A robust detector for rolling element bearing condition monitoring based on the modulation signal bispectrum and its performance evaluation against the Kurtogram

Xiange Tian[1], James Xi Gu[2], Ibrahim Rehab[1], Gaballa M. Abdalla[1], Fengshou Gu[1, 3*], and A. D. Ball[1]
[1]Centre for Efficiency and Performance Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK
[2]School of Engineering, Manchester Metropolitan University, Manchester, M15 6BH, UK
[3]School of Mechanical Engineering, Taiyuan University of Technology, Shanxi, 030024, China

Abstract

Envelope analysis is a widely used method for rolling element bearing fault detection. To obtain high detection accuracy, it is critical to determine an optimal frequency narrowband for the envelope demodulation. However, many of the schemes which are used for the narrowband selection, such as the Kurtogram, can produce poor detection results because they are sensitive to random noise and aperiodic impulses which normally occur in practical applications. To achieve the purposes of denoising and frequency band optimisation, this paper proposes a novel modulation signal bispectrum (MSB) based robust detector for bearing fault detection. Because of its inherent noise suppression capability, the MSB allows effective suppression of both stationary random noise and discrete aperiodic noise. The high magnitude features that result from the use of the MSB also enhance the modulation effects of a bearing fault and can be used to provide optimal frequency bands for fault detection. The Kurtogram is generally accepted as a powerful means of selecting the most appropriate frequency band for envelope analysis, and as such it has been used as the benchmark comparator for performance evaluation in this paper. Both simulated and experimental data analysis results show that the proposed method produces more accurate and robust detection

---

* Corresponding author. Tel: +44-1484-473548; Fax: +44-1484- 473075.
E-mail address: f.gu@hud.ac.uk (Fengshou Gu)

results than Kurtogram based approaches for common bearing faults under a range of representative scenarios.

Keywords: Modulation signal bispectrum; bearing fault detection; envelope analysis; Kurtogram; condition monitoring.

1. Introduction

Bearings are at the heart of almost every rotating machine, and they have received a lot of attention in the field of vibration analysis because they are a common source of machine faults [1][2]. For accurate diagnosis of a bearing fault, a number of techniques have been proposed in recent years to detect and identify specific bearing fault features (bearing frequencies) from within monitored data. Darlow explored the use of a high frequency resonance technique, widely known as envelope analysis [3]. Antoni applied cyclostationary spectral analysis [4][5], and cepstrum analysis, bispectrum analysis and time-frequency analysis have also been used. Ho and Randall investigated the application of self-adaptive noise cancellation in conjunction with envelope analysis to remove discrete frequencies masked within bearing vibration signals [6]. Barszcz applied the same approach to denoise wind turbine vibration signals for bearing outer race fault diagnosis [7]. Sawalhi, Randall and Endo presented an algorithm for enhancing the surveillance capability of spectral kurtosis by using the minimum entropy deconvolution technique. This technique deconvolves the influence of the transmission path and clarifies the impulses, even when they are not separated in the original signal [8]. Zhao applied empirical mode decomposition and the approximate entropy method for severity assessment of a spall-like fault in a rolling element bearing [9]. A recent significant advance in envelope based rolling element bearing fault detection has been the Kurtogram [10] and this has received considerable attention in recent months [11][13]. For this reason, the Kurtogram has been used as the benchmark comparator in this study.

The researchers above, and more, have achieved considerable progress in improving the accuracy of bearing fault detection and diagnosis. Most of the fault detection schemes presented in the literature are based on tracking the amplitude of the characteristic fault frequency but with little attention given to the utilisation of modulation characteristics and noise suppressing which are inherent in measured signals. Recently, Rehab *et al* explored using the MSB to extract fault features from the envelope signal, exploiting its noise suppression capabilities, and in doing so showed more reliable bearing fault severity assessment compared to power spectrum approach [14]. This approach, however, still requires optimisation of the filter's parameters for envelope analysis. In this paper a more straightforward and robust MSB detector is proposed, which does not rely on envelope analysis, and which is shown to provide reliable detection features based only on the demodulation and noise suppression characteristics of the MSB.

Section 2 develops the detector and outlines the theoretical basis for bearing fault diagnosis. Section 3 presents performance studies based on simulated signals, and Section 4 validates the practical application of the detector via two application case studies.

2. The modulation signal based detector

2.1 A bearing vibration signal model

The vibration signature of a rolling element bearing with local defects can be typified by an amplitude modulation process. For a rolling bearing with a local defect of fault characteristic frequency $f_F = 1/T_0$, its vibration acceleration response containing $2M+1$ impulses can be modelled according to [15][16][17], as follows:

$$x(t) = \sum_{m=-M}^{M} A_m(t_i) e^{-\beta(t_i)} \cos(\omega_r t_i)\,\mathrm{u}(t_i) + n(t) \tag{1}$$

where A_m is the amplitude of the m th fault impulse which includes cage and load modulation, M is the number of impulses, $\mathrm{u}(t)$ is a unit step function, T_0 is the time period corresponding to the fault characteristic frequency, β is the structural damping characteristic, ω_r is the excited resonance frequency, $n(t)$ is typical noise which includes both stationary Gaussian noise and aperiodic impulses as would inevitably be encountered in any real measurement environment, and where $t_i = t - (mT_0 + \sum_{i=-M}^{ms} \tau_i)$ in which τ_i represents the effect of random slippage of the rollers as the i^{th} realisation of a zero mean uniformly distributed random variable, with standard deviation within a range <0.02 T_0.

This represents the fault signature of a local bearing defect comprising not only periodic components but also nonlinear modulation effects between fault frequencies, structural resonances and load distribution. Moreover, the signal is contaminated by noise and interference, and this is especially relevant when the fault signature is weak during the early stages of fault development. On this basis, to extract fault signatures effectively, the signal must be both denoised and demodulated.

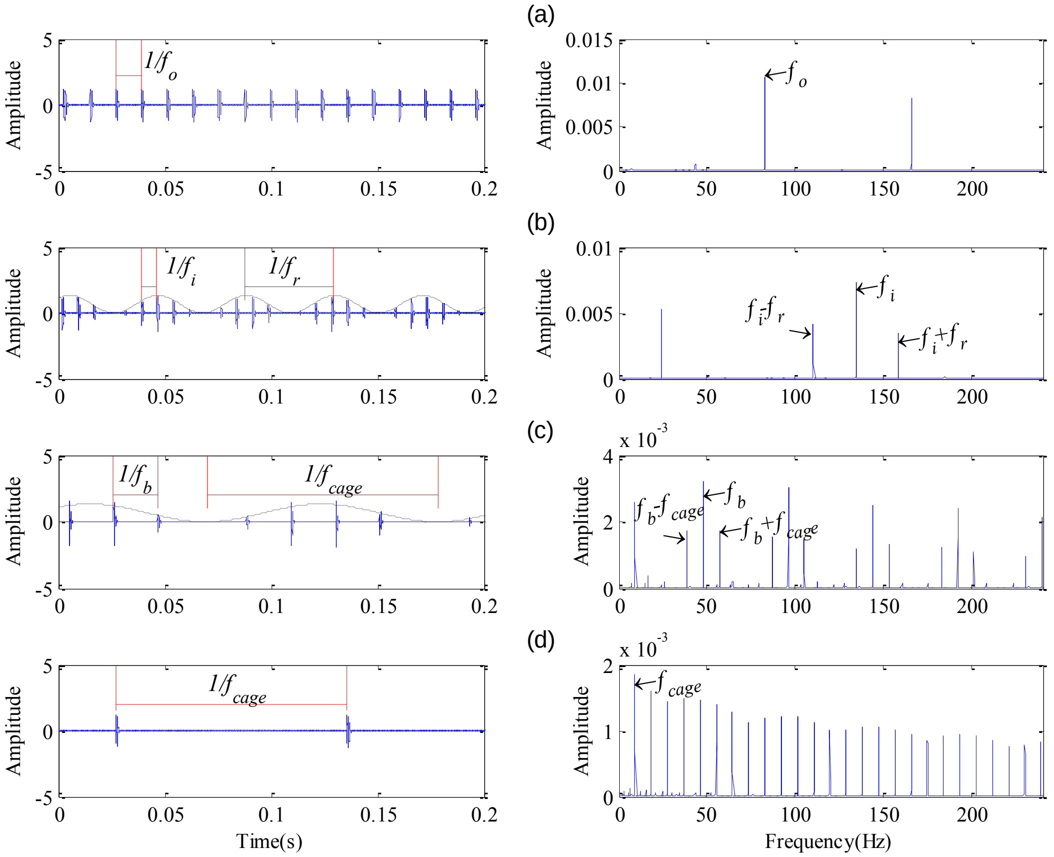

Fig. 1 Simulated fault data and spectra of a rolling element bearing with a localised defect on the (a) outer race, (b) inner race (c) rolling element, and (d) cage

As the effect of random slippage is relatively small, the deterministic part of $x(t)$ in Eq. (1) represents predominately a series of impulse responses to local bearing defects such as a small dent on deferent components of a bearing, with a repetition frequency which reflects the contact of the bearing fault with another part of the bearing (e.g. an area of fatigue damage on a raceway and the periodic interaction of the rolling elements with this), this is called the defect frequency of the bearing. For a typical rolling element bearing there are four possible characteristic defect frequencies and these are determined by the bearing dimensions, the shaft speed and the defect location, in addition to an installation-dependent feature called the contact angle [18]. The repetition frequency for an outer race defect is denoted f_o, that for an inner race defect is f_i, for a rolling element defect is f_b and for a cage defect is f_c. The repetition frequency can be modulated by loaded zone effects on rotating elements, as shown in Fig. 1. For an inner race defect, the modulating frequency is the shaft rotational frequency f_r, but for a rolling element defect it is the ball spin frequency f_{bs} (where $f_{bs} = f_b / 2$). The theoretical characteristic frequencies of a rolling element bearing can be calculated with Eqs. (2)-(5) [18].

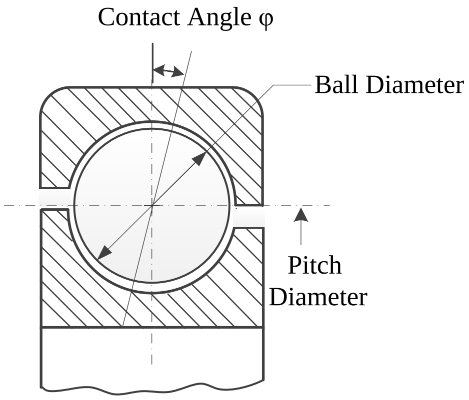

Fig. 2 Contact angle of rolling element bearing

Outer race fault frequency (aka the Ball Pass Frequency for a Fault on the Outer Race - BPFO):

$$f_o = \frac{N_r}{2} f_r (1 - \frac{D_b}{D_c} \cos \varphi) \tag{2}$$

Inner race fault frequency (aka the Ball Pass Frequency for a Fault on the Inner Race - BPFI):

$$f_i = \frac{N_r}{2} f_r (1 + \frac{D_b}{D_c} \cos \varphi) \tag{3}$$

Ball fault frequency (aka the Ball Spin Frequency - BSF):

$$f_b = \frac{D_c}{2D_r} f_r (1-(\frac{D_b}{D_c}\cos\varphi)^2) \tag{4}$$

Cage fault frequency (aka the fundamental train frequency - FTF):

$$f_{cage} = \frac{1}{2} f_c (1-\frac{D_b}{D_c}\cos\varphi) \tag{5}$$

where D_c is the pitch circle diameter, D_b is the roller diameter, φ is the contact angle as shown in Fig. 2, and N_r is number of balls (or rollers).

2.2 The modulation signal bispectrum

To analyse the modulated signals, consideration of the sidebands in the bispectrum was introduced in [19]. The authors [20][22] have found that this approach allows an accurate quantification of modulating components in diagnosing different types of mechanical and electrical faults in machines and is particularly useful in extracting weak fault signatures in motor current signals. In the frequency domain, the modulation signal bispectrum (MSB) of a signal $x(t)$, expressed in the form of the discrete Fourier transform $X(f)$, can be defined as:

$$B_{MS}(f_c, f_x) = E\left\langle X(f_c+f_x)X(f_c-f_x)X^*(f_c)X^*(f_c)\right\rangle \tag{6}$$

where $B_{MS}(f_c, f_x)$ is the bispectrum of signal $x(t)$, $E\langle\ \rangle$ is the expectation operator, f_x is modulating frequency, f_c is the carrier frequency and (f_c+f_x) and (f_c-f_x) are the higher and lower sideband frequencies respectively. This expression takes into account both (f_c+f_x) and (f_c-f_x) simultaneously and hence it is of particular interest to this study, which aims to explore sideband features in the MSB. It enables qualification of modulation effects via the bispectral peak at bifrequency $B_{MS}(f_c, f_x)$. In addition, if random noise components are not coupled but have random distributions, the magnitude of the MSB will be close to zero. In this way, the MSB allows wideband noise and aperiodic components in bearing vibration signals to be suppressed effectively so that the discrete components relating to modulation effects can be revealed more clearly.

To quantify more accurately the sideband amplitudes, the MSB can be modified by removing the substantial influence of carrier frequency (f_c) components via a magnitude normalisation. To differentiate this improvement from the normal MSB, the MSB sideband estimator has been abbreviated to MSB-SE [21], defined as follows:

$$B_{MS}^{SE}\left(f_c, f_x\right) = \frac{B_{MS}\left(f_c, f_x\right)}{\sqrt{\left|B_{MS}\left(f_c, 0\right)\right|}} \tag{7}$$

where $B_{MS}\left(f_c, 0\right)$ is the squared power spectrum estimation at $f_x = 0$ and is equal to $\left|X\left(f_c\right)\right|^4$. This is an important property that can be used as the basis for calculation of individual sideband amplitudes [21].

2.3 An MSB-based robust detector

Fig. 3 shows a typical MSB result from a measured signal. It is clearly evident that the optimal frequency band for detecting a bearing fault is at a specific value of f_c referred to as f_c^{best}, and this band results in a maximum B_{MS}^{SE} peak. In this instance, a single value of f_c, f_c^{best} gives the most significant result, but it will be shown later that more than one value of f_c may give significance peaks at the bearing defect frequency (and its harmonics) labelled

‘*’ in Fig. 3. Based on this approach, fault detection can be implemented directly, avoiding the preliminary processing steps associated with selection of the optimal frequency band during conventional envelope calculation [14][17]. In conjunction with the MSB’s noise and impulse interference suppression capabilities this results in a significantly more robust detection method.

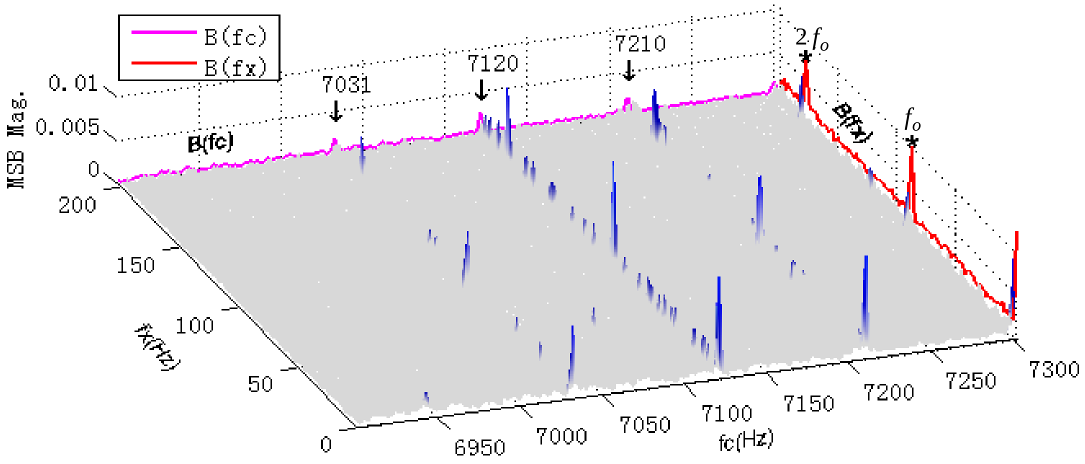

Fig. 3 MSB showing detector $B(f_x)$ formed from slices shown along $B(f_c)$

To achieve even more robust results, the detector can be further improved based on an average of several suboptimal MSB slices such as those with ‘↓’ markers in Fig. 3, and it is this final adaptation that results in what this paper refers to as ‘the robust MSB detector’:

$$B(f_x)=\frac{1}{K}\Sigma_{k=1}^{K}B_{MS}^{SE}\left(f_c^k,f_x\right)\quad f_x>0 \tag{8}$$

where K is the total number of selected f_c suboptimal slices (3, in the case of Fig. 3), the number of which depends on the significance of the peaks themselves. From Fig. 3, it is clear that peaks appear at the fault frequency and its first harmonic in the MSB detector $B(f_x)$ of the simulated signal. The detector is calculated based on several f_c slices, and hence it combines the information of multiple carrier frequencies. In this way, it utilises more wideband characteristics of the impulsive excitations due to a bearing defect, which ensures that the results are more robust because of its increased suppression of strong interferences that can exist in any individual f_c slice.

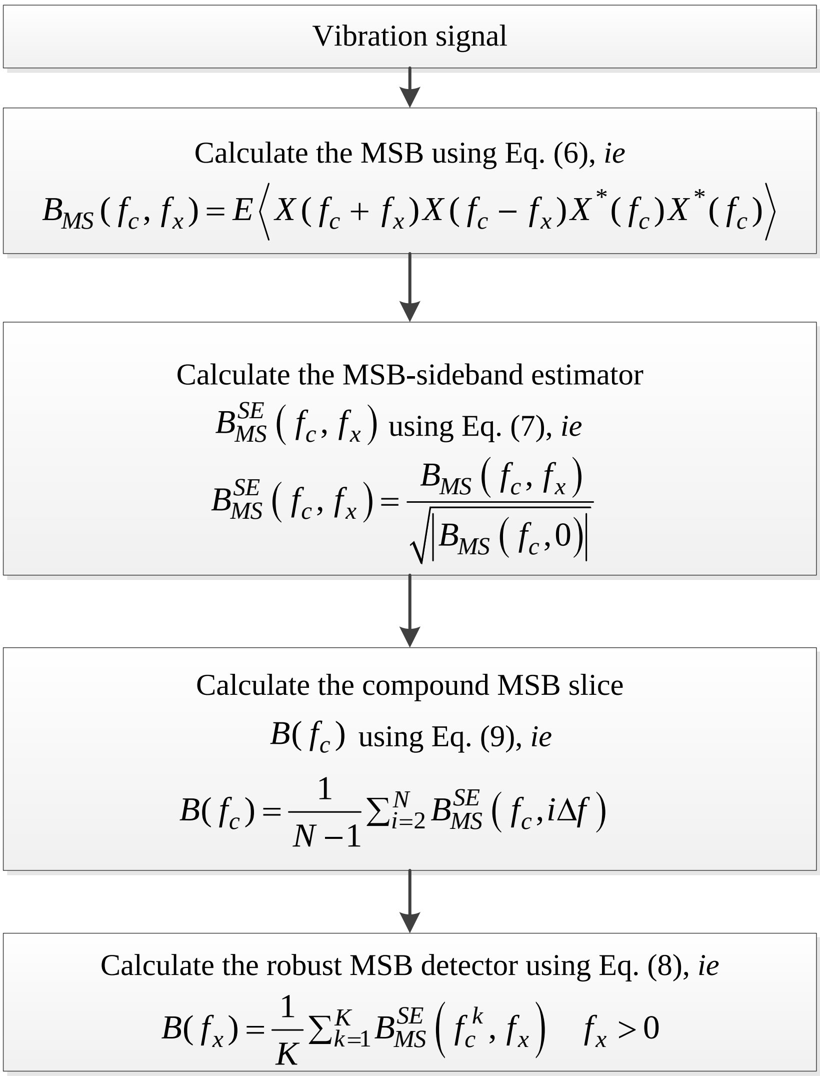

Fig. 4 Flow chart of the robust MSB detector calculation

In order to obtain suboptimal f_c slices, the suboptimal sideband estimator $B_{MS}^{SE}\left(f_c^k, f_x\right)$ can be determined from the compound MSB slice $B(f_c)$, calculated by averaging the significant MSB peaks in the direction of the f_x increment:

$$B(f_c) = \frac{1}{N-1} \sum_{i=2}^{N} B_{MS}^{SE}\left(f_c, i\Delta f\right) \tag{9}$$

where Δf is the frequency resolution in the f_x direction.

In summary then, the robust MSB detector can be implemented using four primary steps shown in Fig. 4:

3. Simulation study

To evaluate the performance of the MSB detector, both no-slippage and slippage signals were produced using a linearised bearing model derived from the nonlinear model presented in [23], and this was validated with different bearing fault cases and different internal clearances. Specifically, it represents three typical resonances in high frequency bands corresponding to the inner race, outer race and seat/sensor, allowing the evaluation to be carried out for different resonant regions. Based on a fundamental period $T_0 = 1/f_o = 1/89$, corresponding to an outer race defect with BPFO=89Hz, a pulse train with no bearing slippage was generated with a fixed time period defined by $t_i = t - mT_0$, whereas an equivalent pulse series with bearing slippage was produced with a random time period by using $t_i = t - (mT_0 + \sum_{i=-M}^{m} \tau_i)$ in which the standard deviation of the time variation was set to either a smaller deviation $\sigma_{\tau 1} = 0.01T_0$ (ie 1%) or a larger deviation $\sigma_{\tau 2} = 0.02T_0$ (ie 2%), to enable the investigation of the influences of bearing slippage at two different levels of occurrence. The pulse series were then used as the input to the model and the output velocity response at the bearing seat was differentiated to obtain the corresponding acceleration signal. This way of signal generation is efficient and also maintains the underlying bearing dynamics;

in comparison it would take a relatively long time to find a numerical solution to a nonlinear model.

Both the no-slippage signals and slippage signals had different levels of extraneous white noise and aperiodic impulsive interferences added to them, creating in total six evaluation scenarios as detailed in Table 1 and denoted A through F; collectively these allowed a range of representative detection circumstances to be assessed.

Table 1. Description of the six evaluation scenarios, denoted A - F

Scenario	**White noise**	**Aperiodic interference**	**SNR1 value**	**SNR2 value**
A - Low noise	Level 1	Low	-14.7dB	-27.9dB*
B - High noise	Level 2	Low	-29.7dB	-42.9dB*
C - Low noise with medium level impact interferences	Level 1	Medium	-14.7dB	-38.9dB**
D - High noise with high level impact interferences	Level 2	High	-29.7dB	-53.9dB**
E - 2% random slippage with low noise	Level 1	Low	-14.7dB	-28.3dB*
F - 2% random slippage with high noise	Level 2	Low	-22.0dB	-35.7dB*

* **SNR2 is calculated from the white noise signal only.**

** **SNR2 is calculated from both the white noise and the impulsive interference.**

To quantify the added noise influences, two means of calculating signal to noise ratio (SNR) were used, one for gauging the power of the stationary noise (referred to as SNR1) and another for gauging the local peaks of the aperiodic impulsive noise cases (referred to as SNR2). For the stationary case, SNR1 is defined as:

$$SNR1 = 10\log_{10}\left(P_s / P_n\right) \tag{10}$$

where P_s and P_n indicate the RMS values of the signal and noise respectively. And for the aperiodic impulsive interference case, SNR2 is defined as:

$$SNR2 = 20\log_{10}\left(P_s / A_n\right) \tag{11}$$

where A_n represents the peak values of the noise or the impulsive interferences. As shown in Table 1, in addition to having SNR1 values, the two random noise-only cases (A and B) also have SNR2 values because the random white noise includes localised spikes of magnitudes several times its RMS value. Nevertheless, these signals are considered as being of low level aperiodic impulsive interference compared to cases C and D where impulsive interferences have deliberately been added. The two white noise cases A and B represent any stationary noise influences from the measurement system, machine operation etc.

It can be seen that the addition of aperiodic impulsive interferences to the white noise cases in A and B result in considerably lower SNR2 values for cases C and D. These two cases allow the study of situations where the signals are contaminated with nonstationary (aperiodic) influences which can arise from electromagnetic effects, looseness, backlash, cavitation etc. Having determined the characteristics of the MSB detector when applied to the noise contaminated no-slippage signals (A – D), slippage signals (E and F) with a high levels of non-stationarity (or randomness) of 2% were then explored, incorporating two different white noise levels; this enabled the accurate assessment of the influence of bearing slippage on the robustness of detection. It is worth noting that although data was simulated for a range of random slippage variations (0.5%, 1%, 2%), only the most severe of these circumstances (2% slippage) was used in the performance evaluation.

The study evaluated if the detector could find bearing fault signatures within the noisy signals and thereafter benchmarked the results with those from the typical fast Kurtogram approach.

3.1 Robustness to white noise

Each simulated signal consisted of 285,715 data points with a sampling rate of 71.5kHz, giving 355 pulses within a time period of 4 seconds. Fig. 5 shows the representative time series and corresponding spectral representations for three different simulated outer race fault signals (note the different amplitude scales for the three time traces). Fig. 5(a) shows the waveform of the simulated rolling element bearing fault signal without any noise. It comprises three primary frequency response regions associated with three resonance frequencies at 3,471Hz, 7,120Hz, and 11,750Hz. Although this is simulated data, these 3 frequencies were chosen because they are representative of those associated with the dominant vibration modes of the inner race, outer race and sensor, in a typical rolling element bearing condition monitoring setup [24]. Fig. 5(b) illustrates the time waveform and the spectrum for the low noise signal with low impulsive interference (Case A). Fig. 5(c) presents the equivalent for the high noise signal again with interference (Case B). The SNR, in both cases, was calculated using Eq. (10). From Fig. 5(b) and (c), it can be seen that for the low noise case, the frequency components around the lowest resonance frequency (3,471Hz) are masked by the noise. For the high noise case, the first two resonance frequencies are themselves completely buried in the noise. It is therefore difficult, in these circumstances, to locate resonance frequencies and hence to implement accurate fault detection if a conventional envelope approach is used. As shown in Fig. 6, the envelope spectrum is unable to reveal the fault components for the high noise case when it uses a band pass filter with a centre frequency of 11,750Hz and a bandwidth of 500Hz (which covers the first 3 harmonics of the inner race fault frequency which is the highest frequency of possible bearing faults).

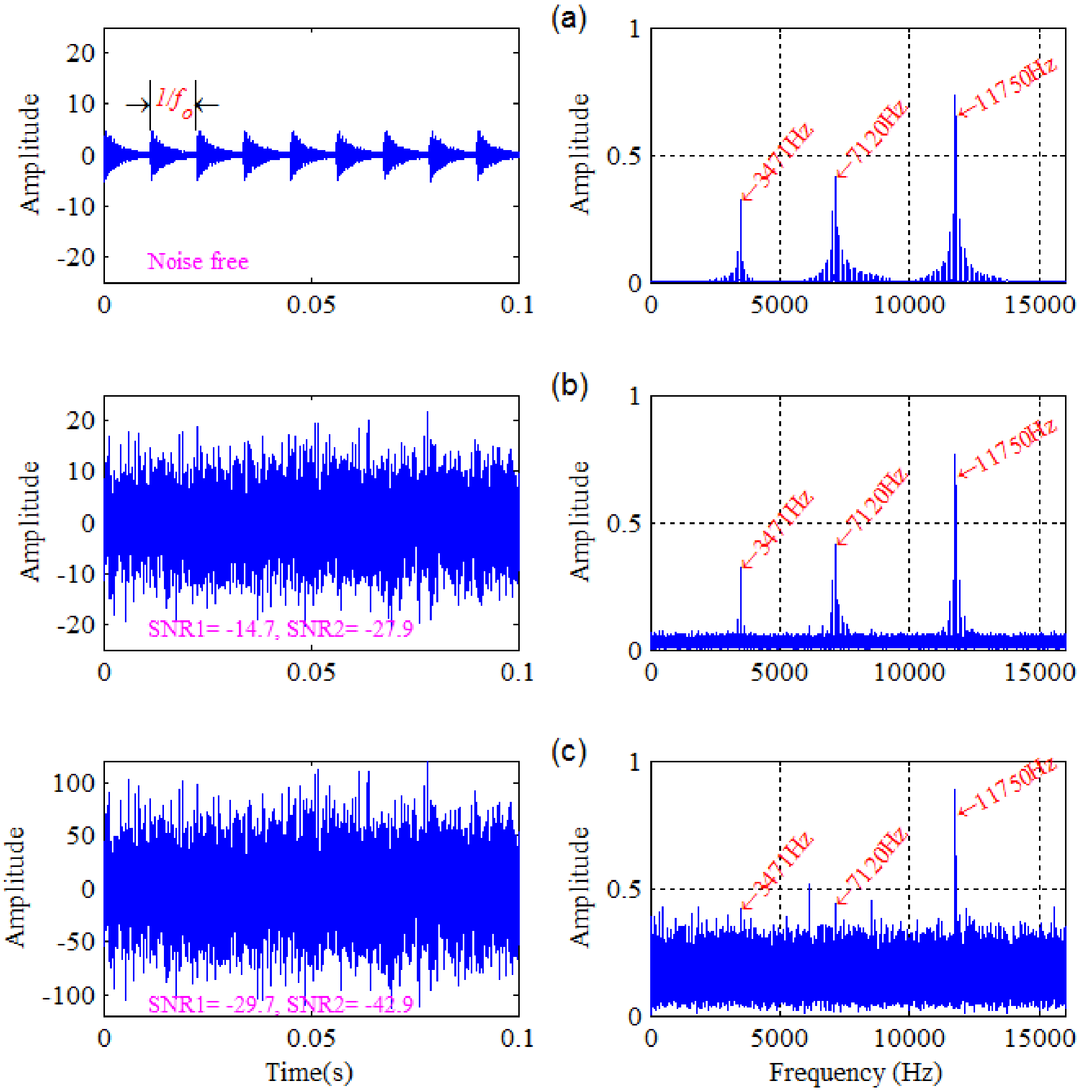

Fig. 5 Time waveforms of the simulated signals and their spectra with different levels of white noise: (a) noise free; (b) low level of white noise; and (c) high level of white noise

Taking the low noise scenario (Case A) as an example to describe the process of calculating the MSB robust detector, the steps are as follows. Firstly, calculate the MSB (the FFT size is 32,768 and the average time is 83) and also calculate the sideband estimator using Eqs. (6) and (7). Then, calculate the compound MSB slice $B(f_c)$ to choose the suboptimal f_c slices and hence to achieve the result displayed in Fig. 7(a). The f_c slices marked by '*' at around 7,000Hz and 12,000Hz have in this case been selected for the calculation of the MSB detector. Subsequently, the robust detector is calculated using Eq. (8), as shown in Fig. 7(c). As a benchmark, the fast Kurtogram algorithm [25] has also been applied to optimise the filter parameters for a narrowband envelope analysis, as shown Fig. 7(b). The envelope spectrum obtained by the Kurtogram is referred to as the Kurtogram-based detector in this paper. The Kurtogram optimised filter centre is at 11,719Hz and the filter bandwidth is 372.02Hz. Both the MSB and Kurtogram-based approaches find the highest resonance frequency at around 11,719Hz as the optimal band, and this was confirmed by multiple simulation runs.

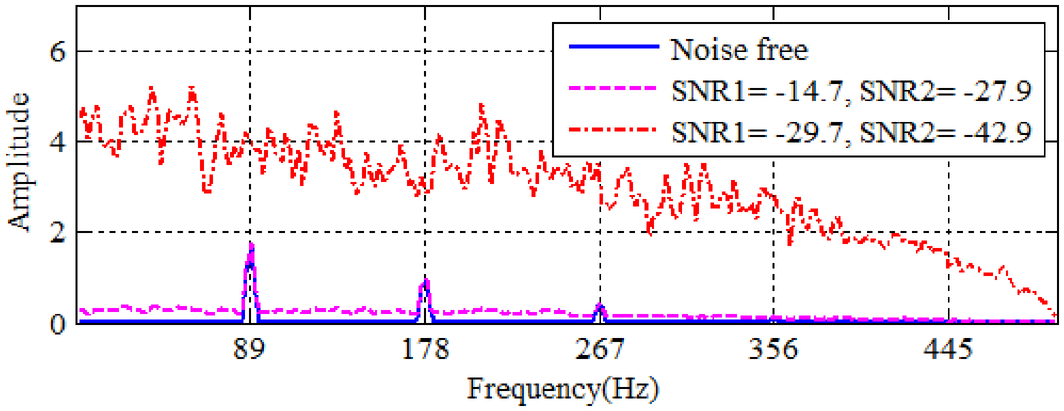

Fig. 6 Detection by conventional envelope spectra

Fig. 7(c) shows the normalised results of the MSB detector and the Kurtogram-based detector for the low noise signal with low impulses and no slippage (Case A). It can be seen that both detectors have distinctive peaks at the outer race fault characteristic frequency and its harmonics, indicating that they are capable of detecting the bearing fault for this low noise case. It is worth noting, however, that the MSB-based detector has lower background noise than that based upon the Kurtogram (which shows more harmonics along with noise).

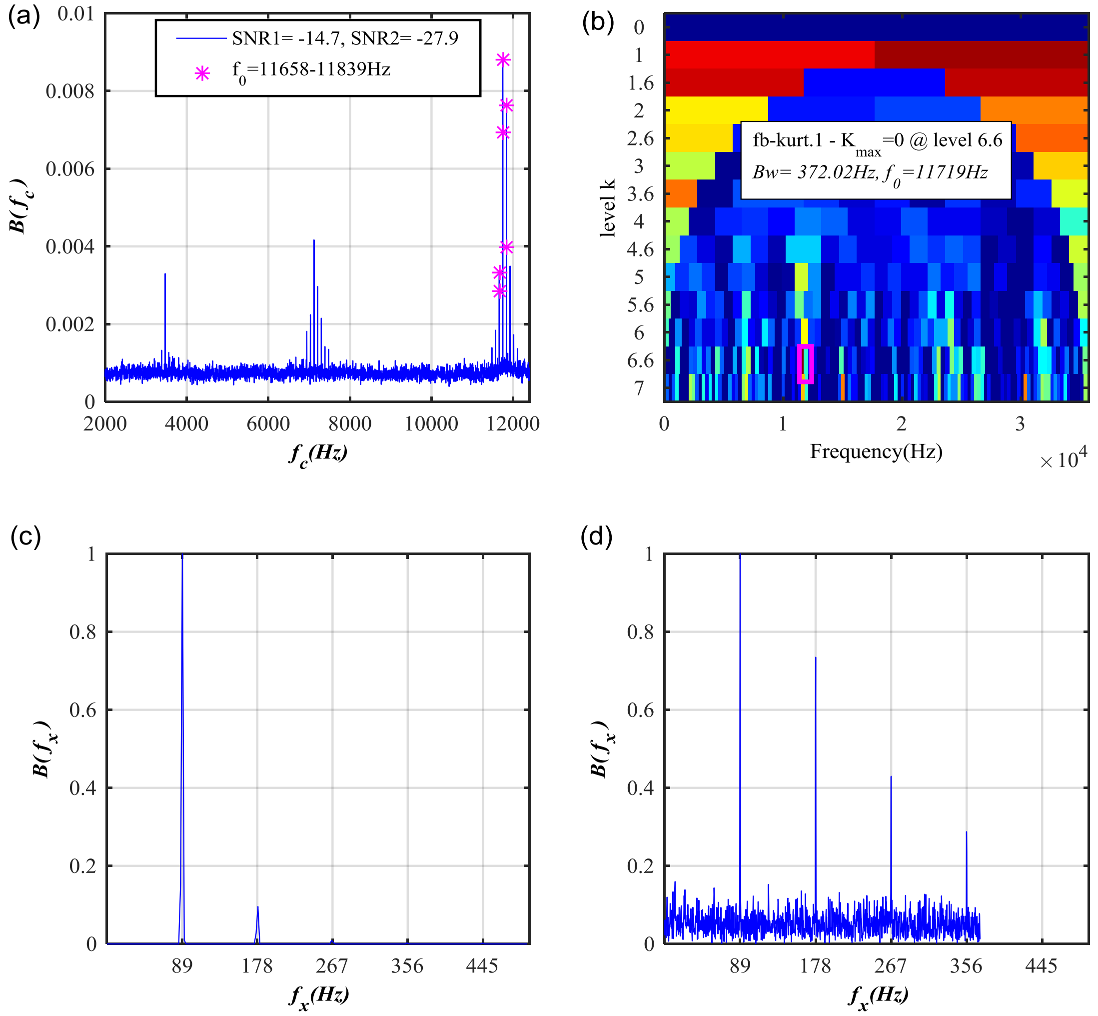

Fig. 7 Results of the MSB robust detector and the Kurtogram-based detector (low noise, low impulsive interference and no slippage, Case A): (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; and (d) Kurtogram based detector

For the high noise, low impulses, no slippage scenario (Case B), the compound MSB slice $B(f_c)$ and the normalised results of the two detectors are shown in Fig. 8. It can be seen from Fig. 8(a) that there are several significant peaks at around 11,719Hz, which is the highest resonance frequency of the simulated signal. In contrast, the central frequency optimised by the Kurtogram is at 26,768Hz and the bandwidth is 17,857Hz, which is not the location of the resonance frequency and bandwidth is too wide.

In summary, the MSB detector can still extract the fault feature frequency even if the noise is very high, whereas the Kurtogram approach does not give any indication of the fault. This greater capability of the MSB detector is attributed to its high performance noise suppression.

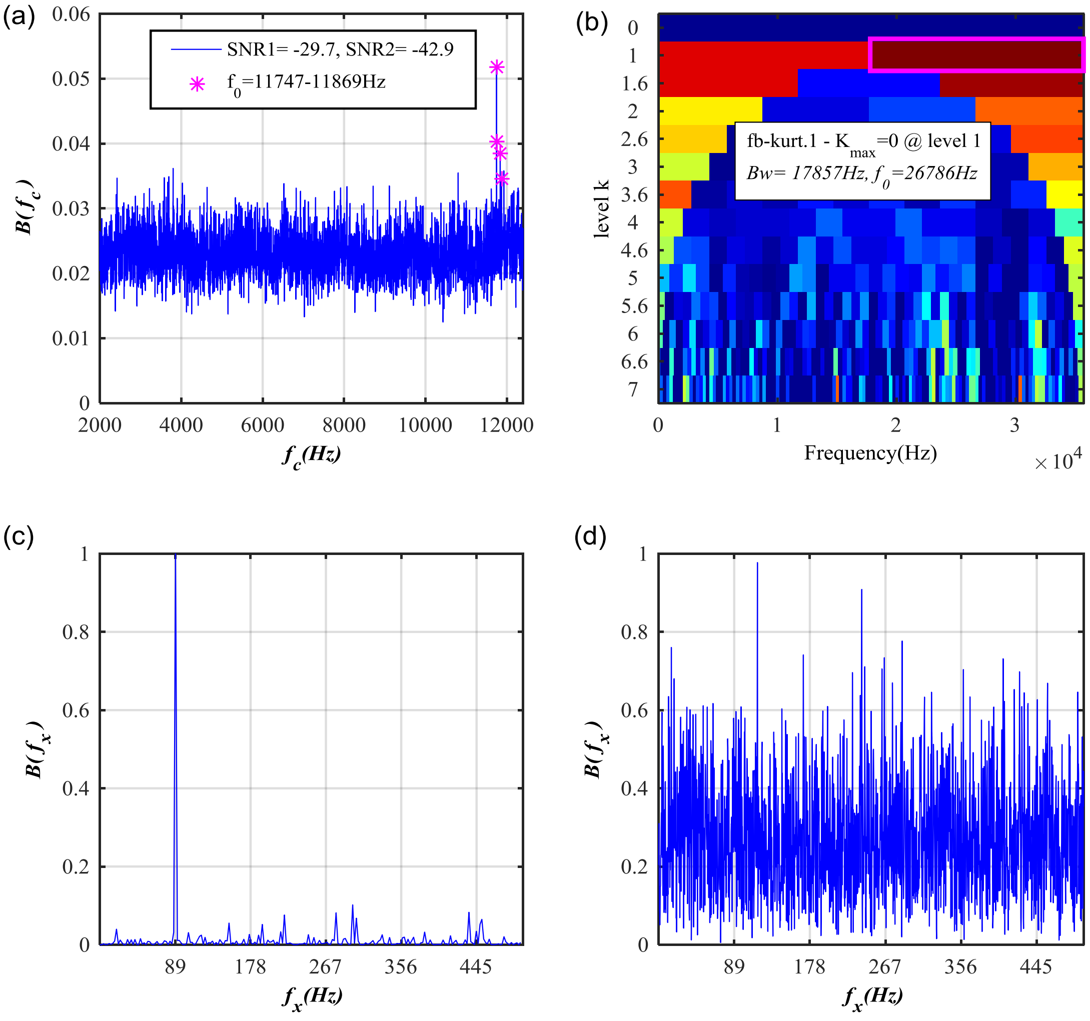

Fig. 8 Results of the MSB robust detector and the Kurtogram based detector (high noise, low impulsive interferences impulses and no slippage, Case B): (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; and (d) Kurtogram based detector

3.2 Robustness to aperiodic impulsive interference in the presence of white noise

To assess the performance of the proposed approach in the presence of aperiodic impulsive interference, two simulated fault signals were produced by adding 20 randomly occurring impulses with random amplitudes to the previous two noisy no-slippage signals in Case A and Case B, thus creating cases C and D. Although the same impulse features were added to cases A and B, the resulting two cases inevitably have different SNRs because of the differences in the random noise levels of A and B. Case C is hence defined as 'low noise with medium level impact interferences' and Case D is defined as 'high noise with high level impact interferences'. Neither Case C nor Case D contains any slippage features.

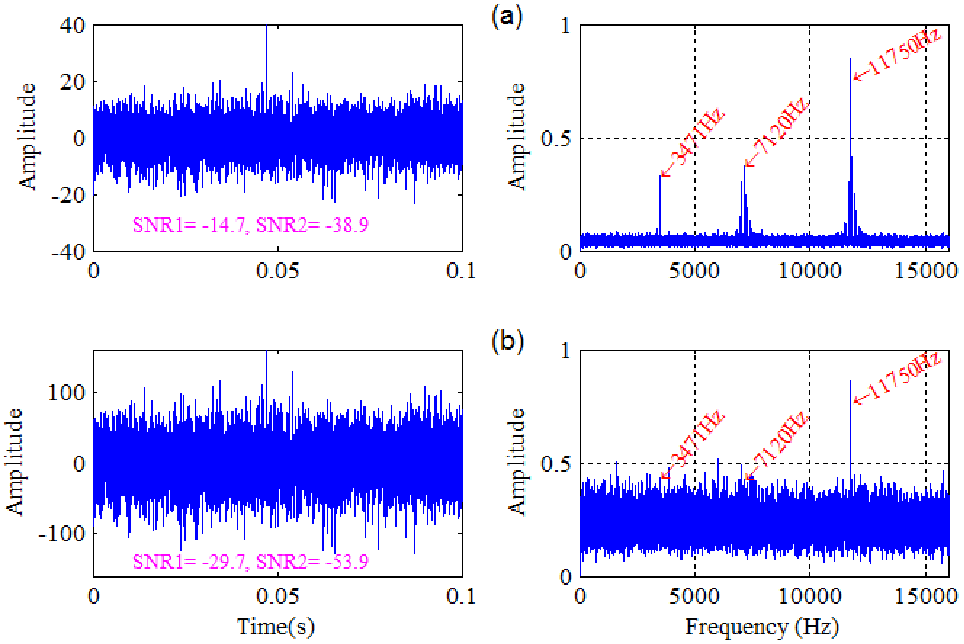

Fig. 9 Signal waveform and spectra with two levels of aperiodic impulses and white noise (cases C and D): (a) low level of aperiodic impulses; and (b) high level of aperiodic impulses

Fig. 9 shows the time waveforms of these simulated signals and their spectra (again, note the different amplitude scales for the two sets of time data). The SNR1 values are the same as those for Fig. 5, however the SNR2 values calculated using Eq. (11) are much lower than those for the stationary noises cases due to the higher localised peaks of the impulses. The resulting spectra show little difference, revealing that that the modulations of the periodic content are still sufficiently significant around the high frequency resonance at 11,719Hz.

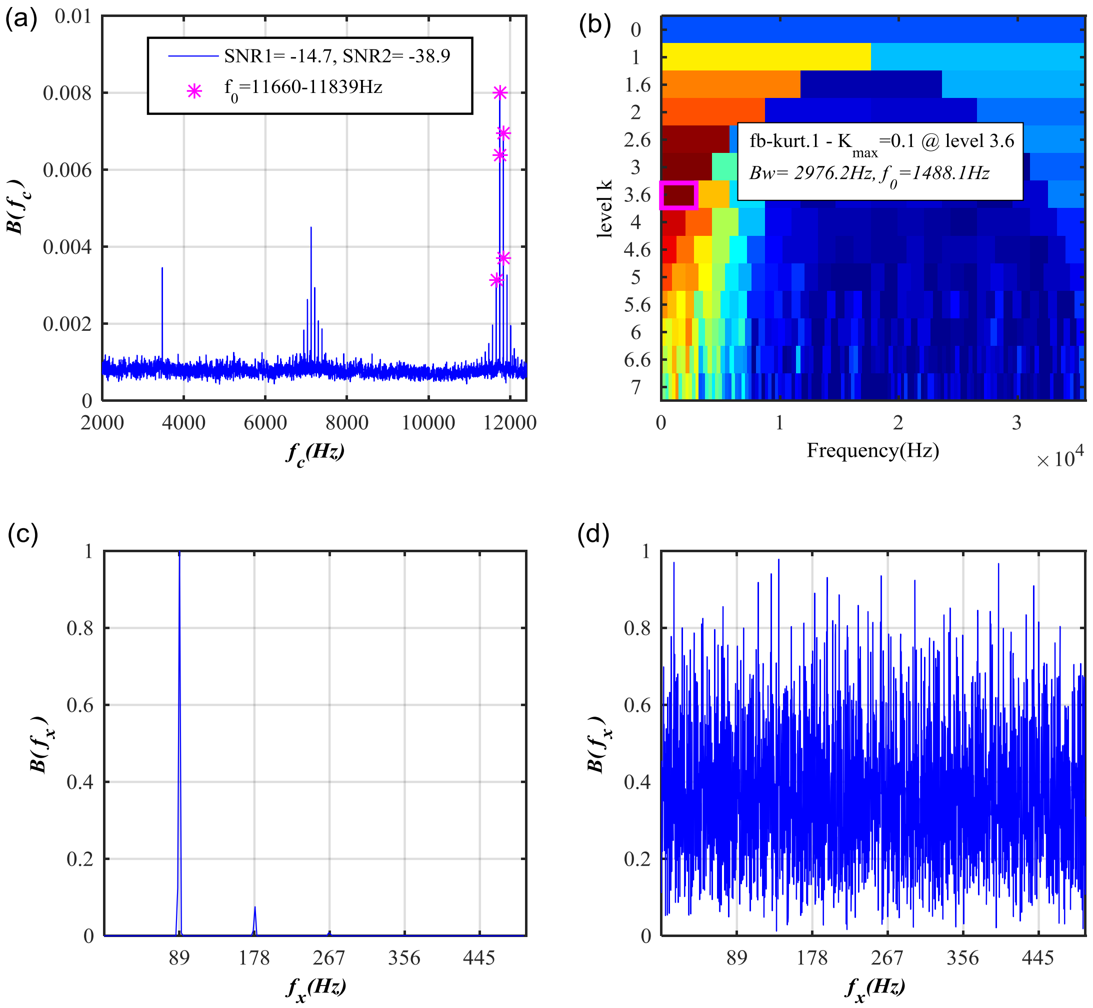

Fig. 10 Results of the MSB robust detector and the Kurtogram-based detector (low noise, medium impulsive interference and no slippage, Case C): (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; and (d) Kurtogram based detector

Using the same calculation parameters, the MSB and Kurtogram detection results were obtained as shown in Fig. 10 and Fig. 11, for Case C and Case D respectively. These figures reveal that the MSB detector is still able to detect the presence of the fault for both Case C and Case D, but because of the influence of local interferences the Kurtogram-based detector cannot find the correct frequency bands and is unable to provide a correct detection, even for the less severe medium level interference of Case C.

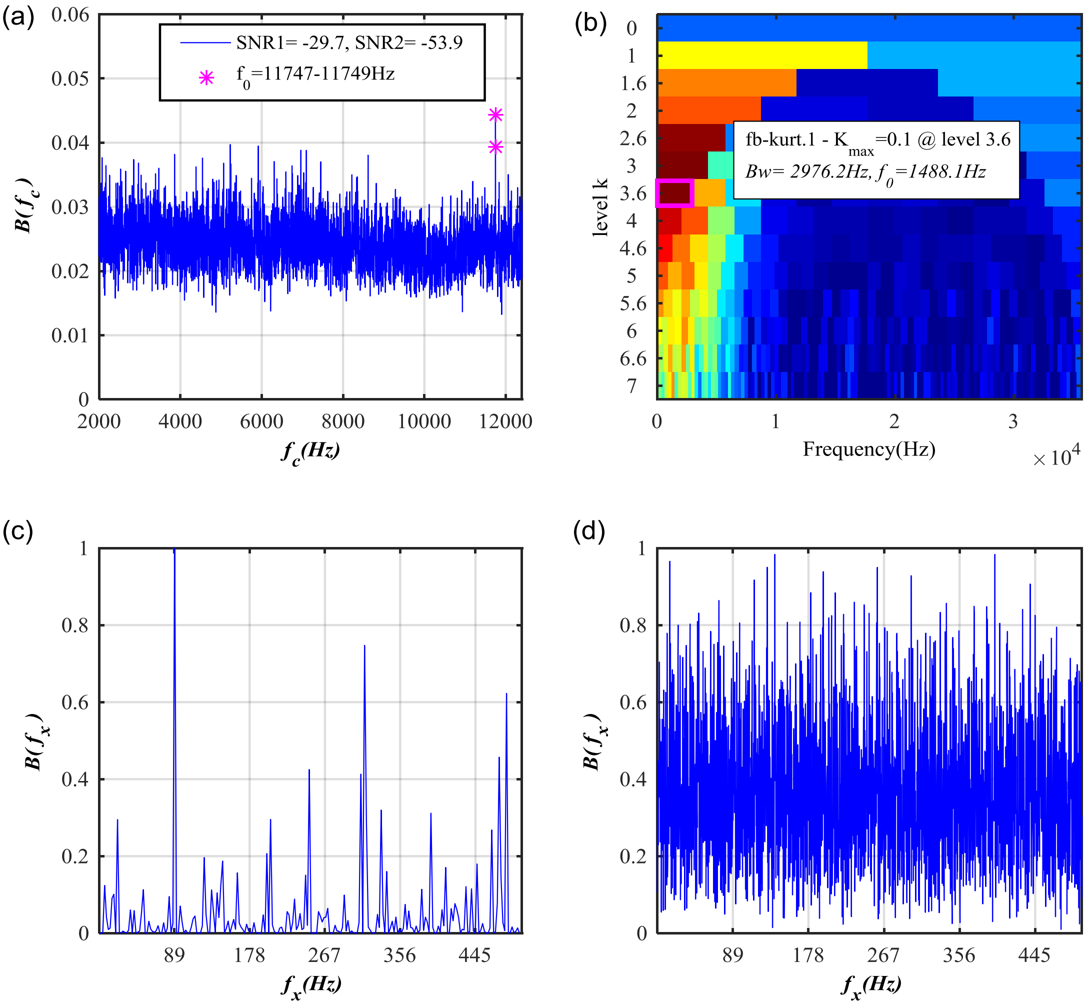

Fig. 11 Results of the MSB robust detector and the Kurtogram-based detector (high noise high impulsive interferences, and no slippage, Case D): (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; and (d) Kurtogram based detector

3.3 Robustness to random roller slippage

To evaluate the influences of rolling element slippage, which occurs randomly in many bearing applications, on the robustness of detection provided by the two methods, two further evaluation signals were produced with differing degrees of slippage, as described at the beginning of Section 3.

Fig. 12 shows the time randomness of 386 fault feature pulses relative to the base period, with $T_0 = 1/89 = 0.112$ (s), and for differing degrees of slippage (0%, 0.5% and 2%). Clearly, with more slippage induced, the deviation from the base period is more significant and consequently the vibration responses contain more random content and less the periodic content. However, as shown in Fig. 12(b), the time domain responses still exhibit clear periodic profiles with respect to the base period for the two slippages cases (0.5% and 2%) although the individual transient processes show observable time shifts from their corresponding baselines. These shifts make the responses in the three resonance bands much smaller and the periodic spectral components significantly smeared. As shown in the magnified spectra of Fig. 12(c), the smearing effects become more serious with increasing slippage and also in the higher frequency resonance bands (because of the multiplying effect of harmonic order). Nevertheless, the responses at the resonances are still significant, showing that periodic content remains and hence provides a basis for bearing fault detection. The random slippage can be considered as an additional form of noise which needs to be suppressed by the detection method to achieve reliable detection results.

To confirm the robustness of the detection to the random slippage, the signal from the most severe 2% slippage was added to each of the white noises case A and B, and the resulting signals E and F were evaluated with the two detection methods. Fig. 13(a) and (c) show that the MSB detector can still find the correct frequency bands and hence can produce a reliable detection result for the Case E signal which has a SNR of -14.7dB. This result agrees with that of the Kurtogram-based detector, the results of which are shown in Fig. 13(b) and (c), in that the optimal frequency band is centralised around $f_0 = 11719Hz$, which is the third resonance band, and also that there are significant components at 89Hz which is the characteristic fault frequency of the outer race. It worth noting that the MSB detector selects a number of continuous peaks from the compound MSB slice $B(f_c)$ because of the spectral smearing effect induced by the random slippage, rather than just a few discrete peaks for no slippage cases. This shows that the MSB detector also offers excellent performance even if the modulation signal has spectral smearing effect due to roller slippage and under moderate random noise conditions (Case E).

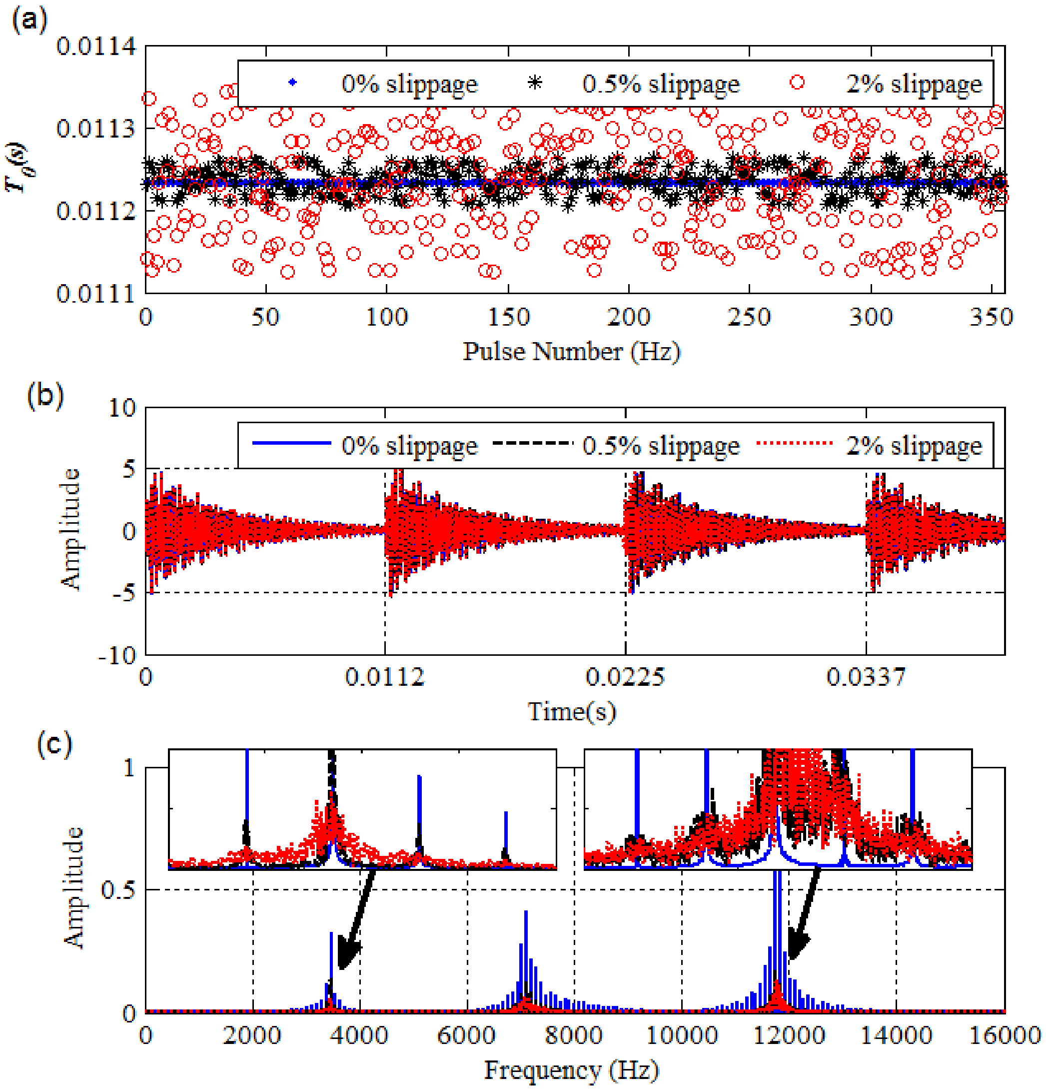

Fig. 12 Impulse responses with random roller slippages: (a) the time randomness of fault feature pulses relative to the base period; (b) the time domain responses; and (c) the spectra

When the SNR1 becomes even lower than -15dB, by adding more white noise, the Kurtogram-based detector can no longer find the optimal bands and is hence unable to produce a correct detection result. In contrast, the MSB detector can still produce a correct detection even with the SNR as low as -22dB (Case F). Fig. 14(a) and (c) show that, in addition to the significant fault component, the MSB detector also produces some observable components in the low frequency range, most probably due to the spectrum smearing effect induced by the random slippage. Although these low frequency components could lead to the potentially spurious detection of faults such as a bearing cage issue or a misalignment, the primary fault on the outer race has much greater amplitude than these low frequency features and can be detected without any difficulty.

The robustness of the MSB detector to noise when slippage effects are present is understandably lower than without slippage because the slippage responses already have substantial random effects contained within them.

From these simulation studies, it can be concluded that in principle the proposed MSB detector is robust to white noise, to aperiodic impulsive interferences and also to random roller slippages, and that in severe noise cases it outperforms the Kurtogram-based approach.

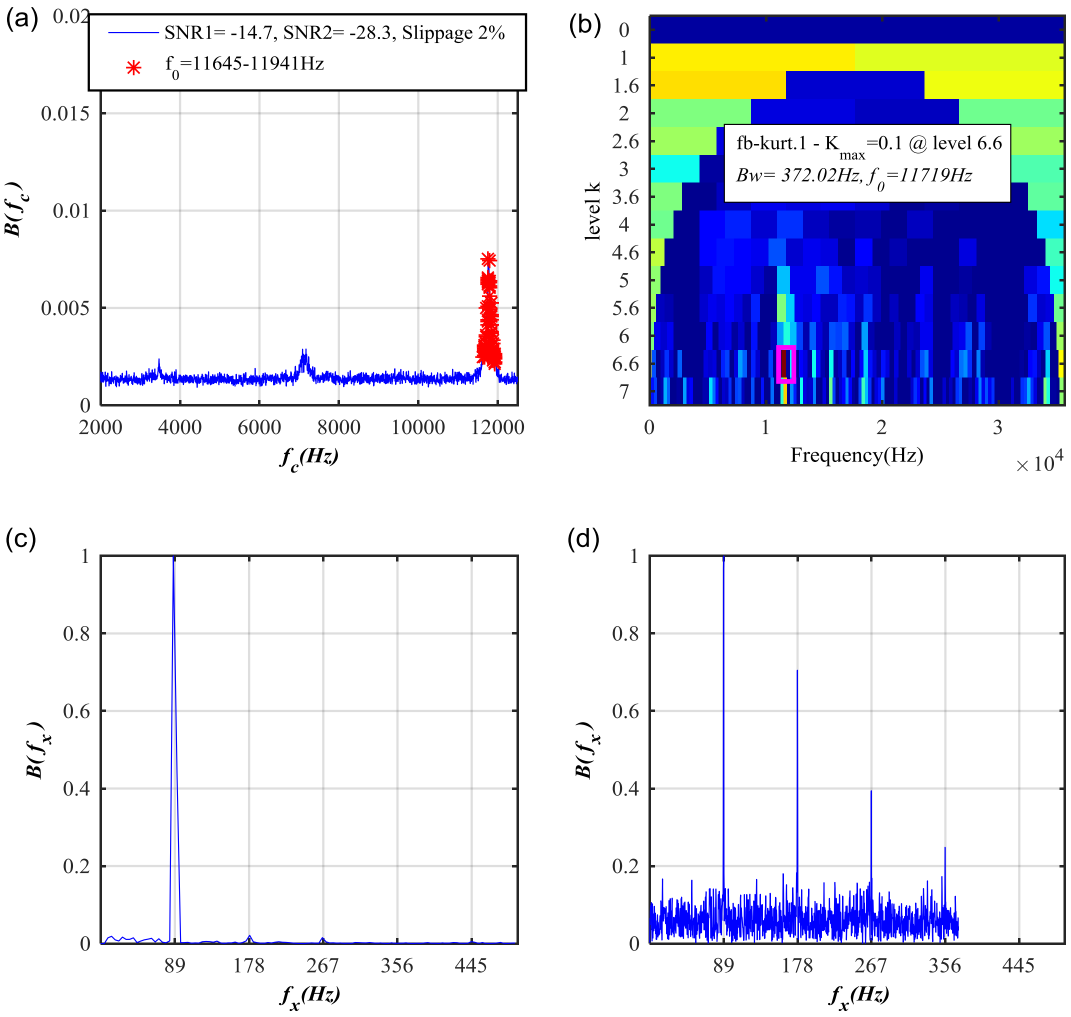

Fig. 13 Detection results from two detectors for the signal with high random slippage (2%), low noise (SNR1=-14.7dB) and low impulsive interferences (SNR2=-28.3dB), Case E: (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; (d) Kurtogram based detector

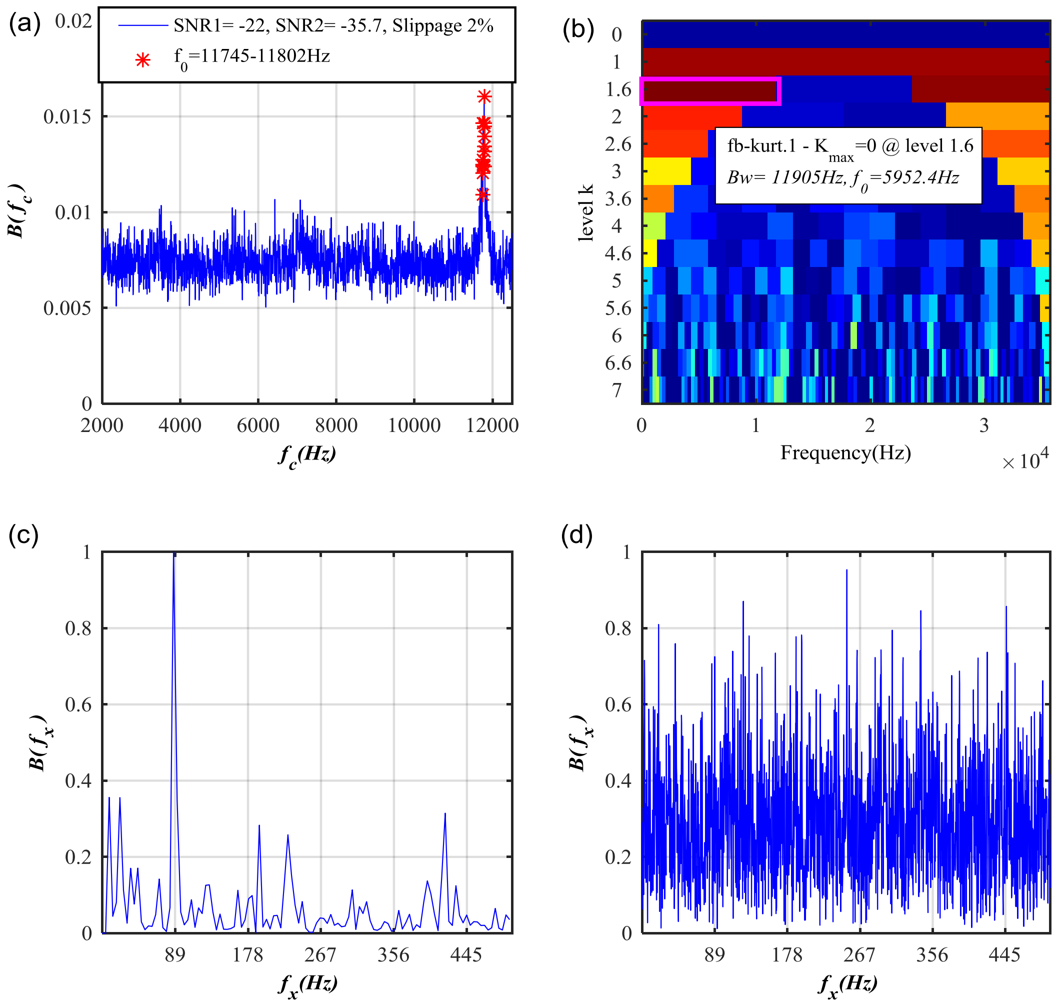

Fig. 14 Detection results from two detectors for high random slippage (2%) high noise (SNR1=-22dB) and low impulsive interferences (SNR2=-35.7dB), Case F (a) the compound MSB slice $B(f_c)$; (b) Kurtogram; (c) MSB robust detector; (d) Kurtogram based detector

4. Application case studies

To explore the practical application of the MSB detector alongside the Kurtogram-based detector, two bearing application cases have been investigated. One is for the bearing in an induction motor and the other is for the bearing in a planetary gearbox. Motor vibration can generally be expected to have lower noise and a narrower bandwidth compared to planetary gearbox vibration, where signals can be expected to contain higher noise levels and a wider bandwidth because of the impulsive excitations caused by the complex gear meshing processes. The monitoring of electric motor driver has always been of interest because of their popularity and importance as prime moves but the condition monitoring of planetary gearboxes has also received much attention in more recent years because of their prominence in wind turbine applications [26] [27].

4.1 Motor bearing fault detection

4.1.1 Experimental setup

The experimental data of the motor bearing analysed in this paper was collected from the bearing test rig illustrated in Fig. 15. It is comprised of a motor, coupling, intermediate shaft, supporting bearings and electrical brake. The vibration sensor was located in the vertical direction on the motor drive end bearing housing. Fig. 16 shows a photograph of the tested bearing, with a small seeded outer race defect.

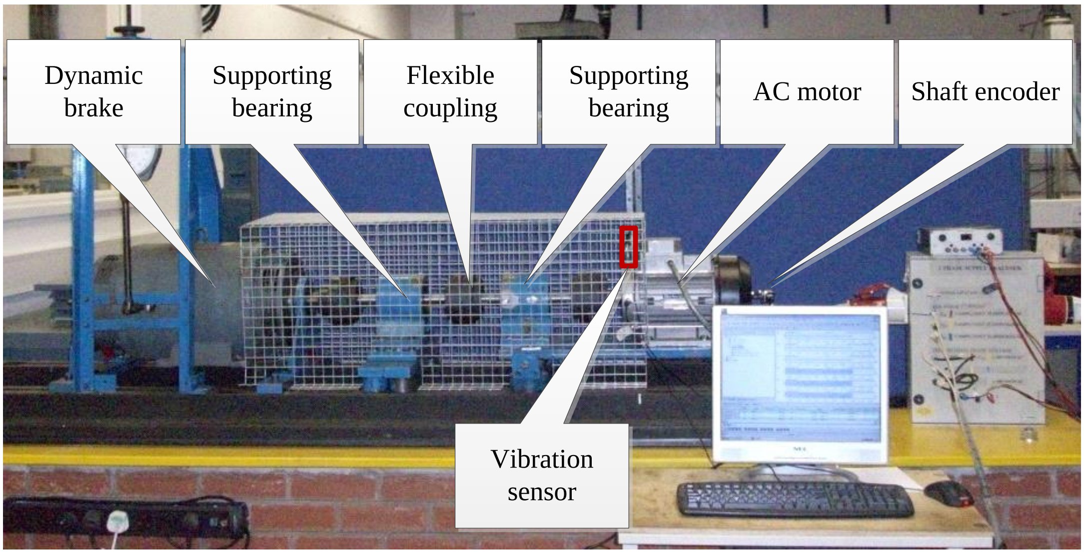

Fig. 15 Photograph of the motor bearing test rig

The tested motor bearing was a NSK Type 6206ZZ deep groove ball bearing with geometry listed in Table 2. The frequency range of the piezoelectric accelerometer used to collect the data was 0.5Hz to 10kHz and the sensitivity was 1.04mV/ms^2. The data was acquired with a sample rate of 96kHz and 24-bit resolution.

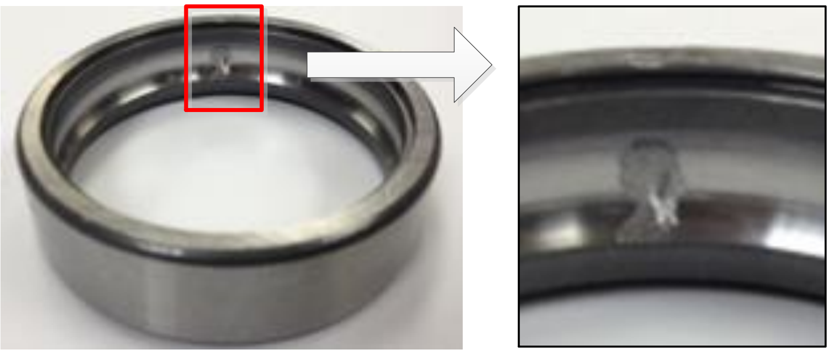

Fig. 16 Photographs of the test bearing with a small seeded outer race defect

Table 2. Specification of NSK Type 6206ZZ deep groove ball bearing

Parameter	Measurement
Pitch Diameter D_c	46.4mm
Ball Diameter D_b	9.53mm
Ball Number N_r	9
Contact Angle φ	0°

Given the horizontal orientation of the shafts in the machine, and the minimal axial load applied to the test bearing, the contact angle (φ) was assumed to be zero.

4.1.2 Detection results and discussion

Fig. 17 shows the vibration time waveform collected on the bearing test rig, along with its associated spectrum. From the vibration spectrum, it can be seen that there are two main resonance frequencies at approximately 2.5kHz and 7kHz. The optimised filter central frequency from the Kurtogram is at 2187.5kHz and bandwidth is 625Hz.

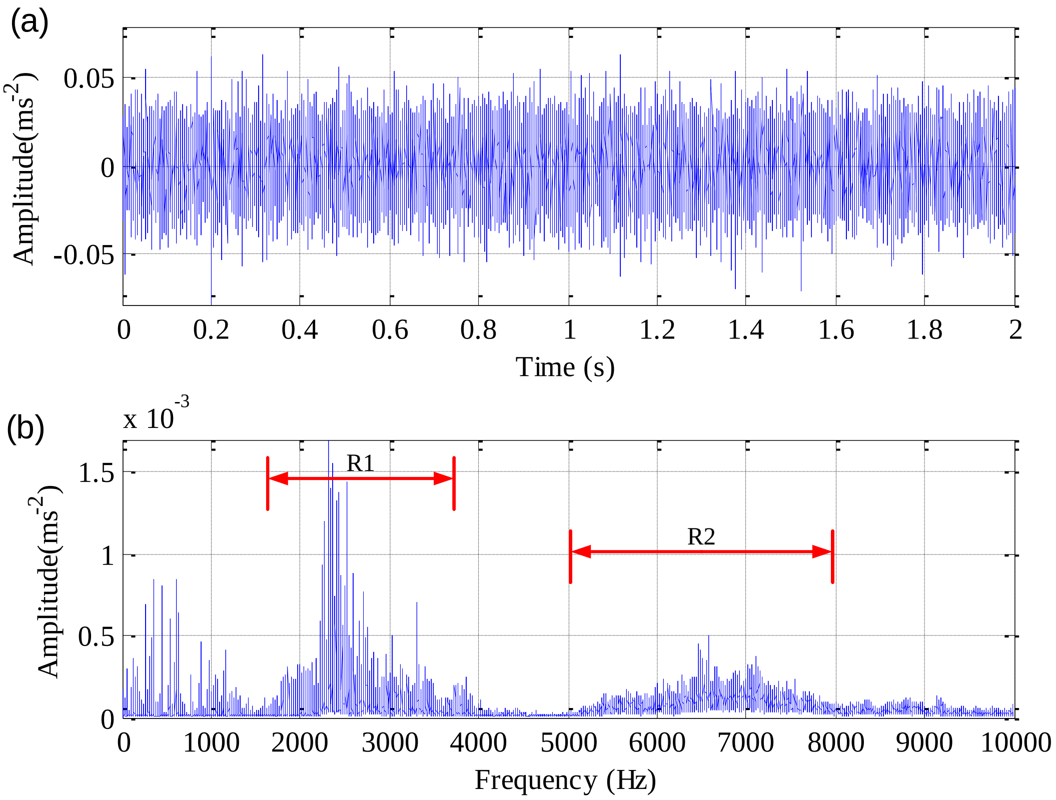

Fig. 17 Waveform of the motor vibration and its spectrum: (a) waveform; and (b) spectrum

Fig. 18 shows the compound MSB slice $B(f_c)$ and the normalised results of the MSB detector. The characteristic frequencies of the tested bearing are marked by different colours of dashed line, with f_i, f_o, f_{cage} and f_b indicating the characteristic frequencies of faults on the inner race, outer race, cage and ball respectively. The suboptimal f_c slices are selected in the range from 6.4kHz to 7.1kHz which corresponds to the second resonance frequency in Fig. 17(b). From the MSB detector results in Fig. 18(b), it can be seen that there is one distinctive peak at the outer race fault frequency. In addition, small peaks appear at the cage fault frequency and its harmonics, but no peaks can be observed at the ball fault frequency or the inner race fault frequency. These results demonstrate that the outer race fault can readily be detected in the test bearing. The presence of the cage fault frequency was unanticipated may be caused by manufacturing effects or inadvertent damage during bearing installation.

The Kurtogram based detector shown in Fig. 18(c) also provides clear indication of the bearing outer race defect.

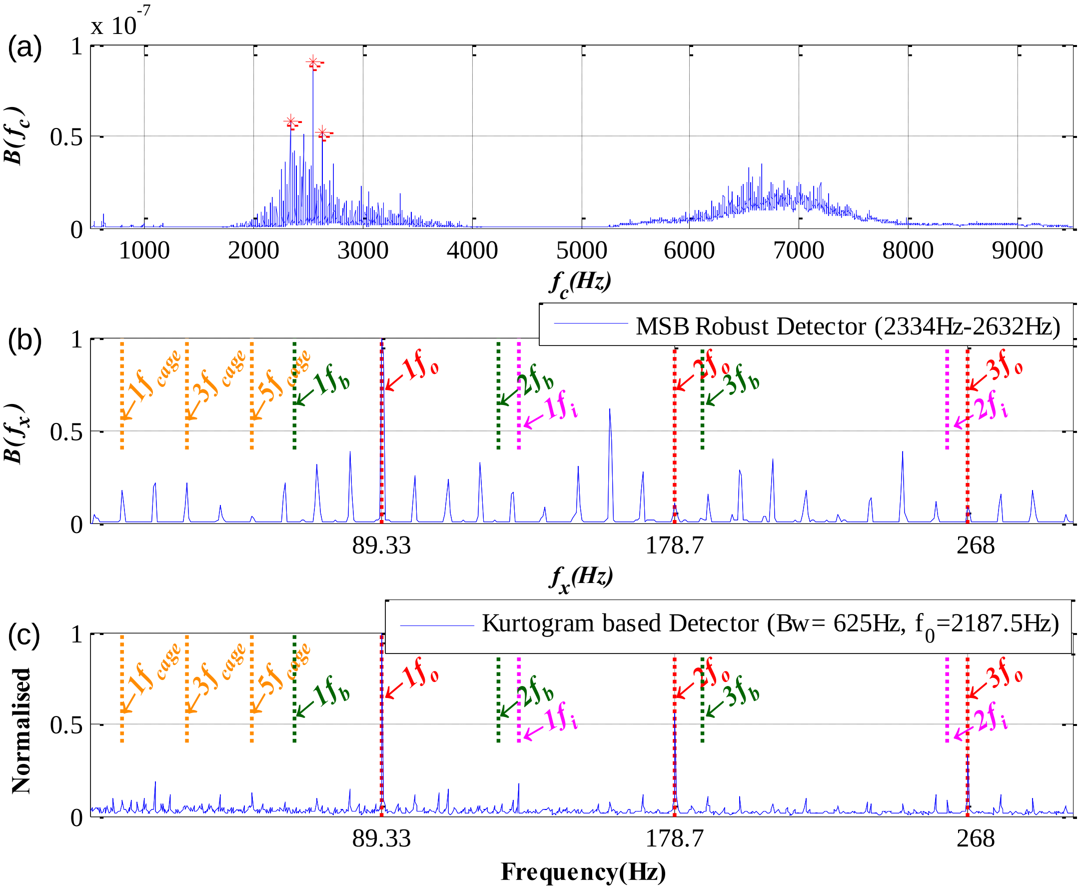

Fig. 18 Results of the MSB robust detector: (a) the compound MSB slice $B(f_c)$; (b) MSB robust detector; and (c) Kurtogram based detector

4.2 Planetary gearbox bearing fault detection

4.2.1 Experimental setup

To assess the effectiveness of the MSB detector for use in low SNR conditions, vibration signals acquired from a planetary gearbox test system were investigated. The test rig shown in Fig. 19, uses a planetary gearbox with a rated torque of 670Nm, and a maximum input speed of 2800rpm with a resulting output speed of 388rpm. The schematic in Fig. 20 illustrates the position of the accelerometer that was mounted on the outer housing of the ring gear along with the location of the test bearing. In contrast to the outer race fault seeded in the motor bearing, in this test an inner race fault was seeded, as shown in Fig. 21.

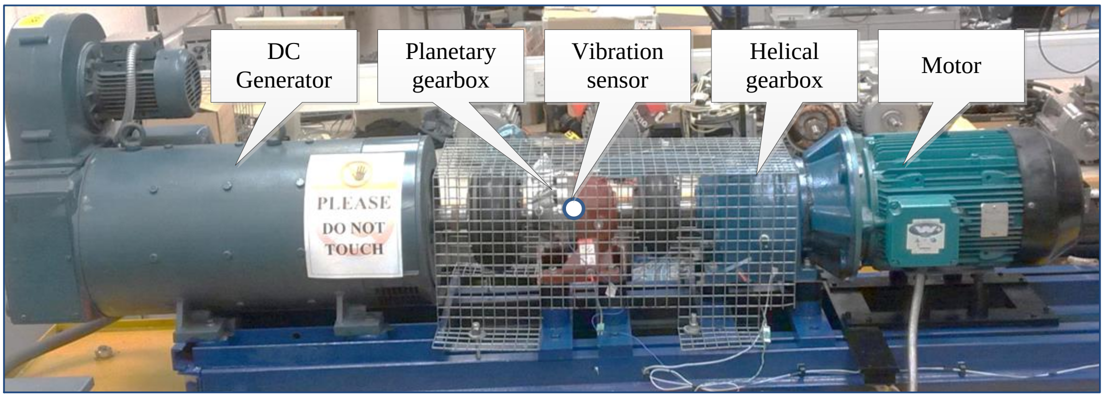

Fig. 19 Photograph of the planetary gearbox test rig

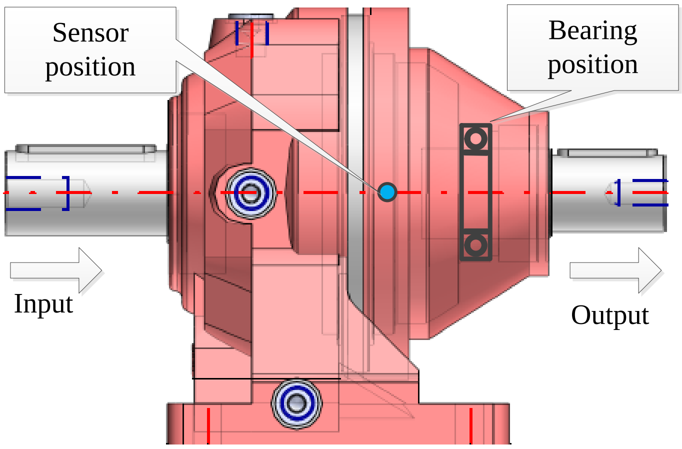

Fig. 20 Schematic for a planetary gearbox

Fig. 21 Inner race defect on bearing

The tested bearing was an SKF Type 6008 deep groove ball bearing with geometry as listed in Table 3. The linear frequency range of the vibration accelerometer used to collect the data was 0.5Hz to 10kHz, with a resonance higher than 35kHz and a sensitivity of 28.7mV/ms^{-2}. The data was again acquired with a sample rate of 96kHz and 24-bit resolution.

Table 3. Specification of SKF 6008 deep groove ball bearing

Parameter	Measurement
Pitch Diameter D_c	54mm
Ball Diameter D_b	7.9mm
Ball Number N_r	12
Contact Angle φ	0°

Given the horizontal orientation of the shafts in the machine, and the minimal axial load applied to the test bearing, the contact angle (φ) was assumed to be zero.

4.2.2 Detection results and discussion

Fig. 22 shows the waveform of the measured vibration signal and its spectrum. From Fig. 22(a), it can be seen that many impulses exist in the vibration signal, generated by the complex rotation and meshing dynamics of the planetary gearbox. Fig. 22(b) shows that there are four possible main resonance frequencies at approximately 1.2kHz, 2kHz. 6kHz and 9kHz. The optimised filter location from the Kurtogram has a central frequency of 5703.1Hz and bandwidth of 156.3Hz, which corresponds to the third resonance frequency.

For the planetary gearbox vibration signal analysis, it is necessary to check the nonlinear degree of the signal. Because the vibration of planetary gearbox is quite complex, it includes both vibration of gear meshing and bearing. The collected vibration is the superposition of all the components, which can reduce the coupling degree between the components. To measure the degree of coupling between three components, a modulation signal bicoherence (MSBc) can be used and calculated as follows:

$$b_{MS}^2(f_1,f_2)=\frac{\left|B_{MS}(f_1,f_2)\right|^2}{E\left\langle\left|X(f_2)X^*(f_2)\right|^2\right\rangle E\left\langle\left|X(f_2+f_1)X(f_2-f_1)\right|^2\right\rangle} \qquad (12)$$

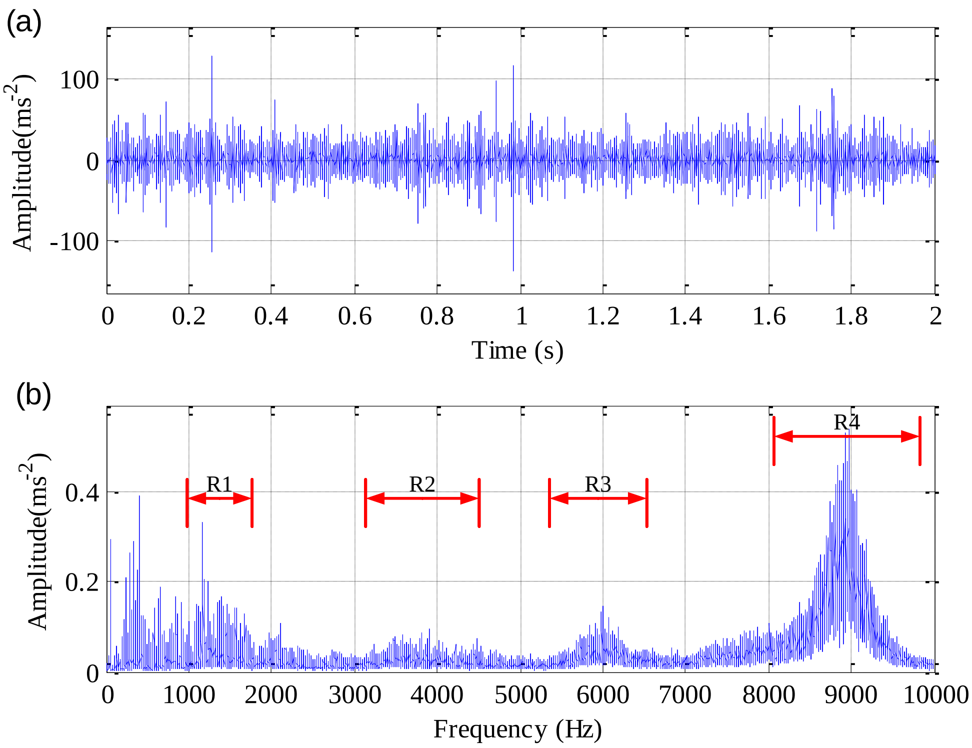

Fig. 22 Time waveform of planetary gearbox vibration and its spectrum: (a) waveform; and (b) spectrum

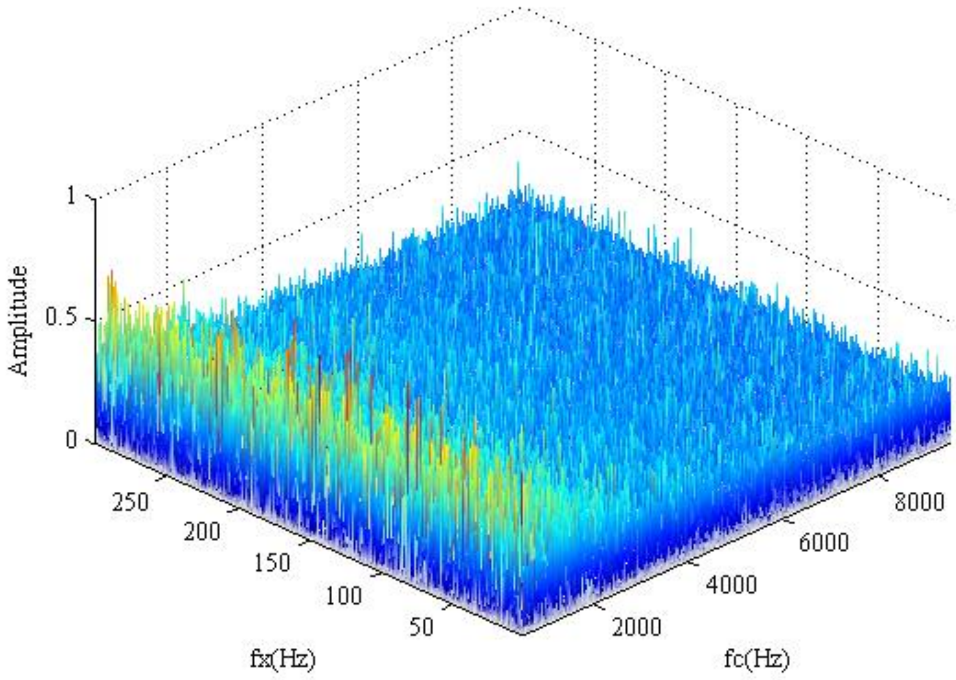

Fig. 23 MSB coherence of planetary gearbox vibration

The MSBc result of the planetary vibration signal is shown in Fig. 23. It can be seen that the coherence is low when f_c is higher than 6kHz, which means the modulation effect is weak and the MSB result is not reliable in this frequency range. Therefore, the highest resonance frequency R4 is excluded for the calculation of the MSB detector.

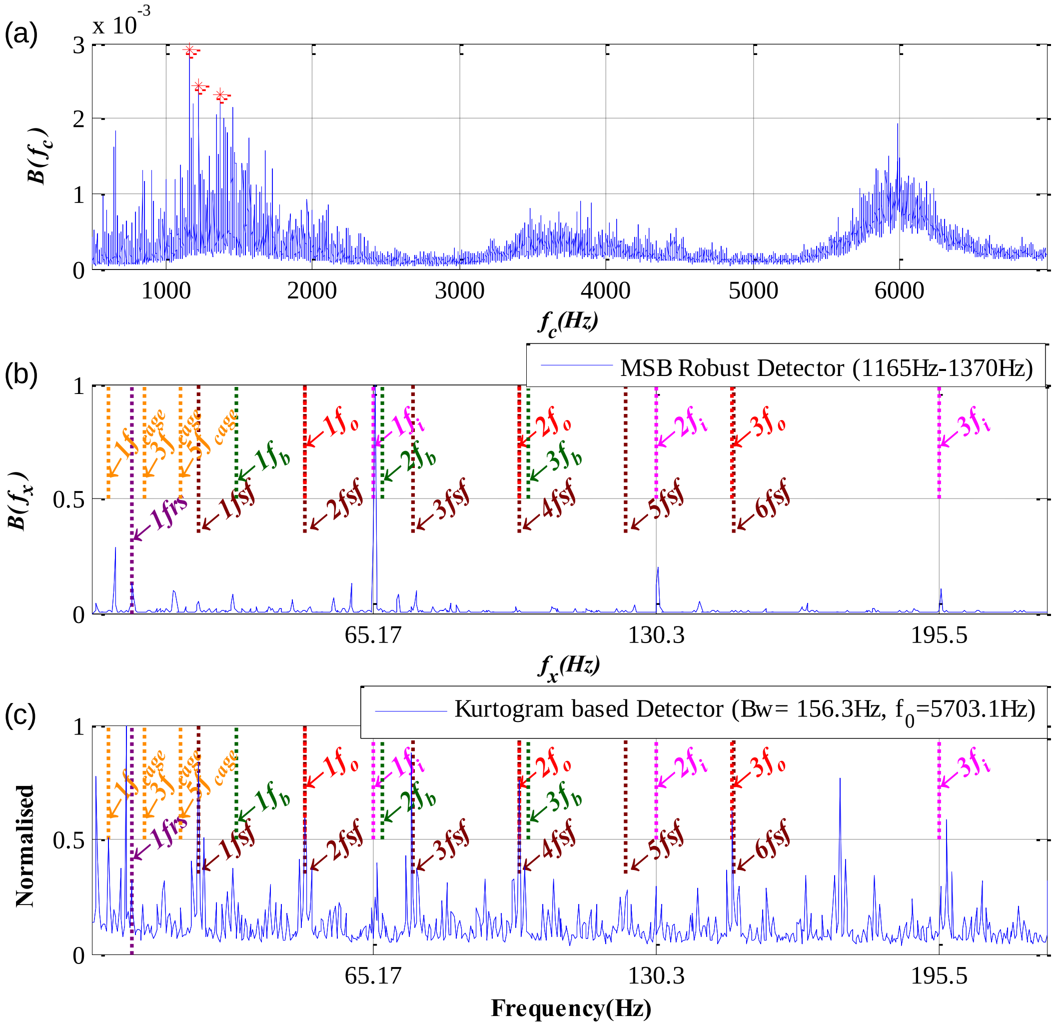

Fig. 24 Results of the MSB detector and the Kurtogram-based detector: (a) the compound MSB slice $B(f_c)$; (b) MSB robust detector; and (c) Kurtogram based detector

$B(f_c)$ is presented in Fig. 24(a), and this shows that the selected suboptimal f_c slices exist in the range from 1.1kHz to 1.5kHz which corresponds to the second resonance frequency of Fig. 24(b). The normalised results of the MSB detector are presented in Fig. 24(b). The characteristic frequency positions for the test bearing f_i, f_o, f_b and f_{cage}

indicate the fault characteristic frequencies of the bearing inner race, outer race, ball, and cage respectively. f_{rs}, f_{rsc}, and f_{sf} denote the shaft rotational frequency, the difference between the shaft rotational frequency and the carrier rotational frequency, and the sun gear fault frequency calculated by Eq. (13) [28] respectively.

$$f_{sf} = K\left(f_{rs} - f_{rc}\right) \tag{13}$$

where K is the number of planetary gears, and f_{rc} is the carrier rotating frequency

Unfortunately, the bearing outer race fault frequency coincides with the sun gear fault frequency, which potentially would make it difficult to distinguish between these two types of fault. From the MSB robust detector shown in Fig. 24(b), it can be seen that there is one distinctive peak at the inner race fault frequency of the bearing along with two small peaks at its second and third harmonics. This result shows that inner race fault can be straightforwardly detected in the test bearing. Small peaks also appear at the sun gear fault frequency or the bearing outer race fault frequency and harmonics thereof. These peaks are also unanticipated and may again be due to manufacturing effects or inadvertent installation damage. This means that the MSB robust detector can achieve accurate and reliable bearing fault diagnosis even if in a low SNR, high impact environment like planetary gearbox vibration.

From the results of the Kurtogram based detector, illustrated in Fig. 24(c), it can be seen that there are numerous frequency peaks associated with the sun gear fault/outer race fault, the inner race fault, and their harmonics. However, it is difficult to identify the bearing fault because the amplitudes of the bearing characteristic frequencies are significantly smaller than other frequency components such as the sun gear fault frequency/outer race frequency, which was also further confirmed by setting filter band around the R4 shown in Fig. 22(b).

5. Conclusions

Based on the proven performance of the MSB in suppressing random noise and decomposing the nonlinear modulation components [11, 15-17], a novel MSB detector has been developed using a number of significant MSB peaks, which are optimal in terms of maximising the modulation contents of bearing fault signals. Simulated signals with different levels of white noise, aperiodic impulsive interference and random roller slippages have been applied to demonstrate the robust performance of the new approach, and its capability has been shown to exceed that of the Kurtogram compared with that of the Kurtogram-based detector. The application to signals from a planetary gearbox shows that the new approach can successfully detect bearing faults in circumstances where the noisy signal contains high levels of modulation due to other impact phenomena.

References

[1] P.D. McFadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, Journal of Sound and Vibration, 96 (1984) 69–82.

[2] Y. Lei, Z. He, Y. Zi, A new approach to intelligent fault diagnosis of rotating machinery, Expert Systems with Applications, 35 (2008) 1593–1600.

[3] M.S. Darlow, R.H. Badgley and G.W. Hogg, Application of high-frequency resonance techniques for bearing diagnostics in helicopter gearboxes, Technical Report, US Army Air Mobility Research and Development Laboratory, (1974) 74–77.

[4] J. Antoni, Cyclic spectral analysis in practice, Mechanical Systems and Signal Processing, 21 (2007) 597–630.

[5] J. Antoni, Cyclostationarity by examples, Mechanical Systems and Signal Processing, 23 (2009) 987–1036.

[6] D. Ho, R.B. Randall, Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals", Mechanical Systems and Signal Processing, Mechanical Systems and Signal Processing. 14 (2000) 763–788.

[7] T. Barszcz, Decomposition of vibration signals into deterministic and nondeterministic components and its capabilities of fault detection and identification, International Journal of Applied Mathematics and Computer Science, 19 (2009) 327–335.

[8] N. Sawalhi, R.B. Randall, H. Endo, The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis, Mechanical Systems and Signal Processing, 21 (2007) 2616–2633.

[9] S. Zhao, L. Liang, G. Xu, J. Wang, W. Zhang, Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method, Mechanical Systems and Signal Processing, 40 (2013) 154–177.

[10] J. Antoni, R.B. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, Mechanical Systems and Signal Processing, 20 (2006) 308–331.

[11] Y. Wang, J. Xiang, R. Markert, M. Liang, Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications, Mechanical Systems and Signal Processing, 66–67 (2016) 679–698.

[12] J. Tian, C. Morillo, M.G. Pecht, Rolling element bearing fault diagnosis using simulated annealing optimized spectral kurtosis, in: 2013 IEEE Conference on Prognostics and Health Management (PHM), 2013, 1–5.

[13] F. Tafinine, K. Mokrani, Real time automatic detection of bearing fault in induction machine using Kurtogram analysis, The Journal of the Acoustical Society of America, 132 (2012) EL405–EL410.

[14] I. Rehab, X. Tian, F. Gu, A.D. Ball, The fault detection and severity diagnosis of rolling element bearings using modulation signal bispectrum, in: Eleventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Manchester, UK, 2014.

[15] M. Liang. Soltani Bozchalooi, An energy operator approach to joint application of amplitude and frequency-demodulations for bearing fault detection, Mechanical systems and signal processing, 24 (2010) 1473-1494.

[16] D. Ho, R.B. Randall, Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals, Mechanical systems and signal processing, 14 (2000) 763-788.

[17] R.B. Randall, J. Antoni, Rolling element bearing diagnostics—A tutorial, Mechanical systems and signal processing, 25 (2011) 485–520.

[18] Briiel & Kjaer, Detection faulty rolling-element bearings, application notes, BO 0210-11.

[19] J.R. Stack, R.G. Harley, T.G. Habetler, An amplitude modulation detector for fault diagnosis in rolling element bearings, IEEE Transactions on Industrial Electronics, 51 (2004) 1097–1102.

[20] A. Alwodai, T. Wang, Z. Chen, F. Gu, R. Cattley and A. D. Ball, A Study of motor bearing fault diagnosis using modulation signal bispectrum analysis of motor current signals, Journal of Signal and Information Processing, 04 (2013) 72–79.

[21] F. Gu, Y. Shao, N. Hu, A. Naid and A.D. Ball, Electrical motor current signal analysis using a modified bispectrum for fault diagnosis of downstream mechanical equipment, Mechanical Systems and Signal Processing, 25 (2011) 360–372.

[22] A. Naid, F. Gu, Y. Shao, S. Al-Arbi, A.D. Ball, Bispectrum analysis of motor current signals for fault diagnosis of reciprocating compressors, Key Engineering Materials, 413-414 (2009) 505-511.

[23] I. Rehab, X. Tian, F. Gu and A.D. Ball, A study of diagnostic signatures of a deep groove ball bearing based on nonlinear dynamic model, In: Proceedings of the 21st International Conference on Automation and Computing, 11-12 September 2015, Glasgow, UK.

[24] T. Harris, Rolling bearing analysis, John Wiley, New York, 2001, 1001.

[25] J. Antoni, Fast computation of the Kurtogram for the detection of transient faults, Mechanical Systems and Signal Processing, 21 (2007) 108–124.

[26] Z. Feng, M.J. Zuo, Vibration signal models for fault diagnosis of planetary gearboxes, Journal of Sound and Vibration, 331 (2012) 4919–4939.

[27] R. Zimroz, W. Bartelmus, T. Barszcz, J. Urbanek, Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings, Mechanical Systems and Signal Processing, 46 (2014) 16–27.

[28] M. Inalpolat and A. Kahraman, A theoretical and experimental investigation of modulation sidebands of planetary gear sets, Journal of Sound and Vibration, 323 (2009) 677-696.