ENHANCED INFORMATION EXTRACTION FROM NOISY VIBRATION DATA FOR MACHINERY FAULT DETECTION AND DIAGNOSIS

XIANGE TIAN

A thesis submitted to the University of Huddersfield

in partial fulfilment of the requirements for

the degree of Doctor of Philosophy

School of Computing and Engineering

The University of Huddersfield

January 2017

COPYRIGHT

ABSTRACT

As key mechanical components, bearings and gearboxes are employed in most machines. To maintain efficient and safe operations in modern industries, their condition monitoring has received massive attention in recent years. This thesis focuses on the improvement of signal processing approaches to enhance the performance of vibration based monitoring techniques taking into account various data mechanisms and their associated periodic, impulsive, modulating, nonlinear coupling characteristics along with noise contamination. Through in-depth modelling, extensive simulations and experimental verifications upon different and combined faults that often occur in the bearings and gears of representative industrial gearbox systems, the thesis has made following main conclusions in acquiring accurate diagnostic information based on improved signal processing techniques:

1) Among a wide range of advanced approaches investigated, such as adaptive line enhancer (ALE), wavelet transforms, time synchronous averaging (TSA), Kurtogram analysis, and bispectrum representations, the modulation signal bispectrum based sideband estimator (MSB-SE) is regarded as the most powerful tool to enhance the periodic fault signatures as it has the unique property of simultaneous demodulation and noise reduction along with ease of implementation.

2) The proposed MSB-SE based robust detector can achieve optimal band selection and envelope spectrum analysis simultaneously and show more reliable results for bearing fault detection and diagnosis, compared with the popular Kurtogram analysis which highlights too much on localised impulses.

3) The proposed residual sideband analysis yields accurate and consistent diagnostic results of planetary gearboxes across wide operating conditions. This is because that the residual sidebands are much less influenced by inherent gear errors and can be enhanced by MSB analysis.

4) Combined faults in bearings and gears can be detected and separated by MSB analysis. To make the results more reliable, multiple slices of MSB-SE can be averaged to minimise redundant interferences and improve the diagnostic performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBRAVIATIONS

ALE	Adaptive Line Enhancer
AM	Amplitude Modulation
ANC	Adaptive Noise Cancellation
AR	Auto-Regressive
CB	Conventional Bispectrum
CVM	Cross Validation Method
CWT	Continuous Wavelet Transformation
DE	Differential Evolution
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
FAM	Fuzzy ARTMAP
FIR	Finite Impulse Response
FT	Filter Tree
FTF	Fundamental Train Frequency
HBM	Harmonics Balance Method
IIR	Infinite Impulse Response
IMF	Intrinsic Mode Function
LCD	Local Characteristic-Scale Decomposition
LMD	Local Mean Decomposition
MSB	Modulation Signal Bispectrum
MSBc	Modulation Signal Bicoherence
MSB-SE	MSB Based Sideband Estimator
MSP	Morphological Signal Processing
NLMS	Normalized Least Mean Square
NLMS	Normalized Least Mean Square
PM	Phase Modulation
PSD	Power Spectral Density
QPC	Quadratic Phase Coupling
RMS	Root Mean Square
SK	Spectral Kurtosis
SNR	Signal To Noise Ratio
SPE	Scale Periodic Exponential
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
TFR	Time-Frequency Representation
TSA	Time Synchronous Averaging
WPD	Wavelet Packet Decomposition

WPT	Wavelet Packet Transform
WVD	Wigner-Ville Distribution

DEDICATION

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Huddersfield. I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Huddersfield.

This work is original except where acknowledgement and references are made to the previous work. Neither this nor any substantially similar dissertation has been or is being submitted for a degree, diploma or other qualification at any other university.

ACKNOWLEDGEMENTS

This work has been carried out in the school of Computing and Engineering, Centre for Efficiency and Performance Engineering (CEPE) at the University of Huddersfield, U. K. It was financially funded by the Fee-waiver Scholarship scheme from the University of Huddersfield. I would like to give my grateful acknowledgements to the University of Huddersfield and CEPE group for their support during the research.

I would like firstly to thank my main supervisor, Professor Andrew Ball, for his guidance and support throughout this research. It is his inspiration which encourages me to cope with all the faced challenges and gain the rewarding research skills.

I am sincerely grateful to my co-supervisor, Dr Fengshou Gu, Principle Research Fellow at the University of Huddersfield, who introduced me to the mechanical engineering. His invaluable guidance and warm-hearted help both in research and life support and encourage me a lot.

I would like to thank all the members of the Centre for Efficiency and Performance Engineering (CEPE) research group, especially Ibrahim Rehab and Gaballa M. Abdalla, for their help and discussion during the experiment and research.

I also appreciate the help from all the friends I have made in Huddersfield. They give me a lot of help and advice in my life and research.

Finally, special thanks extend to my parents and families for their understanding, support, and encouragement.

PUBLICATIONS

2016

1. **Tian, Xiange**, Gu, Xi James, Rehab, Ibrahim, Abdalla, Gaballa, Gu, Fengshou, Ball, Andrew (2016) 'A robust detector for rolling element bearing condition monitoring based on the modulation signal bispectrum and its performance evaluation against the kurtogram'. Submitted to Mechanical Systems and Signal Processing (MSSP), Under review.
2. Rehab, Ibrahim, **Tian, Xiange**, Hu, Niaoqing, Yan, T., Zhang, Ruiliang, Gu, Fengshou and Ball, Andrew (2016) 'A study of two bispectral features from envelope signals for bearing fault diagnosis'. In: IncoME 2016, 30th - 31st August 2016, Manchester Conference Centre, Manchester.
3. Rehab, Ibrahim, **Tian, Xiange**, Zhang, Ruiliang, Gu, Fengshou and Ball, Andrew (2016) 'A Study Of The Diagnostic Amplitude Of Rolling Bearing Under Increasing Radial Clearance Using Modulation Signal Bispectrum'. In: COMADEM 2016, the 29th International Congress on Condition Monitoring and Diagnostic Engineering Management, 20th-22nd August 2016, Empark Grand Hotel in Xi'an, China.
4. Ball, Andrew, Wang, Tie, **Tian, Xiange** and Gu, Fengshou (2016) 'A robust detector for rolling element bearing condition monitoring based on the modulation signal bispectrum'. In: COMADEM 2016, the 29th International Congress on Condition Monitoring and Diagnostic Engineering Management, 20th-22nd August 2016, Empark Grand Hotel in Xi'an, China.

2015

1. **Tian, Xiange**, Gu, Fengshou, Rehab, Ibrahim, Abdalla, Gaballa and Ball, Andrew (2015) An MSB based robust detector for bearing condition monitoring. In: Proceedings of the COMADEM 2015 + X CORENDE, Buenos Aires, Argentina, 1-4 December 2015, ISBN 978-987-23957-6-6.

2. **Tian, Xiange**, Abdalla, Gaballa, Rehab, Ibrahim, Wang, Tie, Gu, Fengshou and Ball, Andrew (2015) Diagnosis of Combination Faults in a Planetary Gearbox using a Modulation Signal Bispectrum based Sideband Estimator. In: Proceedings of the 21st International Conference on Automation & Computing, University of Strathclyde, Glasgow, UK, 11-12 September 2015.
3. Rehab, Ibrahim, **Tian, Xiange**, Gu, Fengshou and Ball, Andrew (2015) 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 & Computing, University of Strathclyde, Glasgow, UK, 11-12 September 2015.
4. Feng, Guojin, Zhen, Dong, **Tian, Xiange**, Gu, Fengshou and Ball, Andrew (2015) A Novel Method to Improve the Resolution of Envelope Spectrum for Bearing Fault Diagnosis Based on a Wireless Sensor Node. In: Vibration Engineering and Technology of Machinery. Springer, pp. 765-775. ISBN 978-3-319-09918-7

2014

1. **Tian, Xiange**, Feng, Guojin, Chen, Zhi, Albraik, Abdulrahman, Gu, Fengshou and Ball, Andrew (2014) The investigation of motor current signals from a centrifugal pump for fault diagnosis. In: Comadem 2014, 16th - 18th September 2014, Brisbane, Australia.
2. Feng, Guojin, **Tian, Xiange**, Gu, James Xi, Yang, Dingxin, Gu, Fengshou and Ball, Andrew (2014) An adaptive envelope analysis in a wireless sensor network for bearing fault diagnosis using fast kurtogram algorithm. In: Comadem 2014, 16th - 18th September 2014, Brisbane, Australia.
3. Abdalla, Gaballa, **Tian, Xiange**, Zhen, Dong, Gu, Fengshou and Ball, Andrew (2014) Misalignment diagnosis of a planetary gearbox based on vibration analysis. In: 21st International Congress on Sound and Vibration, 13th - 17th July, 2014, Beijing, China.
4. Hamomd, Osama, **Tian, Xiange**, Chen, Zhi, Abdulrahman, Albraik, Gu, Fengshou and Ball, Andrew (2014) A new method of vibration analysis for the diagnosis of impeller in a centrifugal pump. In: 21st International Congress on Sound and Vibration, 13th - 17th July, 2014, Beijing, China.

5. Gu, Fengshou, **Tian, Xiange**, Chen, Zhi, Rehab, Ibrahim and Ball, Andrew (2014) Fault severity diagnosis of rolling element bearings based on kurtogram and envelope analysis. International journal of structural analysis & design, 1 (2). ISSN 2372 – 4102

6. Rehab, Ibrahim, **Tian, Xiange**, Gu, Fengshou and Ball, Andrew (2014) 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, 10th - 12th June 2014, Manchester, UK

7. Gu, Fengshou, Wang, Tie, Alwodai, Ahmed, **Tian, Xiange**, Shao, Yimin and Ball, Andrew (2014) A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals. Mechanical Systems and Signal Processing. ISSN 0888-3270

2013

1. **Tian, Xiange**, Wang, Tie, Chen, Zhi, Gu, Fengshou and Ball, Andrew (2013) A transient enhancement method for two-stage helical gearbox fault diagnosis based on ALE. Journal of Signal and Information Processing, 4 (3B). pp. 132-137. ISSN 2159-4481

2. Rehab, Ibrahim, **Tian, Xiange**, Gu, Fengshou and Ball, Andrew (2013) Roller element bearing fault detection and diagnosis based on an optimised envelope analysis. In: Proceedings of Computing and Engineering Annual Researchers' Conference 2013 : CEARC'13. University of Huddersfield, Huddersfield, pp. 176-181. ISBN 9781862181212

3. Elbarghathi, Fathalla, **Tian, Xiange**, Tung Tran, Van, Gu, Fengshou and Ball, Andrew (2013) Multi-stages helical gearbox fault detection using vibration signal and Morlet wavelet transform adapted by information entropy difference. In: COMADEM 2013, 11-13 June 2013, Helsinki, Finland.

4. Wang, Ruichen, Cattley, Robert, **Tian, Xiange**, Gu, Fengshou and Ball, Andrew (2013) A Valid Model of a Regenerative Hybrid Shock Absorber System. In: Proceedings of Computing and Engineering Annual Researchers' Conference 2013: CEARC'13. University of Huddersfield, Huddersfield, pp. 206-211. ISBN 9781862181212

2012

1. **Tian, Xiange**, Gu, Fengshou, Zhen, Dong, Tran, Van Tung and Ball, Andrew (2012) A study on transient enhancement for fault diagnosis based on an active noise control system. In: CM 2012 and MFPT 2012: The Ninth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 12th - 14th June 2012, London, UK.
2. Lin, Jinshan, Chen, Qian, **Tian, Xiange** and Gu, Fengshou (2012) Fault Diagnosis of Rolling Bearings using Multifractal Detrended Fluctuation Analysis and Mahalanobis Distance Criterion. In: Proceedings of the 18th International Conference on Automation and Computing (ICAC) 2012: Integration of Design and Engineering. IEEE, Loughborough, UK, pp. 1-6. ISBN 9781467317221

Chapter 1
Introduction

This chapter provides an introduction to the research background in association with the work presented in this thesis. It describes the research background, introduces common techniques of condition monitoring and presents the motivation for pursuing this research. In addition, it provides the aims and objectives of the research. Finally, it outlines the organisation of the thesis.

1.1 Background of condition monitoring

With the rapid development of modern industry, science and technology, industrial machinery equipment and plants are increasingly becoming large-scale, high-speed, integrated and automated. Due to unforeseen factors, a wide variety of faults and failures could arise in the equipment, resulting in performance decreases, economic losses and even catastrophic accidents. Condition monitoring allows for the early detection of such costly faults and failures, and identifies the condition and nature of faults quickly. Corresponding actions can then be implemented to prolong the machine operation life [1]. Therefore, the research of condition monitoring techniques can help ensure the equipment operation's reliability, obtain greater economic benefit, and detect any abnormal status promptly.

The condition monitoring of a machinery component or parts is achieved by collecting a variety of information (such as vibration, acoustics, pressure, temperature, oil sample, etc.) to achieve the physical and statistical parameters (known as features for brevity) that can reflect the operating status of the equipment, hence allowing accurate decision-making for a timely and effective maintenance [2]. The information is collected by sensors, which could transfer the information into electrical or other physical signals, then applying signal processing to obtain the feature parameters. The main purpose of machine condition monitoring is to get accurate and quantitative information regarding the machinery's current working condition and determine whether the operating status is normal or faulty [3].

The procedure of condition monitoring normally consists of three steps: data acquisition, data processing and decision making [4]. As the middle step, data processing is probably more critical to the success of condition monitoring, which will be discussed in more detail in Section 1.3, as appropriate data processing techniques allow effective noise reduction and correct feature extraction that ensures the decision obtained is reliable and accurate.

1.2 Vibration based condition monitoring

Nowadays, many condition monitoring techniques have been developed, such as vibration analysis [5][6][7], airborne acoustics analysis [8][9], lubricant analysis [10], acoustic emission [7][11][12], temperature [13][14], motor current signal analysis [15][16] etc. Amongst these techniques, vibration analysis is the most popular approach in machine fault diagnosis, because vibration responses are non-destructive, easy to measure, and

sensitive to a wide range of defects. These merits make it very promising as a comprehensive and cost-effective method for condition monitoring applications.

In condition monitoring practice, vibration signals are usually acquired by accelerometers which are non-intrusively mounted on the external surfaces of a machine. The sensor installation encounters rare problems in industry. The signals are amplified and analysed to reveal information about the health of the machine. Although the measuring instruments are of wide frequency responses, high dynamic ranges, high temperature capability and hence high reliability, they are becoming inexpensive due to the rapid development of electronics and manufacturing efficiency.

These signals are analysed using common signal processing techniques such as peak values, root mean square (RMS) values, crest factor, frequency spectrum, cepstrum, wavelet and higher order spectral analysis [17]. Vibration analysis monitors the dynamic response of the machine or equipment to the different excitation forces which are applied to it. These forces exist in any machine and are produced by various physical mechanisms ranging from inevitable friction and wear to repetitive impacts. The response is analysed to reveal information regarding the health of the machine. Vibration analysis has traditionally been recognised as one of the most powerful condition based maintenance technologies, and the keystone of many predictive maintenance programs, since all rotating machines produce vibration in some form. The technique is therefore generally employed for the trouble-shooting and fault diagnosis of rotating machinery and equipment [17]-[20].

The shafts, gears and bearings in a gearbox will generate vibrations during its operation. Vibration signals carry the fault information of a gearbox; if there is any fault occurring in the energy distribution, the vibration signal will change accordingly. Because vibration based fault diagnosis is reliable, cheap, and easy to operate, it has been widely applied in the field of condition monitoring and fault diagnosis.

Defects in gearboxes or bearings (such as surface pitting, broken teeth, gently curved shafts, joint looseness and fatigue spalling in bearings) will produce repetitive impacts, resulting in a modulation phenomenon in vibration signals. In the frequency spectrum, sidebands will appear near the resonance frequency or gear mesh frequencies. Therefore, demodulation is an important process in the signal processing of the gearbox and bearing condition monitoring and fault diagnosis. The commonly applied techniques include spectrum analysis, time synchronous average, Hilbert transform, resonance demodulation method, and envelope analysis [21][22].

In the research areas of failure mechanism and feature extraction of gearboxes, many problems are involved due to the complication in machine structures, harsh working environments, and many different types of interferences. Therefore, this thesis focuses on developing more advanced fault diagnosis methods based on vibration analysis.

1.3 Signal processing techniques for vibration signals

In applications, the vibration signals acquired suffer interference from various noises. As a result, useful information is submerged intensely. Therefore, it is critical to reduce noise and enhance the signal-to-noise ratio (SNR) for more reliable feature extraction. To this end, vibration signal processing techniques have attracted massive attention in the condition monitoring and fault diagnosis research areas, leading to many useful techniques. Based on fundamental signal processing techniques, they may be examined through the following categories:

1.3.1 Time domain analysis

Time domain analysis is directly based on the time series itself. Statistical parameters are the simplest method of feature extractions. There are several popular time domain analysis approaches for the vibration signal analysis of rotating machinery, which have the effect of denoise, such as correlation analysis, adaptive filters and TSA.

1.3.1.1 Statistical parameters

Statistical parameters are indices calculated from the time domain waveforms of vibration signals. They can be divided into dimensional parameters (such as minimum value, maximum value, mean value, variance, standard deviation, root mean square, etc.) and non-dimensional parameters (such as kurtosis, shape factor, crest factor. impulse factor, peak value, etc.). Dimensional parameters are related more to the operating conditions, such as the transmission ratio of the gearbox and the rotational speed of the shaft. Kurtosis and crest factor are sensitive to strong impulses and can be used for abrupt faults diagnosis.

1.3.1.2 Correlation analysis

Correlation analysis includes auto-correlation and cross-correlation analysis.

Autocorrelation is widely used to find periodicity in a noisy signal. The autocorrelation function can be applied for random noise reduction and feature extraction [23]. The autocorrelation enhancement is defined as:

$$R_{xx}(\tau) = F^{-1}\{F_{xx}(\mathbf{x})\} \quad (1.1)$$

where:

$$F_{xx}(\mathbf{x}) = \left|\hat{X}(f)\right|^2 \quad \text{.....(1.2)}$$

$\hat{X}(f)$ is the power spectrum of a signal $x(t)$, F^{-1} denotes the operation of the inverse transform of the Fourier transform.

One typical application is the cyclic autocorrelation analysis for emphasising the correlation between components of weaker energy compared to those of greater energy [24].

The cross-correlation is a measure of similarity between two signals. If both signals contain a periodic component and their period are equal, the cross-correlation function contains the same periodic component. This feature is important in the applications of signal detection and feature extraction.

1.3.1.3 Adaptive filters

The application of adaptive filters for extracting a fault signal from background noise is based on the assumption that the frequency components of the fault signal should be different from the noise. For example, the background noise is continuous while the fault is impulsive. The impulsive behaviour implies that the frequency components will be spread out over many frequency bins due to its impulsive temporal characteristics [28].

On the other hand, adaptive filters have good noise attenuation performance for periodic background noise, whose frequency is low but the amplitude is much higher than the impulsive component. The adaptive filter is capable of tracking the statistics of nonstationary signals if changes in the statistics occurred slowly in comparison with the convergence time of the adaptive filter. However, the impulsive content in the signal is of a short duration and cannot provide sufficient time for algorithm convergence. Therefore, the impulsive component that contains faulty characteristics can be extracted from background noise.

1.3.1.4 Time synchronous average

In experimental vibration signals, there are high levels of random noises which will affect the accuracy of fault diagnosis, especially random impulses which can confuse the periodicity of fault impulses. In such studies, the time synchronous average (TSA) technique can be applied to suppress the random noises. TSA resamples the vibration data synchronously according to the angle of rotation.

When taken over many machine cycles, this technique removes background noise and nonsynchronous events (such as impulses and meshing components from other gear transmission stages) in the vibration signal. It is extremely useful when monitoring multiple shafts which are operating at only slightly different speeds and in close proximity to one another. For the implementation of TSA, a reference signal (usually from a tachometer) is always needed. However, it can be inefficient at reducing some tonal components of the background noise, i.e. those that are commensurate (or nearly commensurate) with the rotational frequency.

1.3.2 Frequency domain analysis

Frequency domain analysis is based on the Fourier transform and commonly applied in rotation machinery fault diagnosis. In the frequency domain, it is easy to get the shaft rotational frequencies, gear mesh frequencies, bearing fault frequencies and their harmonics. Frequency analysis can indicate the fault location in gearboxes and bearings and further analyse the causes of failures. Frequency analysis mainly includes spectral analysis, cepstrum analysis and demodulation spectral analysis.

1.3.2.1 Spectrum analysis

Spectrum analysis is the most commonly used method for machinery condition monitoring and fault diagnosis. It describes the distribution of the original signal in the frequency domain and provides feature information more straightforwardly than in the time domain. Spectrum analysis includes the amplitude spectrum and the power spectrum. The amplitude spectrum represents the amplitude corresponding to each frequency component while the power spectrum shows the distribution of the power.

1.3.2.2 Cepstrum analysis

"*Cepstrum is defined as the power spectrum of the logarithm of the power spectrum.*"[30] It is an important signal processing technique for detecting the periodic components in a complex spectrum in recent decades. By applying the cepstrum analysis to the signal power spectrum, it is capable of identifying the composition of the signal, enhancing the periodic components, and converting the spectral lines in the original spectrum into a single line in cepstrum. Thus, complex periodic components become legible for machinery condition monitoring and fault diagnosis [31]. When a defect appears on bearings or gears, there are multiple clusters of modulation sidebands in the vibration spectra. Then, it is easy to identify the fault frequency and defect location using cepstrum analysis, but it does not have the capability of denoising.

1.3.2.3 Envelope analysis

Envelope analysis is also called demodulation analysis, which can be used to extract the low-frequency signal that modulated on the high-frequency carrier signal. For rolling element bearings, when the rolling elements strike a local fault on the inner or outer race, or a fault on a rolling element strikes the inner or outer race, an impact is produced. These impacts modulate a signal at the associated bearing pass frequencies and then stimulate the resonance vibration of the machine. Envelope analysis can obtain the fault feature frequency from the raw data. The envelope of the analytic signal is calculated by [32]:

$$a(t) = \sqrt{x^2(t) + H^2[x(t)]} \quad \text{.....(1.3)}$$

where $H[x(t)]$ denotes the Hilbert transform of signal $x(t)$.

1.3.3 Time-frequency domain analysis

In signal processing, time-frequency analysis [33] is used for characterising and manipulating signals whose statistics vary in time, such as impulsive components. It is a generalisation and refinement of Fourier analysis, for the case when the signal frequency characteristics are varying with time. Since many signals of interest (such as speech, music, images, and medical signals) have changing frequency characteristics, time-frequency analysis has a broad scope of applications.

1.3.3.1 Short-time Fourier transform

Traditional time-frequency analysis represents the energy or power of signals over both time and frequency jointly, in order to better reveal fault patterns for more accurate diagnostics. The Short-time Fourier transform (STFT) is the most popular time-frequency representation (TFR) as it can be obtained by applying FFT to a short segment signal segregated by a window moving incessantly along the time direction. By examining the frequency content of the signal as the time window moves, a 2D time-frequency distribution is generated. One well-known drawback of the STFT is a resolution limit imposed by the window function, which leads to worse frequency resolution.

The basic idea of STFT [34] is to slice up the signal into suitable overlapping time segments (using windowing methods) and then Fourier analyse each slice to ascertain the frequencies contained in it. It is assumed that frequency information is associated with the time index in the middle of each slice of windowed data. The STFT of a continuous-time signal $x(t)$ is defined as

$$STFT(t, f) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j2\pi f \tau} d\tau \quad \text{.....(1.4)}$$

where $w(t)$ is the window function whose position is translated in time by τ. There are some limitations associated with STFT, the first being the window length. It is obvious that a wide window yields a good resolution in the frequency domain, but the poor resolution in the time domain, and vice versa. So, in practical situations, a compromise between the two resolutions has to be made. Secondly, raw STFT is computationally expensive, but ways of accelerating it by avoiding redundant calculations are now available in the literature. These drawbacks notwithstanding, STFT is an ideal tool in many respects, the most important being its excellent spectrogram structure, which is consistent with our intuition regarding frequency spectra, thereby qualifying as a good visualisation tool.

1.3.3.2 Wigner–Ville distribution

Time-frequency analysis focuses on a distribution of the total energy of the signal at a particular time and frequency. The Wigner-Ville distribution (WVD) [35] is the first to be introduced and has been most widely studied. It was developed in quantum mechanics by Wigner and implemented for signal processing by Ville [36]. The WVD of a signal $x(t)$ is defined as

$$W_x(t, f) = \int_{-\infty}^{\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi f \tau} d\tau \quad \text{.....(1.5)}$$

where * is the conjugate operator and τ the time lag.

The Wigner–Ville spectrum is related to the expected value of WVD, and is introduced for the time–frequency analysis of nonstationary random signals

$$EW_x(t, f) = E\left\{ \int_{-\infty}^{\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi f \tau} d\tau \right\} = \int_{-\infty}^{\infty} R_x(t, \tau) e^{-j2\pi f \tau} d\tau \quad \text{.....(1.6)}$$

where $E\{\ \}$ is the mathematic expectation operator and $R_x(t,\tau)$ the time-varying autocorrelation:

$$R_x(t, \tau) = E\left\{ x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) \right\} \quad \text{.....(1.7)}$$

1.3.3.3 Wavelet analysis

The main purpose of the wavelet transform is to decompose arbitrary signals into localised contributions that can be labelled by a “scale parameter”.

The wavelet transform provides a good means of studying how the frequency contents change with time and consequently is able to detect and localise short-duration impulse components. Wavelet coefficients measure the similarity between the signal and each of its daughter wavelets. The more the daughter wavelet is similar to the feature component, the larger is the corresponding wavelet coefficient.

Wavelet transforms [37] are inner products between signals $x(t)$ and the wavelet family, which are derived from the mother wavelet by dilation and translation. Letting $\psi(t)$ be the mother wavelet, the daughter wavelet will be:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad \text{.....(1.8)}$$

where a is the scale parameter and b is the time translation. By varying the parameters a and b, different daughter wavelets are obtained to constitute a wavelet family. The wavelet transform of a signal $x(t)$ is defined as

$$WT_x(a,b) = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt \quad \text{.....(1.9)}$$

where $WT_x(a,b)$ represents the wavelet transform coefficients, and * denotes the conjugation operation. The factor $\frac{1}{\sqrt{a}}$ is used to ensure energy preservation.

1.3.4 Other signal processing methods

1.3.4.1 High order spectral analysis

Spectral kurtosis

Spectral kurtosis (SK) is a powerful tool for detecting the presence of impulses in a signal, even when they are buried in strong additive noise; this is done by indicating which frequency bands the impulses are taking place. The kurtogram optimisation considers a variety of bandwidths and central frequencies. It is basically a cascade of SK obtained for different values of the STFT window length.

The spectral kurtosis of a signal $x(t)$ may be computed from the STFT, $X(t, f)$, that is the local Fourier transform at the time t obtained by moving a window along the signal. When seen as a function of frequency, the squared magnitude $|X(t, f)|^2$—i.e. the spectrogram—returns the power spectrum at the time t and is a further average over time, $<|X(t, f)|^2>$,

the PSD as computed by the Welch method. When seen as a function of t, $X(t,f)$ may be interpreted as the complex envelope of signal $x(t)$ bandpass filtered around frequency f and its squared magnitude will then indicate how energy is flowing in that frequency with respect to time. If that frequency band happens to carry pulses, bursts of energy will then appear. This may be simply detected by computing the kurtosis of the complex envelope $X(t,f)$ as follows:

$$K(f)=\frac{\left\langle |X(t,f)|^4 \right\rangle}{\left\langle |X(t,f)|^2 \right\rangle^2}-2 \quad \text{.....(1.10)}$$

where the subtraction of 2 is used to enforce $K(f)=0$ in the case $X(t,f)$ is complex Gaussian (instead of 3 for real signals).

Conventional bispectrum

The conventional bispectrum is the third-order cumulant over frequencies. It is a function of two frequencies f_1 and f_2. The equation can be expressed as [38]:

$$B(f_1,f_2)=\frac{1}{M}\sum_{m=1}^{M} X(f_1)X(f_2)X^*(f_1+f_2) \quad \text{.....(1.11)}$$

It is calculated by averaging the triple products of Fourier coefficients over M segments. The bispectrum is capable of detecting the nonlinear interactions between frequency components at f_1, f_2 and f_1+f_2, which represent quadratic phase coupling. This will be discussed in detail in Chapter 2 and Section 3.6.1.

The bispectrum has several important properties [39], which are an important theoretical basis for random noise suppression in mechanical signal processing. These properties are detailed as follows:

(1) The bispectrum of a stationary, zero-mean Gaussian process is zero. Thus a non-zero bispectrum indicates a non-Gaussian process.

(2) The bispectrum suppresses linear phase information or constant phase shift information.

(3) The bispectrum is flat for non-Gaussian white noise and is zero for Gaussian white noise.

Modulation signal bispectrum

Modulation signal bispectrum is developed based on the conventional bispectrum. It is suitable for detecting the nonlinear interactions between the carrier and two sidebands of an amplitude modulation signal. The mathematical expression can be written as [40]:

$$B_{MS}(f_1, f_2) = \frac{1}{M} \sum_{m=1}^{M} X(f_2 + f_1) X(f_2 - f_1) X^*(f_2) X^*(f_2) \quad \text{.....(1.12)}$$

where f_2 is the carrier frequency; $f_2 + f_1$ and $f_2 - f_1$ are the two sidebands of the modulation signal. More details and applications of this method will be introduced in Chapter 2 and Section 3.6.2, in which novel refinements are suggested to enhance the signal of interest.

1.3.4.2 Signal decomposition methods

Empirical Mode Decomposition

EMD is one of the most powerful time-frequency analysis techniques. It is based on the local characteristic time scales of a signal and can decompose the signal into a set of complete and almost orthogonal components called intrinsic mode function (IMF). The IMFs indicate the natural oscillatory mode embedded in the signal and serve as the basis functions, which are determined by the signal itself, rather than pre-determined kernels. Thus, it is a self-adaptive signal processing technique that is suitable for nonlinear and nonstationary processes [41].

Local mean decomposition

Local mean decomposition (LMD) is a data-driven and novel self-adaptive analysis method in the time-frequency domain. It was proposed by Smith in 2005 and firstly applied to electroencephalogram signal successfully [42]. The multi-component signal can be decomposed into a series of mono-components which are product functions (PFs); therefore, each of them is the product of an envelope signal and a purely frequency modulated signal. The instantaneous amplitude of PF can come from an envelope signal, and the well-defined instantaneous frequency can be calculated from a purely frequency modulated signal. In essence, each PF is an amplitude-modulated and frequency-modulated signal (AM-FM signal) [43].

Local characteristic-scale decomposition

Local characteristic-scale decomposition (LCD) was proposed by J. Cheng in 2012. LCD can decompose any complicated signal into a number of intrinsic scale components (ISC), whose instantaneous frequencies have physical meaning. It is especially suitable for processing a multi-component amplitude-modulated and frequency-modulated signal. Since the gear fault vibration signal is a multi-component amplitude-modulated and frequency-modulated signal, the local characteristic-scale decomposition method is especially applicable for gear fault diagnosis [44].

1.4 Motivation of the research

Gearboxes and bearings are fundamental mechanical parts and widely used in various machines such as wind turbines, helicopters and marine propulsions. The reliability of their high performance operation is a critical issue in industry, because the occurrence of unexpected failures may lead to catastrophic accidents and cause severe economic losses. Therefore, it is necessary to conduct condition monitoring and fault diagnosis for gearboxes and bearings in order to reduce equipment operational cost and risk [45].

Defects in gears or bearings, such as surface pitting, broken teeth, gently curved shaft, joint looseness and fatigue spalling, will produce repetitive impacts during rotary motion. During the signal transmission from the source to the transducer, the amplitude modulation will be induced by the resonance frequency of the system, rotation of the shaft, load effect, and dynamic meshing forces. Therefore, the fault features for gears and bearing have some similar characteristics, such as impulsive, periodic, modulation.

In practical applications, some mechanical equipment defects often induce other defects rather than occur alone, resulting in the occurrence of combined faults. A combined fault is composed of two or more defects simultaneously or induced by one defect, which are interrelated and cross-impact. The combined fault could be composed by defects on different components or having different severities on mechanical equipment. Moreover, there are mutual interference and intercoupling between different defects, which makes the signal more complicated. Therefore, combined fault diagnosis is a challenging problem in the field of machinery fault diagnosis. The vibration signals of combined fault in rotating machines also exhibit impulsive, periodic and modulation characteristics, which are similar with gearboxes and bearings. Therefore, the combined fault diagnosis has been studied

based on experimental data. Before the study combined fault diagnosis, several signal processing techniques have been proposed for bearing and gear fault diagnosis separately.

In the signal processing of practical signals, noise is inevitably induced by the data collection process, environmental disturbances, structural distortions and interferences from other machines and related components. Although the To implement reliable fault detection and diagnosis for gears and bearings, this research focuses on the denoising and feature extraction from vibration signals utilising the impulsive, periodic and modulation characteristics of fault mechanisms.

1.5 Aims and objectives of the research

The aim of this research is to develop efficient signal processing techniques capable of enhancing vibration signatures for higher performances of detecting and diagnosing common faults in gears and bearings. Based on a rigorous study of previous investigations, it has regarded largely that the advanced signal processing techniques for condition monitoring are often carried out by either a direct enhancement of signatures or a noise reduction based enhancement. Although each of these two approaches achieves satisfactory performance improvements, they often have the deficiency in signature enhancement and implementations. Moreover, they are rarely combined to jointly achieve the signature enhancement. Therefore, the aim of this research is fulfilled by overcoming these deficiencies, which are then carried out by completing the following refined key objectives:

1. Review the commonly used vibration signal processing techniques for machinery condition monitoring and fault diagnosis.
2. Investigate the vibration signal models of gear defects, bearing defects and the combined faults to study the fault characteristics and provide the primary bases for signal processing.
3. Develop optimal impulse enhancement methods based on ALE and wavelet analysis techniques to improve the SNR of vibration signals for gearbox fault diagnosis.
4. Investigate the kurtogram analysis and develop a filter for obtaining an optimal envelope that enhances the impulsive components due to bearing defects.
5. Develop effective approaches and algorithms for planetary gear fault detection and diagnosis based on MSB analysis to utilise the modulating characteristics of signals.

6. Propose a signal processing approach for diagnosing the combined fault on both the gear and bearing through MSB analysis based on the multiple and asymmetric modulating characteristics.

1.6 Organisation of the thesis

In this section, an outline is given of the contents and relative emphasis of the following chapters for the thesis.

Chapter 2 studies the vibration signal models of the bearing defect, gear defect and combined faults on both the bearing and gear, which provides theoretical foundations for developing signal processing algorithms.

Chapter 3 overviews six signal processing techniques for machinery condition monitoring and fault diagnosis, including ANC, wavelet analysis, kurtogram, conventional bispectrum, MSB, and TSA.

Chapter 4 describes an impulse enhancement method for a two-stage helical gearbox fault diagnosis based on TSA and ALE.

Chapter 5 develops a gear fault diagnostic approach for a two-stage helical gearbox fault diagnosis based on adaptive wavelet analysis.

Chapter 6 details a bearing fault diagnostic method based on kurtogram and envelope analysis.

Chapter 7 develops a robust detector for rolling element bearing condition monitoring based on the MSB and compares its performance evaluation against that of the kurtogram.

Chapter 8 describes a residual sidebands based fault diagnostic approach of a planetary gearbox using MSB analysis.

Chapter 9 proposes a combined fault diagnostic method of a planetary gearbox based on MSB.

Chapter 10 draws conclusions for this research and gives suggestions for future work in related research areas.

Chapter 2
Vibration signal models of bearings and gears

Signal characteristics are of primary concern in selecting effective signal processing techniques. To attain diagnostic features accurately for vibration based monitoring, this chapter examines the vibration signal models of roller bearings and helical gears, which are the most critical and popular components of rotating machines. It starts with presenting common vibration models. Then, two analytical signal models are investigated for bearing faults and helical gear faults, respectively. Finally, a combined signal model for both the bearing and gear faults is developed based on a two-stage helical gearbox model. This chapter provides fundamentals for the follow-up studies on developing signal processing for diagnosing faults on bearings and gears.

2.1 Introduction

To gain an understanding of vibration mechanisms and signal characteristics, vibration models of roller bearings and helical gears are investigated with more attention to local defect induced responses.

As the gear kinetic is simpler compared with that of rolling bearings which have planetary rotations, its dynamic forces can be understood easily. So this chapter starts with investigating gear signal models based on the dynamics of gear transmission. Then, it examines the models for bearings. These then make it easy to understand the vibrations of a planetary gearbox.

In addition to the fact that bearings and gears are used together, both of them can share similar vibration responses to local defects in that they produce local impulsive forces. Dynamically, both of them exhibit a certain degree of nonlinearity due to time-varying contact deformation. Therefore, they are examined jointly in this study.

2.2 Gear fault model based on dynamic analysis

2.2.1 One-stage gearbox vibration model

Commonly, the vibration system of a gear set can be represented as shown in Figure 2.1 [46]. It consists of both rotational motions and translational motions of the pinion (drive) and the gear. According to Newton's law, the vibration governing equations can be established as follows:

the translational motion of the pinion, denoted with a subscript p:

$$m_p\ddot{y}_p + k_{pb}y_p - k_{pg}\left(r_p\varphi_p - r_g\varphi_g - y_p + y_g\right) - c_{pg}\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \quad \text{.....(2.1)}$$

the rotational motion of the pinion p:

$$I_p\ddot{\varphi}_p + r_p k_{pg}\left(r_p\varphi_g - r_g\varphi_g - y_p + y_g\right) + r_p c_{pg}\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \quad \text{.....(2.2)}$$

the translational motion of the driven gear denoted with a subscript g:

$$m_g\ddot{y}_g + k_{pb}y_g - k_{pg}\left(r_p\varphi_p - r_g\varphi_g - y_p + y_g\right) - c_{pg}\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \quad \text{.....(2.3)}$$

the rotational motion of the driven gear g:

$$I_g\ddot{\varphi}_g - r_p k_{pg}\left(r_p\varphi_g - r_g\varphi_g - y_p + y_g\right) - r_p c_{pg}\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \quad \text{.....(2.4)}$$

where y_p and y_g are the vertical displacement of the pinion and gear; I_p and I_g indicate the moment of inertia of the pinion and gear; m_p and m_g are the mass of the pinion and gear; φ_p and φ_g are the angular displacement of the pinion and gear; r_p and r_g denote the base circle radius of the pinion and gear; k_{pb} and k_{gb} are the stiffness of the pinion and gear; k_{pg} and c_{pg} denote the meshing stiffness and damping.

These equations are typical nonlinear parametrically excited vibrations and difficult to find closed form solutions that can be based on for signal processing.

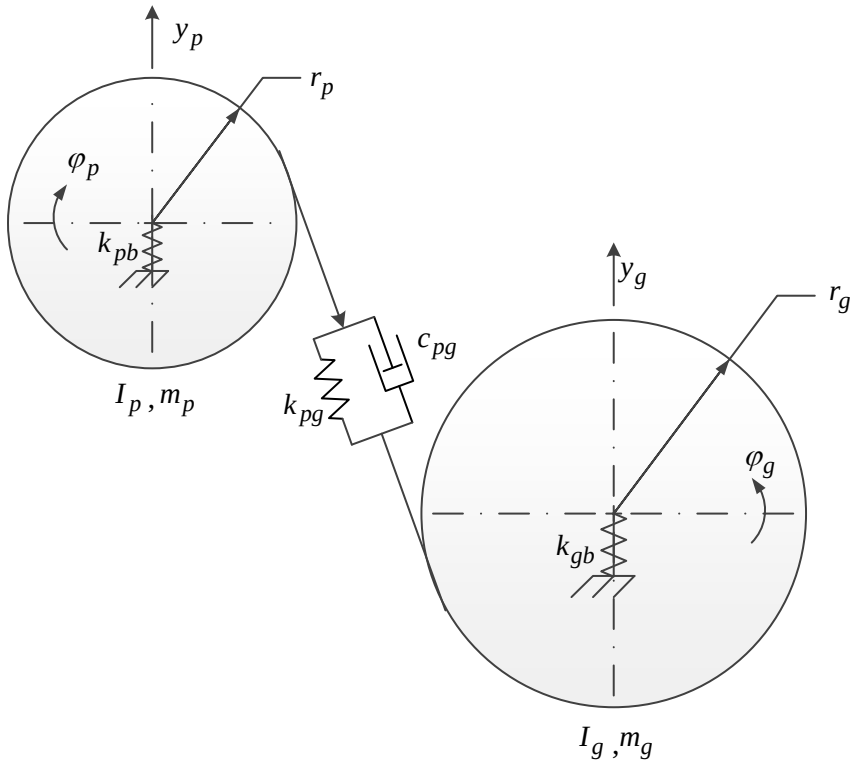

Figure 2.1 Gear transmission model with rotational and translation motion coupling

However, as shown in Equations (2.2) and (2.4), vibration responses in vertical and horizontal directions are mainly due to the tooth stiffness k_{pg}, which is not only time-varying but also coupled with both translational and rotational motions: y_p, y_g, φ_p and φ_g, as illustrated by Figure 2.2. This means that translational motions which usually are measurable on the gearbox case for condition motoring are resulted from a multiple coupling between the translation and the rotational motions. In particular, by representing the time-varying stiffness as $k = \bar{k} + k(t)$, the vibration responses can be represented in the

form of multiple modulations [47], which can be also derived through the harmonics balance method (HBM):

$$y(t)=\sum_{m=0}^{M} I_m[1+a_m(t)]\cos[2\pi f_{me}t+\phi_m+b_m(t)] \quad \text{.....(2.5)}$$

where $y(t)$ is the vibration response, $a_m(t)$ is the amplitude modulation and $b_m(t)$ is the phase modulation due to the local tooth faults. The model shows that vibration signals have the content of both amplitude modulation (AM) and phase modulation (PM) as a consequence of the combined outcomes of both tooth defects such as tooth breakage, tooth wear and manufacturing errors.

Obviously, this model is useful in developing signal processing methods. However, it can be difficult to have effective and stable vibration features for obtaining a reliable diagnostic result due to the multiple modulation and noise effects.

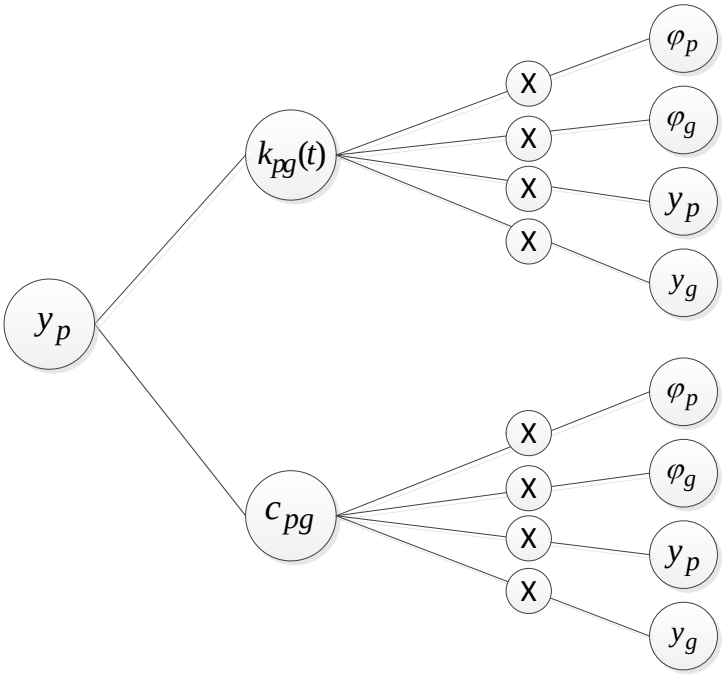

Figure 2.2 Inter-coupling between different motions

2.2.2 Simulated one-stage gearbox vibration signals

The faults localised on one tooth or a few teeth such as cracks and spalls produce modulation effects during the engagement of the fault teeth. Consequently, a large number of sidebands of the tooth-mesh frequency and its harmonics in the spectrum is generated and spread over a wide range, which is spaced by the rotation frequency of the fault gear and characterised by low amplitudes [48].

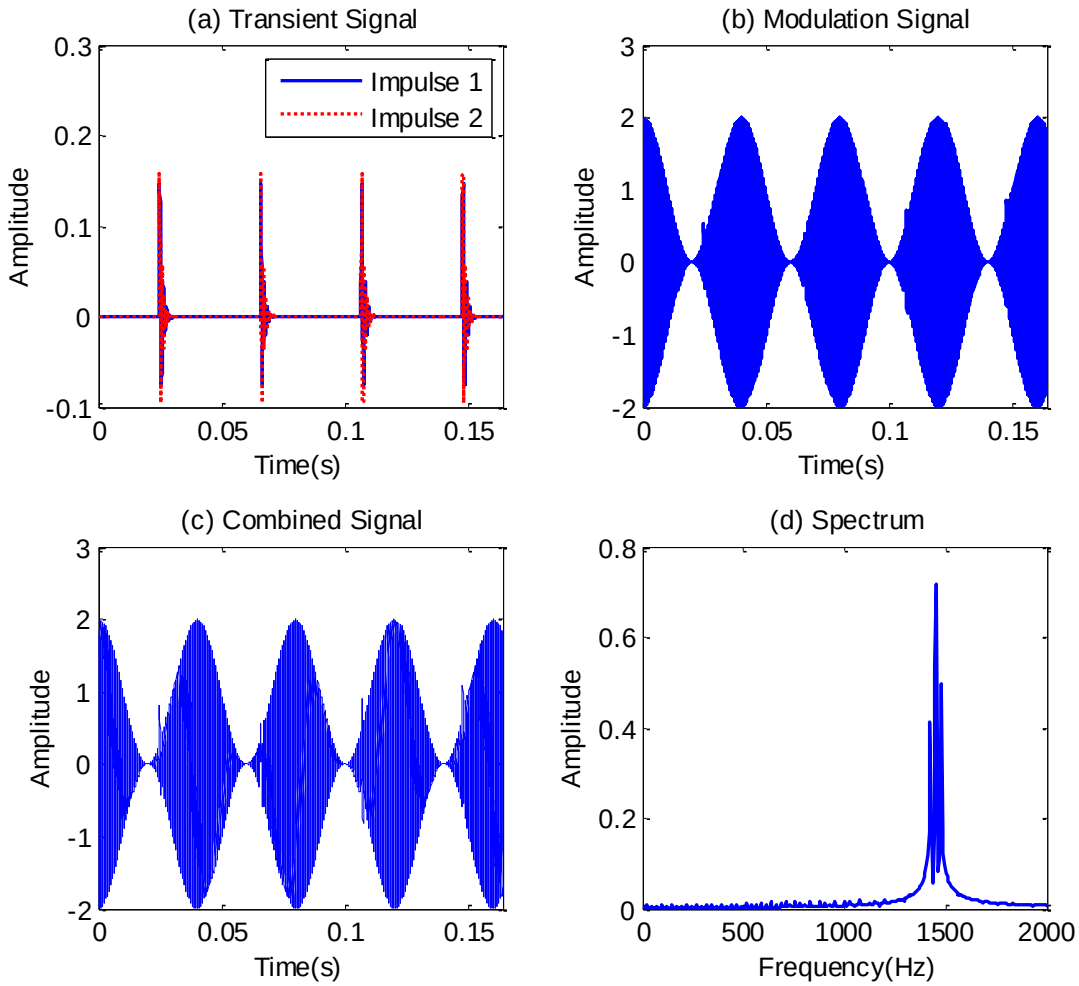

Figure 2.3 Simulated signals and their spectra

A simulated signal is synthesised based on the characteristics of gear vibration signals [49] according to Equation (2.6).

$$s(t) = s_1(t) + s_2(t) + s_3(t) \quad \text{.....(2.6)}$$

where $s_1(t)$ is the modulation component that simulates the modulation between gear mesh frequency $f_{m1} = Z1 * f_r = 1450Hz$ and shaft rotational frequency $f_r = 25Hz$, specifically expressed as:

$$s_1(t) = [1 + \cos(2\pi f_r t)]\cos(2\pi f_{m1} t) \quad \text{.....(2.7)}$$

where the gear tooth number $Z1 = 58$. As shown in Figure 2.3(b), it has very high amplitude, showing that it is the dominant part of vibration signals. Along with this signal there are also two small impulsive components $s_2(t)$ and $s_3(t)$, which are generated by Equations (2.8) and (2.9), respectively, to represent possible impulses from meshing stiffness impacts due to local tooth damages of different amplitudes [50].

$$s_2(t) = 0.2\sum_{n=0}^{3} h(t-nT-\tau_0)e^{-1000(t-nT-\tau_0)}\cos[(2\pi f_1(t-nT-\tau_0)) \quad \text{.....(2.8)}$$

$$s_3(t) = 0.2\sum_{n=0}^{3} h(t-nT-\tau_0)e^{-1000(t-nT-\tau_0)}\cos[(2\pi f_2(t-nT-\tau_0)) \quad \text{.....(2.9)}$$

where $f_1 = 30\,f_r = 750Hz$, $f_2 = 40\,f_r = 1000Hz$, T is the period of shaft rotation, τ_0 is the start time of the impulse and $h(t)$ denotes the step function. As shown in Figure 2.3(a), these components usually have very small amplitudes when the tooth damage is incipient but show clear spikes according to their periods.

However, when these two portions are added together, the repetitive spikes are masked by the large modulation component and are difficult to detect either in the waveform of Figure 2.3(c) or in its spectrum of Figure 2.3(d).

2.2.3 Two-stage gearbox vibration model

Based on modal response mechanisms, the gearbox case responses were also considered as two lumped systems, representing two arbitrary case positions. Thereby it allows the model to be evaluated using measured vibrations at the gearbox case [51].

Following the same procedure used in developing the model for the one-stage gearbox, the vibration responses of a two stage gearbox system can be obtained as:

The rotational motion of the motor rotor can be expressed as:

$$I_m\ddot{\varphi}_1 + k_1(\varphi_1-\varphi_2) + c_1(\dot{\varphi}_1-\dot{\varphi}_2) = M_m \quad \text{.....(2.10)}$$

The rotational motion of the drive gear (Z_1) in first stage is as:

$$\begin{aligned} &I_{p1}\ddot{\varphi}_2 - k_1(\varphi_1-\varphi_2) - c_1(\dot{\varphi}_1-\dot{\varphi}_2) + r_{p1}k_{z1}(r_{p1}\varphi_2 - r_{g1}\varphi_3 - y_{p1} + y_{g1}) \\ &+r_{p1}c_{z1}(r_{p1}\dot{\varphi}_2 - r_{g1}\dot{\varphi}_3 - \dot{y}_{p1} + \dot{y}_{g1}) = 0 \end{aligned} \quad \text{.....(2.11)}$$

The rotational motion of the driven gear (Z_2) in the first stage is given by:

$$\begin{aligned} &I_{g1}\ddot{\varphi}_3 + k_2(\varphi_3-\varphi_4) + c_2(\dot{\varphi}_3-\dot{\varphi}_4) - r_{g1}k_{z1}(r_{p1}\varphi_2 - r_{g1}\varphi_3 - y_{p1} + y_{g1}) \\ &-r_{g1}c_{z1}(r_{p1}\dot{\varphi}_2 - r_{g1}\dot{\varphi}_3 - \dot{y}_{p1} + \dot{y}_{g1}) = 0 \end{aligned} \quad \text{.....(2.12)}$$

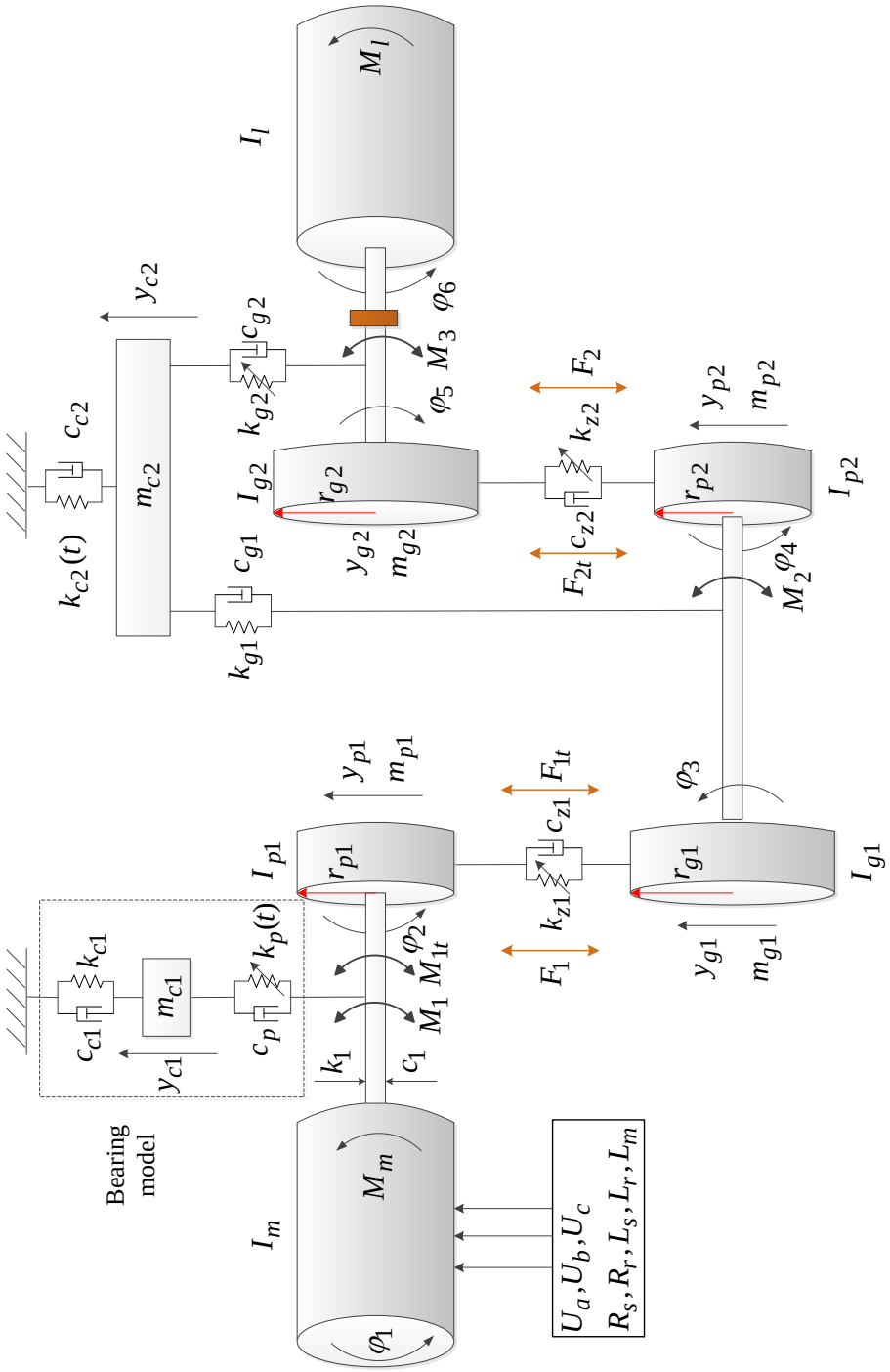

Figure 2.4 The dynamic model of the motor-gearbox and loader system

The rotation of the drive gear (Z_3) in the second stage is given by:

$$I_{p2}\ddot{\varphi}_4 - k_2(\varphi_3 - \varphi_4) - c_2(\dot{\varphi}_3 - \dot{\varphi}_4) + r_{p2}k_{z2}(r_{p2}\varphi_4 - r_{g2}\varphi_5 - y_{p2} + y_{g2}) + r_{p2}c_{z2}(r_{p2}\dot{\varphi}_4 - r_{g2}\dot{\varphi}_5 - \dot{y}_{p2} + \dot{y}_{g2}) = 0 \quad \text{.....(2.13)}$$

The rotation of the driven gear (Z_4) in the second stage is given by:

$$I_{g2}\ddot{\varphi}_5 + k_3(\varphi_5 - \varphi_6) + c_3(\dot{\varphi}_5 - \dot{\varphi}_6) - r_{g2}k_{z2}(r_{p2}\varphi_4 - r_{g2}\varphi_5 - y_{p2} + y_{g2}) \\ -r_{g2}c_{z2}(r_{p2}\dot{\varphi}_4 - r_{g2}\dot{\varphi}_5 - \dot{y}_{p2} + \dot{y}_{g2}) = 0 \quad \text{.....(2.14)}$$

The rotation of the loading rotor (output shaft) is given by:

$$I_l\ddot{\varphi}_6 - k_3(\varphi_5 - \varphi_6) - c_3(\dot{\varphi}_5 - \dot{\varphi}_6) = M_l \quad \text{.....(2.15)}$$

The equation that governs the translational motion of the drive gear (Z_1) in the first gear pair is given by:

$$m_{p1}\ddot{y}_{p1} + k_{p1}(y_{p1} - y_{c1}) + c_{p1}(\dot{y}_{p1} - \dot{y}_{c1}) - k_{z1}(r_{p1}\varphi_2 - r_{g1}\varphi_3 - y_{p1} + y_{g1}) \\ -c_{z1}(r_{p1}\dot{\varphi}_2 - r_{g1}\dot{\varphi}_3 - \dot{y}_{p1} + \dot{y}_{g1}) = 0 \quad \text{.....(2.16)}$$

The translational motion of the driven gear (Z_2) in the first stage is given by:

$$m_{g1}\ddot{y}_{g1} + k_{g1}(y_{g1} - y_{c2}) + c_{g1}(\dot{y}_{g1} - \dot{y}_{c2}) - k_{z1}(r_{p1}\varphi_2 - r_{g1}\varphi_3 - y_{p1} + y_{g1})s \\ -c_{z1}(r_{p1}\dot{\varphi}_2 - r_{g1}\dot{\varphi}_3 - \dot{y}_{p1} + \dot{y}_{g1}) = 0 \quad \text{.....(2.17)}$$

The translational motion of the drive gear (Z_3) in the second stage is given by:

$$m_{p1}\ddot{y}_{p2} + k_{p2}(y_{p2} - y_{c2}) + c_{p1}(\dot{y}_{p2} - \dot{y}_{c2}) - k_{z2}(r_{p2}\varphi_4 - r_{g2}\varphi_5 - y_{p2} + y_{g2}) \\ -c_{z2}(r_{p2}\dot{\varphi}_4 - r_{g2}\dot{\varphi}_5 - \dot{y}_{p2} + \dot{y}_{g2}) = 0 \quad \text{.....(2.18)}$$

The translational motion of the driven gear (Z_4) in the second stage is given by:

$$m_{g2}\ddot{y}_{g2} + k_{g2}(y_{g2} - y_{c2}) + c_{g2}(\dot{y}_{g2} - \dot{y}_{c2}) - k_{z1}(r_{p2}\varphi_4 - r_{g2}\varphi_5 - y_{p2} + y_{g2}) \\ -c_{z2}(r_{p2}\dot{\varphi}_4 - r_{g2}\dot{\varphi}_4 - \dot{y}_{p2} + \dot{y}_{g2}) = 0 \quad \text{.....(2.19)}$$

For the motion of case 1 the governing equation is:

$$m_{c1}\ddot{y}_{c1} + k_{c1}y_{c1} + c_{c1}\dot{y}_{c1} - k_{p1}(y_{p1} - y_{c1}) - c_{p1}(\dot{y}_{p1} - \dot{y}_{c1}) = 0 \quad \text{.....(2.20)}$$

For the motion of case 2 the governing equation is:

$$m_{c2}\ddot{y}_{c2} + k_{c2}y_{c2} + c_{c2}\dot{y}_{c2} - k_{g1}(y_{g1} - y_{c2}) - c_{g2}(\dot{y}_{g2} - \dot{y}_{c2}) = 0 \quad \text{.....(2.21)}$$

where:

I_m = moment of inertia for electric motor;

I_l = moment of inertia for the load system;

I_{p1} = moment of inertia of drive gear one (pinion one) in the first stage;

I_{g1} = moment of inertia of driven gear in the first stage;

I_{p2} = moment of inertia of drive gear (pinion two) in the second stage;

I_{g2} = moment of inertia of driven gear in the second stage;

F_1, F_2 = gearing stiffness forces;

F_{1t}, F_{2t} = gearing damping forces;

M_m = input motor torque;

M_1, M_{1t} = internal moments and coupling damping in first shaft;

M_2, M_{2t} = internal moments and coupling damping in second shaft;

M_3, M_{3t} = internal moments and coupling damping in third shaft;

φ_1 = angular displacement of induction motor;

φ_2, φ_3 = the angular displacements of drive gear Z_1 and driven gear Z_2 in the first stage respectively;

φ_4, φ_5 = the angular displacements of gear one (pinion) and gear two in the second stage respectively;

φ_6 = angular displacement of load system;

r_{p1}, r_{g1} = base circle radius of the drive gear and driven gear in the first stage;

r_{p2}, r_{g2} = base circle radius of the drive gear and driven gear in the second stage;

y_p, y_g =vertical displacement of the drive gear and the driven gear in the first stage;

y_{c1}, y_{c2} = vertical displacement of the upper and lower casings;

m_{p1}, m_{g1} = mass of the drive and the driven in the first stage;

m_{p2}, m_{g2} = mass of the drive and the driven in the second stage;

m_{c2}, m_{c2} = mass of upper and lower casing;

c_p = first stage pinion shaft damping;

c_g = first stage gear shaft damping;

c_1 = shaft one damping; c_2 = shaft two damping; c_3 = shaft three damping;

c_{z1} = first stage gearing damping; c_{z2} = second stage gearing damping;

c_{c1} = upper casing support damping; c_{c2} = lower casing support damping;

k_p = first stage pinion shaft stiffness; k_g = first stage gear shaft stiffness;

k_1 = shaft one stiffness (rotational stiffness);

k_2 = shaft two stiffness;

k_3 = shaft three stiffness and coupling stiffness;

k_{z1} = first stage gearing stiffness (meshing stiffness);

k_{z2} = second stage gearing stiffness (meshing stiffness);

k_{c1} = upper casing support stiffness;

k_{c2} = lower gearing support stiffness.

Similar to the case of the one-stage gear system, the vibration responses of the two-stage gearbox also consist of both translational and rotational motions because of the coupling effect. Moreover, the coupling can also occur between the two stages and the overall vibrations will be a combination of different components between the two stages.

2.2.4 Simulated gearbox vibration signals

To understand the vibration components arising from a multi-stage gearbox, a numerical study was based on the model expressed by Equations (2.10)-(2.21). Figure 2.5 presents the time series of signals and their corresponding spectra. They were obtained based on the two-stage gearbox shown in Figure 2.4 and parameters s in Section 4.3, which is solved numerically when the two stiffness functions are in the form of sinusoidal excitations, which is a basic component to construct a more realistic stiffness function such as a rectangular one.

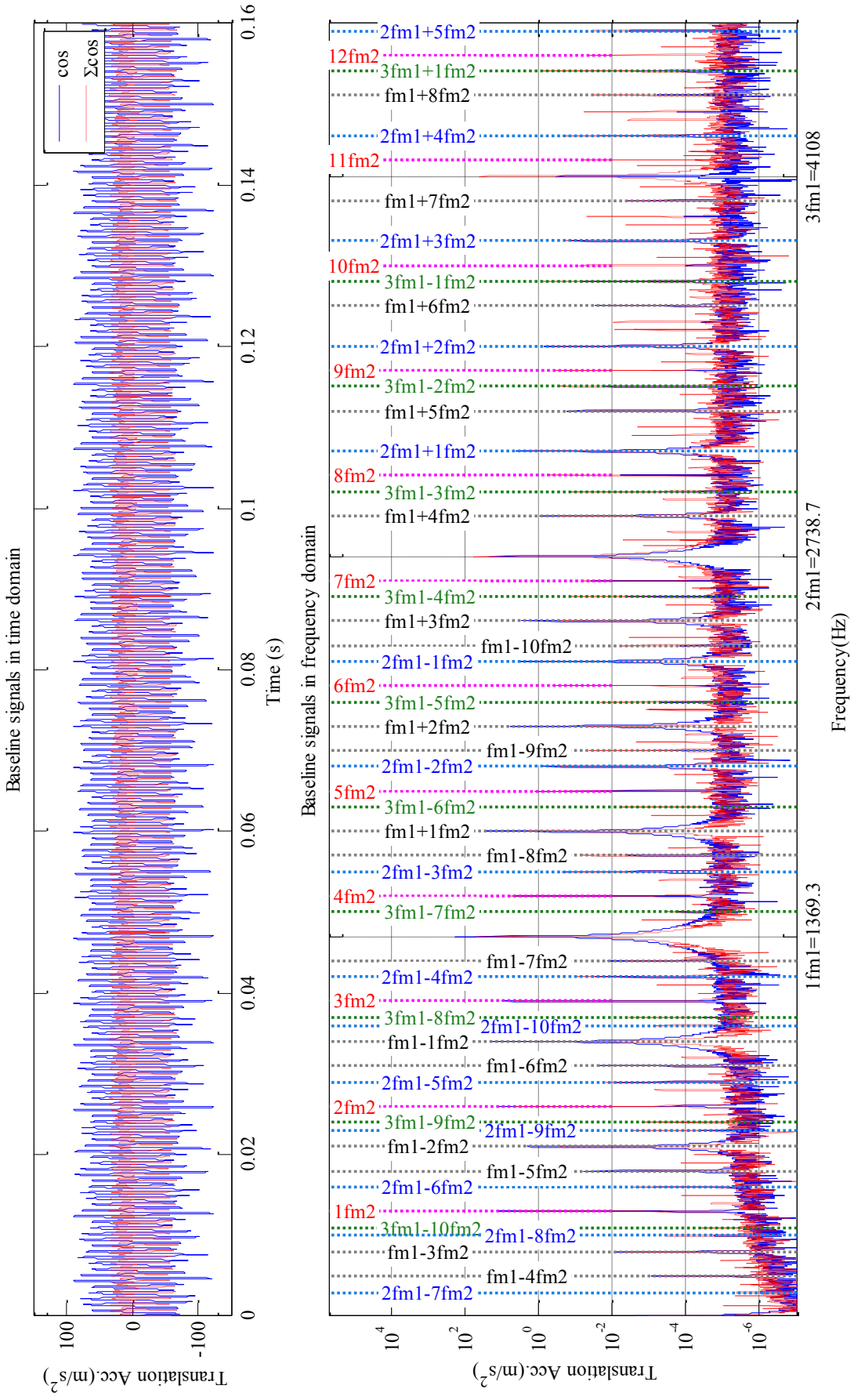

Figure 2.5 Simulated vibration signal of a two-stage gearbox

From the vibration spectra, it can be seen that the signals are composed of not only the first stage mesh frequency component and its harmonics and the component of the second stage mesh frequency and its harmonics, but also a high density of components resulting from a complicated combination of the two mesh frequency components as a consequence of the mutual coupling between the two basic components of the two-stage meshing processes. As an example, the spectral peak at $f_{m1}-7f_{m2}$ is resulted from a modulation between the fundamental components of the first stage with the 7th harmonics of the second stage, showing the coupling between the two stages.

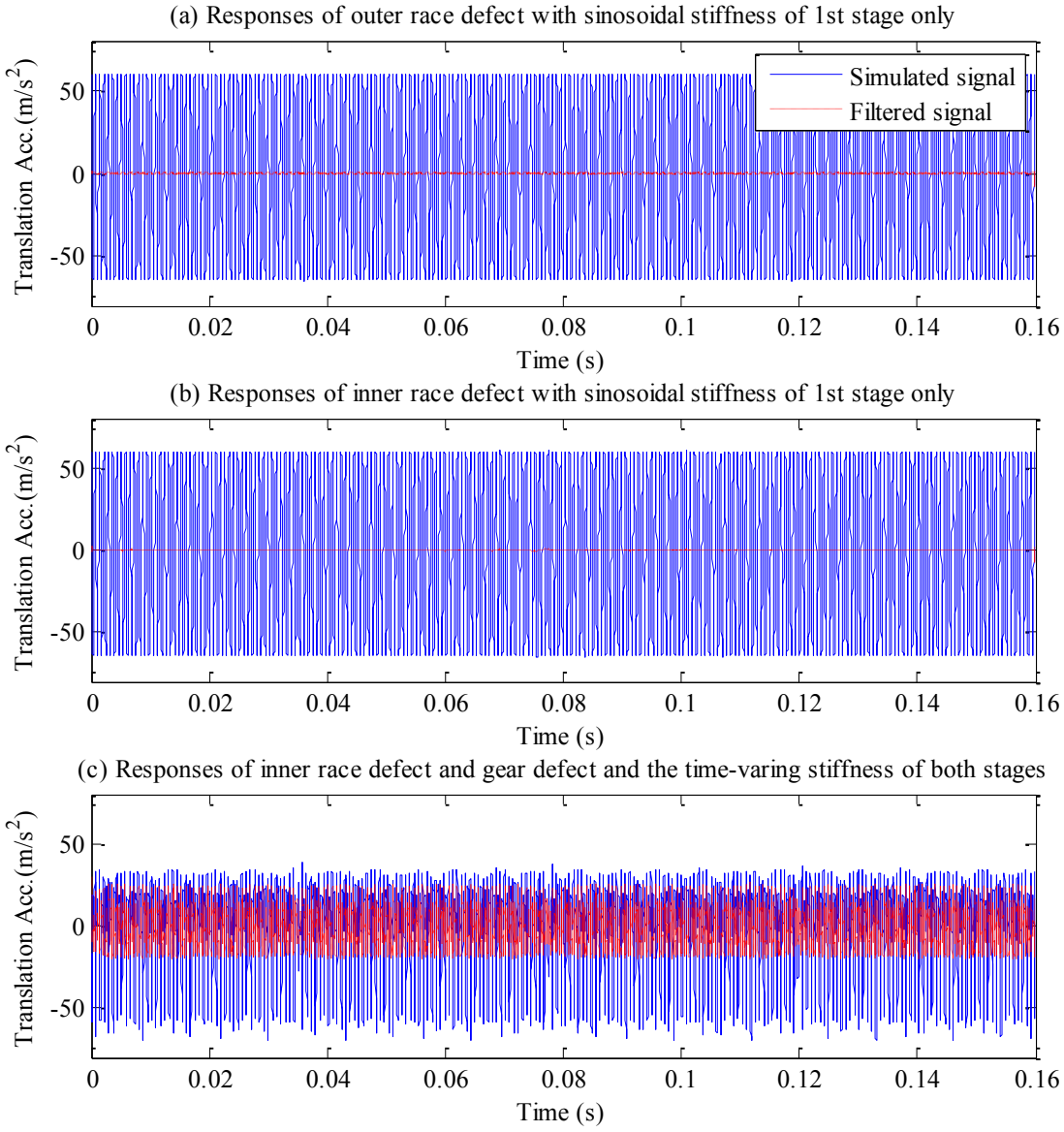

Figure 2.6 Time series of simulated vibration responses of bearing fault and gear fault

It shows that the frequency contents are rich and complicated even if only a sinusoidal stiffness is applied to the system. Therefore, it is necessary to be very careful in identifying

the correct components for diagnostics when analysing signals from real measurements where the stiffness excitation is not only a combination of multiple sinusoidal waveforms but also can be affected by manufacturing errors and noises.

The signal model can be expressed as:

$$x(t) = \sum_{m1=0}^{M} \cos(2\pi f_{m1} t) \sum_{m2=0}^{N} \cos(2\pi f_{m2} t) \quad \text{.....(2.22)}$$

where f_{m1} and f_{m2} denote the mesh frequency of the first and the second stages; M and N indicate the harmonic number of the first and the second stages, respectively.

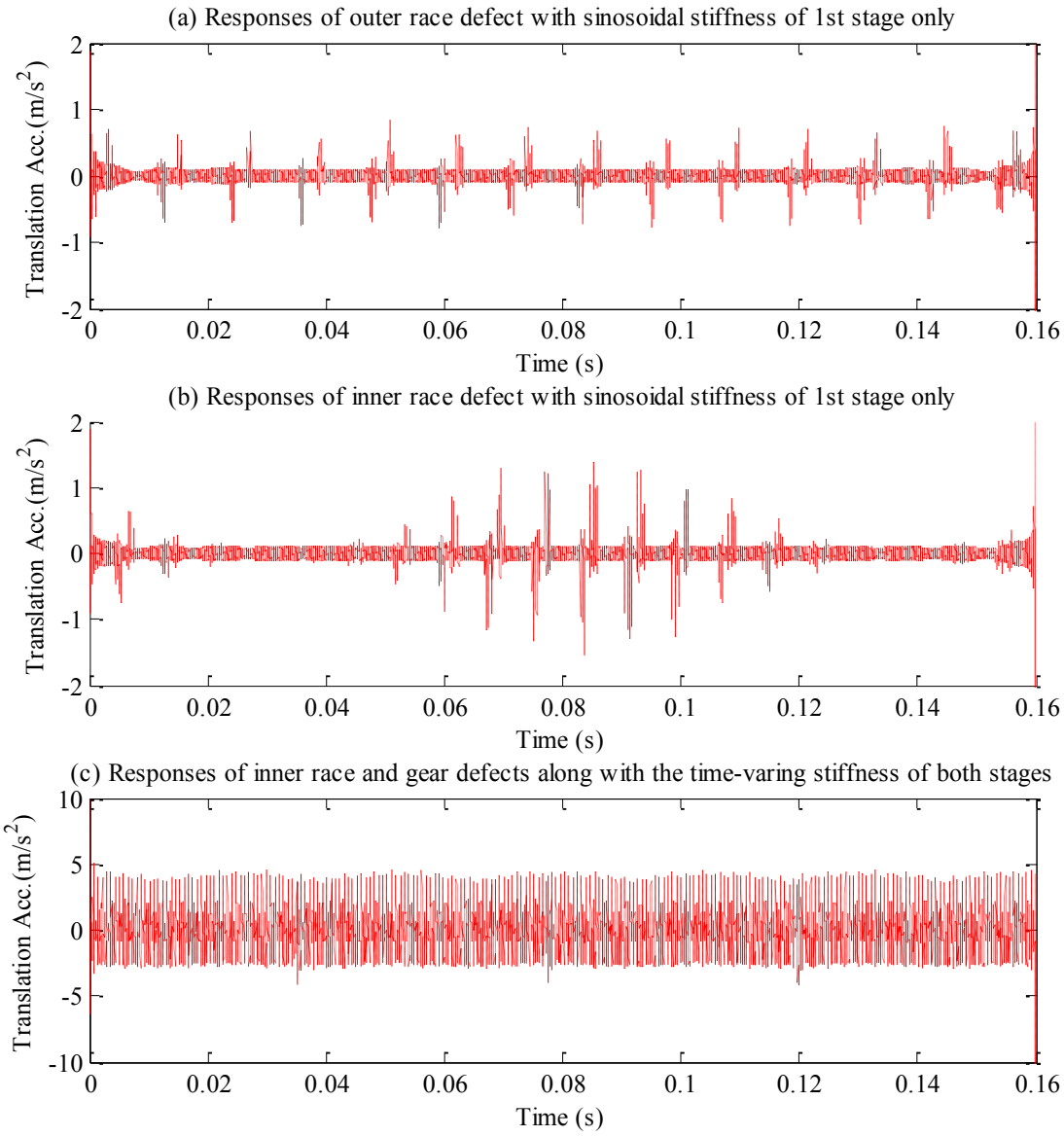

Figure 2.7 Filtered vibration responses of bearing fault and gear fault

In this section, three cases of combined fault are simulated based on the two-stage gear vibration model described in Section 2.2.3. The first one is a bearing outer race defect with sinusoidal stiffness of the 1st stage and constant stiffness of the 2nd stage. The second one is a bearing inner race defect with sinusoidal stiffness of the 1st stage and constant stiffness of the 2nd stage. The third one is a combined fault of bearing inner race defect and gear defect with time-varying stiffness of both stages. From Figure 2.11, it can be seen that the fault components are submerged in the high amplitude meshing frequencies. To view the fault components clearly, a high-pass filter is applied to reduce the meshing frequencies and the results are given in Figure 2.7. It is clear that the signals contain a series of impulses. In the signal processing algorithms, it is critical to extract these impulses for reliable fault detection and diagnosis.

Figure 2.8 gives the spectra of the simulated signals around the mesh frequency of the first stage. It can be seen that the bearing fault frequency is modulated on the gear mesh frequency and its harmonics. The sidebands of bearing fault frequency at cage frequency still exist.

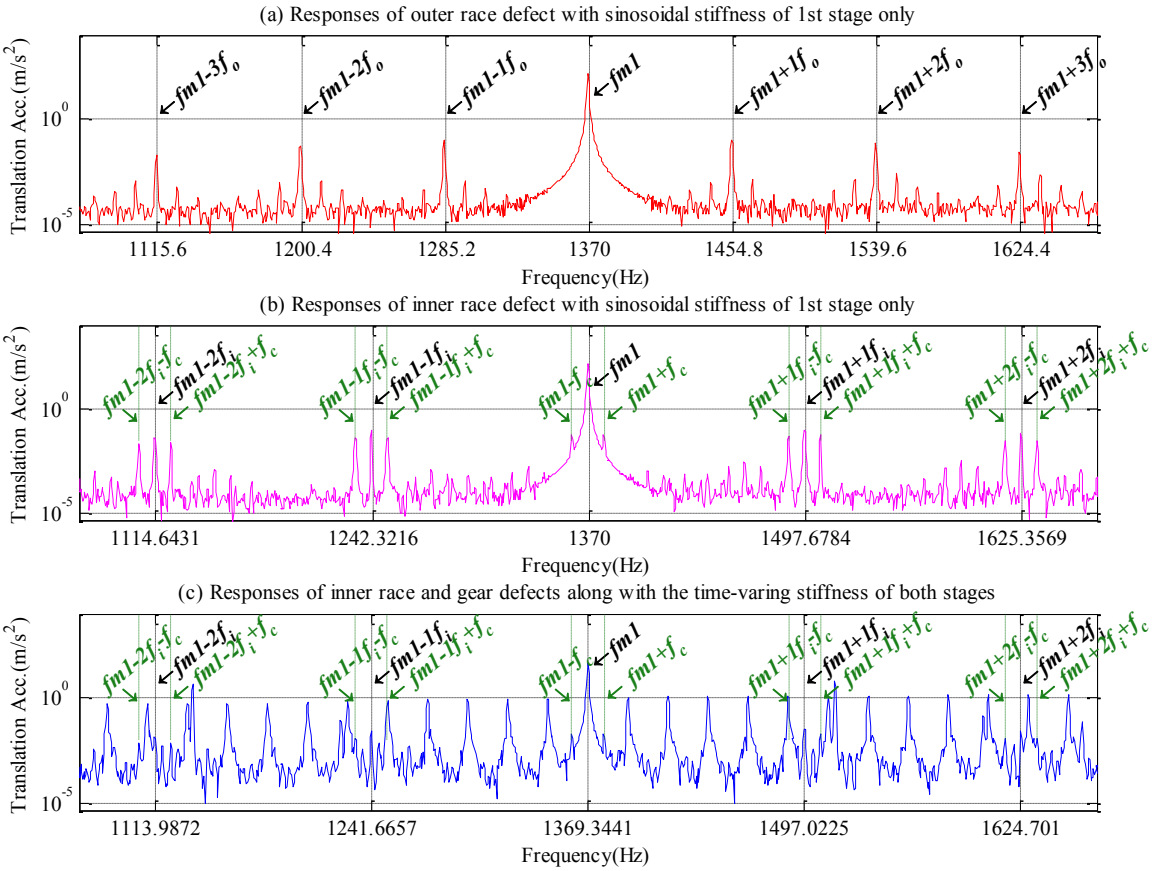

Figure 2.8 Spectra of simulated vibration responses of bearing fault and gear fault

2.3 Bearing fault model based on kinetic analysis

2.3.1 Bearing kinematics

A standard rolling bearing has four fundamental components: the outer race, the inner race, rolling elements and the cage. The rolling elements are guided in the case that ensures uniform intervals between elements and prevents mutual contact. Reference [52] described the bearing kinematics. Figure 2.9 shows a schematic graph of a rolling element bearing and its typical defects, which can be generated by fatigue, wear and poor installation. Due to the planetary motion of the elements, the rotational motions in the bearing are slightly complicated, which will lead to at least four characteristic motions due to defects occurring in the four components.

To derive the fault characteristics of bearing vibration responses, four conditions are assumed in the bearing operating: (1) all rollers have the same diameters; (2) the contact between the rollers, inner race and outer race are in pure rolling contact; (3) no slip occurs between the shaft and the bearing inner race.

From the geometry, assuming a constant operating contact angle φ, the pitch circle diameter of the bearing D_c can be approximated by the inner race diameter D_i and outer race diameter D_o as:

$$D_c = \frac{D_i + D_o}{2} \quad \text{.....(2.23)}$$

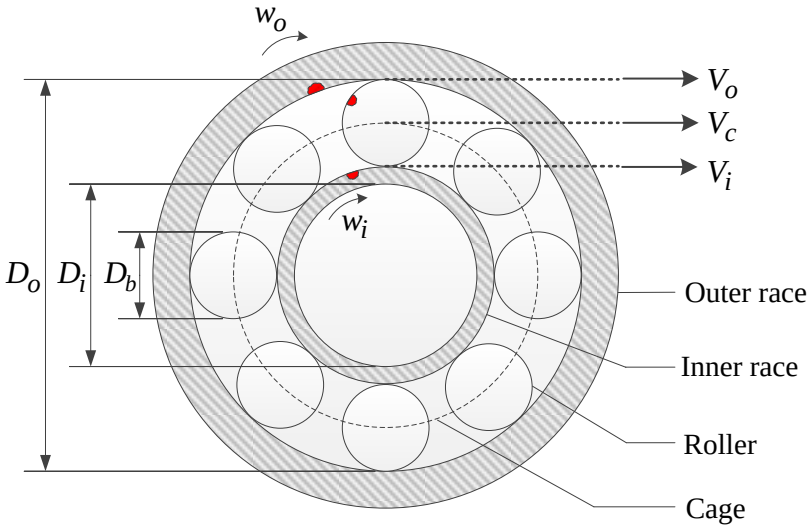

Figure 2.9 Schematic of a rolling element bearing

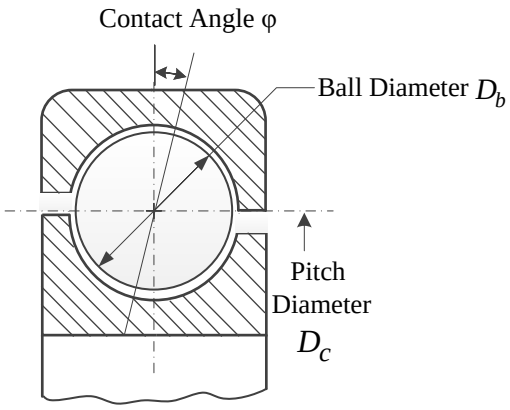

Figure 2.10 Contact angle of rolling element bearing

The race diameters can be expressed with the pitch diameter, the ball diameter and the contact angle as:

$$\begin{aligned} D_i &= D_c - D_b \cos\varphi \\ D_o &= D_c + D_b \cos\varphi \end{aligned} \quad \text{.....(2.24)}$$

The circumferential velocity of the bearing components can be derived in terms of the angular velocity (rad/s) and radius (m). The inner race circumferential velocity can be calculated by:

$$V_i = w_i \frac{D_i}{2} \quad \text{.....(2.25)}$$

and the outer race velocity can be calculated by:

$$V_o = w_o \frac{D_o}{2} \quad \text{.....(2.26)}$$

The velocity of the cage V_c is the average of the inner race velocity and the outer race velocity when no slip occurs.

$$V_c = \frac{V_i + V_o}{2} \quad \text{.....(2.27)}$$

Substituting Equations (2.24)-(2.26) into Equation (2.27), cage velocity V_c becomes:

$$V_c = \frac{w_i (D_c - D_b \cos\varphi)}{4} + \frac{w_o (D_c + D_b \cos\varphi)}{4} \quad \text{.....(2.28)}$$

Conversion of the circumferential velocity from m/s to angular frequency revolutions/sec (Hz) by dividing through by πD_c and rearranging to give the cage frequency in Hz rather than velocity gives:

$$f_c = \frac{f_{in}(1-\frac{D_b}{D_c}\cos\varphi)}{2} + \frac{f_{ot}(1+\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.29)}$$

Equation (2.29) is the theoretical cage or fundamental train frequency (FTF) for rolling element bearings. In many situations, one of the raceway will be stationary, the most common being the outer race. In this case, Equation (2.29) can be further simplified to the familiar form:

$$f_c = \frac{f_{in}(1-\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.30)}$$

The frequency of rotation of the rolling elements with respect to the inner race can likewise be derived by

$$f_{ri} = f_c - f_{in} \quad \text{.....(2.31)}$$

which, by substitution of Equation (2.29), becomes

$$f_{ri} = \frac{f_{ot}(1+\frac{D_b}{D_c}\cos\varphi)}{2} - \frac{f_{in}(1+\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.32)}$$

The ball pass frequency on the inner race can be expressed by:

$$f_i = \frac{N_r(f_{ot} - f_{in})(1+\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.33)}$$

When the outer race is stationary, the ball pass inner race frequency becomes

$$f_i = \frac{-N_r f_{in}(1+\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.34)}$$

The frequency of rotation of rolling elements with respect to the outer race can be calculated by

$$f_{ro} = f_{ot} - f_c \quad \text{.....(2.35)}$$

Substituting Equation (2.29) into Equation (2.35), it becomes

$$f_{ro} = \frac{f_{ot}(1-\frac{D_b}{D_c}\cos\varphi)}{2} - \frac{f_{in}(1-\frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.36)}$$

The ball pass frequency on the outer race can be obtained by multiplying the roller number N_r:

$$f_o = \frac{N_r(f_{ot} - f_{in})(1 - \frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.37)}$$

When the outer race is stationary, the outer race fault frequency can be expressed as:

$$f_o = \frac{-N_r f_{in}(1 - \frac{D_b}{D_c}\cos\varphi)}{2} \quad \text{.....(2.38)}$$

The rotating frequency of the rollers about their own axes can also be derived. The frequency of rotation, assuming no slip, is given by the rotating frequency of the cage with respect to the inner race multiplied by the diameter ratio of the inner race to the ball diameter

$$f_{bs} = f_{ri} * \frac{D_i}{D_b} \quad \text{.....(2.39)}$$

Substitutings Equations (2.24) and (2.32) into Equation (2.39), the roller fault frequency can be expressed as:

$$f_{bs} = \frac{f_{ot} - f_{in}}{2}\frac{D_c}{D_b}(1 - (\frac{D_b}{D_c}\cos\varphi)^2) \quad \text{.....(2.40)}$$

When the outer race is stationary, the roller race fault frequency can be expressed as:

$$f_{bs} = -\frac{f_{in}}{2}\frac{D_c}{D_b}(1 - (\frac{D_b}{D_c}\cos\varphi)^2) \quad \text{.....(2.41)}$$

Equations (2.29), (2.33), (2.37), and (2.40) give the general forms of bearing fault frequencies of the cage, inner race, outer race, and rollers, respectively. Equations (2.30), (2.33), (2.38) and (2.41) show the bearing fault frequencies when the outer race is stationary.

2.3.2 Bearing vibration signal models

The vibration signature of a defective rolling element bearing is typified by amplitude modulations [53]. In a bearing system, the carrier signal will likely be a combination of the resonant frequencies of the bearing and probably of the mechanical system associated with

it. The vibration signal from a rolling element bearing with a local defect can hence be represented as [53]:

$$x(t) = x_f(t)x_q(t)x_{bs}(t) + x_s(t) + n(t) \quad \text{.....(2.42)}$$

where $x_f(t)$ are the periodic impulses produced by the defect, $x_q(t)$ is the modulation effect due to the non-uniform load distribution and $x_{bs}(t)$ is the bearing-induced vibration determined by the bearing structural dynamics, $x_s(t)$ is the machinery-induced vibration determined by the machine structure and related components, and $n(t)$ is the noise which is encountered inevitably in any measurement system. This shows that the fault signatures of a local bearing defect comprise modulation effects between fault frequencies, load distribution and structural resonances. 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. Therefore, to extract fault signatures effectively, the signal must be both denoised and demodulated.

$x_f(t)$ in Equation (2.42) represents 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 [54]. The repetition frequency for an outer race defect is denoted f_o; 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 $x_q(t)$, as shown in Figure 2.11. For an inner race defect, the modulating frequency is the shaft rotational frequency f_r, but for a rolling element defect the modulating frequency mainly includes the cage frequency f_c. The theoretical characteristic frequencies of a rolling element bearing can be calculated by Equations (2.43)-(2.46) [54].

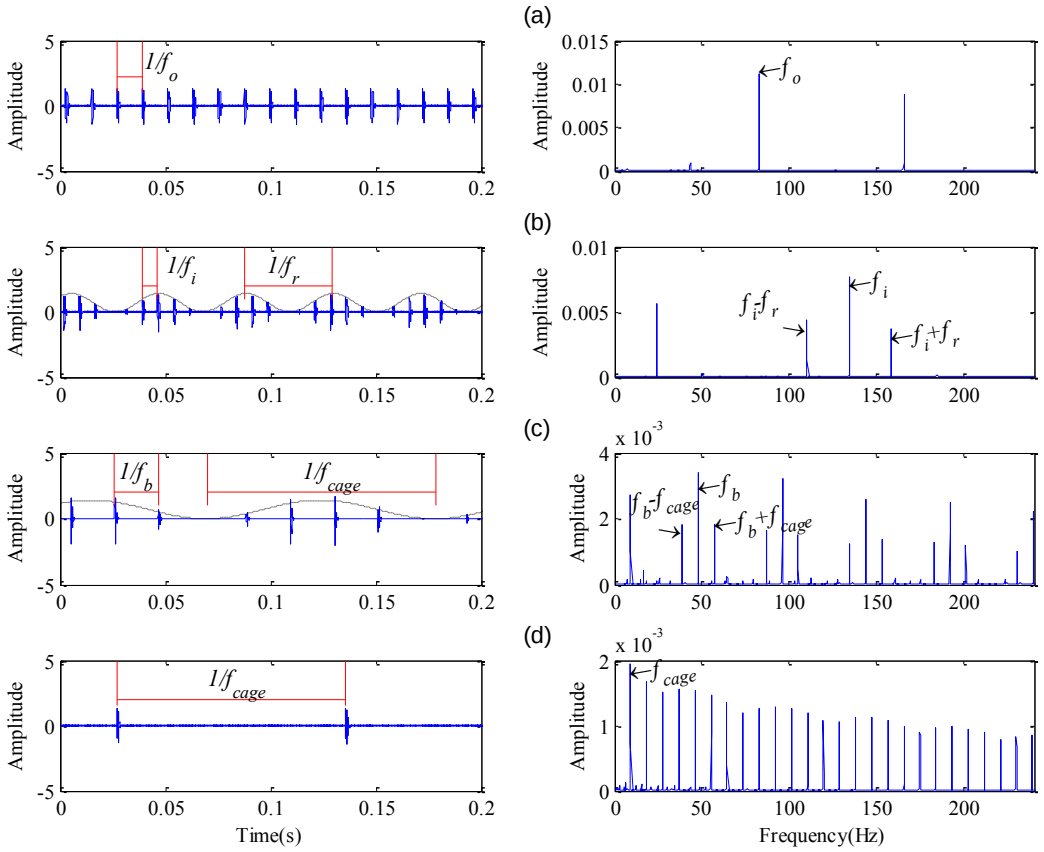

Figure 2.11 Simulated fault data and spectra of a rolling element bearing with a localised defect on the (a) outer race, (b) inner race (c) rolling element, and (d) cage

Outer race fault frequency:

$$f_o = \frac{N_r}{2} f_r (1 - \frac{D_b}{D_c} \cos\varphi) \quad \text{.....(2.43)}$$

Inner race fault frequency:

$$f_i = \frac{N_r}{2} f_r (1 + \frac{D_b}{D_c} \cos\varphi) \quad \text{.....(2.44)}$$

Ball fault frequency:

$$f_b = \frac{D_c}{2D_r} f_r (1 - (\frac{D_b}{D_c} \cos\varphi)^2) \quad \text{.....(2.45)}$$

Cage fault frequency (often called the fundamental train frequency):

$$f_{cage} = \frac{1}{2} f_r (1 - \frac{D_b}{D_c} \cos\varphi) \quad \text{.....(2.46)}$$

where D_c is the pitch circle diameter, D_b is the roller diameter, φ is the contact angle as shown in Figure 2.10, and N_r is the number of balls (or rollers).

From the simulated signal, it can be seen clearly that the bearing signal includes periodic impulses and exhibits amplitude modulation features. The modulation could be induced by the transmission path and load zone effect.

2.4 Combination of gear and bearing models based on planetary gearbox vibration responses

In this section, a combined gear and bearing model is proposed to understand the interaction of the two components.

2.4.1 Planetary gearbox vibration characteristics

A planetary gearbox is composed of a ring gear, a sun gear and multiple planet gears. Usually, the ring gear is stationary, a sun gear rotates around a fixed centre, and planet gears not only spin around their own centres but also revolve around the centre of the sun gear. The planet gears mesh simultaneously with both the sun gear and the ring gear. Due to these complicated gear motions, the vibration signals generated by planetary gearboxes are more complicated than those by fixed shaft gearboxes. In addition, the planet phasing relationship, which is dependent on the number of planets, planet position angles, and the number of teeth of each gear, also adds complexity to vibration signals. In this section, the planetary gearbox vibration signal models will be introduced. The gear damage could produce the amplitude modulation and frequency modulation (AM-FM) effects to the gear mesh vibration at corresponding fault characteristic frequencies [55].

Based on the theoretical analysis in [56], in steady working condition such as constant running load and speed, the vibration perceived by a sensor on the stationary ring can be represented with mutual modulations of both AM and FM phenomena. For a local fault, such as the crack and pitting on one tooth of the sun gear, the signal model for the 1st sinusoidal component can be expressed as:

$$f(t)=[1-\cos(2\pi f_{rs}t)][1-\cos(2\pi f_{rc}t)][1+A\cos(2\pi f_{sf}t+\varphi)] \times\cos[2\pi f_m t+B\sin(2\pi f_{sf}t+\phi)+\theta] \quad \text{.....(2.47)}$$

on the planet gear

$$f(t)=[1-\cos(2\pi f_{rc}t)][1-\cos(2\pi f_{rc}t)][1+A\cos(2\pi f_{pf}t+\varphi)] \times\cos[2\pi f_m t+B\sin(2\pi f_{pf}t+\phi)+\theta] \quad \text{.....(2.48)}$$

and on the ring gear

$$f(t)=[1+A\cos(2\pi f_{rf}t+\varphi)][1-\cos(2\pi f_{rc}t)] \times\cos[2\pi f_m t+B\sin(2\pi f_{rf}t+\phi)+\theta] \quad \text{.....(2.49)}$$

where f_{sf}, f_{pf} and f_{rf} is the fault characteristic frequency of the sun gear, planet gear and ring gear respectively. f_{rc} and f_{rs} are the rotating frequency of the carrier and sun gear. f_m is the gear mesh frequency. θ, φ and ϕ are the initial phases of AM and FM respectively.

Therefore, consider the AM-FM effects with the high orders of fault gear characteristic frequency nf_{xf} as the modulating frequency and with the higher orders of mesh frequency kf_m as the signal carrier frequency and f_{rx} as the corresponding component rotating frequency, the vibration spectral peaks will appear at the frequency locations of $kf_m \pm nf_{xf} \pm f_{rc}$ and $kf_m \pm f_{rx} \pm nf_{xf} \pm f_{rc}$ $(k,n=1,2,3\cdots)$ in the Fourier spectrum. From the analysis of vibration spectra, the gear fault can be detected and located by monitoring the presence of magnitude increase of spectral peaks at the above mentioned frequency locations.

2.4.2 Characteristic frequencies for gear fault detection

According to reference [56], the rotational frequency of the carrier can be calculated as

$$f_{rc}=\frac{Z_s}{Z_r+Z_s}f_{rs} \quad \text{.....(2.50)}$$

the planet gear frequency as

$$f_{rp}=\frac{(Z_p-Z_r)Z_s}{(Z_r+Z_s)Z_p}f_{rs} \quad \text{.....(2.51)}$$

and the mesh frequency as

$$f_m=(f_{rs}-f_{rc})Z_s=\frac{Z_r Z_s}{Z_r+Z_s}f_{rs}=Z_r f_{rc}, \quad \text{.....(2.52)}$$

where f_{rs} is the sun gear rotating speed; Z_r, Z_p and Z_s denote the number of teeth for the ring, planet and sun gears, respectively.

As shown in many previous studies [55][57], detection and diagnosis can be carried out by examining the changes of characteristic frequencies around the mesh frequency f_m and its harmonics. Considering that there are K number of planetary gears moving with the carrier, characteristic frequencies around the mesh frequency can be calculated [58][59] for different local faults occurring on the sun gear

$$f_{sf} = \frac{f_m}{Z_s} = K\left(f_{rs} - f_{rc}\right) \quad (2.53)$$

on the planet gear

$$f_{pf} = 2\frac{f_m}{Z_p} = 2\left(f_{rp} + f_{rc}\right) \quad (2.54)$$

and on the ring gear

$$f_{rf} = \frac{f_m}{Z_r} = Kf_{rc} \quad (2.55)$$

However, as shown in [59][60], only some of these expected sidebands will be apparent in the vibration spectrum when a planetary gearbox has faults due to the effects of constructive superposition of the vibration waves from the three gear sets, whereas other sidebands are hard to see because of the destructive effect of the superposition, and hence the latter have been largely neglected by previous studies when developing methods for fault diagnosis.

2.4.3 Characteristic frequencies for bearing fault detection

A ball bearing consists of an inner race, an outer race, several balls and a cage, which holds the balls in a given relative position. Race surface fatigue results in the appearance of spalls on the inner race, outer race or balls. If one of the races has a spall, it will almost periodically impact with the balls. The fault signature can be represented by successive impulses with a repetition rate depending on the faulty component, geometric dimensions and the rotational speed. The period between impulses is different for all the listed elements and depends on the geometry of the bearing, the rotational speed and the load angle. For a fixed outer race bearing, the theoretical characteristic fault frequencies can be calculated using Equations (2.43)-(2.46).

While the sensor is mounted on the gearbox housing, which is connected to or fastened to the ring gear directly in most situations, the bearing damage induced vibration has two

main paths to go from its source to the sensor through solid mechanical components and their contacts. Through the first path, the vibration signal propagates from its origin to the gearbox casing, and then reaches the sensor. Whereas through the second path, the vibration signal follows a longer path, from its origin to the shaft firstly, then from the shaft go through the sun gear, planet gear and ring gear, after that from the ring gear to the gearbox casing, and finally to the sensor. Therefore, the vibration signal will be amplitude modulated by the sun gear rotating frequency and carrier rotating frequency. Because when the vibration transmitted through the planetary gear, it is also modulated by the gear mesh frequency. Therefore, the vibration signal can be expressed as:

$$f(t) = [1-\cos(2\pi f_{rs}t)]\cos[2\pi f_{bx}t+\alpha] + [1-\cos(2\pi f_{rc}t)]\cos[2\pi f_{bx}t+\alpha] + \sum_{n=1}^{N}\cos(2\pi n f_m t)\cos[2\pi f_{bx}t+\alpha] \quad \text{.....(2.56)}$$

where f_{bx} is the characteristic frequency of bearing, and f_m denotes the planetary gear mesh frequency.

2.5 Summary

In this chapter, vibration signal characteristics for bearing faults, gear faults and combined faults are studied. All of these signals can present periodic, impulsive and modulating content due to their kinetic and dynamic mechanisms, which is the important basis of effective signal processing and feature extraction for accurate fault detection and diagnosis.

Chapter 3
Fundamentals of key signal processing techniques

This chapter provides a more in-depth understanding of key signal processing techniques which are potentially effective in noise reduction and impulsive signature enhancement. In particular, ALE, TSA, wavelet analysis, TFR based kurtogram and bispectrum analysis are focused on as they have been shown to have good performances for impulsive feature enhancement by active investigations in recent years.

3.1 Introduction

Based on the types of data collected or the types of sensors used for data acquisition, there are a variety of technologies that can be used for making maintenance decisions in condition monitoring systems, such as visual inspection, trend monitoring, vibration-based condition monitoring, and acoustic-based condition monitoring. Since mechanical systems or machines account for the majority of plant equipment, vibration monitoring is generally the key technique in most systems. Sometimes, all of these technologies are employed in a comprehensive system in order to make accurate maintenance decision.

Generally, there is considerable noise in the collected vibration signals, not only from measurement but also from the machine's normal operation, which would influence the decision making of fault detection and diagnosis. Especially in the early fault stage, weak fault signals are often submerged in loud background noise. To attain reliable condition monitoring and fault diagnosis results, therefore, noise reduction pre-processing is a critical subject in signal processing. In order to improve the signal-to-noise ratio, numerous techniques have been developed and improved.

In this chapter, the fundamentals of such key techniques are re-examined in more depth regarding their operating mechanisms and performances in achieving noise reduction and impulse signal enhancements. This will pave the ways for developing more advanced techniques for achieving further noise reduction and signal enhancement and hence reliable diagnostic results.

3.2 Adaptive line enhancer

The ALE is a degenerated form of adaptive noise cancellation (ANC), consisting of a single sensor and delay $z^{-\Delta}$ to produce a delayed version input signal, which de-correlates the noise while leaving the target component correlated. Ideally, the output of the adaptive filter in the ALE is an estimate of the noise-free input signal. Hence, the ALE is capable of separating the periodic and stochastic components in a signal [61][62][63].

In 1979 the ALE was introduced as a method for detecting a periodic signal in an incoherent background or conversely of removing periodic interference from a broad-band signal of interest [64]. This technique can be used with any of the adaptive filters classified to date and uses a delay in the input signal to filter out the unnecessary part in it and thus

get the desired response. Naoto Sasaoka et al. applied ALE to reduce the sinusoidal noise in noisy speech signals in 2005 [65].

In 1996 and 1998, to reduce the background noise level of rotating and reciprocating machinery, S. K. Lee and P. R. White exploited a two-stage ALE filter structure in series. The resulting enhanced signals are analysed in the time-frequency domain to obtain simultaneous spectral and temporal information. The techniques developed are applied to the diagnosis of faults within an internal combustion engine and to data from an industrial gearbox [66][67].

In 2006, S. K. Lee and J. S. Lee applied ALE to enhance the impulsive components in gearbox vibration signals and then analyse them by using time-frequency representation [68].

In 2012, J. R. Mohammed et al. presented one noise reduction system based on two stages of operation, with the first stage based on the ALE filters and the second stage on NLMS (Normalised Least Mean Square) filter. The first stage reduces the sinusoidal noise from the input signal and the second stage reduces the wideband noise [69].

According to Widrow [70], ALE is an adaptive self-tuning filter capable of separating the periodic and stochastic components in a signal. As shown in Figure 3.1, the ALE simply uses a single sensor and is therefore easier and more cost-effective to implement in condition monitoring practice.

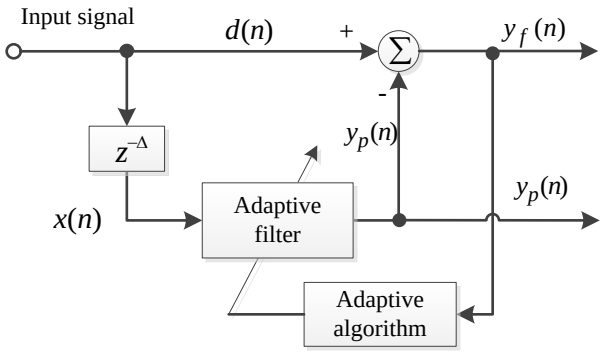

Figure 3.1 Block diagram of adaptive line enhancer

In this case, the reference signal is obtained by delaying the input signal by a certain number of samples Δ. The result is that any stationary input components which are predictable over the delay appear at the filter output $y_p(n)$, whilst the nonstationary output $y_f(n)$ contains those components which are unpredictable over the delay.

The ALE becomes an interesting application in noise reduction because of its simplicity and ease of implementation. However, to obtain the best performance in its computational process, the optimal approach is to execute ALE on a better convergence rate of adaptive algorithm with a less complex structure.

The most widely used adaptive filtering technique is a version of the LMS algorithm, initially proposed by Widrow and Hoff. The LMS is based on the steepest descent method, a gradient search technique to determine filter coefficients that minimise the mean square prediction of a transversal filter. The NLMS converges faster than the conventional LMS because it employs a variable step size parameter aimed at minimising the instantaneous output $y_f(n)$ [72] [73]. NLMS algorithm can be summarised as shown below.

The output of adaptive filter is

$$y_p(n) = \mathbf{x}^{\mathbf{T}}(n)\mathbf{w}(n) \quad \text{.....(3.1)}$$

while $x(n)$ and $w(n)$ are given in Equations (3.2) and (3.3)

$$\mathbf{x}(n) = [x(n) \cdots x(n-L-1)]^{\mathbf{T}} \quad \text{.....(3.2)}$$

$$\mathbf{w}(n) = [w_0(n) \cdots w_{L-1}(n)]^{\mathbf{T}} \quad \text{.....(3.3)}$$

The adaptive coefficients are updated by the following formula

$$\mathbf{w}(n) = \mathbf{w}(n-1) - \frac{\mu}{\|\mathbf{x}(n)\|^2 + \alpha}\mathbf{x}(n)y_{ap}(n) \quad \text{.....(3.4)}$$

$$y_{ap}(n) = d(n) - y_p(n) \quad \text{.....(3.5)}$$

where the non-stationary output signal $y_{ap}(n)$ is given by Equation (3.5), μ is the step size parameter which controls the convergence speed and the stability of the filter and α is a small constant in order to prevent division by zero in case no input signal is present.

3.3 Wavelet analysis

Wavelet transforms are effective in processing non-stationary vibration signals and have been successfully used for revealing the inherent information for fault detection and diagnosis. The applications of the wavelet transform in machine fault diagnosis include the following main arears: the time-frequency analysis of signals, fault feature extraction, the

denoising and extraction of weak signals, the compression of vibration signals and system identification [74]-[78].

Wavelet transforms are capable of providing both time-domain and frequency-domain information simultaneously. Similar to a wavelet function, the impulsive feature components in vibration signals have local energy distributions in both the time domain and the frequency domain. Wavelet functions can be used for the detection of impulsive feature components because they have similar time-frequency structures. However, there are different types of wavelet functions each of which have different time-frequency structures; it is obvious that using an optimal wavelet function, whose time-frequency structure best matches that of the impulsive component, can achieve the best performance in impulsive component detection. To find such an optimal wavelet, many algorithms have already been proposed.

In 2000, J. Lin and L. S. Qu used the wavelet entropy as a rule to optimise the parameters of the wavelet function [76]. The vibration signals from the rolling bearing and the gearbox were denoised based on the Morlet wavelet, which was optimised by minimum wavelet entropy. The results of the application in rolling bearing diagnosis and gearbox diagnosis are satisfactory.

In 2003, J. Lin and M. J. Zuo introduced a Morlet wavelet parameters optimisation method based on the kurtosis maximisation principle [79]. The adaptive wavelet filter is found to be very effective in the detection of symptoms from vibration signals of a gearbox with early fatigue tooth crack.

In 2009, W. S. Su et al. optimised Morlet wavelet parameters by genetic algorithm for rolling element bearing fault diagnosis [80]. In this paper, to eliminate the frequency associated with interferential vibrations, the bearing vibration signal is filtered with a band-pass filter determined by a Morlet wavelet function whose parameters are optimised by genetic algorithm.

In 2010, W. He, Z. N. Jiang and Q. Qin developed a joint adaptive wavelet filter and morphological signal processing method for weak mechanical impulse extraction [81]. The optimisation algorithm of the proper wavelet centre frequency and bandwidth of filter is based on differential evolution (DE). This method was applied in bearing vibration signal analysis combined with morphological signal processing (MSP).

In 2011, W.Y. Liu et al. proposed a hybrid time-frequency method based on the improved Morlet wavelet and auto terms window [82]. The shape parameters of Morlet and the appropriate scale parameter for continuous wavelet transformation (CWT) are optimised using the cross validation method (CVM) and the minimum Shannon entropy method. The method was utilised in the gearbox vibration impulse detection for fault diagnosis and the results show that the gearbox fault feature is much clearer than the original signal.

In 2011, Y. H. Jiang et al. put forward another denoising method based on adaptive Morlet wavelet and singular value decomposition (SVD) [83]. Modified Shannon wavelet entropy is used to optimise central frequency and bandwidth parameter of the Morlet wavelet in order to achieve optimal match with the impulsive components. Then the scale periodic exponential (SPE) spectrum based on SVD is utilized to select the appropriate transform scale.

There are two kinds of wavelets: real analytic wavelet and complex analytic wavelet. The complex analytic wavelets can separate the amplitude and phase component and be used to measure the time evolution of the frequency transitions. The real analytic wavelets can be used to detect the impulsive components. Previous studies [80]-[83] have shown the Morlet wavelet gives superior results in detecting impulse components in different types of vibration signals. Therefore, the real-part of the complex Morlet wavelet is also used in this study for optimal detection and diagnosis.

The real Morlet wavelet is defined in the time domain as a sinusoidal wave multiplied by a Gaussian function [83]

$$\psi(t) = \left(\pi f_b\right)^{-0.5} \exp\left(-\frac{t^2}{f_b}\right) \exp\left(j2\pi f_c t\right) \quad (3.6)$$

where f_b is the bandwidth parameter and f_c is the wavelet centre frequency. A feature of this wavelet is that its Fourier spectrum is a Gaussian function:

$$\psi(f) = \exp\left(-\pi^2 f_b \left(f - f_c\right)^2\right) \quad (3.7)$$

It is obvious from Equation (3.7) that the shape of mother wavelet is controlled by f_b and f_c. Based on understandings from previous studies, Morlet wavelet parameters have to be optimised to obtain the best detection of impulse components. Entropy technique based optimisation has been applied in adaptive machine fault detection and diagnosis and shows

promising results in many industrial fields, such as system parameter recognition, structure optimisation, device operation-state detection and fault diagnosis [84]-[88].

3.4 Time synchronous average

One of the most powerful algorithmic tools for vibration analysis is the TSA. TSA resamples the vibration data synchronously with a particular shaft in a complicated drivetrain, and is the basis of numerous gear and shaft condition indicator algorithms.
In essence, TSA is a signal processing technique that extracts periodic waveforms from noisy data based on a scheme of signal phase alignment, as reviewed in Section 1.3.1. Thus, TSA is well suited for multistage gearbox analysis, where it allows the vibration signature of the gear under analysis to be separated from other gears and noise sources in the gearbox that are not synchronous with that gear. Additionally, variations in shaft speed can be corrected, which results in the spreading of spectral energy into an adjacent gear mesh bin. In order to do this, a signal phase-locked with the angular position of a shaft within the system is needed.

This phase information can be provided through an n-revolution tachometer signal (such as a Hall sensor or optical encoder, where the time at which the tachometer signal crosses from low to high is called the zero crossing) or through the demodulation of gear mesh signatures [89]. If n is 1 for the shaft under analysis, a balance solution can be obtained. If n is greater than one, or the shaft/gear under analysis is not the shaft generating the tachometer signal, then no balance solution is available (phase relative to the imbalance vector is ambiguous).

In general, TSA enhances signals by signal phase alignments. Signal components with the same phase as the tachometer signal, which includes the integer multiples of the shaft frequency from the tachometer signal, will be enhanced whereas those out-phase components of the tachometer signal, which includes any components of fractional orders of the shaft frequency, are suppressed. This shows that the use of phase information is very effective in obtaining a signal with minimal noise contamination.

However, to achieve reliably TSA of vibration signals, additional tachometer signals must be collected simultaneously, which not only requires more hardware investment but also it is often difficult, for example, to install a sensor close to the rotating system. In addition, TSA is probably not very effective in analysing bearing signals as the characteristic

frequencies of bearing faults, which have been discussed in Section 2.3, are in the fractional orders of shaft frequency.

Assuming a signal $x(t)$ consists of a periodic signal $x_T(t)$ and a noise component $v(t)$, the period of $x_T(t)$ is T_0 whose corresponding frequency is f_0, thus the signal can be expressed as [90]:

$$x(n) = x_T(n) + v(n) \quad (3.8)$$

The synchronous average of the signal $x(t)$ by using TSA can be expressed as:

$$y(n) = \frac{1}{M} \sum_{i=1}^{M-1} x(t + iT_0) \quad (3.9)$$

where M is the number of the average segments, $y(t)$ is the averaged signal.

3.5 Time-frequency representation and kurtogram

The SK was first introduced by Dwyer, as a complement to the power spectral density (PSD). It was applied to overcome the inefficiency of PSD to detect and characterise impulses in a signal [91]. The definition of the "kurtogram" was first introduced by J. Antoni in order to generalise the so-called "spectral kurtosis" (SK) to a wider class of non-stationary signals [92]. It is a fourth-order spectral analysis tool based on spectral kurtosis (SK) recently introduced for detecting and characterising non-stationarities in a signal. SK is a sensitive tool for detecting impulses in a signal and can indicate the frequencies that the impulses occur.

In recent years, many kinds of implementation methods for the kurtogram have been proposed, such as those based on STFT, Filter Tree (FT), Discrete Wavelet Transform (DWT), Wavelet Packet Decomposition (WPD), etc. A brief introduction to these methods is summarised regarding their implementation and performance on noise reduction and signal enhancement.

In the previous paper, the authors demonstrated the high potential of the spectral kurtosis (SK) to detect and characterise non-stationary signals. The present paper brings together these ideas and shows how the SK can be efficiently used in the vibration-based condition monitoring of rotating machines. First, and in contrast to classical kurtosis analysis, the SK provides a robust way of detecting incipient faults even in the presence of strong masking noise. Second, the SK offers an almost unique way of designing optimal filters for filtering

out the mechanical signature of faults. The first property is of practical importance for monitoring purposes, whereas the second one proves very useful in diagnostics.

In 2006, J. Antoni firstly introduced the concept of the kurtogram based on STFT, from which optimal band-pass filters can be deduced; for instance, as a prelude to envelope analysis [92].

In 2007 J. Antoni [93] proposed a fast algorithm of the kurtogram in order to make it a tool with potential on-line industrial applications. The presented algorithm has a computational complexity similar to that of the FFT. As a by-product, the fast kurtogram also returns the complex envelopes of the signal in selected frequency bands. These complex envelopes are ready for further processing, such as through envelope spectrum analysis.

A kurtogram based on STFT limits the accuracy improvement of the kurtogram in extracting impulsive characteristics from a noisy signal. To overcome these shortcomings, in 2011 Yaguo Lei et al. [94] introduced WPT into the kurtogram to overcome its shortcomings and to further enhance its accuracy in discovering characteristics and detecting faults. The results from the experiment demonstration verified the effectiveness of the method in extracting the fault characteristics and diagnosing the faults of rolling element bearings.

In 2011, Barszcz and JabŁoński [95] found that temporal signal-based kurtosis can be considerably affected by noise, and proposed a novel method called the protrugram, which calculated the kurtosis of envelope spectrum amplitudes. It is logical to measure kurtosis in the frequency domain. When a bearing is healthy, its envelope spectrum is randomly distributed over that of the whole frequency.

In 2013, Dong Wang and Peter W. Tse [96] proposed an enhanced kurtogram, the major innovation of which is kurtosis values calculated based on the power spectrum of the envelope of the signals extracted from wavelet packet nodes at different depths. The power spectrum of the envelope of the signals defines the sparse representation of the signals and kurtosis measures the protrusion of the sparse representation. The enhanced kurtogram helps to determine the location of resonant frequency bands for further demodulation with envelope analysis. The frequency signatures of the envelope signal can then be used to determine the type of bearing fault.

3.6 Bispectrum analysis

3.6.1 Conventional bispectrum

Due to its inherent properties of nonlinearity detection and noise suppression, CB has been widely applied for suppressing noise and extracting the nonlinear fault information in the machinery fault detection and diagnosis areas. L. Saidi et al. [97] used CB to analyse the motor current signals for the fault detection of bearing defects, broken rotor bars, and their combined fault in induction machines. S. Guoji et al. [98] verified the prospects of CB for gearbox fault diagnosis by theoretical analysis of modulated vibration and practical vibration signals from a helicopter gearbox. H. Zhao et al. [99] applied CB to the vibration signals of engine crankshaft bearings for suppressing the interference of non-Gaussian noise. J. Huang et al. [100] implemented the feature extraction from vibration signals with CB for gearbox fault diagnosis. Y. Liu et al. [101] presented a method for gear pitting fault diagnosis using CB and 1(1/2)-dimension spectrum. Y. Xiao et al. [102] used CB to process the vibration signal of the diesel engine for diagnosis of the matching clearance between the piston pin and piston pin boss.

For a discrete time current signal $x(n)$, its Discrete Fourier Transform (DFT) $X(f)$ can be defined as[103][106]:

$$X(f)=\sum_{t=-\infty}^{+\infty} x(t)e^{-2j\pi ft} \quad \text{.....(3.10)}$$

In a complex number format, $X(f)$ can be rewritten in the format of magnitude $|X(f)|$ and phase $\phi(f)$:

$$X(f)=|X(f)|e^{j\phi(f)} \quad \text{.....(3.11)}$$

and the second-order measure (power spectrum) of $x(t)$, can be computed by the formula

$$P(f)=E\left\langle X(f)X^{*}(f)\right\rangle \quad \text{.....(3.12)}$$

where the complex $X^{*}(f)$ is a conjugate of $X(f)$ and $E\langle\ \rangle$ is the statistical expectation. The power spectrum is a linear transform and is a function of the frequency f. Extending this definition to the measures of order 3 gives rise to the conventional bispectrum $B(f_1, f_2)$ and it can be defined in the frequency domain as [107][108]:

$$B(f_1, f_2) = \mathrm{E}\left\langle X(f_1)X(f_2)X^*(f_1+f_2)\right\rangle \quad \text{.....(3.13)}$$

where f_1, f_2 and f_1+f_2 are the individual frequency components obtained from the Fourier transform integral. Note that second-order measures differ from third-order measures in that the latter are complex quantities containing both magnitude and phase information about the original time signal. Thus, the bispectrum detects the presence of quadratic phase coupling (QPC) [109].

If the frequency components at f_1, f_2 and f_1+f_2 are independent of each other, each frequency will be characterised by statistically independent random phases distributed over $(-\pi,\pi)$. Upon statistical averaging, denoting expectation operator by $E\langle\ \rangle$ in (3.13), the bispectrum will tend towards zero due to the random phase mixing effect. In this way, random noise can be suppressed significantly. On the other hand, if the three spectral components: f_1, f_2 and f_1+f_2 are non-linearly coupled to each other, the total phase of the three components will not be random at all, even though each of the individual phases is random. In particular, the phases will have the following relationship [108].

$$\phi(f_2)+\phi(f_1)-\phi(f_2+f_1) = const. \quad \text{.....(3.14)}$$

Consequently the statistical averaging will not lead to a zero value in the bispectrum. This non-linear coupling is indicated by a peak in the bispectrum at the bifrequency $B(f_1, f_2)$. To measure the degree of coupling between coupled components, a normalised form of the bispectrum or bicoherence is usually used and is defined as [107]:

$$b^2(f_1, f_2) = \frac{\left|B(f_1, f_2)\right|^2}{E\left\langle\left|X(f_1)X(f_2)\right|^2\right\rangle E\left\langle\left|X(f_1+f_2\right|^2\right\rangle} \quad \text{.....(3.15)}$$

3.6.2 Modulation signal bispectrum

In 2004, Jason R. Stack et al. [40] developed an amplitude modulation detector for detection of signal-point defects in rolling element bearings in an incipient stage. For the application of amplitude modulation signals, the conventional bispectrum only considers the higher sideband and neglects the lower sideband. Furthermore, in the bearing vibration acceleration signals, the carrier is the resonance frequency which has high amplitude, and the modulating frequency is the fault frequency of bearing which has very low amplitude compared to the carrier. The conventional bispectrum is not suitable for this kind of signal.

To solve this problem, Jason R. Stack et al. proposed a new AM detector, which is referred to as a modulation signal bispectrum in this thesis.

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 [107][108][110]. It also has been applied to analyse the vibration signals [56][104][105] for bearings and gearbox fault detection and diagnosis.

In the frequency domain, the MSB of a signal $x(t)$ can be defined in the form of the DFT as:

$$B_{MS}(f_c, f_x) = E\left\langle X(f_c + f_x)X(f_c - f_x)X^*(f_c)X^*(f_c)\right\rangle \quad \text{.....(3.16)}$$

where $B_{MS}(f_c, f_x)$ is the bispectrum of the signal $x(t)$, $E\langle\ \rangle$ is the expectation operator, f_x is the modulating frequency, f_c is the carrier frequency, $(f_c + f_x)$ 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 the 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.

The magnitude and phase of MSB can be expressed as Equations (3.17) and (3.18) respectively:

$$A_{MS}(f_c, f_x) = E\left\langle \left|X(f_c + f_x)\right|\left|X(f_c - f_x)\right|\left|X^*(f_c)\right|\left|X^*(f_c)\right|\right\rangle \quad \text{.....(3.17)}$$

$$\varphi_{MS}(f_c, f_x) = \varphi(f_c + f_x) + \varphi(f_c - f_x) - \varphi(f_c) - \varphi(f_c) \quad \text{.....(3.18)}$$

To measure the nonlinear effects of modulation signals accurately, both higher sideband $f_c + f_x$ and lower sideband $f_c - f_x$ are considered simultaneously in Equation (3.16). If they are due to the nonlinear effect between f_c and f_x, there will be a bispectral peak at bifrequency $B_{MS}(f_c, f_x)$. On the other hand, if these components are not coupled but have random distribution the magnitude of MSB will be close to nil. In this way it allows the

wideband noise in bearing vibration signals to be suppressed effectively so that the discrete components can be obtained more accurately.

To measure the degree of coupling between three components, a modulation signal bicoherence (MSBc) can be used and calculated by Equation (3.19).

$$b_{MS}^2(f_c, f_x) = \frac{\left|B_{MS}(f_c, f_x)\right|^2}{E\left\langle \left|X(f_c)X^*(f_c)\right|^2 \right\rangle E\left\langle \left|X(f_c+f_x)X(f_c-f_x)\right|^2 \right\rangle} \quad \text{.....(3.19)}$$

In addition, it can also be seen that for the case of $f_x = 0$, the MSB is degraded to power spectrum as shown in Equation (3.20).

$$PS(f_c) = \sqrt{B_{MS}(0, f_c)} = \sqrt{E\left\langle X(f_c)X^*(f_c)X(f_c)X^*(f_c)\right\rangle} \quad \text{.....(3.20)}$$

The MSB analysis is also suitable for harmonic components in vibration signals.

3.6.3 Phase relationship of MSB

In this section, the phase relationship of MSB is derived for the application to amplitude modulation signals and harmonic signals.

For an amplitude modulation signal

To produce an amplitude modulation signal, suppose that there are two signals with different amplitudes, angular frequencies and initial phases:

$$\begin{aligned} x_1(t) &= A_1\cos(\omega_1 t + \phi_1) \\ x_2(t) &= A_2\cos(\omega_2 t + \phi_2) \end{aligned} \quad \text{.....(3.21)}$$

where $\omega_2 > \omega_1$, $x_2(t)$ is the carrier signal. Then, the instantaneous phase of the two signals can be expressed as:

$$\begin{aligned} \varphi\left[x_1(t)\right] &= \omega_1 t + \phi_1 \\ \varphi\left[x_2(t)\right] &= \omega_2 t + \phi_2 \end{aligned} \quad \text{.....(3.22)}$$

The amplitude modulation signal of these two signals can be written as:

$$\begin{aligned} S_m(t) &= [1 + x_1(t)]x_2(t) \\ &= A_1A_2[1+\cos(\omega_1 t + \phi_1)]\cos(\omega_2 t + \phi_2) \\ &= A_1A_2\cos(\omega_2 t + \phi_2) + A_1A_2\cos(\omega_1 t + \phi_1)\cos(\omega_2 t + \phi_2) \end{aligned} \quad \text{.....(3.23)}$$

To expand the second term of Equation (3.23), the sum-to-product formula in Equation (3.24) will be applied.

$$\cos A\cos B = \frac{1}{2}[\cos(\mathrm{A}-\mathrm{B}) + \cos(\mathrm{A}+\mathrm{B})] \quad \text{.....(3.24)}$$

Then, the expression of the signal can be written as:

$$\begin{aligned} S_m(t) &= A_1 A_2 \cos(\omega_2 t + \phi_2) + \frac{A_1 A_2}{2}\left[\cos\left((\omega_2 - \omega_1)t + (\phi_2 - \phi_1)\right) + \cos\left((\omega_2 + \omega_1)t + (\phi_2 + \phi_1)\right)\right] \\ &= s_1(t) + s_2(t) + s_3(t) \end{aligned} \quad \text{.....(3.25)}$$

where

$$\begin{aligned} s_1(t) &= A_1 A_2 \cos(\omega_2 t + \phi_2) \\ s_2(t) &= \frac{A_1 A_2}{2}\cos\left((\omega_2 - w_1)t + (\phi_2 - \phi_1)\right) \\ s_3(t) &= \frac{A_1 A_2}{2}\cos\left((\omega_2 + \omega_1)t + (\phi_2 + \phi_1)\right) \end{aligned} \quad \text{.....(3.26)}$$

$s_1(t)$ is the carrier; $s_2(t)$ is the lower sideband and $s_3(t)$ is the upper sideband.

Therefore, the phase for the three components can be written as:

$$\begin{aligned} \varphi[s_1(t)] &= \omega_2 t + \phi_2 \\ \varphi[s_2(t)] &= (\omega_2 - \omega_1)t + (\phi_2 - \phi_1) \\ \varphi[s_3(t)] &= (\omega_2 + \omega_1)t + (\phi_2 + \phi_1) \end{aligned} \quad \text{.....(3.27)}$$

According to Equation (3.18), the phase of MSB can be expressed as:

$$\begin{aligned} \varphi_{MS}(f_c, f_x) &= \varphi(f_c + f_x) + \varphi(f_c - f_x) - \varphi(f_c) - \varphi(f_c) \\ &= \varphi[s_3(t)] + \varphi[s_2(t)] - \varphi[s_1(t)] - \varphi[s_1(t)] \\ &= [(\omega_2 + \omega_1)t + (\phi_2 + \phi_1)] + [(\omega_2 - \omega_1)t + (\phi_2 - \phi_1)] - [\omega_2 t + \phi_2] - [\omega_2 t + \phi_2] \\ &= 0 \end{aligned} \quad \text{....(3.28)}$$

Hence, the MSB phase of an amplitude modulation signals is always zero.

For a harmonic signal

If only the first three components of a harmonic signal are considered, they are defined as:

$$\begin{aligned} y_1(t) &= A_1 \cos(\omega_1 t + \phi_1) \\ y_2(t) &= A_2 \cos(\omega_2 t + \phi_2) \\ y_3(t) &= A_3 \cos(\omega_3 t + \phi_3) \end{aligned} \quad \text{.....(3.29)}$$

where A_1, A_2, and A_3 denote their amplitudes; ω_1, ω_2, and ω_3 denote their frequencies, ϕ_1, ϕ_2, and ϕ_3 denote their initial phases, respectively. Furthermore, their frequencies have the relationship $\omega_2 = 2\omega_1, \omega_3 = 3\omega_1$, then $2\omega_2 = \omega_1 + \omega_3$.

The harmonic signal can be expressed as

$$\begin{aligned} S_h(t) &= y_1(t) + y_2(t) + y_3(t) \\ &= A_1\cos(\omega_1 t + \phi_1) + A_2\cos(\omega_2 t + \phi_2) + A_3\cos(\omega_3 t + \phi_3) \end{aligned} \quad \text{.....(3.30)}$$

According to Equation (3.18), the phase of MSB can be expressed as

$$\begin{aligned} \varphi_{MS}(\omega_1, \omega_2, \omega_3) &= \varphi(f_c + f_x) + \varphi(f_c - f_x) - \varphi(f_c) - \varphi(f_c) \\ &= \varphi(y_3) + \varphi(y_1) - \varphi(y_2) - \varphi(y_2) \\ &= (\omega_3 t + \phi_3) + (\omega_1 t + \phi_1) - (\omega_2 t + \phi_2) - (\omega_2 t + \phi_2) \\ &= \phi_1 + \phi_3 - 2\phi_2 \\ &= const. \end{aligned} \quad \text{.....(3.31)}$$

Hence, the MSB phase of a harmonic signal is a constant, which means MSB is also suitable for analysing harmonic signals while the angular frequencies of the harmonic components meet the condition $\omega_2 - \omega_1 = \omega_3 - \omega_2$. It means the MSB can be extended to higher harmonic components.

3.7 Summary

The chapter reviews the fundamentals of key signal processing methods that will be further studied in this thesis. TFR, wavelet based analysis and their derived kurtogram analysis are particularly effective for enhancing impulsive components but they need considerable optimisation for achieving an optimal analysis. Nevertheless, they do not have the mechanism of noise reduction; rather, they just highlight the signals of interest.

ALE has the capability of noise reduction and can be implemented with relatively simple iterative approaches in real-time. However, its noise reduction performance can be limited by the instability of existing methods.

TSA is the most effective method for noise reduction and periodic signal enhancement. The major drawback is that it needs additional channels for collecting the angular reference signals and sensor installation is often restricted in practice because of the connection with rotating components.

Bispectrum analysis including both CB and MSB achieves noise reduction through signal phase alignments without the need for the reference signals. In addition, bispectrum has the capability of nonlinear characterisation and leads to the sparse representation of complicated signals for more reliable feature extraction.

These understandings of information enhancement mechanisms provide a concrete basis for developing more effective methods to process the signals from various cases in this study. Details of development will be presented in Chapter 4-Chapter 9.

Chapter 4
Impulse enhancement for a two-stage helical gearbox fault diagnosis

The periodic impulsive component is one of the typical fault characteristics in vibration signals from gearboxes with local tooth defects. However, this component exhibits very small amplitude in the early stage of the fault and is often masked by various noises, which makes it difficult to extract accurately for fault detection. In the gearbox vibration signals, the two main types of noise are gear mesh components and random noise. Gear mesh components are inherent noise generated by the machinery's running process and random noise is inevitably induced by the measurement system. In this chapter, to enhance the impulses for reliable fault diagnosis especially for incipient fault, TSA and ALE are applied to reduce the random noise and the periodic gear mesh components.

4.1 Introduction

Impulsive vibration signals in machinery are often caused by component impacts which are commonly associated with component faults. It has long been recognised that the presence of a fault is often indicated by the presence or increase in impulsive components in measured signals. However, it tends to be difficult to make objective measurements of impulsive components due to the existence of high levels of background noise. The detection of these impulsive components is hampered by the presence of the signals associated with the normal running of the machine, with the consequence that the detection of the weak impulsive components, which are especially associated with incipient faults, is difficult [71]. It is the 'normal' signals which form the background noise environment against which the detection of fault induced impulsive components must be conducted. To improve the precision of fault diagnosis, it is valuable to enhance the impulsive components by suppressing this background noise prior to further processing.

De-noising and extraction of such faulty signals are very important for fault diagnostics, especially for early fault detection, in which the fault features are often very weak and embedded in noise. Therefore, it is necessary to enhance the data reliability and improve the accuracy of the signal analysis. After the successful pre-processing, the signal has an increased SNR, making it more amenable to one of a gamut of signal processing tools that can characterise the signal, such as Auto-Regressive (AR) modelling, kurtosis evaluation, cepstrum analysis, time-frequency analysis and higher order spectra analysis [111].

The ALE was introduced as a method for detecting a periodic signal in an incoherent background or conversely for removing periodic interference from a broadband signal of interest [112]. Naoto Sasaoka etc. applied ALE to reduce the sinusoidal noise in noisy speech signal [113]. J. R. Mohammed etc. presented a noise reduction system based on two stages of operation with the first stage based on the ALE filters and the second stage on NLMS (Normalised Least Mean Square) filter. The first stage reduces the sinusoidal noise from the input signal and the second stage reduces the wideband noise [114]. S. K. Lee and P. R. White exploit two ALE filters in series to reduce the level of background noise. The resulting enhanced signals are analysed in the time-frequency domain to obtain simultaneous spectral and temporal information. The technique has been successfully applied to the vibration signals from an internal combustion engine and an industrial gearbox [111].

However, the scheme proposed by S. K. Lee and P. R. White applies two ALE filters which have different structures make it more complex and has limited capability in reducing random noise. Therefore, TSA is used to suppress the random noise in vibration signals which has been widely applied to gearbox vibration signals. In this chapter, a signal processing scheme which combines TSA and ALE is proposed to reduce the random noise induced by the measurement process and gear mesh components respectively. In such a way, the impulses containing fault information could be highlighted to produce more reliable detection and diagnosis results. Because TSA has been introduced in Section 3.4, the principle and parameter selection method for ALE are described in details in this chapter.

4.2 Signal processing method

4.2.1 Denoising scheme

The principle of the ALE [115] used to reduce sinusoids will be described in this section based on the block diagram given in Figure 3.1. The primary input and reference input signals of the system are given by

$$d(n) = s(n) + \eta(n) \quad \text{.....(4.1)}$$

$$x(n) = d(n-\Delta) = s(n-\Delta) + \eta(n-\Delta) \quad \text{.....(4.2)}$$

where $s(n)$ is the fault impulse signal, $\eta(n)$ represents the sinusoidal noise and Δ is the time delay factor. Due to the short time duration of impulses, the autocorrelation of impulses fades as Δ increases. On the other hand, the delayed sinusoidal noise $\eta(n-\Delta)$ is correlated with $\eta(n)$. Thus, when the adaptive filter is stable, $\eta(n)$ may be estimated by the stationary output of adaptive filter $y_p(n)$. The impulsive components in ALE non-stationary output $y_f(n)$ are enhanced compared with the input signal $d(n)$.

$$\begin{aligned} y_f(n) &= d(n) - y_p(n) \\ &= d(n) - \hat{\eta}(n) \\ &= s(n) + \eta(n) - \hat{\eta}(n) \end{aligned} \quad \text{.....(4.3)}$$

The cost function of the NLMS algorithm is defined as:

$$\begin{aligned} E\left[y_f^2(n)\right] &= E\left[d(n)-y_p(n)\right]^2 \\ &= E\left[s(n)+\eta(n)-y_p(n)\right]^2 \\ &= E\left[s(n)\right]^2 + 2E\left[s(n)(\eta(n)-y_p(n))\right] \\ &\quad + E\left[\eta(n)-y_p(n)\right]^2 \end{aligned} \quad \text{.....(4.4)}$$

Since the fault signal is uncorrelated with noise, $E\left[s(n)(\eta(n)-y_p(n))\right]=0$. Then, Equation (4.4) becomes:

$$E\left[y_f^2(n)\right] \approx E\left[s(n)\right]^2 + E\left[\eta(n)-y_p(n)\right]^2 \quad \text{.....(4.5)}$$

Minimising $E[y_f^2(n)]$ is equivalent to minimising $E[\eta(n)-y_p(n)]^2$. Thereafter, the convergence of the algorithm will ensure $y_p(n)$ achieves the minimum mean-square estimation of $\eta(n)$. The signal is separately into two parts: $y_f(n)$ contains mainly non-stationary components (*e.g.* impulses) and $y_p(n)$ contains mainly stationary components (*e.g.* gear mesh frequencies). The fault impulses remain in $y_f(n)$, which is used for the following feature extraction.

There are three important parameters in ALE: the step size μ , the length of filter and the number of delay samples. μ is the step size parameter controlling the convergence rate within its suitable range. The step size value affects the convergence behaviour of an LMS filter; too small a value of μ leads to extremely long convergence time of the algorithm, whereas too large a value of μ causes the algorithm to diverge, thus degrading the performance of the adaptive filter. Therefore, choosing a suitable value for the step size is necessary when implementing the LMS algorithm as an adaptive filter.

The periodic nature of the impulsive component lays open the possibility that the first stage of the scheme will identify them with the narrowband components and in doing so attenuate them. To avoid such an eventuality, care over the choice of the parameters Δ and L must be exercised. From reference [114], the length of the filter must satisfy the condition in Equation (4.6), where T_p is the period of the signal.

$$L < T_p - 2\Delta \quad \text{.....(4.6)}$$

4.2.2 Parameters selection method

In this part, the selection method of ALE parameters applied in this chapter is described. Kurtosis is commonly used in engineering for the detection of fault symptoms because it is sensitive to sharp variant structures, such as impulses. The bigger the impulse in a signal, the larger the kurtosis is. The kurtosis value comparison of signals after the ALE processing based on different parameter sets is shown in Figure 4.1. The green line indicates the kurtosis value of the signal before ALE processing. Different line colours and style represent different μ values, and L indicates the length of the adaptive filter.

From Figure 4.1, it is obvious that the kurtosis values are increased after ALE. To compare the impulse enhancement effects, the parameter sets corresponding to the maximum kurtosis under each L are selected for comparison. Then, the signals are analysed in the time-frequency domain.

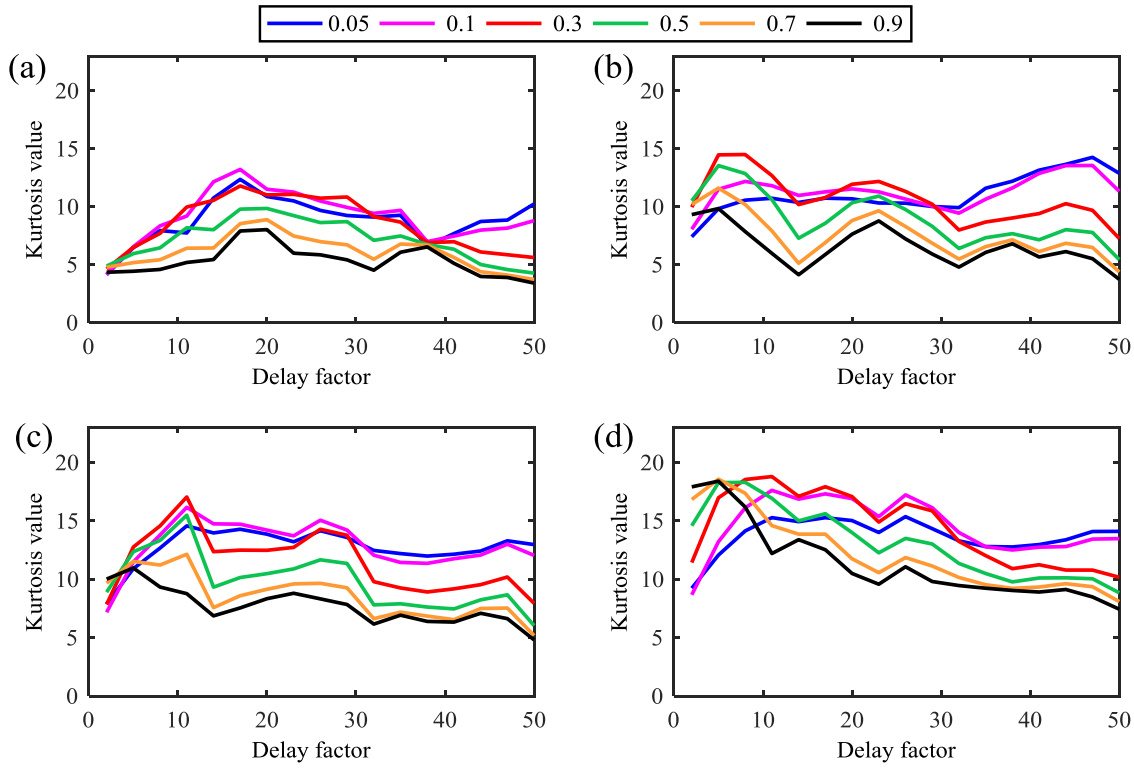

Figure 4.1 Kurtosis values under various parameters

(a) L=16; (b) L=32; (c) L=64; and (d) L=128

Figure 4.2 illustrates the time-frequency analysis comparison between the signal and signals after ALE with four different parameter sets. Figure 4.2(a) shows the raw signal and Figure 4.2(b)-(e) presents the ALE results while the parameters are selected by maximum kurtosis criteria in Figure 4.1 with the four filter lengths 16, 32, 64 and 128. It

can be seen that the main periodic components at 1500Hz and 2500Hz are reduced greatly after ALE and the impulsive components, which have very small amplitudes can be observed clearly. It is obvious that the result in Figure 4.2(e), which corresponds to the maximum kurtosis value parameter set in Figure 4.1, can reveal the impulsive components much better than the other three parameter sets. Thus, the kurtosis maximising criterion is applied to choose the ALE parameters in the following machine fault diagnosis.

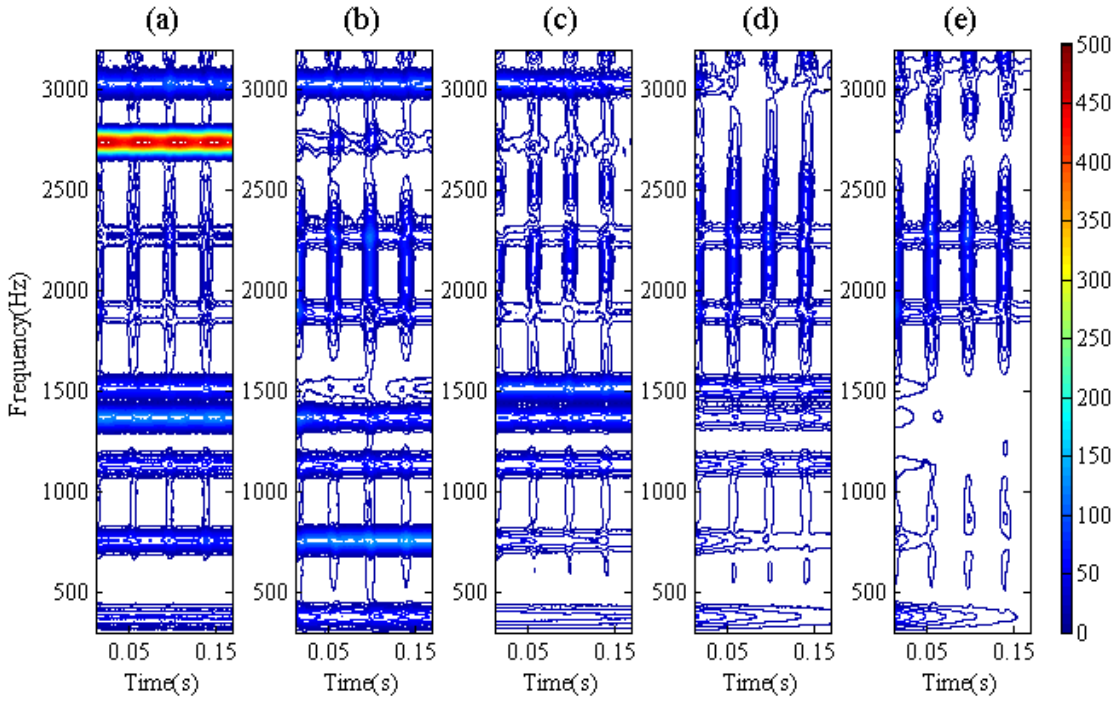

Figure 4.2 Time-frequency analysis results

(a) TFR of signal; (b) TFR of ALE output (L=16); (c) TFR of ALE output (L=16); (d) TFR of ALE output (L=32); and (e) TFR of ALE output (L=64).

4.3 Experimental setup

This experimental verification is based on a test system consisting of a 3-phase induction motor, two permanent helical gearboxes connected back to back, couplings, load devices, and a resistor bank, as illustrated in Figure 4.3, of which the main subsystems are detailed in the schematic diagram in Figure 4.4.

The gearbox under testing is a two-stage helical gearbox, detailed in Table 4.1 and Figure 4.5, which is a standard industrial gearbox used for the power transmission of different machines such as compressors and pumps. Using such a commercial gearbox allows signals to be more realistic and hence the verification can be more reliable.

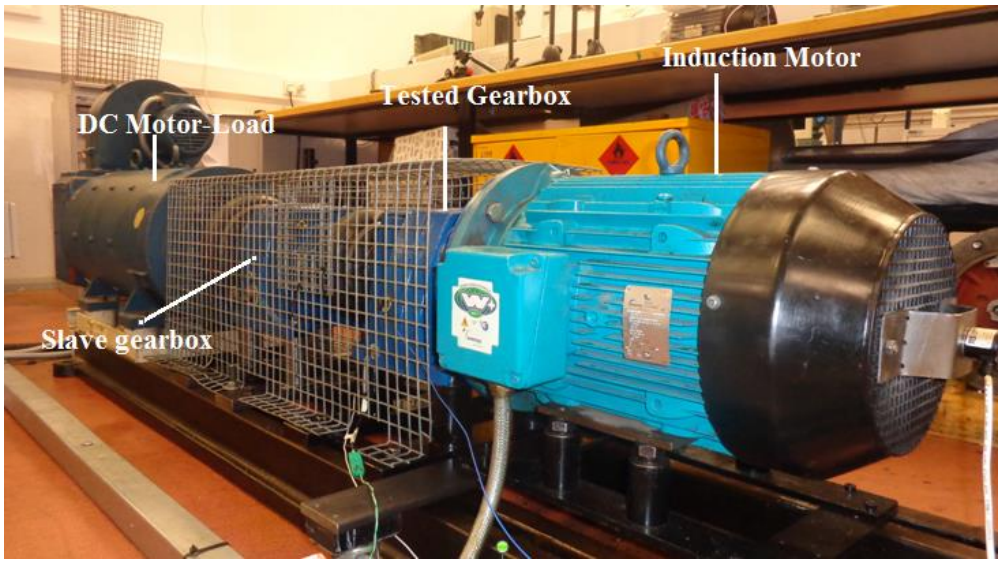

Figure 4.3 Gearbox test rig used in this research

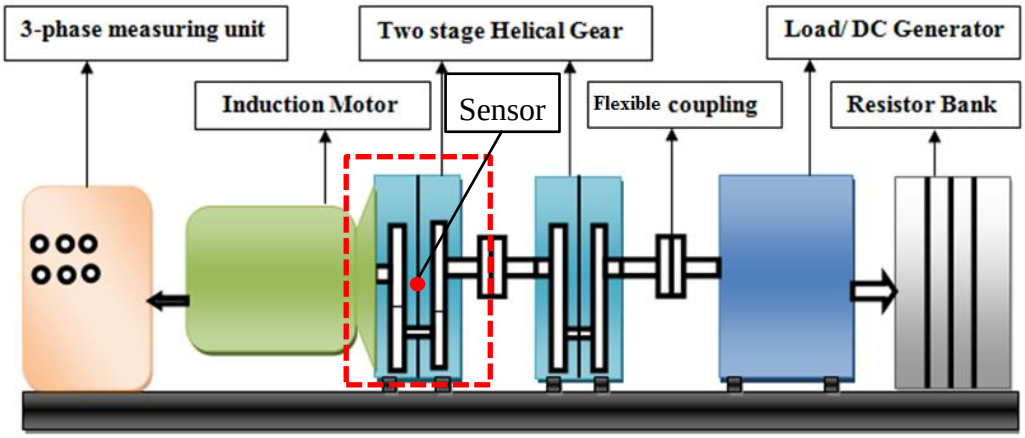

Figure 4.4 Schematic layout of the experimental gearbox test rig

Table 4.1 Gearbox specification

Model: M07223.6BMCE1A11.A	Power capacity: 10kW	Speed ratio: 3.667
Description	1st stage PG0740.8/M07E	2nd Stage M07-24.5B-C
Tooth number	Z1/Z2=58/47	Z3/Z4=13/59
Speeds of shaft	24.42 Hz (input) at full load	6.64 Hz (output) at full load
Meshing frequency	$f_{m1} = 1416Hz$	$f_{m2} = 391Hz$

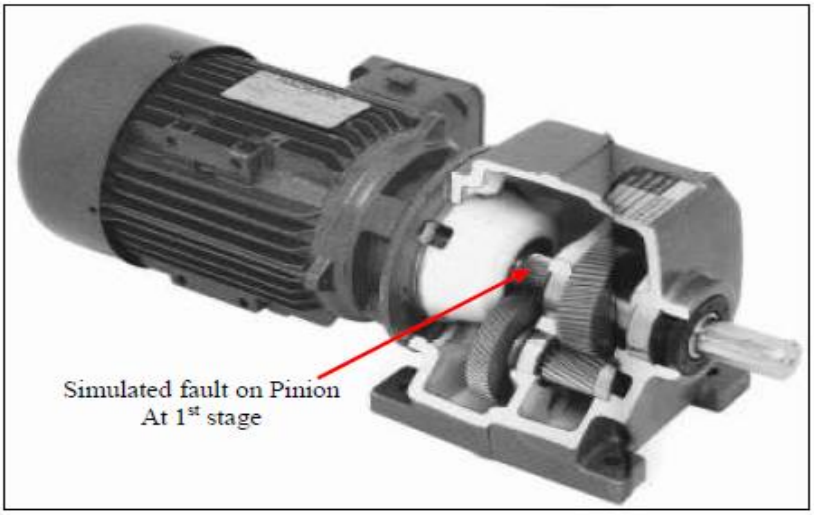

Figure 4.5 The illustrative construction of testing gearbox

Three fault cases (30%, 60%, and 100% tooth damage) were simulated on the pinion gear at the first stage. The first is a tooth breakage of 30% on one of the 58 teeth of the pinion gear; the second is 60% tooth breakage; and the third is 100% tooth breakage on the same teeth. The 30% and 100% tooth damage gears are illustrated in Figure 4.6.

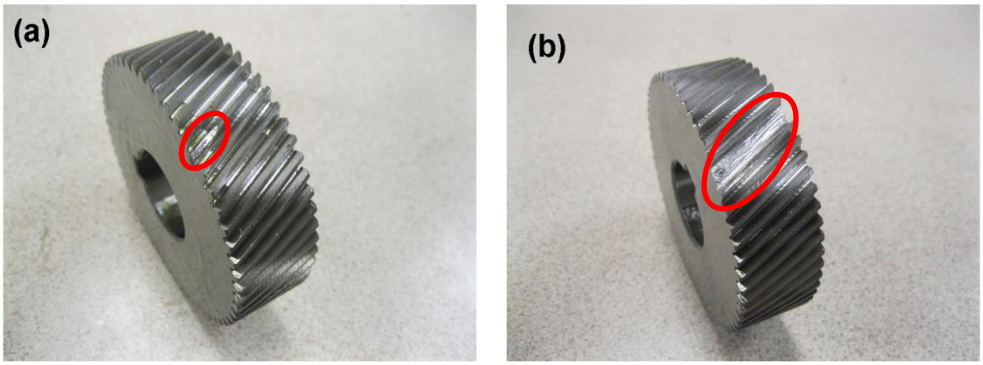

Figure 4.6 Gear faults: (a) 30% tooth damage and (b) 100% tooth damage

The position of the vibration sensor is illustrated in Figure 4.4. Vibration signals were measured by a type PCB 338C04 accelerometer with a sensitivity of 100mv/g, and the frequency response range is from 0 Hz to 20 kHz. It is mounted in the horizontal direction of the gearbox housing, as shown in Figure 4.4. An encoder was installed at the end of the motor shaft to measure the instantaneous angular speed and record the initial phase of the input gear per revolution.

Vibration signals from an accelerometer were collected from the same gearbox in which the two broken gears were tested separately, after a baseline signal was collected. All of the tests were performed under an operating condition of 90% load. The data was sampled

at a rate of 100 kHz, at which the narrow pulse signal from the shaft encoder was also acquired simultaneously for implementing TSA.

4.4 Results and discussion

The practical gearbox experimental vibration signals were collected when the gearbox operated under five different loads and four conditions (Healthy, 30% tooth breakage, 60% tooth breakage and 100% tooth breakage).

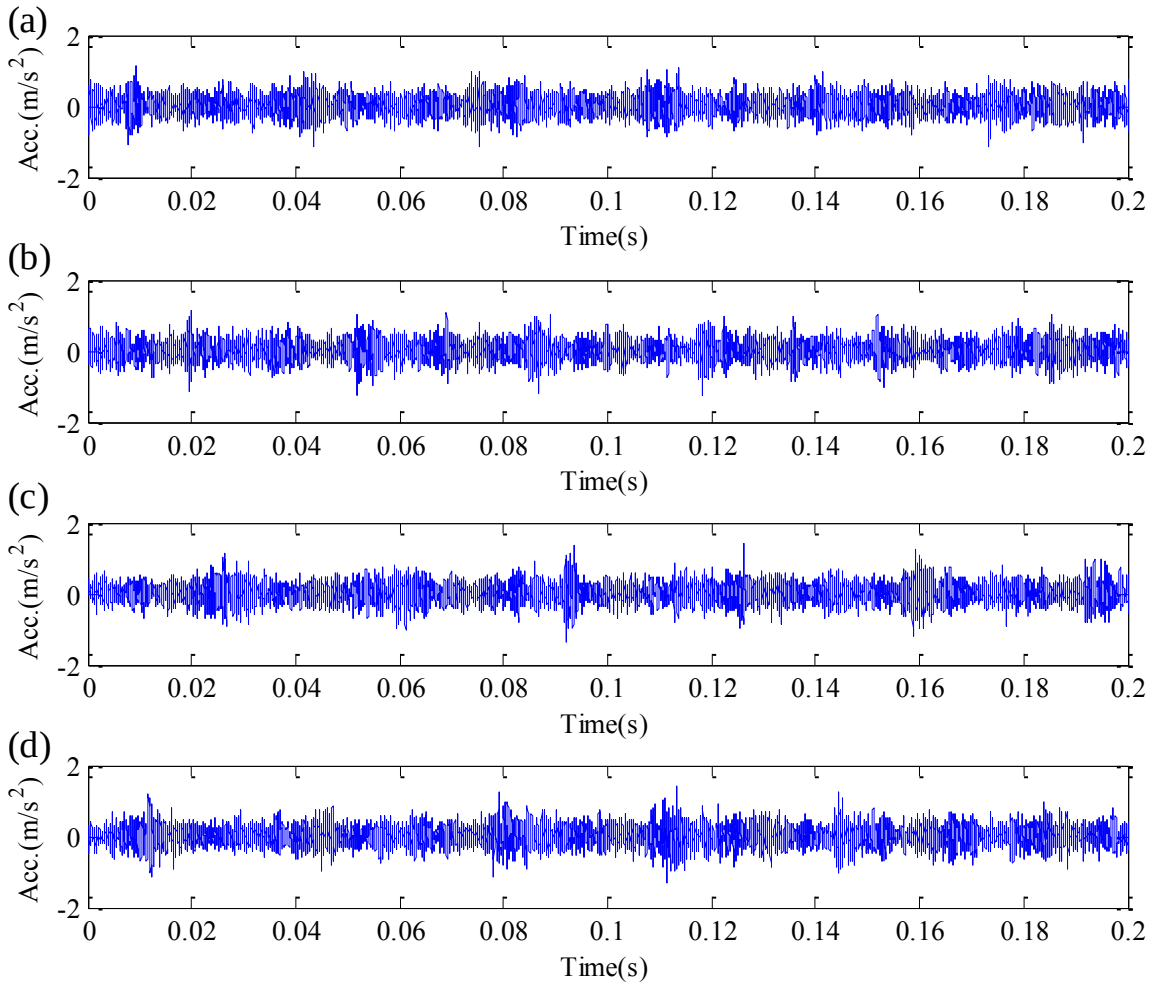

Figure 4.7 Raw vibration signals in time domain for (a) baseline; (b) 30% damage; (c) 60% damage; and (d) 100% damage

Figure 4.7 gives out the raw vibration signals for four cases under the 3rd load and the baseline signal is collected under healthy condition. Figure 4.8 shows the TSA results of the vibration signals given in Figure 4.7 for five revolutions of the input shaft. It can be seen that the impulsive components in the vibration signals are highlighted for all of the test conditions. TSA signals show a much clearer indication of the 60% and 100% tooth

damage compared with the baseline and 30% tooth damage. However, the signals between the baseline and 30% tooth damage cannot be observed with noticeable differences.

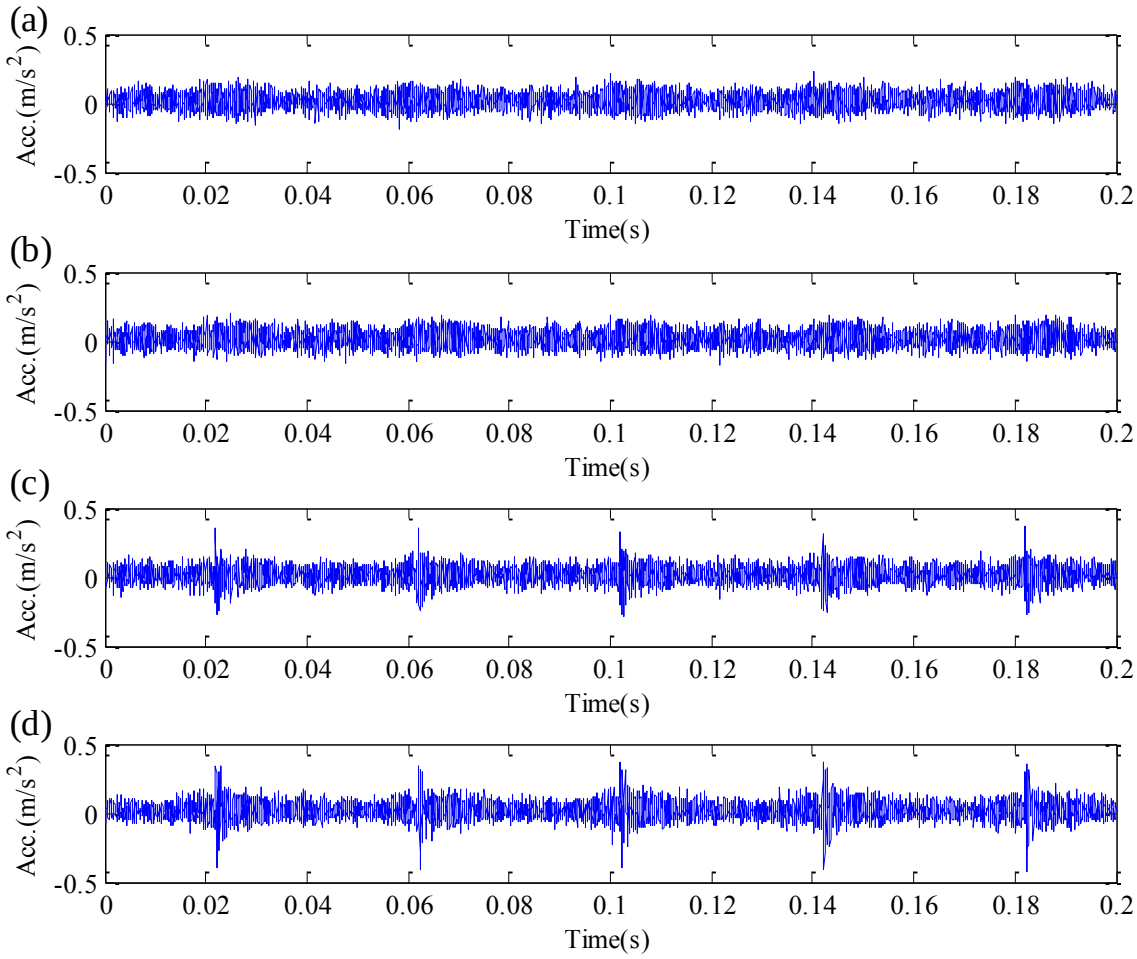

Figure 4.8 TSA vibration signals in time domain for (a) Baseline; (b) 30% damage; (c) 60% damage; and (d) 100% damage

To evaluate the denoise effect of the ALE algorithm, Figure 4.9 and Figure 4.10 present the time-frequency analysis results for comparison. Taking the signals under the 4th load as an example, signals under other loads have similar performance. As shown in Figure 4.9, there are several periodic meshing frequencies generated by rotation of the machine at about 1500Hz, 3000Hz and 4500Hz and can be treated as noise for fault detection. The purpose of ALE is to reduce these frequencies. The TSA signals also contain impulsive components which are of short duration and periodic contents for certain frequencies. These components carry fault information of gearbox tooth breakage. Figure 4.10 shows the results after the ALE processing. It can be seen that the mesh components are reduced significantly. The results show that ALE is effective in reducing periodic components and highlights impulses by reducing the gear mesh components.

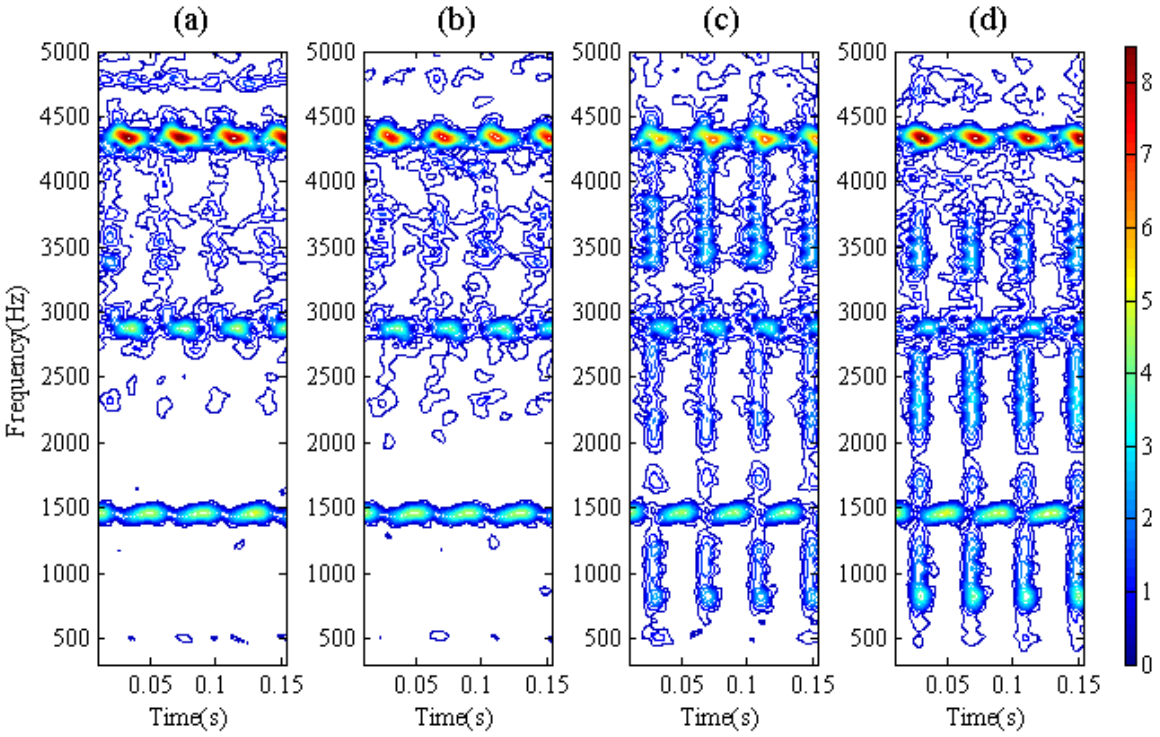

Figure 4.9 TSA vibration signals in time-frequency domain for (a) Baseline; (b) 30% damage; (c) 60% damage; and (d) 100% damage

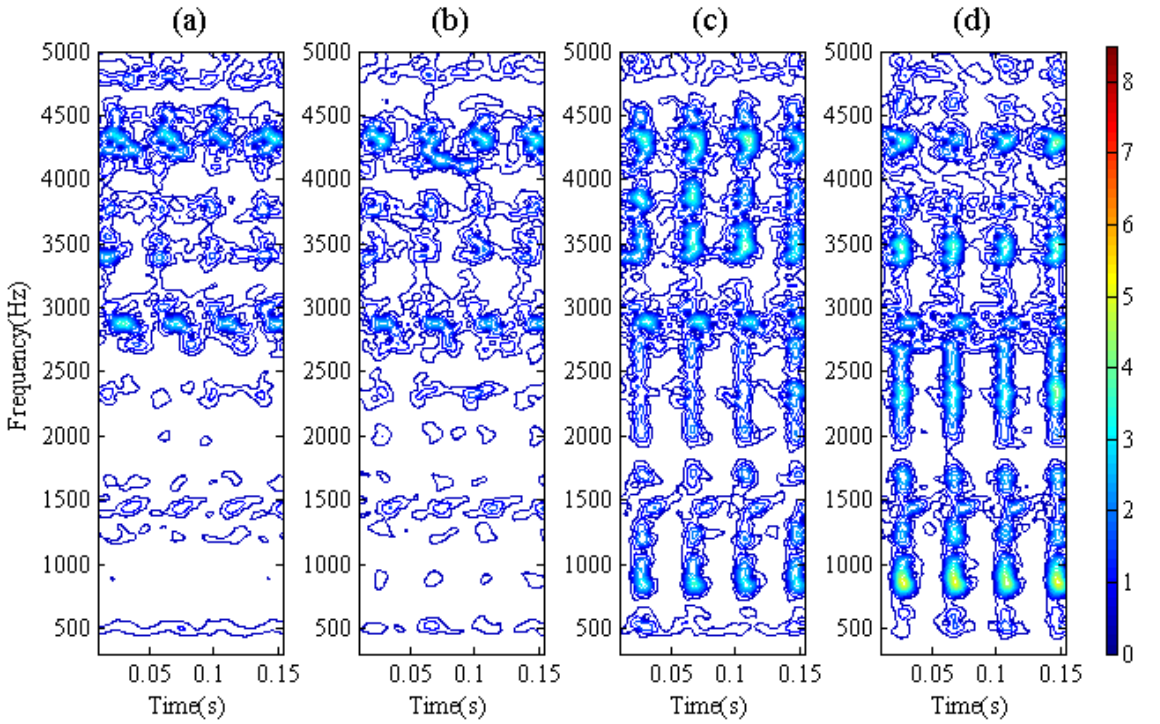

Figure 4.10 Signals after ALE in time-frequency domain for (a) Baseline; (b) 30% damage; (c) 60% damage; and (d) 100% damage

For a detailed comparison, two common feature parameters (root mean square (RMS) and kurtosis) are calculated from the TSA vibration signals and the ALE signals. As shown in

Figure 4.11(a), RMS is not able to separate the four cases under different loads. From Figure 4.11(b), it can be seen that the kurtosis values of 60% tooth damage and 100% tooth damage have been separated from the other two cases except for the first load. However, the difference between the baseline and 30% tooth damage is not very obvious for fault separation. Figure 4.11(c) and Figure 4.11(d) illustrate the RMS and kurtosis value comparison from the signal after the ALE denoising processing. From the RMS value comparison of the ALE output signal, there are differences under all loads except for the first and all four condition cases, which gives a clear difference compared with the results of the TSA signal. The kurtosis value of the ALE signal is much higher than that of the TSA signal, which means the impulsive components are better enhanced. But the difference between the kurtosis values of different condition cases is not as clear as the RMS value of ALE signals. Figure 4.12 illustrates the results obtained from another set of ALE parameters; the results for 30% tooth damage are worse, while for higher severity faults the results are better.

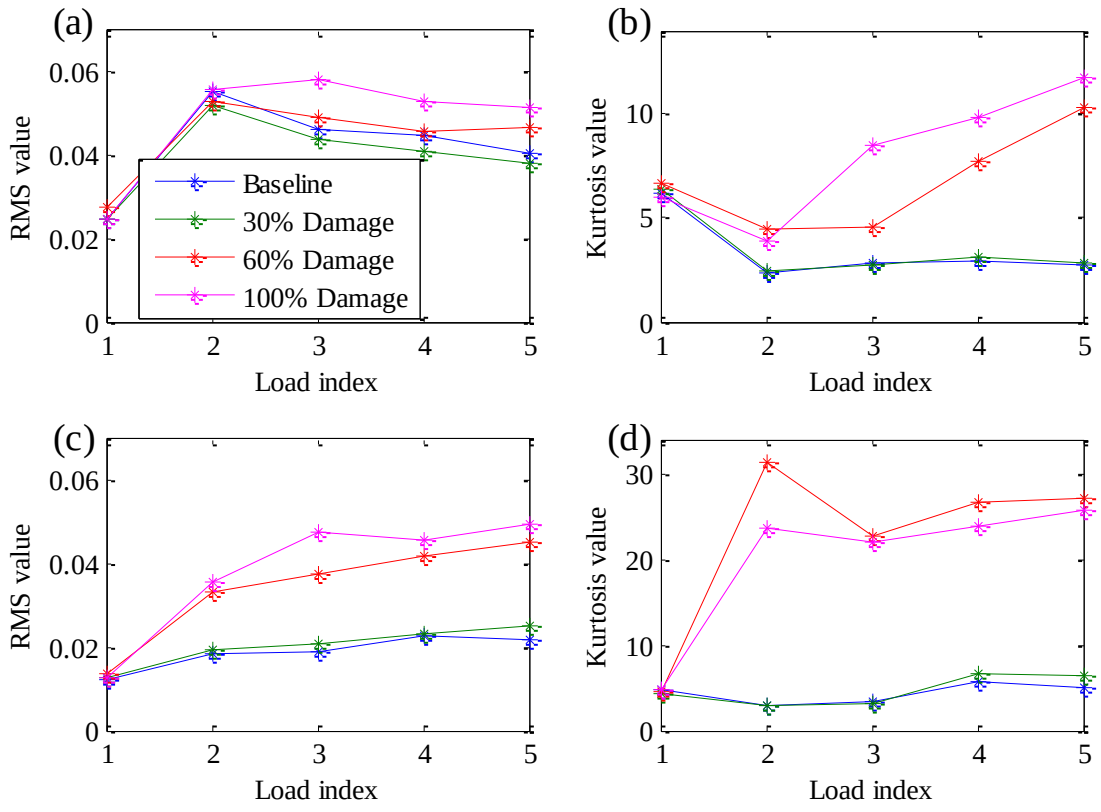

Figure 4.11 RMS and kurtosis value comparison of TSA and ALE signal (L=20, μ =0.075, Δ=12)

(a) RMS of TSA signal; (b) Kurtosis of TSA signal; (c) RMS of ALE output; and (d) Kurtosis of ALE output

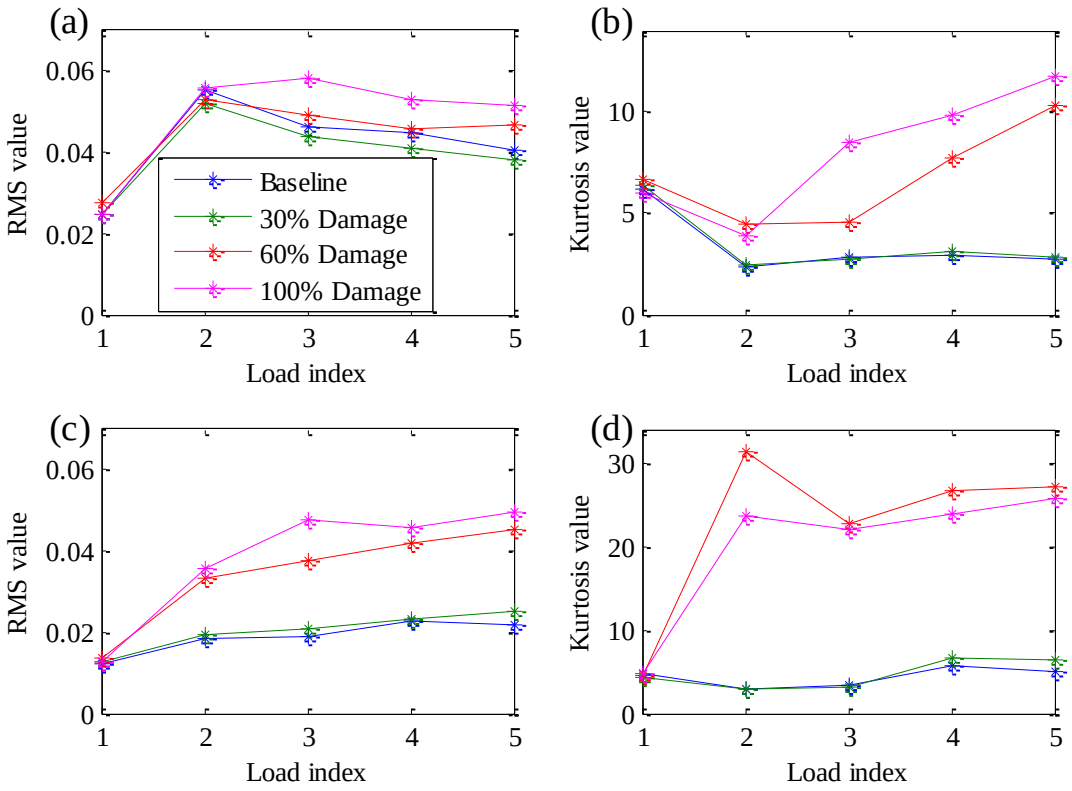

Figure 4.12 RMS and kurtosis value comparison of TSA and ALE signal (L=32, μ =0.0375, Δ=60)

(a) RMS of TSA signal; (b) Kurtosis of TSA signal; (c) RMS of ALE output; and (d) Kurtosis of ALE output

4.5 Summary

In this chapter, ALE has been examined for the tooth breakage detection and diagnosis of a two-stage helical gearbox based on noisy vibration signals. Firstly, TSA is used to suppress the random noise, and then ALE is applied to reduce the inherent periodic mesh components, at last TFR further highlights the impulsive fault features in the time-frequency domain. In this way, the incipient tooth breakage fault can be detected successfully. The results obtained from simulating and experimental vibration signals of the two-stage helical gearbox have confirmed that the ALE method is effective in reducing the periodic gear mesh components and thereby enhances the impulsive fault features in noisy vibration signals. The results show a clear difference even between the baseline and a small percentage (30%) tooth breakage of a helical gear.

Chapter 5
Gear fault diagnosis based on an optimised wavelet analysis

Motivated by the mechanism adopted in ALE, a Shannon entropy difference based method is proposed to select an optimal wavelet scale for detecting gear faults. The results obtained based on vibration signals from the two-stage helical gearbox have verified that the proposed method is effective for extracting the small impulsive component to show a clear difference between the baseline and a small fault i.e. 30% tooth breakage of a helical gear. Similar to Chapter 4, TSA is applied for the signal pre-processing before the wavelet analysis.

5.1 Introduction

In machinery fault diagnosis, impulses often appear in vibration signals, which means localised fault occurs. The behaviour of a impulsive signal is non-stationary, changing according to time, so it is important to analyse time and frequency information simultaneously.

Wavelet transform is capable of providing time-domain and frequency-domain information simultaneously. Similar to a wavelet function, the impulsive feature components of vibration signals have local energy distributions in both time domain and frequency domain. However, because there are different types of wavelets each of which has different time-frequency structures, it is obvious that using an optimal wavelet, whose time-frequency structure matches that of the impulsive component best, can achieve the best performance in the impulsive component detection.

As reviewed in Section 3.3, wavelet analysis has been successfully applied to vibration signals from bearings and gearboxes. Many parameter optimisation schemes have been proposed, such as minimum wavelet entropy, maximum kurtosis, genetic algorithm, etc. Shannon entropy in the time domain can measure signal or system uncertainty. Spectrum entropy based on Shannon entropy can be taken as a measure of signal or system complexity. Wavelet entropy measures obtained based on wavelet analysis can signify the complexity of unsteady signal or system in both time domain and frequency domain, because it is a method based on time-frequency analysis, featured with multi-resolution analysis [117]. The appearance of impulsive components in a signal will lead to the reduction of entropy value. Therefore, in the wavelet spectrum, the entropy of coefficients has lower value at the scale where the impulses account for a larger proportion. The purpose this chapter is to optimise the wavelet parameters for the analysis of the fault feature impulse component of the vibration signal by a maximum Shannon entropy difference method, and applying it to the small tooth break defect detection of a gear box.

5.2 Signal processing method

5.2.1 Characteristics of Morlet wavelet transform

The definition of the real Morlet wavelet in the time domain and its Fourier spectrum have been introduced in Section 3.3. The shape of the mother wavelet is controlled by f_b and f_c. In particular, f_c is related to the oscillation frequency, whereas f_b controls the decay

rate of the exponential envelope in the time domain. The trade-off of two parameters will thus regulate the time and frequency resolution of wavelet analysis. The frequency resolution will increase with the increase of f_b, whereas time resolution will decrease. When f_b tends to be 0, the Morlet wavelet becomes a Dirac function with the finest time resolution. In contrast, when f_b tends to be infinity, the Morlet wavelet degrades to a cosine function which has the finest frequency resolution. The central frequency f_c controls the oscillatory frequency of the Morlet wavelet. When f_c is larger, the oscillatory frequency of the Morlet wavelet is higher and will lead to an increase in the frequency resolution of the Morlet wavelet.

If a daughter wavelet is considered as a filter, the wavelet transform of Equation (1.9) is actually a filter bank [84]. Choosing a suitable combination of bandwidth f_b and centre frequency f_c will lead to the optimal result that represents the key characteristics of the signals.

The faults localised on one tooth or a few teeth such as cracks and spalls produce the modulation effects during the engagement of the fault teeth. Consequently, a large number of sidebands of the tooth-meshing frequency and its harmonics in the spectrum is generated and spread over a wide frequency range. The sidebands are spaced by the rotation frequency of the fault gear or the gear shaft and characterised by low amplitudes [120]. To show the key characteristics of the Morlet wavelet transform and the influences of its parameters, a simulated signal is synthesised based on the characteristics of gear vibration signals, as described in Section 2.2.2.

Figure 5.1(a)-(c) show contour plots of the CWT coefficient of the simulated signal with three sets of different wavelet parameters. It is quite clear that the repetitive spikes, which are illustrated by straight patches across a wide number of scales, can be observed for all sets of the parameters used. This demonstrated that the Morlet wavelet has superb performance in enhancing the feature from the small impulses, allowing the easy separation of the small impulse and hence the high capability of fault detection.

However, a careful study of these three results has found that there are slight differences between different parameter sets. For smaller f_c CWT of Figure 5.1(a) shows as being coarser in scale direction compared with that of Figure 5.1(b) obtained from a higher f_c

value. For a larger f_b the scale resolution is further improved, as shown in Figure 5.1(c) when the f_b is 7. However, this higher f_b reduces the performance in discriminating the impulse in the time direction.

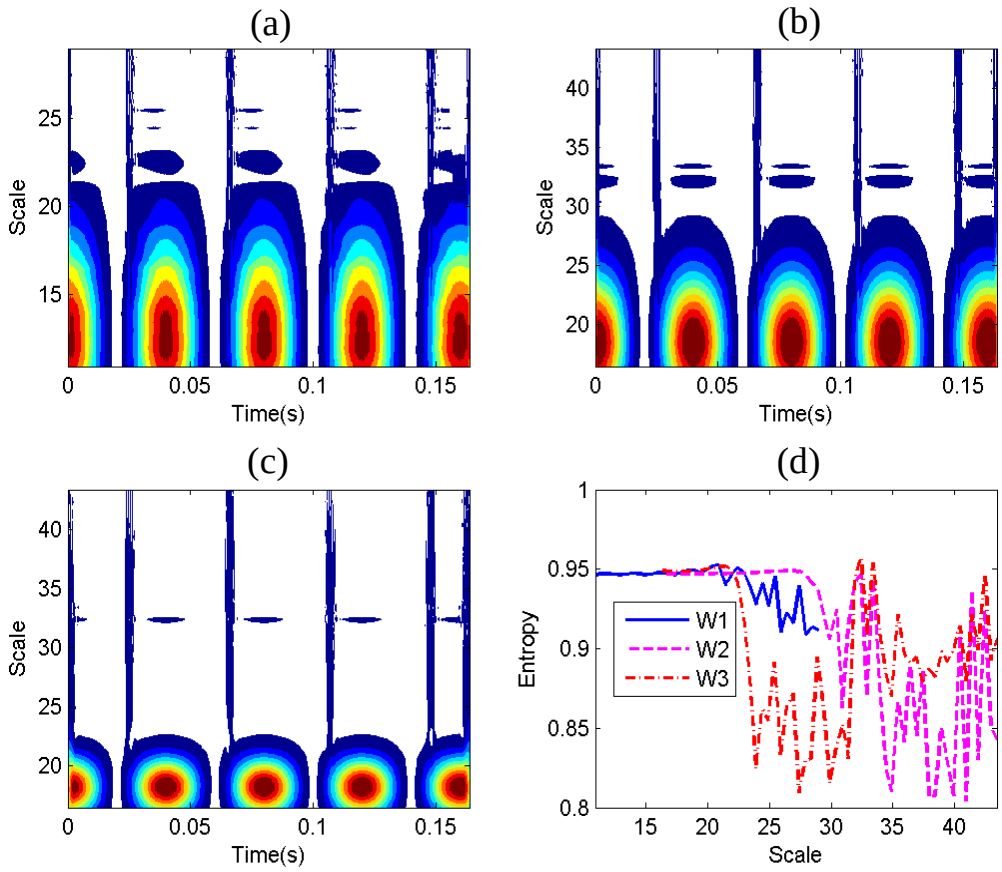

Figure 5.1 Wavelet coefficients of simulated signal using different parameters (a) W1:fc=0.7, fb=1; (b) W2: fc=1.05, fb=1; (c) W3: fc=1, fb=6; and (d) Entropy vs. scales

To give an accurate measure of the performance in detecting these impulses, the Shannon wavelet entropy, which reflects the sparsity of the wavelet transform performance and used in [82][83], is calculated at different scales by:

$$En(a_i) = -\sum_{j=1}^{N} p_j(a_i) \log p_j(a_i) \quad \text{.....(5.1)}$$

and the probability density of the wavelet coefficient along the time direction at a given scale a_i is:

$$p_j(a_i) = WT(a_i, t_j) / \sum_{j=1}^{N} WT(a_i, t_j) \quad \text{.....(5.2)}$$

where $\{WT(a_i, t_j)\}_{j=1,...N}$ is the class of the coefficients at scale $a_i, i = 1,...,M$; and N is the sample number along the time direction. Figure 5.1(d) shows the entropy values over different scales for the three wavelet parameter sets. It can be seen that entropy values are high at low scales for all parameters set, showing that these scales give low features for the impulse signals, which are clearly indicated by the high levels and crowded contours in all three plots. On the other hand, the entropy amplitude becomes lower in high scale ranges, showing closer consistency with that in the contour in presenting the discrimination performance of the impulses.

Moreover, the minimum values of entropy are different for three different parameter sets. The parameter set: $f_c = 1.05$ and $f_b = 1$ gives the minimal entropy, showing it is a better selection of wavelet parameters. Furthermore, this minimum corresponds to a high scale value. This is very close to the frequency content of the impulses induced in the simulated signals.

5.2.2 Optimisation of Morlet wavelet by maximising wavelet entropy difference

Based on understandings from previous studies, the Morlet wavelet parameters have to be optimised to obtain the best detection of impulse components. Entropy technique based optimisation has been applied in adaptive machine fault detection and diagnosis and shows good results in many industrial fields, such as system parameter recognition, structure optimisation, device operation-state detection and fault diagnosis [85]-[88]. Based on the wavelet optimisation procedure suggested in [80][83], the f_b and f_c in Morlet wavelet are optimised for minimising the entropy of two simulated signals according to Equation (2.6). The first signal has smaller impulse components, .i.e. 90% of that in $s_2(t)$ and $s_3(t)$, whereas the second has larger impulse components; i.e. the sample amplitude calculated by Equation (2.7).

But these methods may have problems in extracting useful features when signals have inherent components with high amplitudes. Particularly, vibration signals from a multistage gearbox have modulating components and mark the small impulse component due to a small fault, which will be demonstrated in the next section. Some big impulses generated by machine rotation movement will make the result worse. On the other hand, the impulse for a small fault is generally so small that makes it difficult to recognise the fault components. To solve this problem, a new optimisation method based on maximum Shannon entropy difference is proposed, which takes the entropy difference between the

fault signal and the baseline as the criterion in applying wavelet analysis, and the constant components such as those of meshing frequencies could be reduced, making it much easier to extract fault feature.

Assuming the CWT results of the two signals $x(t)$ and $y(t)$ are denoted with $X(a)$ and $Y(a)$, where a denotes the scale in wavelet analysis, then the entropies $EnX(a_i)$ and $EnY(a_i)$ are calculated along the time direction at the scale a_i, ($i = 1,...,M$), separately. Signal $x(t)$ is deemed as the reference signal.

Then the entropy difference between the two signals is obtained by:

$$EnDiff(a_i) = EnY(a_i) - EnX(a_i) \quad (5.3)$$

And the maximum entropy difference is defined as:

$$MaxEnDiff = \max[EnY(a_i) - EnX(a_i)] \quad (5.4)$$

The procedure of the new adaptive method is described as follows:

(1) Choose a proper range for bandwidth f_b, centre frequency f_c and scale a_i to ensure the fault feature exists in the analysis range, which is around the first few meshing components.

(2) Firstly, perform the processes for the baseline signal because this is the reference for fault detection. Select a couple of f_b and f_c, and compute the Shannon entropy for each scale as $EnR(a_i)$. Follow the same procedure to compute the entropy of a small fault signal as the baseline signal and get $EnS(a_i)$.

(3) Then, compute $EnDiff(a_i) = EnR(a_i) - EnS(a_i)$ to get the entropy difference between the two faults and the baseline signal. Also, find the maximum entropy difference along the scale and save as $MaxEnDiff(fb, fc) = \max(EnDiff(a_i))$, so there is a wavelet entropy difference series $EnDiff(a_i)$ for each combination of (f_b, f_c).

(4) Then, repeat steps (2) and (3) for each combination of f_b and f_c. Then, a set of $MaxEnDiff(f_b, f_c)$ for every f_b and f_c is obtained.

(5) Next, find the best parameters f_{b0} and f_{c0}, which is the one where the entropy difference $MaxEnDiff(f_b, f_c)$ is the maximum.

(6) The best scale a_{i0} is the one making $EnDiff(a_i)$, which corresponds to f_{b0} and f_{c0}, getting the maximum value.

The appearance of impulsive fault components could result in a decrease of the entropy value, compared with the baseline signal the more it decreases, the more impulse components are taking place. Therefore, the scale selected with maximum entropy difference criteria is more reliable for fault detection. This optimising method will reduce the influence of periodic components generated by rotation machine mechanism and enhance the fault feature of impulsive components.

5.3 Experimental setup

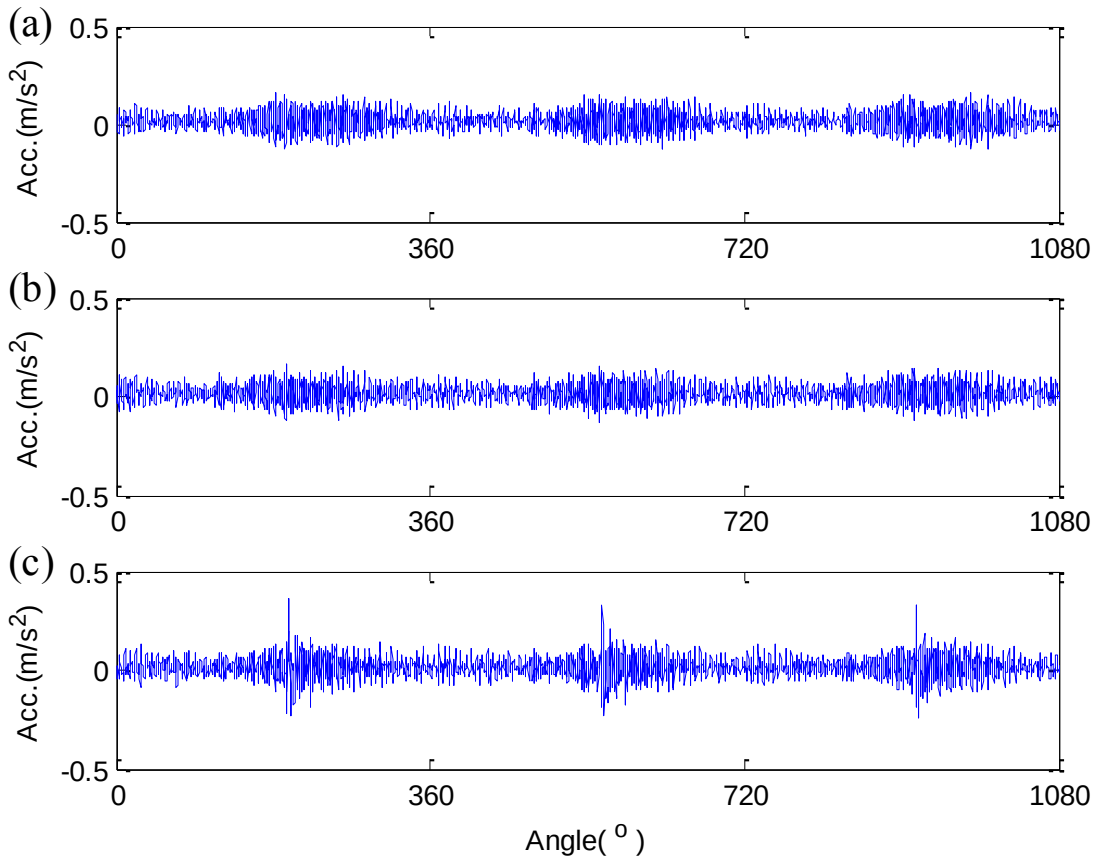

Figure 5.2 Three condition signals in the time domain

(a) Baseline; (b) Smaller fault; and (c) Larger fault

To examine the sensitivity of wavelet analysis, the three cases (healthy, 30% tooth breakage and 100% tooth breakage) of experimental data are applied as in Chapter 3. Because the experimental setup is the same as in Chapter 3, it will not be repeated. The two degrees of tooth breakage (30% and 100%) are known as the smaller and larger fault for brevity, respectively. The larger fault is for understanding the potential characteristics

of wavelet transform and the smaller tooth breakage is interesting in this study to evaluate the performance of the proposed wavelet analysis method.

5.4 Results and discussion

The vibration signals under three kinds of healthy condition (baseline, smaller fault and larger fault) are applied to verify the effect of the proposed method. Figure 5.2 shows the three kinds of signals in the time domain and Figure 5.3 shows the corresponding spectra. It is difficult to see the fault feature, especially for the baseline and smaller fault.

Figure 5.4 and Figure 5.5 give out the adaptive wavelet results for the smaller fault. Figure 5.4 is the adaptive wavelet spectrum for the baseline and small fault signal. Figure 5.5 plots the entropy of the two signals and the entropy difference between them according to scale. The entropy difference reaches the top point when the scale is 34. This indicates the fault feature is best revealed at scale 34.

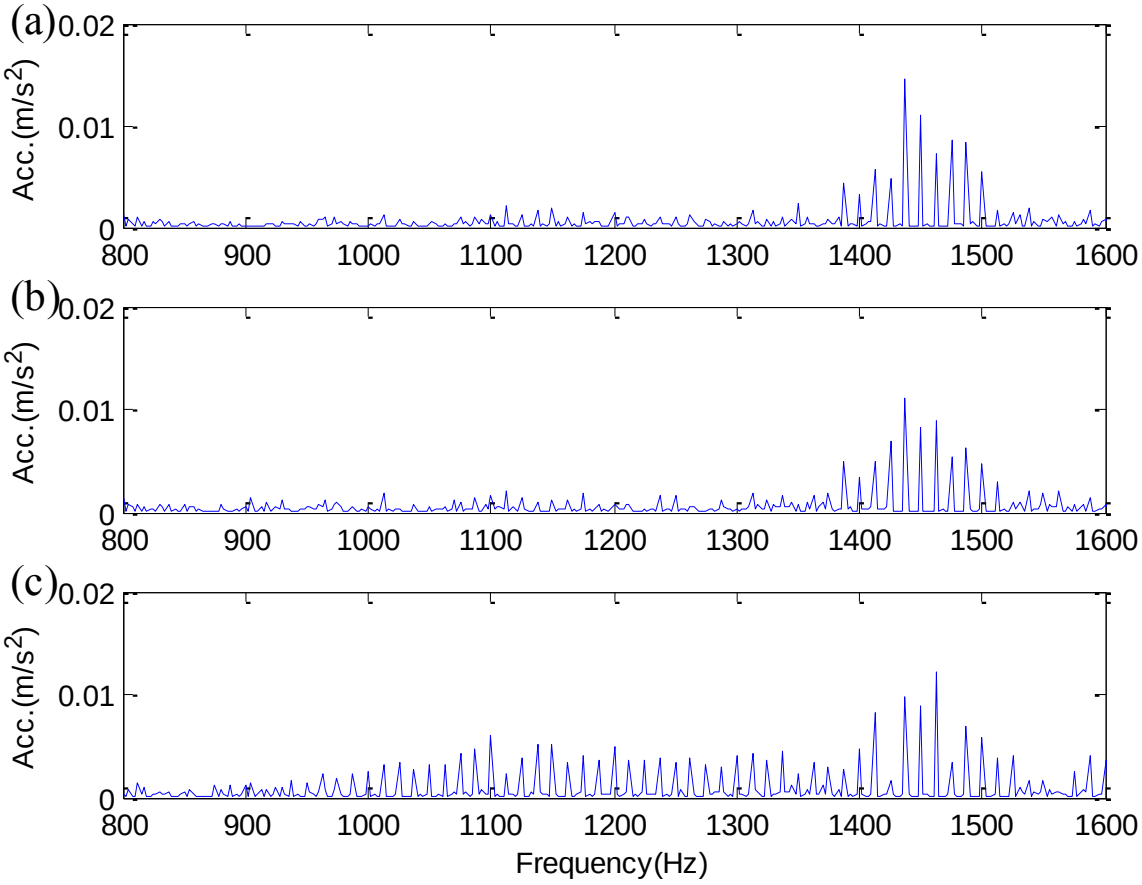

Figure 5.3 Three condition signals in the frequency domain

(a) Baseline; (b) Smaller fault; and (c) Larger fault

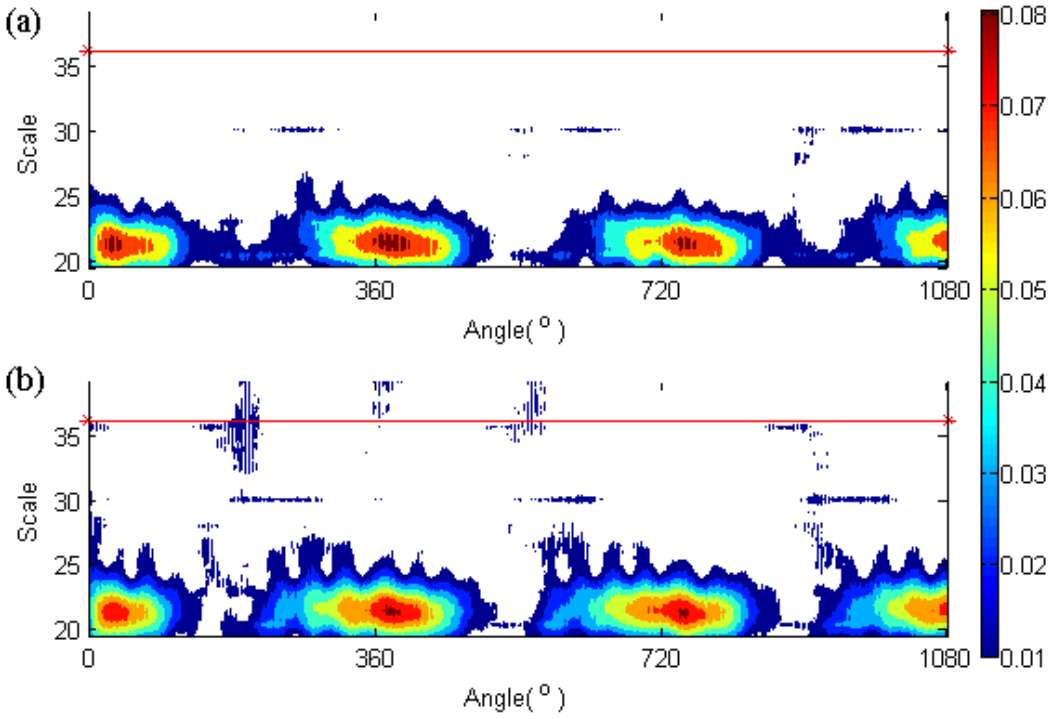

Figure 5.4 Adaptive wavelet results for the smaller fault

(a) Baseline and (b) Smaller fault

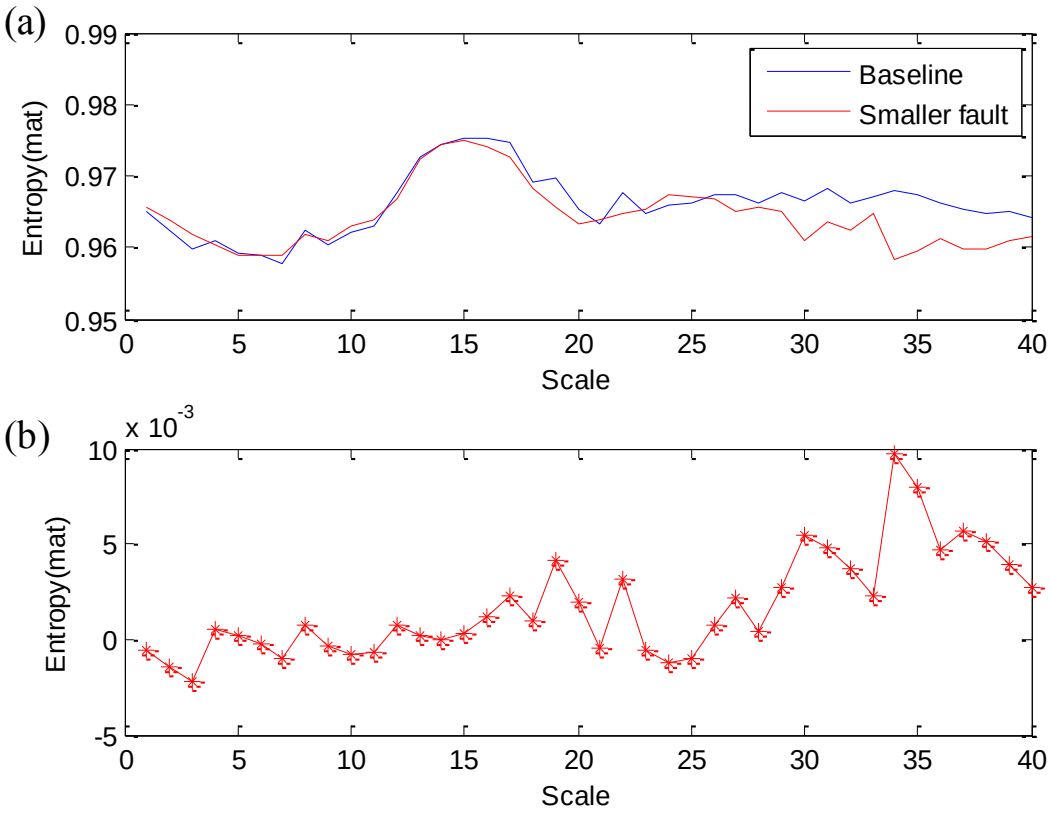

Figure 5.5 Wavelet entropy values for the smaller fault

(a) Entropy and (b) Entropy difference

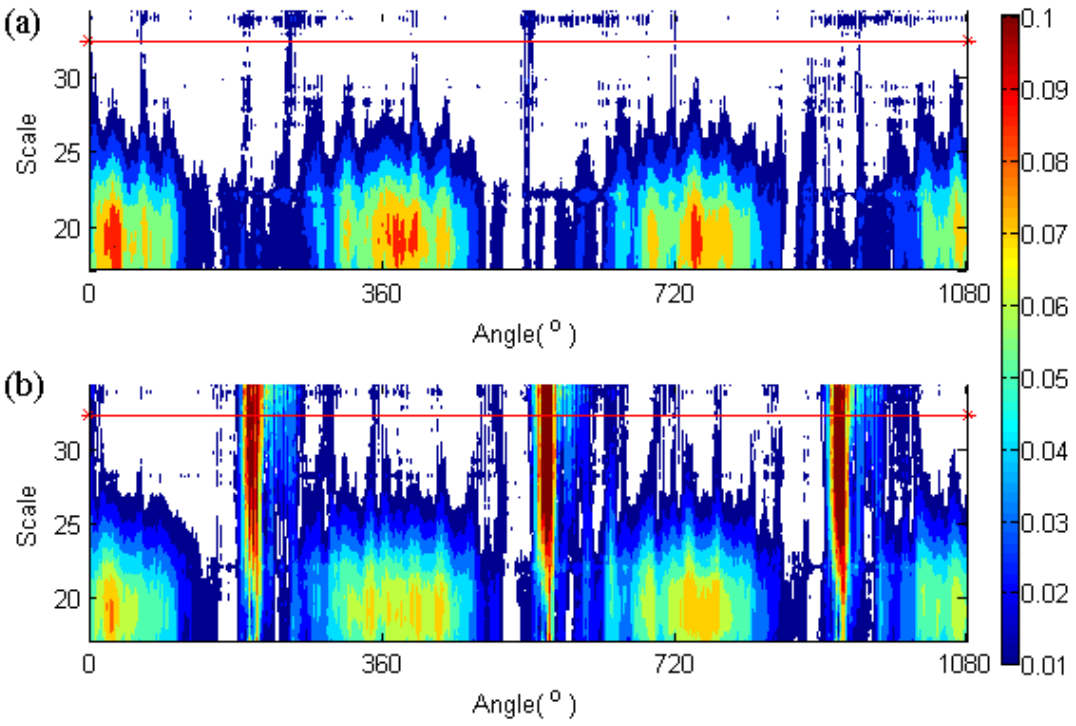

Figure 5.6 Adaptive wavelet results for the larger fault

(a) Baseline and (b) Larger fault

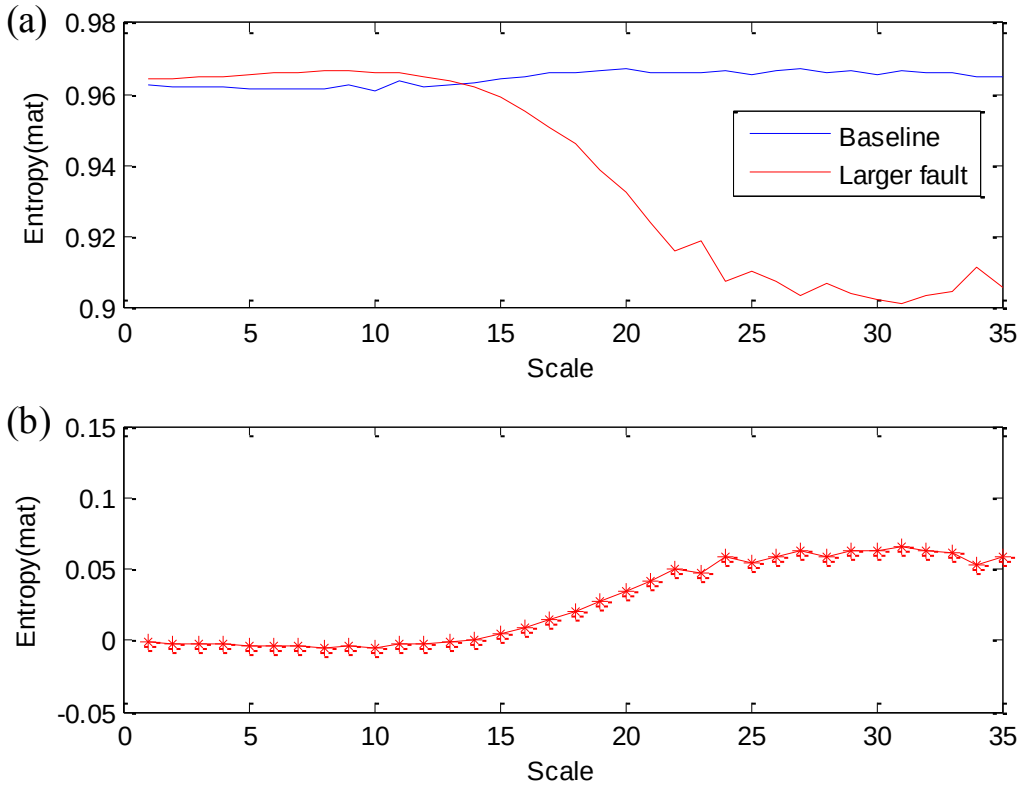

Figure 5.7 Wavelet entropy values for the larger fault

(a) Entropy and (b) Entropy difference

Figure 5.6 and Figure 5.7 present the adaptive wavelet results for the larger fault. Figure 5.6 is the adaptive wavelet spectrum for both the baseline and the larger fault signal. The fault features are very obvious and are spread along the scale direction. Figure 5.7 plots the entropy of the two signals and the entropy difference between them for different scales. The entropy difference is relatively high in the scale range from 25 to 35. This means the fault feature is obvious in this scale range. Any one of these scales in this range can be used for detection. The entropy difference reaches the top when the scale is 31, at which the best detection of impulsive components can be achieved.

From the experimental results, it can be concluded that the adaptive wavelet method is effective in enhancing weak impulsive components. The method proposed in this chapter has the ability to extract small faults from complicated gearbox signals.

5.5 Summary

In this chapter, an adaptive wavelet analysis technique based on the Shannon entropy difference is developed and applied to gearbox vibration signals for detecting incipient gear faults. The optimisation is implemented on both the underline signal and a reference signal, such as a baseline measurement or a simulated signal. In this way, the effects of dominated components (e.g. gear mesh frequencies) are reduced and new impulsive responses can be enhanced significantly to obtain reliable fault detection.

This method can also be potentially applied to bearing fault analysis. However, the drawback of this method is that it needs a baseline signal as a reference, which may not be available. This could limit the application of this method as it needs additional efforts to obtain and keep the reference signal.

Chapter 6
Bearing fault severity diagnosis based on kurtogram and envelope analysis

Envelope analysis is a widely used method for bearing fault detection. To obtain high detection accuracy, it is critical to determine an optimal filter so that the envelope signal obtained highlights more the impulsive components due to the fault. Kurtogram is a popular and effective method for automatic filter band selection for bearing fault detection in recent years. This chapter focuses on the study of the bearing fault severity diagnosis based on kurtogram and envelope analysis. It was carried out with simulated vibration signals from bearings with defects of constant width on outer race, inner race and roller, respectively. Then, an experimental study follows to evaluate this model.

6.1 Introduction

Rolling element bearings are at the heart of almost every rotating machine. Therefore, they have received a lot of attention in the field of vibration analysis as they represent a common source of faults [121]. In order to keep machinery operating at its best performance and avoid catastrophic failure, financial cost and personal injuries, different methods bearing fault diagnosis have been developed and used effectively to detect machine faults at an early stage, among which vibration signal processing is the most frequently applied one [122].

Vibration based condition monitoring have been widely used for detection and diagnosis of bearing defects for several decades [123]. A limited amount of work has been undertaken in using the vibration signals to investigate the bearing fault severity levels. Ocak et al. [124] introduced a bearing fault detection and diagnosis scheme based on hidden Markov modelling of the vibration signal, the data was gathered for four different conditions (normal, inner race fault, outer race fault and ball fault) with three different severities. Ocak et al. [125] presentd two separate algorithms for estimating the running speed and the bearing key frequencies of an induction motor using vibration data of ball bearing with inner and outer race defects with two different severity levels. Xu et al. [126] presented different bearing fault conditions feature parameters using a modified fuzzy ARTMAP (FAM) network model based on the feature-weight learning of inner race, outer race and ball faults using four different defect diameters. Zhang et al. [127] proposed a new method based on multi-scale entropy and adaptive neuro-fuzzy interference system conducted on electric motor bearings with three different fault categories of outer race, inner race and ball faults and four levels of fault severity. De Moura et al. [128] combined signal processing and pattern recognition techniques to diagnose three severity levels of the outer fault as well as no-fault class was also considered. Muruganatha et al. [129] proposed a simple time series method for bearing fault feature extraction using singular spectrum analysis of the vibration signal for the inner race, outer race and ball defects with four different fault sizes. Zhu et al. [130] proposed a fault feature extraction method based on the intrinsic mode function, envelope sample entropy for rolling bearings fault diagnosis. Single point faults with four different fault diameters were introduced into the test bearings with four conditions separately, which are healthy, with the inner race fault, with the outer race fault and with the ball fault. Jin et al. [131] presented a fuzzy ARTMAP (FAM) ensemble approach based on the improved Bayesian belief method and applied to the fault

diagnosis of rolling element bearings with four different defect diameters and introduced into the inner race, the ball and outer race.

However, most of these works are for fault detection and fault type diagnosis and limited investigations have been conducted on fault severity diagnosis, which is critical to make decisions for maintenance actions. To improve current diagnostic capability, this chapter pays more attention to bearing fault severity diagnosis.

6.2 Vibration responses to different sizes of fault

The bearing frequency equations provide a theoretical estimate of the frequencies to be expected when various defects occur on the bearing elements, based on the assumption that an ideal impulse will be generated whenever a bearing element encounters a defect. For localised bearing faults such as spalling and pitting, sharp force impacts will be generated. These impacts will excite structural resonances and the resulting vibration will be measured by the transducer mounted externally on the machine casing [132].

However, due to the different geometry of the contact between the localised fault and the bearing component, the contact stiffness can change because of the different geometrical properties in contact zones. On the other hand, a damaged bearing (particularly a small amount of damage at an early stage of damage development) usually produces small amplitudes of vibration in high frequency bands due to impulsive impacts [133].

The contact deformation is composed of geometric deformation and elastic deformation. Elastic deformation occurs along the contact surfaces of a bearing's rolling elements and raceway surfaces under loading. Geometric deformation caused by defect is related to defect location and size, while elastic deformation is related to load and defect size.

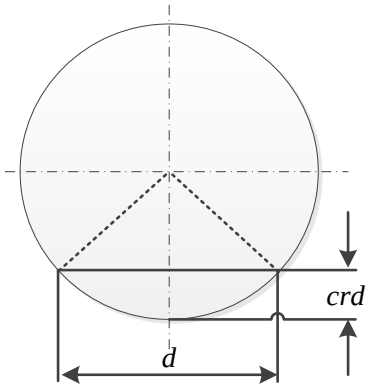

Figure 6.1 Schematic diagram of geometric deformation

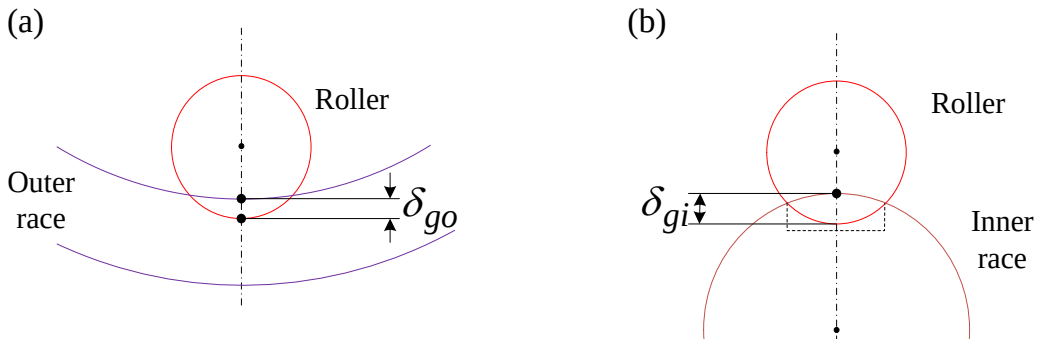

Figure 6.2 Geometric deformation for two kinds of contact

(a) Roller-Outer race (concave-convex); and (b) Roller-Inner race (convex-convex)

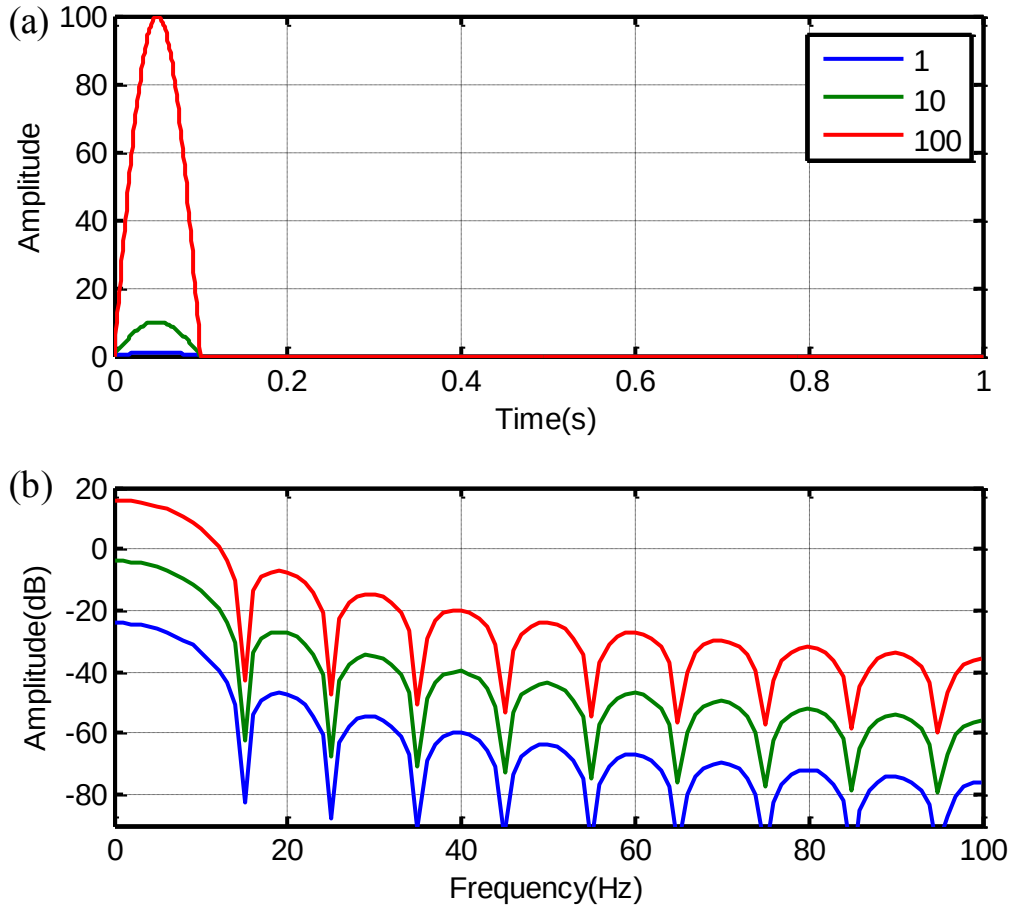

Figure 6.3 Frequency responses of difference pulse heights

(a) Half-sine pulse of different heights; and (b) Frequency response of different pulse heights

The total deformation Δ includes geometric deformation δ_g and elastic deformation δ_e.

$$\Delta = \delta_g + \delta_e \quad \text{.....(6.1)}$$

As shown in Figure 6.1, if the width of fault is d and radius of circle is r, the chord height can be expressed as

$$crd = r - \sqrt{r^2 - \left(\frac{d}{2}\right)^2} \quad \text{.....(6.2)}$$

There are two kinds of contact model between bearing components, as shown in Figure 6.2, which are contact between a roller and outer race (concave-convex) and contact between a roller and inner race (convex-convex). If the chord height for the roller, outer race and inner race are defined as crd_r, crd_o and crd_i, respectively, then the geometric deformation of the roller-outer race δ_{go} and the roller-inner race δ_{gi} are given by Equation (6.3) and Equation (6.4), respectively.

$$\delta_{go} = crd_r - crd_o \quad \text{.....(6.3)}$$

$$\delta_{gi} = crd_r + crd_i \quad \text{.....(6.4)}$$

Furthermore, considering that the fault on the inner race creates convex-convex contact whereas the fault on the roller has both concave-concave and concave-convex contact. The vibration impact from the inner race defect may create the highest responses when the sizes of the faults are the same on different races.

From the relationship, it is easy to understand that geometric deformations on different components have a relationship as shown in Equation (6.5).

$$\delta_{go} < \delta_{gr} < \delta_{gi} \quad \text{.....(6.5)}$$

Based on this relationship, Figure 6.3 illustrates the half-sine pulses of three different heights: 1, 10 and 100 and their frequency responses. It is obvious that when the amplitude is higher, the frequency response increases. This shows that when the fault size is the same, the fault on the inner race may produce the highest responses, whereas the fault on the outer race will cause the lowest responses.

Taking into account the elastic deformation, it is easy to understand the impulsive differences between different fault severities. As fault degree increases, the load area will decrease while elastic deformation δ_e will increase, which will result in the growth of impulse height.

6.3 Signal processing based on kurtogram

Vibration signals from a defective bearing with a localised fault contain a series of impulse responses, which result from the impacts of the defective part(s) with other elements. These impulses are generated almost periodically and their characteristics depend on the location of the defect, such as on the inner race, outer race or rolling elements.

The envelope spectrum is a very efficient diagnostic tool for the aforementioned faults, as the information about the fault is extracted from the spacing between impulses but not from the excited frequencies. The process of obtaining the envelope spectrum is often named as signal demodulation. However, the quality of the demodulated signal depends on the frequency band selected for the demodulation, which requires two parameters bandwidth and central frequency [134].

Because of the high values it takes at those frequencies where an impulsive bearing fault signal is dominant and its theoretical nullity where there is stationary noise only, it makes sense to use the SK as a filter function to filter out that part of the signal with the highest level of impulsiveness.

In the case of a rolling element bearing signal $x(t)$ modelled as a series of impulse responses $g(t)$ excited by impulses X at time τ_k:

$$x(t) = \sum_k g(t - \tau_k) X(\tau_k) \quad \text{.....(6.6)}$$

Then, if an impulsive component $x(t)$ is buried in additive stationary noise $n(t)$, the resulting measurement signal $y(t) = x(t) + n(t)$ has spectral kurtosis

$$K_y(f) = \frac{K_x(f)}{\left[1 + S_n(f)/S_x(f)\right]^2} \quad \text{.....(6.7)}$$

where $K_x(f)$ is the spectral kurtosis of $x(t)$ and $S_n(f)/S_x(f)$ denotes the SNR. This suggests that the optimal filter that minimises the similarity between the filtered component and the true noise-free signal is the square root of the spectral kurtosis. Following similar lines, the optimum filter, which minimises the SNR of the filtered signal, can be selected regardless of its shape. The matched filter is a narrow band filter at the maximum value of spectral kurtosis.

As previously pointed out, the spectral kurtosis, and therefore the optimal filter which can be obtained from it, will critically depend on the choice of the STFT window length or, as

also stated, on the bandwidth of the band-pass filter that outputs the complex envelope $X(t, f)$. One solution is to display the spectral kurtosis also as a function of the latter parameter, thus giving rise to a two-dimensional representation called "kurtogram". The equation for the kurtogram based on STFT can be presented as

$$K_x(f,n) = \frac{\left\langle \left| X(t,f)_n \right|^4 \right\rangle}{\left\langle \left| X(t,f)_n \right|^2 \right\rangle^2} - 2 \quad \text{.....(6.8)}$$

where n denotes the window length of STFT.

6.4 Simulation study of the kurtogram with bearing vibration signal

To evaluate the efficiency of the kurtogram in determining optimal filter parameters, four signals were generated by the bearing model presented in Chapter 2. They correspond to four bearing fault cases of outer race fault, inner race fault, roller fault and a combination of the above 3 faults, respectively. The resonance frequency or carrying frequency set in the bearing model is around 3.9kHz with a sampling frequency of 50kHz. The impulsive rates are 82.5Hz, 134.0 and 48.0Hz, representing faults on the outer race, inner race and rolling element respectively.

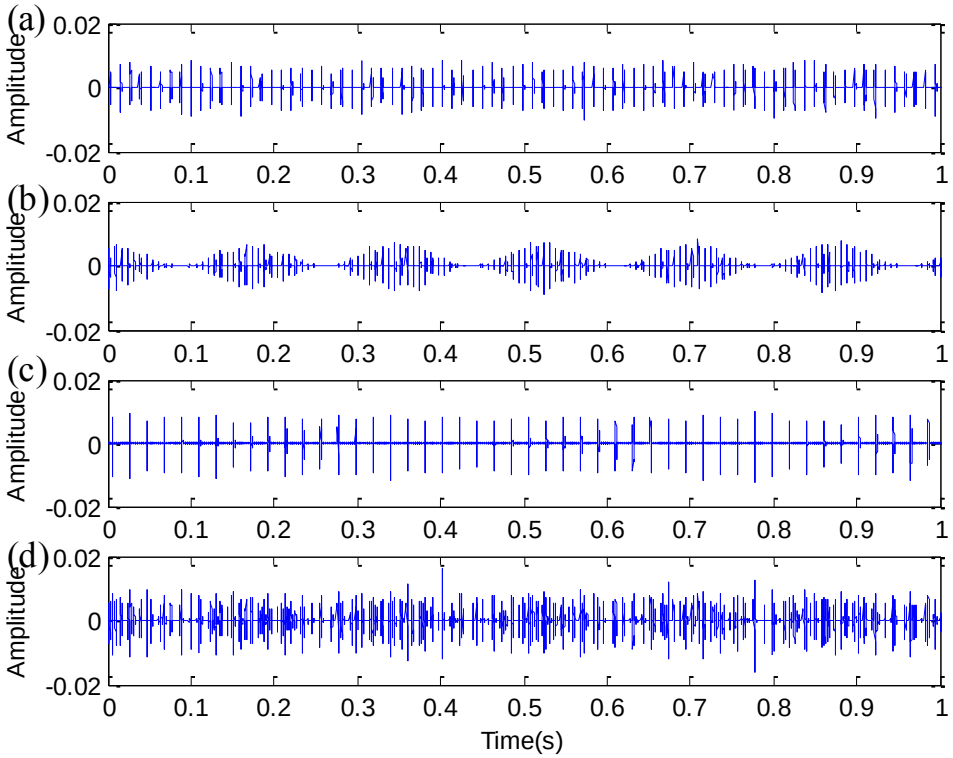

Figure 6.4 Simulated bearing signals for different fault cases

(a) Outer race fault; (b) Inner race fault; (c) Roller fault; and (d) Combined fault

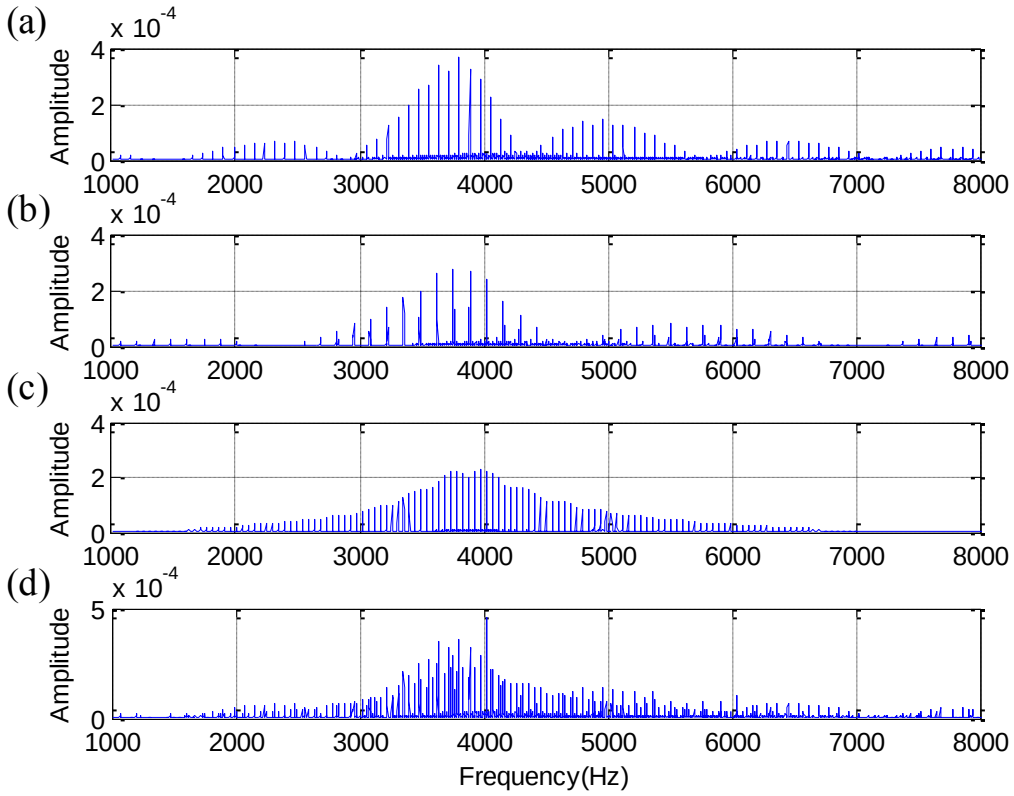

Figure 6.5 Simulated bearing signals in the frequency domain

(a) Outer race fault; (b) Inner race fault; (c) Roller fault; and (d) Combined fault

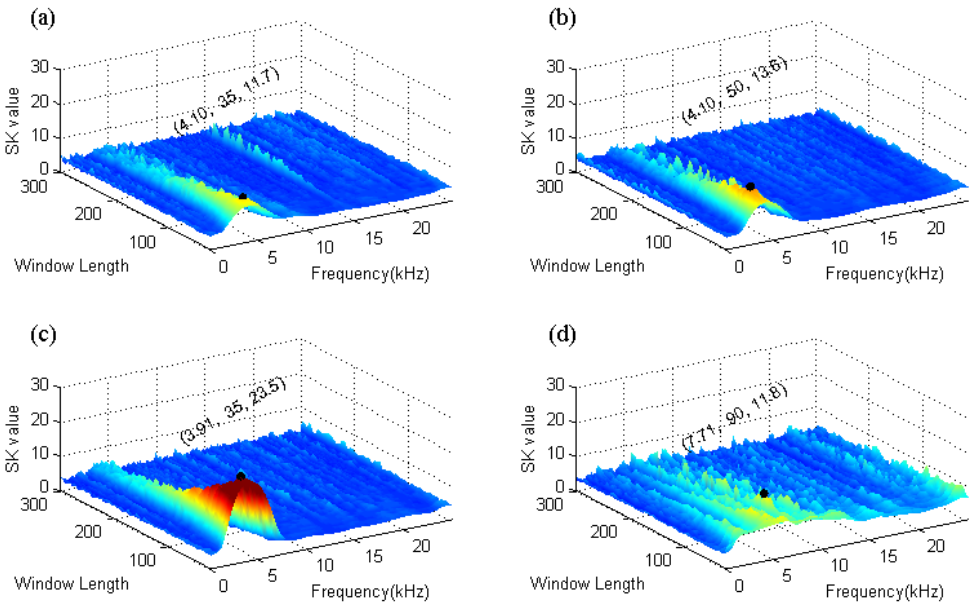

Figure 6.6 Kurtogram of bearing simulation signals

(a) Outer race fault; (b) Inner race fault; (c) Roller fault; and (d) Combined fault

Figure 6.4 shows bearing simulation signals in the time domain and Figure 6.5 presents the corresponding signal spectra. From Figure 6.5, it is obvious that the signal energy is

mainly near 4kHz, which is the resonance frequency of the bearing model. To get accurate envelope analysis results, it is critical to find out the resonance frequency. However, it is difficult to verify the resonance frequency from a practical experiment signal due to the complexity of the practical system model. The vibration signal distributes in the whole frequency range. Therefore, to develop a method for detecting the frequency band, which has the highest signal-to-noise ratio is important for bearing envelope analysis.

Figure 6.6 gives kurtogram based on the FFT of the bearing simulation signal. From the figure, it is obvious that the optimised central frequency of the bandpass filter is 4.1kHz, 4.1kHz, 3.91kHz, and 7.81kHz, respectively. The optimised central frequency of the three signal faults is near the resonance frequency of system 3.9kHz. However, the central frequency for the combination of 3 faults is twice that of the system resonance frequency. One thing worth mentioning is that it is necessary to add some random noise to the signal before computing the kurtogram. Otherwise, the kurtogram cannot find the right central frequency.

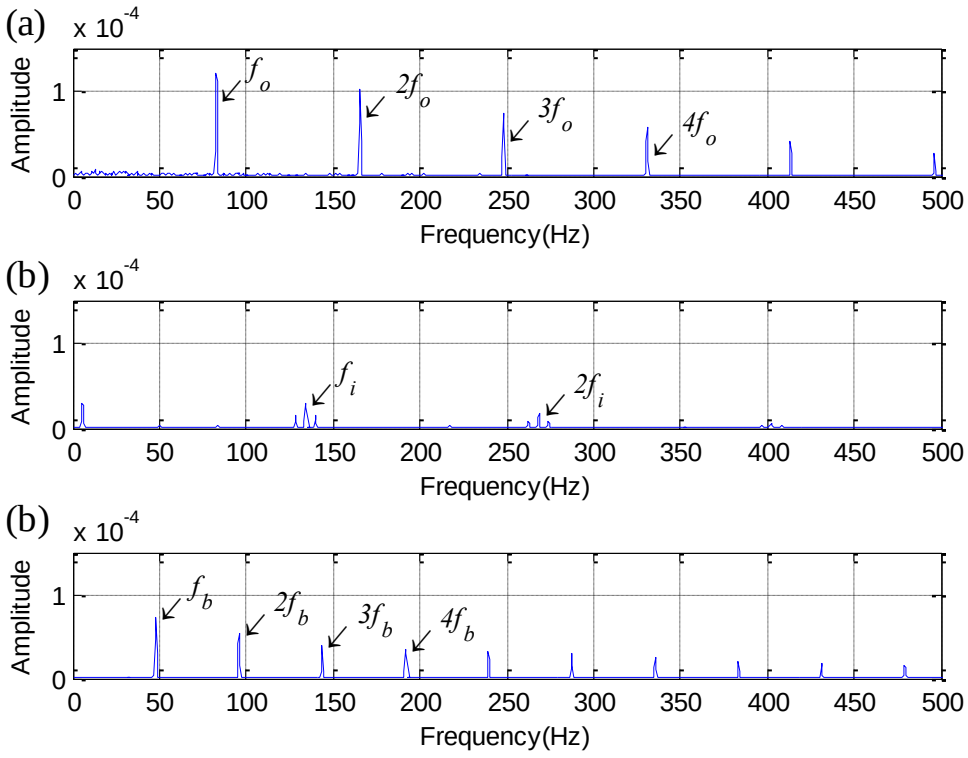

Figure 6.7 Envelope of simulated bearing single fault signals based on the kurtogram

(a) Outer race fault; (b) Inner race fault; (c) Roller fault; and (d) Combined fault

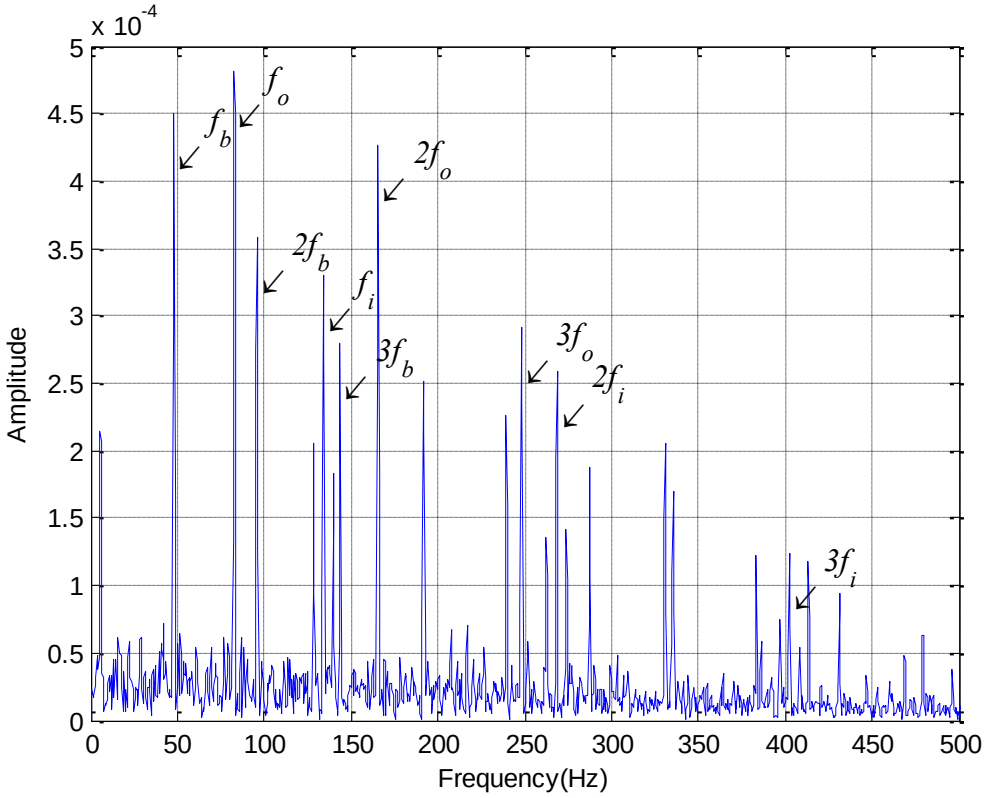

Figure 6.8 Envelope of simulated bearing combined fault signals based on the kurtogram

Then, the four signals are filtered by the bandpass filter designed above and processed with envelope analysis. Figure 6.7 and Figure 6.8 illustrate the envelope of bearing simulation signals filtered by the kurtogram optimised filter. Figure 6.7 presents the results of three fault signals. It can be seen that the fault characteristic frequencies and their harmonics are obvious. The fault frequencies of outer race fault, inner race fault and roller are at 82.5Hz, 134Hz and 48Hz, respectively. Figure 6.8 displays the result of a combination of 3 faults. It is easy to see the three fault frequencies. The result indicates that all of the three kinds of fault exist in the signal.

From all the above simulation results, it is concluded that the kurtogram is effective in filter optimisation in bearing signal processing.

6.5 Test rig facility and fault simulation

The experimental data analysed in this chapter was collected from the bearing test rig illustrated in Figure 6.9. It is composed of a motor, coupling, shaft, bearings and brake.

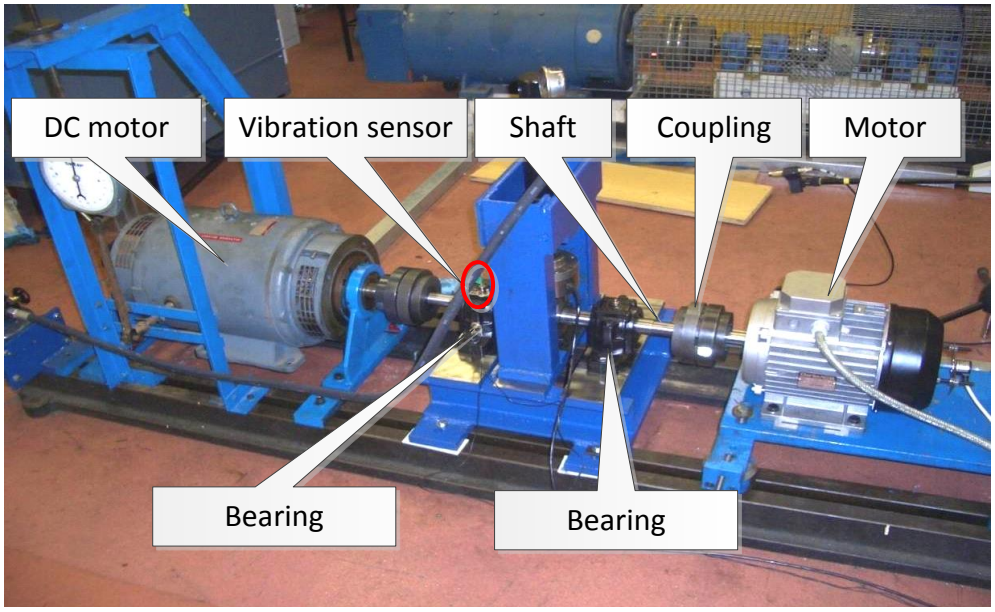

Figure 6.9 Photograph of bearing test rig

The bearing type is NSK N406 cylindrical roller bearing and its geometric dimensions are listed in Table 6.1. One Sinocera piezoelectric accelerometer is mounted on the housing of the N406 bearing vertically to measure the vibration. The frequency range of the accelerometer is from 0.5Hz to 10 kHz and the sensitivity is 8.08mv/ms^2. Defect frequencies in the experiment calculated according to Equations (4)-(9) are listed in Table 6.2.

Table 6.1 Specification of NSK type N406 cylindrical roller bearing

Parameter	Measurement
Pitch Diameter	59 mm
Bore Diameter	30 mm
Roller Diameter	14 mm
Roller Number	9
Contact Angle	0

Table 6.2 Fault characteristic frequencies

Fault type	Defect frequency (Hz)
Outer race	85.8
Inner race	139.2
Roller	49.7

The experiment was carried out based on ten different rolling bearings. One bearing is healthy and was taken as the baseline for comparison. Three bearings have induced outer

race faults with three different fault severities, which have constant size but three different lengths, 30%, 60% and 100% of the bearing outer race width. In the same way, inner race faults and roller faults are induced to the other two groups of bearings, and each group includes three bearings for three different fault severity levels. Figure 6.10 gives the photograph of defect rolling bearing with 30% roller fault, 60% inner race fault and 100% outer race fault, respectively.

Figure 6.10 Photographs of fault bearings

(a) 30% roller fault; (b) 60% inner race fault; and (c) 100% outer race fault

6.6 Results and discussion

In this section, experimental data processing results are discussed based on the mechanical vibration model. The same baseline signal is applied to the outer race, inner race and roller fault for comparison.

Figure 6.11 shows the raw vibration signals for the baseline, inner race and outer race at four kinds of conditions, respectively. It can be seen that the vibration amplitudes of the outer race are smaller than the inner race and roller for three kinds of fault severities. In the meantime, the vibration amplitude of the roller is higher than the inner race at 100% damage but similar at both 30% damage and 60% damage.

Figure 6.12-Figure 6.14 give the envelope analysis results of three kinds of fault. From the figures, it is obvious that the characteristic frequencies and their harmonics are significant, which verifies that filter parameters can be optimised by kurtosis maximum principle effectively.

The amplitude of the inner race fault frequency grows with the damage severity while the outer race and roller fault do not have such a good trend. Envelope analysis results indicate that it can be applied for fault detection and can recognise the fault type by fault characteristic frequencies.

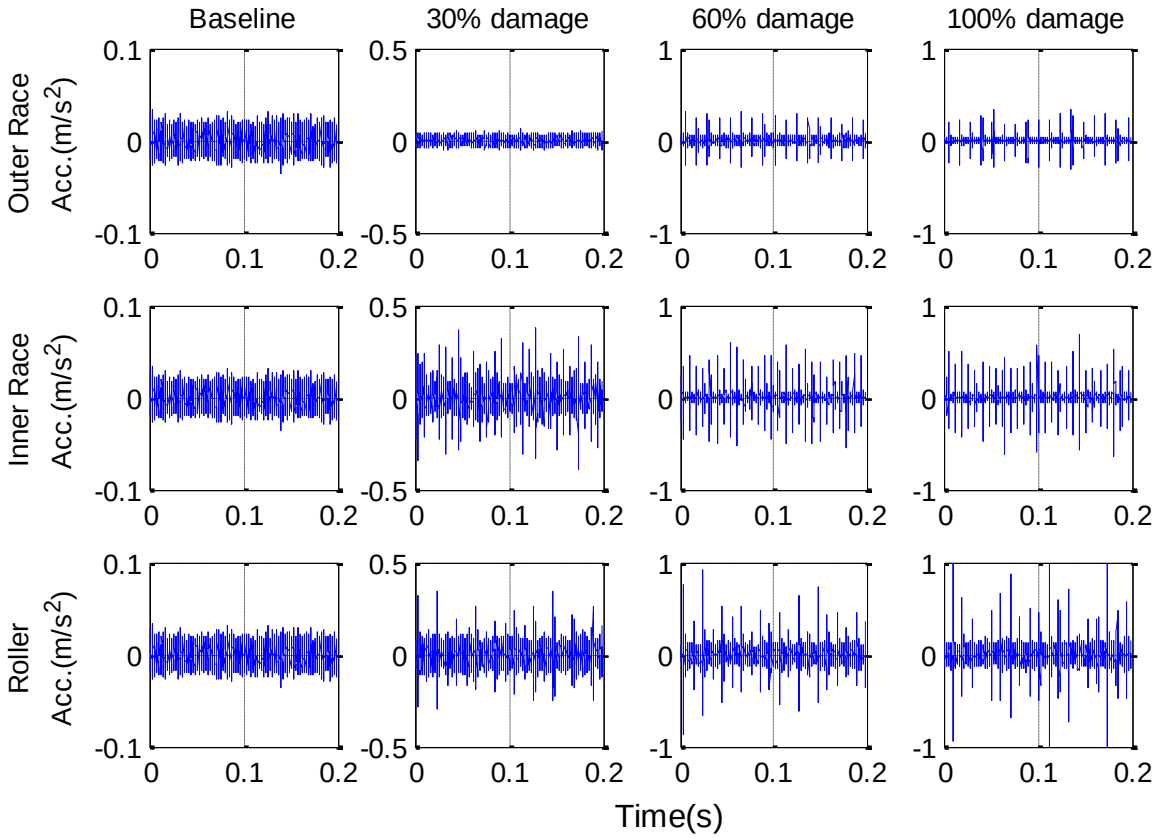

Figure 6.11. Vibration signals in the time domain

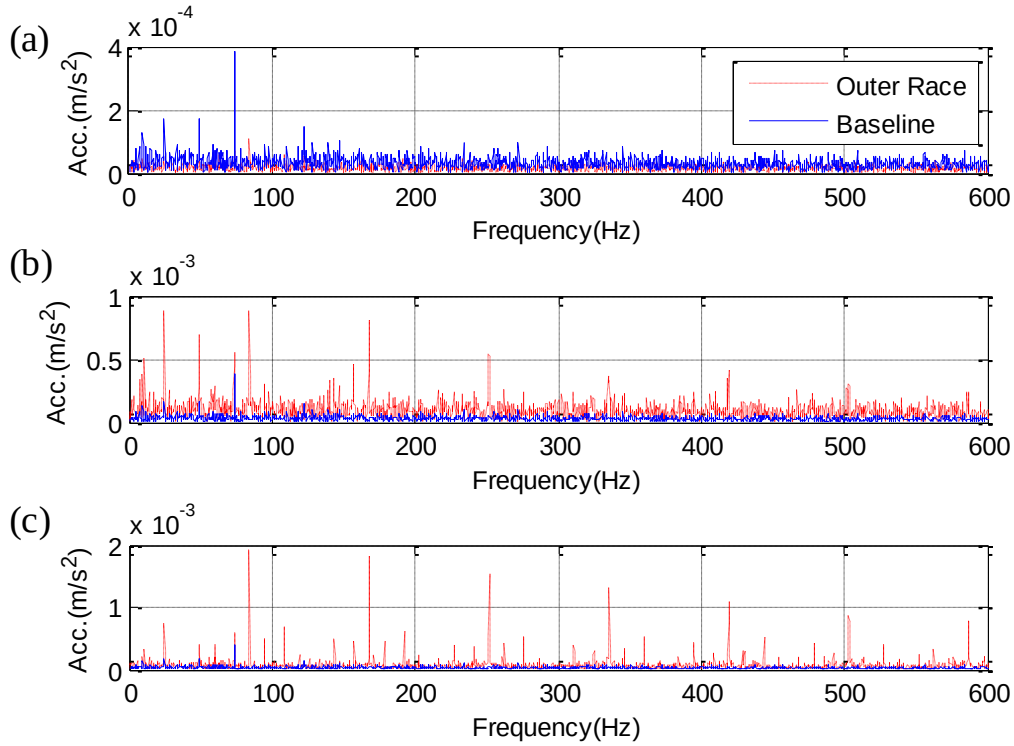

Figure 6.12 Envelope analysis results for outer race defect

(a) 30% damage; (b) 60% damage; and (c) 100% damage

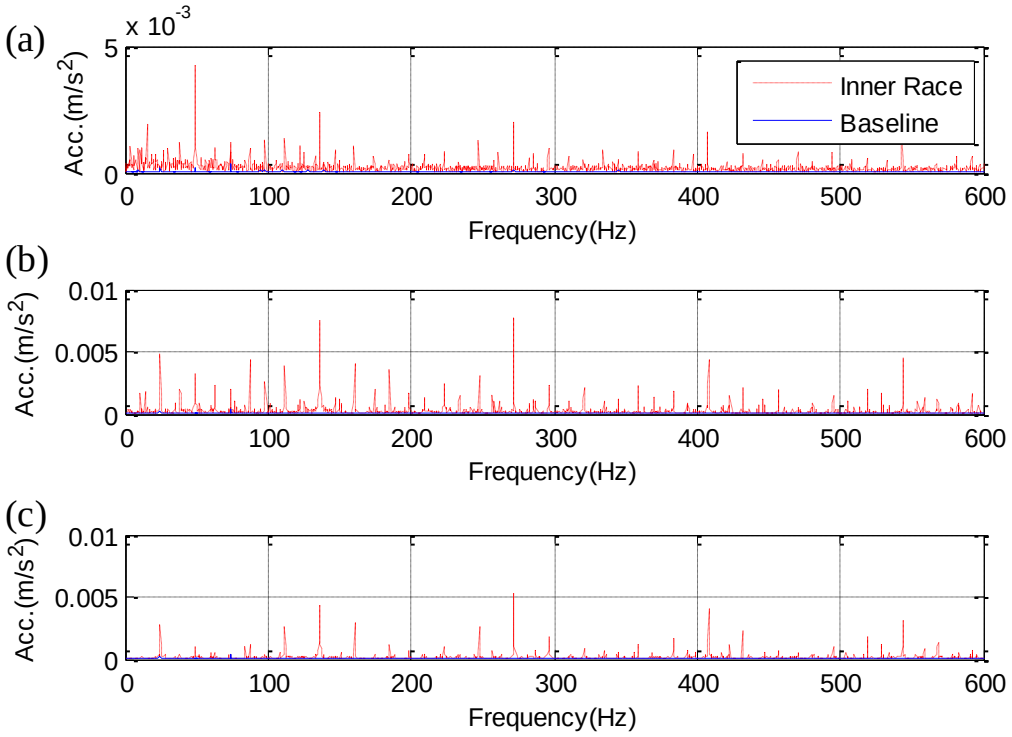

Figure 6.13 Envelope analysis results for inner race defect

(a) 30% damage; (b) 60% damage; and (c) 100% damage

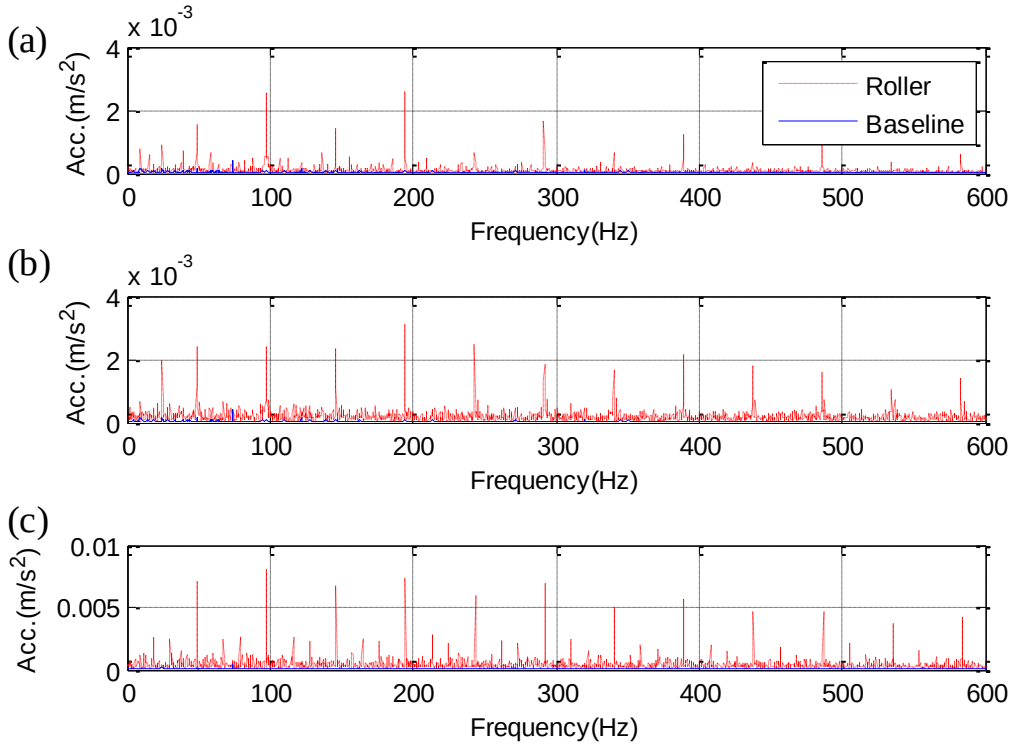

Figure 6.14 Envelope analysis results for roller defect

(a) 30% damage; (b) 60% damage; and (c) 100% damage

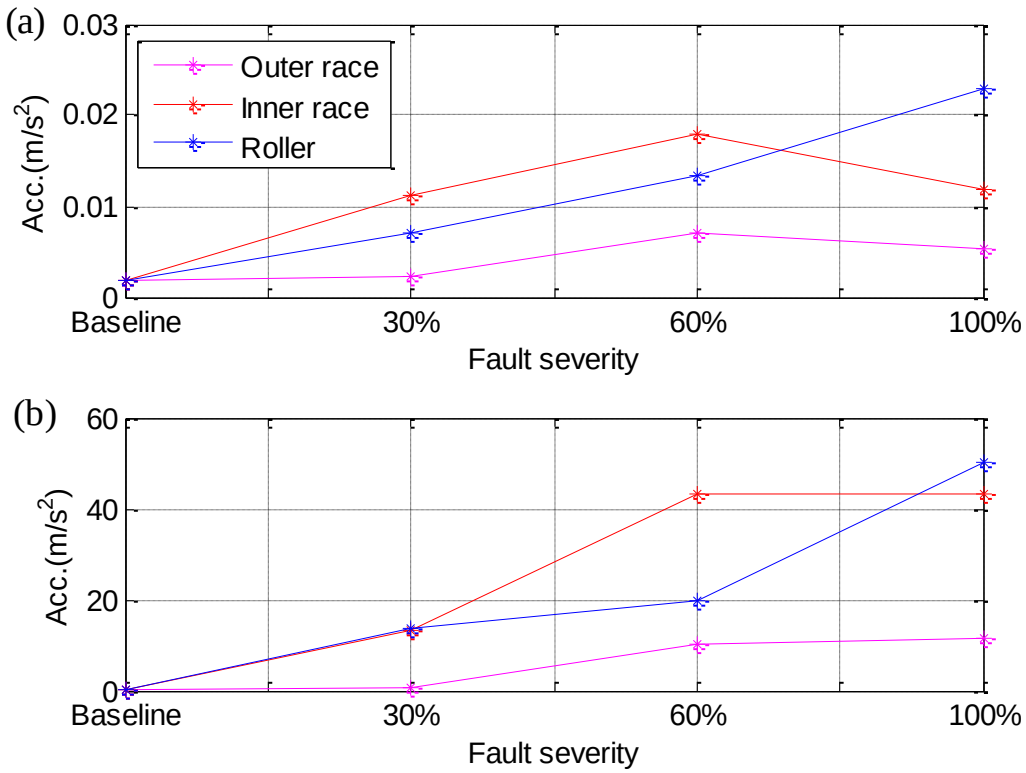

Figure 6.15 Results comparison

(a) Mean RMS of three tests; and (b) Mean kurtosis of three tests

Figure 6.15 demonstrates the mean RMS and kurtosis comparison of three tests. The RMS value increases greatly with the damage severity and shows an obvious difference between three kinds of faults. But 100% damage does not follow the prediction because the motion between the bearing components includes slippage. In addition, the kurtosis results can also separate the outer race fault. However, the difference is tiny between the inner race and roller fault, showing that the RMS is the better choice for fault diagnosis.

When there is the same size fault on the outer race, inner race and roller, the fault impulse amplitude of the outer race is constant and smaller and the waveform impulse is less spiky, while the inner race fault frequency is modulated at the shaft frequency and the roller fault frequency is modulated at the cage frequency. With the same defect size on the three components separately, the outer race impulse has the lowest “peakedness” compared with the inner race fault. On the contrary, the roller fault has the largest deformation and should have the highest “peakedness” which is not the case in the experimental results. This may be caused by the high level of noise in the roller vibration which will impact the kurtosis value.

The experimental data analysis results show that the RMS and kurtosis values can be used to predict bearing fault severity, which provides a decisive reference for taking maintenance actions.

6.7 Summary

In this chapter, a bearing fault diagnosis approach has been developed by combining the TFR based kurtogram with envelope analysis. The kurtogram is utilised to automatically select a filter band which has high SNR in the vibration signal for envelope analysis. Then, the fault location diagnosis and detection feature is extracted from the envelope spectrum of the filtered vibration signal, and the RMS and kurtosis values are used for fault severity diagnosis. The method has been confirmed by both simulated and experimental signals. Particularly, the both RMS and kurtosis of the optimised envelope signals provides a consistent indication of the fault severity for common faults occurring on the outer race, inner race and roller of a bearing.

Chapter 7
A robust detector for bearing condition monitoring based on the MSB

The method applied in the last chapter has a significant disadvantage in that it is sensitive to random noise and aperiodic impulses which normally occur in practical applications. For random noise suppression, averaging is an effective method. However, it is necessary to align the phases between multiple segments of data. For example, an encoder signal is required to indicate the start point of each revolution in the TSA analysis, which needs additional devices to collect the signal and these are sometimes difficult to obtain. To solve this problem, a novel MSB based robust detector for bearing fault detection is proposed to denoise and extract the modulation fault features in this chapter. Because of its inherent noise suppression capability, the MSB allows for the 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.

7.1 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 [135][136]. For the 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 [137]. Antoni applied cyclostationary spectral analysis [138][139], 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 [140]. Barszcz applied the same approach to denoise wind turbine vibration signals for bearing outer race fault diagnosis [141]. 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 [142]. Zhao applied empirical mode decomposition and the approximate entropy method for the severity assessment of a spall-like fault in a rolling element bearing [143]. A recent significant advance in envelope based rolling element bearing fault detection has been the kurtogram [144] and this has received considerable attention in recent months [145][146]. 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 or 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 the power spectrum approach [104] This approach, however, still requires optimisation of the filter's parameters for envelope analysis. In this chapter, 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.

7.2 The modulation signal based detector

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 [108], defined as follows:

$$B_{MS}^{SE}(f_c, f_x) = \frac{B_{MS}(f_c, f_x)}{\sqrt{|B_{MS}(f_c, 0)|}} \quad \text{.....(7.1)}$$

where $B_{MS}(f_c, 0)$ is the squared power spectrum estimation at $f_x = 0$ and is equal to $|X(f_c)|^4$. This is an important property that can be used as the basis for calculation of individual sideband amplitudes [108].

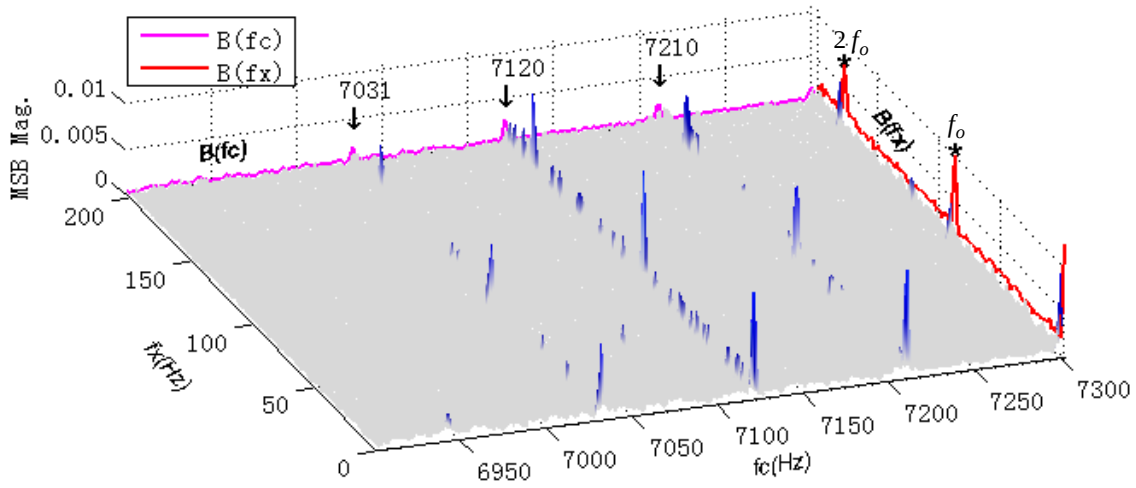

Figure 7.1 MSB showing detector $B(f_x)$ formed from slices shown along $B(f_c)$

Figure 7.1 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 Figure 7.1. Based on this approach, fault detection can be implemented directly, avoiding the preliminary processing steps associated with the selection of the optimal frequency band during conventional envelope calculation

[104][134]. In conjunction with the MSB's noise and impulse interference suppression capabilities, this results in a significantly more robust detection method.

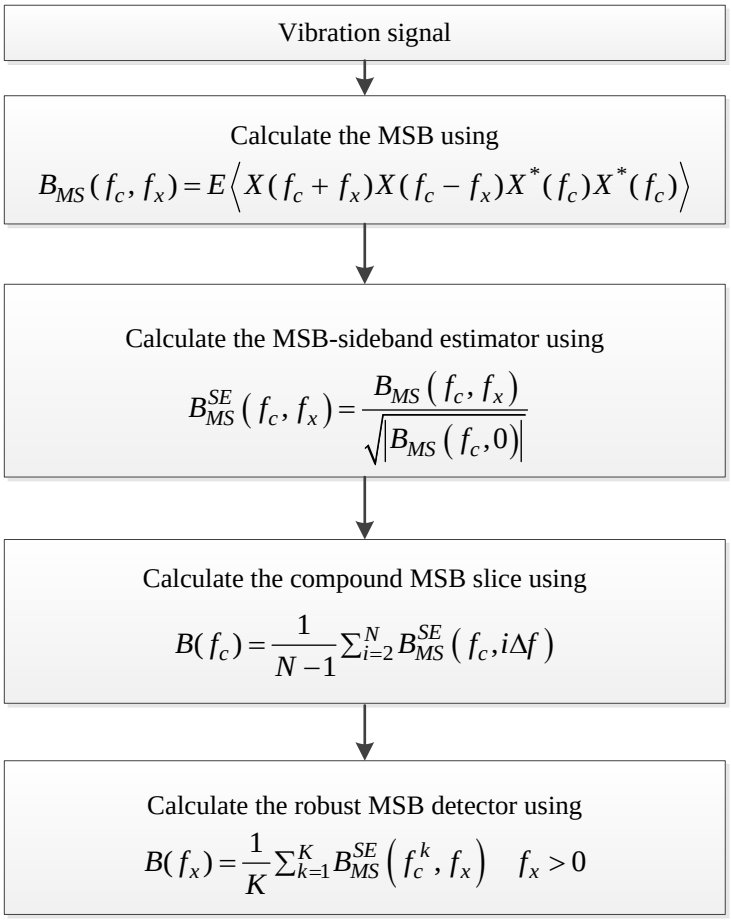

Figure 7.2 Flow chart of the robust MSB detector calculation

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 Figure 7.1, and it is this final adaptation that results in what this paper refers to as 'the robust MSB detector':

$$B(f_x) = \frac{1}{K}\sum_{k=1}^{K} B_{MS}^{SE}\left(f_c^k, f_x\right) \quad f_x > 0 \qquad (7.2)$$

where K is the total number of selected f_c suboptimal slices (3, in the case of Figure 7.1), the number of which depends on the significance of the peaks themselves. From Figure 7.1, 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.

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:

$$B(f_c) = \frac{1}{N-1}\sum_{i=2}^{N} B_{MS}^{SE}\left(f_c, i\Delta f\right) \quad (7.3)$$

where Δf is the frequency resolution in the f_x direction.

In summary, the robust MSB detector can be implemented using four primary steps shown in Figure 7.2:

7.3 Simulation study

To evaluate the performance of the MSB detector, simulated signals with known white noise and impulsive interference levels were used in a comparative study between the proposed MSB approach and the identified benchmark method, the kurtogram. Firstly, a noise free signal was produced using a nonlinear bearing model [147] that included a defect on the outer race, along with nonlinear loading effects. Then, differing levels of white noise and aperiodic impulsive interference were added to the signal, thus creating the four scenarios detailed in Table 7.1. To quantify noise influences, two means of calculating SNR were used, one for the stationary noise and another for the aperiodic impulsive noise case. For white noise, SNR is defined as:

$$Type\ 1\ SNR = 10\log_{10}\left(P_s / P_n\right) \quad (7.4)$$

where P_s and P_n indicate the power of the signal and noise respectively. And for the aperiodic impulsive interference, SNR is:

$$Type\ 2\ SNR = 20\log_{10}\left(A_s / A_n\right) \quad (7.5)$$

where A_s and A_n represent the amplitude of the signal and the noise, respectively.

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

Table 7.1 Description of the four noise scenarios

Scenario	White noise	Aperiodic interference	Type 1 SNR value	Type 2 SNR value
Low noise signal without interferences	Level 1	None	-15dB	n/a
High noise signal without interferences	Level 2	None	-30dB	n/a
Low noise signal with low level interferences	Level 1	Level 1	-15dB	-22dB
High noise signal with high level interferences	Level 1	Level 2	-22dB	-48dB

7.3.1 Robustness to white noise

Figure 7.3 shows the time series and corresponding spectral representations for three different simulated outer race fault signals (note the different amplitude scale for the three time traces). Figure 7.3(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 [148]. Figure 7.3(b) illustrates the time waveform and the spectrum of the low noise signal without impulsive interference. Figure 7.3(c) presents the equivalent for the high noise signal again without interference. The SNR, in both cases, is calculated using Equation (7.4). From Figure

7.3(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 to locate resonance frequencies and hence to implement accurate fault detection if the conventional envelope approach is used.

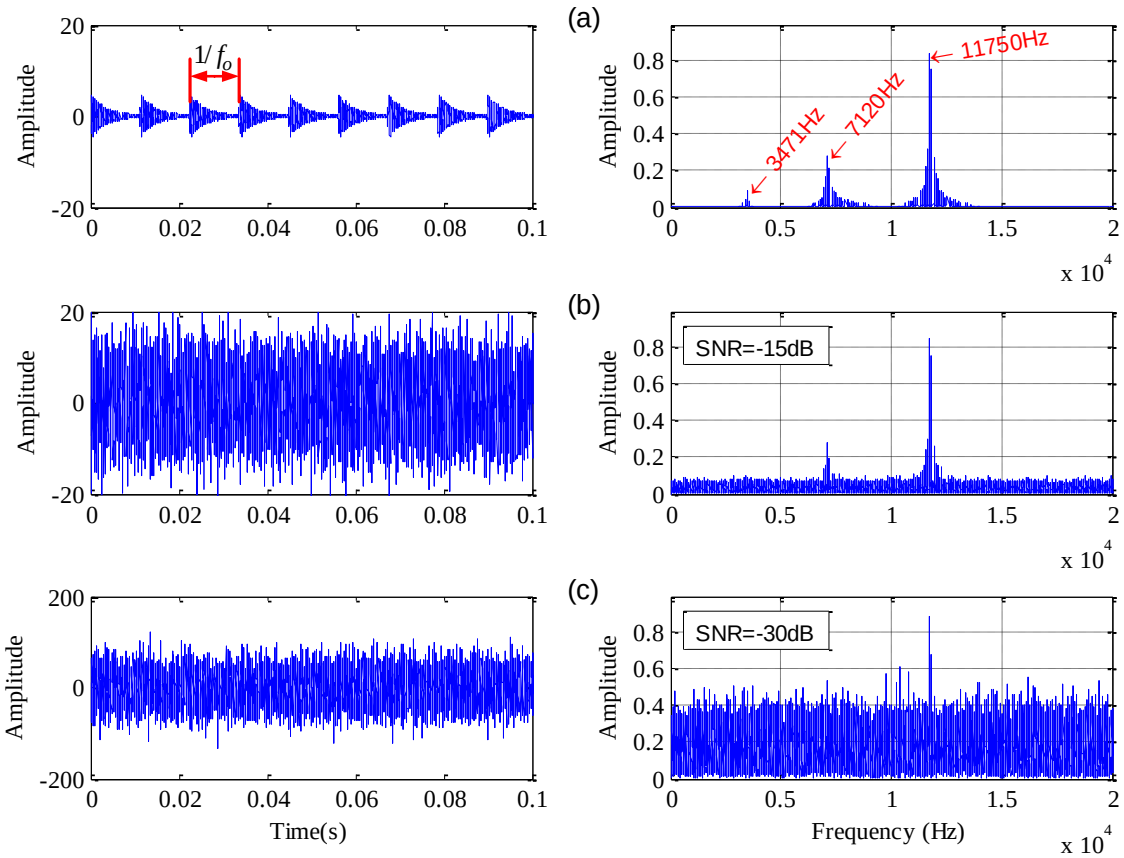

Figure 7.3 Time waveforms of the simulated signals and their spectra

(a) No noise; (b) Low noise; and (c) High noise

If one takes the low noise scenario as an example to describe the process of calculating the MSB robust detector, then the steps are as follows. Firstly, calculate the MSB (the FFT size is 262144 and the average time is 32) and sideband estimator using Equations (3.16) and (7.1). Then, calculate the compound MSB slice $B(f_c)$ to choose the suboptimal f_c slices and hence to achieve the result displayed in Figure 7.4(a). The f_c slices marked by '*' at around 7,000Hz (R2) and 12,000Hz (R3) have in this case been selected for the calculation of the MSB detector. Subsequently, the robust detector is calculated using Equation (7.2), as shown in Figure 7.4 (b). As a benchmark, the fast kurtogram algorithm [149] has also been applied to optimise the filter parameters for a narrowband envelope

analysis. 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 520.8Hz. Both the MSB and kurtogram-based approaches find the highest resonance frequency at around 11,719Hz. However, the MSB detector also finds the middle resonance frequency, which the kurtogram does not do, meaning that it is more flexible and potentially more capable in that it can detect multiple resonance frequencies, whilst the kurtogram can only select one optimal filter band.

Figure 7.4(b) shows the normalised results of the MSB detector and the kurtogram-based detector for the low noise signal. 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 the low noise case, although the MSB-based detector has lower background noise than that based upon the kurtogram.

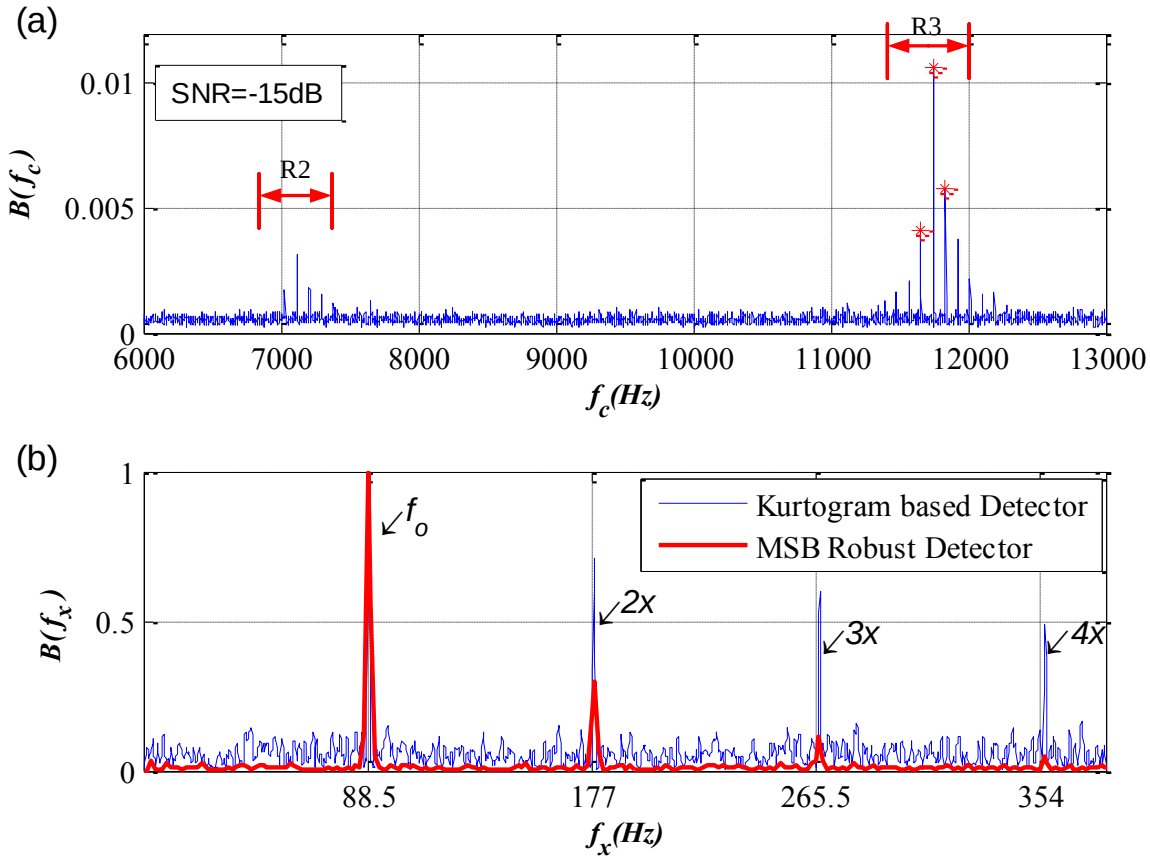

Figure 7.4 The compound MSB slice and results comparison (low noise, no impulsive interference)

(a) The compound MSB slice $B(f_c)$; and (b) Results of the MSB robust detector and the kurtogram based detector

For the high noise scenario, the compound MSB slice $B(f_c)$ and the normalised results of the two detectors are shown in Figure 7.5. It can be seen from Figure 7.5(a) that there are several significant peaks at around 11,800Hz, which is the highest resonance frequency of the simulated signal. In contrast, the central frequency optimised by the kurtogram is at 195.3Hz and the bandwidth is 390.6Hz, which is not the location of a resonance frequency. 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.

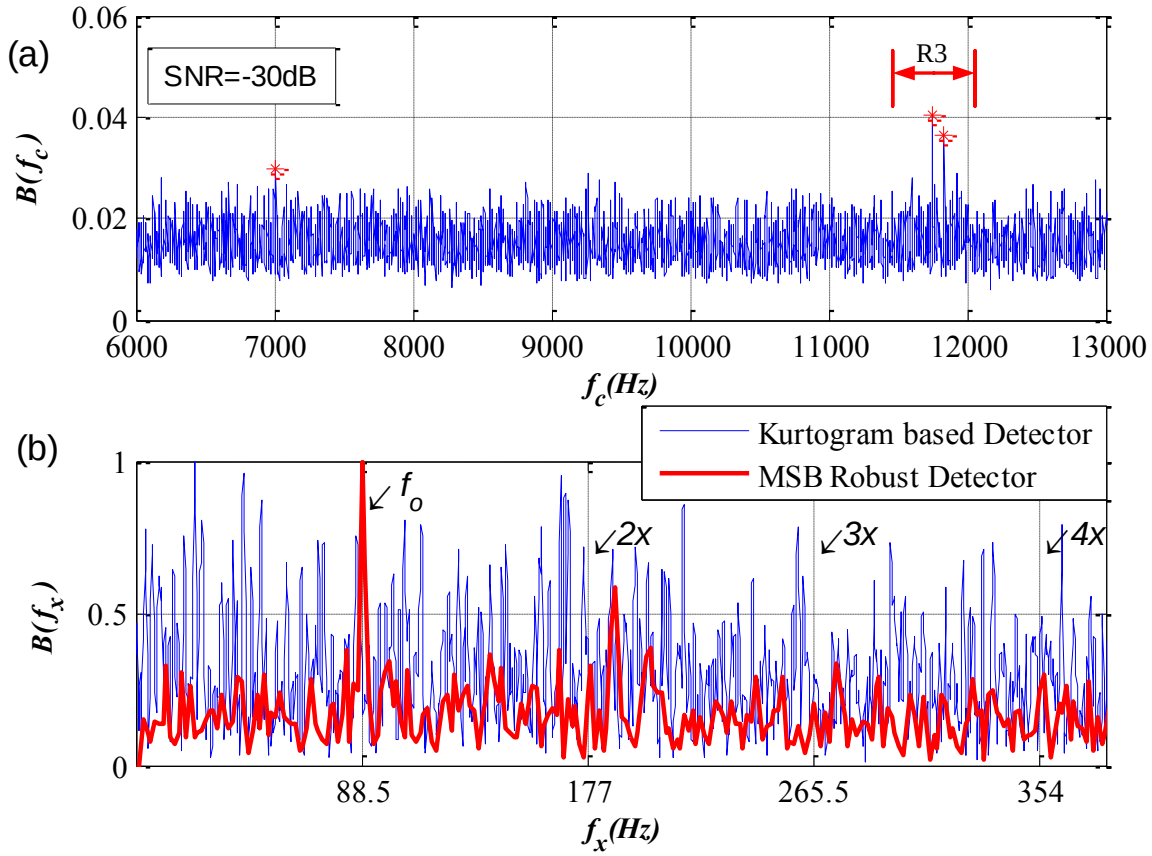

Figure 7.5 The compound MSB slice and results comparison (high noise, no impulsive interference)

(a) The compound MSB slice $B(f_c)$; and (b) Results of the MSB robust detector and the kurtogram based detector

7.3.2 Robustness to aperiodic impulsive interference with the presence of white noise

To examine the performance of the detector in the presence of aperiodic impulsive interference, a simulated fault signal was produced by adding random impulses to the previous low noise signal (again, note the different amplitude scales for the two sets of time data). Figure 7.6 and Figure 7.7 show the time waveforms of the simulated signals and their spectra. The SNR in the figure was calculated using Equation (7.5).

For the signal with low levels of amplitude interference in Figure 7.4(a), the optimised filter central frequency obtained by the kurtogram is at 11,719Hz and the filter bandwidth is 520.8Hz, meaning that the optimised filter is located at the third resonance frequency of the simulated signal. However, for the signal with high levels of interference, the optimised filter central frequency by kurtogram is at 4,166.7Hz and the filter bandwidth is 8,333Hz, which includes the third resonance frequency but has such a wide bandwidth that it results in a very low SNR for the filtered signal.

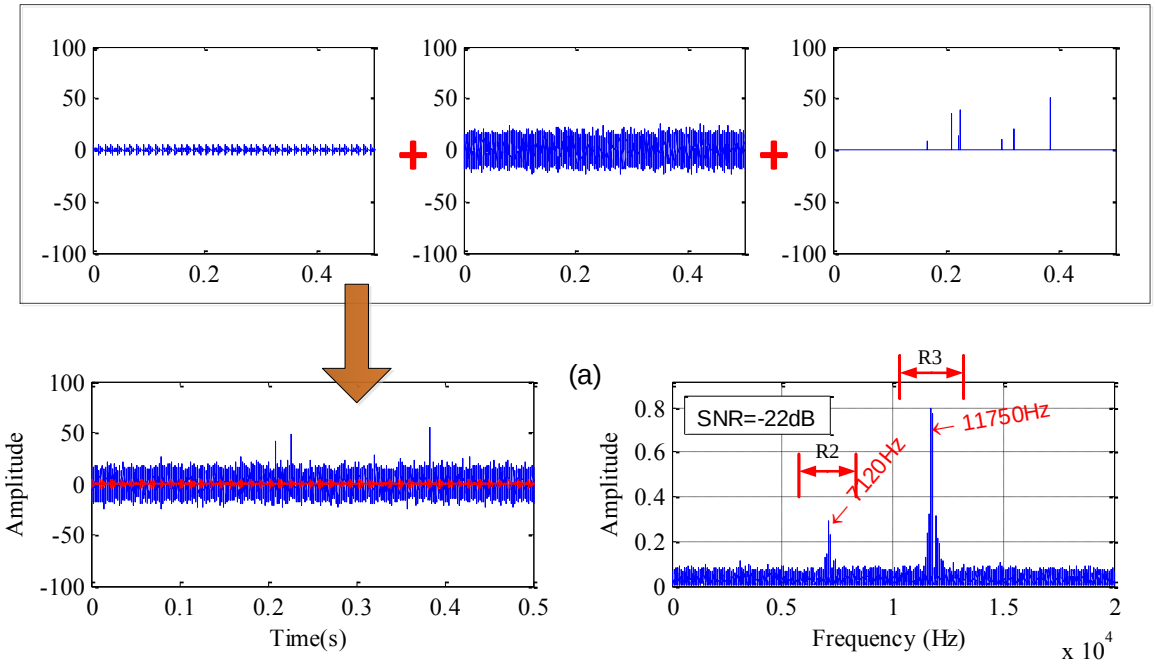

Figure 7.6 Signal waveform and spectra with low level of impulsive interference

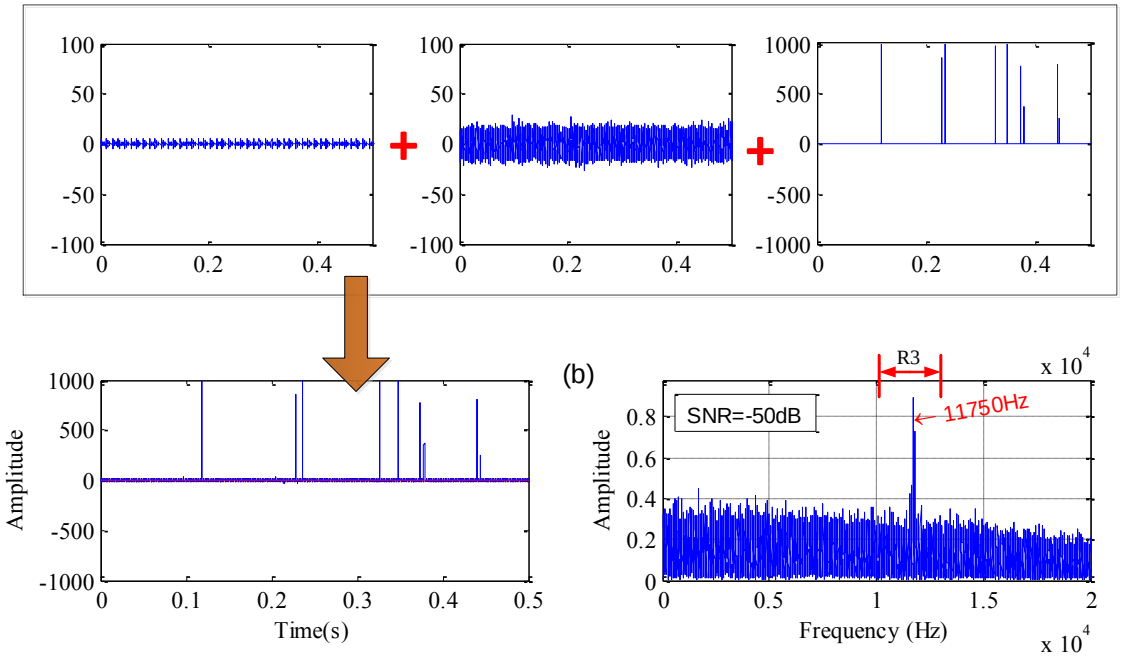

Figure 7.7 Signal waveform and spectra with high level of impulsive interference

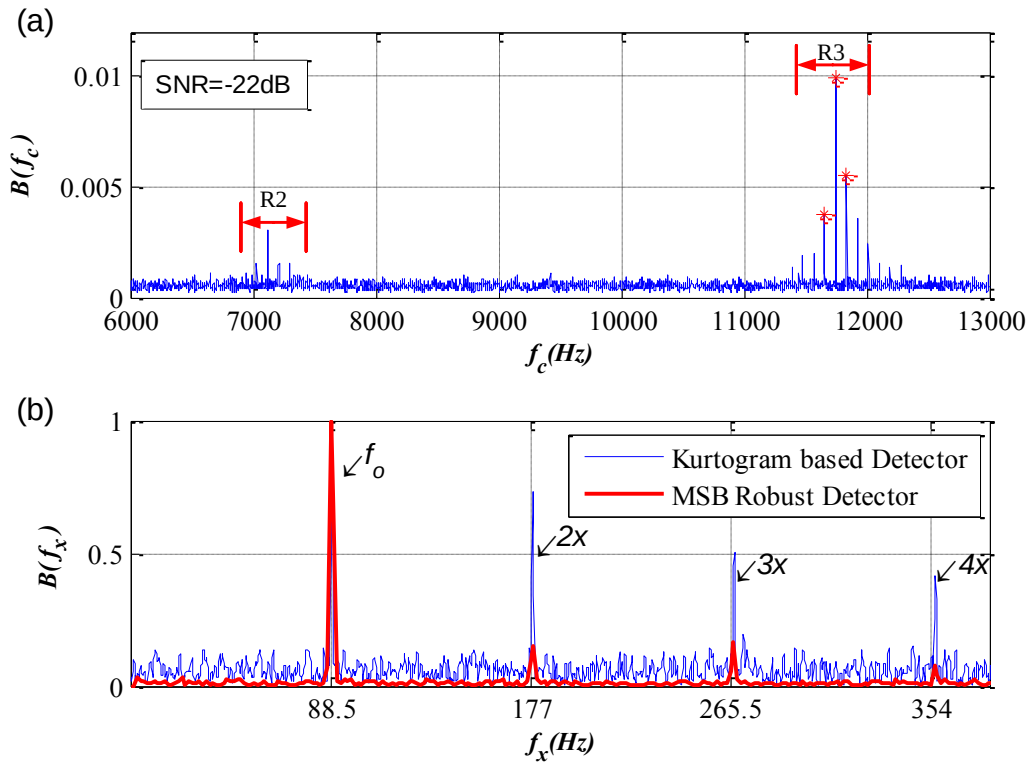

Figure 7.8 The compound MSB slice and results comparison (low noise, low level impulsive interference)

(a) The compound MSB slice $B(f_c)$; and (b) Results of the MSB robust detector and the kurtogram based detector

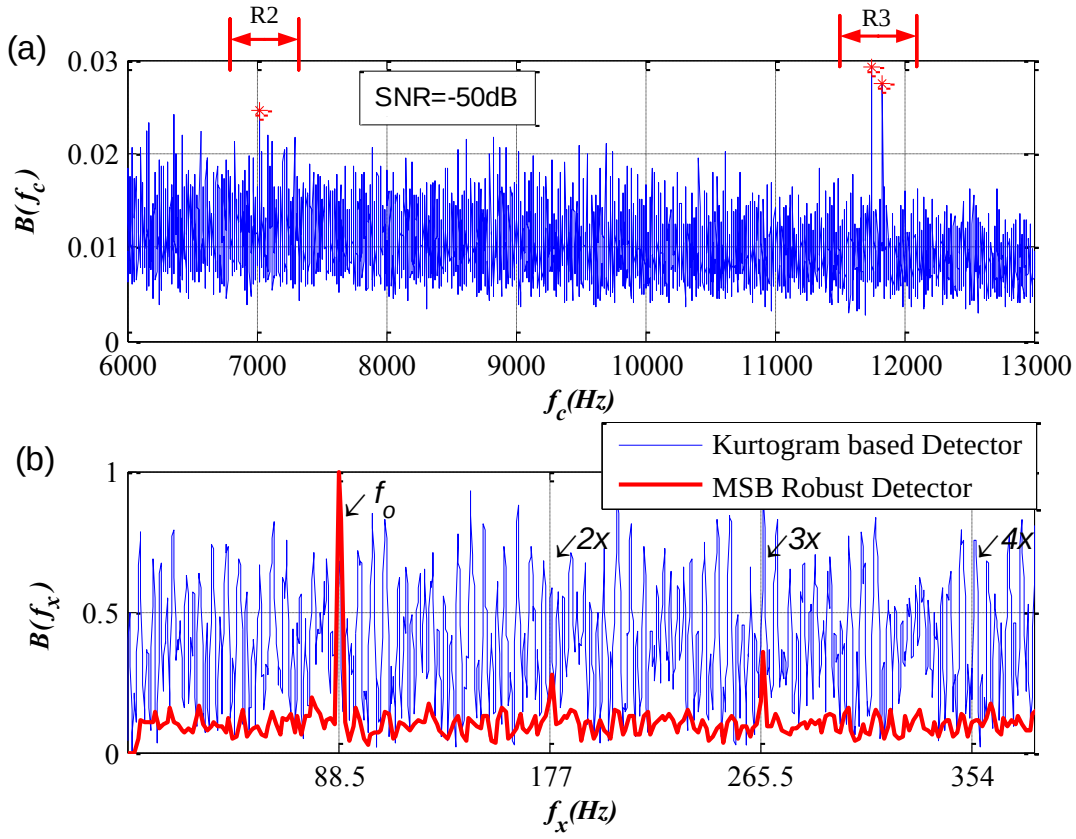

Figure 7.9 The compound MSB slice and results comparison (low noise, high level impulsive interference)

(a) The compound MSB slice $B(f_c)$; and (b) Results of the MSB robust detector and the kurtogram based detector

Figure 7.8 and Figure 7.9 show the results of $B(f_c)$ for both the MSB detector and the kurtogram-based detector, for the two levels of impulsive interference. For the low interference signal, Figure 7.8 illustrates that both detectors are effective but that the MSB-based robust detector has lower background noise. For the high interference signal, Figure 7.9 illustrates that the MSB robust detector still can detect the fault characteristic frequency whereas the kurtogram-based detector contains too much noise and is not able to reveal the presence of the fault.

From the simulation studies, it can be concluded that the proposed MSB detector is robust to white noise and also to aperiodic impulsive interferences, and that in severe noise cases significantly outperforms the kurtogram-based approaches.

7.4 Two 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 mesh processes. The monitoring of the electric motor driver has always been of interest because of its popularity and importance but the condition monitoring of planetary gearboxes has also received much attention in more recent years because of their prominence in wind turbine applications [55][150].

7.4.1 Motor bearing fault detection

7.4.1.1 Experimental setup

The experimental data of the motor bearing analysed in this paper was collected from the bearing test rig illustrated in Figure 7.10. 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. Figure 7.11 shows a photograph of the tested bearing, with a small seeded outer race defect.

The tested motor bearing was an NSK Type 6206ZZ deep groove ball bearing with geometry listed in Table 7.2. 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.

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.

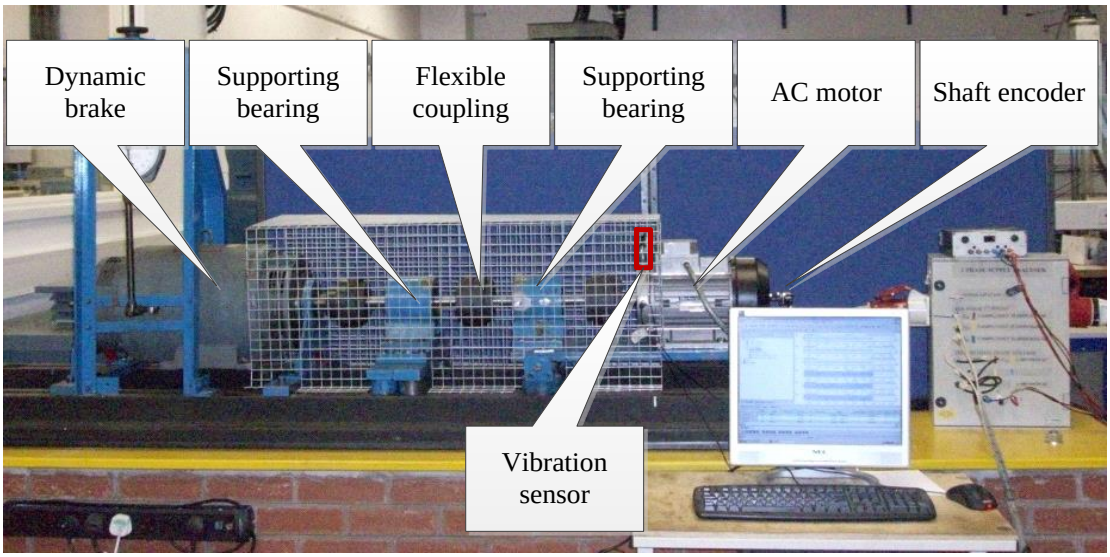

Figure 7.10 Photograph of the motor bearing test rig

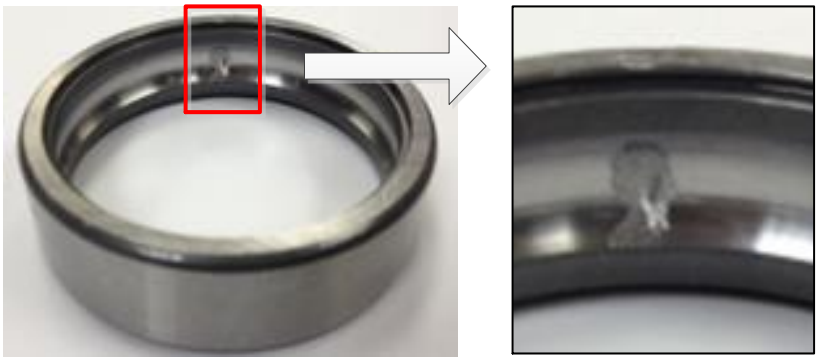

Figure 7.11 Photographs of the test bearing with a small seeded outer race defect

Table 7.2 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˚

7.4.1.2 Detection results and discussion

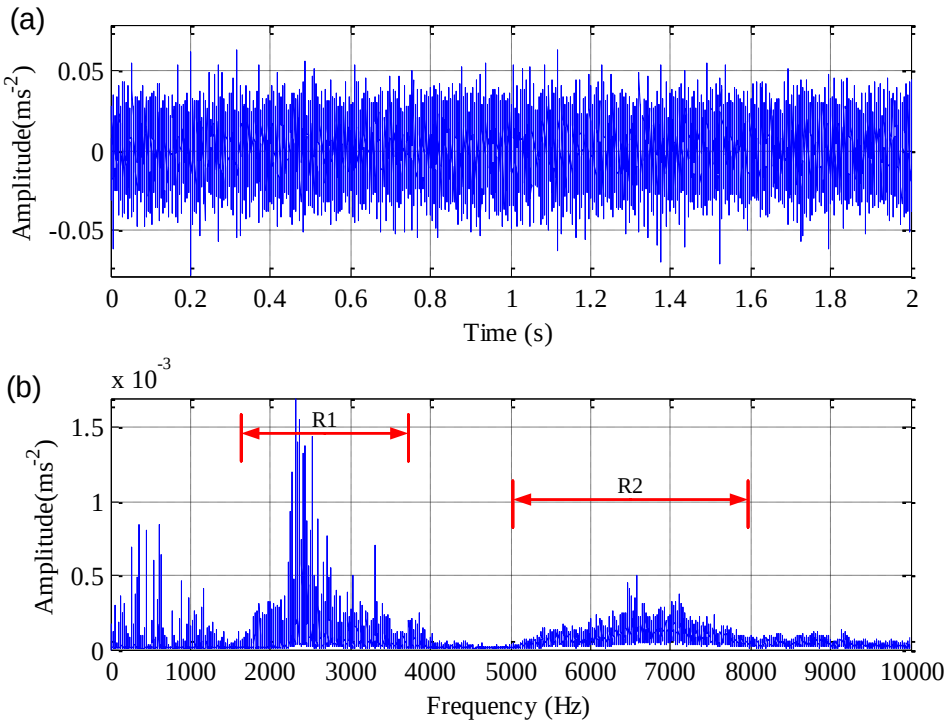

Figure 7.12 Waveform of the motor vibration and its spectrum (a) Waveform; and (b) Spectrum

Figure 7.12 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.

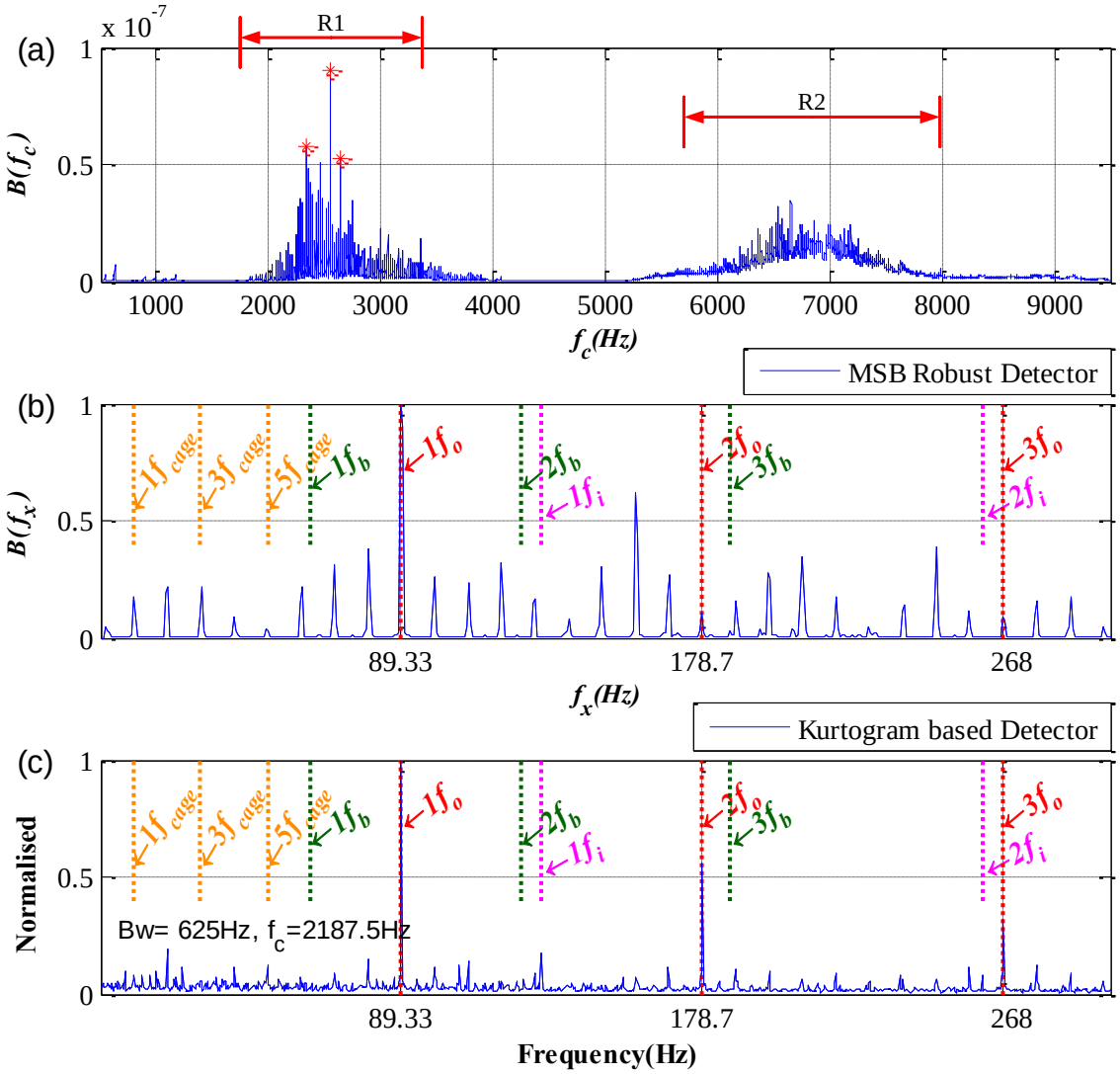

Figure 7.13 Signal processing results

(a) The compound MSB slice $B(f_c)$; (b) Result of MSB robust detector; and (c) Result of kurtogram based detector

Figure 7.13 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 Figure 7.12(b). From the MSB detector results in Figure 7.13(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 and may be caused by manufacturing effects or inadvertent damage during bearing installation. The kurtogram based detector shown in Figure 7.13(c) also provides a clear indication of the bearing outer race defect.

7.4.2 Planetary gearbox bearing fault detection

7.4.2.1 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 Figure 7.14 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 Figure 7.15 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 Figure 7.16.

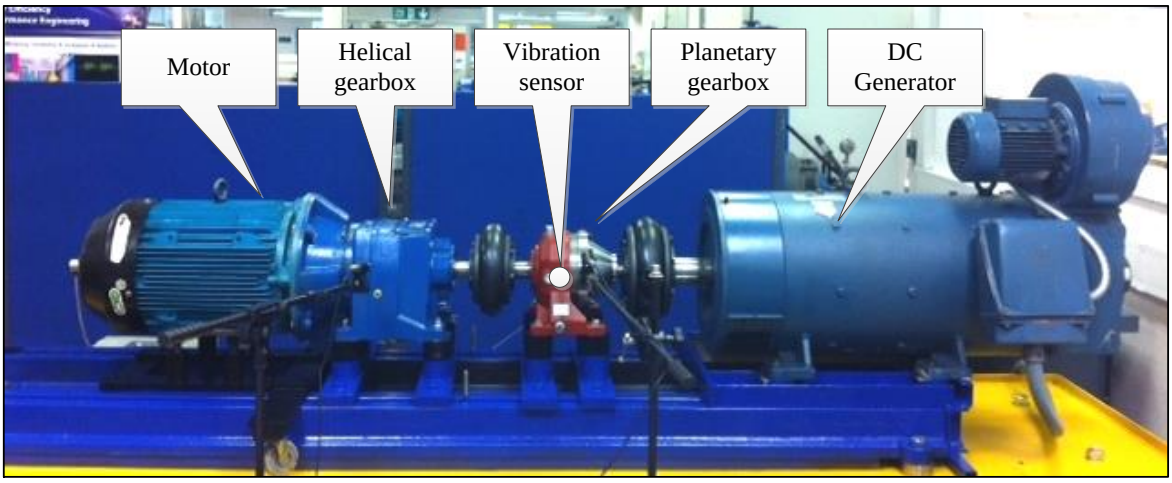

Figure 7.14 Photograph of the planetary gearbox test rig

The tested bearing was an SKF Type 6008 deep groove ball bearing with geometry as listed in Table 7.3. 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 7.3 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˚

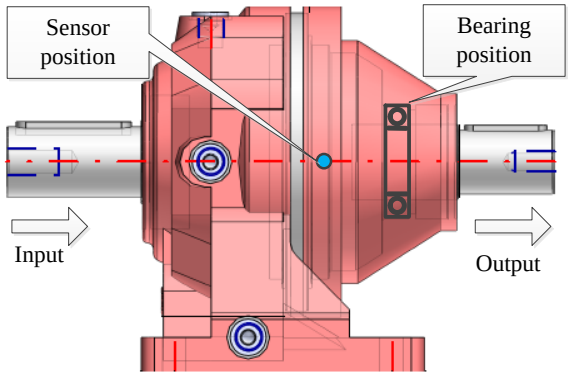

Figure 7.15 Schematic for a planetary gearbox

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.

Figure 7.16 Inner race defect on bearing

7.4.2.2 Detection results and discussion

Figure 7.17 shows the waveform of the measured vibration signal and its spectrum. From Figure 7.17(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. Figure 7.17(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.

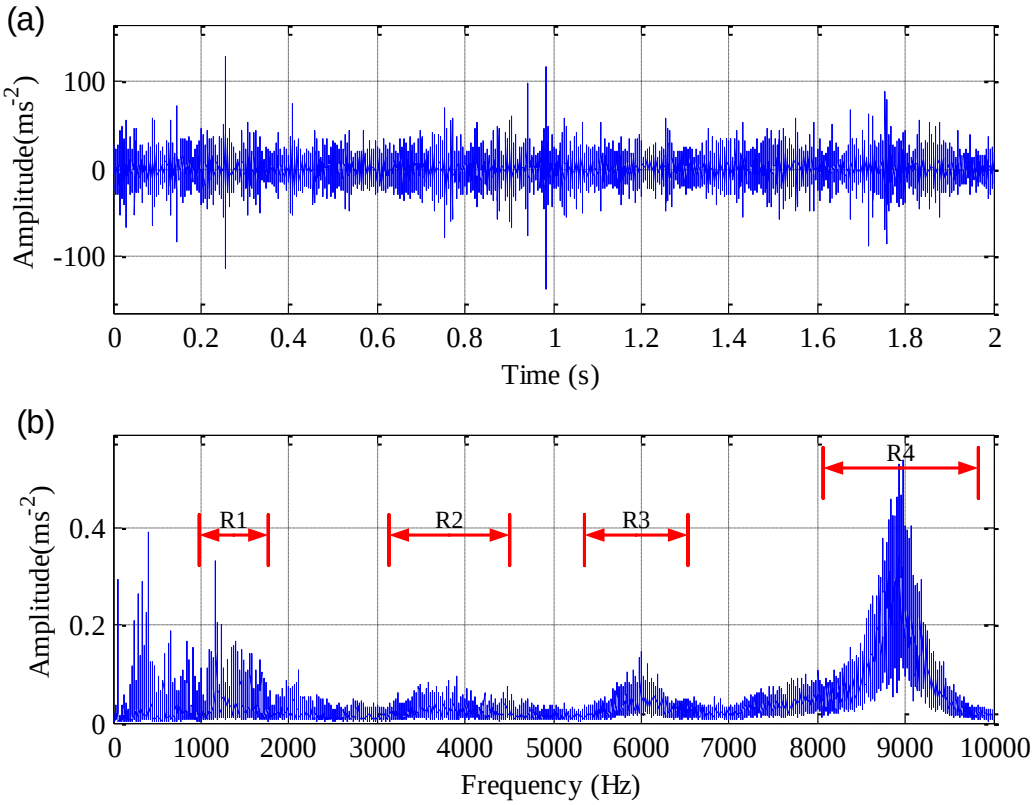

Figure 7.17 Waveform of planetary gearbox vibration and its spectrum

(a) Waveform; and (b) Spectrum

For the planetary gearbox vibration signal analysis, it is necessary to check the nonlinear degree of the signal. Because the vibration of the planetary gearbox is quite complex, it includes both vibrations of gear mesh 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:

$$b_{MS}^2(f_1,f_2)=\frac{\left|B_{MS}(f_1,f_2)\right|^2}{E\left\langle\left|X(f_2)X(f_2)X^*(f_2)X^*(f_2)\right|^2\right\rangle E\left\langle\left|X(f_2+f_1)X(f_2-f_1)\right|^2\right\rangle} \quad \text{.....(7.6)}$$

The MSBc result of the planetary vibration signal is shown in Figure 7.18. 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 from the calculation of the MSB detector.

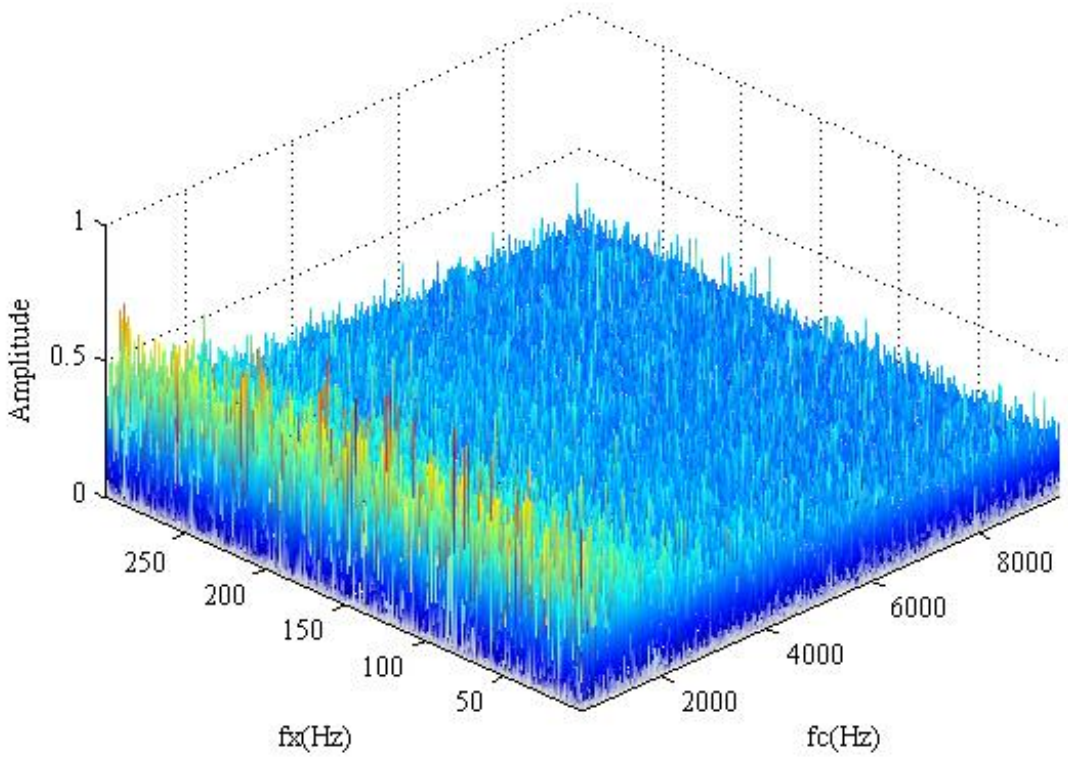

Figure 7.18 MSB coherence of planetary gearbox vibration

$B(f_c)$ is presented in Figure 7.19(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 Figure 7.19(b). The normalised results of the MSB detector are presented in Figure 7.19(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 Equation (7.7) [60] respectively.

$$f_{sf} = K\left(f_{rs} - f_{rc}\right) \quad (7.7)$$

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

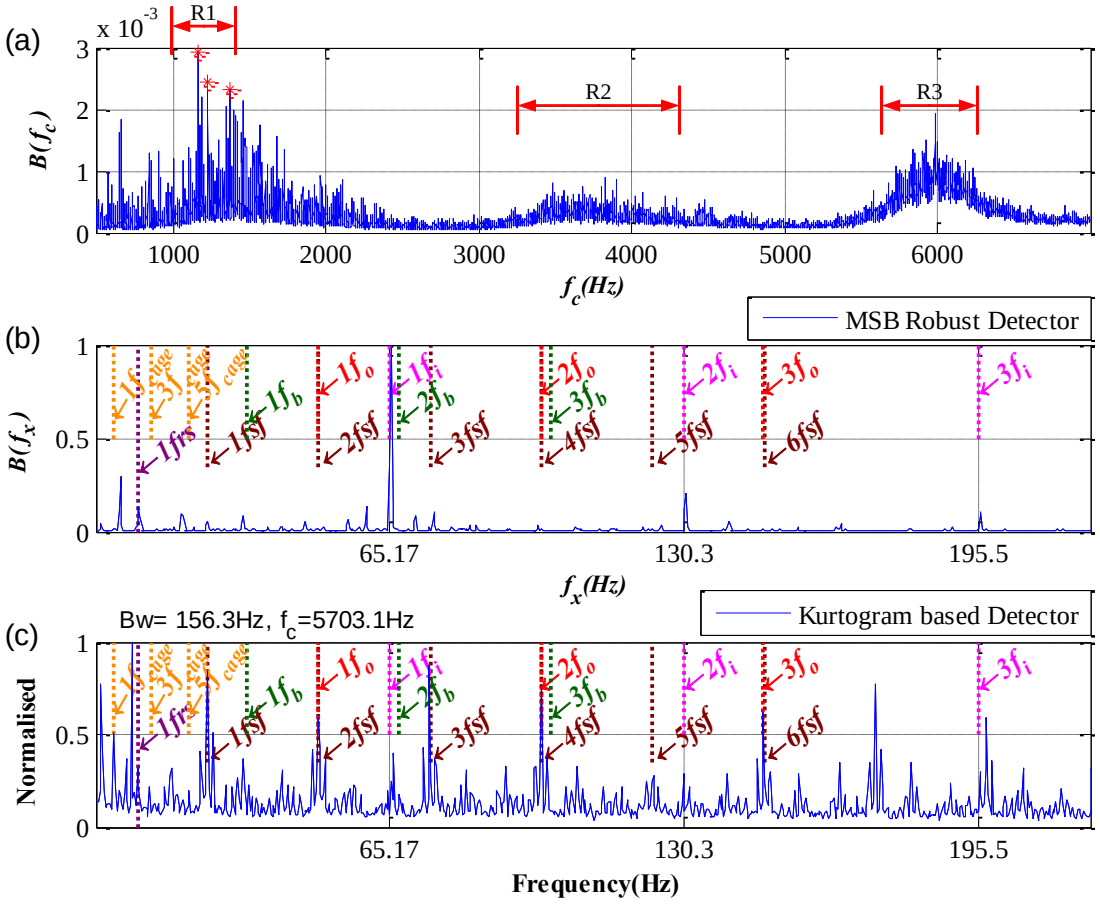

Figure 7.19 Signal processing results

(a) The compound MSB slice $B(f_c)$; (b) Result of MSB robust detector; and (c) Result of kurtogram based detector

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 Figure 7.19(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 the 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 in a low SNR, high impact environment such as the planetary gearbox vibration.

From the results of the kurtogram based detector, illustrated in Figure 7.19(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.

7.5 Summary

Based on the proven performance of the MSB in suppressing random noise and decomposing the nonlinear modulation components [40][53][54][145] a novel MSB detector has been developed using a number of discrete MSB peaks, which are optimal in terms of maximising the modulation contents of bearing fault signals. Simulated signals with different levels of white noise and aperiodic impulsive interference have been applied to demonstrate the robust performance of the new approach, and its capability has been shown to exceed 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 signal contains high levels of modulation due to other impact phenomena.

Chapter 8
A novel residual sideband approach to the fault diagnosis of a planetary gearbox

This chapter presents a novel approach to the fault diagnosis of planetary gearboxes based on the estimation of residual sidebands using MSB. The residual sideband is found to be less influenced by gear errors compared with the in-phase sideband due to the superposition of concurrent mesh excitations of the gearbox. Therefore, the diagnostic feature extracted by applying MSB to the residual sideband can produce an accurate and consistent diagnosis. This has been evaluated by both simulation and experimental studies. The conventional spectrum analysis was also implemented based on commonly used in-phase sidebands for the purpose of performance benchmarking.

8.1 Introduction

Planetary or epicyclic gearboxes are widely used for the power transmission of important machines such as helicopters, automobiles, aircraft engines and marine vehicles due to their large transmission ratios and strong load-bearing capacity. In order to ensure reliable operation, their condition monitoring techniques have been investigated for many years. As shown in the comprehensive review papers by Lei et al. [58], Samuel et al. [151], and Yuksel et al. [152], numerous researches have been carried out on the investigation of vibration characteristics for monitoring various faults such as tooth defects including gear pitting, crack and wear. In particular, many novel signal processing methods in the time domain, frequency domain, time-frequency domain and advanced intelligent methods have been applied to analyse the complicated vibration signals for the purpose of suppressing random noise and interfering components and hence defining accurate and reliable diagnostic features, of which the most representative are the fast dynamic time warping [153], spectral kurtosis [154], blind deconvolution denoising [155], ensemble empirical mode decomposition [156], adaptive Morlet wavelet and singular value decomposition [83], cyclo-stationary analysis [157] and adaptive stochastic resonance signal enhancement [158]. Moreover, the majority of these have developed the diagnostic parameters based on apparent vibration components which have large amplitudes presenting in vibration spectrum and also can be correlated explicitly to fault dynamics. Particularly, studies such as in [59][151] have shown that these large spectral components are the in-phase sidebands due to the effect of a constructive wave superposition from the concurrent meshing sources between the sun and multiple planet gears.

However, due to common manufacturing errors such as the accumulative pitch deviations, eccentricities etc. there will be inherent asymmetry between the multiple meshing sources in an operational planetary gearbox. Therefore, the wave superposition will also result in observable sidebands at the characteristic frequencies which are different from the in-phase sidebands. These sidebands may contain high quality information for diagnostics. Unfortunately, they have been overlooked previously because of their small amplitudes.

In addition, higher order spectrum analysis approaches have also not been applied to the vibration signals from planetary gearboxes, although they have been demonstrated to be effective for denoising the vibration signals from fixed shaft gearboxes [159] and estimating weak sidebands in motor current signals [160].

Therefore, this study will focus on characterising the small sidebands which are usually not as noticeable in the spectrum of planetary gearbox vibration. It will show that they contain reliable diagnostic information because they are much less influenced by inherent manufacturing errors. Firstly, the chapter demonstrates the existence and characteristic of these small sidebands by numerical studies. Then, an MSB based approach is applied to extract these sidebands accurately which are easily contaminated by random noise due to their small amplitudes. Finally, an experimental study is carried out to verify the effectiveness and performance of this approach in diagnosing different types of gear faults.

8.2 Planetary gearbox vibration based diagnosis

8.2.1 Characteristic frequencies for fault detection

As shown in Figure 8.1, a common planetary gearbox consists of three planet gears of the same size, meshed with one sun gear and ring gear concurrently. The carrier is floating and affixed to the output or input shaft by means of splines which allow it to move axially as required for uniform load sharing between the three planetary gears. In this chapter, an industrial planetary gearbox, detailed in Table 8.1, is focused on for fault diagnosis study.

As shown in many previous studies, detection and diagnosis can be carried out by examining the changes of characteristic frequencies around mesh frequency f_m and its harmonics. Considering that there are *K* number of planetary gears moving with the carrier, characteristic frequencies around the meshing frequency can be calculated [58][59] for different local faults occurring on the sun gear

$$f_{sf} = \frac{f_m}{z_s} = K(f_{rs} - f_{rc}), \quad (8.1)$$

on the planet gear

$$f_{pf} = 2\frac{f_m}{z_p} = 2(f_{rp} + f_{rc}) \quad (8.2)$$

and on the ring gear

$$f_{rf} = \frac{f_m}{z_r} = Kf_{rc} \quad (8.3)$$

However, as shown in [59][161], only some of these expected sidebands will be apparent in the vibration spectrum when a planetary gearbox has faults due to the effects of constructive superposition of the vibration waves from the three gear sets, whereas other

sidebands are hard to see because of the destructive effect of the superposition, and hence the latter have been largely neglected by previous studies when developing methods for fault diagnosis.

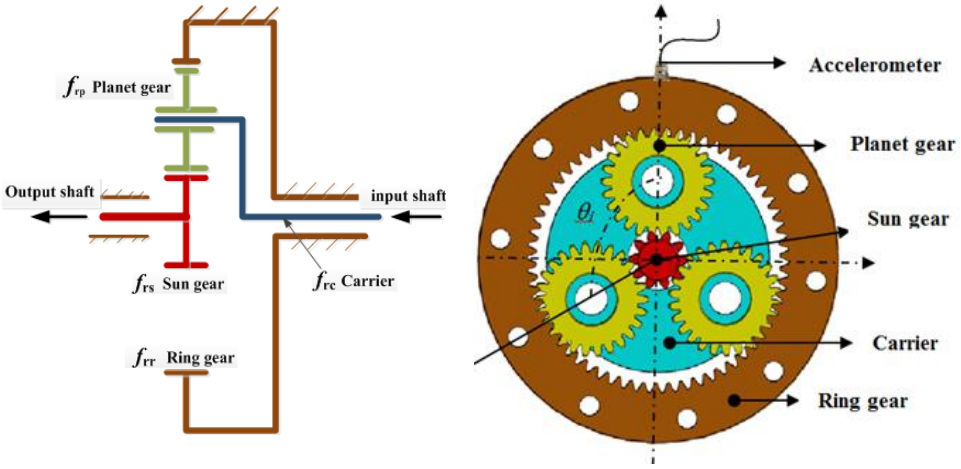

Figure 8.1 Schematic of a planetary gearbox

Table 8.1 Planetary gearbox specification

No	Gear	Teeth	Frequency
1.	Ring	$Z_r = 62$	$f_{rr} = 0$
2.	3×Planets	$Z_p = 26$	f_{rp}
3.	Sun	$Z_s = 10$	f_{rs}
4.	Input shaft		f_{rc}
5.	Output shaft		f_{rs}
6.	Transmission ratio	7.2	
7.	Maximum torque	670 Nm	
8.	Maximum input speed	2800 rpm	
9.	Maximum output speed	388 rpm	

8.2.2 Characteristics of residual sidebands

According to the signal model developed in [59][60][161], a vibration spectrum can be simulated for the case of a tooth defect on the sun gear of a typical industrial planetary gearbox whose key specification is provided in Table 8.1. Firstly, the spectrum of a single pair of gears is presented in Figure 8.2 in order to understand the spectrum characteristics for the combination of multiple gear pairs. The spectrum includes the components in the frequency range up to the third harmonics of the mesh frequency when a triangle wave was employed as the fundamental meshing source and a sinusoidal wave as the excitation of accumulative pitch errors of different gears. It shows clearly that different sidebands due to the errors are located around the mesh frequency and its 2nd and 3rd harmonics in a nearly symmetrical pattern. In particular, the simulation includes the 3rd, 4th , 5th and 7th harmonics of the accumulative pitch error on the sun gear to show the characteristics of the 3rd and 6th order of sidebands, as predicted by Equation (8.1). As shown by 'x','+' and '*' markers which denote the 3rd and 6th sidebands corresponding to the three mesh frequencies respectively. The 1st order sidebands can exhibit relatively symmetrical even if the interferences exist between the sideband at $3f_{sf}$ of the 1st mesh frequency and the sideband at $-7f_{sf}$ of the 2nd mesh frequency. Meanwhile, the simulation also includes only the 1st harmonic component of the error on the planet gear and carrier. However, multiple symmetrical sidebands can be observed at $mf_m \pm n(f_{rp} + f_{rc})$ and $mf_m \pm c\,f_{rc}$, illustrated with 'o' and 'square' markers respectively, because of the inclusion of phase modulations during the simulation.

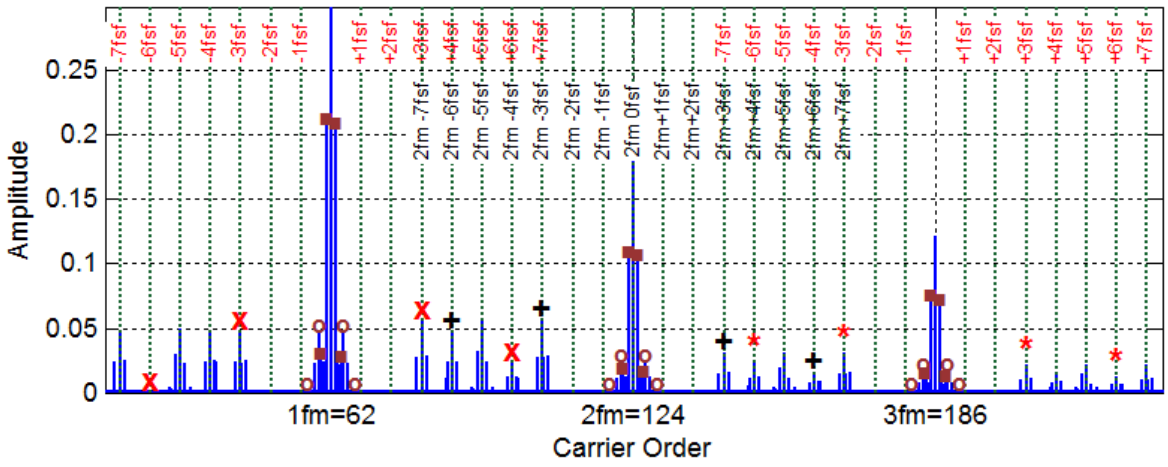

Figure 8.2 Spectrum of single pair of teeth for gears with errors

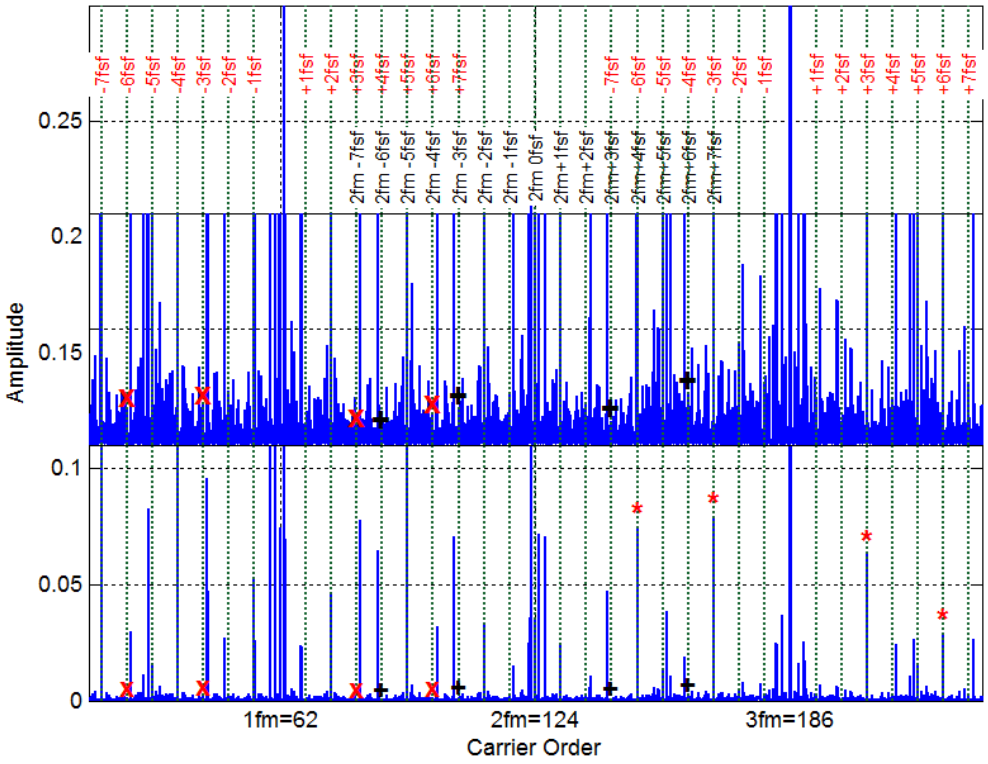

Figure 8.3 Spectrum of superimposed signal with gear error

When superposing over three such signals with equal phase differences of 120°, the resultant spectrum shows very asymmetrical sidebands. As shown in Figure 8.3, the spectral lines with high amplitudes at $1f_m - 7f_{sf}$, $1f_m - 6f_{sf} + 1f_{rc}$, $1f_m - 5f_{sf} - 1f_{rc}$ etc. and at $3f_m \pm 3f_{sf}$ and $3f_m \pm 6f_{sf}$, $3f_m \pm 6f_{sf}$ become larger due to the constructive effect of the in-phase superposition between multiple vibration sources, compared with that of the single pair of teeth meshing process. This spectral pattern agrees with that predicted in previous studies such as [60] [161] for indicating the local fault of the sun gear. However, there are still many other observable spectral components which have not been paid attention previously. Particularly, the components that are located symmetrically around the first and the second mesh frequencies at $1f_m \pm 3f_{sf}$, $1f_m \pm 6f_{sf}$, $2f_m \pm 3f_{sf}$ and $2f_m \pm 6f_{sf}$, illustrated by 'x' and '+' markers, are also clearly visible. These components are well-known diagnostic features which are used widely for the fault diagnosis of fixed shaft gearboxes.

In order to examine the usefulness of these small sidebands in diagnosing faults in a planetary gearbox, they are referred to as residual sidebands in this study because they result from the incomplete superimposition of the out-phase waves and have negligible amplitudes in the spectrum. In contrast, the high amplitude sidebands, such as those around

the third mesh frequency at $3f_m \pm 3f_{sf}$ and $3f_m \pm 6f_{sf}$, which have been explored widely previously, are referred to as in-phase sidebands.

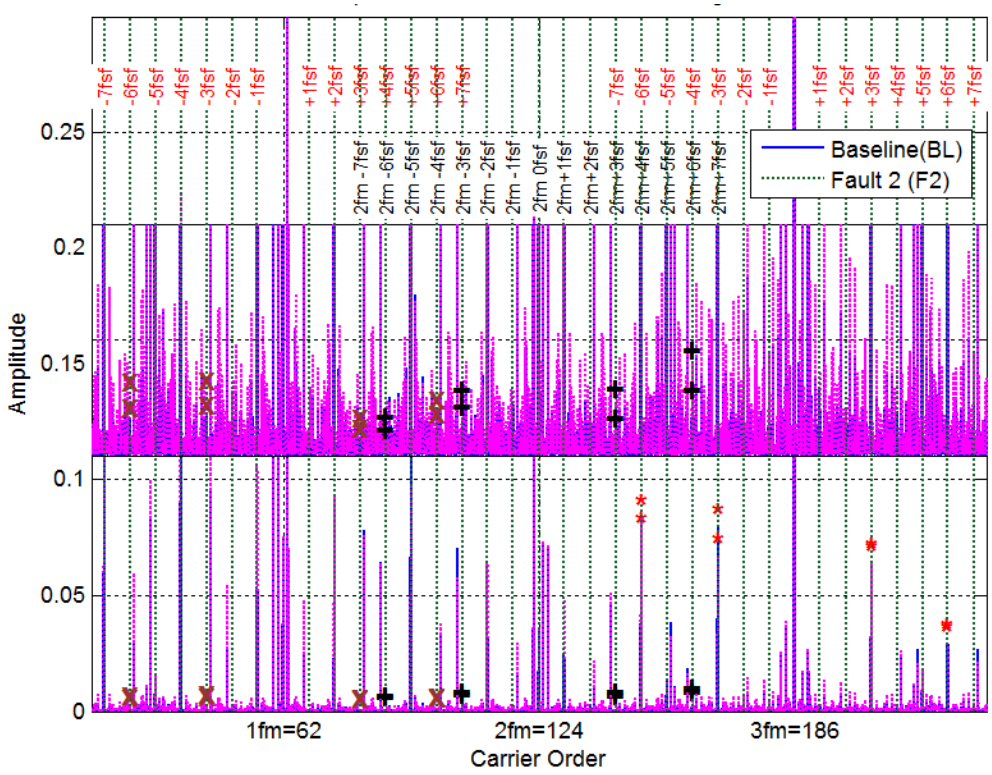

Figure 8.4 Spectrum change for the case of sun gear fault with gear errors

In fact, it is straightforward to understand the existence of the residual sidebands. The development of the existing theory such as [59] for predicting fault components is based on the assumption that the system is symmetrical and meshing between perfecting gears. Thus, it will predict zero amplitudes for the residual sidebands. However, considering a local fault on any of the gears and inevitable manufacturing errors, the assumption does not hold true anymore. Consequently, the superimposition is not impeccable and will result in non-zero components of the residual sidebands and at many other frequencies, as shown in the magnified spectrum of the middle plots in Figure 8.3. In practice, the system can never be symmetrical because of inevitable manufacturing and installation errors. This means that both the residual sidebands and the in-phase sidebands coexist in measured signals and will be influenced by the errors.

To examine the influences of the errors on both types of sidebands and show their effectiveness for fault diagnosis, further simulations were carried out for three cases, denoted as baseline (BL), Fault 1 (F1) and Fault 2 (F2) corresponding to the baseline condition and two increments of tooth breakages on the sun gear respectively. Moreover,

these three cases all include the same effect of accumulative pitch errors, as shown in Figure 8.1. Figure 8.4 shows a comparison of spectra between the baseline and the case of F2. The in-phase components, illustrated by '*' markers with high amplitudes, clearly show differences between the two cases. Obviously, the differences of the in-phase sidebands around $3f_m$, which are illustrated by the differences between the two '*' markers at the same spectral lines, can be used to differentiate these two cases. In the meantime, the residual sidebands, illustrated by the low amplitude arrows in the magnified spectrum, also show sufficient differences to make a clear difference between the two cases. These demonstrate that both types of sidebands can be good indicators for differentiating the faults.

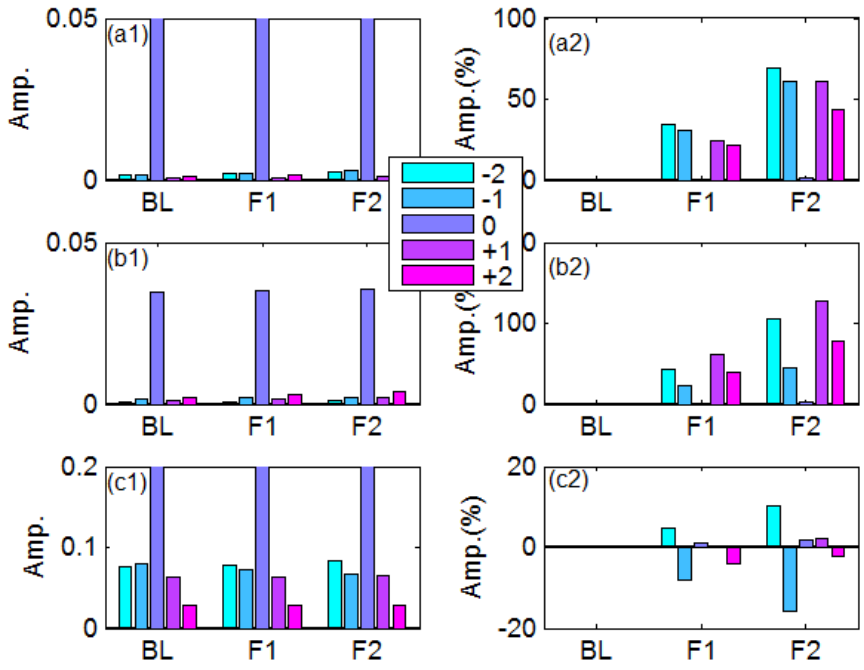

Figure 8.5 Comparison of sideband changes due to faults on sun gear

(a1) Spectral amplitude at $1f_m \pm nf_{sf}$; (b1) Spectral amplitude at $2f_m \pm nf_{sf}$; (c1) Spectral amplitude at $3f_m \pm nf_{sf}$;(a2) Spectral change at $1f_m \pm nf_{sf}$; (b2) Spectral change at $2f_m \pm nf_{sf}$; and (c2) Spectral change at $3f_m \pm nf_{sf}$.

To show accurately the performance of using the residual sidebands for diagnosing the fault on the sun gear, their amplitudes at up to 2nd order components are extracted from the spectrum and compared with those of the in-phase sidebands around the third harmonics. Figure 8.5 presents these amplitudes against the fault cases and their relative changes with respect to their baselines. It shows that the residual sidebands around the first and the

second mesh frequencies increase monotonically with the fault severities, whereas the in-phase sidebands around the third mesh frequency exhibit a non-uniform change. Moreover, the relative changes of the residual sidebands are about five times higher than those of the in-phase sidebands. This means that the residual sidebands are more sensitive and reliable to the fault changes and give rise to a much better result in detecting and diagnosing the fault than the in-phase sidebands. This is mainly because the in-phase superimposition also accumulates the effect of manufacturing errors simultaneously while it accumulates the effect of faults over the three gear sets. For small changes in sideband amplitudes caused by faults, the in-phase sidebands will show a relatively small change. Moreover, the changes can be either incremental or decremental, depending on the consistency of phases between the sidebands due to errors and due to the fault. On the other hand, because the effect of errors is minimised at the residual sidebands, any small changes due to faults will show effectively corresponding changes which are both significant and consistent. Therefore, the residual sidebands can potentially produce a consistent result in agreement with the fault severity and hence can give a more correct diagnosis.

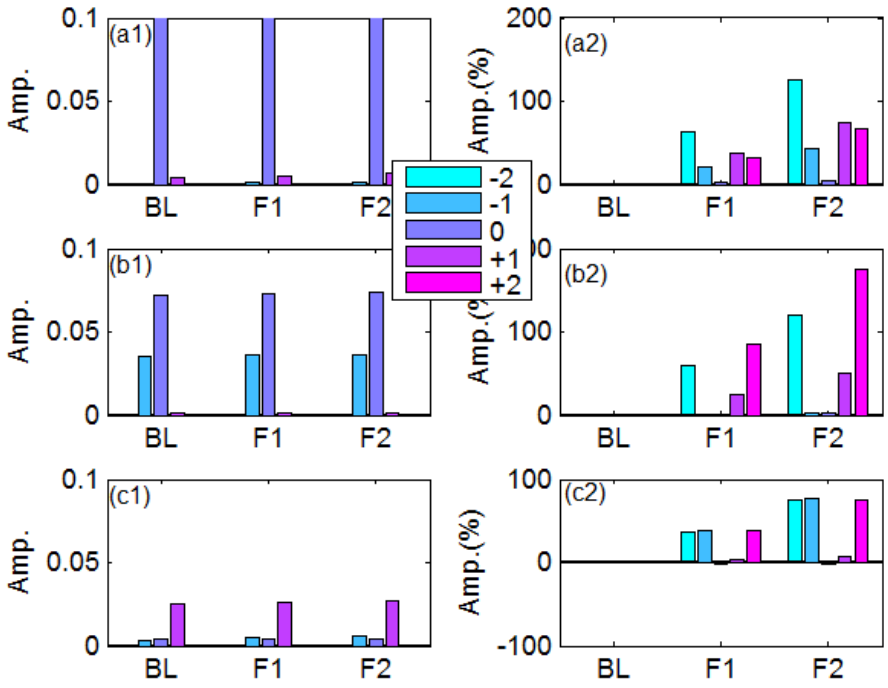

Figure 8.6 Comparison of sideband changes due to faults on one planet gear

(a1) Spectral amplitude at $1f_m + f_{rc} \pm nf_{pf}$ **; (b1) Spectral amplitude at** $2f_m + f_{rc} \pm nf_{pf}$ **; (c1) Spectral amplitude at** $3f_m + f_{rc} \pm nf_{pf}$ **; (a2) Spectral change at** $1f_m + f_{rc} \pm nf_{pf}$ **; (b2) Spectral change at** $2f_m + f_{rc} \pm nf_{pf}$ **; and (c2) Spectral change at** $3f_m + f_{rc} \pm nf_{pf}$ **.**

In the fault case of the planet gear, the residual sidebands are more significant around $mf_m + f_{rc}$ due to the superimposition effect. As shown in Figure 8.6, their amplitudes close to the first, second and third harmonics of the mesh frequencies all show monotonically increasing trends. However, the sidebands close to the first one are more symmetrically distributed around their carrying frequency, indicating less error influences and hence allowing better diagnostic results to be obtained, compared with the other two sets of residual sidebands.

8.2.3 Modulation signal bispectrum

As discussed previously, the residual sidebands have relatively low amplitudes and hence can be easily contaminated by various noises. To extract them this study uses an MSB based sideband estimator (MSB-SE), which has been developed recently by the authors [162][163] in characterising the effects of modulation signals with small amplitudes. For a vibration signal $x(t)$ with corresponding Fourier Transform $X(f)$, MSB-SE can be obtained by

$$B_{MS}^{SE}(f_s, f_c) = E[X(f_c + f_s)X(f_c - f_s)X^*(f_c)X^*(f_c)] / E[|X(f_c)|^2] \quad \text{.....(8.4)}$$

where the product between the upper sideband $X(f_c + f_s)$, the lower sideband $X(f_c - f_s)$ and the normalised carrier component $X^*(f_c)X^*(f_c)/|X(f_c)|^2$ allows the sideband effect to be combined and quantified without the effect of the carrier amplitude. Moreover, because of the average operation, denoted by the expectation operator $E[\ ]$ in Equation (8.4), the sideband products which associate with a constant phase value can be enhanced, while the noise and interfering components with random phases are suppressed effectively. This MSB based approach has been shown to yield outstanding performance in characterising the small modulating components of motor current signals for diagnosing different electrical and mechanical faults under different load conditions [160][163][164][165]. Therefore, it is also evaluated in this study to extract the residual sidebands of vibration signal for the purpose of gear fault diagnosis.

8.3 Experimental setup

8.3.1 Test facilities

To verify the effectiveness of using residual sideband based diagnosis, vibration signals were acquired from an in-house planetary gearbox test system, as shown in Figure 8.7. The

test system consists of a three-phase induction motor of 11kW at 1465rpm, flanged in a cantilever type arrangement to a two-stage helical gearbox, the planetary gearbox, two flexible tyre couplings, and a DC generator for applying load to the gearbox. The helical gearbox is used as a speed reducer with a transmission ratio of 3.6 whereas the planetary gearbox is as an increaser with transmission ratio 7.2 so that the system can be loaded effectively by the DC motor.

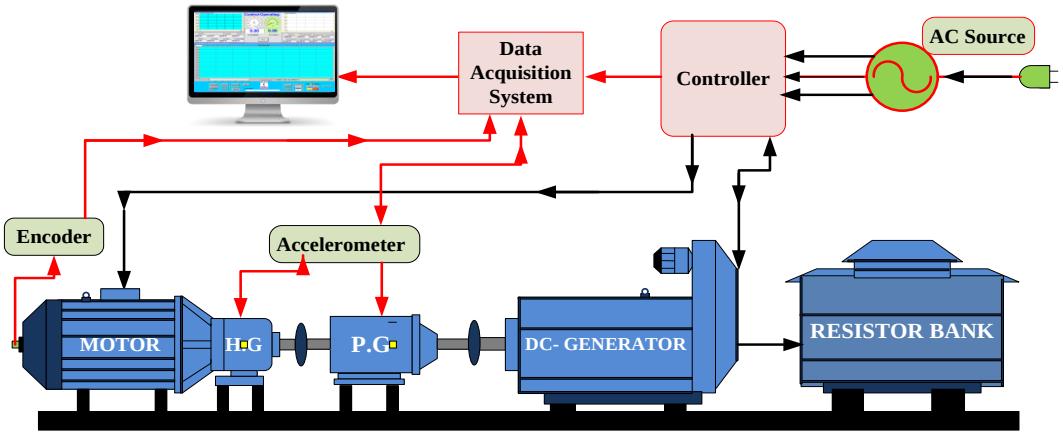

Figure 8.7 Planetary gearbox test system

The schematic in Figure 8.7 also shows the position of the accelerometer mounted on the outer surface of the ring gear. The shaft encoder at the free end of the induction motor produces 100 pulses per revolution for measuring the speed of the system, which is based for identifying characteristic frequencies of gearbox vibration. Furthermore, a variable speed controller is attached to the test rig between the AC power source and the motor to control the speed of the testing system, allowing any specific operating conditions between 0 to 1465rpm and between 0% and 90% of the motor full load to be programmed.

8.3.2 Data acquisition

To examine the influence of the operating condition on fault diagnosis performance, the tests were carried out when the speed of the AC motor was at 40% of the full speed. Considering the transmission ratios, the planetary gearbox will operate at 80% of its full speed correspondingly. At each speed, the system is loaded under five successive loads (0%, 25%, 50%, 75% and 90% of the full load). These operating conditions will allow an exploration of different load influences on vibration contents for developing a reliable diagnostic method.

The vibration is measured by a general purpose accelerometer with a sensitivity of 100 mv/g and frequency response ranges from 1Hz to 10kHz. All of the data were logged

simultaneously by a multiple channel, high-speed data acquisition system operating at 100kHz sampling rate and 16-bit resolution.

8.3.3 Gear faults

Two tests were carried out to examine the tooth faults on the sun and planet gears respectively. The first test is for testing different sizes of defects on one tooth of the sun gear and the second is for testing different sizes of defects on the planet gear tooth. A baseline measurement was firstly carried out when the gears were healthy i.e. without any defects induced. Then, the second measurement was completed when a defect on the sun gear was created manually by damaging a single surface of the tooth by about 30% of the full tooth width. The third measurement was for the case of increasing the defects up to 60%. For the convenience of discussion, these three cases are denoted as BL, F1 and F2, respectively. Figure 8.8(a) shows the two defects respectively.

Likewise, one of the planetary gears was also tested under its baseline and two increments of tooth defect. Figure 8.8(b) shows the two defects respectively, which were simulated on one side of the gear. In addition, both of the tests for the sun gear fault and planet gear fault were conducted on the same gearbox casing but with two different sets of gears for accurate comparison.

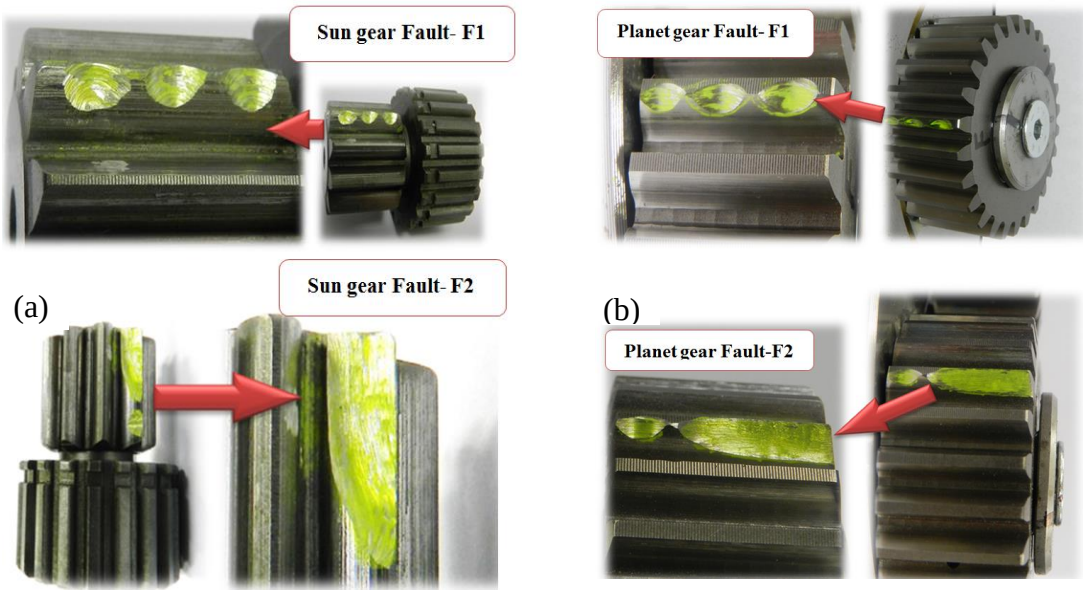

Figure 8.8 Tooth defects simulated on the sun and planet gears

(a) Sun gear defects; and (b) Planet gear defects

8.4 Diagnostic results and discussion

8.4.1 Spectrum features of vibration signals

Figure 8.9 shows the typical spectra from three cases of the sun gear test. It can be seen that three distinctive peaks close to the first three mesh frequencies appear at $f_m + f_{rc}$, $2f_m - f_{rc}$ and $3f_m$ respectively, which agrees with the model prediction and the finding of previous studies [59][60][161]. However, there are also many distinctive peaks between two mesh frequencies. For example, the components at $2f_m - 6f_{sf} - 1f_{rc}$, $2f_m + 7f_{sf}$ etc. should not appear for a healthy planetary gearbox. However, according to the simulation studies in Section 8.2, these peaks may indicate that the gearbox has significant manufacturing and installation errors. This means that the diagnosis based on the in-phase sideband may not be accurate, which can be demonstrated by examining the changes at $3f_m \pm 3f_{sf}$ and $3f_m \pm 6f_{sf}$ between the three cases.

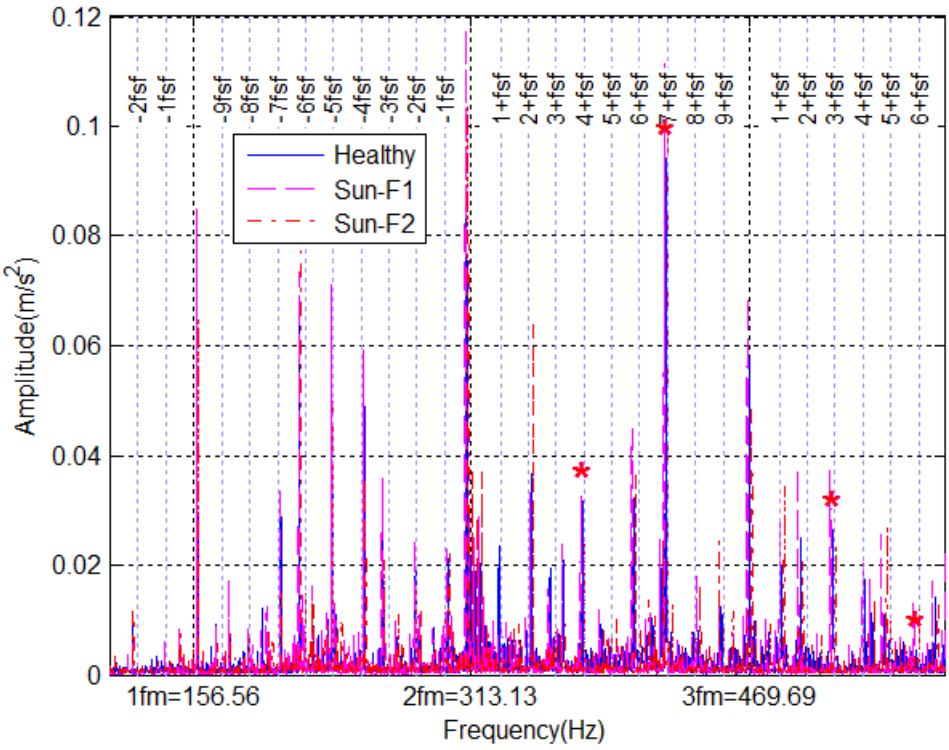

Figure 8.9 Spectrum for different fault cases of the sun gear at 1114 rpm and 75% load

To examine the changes of the residual sidebands, their amplitudes, including those of the in-phase components, are extracted and presented in Figure 8.10. Because of the difference of frequency characteristics due to vibration transfer paths, the spectrum shows very low amplitude below $1f_m$, which is also the general feature of acceleration responses. So the

residual sidebands around $1f_m$ are not explored because of their low SNRs. Based on the changes of these sidebands, it can be found that the amplitudes of the second-order residual sidebands around $2f_m$ and those at $2f_m$ may be good indicators of the fault severity as their amplitudes show an increasing change which agrees with the fault severities under high load conditions.

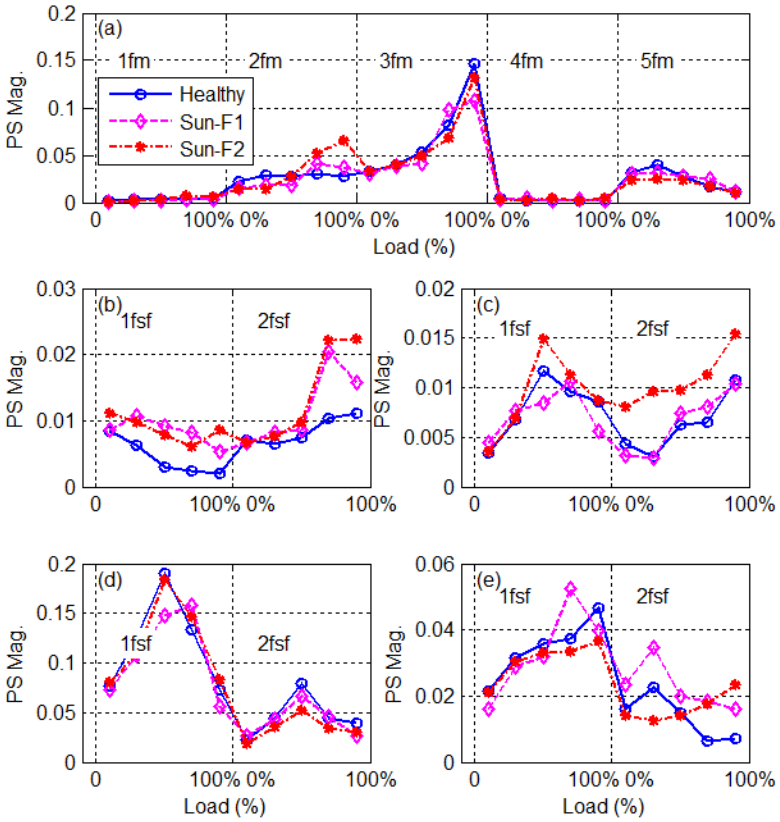

Figure 8.10 Spectral amplitude comparison between the cases of sun gear faults

(a) Spectral amplitudes at mesh frequencies; (b) Lower sidebands at $2f_m$; (c) Higher sidebands at $2f_m$; (d) Lower sidebands at $3f_m$; and (e) Higher sidebands at $3f_m$.

In contrast, the amplitudes of other sidebands and mesh components exhibit high fluctuations between fault cases and loads, so they cannot be used for obtaining a consistent diagnostic result. In particular, the in-phase sidebands around $3f_m$ have higher amplitudes but show small relative changes between different fault cases and load conditions. It means that they cannot be effective indicators of the faults.

In general, the spectrum analysis shows that the residual sidebands allow slightly better diagnostic results, compared with the in-phase sidebands. However, as shown in Figure 8.10 (b) and (c), these residual sidebands still cannot differentiate F1 from the baseline or

make consistent separation appropriately because of their low amplitudes and high noise influences.

8.4.2 MSB features of vibration signals

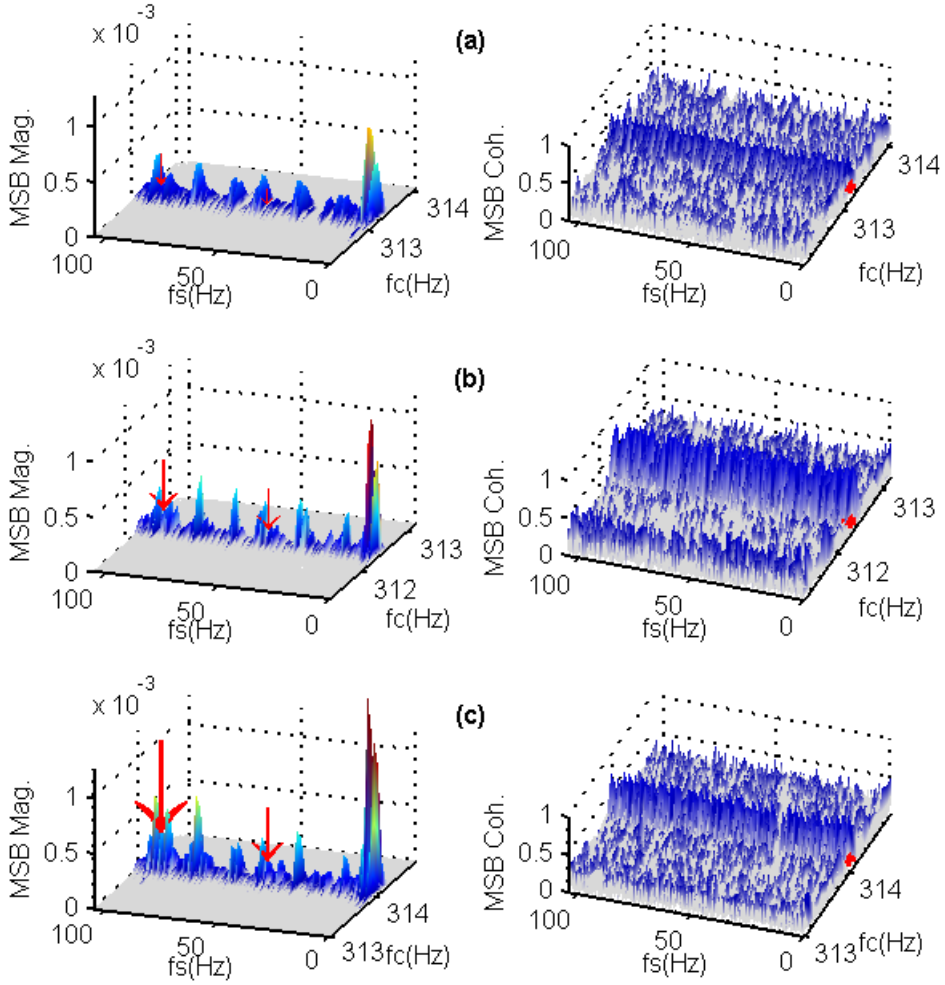

Figure 8.11 MSB results for different cases of the sun gear tests under 75% load

(a) MSB slices around $2f_m$ for baseline; (b) MSB slices around $2f_m$ for sun gear Fault-F1; and (c) MSB slices around $2f_m$ for sun gear Fault-F2.

To improve the diagnostic results, the MSB analysis is applied to corresponding signals. Figure 8.11 shows a typical MSB result for three cases of sun gear conditions under 75% load. To show a clear change of the residual sidebands around mesh frequency $2f_m = 313Hz$, MSB and its corresponding coherence results are presented in the bifrequency domain in the region of $f_c \le 2f_m \pm 1 = 313 \pm 1Hz$ and $f_s < 100Hz$ to include the sidebands up to $6f_{sf}$. It can be seen in Figure 8.11 that the magnitude of many MSB peaks increases with the fault severities. In particular, peaks at bifrequencies $(2f_m, 5f_{sf})$ and

$(2f_m, 5f_{rc})$ show very distinctive amplitudes. In addition, they also vary consistently with the faults. However, these two peaks are also the in-phase sidebands and influenced more by errors, which will be shown in further content. Therefore, they are not considered in this study. Instead, it is the residual sidebands at $(2f_m, 3f_{sf})$ and $(2f_m, 6f_{sf})$ that are interesting in this study for diagnosing the fault severity and location. As illustrated by the size of the arrows, although these sidebands have small amplitudes, they show consistent changes with the sizes of faults induced. Furthermore, these small amplitudes are fully supported by the high amplitudes of their corresponding coherence results.

Moreover, these changes can show a clear difference between the F1 and baseline, which could not be separated based on the spectral amplitudes, as shown in Figure 8.11. This demonstrates that MSB analysis allows more accurate sideband estimation, resulting from its high performance of noise suppression.

8.4.3 Diagnosis of sun gear faults

To explore the performance of the residual sidebands obtained by MSB analysis, the peak values at the corresponding characteristic frequencies for planetary gearbox diagnosis are extracted from MSB results and presented in Figure 8.12. From the results of residual sidebands obtained from the MSB slice at $2f_m$, it can be seen that the amplitudes at the sun gear fault frequency f_{sf} show a good increasing trend with loads, which agrees with the load characteristics of gear transmissions. Moreover, these amplitudes show clear incremental differences between three tested cases. Therefore, they can be used for obtaining fault diagnosis reliably.

Meanwhile, the amplitude changes for other characteristic frequencies are also provided to assure the diagnostic results. These changes exhibit high fluctuations with the fault progression and the load increases, which are not consistent with the gear dynamic characteristics in that the fault usually causes higher vibrations and also increases with load. Therefore, they cannot be used to indicate the corresponding faults but are caused by refitting errors.

Therefore, the fault location can be identified by checking the feature that the increase in residual sidebands occurs over several different loads simultaneously.

For performance comparison, the results from in-phase sidebands around $3f_m$ are also presented in Figure 8.13. As shown by the amplitudes at f_{sf}, these in-phase sidebands

cannot indicate the incremental changes caused by the faults consistently, which agrees with the results from the simulation in Section 8.2.

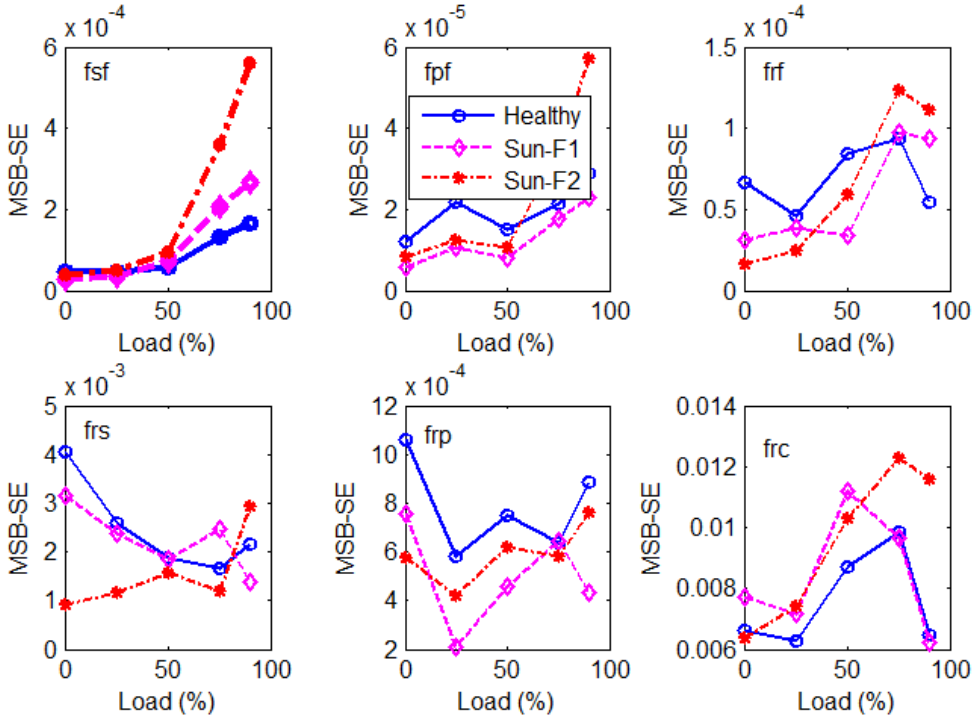

Figure 8.12 MSB diagnosis results of the sun gear faults from the sidebands around $2f_m$

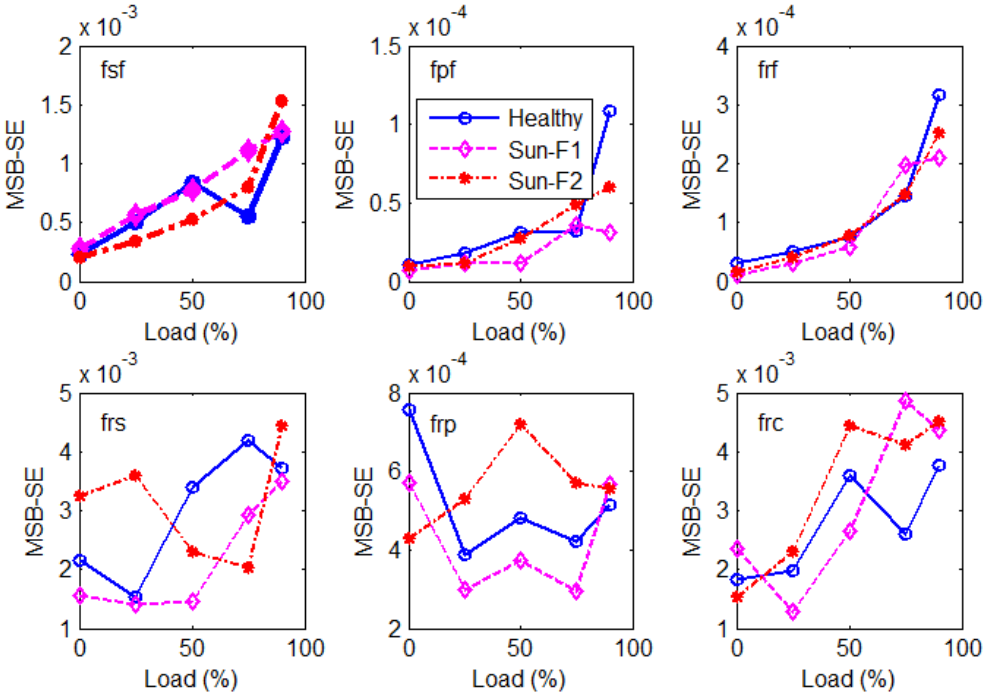

Figure 8.13 MSB diagnosis results of the sun gear faults from the in-phase sidebands around $3f_m$

8.4.4 Diagnosis of planet gear faults

From the above discussion and diagnostic results, it can be concluded that the residual sidebands based method is reliable for planetary gearbox diagnosis. To further evaluate its performance, the method is also applied to the datasets from the planet gear test. Figure 8.14 shows the spectrum of the vibration signals for the three cases of planet gear fault. Compared with the spectrum in Figure 8.9, they show very different spectral characteristics. Particularly, the spectral amplitudes at $2f_m - 6f_{sf} - 1f_{rc}$, $2f_m + 7f_{sf}$ are significantly lower than those in Figure 8.9. This indicates that this gear set may have a lower error distribution relating to the sun gear, even though the results from the conventional spectrum based method cannot give a correct diagnostic result that agrees with the fault severities induced as explained in Section 8.3.3.

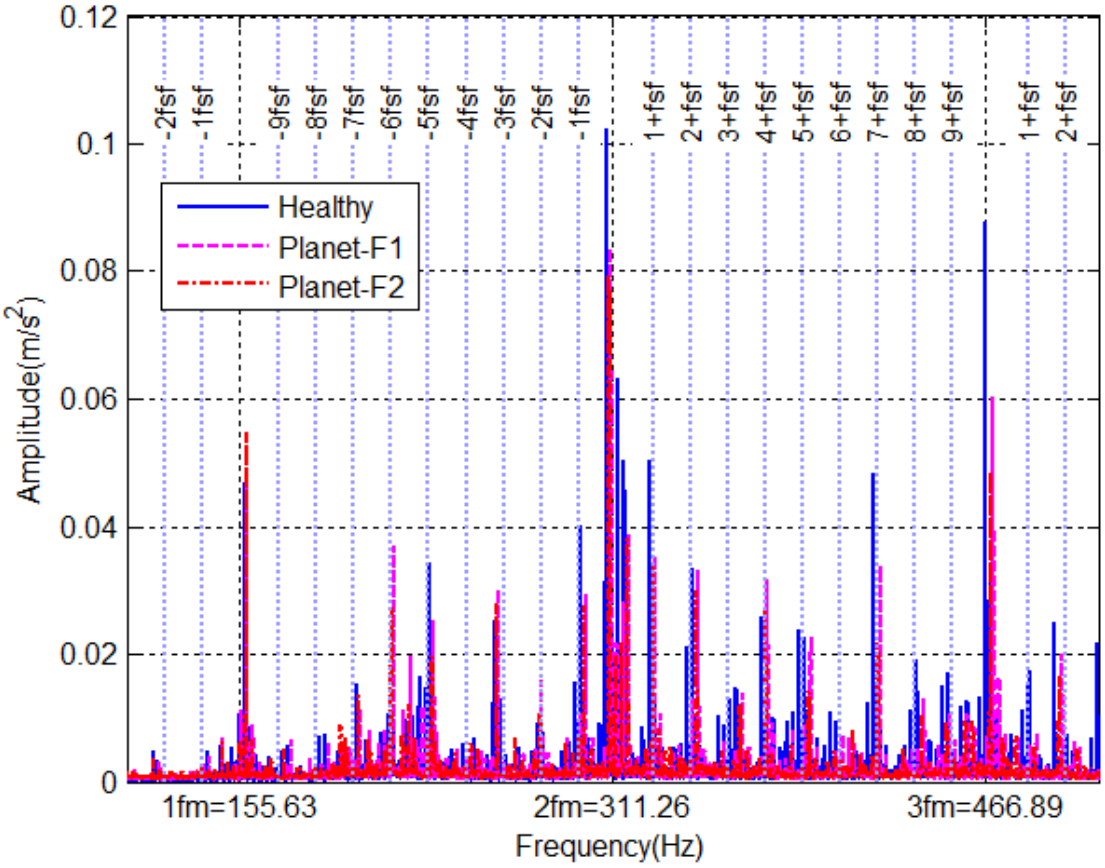

Figure 8.14 Spectrum for different fault cases of the planet gear at 1108 rpm and 75% load

Figure 8.15 presents the MSB based diagnostic results obtained from the residual sidebands around $1f_m$, rather than that of $2f_m$ which is used for diagnosing the sun gear fault case in Section 8.4.3. This is because these sidebands exhibit the features of residual

sidebands much more, i.e. lower amplitudes and less sensitivity to the gear errors. Moreover, they also have less interfering components compared with those around $2f_m$. As shown in the figure, the MSB peaks at planet gear fault frequency f_{pf} show results which increase with fault severity and load, resulting incorrect diagnostic results. In the meantime, the sidebands at f_{rp} also exhibit an increasing change, which can be a secondary feature for the fault on the planetary gear. However, the amplitudes of other characteristic frequencies show increasing or decreasing changes with loads and fault severities which cannot be used to indicate the faults corresponding to the characteristic frequency.

On the other hand, the in-phase sidebands around $3f_m$, shown in Figure 8.16, are also able to separate the faults only when the faults become serious i.e. when the effect of gear error is less than that of the planet gear fault case. However, they cannot separate the smaller fault. This further proves that the residual sidebands integrating with MSB analysis can provide more reliable and accurate diagnostic information than the in-phase sidebands.

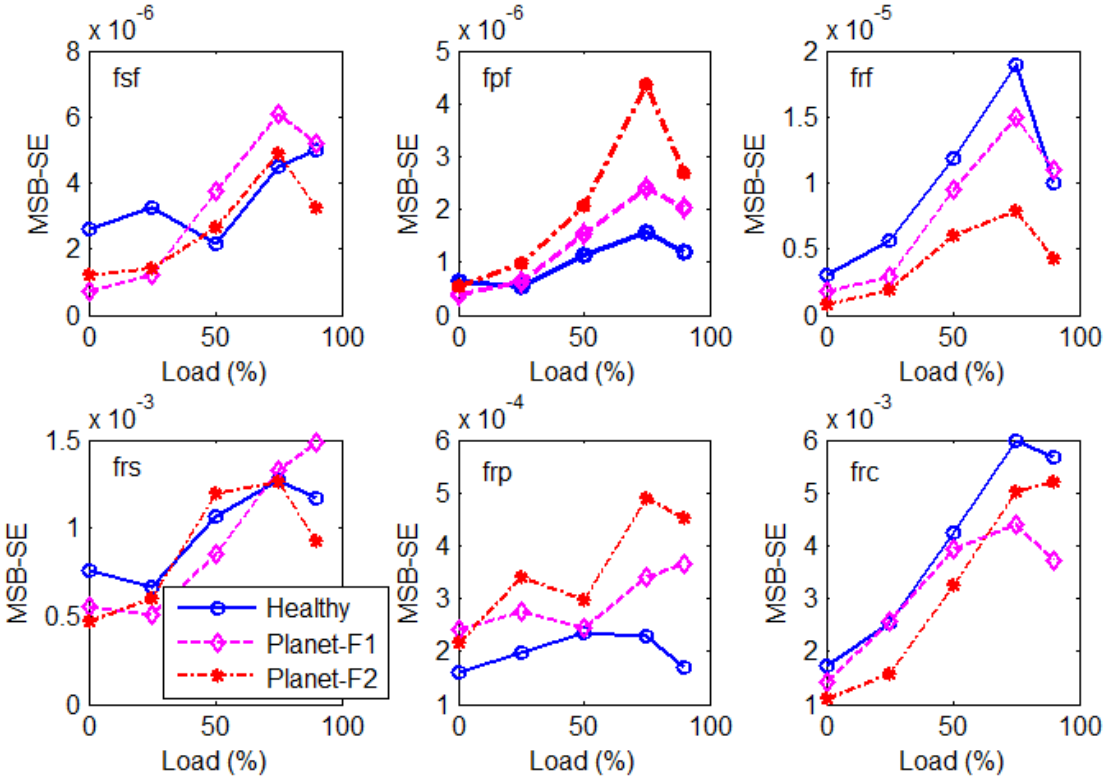

Figure 8.15 MSB diagnosis results for different planet gear faults from the residual sidebands around $1f_m$

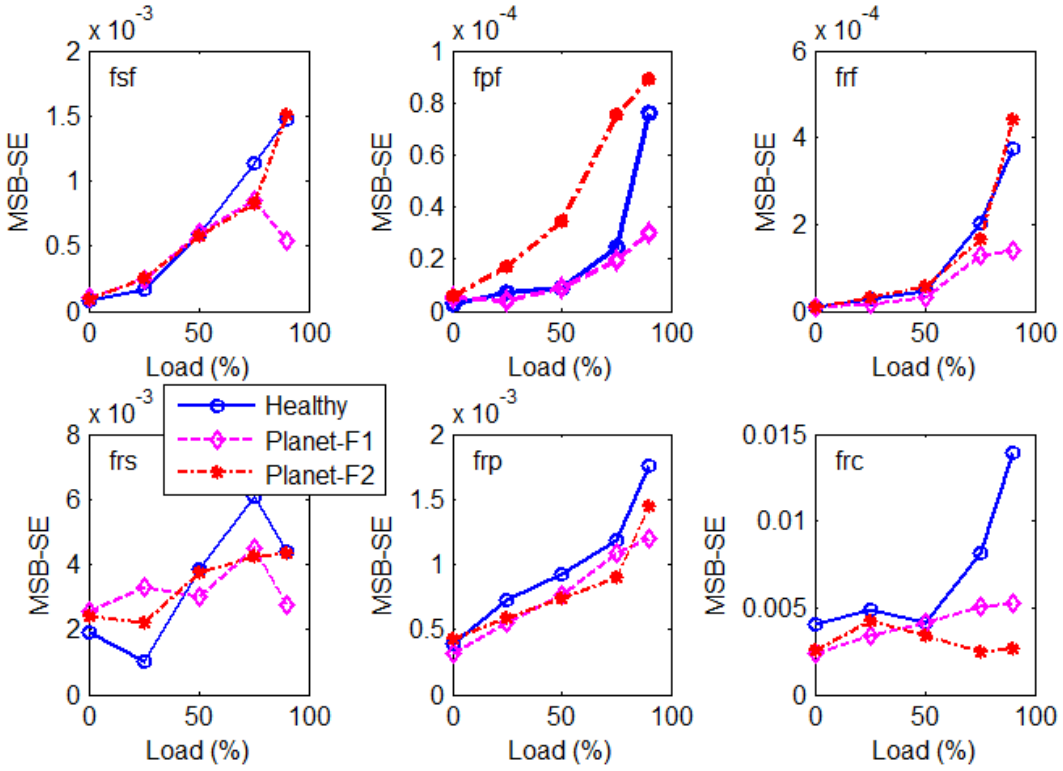

Figure 8.16 MSB diagnosis results for different planet gear faults from the in-phase sidebands around $3f_m$

8.5 Summary

Residual sidebands found in spectrum result from the imperfect superposition of multiple mesh sources in an operational planetary gearbox with inherent errors. Although their amplitudes are relatively small compared with those of the in-phase ones, they are much less influenced by the gear errors. Incorporated with MSB, these small sidebands can be estimated with a good degree of accuracy, which leads to more accurate and consistent diagnostic results. Both simulation and experimental results have shown that this residual sideband based method can provide correct and hence reliable diagnostic results for different sizes of faults on either the sun gear or the planet gear under different loads. On the other hand, the in-phase sidebands can give correct diagnostic results only when the fault amplitude is larger than the effect of errors under high load operations. In addition, it also shows that the MSB based method outperforms the conventional spectrum analysis significantly for diagnosing faults in planetary gearboxes.

Chapter 9
Diagnosing combined faults in a planetary gearbox

This chapter presents a novel method for diagnosing the combined faults of the bearing and the sun gear in a planetary gearbox. Vibration signals measured on the gearbox housing exhibit complicated characteristics because of multiple modulations due to concurrent excitation sources, signal paths and inevitable noise. To separate these modulations accurately, MSB analysis is applied to achieve a sparse representation for the complicated signal contents, which allows for the effective enhancement of various sidebands for extracting the diagnostic information. The proposed method has been applied to diagnose an industrial planetary gearbox in which coexists both a sun gear fault and bearing inner race fault (two kinds of bearing fault severities). The gearbox was operated under different load and speed conditions.

9.1 Introduction

There are some studies on combined fault diagnosis of bearing and gear signals or bearing fault diagnosis in a gearbox. Sawalhi et al. [166] proposed a method based on the TSA. Firstly, it isolates and then removes the deterministic components corresponding to each gear in the system by synchronous averaging, leaving a residual stochastic signal which should be dominated by bearing faults in some frequency bands. Then, the residual signal is applied to cepstrum pre-whitening for bearing fault detection. Vishwash et al. [167] used a multi-scale slope feature extraction technique, which is based on wavelet multi-resolution analysis, discrete wavelet transform (DWT) and wavelet packet transform (WPT), for fault diagnosis of gears and bearings. For planetary bearing fault diagnosis, Bonnardot and Randall et al. [168] presented an enhanced unsupervised noise cancellation that uses an unsupervised order tracking algorithm to perform the noise cancellation in the angular domain. To extract fault features of the rolling element bearing from the masking faulty gearbox signals, Tian et al. [169] explored a method which combines WPT, Pearson correlation coefficient and envelope analysis for bearing and gear fault detection, but did not consider different fault levels and operating conditions. Elasha et al. [170] developed a method for defective bearings in a planetary gearbox by applying an adaptive filter, spectral kurtosis and envelope analysis to both AE and vibration signals. These efforts in improving data quality have shown different degrees of success in diagnosing fault types and severities.

However, little attention has been paid to multiple faults occurring concurrently, with different defect levels and under various operating conditions, which are more significant as the structures of rotating machinery become of larger scale, of higher speed, and more complicated [171]. In addition, these studies have usually focused more on noise reduction and feature extraction but with limited efforts have been made utilising of multiple modulation characteristics to extract the diagnostic information.

To fill these gaps, this chapter presents a new method for combined fault detection of gear and bearing based on MSB-SE analysis of vibration signals, which has been demonstrated to be particularly effective in highlighting sidebands and hence diagnosing faults on gears only [105].

9.2 Theoretical background for diagnosing combined faults

9.2.1 Spectrum features of vibration signals

Figure 9.1 shows the typical spectra for the three cases under the same load. They exhibit complicated patterns and high density of the spectral component, which needs careful examination to find the components of interest. Three distinctive peaks close to the first three mesh frequencies appear at $f_m + f_{rc}$, $2f_m - f_{rc}$ and $3f_m$ respectively, which agrees with the model prediction and that of previous studies [59][60][161]. However, there are also many distinctive peaks between two mesh frequencies. For example, the components at $2f_m - 6f_{sf} - 1f_{rc}$, $2f_m + 7f_{sf}$ etc. should not appear for a healthy planetary gearbox. The presence of these peaks may be from gearbox manufacturing and installation errors. The green dashed lines show the bearing inner race fault frequency and its harmonics. It is obvious that their amplitudes are quite small compared with the other components. This makes it difficult for reliable bearing fault diagnosis. Therefore, the modulation effects between the bearing fault frequency and other characteristic frequencies such as f_{rs} and f_{rc} are used for bearing fault diagnosis.

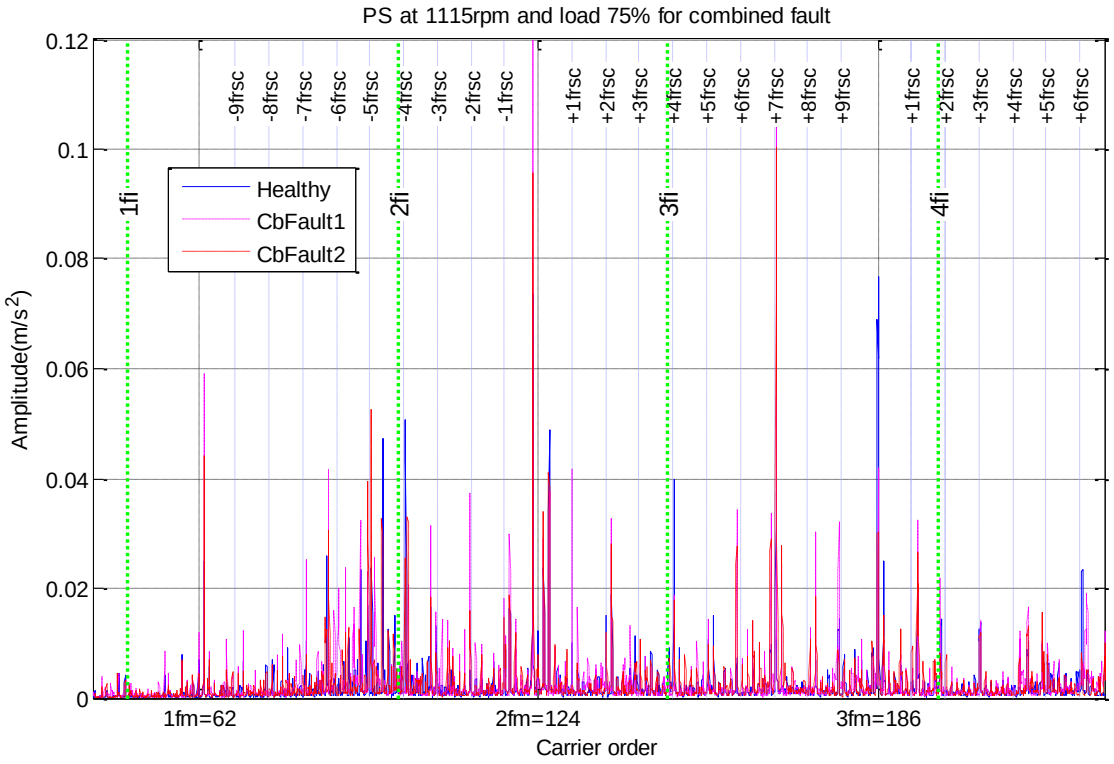

Figure 9.1 Spectra for different fault cases of the gearbox at 1115 rpm and 75% load

9.2.2 MSB features of vibration signals

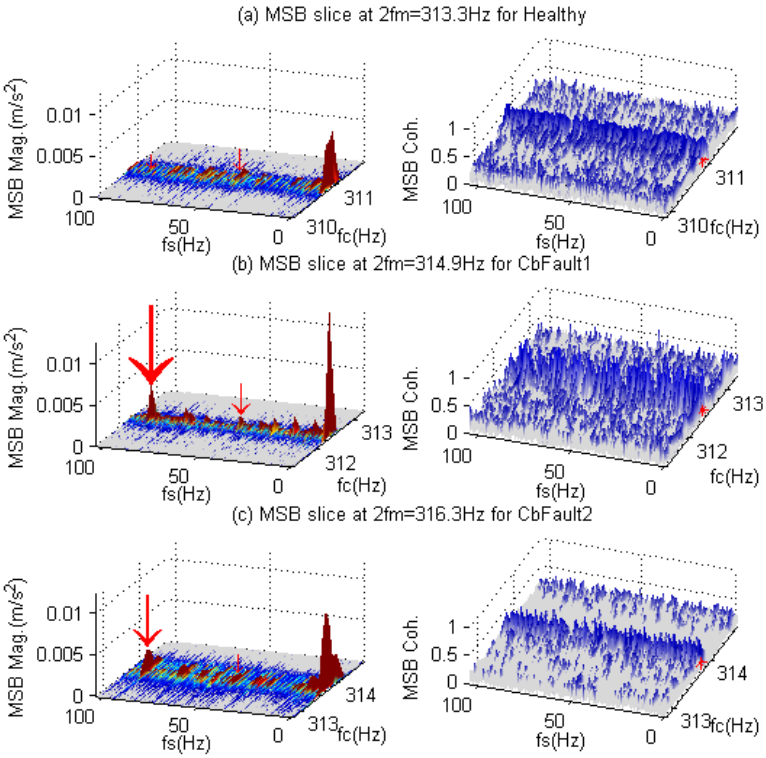

Figure 9.2 MSB results for different cases of the tests under 75% load

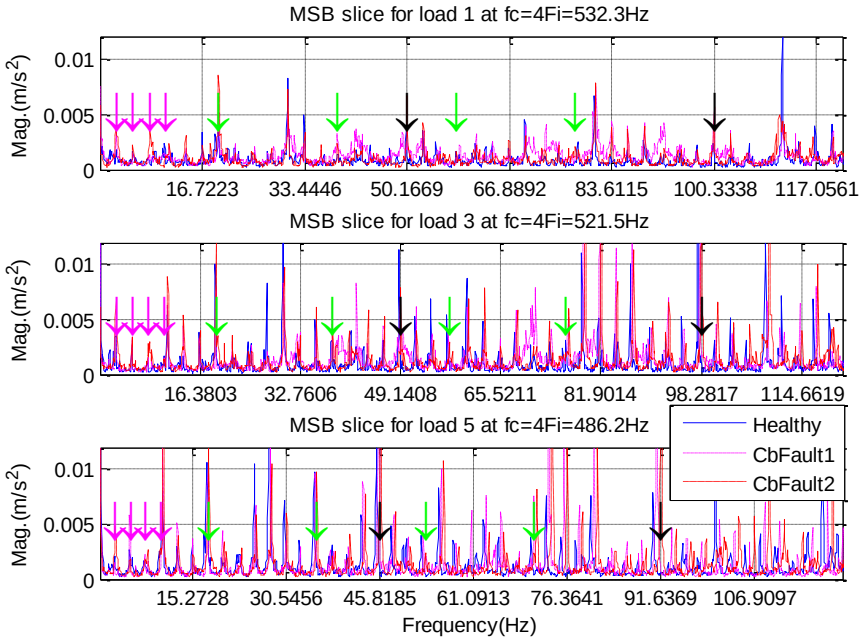

Figure 9.3 MSB slice for different cases at $f_c = 4f_i$

Figure 9.2 shows typical MSB results for the three cases of test under 75% load. To show a clear change of the residual sidebands around the harmonic of the mesh frequency $2f_m = 313Hz$, MSB and its corresponding coherence results are presented in the

bifrequency domain in the region of $f_c \le 2f_m \pm 1 = 313 \pm 1Hz$ and $f_s < 100Hz$ to include the sidebands up to $6f_{sf}$.

Figure 9.3 illustrates the MSB slices at $f_c = 4f_i$ for bearing fault detection. The pink, green and black arrows show the sidebands at f_{rc}, f_{rs} and f_{sf} respectively. These sidebands show differences between the baseline, small bearing inner race defect and large bearing inner race defect cases.

9.3 Experimental setup

To verify the effectiveness of MSB-SE based diagnosis, vibration signals were acquired from an in-house planetary gearbox test system as shown in Figure 9.4. The maximum torque of the planetary gearbox is 670Nm, the maximum input speed is 2800rpm and the maximum output speed is 388rpm. The maximum speed of the motor is 1470rpm. The helical gearbox is used as a speed reducer and the transmission ratio is 3.6. The planetary gearbox is used as a speed increaser and the transmission ratio is 7.2. The maximum speed of the DC generator is 1750rpm. Overall, the maximum speed of the DC generator limits the operating speed of the AC motor. Considering the safety of the system, the maximum speed of the experiment is limited to 80% of the DC generator's maximum speed. Therefore, the corresponding maximum speed is at 40% of the motor's full speed. The schematic in Figure 9.5 shows the position of the accelerometer that was mounted on the outer surface of the ring gear and the position of the studied bearing.

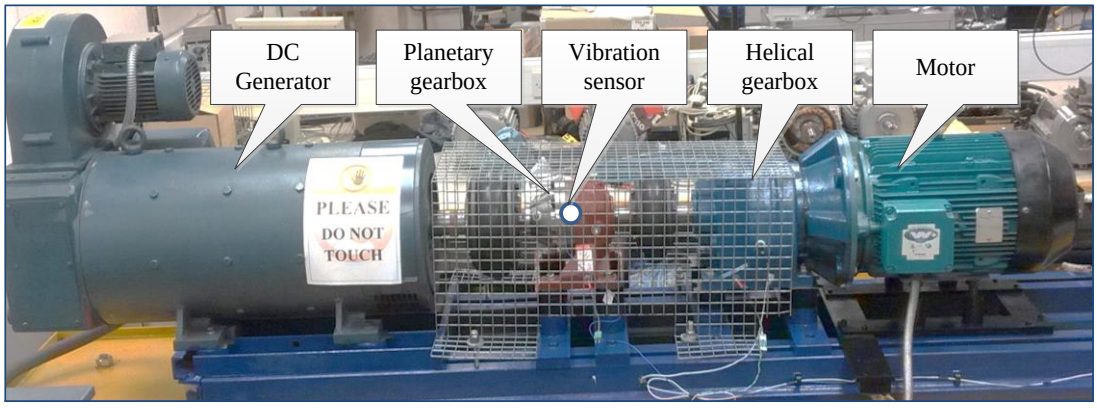

Figure 9.4 Photograph of planetary gearbox test rig

In the experiment, the planetary gearbox operates at 80% of its full speed under 5 load conditions (0%, 25%, 50%, 75% and 90% of the full load). The load setting allows fault diagnoses to be examined with variable load operations, which are the cases for many

applications such as wind turbine, helicopters etc. The vibration is measured by a general purpose accelerometer with a sensitivity of 31.9 mv/(ms^{-2}) and frequency response ranging from 1Hz to 10kHz. All of the data were logged simultaneously by a multiple-channel, high-speed data acquisition system with 100kHz sampling rate and 16-bit resolution.

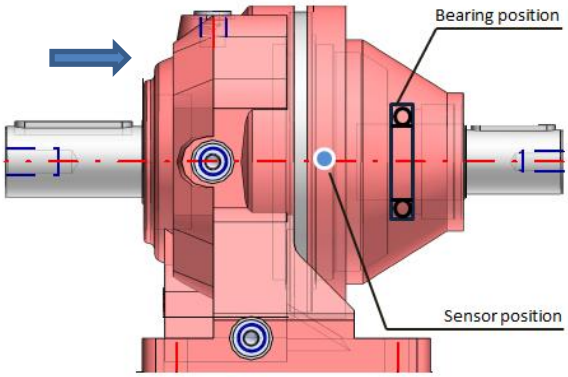

Figure 9.5 Schematic for a planetary gearbox

Figure 9.6 Tooth defects simulated on the sun gear and two kinds of inner race defect on deep groove ball bearing

Table 9.1 Specification of the 6008ZZ deep groove ball bearing

Parameter	Measurement
Pitch Diameter	54 mm
Ball Diameter	7.398 mm
Ball Number	12
Contact Angle	0

Three cases of test were carried out to examine the combined faults. The first is the healthy case, in which there is no defect on either the gear or the bearing. The second is for the combined fault of a small bearing inner race defect and sun gear tooth defect. The third is for the combined fault of a large bearing inner race defect and sun gear tooth defect. For convenience of discussion, these three cases are denoted as Healthy, CbFault1 and

CbFault2, respectively. Figure 9.6 shows the defects on the sun gear and bearing inner races.

The experiment bearing is located on the output shaft of the planetary gearbox, which is close to the sun gear. It is a 6008ZZ deep groove ball bearing and its geometric dimensions, which are needed for fault frequency calculation, are listed in Table 9.1.

9.4 Combined fault diagnosis at high speeds

Gearboxes probably work under different speeds. To simulate different work conditions, two kinds of speeds are applied during the experiment. The sun gear fault is diagnosed with the sidebands around mesh frequencies. The bearing fault is detected with the sidebands of two kinds of carrier frequencies: one is the gear mesh frequency and another is the bearing fault frequency.

9.4.1 Diagnosis of sun gear fault

The diagnostic result of the MSB for the sun gear fault is presented in Figure 9.7. From the MSB results obtained from the MSB slice at $2f_m - f_{rc}$, it can be seen that all of the fault frequencies and rotational frequencies show good increasing trends. The diagnostic results for the sun gear are presented in Figure 9.8. From the results of the residual sidebands obtained from the MSB slice at $2f_m - f_{rc}$, it can be seen that only the amplitudes at the sun gear fault frequency f_{sf} show a good increasing trend with loads, which agrees with the load characteristics of gear transmissions. Moreover, these amplitudes show clear incremental differences between three tested cases under high load.

The corresponding MSB coherence results are printed in Figure 9.9, which can be used to assure the reliability of the MSB-SE results. From the figure it can be seen that the MSB coherences are low at the planet gear fault frequency f_{pf} and the ring gear fault frequency f_{rp}, which indicates that there is no significant modulation phenomenon at these two sidebands. Also, it means that there is no fault on the planetary gear or ring gear. Meanwhile, the amplitude changes for other characteristic frequencies are also provided to assure the diagnostic results. These changes exhibit high fluctuations with the fault progression and load increases, which are not consistent with the gear dynamic characteristics in that the fault usually causes higher vibrations and also increases with load. Therefore, they cannot be used to indicate that the corresponding faults are only

caused by refitting errors. The fault location can be identified by checking the feature that the increase in residual sidebands occurs over several different loads simultaneously.

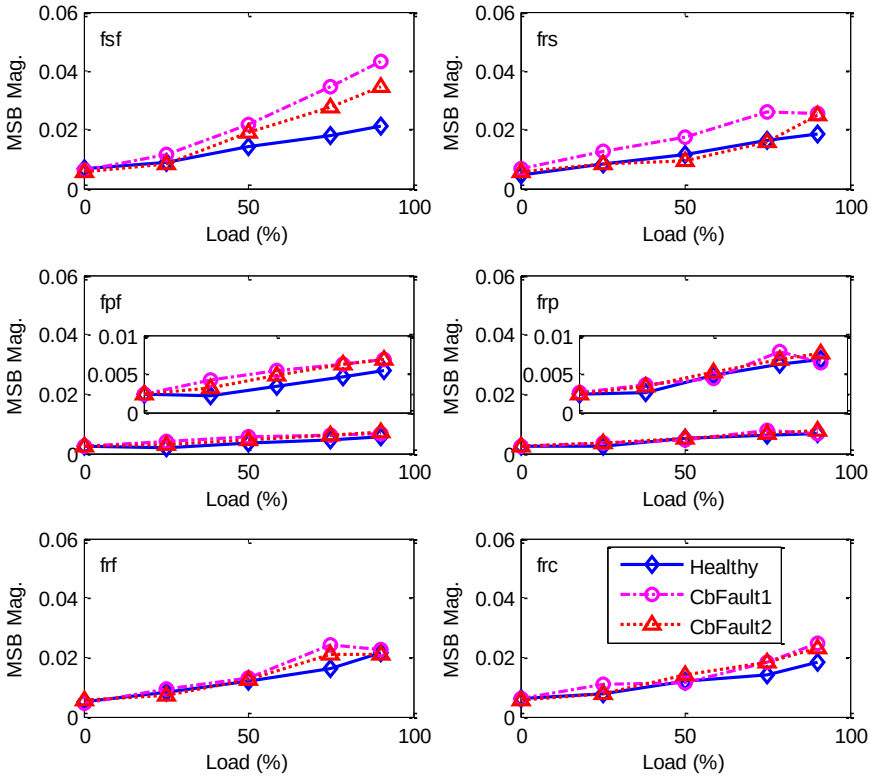

Figure 9.7 MSB diagnosis results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

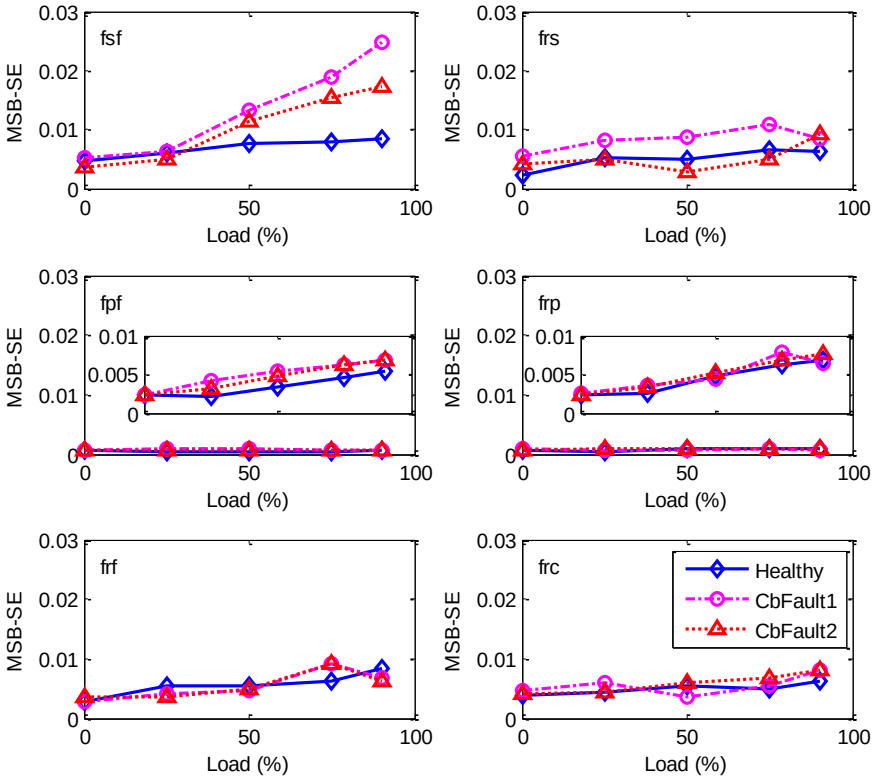

Figure 9.8 MSB-SE diagnosis results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

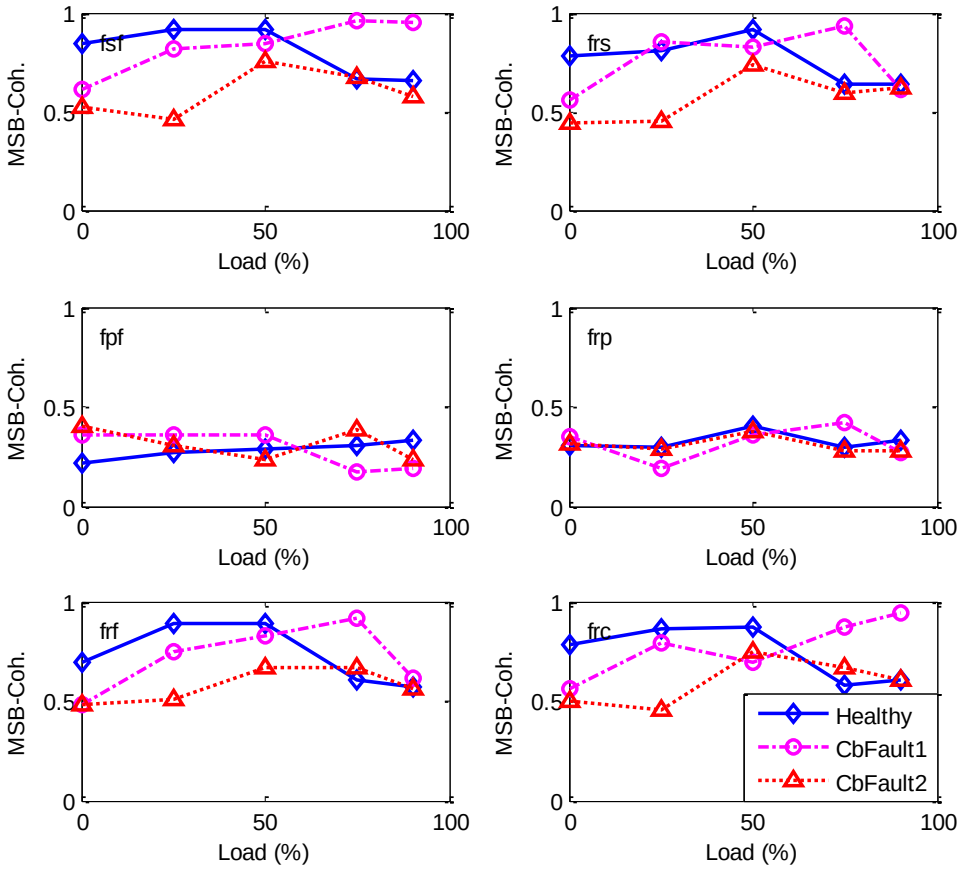

Figure 9.9 MSB-Coh. results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

9.4.2 Diagnosis of bearing fault based on f_m

The bearing characteristic frequencies can be modulated by the gear mesh frequency during the transmission from the vibration sources to the gearbox house. Therefore, it is possible to detect the bearing frequencies around the gear mesh frequency and its harmonics in theory. To evaluate this method, the bearing fault detection is investigated using the sidebands around the gear mesh frequency and its harmonics. Figure 9.10 and Figure 9.11 show the MSB-SE and MSB coherence results obtained from the slice at $4f_m$, respectively. Although the MSB coherence results show that all the feature frequencies have high amplitudes, only the amplitudes of sidebands at the bearing inner race fault frequency f_i show good increasing trends with the load. This means that the defect on the bearing inner race can be detected. However, the two kinds of inner race fault severities cannot be separated.

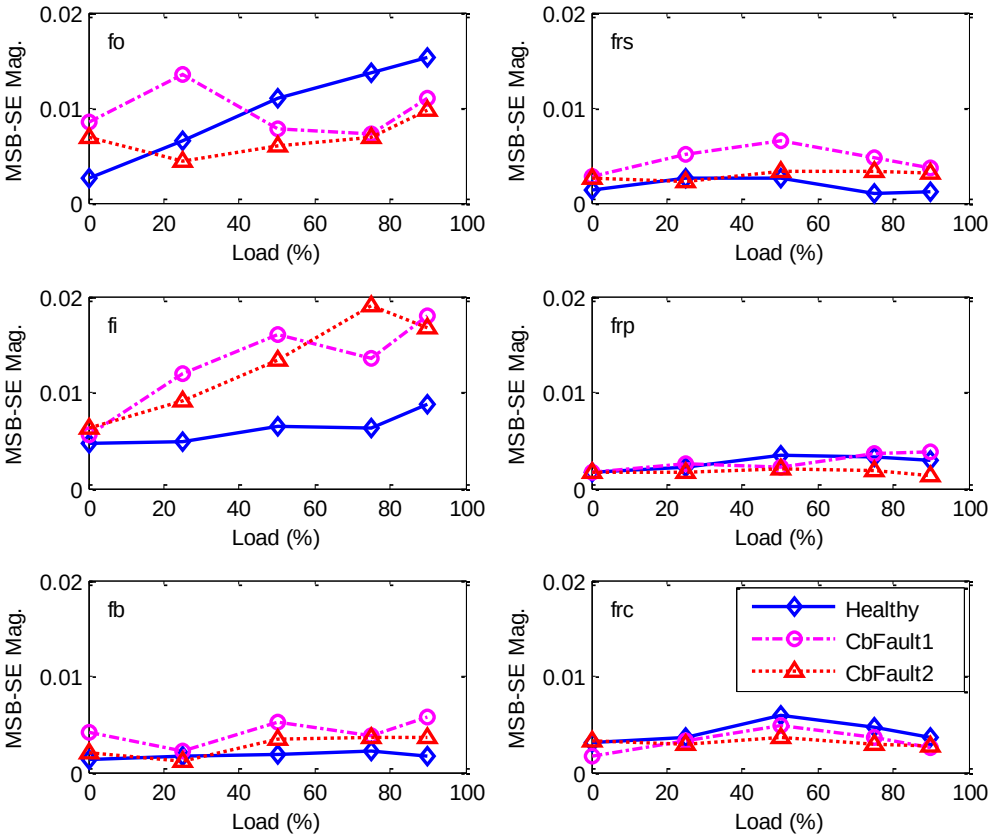

Figure 9.10 MSB-SE diagnosis results of the bearing faults using the slice at $f_c = 4 f_m$

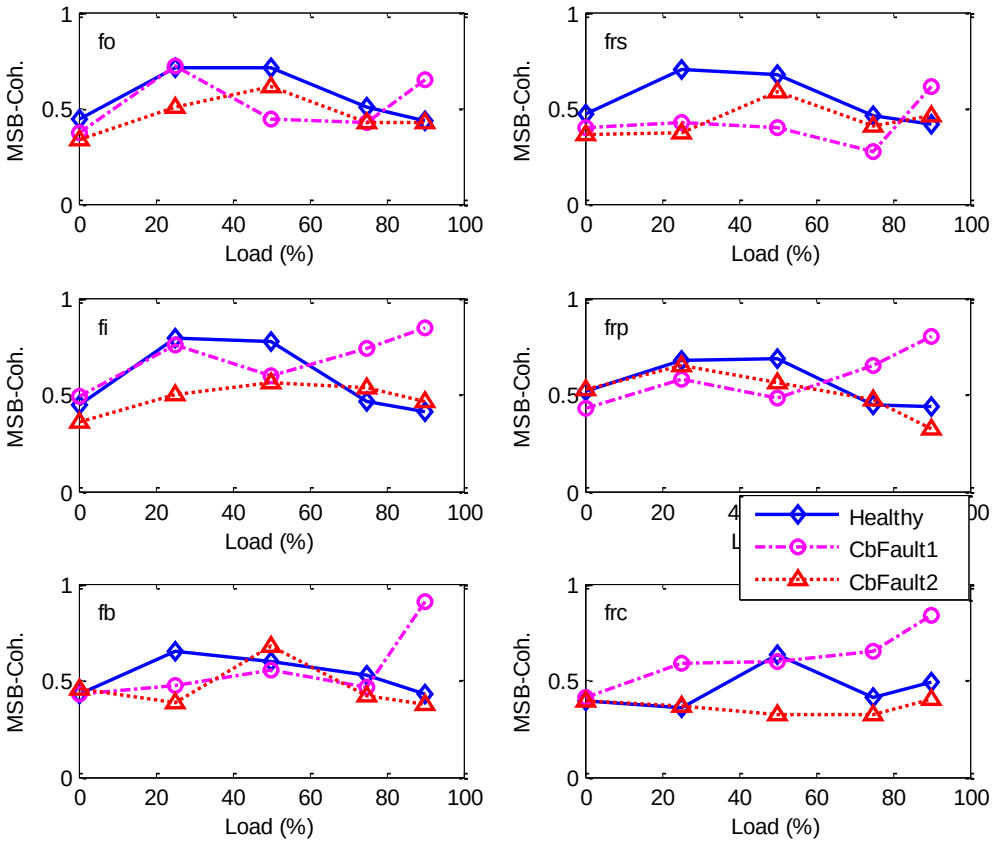

Figure 9.11 MSB-Coh. results of the bearing faults using the slice at $f_c = 4 f_m$

9.4.3 Diagnosis of bearing fault based on f_i

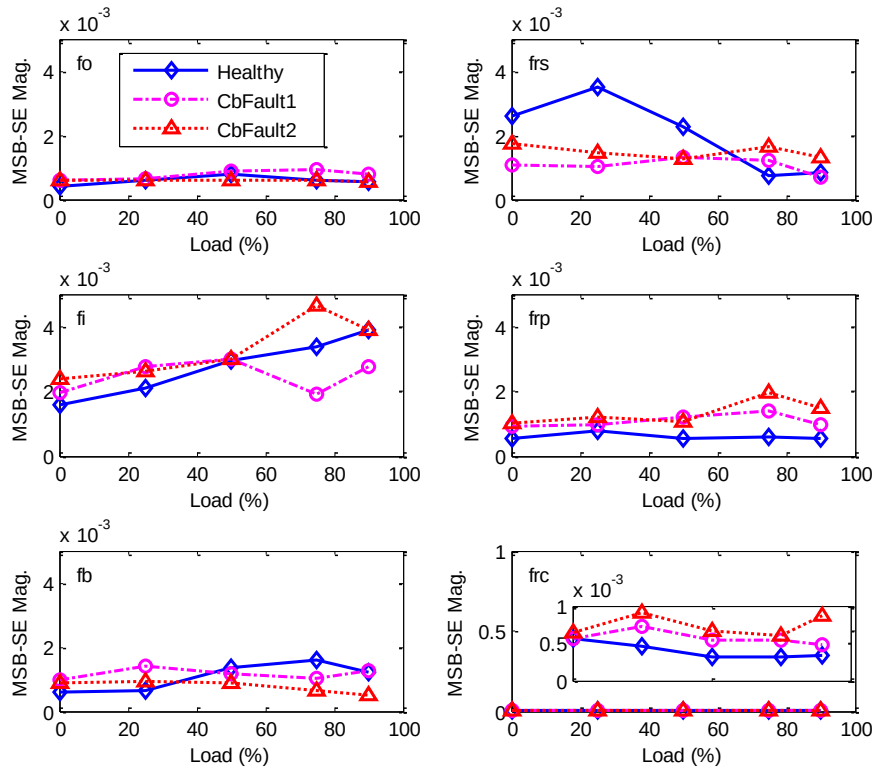

Figure 9.12 Averaged MSB-SE diagnosis results of the small bearing faults using the slices at $f_c = f_i$ and $f_c = 4f_i$

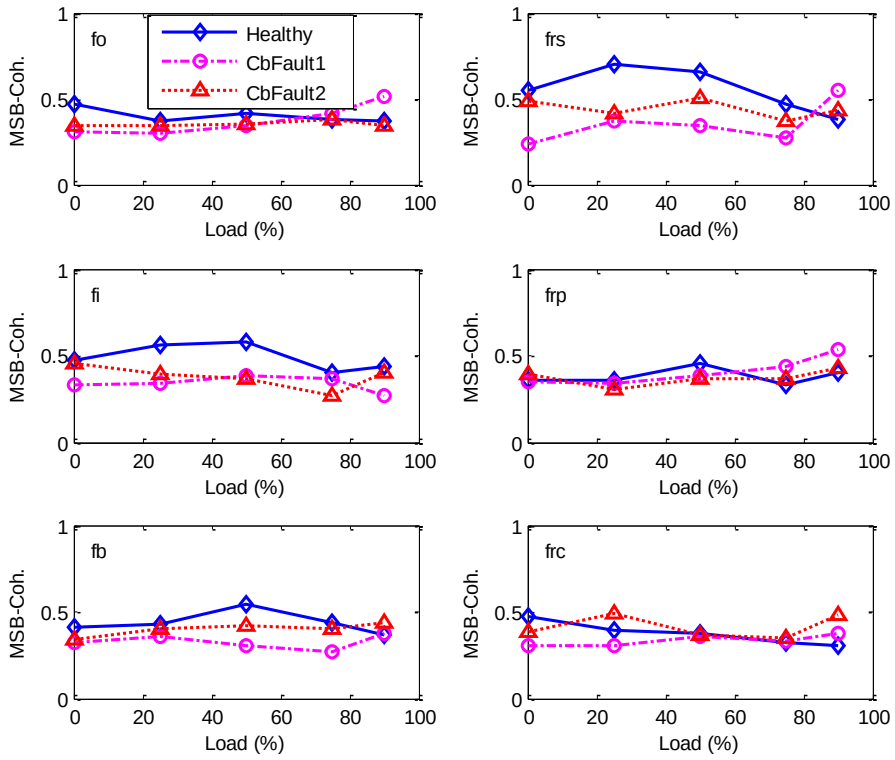

Figure 9.13 Averaged MSB-Coh. results of the small bearing faults using the slice at $f_c = f_i$ and $f_c = 4f_i$

From the spectrum in Figure 9.1 it can be seen that the characteristic frequency of bearing fault and its harmonics suffer interference from the complex gearbox frequency components. The frequency amplitudes will be greatly reduced, which is not conductive to bearing fault diagnosis. Therefore, only the harmonics with high coherence values will be selected for bearing fault diagnosis.

In this section, MSB slices at $f_c = f_i$ and $f_c = 4f_i$ are selected for fault detection because of their high coherences. The averaged MSB-SE and MSB coherence results are presented in Figure 9.12 and Figure 9.13, respectively. From these two figures, it can be seen that the sidebands at f_{rc} can separate the small inner race fault from the larger one, while the sidebands at f_{rp} only show better results under the high load conditions where the modulations are stronger.

9.5 Combined fault diagnosis at low speeds

9.5.1 Diagnosis of sun gear fault

The diagnostic results of MSB, MSB-SE and MSB coherence for the sun gear fault are presented in Figure 9.15, Figure 9.15 and Figure 9.16, respectively. From the MSB results obtained from the MSB slice at $2f_m - f_{rc}$, it can be seen that all of the fault frequencies and rotational frequencies show good increasing trends. From the MSB-SE results shown in Figure 9.15, it can be seen that only the amplitudes at sun gear fault frequency f_{sf} and shaft rotational frequency f_{rs} show good increasing trends with loads, which agrees with the load characteristics of gear transmissions. Figure 9.16 shows that the MSB coherence has high amplitudes at f_{sf} and f_{rs}, which indicates that the modulation effect is stronger at these frequencies and the results are more reliable. In the experiment, only the sun gear has a fault. The MSB-SE results give clear indication of sun gear fault while the MSB results give fault feature of all fault frequencies. Therefore, comparing the diagnostic results of the MSB and the MSB-SE, it can be concluded that the MSB-SE is more reliable because it has removed the effect of the meshing frequency which has high amplitude but contains little fault information. Moreover, these amplitudes show clear incremental differences between the three tested cases under high loads. Therefore, they can be used for obtaining fault diagnosis reliably.

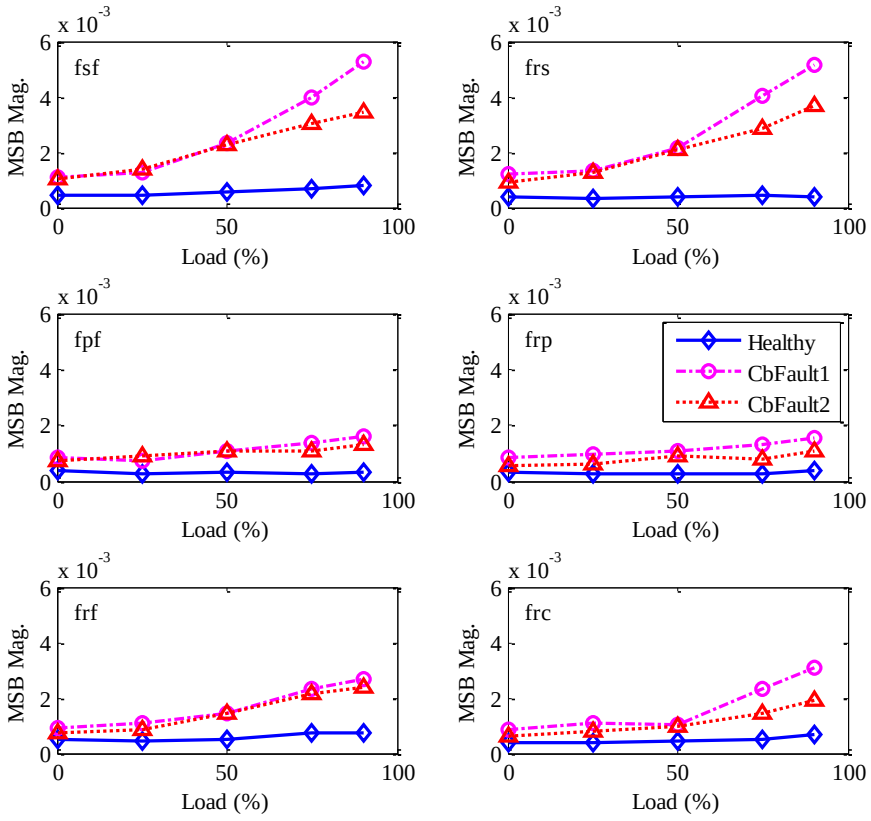

Figure 9.14 MSB diagnosis results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

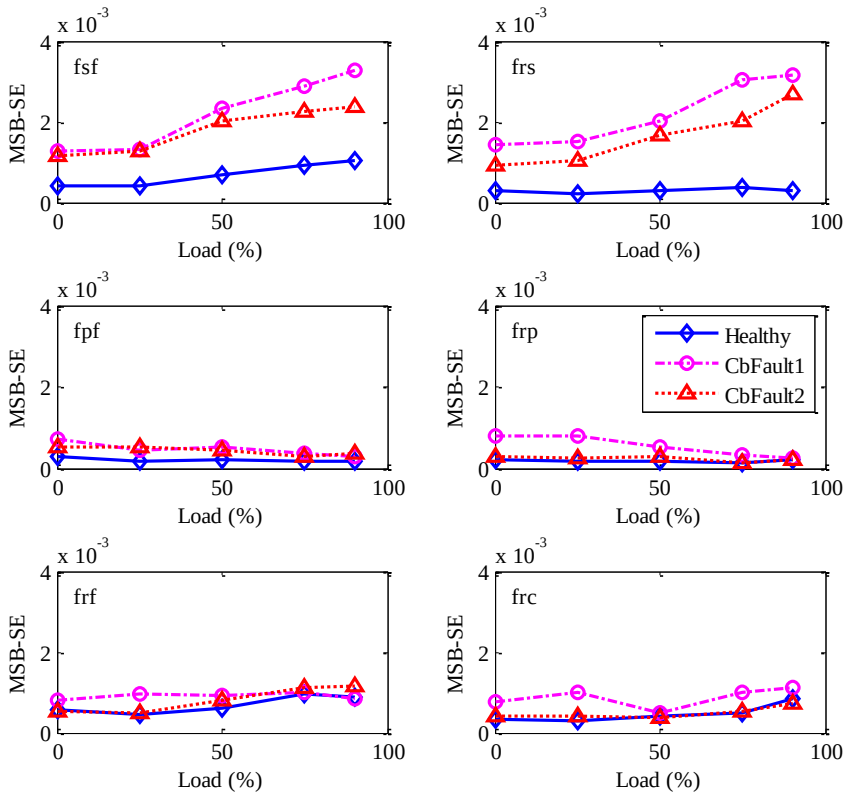

Figure 9.15 MSB-SE diagnosis results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

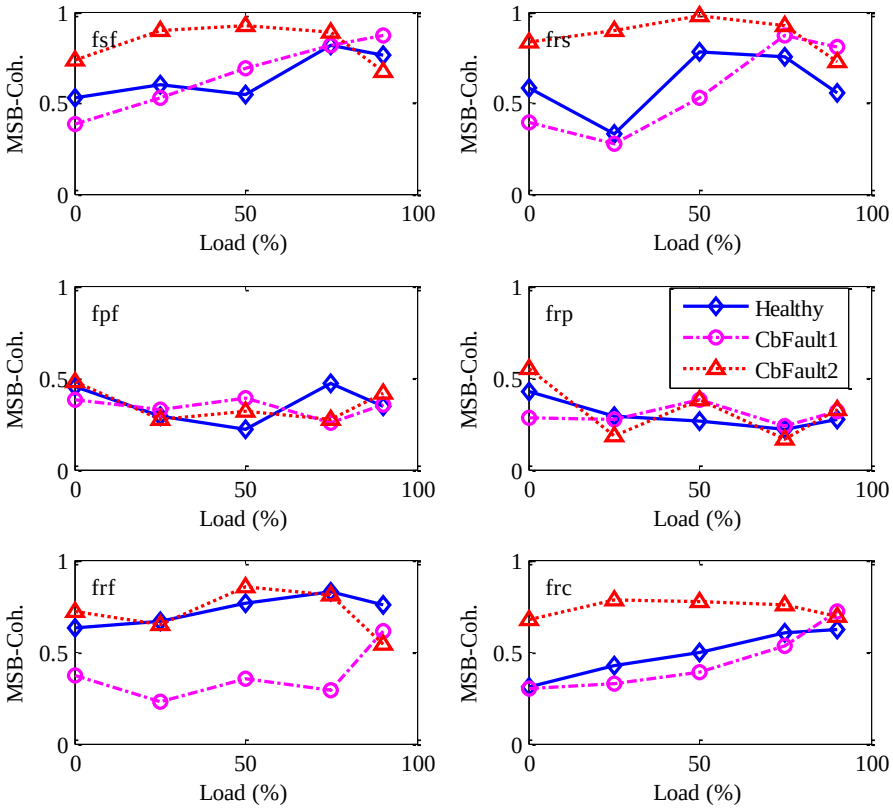

Figure 9.16 MSB-Coh. results of the sun gear faults using the slice at $f_c = 2f_m - f_{rc}$

9.5.2 Diagnosis of bearing fault based on f_m

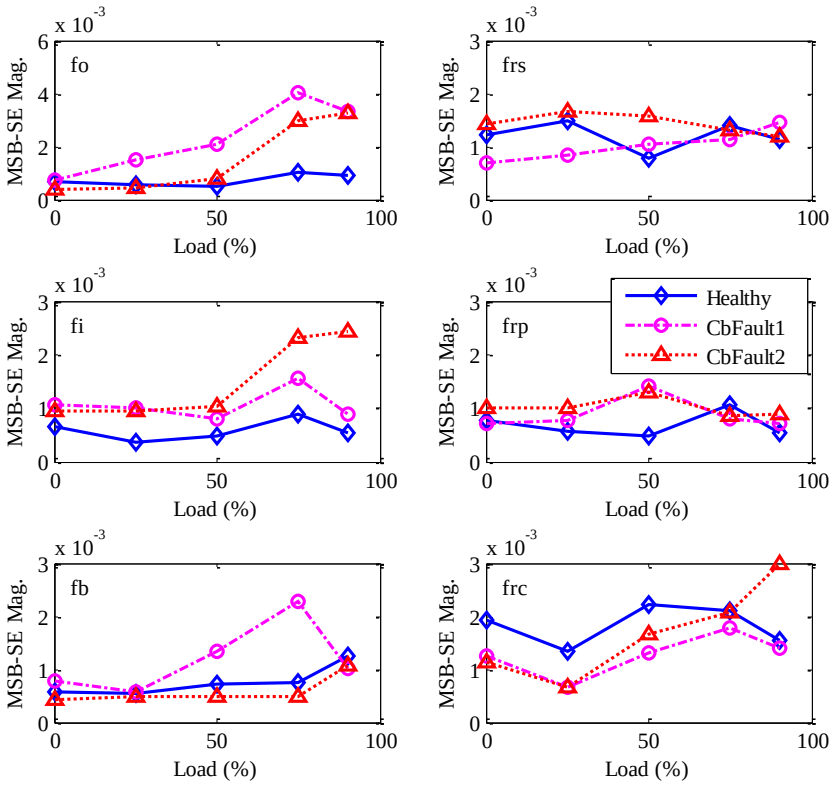

Figure 9.17 MSB-SE diagnosis results of the bearing faults using the slice at $f_c = 4f_m$

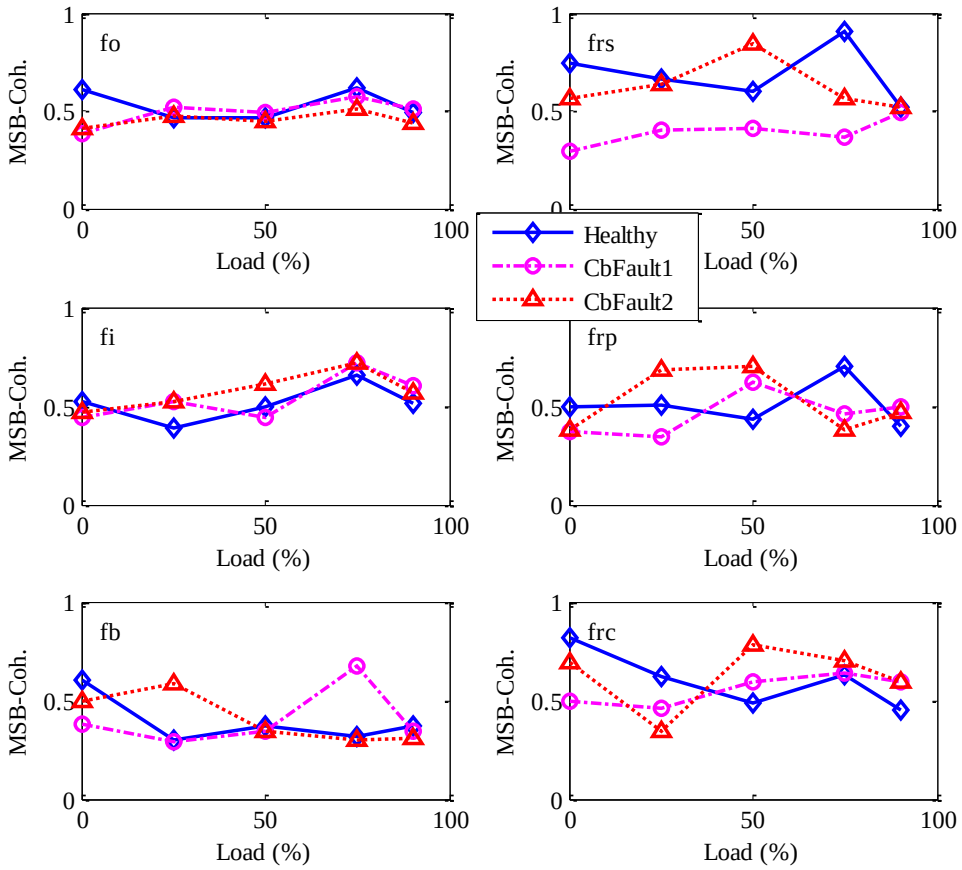

Figure 9.18 MSB-Coh. results of the bearing faults using the slice at $f_c = 4 f_m$

Figure 9.17 and Figure 9.18 show the MSB-SE and MSB coherence results obtained from the MSB slices at $4 f_m$, respectively. From Figure 9.17, it can be seen that the amplitudes of sidebands show good increasing trends with the load at the inner race bearing fault frequency. It could separate the small and large inner race faults at high load conditions. Although the coherence amplitudes are not as high as those in gear fault detection, they are acceptable because the slippage in the bearing can influence the coherence amplitudes.

9.5.3 Diagnosis of bearing fault based on f_i

In this section, the MSB slice at $f_c = 3 f_i$ is selected for fault detection because of its high coherences. The MSB-SE and MSB coherence results are presented in Figure 9.19 and Figure 9.20, respectively. From the figures, it can be seen that the sidebands at f_{rc} and f_i can separate the small inner race fault from the larger fault.

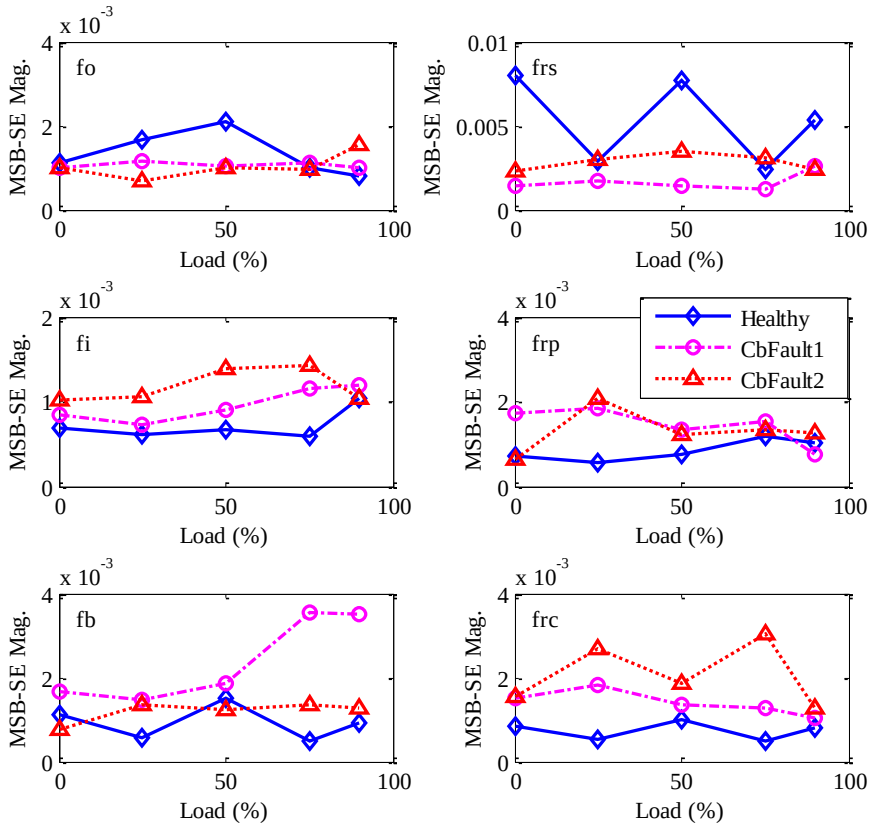

Figure 9.19 MSB-SE diagnosis results of the bearing faults using the slice at $f_c = 3f_i$

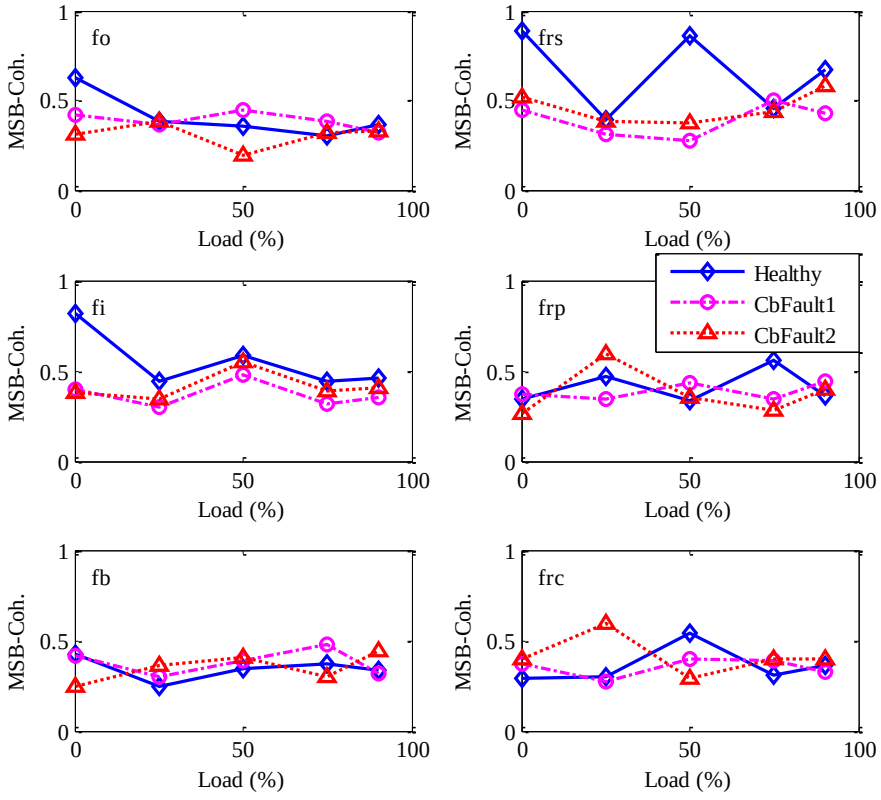

Figure 9.20 MSB-Coh. results of the bearing faults using the slice at $f_c = 3f_i$

9.6 Summary

In this chapter, a combined fault diagnosis method based on MSB-SE is proposed for detecting and diagnosing the combined faults of the bearing and gear in a planetary gearbox. The sideband extracted by MSB-SE at the MSB slices relating to characteristic frequencies contain more information on bearing and gear faults than MSB, because it has removed the influence of carrier frequency (e.g. gear mesh frequencies). Reliable results are obtained by averaging multiple slices of MSB-SE to minimise complicated interferences and noises from the concurrent sources and path distortions in a planetary gearbox.

The proposed method was verified with experimental vibration signals from a planetary gearbox with combined gear and bearing faults. The diagnostic results show that not only the types of combined faults (defects on bearing inner race and tooth breakages of sun gear) but also the severity of the two bearing faults can be separated successfully under various load and speed conditions.

Chapter 10
Conclusions and future work

This chapter summarises the achievements of the research work and explains how the objectives stated in Chapter 1 were achieved. A summary of the author's contributions to knowledge is also included. Finally, recommendations for the continuation of this vibration signal processing research are presented.

10.1 Review of objectives and achievements

The purpose of this research is to develop improved vibration signal analysis technologies for reliable condition monitoring and fault diagnostics of bearings and gearboxes. The research has focused on the refinement of a number of typical signal processing algorithms so that they can achieve simultaneous noise reduction and feature extraction in noise contaminated and complicated vibration signals.

All the proposed objectives of this research programme have been fulfilled. The work performed in obtaining these objectives is summarised according to the order in Section 1.5.

Objective 1: Review the commonly used vibration signal processing techniques for machinery condition monitoring and fault diagnosis.

Achievement 1: Representative signal processing techniques for machinery fault diagnosis are reviewed according to historic progress of the machinery condition monitoring, which has identified that ANC, ALE, TFR, Wavelet transform and kurtogram adaptive wavelet are representative techniques for achieving either noise suppression or signature enhancement, whereas TSA and bispectrum analysis allows for simultaneous noise reduction and signature enhancement which is the paradigm identified in this thesis to be focused on in improving the signal processing techniques.

Objective 2: Investigate the vibration signal models of gear defects, bearing defects and the combined faults to study the fault characteristics and provide the primary bases for signal processing.

Achievement 2: Vibration signal characteristics for bearing faults, gear faults and combined faults are studied based on fault generation mechanisms: kinematic and dynamic nonlinear couplings. It has understood that all the resulted vibration signals exhibit periodic, impulsive and modulating phenomena, which are important bases for further improving signal processing and feature extraction methods.

Objective 3: Develop optimal impulse enhancement methods based on ALE and wavelet analysis techniques to improve the SNR of vibration signals for gearbox fault diagnosis.

Achievement 3: ALE has been optimised upon the noisy vibration signals for the fault detection and diagnosis of a two-stage helical gearbox. TSA is used to suppress the

random noise firstly and then applying ALE is to reduce the continuous mesh components by using minimising wavelet entropy difference for optimal filter selection, which allows the impulsive features to be enhanced significantly. In this way, the incipient tooth damage fault can be detected reliably.

Objective 4: Investigate the kurtogram analysis and develop a filter for obtaining an optimal envelope that enhances the impulsive components due to bearing defects.

Achievement 4: An effective adaptive filter technique has been developed by combining SK with envelope analysis for rolling bearing fault detection and diagnosis. The adaptive filter is applied to improve the SNR. The filter parameters including bandwidth and central frequency are optimised by a maximal SK criterion. The effectiveness of the proposed method has been evaluated based on comprehensive experimental datasets from three types of faults and three levels of damage severities. The diagnostic results show that not only the types of fault can be identified correctly but also the severity is estimated with a good degree of accuracy.

Objective 5: Develop effective approaches and algorithms for planetary gear fault detection and diagnosis based on MSB analysis to utilise the modulating characteristics of signals.

Achievement 5: Based on vibration generation and propagation mechanisms in a planetary gearbox, an MSB based residual sideband analysis approach is developed. Both simulative and experimental results have shown that this new approach can provide correct and hence reliable diagnostic results for different sizes of faults on either the sun gear or the planet gear under different loads.

Objective 6: Propose a signal processing method for diagnosing the combined fault on both the gear and bearing through MSB analysis based on the multiple and asymmetric modulating characteristics.

Achievement 6: A combined fault diagnosis method based on MSB-SE is developed for monitoring faults on both the bearings and gears in the planetary gearbox. It matches and averages multiple MSB slices that correlate to fault characteristic frequencies for obtaining more reliable diagnostic features. The verification with experimental data from the planetary gearbox shows that not only the types of the combined faults: defects in bearing inner race and tooth breakages of sun gear can be

separated but also the severity of the two faults can be estimated successfully under different load conditions.

10.2 Conclusions

Based on the investigations described in the former chapters, the key findings covered in this thesis can be concluded as following:

Conclusion 1: There are many techniques in development for gear and bearing diagnostic. Amongst them, ALE, TSA and bispectrum analysis are potential because they have the inherent capability for noise reduction and signal enhancement. However, bispectrum is more promising because it is easy and reliable for implementation in practice.

Conclusion 2: The vibration signals from bearings and gears exhibit periodic, impulsive and modulating characteristics that correspond to different degrees of local defects. These characteristics are important for signal analysis techniques development. Moreover, the common characteristics of bearings and gears make it feasible to implement fault detection and diagnosis using the same types of techniques.

Conclusion 3: Optimised wavelet analysis allows the impulsive components of fault gears to be extracted at selected scale for fault detection. As wavelet has not had the true noise reduction capability, TSA has to be applied for pre-noise reduction. In addition, it also needs a reference signal measured from a tachometer, which increases complexity and cost due to additional hardware to collect the reference signal.

Conclusion 4: Kurtogram is particularly effective for determining a band-pass filter that allows optimal envelope analysis and results in excellent performance of bearing fault detection and fault severity diagnosis. However, because Kurtosis values are very sensitive to aperiodic impulses and without any noise reduction mechanisms, kurtogram analysis is not robust to the interferences of such impulses including strong random noise, which can lead to inadequate detection and diagnosis.

Conclusion 5: The proposed MSB-SE based robust detector can achieve both optimal band selection and envelope spectrum simultaneously and shows more reliable results for bearing fault detection and diagnosis because it achieves noise reduction and signal enhancement jointly, which is realised by accounting the full characteristics of

vibration signals, rather than just the impulsive character accounted in wavelet and kurtogram.

Conclusion 6: The proposed residual sideband based approach is based on new findings that small sidebands around gear mesh components are due to an imperfect superposition of concurrent gear mesh vibration sources in an operational planetary gearbox with inherent errors, and that the MSB is particularly effective in extracting such small sidebands. As a result, it yields more accurate and consistent diagnostic results in a wide range of operating conditions, compared with that of the conventional spectrum analysis that is often based on the in-phase sidebands which are more affected by manufacturing errors.

Conclusion 7: Combined faults in bearings and gears can be separated by MSB analysis. To make the results more reliable, multiple slices of MSB-SE should be averaged to minimise complicated interferences and noises from the concurrent sources and path distortions in a planetary gearbox.

10.3 Research contributions to knowledge

The researches carried out have brought a number of new understandings on the subject of developing effective signal analysis of vibration signals for accurate detection and diagnosis of faults on bearings and gears. For convenience, key contributions to knowledge in this thesis are outlined below:

Contribution 2: The ALE algorithm based on an adaptive LMS filter allows the stationary periodic gear mesh components to be suppressed significantly in the time domain, achieving the enhancement of impulsive components for effective fault detection and diagnosis in a multistage gearbox. This has extended the knowledge that ALE can be applied to the vibration signals from a multi-stage helical gearbox for condition monitoring.

Contribution 3: Like ALE, the optimal wavelet analysis is effective for extracting impulsive fault signals. However, because it does not have any mechanisms for noise suppression, TSA has to be used to suppress the meshing components in order to achieve an adequate detection and diagnosis performance for the faults on the multistage gearbox. These understandings have not been found in previous studies.

Contribution 4: The TFR based kurtogram shows high performance in highlighting impulsive components and achieving outstanding diagnostic results for bearing faults. However, it is deficient in dealing with noise interferences and hence not robust, and insensitive to incipient bearing defects. This deficiency has been realised for the first time in this study.

Contribution 5: A novel MSB based detector has been suggested for the first time in detecting the incipient bearing faults. MSB takes into account more comprehensively the fault signal characteristics so that it can suppress both the stationary and aperiodic noises to enhance the periodic impulses due to bearing faults.

Contribution 6: A novel approach for the fault diagnosis of planetary gearboxes has been proposed based on an accurate estimation of residual sidebands using MSB analysis. Especially, the utilisation of asymmetric sidebands around the mesh frequency has shown outstanding performances as the sidebands are less influenced by gear errors, compared with that of symmetric sidebands that have used traditionally for diagnosing planetary gearboxes.

Contribution 7: As MSB-SE based analysis allows not only noise reduction but also sparse representation for the complicated modulations in the signals from the combined fault of the planetary gearbox, it then optimises the fault related content and allows more efficient fault feature exactions. Especially, the average of multiple MSB slices is used to obtain more robust results, which is realised the first time in this thesis.

10.4 Suggestions for further research

A great deal of preliminary research work has been undertaken as part of this study to investigate signal processing techniques for the analysis of non-stationary signals in machinery condition monitoring. If a further study will be undertaken, some key recommendations could be considered in this research area.

Recommendation 1: Improve the kurtogram based bearing fault diagnosis by utilising multiple bands associated different resonances. Resonance regions in the kurtogram often exhibit certain patterns, so it is possible to recognise the number of resonance regions and extract the feature separately for a final average to improve the robustness of such method.

Recommendation 2: Investigate the fault mechanisms for developing more effective signal processing methods to detect and diagnose more common combined fault scenarios such as multiple defect locations, different combinations of fault severity, and multiple fault types, which are often cases in practical applications.

Recommendation 3: Optimise the MSB based feature extraction using clustering algorithm for automated feature extraction of combined fault diagnosis. The MSB slices carry fault information for difference components and severity of a machine, so the classification of the slices is useful for reliable and accurate feature extraction.

Recommendation 4: Carry out more studies of diagnosing planetary bearing faults because this kind of faults occurs commonly due to the unbalance forces and its associated vibration signal is more difficult to be resolved due to small in amplitude and non-stationary in the wave shapes.

REFERENCES

[1] S. Vidyasankar, Importance of Condition Monitoring Unveiled, July 2005

[2] M. Neal, and Associates, *Guide to the condition monitoring of machinery*, London: H.M. Stationery Off., 1979.

[3] C. Mechefske, "Machine Condition Monitoring and Fault Diagnostics," *Vibration and Shock Handbook*, W. De Silva (ed.), Chapter 25, pp. 25/1-25/35, CRC Press, 2005.

[4] A. Jardine, D. Lin and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510, 2006.

[5] R. Randall, *Vibration-based condition monitoring*. Chichester, West Sussex, U.K.: Wiley, 2011.

[6] E. P. Carden and P. Fanning, "Vibration Based Condition Monitoring: A Review", *Structural Health Monitoring*, vol. 3, no. 4, pp. 355-377, 2004.

[7] A. Al-Ghamd and D. Mba, "A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size", *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1537-1571, 2006.

[8] N. Tandon and A. Choudhury, "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings", *Tribology International*, vol. 32, no. 8, pp. 469-480, 1999.

[9] G. Ramroop, K. Liu, F. Gu, B. S. Payne, and A. D. Ball, "Airborne Acoustic Condition Monitoring of a Gearbox System." In : *2001 5th Annual Maintenance and Reliability Conference (MARCON 2001)*. May, 2001.

[10] J. Zhu, J. Yoon, D. He and E. Bechhoefer, "Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines", *Wind Energ.*, vol. 18, no. 6, pp. 1131-1149, 2014.

[11] W. Caesarendra, B. Kosasih, A. Tieu, H. Zhu, C. Moodie and Q. Zhu, "Acoustic emission-based condition monitoring methods: Review and application for low speed slew bearing", *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 134-159, 2016.

[12] D. Mba, "Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines: Bearings, Pumps, Gearboxes, Engines, and Rotating Structures", *The Shock and Vibration Digest*, vol. 38, no. 1, pp. 3-16, 2006.

[13] H. de Azevedo, A. Araújo and N. Bouchonneau, "A review of wind turbine bearing condition monitoring: State of the art and challenges", *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 368-379, 2016.

[14] Z. Amira, B. Mohamed and E. Tahar, "Monitoring of temperature in distributed optical sensor: Raman and Brillouin spectrum", *Optik - International Journal for Light and Electron Optics*, vol. 127, no. 8, pp. 4162-4166, 2016.

[15] S. Nandi, H. Toliyat and X. Li, "Condition Monitoring and Fault Diagnosis of Electrical Motors—A Review", *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719-729, 2005.

[16] P. Tavner, L. Ran, J. Penman, and H. Sedding, *Condition Monitoring of Rotating Electrical Machines*. The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK: IET, 2008.

[17] S. Goldman, Vibration spectrum analysis: a practical approach, New York: Industrial Press, 1999.

[18] R. Willetts, *Holistic condition monitoring*, EngD thesis, University of Manchester, 2002.

[19] Michael, R. J., Practical applications of On-line Vibration Monitoring. Bretech Engineering Ltd, Saint John, NB, Canada, E2L 3V6

[20] J. Trampe Broch, *Mechanical vibration and shock measurements*. Denmark: Bruel and Kjaer, 1984.

[21] P. McFadden and J. Smith, "Vibration monitoring of rolling element bearings by the high-frequency resonance technique-a review", *Tribology International*, vol. 17, no. 1, pp. 3-10, 1984.

[22] A. Rai and S. H. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings", *Tribology International*, vol. 96, pp. 289–306, Apr. 2016.

[23] G. Luo, D. Osypiw and M. Irle, "Real-time condition monitoring by significant and natural frequencies analysis of vibration signal with wavelet filter and autocorrelation enhancement ", *Journal of Sound and Vibration*, vol. 236, no. 3, pp. 413-430, 2000.

[24] B. Kilundu, X. Chiementin, J. Duez and D. Mba, "Cyclostationarity of Acoustic Emissions (AE) for monitoring bearing defects", *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2061-2072, 2011.

[25] G. Chaturvedi and D. Thomas, "Bearing Fault Detection Using Adaptive Noise Cancelling", *Journal of Mechanical Design*, vol. 104, no. 2, pp 280-289, 1982.

[26] R. H. Bannister, "A review of rolling element bearing monitoring techniques", in *I.MECH.E Conference on Condition Monitoring of Machinery and Plant*, pp 11-24, 1985.

[27]]C. Tan, "Adaptive noise cancellation of acoustic emission noise in ball bearings", in *Proceedings of the Asia-Pacific Vibration Conference*, Melbourne, Australia, pp. 5.28-5.35, 1991.

[28] S. Elliott, *Signal processing for active control*. San Diego, Calif.: Academic, pp 77, 2001.

[29] D. Zhen, Y. Fan, F. Gu, and A. Ball, "A Study of Sound Source Characteristics in Active Noise Control and Fault Diagnosis", in *Inter noise 2010-39th International congress on Noise Control Engineering*, Lisbon, Portugal, 2010.

[30] R.B. Randall, "Cepstrum Analysis and Gearbox Fault Diagnosis", application notes, B&K Instruments, Inc.

[31] N. Tandon, "A comparison of some vibration parameters for the condition monitoring of rolling element bearings", *Measurement*, vol. 12, no. 3, pp. 285-289, 1994.

[32] E. Bechhoefer, P. Menon, and M. Kingsley, "Bearing envelope analysis window selection Using spectral kurtosis techniques", *in 2011 IEEE Conference on Prognostics and Health Management (PHM)*, 2011, pp. 1–6.

[33] P. Flandrin, *Time-frequency/time scale analysis*. San Diego: Academic Press, 1999.

[34] L. Cohen, *Time frequency analysis*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[35] Y. Zhou, J. Chen, G. Dong, W. Xiao and Z. Wang, "Wigner-Ville distribution based on cyclic spectral density and the application in rolling element bearings diagnosis", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 225, no. 12, pp. 2831-2847, 2011.

[36] I. Mayes, A. Steer, and G. Thomas. "The application of vibration monitoring for fault diagnosis in large turbo-generators", in *Proceedings of the Sixth Thermal Generation Specialists Meeting*, Madrid, 1981.

[37] R. Kronland-martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 01, no. 02, pp. 273-302, 1987.

[38] T. Ning, Y. Kung, F. Wei, "Detection of distributed gear faults with a new bispectral analysis", in *Proceedings of the 1996 IEEE IECON 22nd International Conference on Industrial Electronics, Control, and Instrumentation*, vol.3, pp. 1960–1965, 1996.

[39] C. Nikias and M. Raghuveer, "Bispectrum estimation: A digital signal processing framework", *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869-891, 1987.

[40] J. Stack, R. Harley and T. Habetler, "An amplitude modulation detector for fault diagnosis in rolling element bearings", *IEEE Transactions on Industrial Electronics*, vol. 51, no. 5, pp. 1097-1102, 2004.

[41] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung and H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis", *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903-995, 1998.

[42] J. Smith, "The local mean decomposition and its application to EEG perception data", *Journal of The Royal Society Interface*, vol. 2, no. 5, pp. 443-454, 2005.

[43] Z. Meng, and Y. Wang, "Rotating machinery fault diagnosis method based on the differential local mean decomposition", *Journal of Mechanical Engineering*, vol. 50, no. 11, pp. 101-107, 2014.

[44] J. Cheng, Y. Yang, and Y. Yang, "Local characteristic-scale decomposition method and its application to gear fault diagnosis", *Journal of Mechanical Engineering*, vol. 48, no. 09, pp. 64-71, 2012.

[45] Q. Miao and Q. Zhou, "Planetary Gearbox Vibration Signal Characteristics Analysis and Fault Diagnosis", *Shock and Vibration*, vol. 2015, pp. 1-8, 2015.

[46] S. Jia, I. Howard and J. Wang, "The dynamic modeling of multiple pairs of spur gears in mesh, including friction and geometrical errors", *International Journal of Rotating Machinery*, vol. 9, no. 6, pp. 437-442, 2003.

[47] R. Randall, "A New Method of Modeling Gear Faults", *Journal of Mechanical Design*, vol. 104, no. 2, pp. 259-267, 1982.

[48] M. Hoseini, M. Mandal, M. Zuo, and G. Mani, "Gearbox fault diagnosis using Hilbert transform and segmented regression", in *Proceedings of the Fifth International Conference on Condition Monitoring & Machinery Failure Prevention Technologies*, pp. 650-656, Edinburgh, Scotland, UK, 2008.

[49] X. Fan and M. Zuo, "Gearbox fault detection using Hilbert and wavelet packet transform", *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 966-982, 2006.

[50] E. Schukin, R. Zamaraev and L. Schukin, "The optimisation of wavelet transform for the impulse analysis in vibration signals", *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1315-1333, 2004.

[51] W. Bartelmus, "Mathematical modelling and computer simulations as an aid to gearbox diagnostics," *Mechanical systems and signal processing*, vol. 15, pp. 855-871, 2001.

[52] Ian Howard, "A Review of Rolling Element Bearing Vibration “Detection, Diagnosis and Prognosis”", *Defence science and technology organization (DSTO-RR-0013), Airframes and Engines Division*, 1994.

[53] Z. Kıral and H. Karagülle, "Vibration analysis of rolling element bearings with various defects under the action of an unbalanced force", *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 1967-1991, 2006.

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

[55] Z. Feng and M. Zuo, "Vibration signal models for fault diagnosis of planetary gearboxes", *Journal of Sound and Vibration*, vol. 331, no. 22, pp. 4919-4939, 2012.

[56] F. Gu, G. Abdalla, R. Zhang, H. Xu and A. D. Ball, "A Novel Method for the Fault Diagnosis of a Planetary Gearbox based on Residual Sidebands from Modulation Signal Bispectrum Analysis", in *Comadem 2014*, Brisbane, Australia, 2014.

[57] Q. Miao and Q. Zhou, "Planetary Gearbox Vibration Signal Characteristics Analysis and Fault Diagnosis", Shock and Vibration, vol. 2015, pp. 1-8, 2015.

[58] Y. Lei, J. Lin, M. Zuo and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review", *Measurement*, vol. 48, pp. 292-305, 2014.

[59] L. Hong, J. Dhupia and S. Sheng, "An explanation of frequency features enabling detection of faults in equally spaced planetary gearbox", *Mechanism and Machine Theory*, vol. 73, pp. 169-183, 2014.

[60] M. Inalpolat and A. Kahraman, "A theoretical and experimental investigation of modulation sidebands of planetary gear sets", *Journal of Sound and Vibration*, vol. 323, no. 3-5, pp. 677-696, 2009.

[61] R. Ramli, A. Abid Noor, and S. Samad, "A Review of Adaptive Line Enhancers for Noise Cancellation", *Australian Journal of Basic & Applied Sciences*, vol. 6, no. 6, pp. 337-352, 2012.

[62] B. Widrow and S. Stearns, *Adaptive signal processing*. Englewood Cliffs, N.J.: Prentice-Hall, 1985.

[63] R. Campbell, N. Younan and J. Gu, "Performance analysis of the adaptive line enhancer with multiple sinusoids in noisy environment", *Signal Processing*, vol. 82, no. 1, pp. 93-101, 2002.

[64] J. Treichler, "Transient and convergent behavior of the adaptive line enhancer", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, no. 1, pp. 53–62, 1979.

[65] N. Sasaoka, K. Sumi, Y. Itoh and K. Fujii, "A new noise reduction system based on ALE and noise reconstruction filter", in *proceeding of: International Symposium on Circuits and Systems*, Kobe, 2005.

[66] Lee, S. K.; White, P. R., "Fault diagnosis of rotating machinery using a two-stage adaptive line enhancer", *IEE Colloquium on Modeling and Signal Processing for Fault Diagnosis (Digest No.: 1996/260)*, pp. 6/1-6/6, Leicester, 1996.

[67] S. Lee and P. White, "The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery", *Journal of Sound and Vibration*, vol. 217, no. 3, pp. 485-505, 1998.

[68] S. Lee and J. Lee, "Design of Adaptive Filter for Health Monitoring on a Gearbox", *Key Engineering Materials*, vol. 321-323, pp. 1237-1240, 2006.

[69] J. Mohammed and M. Shafi, "An Efficient Adaptive Noise Cancellation Scheme Using ALE and NLMS Filters", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 2, no. 3, pp. 325–332, 2012.

[70] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, J. Eugene Dong and R. Goodlin, "Adaptive noise cancelling: Principles and applications", *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692-1716, 1975.

[71] R. Ramli, A. Abid Noor, and S. Samad, "A Review of Adaptive Line Enhancers for Noise Cancellation", *Australian Journal of Basic & Applied Sciences*, vol. 6, no. 6, pp. 337-352, 2012.

[72] H. Simon, *Adaptive filter theory*, Prentice Hall, 2, pp.478-481, 2002.

[73] P. Diniz, Adaptive filtering: algorithms and practical implementation, Springer, 2013

[74] Z. Peng and F. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography", *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199-221, 2004.

[75] V. Chancey and G. Flowers, "Identification of transient vibration characteristics using absolute harmonic wavelet coefficients", *Journal of Vibration and Control*, vol. 7, no. 8, pp. 1175-1193, 2001.

[76] J. Lin and L. Qu, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnostics", *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135-148, 2000.

[77] M. Xu, J. Zhang, G. Zhang, and W. Huang, "Method of data compressing for rotating machinery vibration signal based on wavelet transform", *Journal of Vibration Engineering*, vol. no. 4, pp. 531-536, 2000.

[78] J. Ma, Y. Liu, and Z. Zhu, "Study on identification of dynamic characteristics of oil-bearings based on wavelet analysis", Chinese Journal of Mechanical Engineering, vol. 36, no. 05, pp. 81-83, 2000.

[79] J. Lin and M.J. Zuo, "Gearbox fault diagnosis using adaptive wavelet filter", *Mechanical Systems and Signal Processing*, vol. 17, no. 6, pp. 1259-1269, 2003.

[80] W. Su, F. Wang, H. Zhu, Z. Zhang and Z. Guo, "Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement", *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1458-1472, 2010.

[81] W. He, Z. Jiang and Q. Qin, "A joint adaptive wavelet filter and morphological signal processing method for weak mechanical impulse extraction", *Journal of Mechanical Science and Technology*, vol. 24, no. 8, pp. 1709-1716, 2010.

[82] W. Liu and B. Tang, "A hybrid time-frequency method based on improved Morlet wavelet and auto terms window", *Expert Systems with Applications*, vol. 38, no. 6, pp. 7575-7581, 2011.

[83] Y. Jiang, B. Tang, Y. Qin and W. Liu, "Feature extraction method of wind turbine based on adaptive Morlet wavelet and SVD", *Renewable Energy*, vol. 36, no. 8, pp. 2146-2153, 2011.

[84] J. Lin and M. Zuo, "Extraction of Periodic Components for Gearbox Diagnosis Combining Wavelet Filtering and Cyclostationary Analysis", *Journal of Vibration and Acoustics*, vol. 126, no. 3, p. 449, 2004.

[85] C. Yu, J. Li and S. Huang, "Vibration Fault Diagnosis of Rotating Machine Based on the Principle of Entropy Increase", *Advanced Materials Research*, vol. 530, pp. 109-114, 2012.

[86] C. Zhang, L. Yang, J. Chen, Fault diagnosis of gear cases based on entropy of singular values and support vector machine, in *Electric Information and Control Engineering (ICEICE), 2011 International Conference on.*, pp. 4207-4211, 2011.

[87] Y. Chen, C. Wan, and L. Lee, "Entropy-Based Feature Parameter Weighting for Robust Speech Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006*, 2006.

[88] A. Asl, and H. Zadeh, "Constrained optimization of nonparametric entropy-based segmentation of brain structures", in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 41-44, 2008.

[89] F. Combet and L. Gelman, "An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor", Mechanical Systems and Signal Processing, vol. 21, no. 6, pp. 2590-2606, 2007.

[90] W. Wu, J. Lin, S. Han and X. Ding, "Time domain averaging based on fractional delay filter", *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1447-1457, 2009.

[91] R. Dwyer, "Detection of non-Gaussian signals by frequency domain Kurtosis estimation", in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP 83*, vol. 8, pp. 607–610, 1983.

[92] J. Antoni and R. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines", *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308-331, 2006.

[93] J. Antoni, "Fast computation of the kurtogram for the detection of transient faults", *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 108-124, 2007.

[94] Y. Lei, J. Lin, Z. He and Y. Zi, "Application of an improved kurtogram method for fault diagnosis of rolling element bearings", *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1738-1749, 2011.

[95] T. Barszcz and A. JabŁoński, "A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram", *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 431-451, 2011.

[96] D. Wang, P. Tse and K. Tsui, "An enhanced Kurtogram method for fault diagnosis of rolling element bearings", *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 176-199, 2013.

[97] L. Saidi, F. Fnaiech, G.A. Capolino, H. Henao, "Stator current bi-spectrum patterns for induction machines multiple-faults detection", *in: IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, 2012: pp. 5132–5137.

[98] S. Guoji, S. McLaughlin, X. Yongcheng and P. White, "Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears", *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 76-89, 2014.

[99] H. Zhao, H. Shen, Y. Fu, and G. Wang, "Using singular value decomposition and high order spectrum for bearings fault diagnosis", *in Transportation Electrification Asia-Pacific (ITEC Asia-Pacific), 2014 IEEE Conference and Expo*, pp. 1-4, 2014.

[100] J. Huang, S. Bi, H. Pan, and X. Yang, "The research of higher-order cumulant spectrum for vibration signals of gearbox", *in 2006 IEEE International Conference on Information Acquisition*, pp. 1395-1399, 2006.

[101] Y. Liu, K. Yan, Y. Zhou, and H. Xu, "Bispectral analysis for diagnosis of gear pitting fault", *in 7th World Congress on Intelligent Control and Automation, 2008. WCICA 2008*, pp. 6847–6851, 2008.

[102] Y. Xiao, J. Mei, R. Zeng, H. Zhao, L. Tang, and H. Huang, "Mechanical fault diagnose of diesel engine based on bispectrum and Support Vector Machines", *in 2nd IEEE International Conference on Computer Science and Information Technology, 2009. ICCSIT 2009*, pp. 425–429, 2009.

[103] Z. Zheng, H. Chen, and X. Li, "Bispectrum Based Gear Fault Feature Extraction and Diagnosis", *Journal of Vibration Engineering*, vol. 15, no. 3, pp. 354-358, 2002.

[104] I. Rehab, X. Tian, F. Gu, and A. 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.

[105] F. Gu, G. Abdalla, R. Zhang, H. Xu and A. Ball, "A Novel Method for the Fault Diagnosis of a Planetary Gearbox based on Residual Sidebands from Modulation Signal Bispectrum Analysis", in *Comadem 2014*, Brisbane, Australia, 2014.

[106] F. Feng, A. Si, and P. Jiang, "Fault Diagnosis for Crankshaft Bearing in Diesel Engine Based on Bispectrum Analysis," *Technical Bulletin of Nippon Kaiji Kyokai*, vol. 7, pp. 1033-1038, 2013.

[107] A. Naid, F. Gu, Y. Shao, S. Al-Arbi and A. Ball, "Bispectrum analysis of motor current signals for fault diagnosis of reciprocating compressors ", Key Engineering Materials, vol. 413-414, pp. 505-511, 2009.

[108] F. Gu, Y. Shao, N. Hu, A. Naid and A. Ball, "Electrical motor current signal analysis using a modified bispectrum for fault diagnosis of downstream mechanical equipment", *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 360-372, 2011.

[109] J. Fackrell, "Bispectral analysis of speech signals", PhD thesis, University of Edinburgh, UK, 1996.

[110] A. Alwodai, T. Wang, Z. Chen, F. Gu, R. Cattley and A. Ball, "A Study of Motor Bearing Fault Diagnosis using Modulation Signal Bispectrum Analysis of Motor Current Signals", *Journal of Signal and Information Processing*, vol. 04, no. 03, pp. 72-79, 2013.

[111] S. K. Lee and P. R. White, "The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery", *Journal of Sound and Vibration*, vol. 217, no. 3, pp. 485-505, 1998.

[112] J. Treichler, "Transient and convergent behavior of the adaptive line enhancer", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 1, pp. 53-62, 1979.

[113] N. Sasaoka, K. Sumi, Y. Itoh and K. Fujii, "A new noise reduction system based on ALE and noise reconstruction filter", in *proceeding of: IEEE International Symposium on Circuits and Systems ISCAS 2005*, pp. 272-275, Kobe, 2005.

[114] J. Mohammed and M. Shafi, "An Efficient Adaptive Noise Cancellation Scheme Using ALE and NLMS Filters", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 2, no. 3, pp. 325-332, 2012.

[115] B. Widrow and S. D. Stearns, Adaptive Signal Processing, Englewood Cliffs, NJ: Prentice Hall, 1985.

[116] J. Piñeyro, A. Klempnow and V. Lescano, "Effectiveness of new spectral tools in the anomaly detection of rolling element bearings", *Journal of Alloys and Compounds*, vol. 310, no. 1-2, pp. 276-279, 2000.

[117] Z. He, X. Chen and Q. Qian, "A study of wavelet entropy measure definition and its application for fault feature pick-up and classification", *Journal of Electronics (China)*, vol. 24, no. 5, pp. 628-634, 2007.

[118] J. Lin, L. Qu, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis", *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135-148, 2000.

[119] A. Najmi, J. Sadowsky, "The continuous wavelet transform and variable resolution time-frequency anaysis", *Johns Hopkins APL Technical Digest*, vol. 18, no. 1, pp. 134-140, 1997.

[120] M. Hoseini, M. Mandal, M. Zuo, and G. Mani, "Gearbox fault diagnosis using Hilbert transform and segmented regression", *in Proceedings of the Fifth International Conference on Condition Monitoring & Machinery Failure Prevention Technologies*, pp. 650-656, Edinburgh, Scotland, UK, 2008.

[121] P. McFadden and J. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing", *Journal of Sound and Vibration*, vol. 96, no. 1, pp. 69-82, 1984.

[122] Y. Lei, Z. He and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery", *Expert Systems with Applications*, vol. 35, no. 4, pp. 1593-1600, 2008.

[123] N. Tandon and A. Choudhury, "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings", *Tribology International*, vol. 32, no. 8, pp. 469-480, 1999.

[124] Ocak, H. and K. Loparo. "A new bearing fault detection and diagnosis scheme based on hidden Markov modeling of vibration signals", in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 5, pp. 3141-3144. IEEE, 2001.

[125] H. Ocak and K. Loparo, "Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data", *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 515-533, 2004.

[126] Z. Xu, J. Xuan, T. Shi, B. Wu and Y. Hu, "Application of a modified fuzzy ARTMAP with feature-weight learning for the fault diagnosis of bearing", *Expert Systems with Applications*, vol. 36, no. 6, pp. 9961-9968, 2009.

[127] L. Zhang, G. Xiong, H. Liu, H. Zou and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference", *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077-6085, 2010.

[128] E. de Moura, C. Souto, A. Silva and M. Irmão, "Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses", *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1765-1772, 2011.

[129] B. Muruganatham, M. Sanjith, B. Krishnakumar and S. Satya Murty, "Roller element bearing fault diagnosis using singular spectrum analysis", *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 150-166, 2013.

[130] K. Zhu, X. Song, and D. Xue, "Fault Diagnosis of Rolling Bearings Based on IMF Envelope Sample Entropy and Support Vector Machine", *Journal of Information & Computational Science*, vol. 10, no. 16, pp. 5189-5198, 2013.

[131] M. Jin, R. Li, Z. Xu and X. Zhao, "Reliable fault diagnosis method using ensemble fuzzy ARTMAP based on improved Bayesian belief method", *Neurocomputing*, vol. 133, pp. 309-316, 2014.

[132] T. Igarashi and H. Hamada, "Studies on the Vibration and Sound of Defective Rolling Bearings: First Report : Vibration of Ball Bearings with One Defect", *Bulletin of the Japan Society of Mechanical Engineers*, vol. 25, no. 204, pp. 994-1001, 1982.

[133] N. Tandon and A. Choudhury, "An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect", *Journal of Sound and Vibration*, vol. 205, no. 3, pp. 275-292, 1997.

[134] R. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial", *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485-520, 2011.

[135] P. McFadden and J. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing", *Journal of Sound and Vibration*, vol. 96, no. 1, pp. 69-82, 1984.

[136] Y. Lei, Z. He and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery", *Expert Systems with Applications*, vol. 35, no. 4, pp. 1593-1600, 2008.

[137] M. Darlow, R. Badgley and G. Hogg, "Application of high-frequency resonance techniques for bearing diagnostics in helicopter gearboxes", Technical Report, *US Army Air Mobility Research and Development Laboratory*, pp.74-77, 1974.

[138] J. Antoni, "Cyclic spectral analysis in practice", *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 597-630, 2007.

[139] J. Antoni, "Cyclostationarity by examples", *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 987-1036, 2009.

[140] D. Ho, R. Randall, "Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals", *Mechanical Systems and Signal Processing*, vol. 14, no. 5, pp. 763-788, 2000.

[141] 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*, vol. 19, no. 2, 2009.

[142] N. Sawalhi, R. Randall and 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*, vol. 21, no. 6, pp. 2616-2633, 2007.

[143] S. Zhao, L. Liang, G. Xu, J. Wang and 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*, vol. 40, no. 1, pp. 154-177, 2013.

[144] J. Antoni and R. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines", *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308-331, 2006.

[145] Y. Wang, J. Xiang, R. Markert and M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications", *Mechanical Systems and Signal Processing*, vol. 66-67, pp. 679-698, 2016.

[146] F. Tafinine and K. Mokrani, "Real time automatic detection of bearing fault in induction machine using kurtogram analysis", *The Journal of the Acoustical Society of America*, vol. 132, no. 5, pp. EL405–EL410, 2012.

[147] 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*, Glasgow, UK, 2015.

[148] T. Harris, *Rolling bearing analysis*. New York: John Wiley, 2001.

[149] J. Antoni, "Fast computation of the kurtogram for the detection of transient faults", *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 108-124, 2007.

[150] R. Zimroz, W. Bartelmus, T. Barszcz and 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*, vol. 46, no. 1, pp. 16-27, 2014.

[151] P. Samuel and D. Pines, "A review of vibration-based techniques for helicopter transmission diagnostics", *Journal of Sound and Vibration*, vol. 282, no. 1-2, pp. 475-508, 2005.

[152] C. Yuksel and A. Kahraman, "Dynamic tooth loads of planetary gear sets having tooth profile wear", *Mechanism and Machine Theory*, vol. 39, no. 7, pp. 695-715, 2004.

[153] L. Hong and J. Dhupia, "A time domain approach to diagnose gearbox fault based on measured vibration signals", *Journal of Sound and Vibration*, vol. 333, no. 7, pp. 2164-2180, 2014.

[154] T. Barszcz and R. Randall, "Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine", *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1352-1365, 2009.

[155] Bin Zhang, T. Khawaja, R. Patrick and G. Vachtsevanos, "Blind Deconvolution Denoising for Helicopter Vibration Signals", *IEEE/ASME Transactions on Mechatronics*, vol. 13, no. 5, pp. 558-565, 2008.

[156] Z. Feng, M. Liang, Y. Zhang and S. Hou, "Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation", *Renewable Energy*, vol. 47, pp. 112-126, 2012.

[157] R. Zimroz and W. Bartelmus, "Gearbox condition estimation using cyclo-stationary properties of vibration signal", *KEM*, vol. 413-414, pp. 471-478, 2009.

[158] Y. Lei, D. Han, J. Lin and Z. He, "Planetary gearbox fault diagnosis using an adaptive stochastic resonance method", *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 113-124, 2013.

[159] S. Guoji, S. McLaughlin, X. Yongcheng and P. White, "Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears", *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 76-89, 2014.

[160] Z. Chen, T. Wang, F. Gu, H. Mansaf, and A. D. Ball, "Gear transmission fault diagnosis based on the bispectrum analysis of induction motor current signatures", *Journal of Mechanical Engineering*, vol. 48, no. 21, pp. 84, 2012.

[161] P. McFadden and J. Smith, "An explanation for the asymmetry of the modulation sidebands about the tooth meshing frequency in epicyclic gear vibration", *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 199, no. 1, pp. 65-70, 1985.

[162] F. Gu, T. Wang, A. Alwodai, X. Tian, Y. Shao and A. Ball, "A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals", *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 400-413, 2015.

[163] A. Alwodai, F. Gu and A. Ball, "A comparison of different techniques for induction motor rotor fault diagnosis", *Journal of Physics: Conference Series*, vol. 364, 012066, 2012.

[164] M. Haram, T. Wang, F. Gu, and A. Ball, "An Investigation of the electrical response of a variable speed motor drive for mechanical fault diagnosis", *in Proceedings of the*

24th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM 2011), pp. 867-874, 2011.

[165] A. Alwodai, X. Yuan, Y. Shao, F. Gu, and A. Ball, "Modulation signal bispectrum analysis of motor current signals for stator fault diagnosis", *in 18th International Conference on Automation and Computing (ICAC), IEEE*, Loughborough, UK, 2010.

[166] N. Sawalhi, R. Randall and D. Forrester, "Separation and enhancement of gear and bearing signals for the diagnosis of wind turbine transmission systems", *Wind Energy*, vol. 17, no. 5, pp. 729-743, 2013.

[167] B. Vishwash, P. Pai, N. Sriram, R. Ahmed, H. Kumar and G. Vijay, "Multiscale Slope Feature Extraction for Gear and Bearing Fault Diagnosis Using Wavelet Transform", *Procedia Materials Science*, vol. 5, pp. 1650-1659, 2014.

[168] Bonnardot, F., R. B. Randall, J. Antoni, and F. Guillet. "Enhanced unsupervised noise cancellation using angular resampling for planetary bearing fault diagnosis", *International journal of acoustics and vibration*, vol. 9, no. 2, pp. 51-60, 2004.

[169] J. Tian, M. Pecht and C. Li, "Diagnosis of rolling element bearing fault in bearing-gearbox union system using wavelet packet correlation analysis", *Dayton, OH*, pp. 24-26, 2012.

[170] E. Faris, M. Greaves and D. Mba, "Diagnostics of a defective bearing within a planetary gearbox with vibration and acoustic emission", in *The 4th International Conference on Condition Monitoring of Machinery in Nonstationary Operations (CMMNO 2014)*, Lyon, France, 2014.

[171] Q. Zhang, Q. Hu, G. Sun, X. Si and A. Qin, "Concurrent Fault Diagnosis for Rotating Machinery Based on Vibration Sensors", *International Journal of Distributed Sensor Networks*, vol. 2013, pp. 1-10, 2013.