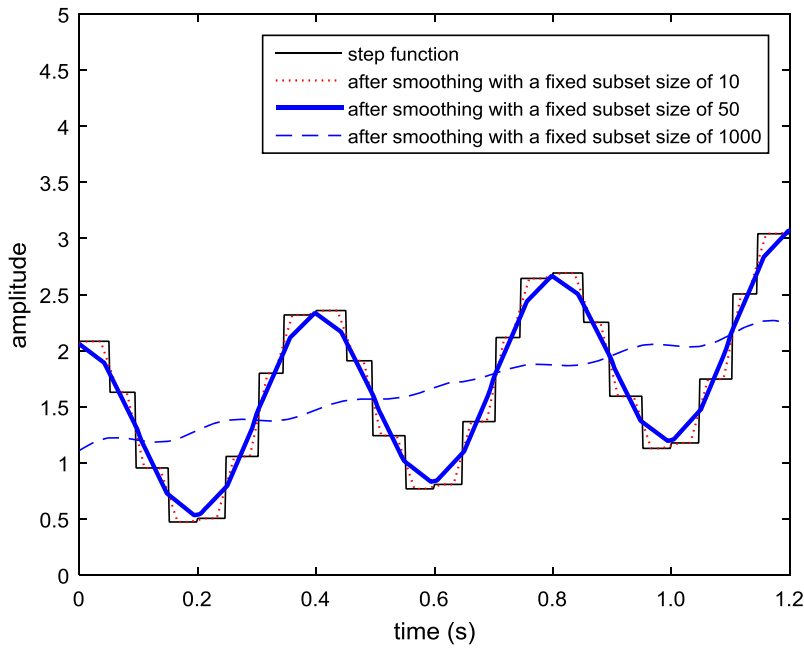


Fig. 2. Smoothing a step function with three fixed subset sizes.

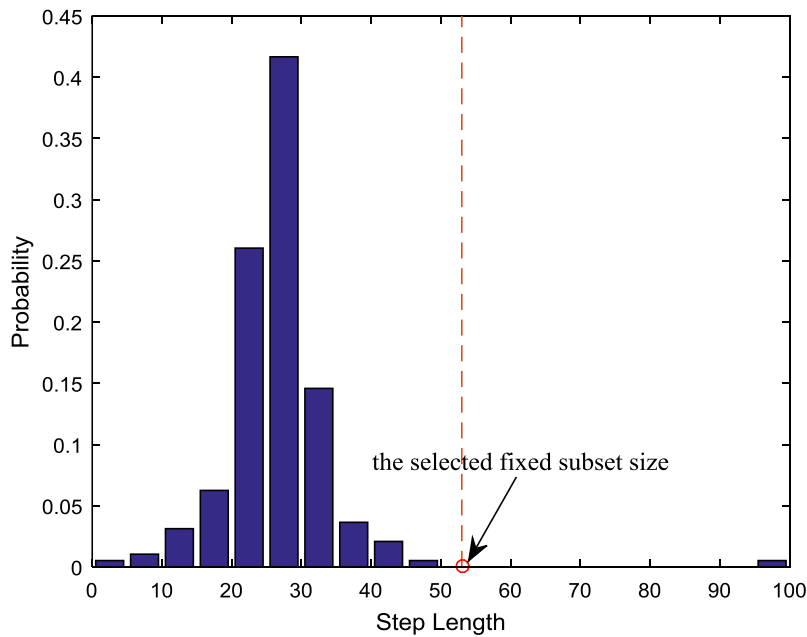


Fig. 3. Fixed subset size selection illustration.

$$\text{and } \delta_s = \sqrt{\sum_{k=1}^{N_b} (s(k) - \mu_s)^2 S(k)}, \quad (8)$$

where $s(k) = (\text{edge}(k) + \text{edge}(k+1))/2$, μ_s is the step length center, σ_s is the standard deviation of the step length, and N_b is the number of bins in the histogram.

Finally, the selected subset size λ^* is calculated by

$$\lambda^* = \text{odd}(\mu_s + 3 \times \delta_s), \quad (9)$$

where $\text{odd}()$ is a function that rounds its input to the nearest odd integer greater than or equal to the input.

We interpret Eq. (9) as follows. Fig. 3 shows an example of the selected fixed subset size determined by the proposed method. In this example, we have μ_s and σ_s equal to 26.64 and 8.13, respectively, and thus the fixed subset size is 53 by substituting the two values into Eq. (9). From the figure, the selected size is greater than most of the step lengths. This is not surprising because we introduce the three standard deviations principle into Eq. (9). A relatively large value of λ can guarantee that (1) no (almost no) flat step exists in the smoothed envelope signal, as shown in Fig. 2, and (2) local information retains in the smoothed envelope signal.

3.1.3. Sifting stopping criterion

Smith [2] defined the principle of sifting stopping criterion in Eq. (3). According to this principle, two challenges need to be solved: (1) how to describe the target signal in a practical way under the original principle and (2) how to realize a self-adaptive mechanism for the criterion. We introduce the solution as follows.

The first challenge is considered as follows. As the target signal $a_{1p}(n)$ has a baseline of one, we get a zero-baseline envelope signal by subtracting 1 from $a_{1p}(n)$. We describe the zero-baseline envelope signal from both global and local viewpoints. From a global viewpoint, root mean square (RMS) is used to measure overall energy in the zero-baseline envelope signal. RMS is a statistical measure defined as the square root of the arithmetic mean of the squares of a set of numbers. If the envelope signal is approaching one, the zero-baseline envelope signal is approaching zero, and its RMS value is approaching zero as well. Based on the original principle, RMS is expected to be as small as possible. In addition, excess kurtosis (EK) is used to measure local anomaly from a local viewpoint. EK is an adjusted version of conventional kurtosis, which is kurtosis minus three. If the envelope signal is approaching zero, the EK of the zero-baseline envelope signal is approaching zero as well. Based on the original principle, the EK is expected to be as small as possible.

The above two viewpoints are both important and necessary to implement a relatively reasonable estimation of the zero-baseline envelope signal. If only the global characteristic is considered, it may lead the envelope signal to stop at a case in which most samples are zeros, whereas a very few outliers have extremely large amplitudes. If only the local characteristic is considered, it may lead the envelope signal to stop at a case in which all samples are very similar, whereas their amplitudes are far beyond zero. Both of the two cases can deteriorate a sifting process and finally result in poor performance of LMD.

Based on the above discussions, the proposed objective function to describe the zero-baseline envelope signal is defined as follows:

$$f = \text{RMS}(z(n)) + \text{EK}(z(n)), \quad (10)$$

where $z(n) = a(n) - 1$, i.e., the zero-baseline envelope signal,

$$\text{RMS} = \sqrt{\frac{1}{N_s} \sum_{n=1}^{N_s} (z(n))^2}, \quad (11)$$

$$\text{and EK} = \frac{\frac{1}{N_s} \sum_{n=1}^{N_s} (z(n) - \bar{z})^4}{\left(\frac{1}{N_s} \sum_{n=1}^{N_s} (z(n) - \bar{z})^2\right)^2} - 3. \quad (12)$$

The second challenge is as follows. As most of the reported methods in Section 1 have done, we can simply use the objective function in Eq. (10) with a predefined threshold to establish a sifting stopping criterion. However, this straightforward approach still requires a prior knowledge of the target signal. Different thresholds may lead to a quite diverse performance of LMD, and an improper one may also lead to a very slow stop for a sifting process.

Usually, a maximum number of sifting iterations forcibly stops the sifting process; however, this kind of sifting stop introduces potential risks, mostly caused by envelope estimation. LMD must estimate the local mean $m(n)$ and local magnitude $a(n)$ in each sifting process, and thus there is a difference between the estimated and the true signals. That is, errors always exist for each time of estimation. These errors can be cumulated and amplified along with the number of sifting processes in LMD. In addition, numerical computation and the division operation used in LMD could make the sifting process even worse. We simply call this issue the *error accumulation effect* of LMD. All above aspects related to error accumulation may lead to a divergence in the sifting process. However, the objective function in Eq. (10) and the reported sifting stopping criteria do not consider a negative effect on LMD performance for a large number of sifting processes.

With the objective function, the proposed sifting stopping criterion is described as follows. For a sifting process of the i th PF, $a_{ij}(n)$ is the smoothed local magnitude generated from the j th iteration. In each iteration, the objective function value f_{ij} can be calculated by its definition in Eq. (10). The proposed sifting stopping criterion makes its decision depending on f_{ij} , f_{ij+1} , and f_{ij+2} obtained from three successive iterations. If $f_{ij+1} > f_{ij}$ and $f_{ij+2} > f_{ij+1}$, the sifting process stops and returns the corresponding results at the $(j-1)$ th iteration; otherwise, the sifting process continues until the iteration number reaches a predefined value that indicates the maximum iteration allowed in each sifting process.

3.2. Time-frequency representation based on the robust LMD

The robust LMD gives an envelope signal $a_i(n)$ and a pure FM signal $s_i(n)$ for each PF_i . Obviously, the instantaneous amplitude for the i th PF is the envelop signal $a_i(n)$. Based on the FM signal $s_i(n)$, the instantaneous phase and frequency for the i th PF can be derived from Eqs. (13) and (14), respectively.

$$\phi_i(n) = \arccos(s_i(n)). \quad (13)$$

$$f_i(n) = \frac{d\phi_i(n)}{2\pi T_s}. \quad (14)$$

Notably, the signal $s_i(n)$ must be in the range of 1 to -1 . For practical application of the robust LMD, the extreme values of the signal $s_i(n)$ can be compulsively set to ± 1 because its corresponding estimated envelope signal is approximately equal to one [2].

TFR based on the robust LMD is introduced as follows. With time scale as the horizontal axis and frequency scale as the vertical axis, a null time-frequency matrix is created. In this null matrix, indexes of the column are regarded as new “frequency” series that are discretized again, and indexes of the row are regarded as new “time” series multiplied by the sampling frequency. For all instantaneous amplitude data of each PF_i , traverse all entries in the matrix and add them to the time-frequency matrix one by one. Finally, the complete time-frequency matrix is generated and can be viewed as a spectrum.

In summary, as shown in Fig. 4, TFR based on the robust LMD can be viewed as three blocks, i.e., signal separation, demodulation analysis, and TFR. Signal separation is used to extract monocomponent AM-FM signals from multicomponent ones. Demodulation analysis is used to obtain instantaneous amplitude and frequency data for each monocomponent AM-FM signal. These two blocks are actually implemented simultaneously by the robust LMD. The block of TFR is used to fuse all the information and present it in the time-frequency domain for multicomponent signals.

4. Numerical validation with simulated multicomponent signals

In this section, we use simulated multicomponent AM-FM signals to demonstrate TFR's performance based on the robust LMD. We compare the proposed method with three reported and recognized methods: conventional LMD implemented by Smith [2], HHT implemented by Rilling [6], and short time Fourier transform (STFT) implemented by MATLAB. We randomly select four distinguishing cases of AM and FM combinations for this comparison, and the details are described as follows.

(1) Case 1: A multicomponent signal with two pure AM components

In the first case, we generate a simulated signal including two pure AM signals. The signal is expressed in Eq. (15). The signal is sampled at a frequency of 5000 Hz, and has a length of 5000 data points.

$$x(t) = 5 \times (1 + \cos(2\pi t \times 20)) \cos(2\pi t \times 500) + 15 \times (3 + \cos(2\pi t \times 20)) \cos(2\pi t \times 1500) + n(t) \quad (15)$$

where $n(t)$ is the Gaussian white noise with a signal-to-noise ratio (SNR) of 5 dB generated by the built-in function of *awgn* in MATLAB.

Figs. 5–8 show the TFR plots from the four methods. In these figures, the x-axis and y-axis represent time and frequency, respectively, and the color represents the instantaneous amplitude of a frequency at a time. Our observations from the four figures are as follows. STFT, the conventional LMD, and the robust LMD can successfully locate the 2 AM components at 500 Hz and 1500 Hz, which correspond to carriers of the two pure AM components in Eq. (15). We also observe the AM effect at the two frequencies. STFT shows a low time-frequency resolution in its TFR spectrum, the robust LMD shows a high resolution in its TFR spectrum, and the conventional LMD is in the middle. Unfortunately, HHT cannot locate the correct

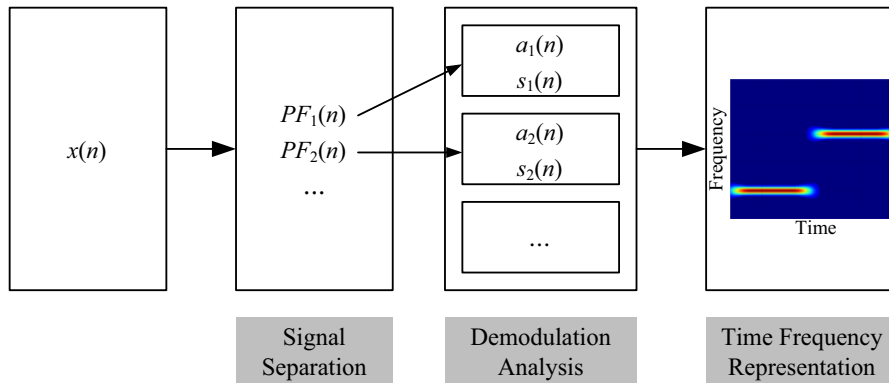


Fig. 4. Time-frequency representation based on the robust LMD.

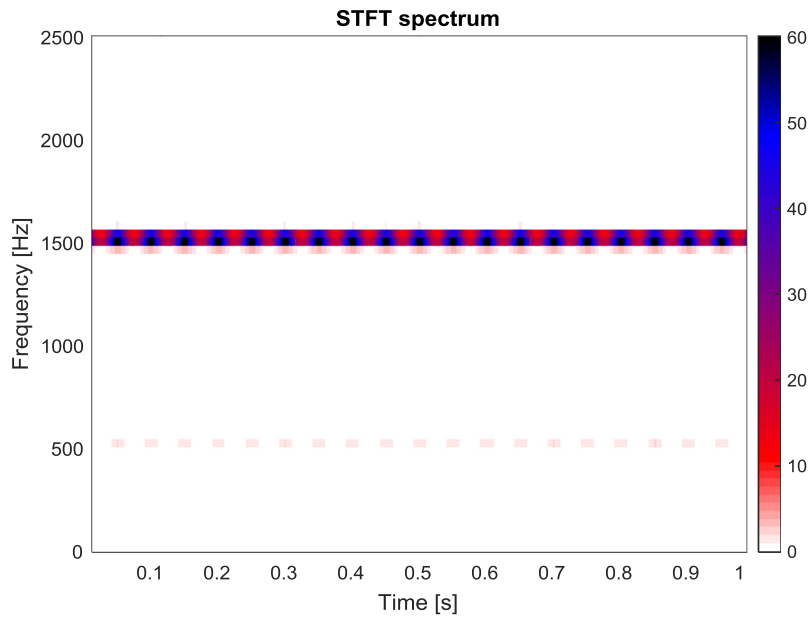


Fig. 5. STFT spectrum in case 1.

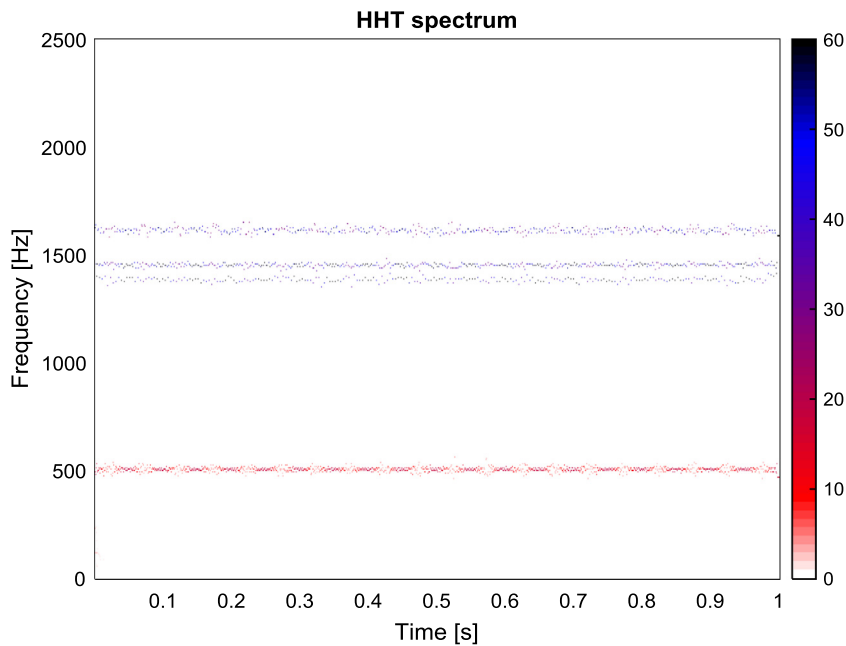


Fig. 6. HHT spectrum in case 1.

place for the second AM component in this case. This may be caused by cross interference among the 2 AM components. From Fig. 6, HHT generates three unexpected components, i.e., 1600 Hz, 1450 Hz, and 1350 Hz, and it shows a large fluctuation in its TFR plot. Both the conventional and robust LMD methods work better in decomposition accuracy and time-frequency resolution than the other two comparison methods, which shows the natural advantages of LMD theory to process multicomponent AM-FM signals. From the above comparison in the case 1, the robust LMD not only finds the right components but also has a high time-frequency resolution with fewer fluctuations (approximately 10 Hz).

(2) Case 2: A multicomponent signal with a pure AM component and a pure FM component

In the second case, we generate a simulated signal in Eq. (16) with a pure AM component and a pure FM component. The sampling configuration and noise setting are the same as in the first case.

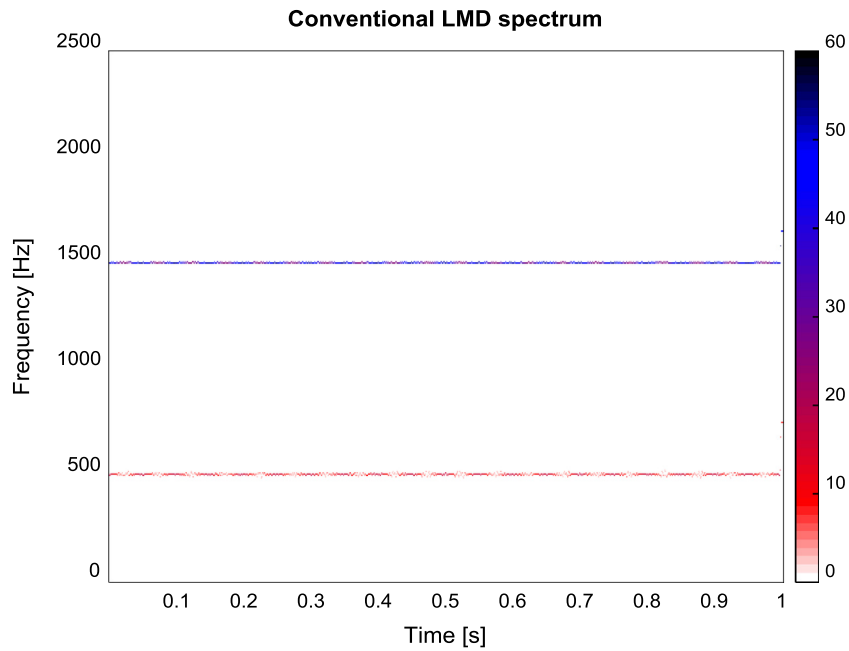


Fig. 7. Conventional LMD spectrum in case 1.

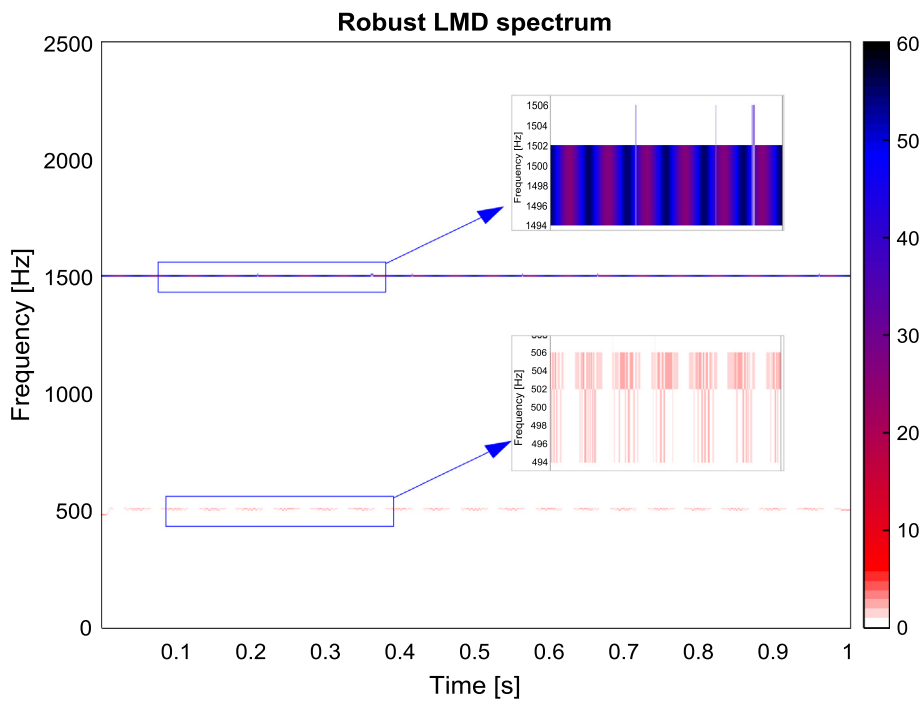


Fig. 8. Robust LMD spectrum in case 1.

$$x(t) = 5 \times (5 + \cos(2\pi t \times 20)) \cos(2\pi t \times 1000) + 11 \times \cos(2\pi t \times 400 + 2 \times \cos(2\pi t \times 10)) + n(t) \quad (16)$$

Figs. 9–12 show the TFR plots from the four methods, and our observations from the four figures are as follows. From Fig. 9 and Fig. 12, STFT and the robust LMD can successfully locate the two components at 1000 Hz and 400 Hz, which correspond to carriers of the two components in Eq. (16). STFT shows a much lower time-frequency resolution than the robust LMD. From the robust LMD spectrum, we can clearly observe the pure FM component with a frequency-modulation sinusoidal

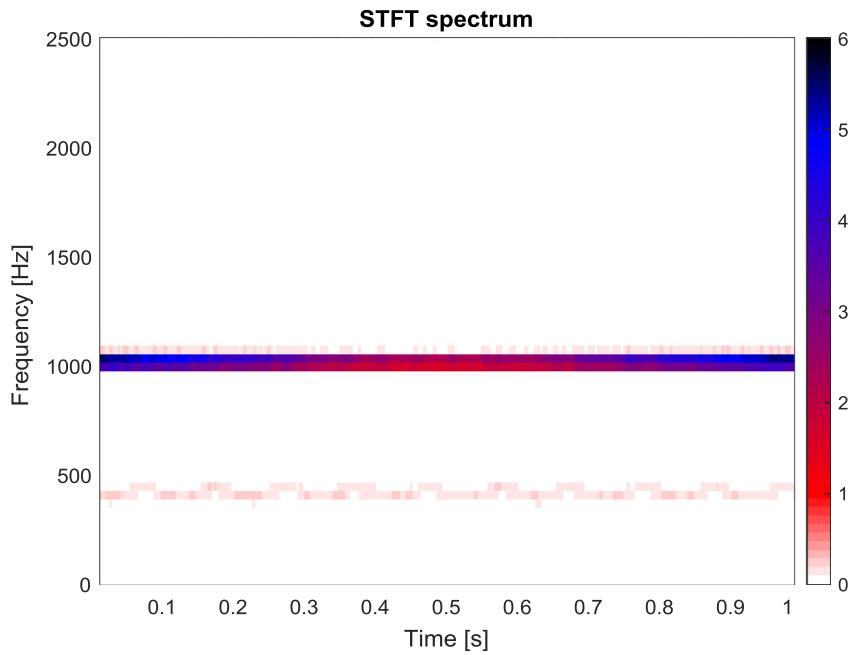


Fig. 9. STFT spectrum in case 2.

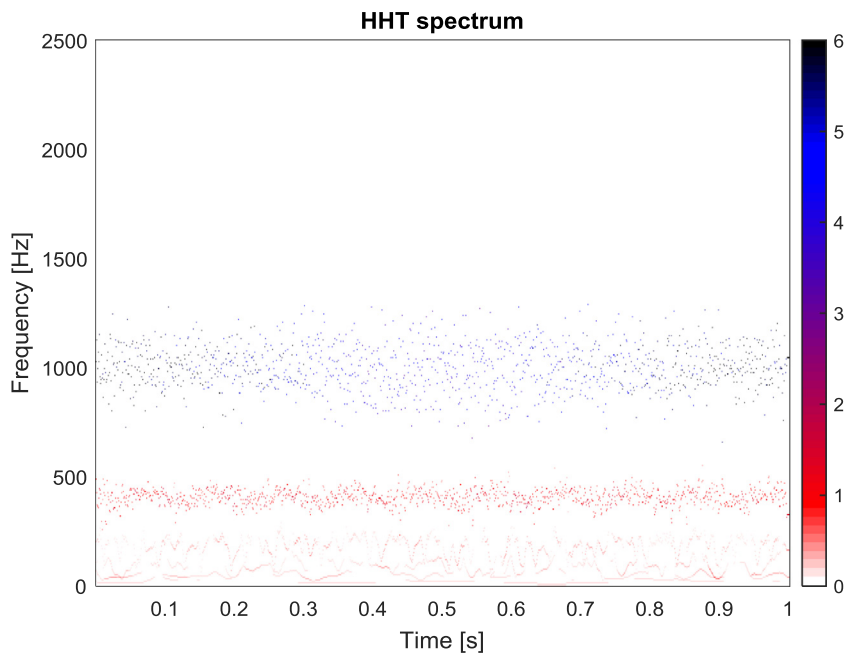


Fig. 10. HHT spectrum in case 2.

signal that has a period of 0.1 s. From Fig. 10 and Fig. 11, HHT and the conventional LMD can roughly find the right two components, whereas their energy is quite scattered around the two carrier frequencies in their TFR plots. For example, the first component scatters between a large frequency range from approximately 800 Hz to 1200 Hz for HHT. Moreover, some confused mess components also appeared between 0 and 200 Hz in the two plots. This reflects cross interference with TFR after introducing a FM component that often shows more complex characteristics. This example shows that the robust LMD has the best performance in component recognition and time-frequency resolution among the four methods.

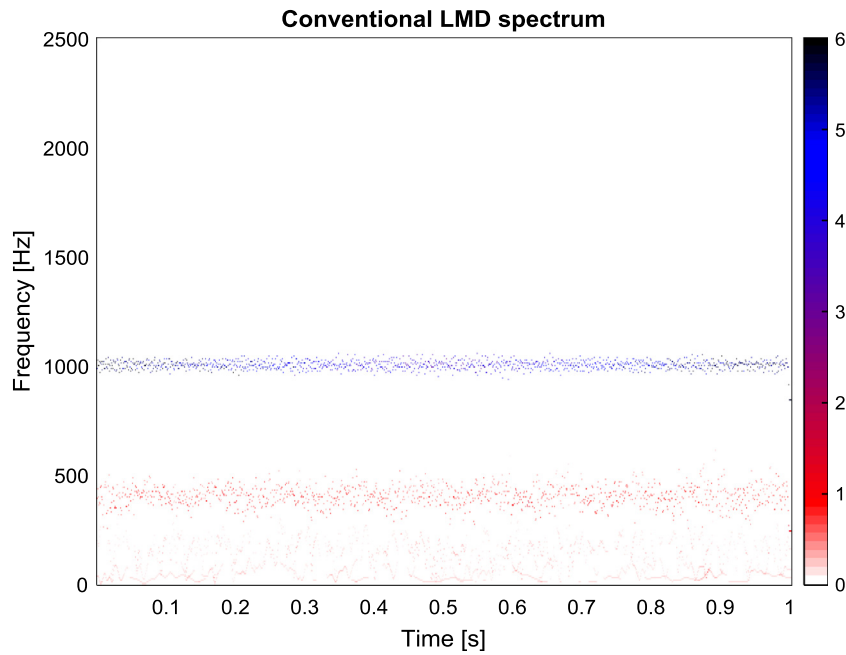


Fig. 11. Conventional LMD spectrum in case 2.

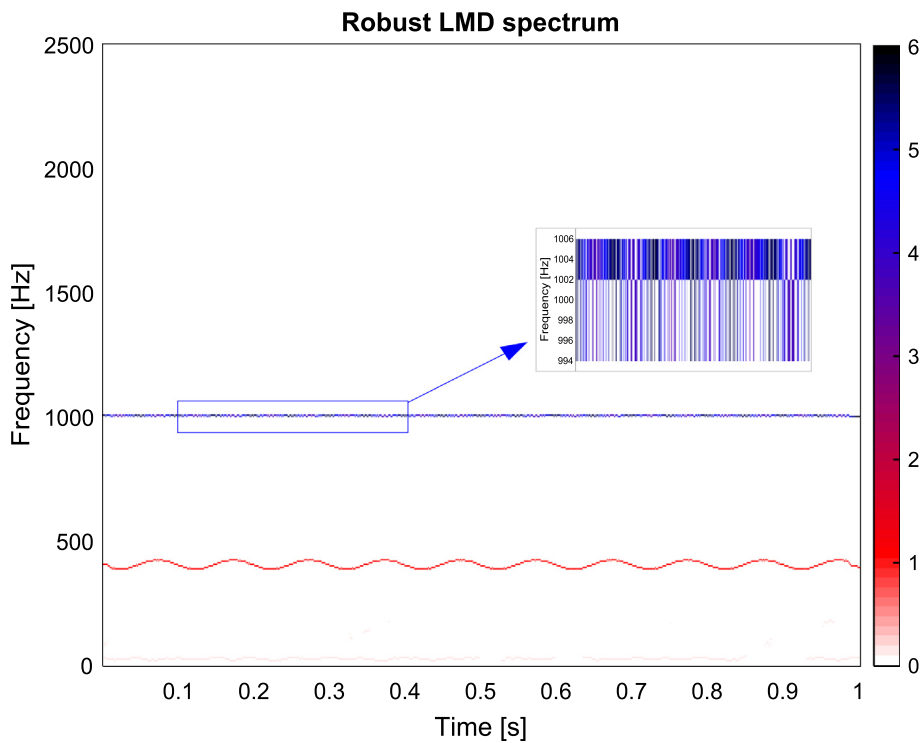


Fig. 12. Robust LMD spectrum in case 2.

(3) Case 3: A multicomponent signal with two pure FM components

In the third case, we generate a simulated signal in Eq. (17) with two pure FM components. The sampling configuration and noise setting are the same as in the first case.

$$x(t) = 5 \times \cos(2\pi t \times 600 + \cos(2\pi t \times 20)) + 15 \times \cos(2\pi t \times 1200 + 4 \cos(2\pi t \times 20)) + n(t) \quad (17)$$

Figs. 13–16 show the TFR plots from the four methods, and our observations from the four figures are as follows. The first FM component with a carrier frequency of 600 Hz can be recognized in the STFT spectrum and the robust LMD spectrum, but it is hardly recognized in the HHT spectrum and the conventional LMD spectrum. The second FM component with a carrier frequency of 1200 Hz can be clearly observed in all spectrums. Again, this example shows that the robust LMD keeps the best performance in component recognition and time-frequency resolution among the four methods.

(4) Case 4: A multicomponent AM-FM signal of gear vibration

In the final case, we apply both the proposed method and HHT to a simulated gear vibration signal in Eq. (18). The vibration signal from a faulty gear usually shows an AM-FM characteristic; that is, TFR to the AM-FM signals is very important in accurate fault diagnosis of gears.

$$y(t) = \sum_{m=1}^M X_m (1 + d_m(t)) \cos(2\pi t \times mzf_r + \phi_m + b_m(t)) \quad (18)$$

where f_r is the rotation frequency of the shaft, z is the number of the teeth in the gear, X_m is the amplitude of the m -th order harmonic component, ϕ_m is the initial phase of the m -th order harmonic component, and $d_m(t)$ and $b_m(t)$ correspond to the amplitude-modulation function and phase-modulation function of the m -th order harmonic component.

According to Eq. (19), we generate a simulated gear fault vibration signal including two order harmonic components that are shown in Eq. (19). This means that the number of gear teeth is 30 and $f_r = 20$ Hz. The sampling configuration and noise settings are the same as in the first case.

$$x(t) = 5 \times (1 + \cos(2\pi t \times 20)) \cos(2\pi t \times 600 + \cos(2\pi t \times 20)) + 15 \times (1 + \cos(2\pi t \times 20)) \cos(2\pi t \times 1200 + 4 \cos(2\pi t \times 20)) + n(t) \quad (19)$$

Figs. 17–20 show the TFR plots from the four methods, and our observations from the four figures are as follows. We can roughly recognize two sinusoidal waves at approximately 600 Hz and 1200 Hz in all four plots, and they correspond to the two components in Eq. (19). The color changes along each sinusoidal wave reflect amplitude-modulation effects. This case has the most complex signal structure among the four cases, and it is also very challenging to the robust LMD. Fortunately, from Fig. 20, the robust LMD successfully extracts the right AM signal with a frequency of 20 Hz and the right FM signal with a frequency of 20 Hz for the simulated signal. Even though it shows a larger fluctuation than that in the first three cases, the proposed method is still the best method for this case.

(5) Summary

To summarize the above four cases, the proposed robust LMD consistently shows the best performance in both decomposition accuracy and time-frequency resolution over STFT, HHT, and the conventional LMD. STFT is stable but has a low

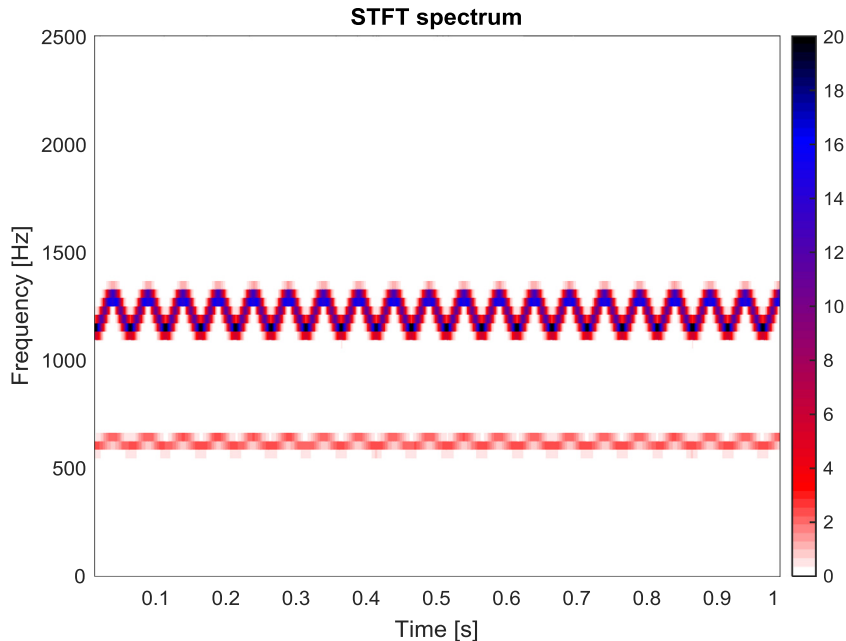


Fig. 13. STFT spectrum in case 3.

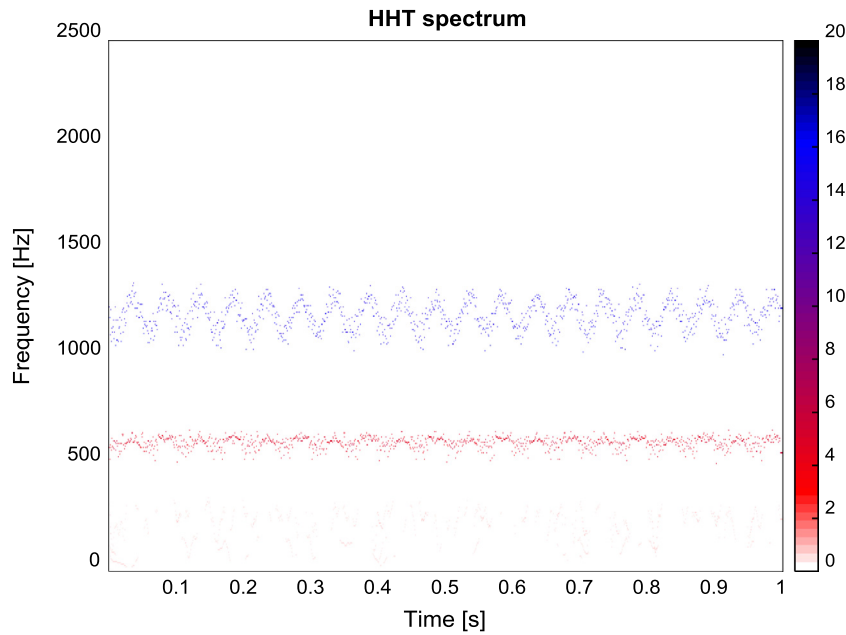


Fig. 14. HHT spectrum in case 3.

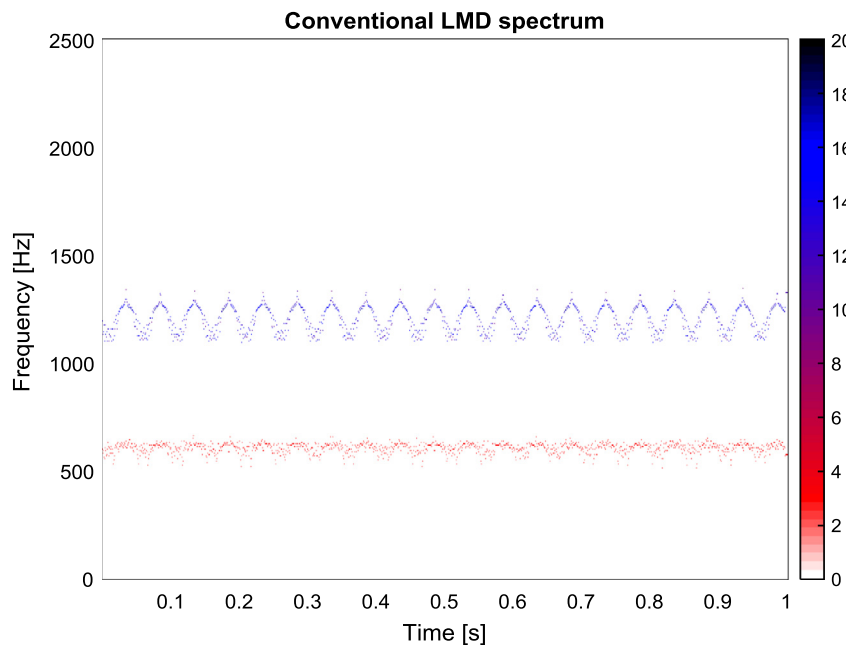


Fig. 15. Conventional LMD spectrum in case 3.

time-frequency resolution. HHT shows the worst performance in the four cases because it is not designed for multicomponent AM-FM signals. The conventional LMD can roughly locate the right component places in the four cases; however, it is strongly affected by cross interference among multiple components and between the AM and FM parts. Besides, it shows up a low resolution with a relatively large fluctuation. This motivates us to further explore the robust LMD. In addition, the robust LMD greatly improves the efficiency of LMD. The following data support this conclusion. The elapsed CPU times of the conventional LMD are 651.6885, 615.4055, 2262.1773, and 2687.5211 s for the four cases. The elapsed CPU times of the robust LMD are 3.2261, 6.2656, 2.2235, and 2.3174 s for the four cases. CPU time is collected by two functions of *tic* and *toc* in MATLAB R2013a (64-bit) on a computer with a Intel Core i5-3337U processor (1.8 GHz, 1.8 GHz), 8 GB RAM and an operating system of Windows 8 (64-bit). In conclusion of this section, the robust LMD significantly improves the performance of LMD in terms of both accuracy and efficiency.

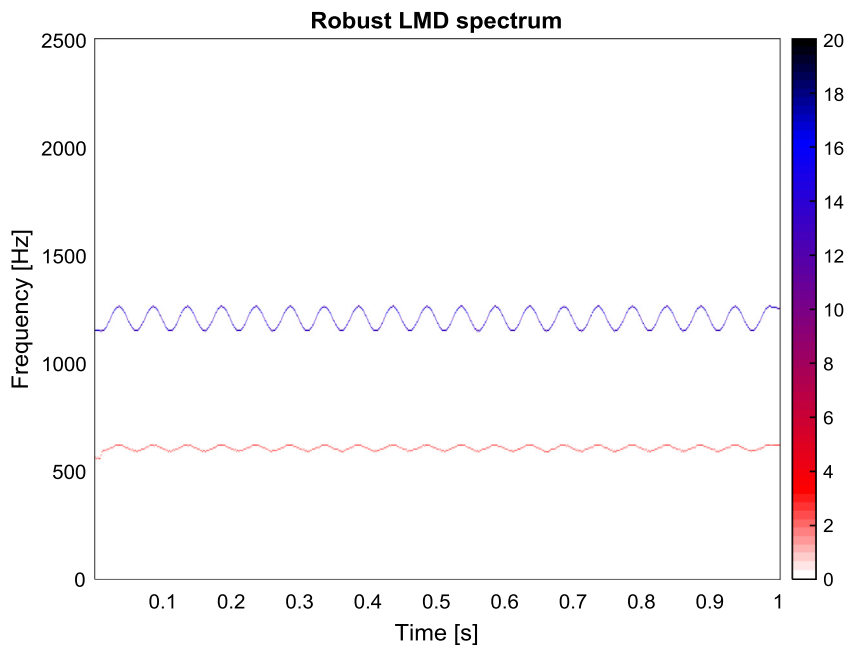


Fig. 16. Robust LMD spectrum in case 3.

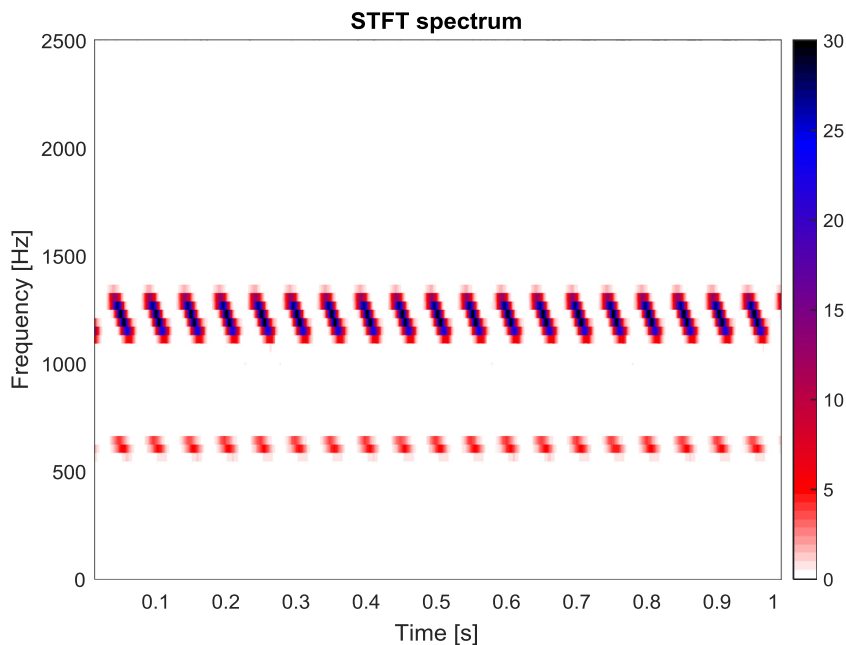


Fig. 17. STFT spectrum in case 4.

5. Application to machinery fault diagnosis

In this section, we apply the proposed robust LMD in conjunction with fast Kurtogram to perform fault diagnosis of a bearing fault that usually presents an AM-FM characteristic. The fast kurtogram proposed by Antoni [24] is a fourth-order spectral analysis tool for detecting and characterizing non-stationarities in a signal and is a powerful analysis tool to detect incipient transient faults [25]. It is worth pointing out that TFR is a crucial component of the fast kurtogram, and it is usually implemented by finite impulse response (FIR) filters and STFT. However, FIR filters cannot guarantee the best fit for any field signals. STFT has a limitation in its time-frequency resolution capability due to the uncertainty principle. Considering this

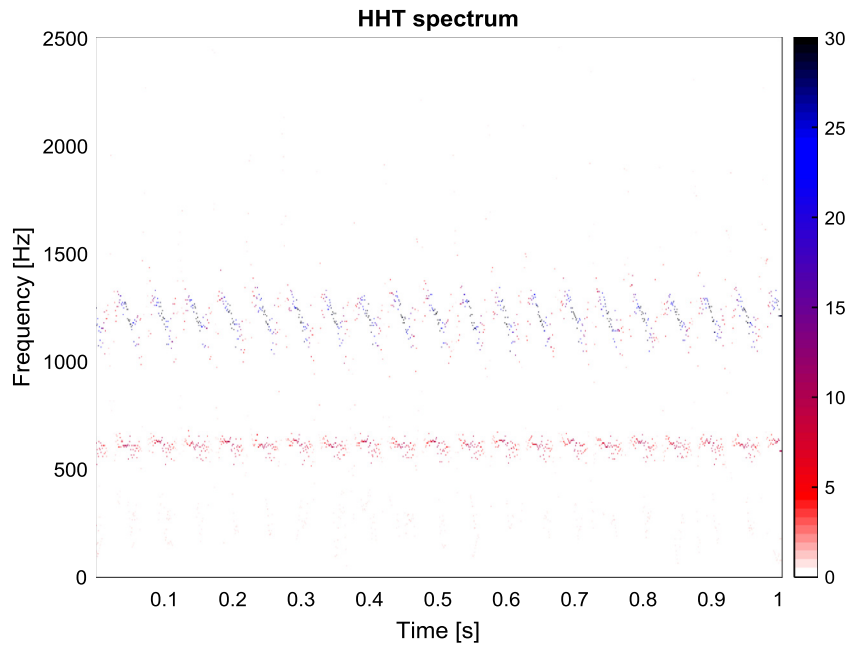


Fig. 18. HHT spectrum in case 4.

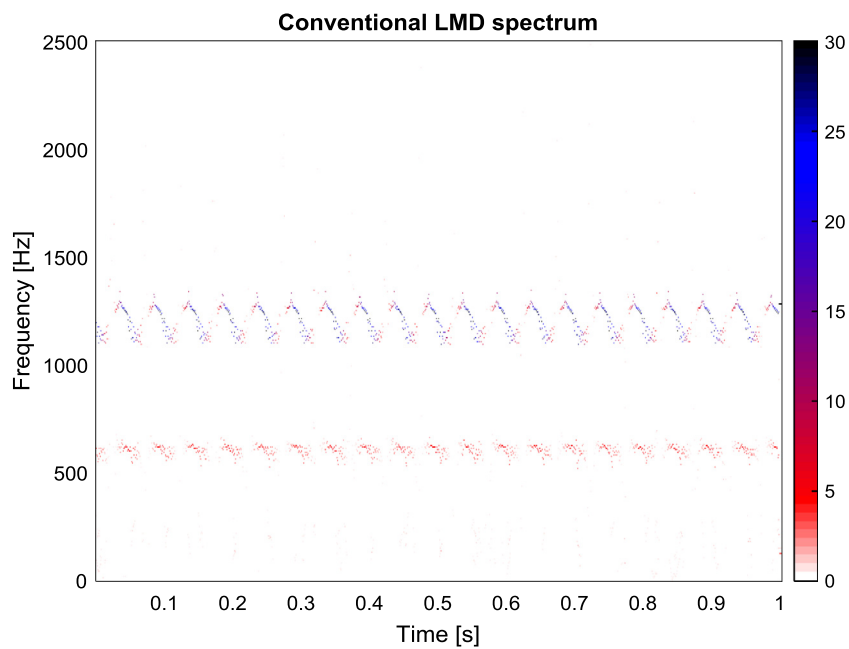


Fig. 19. Conventional LMD spectrum in case 4.

challenge in the fast kurtogram, we try to apply the robust LMD as an alternative TFR tool for the fast kurtogram to improve the accuracy of bearing fault diagnosis.

The bearing fault data are from an accelerated degradation test (ADT) experiment that is conducted on PRONOSTIA shown in Fig. 21. More details about the PRONOSTIA test rig are described in [26]. With this test rig, 17 ADT datasets were collected and opened on the IEEE 2012 prognostic and health management (PHM) data challenge competition. We selected one dataset, namely bearing3_1, for the purpose of demonstrating the proposed method. The dataset of bearing3_1 was collected by two acceleration sensors in the vertical and horizontal directions at a sampling frequency of 25,600 Hz under a consistent

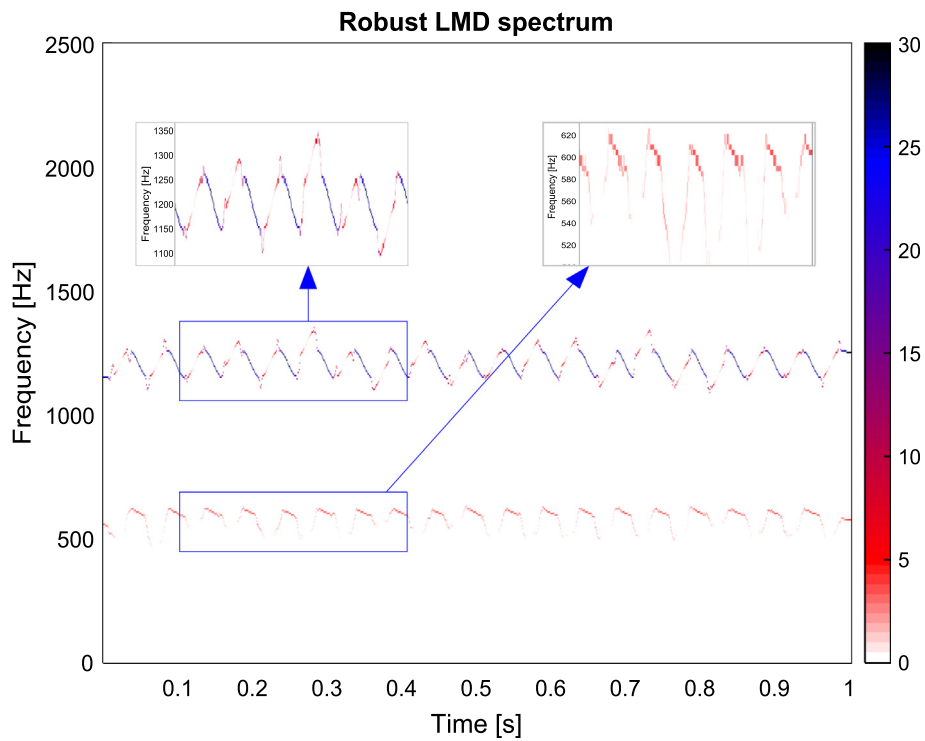


Fig. 20. Robust LMD spectrum in case 4.

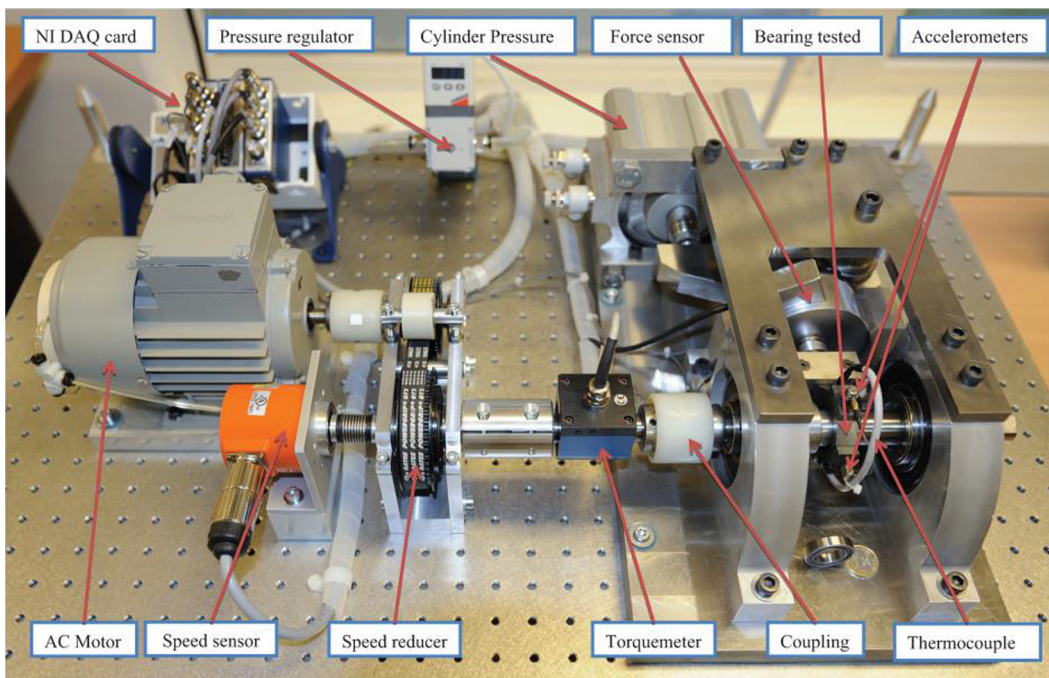


Fig. 21. The PRONOSTIA Test Rig [26].

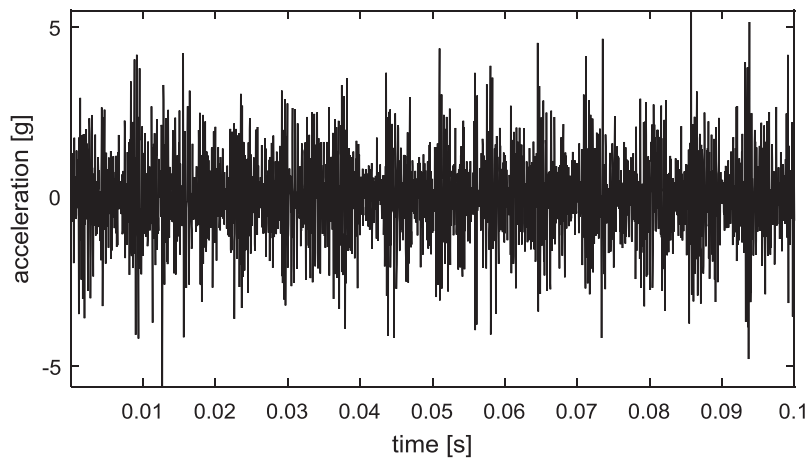


Fig. 22. Time-domain waveform in the vertical direction.

working condition that has a 5000 N radial load and a 1500 rpm speed. There are 515 time records in this dataset, and we selected the 496th time record with a 0.1-s duration in the vertical direction as the test bearing fault data. Fig. 22 shows the temporal waveform of this time record. The selected data concern the outer race fault [27] that is naturally generated rather than artificially created. Based on the parameters of this test bearing [26], the characteristic frequency for the outer race fault (f_{BPFO}) is $5.6113 \times f_n$, where f_n is the shaft rotating frequency (i.e., 25 Hz for the test data).

Figs. 23–25 show kurtograms with three different TFR approaches, i.e., FIR filter, STFT, and the robust LMD. From the three figures, we obtain an optimal frequency band. For example, in Fig. 23, the fast kurtogram find the optimal frequency band with a center frequency (f_c) of 400 Hz and a bandwidth (B_w) of 800 Hz corresponding to the maximum kurtosis (K_{max}) of 7.4 at the level 4. Most importantly, the robust LMD significantly increase the K_{max} value of the fast Kurtogram to 60. This reflects that the robust LMD is more sensitive to abnormal signals than FIR filter and STFT do for fast kurtogram by means of the proposed scheme. Next, the temporal signal is filtered with the optimal center frequency and bandwidth obtained from the kurtogram. The envelope spectrum is then generated by FT for the squared envelope of the filtered signal. Fig. 26 shows three envelope spectrums derived from FIR filters in Fig. 23, STFT in Fig. 24, and the robust LMD in Fig. 25. From Fig. 26, the characteristic frequency for the outer race fault ($f_{BPFO} = 140$ Hz) in the robust LMD–kurtogram–Envelope shows a large amplitude, which mostly indicates that an outer race fault occurring in the test bearing. Unfortunately, such increase at f_{BPFO} is not

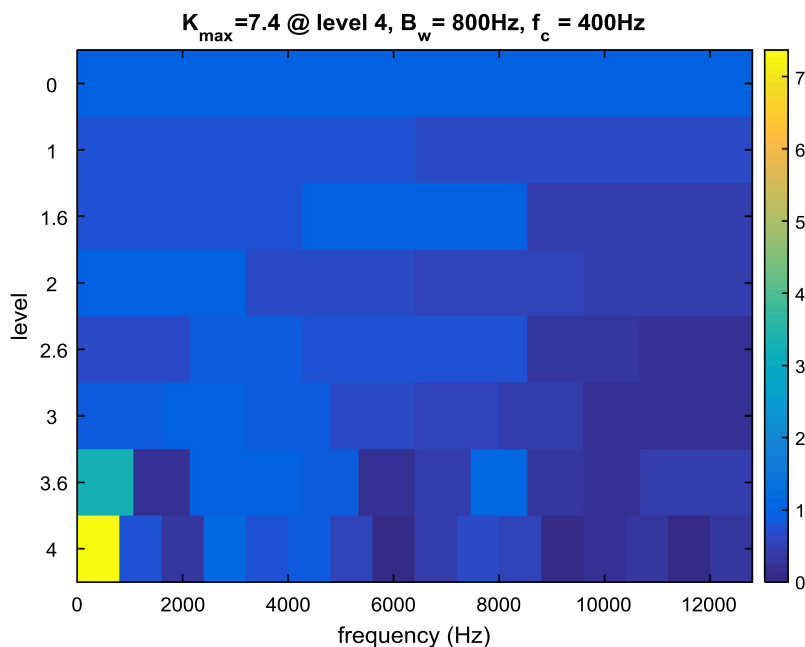


Fig. 23. Fast kurtogram with FIR filter.

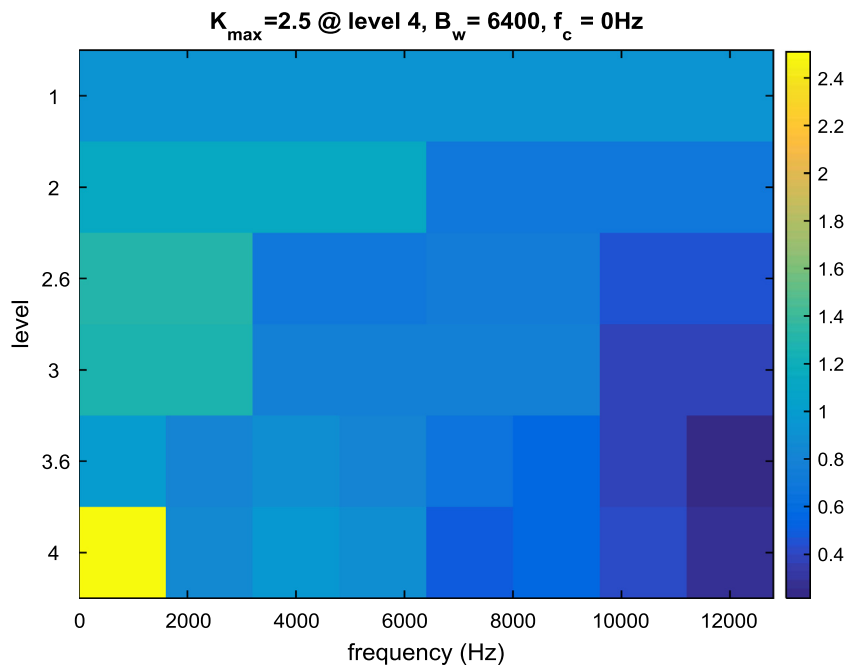


Fig. 24. Fast kurtogram with STFT.

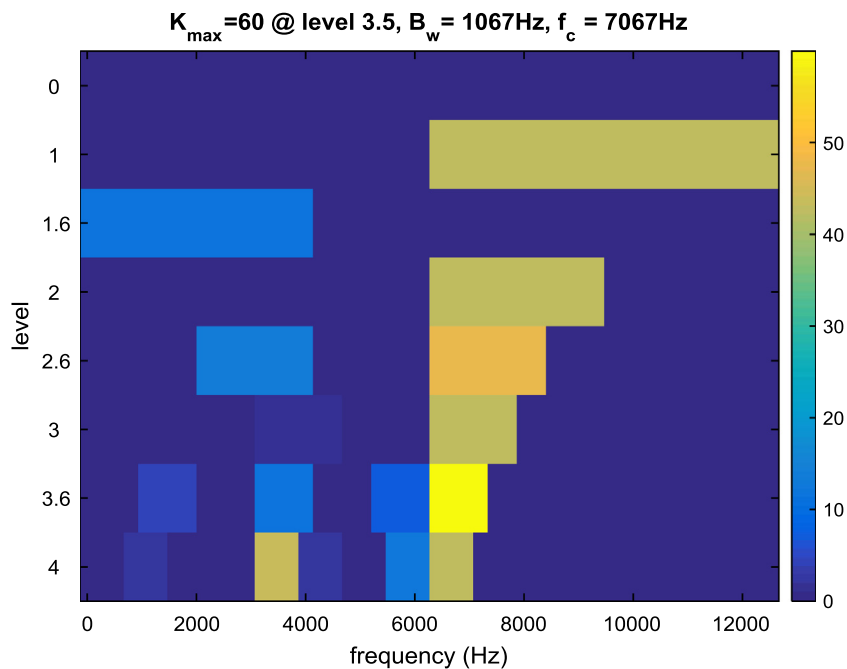


Fig. 25. Fast kurtogram with the robust LMD.

apparent in the FIR filter-kurtogram-Envelope and the STFT-kurtogram-Envelope, which means no evidences of bearing fault diagnosis. Recall Fig. 25, we can conclude that the large $K_{\max}$ with the value of 60 is mainly contributed by the outer race fault, and the robust LMD in conjunction with the fast kurtogram can successfully detect this fault. Although it is difficult to know the real multiple AM-FM components in the collected signal, we can say that the proposed method provides a better TFR estimation than FIR filter and STFT do, and thus greatly enhance performance of the fast kurtogram for bearing fault diagnosis.

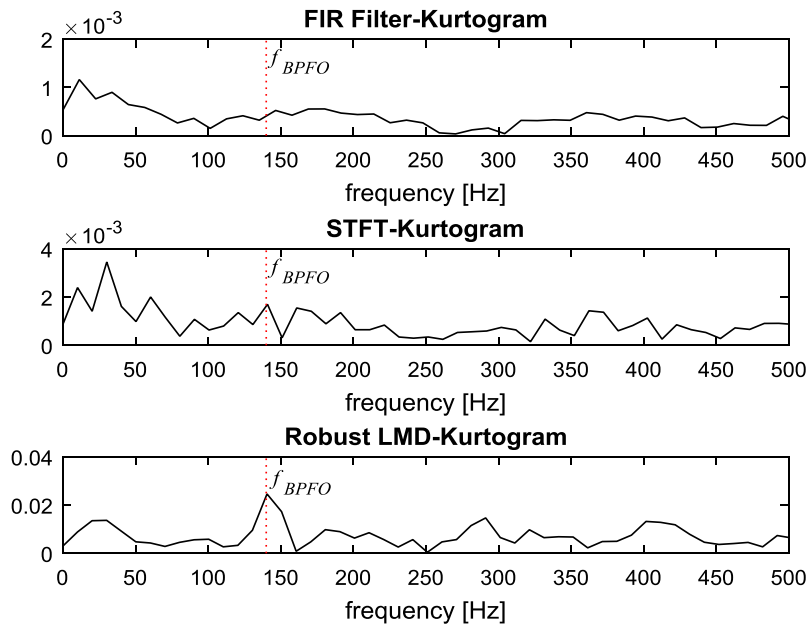


Fig. 26. Amplitude spectrum of the squared envelope.

6. Conclusions

Extracting and presenting multicomponent AM-FM signals are important aspects of signal processing. This paper presents a TFR method based on robust LMD with a novel scheme to solve the end effect and mode mixing problems. The proposed scheme focuses on the improvement for three parameters, i.e., boundary condition, envelope estimation, and sifting stopping criterion. All three parameters show their great influence to TFR performance for multicomponent AM-FM signals. The mirror extension method, fixed subset size selection method, and self-adaptive sifting stopping criterion have been proposed for a better selection of the above three parameters, respectively. The results of four simulation cases show that the robust LMD outperforms the conventional LMD, HHT, and STFT in terms of decomposition accuracy and time-frequency resolution for multicomponent AM-FM signal analysis. Last but not the least, the improved fast kurtogram with the robust LMD successfully finds a natural outer race fault for a test bearing. The experimental results of both simulated and experimental signals demonstrate that the robust LMD shows a significant improvement over the conventional LMD, and it can be used as a promising tool of TFR to characterize multicomponent AM-FM signals.

Acknowledgement

This work was supported by National Natural Science Foundation of China (51505066, 51375078), The National Key Research and Development Program of China (2016YFB1200401), State Key Laboratory of Traction Power, Southwest Jiaotong University (TPL1608), the Fundamental Research Funds for the Central Universities (ZYGX2015J081), and China Postdoctoral Science Foundation Funded Project (2015M570774 and 2016T90842).

References

- [1] Y. Pantazis, Decomposition of AM-FM signals with applications in speech processing, vol. 2010 Crete, University of Crete, Greece, 2010.
- [2] J.S. Smith, The local mean decomposition and its application to EEG perception data, *J. R. Soc. Interface* 2 (2005) 443–454.
- [3] Y. Wang, Z. He, Y. Zi, A comparative study on the local mean decomposition and empirical mode decomposition and their applications to rotating machinery health diagnosis, *J. Vibration Acoust. Trans. ASME* 132 (2010) 1–10.
- [4] Z. Feng, M. Liang, F. Chu, Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples, *Mech. Syst. Signal Process.* 38 (2013) 165–205.
- [5] J. Cheng, Y. Yang, Y. Yang, A rotating machinery fault diagnosis method based on local mean decomposition, *Digital Signal Process.* 22 (2012) 356–366.
- [6] G. Rilling, P. Flandrin, P. Goncalves, On empirical mode decomposition and its algorithms, in: *Proceedings of The 6th IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing*, Grado, Italy, 2003.
- [7] Y. Wang, Z. He, Y. Zi, A demodulation method based on improved local mean decomposition and its application in rub-impact fault diagnosis, *Meas. Sci. Technol.* 20 (2009) 1–10.
- [8] D. Ren, S. Yang, Z. Wu, G. Yan, Research on end effect of LMD based time-frequency analysis in rotating machinery fault diagnosis, *China Mech. Eng.* 8 (2012) 951–956.
- [9] J. Cheng, M. Shi, Y. Yang, Roller bearing fault diagnosis method based on LMD and neural network, *J. Vibration Shock* 29 (2010) 141–144.
- [10] K. Zhang, J. Cheng, Y. Yang, Processing method for end effects of local mean decomposition based on self-adaptive waveform matching extension, *China Mech. Eng.* 2 (2012) 457–462.

- [11] W.Y. Liu, Q.W. Gao, G. Ye, R. Ma, X.N. Lu, J.G. Han, A novel wind turbine bearing fault diagnosis method based on integral extension LMD, *Measurement* 74 (2015) 70–77.
- [12] Y. Li, M. Xu, Z. Haiyang, Y. Wei, W. Huang, A new rotating machinery fault diagnosis method based on improved local mean decomposition, *Digital Signal Process.* 46 (2015) 201–214.
- [13] L. Deng, R. Zhao, An improved spline-local mean decomposition and its application to vibration analysis of rotating machinery with rub-impact fault, *J. Vibroeng.* 16 (2014) 414–433.
- [14] Y. Zhang, Y. Qin, Z. Xing, L. Jia, X. Cheng, Roller bearing safety region estimation and state identification based on LMD–PCA–LSSVM, *Measurement* 46 (2013) 1315–1324.
- [15] M. Wang, L. Zhang, W. Liang, L. Duan, Local mean decomposition method based on B-spline interpolation, *J. Vibration Shock* 29 (2010) 73–78.
- [16] H. Li, L. Guo, T. Chen, L. Yang, X. Wang, Iris recognition based on PCHIP-LMD, *Opt. Precision Eng.* 21 (2013) 197–206.
- [17] X. Cheng, K. Zhang, Y. Yang, Ensemble local mean decomposition method based on noise-assisted analysis, *J. Mech. Eng.* (2011) 55–62.
- [18] N.E. Huang, Z. Shen, S.R. Long, M. Wu, H.H. Shih, Q.N. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Roy. Soc. A-Math. Phys. Eng. Sci.* 454 (1998) 903–995.
- [19] Q. Xie, B. Xuan, J. Li, W. Xu, H. Han, EMD algorithm based on bandwidth and the application on one economic data analysis, in: *Proceedings of 15th European Signal Processing Conference*, Poznan, Poland, 2007, pp. 2419–2423.
- [20] L. Lin, H. Ji, Signal feature extraction based on an improved EMD method, *Measurement* 42 (2009) 796–803.
- [21] K. Zhang, J. Chengli, Y. Yang, Research on the product function criterion in local mean decomposition method, *J. Vibration Shock* (2011) 84–88.
- [22] Y. Li, M. Xu, Y. Wei, W. Huang, A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree, *Measurement* 77 (2016) 80–94.
- [23] W. Jiang, Z. Zheng, Y. Zhu, Y. Li, Demodulation for hydraulic pump fault signals based on local mean decomposition and improved adaptive multiscale morphology analysis, *Mech. Syst. Signal Process.* 58–59 (2015) 179–205.
- [24] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, *Mech. Syst. Signal Process.* 20 (2006) 282–307.
- [25] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, *Mech. Syst. Signal Process.* 21 (2007) 108–124.
- [26] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, RONOSTIA: an experimental platform for bearings accelerated degradation tests, in: *IEEE International Conference on Prognostics and Health Management* Denver, Colorado, United States, 2012, pp. 1–8.
- [27] T. Wang, Bearing life prediction based on vibration signals: a case study and lessons learned, in: *IEEE Conference on Prognostics and Health Management*, Denver, Colorado, USA, 2012, pp. 1–7.