FAULT DETECTION AND DIAGNOSIS OF A MULTISTAGE HELICAL GEARBOX USING MAGNITUDE AND PHASE INFORMATION FROM VIBRATION SIGNALS

J X GU

PhD 2016

FAULT DETECTION AND DIAGNOSIS OF A MULTISTAGE HELICAL GEARBOX USING MAGNITUDE AND PHASE INFORMATION FROM VIBRATION SIGNALS

James Xi Gu

A THESIS SUBMITTED TO THE MANCHESTER METROPOLITAN UNIVERSITY IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

SCHOOL OF ENGINEERING

FACULTY OF SCIENCE AND ENGINEERING

MANCHESTER METROPOLITAN UNIVERSITY

MARCH 2016

ABSTRACT

Vibration generated by a gearbox carries a great deal of information regarding its health condition. This research aims primarily on the detection and diagnosis of tooth defects in a multistage gearbox based on advanced vibration analysis. Time synchronised averaging (TSA) analysis is effective at removing noise but it is inefficient in implementation and in diagnosing different types of faults such as bearing defects other than gears. Conventional bispectrum (CB) can eliminate Gaussian noise while it preserves the signal's phase information, however its overpopulated contents can still provide inaccurate information regarding to different types of gear faults. Recently developed modulation signal bispectrum (MSB) has the high potential to lead to the high accuracy of diagnostics of gearboxes as it more effectively characterises modulation signals such as gearbox vibrations. Therefore, the research takes MSB as the fundamental tool for analysing gearbox vibration signals and developing accurate diagnostic techniques.

Firstly, it has realised that conventional techniques often ignore the effect of phase information in gearbox diagnostics. This thesis then focuses on developing CB and MSB based techniques for detecting and diagnosing of gearbox faults.

Secondly, it has found that vibration responses from a multiple stage gearbox have high interferences between amplitude modulation (AM) and phase modulation (PM) which can be formalised from both gear faults and inherent manufacturing errors. However, the faults can induce wider bandwidth vibrations. Correspondingly, optimal component based schemes are also developed based on the use of MSB coherence results.

Then the proposed MSB method allows an effective gearbox diagnosis using the signals in a narrower frequency band that is below twice the rotational frequency plus the highest meshing frequency amongst different gear transmission stages, being more suitable for wireless network condition monitoring systems.

It has also found that the signals at resonance frequencies has a higher signal-to-noise ratio and more effective for obtaining accurate diagnosis. Also software encoder based TSA was found to be not robust and accurate due to the influences of noise and referencing components on obtaining a reliable phase signal for implementing TSA.

Finally, the diagnostics carried out upon different fault cases using both CB and MSB have verified the proposed approaches can provide accurate diagnostic results, and with the new MSB based detector and estimator being more effective in differentiating between diffident fault locations for two local and one non-uniformly distributed tooth damages in a two stage helical gearbox.

.

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

LIST OF ABBREVIATIONS

Time synchronised averaging	(TSA)
Conventional bispectrum	(CB)
Modulation signal bispectrum	(MSB)
Amplitude modulation	(AM)
Phase modulation	(PM)
Condition monitoring	(CM)
Microelectromechanical systems	(MEMS)
Root-mean squared	(RMS)
Acoustic emission	(AE)
Alternating current	(AC)
Direct current	(DC)
Higher order spectra	(HOS)
Hilbert transform	(HT)
Short time Fourier transform	(STFT)
Wigner-Ville distribution	(WVD)
Wavelet transforms	(WT)
Crest factor	(CF)
Fast Fourier transform	(FFT)

Signal-to-noise ratio	(SNR)
Fourier transform	(FT)
Discrete Fourier transform	(DFT)
MSB sideband estimator	(MSB-SE)
Modulation power index	(MPI)
Power spectrum	(PS)
Current transformer	(CT)
Printed circuit board	(PCB)
Analogue to digital converter	(ADC)
Data acquisition system	(DAS)
Instantaneous angular speed	(IAS)
Integrated circuit piezoelectric	(ICP)

NOMENCLATURE

A	Signal magnitude for amplitude modulation
B	Bispectrum
E	Gear error
P	Power spectrum
T	Torque
X	Signal in the frequency domain
Z	Tooth number
Z^2	Modulation power index
I_g	Moment of inertia of driven gear in the first stage
I_p	Moment of inertia of driving gear one (pinion) at the first stage
N	Total number of samples
a, b	Time shift
b^2	Bispectrum coherence
c_{pg}	First stage meshing damping
e	Gear errors including different gear fault modes
f	Frequency
f_c	Carrier frequency
$f_{m,}$ f_x	Modulating frequencies
f_{me}	Meshing frequency
i	$i=1,2$, subscription for gear 1 and 2 respectively
m, n	Time index
x, y, s	Time domain signal
t	Time
$r_{p,}$	Base radius of the driving gear in the first stage
r_g	Base radius of the driven gear in the first stage

y_p	Vertical displacement of the driving gear the first stage
y_g	Vertical displacement of the driven gear the first stage
m_p	Mass of the driving gear in the first stage
m_g	Mass of the driven gear in the first stage
k_{pg}	First stage time varying meshing stiffness
k_{pb}	Driving gear bearing stiffness
k_{gb}	Driven gear bearing two stiffness
β	Magnitude of phase modulation
δ	Signal phase
φ	Angular displacement
φ_p	Angular displacement of driving gear;
φ_g	Angular displacement of driven gear
γ	Gear error amplitude
ρ	Eccentric error amplitude
σ^2	Variance
μ	Mean value
τ	Time delay
ω	Angular frequency

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

I dedicate this work to my dear mother and father for the lifelong and boundless love that I received from them as well as his unwavering supports. Without their supports, reaching this point would not have been possible.

Few goals are ever achieved without the help of teachers, mentors, colleagues, family, and friends. For the assistance, support, and guidance I have received while completing my postgraduate studies I am indebted to Dr Al-Hussein Albarbar, I would like to express my sincere gratitude to him. I would like to acknowledge invaluable input and countless enlightening meetings, discussions and conversations with them.

Thanks to Dr Mansaf Haram for preparing gear samples and supporting vibration dataset acquisitions during the full course of the run-to-failure gear test.

My special thanks go to my colleagues and friends, I was fortunate to have nice peers in my research group.

My thanks also go to the technical team for their kind help and cooperation.

CONTENTS

CHAPTER 1.

INTRODUCTION

1.1 Background

All machines with rotating parts give rise to both vibration and signal. Every mechanical rotating system has its own unique specific vibration characteristics, which is related to both the condition and the manufacture of the system. If the current condition of the machine changes, the vibration signals produced by the machine will also be affected. The change in the vibration signature can be analysed to detect incipient defects before they the occurrence of any serious failure. This is the principle in which the majority condition monitoring (CM) techniques are based upon.

CM saves resources and operating costs through increased maintenance efficiency [1], which will decrease the machinery downtime. It will also reduce the chance of machined breakdowns, total failures and any potential accidents.

The fundamentals of condition monitoring has already been applied and used in industry for an extensive amount of time. An experienced operator can monitor the state of rotating machinery by simply listening to the vibrations from its casing. However, it often takes up to many years for an operator to develop the necessary skills and experience, with some defects appearing very seldom. Therefore, it would be more advantageous for the monitoring process to be more automated.

Gearboxes are a critical component of many industrial machineries such as power generators, wind turbines, and mechanical vehicles [2]. Failures of the gearbox can result in deficient, dangerous and unstable operating environments which can lead to injury of the operator and

system shutdown leading to financial losses. As such, early diagnosis of the gearboxes' condition has been a focus leading to the development many different diagnostic approaches developed in the condition monitoring community, primarily to avoid any negative influences and potentially damaging accidents [3].

Common faults of gearboxes including manufacturing defects, mounting defects and defects appearing during transmission [4] generate their own unique fault features, different types of signals, such as vibration, motor current output, sound, lubricant analysis, temperature, and instantaneous angular speed can be used for detection and diagnosis of faulty gearboxes [1][3][4].

Vibration signals are most commonly used in maintenance strategies for gearbox CM, the vibration signal is easily acquired using accelerometers, attaining information that can reveal the primary excitations of gearbox dynamics. Airborne acoustics or noises are also closely correlated to vibration, as such, acoustic emissions have also been investigated vigorously for fault detection and diagnosis of gearbox in the past few decades. Regardless, both vibration and acoustic signals which is collected also contain noises and undesirable information. Therefore, detailed analysis using advanced signal analysis techniques are required in order to extract the required useful information for an accurate fault diagnosis.

Most current techniques investigated for gearbox faults detection and diagnosis have mostly focused around time-domain and frequency-domain based analysis, which include applies techniques based on statistical parameters of waveforms, spectra, amplitude and phase demodulation. Most conventional techniques are able to reliably detect and provide an indication of the presence of faults. However, they are unable to provide any detailed information regarding location or severity of the fault, since they are not suitable for application to non-stationary components [5]. Gearbox vibration signals commonly contain

significant noise and possess properties, which are of a non-stationary nature. As a result, it can be difficult to detect the early symptoms of a potential failure without an appropriate analysis tools.

To capture better the non-stationarity, the joint time-frequency analysis is developed for a more reliable and effective method for rotating machinery condition monitoring. Wavelet transform (WT) is also a standard and powerful time-frequency signals processing tool, WT is already commonly utilised in industry for machinery fault diagnosis and CM. WT has the capability to reveal the time-frequency characteristics of signals, thus it is more efficient at locating any small transients and further enhance any existing peaks [6].

The development of more complex and reliable CM techniques has also lead to a more automated embedded CM system have been developed recently. The system will firstly measure the real physical acceleration and extract key features using advanced signal processing techniques, which will then be compared to a set of know predetermined values in order to detect faults. The benefit of an embedded CM solution is that it allows the possibility of fully automated fault detection eliminating human error. Another advantage is that the real time status of machinery can be provided by deploy a low cost effective automated sensing CM strategy.

Real time fault detection at the source will reduce the amount of data that needs to be transmitted to the central processing computer; this will allow the system to send the data using a low power and low cost on board wireless communication system. Albarbar et al [7] performed a study which proved that MEMS accelerometers are of sufficient quality to be utilised for rotating machinery health and condition monitoring purposes. This has led to a new area of research within the field, allowing cheap sensor devices to be created which will enable industry to move toward condition-based maintenance (Jardine et al [8]) rather than

periodic or breakdown based maintenance. This will reduce maintenance costs, and avoid accidental shutdowns in the critical areas of the plant.

1.2 Research Motivation

Most previous researches focus on the utilisation of signal magnitude at specific frequency components for fault detection and diagnosis. It means that signal phase information, as an essential factor of a signal, has been largely ignored, which is due to that it is difficult to obtain phase values accurately in signal processing. However, as the one of fundamental parameters, the signal phase plays an indispensable role in characterising complicated vibration responses and hence it is necessary to examine phase responses in conjunction with amplitude and frequency values in order to achieve accurate characterisation of measured vibration signals for fault diagnosis.

A technique developed by Yang et al [9] concluded that using a combination of power spectrum, bispectral and bicoherence vibration analyses provides promising results for detecting the condition of bearings. A similar method was also used by Arthur et al [10] for CM of induction machines. Especially, these analyses have demonstrated the usefulness of phase information embedded in the bispectrum estimation for fault diagnosis. However, there have so far been little research regarding its application for gearbox CM. In particular, conventional bispectrum (CB) and an enhanced modulated signal bispectrum (MSB) have not been evaluated systematically onto vibration data collected from a multistage gearbox.

1.3 Research Aim and Objectives

To fill the gap that limited works on using signal phases for diagnosis, this research aims at on developing detection and diagnosis approaches based on not only signal amplitude and frequency information, but also utilizing phase characteristics as much as possible. To achieve it, the research is carried out to prioritise according to following main objectives:

1. To review previous researches on vibration based condition monitoring techniques with highlights on signal processing techniques.
2. To examine typical signal processing methods in analysing vibration signals from a multistage gearbox with combination of information regarding amplitude, frequency and phase.
3. To develop signal models for characterizing gearbox vibration responses under different fault cases.
4. To evaluate the performance of synchronous average (TSA) based diagnosis.
5. To evaluate the performances of conventional bispectrum (CB) analysis based diagnosis.
6. To evaluate the performances of modulation signal bispectrum (MSB) analysis based diagnosis.
7. To investigate effective features based on MSB analysis for gearbox diagnosis.
8. Comparing the effectiveness of the TSA, CB and MSB analysis methods for diagnosis of a gearbox's condition.

1.4 Organisation of Thesis

Chapter 1 briefly explains the rationale of the thesis. It follows the research motivation and the establishment of this research's aim and overall objectives. Besides, it also outlines other chapters in the rest of the thesis for the ease of following up the presentation of the research contents.

Chapter 2 is a general literature review regarding the need for condition monitoring for all critical components of a machine and different measurement based technologies. It begins with a brief background of condition monitoring followed by its developments and growth.

After that there is a review of where a lot of condition monitoring is currently used, with more information on the rise of analysing the vibration signals generated by machinery. There is a view regarding the modernisation of condition monitoring techniques by utilising emerging microelectromechanical systems (MEMS) sensors and wireless networks.

Chapter 3 focuses on overviewing current data processing techniques which are applied vibrations signals for the purpose of condition monitoring. Techniques that are potentially able to include phase information are focused more.

Chapter 4 starts with a brief description of the operations of a multi-stage gearbox along with the common faults. A mathematical model of a two-stage gearbox is then developed and evaluated in Matlab platform numerically and with measured data experimentally. Furthermore, by taking into the phase information it suggests a new diagnostic framework that highlights the utilisation of both amplitude and phase modulations along with efficient noise suppression.

Chapter 5 develops MSB based techniques for gearbox CM under the guidance of the framework. These techniques are then evaluated by applying them to signals generated based the gearbox model and benchmarked with corresponding techniques based on TSA and CB analysis.

Chapter 6 provides information regarding the experimental rig and the dataset. That which consists of three key components, an AC generator, a two stage helical gearbox and a DC generator to dissipate the load in a resistor bank. Three sets of data were collected, firstly when the gearbox was healthy, the second set of data was seeded with local tooth faults on the meshing gears, and finally in the last set of data contained, the gear possessed non-uniform distributed faults as a result of long operational fatigue.

Chapter 7 presents implementation of TSA and the results obtained by applying TSA analysis to the data collected in chapter 6.

Chapter 8 uses the conventional bispectrum analysis for detecting and diagnosing faults in a gearbox. This method combines the features from both the lower and higher sidebands to minimise the effects of Gaussian noise and highlight the faults.

Chapter 9 applies the MSB to signals from a gearboxes operating with 30%, 60% and 100% tooth breakage faults. Results showing that MSB is more reliable for fault detection and diagnosis compared to other techniques based on the same datasets.

Finally, main conclusions drawn and future works identified by this research are presented in chapter 10.

CHAPTER 2.

CONDITION MONITORING

2.1 Introduction

Decades ago, the majority of rotating machinery were maintained by reactive maintenance, for example compressors and turbines will keep operating until a fault develops [11]. However, the unexpected downtimes associated with this type of maintenance (reactive or run to failure based maintenances strategy) is undesirable and costly. Run to failure also possess safety issues and may lead to a catastrophic failure of the entire system [12]. Periodic maintenance minimises the risk of failure, but scheduling regular inspections is wasteful if it is not required [13].With the help of proper condition monitoring (CM) and fault diagnostic techniques, predictive maintenance (PDM) and condition-based maintenance (CBM) have become increasingly adopted in industries [14][15][16].

CM can be defined as continuous online inspection regarding the health of the machinery throughout its operational life [17]. It is important to be able to detect faults while they are still developing, this is known as incipient failure detection [18]. The incipient detection of the critical component failures also prevents catastrophic failures and allows for a safe operating environment. It is becoming ever more important to deploy comprehensive CM schemes for continuous assessment of the rotating components in machines [19].

CM can provide advanced warning of imminent failures, detailed information regarding the health of a machine makes it possible to better schedule future preventive maintenance and repair work as well as allowing the machine to operate at optimal working condition. This can result in minimum down time and optimum maintenance schedules. CM and fault diagnosis scheme allows the machine operator have key knowledge regarding to the type of

fault which needs fixing, thus the repair engineer can have the necessary spare parts before the machine is stripped down, thereby reducing outage times and cost induced.

2.2 Condition Monitoring Techniques

Recent advances in electronic processing capabilities which are compatible with advanced signal processing and data acquisition systems has reduced the overall equipment cost associated with CM, as a result online monitoring and diagnostic systems becoming available to all industrial production. These techniques have been proven to be capable of detecting and localising a faulty component at an early stage to prevent the system from full failure [20].

The choice of CM technique depends on several factors, such as the resources available, the component being examined and the desired outcome of the maintenance service. Human sense such as an eye inspection or smell could be satisfactory evaluation for the condition of an electrical or mechanical machine when it is near to be failure; on the other hand an advanced CM method may be demanded for more advanced action courses. Selecting the correct strategy is a task in itself, optimising between cost of implementing the maintenance strategy and cost of breakdowns [21]. The most commonly used CM techniques in modern industry are provided in the following sections.

2.3 Torque

Torque measurement has also been utilised for drive train fault detection. The rotor faults may cause either a torsional oscillation or a shift in the torque-speed ratio. Such information can be used to detect rotor faults, e.g. mass imbalance. Also, shaft torque has a potential to be used as indicator for decoupling the fault-like perturbations due to higher load [22]. However, inline torque sensors are usually highly expensive and difficult to install.

Therefore, using torque measurement for drive train fault diagnosis and condition monitoring is still not practically feasible.

Air gap eccentricity is a common fault that exist in many electric motors, they produce sidebands at special frequencies in the air gap torque [22]. However, it is not feasible for the air gap to be measured directly. The difference between the estimated torques from an ideal model allows for the early diagnosis of broken bars. From the input terminals, the instantaneous power includes the charging and discharging energy in the windings. Therefore, the instantaneous power cannot quite represent the instantaneous torque. From the output terminals, the rotor, shaft, and mechanical load of a rotating machine constitute a torsional spring system with its own natural frequency. As such, the harmonic attenuations of the components from the air gap torque transmitted through the torsional spring system will differ to that of the torque components.

2.4 Temperature Monitoring

Thermal monitoring (thermography) is already a well-established CM technique. The emitted heat energy is used for CM in this technique. The temperatures and machine thermal patterns can be monitored during its operation, the information can be readily compared to previous known parameters for fault diagnosis.

However, thermography is more suitable for mechanical machines, as it can quickly indicate a problem area. Typically, the heat must transfer from machine part (gear) to the surface of the machine and appears as a pattern to be detected by infrared camera [23]. Any deviation from the normal machine thermal signature (increase or decrease) is considered to be an indication of a possible error [24].

Thermal imaging increases the ability to expect machine component failure, before personal injury, equipment damage, or a costly shutdown occurs. One of the advantages of

thermography is that it can be applied to a wide range of machines such as motors, pumps, pulleys, bearings, conveyors, fans, drives etc. One of the drawbacks of this method is that care must be taken be aware of the atmosphere where the machine is working, because any change in atmospheric temperature factors can have a direct influence on the accuracy of the measurements [25]. Also, because there are numerous components that are heat sources in a real mechanical system, thermal CM is unable to specify from which source the heat is radiating from. In addition, additional heat usually means that the faults already prevalent, thereby temperature based monitoring cannot give sufficient advanced indication of faults for undertaking maintenance actions.

2.5 Oil Analysis and Wear Debris Monitoring

Oil analysis and wear debris technique for rotating machinery is one of CM techniques that is currently very popular. The oil is essential for smooth operation of all rotating machinery, oil analysis can diagnose the machine condition by investigating wear elements in the oil [26]. In particular, it monitors the three key components that make up the base structure of the oil (the base oil, lubricant additives and contaminants). Continuous gear wear influences the machine performance and changes the characteristics of the oil.

Oil analysis based CM provides an early notice about machine health. In addition, it can sense risk earlier than vibration methods[27]. A sample of lubricants from rotating machine equipment is taken and then analysed it for relevant information regarding the machine under inspection. One weakness of this method is that the oil test is not sensitive to the element sizes related to abnormal wear in rotating machinery, and only provides an indication of when the oil should be changed [27]. In addition, the instrumentation cost is too high to be affordable for many industries and operational stability is poor to give consistent indication for machine conditions.

2.6 Acoustic Emission Monitoring

Acoustic emission (AE) based method for detection and diagnosis of machines faults in CM has been widely regarded as a powerful non-destructive CM technique. When rotating components are in operation, its discontinuities create mechanical stress, as a result energy is released in the form of both vibration and noise. AE investigates the generated signals of acoustics and noise by using appropriate signal processing techniques, which are picked up by microphones, the operating range of AE usually vary between 100KHz and 2MHz[29]. The lower frequency limit is due to numerous background acoustical interference such as friction or noise generated from the operating process. At the upper limit, the acoustic signal is likely to be attenuated and will require a sophisticated detection unit.

Acoustic analysis offers non-intrusive data collection which is easy to mount, inexpensive, and with sensitive microphones is a good choice for online CM, including bearings, gearboxes, and engines. Acoustic monitoring has significant advantages particularly when the noise problem is under investigation; as it delivers direct clarification of noise sources. Newer AE signal techniques such as those developed by Albarbar et al [30] have been successful in isolating low acoustic areas interest in a high acoustic emitting environment. However, the effect of the acoustic atmosphere certainly plays a vital part in machines condition monitoring, one of the main problems being potential contamination by background noise, such as various contributions from close AE sources in its close vicinity, for example the loading, cooling fan, and driving motor [31]. The other problem is that AE data storage and processing is still a issues as it is acquired at a very high sampling rate, typically over 10MHz.

2.7 Motor Current Signature Analysis

Motor current signature analysis (MCSA) is considered to be amongst the most popular CM technique in detecting and diagnosing the existence of electrical and mechanical faults in

electrical machines such as those in induction motors. In addition, it has been revealed to be a powerful technique for monitoring the condition of downstream driven mechanical components such as gears, bearings and shafts.

MCSA is based on the idea that an alternating current (AC) or direct current (DC) electric motor driving a mechanical load also operates as a permanent transducer in itself, sensing load and speed variations of the mechanical system then converting them to a variable induced current in the motor windings.

These alterations in the motor current are carried to a hall effect current sensor via the motor power cables, and can be detected by this current transducer and then processed by appropriate equipment to obtain information regarding the condition of the machine [32]. This process is also non-intrusive and can be accomplished without disturbing either the motor drive or the driven system.

From these indicators taken at suitable time intervals, early information can be extracted about the health of the equipment being monitored in both the time and frequency domains. It is this process of sampling the motor current and performing signal analysis on it that is recognised as MCSA. Equipment for collecting and performing signal analysis relatively obtainable and affordable [33]. Modern CM systems easily allows the motor current can be analysed remotely in real time from a control panel, providing an easy way to monitor inaccessible equipment.

Although MCSA allows the implementation of a CM system to be affordable, remote and system level monitoring, its sensitivity is questionable for rotating machines with large initial moments. Particularly, it may not be able to give fault indications to individual components or sub-systems located distant to the driving motor.

2.8 Vibration Monitoring

Vibration analysis technique is widely used for early faults detection in rotating machineries, such as gearbox, motor, pump diagnosis. The vibration signals of a gearbox are composed of signatures based on its gears, the detection of the fault is achieved by analysing the signal of each individual faulty component. Many different possible signal processing methods can be applied to the gearbox vibration signal to provide a more accurate and detailed diagnosis.

The response of the machine or equipment is monitored by vibration analysis at different forces applied on it. Then this response (measured signal) is analysed to obtain clear information about the machine health conditions [34]. Vibration technique has been known as one of the most comprehensive condition monitoring techniques, and it is an essential practice of various predictive maintenance programs**.** While all machines generate vibration in varying forms and sizes, vibration analysis is most commonly deployed for fault severity diagnosis in rotating machinery [35].

In practice, accelerometers and other displacement based sensors that are used to collect vibration data with a wide band of frequency, the raw vibration signal will contain information from each rotating component of the machine. The collected data are investigated by using numerous signal processing methods such as frequency content, root mean square (RMS), peak values, crest factor, cepstrum and higher order spectral (HOS) analyses [36].

Many types of machine faults may not create enough significant changes in the signature of vibration compared with its baseline, so vibration monitoring technique is not adequate for all CM [37]. In general, the vibration signal produced from a machine is the sum of many components and structures of vibration signals to which they may are coupled. However, rotating machines faults produce different vibration signals (components) at different frequencies and locations that may often be linked to exact machine fault conditions [38].

Despite that fact, it still remains that vibration based CM techniques often provide the best reporting of all failure prevention techniques, it has since been commonly used for online rotating machines monitoring such as bearings, shafts, pumps, gearboxes, motors, etc. [39], [40]. However, careful consideration must be taken into account for the position of where the accelerometer is attached, since the collected vibration signals could be a combination of various sources of signals. As the consequence of multiple sources together with coupling effects, the signal can have high information contents but might be a function of non-linear and non-stationary sources, and may be affected by transmission. Such weakness could make vibration monitoring technique not the best technique to be selected [38] unless effective decoupling and separation methods can be available.

Therefore, it utilises the full information included in measured vibration signals for accurate diagnosis, vibration based technique needs to concentrate its efforts on improving signal processing methods in order to make significant advancements for solving problems with multiple source coupling and noise contamination.

2.9 Chapter Summary

Based on this review, typical techniques for mechanical condition monitoring are summarised in Table 2.1, with the target components, advantages and disadvantages listed for each, which may provide references for technique selection in applications and research efforts. In addition, their feasibilities are also evaluated based on a multistage gearbox system.

Table 2.1 shows the common parameters to analyse for condition monitoring, in comparison with other methods, analysing the vibration data for the gearbox, bearing and shaft is seen as reliable, however its main disadvantage is that it is a relative expensive system to set up requiring complex transducers and expensive signal processing systems based on conventional measurement systems. However recent advances in signal processing units and cheap MEMs accelerometers means that even complex diagnostic systems are more viable.

There are also studies where a combination of techniques has been proven useful for detection and diagnosis of faults in machines. For example, Loutas et al [41] utilises a combination of vibration, AE and oil debris analysis for condition monitoring of rotating machinery. Based on the findings from [42][43], it was suggested that vibration signal analysis proved to be superior to AE for detection of faults in gearboxes. However current methods for vibration based CM systems such as wavelet filtering requires high sampling rates [44],

Therefore, a new novel MSB based signal analysis technique is developed in this thesis to provide reliable fault detection and diagnosis. This novelty requires lower bandwidth than existing techniques (at below twice the rotational frequency plus the meshing frequency).

Table 2.1 Summary common techniques used in condition monitoring

Measured Parameters	Monitored Components	Advantages	Disadvantages
Vibration	Gearbox, Bearing, Shaft	Reliability. Standardisation.	Expensive More subject to sensor failures Limited performance at low speeds
Torque	Gearbox, Rotor	Direct measurement of rotor load	Expensive Intrusive
Oil/Debris Analysis	Gearbox, Bearing	Direct characterization of bearing condition	Limited to bearing with a closed-loop oil supply system Expensive for online operation
Temperature	Gearbox, Bearing	Standardisation. Can be compared to previous known values. Straightforward technique.	Embedded temperature gauge required Many other factors can cause temperature variations
Acoustic Emission	Gearbox, Bearing	Able to detect faults at an early stage Good for low-speed operation High signal-to-noise ratio Frequency range far from load perturbation	Expensive Requires very high sampling rate
Motor stator currents and voltages	Gearbox, Bearing	Only applicable to systems with electric motors or generators No additional sensors required Inexpensive Non-intrusive Easy to implement	Angular displacement rather than force based Difficult to detect incipient faults Low signal to noise ratio.

CHAPTER 3.

OVERVIEW OF SIGNAL PROCESSING TECHNIQUES

3.1 Introduction

Signal analysis is used for extracting key information hidden in the vibration to determine the condition of the machine as well as diagnosing any faults that may exist at different locations and severities. For example, in a healthy the vibration signal generated by the gearbox is dominated by its meshing frequency and its harmonics [45]. A fault such as local tooth breakage or a distributed tooth wear will exhibit a change in the vibration energy signatures at these frequencies, in particular at the sidebands which are spaced at different frequencies of interest.

In this chapter, the vibration waveforms are examined under the most common signal processing techniques. Firstly, the peak value and root mean square (RMS) are evaluated, they are commonly used to show the amplitude and strength of a signal, the peak value represents the maximum value of the waveform, whereas the RMS provides the effective value of the varying waveform. Additionally, the Kurtosis of a signal provides an indication of the "tailedness" of the waveform. Another commonly used analysis tool is the crest factor (CF), CF is ratio of a signals peak value to its RMS value, this shows the amount of impacting or the extremeness of a waveform. Both kurtosis and peak value are calculated to reflect signal structures with dimensionless parameters to avoid the influence of operating parameters. These common waveform parameters are used overwhelmingly in the field of CM.

This overview of time domain based analysis methods is then followed by a general outline of different frequency domain based signal processing techniques. They are the most commonly used signal processing method, although they ignore important phase information.

Furthermore, this review covers advanced techniques including typical time-frequency analysis methods such as short time Fourier transform (STFT), Wigner-Ville distribution (WVD) and wavelet transforms (WT) which are suitable for non-stationary signals and take into account phase information to a certain degree. TSA include explicitly phase information and is critically effective for noise separation. Bispectrum is particularly interesting as it takes into account phase information implicitly and highlights non-linear effects.

On the other hand, there are also many other methods which have been actively investigated for processing gearbox vibration signals. This class of methods include cyclic spectrum analysis [46][47], empirical mode decomposition [48][49] etc. They are deemed to be implemented with either little consideration of signal phases, or with limited noise reduction capability, or the needs of considerable pre- or post-processing, or with less meaningful to physical systems. Therefore, they will be not discussed further.

3.2 Time Domain Based Amplitude Analysis

Figure 3.1 shows a simulation of typical vibration signals that can be collected from a transducer mounted on the gearbox casing, the simulated signals were generated using Matlab. The first case shows a healthy signal with a meshing frequency of 120Hz, the second signal contains a local tooth fault and the final signal contains a fault with some Gaussian noise. The figure below is used to provide a brief overview of the typical signal that requires analysing. Further information on the construction of the simulated signals and their explanations can be found in Chapter 4.

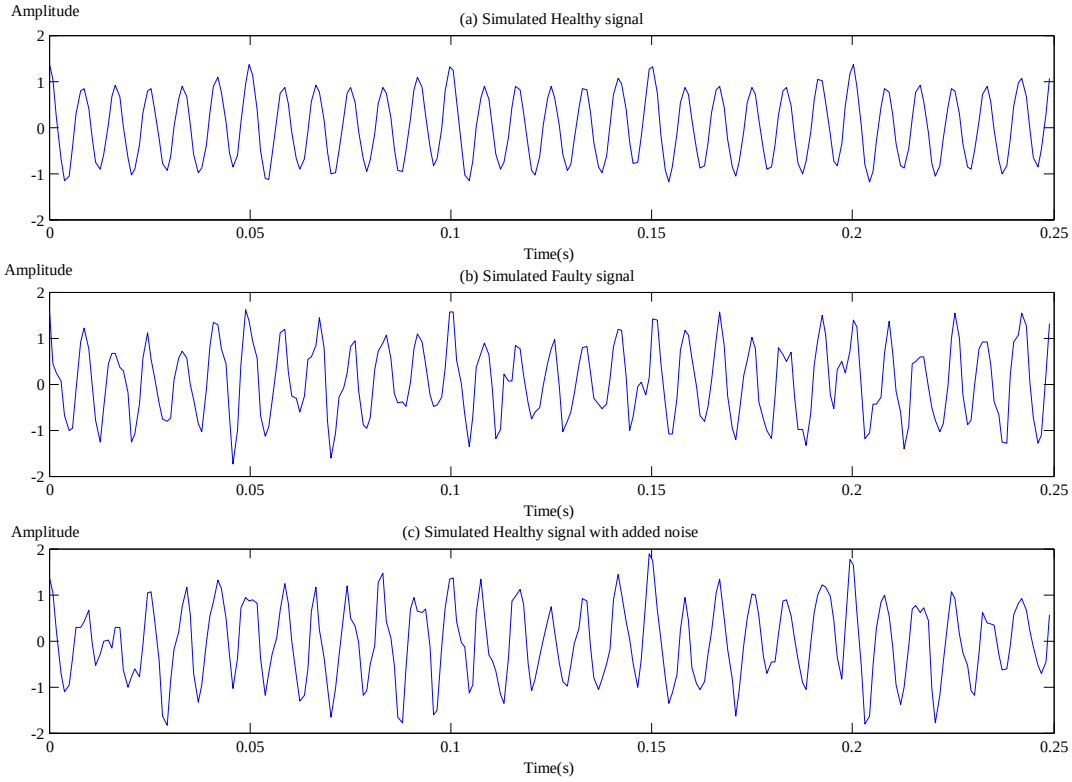

Figure 3.1- Matlab simulated gearbox vibration signals of three different cases, healthy, damaged, and the damaged case with background Gaussian noise.

Analysing the vibration signal of a system is often one of the quickest and simplest methods used for fault detection [50]. Time-domain based analysis techniques uses the information provided by the time-based variation of the amplitude contained within a typical gearbox vibration signal to detect the presence of faulty gears. Fluctuation of the amplitude can provide early indication that a fault is likely present, the frequency of the impulses of interest can then provide information relating to the likely faulty meshing gear pairs [51].

Time domain based parameters are often applied if the periodicity of the vibration signal is predictable, the fault type that in existence also need to be detectable impulses. Analysing the

signal in its time domain enables changes in the vibration signature easily recognised, but it is incapable of locating the source of a fault nor can it distinguish the severity of a fault.

Most rotating mechanical systems will generate sufficiently high vibration when operating to provide a clear waveform in the time domain which can be picked up by transducers. However, when a fault begins to develop, the change in vibration is usually be very small and therefore may not be detectable with simple processing methods. As the severity of the fault progresses, the size of the change in its vibration signal becomes greater and more notable, but still may prove insufficient for reliable fault diagnosis [52].

Healthy mechanical systems are generally considered to be deterministic, its vibration signal can be predicted based on its past operational properties. However, systems such as a gearbox with a tooth breakage fault possesses characteristics which are difficult to predict, and are characterised as non-deterministic signals. The values can be estimated by recognising some of its statistical parameters which can be used to predict fault progression [53].

Time-domain based parameters are already commonly used for various rotationary related mechanical fault detection include: Peak Values, RMS, Kurtosis and CF [54][55][56]. These parametric indicators are collectively known as “condition indices”, [57]. Given prior investigation into the values of healthy operational parameters, these condition indices can provide a rudimentary fault detection system. The values for the condition indices is generally related to the severity of the fault; this allows the condition index to provide an indication of the status of a fault and whether or not it is growing.

Because of the simplicity of time domain based parameters, a maintenance engineering using simple visual inspection of the vibration waveform over time can quickly apply most of the

time domain based techniques. However, processing the signal will provide a more accurate statistical value that can give an earlier indication regarding the severity of the fault and a more accurate diagnosis.

3.3 Signal Strength Based Diagnosis (Peak Value and Root Mean Square)

The peak value and RMS are both simple analytical values, both are commonly used to measure the overall size of the vibration signal. The peak value is calculated for a time domain signal $x(n) \quad n = 1, 2, \cdots N$, by:

$$PV = \max(x) \tag{3.1}$$

The RMS value is a measure of the power within the vibration signal. It is calculated from:

$$RMS = \sqrt{\frac{\sum_{n=1}^{N} x_n^2}{N}} \tag{3.2}$$

This feature provides a more accurate value with noise presence compared to the peak value. An increase in either peak value or RMS can be as a result of a fault which causes more energy release. However, any peaks of a short time duration will be averaged out if analysed by using the RMS value, therefore it is unable to detect any local faults. Also, neither indices can provide an indication relating to the source of a fault.

3.4 Crest Factor

Crest Factor is the ratio between peak value of the input signal and its RMS value. It is defined simplyas:

$$CF = \frac{PV}{RMS} \tag{3.3}$$

Given a localised fault, an impulsive peak is generated at the rotational frequency of the fault, increasing its PV depending on the severity of the fault, whilst its RMS value will increase by

only a small amount. Subsequently, this will also increase its Crest Factor. Most values for the Crest Factor lie between 2 and 6 [55]. Crest Factor values that are greater than 6 is a good indicator of some local fault.

3.5 Kurtosis

Kurtosis is the fourth order of a distribution and measures its “tailedness”, which is also known as the peakedness or flatness of the distribution in comparison to a normal distribution. Kurtosis is used as an indicator of major peaks in the data set.

The formula for calculating kurtosis is :

$$Kur = \frac{\sum_{n=1}^{N}[x(n)-\mu]^4}{[\sigma^2]^2} \tag{3.4}$$

Where N is the number of data points, $x(n)$ is the time series of the data, μ is the mean value of the data and $\sigma^2 = \sum_{n=1}^{N}[x(n)-\mu]^2$ is the variance of the data.

A high kurtosis value is an indication of the presence of a sharper peak with long tails, whereas a lower value represents a rounded peak. If the signal consists of just random noise made up of a Gaussian (Normal) distribution, it has a kurtosis value of 3, whilst a pure sine wave has a kurtosis value of Kur = 1.5. Kurtosis values of greater than 3 indicates that the distribution approach zero slowly and is relatively flat.

Vibration signals produced from a healthy gearbox will show a mostly consistent pattern, which has a kurtosis value of close 3.0, or lower. As a gear fault develops and becomes more severe, the impulses caused by local faults will also increase, this causes the kurtosis value to also increase, indicating that the distribution of the vibration is no longer a Gaussian distribution [54]. Given its robustness against noise, kurtosis is commonly used as a parameter for the detection of gear faults [57].

3.6 Frequency Domain (Spectrum) Analysis

Frequency-domain based signal processing is a well-adopted conventional technique used for vibration analysis, analysis of the signal in its frequency spectrum is extremely useful in the detection and diagnosis of faults in rotating machinery [58][59]. The waveform is transformed into can be broken down into several frequency components and the source of each component can usually be accounted for. It is therefore easy for analysts to focus on the frequency components which provide invaluable information for fault detection and diagnosis. Frequency domain based fault detection is most commonly used for many types of signals including vibration [60] and acoustic emissions [61].

Because the overall vibration signal is actually a combination of vibration produced over a broadband of frequencies; breaking the signal down in the frequency domain and displays the size the components in discrete neighbouring narrow frequency bands. Therefore, a common approach for detecting faults in most rotating machinery is to use the Fourier transform (FT) to present the vibration signal in its frequency domain.

As an efficient computation method, Fast Fourier transform (FFT) is commonly used to obtain FT if the machine rotates at a constant speed so that the measured signal's frequency components are not changing over time. The magnitudes at each frequency is as a direct result of the vibrations produced by each component by the machine. Knowing information regarding the details of rotating parts can allow each frequency component to be traced back to its source. Therefore, changes in vibration levels within a particular frequency band can be traced back to a particular component. Analysis of the component energy levels at different frequency bands can provide some fault diagnosis information and the fault location information.

The vibration signal's sidebands are generated by either amplitude modulation or frequency modulation, and they can often provide useful information relating to the type and severity of a fault. For example, an amplitude modulation can be used to diagnose tooth breakage in a particular gear, the frequency shift with the AM is also used to calculate the gear's rotational speed. This gives an indication of the location of the damaged gear within the gearbox. Errors which are as a result of manufacturing error, system set up error and non-uniform distributed fatigue related faults will cause a frequency modulation. As mentioned earlier, *Randall* has shown that the first three gear meshing harmonics and their sidebands provide sufficient information to accurately diagnose and identify faulty gears [62].

Since a signal's fluctuations in the amplitude of particular frequency peaks, or its sidebands gives an early detection of potential gear failure, it is potentially a very robust diagnostics tool. The spacing of the sidebands is dependent on the rotational properties of each component within the system and of the loading and on the transmission path. In real vibration signals, due to the number and complexity of such components, it can be difficult to distinguish between the multitudes of sources, therefore it is unreliable when trying to extract useful features purely from the vibration signal's frequency spectra if using only on a standard FFT. A low signal-to-noise ratio (SNR) will further exasperate such problems and it becomes near impossible to provide an accurate fault diagnosis.

3.7 Time-Frequency Based Analysis

Real signals are generally non-stationary in nature. Such machines typically generate a distinct frequency characteristic at a particular given instant of time. Therefore, in such cases, it is more preferable to use a technique that offers joint time-frequency analysis, therefore providing an indication of the way that the frequency varies against time.

Sinusoidal components which make up the signal will differ over the entire time domain and are not concentrated in time. Conventional FFT does not provide any signatures of the change in frequency components over time. Therefore, it would be useful to examine a signal simultaneously in both the time and frequency domains, allowing the inspection of their areas of interest that are concentrated in both its time and frequency domains, such as the frequency modulated Gaussian functions [63]. Thus, in order to overcome the loss of information using classical Fourier analysis, the easiest way is simply to compare the signal with its elementary functions that are localised in both the time and frequency domain simultaneously. Three such representative signal processing techniques which are explained in this chapter are:

- Short-Time Fourier transform (STFT);
- Wigner-Ville Distribution (WVD)
- Wavelet transform (WT).

3.7.1 Short Time Fourier Transform

The STFT is one of the most widely used method for examining non-stationary signals. The basic principle of STFT is to break up the initial signal into small time segments and apply the Fourier transform to each time segment to ascertain the frequencies that existed in that segment. The totality of such spectra indicates how the spectrum is varying in time.

The representation of the STFT using localised elementary functions is:

$$STFT(\tau, f) = \int_{-\infty}^{\infty} g(t-\tau)x(t)e^{-j2\pi ft}dt \qquad (3.5)$$

where the sliding window function *g(t)* is a Gaussian window defined by

$$g(t) = e^{-(t/a)^2} \tag{3.6}$$

where *a* represents the length of the window.

The formula (3.5), will provide the Fourier transform based on the function *F(t),* which is windowed by *g(t)* centred at time, *τ*. By repeating the same procedure at multiple translated time locations of *τ*, a portrayal of how the frequency components of the signal evolves with time can then be obtained [63], allowing both time and frequency features to be investigated concurrently.

The early application of time-frequency methods to gear faults began with the work of *Forrester,* STFT method was applied to the detection of gear failures [50]. The disadvantage of the STFT application is that the window width remains constant during the entire duration of the analysis. This is a limitation a high resolution is required to identify impulses which give a sudden change over time. Also the window's width is fixed after it is chosen, and therefore it cannot be changed during the rest of the transform.

The limitation of the STFT lies in the inability to achieve more detailed time localisation by using a smaller window functions. This is because the information from the resulting spectrum decreases as the time window shortens.

Eventually, by continuously decreasing the duration of the time window's interval, the answer that is obtained for that particular spectrum becomes meaningless. Moreover, similar to spectrum analysis, STFT does not have any noise suppression capabilities, as its averaging representation of the spectrogram is only the magnitude information of the signal under analysis.

3.7.2 Wigner-Ville Distribution

The Wigner-Ville distribution (WVD) is a more accurate joint time-frequency signal analysis method. WVD is a commonly used time-frequency analysis techniques, as it provides a very high resolution that allows for an accurate portrayal in both time and frequencies domains. However, WVD is particularly prone to aliasing related errors arising during the computation of WVD and to its nonlinear behaviour. In order to prevent aliasing occurring, the original real signal has to be transformed into a more complex analytical signal.

The Wigner distribution was originally developed by Eugene Wigner in 1932 [60], its main purpose was to study statistical equilibrium in quantum mechanics, the Wigner distribution was first used in signals processing Ville in 1947. It later became known as the Wigner-Ville distribution (WVD) amongst the signals processing community.

$$W(t, f) = \int s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{-j2\pi f \tau} d\tau \tag{3.7}$$

where $s*$ represents the complex conjugate of the analytic signal s which is calculated by Hilbert transform of the time domain signal s.

The behaviour of the frequency of a signal can be classified into two types: linear representation such as the Fourier transform (Short time Fourier transform) and quadratic representation such as the Wigner-Ville distribution.

The WVD provides a major advantage over STFT in that it can provide a better resolution in both time and frequency [61]. However, the downside with the WVD is that it natural possesses non-linear behaviour that usually creates significant interference (the cross-terms) between different signal components. In addition, it does not possess any noise suppression capabilities.

3.7.3 Wavelet Transform

The wavelet transform (WT) is a joint time-frequency analysis method similar to WVD but even more powerful. WT is independent of time so therefore it can also provide analysis regarding the local behaviour of the signal. STFT uses complex sine and cosine functions in order to map the time signal into its frequency domain. WT instead uses wavelets, which is a set of elementary functions which are independently widened and shifted as a function of time.

More recently, several WT analysis based techniques has been widely accepted as suitable signal processing methods for rotating machine fault detection and diagnosis. Decomposing a time domain series into time-frequency spectrum reveals the magnitude of the signal's frequencies as well as the time duration of each individual frequency [62]. This is especially useful for processing vibration signals from faulty rotating machinery, where the vibration changes due to a fault can vary significantly depending on whether it is either a distributed or local fault [63]. In particular, when examining the condition of a gearbox, WT is commonly used for the identification of all the possible transients in vibration signals that are generated by different faults. WT possesses multiple resolutions for localisation of the short time components, allowing many different types of gear faults to be highlighted with a single time-scaled distribution analysis from the transform [64][65].

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int x(t) \psi * \left(\frac{t-b}{a} \right) dt \qquad (3.8)$$

where $\psi *$ is also called the mother function. *a* represents the scale index and *b* represents the time shifting. The quantity *W(a,b)* reflects the signal's behaviour in around the region of (b, ω_0/a).

The most important characteristic of WT is its flexibility of the time window which can be narrowed or widened the time window accordingly, depending on a higher or lower signal

frequency[66]. There also exists a strong correlation between wavelet's scale and frequency, which is revealed by analysing the wavelets. This allows for the detection of extremely short transient signals in the time dimension, therefore it is able detect the non-stationary vibration signals[67]. As such, the WT is a processing method that is very well suited to the detection and diagnosis of early failures in gearboxes.

However, WT requires prior knowledge regarding the noise characteristics during the initial denoising process. In practice, this knowledge is usually unknown.

3.8 Time Synchronous Averaging

Time synchronous averaging (TSA) or time domain averaging, is a well utilised signal processing technique, TSA is often used to extract the periodic waveforms from signals which contain heavy noise [68][69]. It has regarded as an essential algorithmic tool for processing vibration signals in examining the condition of rotating equipment [70].

To compute a synchronous average signal *y(t)* from a time based signal *x(t)*, it needs to have an additional trigger signal such as pulse trains from a shaft encoder.

Where the frequency f_t is equivalent to its convolution [71]:

$$y(t) = c(t) * x(t) \tag{3.9}$$

c(t) is a train of N impulses each of amplitude 1/N, the impulses are spaced at intervals $T_t = 1/f_t$, and is given by:

$$c(t) = \frac{1}{N} \sum_{i=0}^{N-1} \delta(t + iT_t) \tag{3.10}$$

At the frequency domain, this is equivalent to the multiplying $X(\mathrm{f})$ and *C(f)* together, which are the FTs of the signal and pulse train

$$Y(f) = C(f)X(f) \tag{3.11}$$

With $C(f)$ defined as

$$C(f) = \frac{1}{N}\frac{\sin(\pi N T_t f)}{\sin(\pi T_t f)} \tag{3.12}$$

which is essentially a comb filter. For very large number of averages *N*, the only frequencies that are passed are exact multiples of the trigger frequency *f (t)*. Therefore, if the trigger is set to be frequency $f = ft - f$, only frequency lines at integer multiples of $ft - f$ will be kept and other components can be removed effectively.

Conversely, processing signals using TSA can often be too difficult and complex for obtaining reliable analysis in real applications. This is due primarily to the difficulties to have a reliable trigger signals, either from hardware such as shaft encoders or demodulating meshing component for a phase signal.

In addition, although the TSA is a powerful tool for isolating gear and shaft components synchronous with a particular shaft, it fails to isolate random vibration components because the subtraction of a single TSA signal from the original signal results in a combination of random signals and other shaft-synchronous components. To solve this problem, a frequency domain method was proposed by computing the FFT of the entirety of a signal, and simply removing spectral peaks at the discrete harmonics of the known periodic signals [72]. It also tried in the time domain method by consisting of multiple resampling and subtraction of time synchronous averages from a signal [73]. However, both of these issues have shortcomings. The frequency-domain approach must deal with the fact that the frequency of the harmonics leaks into adjacent windows, except for the very special case in which the numbers of sample per period and the number of periods are powers of two. In the case where there is leakage, removal of discrete peaks at shaft harmonics will not completely remove the periodic signal.

The time-domain approach described in has the issue that there are certain time-domain features that are common to the time synchronous average of two or more shafts, the signal corresponding to the gear mesh frequency being a universal example. Thus, repeated subtraction of individual TSAs will “over-subtract” certain features of the periodic signals [74].

In general, TSA can effectively suppress noise but it is not able to characterize nonlinearity. In addition, it still has many problems to implement an encoder-free TSA even though a great deal studies have careered out in different cases.

3.9 Bispectrum based Analysis

Detection and diagnosis information in gear vibration signals are always influenced by the presence of various noises including random noise, uninteresting vibration associated components, as investigated in [75] and [76] and even interferences. In addition, vibration of gearbox often exhibits nonlinear coupling mechanisms such as modulations. To overcome these problems, bispectrum analysis techniques have been investigated to analyse gear acoustic signals in [77] for more reliable features for gear fault detection [78]. Recently, a more in-depth study [79] has given theoretical framework on applying bispectrum to gear vibrations signals. These interesting studies have demonstrated that bispectrum analysis is potentially useful for characterising nonlinear fault features in line with noise suppression, and thereby for more accurate diagnostic features.

However, the authors in [79] have not shown a general diagnostic features that can be used for different operating conditions. This may be due to the fact that conventional bispectrum can results in massive spurious components for multiple stage gearbox.

3.10 Chapter Summary

The peak value, RMS, CF and kurtosis can be investigated easily from a time domain based vibration signal for fault detection and diagnosis. These signal processing techniques have already been adopted to be used under various rotational speed and load scenarios. The investigation shows that most statistical based parameters are highly affected by operating conditions and gear condition. By simply only measuring the peak value and RMS of a vibration signal, and comparing these values to a predetermined set of values measured from the vibrations from a normal healthy gear, the diagnosis of any defects and the severity cannot be performed. Even if the peak value and RMS are combined and using statistical analysis based on CF and kurtosis, it still lacked the ability to accurately detect gear defects in its earlier stages. The restrictions of such techniques often leads to an inaccurate diagnosis.

Amongst different techniques, bispectrum analysis combines both magnitude, frequency and phase information to achieve noise reduction and nonlinear characterisation. Therefore, this approach is potentially enabling to provide more reliable foundations for achieving accurate diagnosis, especially compared with techniques such as time-frequency analysis methods and wavelet transforms that have little effect on noise reduction, rather they only enhance the desired components through reassigning signal energy on time-frequency planes.

CHAPTER 4.

MODELLING GEARBOX VIBRATION RESPONSES

4.1 Introduction

Ever increasing demand to minimise the cost of maintenance whilst improving its quality has meant that CM of gearbox systems is a commonly investigate subject area. Gears operate under much more extreme conditions in comparison to most other components of the machine, this means that they are often the most common cause of failure. Their high operational requirements, especially at their teeth, where there is the most stress applied, means that gears require regular servicing or replacing [80][81].

There are three typical categories of gear failure that cause transmission error and gearbox failure: manufacturing faults (i.e. error in the tooth profile), installation faults (i.e. the alignment of the gears) and fatigue faults which occur during the work process (i.e. cracking of teeth) [82]. Sudden changes in the load of the gear teeth during operation is the primary reason for creating gear fatigue cracks appearing at the roots of the teeth, which weakens the overall structural durability of the gear.

Tooth breakage is often considered to be one of the most serious problem for gears because it ultimately leads to complete failure of the gearbox. A small initial fatigue crack at the base of a tooth will is not usually considered to be a major concern, however as the crack develops and propagates, the damage increases exponentially and so can often lead to complete failure of the tooth [83]. If the crack can be detected early and its development analysed, the gear can be replaced before severe breakage occurs.

As the pair of gears' teeth mesh and rotate there will be continuous periodic variations in torque and force, the primary vibration signal that is generated will be at the tooth meshing frequency [84]. There are also various friction forces and other noise throughout the gearbox that will be picked up by the vibration sensors at the gearbox case, this again creates more random components that distort the vibration data that is collected.

Martin et al have shown that the amplitude of the vibration signal will be primarily dependent on both the size of the forces and the physical properties and as well as the geometries of the gears themselves. The amplitude is also affected by other factors such as the tooth stiffness, shaft misalignments, errors in the manufacturing profile and pitch, gear eccentricity and load that is distributed onto each tooth as an overall function of the angle of rotation [85]

When the teeth mesh together, the vibrations will oscillate from the gear onto its mounting shaft, then onto the shaft bearing and finally to the casing of the gearbox where the signal is picked up by an accelerometer. As the vibration travels from the point of origin to the location where it is picked up by the accelerometer on the gear casing, it will be affected by vibrations from a multitude of other sources, which can either attenuate or resonate the signal. Therefore the detected signal will initially require processing by using a filter to remove any high frequency noise picked up [86].

Analysis of vibration signals is particularly useful for CM of gearboxes since any change in the vibration signal of a gearbox is due to a change in the condition of the gearbox itself. Any defects on a gear will alter its meshing behaviour and therefore cause a change in both its amplitude and phase modulations in its vibration signal. Therefore the changes in a gearboxes' vibration signal can be analysed to diagnose its current condition.

Generally, using spectral analysis of the vibration signal in its frequency-domain is the most commonly used method used for CM of gearboxes. This is because the most important CM elements in the vibration signal of gears are the tooth meshing frequency, its harmonics and the sidebands which are located on either side of the gear tooth meshing frequency. The sidebands are separated by integer multiples of the gear rotation frequency. The behaviour of these sidebands can be strongly indicative of the presence of a fault, for example, through an increase in the number of sidebands and their relative amplitudes is often due to the meshing of damaged tooth. The information contained in the first three meshing frequencies of the gear and its sidebands can give indicate the condition of the gear [87].

Although theoretically, analysing the changes in the amplitude of particular sidebands in the frequency domain can be used to diagnose the condition of a gearbox. In practice, if only using a simple Fourier transform (FT) to transform the signal from time domain to frequency-domain, it will still be difficult to extract meaningful information from the vibration signal. In the early stages of fault development, important frequency components are of relatively low amplitude and which are masked by other vibrational sources or affected by background noise [88]. This is especially relevant since the individual vibration impulses generated by gear defects typically have short durations and causes the matching frequency response to be spread over a wide frequency band with a small amplitude. It is also complex in identifying the specific component that indicates a fault amongst a multitude of various other frequency components.

The combination of time-frequency analysis is becoming more and more popular for diagnosing gearbox faults, it is gradually replacing conventional time-domain based and frequency-domain based analysis. With the signal analysed in both its time and frequency domains concurrently is a useful tool for examining the non-stationary vibration elements in

the signal, and their source is relatively simple to conclude. Overall it is not very difficult to characterise the local features of the signal, and all of the distinct components around the meshing frequency and its harmonics, also their causality and the phase change can modelled [89].

4.2 Vibration Model of a Gear Pair

Commonly, the vibration system of a gear set can be represented as shown in Figure 4.1 [90]. 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 developed based on the symbols defined in Table 4.1.

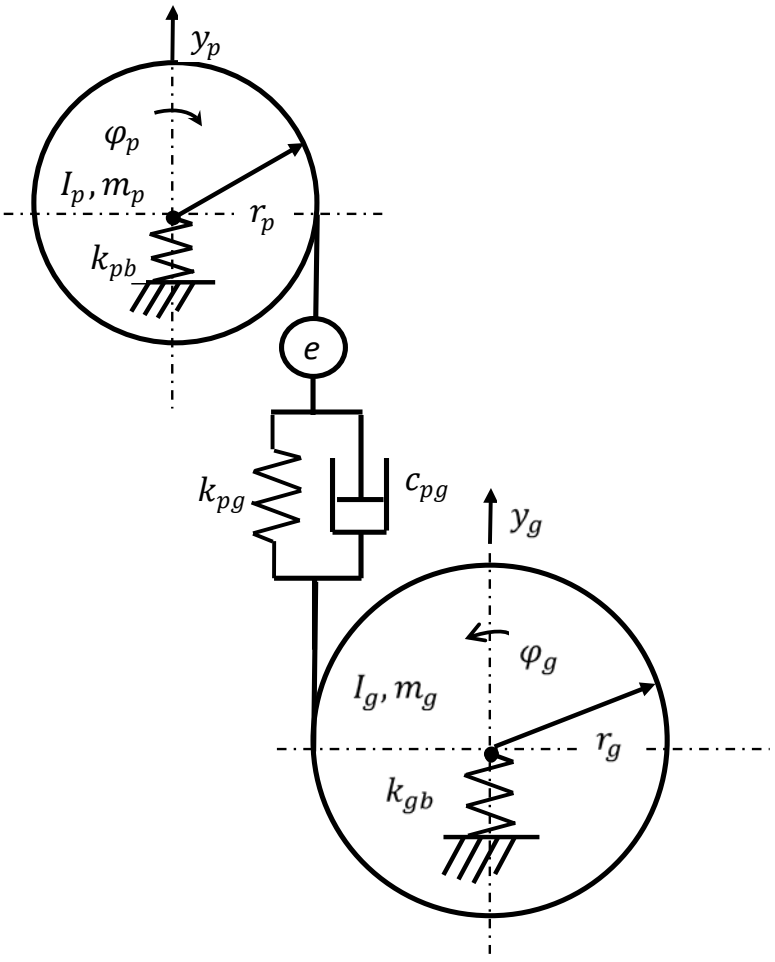

Figure 4.1 Vibration model for a gear pair

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

$$m_p\ddot{y}_p + k_{pb}y_p - k_{pg}(t)\left(r_p\varphi_p - r_g\varphi_g - y_p + y_g + e\right) - c_{pg}(t)\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \tag{4.1}$$

The translational motion of the driven gear g in the first stage is given by:

$$m_g\ddot{y}_g + k_{gb}y_g - k_{pg}(t)\left(r_p\varphi_p - r_g\varphi_g - y_p + y_g - e\right) - c_{pg}(t)\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \tag{4.2}$$

The rotational motion of the drive gear or pinion p in first stage is as:

$$I_p\ddot{\varphi}_p + r_p k_{pg}(t)\left(r_p\varphi_g - r_g\varphi_g - y_p + y_g + e\right) + r_p c_{pg}(t)\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \tag{4.3}$$

The rotational motion of the driven gear g in the first stage is given by:

$$I_g\ddot{\varphi}_g - r_p k_{pg}(t)\left(r_p\varphi_g - r_g\varphi_g - y_p + y_g - e\right) - r_p c_{pg}(t)\left(r_p\dot{\varphi}_p - r_g\dot{\varphi}_g - \dot{y}_p + \dot{y}_g\right) = 0 \tag{4.4}$$

As the tooth stiffness $k_{pg}(t)$ and $c_{pg}(t)$ is of time-varying function, these are nonlinear parametrically excited differential equations and difficult to find closed form solutions. Nevertheless, it shows that translational motions $\ddot{y}_p$ and $\ddot{y}_g$, which can easily transfer to the gearbox case due to high stiffness of bearings and then are measurable by accelerometers

mounted on the case for condition motoring, are coupled with the rotational motions. It means that the measured vibrations are a combined effect of the translational and rotational motions. The translational motion can be the amplitude modulator while the rotational motion can be the angular modulator. Therefore, the vibration signal exhibit combined modulation characteristics.

Furthermore, combination with multiple stage vibrations, inherent manufacturing errors and fault excitations will result in much more complicated modulation. This makes it difficult to have effective and stable vibration features for obtaining a reliable diagnostic result.

Table 4.1 Nomenclature of gear vibration model

I_p	moment of inertia of driving gear one (pinion) at the first stage
I_g	moment of inertia of driven gear in the first stage
φ_p	angular displacement of driving gear;
φ_g	angular displacement of driven gear
$r_{p,}$	base radius of the driving gear in the first stage

r_g	base radius of the driven gear in the first stage
y_p	vertical displacement of the driving gear the first stage
y_g	vertical displacement of the driven gear the first stage
m_p	mass of the driving gear in the first stage
m_g	mass of the driven gear in the first stage
c_{pg}	first stage meshing damping
k_{pg}	first stage time varying meshing stiffness
k_{pb}	driving gear bearing stiffness
k_{gb}	driven gear bearing two stiffness
e	gear errors including different gear fault modes

4.3 Vibration Model of a Multiple Stage Gear Transmission

To model the vibration responses of a multiple stage gearbox, multiple models shown Figure 4.1 can be linked through a shaft with a torsional stiffness k_{12} as shown in Figure 4.2. In this way, the rotational motions are coupled between two stages by revising Equation (4.4) to include a relative torsional motions. For example for a two stage gearbox the rotational motion of the gear at the 1st stage will be revised as:

$$I_{g1}\ddot{\varphi}_{g1} - r_{p1}k_{pg1}(t)\left(r_{p1}\varphi_{g1} - r_{g1}\varphi_{g1} - y_{p1} + y_{g1} - e_1\right) - r_p c_{pg1}(t)\left(r_{p1}\dot{\varphi}_{p1} - r_{g1}\dot{\varphi}_{g1} - \dot{y}_{p1} + \dot{y}_{g1}\right) - k_{12}(\varphi_1 - \varphi_2) = 0 \quad (4.5)$$

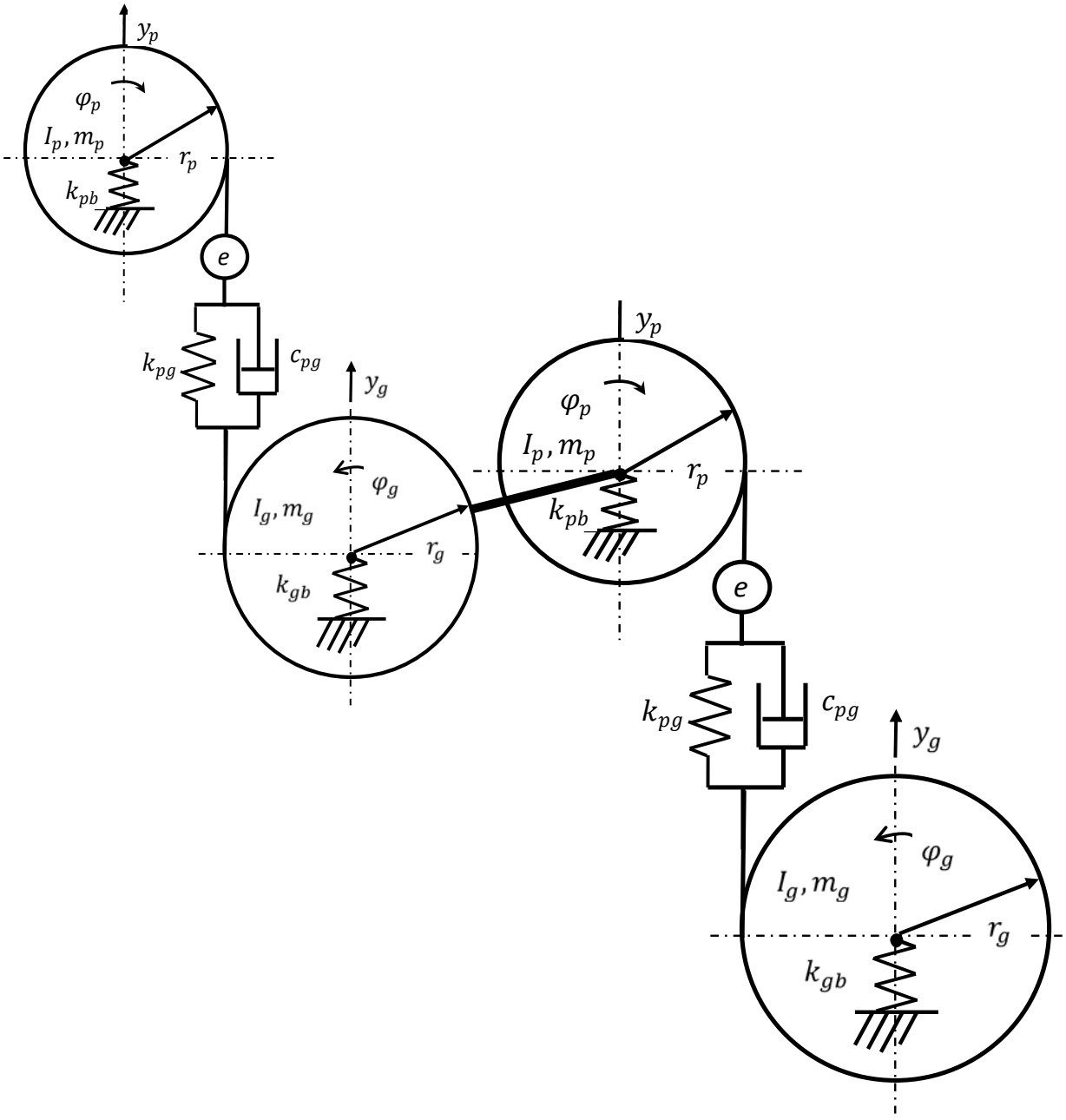

Figure 4.2 Vibration model for two gear pairs in motion

As shown in Equation (4.5), the vibrations between the two stages are coupled together. As a result, the translational vibration at the 1st stage which are measurable on the case will have additional components such as the mesh frequencies of the 2nd stage. Internally, these vibrations will modulate the stiffness functions to results in multiple modulations and

consequently more complicated components and spectrum patterns. This will cause further difficult to make differences between each components for the purpose of fault detection diagnosis.

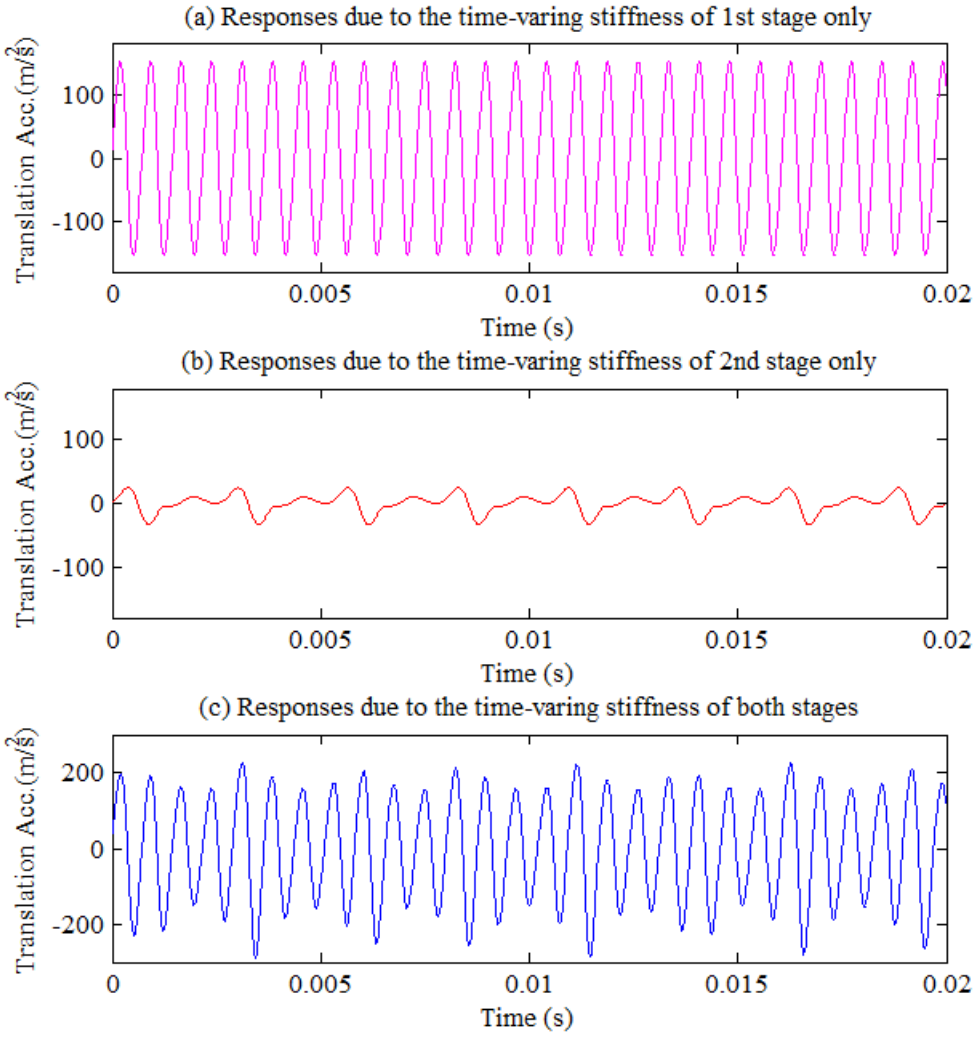

Figure 4.3 Vibration response to sinusoidal stiffness at the 1st and 2nd stages of a two stage gearbox

Figure 4.3 shows the vibration waveforms of translational responses for three cases of excitations in a form of a sinusoidal mesh stiffness k_{pg}. They were obtained numerically based on a two stage gearbox whose parameters are detailed in Table 6.1. The response of

translation vibration to the stuffiness excitation at the 1st stage alone exhibits stationary waveforms corresponding to the faster variation at the first stage, whereas the response of the 2nd stage also exhibit stationary but with a slower variation, which are show in Figure 4.3 (a) and (b) respectively. However, even though the excitation is of a pure sinusoidal input the waveform have multiple harmonics, showing typical nonlinear effect of the vibration system due to multiple couplings between the translational and rotational motions.

This nonlinear effect become more obvious and significant when both stages have the sinusoidal excitations applied simultaneously. As shown in Figure 4.3 (c), the waveforms to such multiple excitations exhibit clear modulation characteristics in that the slower oscillations modulates the faster oscillations or visa verse, making the responses more complicated to be separated appreciably.

To further shows the mutual coupling effect between the vibrations of the two stages, the spectra corresponding to the waveforms are presented in Figure 4.4. As shown in Figure 4.4 (a) and (b), the sinusoidal excitation induces only mesh components and their higher harmonics when the excitation applies to the system individually. However, when the excitations coexist in both stages, the spectrum in Figure 4.4 (c) become much more completed than just a simple summation of the two spectra in Figure 4.4 (a) and (b). Instead, it shows clear modulation characteristics due to the effect that the two sets of harmonic components are modulated each other.

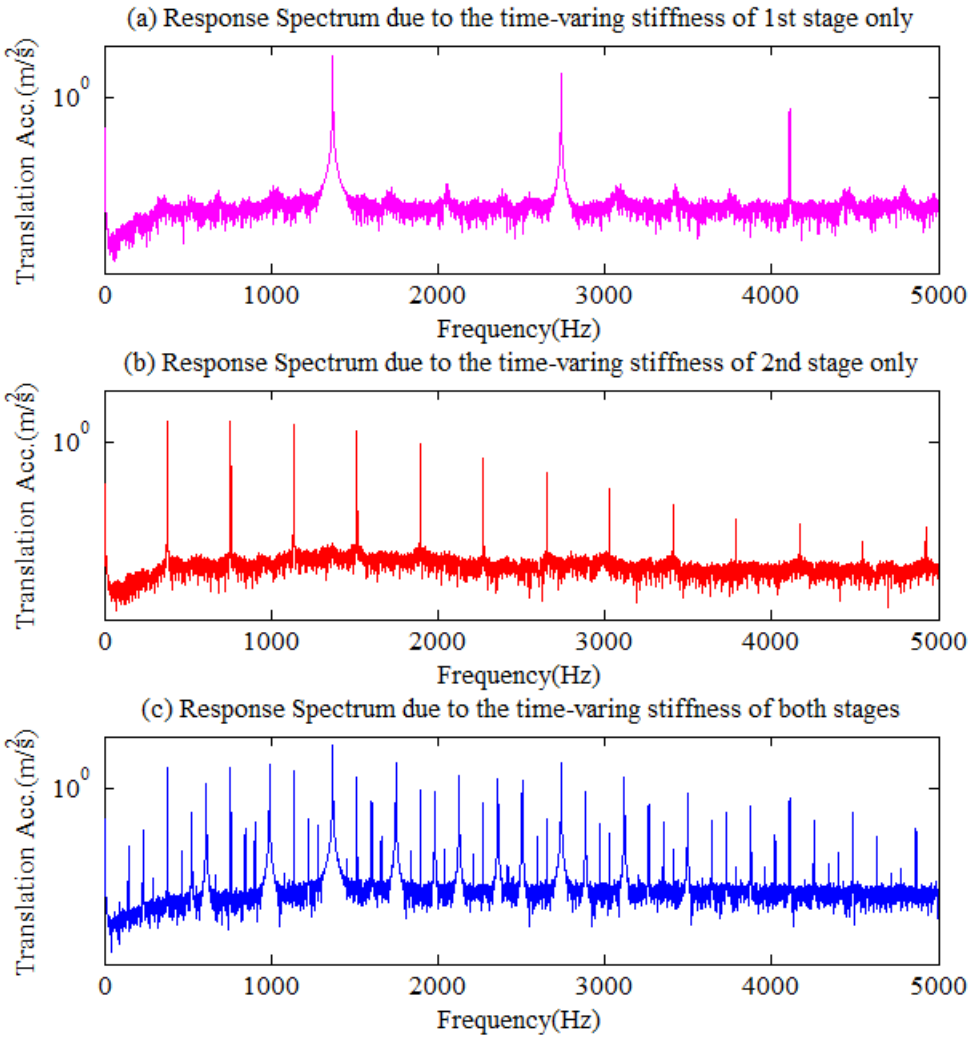

Figure 4.4 Vibration spectra for sinusoidal stiffness at the 1st and 2nd stages of a two stage gearbox

Obviously, because the complexity of multiple components modulation, it is not very easy to separate these components based on this spectrum analysis which contains only amplitude and frequency information. It is expected to have more chance to separate them when also include phase information.

4.4 Common Imperfectness of Gears

Gears always have various manufacturing and installing errors. They can cause significant contribution to overall vibration responses of a perfect gear pair. For brevity, three common errors are modelled and their vibration characteristics are examined.

4.4.1 Transmission Error

One of the most common errors is the run-out transmission error. For simplicity and effectiveness, only one component of the first types of error E_{pi} , which is often more significant due to the effect of gear cutting process, is modelled as:

$$E_{pi} = \gamma_i cos(\varphi_i + \varphi_{i0}) \tag{4.6}$$

where γ_i is the maximum error amplitude, the i=1,2 denoting engaging gears 1 and 2 respectively, φ is the angular displacement and φ_{i0} is the initial phase displacement for the error which will be included to reflect installation angular positions. Thus the overall stiffness of an engaging gear pair will be:

$$k_{z1} = k_{z1}\left[\left(\varphi_1 + \gamma_1 cos(\varphi_1 + \varphi_{10}\right) + \gamma_2 cos(\varphi_2 + \varphi_{20})\right] \tag{4.7}$$

4.4.2 Eccentricity Error

The influence of any eccentricities can also be investigated by incorporating it into the error function (4.7). The eccentricities can be modelled with its first order as in [90]:

$$E_i = \rho_i cos(\varphi_i + \varphi_{0i}) \tag{4.8}$$

where ρ_i is the maximum error amplitude of the eccentricity for the ith gear. Therefore, the stiffness combined with eccentricity error becomes

$$k_{z1} = k_{z1}\left[\left(\varphi_1 + \gamma_1 \cos(\varphi_1 + \varphi_{10}\right) + \gamma_2 \cos(\varphi_2 + \varphi_{20}\right) + \rho_1 \cos(\varphi_1 + \varphi_{01}) + \rho_2 \cos(\varphi_2 + \varphi_{02})\right] \tag{4.9}$$

4.4.3 Unbalanced Mass Error

Unbalanced mass m_i is modelled as a dynamic torque:

$$T_i = m_i e_i \omega^2 cos(\varphi_i + \varphi_{0i}) \tag{4.10}$$

where e is displacement of the mass centre

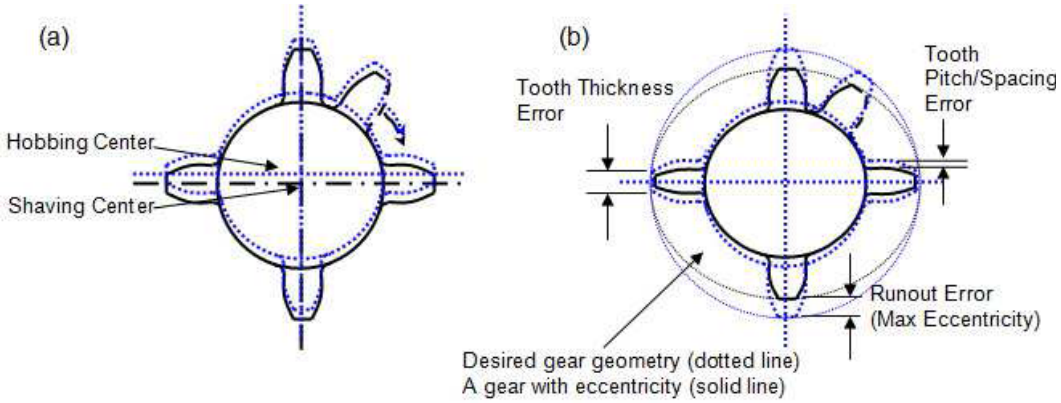

Figure 4.5 Eccentric gear. (a) Eccentricity in a gear caused by manufacturing errors, (b) Resulting errors in gear geometry [90]

In addition, the motor shaft and gearbox input shaft are often connected by a key-sleeve mechanism or coupling. Considering the clearance of the sleeve or coupling, the misalignment between these two shafts can happen, which is modelled by refereeing Figure 4.6 as:

$$r_2^2 = r_1^2 + e^2 - r_1 e cos\left(\theta\right) \tag{4.11}$$

$$\omega_2 = \frac{\omega_1 r_1}{r_2} = \frac{\omega_1 r_1}{\sqrt{r_1^2 + e^2 - r_1 e\cos(\omega_1 t)}} \tag{4.12}$$

For simplification the modelling of this effect is represented by a small speed variation added onto ω_2

$$\Delta\omega_2 = e\cos(\varphi_1 + \varphi_0) \tag{4.13}$$

where $\Delta\omega_2$ is the speed variation with respect to speed of operation.

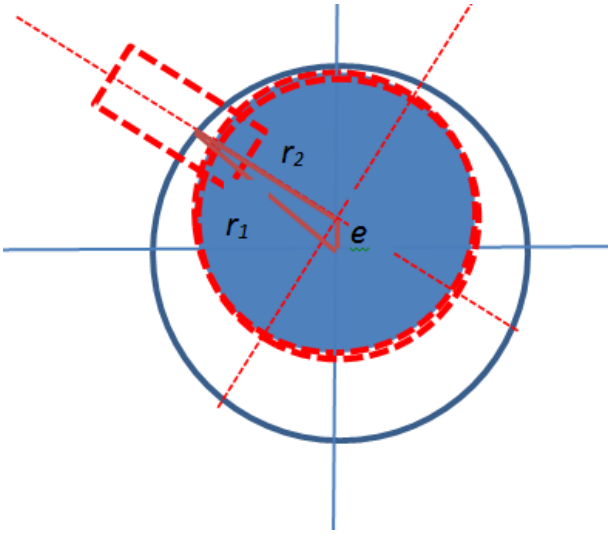

Figure 4.6 Misalignment between 2 shafts connected by key-sleeve

4.5 Model for Gear Tooth Breakage

In a gear pair the system vibration response is due to the deformation and elasticity of the contacting components and not due to the force or the response time history. So a piece-wise linear function was used rather than a rectangular or a half-sine function. This linear function can be easily modified to include different effects including tooth breakage and tooth profile wear.

To simplify the model, just the major factors, the mesh stiffness, $k(t)$ $k(t)$ has been considered in this study The mesh stiffness is defined to be the ratio between the force acting along the line of action, and the tooth displacement along the same line. In order to obtain a mathematical model, the stiffness for a single pair of meshing teeth was first investigated. The stiffness for the meshing of a gear pair is affected by several characteristics [91], firstly the stiffness due to the Hertzian contact, the axial stiffness between both teeth and the bending stiffness of the two teeth from the drive Z_1 and the driven gear Z_2. All these factors can lead to an increase in the stiffness amplitude in the middle phase of the meshing time.

Due to the complexity in precisely predicting the actual pattern of the variation in stiffness in real cases, approximations are usually used. Therefore, this research study presumes that the tooth mesh stiffness has the combined profile shown in Figure 4.5 (a).

As mentioned above, a piece-wise linear function is used to represent the transient value of the stiffness during the gear mesh-in and gear mesh-out of a gear pair. The stiffness is assumed to have a constant amplitude variation in the middle part of the engagement of the tooth pair. Maximum stiffness value of a single gear pair is usually in the order of 10^7 N/m [92].

In a single pair of teeth contact, the meshing stiffness stays mostly unchanged, since the reduction of stiffness in one of tooth in the pair is then compensated by an increase in stiffness of the other mating tooth. When the gears rotate the point of contact between a tooth pair will be higher in one tooth and lower in the other tooth of the matching pair.

While the gears are rotating, the number of teeth in contact also varies, as a consequence the effective length of the contact line alters creating variations in overall mesh stiffness. For

spur gears, where the contact ratio is lower, such variations are mainly because of a load transfer occurring over a single tooth in one double tooth pairing. However, in the case of a helical gear pairs, the meshing change of the overall length in the line of contact line is relatively much smaller since there is a higher contact ratio which is distributed between three or four teeth at a time. Therefore, the mesh overall meshing stiffness variations in helical gears are considerably smaller compared against that of spur gears. Thus models of helical gear dynamics consider meshing stiffness as a constant value that uses its time averaged result [93].

In order to simplify this research model, the meshing stiffness is assumed as a function of angular displacement to depend upon two factors: overlap ε_α and contact ratio ε_c. The first parameter measures the overlap of adjacent teeth in the axial or face-width direction and the other parameter measures the contact between adjacent teeth in a diagonal section.

The overall profile of a single pair stiffness can be interpreted as linear piecewise, which is shown in Figure 4.5 (a).

Figure 4.5(b) shows the x-axis denoting the number of base pitches, which needs be multiplied by ($2\pi/Z_1$) to provide the angular variation. The total meshing stiffness can then be calculated by summing the overall stiffness of a single tooth pair, and incrementing by one tooth at a time for the total number of Z_1 teeth.

Further, in Figure 4.7 (b) illustrates the resulting meshing stiffness waveform for the underlying gear pair of this research study. The minimum number of teeth in contact is three and the maximum number is four. This variation between the minimum and maximum number of teeth in contact is the main source of the gear vibration. The period of the

fluctuation shows that even in a healthy condition (given that there are manufacturing faults or faults) a gear pair will still produce vibrations. With the existence of manufacturing faults, the variations in the stiffness will be non-uniform. The minimum and maximum number of contacting teeth pairs will vary depending on the shaft angular position, as does the resulting error's amplitude, typically causing greater variations in the stiffness amplitude and resulting in vibrations with higher amplitudes.

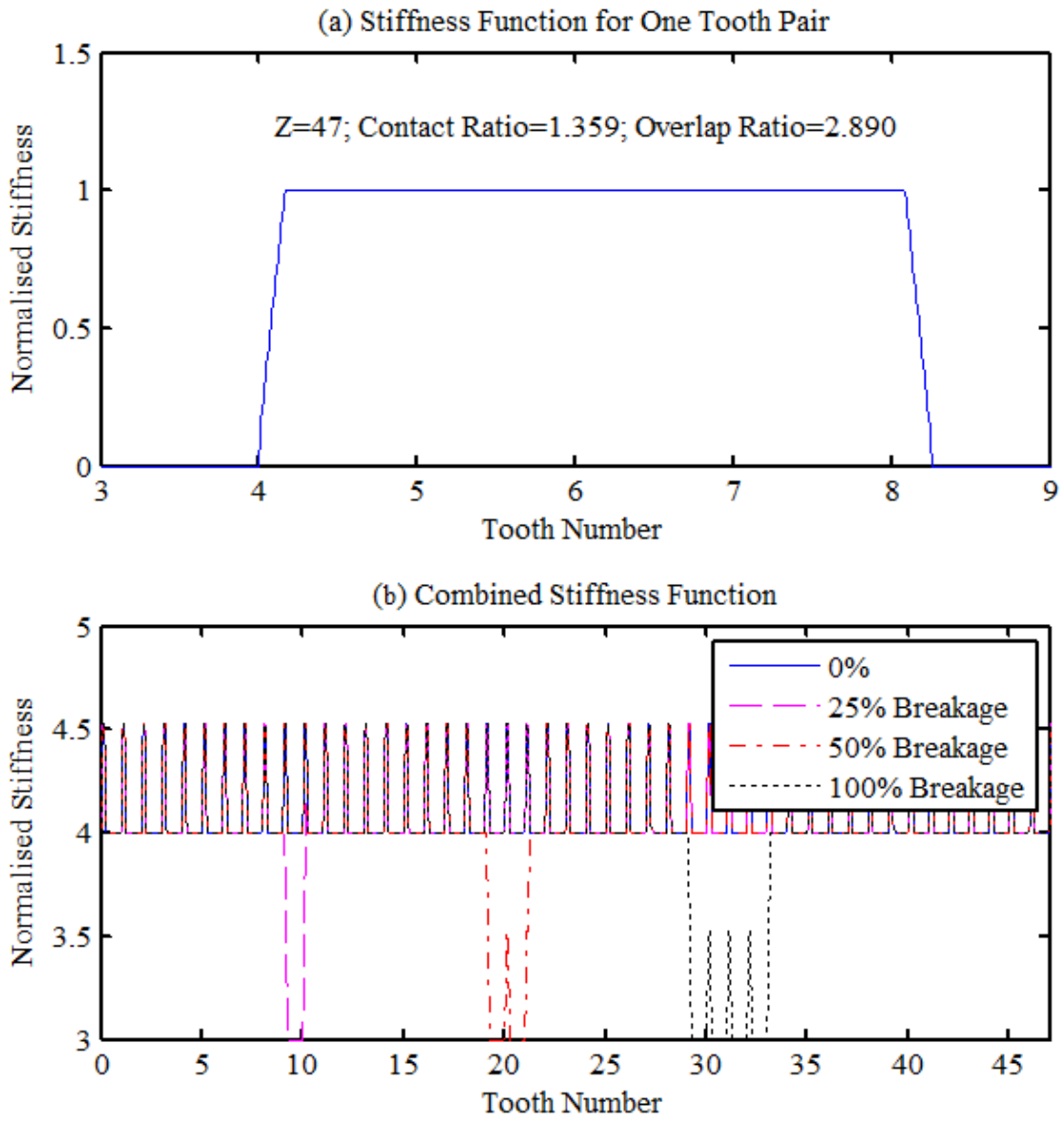

Figure 4.7 Schematic of tooth stiffness variation under different fault cases

It is well known and mentioned in chapter three that tooth breakage (failure) is a common fault in helical gears. This kind of fault begins with small damages to the structure of the tooth, which then leads to a change in its meshing stiffness. The colours in Figure 4.7 (b), blue, pink, red and black, represents the influence of a single tooth breakage on the overall meshing stiffness. The tooth breakage are missing 25%, 50%, 75% and 100% of a complete single tooth. The figure shows that even local faults can result in a sudden drop in the stiffness amplitude. Additionally, at lower stiffness values, the width of the interval increases along with the severity of the fault, however the low amplitude value remains the unchanged. The sudden drops in the meshing stiffness value should then create an equivalent rise in the overall amplitude of the system vibration.

4.6 Model for Tooth Wear

4.6.1 Tooth Surface Material Loss

Due to poor lubrication and high stresses, gear tooth profiles are subject to slow wear during their lifetime in which asperities of contacting tooth surfaces are gradually worn out until very fine, smooth, conforming surfaces develop. This light wear can occur by either abrasive or adhesive mechanisms when thin oil films or what is known as boundary lubrication conditions prevail usually on low speed operations. This means the loss of material from the contact surface of a gear and reduction of gear tooth thickness which subsequently alter stiffness, and damping and friction characteristics, leading to higher levels of vibration.

The tooth wear of gears, combined with the manufacturing errors, are amongst the cause of a change in the profile outline of the teeth, and is the main factor in gear failure, leading to the change in the instant gear ratio during the range of the angular motion [94]. In particular, the change in the gear ratio caused by the changing stiffness and the eventual damage of the

tooth causes an increase in the dynamic loading, even if the teeth is undamaged. Moreover, dynamic loading is at its maximum at the start and at the end of the engagement in the pairing.

Surface wear of the teeth is one of the major causes of failure in gearbox transmission systems [95]. Slow wearing of tooth profiles can result in a unique surface material loss [96], which consequently alters the gear mesh excitations errors in its kinematic motion, further enhancing the undesirable dynamic effects.

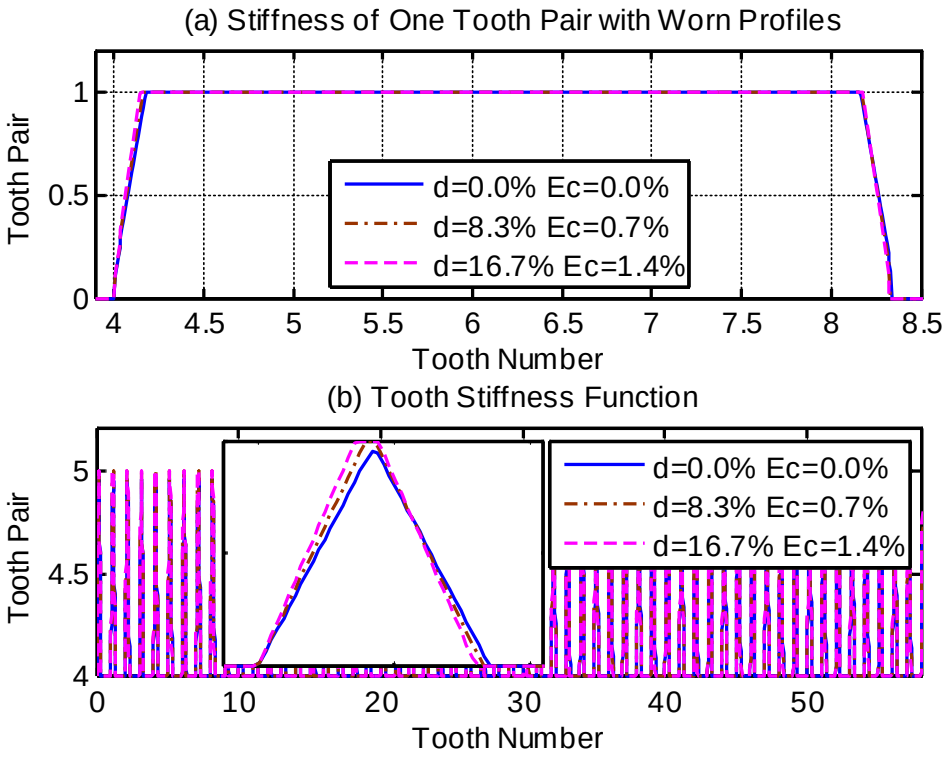

Figure 4.8 (a) Stiffness of a single pair of teeth and (b) The meshing stiffness function k1 and k2 for the pinion and the gear of the first pair of gears (within one shaft period) shown in different colours as a function of angular displacement θ, with severity of faults due to a broken tooth blue (healthy gear) brown (25%) pink (50%), red (75%), and black (100%)

In order to understand the effect of tooth wear on the vibration response, the meshing stiffness function is combined with a reduced width of the stiffness function for a single tooth pair. As shown in the Figure 4.8(a), the stiffness function is modified to be sharper at the meshing-in and meshing-out stages, to simulate the observation that the maximum dynamic

loading takes place in the beginning and at the end of the engagement. In addition the overall duration of the tooth stiffness function is also reduced by reducing the contact ratio to take into account the reduction of tooth thickness, as shown by the magnified tooth excitation profile in Figure 4.8(b).

In the simulation, two different degrees of tooth wear are simulated at the first stage. Assuming that the wear occurs on both the drive gear and the driven gear and that the effect of wear on tooth bending deformation is negligible, the tooth stiffness thus is affected only if the tooth dimension changes due to wear, which leads to a change of stiffness profile. To model this effect and obtain the corresponding dynamic responses, the simulation is implemented by increasing the sharpness and decreasing the contact ratio simultaneously. The amplitude for the sharpness increase is 8% and 17% and the reduction of contact ratio is 7% and 14% for the smaller wear and larger wear respectively.

4.6.2 Error Enlargement

Moreover, the effect of manufacturing errors influences the characteristics of the progression of tooth wear. Also, the errors of the instantaneous contact forces are uniform over a revolution of the gear. Obviously teeth under a greater contact force due to misalignment or gear eccentricity will produce more severe wear. To include this characteristic into the model, a small increment is added to the error amplitude as

$$E_w = (\rho_1 + \rho_{w1}) sin(\varphi_1) + (\rho_2 + \rho_{w2}) sin(\varphi_2) \tag{4.14}$$

where ρ_{w1} is the wear depth of driving gear (Z_1) and ρ_{w2} is the wear depth of driven gear (Z_2).

4.7 Signal Model for the Vibration of Perfect Gearboxes

In a perfect operating scenario, the vibration signals created by a gearbox can be mathematically represented by simply by summing the vibration that is generated by its each individual component. Generally, if the gear is in a good working condition, the vibration signal will be dominated when the teeth of the gears come in to contact with one another, the vibration excitation forces are a direct result of the meshing forces between the teeth of the gears. Since the gears behave in a rotatory manner, the vibration signal's components are periodic. Therefore, the overall vibration signal can be simulated by using a mathematical equation which sums together each of the individual components.

Firstly the fundamental meshing frequency, f_{me} for a particular gear is calculated by using the following equation:

$$f_{me} = N.f_r \tag{4.15}$$

Where N represents the number of teeth on the gear and f_r represents rotational frequency of the gear. Since the overall vibration signal consists of a meshing frequency for each set of gears, the source of the meshing frequency is easily distinguishable.

The meshing vibration signal, $x(t)$ is comprised of M harmonic components, it can be expressed as:

$$x(t) = \sum_{m=0}^{M} I_m \cos(2\pi f_{me} t + \phi_m) \tag{4.16}$$

where I_m represents the amplitude of the m^{th} harmonic, f_{me} is the meshing frequency and ϕ_m is the phase angle.

4.8 Signal Model for the Vibration of Faulty Gearboxes

Broken teeth is a very common cause failure in helical gears. The gear fault starts off appearing as there is relatively little damage to the particular tooth, however this creates a difference in the meshing stiffness that will further magnify the damage if unresolved. A local fault will cause a sudden drop in its stiffness amplitude

If there is a fault or error in the gear, there will be amplitude and phase as a consequence of the damaged tooth meshing.

$$y(t) = \sum_{m=0}^{M} I_m[1 + a_m(t)]\cos[2\pi f_{me} t + \phi_m + b_m(t)] \qquad (4.17)$$

where $a_m(t)$ is the amplitude modulation and $b_m(t)$ is the phase modulation both of which is due to the local tooth fault. This is a very generic signal model proposed in [84] and has been referred widely by many researchers such as the latest study of exploring conventional bispectrum for the analysis of gear fault vibrations [99]. 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. Specifically, the rotational effect may be represented more by the PM terms and the translational effect is represented more by the AM terms.

In order to understand the mechanism for sideband interactions, the linear solutions of either Equation (4.1) or (4.2) is investigated by considering only the 1st order term in each of three components AM, PM and the carrier. This simplification avoids the complicity of mathematic operations but allow gaining an in-depth understanding of the interactions between AM and PM without the loss of generality represented in Equation (4.17).

With simplification Equation (4.17) is approximated as

$$x(t) = I_m[1 + \mathrm{a}_m(t)]\cos[2\pi f_c t + b_m(t)] \tag{4.18}$$

with a time increment phase that consists of a linear term and a oscillatory term to the angular changes of the mesh process as:

$$\theta_m(t) = 2\pi f_{me}(t)t = 2\pi(f_{me} + b_m(t))t \tag{4.19}$$

in which assuming zeros phase relative to the carrier waves, AM signals $\mathrm{a}_m(\mathrm{t})$ is :

$$a_m(t) = A\cos[2\pi f_F t] \tag{4.20}$$

where f_F is the frequency of the fault. By introducing a phase modulation index or shaft fluctuation rate β :

$$\beta = \frac{\Delta f_F}{f_F} << 1 \tag{4.21}$$

to express small frequency deviation Δf_F in a pseudo steady operation of a gearbox associated rotation system that always has speed fluctuations. Thus the phase modulation signal $b_m(t)$ is defined for the 1st order as:

$$\mathrm{b}_m(t) = \beta\cos[2\pi f_F t - \varphi] \tag{4.22}$$

which includes an angle φ to represent the phase difference relative to the meshing components or the carrier, allowing the expression of possible delays or shifts between the translational and rotational excitations as a consequence of various effects including gear errors, fault cases and nonlinear vibration transfer paths. These expresses yields

$$x(t) = I_m(1 + \mathrm{a}_m(\mathrm{t}))\cos[2\pi f_c t + b_m(\mathrm{t})] \tag{4.23}$$

$$
\begin{aligned}
y(t) &= I_m(1+a_m(t))\cos[2\pi f_c t + b_m(t)] \\
&\approx I_m(1+a_m(t))[\cos(2\pi f_c t) - b_m(t)\sin(2\pi f_c t)] \\
&\approx I_m\{\cos(2\pi f_c t) + a_m(t)\cos(2\pi f_c t) - b_m(t)\sin(2\pi f_c t) - a_m(t)b_m(t)\sin(2\pi f_c t)\}
\end{aligned}
\tag{4.24}
$$

In which the approximation is made as $\beta = \Delta f_F / f_F << 1$, $|b_m(t)| << 1$, Therefore, $\cos(b_m(t)) \approx 1$ and $\sin(b_m(t)) \approx b_m(t)$. This expansion of Equation (4.17) allows the examination not only magnitude characteristics of modulations using the expression of Bessel function [99] but also it can gives details of phase influences that result in interferences between AM and PM components.

In addition, $|a_m(t)b_m(t)| << 1$. This approximation leads to more simplified expression:

$$
y(t) \approx I_m\{\cos(2\pi f_c t) + a_m(t)\cos(2\pi f_c t) - b_m(t)\sin(2\pi f_c t)\} \tag{4.25}
$$

substituting $a_m(t)$ and $b_m(t)$ into (4.24) yields the expression of the vibration response in the format of sideband components as:

$$
\begin{aligned}
y(t) &\approx I_m\{\cos(2\pi f_c t) + A\cos(2\pi f_F t)\cos(2\pi f_c t) - \beta\cos(2\pi f_F t - \varphi)\sin(2\pi f_c t)\} \\
&= I_m\{\cos(2\pi f_c t) + A\cos(2\pi f_F t)\cos(2\pi f_c t) - \beta\cos(2\pi f_F t - \varphi)\sin(2\pi f_c t)\} \\
&= I_m\{\cos(2\pi f_c t) \\
&\quad + \frac{A}{2}[\cos(2\pi(f_c - f_F)t) + \cos(2\pi(f_c + f_F)t)] \\
&\quad - \frac{\beta}{2}[\sin(2\pi(f_c - f_F)t + \varphi) + \sin(2\pi(f_c + f_F)t - \varphi)\}
\end{aligned}
\tag{4.26}
$$

combining the sine and cosine terms in (4.26) to obtain sideband amplitudes:

$$
\begin{aligned}
y(t) &= I_m\{\cos(2\pi f_c t) \\
&\quad + \frac{A}{2}[\cos(2\pi(f_c - f_F)t) + \cos(2\pi(f_c + f_F)t)] \\
&\quad - \frac{\beta}{2}[\sin(2\pi(f_c - f_F)t)\cos(\varphi) + \cos(2\pi(f_c - f_F)t)\sin(\varphi) \\
&\quad + \sin(2\pi(f_c + f_F)t)\cos(\varphi) - \cos(2\pi(f_c + f_F)t)\sin(\varphi)]\} \\
&= I_m\{\cos(2\pi f_c t) \\
&\quad + \frac{A - \beta\sin(\varphi)}{2}\cos(2\pi(f_c - f_F)t) - \frac{\beta\cos(\varphi)}{2}\sin(2\pi(f_c - f_F)t) \\
&\quad + \frac{A + \beta\sin(\varphi)}{2}\cos(2\pi(f_c + f_F)t) - \frac{\beta\cos(\varphi)}{2}\sin(2\pi(f_c + f_F)t)\}; \\
&= I_m\{\cos(2\pi f_c t) + A_L\cos(2\pi(f_c - f_F)t - \delta_L) + A_H\cos(2\pi(f_c + f_F)t - \delta_H)\}
\end{aligned}
\tag{4.27}
$$

where the magnitudes of the lower and higher sidebands are

$$
\begin{aligned}
\mathrm{A}_L &= \sqrt{\left(\frac{A - \beta\sin(\varphi)}{2}\right)^2 + \left(\frac{\beta\cos(\varphi)}{2}\right)^2} \\
&= \frac{1}{2}\sqrt{A^2 + \beta^2 - 2A\beta\sin(\varphi)}
\end{aligned}
\tag{4.28}
$$

$$
\begin{aligned}
\mathrm{A}_H &= \sqrt{\left(\frac{A + \beta\sin(\varphi)}{2}\right)^2 + \left(\frac{\beta\cos(\varphi)}{2}\right)^2} \\
&= \frac{1}{2}\sqrt{A^2 + \beta^2 + 2A\beta\sin(\varphi)}
\end{aligned}
\tag{4.29}
$$

And the associated phase angles are

$$
\tan(\delta_L) = \frac{-\beta\cos(\varphi)}{A - \beta\sin(\varphi)} \tag{4.30}
$$

and

$$
\tan(\delta_H) = \frac{-\beta\cos(\varphi)}{A + \beta\sin(\varphi)} \tag{4.31}
$$

It shows that the magnitudes of the lower sideband and the higher sideband are different in amplitudes, instead, they can change with not only the modulation levels but also the phase difference between AM and PM. In addition, the phase also exhibits the same characteristics as the magnitudes[97][98]. Consequently, diagnostic results using sideband magnitudes directly will be unstable or unreliable if the phase difference is not taken into account adequately.

4.9 The Estimation of Modulation Magnitudes

For diagnosing the severity of a fault, it is obvious to include the entire amplitude of the modulation from both AM and PM. As shown in Equations (4.27) and (4. 28) the maximum value of sideband amplitude can be the combination of the sum of powered amplitudes, which ignores the effect of phase interference. Therefore, a modulation power index (MPI) is proposed, the MPI is defined as:

$$Z^2 = A^2 + \beta^2 \tag{4.32}$$

to represent the full effect of both PM and AM. So it can be more reliable comprehensive for fault severity assessment by using this quantity. Alternatively, it means that the amplitude reduction due to phase influences should be removed in quantifying the combined modulation for gear diagnosis.

By power spectrum analysis, the amplitude of sidebands can be estimated to be $\hat{\mathrm{A}}_L^2$ and $\hat{\mathrm{A}}_H^2$ which appear at the lower and higher sideband frequencies respectively. According to Equations (4.28) and (4.29), it can be known that:

$$4\hat{\mathrm{A}}_L^2 = Z^2 - 2A\beta\sin(\varphi) \tag{4.33}$$

$$4\hat{\mathrm{A}}_H^2 = Z^2 + 2A\beta\sin(\varphi) \tag{4.34}$$

By adding these two equations, it is straightforward to yield the MPI from power spectrum as

$$Z^2 = 2(\hat{A}_L^2 + \hat{A}_H^2) \tag{4.35}$$

This means that using conventional spectrum the AM-PM effect can be quantified using (4.35). Unfortunately, many developed fault diagnostic features have not adopted this relationship appropriately, leading to a considerable number of parameters which needs dimensionality reduction techniques in order to implement automated gear fault diagnosis [99].

Moreover, measured signals are inevitably contaminated by various noises such as stationary white noise and random aperiodic impulses[100]. Power spectrum analysis has no means to remove such noises. Therefore, power spectrum based estimation often results in unsatisfactory results for small signals due to incipient faults.

4.10 Chapter Summary

In this chapter gear vibration responses have been examined through developing models that takes into account varying tooth stiffness, various errors and fault failures. It is concluded in theory that gear vibration signals consists of high frequency meshing components and its higher order harmonics AM, PM and low frequency harmonics that induces the AM and PM components.

To separate and quantify these components for the purpose of diagnostics, the overall power of PM and AM, defined as modulation power index (MPI) is necessary. Power spectrum analysis can provide information for calculate MPI but it will result in low accuracy because power spectrum does not include signal phase effect to suppress the influences of various noise including different gear errors.

Therefore, it is essential to suppress these noises for achieving accurate diagnosis of gear faults.

CHAPTER 5.

MODULATION SIGNAL BISPECTRUM BASED GEAR FAULT DETECTION AND DIAGNOSIS

5.1 Introduction

To suppress noise and interferences, TSA analysis has been long regarded as one of the most effective approaches in CM fields. However, it has not been found that TSA is examined rigorously to characterise AM-PM effect under high noise conditions. In addition, the TSA analysis implemented based on meshing phase calculation can have the shortage of sensitive to noise. Moreover, TSA results discard the useful information for diagnosing gearbox bearing problems as bearing characteristic frequencies are usually not integer times of shaft frequency.

To obtain accurate diagnosis results this chapter examines conventional bispectrum that has the high performance of noise reduction and nonlinearity characterisation as overviewed during literature review. However, it has found that it is unable to effectively characterise the modulations in a low frequency range. In other words, it needs a frequency bands as wide as $> 2 f_{me} + 5 f_r$ for getting diagnostic information. This wide band is not very appropriated as wide frequency spans mean more interferences and noise from vibration transmission path.

Therefore, the modulation bispectrum is examined more extensively as it has been shown to have the same performances as CB for noise suppression but needs narrower bandwidth to analyse the vibrations with modulation nonlinearity. In addition, it allows a more sparse representation of the AM-PM signals.

The chapter starts with an outline of power spectrum, CB and MSB, then a new signal processing technique based on MSB is developed for more accurate modulation power estimation. Finally, it verifies the new estimator by a comprehensive simulation study under different noise scenarios.

5.2 Conventional Bispectrum Based Diagnosis

For a discrete time gear vibration signal $x(t)$ its Discrete Fourier transform (DFT) $X(f)$ is in the form of:

$$X(f) = \sum_{t=-\infty}^{\infty} x(t)e^{-2j\pi t} \tag{5.1}$$

which usually can be obtained efficiently by the well-known algorithm – fast Fourier transform (FFT). With the DFT results, the second-order measure of power spectrum (PS) of $x(t)$ can be estimated by the formula:

$$P(f) = E\left\langle X(f)X^{*}(f)\right\rangle \tag{5.2}$$

where $X^{*}(f)$ is the complex conjugate of $X(f)$ and $E\langle\ \rangle$ is the statistical expectation. The power spectrum is the popular method for vibration signal analysis in condition monitoring because it can be particularly effective in characterising the inherent periodicity in vibration signals from rotation machines and produce a quick result of different frequency components.

However, it contains only magnitude information of individual component f and ignores the effects of signal phases, leading to random noise inclusion and consequently unsatisfactory diagnosis performance.

Extending this definition to the measures of order three gives rise to the estimation of conventional bispectrum (CB) [101][102]:

$$B(f_1, f_2) = E\left\langle X(f_1)X(f_2)X^*(f_1 + f_2)\right\rangle \tag{5.3}$$

where f_1, f_2 and $f_1 + f_2$ .are three individual frequency components in a vibration signal. Note that, unlike the power spectrum which is the second-order measures of vibration statistics, this third-order measurement is a complex quantity which contains both magnitude and phase information regarding the original time domain based signal $x(t)$. If the three spectral components f_1, f_2 and $f_1 + f_2$ are non-linearly coupled with one another, the overall phases of the three components no longer remain random. However, the individual phases of each component remains random, in particular, the relationship between phases for three components are shown in the following formula:

$$\varphi(f_2) + \varphi(f_1) = \varphi(f_2 + f_1) \tag{5.4}$$

As a result, statistical averaging will result in a non-zero value in the overall, bispectrum but a distinctive magnitude. A peak in the bispectrum at the bifrequency pair indicates the nonlinear coupling (f_1, f_2).

In contrast, if the frequency components f_1, f_2 and $f_1 + f_2$ are all independent components, each frequency will be characterised by the statistically independent random phases distributed between –π and π. Upon statistical averaging denoted by the expectation operator $E\langle\ \rangle$ in Equation. (5.3), their bispectrum magnitude $|B(f_1, f_2)|$ will tend towards zero due to the random phase mixing effect. In this way, random noise or vibration contents can be suppressed effectively.

To measure the degree of coupling between coupled components, a normalised form of the bispectrum or coherence is usually, which is defined as

$$b^2(f_1, f_2) = \frac{|B(f_1, f_2)|^2}{E\left\langle |X(f_1)X(f_2)|^2 \right\rangle E\left\langle |X(f_1 + f_2)|^2 \right\rangle} \tag{5.5}$$

The bicoherence is independent of the amplitude of the triple product of the DFT amplitudes and its values are bounded between 0 and 1. The bicoherence is close to 1 if there are nonlinear interactions among frequency combinations f_1, f_2 and $f_1 + f_2$. Conversely, a value of near 0 implies an absence of interactions between the components or noise content only. The possible amplitudes in the latter case may suggest that the components are originated independently from a system or resulted from random noise. Therefore, based on the amplitude of bicoherence the nonlinear interactions can be detected and the interaction degrees can be also measured between the coupling components, giving a confident quantitative assessment on the component of interest.

It has been shown recently in [103] that vibration signals for tooth breakages can be analysed by using CB for diagnostics. Especially, it has been shown that bispectrum peaks will exist at bifrequencies $(\mathrm{m}\, f_{me}, \mathrm{n}\, f_m)$. However, the diagnostic results were achieved through the components which have phase value with zero or π in order, being attempt to meet the condition when CB is originally applied to detect the effect of quadratic phase coupling (QPC). However, this outlook the fact that the combined AM-PM modulation of gear vibration often does not follow the phase relations for pure AM or PM modulation process. The restriction of using zero or π may be not an adequate choice as it can discard majority useful information that comes from the common situation of the combination between AM and PM. As shown in Equation (5.4), CB magnitude peaks can appear as long as the biphase $\varphi(f_2) + \varphi(f_1) = \varphi(f_2 + f_1) = \text{constant}$. This means that the combined modulation can be

indicated by CB as the coupled signals keep constant phase values across different measurement instants or measurement times, which is opposite to the fact that the random content of a signal will have a nil CM magnitude as its phase changes with different measurement instants. Therefore, there are many CB peaks that could be used for gear fault diagnosis rather just than that with phase values close 0 or π.

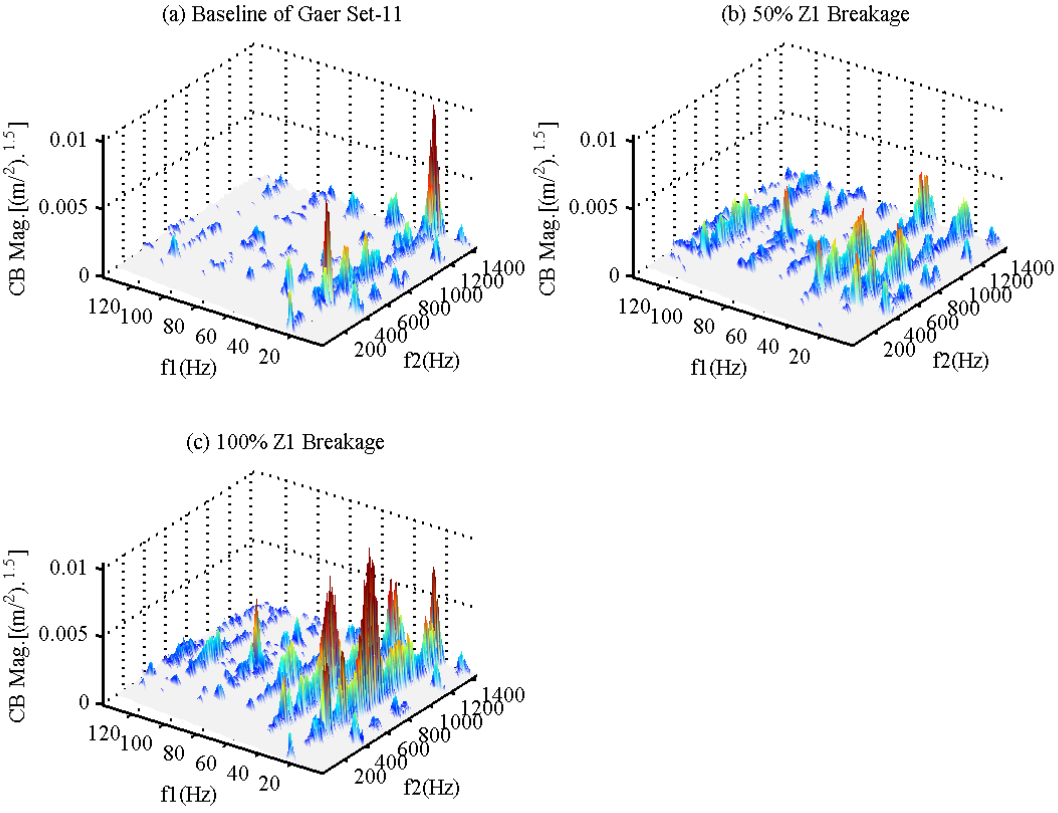

Figure 5.1 CB magnitudes for tooth breakage diagnosis.

Figure 5.1 presents a typical analysis results of CB for different degrees of tooth breakage in which a meshing frequency is 1380Hz and the shaft frequency is at $f_1 = 23.8$Hz. It can be seen that distinctive CB magnitude peaks appears across a wide bifrequency range, especially f_2 being as wide as that from $f_2 = 400Hz$ to $f_2 = 1400Hz$. Moreover, The CB magnitudes in several sub-frequency frequency bands can be based on for fault diagnosis. At bifrequencies

such as (23.8, 600-800), (2×23.8, 600-800) etc, which are all lower than the meshing frequency, and mainly of the terms of shaft frequencies and its higher order harmonics i.e. for the signal model case when meshing frequency $f_{em} = 0$, i.e.

$$x(t) = \sum_{m=0}^{M} I_m[1 + a_m(\mathrm{t})]\cos[\phi_m + b_m(\mathrm{t})] \tag{5.6}$$

The clear increase in amplitudes with fault severity can be used to quantify the severity of the fault and 23.8Hz and its harmonics shows that the fault is from the gear rotating at a shaft frequency of 23.8Hz. This correct diagnosis shows the CB can provide reliable and rich information for gear conditions.

Moreover the significance of the CB peaks is fully supported by CB coherence results. As shown in Figure 5.2, high coherence values appearing at the bifrequency bands correspond to the high peaks in CB magnitude results of Figure 5.1. It shows that vibration components are coupling each other with a wide frequency range. This coupling effect enhances the definitive components such as those relating to the faults while the random noise contents are suppressed significantly. This effect is a very outstanding property compared with power spectrum which is unable to suppress noise. Therefore, the CB results are more reliable.

In addition, the number of coherence peaks is also increasing more with the increase of fault severity. This means that that CB coherence can be also employed for fault severity diagnosis.

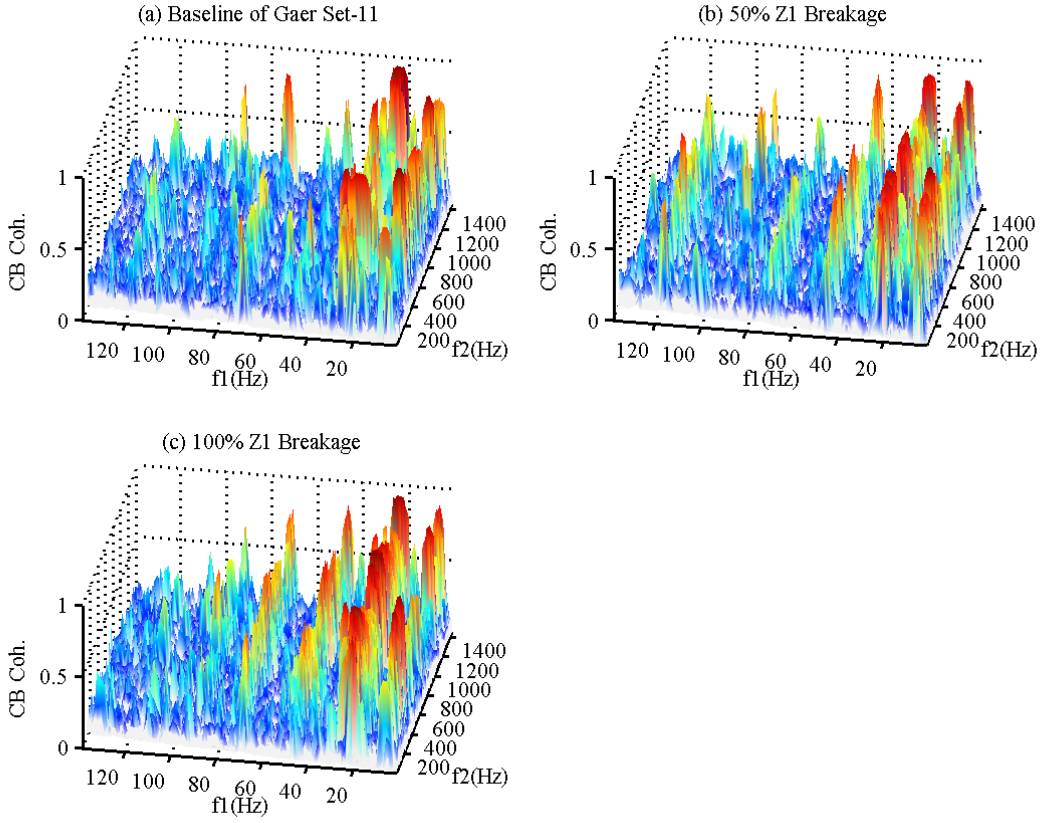

Figure 5.2 CB coherence magnitudes for tooth breakage diagnosis.

Figure 5.3 shows the CB phase results corresponding to the three cases. For highlighting the important components, the results are presented for the CB peaks that have coherence values higher than 0.8. This range can reflect the full changes of CB with respective to the fault cases. However, the phase values are spreading between 0 and π in a nearly random matter. This shows that phase may be used for location diagnosis but not for severity. Moreover, it shows that phase values other than 0 and π also contains the useful detection information and the ignorance of phase related component would lead to lower detection performance.

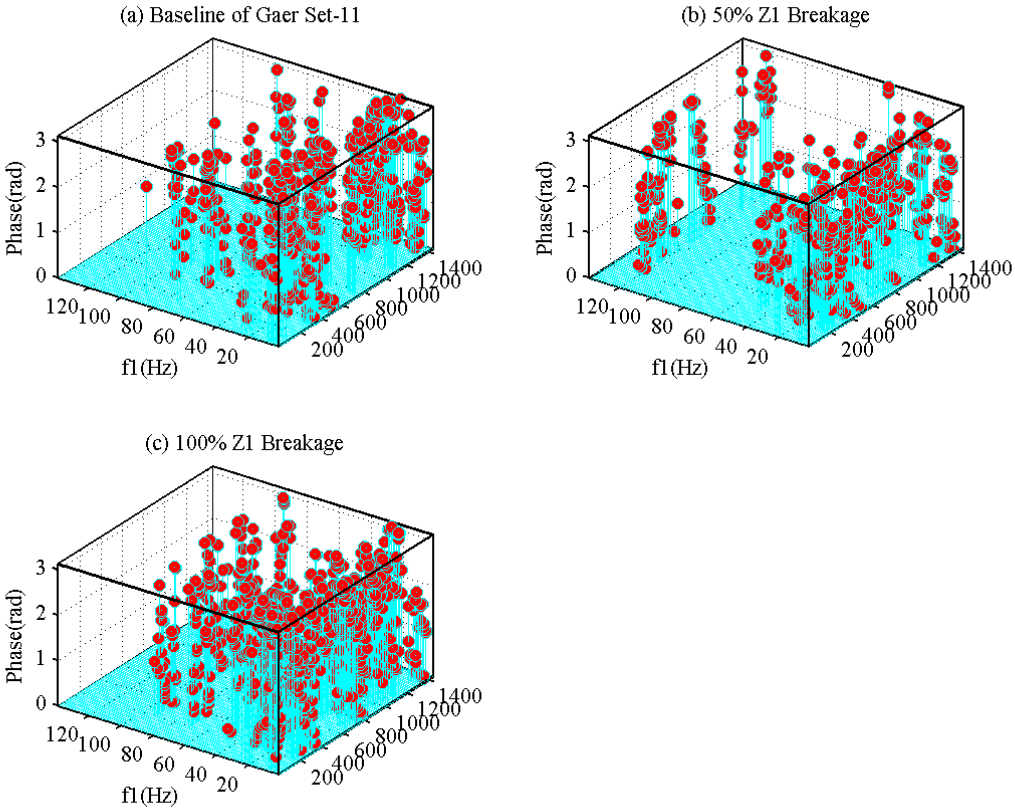

Figure 5.3 CB coherence phases for tooth breakage diagnosis.

Nevertheless, CB magnitude and coherence already provide sufficient diagnostic information. From this demonstration it can be drawn that CB has superior performance for gear fault diagnosis, which also shows that it is very effective to suppress noises.

However, it has also noticed that CB has a shortage of over populated bispectrum peaks. These superfluous peaks can cause confusions and make it very difficult to find the optimal one for diagnostics. This may stem from that CB highlights only the presence of nonlinearity from the harmonically related frequency components: f_1, f_2 and $f_1 + f_2$. As gear meshing process can excite components over a wide frequency, the coupling of these components thus can result in a very dense CB peaks. This becomes even worse for a scenario when a multi stage gearbox is concerned to be under a local fault such as tooth damage which will bring more frequency contents.

5.3 Modulation Bispectrum

As examined in previous section, CB highlights sideband components at $f_1 + f_2$ but overlooks the possibility that the occurrence of $f_1 - f_2$ may be also so due to the nonlinearity between f_1 and f_2. Because of this, it is not very effective to describe modulation signals which are the most important characteristics of gear signals discussed in Chapter 4. As a result CB produces the superfluous peaks.

To overcome this shortage, the formalisation of including both sidebands in bispectrum analysis was introduced in [104]. The authors [105][106][107] have found that the formalisation is particularly useful in extracting weak fault signatures in motor current signals. It also allows an accurate quantification of quantify modulating components in diagnosing different types of mechanical and electrical faults in a machine. Subsequently it has named as MSB to highlight the unique properties of characterising modulating component and make difference from the conventional bispectrum. In the frequency domain, MSB of a signal $x(t)$ in the form of the Discrete Fourier transform (DFT) $X(f)$ is defined as:

$$B_{MS}(f_x, f_c) = E\left\langle X(f_c + f_x)X(f_c - f_x)X^*(f_c)X^*(f_x)\right\rangle \tag{5.7}$$

The phase relationship of MSB is

$$\varphi_{MS}(f_x, f_c) = \varphi(f_c + f_x) + \varphi(f_c - f_x) - \varphi(f_c) - \varphi(f_x) \tag{5.8}$$

where is f_x is 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. It takes into account both $f_c + f_x$ and $f_c - f_x$ simultaneously in Equation. (5.7) for quantifying the nonlinear effects of modulation

signals. If they are due to the modulation effect between f_c and f_x, there will be a bispectral peak at bifrequency $B_{MS}(f_x, f_c)$. On the other hand, if these components such as various noises are not coupled but have random distribution, their magnitude of MSB will be close to nil. In this way, it allows the wideband noise and aperiodic components in gear vibration signals to be suppressed effectively so that the discrete components relating modulation effects can be represented sparsely and characterised more accurately. In the meantime, superfluous peaks from CB can be also restricted significantly.

Similar to CB coherence in Equation. (5.5), a normalized form of MSB, also named as modulated signal bicoherence, is introduced as

$$b^2{}_{MS}(f_x, f_c) = \frac{|B_{MS}(f_x, f_c)|^2}{E\left\langle |X(f_c)X(f_x)X^*(f_c)X^*(f_x)|^2 \right\rangle E\left\langle |X(f_c+f_x)X(f_c-f_x)|^2 \right\rangle} \quad (5.9)$$

to measure the degree of coupling between three components against nose influences in the same way as the conventional bicoherence.

According to Equation (4.25), the MSB magnitude of the first harmonic component for the faulty gear is

$$\begin{aligned} |B_{MS}(f_F, f_c)| &= I_m^4 A_L A_H \\ &= \frac{I_m^4}{4}\sqrt{A^2+\beta^2-2A\beta\sin(\varphi)}\sqrt{A^2+\beta^2+2A\beta\sin(\varphi)} \end{aligned} \quad (5.10)$$

And phase

$$\varphi_{MS}(f_F, f_c) = -2\delta \quad (5.11)$$

This shows that MSB combines the carrier and modulating components effectively to reflect the characteristics of the modulated gear signals. The magnitude shows the strength of interaction between modulating component and the carrier while the phase provides

information upon if the modulation is PM and AM individually or in the form of combination.

5.4 Modulation Signal Bispectrum Based Sideband Estimation

To quantify the effect of the modulating components or sideband amplitudes only, MSB can be further improved by removing the high magnitude influence of f_c components through a magnitude normalisation. To differentiate this improvement from a normal MSB it is known as the MSB sideband estimator to highlight its capability of the reflection of only modulating component and also abbreviated to MSB-SE [106].

$$B_{MS}^{SE}(f_m, f_c) = \frac{B_{MS}(f_m, f_c)}{\left|B_{MS}(0, f_c)\right|^{0.5}} \tag{5.12}$$

where $B_{MS}^{SE}(f_c, 0) = \left|X(f_c)\right|^4$ is the squared power spectrum estimation at $f_x = 0$. According to Equation (5.2), the power spectrum can be obtained through MSB as

$$PS(f) = \left|B_{MS}(0, f_c)\right|^{0.5} \tag{5.13}$$

which is also an important property that can be based on to calculate individual sideband amplitude in assessing the modulation effect. This scheme of modulation sideband estimation has been shown in [106] to give more accurate result in diagnosing the small rotor bar breakage in electrical machines. By this normalisation, the modulating component of the fault gear signal can be

$$\left|B_{MS}^{SE}(f_F, f_c)\right| = \frac{I_m^2}{4}\sqrt{A^2 + \beta^2 - 2A\beta\sin(\varphi)}\sqrt{A^2 + \beta^2 + 2A\beta\sin(\varphi)} \tag{5.14}$$

to show the modulating strength due to gear fault contributions. In the meantime, MSB hpase stile the same.

5.5 Modulation Signal Bispectrum Based Detector

According to Equations. (5.7) or (5.10), it is straightforward to know that the best band for detecting a gear fault is at a f_c slice which has the highest peak: $B_{MS}^{SE}\left(f_m, f_c^*\right)$ when $f_m > 0$ in line with that it consists of significant peaks at characteristic frequencies corresponding to higher order harmonics in the direction of f_m increment. Figure 5.4(a) shows the result of typical gear vibration signal. As illustrated by '*' markers in Figure 5.4 (b) and (c), four f_c slices can be taken as optimal bands to characterise the modulation jointly as shown in Figure 5.4(d). Based on this observation, fault detection can be straightforwardly implemented. This avoids preliminary processing steps for selecting the optimal band in the envelope calculation, such as the situations for bearing fault diagnosis in which the selection of an adequate band has been investigated extensively [108][109]. Moreover, MSB significantly suppresses random noise and aperiodic impulse interferences in line with the enhancement of the modulation components. Therefore, it is very potential to result in a robust and accurate detection and diagnosis result.

Furthermore, to achieve more robust results, the MSB based detection scheme can be improved based on an average of several suboptimal f_c slices such as those with '*' markers in Figure 5.4, which leads to a robust MSB detector:

$$B(f_m) = \frac{1}{K}\sum_{k=1}^{K}\frac{B_{MS}^{SE}\left(f_m, f_c^k\right)}{\max\left(\left|B_{MS}^{SE}\left(f_m, f_c^k\right)\right|\right)} \tag{5.15}$$

where K is the total number of the selected f_c suboptimal slices, which can be the first several significant MSB peaks which depends on the values of MSB coherence. In this way, it can utilise more wideband characteristics of the impulsive excitations due to gear defects,

which ensures results to be more robust by suppressing possible strong interferences presenting in an individual f_c slice due to possible interferences.

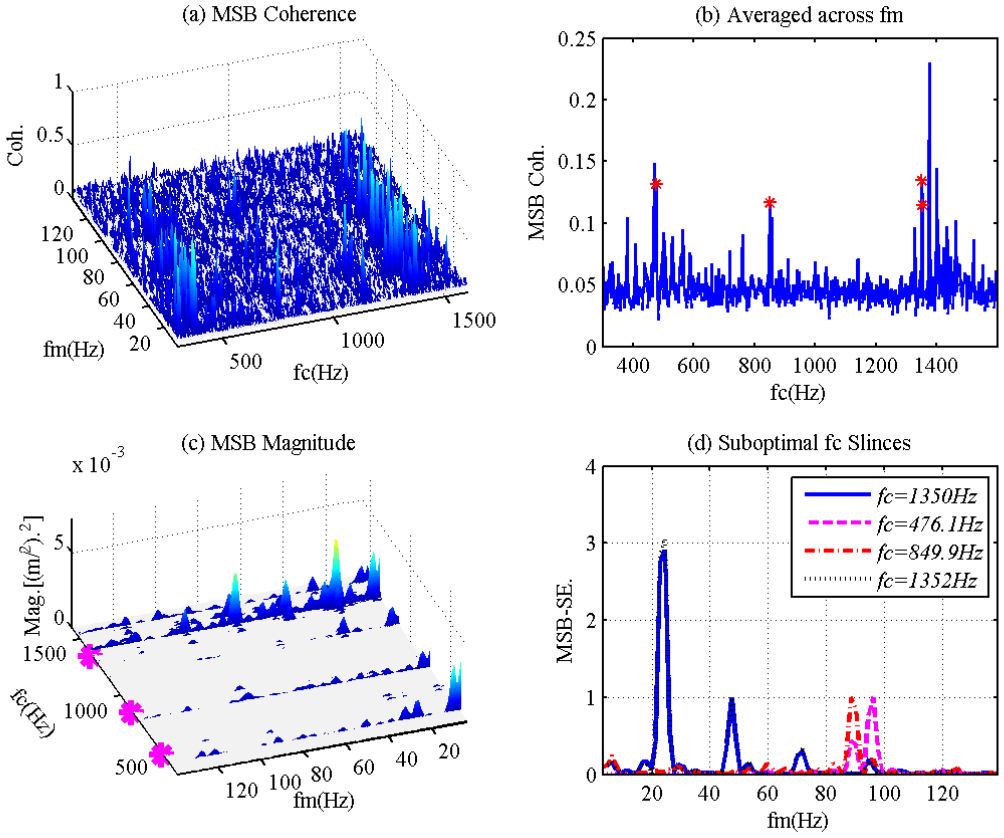

Figure 5.4 MSB and robust detector

In order to obtain the reliable f_c slices, a number of suboptimal $B_{MS}^{SE}\left(f_c^k, f_c\right)$ can be determined based on a MSB coherence slices $B(f_c)$ which is calculated by averaging MSB coherence $b^2{}_{MS}(f_m, f_c)$ in the direction of f_c increment and expressed as:

$$\bar{b}_{MS}^2(f_c) = \frac{1}{N-1}\sum_{i=2}^{N} b_{MS}^2\left(f_m^i, f_c\right) \tag{5.16}$$

It is worthy note that the average start with 2nd index of f_m indices to avoid the inclusion of the power spectrum connected slice in Equation (5.16) because its coherence are close to 1 or it has no effect on noise reduction.

Specifically, the robust MSB detector can be implemented according to five main steps:

1) Calculate MSB results using Equations. (5.7) and (5.9);
2) Calculate reliable MSB coherence slice $B(f_c)$ by Equation (5.10);
3) Select the significant peaks as the suboptimal slices by setting a threshold such as 0.1;
4) Calculate MSB-SE $B_{MS}^{SE}(f_m, f_c)$ by Equation (5.15); and
5) Calculate the detector by Equation (5.18).

The determination of threshold do not needs many trials. A true RMS value of $B(f_c)$ can be an effective one for the separation of several distinctive peaks from $B(f_c)$ as MSB already has suppressed most of the components that are not coupled but only those left are correlated in the form of modulations.

5.6 Modulation Signal Bispectrum Based AM-PM Estimator

Chapter 4 has introduced the use of modulation power index (MPI) for the measure of combined modulation effects of AM and PM and shown that the calculation of MPI estimator can be from power spectrum analysis. In the meantime, it also points out that power spectrum includes noise effect, and the estimator is developed thus is an overestimation which will results in low accuracy and unstable diagnostic results.

To improve the modulation estimator, MSB sideband estimator can be very useful. According to Equation (5.10) as well as (5.11), the MSB-SE can be expressed as a product between the lower sideband A_L and higher sideband A_H. In terms of square power the connection between MSB-SE and sidebands is

$$B_{ms}^{SE}|^2 = A_L^2 A_H^2 \tag{5.17}$$

In addition, the difference between the squared sideband magnitudes is also defined as

$$\Delta = A_L^2 - A_H^2 \tag{5.18}$$

where A_L^2 and A_H^2 can be obtained using MSB results through Equation. (5.11). It means that it is without the need to have the storage space for the power spectrum as shown in Equation (5.2). It is worthy to note that Δ is a noise minimised estimation because it is very common that the noise content are additive in measured vibration signals and that the bandwidth between two sidebands is narrow.

Letting $A_H^2 = A_L^2 - \Delta$ and substituting it into Equation (5.20) yield a quadratic equation:

$$A_L^4 - \Delta A_L^2 - |B_{ms}^{SE}|^2 = 0 \qquad (5.19)$$

Equation. (5.22) is the standard form of quadratic equation in which the coefficients of the quadratic formula are as

$$a = 1; b = -\Delta; c = -|B_{ms}^{SE}|^2 \qquad (5.20)$$

Importantly, both Δ and $|B_{ms}^{SE}|^2$ are both noise minimised quantities. Solutions of A_L^2 and A_H^2 by using Equation (5.23) are more accurate and reliable. Therefore, the modulation power index for reflecting combined modulation effects of AM and PM:

$$Z^2 = 2(A_L^2 + A_H^2) \qquad (5.21)$$

will be more accurately quantify the effect of modulations to achieve accurate diagnostic results.

5.7 Characteristics of Simulated Modulation Signals

To evaluate the performance of the MSB based estimator, it has been applied to simulated signals with different degrees of phase variation and noise levels. In the meantime, as the effective methods accepted widely, TSA, CB analyses are also applied to the signals in order

to benchmark the performance that are resulted from MSB in characterising the interference of AM-PM modulation and in suppressing noise influences.

The signal is based on the model in Chapter 4 but it consists of only several lower orders of modulation components, for being clear in presentation and explanation. Moreover it includes the all necessary components of a gear vibration signal which includes the low frequency harmonic components, AM and PM components, meshing components and noise content. Specifically, the signal has following terms:

$$x(t) = a(\mathrm{t}) + [1 + a(t)]\cos[2 \times 120\pi t + b(\mathrm{t})] + n(\mathrm{t}) \tag{5.22}$$

in which $n(t)$ is Gaussian noise with a standard deviation (ST) 0.35. This value is used, as a higher value than that would cause TSA implementation unstable and make effective comparisons with MSB analysis. It results in a SNR of about 8dB for the signal as a whole, which can vary with modulation interference levels. The carrier is at 120Hz, and the modulating signals $a(\mathrm{t})$ and $b(t)$ are

$$a(t) = A\cos[2\pi 20t] + 0.8A\cos[4\pi 20t] + 0.5A\cos[6\pi 20t] \tag{5.23}$$

$$\mathrm{b}(t) = \chi A\cos[2\pi 20t - \varphi] \tag{5.24}$$

Where A=0.1, being 10% of the carrier and common cases for small faults. PM coefficient varies in the range $\chi = 0, 0.1, 0.2, 0.3, 0.4, 0.5$. It means that the maximum PM magnitude is 0.05 or 5% of variation of shaft speed which is the common of tolerance for most industrial motor systems. The carrier frequency is the result of 20Hz×6=120Hz. It implies that the signal is from a system with a baseline shaft frequency of 20Hz on which is installed a gear with 6 meshing teeth.

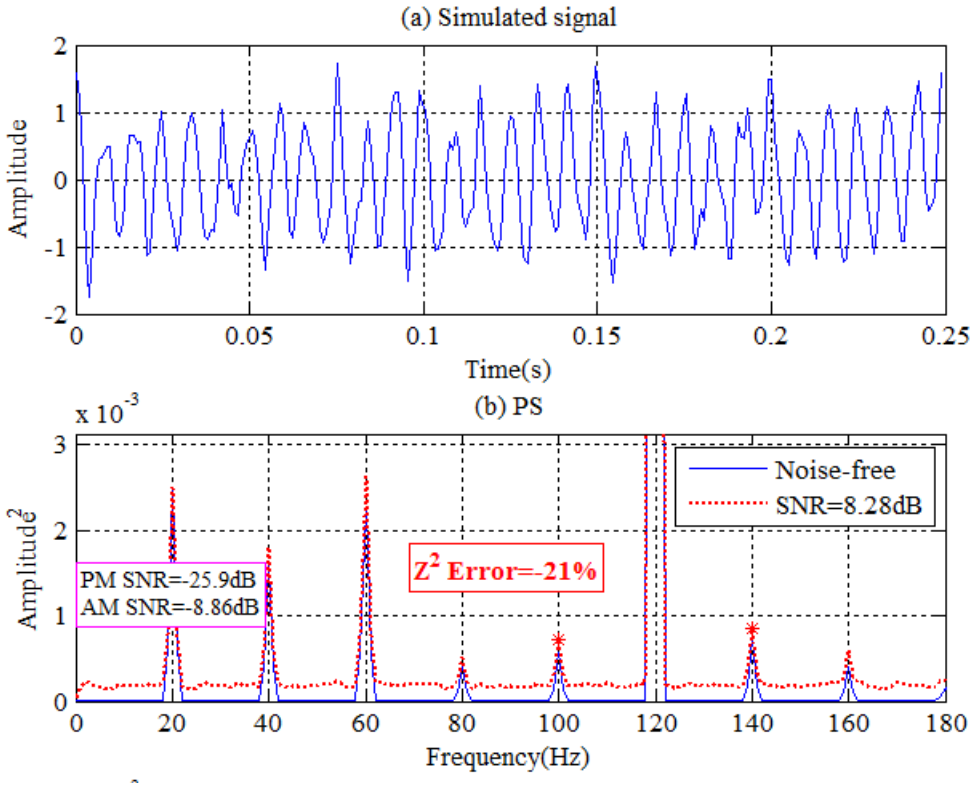

Figure 5.5 A typical signal simulated under $\chi = 0.1$ and $\varphi = \pi/10$, with its power spectrum

The AM signal is created to have three components in order to show the effect of interferences between modulation and harmonic components, which is common cases gear signals especially that of multistage gearboxes as they have multiple meshing and error components in the signal. The PM signal only has one contents but its amplitude and phase vary in a wide range to show the influences of noise and interaction on the estimation accuracy. In addition, $A = 0.1$ which 10% of the carrier for reflecting the small modulation content from incipient faults in measured gear vibration signals.

The simulation is carried out based on a sampling frequency of 1024Hz, and 1024 points is for all FFT calculations including the power spectrum and CB, MSB. This FFT frame allows the avoidance of error due to spectral leakage in DFT calculations. In addition the overall

length of signal is 51712 points in order to obtain sufficient averages for noise reduction in implementing CB and MSB approaches.

Figure 5.5 illustrates a typical subsection of the signal simulated with both PM magnitude and phase effect interferences when $\chi = 0.1$ and $\varphi = \pi / 10$. As the magnitude of PM only 10% of that of AM its SNR is only -25.9dB, compared with the noise content added, whereas the AM component has a SNR of -8.85dB. Although these additional components including the harmonic components are have small amplitude, simulated signals exhibit clear modulation phenomena. However, the signal still shows relatively clean as the noise content is much lower compared the sinusoidal wave at the carrier frequency.

Meanwhile, its power spectrum also can allow these small modulations observable in the magnified graph, showing the powerful effectiveness of spectrum analysis. As shown in Figure 5.5 (b), all the sidebands around 120Hz and the harmonics in the low frequency range can be seen clearly. However, they show higher amplitudes compared with its noise-free case. Because of the noise addition, the amplitude estimation from the average of the sideband pairs will have an overestimation. As shown in the figure, the magnitudes at the sideband pair at (120-20)Hz and (120+20)Hz is slightly higher than that of noise free one. Consequently, the calculation of MPI or Z-square at the 1st set of sidebands due to the sidebands give an error of over 20%, being significantly lower than that of the induced components for both PM and AM. Obviously such an error can lead to a poor diagnosis.

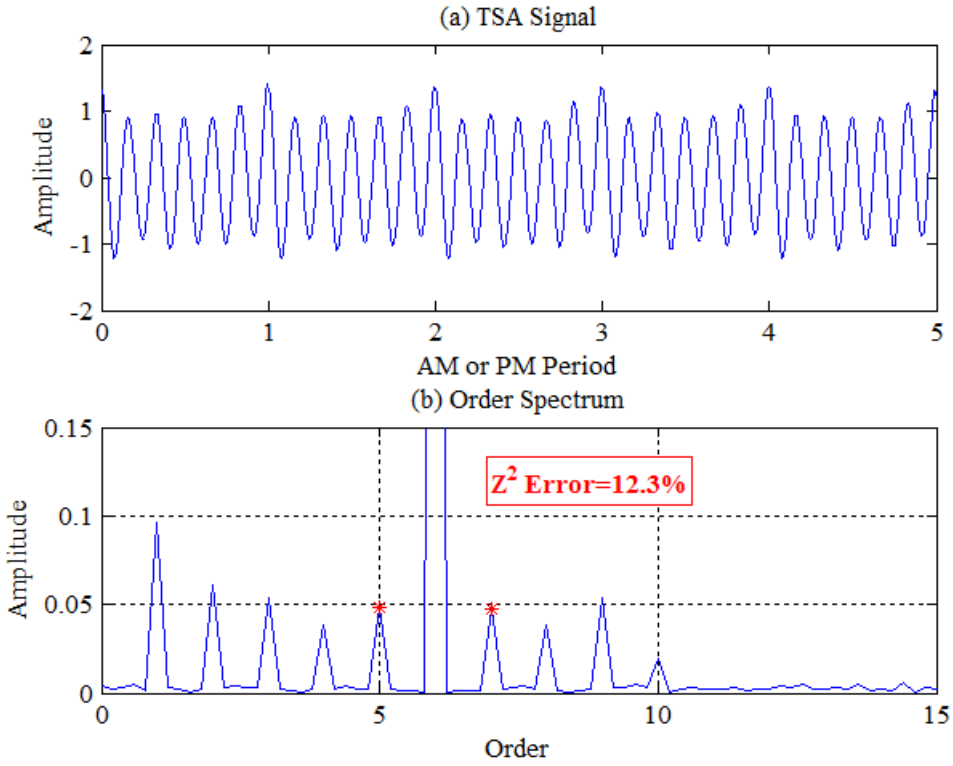

Figure 5.6 TSA signal and its order spectrum for modulation analysis

5.8 TSA based analysis

To examine the effectiveness of TSA analysis upon this signal, the increment phase signal of the carrier at 120Hz is calculated using Hilbert transform, the raw signal is then resampled to be decomposed into the phase signal to obtain signals in the angular domain. Following this, the resampled signal is divided into smaller segments with a length of integer periods of the PM signal, and finally the average can be performed across different segments to obtain the TSA signal. For this case the segment length is of 5 periods of the baseline frequency 20Hz. This TSA analysis attains 200 averages and adequate resolution for separating each component induced in the order domain which is realised by applying FFT to the TSA signal. Figure 5.6 shows the results from TSA analysis of the same signal as that of implementing power spectrum analysis.

The TSA signal in the angular domain (expressed in PM period) in Figure 5.6 (a) presents similar information to the raw signals, which tells that the signal consists of mainly the carrier component and certain degrees of modulation. However, its order spectrum in Figure 5.6 (b) exhibits distinctive sidebands and much lower noise floor. As this TSA signal spectrum shows typical modulation sidebands, its corresponding MPI can be calculated according to Equation (5.24) for measuring the combined modulation of AM and PM. For this lower noise sideband, however, it produces a result 12.3% higher than that of the induced components for both PM and AM. Although the error is lower than that of PS based estimation, it is still very significant to influence diagnostic result.

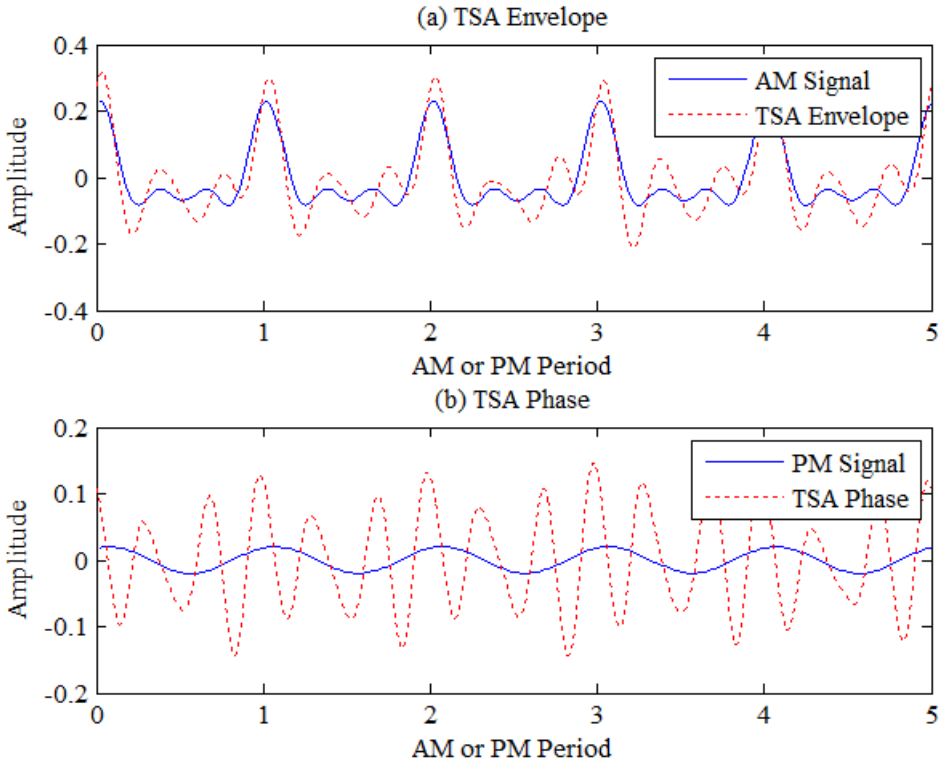

Figure 5.7 Envelope and phase by TSA analysis

Further examination of the reasons behind of this high error resulted from TSA by changing the noise levels added to the signal has found that the high error value is originated from the

low robustness of the algorithm. When the noise level is higher than that of the PM amplitudes it is difficult to obtain reliable phase estimations and consequently, the TSA result cannot be accurate. Figure 5.7 shows both the envelop signal and the phase signal obtained in the process of TSA analysis. They have very significant distortion to the original signals. Especially, the phase distortion on which TSA is based for implementing the resampling using a cubic spline data interpolation can cause significant errors to the resampled signals. Therefore, the TSA analysis of high noisy signals can lead to high errors.

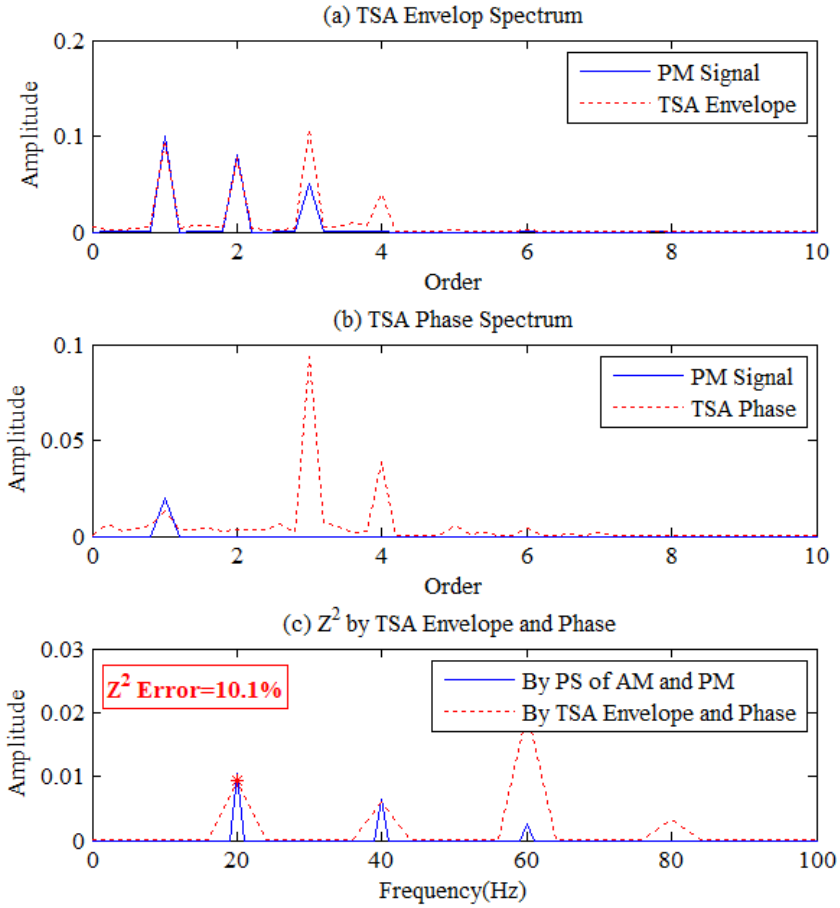

Figure 5.8 Envelope and phase by TSA analysis

Furthermore, examining the phase and envelop signals in the frequency domain has found that the distortion appears mainly in high frequency or high order components, as shown in

Figure 5.8. These are resulted from the interference between the sideband and the higher order harmonics due to the first term in Equation (5.19). It seems that a narrow band filter can be applied to eliminate these components. In practice, it is difficult to determine the bandwidth as gear faults could induce components in a very wide band. Nevertheless, the selection of an optimal band for implementing TSA for gear diagnosis has received considerable attention as outlined in Chapter 3.

However, it is worth noting that the error of MPI by using TSA envelop and phase signals is 10.1%, being a slight improvement compared using TSA signal directly. Even though, this inaccuracy can be still too high to obtain accurate diagnosis. In addition, more numerical investigation by adding more noise such as at ST= 0.4, TSA operation is unstable, resulting in extreme deviation in MPI. In this sense, it is not very robust. Therefore, the use of TSA needs more attention in selecting appreciate parameters for reliable implementation, which is still an open problem deserving more studies

5.9 CB analysis

Figure 5.9 presents results from CB analysis for the same signal case when the PM is 2% of AM and the phase difference is $\varphi = \frac{\pi}{10}$. The calculation uses a 1024 point FFT which produces a 1Hz frequency resolution. By overlapping 50% between FFT data frames, which are applied by Hanning windows, 100 averages are achieved during the calculation. For comparison, the results are presented only in the same frequency bandwidth as before in order to examine their performance in a low or a narrower frequency range. In addition, it ignores the redundant contents as they are symmetrical to that presented in Figure 5.9.

Aiming at separating modulation effect, CB results are mainly addressed regarding to components relating to sidebands. Thus distinctive components in CB magnitude results such

as three main peaks at bifrequencies (20, 120), (40,120) and (60, 120) Hz are from the coupling between the carrier and one sideband only, rather than the matching pair. For instance, the peak at bifrequency (20, 120) Hz is the product between 20Hz, 120Hz and 140Hz, which does not include the lower sideband of 100Hz, It means that these peaks are not capable of obtain the full information about the modulation. Therefore, it is impossible to calculate the MPI for quantifying modulations adequately.

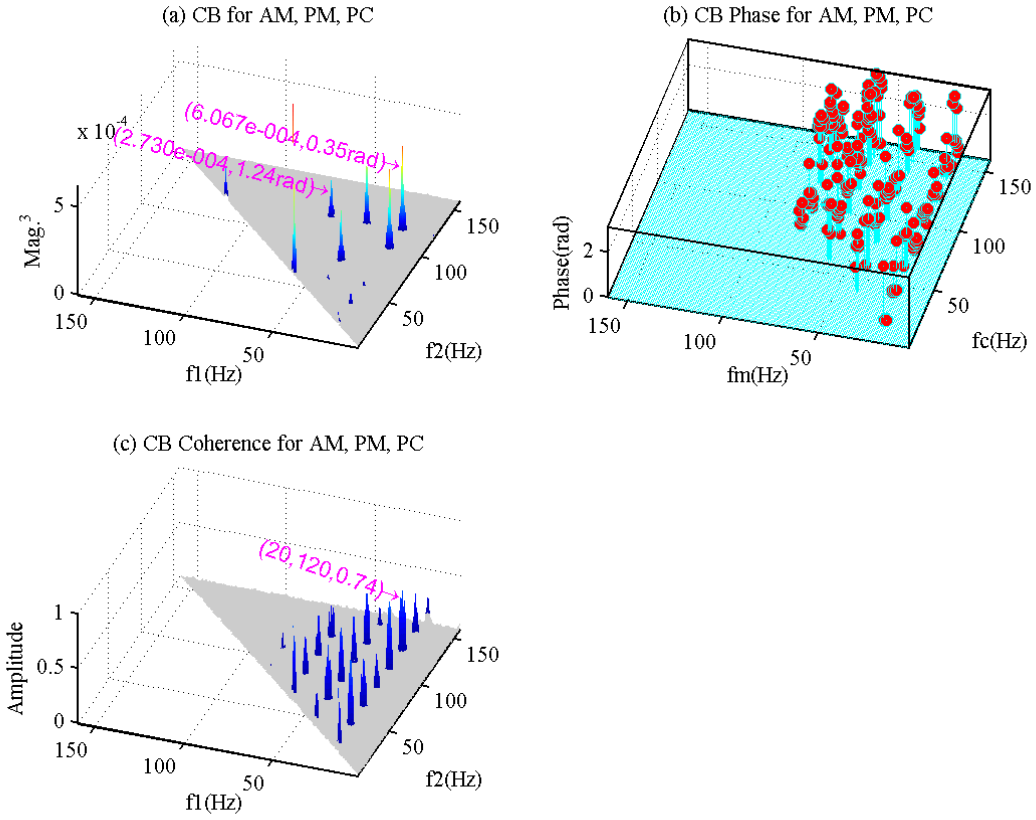

Figure 5.9 Results from CB analysis

In addition, the high density of coherence peaks shows that the non-modulation components: 20Hz, 40Hz and 60Hz can lead to interactions with themselves and the modulating and carrier components. These surplus components can make it even more difficult to select useful components.

The modulation effect in CB could be shown in a higher frequency range such as 2×fc=240Hz for the closest one. This means that it needs a wider band to show full modulation effect, and hence consume more computational resources. This can be a considerable deficiency for real time diagnosis. Not only more computation is needed for CB analysis but the resulted components are more likely to be inadequate representation of modulation because the vibration responses of a system can be very different between two components with distant frequency bands, which will be examined further in experimental study contents.

5.10 MSB analysis

By using the same analysing parameters as CB, MSB produces corresponding results as shown in Figure 5.10. From the magnitude results, it is clearly shown that MSB is very selectable for the modulation components of interest, enabling enhancing the three components of AM+PM by a MSB slice at fc=120Hz and sparse representation.

Figure 5.10(a) shows three distinctive peaks appearing at bifrequencies (20, 120), (40,120) and (60, 120) Hz respectively. According to the definition of MSB these peaks reflect the modulation of three AM components, one PM component with the carrier respectively. However, because of the interferences between the PM and the first AM components, the phase peak at (20, 120) Hz is 0.51rad instead of zero. So does the phase of the peak at (60, 120) because the interference occurs between the component at frequency (120-60) Hz of AM modulation and at 60Hz of the third harmonics of AM when taking as the low frequency harmonics. As a result of these interferences, the magnitudes do not represent the real degree of modulations corresponding to the levels induced to the signal.

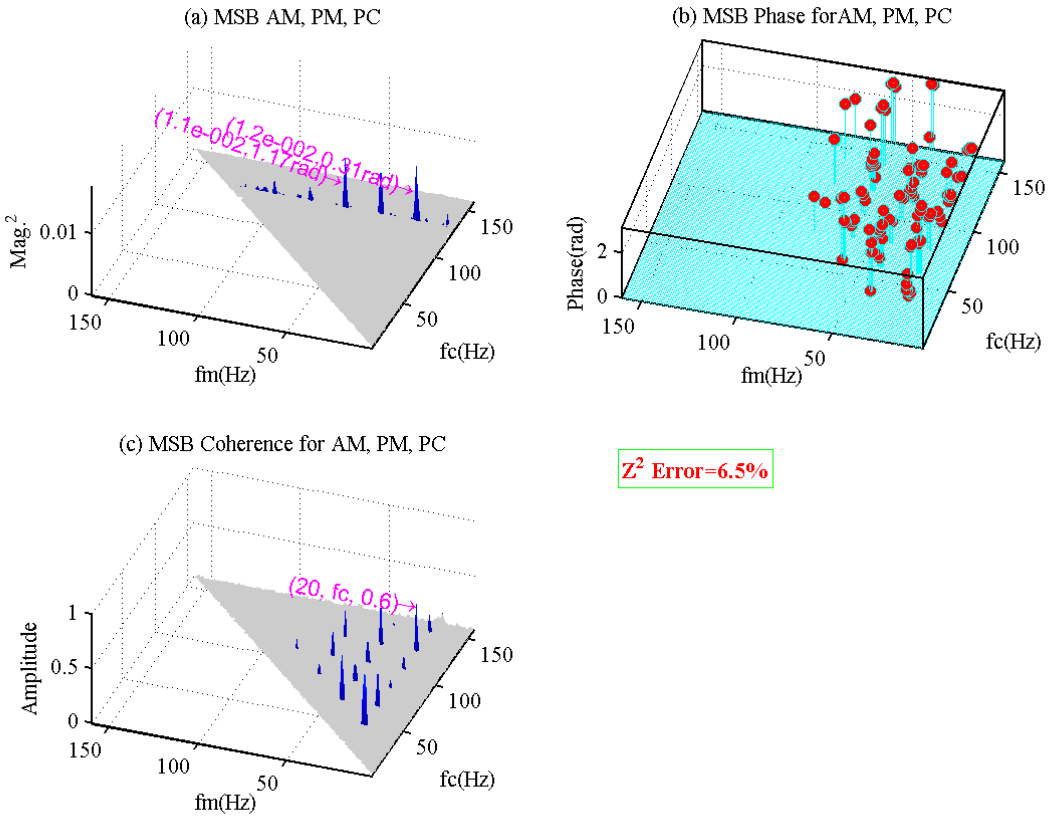

Figure. 5.10 Results from MSB analysis (*fm*>1)

Moreover, these MSB peaks are ensured by their corresponding coherence peaks as shown in Figure 5.10 (c). These high coherence peaks indicate that these peaks are very outstanding from the noise levels and corresponding magnitudes can be trusted for further analysis. However, the coherence still shows many other peaks due to the interactions between any spectral peaks with their neighbour peaks each side. For example the AM harmonic peak at 40Hz can produces interactions with 20Hz and 60Hz, leading to a MSB peaks at (20, 60). Although the number of peaks is much less than that of CB, they can also cause troubles to be separated from true modulations for diagnosis. Nevertheless, these peaks all stem from the same excitations as that of modulation. The involvement of them as the ways for modulation would increase the reliability for detection.

Moreover, MSB results shows that the modulation effect can be fully reflected at the slice of fc=120Hz i.e. the meshing component. This will simplify the understanding and as well as computational consumption.

Based on these observations, MPI can be obtained by following steps:

1) Select the maximum or mean of coherence values across different f_{me} slices, leading to Figure 5.11 (a);
2) Calculate MSB-SE at 120Hz slices for the detector ;
3) Calculate MPI at those frequencies by using a coherence threshold such as higher than 0.1, which results in Figure 11(b).

Compared with the TSA and PS, the MPI from MSB analysis at the first sideband set has an error 6.5%, showing a significant improvement. This shows more that MSB is more effective in characterising the modulations.

It is worthy mention that phase results for both CB and MSB have found to be not as useful as the magnitude and coherence for the purpose of modulation quantification. However, as discussed previously it can help to differentiate whether a MSB peak is from single type of modulation or both. As shown in the Figure 11 (c), because the second peak is from pure AM, its phase is close to zero. Others peaks having phase values higher than zeros may show that they are from the combination between AM and PM modulations. This will be helpful in explaining mechanisms for the MSB peaks identified. Moreover, phase characteristics associated with modulating components have already played decisive roles to suppress noises and enhance modulations during the MSB calculation process.

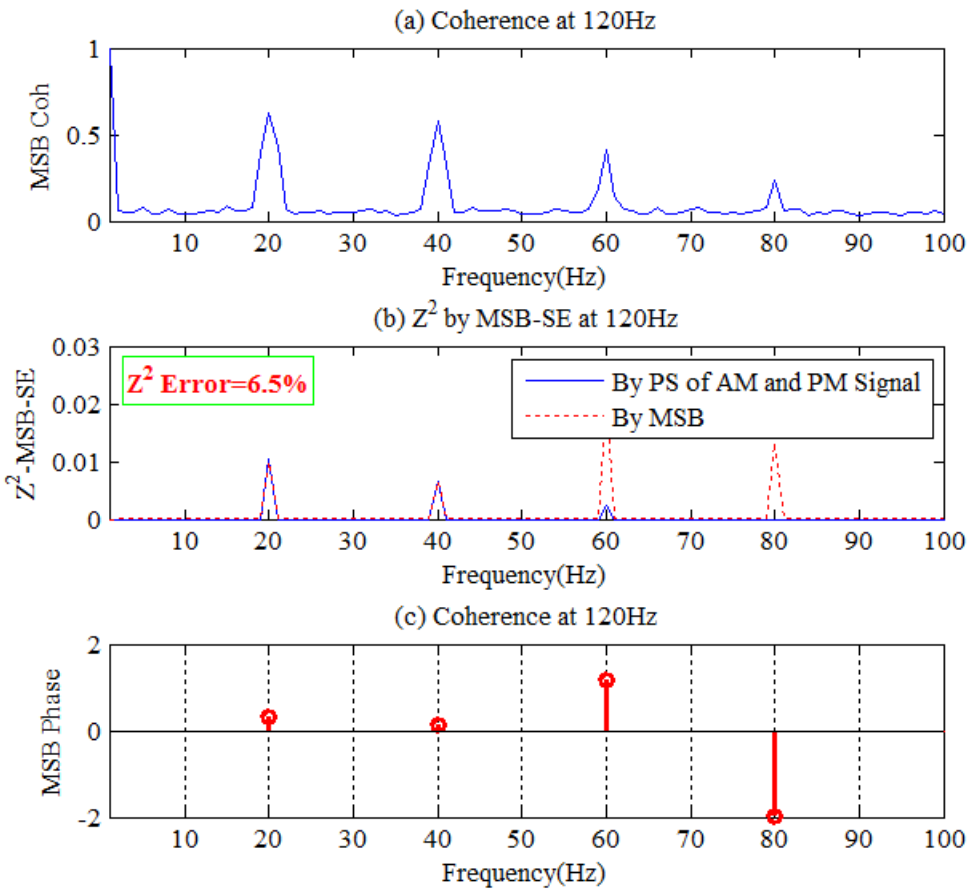

Figure. 5.11 Results from MSB analysis for $f_m > 20$ Hz

5.11 Performance Comparison

As CB has the deficiency in representing modulation effects inadequately, a full performance benchmark of MSB is made against just TSA analysis and power spectrum analysis. Simulation studies were conducted to obtain MPIs under different interferences between AM and PM under the noise level SD=0.35 at which TSA can be relatively stable to allow the results at different cases to be achieved. As shown in Section 5.8 and Section 5.9 the MSB based estimator is achieved by using the MSB-SE slice at the $f_c = 120$ Hz and that of $f_m = 20$ Hz. All methods are compared for the estimation of the AM-PM component at 20Hz

and the result is ensured by over 20 randomised Monte Carlo tests, where the noise was different for every case.

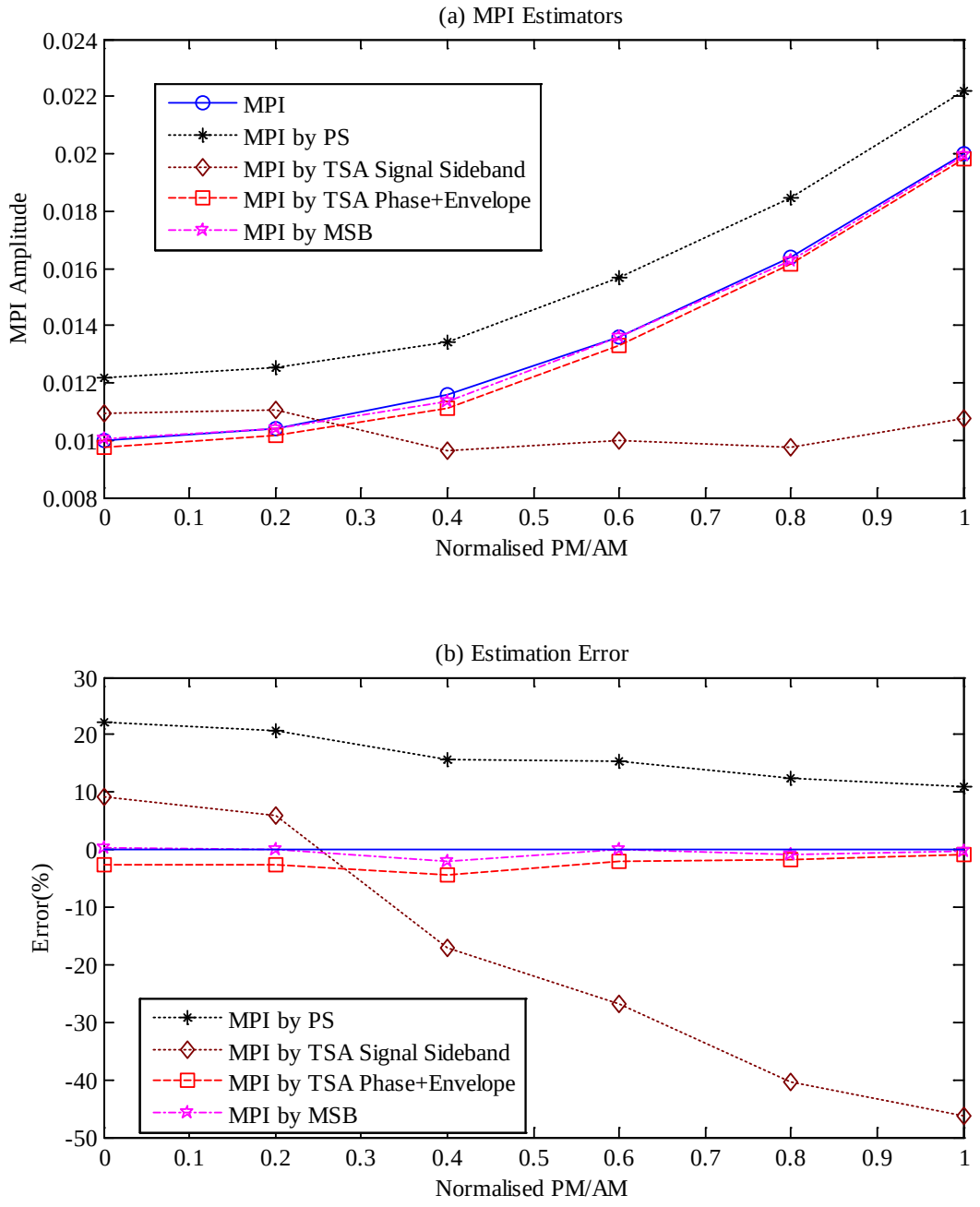

Figure 5.12 Comparison between MSB based estimation and TSA based estimation of the AM-PM effect.

Figure 5.12(a) shows - the comparison between different estimators and against the theoretical MPI. It shows that MPI values increase with more PM content induced even if the PM phases vary between 0 and π. It shows that MPI gives full measures of the combination of PM and AM.

Compared with noise free case, PS based method produces the highest deviations whereas MSB based method gives the best estimation. The estimation from TSA signal sideband fluctuated around, showing that the phase errors can be significantly wide deviations due to noise influences.

TSA envelope and phase signals can jointly produce results closer to the theoretical but has error showing that the estimation is slightly lower. This shows that this method can be also accurate if phase noise is filtered effectively.

Moreover, as shown in Figure 5.12 (b) the errors variation over different PM interferences has little influences on the accuracy of MSB based estimation. Overall, its average error is as low as less than 1%, which will allow produce accurate estimation of modulation effects for vibration based fault severity diagnosis, verifying that MSB based analysis outperforms the others.

In addition, Figure 5.13 (a) also shows that MSB based estimate can show stable results for different PM phases. It means that the MSB estimator is not only robust to noise but also to the influences of phase interferences.

On the other hand, the estimation directly from MSB peaks shown in Fig 5.13 (b) shows highly fluctuation. It means that it is not possible to obtain a stable result for diagnosing fault severity as the PM effect in gearbox vibrations can exhibit different phases relative to AM

due to possible changes in gear fault severity, operating load, running speed and noise contents between different measurements.

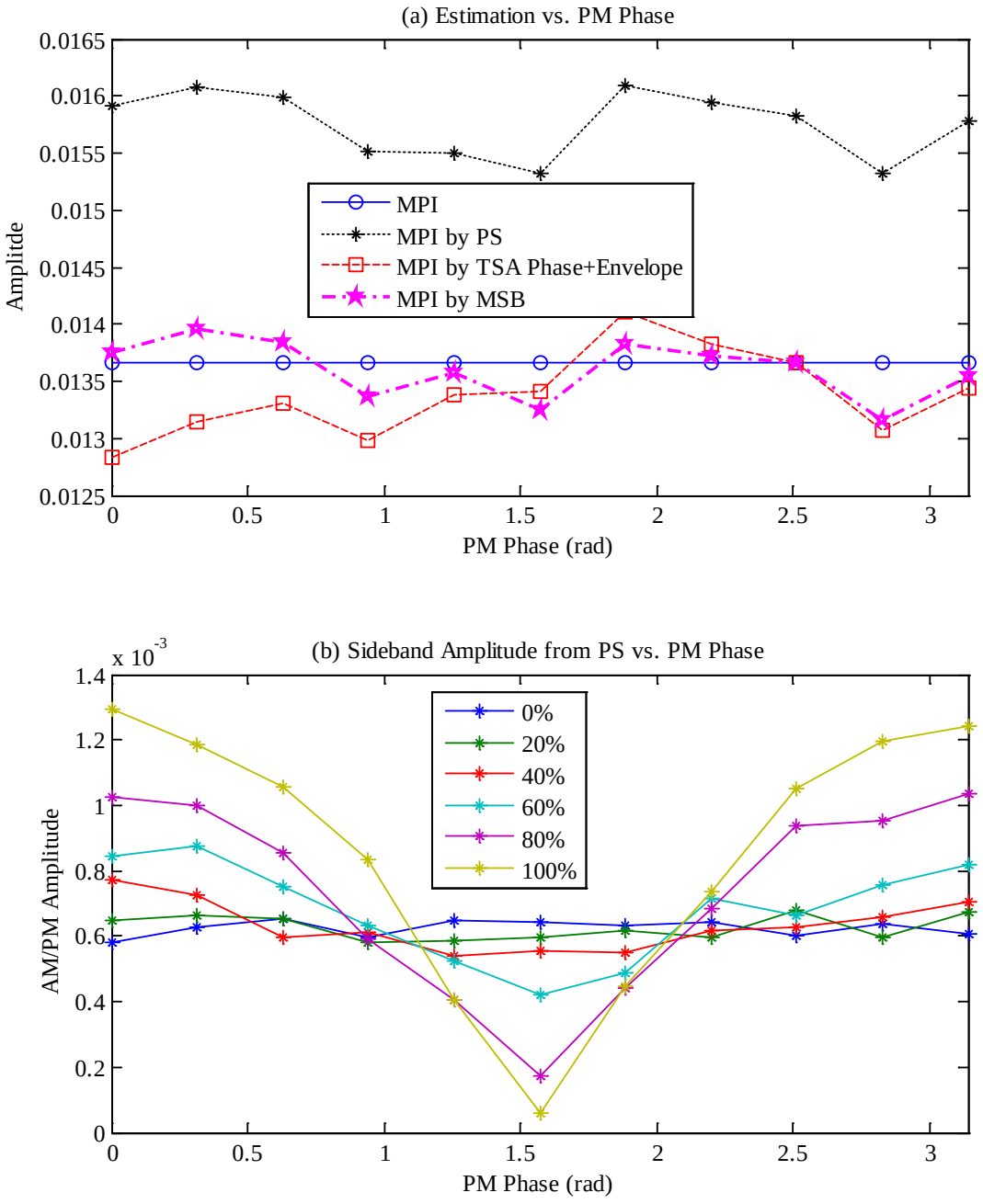

Figure 5.13 The estimation of AM-PM effect by PM phase variation.

5.12 Chapter Summary

Based on the characteristics of MSB analysis, a new indicator, referred to as modulation power index (MPI) is developed to accurately reflect the content of both AM and PM in gear vibration signals. Moreover, it shows that MSB based analysis outperforms TSA and CB based on simulated signals generated from the gearbox model.

CHAPTER 6.

GEARBOX VIBRATION DATASETS FOR EXPERIMENTAL STUDY

6.1 Introduction

To verify approaches in previous sections, in particular the simulated faults in Chapter 5, vibration datasets were acquired based on a typical industrial gearbox operating under different gear conditions including local tooth defect and uniform tooth wear, which are common failure modes in gear transmission [110][111].

The details of the test system and its simulated faults are outlined in this chapter. The type of gearbox used is two staged with helical gears (the gearbox is commonly used commercially, so data collected is more realistic and relatable to industrial applications). Firstly, an overall layout of the test rig is provided, then the chapter explains each of the components which make up the test rig. Then information is provided on the type and severity of tooth breakage faults which were used and how they were created. The chapter also looks at how the vibration data was processed and collected, the position of the accelerometers and the different operating conditions under which they were collected. Finally, the layout of the data set is given, the data collected will be used to test the advanced signal processing techniques in order to find the simulated faults.

6.2 Experimental Rig Setup

A complex test rig is required to collect enough accurate data that can be used to test complex signal processing techniques. Experimental work had to be conducted using a test

rig developed at the University of Huddersfield in order to gather an accurate and reliable data set.

First of all a 3-phase induction motor which provides drive to two permanent helical gearboxes connected back to back, couplings, load devices, and a resistor bank, as shown in Figure 6.1, the schematic diagram of the rig is shown in Figure 6.2.

The main aim of these experimental studies is to monitor the condition of a healthy gearbox operating under different loads, piezoelectric accelerometers are placed on the casing to provide accurate vibration data.

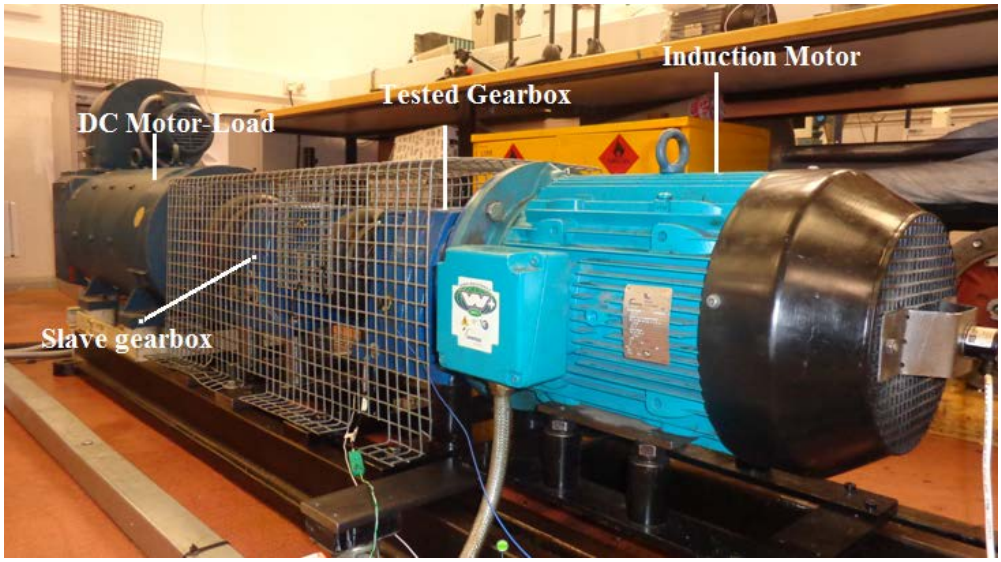

Figure 6.1 The experimental test rig with motor, load and both gearboxes

The faults induced in the first gearbox are local specific teeth removal fault and general fatigue gear teeth wear from prolong use. The faults were firstly a percentage of breakage in a specific gear tooth, this local breakage fault was induced artificially, and secondly a fatigue fault of the gear was created from the prolonged operation of the gearbox that was being

tested. Vibration data is collected from both faults are to be used for CM technique application.

This experiment set up was used as it allows for various faults to be artificially created in the gears, it also allows for the data to be collected relatively easily. Also since the setup is commonly used in industry, it has practical applications.

6.3 Experimental Rig Layout

The schematic of the experimental rig for this research is shown in Figure 6.2.

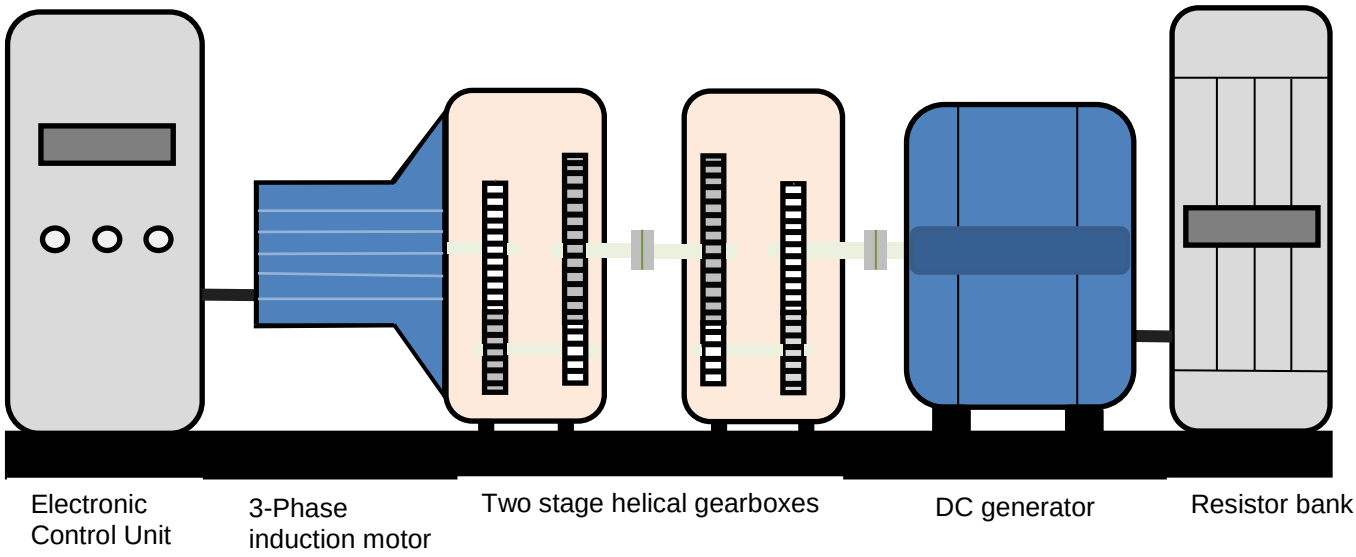

Figure 6.2 Schematic layout of the experimental gearbox test rig

The main gearbox (test gearbox) acts as a speed reducer flanged in a cantilever type arrangement with the induction motor made by Electro-drives Company. The two gearboxes are both two stage helical gearboxes manufactured by David Brown Radian Limited, the

gearboxes are connected back-to-back. Mechanical specifications of both gearboxes are provided in Table 6.1 below.

Table 6.1 Performance specification of the two-stage helical gearbox

Description	First Stage (Input Shaft I)		Second Stage (Output Shaft III)	
Number of teeth	*Z1/Z2*=58/47		*Z3/Z4=13/59*	
Shaft speeds*(Hz)* at rated load and speed	$f_{r1} = 24.42$ (input)	=30.16 (middle)		=6.64 (output)
Meshing frequency	$f_{me12} = 1416.36$ Hz		$f_{me34} =$ =391.76 Hz	
Contact ratio	1.450		1.469	
Reduction ratio	0.8103		4.5385	
Shaft	42mm		40mm	

First of all, using a control panel, a three phase system provides power to the AC motor. The AC motor is already integrated with the gearbox, in its output is directly connected to the input of the testing gearbox. This minimises the need for extra components thus reducing the cost and also improves reliability.

The output of the main gearbox is connected to the second gearbox via a flexible coupling, another flexible coupling is again used to connect the output of the second gearbox to the DC generator. The combination of the second gearbox and the DC generator acts as a mechanical load for the output of the experimental gearbox, which operates at a low rpm and high torque. Since the DC generator requires a high rotational speed of around 1750 rpm for its input, a second identical gearbox that is reversed is required to increase the rotational speed again so that it will fulfil the requirements of the DC generator.

Finally, the mechanical energy supplied from the output of the second gearbox is converted back to electrical energy by the DC generator, this electric energy is finally dissipated in the

resistor bank. By controlling the resistor values of the resistance bank, different load values for the system can be simulated.

6.4 AC Motor

The power supplied to the rig is a three phased AC motor, a photo of the power generator is shown in Figure 6.23. The AC motor is rated at 11kW and has 2 pairs of poles, its output is 1465rpm at full load, further specifications can be found in Table 5.2.

Figure 6.3 Photograph of an AC motor/gearbox combination

The operation of an AC asynchronous motor starts with the stator winding coil being energized by the AC power supplied to it a 3-phase electric current passes through the windings of the stator will induce a 3-phase electromagnetic field [112], with each of the phases being equal in magnitude. This creates magnetic poles that fluctuate back and forth synchronised with the direction of the current. These electromagnets are positioned so that

when active the magnetic poles appear to rotate around the rotor, this rotation speed is also known as the synchronous speed (N_s), and as a consequence a magnetic field will be induced in the rotor which will create a tangential magnetics force. The rotor can then rotate if this magnetics force is greater than the torque of the motor.

Table 6.2 Induction motor electrical specification

Type of Motor	Induction
No. of phases	3
No. of poles	2-Pair
Rated power	11kW
Rated voltage	415 V
Current	22A @Full Load
Rated speed	1465 rpm @Full Load
Winding	Y Star To Δ delta
Number of stator slots	48
Number of Rotor slots	40

The rotational speed of the rotor,N_r is determined by the supply frequency of the rotor and its number of poles. By using (6.1) and (6.2), the motor's synchronous speed can calculated from the main supply frequency multiplied by the phase difference between the stator coils (120°), the number of poles and the slip of the motor [113]. Once the AC motor is manufactured, the number of poles and its slip are fixed and cannot be changed.

$$N_s = \frac{120.f_{main}}{P} \tag{6.1}$$

$$N_r = N_s(1-S) \tag{6.2}$$

$$N_s = \frac{120f}{P} \tag{6.3}$$

Where: N_s is the synchronous speed; N_r represents the rotor speed; f_{main} is the main supply frequency; P is the number of poles, and finally S is the slip.

6.5 Gearbox

The gearbox was supplied by a major manufacturer, it is connected in a cantilever type arrangement with the AC induction motor as shown in Figure 6.3. The major mechanical specifications of the helical gearbox are found in table 6.1. The gearbox has two stages, in the first stage the first pinion is attached to the first gear, where the ratio is reduced by 0.8103 (or an increase of 1.2341). The first gear lies on the same shaft as the second pinion, which is attached to the second gear, this reduction ratio is 4.385. The second gear's output is connected via a coupling to the second slave gearbox which reverses the ratios.

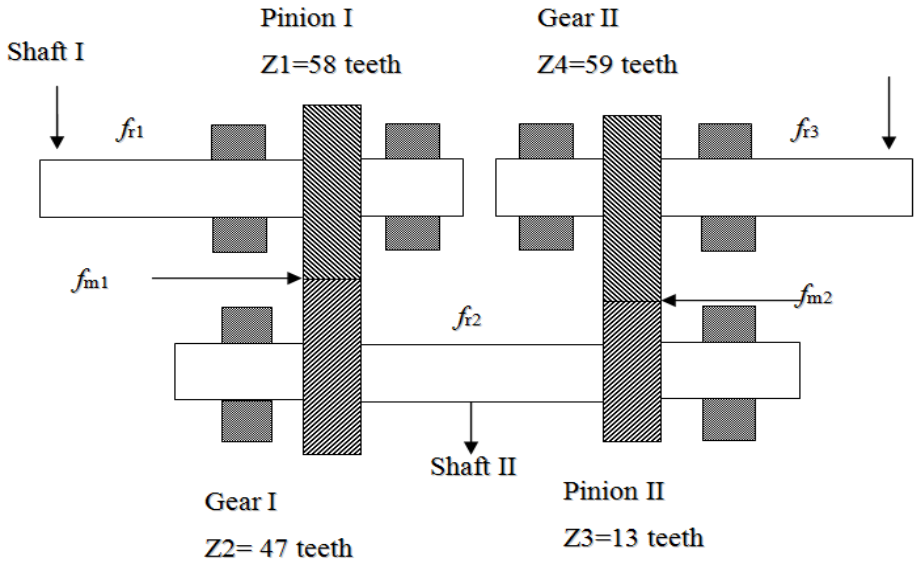

Figure 6.4 the schematic diagram of the two-stage helical gearbox.

Consumer demand for lower rotational speed and higher torque output has increased the popularity of AC motors already combined with a gearbox. Thus there is no longer any requirement for an extra coupling to connect the output shaft of the AC motor and the input shaft of the gearbox, this prevents the occurrence any misalignment errors at the coupling. The motor and gearbox combination is constructed by installing the front flange plate of the gearbox directly onto the flange plate of the AC motor, with no gaps in between the two plates.

6.6 DC Generator (The Mechanical Load)

The experimental set up uses a DC motor as the generator. A DC motors converts the mechanical rotational energy into electrical energy. When the shaft is rotating, the stator windings of the stator produce electro-magnetic flux lines that flow around the stator [114]. Figure 6.5 shows a simplified diagram of the components that make up a typical DC generator.

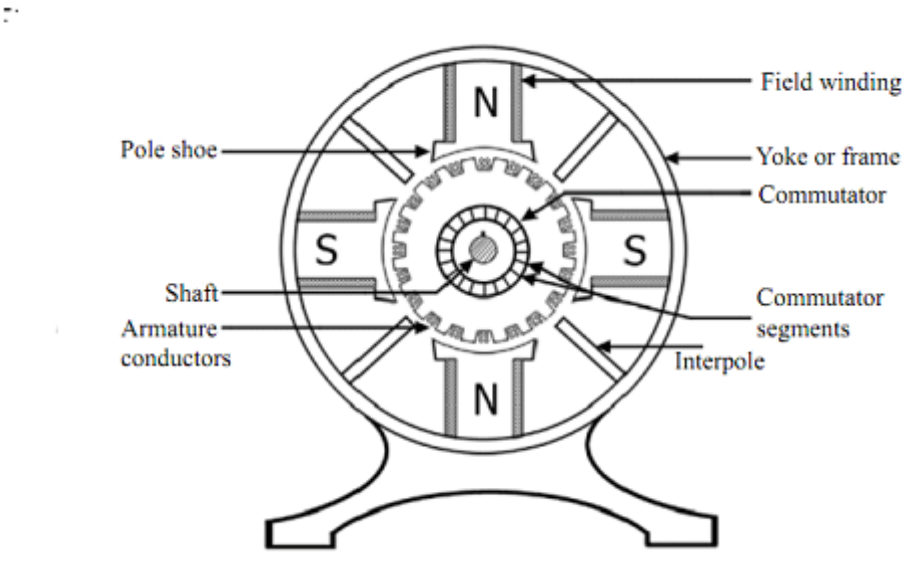

Figure6.5 Outer and Rotor Design of a DC Motor

Between the pole and the armature is a small air gap, the magnetic flux lines cut the armature windings. The field current is relatively low as there are multiple coils generating the magnetic field, it will increase slightly if a given large increase in the DC motor power is provided.

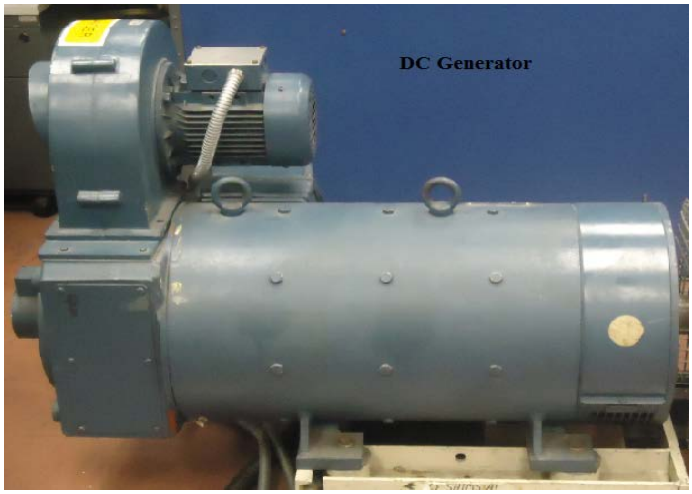

Figure6.6 DC generator (mechanical load)

The specific DC generator motor that is used in the experiment setup is shown in Figure 6.6. The generator can produce 85kW at 1750 rpm. This is rated much higher than the output of the AC motor, therefore the generator needs to be adapted so it can operate at the necessary loads and rotational speeds. The setup utilises a resistor bank to dissipate the electrical energy generated by the motor so that it stays an open system, this prevents any electrical noise from affecting the rest of the experiment setup[115].

Classification	DC (shunt)
Number of phases	3
Number of poles	2pair
Rated power	85 kW
Voltage (excitation)	360 V
Current	4.7A
Rated speed	1750 rpm

6.7 Test System Electronic Control Unit

An electronic control unit was employed to ensure accurate operation environment, thereby gathering reliable data and minimising any discrepancies. It will allow the collection of obtain multiple repeated testing whilst maintaining the identical operating speed and load profile consistently throughout the experiment's run [116]. The electronic unit monitors details of the operating condition which includes the speed, load and temperature. These factors are kept the same to ensure that they do not influence the data after the gearbox faults are introduced

For the controlled test, 5 repeating AC and DC motor settings were used. The AC motor ran at 100% speed throughout these 5 tests, whereas the DC motor was loaded at 20%, 40%, 60% 80% and 100% of its full load.

Figure 6.7 Test rig control panel

6.7.1 Electrical Drive Specification

A major advantage of an AC drive is that the speed of the induction can be controlled precisely. There are several methods by which induction motor speed control is achieved:

- Open loop (no feedback)
- Closed loop (a feedback device is mounted on the induction motor to supply a feedback signal representing the actual shaft rpm)
- Slip compensation (uses the output current transducers for current feedback to estimate the slip of the induction motor).

The present test rig was operated with an open loop system and used simple push buttons for start and stop conditions.

6.7.2 Load Control

The required operating condition for the induction motor used in this research is the motor mechanical load.

This is provided by the DC motor. It is not controlled by a DC drive -- as an alternative the motor dissipates the generating power of the loading in form of heat by a resistor bank connected to the test rig. The DC motor field current is controlled by a DC drive connected to two supply phases that are organized to provide a current which is pre-set by a potentiometer. This provides a wide range of test rig load conditions (from 0 to 100% load) to be applied to the AC motor.

6.7.3 Speed control

Figure 6.8 illustrates the layout of the test rig. The inverter is connected to the mains (the three-phase 415 VAC) supply to attain speed control. Due to the speed set point, the inverter changes the supply frequency in order to generate new supply voltages for the AC motor. Then a new voltage is applied to the AC motor and the system will operate at a required speed.

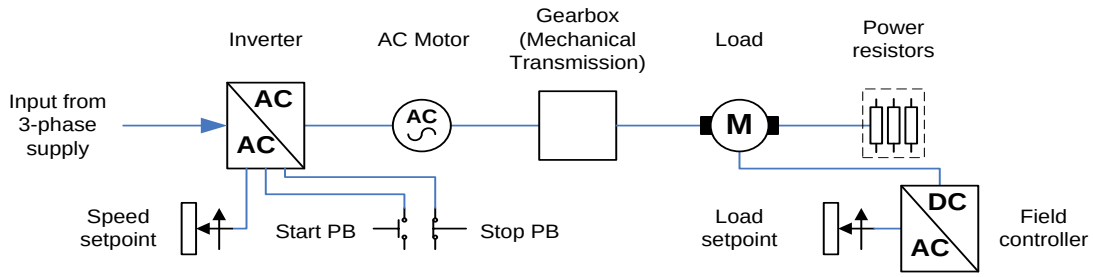

Figure 6.8 Block diagram of the experimental test rig

6.7.4 Test Rig Measurement Devices

A photo and diagram of the three-phase current measurement unit are presented in Figure 6.9 and Figure 6.10 respectively. The motor current in each phase was measured using a Printed circuit board (PCB) mounted Hall Effect Current Transformer (CT). The measured value for

the current in each line is fed into the data acquisition system (DAS) unit, which converts this into a voltage measurement, provides appropriate filtering and anti-aliasing, and feeds the signals to the data collection channel of a data collector/analyser. Thus this device can be used to measure the instantaneous current in each of the three phases, the instantaneous voltage of each of the three phases and the instantaneous electrical power supplied by each of the three phases.

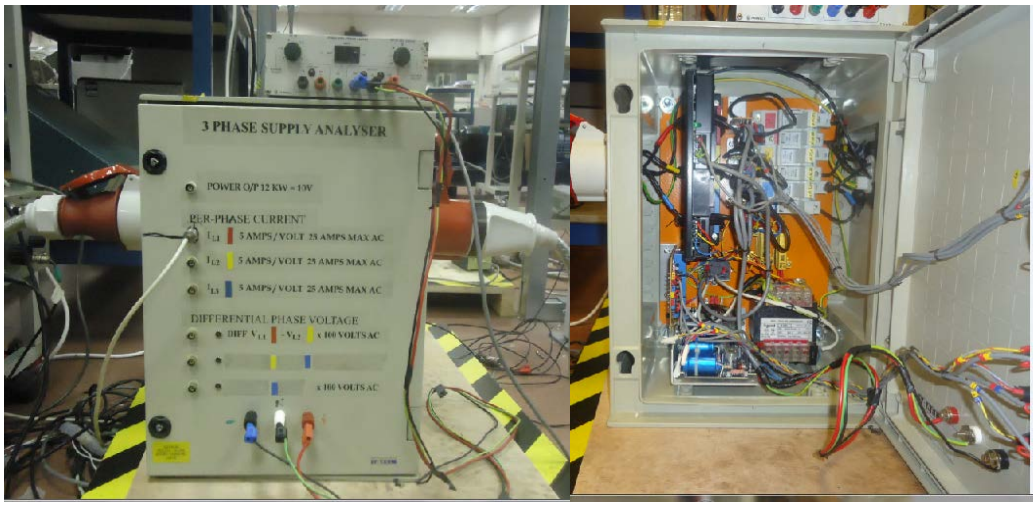

Figure 6.9 Three-phase current measuring unit

Table 6.3 Specification of three-phase measurement unit

Classification	Specification
Phase current	Max. voltage reading: 12.8V at 5A/Volt
Line to Line voltage	Max. voltage reading: 12.8V at 100V/Volt
Total power	Max. power reading 10 VDC at 2kw/Volt

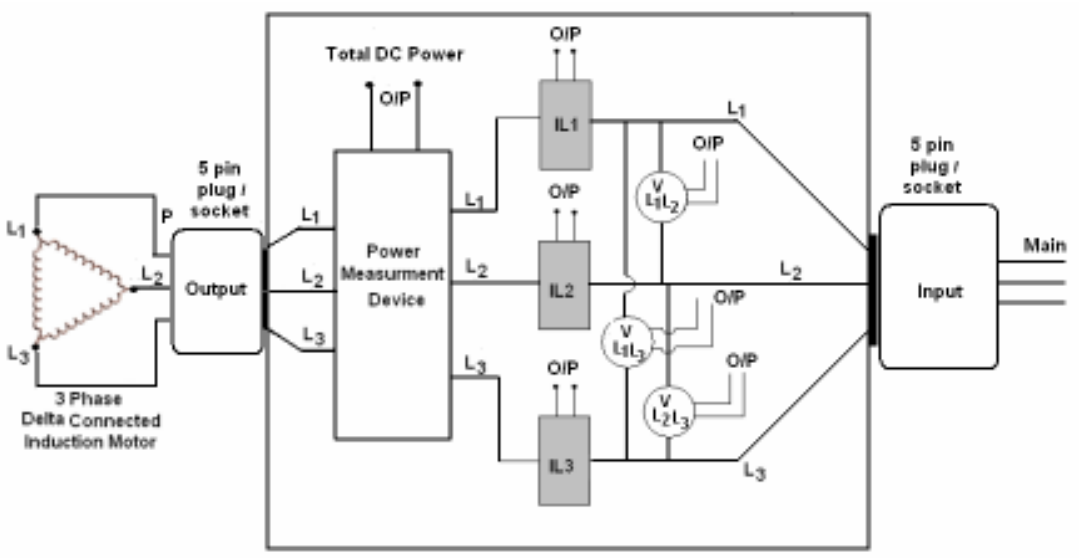

Figure 6.10 Diagram of three phase measurement unit

Where:

VL1L2, VL2L3 and VL1L3 are line to line voltages,

IL1, IL2 and IL3 are the three line currents,

Ⓥ = Hall Effect Voltage Transducer

▢ = Hall Effect Current Transducer

6.7.5 Speed Measurements

An incremental optical encoder (Manufacturer: HENGSTLER; Type: RS-32-0/100ER .11KB, 30VDC=/0.1A) fitted at the end of the AC shaft as seen Figure 6.11 has been used to calculate the Instantaneous Angular Speed (IAS) and shaft phases in order to calibrate the current signal and implement TSA. The optical encoder output is connected to the data acquisition system (DAS), which is connected directly to a host computer. The encoder produces a square pulse output, being 100 pulses per revolution and allowing IAS to be obtained up to 50 orders of the shaft frequency. This is sufficiently high for the test system.

In addition, it also ensures that the signal average produced is synchronised with a known shaft position, and that the location of the broken tooth can be determined [117].

Figure 6.11 Mounting of motor encoder

6.7.6 Temperature Measurement

Measurement of the oil temperature inside the gearbox was obtained by installing thermocouples inside the gearbox casing through a hole drilled in the oil drain tab as shown in Figure 6.12. These were K-type thermocouples with a linear response from –20^{o}C to 220^{o}C. The main idea behind gearbox temperature monitoring was not only for safety considerations but also to keep the gearbox temperature in the same range during all test steps. This avoids any change in gearbox vibration due to increase/decrease in its temperature. In addition, the gearbox temperature was measured to ensure the safe operation of other measurement sensors.

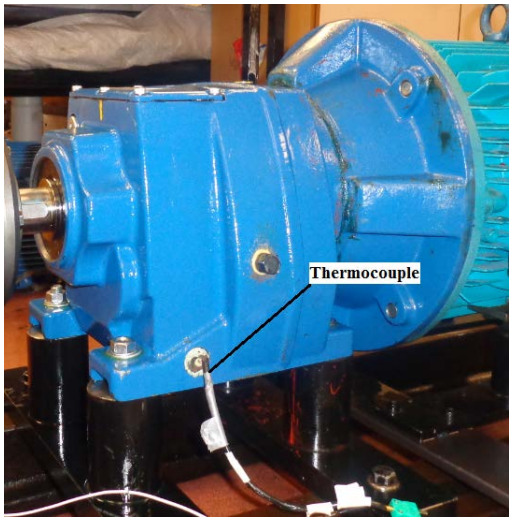

Figure 6.12 Thermocouple sensor placement

6.7.7 Data Measurement and Acquisition

Having converted the collected electric current by current transducers into low voltage signals using the analyser, the signals need to be recorded on to a hard disk for further processing by a DAS, which consists of a PD2-MF-16-500/16L Analogue to Digital converter (ADC), the electrical power analyser, a speed measurement device, two channels for vibration sensors, and one channel for the temperature device. In addition, the DAS runs a computer program to control the data acquisition process.

The data acquisition card used in this research has the major specifications shown in Table 6.4. It has 16 input channels which allow the electrical current, speed, vibration and temperature to be measured simultaneously for more accurate analysis. Moreover, it has a 16 bit data resolution which is equivalent to a dynamic range as high as 90dB and allows weak signals such as the modulated components to be measured with high accuracy.

The program of the data acquisition control is developed on Lab-windows and consists of a main data check panel and a parameters panel. It has the ability to execute on-line data checking. Required settings such as the number of channels, data length, sampling frequency and filenames can be collected and recorded on a separate set-up page of the software.

Table 6.4 Main specification of data acquisition system

Parameter	Performance
Number of channels	16 differential, 16 single ended
Data resolution	16 bits
Sampling rate (maximum)	500 k samples/s for 16 channel
Maximum input voltage	±10 V

The experimental tests were conducted on a test rig system, where 1.6 x 10^6 data points were collected at a sampling rate of 100 kHz, so a frequency resolution of 0.000625Hz was obtained. The data points were collected and recorded over 20 seconds. The raw data were measured and recorded six times, three for healthy conditions and three for faulty conditions, and saved in separate files at the same operating conditions.

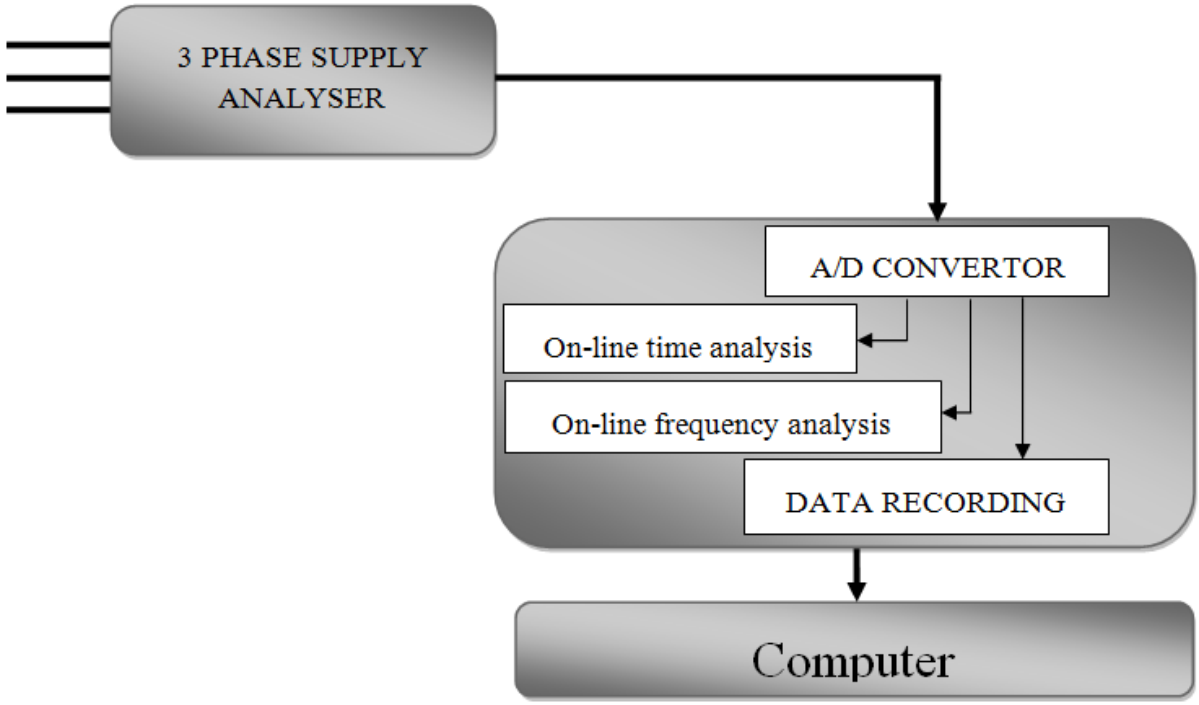

Figure 6.13 Schematic layout of data acquisition system

The data was anti-alias filtered before passing through an A/D converter, including the optic triggering signals. Although the dynamic signal analyser was capable of calculating the FFT and simple statistical parameters of the signals, it was only used for data acquisition and recording. More complicated signal processing techniques were carried out in MATLAB. A block diagram in Figure 6.13 shows a schematic diagram of the instrumentation connections.

6.7.8 Software: Lab Windows

The software of the DAS is improved using National Instruments Lab Windows TM/CVI Version 5.5. A programme in C language was used to write the interface environment [118] because it has a large set of run-time libraries for analysis, data acquisition and. instrument control. The Lab Windows development environment offers such features as automatic code generation, making measurement tasks much easier than in traditional C or C++ environments. The data acquisition software can control the data acquisition process so that multiple channels of dynamic data can be acquired at different rates and data lengths. It can be used to acquire current, vibration, instantaneous angular speed (IAS), noise, and temperature measurements. The software package has an acquisition set-up panel so that the user can make modifications to e.g. the sampling frequency and sample data length. The control commands and status indicators are listed on the screen. This will ensure an optimal data set is collected for off-line data analysis.

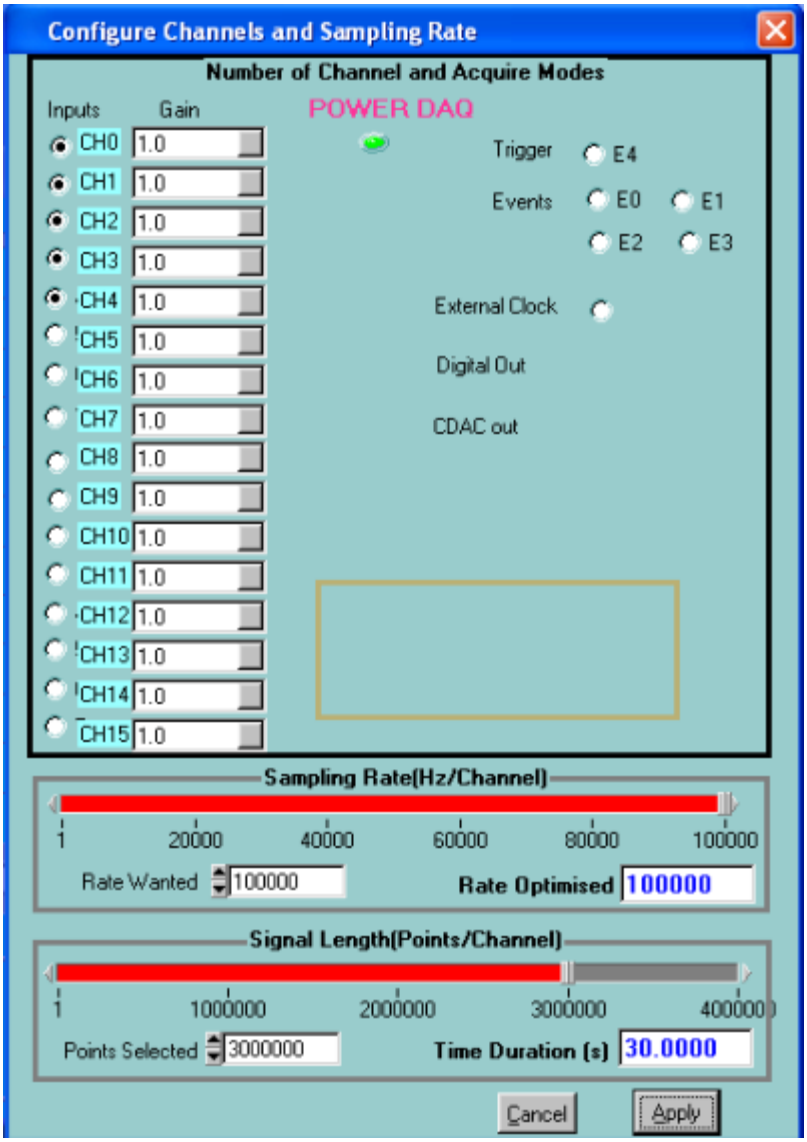

Figure 6.14 Set-up screen of data acquisition software

Figure 6.14 shows the set-up screen: here the sampling frequency is set at 100 kHz (this enables the high frequencies associated with transient events to be collected) and the data length is set at 300,000 samples. The time duration of the data point collection = number of samples ÷ sampling frequency, so the real time duration of the data record is 30.0 seconds.

The panel created to display the data collected for gearbox monitoring. It included several scopes that displayed sensor signals and digital indicators that displayed numerical values of

the monitored parameters. When the test rig starts operation the DAS starts data collection so that each pulse is always measured at exactly the same input shaft position to ensure accurate time domain averaging. Data files were saved in binary format to save disk space for further analysis in order to obtain the results and the Figures in the later chapters of this research work in the MATLAB environment

6.7.9 Signal Channels and Screenshots of Data Acquisition

The data acquisition system can be configured to take a maximum of 16 channels, of which a total of five were actually used for this gear condition monitoring experiment. A PC card is connected to a spare PCI (Peripheral Card Interface) slot in the computer. This configuration is shown in Figure 6.16 below:

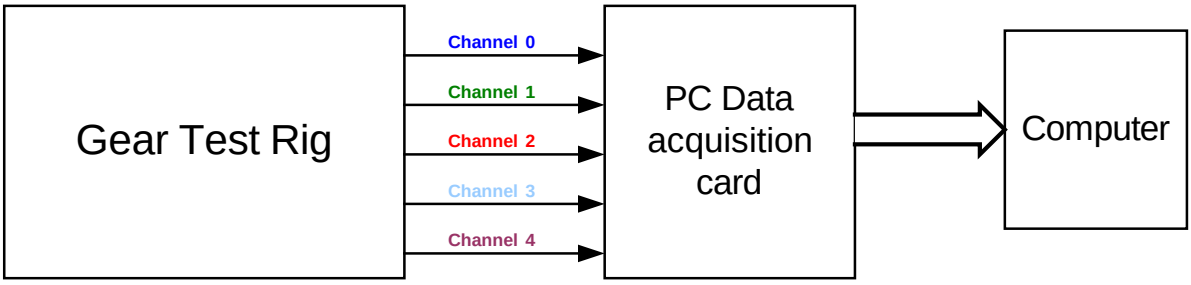

Figure 6.15 Diagram of measured signals

Channel 0 – Blue: signal from shaft encoder (1 pulse / revolution) to determine average speed calculation

Channel 1 – Dark Green: Signal from shaft encoder (100 pulses / revolution) for Instantaneous Angular Speed

Channel 2 – Bright Red: Motor current

Channel 3 – Yellow: Vibration signal from gearbox

Channel 4 – Dark magenta: Vibration signal from motor flange.

These five signals are measured from the test rig and sent to the data acquisition system which is connected to a software package allowing the display of virtual instruments on the computer screen. The data acquisition system display is shown in Figure 6.16:

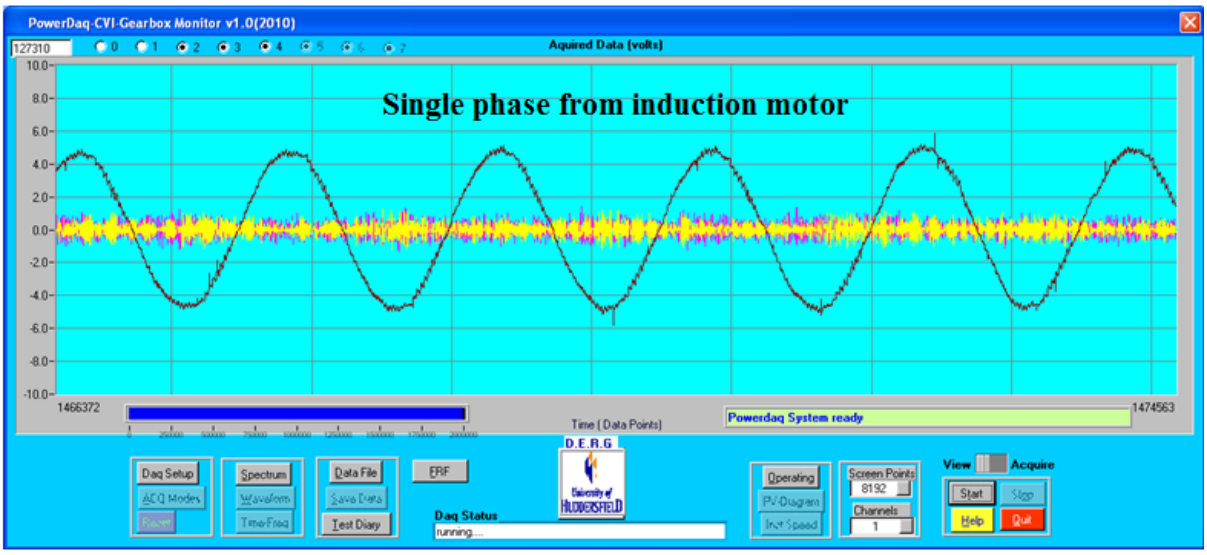

Figure 6.16 Screenshot of measured signals

Also, the software allows the display of operating conditions through a real-time ‘dashboard’ shown in Figure 6.17, allowing basic status to be monitored and taking correcting actions if any abnormality found in these parameters:

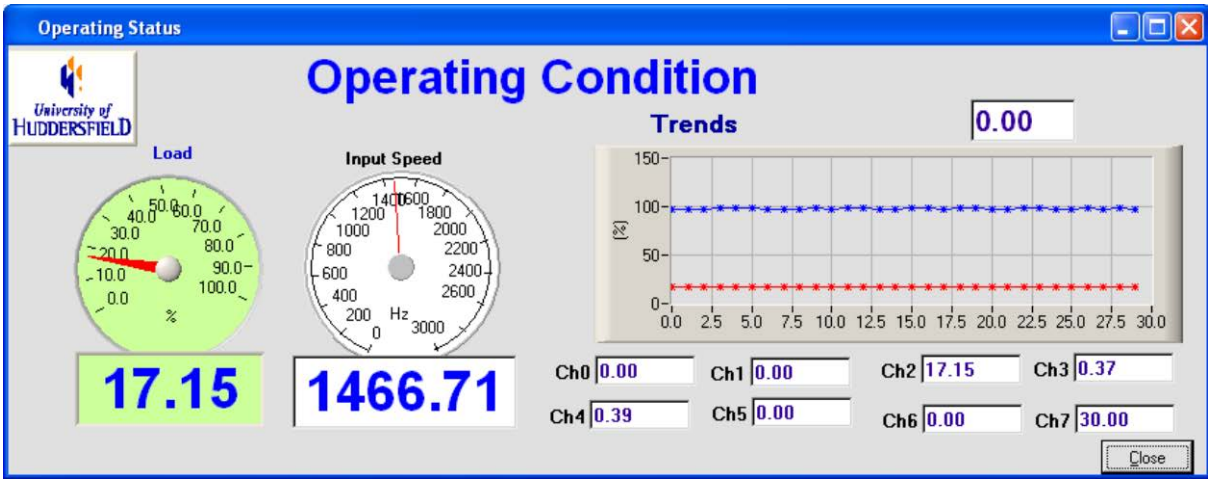

Figure 6.17 Screenshot of operating status through 'dashboard'

6.8 Vibration Measurement (Accelerometers)

An accelerometer was attached to the casing of the gearbox to collect the vibration data, the piezoelectric crystal generates a small charge when a small acceleration is applied which is proportional to the size of the acceleration. The type of accelerometer used to collect the data in this thesis is an integrated circuit piezoelectric (ICP) accelerometer, specifically the model 338C04 manufactured by PZB Piezotronics as showing in Figure 6.18 below.

Figure 6.18 PZB 338C04 ICP accelerometer

The accelerometers has a sensitivity of 10.2mV/ms^{-2}, with a measurement range of up to ±490.5 ms^{-2} peak-to-peak. The accelerometers are directly attached the side of the gearbox itself and at the top of the gearbox. The accelerometers are connected to the casing using a screw-threaded brass stud, ceramic cement is also used to prevent any over-heating.

6.9 Fault Types of the Testing Gearbox

In this thesis, two types of fault were examined: a gear wear produced from a prolonged operation of the gearbox and a gear tooth breakage with two percentages (50%, 100%) of tooth breakage (artificially induced) in the drive and driven gear at the first stage of the helical gearbox.

Gear fatigue failure is an uniformly distributed gear fault, it can be broadly splits into two major categories: bending fatigue and surface contact fatigue bending is created by the expansion of a crack near the root of a teeth leading to partial or a complete tooth fracture [119][120]. Gear teeth surface contact fatigue (wear) is considered to be one of the most common causes of gear failures. This is because surfaces that are constantly exposed to rolling and sliding contact under varying load conditions results in local Hertzian contact, as a result, the gear will often suffer from contact fatigue [121]. The fault can be further categorised into two main types: pitting and spalling. Pitting arises with meshing gears along the contact region as a result of a subsurface crack. On the other hand spalling is associated with the existence of deep craters on the tooth surface [122][123]. Spalling and pitting are generally the most common causes of fatigue failure in gears [124][125].

All the studies reviewed so far have confirmed the fact that the uniform wear of gear teeth is another common kind of gear failure [126]. Many gears, especially those that are installed in industrial machinery, are commonly subjected to continuous prolonged running with a heavy load. Under these circumstances their pitting durability is decreased, their bending strength and also their operating life are also affected. The evaluation of gear teeth conditions after a definite period of their utilization has great importance in avoiding any unexpected and severe failures. As a consequence of gear teeth wear their shape is altering continuously and this causes the alteration of gear ratios -- leading to angular acceleration and an additional dynamic load in the region of gear teeth contact and further increases in the gear teeth wear level. This process is a continuous one with various levels of severity during the whole period of utilisation.

In recent years, there have been a signification amount of publications on gear wear failures. Ma and Li found that up to 65% of gearbox failures are caused by gears faults, and

approximately all gear faults are caused by fatigue induced fractures [126]. Further Lin and Zuo have demonstrated that tooth breakage is the most serious gear defect because it will lead to the completely failure of the gearbox [127]. Initially a fatigue crack developing at a particular tooth is not considered to be a serious fault, but as it accelerates, further consequences such as tooth failure can appear, which will often prove catastrophic to the system. In fact, fatigue will often result from continuous gear working and fluctuation of load and speed application conditions, poor construction, installation and maintenance will often exasperate the problem further.

6.9.1 Localised Gearbox Faults

Localised gear faults such as tooth breakages are a very common occurrence in gear transmission because of high bending loads on gear tooth which cause high stress on gear roots. This type of fault starts with individual tooth which may have manufacturing or material defects. Moreover, this is the ultimate type of gear failure that could cause high negative consequences. Therefore, it has been studied extensively in CM. Because of the high overlap ratio of the helical gears, this type of fault only has a small impact on the overall power transmission of the gearbox and gear dynamics. Early studies relating to tooth breakage in gearboxes use conventional techniques such as vibration signal analysis to diagnose the faults, however such techniques have often proved unreliable to be in detecting the presence and severity of tooth breakage faults [128].

The two localised faults were artificially introduced firstly to the drive gear (Z_1) and then to the driven gear (Z_2) as well, both faults were introduced only to the first stage in the gearbox. And both faults were tested independently. In total 3 sets of gears were used for local tooth fault data collection, one where there were no faults, one set where there was a

tooth breakage fault in the driven gear and another set where the fault was on the driving gear. The test data used in this thesis are from seeded faults of tooth breakages of two scenarios: faults of 50% and 100% for the drive gear (Z_1), and 30%, 60% and 100% for driven gear (Z_2) , These allow for rigorous evaluation in which more advanced techniques can better perform the diagnosis. Especially, the driven gear is seeded with a small fault (30%). This can evaluate the detection sensitivity and diagnosis accuracy of TSA, CB and MSB in differentiating incipient faults.

6.9.2 Fatigued Gearbox Faults

In order to investigate the consequences of the prolonged operation of the gearbox a long-term fatigue test was conducted on the gearbox.

The gears which were used in this thesis are helical gears that relatively wide, they also have a high average number of teeth, therefore contact up to 4 teeth are in contact with one another during the meshing cycle. Since it would be impractical to conduct a long enough run time for the gear to indicate signs of fatigue wear in the lab, the process needs to be accelerated. As seen in Figure 6.19, the width of the gear being test (Z_1), was reduced by half (50%) before it was ran to simulate fatigue wear. Then gear (Z_1) was installed in the gearbox system as it appears in the schematic diagram Figure 6.20, which shows how gear (Z_1) meshes with gear (Z_2) in the first stage.

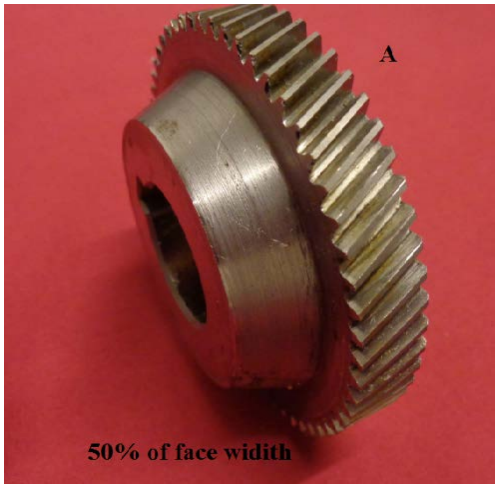

Figure 6.19 - Simulated fault on Z1

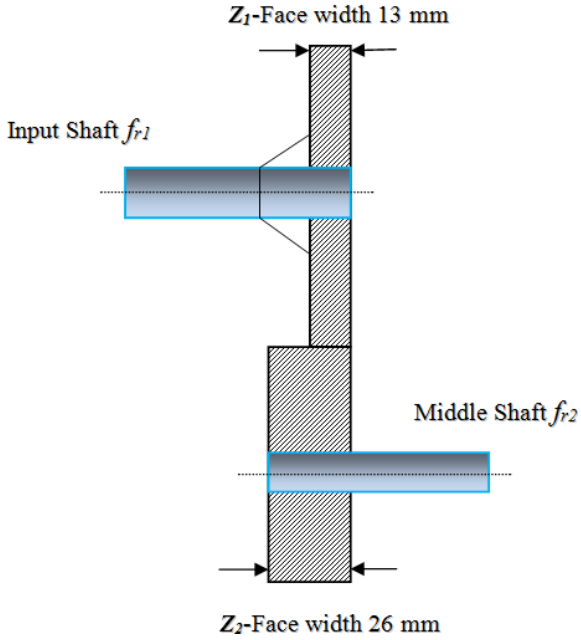

Figure 6.20 - The first gear (Z_1) with the simulated fault mesh with the second figure the first stage

Before beginning the test, the gearbox was run for 4 hours at 95% of its full load (19A), this allowed the gearbox warm up to reach its operational temperature of 55°C.

The 50% reduction in the width of the gear has meant that the tooth overlaps has also reduced by 50%, therefore the force applied to the tooth is now doubled, this will result in

shortening the service life of the gear due to higher contact and bending stresses. The gearbox vibration signals were visually monitored on a daily basis while the signals were recorded at approximately one hour intervals and continuously analysed using RMS and kurtosis techniques, this was done to avoid any catastrophic breakdowns that might occur. The load was also continuously varied to simulate applications in real industries such as automotive transmissions, centrifugal pumps and steel rolling machines.

However, the test was terminated after 470 hours, due to a significant increase in monitored electrical current amplitude, vibration monitoring signals as well as a high noise level, indicating there was clearly an abnormality occurring in the system. The gearbox was dismantled and the presence of a distributed wear fault could be seen on the pinion teeth, illustrated in Figure 6.21 and Figure 6.22 in which the tooth surfaces have clear fatigue pitting dents.

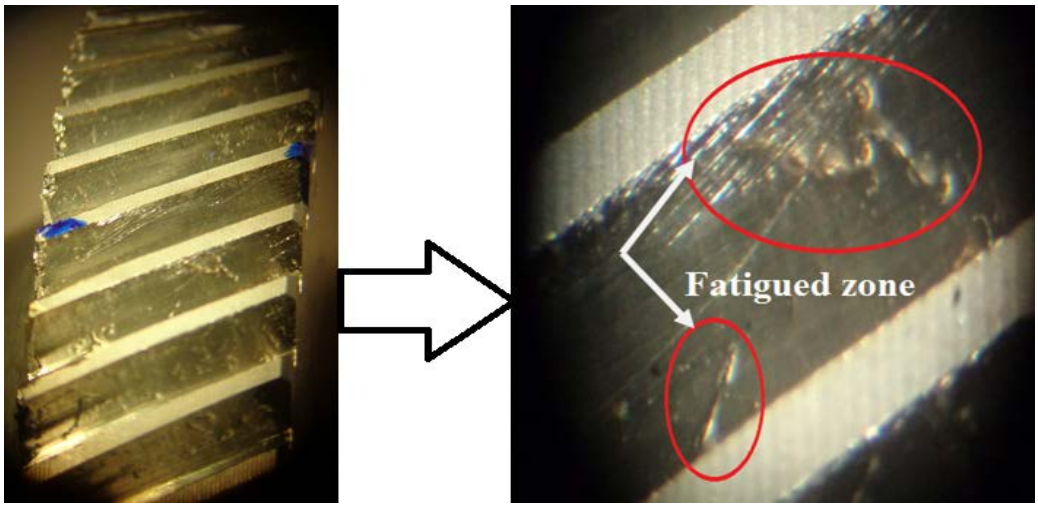

Figure 6.21 Pitted areas after 470 hours of operation

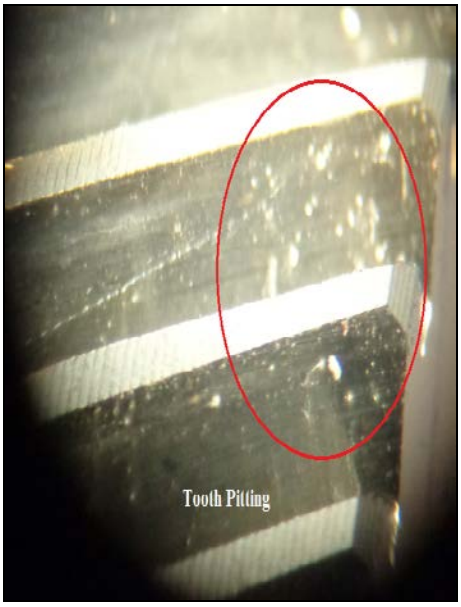

Figure 6.22 Zoomed pitted area for tested gear after 470 hours of operation

6.10 Chapter Summary

The experimental rig that was used is made up of 3 key components. Firstly, an 11kW AC motor provides drive to the system, its output was 1465rpm. Two gearboxes were used as part of the overall setup, only the first gearbox had faults introduced to it and was monitored. Finally, a DC generator was used to act as a load for the system. There is also an extensive control unit in place to provide continuous monitoring ensuring that the experiment is running effectively without any problems. This is to provide accurate and reliable data with minimal anomalies.

Two types of faults were added to the gearbox separately. Firstly, a local fault was created, this was done by wearing down a tooth on the gear. A 50% and 100% tooth damages were created sequentially. Because of the high contact ratios of the testing gear, a local tooth damage does not have a significant impact on the transmission performances, therefore, the only a significant local fault is detectable.

Another common cause of the failure of gearboxes is fatigue wear of a particular gear under prolonged operation, this is a non-localised fault and the damage to the gear is distributed evenly. Because it would require running the test rig for a significant amount of time to simulate the gear damage, the width of the gear was reduced by 50% to speed up this process. The gear used for the fatigue fault was in operation for 470 hours at full load in which datasets were taken at every 50 minutes.

These two datasets will be utilised for evaluating TSA, CB and MSB analysis in the following chapters.

CHAPTER 7.

LOCAL GEAR FAULT DETECTION AND DAIGNOSIS

7.1 Introduction

To verify MSB based diagnosis approaches proposed the first two experimental datasets are employed. These two sets of gears depicted in Chapter 6 are from the same type of gear defect modes-local faults, but with different fault severity and locations, and as well as different gear error characteristics. Each of them was induced manually with local tooth breakages of different severity. The fault for first set has the fault induced to the pinion gear on the first shaft, referred as to Gear-11 (11 is one of 20 gears tested in one experiment) or Z1 alternatively as necessary, whereas the second has the fault on the driven gear which is mounted on shaft 2, denoted as Gear-D (D is one of 8 gears tested in another experiment) or Z2. These fault scenarios allow the diagnosis performance to be evaluated not only on the fault severities but also the fault locations. Obviously these details are critical for take maintenance actions accurately. The notations for gear sets are due to experimental convenience in testing multiple gear sets in the rig.

Moreover, as the two sets of gears were obtained directly from supplier at different time periods, they will have slightly different manufacturing accuracy. Therefore, it helps to verify the robustness of upon the error influences and develop more effective diagnosis scheme based MSB. Nevertheless, during the test the re-installation of testing gears after creating different degrees of the tooth defects will also bring errors due to changes in different fitting clearances between the shafts and gears.

The evaluation starts with investigating the characteristics of MSB results. Then shows how a more reliable MSB based MPI or Z-square value is calculated to achieve accurate diagnostic results. Subsequently, CB and TSA results are also presented for each set of the gears to make comparisons between different methods in terms of information provisions and diagnostic performance and implementation requirements.

7.2 Representative Power Spectrum of Vibration Responses

Before applying MSB analysis, a typical power spectrum analysis is performed to show general properties of vibration content from the testing of gearbox acceleration signals. As shown in Figure 7.1, where the frequency range was examined up to 1500Hz, being less than (f_{me12}+5×f_{r2}) as depicted in Table 6.1, the baseline spectrum have many observable components including the meshing components of the first meshing component f_{me12} from the higher speed stage, or the first stage, and three components from the second stage or lower speed stage . In addition, the baseline also has sidebands at two high frequency meshing components: f_{me12} = 1416.36 Hz and 3× f_{me34}=3×391.76Hz, showing the effect of manufacturing errors. Especially since the sidebands are asymmetric, they indicate the co-existence of AM and PM and their interference effects [129][130]. Therefore, it is not possible to make an accurate diagnosis of the fault severity by using these components directly.

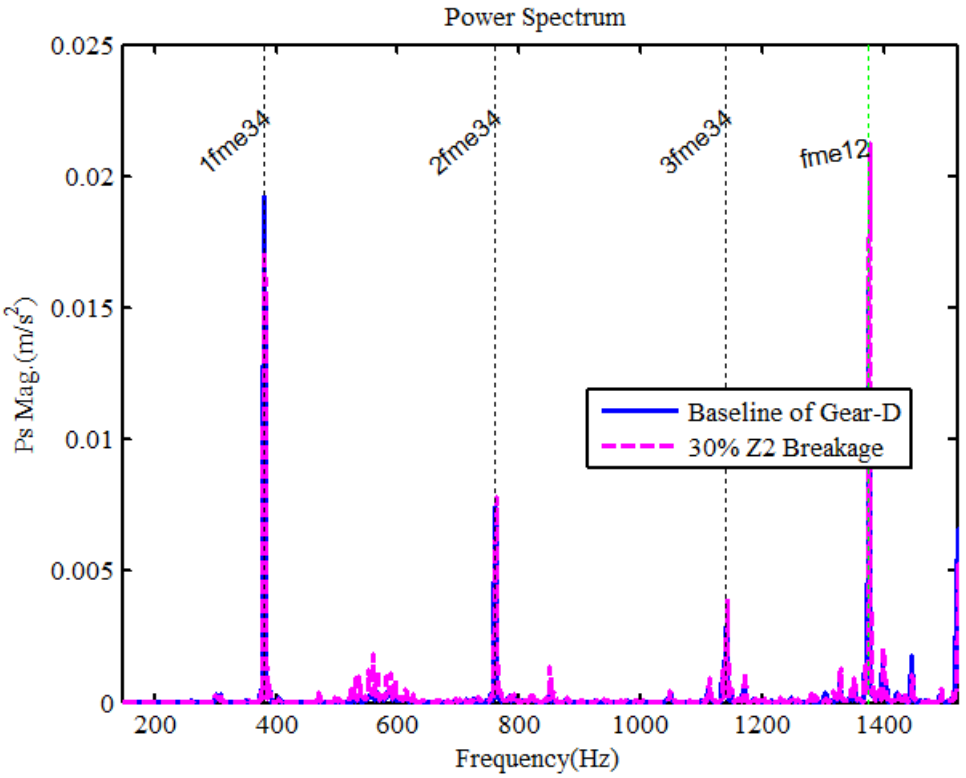

Figure 7.1– Power spectrum of Gear-D for baseline and small fault case

However, the fault causes a clear increase of spectrum amplitudes in the frequency band around 590Hz. This band is far away from the meshing frequency of f_{me12} =1416.36Hz at which the fault should exhibit more sidebands. Clearly, this changing pattern in spectrum is not consistent with that common knowledge of diagnosing the local faults based on spectral changes around meshing frequencies, which is published previously in numerous papers [131][132].

Nevertheless, the spectrum shows agreeable features as discussed in the contents of vibration signal models and fault diagnosis theories development in Chapter 4 and 5 respectively.

7.3 MSB Characteristics of Tooth Breakage on Drive Gear

As the length of each data record is 3,000,000 samples and the sampling rate is 100kHz. To achieve a frequency resolution of 1.5Hz for the spectrum analysis, allowing each possible

sidebands at about 24Hz, 30Hz and 6Hz (varying slightly due to speed accuracy achiable by the speed controller) to be seperated, FFT frame is 65536 points. With 50% data fame overlapping between different frames of FFT calculations, the data length allows 90 averages for MSB and spectrum calculation. Obviously, the greater the number of averaging, the lower the noise level obtained. Howerer, more averages need more calculational consumption, which is a drawback for online monitoring. In this study, it has tested that 50 averages would produce a result that is close to that presented in the thesis. However, for higher robustness and more averages required by other methods, 90 averages achiveable are adopted in calculating CB and MSB.

In addition, CB results are presented for its principal frequency ranges and MSB is in the range for postive f_m reanges to avoid any information reduntances. Magnitudes of MSB have been root-squared to promote graph displays of small components in MSB results, which also can be matched approximately with those of power spectrum in order to make direct comparisons between the two. The characterstic frequencies of gear vibrations are particularly labelled on the bifrequency plane for highliting MSB contents of interest. It is worth noting that these frquency values can be different slightly due to errors in opearting speed control. In addition, the MSB phases are adjusted to show only the values with coherences higher than 0.1 to show only significant components with cleaner graph presentations.

7.3.1 Baseline of the Tooth Breakage on Z1

Figure 7.2 shows the MSB magnitude results for the tooth breakages on gear Z1 in different frequency bands, which are divided according to the carrier frequency of f_c. For the baseline case shown in Figure 7.2(a), it can be observed that distinctive peaks present in four incrementing bands i.e. f_c slices which are at or around 381.9Hz, 763.7Hz, 1146Hz, and

1381Hz, the first four harmonics of f_{me34}. The first three frequencies are the meshing component of the second stage and its 2nd and 3rd harmonics, whereas the fourth one is the meshing frequency of the first stage transmission. These shows that the meshing content are the main vibration responses, being agreeable with theoretical prediction in Chapter 4.

In addition, there is one more significant peak at 590Hz. It may indicate the occurrence of a resonance.

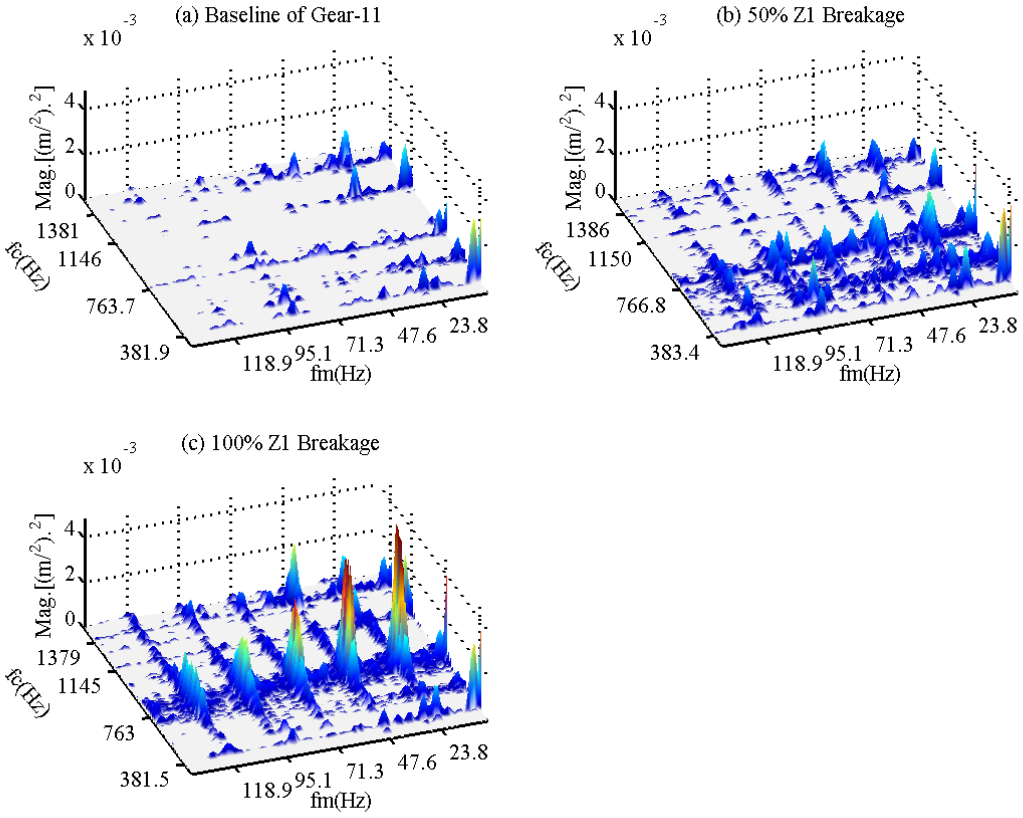

Figure 7.2 MSB magnitude results for Gear 11 under different fault severities

(in which gear characteristic frequencies are labeled with $f_c=1f_{me34}$, $2f_{me34}$, $3f_{me34}$ *and* $1f_{me12}$ and $f_m=1f_{r2},...4f_{r2}$)

However, among these four meshing frequency slices at 381.9Hz, 763.7Hz, 1146Hz, and 1381Hz, only the first meshing component at 381Hz and 1381Hz for the first stage and

second stage respectively can be ensured to be significant because of their corresponding coherence values are correspondingly high as shown in Figure 7.3 (b). These high coherence peaks convince more that the signal at these components are strong and hence can be relied on to evaluate possible modulation effect due to manufacturing errors. Together with MSB magnitude results, it can be seen that the baseline of this gear set has high levels of errors on the first gear or Z1 because the clear peak at bifrequencies (23.8, 1381) Hz and (2×23.8, 1381) Hz. In addition, as the error is high it also modulates the third meshing harmonics of the second stage, leading to the MSB peak at bifrequency (23.8, 1146.0) Hz.

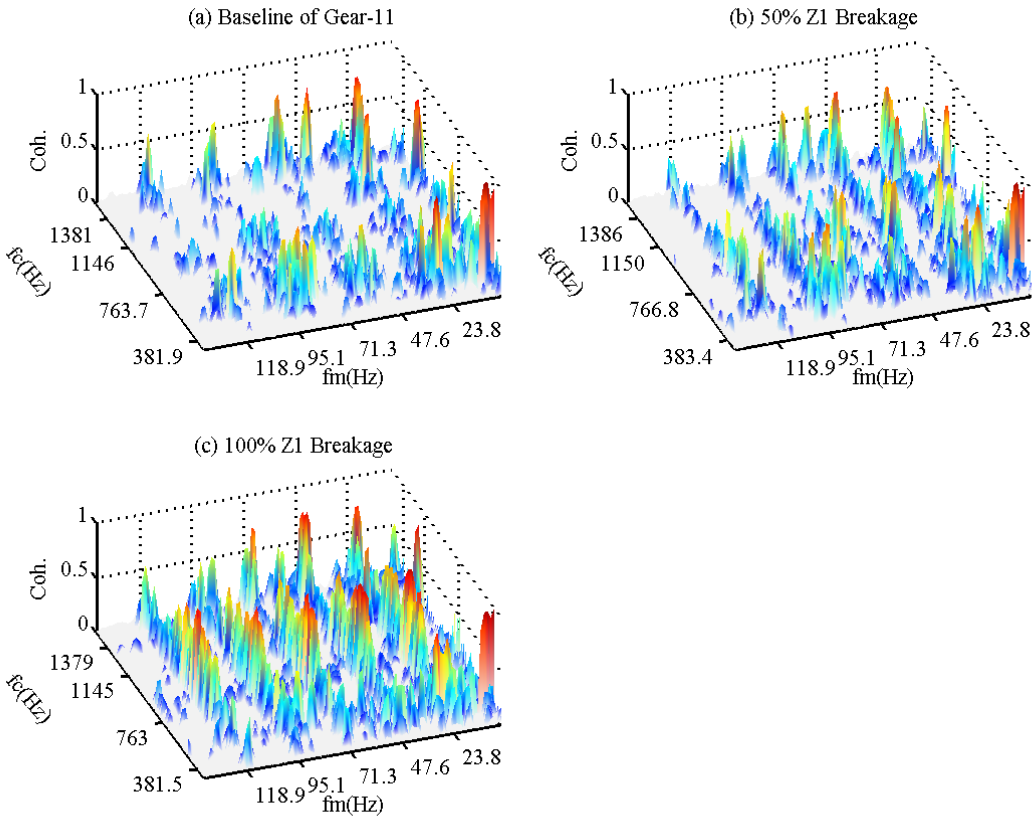

Figure 7.3 MSB coherence results for Gear 11 under different fault severities

In the meantime, baseline phase results in Figure 7.4 (a) have shown that these MSB peaks with high coherence values in Figure 7.3(a) have scattered widely between 0 and π. This shows that the modulations are more likely to be combined between AM and PM and indicates that MSB magnitude peaks are not very accurate to represent the modulation levels. Therefore, it needs to further demodulation in order to obtain MPI.

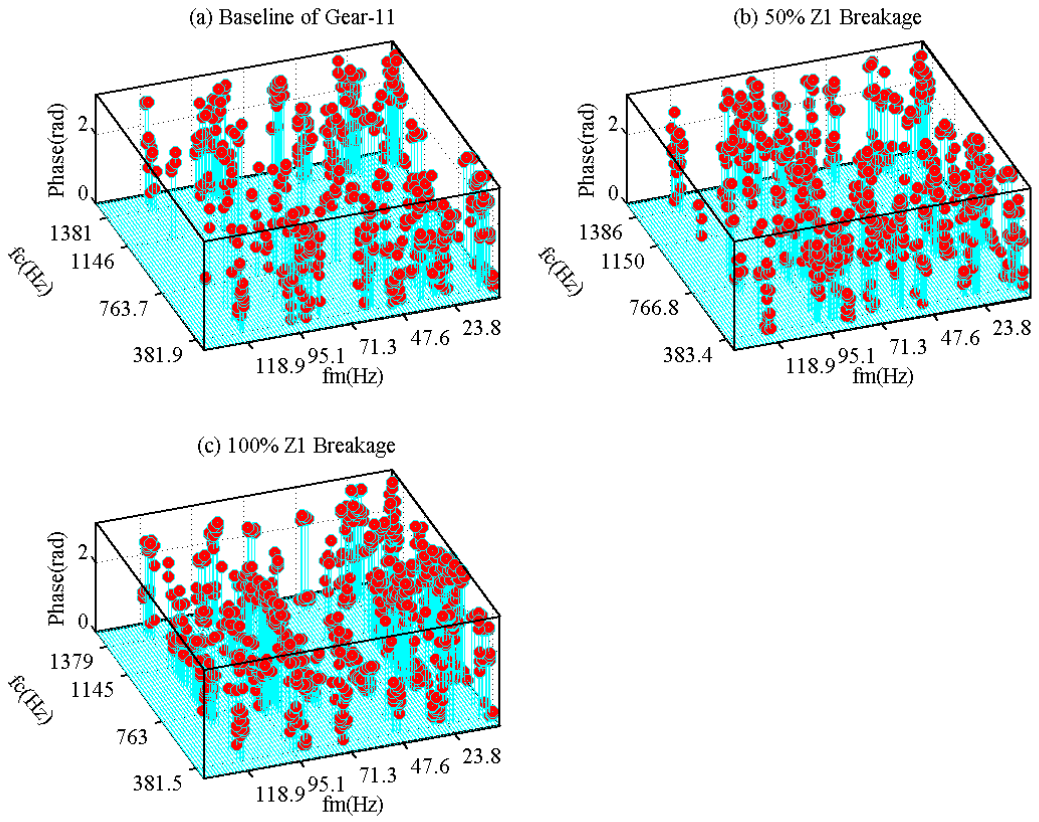

Figure 7.4 MSB phase results for Gear-11 under different fault severities

7.3.2 Fault Detection and Diagnosis of Tooth Breakage on Z1

Although the occurrence of AM-PM interferences indicated by phase results, both MSB magnitude and coherence can provide sufficient information for fault severity and location diagnosis. Comparing MSB magnitude results between the baseline and the two fault cases, it is obvious that there are many MSB slices associating with different carrier f_c that show

consistently increasing peaks with fault severity and hence they can indicate the fault severity. Moreover, these peaks appear at bifrequencies relating to 23.8Hz and its higher harmonics. This indicates that the fault occurs on the first gear. Therefore, the faults can be differentiated by using both MSB magnitude and coherence without any uncertainties. In addition, this also demonstrates that the frequency bandwidth up to $f_{me12}+5\times f_{r2}$ is sufficiently wide to cover full diagnostic information.

It is worthy note that the peak at bifrequency (23.8, 1183) Hz shows lower magnitude for the 50% tooth breakage, compared with that of baseline. This could lead to incorrect detection result if this meshing component is used for diagnosis. However, common understandings are that the component coupled with the tooth meshing 1183Hz and can have good SNR and hence accurate diagnosis. Obviously this is not true for this test case. Instead the peaks can be reduced by both the out-phase modulations or interference between AM and PM as pointed out in the studies made in Chapter 4 and 5. This out-phase modulations can be from the fault effect or the errors as a result of reinstallation after the fault was created offline. This will be clarified in the section addressing accurate diagnostics.

In addition, the significant peaks of MSB magnitudes in the frequency ranges appearing around f_c=766Hz and f_c =550Hz in Figure 2 (b) and (c) can be diagnosed as a resonance of the system. This is because that of impulsive excitation of the local tooth breakage fault can result in wideband components that cause higher responses in the resonance range. Especially, the 766Hz is the second harmonics of fm34 and very distinctive. It can indicate that there is a significant resonance at this band, being likely the one of the modes from the second shaft which needs to be confirmed by further studies. Nevertheless, this resonance magnification helps to obtain a high SNR for detection and diagnosis. Moreover, when the occurrence of a resonance, either AM or PM will be more dominant, i.e. one of them is much

larger than the other, leading to interference between AM and PM insignificance. Therefore, the diagnosis can also be more consistent.

Overall, MSB analysis can provide plenty of accurate components for obtaining reliable performance of both localising the fault and quantifying fault severity. However, it remains question that which of them are optimal. Obviously the solution will lead to more success to detect and diagnose even smaller fault.

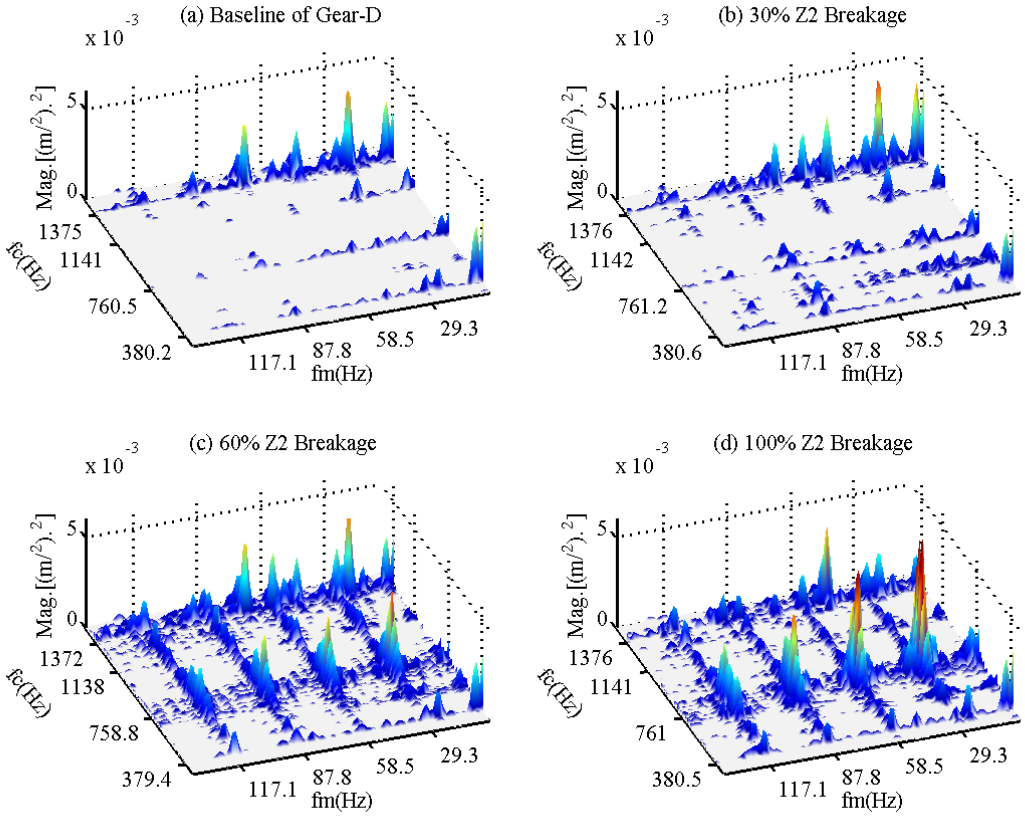

Figure 7.5 MSB magnitude results for Gear-D under different fault severities

7.3.3 Fault Detection and Diagnosis of Tooth Breakage on Z2

MSB results for the gear on the second shaft shows more modulation for the baseline case, compared with that of Z1 tooth breakage. As shown in the magnitude and coherence results in Figures 7.5(a) and 7.6(a) respectively the bifrequencies of (29.3, 1375) Hz, (2×29.3, 1375) Hz, and (3×29.3, 1375) Hz all have distinctive peaks. Moreover, they have higher

magnitudes than that of Z1 tooth breakage. This indicates that this gear set have higher errors which can influence the accuracy of fault diagnosis. In fact, if a diagnosis is based on the meshing component slice at $f_c = f_{me12} = 1375$Hz, the results are not consistent with the fault severities induced, which are clearly shown by the irregular variation of the magnitudes between different fault cases for this gear. Therefore, it is further proof that the meshing components are not very suitable for the purposes of fault detection and diagnosis because of the error influences.

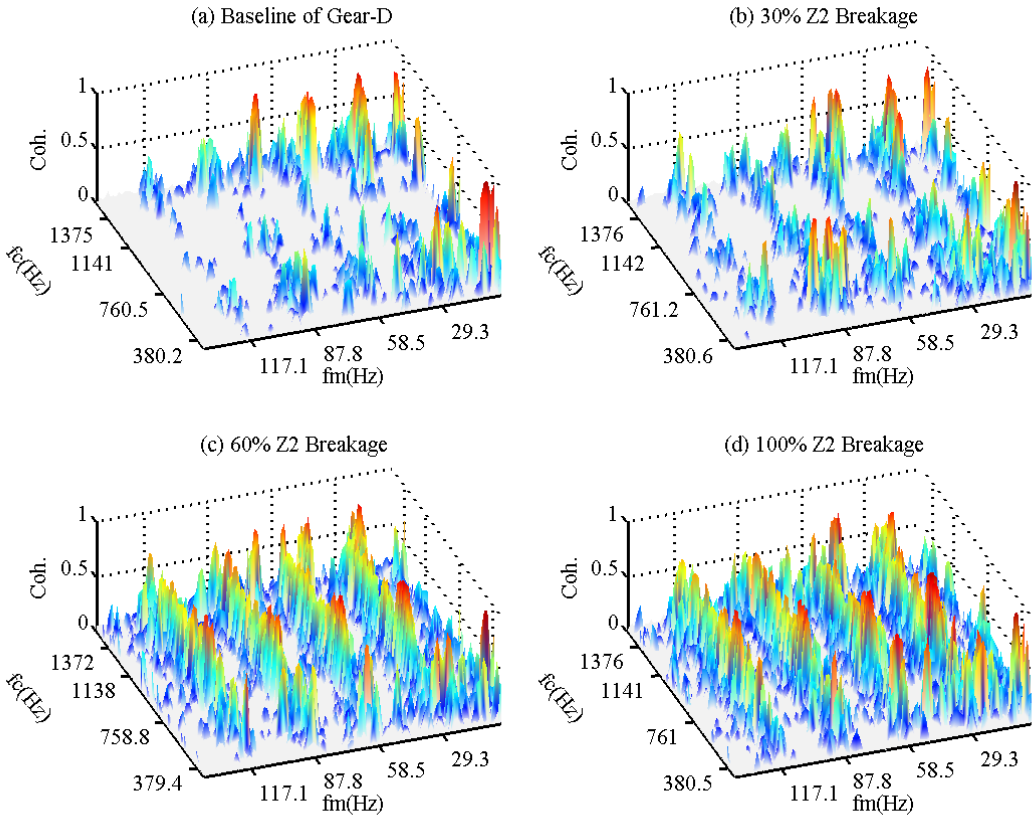

Figure 7.6 MSB coherence results for Gear-D under different fault severities

Apart from the meshing component at f_{me12}, the other bands can show similar increasing changes to Z1 breakages in that the magnitudes and coherence increase consistently with

fault severity. In particular, the smaller fault case of 30% tooth breakage can be diagnosed by the clear increase in many MSB peaks at such bifrequencies (29.3,380.6) Hz, (29.3, 761.2) Hz, (29.3, 1142) and also their higher harmonics relating to 29.3Hz, which is the frequency that the gear operates with the tooth breakage. Therefore, it shows that MSB performs outstandingly in diagnosing this incipient fault.

Moreover, for the two large fault cases, the resonance induced responses at about $2\times f_{me34}$ =760Hz are still very dominant. MSB magnitudes shows agreeable increments across different severities, showing more support that higher SNR signals at resonance bands give a more accurate diagnosis

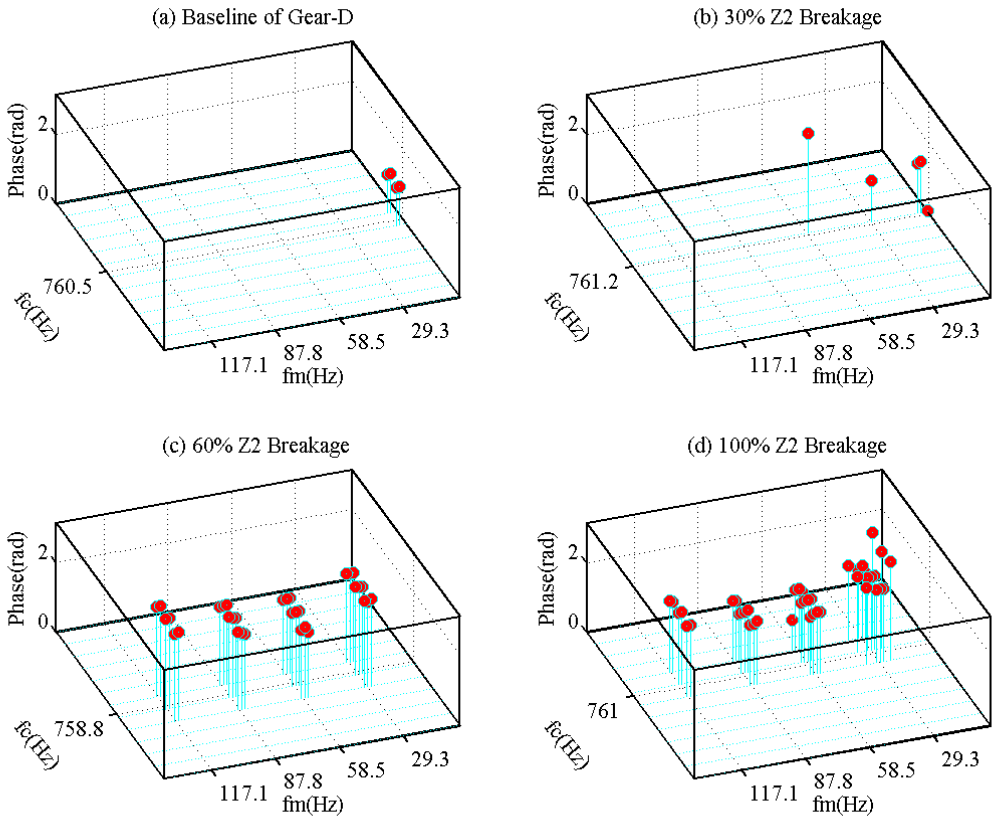

Figure 7.7 MSB phase results at a resonance for Gear-D under different fault severities

More interestingly, the phase results give more convincible supports that either AM or PM is dominant at a resonance but dese not interfere with each other. As shown in Figure 7.7, phase values in Figure 7.7 (c) and (d) are relatively stable across different f_m components, indicating that only type of modulations occurs. On the other hand, AM and PM interferences exhibit large phase differences between different components even if they are close each other as shown in Figure 7.7 (b).

Obviously, system identification is one of important and yet open subjects in dynamics and has been investigated actively for many decades. An in-depth and systematic investigation of using MSB for system identification deserves attention but may deviate too much from main subject addressed in this thesis. Nevertheless, it is believed that MSB can help to obtain more accurate results because of its superior noise suppression performance.

In general, this second verification of MSB performance has found that MSB can detect even smaller fault. This is achieved successfully by observing the changes in the components other than the meshing components where the fault induced to. Instead the components which are insignificant under baseline have been much less influenced by errors. Therefore, they are likely to be more apparent once a fault has small amplitude and hence can be based on for achieving more reliable and accurate detection and diagnosis.

7.4 MSB based Local Fault Detection and Diagnosis

As discussed, MSB can produce good results for gear fault detection and diagnosis. For more accurate diagnostics, MSB results can be further improved by combining AM and PM effects together. In addition, it has shown that the meshing component has been influenced more by inherent gear errors. This suggests that another band other than the meshing components need to be identified for obtaining optimal diagnosis. Amongst different bands examined that can provide consistent diagnosis, the resonance bands can provide better results because of

the magnification of modulation effect. However, this resonance probably only happens in limited scenarios, not the general characteristics for every industrial gearbox. Furthermore, an automated implementation MSB analysis always is required for high efficient diagnosis. To deal with these issues, a comprehensive diagnostic procedure needs to be developed.

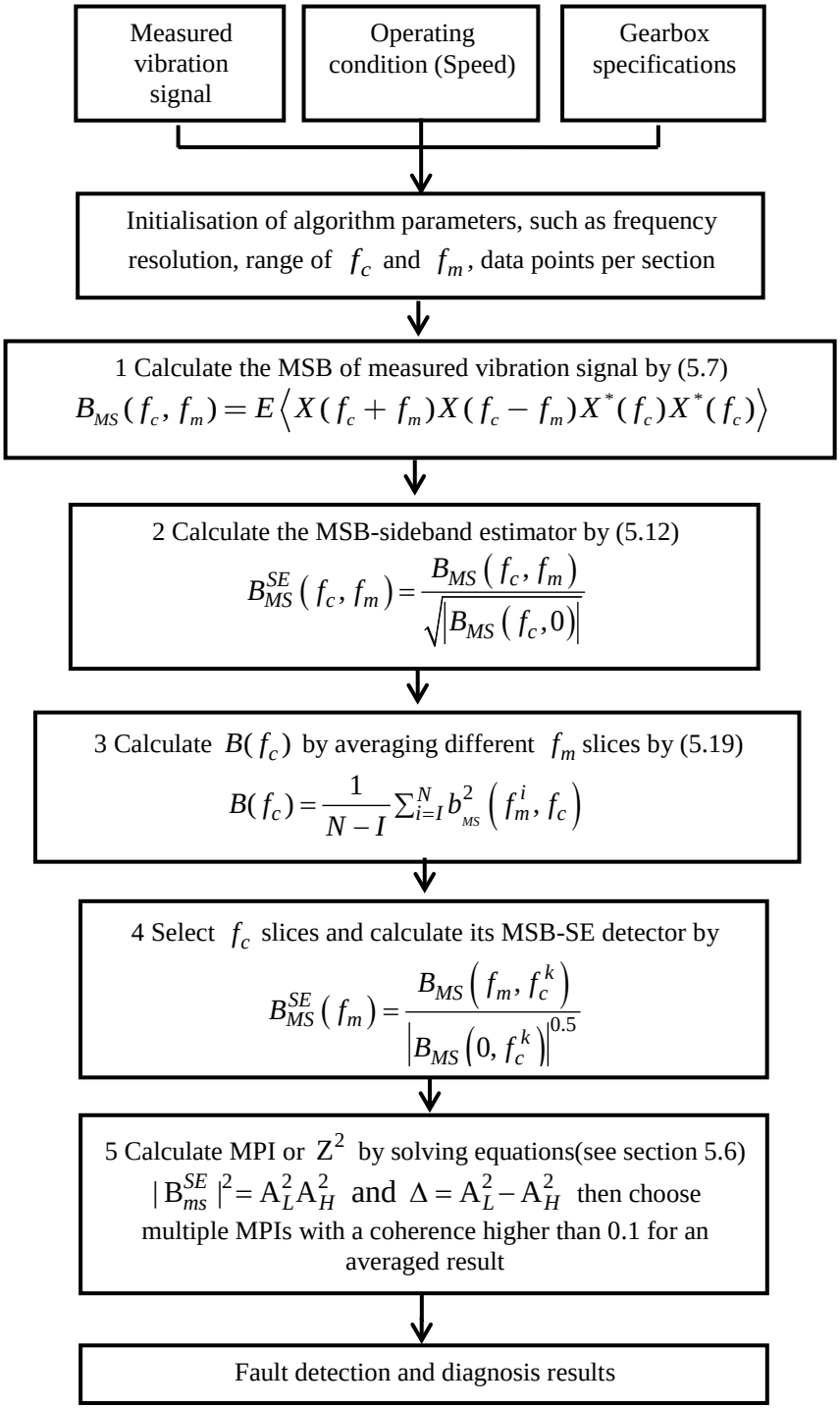

Figure 7.8 A flow chart of automated diagnosis using MSB

Firstly, to avoid or minimize error influences, a frequency band for the f_c in MSB can be used that does not include the first meshing frequency. For current gearbox the band can be from 500Hz to 1300Hz in which there are only two higher order harmonics of f_{me34} but not the first from both f_{me34} and f_{me12}. Obviously, it not only includes the resonant bands but a sufficiently wide range of band that potentially appear new components when a fault occurs.

Secondly, the frequency band for f_m can be set based on the same principle as that of f_c for minimising manufacturing errors. This means that the first few components of different shaft frequencies need to be avoided as it is the most common error can occur in a gearbox due to not only manufacturing accumulative pitch problems, but also due to misalignment and unbalances during installation. In addition the highest can be 4 or 5 times of f_r which is sufficiently wide to include most sideband components of interest.

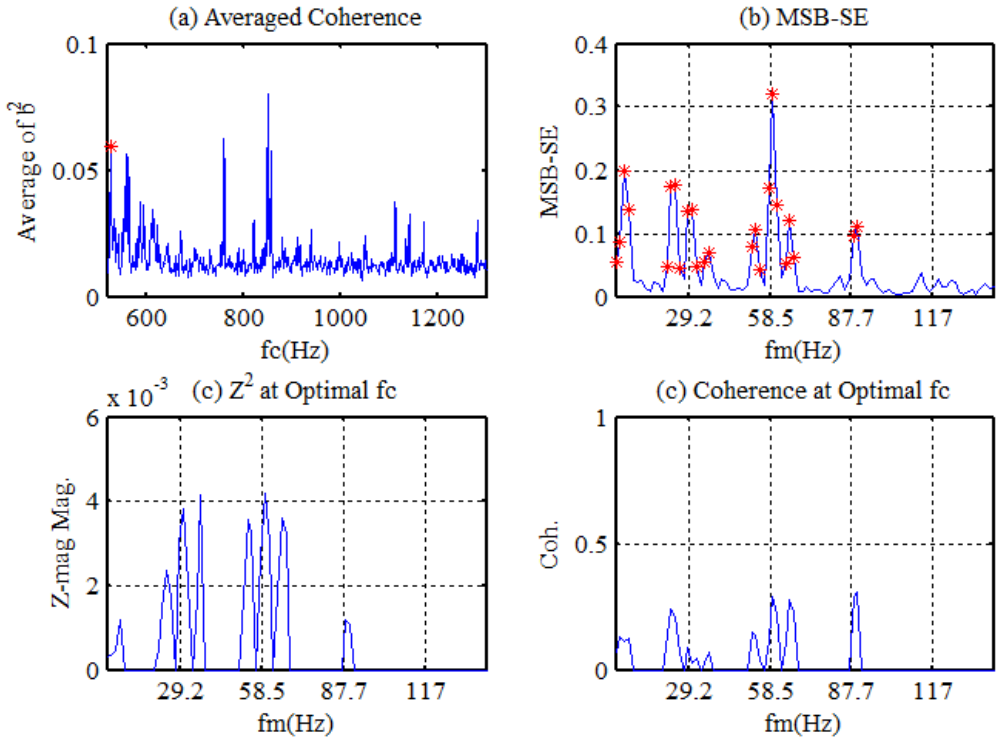

Figure 7.9 Illustration of optimal MSB components

Once theses bandwidth is determined, an automatic procedure can be implemented to select optimal MSB components inside the band for diagnosis, which is depicted in the flow chart of Figure 7.8. As it can be seen in the procedure, a critical step of the procedure is required to obtain averaged MPI for more accurate and robust detection and diagnosis. Figure 7.9 exemplifies a typical implementation for determining the critical steps. Figure 7.8(a) is the result from Step 3, in which the first index I is set to an f_m index values equivalent to f_r =32Hz to void the first harmonic. The MSB-SE is from a peak in the averaged coherence. This value is away from the resonance to show the effect of MPI in Figure 7.9(b). Compared with MSB-SE, the MPI values are enhanced around 29.2 Hz and 58.5Hz with different degrees. As a result, the value at 58.5Hz seems lower but it will give more correct estimation. In addition, the peak at 29.2Hz has very low coherence, indicating that it is influenced more by noise and will be discarded in the average processing.

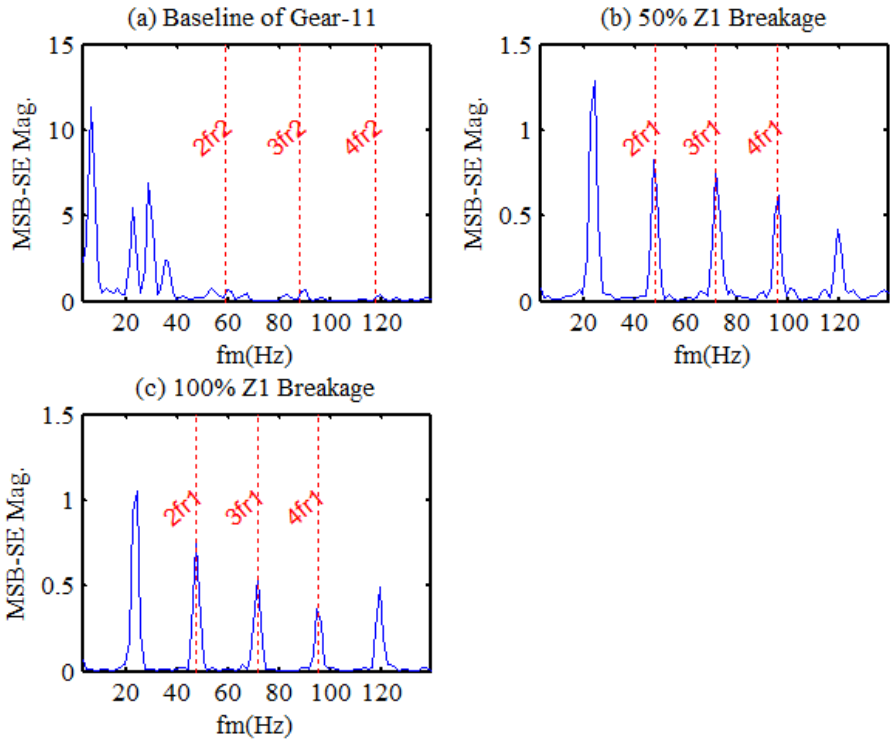

Figure 7.10 Location diagnosis based MSB Detector for the breakage Z1

Figures 7.10 and 7.11 present the diagnostic results for both fault location and severities for the Z1 tooth breakage cases. And Figures 7.12 and 7.13 present the diagnostic results for both fault location and severities for the Z2 tooth breakage cases. These results are fully consistent with the fault scenarios induced during data acquisition. Therefore, it shows that:

1) MPI suggested is accurate to quantify AM and PM combination;

2) MSB-SE detector developed is reliable to select optimal bands;

3) The proposed automatic procedure is feasible and reliable; and

4) The scheme for using frequency band other than meshing component is correct.

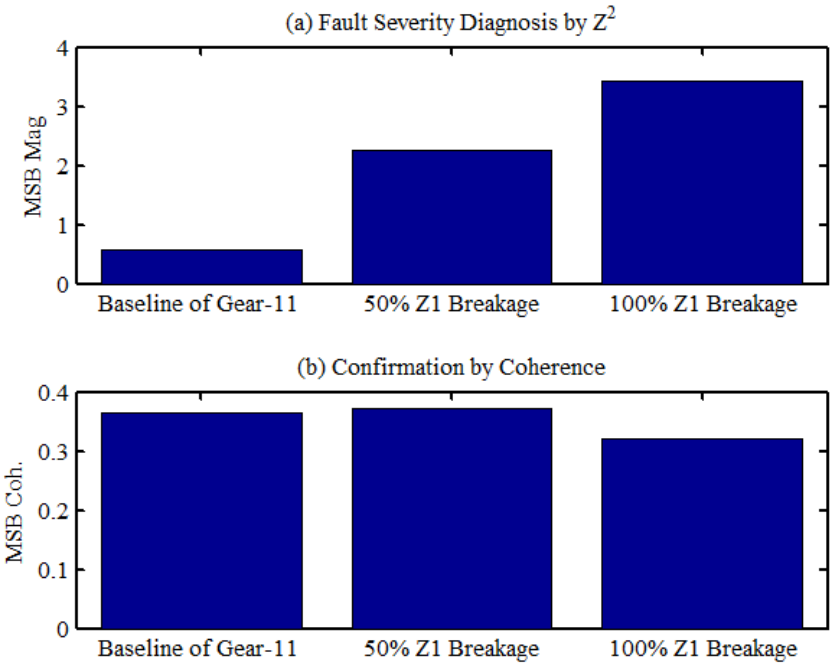

Figure 7.11 Severity diagnosis based MSB-MPI for the breakages on Z1

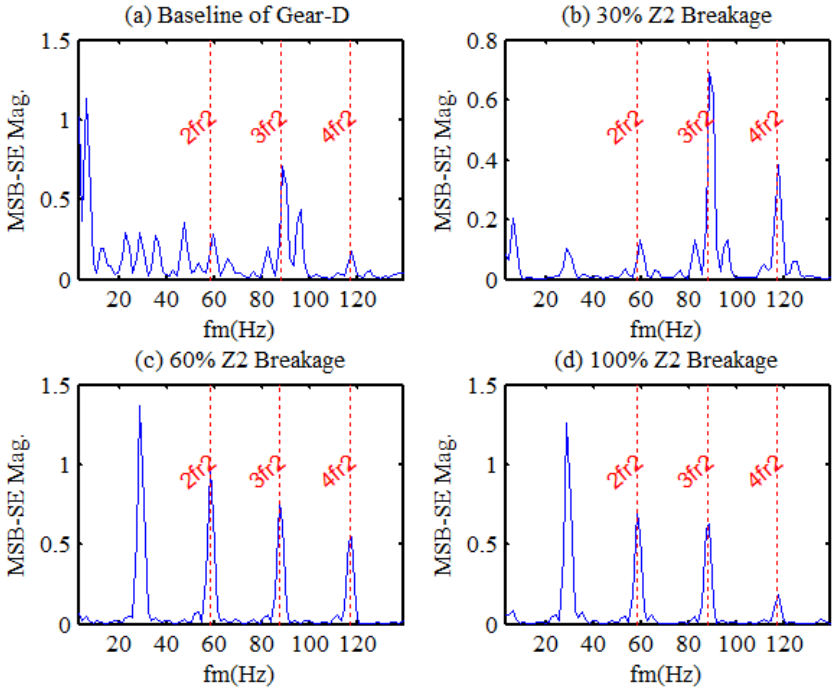

Figure 7.12 Location diagnosis based MSB Detector for the breakage Z2

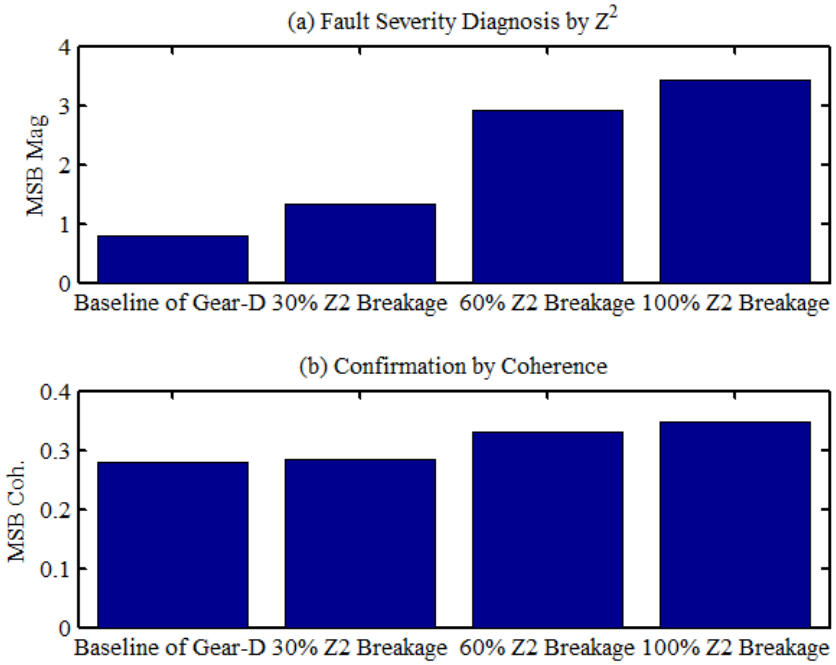

Figure 7.13 Severity diagnosis based MSB MPI for the breakages on Z2

7.5 CB Characteristics of Tooth Breakage

Figure 7.14 and 7.15 show the results from CB analysis for the two cases of tooth breakages respectively examined in Section 7.3. Based on CB magnitude increase characteristics it can be seen that they can detect and diagnose the fault locations and severity when the faults are very severe i.e. for the cases of 100% tooth breakages. These inferior results are due to that CB results have combined changes due to the faults with the error effects from the lower order components. As discussed in Section 7.2, these errors can also cause problems with less accuracy in estimating the meshing components, which in turn also causes CB to produce incorrect diagnostic results. Therefore, in this frequency range of interest CB performs poorly in diagnosing tooth breakages, compared with MSB.

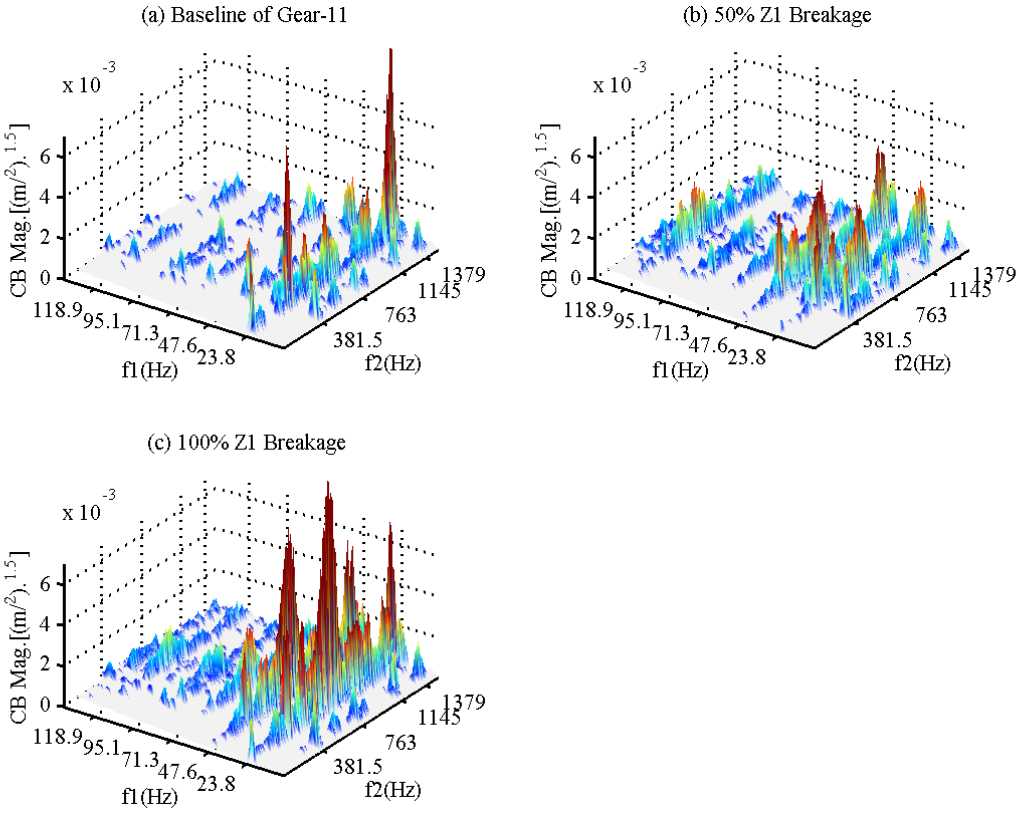

Figure 7.14 CB results for tooth breakages on Z1

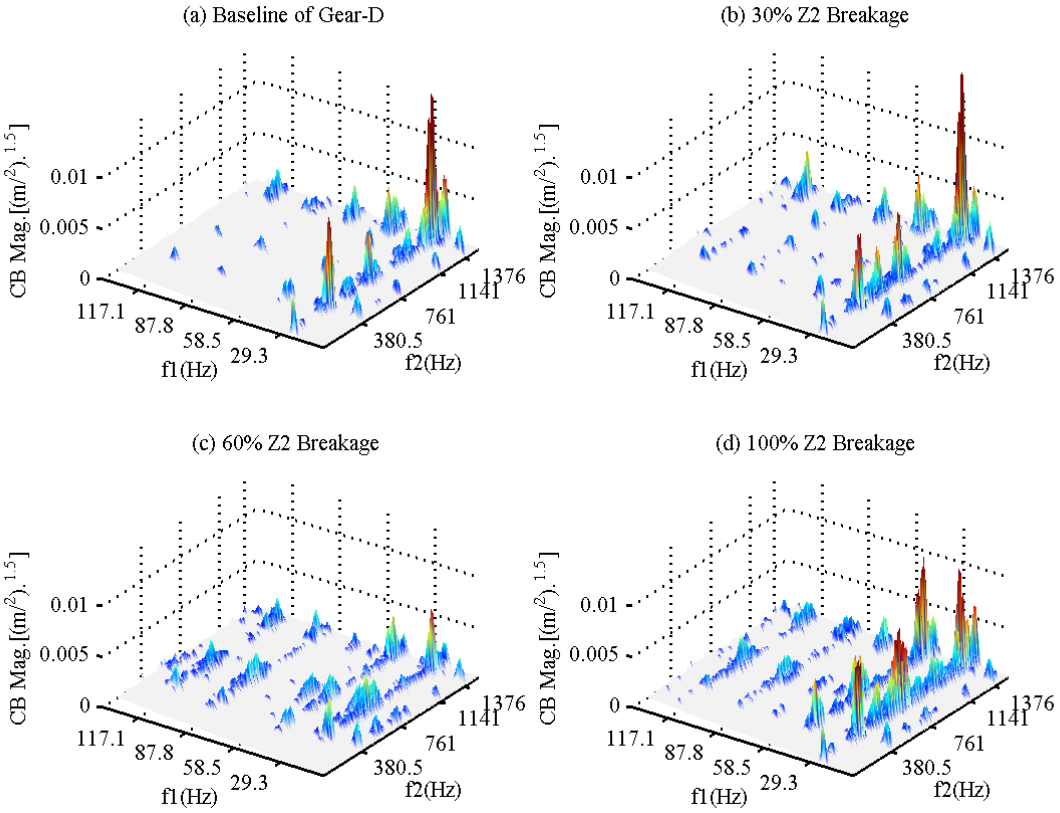

Figure 7.15 CB results for the tooth breakages on Z2

7.6 TSA Signal Characteristics of Tooth Breakage Vibration Responses

For more benchmarks, TSA signals for the case of Z1 tooth breakage are also obtained as shown in Figure 7.16. The signals resulted from the phase signal of shaft 1, denoted as tsa1, shows clear spikes with the increase in fault severity. In the meantime, TSA signals obtained based on the phase of other two shaft exhibit stationary waves, showing they are in healthy statues. This demonstrates that TSA signal allows this fault case to be separated correctly.

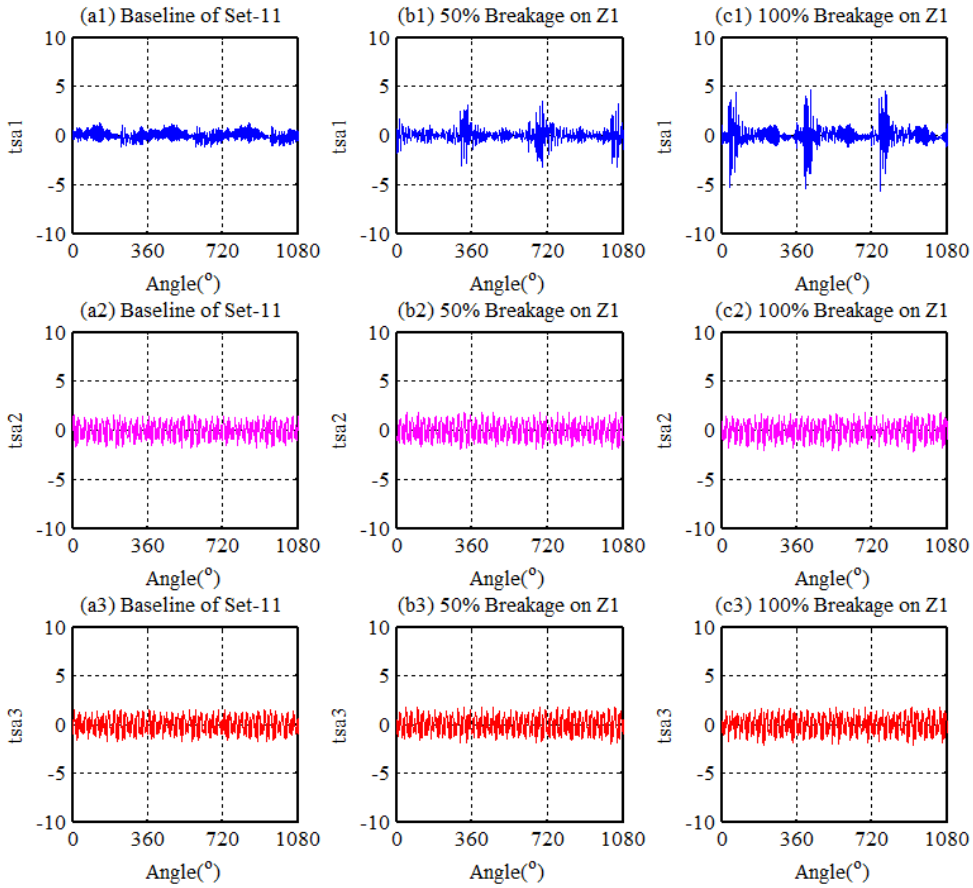

Figure 7.16 Location and severity diagnosis based on TSA signals for the breakage on Z1

However, diagnosing the fault on Z2 based on TSA signals results in lower performances. As it can be seen in Figure 7.17 that the spike signature appears in tsa2 signals until the fault severity is at 60%. It means that it is not sufficiently sensitive to the small fault. This deficiency may be still from the influences of phase calculation error which has discussed in the simulation studies. As gained from MSB analysis, this gear has higher manufacturing errors. These errors may induce more phase deviation and makes the overall results are poorer.

In addition, TSA signals for the third shaft also show similar spikes to that of tsa2 signal. This can lead to a completely incorrect diagnosis.

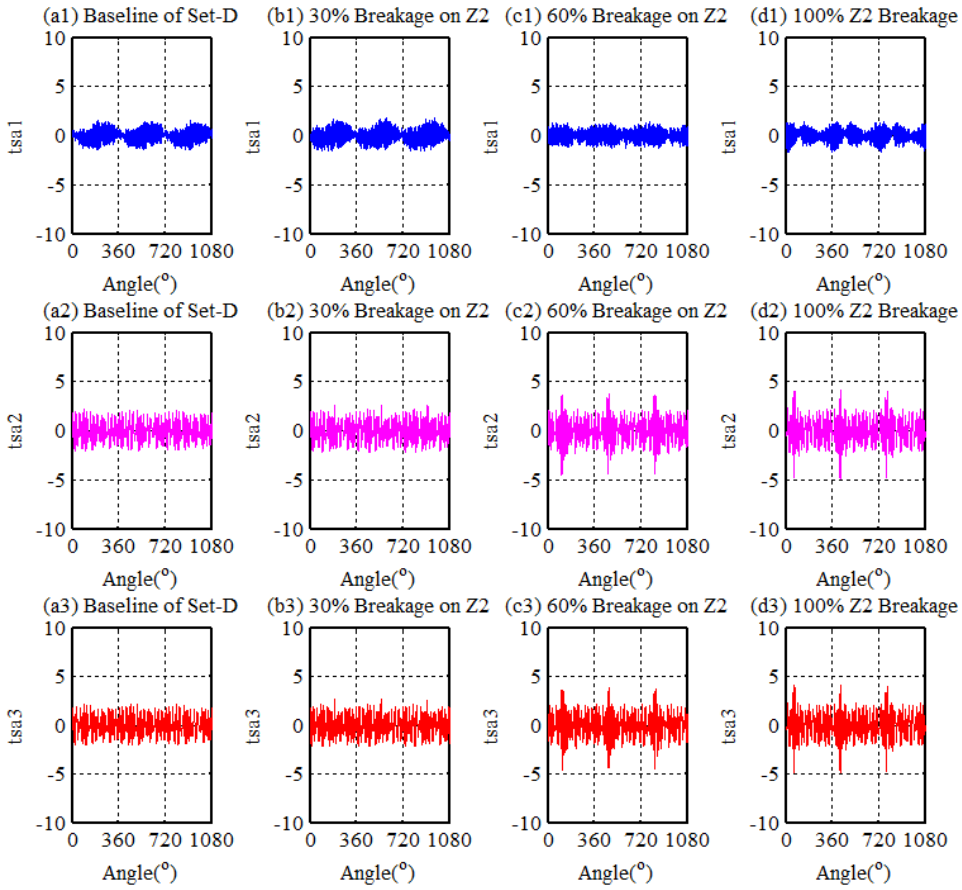

Figure 7.17 Location and severity diagnosis based on TSA signals for the breakage on Z2

Furthermore, characteristic frequencies of roller bearing are usually fractional values of shaft frequencies and their amplitudes are much smaller than meshing components. This means that TSA based on meshing frequency phases can remove these bearing contents, as they are asynchronous to any shaft components. This is another shortfall of TSA, which means that TSA is not particularly suitable for implementing a comprehensive gearbox condition monitoring system.

7.7 Chapter Summary

MSB based magnitude and coherence showed that it can provide an accurate diagnosis of the fault severity and also the location of the fault. However, when CB is applied to the same signals, the error at lower order components results in less robust performances for distinguishing fault locations. TSA can again produce erroneous results and the amplitude of the fault frequencies are dominated by the magnitude of the meshing frequencies.

CHAPTER 8.

DISTRIBUTED GEAR FAULT DETECTION AND DIAGNOSIS

8.1 Introduction

With the same bandwidth and analysis parameters as in Chapter 7, the diagnosis procedure is applied the datasets from the run to failure test. Only typical results are presented which are identified to indicate important deterioration events occurring on the testing gear set.

In addition, the test has not resulted in local tooth breakages but instead has led to clear tooth surface wear phenomena such as pitted areas, fatigue zones and micro scratch markers as illustrated in Chapter 6. These defects are found across most teeth of the meshing gears. Therefore, this gear set is regarded as distributed faults, which often is found more difficult to be detected and quantified.

Furthermore, as the testing gear was operated by keeping the test rig under the same configuration during the full test course. It means that errors due to reinstallation are avoided, which can happen for the previous tests as the tooth breakages need to be created by taking the gear out of the case each time. This means that any errors observed in this test would only the inherent manufacturing errors. Moreover, it is interesting to know how MSB responses to the gradual changes while the gears are deteriorating.

8.2 MSB Characteristics of Asymmetric Tooth Wear

Figures 8.1, 8.2 and 8.3 present MSB magnitude, coherence and phase results respectively for six representative operating time instants. These time instants are identified by referring previous analysis results using motor current signature analysis (MCSA) applied to the same

test by Haram et al [133]. Specifically, by analysing the sidebands relating to the first shaft which are extracted from motor current signals, the results for operating period from 3.1 hours to 210 hours has no clear indications of change. Then it has found that there is a gradual increase trend of the sideband which allows the deterioration of the gearbox system to be detected. However, it has the concerns that the results may also include the influences of other part of the test rig such as the drive motor.

However, by examining these MSB results, it is difficult to find a meaningful component that can be consistent with the deterioration process positively. Especially, the meshing components at $f_{me}12$ exhibit a large degree of oscillations across different cases, showing that is highly unstable stable and yet again proves the claim that the meshing frequency is not a very effective band for diagnosis.

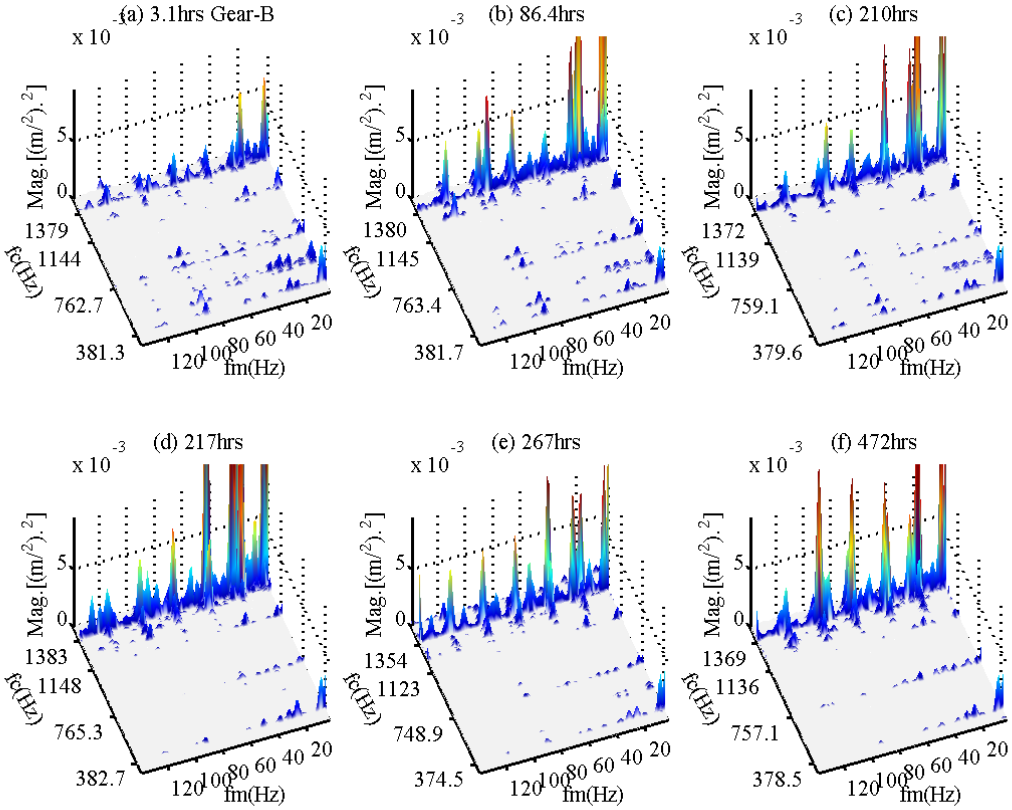

Figure 8.1 MSB magnitude results for Gear-B under different fault severities

In addition, it is not very clear to show the occurrence of the resonances. This can be well understood in that such distributed faults often induce additional vibration in narrower bands because of relatively smoother interference of the abnormal meshing process.

However, a closer examination has found that there are still many peaks showing gradual increase with operation times even though they appear in different frequency bands.

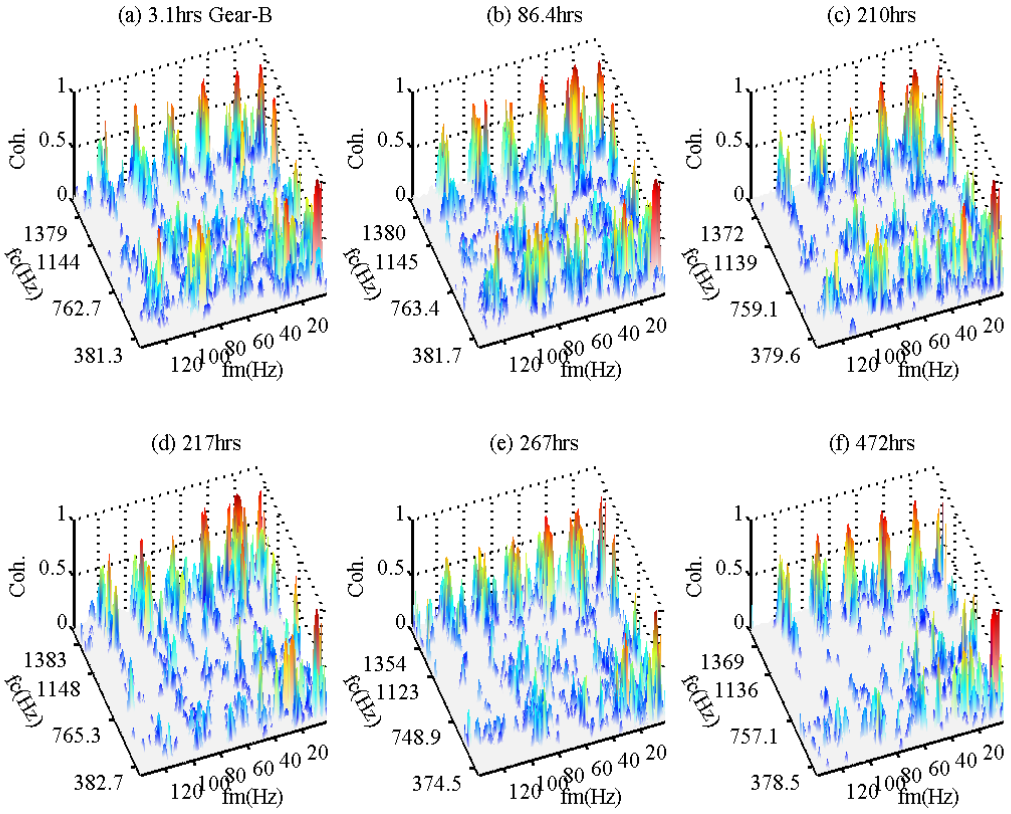

Figure 8.2 MSB coherence results for Gear-B under different fault severities

Also it is found that from the coherence and phase results that the number of MSB peaks show gradual decrease. This may indicate that the AM and PM interferences are relatively stable to suite to a particular status of modulation as the gears are deteriorated more.

In general, these observations are not very convincing in making some definitive decisions on the exact condition of the gears.

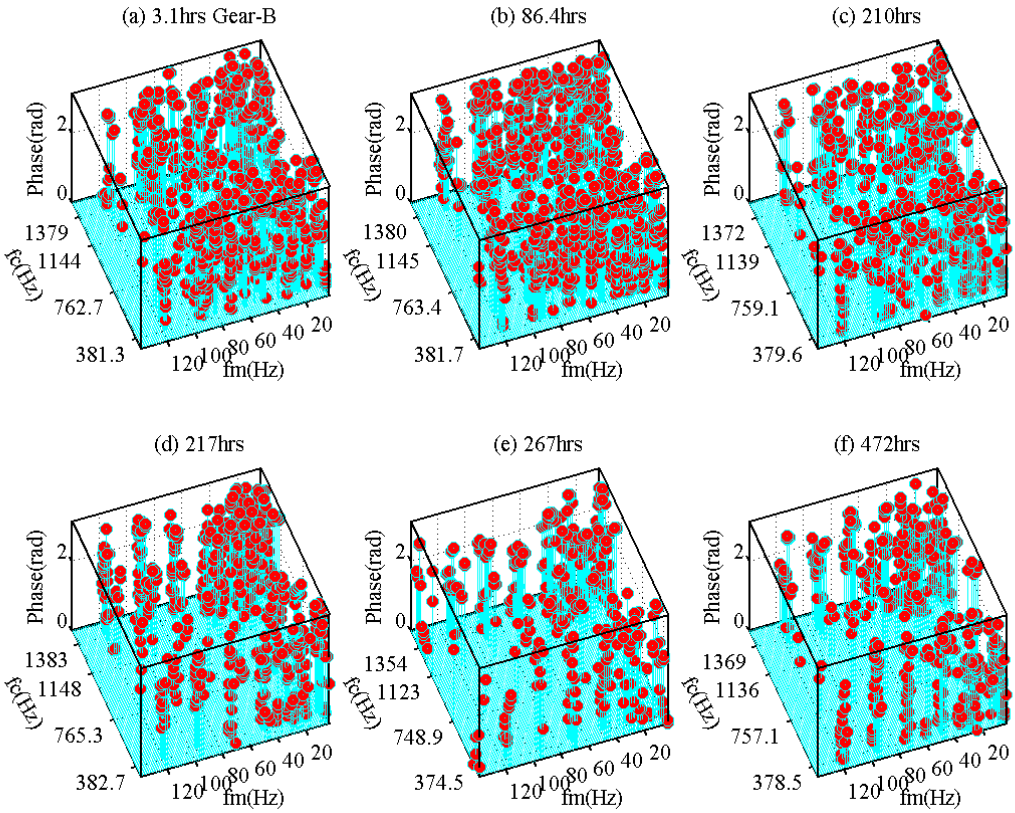

Figure 8.3 MSB phase results for Gear-B under different fault severities

8.3 Detection and Diagnosis of Gear Tooth Wear based on MSB

However, by applying the automated detection and diagnosis procedure developed with the same parameters it has obtained positive results as shown in Figure 8.4 and 8.5 regarding to location and severity respectively.

Firstly, the location diagnosis shows that the gear deterioration process has a clear transection phase from Z2 to Z1. At early stage Z2 may has higher errors than Z1. This has been shown by the detection results from Figure 8.4(a) to 8.4 (c) in which the more modulation effects at

Z2 appear slightly more than that of Z1. However, with more operation progression, Z1 is subjected to the effect of higher misalignments from the drive motor it connects to. As a result, the wear degree on Z1 becomes more significant and asymmetric, leading to greater modulation effect appear on Z1. This typical asymmetric wear has been found in the majority gear systems.

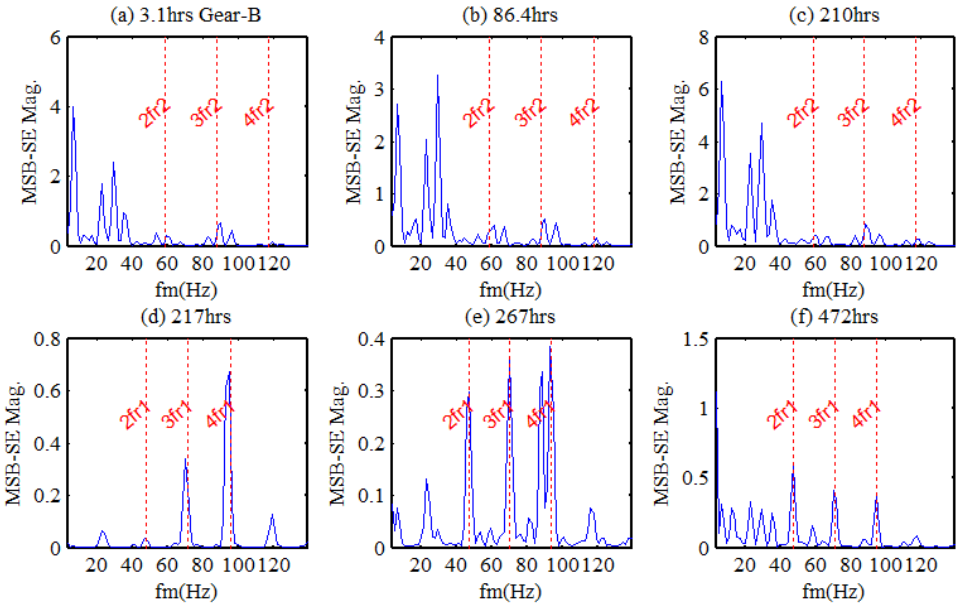

Figure 8.4 Location diagnosis based MSB Detector for Gear-B under the different fault severities

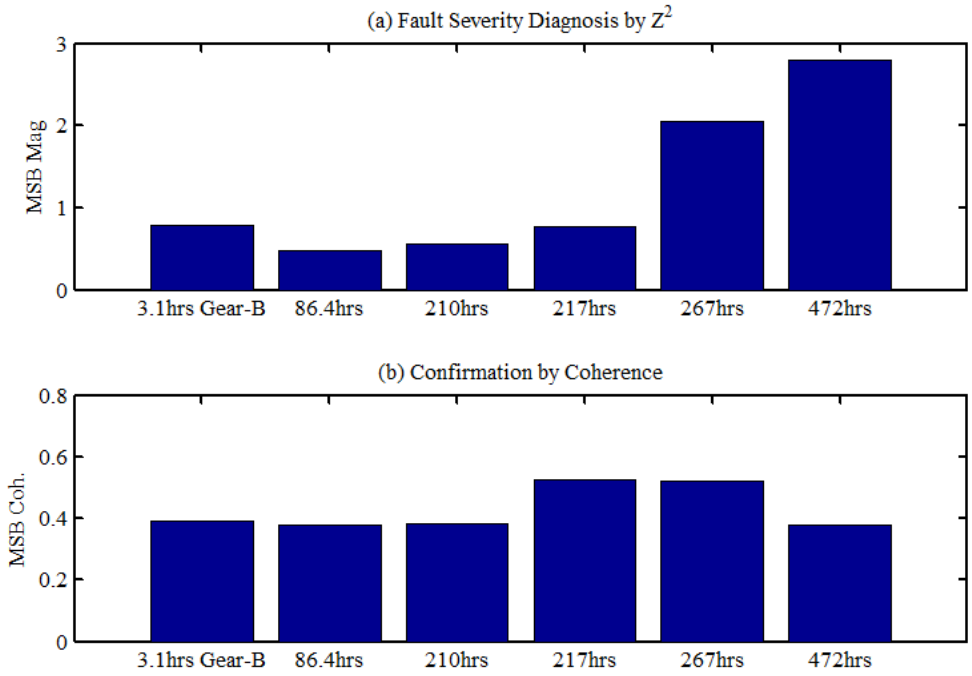

Figure 8.5 Severity diagnosis using MSB MPI for Gear-B

In the meantime, the MPI shows consistent results with the operation progress. However, in the beginning, MPI value is slightly higher, which is mainly due to the error effect or commonly known as running-in effect. After which, MPI shows an increase trend to reflect the severity correctly. Furthermore, the diagnosis results are ensured by high coherence values as shown in Figure 8.5(b) which are in the similar ranges as those in previous cases.

Overall, thorough and accurate diagnosis results achieved for this difficult case shows that this proposed diagnosis procedure is powerful and reliable.

8.4 CB Characteristics of Tooth Wear

Because of being sensitive to errors at Z1, CB results in Figures 8.6 and 8.7 cannot give very meaningful indication of the gear deterioration process due to inconsistent changes in magnitudes over the time instants. This confirms that more that the CB is not very effective for gear wear diagnosis in this lower frequency bands.

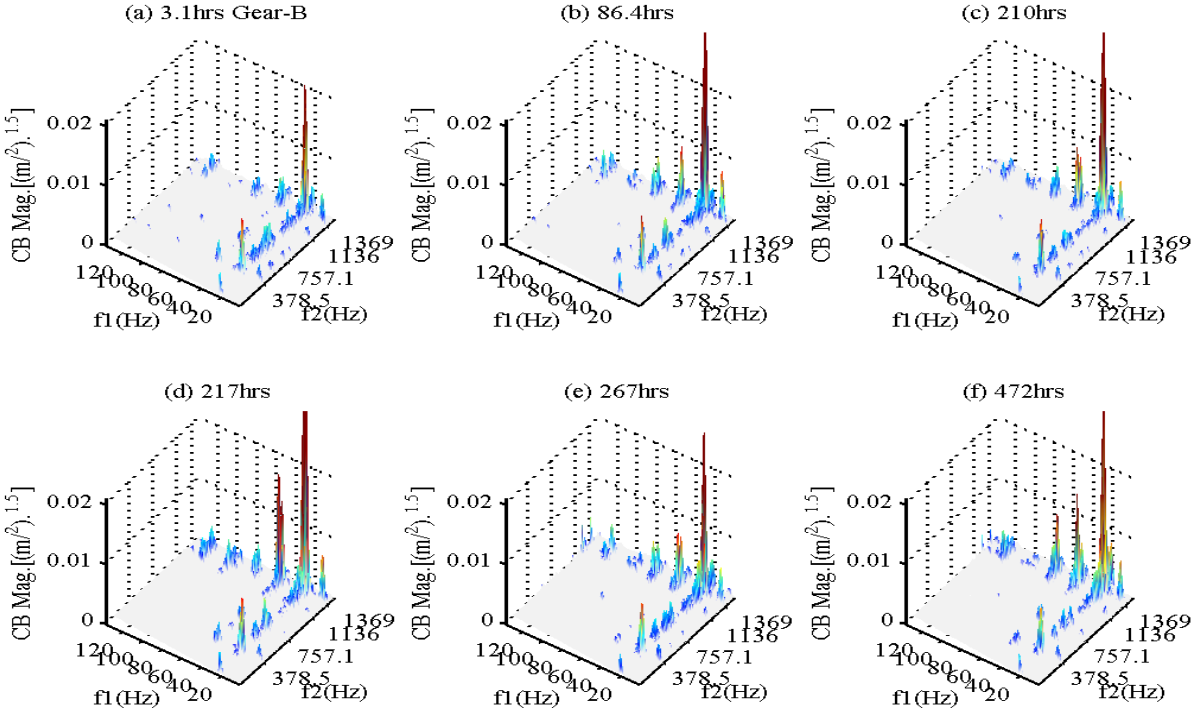

Figure 8.6 CB magnitude results for the gear tooth wear

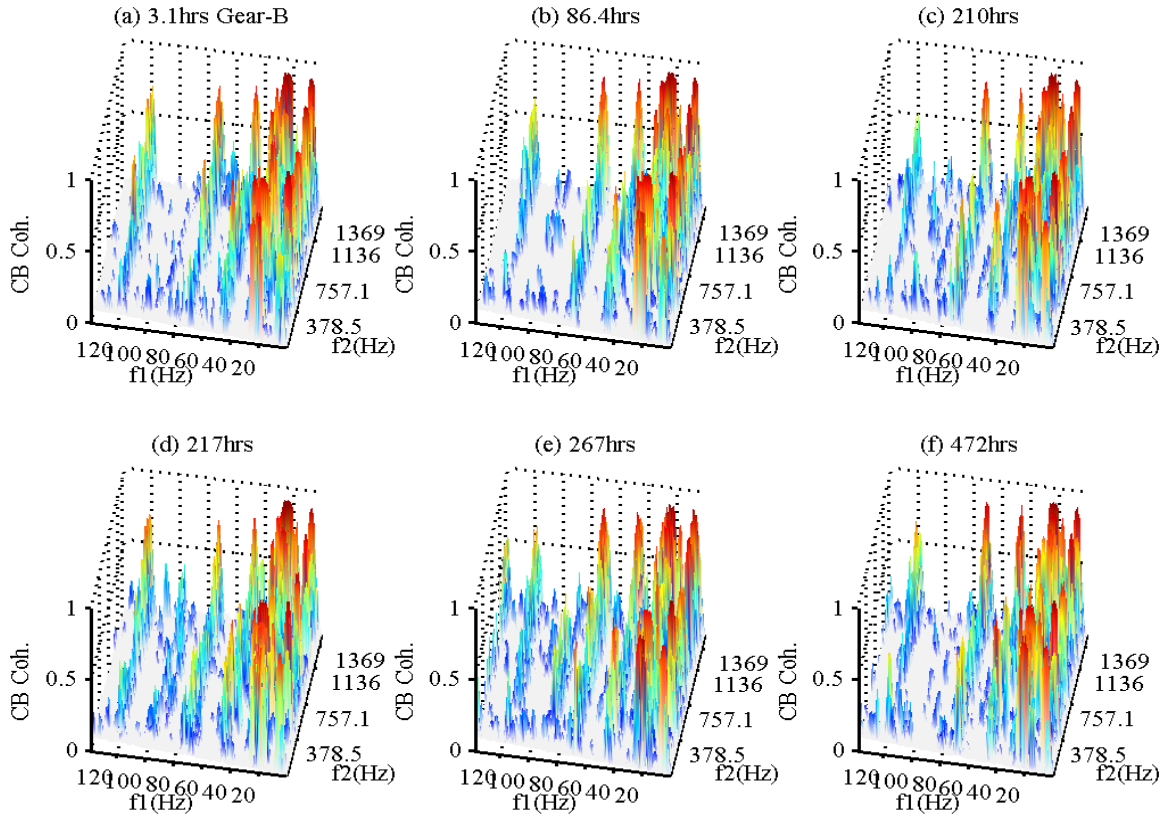

Figure 8.7 CB coherence results for the gear tooth wear

8.5 TSA Signal Characteristics of Tooth Wear

For the similar reasons, the inherent error of this gear set also leads to incorrect diagnosis by using TSA signals. As shown Figure 8.8, tsa1 signals show a clear envelope over the rotational period of the first shaft. This means that the error of asymmetry on Z1 is very significant at early stages but it become lower at late stages because of the effect of asymmetric tooth wear.

TSA signals for tsa2 and tsa3 also show inconsistent changes and cannot be based on for making diagnostic decisions positively.

Probably, because of the clear periodic envelopes on of tsa1 signals, it can lead to a diagnosis that Z1 is the main problem of the system. However, it is not convincible because the

amplitude decreases with operation progress. Therefore, TSA approaches are also very ineffective in diagnosing this distributed fault.

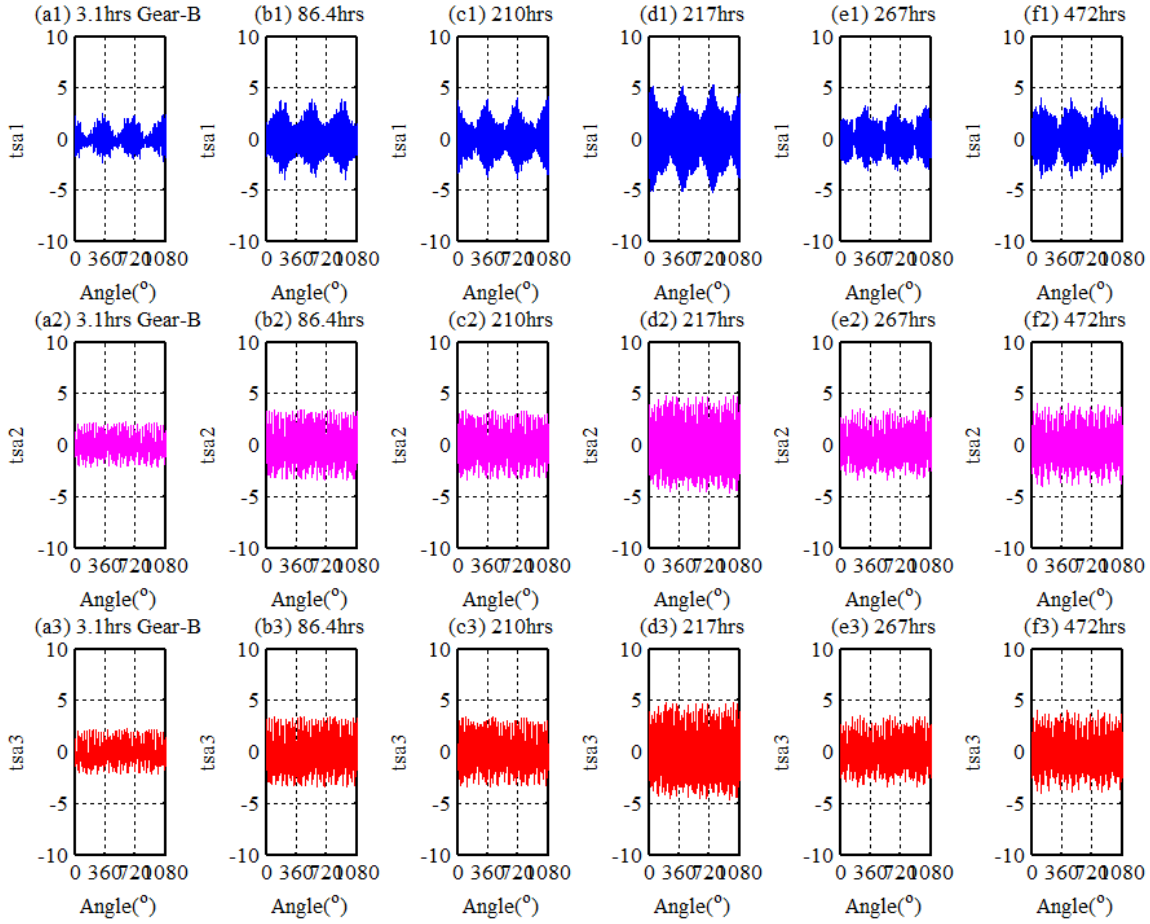

Figure 8.8 Diagnosis based on TSA signals for gear tooth wear

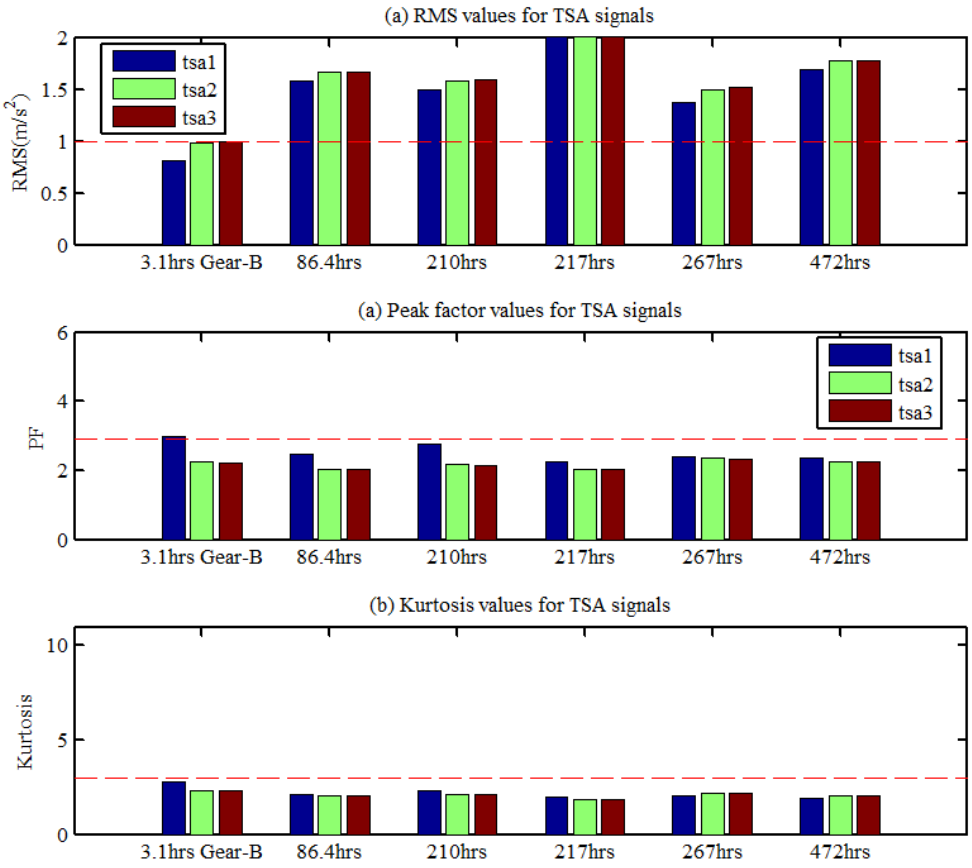

Figure 8.9 Time domain based diagnosis based on TSA signals for gear tooth wear

Figure 8.9 shows further that although there is a difference between the RMS values for the three TSA signals at the gearbox's healthy condition compared to the rest of the RMS values, it cannot provide any more inconsistent information other than that the gearbox is not at its ideal condition.

8.6 Chapter Summary

The progression of fatigue errors cannot be reliably diagnosed using CB since the method is not sensitive enough to detect the small changes over time. TSA is also unsuitable for usage as a fatigue diagnostic tool as the results are inconsistent with time. Whereas MSB coherence and phase results can confirm that as the severity of the fatigue wear increases, so does its MPI value, proving that it is the most accurate technique for diagnosing small fatigue faults in gearboxes.

CHAPTER 9.

CONCLUSIONS AND FURTHER WORK

Vibration generated by a gearbox provides a great deal of information regarding its heathy condition. However, the influences of noise and interference content make it very difficult to extract the useful information for achieving a reliable and accurate diagnosis of gearbox faults. Therefore, it is critical to investigate novel approaches to effectively suppress noise and decouple the interferences. Of numerous signal processing techniques investigated in recent years, TSA, CB and MSB have been focused in this thesis for the purpose of noise suppression and interference decoupling because they have the capability of noise reduction. Especially, MSB can additionally highlight modulation characteristics that are of the primary properties of vibrations in a multi stage gearbox.

TSA is very effective to extract vibration content that are synchronous with a particular gear and exclude any non- synchronous contents such as from other gears and noise. However, it has found that TSA signal also destroy the PM content. Consequently, it does not take the full information of AM-PM effect and result in low diagnostic performance, which is because the phase signal relied on by TSA is numerically unstable due to strong noise effect. In addition, TSA analysis also removes the diagnostic information for roller bearings, which make it impossible to achieve a comprehensive system for gearbox condition monitoring.

Convention bispectrum (CB) can reduce Gaussian noise in signals by highlighting components that are coupled in both frequency and phase. It is particularly useful for extracting QPC contents. However, it is inefficient in extract AM-PM components in that it

can results in overpopulated CB components due to the multiple couplings between low frequency components including their harmonics and high frequency meshing components which are present inherently in multiple stage gearboxes. This make it difficult to extract correct information for implementing diagnosis. In addition, the low frequency components can be easily influenced by distant vibrations and gear errors. So the components are not very reliable for gear diagnosis. It is feasible to have AM-PM contents in the high frequency range from the first to second harmonics of meshing frequency. However, the frequency span between the two mesh harmonics often is sufficiently wide to cover different structural vibration modes, resulting in a poor estimation of AM-PM effect by using two demodulated components from two meshing harmonics.

On the other hand, recently focused MSB can demodulate the modulation components around just one meshing components, which can improve estimation significantly, even if the meshing frequency is close to a resonant band. In the meantime MSB maintains the properties of noise suppression and inclusion of power spectrum content. Therefore, it is more effective in characterising the gear vibration signals for accurate diagnostic results.

Specifically, MSB combines the carrier and modulation components as the products between the two sidebands and the carrier. It allows the small modulation content highlighted due to the distinctive peaks of the carrier. This means that the week modulation can be enhanced to show up by the high amplitude of the meshing content. Phase To separate between the modulation and the meshing carrier, the MSB-estimator can be used to remove the carrier magnitude, leaving only the product of just two sidebands, which can be sufficient to quantify the modulation levels due to gear faults.

However, because AM-PM combination can be resulted from various phase differences between AM and PM, the sideband product is not necessarily a complete measure of AM and

PM. Therefore, based on the model of AM-PM formalisation in gearbox vibration signals, a MPI is suggested to reflect the combination of the power in AM and PM to reflect the maximum consequence of the interfered modulation that can favourite the early fault detection.

Subsequently, a new scheme to calculate the MPI is developed based on MSB results. Using the sideband product from MSB component of interest and the sideband difference from MSB at zero frequency of modulation, an optimal estimation of MPI can be obtained by solving a quadratic equation. As the sideband product and difference are the noise minimisation components, the estimator produce a more accurate and reliable results for qualifying the gear modulations.

Besides, MSB phase results can show the interference degrees between AM and PM. Any phase value apart from 0 and π means that the occurrence of AM-PM interference and the proposed scheme is required for decoupling them. However, this approach of detecting interference often not very efficient and accurate as the angle calculation often suffer from numerical instability, which has had significant influences on TSA. To avoid this MSB coherence results are employed. It can be based on to select MSB components which exhibit high coherence values to be the candidates for applying the scheme, leading to a MSB-detector. Moreover, high coherence results are referenced as the conditions to determine significant optimal carrier ranges for achieving reliable AM-PM decoupling. Finally, coherence can be an important reference to confirm the detection, the cases in experimental evaluation studies carried out in this research should possess coherence values of at least higher than 0.1.

Vibration datasets ware collected under different fault conditions including local tooth faults and asymmetric tooth wear in a two stage helical gearbox which is from a major gear

company in U.K. supplying such gearboxes to different power transmission systems in industry available. These datasets can be representative and the evaluation is reliable.

Two cases of local faults on the gearbox are examined, one is at the input shaft and other is at the middle shaft. Diagnostic results show that the new MPI gives a monochromatic increase trend that is consistent with the fault severity induced to corresponding to the faulty gear locations, showing that the detector and estimator give correct diagnostic results. Specially, the 30% tooth defect cannot be separated correctly by previous know methods. Therefore, it is can be deemed that the new scheme is more reliable and accurate in diagnosing tooth breakages.

To further examine the performance, the datasets from a run-to-failure test are utilised with the same parameters setting to the MSB methods in detecting tooth breakage. The diagnosis obtained by the dataset has identified that the gear condition has deteriorated significantly at the middle phase of the test operation, providing an early warn that the over-worn gear set occurs more at the first stage of the gearbox. This confirms that the developed scheme is also effective to diagnosing distributed gear fault which is a subject that has received considerable researches.

9.1 Conclusions

Overall the research carried out through theoretical analysis and experimental studies has yielded following key findings and main conclusions which are also deemed to be novel as they are obtained for the first time in this thesis:

1. The key issues need to be tackled for achieving accurate gear diagnostics are both noise reduction and AM-PM decoupling.

2. The decoupling can be achieved by power spectrum analysis but not accurate enough to resolve the low modulation of incipient faults.
3. TSA analysis is unable to reliably decouple AM-PM in gear signals for diagnosis because its phase signal is sensitive to noise;
4. CB requires a wide bandwidth to achieve demodulation and decoupling which result in not only poor results and but also not suitable for emerging technologies that can be used for condition monitoring
5. MSB is particularly suitable for gear vibration analysis because it include phase information in achieving noise reduction and modulation enhancements;
6. The developed MSB detector and estimator are reliable and accurate for diagnosing different gear faults.
7. The optimal bandwidth can be the range that has few components observable in MSB results of baseline signals.
8. The first modulation component needs to be avoided to reduce the influences of possible gear manufacturing errors.
9. The commonly recognised bandwidth around meshing frequencies is not the optimal one as it can be influenced significantly by manufacturing errors.
10. This MSB based diagnosis scheme is more suitable to be embedded into emerging techniques for cost-effective real time monitoring.
11. MSB estimated by FFT and Welch's estimator is sufficiently accurate to produce correct diagnosis.

9.2 Future Works

To further improve the diagnostic performance of multistage gearboxes using the vibrations, researches in the direction of noise reduction and modulation decoupling can be further

carried out in the two approaches: MSB based methods and new signals processing method based methods:

9.2.1 MSB based Methods

1 To improve computational efficiency of MSB in order to allow for real-time monitoring and embedded devices;

2 To investigate applying the MSB based detector and estimator to diagnosing gearbox bearing faults and separating gear and bearing faults which occur concurrently;

3 To confirm and improve MSB based system resonance identification in theory and by experiments;

4 To identify MSB components from non-modulation contents to further achieve sparse representation;

5 To develop sensing systems embedded with MSB analysis.

9.2.2 New Signal Processing Framework based Methods

1 To investigate compressive sensing techniques such as random or sparse sampling for gearbox real time monitoring;

2 To investigate nonlinear model based signal processing techniques such as manifold learning and stochastic resonance approaches;

3 To investigate artificial intelligent based signal technologies which include genetic algorithms, particle swarm algorithms for optimising diagnostic accuracies and various neural networks for the modelling and classification of completed datasets;

4 As a fast devolvement technique wavelet analysis can be effective in de-noising stationary random and highlight nonstationary content of local gear faults, which can be a prepressing method for MSB analysis in order to further improve the diagnostic accuracy.

REFERENCES

[1] Rao, B. K. N. *Handbook of condition monitoring*. Elsevier, (1996).

[2] Dalpiaz, G., A. Rivola, and R. Rubini. "Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears." *Mechanical Systems and Signal Processing*, (2000).

[3] Randall, R.B., "Frequency Analysis", *Brüel and Kjær*, 3rd edition, (1987).

[4] Choy, F. K., et al. "Analysis of the effects of surface pitting and wear on the vibration of a gear transmission system." *Tribology International*, (1996).

[5] Bartelmus, W., and R. Zimroz. "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions." *Mechanical Systems and Signal Processing* (2009).

[6] Peng, Z. K., and F. L. Chu. "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography." *Mechanical systems and signal processing*, (2004): 199-221.

[7] Albarbar, A, et al. "Suitability of MEMS accelerometers for condition monitoring: An experimental study." *Sensors 8.2* (2008): 784-799.

[8] Jardine, A KS, Daming Lin, and Dragan Banjevic. "A review on machinery diagnostics and prognostics implementing condition-based maintenance." *Mechanical systems and signal processing* (2006): 1483-1510.

[9] Yang, D.M., et al. "Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks." *Mechanical systems and signal processing* (2002): 391-411.

[10] Arthur, N, and Penman, P. "Induction machine condition monitoring with higher order spectra." *Industrial Electronics*, IEEE Transactions (2000): 1031-1041.

[11] Thompson, T., and M. Granger. "What Price Preventive Maintenance?" *Proceedings of the Water Environment Federation*, (2004).

[12] Arts, R. H. P. M., Gerald M. Knapp, and Lawrence Mann Jr. "Some aspects of measuring maintenance performance in the process industry." *Journal of Quality in Maintenance Engineering*, (1998).

[13] C. A. Walford, “Wind turbine reliability: understanding and minimizing wind turbine operation and maintenance costs,” *Sandia National Laboratories, Rep. SAND2006-1100,* (2006).

[14] Crabtree, C.J.; Feng, Y.; Tavner, P.J. Detecting Incipient Wind Turbine Gearbox Failure: A Signal Analysis Method for On-line Condition Monitoring. *In Proceedings of European Wind Energy Conference (EWEC),* (2010).

[15] Rogers, L. M., The application of vibration analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings. *Tribology International,* (1979): 51-59.

[16] Tavner, Peter J., and James Penman. Condition monitoring of electrical machines. *Vol. 1. Research Studies Pre*, (1987).

[17] McFadden, P. D., and J. D. Smith. "Model for the vibration produced by a single point defect in a rolling element bearing." *Journal of sound and vibration 96.1* (1984): 69-82.

[18] Beguenane, R. Benbouzid, M. Induction Motors Thermal Monitoring by Means of Rotor Resistance Identification. *IEEE Transactions on Energy Conversion, ,* (1999).

[19] Rgeai, M. N. “Helical gearbox fault detection using motor current signature analysis” Doctoral dissertation, University of Manchester, (2007)

[20] Willetts, R., et al. "Generating adaptive alarms for condition monitoring data." *International Journal of COMADEM 8.3* (2005): 26-36.

[21] Ekici, Sami, Selcuk Yildirim, and Mustafa Poyraz. "Energy and entropy-based feature extraction for locating fault on transmission lines by using neural network and wavelet packet decomposition." *Expert Systems with Applications 34.4* (2008): 2937-2944.

[22] Hsu, John S. "Monitoring of defects in induction motors through air-gap torque observation." Industry Applications, IEEE Transactions on 31.5 (1995): 1016-1021.

[23] Crocker, Malcolm J. *Handbook of noise and vibration control*. John Wiley & Sons, 2007.

[24] Milanfar, Peyman, and Jeffery H. Lang. "Monitoring the thermal condition of permanent-magnet synchronous motors." *Aerospace and Electronic Systems, IEEE Transactions on 32.4* (1996): 1421-1429.

[25] Booth, Campbell, and Jim R. McDonald. "The use of artificial neural networks for condition monitoring of electrical power transformers." *Neurocomputing 23.1* (1998): 97-109.

[26] Chambers, K. W., M. C. Arneson, and C. A. Waggoner. "An on-line ferromagnetic wear debris sensor for machinery condition monitoring and failure detection." *Wear 128.3* (1988): 325-337.

[27] Peng, Z., and N. Kessissoglou. "An integrated approach to fault diagnosis of machinery using wear debris and vibration analysis." *Wear 255.7* (2003): 1221-1232.

[28] Miller, Jerry L., and Duka Kitaljevich. "In-line oil debris monitor for aircraft engine condition assessment." *Aerospace Conference Proceedings, 2000 Vol. 6,*. IEEE, (2000).

[29] Mba, David, and Raj BKN Rao. "Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines: Bearings, Pumps, Gearboxes, Engines and Rotating Structures." *COMADEM*, (2006).

[30] Albarbar, A., Gu, F. and A. D. Ball. "Diesel engine fuel injection monitoring using acoustic measurements and independent component analysis." *Measurement* 43.10 (2010): 1376-1386.

[31] Li, C. James, and S. Y. Li. "Acoustic emission analysis for bearing condition monitoring." Wear 185.1 (1995): 67-74.

[32] Haynes, H. and R. Kryter. “Condition monitoring of machinery using motor current signature analysis”. International Machinery Monitoring and Diagnostic Conference, 1st. (1989).

[33] Pillay, P. and Xu. Z, “Motor current signature analysis.” in Industry Applications Conference, 1996. Thirty-First IAS Annual Meeting, IAS'96., Conference Record of the 1996 IEEE. (1996).

[34] Smith, J.D. - Gears and their Vibration: A Basic Approach to Understanding Gear Noise, NY: Marcel Dekker, Inc., ISBN: 082471797X, (1983).

[35] Geartech- “Gear Failure Atlas.” Townsend, MT, (1999).

[36] Choy,F., Zakrajsek, J., Handschuh, R., and Townsend. D.- “Vibration Signature Analysis of a Faulted Gear Transmission System”, *Journal of Propulsion and Power 12*, (1996): 289-295,.

[37] Barbar, A. –“Handbook of Noise and Vibration Control, sixth Edition”, Elsevier Science Publisher Ltd., UK, (1992).

[38] Smith, J. D. – “Gear Noise and Vibration,” Marcel Dekker, NY, 1999.

[39] McFadden, P. D., and M. M. Toozhy. "Application of synchronous averaging to vibration monitoring of rolling element bearings." *Mechanical Systems and Signal Processing 14.6* (2000): 891-906.

[40] Bartelmus, W., and R. Zimroz. "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions." *Mechanical Systems and Signal Processing 23.5* (2009): 1528-1534.

[41] Loutas, T. H., et al. "Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements." *Applied Acoustics 70.9* (2009): 1148-1159.

[42] Loutas, T. H., et al. "The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery." *Mechanical systems and signal processing 25.4* (2011): 1339-1352.

[43] Lin, Jin, and M. J. Zuo. "Gearbox fault diagnosis using adaptive wavelet filter." *Mechanical systems and signal processing* 17.6 (2003): 1259-1269.

[44] Zheng, Haibao, Z. Li, and X. Chen. "Gear fault diagnosis based on continuous wavelet transform." *Mechanical systems and signal processing* 16.2 (2002): 447-457.

[45] McFadden, P. D. "Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration." *Journal of Vibration and Acoustics 108.2* (1986): 165-170.

[46] Antoni, J. "Cyclic spectral analysis in practice." *Mechanical Systems and Signal Processing 21.2* (2007): 597-630.

[47] Raad, A, Antoni, J and Sidahmed,M. "Indicators of cyclostationarity: Theory and application to gear fault monitoring." *Mechanical Systems and Signal Processing 22.3* (2008): 574-587.

[48] Chen, X, et al. "Research of weak fault feature information extraction of planetary gear based on ensemble empirical mode decomposition and adaptive stochastic resonance." *Measurement 73* (2015): 55-67.

[49] Amarnath, M., and IR Praveen Krishna. "Detection and diagnosis of surface wear failure in a spur geared system using EEMD based vibration signal analysis." *Tribology international 61* (2013): 224-234.

[50] Forrester, D.B. “Advanced vibration analysis techniques for fault detection and diagnosis in geared transmission systems”, Swinburne University of Technology, 1996.

[51] McInerny, S.A., and Yi, D. "Basic vibration signal processing for bearing fault detection." *Education, IEEE Transactions on 46.1* (2003): 149-156.

[52] Samanta, B., and Al-Balushi, A.D. "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features." *Mechanical systems and signal processing* 17.2 (2003): 317-328.

[53] Lin, J., & Zuo, M.J. (2003). Gearbox fault diagnosis using adaptive wavelet filter. *Mechanical systems and signal processing, 17*(6), 1259-1269.

[54] Antoni, J, and Randall,R.B. "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines." *Mechanical Systems and Signal Processing 20.2* (2006): 308-331.

[55] Lai, T. M. et al. "High-impedance fault detection using discrete wavelet transform and frequency range and RMS conversion." *Power Delivery, IEEE Transactions on 20.1* (2005): 397-407.

[56] Heng, Aiwina, et al. "Rotating machinery prognostics: State of the art, challenges and opportunities." *Mechanical Systems and Signal Processing 23.3* (2009): 724-739.

[57] Barszcz, Y, and Randall, R. "Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine." *Mechanical Systems and Signal Processing 23.4* (2009): 1352-1365.

[58] Rai, V. K., and Mohanty A. R. "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform." *Mechanical Systems and Signal Processing 21.6* (2007): 2607-2615.

[59] Gani, A, and Salami, M. J. E. "A LabVIEW based data acquisition system for vibration monitoring and analysis." *Research and Development, SCOReD. Student Conference on. IEEE,* 2002.

[60] Bianchini, C, et al. "Fault detection of linear bearings in brushless AC linear motors by vibration analysis." *Industrial Electronics, IEEE Transactions on 58.5* (2011): 1684-1694.

[61] Hamel, M, Abdulmajid A, and David M. "Investigation of the influence of oil film thickness on helical gear defect detection using Acoustic Emission." *Applied acoustics 79* (2014): 42-46.

[62] Randall, R.B., and Antoni, J. "Rolling element bearing diagnostics—a tutorial." *Mechanical Systems and Signal Processing 25.2* (2011): 485-520.

[63] Qian, S, and Chen, D. "Joint time-frequency analysis." *Signal Processing Magazine,* IEEE 16.2 (1999): 52-67.

[64] Wigner, E. "On the quantum correction for thermodynamic equilibrium." *Physical Review 40.5* (1932): 749.

[65] Rioul, O, and Vetterli, M. "Wavelets and signal processing." *IEEE signal processing magazine* 8.LCAV-ARTICLE-1991-005 (1991): 14-38.

[66] Daubechies, I. "The wavelet transform, time-frequency localization and signal analysis." *Information Theory, IEEE Transactions on 36.5* (1990): 961-1005.

[67] Li, C. J, and Jun M. "Wavelet decomposition of vibrations for detection of bearing-localized defects." *NDT & E International 30.3* (1997): 143-149.

[68] Vetterli, M, and Kovacevic, J. "Wavelets and subband coding." *LCAV-BOOK*. Prentice-hall, 1995.

[69] Braun S., "The extraction of periodic waveforms by time domain averaging, *Acoustica*", 32 (1975) 69–77.

[70] McFadden P.D., "A revised model for the extraction of periodic waveforms by time-domain averaging", *Mechanical Systems and Signal Processing 1 (1)* (1987) 83–95.

[71] Bechhoefer, E., Michael Kingsley, **"**A Review of Time Synchronous Average Algorithms", *Annual Conference of the Prognostics and Health Management Society*, (2009).

[72] McFadden P.D. and Toozhy M., "Application of synchronous averaging to vibration monitoring of rolling element bearing". *Mechanical systems and signals processing, 14*, (2000).

[73] Komgom N. C, Njuki M., Aouni L., Marc T, "On the use of time synchronous averaging, independent component analysis and support vector machines for bearing fault diagnosis", *First International Conference on Industrial Risk Engineering*, (2007),.

[74] Qiu, H., Lee, J., Lin, J., & Yu, G. "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics". *Journal of sound and vibration*, *289*(4), (2006): 1066-1090.

[75] Wuxing L., et al ,"Classification of gear faults using cumulants and the radial basis function network", *Mechanical Systems and Signal Processing, 18 (2) (*2004*)*, pp. 381–389.

[76] Jafarizadeh, R. et al, "Asynchronous input gear damage diagnosis using time averaging and wavelet filtering", *Mechanical Systems and Signal Processing*, 22 (1) (2008): 172–201.

[77] Sato, T. et al, “Real-time bispectral analysis of gear noise and its applications to contactless diagnosis”, *Journal of the Acoustical Society of America 62(2*), (1977): 382–387.

[78] Parker Jr., B.E. et al, “Fault diagnostics using statistical change detection in the bispectral domain”, Mechanical Systems and Signal Processing 14(4) (2000):561–570.

[79] G., Shen, M. Stephen, Y. Xu, W. Paul, “Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears”, *Mechanical Systems and Signal Processing., Mechanical Systems and Signal Processing 43*, (2014): p76:

[80] Kia, S.H., Henao, H. and Capolino, G.A. “Gearbox monitoring using induction machine stator current analysis”. Diagnostics for Electric Machines, *Power Electronics and Drives, 2007. DEMPED 2007*. IEEE International Symposium on. (2007).

[81] Drago, R.J., “Fundamentals of Gear Design”, *Butterworths, Boston*, MA, (1988).

[82] Alfredsson, B., “A Study on Contact Fatigue Mechanisms, in Solid Mechanics.” *Royal Institute of Technology Stockholm*. (2000).

[83] Hajnayeb, A., et al. "Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis." *Expert Systems with Applications 38.8*, (2011): 10205-10209.

[84] Randall, R. B. "A new method of modelling gear faults." *Journal of Mechanical Design 104.2* (1982): 259-267.

[85] Combet, F, and Gelman, L. "An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor." *Mechanical systems and signal processing 21.6* (2007): 2590-2606.

[86] Tuma, J. "Gearbox noise and vibration prediction and control." *International Journal of Acoustics and Vibration 14.2* (2009): 99-108.

[87] Wang W. et al "Assessment of gear damage monitoring techniques using vibration measurements." *Mechanical Systems and Signal Processing 15.5* (2001): 905-922.

[88] Toutountzakis, T., and Mba D., "Observations of acoustic emission activity during gear defect diagnosis." *NDT & E International 36.7* (2003): 471-477.

[89] Staszewski, W. J., K. Worden, and G. R. Tomlinson. "Time–frequency analysis in gearbox fault detection using the Wigner–Ville distribution and pattern recognition." *Mechanical systems and signal processing 11.5* (1997): 673-692.

[90] Bartelmus, W., “Mathematical modelling and computer simulations as an aid to gearbox diagnostics” *Mechanical systems and signal processing*, 2001. 15(5): 855-871.

[91] Amabili, M. and Rivola, A. “Dynamic analysis of spur gear pairs: steady-state response and stability of the SDOF model with time-varying meshing damping”. *Mechanical systems and signal processing*, (1997): 375-390.

[92] Cai, Y. "Simulation on the rotational vibration of helical gears in consideration of the tooth separation phenomenon (a new stiffness function of helical involute tooth pair)." *Journal of Mechanical Design 117.3 (1995): 460-469.*

[93] Draca, S., “Finite element model of a double-stage helical gear reduction”, *University of Windsor:* University of Windsor (2006)

[94] Wojnarowski, J. and Onishchenko, V., “Tooth wear effects on spur gear dynamics. Mechanism and machine theory”, (2003): 161-178.

[95] Yuksel, C. and Kahraman A., “Dynamic tooth loads of planetary gear sets having tooth profile wear”, Mechanism and machine theory, (2004): 695-715.

[96] Flodin, A. and Andersson, S., "A simplified model for wear prediction in helical gears." *Wear 249.3* (2001): 285-292.

[97] Shen, G. et al, “Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears”, *Mechanical Systems and Signal Processing, Volume 43, Issues 1–2, 3*(2014): 76-89.

[98] Jia, S., Howard I. and Wang, J., “The Dynamic Modelling of Multiple Pairs of Spur Gears in Mesh, Including Friction and Geometrical Errors”, *International Journal of Rotating Machinery*, (2003): 437–442.

[99] Rivola A, and White P.R. “Bispectral analysis of the bilinear oscillator with application to the detection of fatigue cracks”, *Journal of Sound and Vibration*. (1998): 889-910.

[100] Aherwar, A, and Saifullah, K. "Vibration analysis techniques for gearbox diagnostic: a review.", *International Journal of Advanced Engineering Technology 3.2* (2012): 4-12.

[101] Kim Y. C, and Powers E. J, “Digital bispectral analysis and its applications to nonlinear wave interactions,” *IEEE Transactions on Plasma Science, vol. PS-7, no. 2*, (1979): 120–131.

[102] Collis, W. B., White, P. R. Hammond, J.K. “Higher-Order Spectra: The Bispectrum and Trispectrum”, *Mechanical Systems and Signal Processing, vol. 12(3),* (1998): 375-394.

[103] Shen, G. et al, “Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears,” *Mechanical Systems and Signal Processing, Volume 43, Issues 1–2,* (2014): 76-89.

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

[105] Alwodai, A. et al,” A Study of Motor Bearing Fault Diagnosis using Modulation Signal Bispectrum Analysis of Motor Current Signals,” *Journal of Signal and Information Processing. 04* (2013) 72–79.

[106] Gu, F. et al “Electrical motor current signal analysis using a modified bispectrum for fault diagnosis of downstream mechanical equipment,” *Mechanical Systems and Signal Processing. 25,* (2011) 360–372.

[107] Naid, A. el al, “Bispectrum Analysis of Motor Current Signals for Fault Diagnosis of Reciprocating Compressors*” Key Engineering Materials,* (2009) 413-41.

[108] Barszcz, T. and JabŁoński, A. “A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram,” *Mechanical Systems and Signal Processing. 25,* (2011): 431–451

[109] Rehab, I. et al, “The fault detection and severity diagnosis of rolling element bearings using modulation signal bispectrum,” *Eleventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Manchester, UK, (*2014).

[110] Lyman, T., “ASM Metals Handbook, Failure of Gears, Vol. 10”**,** *Failure Analysis and Prevention (8th edn),* (1975): 507-524.

[111] Becker, W. T., and Shipley, J. R., "Failure analysis and prevention, ASM handbook, vol. 11." *ASM International* (2007): 718-20.

[112] Brown, J. E., and Jha, C. S. "The starting of a 3-phase induction motor connected to a single-phase supply system." *Proceedings of the IEE-Part A: Power Engineering* 106.26 (1959): 183-190.

[113] El-Hachemi B., "A review of induction motors signature analysis as a medium for faults detection." *Industrial Electronics, IEEE Transactions on* 47.5 (2000): 984-993.

[114] Fitzgerald, A. E. Kingsley, C. and Umans. S.D. “Electric machinery.” *Vol. 5. New York: McGraw-Hill*, (2003).

[115] Construction of DC Machine, 14th July 2014 [cited 31st October 2015]; Available from: http://www.polytechnichub.com/construction-dc-machine/

[116] Li, M., “Using Ac Motor as a Transducer for Detecting Electrical and Electromechanical Faults.” *University of Huddersfield*, (2011).

[117] White C.J, “Detection of gear failure.” Workshop in Online-condition Maintenance. Southampton.(1972).

[118] Naid, A. “Fault Detection and Diagnosis of Reciprocating Compressors using Motor Current Signature Analysis.” *Diagnostic Engineering. The University of Huddersfield: Huddersfield*. (2009): 201.

[119] Kia, S.H., Henao, H. and Capolino, G.A. “Gearbox monitoring using induction machine stator current analysis.” *Diagnostics for Electric Machines, Power Electronics and Drives, 2007. SDEMPED 2007. IEEE International Symposium on*. 2007. IEEE.

[120] Drago, R. J., Fundamentals of Gear Design. Butterworths, Boston, MA, 1988

[121] Alfredsson, B, A “Study on Contact Fatigue Mechanisms”, *Solid Mechanics Royal Institute of Technology Stockholm*. (2000).

[122] Eftekharnejad, B., “Condition monitoring of gearboxes using acoustic emission”. 2010, Cranfield Unversity: School of Engineering (SoE).

[123] Asi, O., “Fatigue failure of a helical gear in a gearbox.*” Engineering Failure Analysis*, (2006): 1116-1125.

[124] Tezcan, S. and Kovan. V., "Pitting failure of truck spiral bevel gear." *Engineering Failure Analysis* 14.4 (2007): 614-619.

[125] Chaari, F., et al. "Effect of spalling or tooth breakage on gear mesh stiffness and dynamic response of a one-stage spur gear transmission." *European Journal of Mechanics-A/Solids* 27.4 (2008): 691-705.

[126] Fakhfakh, T., Chaari, F. and M. Haddar, M. “Numerical and experimental analysis of a gear system with teeth defects.” T*he International Journal of Advanced Manufacturing Technology*, (2005): 542-550.

[127] Ma, J. and Li. C., “On localized gear defect detection by demodulation of vibrations-a comparison study.” *Proceedings of the ASME international mechanical engineering congress and exposition, Part 1, San Francisco, CA*. (1995): 565-576.

[128] Lebold, M, et al. "Review of vibration analysis methods for gearbox diagnostics and prognostics." *Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention Technology*, (2000).

[129] McFadden, P. D., and Smith, J.D. "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 199.1 (1985): 65-70.

[130] Inalpolat, M., and Kahraman, A. "A theoretical and experimental investigation of modulation sidebands of planetary gear sets." *Journal of Sound and* Vibration 323.3 (2009): 677-696.

[131] Ebersbach, S., Peng, Z. and Kessissoglou, N. J. "The investigation of the condition and faults of a spur gearbox using vibration and wear debris analysis techniques." *Wear 260.1* (2006): 16-24.

[132] Staszewski, W. J. "Wavelet based compression and feature selection for vibration analysis." *Journal of sound and vibration 211.5* (1998): 735-760.

[133] Haram, M. Condition monitoring of a multistage helical gearbox using higher order spectra analysis of motor current signals, PhD Thesis, *University of Huddersfield*, (2015).

APPENDICES

A.1 MEMS Sensor

The MEMS accelerometer, which includes both the signal conditioning circuitry and the sensor are fabricated together on a single monolithic chip. The output signals are analogue voltages that are proportional to acceleration. It uses sensing change in capacitance as shown in Figure 2. The deflection of the internal mass changes the capacitance between the finger and the adjacent cantilever beams. The sensor has an electronic device that will convert the capacitance change due to acceleration into a voltage.

Its major benefit is that it is cheap and has low power consumption, this means that measuring vibration no longer needs to be expensive, which is a major disadvantage as summarised in Table 11.1.

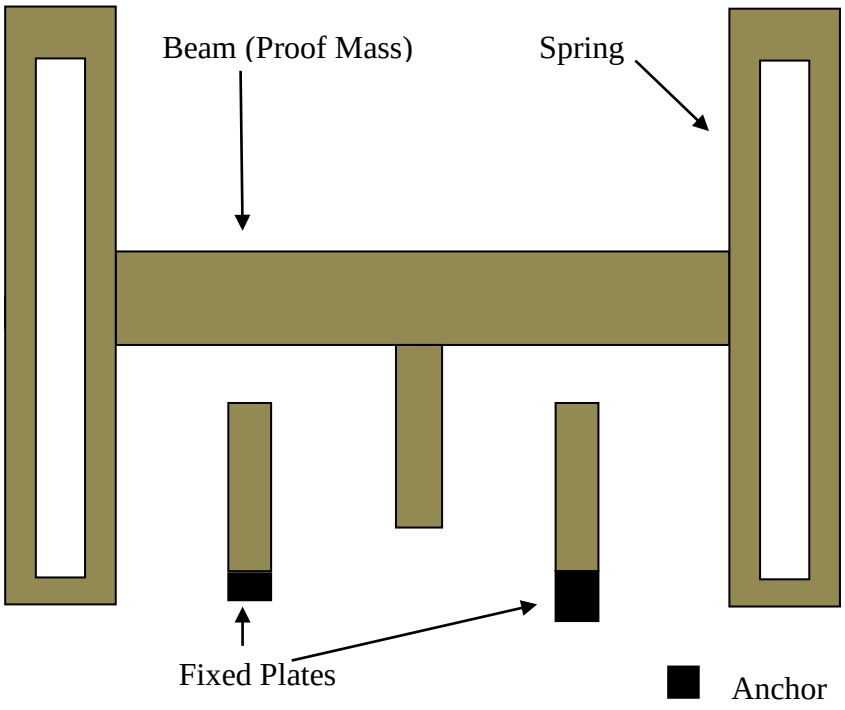

Figure A.1 Capacitive MEMS accelerometer

The choice of accelerometers depends on several factors and some of them are listed below:

Static characteristics of sensors

- Input range : interval between the maximum and minimum of input
- Output range : interval between the maximum and minimum of output
- Span : the interval of output range of a measurement system Span= O_{max} - O_{min}
- Zero: the system output corresponding to a zero input.
- Resolution: is the smallest detectable incremental change of input that can be detected in output signal.
- Sensitivity: is the rate of change in output corresponding to the rate of change in input.
- Repeatability: inability of a sensor to represent the same value under identical conditions.
- Deadband: range of input in which the output remains at zero.
- Hysteresis: the delay phenomenon in output due to energy dissipation.

Dynamic characteristics of sensors

A measure of a sensor's capability of the following rapid change in input:

- Delay (response time)
- Rise time

- Overshoot
- Settling time

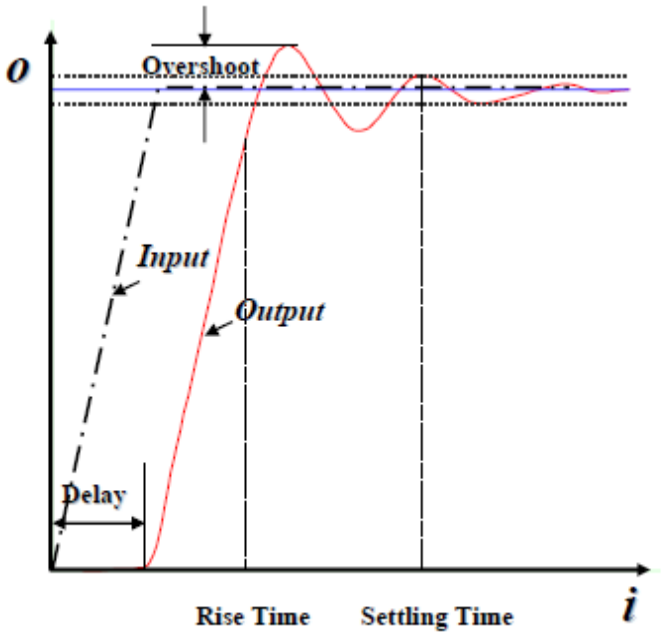

Figure A3 dynamic characteristics of sensor

There are more factors to consider when choosing the sensor such as cost, size, weight, linearity, and environment (temperature, shock, vibration, etc.). Based on these factors, 3 different MEMs sensors were chosen for further testing.

A.2 DsPIC Processor

Due to the cheap cost of DsPIC, it is economically feasible to embed any required number of them within a machine or process[A.1]. Each device may be specially configured to act a sensing system for a particular machine element or process function. Then, they can be linked together to form a distributed sensor network. Digital signal processing applications seek to extract useful information from an incoming set of signals, represented by sequence

of numbers. The algorithms, or computational prescriptions, for extracting this information are originally invented and developed in general-purpose computing environments in which the speed of modifying and retesting the algorithm is more important than the speed of executing the algorithm [A.2].

In a real time signal processing application, the samples of the digital signals arrive at the input of the processing system. The samples arrive over some interval of time, until a group of samples, known as a frame, have arrived that are sufficient to begin processing. Processing then begins to obtain a more information rich version of the sample data. Performing these computations on this sample requires a certain computational time.

Real time signal processing completes the computation of the output associated with a frame of input signal samples in a time that does not exceed the duration of that time [A.3]. In generally, for a real time application, portable implementation is necessary if that algorithm is

to find practical use. The dsPIC microcontrollers can be able to implement such implementations. In this paper, a dsPIC microcontroller based device which allows monitoring and processing of analogue signals such as physiological originated signals (ECG, EMG, EEG, etc.), is proposed. This device consists of following hardware stages; a microcontroller adaptor board, dsPIC development board, (128×64) graphic LCD, and (4×4) keypad. Also, there are five different software modules which are ADC, FFT, graphic LCD, keypad, and filter programs. The proposed device takes analogue signals via an ADC module and processes them with dsPIC by using keypad commands and then allows to plot at graphic LCD.

A.3 Software Modules

In the system design, the speed of computation and memory capacity are considered as two most important characteristics. Since dsPIC30F6010 device has these properties, it has been chosen for our design. This chip has the following specifications;

- 30 MIPs processor speed
- 10 bit ADC
- 4 Kbyte EEPROM
- 8 Kbyte SRAM
- 144 Kbyte program memory
- 24 bit instruction bus
- 16 bit data bus
- 1 clock cycle DSP processing
- Optimised instruction architectures featuring flexible addressing modes.
- Microchip MPLAB v2.50 is used for software modules. Then, C codes are compiled by using C30 compiler. The following sub-modules constructed the main software module;
- ADC program
- FFT program
- Graphic LCD program
- Keypad program
- Filter application program

A.4 ADC Program

To process the analogue signals coming from some physical originated sources, those signals must be converted to the digital form via ADC. For this aim, we used a 10 bit A/D converter module [A.3]. The conversion speed is adjusted by SAMC (Auto-Sample Time Bits) bits which are located in ADCON3 (A/D Control Register 3). After that, they obtained 16 bits binary value is written in ADCBUFO register. The ADC timing diagram is given in Figure 1. Since dsPIC signal processing function were written with signal fractional numerical format, we have used this format for all signal processing functions that we developed. For some applications, we have formed signal sequence with different lengths. For example, the FFT calculation uses a sequence with 512 samples.

A.5 FFT Program

In order to adjust FFT length, FFT_BLOCK_LENGTH and LOG_BLOCK_LENGTH coefficient must be defined in "parameter.h" user defined library. This module can be able to calculate a FFT process with 64, 128, 256, and 512 sample lengths. DSP processing cycles are given in Table 2 and FFT processing can be implemented as following way [A.4]:

- Initialization: Twiddle factor coefficients are defined.
- Input signal is scaled to be located at between the -0.5 and 0.5 interval.
- Butterfly computation is implemented by calling "FFTComplexIP ( )" function from microchip signal processing library.
- Bit-reversed RE_ordering is a part of FFT algorithm. In this routine, FFT sequence is re-ordered in bit-reversed manner.
- Square the magnitude computation of every complex entity for the FFT outputs.

- The peak value searching is implemented by “vector max ( )” function.

Table A.2 DSP Processing cycles [A.4].

Function	Cycle Count Equation	Conditions*	Number of Cycles	Execution Time @30 MIPS
Complex FFT**	—	N=64	3739	124.6 µs
Complex FFT**	—	N=128	8485	282.8 µs
Complex FFT**	—	N=256	19055	635.2 µs
Single Tap FIR	—	—	1	33 ns
Block FIR	53+N(4+M)	N=32, M=32	1205	40.2 µs
Block FIR Lattice	41+N(4+7M)	N=32, M=32	7337	244.6 µs
Block IIR Canonic	36+N(8+7S)	N=32, S=4	1188	39.6 µs
Block IIR Lattice	46+N(16+7M)	N=32, M=8	2350	78.3 µs
Matrix Add	20+3(C*R)	C=8, R=8	212	7.1 µs
Matrix Transpose	16+C(6+3(R-1))	C=8, R=8	232	7.7 µs
Vector Dot Product	17+3N	N=32	113	3.8 µs
Vector Max	19+7(N-2)	N=32	229	7.6 µs
Vector Multiply	17+4N	N=32	145	4.8 µs
Vector Power	16+2N	N=32	80	2.7 µs
PID Loop Core	—	—	7	231 ns
*C= #columns, N=# samples, M=#taps, S=#sections, R=#rows **Complex FFT routine inherently prevents overflow				
1 cycle = 33 nanoseconds @ 30 MIPS				

A.6 Wireless Sensor Network

Wireless Sensor Networks (WSN) are a trend of the last few years due to the advances made in the wireless communication, information technologies and electronics field. The developments of low-cost, low-powered, multifunctional sensors have received increasing attention from various industries [A.6]. WSN is a wireless network composed of autonomous and compact devices called sensor nodes or motes.

A sensor network is designed to detect desired phenomena, then collect, process the data and transmit this information to users. Sensor nodes or motes in WSNs are small sized and are capable of sensing, gathering and processing data while communicating with other connected

nodes in the network, via radio frequency (RF) channel. The sensor nodes scattered in a sensor field where each sensor nodes collects data and route the data back through a multi-hop hybrid wireless communications. The design of the sensor network is influenced by factors including scalability, operating system, fault tolerance, sensor network topology, hardware constrains, transmission media, and power consumption [A.7]. The developments is low-cost, low-powered, multifunctional sensor have received increasing attention from various industries [A.6].

There are two kinds of sensor used in the network. First is the normal sensor node deployed, the other is a gateway node that interfaces sensor network to the external world. Sensors which measure vibration, acoustic emissions, light and temperature are among the types of sensor being used depending on the application. The system architecture of a sensor node consists of a radio transceiver or optical as the communication unit, microcontroller for the processing unit, sensor as the sensing unit and battery as the power unit. The hardware device in the sensing unit may consist up to several sensors. This device produces measurable response to change which acts as an interface between motes to the environment. The processing unit or control unit (CU) is responsible for the collecting and processing of the captured signal from the sensor unit. These signals are then transmitted to the network. It determines both energy consumption as well as computational capabilities of the sensor node. The power unit consisting of battery supplies power to the sensor node. It is important to choose the battery type since durability will affect the design of sensor node [A.7].

WSNs has been used in high-end application such as radiation and nuclear-threat detection systems, weapons sensors for ships, biomedical applications, habitat sensing and seismic monitoring [A.8]. Measurable changes are vibration, temperature, sound, motion, pollutants or pressure in environmental conditions [A.7]. There is also the possibility of a ultra-

compact, high data-rate wireless sensor node is constructed using eCAM miniature camera, plugging a VGA video camera to an Eco node. The purpose is to show that high data-rate wireless capability and expandability of a platform can came in a miniature package. The camera module itself already performs compression in hardware thus represents an optimized subsystem in terms of its power consumption and bandwidth demand [A.9], [A.10]. This is however expensive and is beyond what is proposed for this project.

A low cost, more readily available solution is MiWi, its protocol is based on the MAC and PHY layers of the IEEE 802.15.4 specification [A.11], and is tailored for simple network development in the 2.4 GHz and sub-GHz ISM frequency bands. The protocol provides the features to find form and join a network, as well as discovering nodes on the network and route to them. It does not cover any application-specific issues, such as how to select which network to join to, how to decided when a link is broken or how often devices should communicate. This flexibility can minimise power consumption if properly implemented. The low cost and power consumption of the MiWi protocol makes it ideal for providing wireless communication in the WSN allowing several smart sensors on a wind farm to communicate with each other.

References for Appendices

[A.1] R.A. Siddiqui, W. Amer, Q. Ahsan, R.I. Grosvenor, P.W. Prickett, "Multi-band infinite impulse response filtering using microcontrollers for e-monitoring applications", Microprocessors and Microsystems, 2007, Vol:31, pp. 370-380.

[A.2] J.G. Ackenhusen, Real time signal processing, Prentice Hall, New Jersey, 1999.

[A.3] Microchip Technology Inc., DS70119D – dsPIC30F6010 Datasheet.

[A.4] R.G. Lyons, “Understanding Digital Signal Processing”, The Fast Fourier Transform, 1997,pp. 129-154,.

[A.5] Microchip Technology Inc., DS51456C - 16-Bit Language Tools Libraries, pp. 58-72. [33] Y.M.Yussoff, R.A.Rahman, N.H.Abidin, F.H.Yahaya, “Performance Analysis of Wireless Sensor Network”, Conference on Signal Processing and Applications (CSPA 2009).

[A.6] I. Mahgoub, M. Ilyas, ‘Sensor Network Protocols’, CRC Press, 2006, pp.1-248

[A.7] W. S. Lan F.Akyildiz, yogesh Sankarasubramaniam, Erdal Cayirci, “A Survey on Sensor Network,” IEEE Communication Magazine, pp. 102, August 2002, 2002.

[A.8] K. Sohraby, D. Minoli and T. Znati, ‘Wireless Sensor Network: Technology, Protocols and Applications’, John Wiley & Sons, 2007,pp.1-300 http://www.ulb.tu-darmstadt.de/tocs/195087267.pdf

[A.9] COMedia. C328-7460 JPEG compression camera module. http : // is6.pacific. net. hk / comedia / C328-7640. PDF.

[A.10] Chulsung Park, Pai H. Chou, “ Eco; Ultra-Wearable and Expandable Wireless Sensor Platform, In Third International Workshop on Body Sensor Network (BSN 2006), April 2006.

[A.11] D.Flowers and Y.Yang, "MiWi Wireless Networking Protocol Stack," 2008.