POLITECNICO DI TORION

Master course in

MECHANICAL ENGINEERING

Master's thesis

Damage detection on rolling element bearings based on different numerical analyses

Supervisor

Prof. Fasana Alessandro

Candidate

Bingzheng Lu

Academic year 2018/2019

[Abstract] In monitoring and maintaining rotary machines, it is extremely important to identify the defects of rolling element bearings (REBs), which may influence the production efficiency and regular service. The presentation of the prevalent defects of REBs and common detection methods are briefly introduced in this thesis. And several methods of enhancing bearing signal namely envelope analysis, spectral kurtosis and so on are presented in detail. A recently proposed pre-processing methodology to extract the novel signal, a signal that contains information that is not present in the historical reference data is described emphatically. Also, the procedure of verifying the sensibility of novel signal to damage is performed by numerical modelling consisting of signal generation, novel signal extraction and envelope analysis. The results indicate that the novel vibration signal is not always so sensitive to damage, which provides a promising diagnostic potential for detection of bearing defects. The final conclusion is drawn based on comparison among Fast kurtogram, Autogram and the proposed method.

[Key Words] Rolling element bearings; defects; novel information; Fast kurtogram; Autogram

Acknowledgement

Foremost, I would like to express my profound gratitude to my supervisor Fasana Alessandro for guiding me into a constructive journey of thesis with patience and motivation. The door of his office was always open for me to get over a difficulty throughout the whole process of researching and revising the thesis.

My sincere thanks also go to my parents for their unconditioned love and perpetual encouragement which makes me the person who I am. When encountering with big decision concerning the trajectory of my life, I could invariably follow my heart and chase my dream with the continuous support of them. I wish to accompany through the rest of my life in reward for their devotion of bringing me up.

Last but not the least, I would seize the opportunity to convey my deep sense of gratitude to Xiaoying Chen for accompanying me to study, helping me to debug the programs, communicating about musicals and operas, and cooking tasty desserts. It was my fortune to have a friend understanding and supporting me during plenty of tough times, to whom I am appreciatively indebted.

Contents

1. Introduction

Rolling element bearings are critical mechanical components in rotating machinery, existing in a broad range of applications across almost all industries, which enable rotational or linear movement, besides reduce frictions and handle stresses. Therefore, the condition monitoring and fault diagnosis of the rolling element bearings is vital to maintenance strategy and operational safety. Any defects in the bearing must be identified in time to avoid increase in downtime in case of catastrophic failure.

The layout of this chapter is organized in a systematic way to cover general information concerning REB, namely essential structure, working principle and common defects. In section 1.3 and 1.4 the characteristics of different types of defects and the features of generated signals are introduced theoretically. Section 1.5 relates the reason why local faults are chosen to be analyzed in the thesis and the detailed process of diagnostic analysis is demonstrated in Chapter 4 and 5.

1.1 REB

As defined in Wikipedia, a rolling element bearing is a bearing which carries a load by placing rolling elements (such as balls or rollers) between two bearing rings called races. The terms rolling element bearings, antifriction bearings and rolling bearings are used to describe that class of bearings in which the main load is transferred through elements in rolling contact rather than in sliding contact. The robustness of bearing is of importance for the operation condition where bearings are subjected to heavy and dynamic loadings generated by machines or transmitted through the components of REBs.

In general, a bearing consists of outer ring, inner ring, cage and rolling elements. The three principal dimensions are outside diameter, bore size and width which are demonstrated in Fig. 1 together with main components. A variety of bearings are designed for all kinds of applications with diverse advantages and disadvantages. The main criteria for selecting the types are motions and loads a bearing can preferably support. As depicted in Fig. 2 the load zone is distributed associated with a unidirectional vertical load (outer race is fixed).

1.2 Defects of REBs

Nomenclature

REB	Rolling element bearing
DOF	Degree of Freedom
EHL	Elasto-hydrodynamic Lubrication
FCF	Fault Characteristic Frequency
IAS	Instantaneous Angular Speed
BPOO	Ball Pass Order on the Outer Race
IFCF	Instantaneous Fault Characteristic Frequency
FCO	Fault Characteristic Order
AE	Acoustic Emission
CNN	Convolutional Neutral Networks
FBG	Fiber Bragg Grating
SNR	Signal to Noise Ratio
SES	Squared Envelope Spectrum
SK	Spectral Kurtosis
STFT	Short-Time Fourier Transfer
ELMD	Ensemble Local Mean Decomposition
SC	Spectral Correlation
CMS	Cyclic Modulation Spectrum
Fast-SC	Fast Spectral Correlation
FFT	Fast Fourier Transform
ACP	Averaged Cyclic Periodogram
CPU	Central Processing Unit
FK	Fast Kurtogram
MODWPT	Maximal Overlap Discrete Wavelet Packet Transform
AC	Autocorrelation
OT	Order Tracking
cICA	Constrained Independent Component Analysis
PDF	Probability Density Function

Through the whole service life of REBs a large number of factors may exert an influence on the premature bearing failure. And the defects are going to be expounded in this section.

1.2.1 Common faults

Defects of bearings may vary in different ways and during different stages. In the early phase bearing faults usually start as small pits or spalls, and give sharp impulses covering a wide range of frequency. Besides, the faults can be caused by excessive load, true or false brinelling,

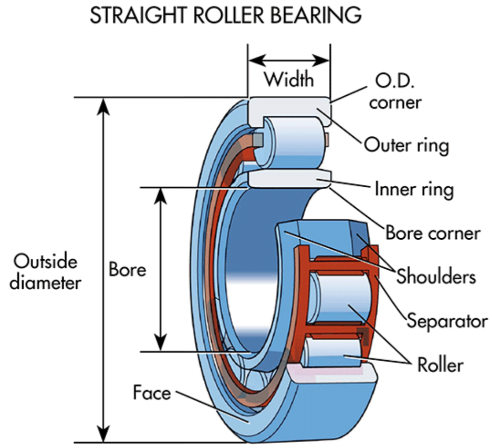

Fig. 1. Fundamental components and principle dimensions for a straight roller bearing.

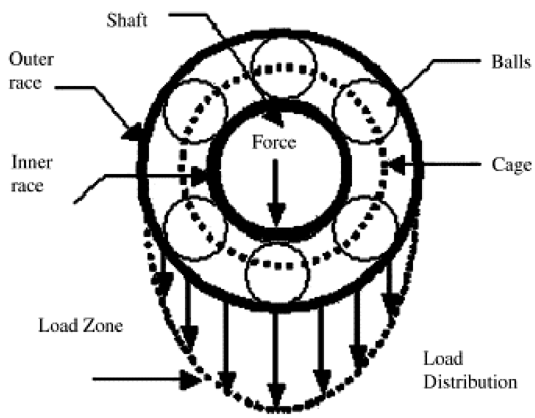

Fig. 2. Load distribution of a bearing under a unidirectional vertical load

misalignment and incorrect design. Near the end, the bearing failure is accounted for corrosion, contamination, abrasive wear, overheating and poor lubrication. It is quite necessary to identify these defects and to analyze the vibration signals generated by specific components of bearing.

1.2.2 Cage damage

Cages or retainers are usually made of mild steel, bronze or brass and can be easily damaged, which may result in bearing premature problems. On account of the low mass, the defects of cages are not visible unless occurring in manufacturing process. According to the manual guide [7], the wear of cage is susceptible to starved lubrication and contamination, excessive speed, roller skewing and tilting. Moreover, fracture of cage connection, cage fracture and the damage due to incorrect mounting are also the usual inducements for cage damage. During cage failure, signals obtained are in the form of random vibration bursts.

1.2.3 Localized and extended

Simply, bearing defects can be categorized as localized and extended faults. For local types pits, cracks and spalls are included as a result of fatigue over the rolling surface, of which spalling can be obviously discovered. Spalling is the pitting or flaking away of bearing material, which can be classified as geometric concentration spalling, point surface origin spalling and inclusion origin spalling. Once initiated this type of failure will spread and propagate to a larger area. Bentley [10] states that 90% of the total bearing faults involve damage of the inner ring, outer ring and rolling elements due to the localized defects. Fig. 3 illustrates typical modulation patterns for unidirectional (vertical) load on the bearing, at shaft speed for inner race faults, and cage speed for rolling element faults (from Ref. [1]), where D is the pitch diameter and d is the ball diameter.

In contrast, the distributed faults encompass surface roughness, waviness, and misaligned races and off size balls. Primary causes involve manufacturing error, abrasive wear, and improper installation which take place in production and operation. In addition, when local defects grow along the periphery of raceway under variable loading, it is thought as extended raceway defect by Sham Kulkarni [9], on which little attention has been captured. In the paper, the feature is extracted by using time domain analysis.

1.3 Defect frequency

The frequencies related to the repetition of impulses generated by concentrated defects are called characteristic defect frequencies. Every bearing has its own characteristic frequency which depends on rotating speed and the size of components. Defects in rolling element bearings give rise to impulses as the elements interact with the fault and the typical vibration is produced. The impulses are inclined to be generated almost periodically and their characteristics diversely depend on the location of the defect and the position of the load zone.

The formula for the various frequencies shown in Fig. 3 are as follows:

Ballpass frequency, outer race:

$$\mathrm{BPFO} = \frac{\mathrm{n} \cdot f_r}{2}\left(1 - \frac{\mathrm{d}}{\mathrm{D}}\cos\phi\right)$$

Ballpass frequency, inner race:

$$\mathrm{BPFI} = \frac{\mathrm{n} \cdot f_r}{2}\left(1 + \frac{\mathrm{d}}{\mathrm{D}}\cos\phi\right)$$

Fundamental train frequency (cage speed):

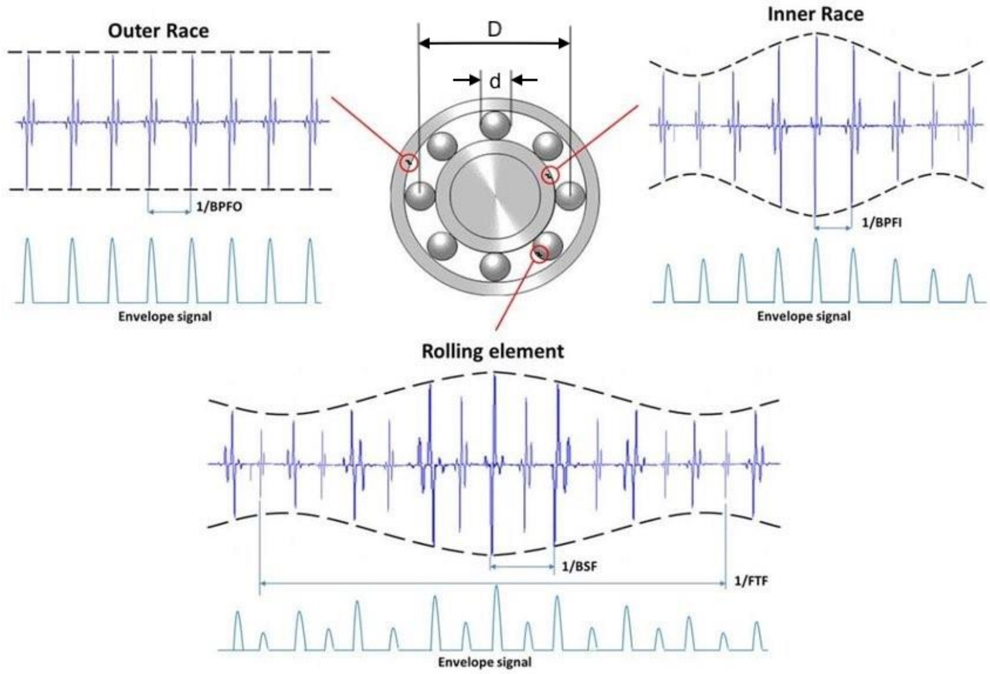

Fig. 3. Typical signals and envelope signals from local faults in rolling element bearings from Ref. [1]

$$\mathrm{FTF} = \frac{f_r}{2}\left(1 - \frac{\mathrm{d}}{\mathrm{D}}\cos\phi\right)$$

Ball (roller) spin frequency:

$$\mathrm{BSF(RSF)} = \frac{\mathrm{D}}{2\mathrm{d}}\left[1 - \left(\frac{\mathrm{d}}{\mathrm{D}}\cos\phi\right)^2\right]$$

where f_r is the shaft speed, n is the number of rolling elements, and ϕ is the angle of the load from the radial plane.

1.4 Bearing fault models and cyclostationarity

Numerical modelling based on the real prototype of bearing defects serves as a valid tool in bearing diagnostics to reconstruct representative signals for further operation. However, the effectiveness of the model depends on how refined the building process is and whether the real case is in accordance with the considered. In this section, bearing models are going to be described respectively regarding the development history of them. Furthermore, a kind of characteristic signal applicable for signal analysis, cyclostationary signal, is introduced together with the application in bearing defects.

1.4.1 Bearing faults modelling

1.4.1.1 Mathematical (analytical) model

A great number of scholars have come up with various models of REBs in history. Sunnersjo proposed the first mathematical bearing vibration model where a 2 DOF system was constructed, which provides the load-deflection according to Hertzian contact theory in 1978[11]. The mass and inertia of the rolling elements were ignored. McFadden and Smith built the most utilized mathematical model for rolling element bearing localized faults in 1984[13], which is recognized as the first valid model. Referring to the model local bearing faults comprised a sequence of high frequency bursts, which represent the impulse response of the signal transmission path within short duration. The impulses repeat at a rate given by the fault interacting with the rolling elements, whether it is on the inner race, outer race or rolling elements. Su & Lin (1992) [14] developed the previous model by considering variable load due to shaft and roller errors. Those variations however resulted in some different effects in the bearing signals from those previously obtained. Tandon and Choudhury [15] presented an analytical model to predict the discrete spectrum which has peaks at the characteristic defect frequencies and their harmonics. Ho & Randall (2000) [16] modelled the bearing fault vibrations as a series of impulse responses of a single-degree-of-freedom system which varied the spacing between the bursts randomly by a small percentage. Randall and Antoni introduced the slip between the rolling elements into the vibration fault signal model [1]. This slip will triger a random fluctuation among the impulses due to the defect in the bearings. In 2011 Tadina [17] developed an improved bearing model in order to investigate the vibrations of a ball bearing during run-up.

1.4.1.2 Dynamic model

In a discrete dynamic system, basic elements are mass, stiffness, damping and external forces. Hence to build a representative model these components are required to be considered and incorporated. The first publication to complete a dynamic model of rolling element bearings was issued by Gupta (1975) [18] through solving the generalized differential equations of motion of the balls in an angular contact ball bearing. In the paper critical speed, mode shape and unbalance response of a dual-rotor system was studied via experiment of a dual-rotor rig with an inter-shaft bearing. Fukata et.al (1985) [12] introduced a comprehensive model with 2 DOF system, which provides the load deflection relationships. It addressed the non-linearity and the time variant characteristics of rolling element bearings under condition of ignoring the mass and the inertia of the rolling elements. Fukata's model was further developed to a bearing–pedestal model (4 DOF model) by Feng et. al (2002) [19] considering the effect of slippage of the cage and rolling elements as well as the effect of localized faults (spalls) in the inner and outer race. And then Sawalhi et.al (2006) [20,21] updated Feng's model by introducing an extra degree of freedom (sprung mass system to excite a typical high frequency resonance of the system) and by changing the way of modelling the spall. Tiwari et al. [22,23] studied the effect of the ball bearing clearance on the dynamic response of a rigid rotor system. In 1999 Wijnant et al. [24] considered the effect of the elasto-hydrodynamic lubrication (EHL) and developed computational models for both EHL problem and the structural dynamics problem to explore the influence of the EHL on the REB dynamics. Sopanen & Mikkola's model [25,26] contained the effect of different geometrical faults (surface roughness, waviness and localized and distributed effects) and the EHL. This is a six DOF dynamic model of the deep-groove ball bearing, of which both the non-linear Hertzian contact deformation and the elasto-hydrodynamic fluid film are included.

1.4.2 Cyclostationarity

Broadly speaking, vibration signals can be divided by deterministic (i.e., whose behavior can be described exactly by an equation) or random (i.e., whose behavior cannot be predicted exactly), or a combination of both. Deterministic signals are further categorized as periodic and non-periodic, and random signals as stationary (whose average properties do not change with time) and non-stationary. Cyclostationarity represents a further category including signals which, although not necessarily periodic, are produced by a hidden periodic mechanism. This consists of periodic signals as a special case, but also stationary signals and non-stationary signals which exhibits periodicity after passing through a non-linear transform. As such, cyclostationarity is capable of comprising most signals generated by rotating and reciprocating machines.

Cyclostationary process was firstly put forward by Gardner [27], who describes a random process with a periodic autocorrelation function. An approach based on the properties of cyclostationary processes has been suggested by Antoni and Randall (2002) [28], which showed how to distinguish the modulation associated with a bearing fault from that related to a gear fault.

Degrees of cyclostationarity higher than one and two is usually referred to as ''higher-order'' Strictly speaking, a *n*th order cyclostationary signal is one whose *n*th order statistics are periodic or, equivalently stated, one which produces a peak in its Fourier transform after passing through any non-linear transformation involving nth power. First-order cyclostationary signal is simply a signal which contains periodic components, and its mean value in the ensemble average sense reproduces that properly. Similarly, a second-order cyclostationary signal is one whose autocorrelation function is a periodic function of time.

Bearing signals with local faults are described as 2^{nd} order pseudo-cyclostationary due to the fact that they don't have a defined mean period, and consequently their autocorrelation function is not truly periodic, because of the non-stationarity in the inter-arrival times of the successive impacts. As for extended faults existing on the inner race, it periodically enters and exits the load zone, and the resulting signal is modulated by the shaft speed. This has been described as a purely cyclostationary process as opposed to the former. In this way, both fault types give rise to signals that can be treated as cyclostationary and the statistics are obtained by ensemble averaging over an ensemble of realizations. In practice the optimum way to analyze a faulty bearing signal depends on the type of fault present.

1.5 What we concern: detecting local faults

According to the above elaboration, signals generated from local faults can be detected more feasibly than distributed ones. Herein, the signal analysis and numerical modelling will be focused on local faults in Chapter 4 and 5 so as to derive clear results.

2. Damage detection of REB

Varying symptoms can be manifested while REBs are damaged due to various causes, which give us an idea to detect the damages in specific ways. For bearings working in high speed conditions, temperature rise occurs frequently as a result, which has been evaluated and modeled by a number of researchers. And the lubricant performance is prone to be destroyed by grease degradation. On top of that, a standard test is also described in Section 2.3. At last several techniques for bearing diagnosis are introduced in section 2.4.

2.1 Temperature rise

In a general way, a bearing should be working in an applicable condition where the temperature is lower than 93.3°C (200°F). An excessively high temperature rise has a great influence on bearing life and reliability. Therefore, it is necessary to study heating mechanism, heat transfer process, curvature coefficient of the inner ring f_2 (radius of inner race over ball diameter) and temperature distribution of bearings. Referring to the factors that affect the temperature rise, rotational speed, preload, operating condition, lubricant viscosity and so on are normally considered.

As stated in [35], under the conditions of high speed and light load, skidding, a tribo-dynamic phenomenon often occurs. With a significant effect on the thermal distribution and service reliability of the bearings, it is caused by the sliding of rollers in the direction of motion, while rollers enter the high-load contact zone with insufficient lubrication. According to results of Junning Li's model revealing relationship between skidding and thermal distribution, the inner ring raceway has the highest temperature, whereas the cage has the lowest.

In order to exactly evaluate the heat generation and temperature in the bearing, a model for computing the temperature of the high-speed ball bearing with axisymmetric load is set up by L Q Wang [36]. In this model, heat sources in the raceways are treated as moving and the heat produced by the ball and cage is averaged on the circumferential zone where the heat source passes. It turned out that the contact surfaces in the raceways and the ring land (shoulder) are the high temperature zone. The curvature coefficient of the inner ring also has a great influence on both the heat generated owing to spin motion and the maximum temperature rise in the inner ring.

2.2 Grease degradation

Greases are widely used in ball and roller bearings for lubricating moving surface. In terms of ingredients, there are three components that form lubricating grease: oil, thickener and additives. And feasibility of regular work is assured by several characteristics, for example, pump ability, water resistance, dropping point, oxidation stability, temperature effect and so on. As in inadequate lubrication condition, there isn't a sufficient amount of bearing lubricant to separate the rolling and sliding contact surfaces during service, which may give birth to negative effects, e.g., discoloration, scoring and peeling, excessive roller end heat, total lockup and so forth.

Lubricating grease suffers severe physical and chemical degradation during operation in a bearing [38]. These changes are not solely due to high temperature but are rather the result of combined thermal and mechanical effects, compounded by the presence of metal debris and moisture. As a result, the grease performance can deteriorate and under severe conditions this can lead to failure. Temperature in excess of 204.4°C (400°F) can anneal the ring and ball materials, which is fatal to the regular service. The monitoring under caution temperature 82.2-93.3°C (180-200°F) has been considered unreliable because in the meanwhile so many variables such as ambient temperature, speed, load and runtime have a pronounced influence on bearing temperature.

P. M. CANN [39] has conducted an experiment in the modified DIN 51 806 test rig R2F(M) to test the performance of grease which provides detailed information. After running for different temperature and speed conditions up to 300 hours, lubricant remaining in the cage pocket region was heavily degraded and contained very little thickener. the result shows that the grease on the seals contained different amount of thickener located in the seal position. The lubricant remaining on the inner raceway surface was predominately base oil although there was some thickener in existence. Specially, the technique infrared spectroscopic was used to characterize the degree of oxidation and the degradation of the grease both in the bulk sample as well as from thin grease layers remaining on the bearing surfaces.

2.3 Standard bearing test

Mobil [5] proposed a better way to define high-temperature performance capability which can lead to a standardized bearing test. The test is conducted under accelerated operating conditions to promote grease ageing process. Factors limiting grease high-temperature performance include the degradation resulting from thickener as well as base oil oxidation and the loss of base oil due to grease bleed or evaporation. To evaluate the high-temperature limit, bearings mounted in five identical rigs are run in parallel. The hours to grease failure in each rig can be treated using Weibull statistics to determine the time at which 50 percent of the bearings are expected to fail. This defines the "L50" life of the candidate grease at the corresponding test temperature, which has something in common as "L10" life of bearings i.e., the life at which ten percent of the bearings in that application can be expected to have failed due to classical fatigue failure.

2.4 Techniques to detect damages

Considerable effort has been devoted to detect damages in bearing by a host of scholars. This section explains why acceleration signal is acquired for diagnostic analysis and then it goes with a couple of techniques widely used in damage detection. Two methods for speed variation conditions are introduced in Section 2.4.2 followed by some other newly presented techniques.

2.4.1 Acceleration signal

The vibration signals collected from bearings contain rich information with respect to machine health conditions, which is worth deeply studying. According to the guide [40], apart from detecting vibratory acceleration signal, it’s optional to measure velocity and displacement. Whereas the choice of parameter is important if the signal covers a large span of frequency.

Measurement of displacement gives the most weighted low frequency components and conversely acceleration measurement weighs the level towards the high frequency. Generally, with electronic integrators we can convert the acceleration signal to velocity or displacement by integrating in time domain.

Therefore, the best way to collect the original data is using an accelerometer to measure the raw acceleration signal. There are a certain key factors that influence the success of acceleration measurement. Firstly, the characteristics of accelerometer should be carefully selected in a suitable way ahead of measurement. It's indispensable to take sensitivity, mass and frequency range into account. Secondly, it is highly recommended to mount the accelerometer as close to the bearing as possible, preferably on a flat, clean surface to guarantee consistent results. Finally, once accelerometer has been installed and calibrated, data acquisition should then be conducted with regular intervals over a period of time to improve the accuracy of measurement. In addition, it's better to eliminate the interference from temperature, cable noise, base strains and so on.

2.4.2 Speed variation condition

In constant speed condition, the transient impulses of the vibration signal excited by a localized fault behave like periodic or quasi-periodic. The repetition frequency of the impulses is called fault characteristic frequency (FCF). A great many methods have been proposed to diagnose bearings based on stationary assumption, thus they are not applicable to the fault diagnosis of under rotating speed variation conditions. This constraint significantly limits the bearing diagnosis in industrial production. What follows in the section is the application of two methods to extend the bearing diagnosis to a more generalized case.

For many years, measurement of Instantaneous Angular Speed (IAS) signal has revealed a great sensitivity to different types of defects such as bearing or gear faults over a large bandwidth of orders. The main advantage lies on the fact that the signal is sampled angularly rather than with a constant time step, which allows very long duration measurement without disturbances from non-constant operating speed of the machine. Moreover, this signal investigates different transfer path from excitation to response of the rotating machine. On the basis of the previous method, Adeline Bourdon [42] reconstructed the IAS variations signal to analyze the shape of rotating speed fluctuations introduced by bearing faults. In the analysis, modifications due to operating conditions and size of the defect were clearly separated. With the assistance of a signal processing tool based on the use of a filter defined in the angular frequency domain, magnitude of the speed variations related to the BPOO (Ball Pass Order on the Outer race) of the bearings under monitoring was quantified. Furthermore, the speed variations due to an evolution of the defect and those due to variations in operating conditions were distinguished. As shown in the result, the proposed tool could be used in stationary or non-stationary operating conditions.

On consideration of removing the instantaneous rotating speed information from instantaneous fault characteristic frequency (IFCF) to uncover fault characteristic order (FCO), Yi Wang [43] proposed the method of rotating speed isolation. The results of simulation and experiments displayed that bearing faults can be detected under speed variation condition without the use of tachometers and the method outperformed even the conventional envelop analysis. In this way, it can serve as a promising approach for bearing faults detection.

Besides, inner race defect gives better response to speed variation, and defect from outer race exhibits higher sensitivity to high load [44].

2.4.3 Acoustic emission

Acoustic emission(AE) is the phenomenon of radiation for elastic waves in solids that occurs when a material undergoes mechanical or thermal stresses. Born of the characteristic, AE technique, a nondestructive inspection technique that permits the evaluation of the states of materials by measuring elastic stress waves, can be applied to detect damages in bearing. With high sensitivity to deformation and fracture, it behaves a significant tool for condition monitoring, whose instruments consist of a transducer, mostly of the piezoelectric type, a pre-amplifier and a signal processing unit. By measuring and analyzing AE signals produced by tribological processes, the state of sliding surfaces on a machine can be identified and evaluated. Besides, AE signals can also capture essential diagnostic information from low-energy signals.

Alan Hase et al [45] have done an experiment by measuring high-frequency component of AE signals that originate from adhesion higher than 1 MHz to detect the early seizure. In the study, friction and wear in a high-speed sliding bearing were examined using a test rig on behalf of a real machine. Changes in the amplitude and frequency components of AE signals detected during the rupture of the lubricating film, the progress of wear, and phenomena preceding seizure are described. This was a fundamental study on the lifetime assessment of a sliding bearings, and it proved to be an effective way on early detection of seizure in machineries.

It is also feasible to combine AE signals with other techniques to constitute a bearing diagnostic method. For example, Md Junayed Hasan [46] presented a reliable fault diagnosis scheme based on acoustic spectral imaging (ASI) of acoustic emission(AE) signals, which involved transfer learning technique, convolutional neutral networks (CNN), and spectrum imaging. Thus it turned out to be a robust technique with high diagnostic accuracy, which is validated both in simulation and experiment. The effectiveness of AE technique has been investigated continuously by researchers, which still can be explored in other fields.

2.4.4 Fiber-optic sensor

A pioneering research on bearing diagnosis was provided by Hasib Alian [48] to measure strains inside the rings with Fiber Bragg Grating (FBG) fiber-optic sensors. There are at least two excellent capabilities for FBG sensors, for example, the damage size of small spalls both in inner or outer races can be measured accurately, and the sensors are small enough to mount in places close to the bearing. The disturbance effect on transmission path can be minimized so that enhanced signal-to-noise ratio is guaranteed. The parameters in the experiment comprised loads, rotation speeds, sensor locations, faults of different sizes, both on the inner and outer races and so on. Ultimately, the result of experiment demonstrated that the sensor appeared to be a good tool for bearing detection with great discrimination power.

2.4.5 Wear debris detection

Wear debris analysis, an outdated technique from the perspective of today, was recognized as an important and direct indicator of the wear state of components such as gears and bearings. Herein only the fundamental principles are presented in short. Kuhnell BT [49] indicated in the paper that besides using sensors to detect metallic particles in the lubricant, spectrographic analysis of different metallic elements can facilitate the location of fault. Commonly, the wear process of a machine occurs owing to diverse reasons concerning operating environment, load condition and dynamic behavior. In that case, certain parts may break or disturb the normal work of the machine, which has negative effect on performance. Hence it is of significance to monitor the amount, size, appearance of wear debris particles in lubricant.

3. Signal analysis techniques

As discussed in the last chapter, bearing damage can be detected by analyzing acceleration signal generated by specific components. The vibration signals produced by faults have been widely studied, and intensely powerful diagnostic techniques are now available. Generally, these techniques can be classified according to different rules. In some papers, the main techniques to diagnose and examine REB defects are categorized into time domain approach, frequency domain approach, time-frequency domain approach and envelope analysis. When analyzing machine vibrations in frequency domain normally a number of prominent periodic frequency components which are directly related to the fundamental movements of various parts of the machine can be discovered. With frequency analysis we are therefore able to trend the source of undesirable vibration. There are also tutorials dealing with, for example, separation of bearing signals from discrete frequency noise and enhancement of the bearing signals. This chapter will be organized by listing a few universal techniques and concentrating on novel information extraction in the end. The last three methods are going to be applied for analyzing the bearing simulation signal in the coming chapters.

3.1 Envelope analysis

The envelope detection technique arose in 1974, and it was originally called high frequency resonance technique by Darlow et al. [51]. And then McFadden and Smith [52,53] put forward the concept of envelope analysis for REBs, which narrated that an impulse signal is generated each time a localized defect contacts with another surface in the bearing. At that time, envelope detection was carried out by using an analogue circuit to band-pass filter the analogue vibration signal around a structural resonance, afterwards, rectificating full or half wave to reconstruct the estimated envelope signal. Benefiting from the development of digital processing techniques, envelope analysis experienced considerable improvement. Nowadays, it has become a benchmark method to extract bearing diagnostics from a vibration signal. As stated in [1], the technique tells that a signal is band-pass filtered in a high frequency band in which the fault impulses are amplified by structural resonances. It is then amplitude demodulated to form the envelope signal, whose spectrum contains the desired diagnostic information in terms of both repetition frequency (defect frequency) as well as modulation by the appropriate frequency at which the fault is passing through the load zone.

The Fig. 4 depicts the procedure for envelope analysis using the ‘‘Hilbert transform’’ method which was presented by Ho and Randall [16], where a one-sided spectrum (positive frequencies only) is inversely transformed to the time domain. This gives a complex time signal (a so-called

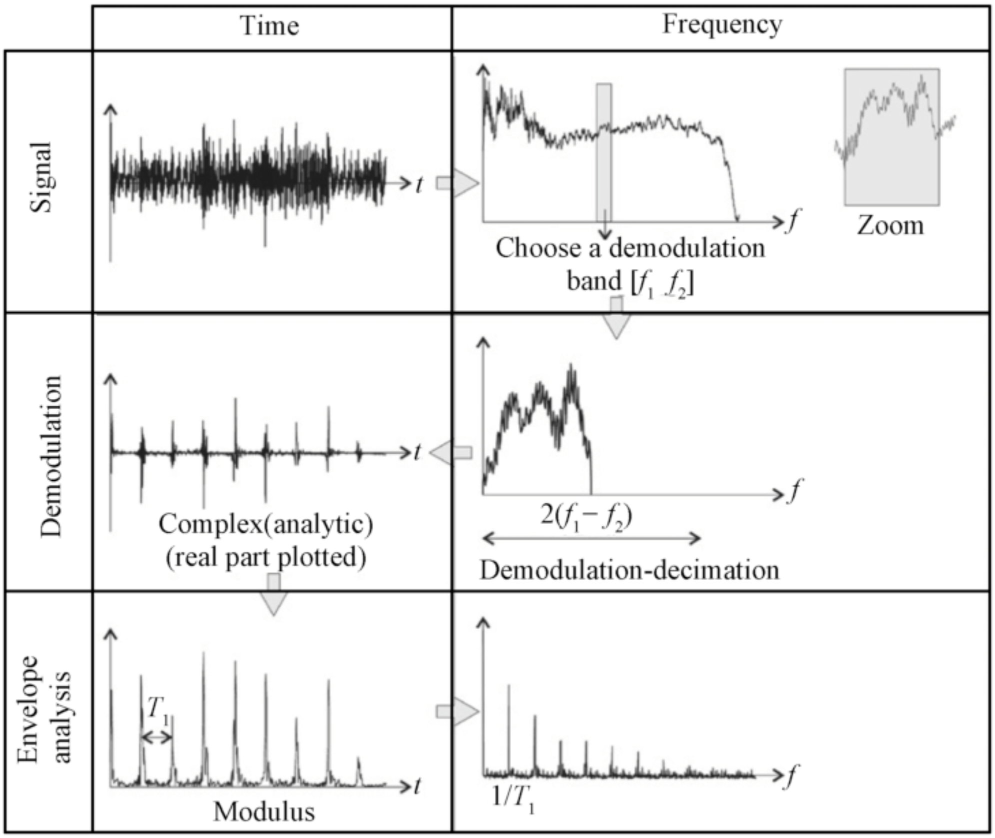

Fig. 4. Procedure for envelope analysis using the fractional Hilbert transform method

''analytic signal'') whose imaginary part is the Hilbert transform of the real part. An immediate benefit is that the extraction of the section of spectrum to be demodulated is effectively by an ideal filter, which thus can separate it from adjacent components that might be much stronger.

Unfortunately, the method suffers from certain restrictions. For example, it is often performed for a certain frequency region where the signal-to-noise ratio (SNR) is high, and to ensure a one-sided spectrum, the frequency band is padded with zeros to double the length in order to set the negative frequency components to zero. Therefore, Ho and Randall (2000) suggested that it was a better choice to analyze the squared envelope rather than the envelope. The main advantage manifested that the higher harmonics components in the envelope spectrum would be not included. This is because the square root process produces higher harmonics, and in addition some of these are aliased if they locate above the Nyquist frequency. Although the effect cannot be removed from digital signals by low-pass filtering, it still can be reduced by introducing a higher sampling frequency.

It is noted in Ref. [54] that since the squaring doubles the frequency content of a signal, the sampling frequency should be doubled before the signal is squared or rectified digitally. In addition, squared envelope spectrum (SES) also exhibits simplicity, low computational cost and implementation easiness.

To sum up, the envelope analysis can discover more pronounced diagnostic information than the analysis of raw signals, including the repetition rate of the fault and potential modulations. It appears to be an advantageous tool in very early detection and fault symptom identification.

3.2 Kurtosis, spectral kurtosis and kurtogram

The concept of kurtosis was firstly proposed by Dyer & Stewart (1977) [55], which acted as a tool for measuring the severity of faults in machines. Whereas a detailed account was not given and only a suggestion which stated that to obtain clearer results selecting proper frequency bands for filtering is highly required. Until 1983, Dwyer [56] detected impulsive events in sonar signals by using spectral kurtosis (SK), which grasped the impulsiveness in frequency domain from a signal based on short time Fourier transform (STFT). Later Antoni [57] gave a comprehensive formalization and definition of SK for nonstationary processes, at the same time, linked theoretical concepts with practical applications to detect the existence of faults in REBs.

The spectral kurtosis extends the concept of the kurtosis into a function of frequency which indicates how the impulsiveness of a signal is distributed in the frequency domain. Given a signal $\mathrm{x(t)}$, the local Fourier transform at time t is easy to compute by moving a window along the signal, denoted as $\mathrm{X(t,f)}$. then SK can be defined as follows:

$$\mathrm{K(f)} = \frac{\langle|\mathrm{X(t,f)}|^4\rangle}{\langle|\mathrm{X(t,f)}|^2\rangle} - 2$$

Where $\langle\cdot\rangle$ is the time-averaging operator. And the calculation of SK from the STFT for a simulated bearing fault signal is displayed in Fig. 5 (cited from Ref. [1]). In the equation, a fourth-order statistic is used. In this way, the presence of transients in a signal can still be detected even when they are buried in strong additive noise. Hence SK endows a powerful tool by seeking out the frequency bands where the faults take place.

Kurtogram displays the spectral kurtosis in the form of two-dimension map involving frequency and window length. Hence, the optimal central frequency and bandwidth of the band-pass filter

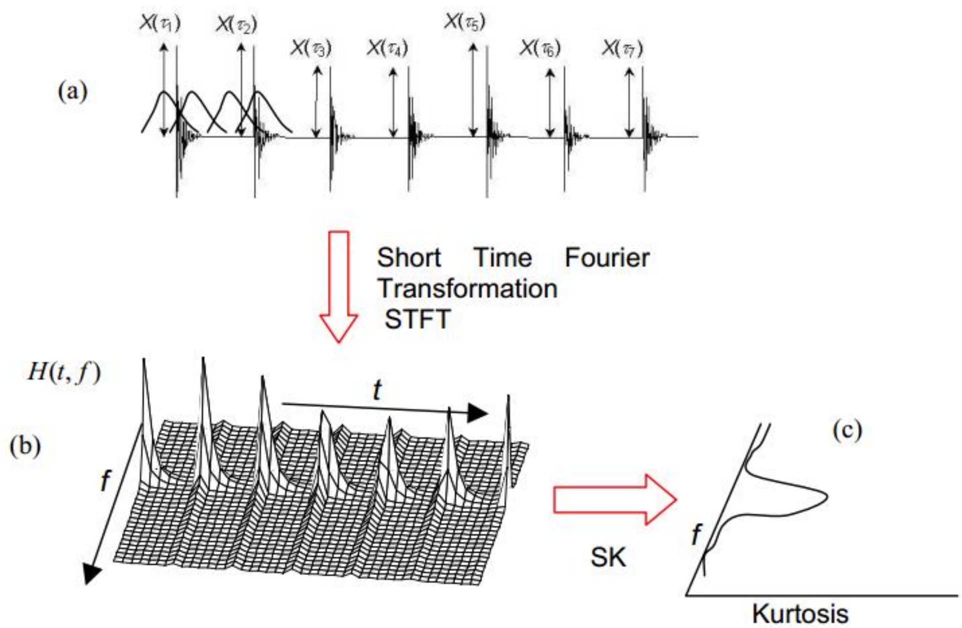

Fig. 5. Calculation of SK from the STFT for a simulated bearing fault signal: (a) time signal, (b) STFT and, (c) spectral kurtosis

can be found to maximize the kurtogram, which makes it an effective tool to detect and characterize non-stationarities in a signal. However, the disadvantage locates in situations where signal to noise ratio is very low or in presence of non-Gaussian noise such that the method will fail.

Apparently, it is costly and lacks of practical convenience to compute kurtogram for all possible combinations of center frequencies and bandwidths. One solution may be feasible by subdivision of the bandwidths into rational ratios that permit the use of fast multi-rate processing, which is called fast kurtogram (FK). A comparison between the fast kurtogram and the full kurtogram of a signal is illustrated in Fig. 6. Although the former is suboptimal due to its coarser resolution, it happens to return virtually the same result as the latter in terms of transient locations. FK is order of magnitude faster than full kurtogram and then it appears suitable for industrial applications ideally.

Thanks to the feature of detecting the frequencies where an impulsive bearing fault signal is dominant and nullifying the parts where there is stationary noise only, SK can be qualified as a filter to select those components of the signal with the highest level of impulsiveness. Antoni [1] stated that it is possible to define both Wiener and matched filters in terms of the SK for an

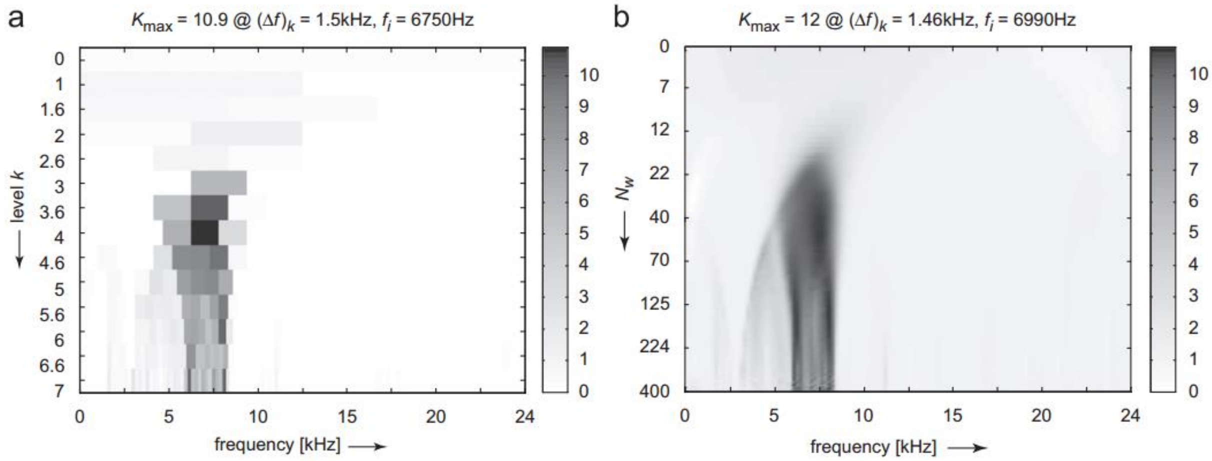

Fig. 6. Comparison of the (a) fast kurtogram with the (b) full kurtogram for an impulsive signal [63]

impulsive signal, such as signal owing to bearing faults masked by stationary random noise. As for Wiener filter, proportional to the square root of the spectral kurtosis, it maximizes the similarity between the filtered component and the clear signal without noise. Whereas matched filter, a narrow band filter, maximizes the SNR of the filtered signal irrespective of its shape.

Concerning the application of fast kurtogram to REB diagnostics, Antoni [58] gave an example to illustrate its function in seeking the repetitive-like transient faults. To sum up, it tells the presence of abnormal transients, indicates the corresponding frequency bands and addresses the maximum SNR region. It is also pointed out that it actually returns the critically sampled complex envelope in the selected frequency band.

Much effort has been made to the enhancement of SK. Considering SK's susceptibility to non-Gaussian noise, Yongxiang Zhang [59] proposed a method to design the parameters for optimal resonance demodulation combining FK for initial estimation and a genetic algorithm for final optimization, which gives fast convergence to the solution over the whole zone. The algorithm can find the best filter in narrow scope benefiting from the best spectral kurtosis given by FK. Compared to traditional envelope analysis and FK alone, this method provided more flexibility for diagnosis of REBs. Moreover, based on ensemble local mean decomposition (ELMD) and FK, Lei Wang et al. [60] presented a time-frequency domain method to diagnose faults in rotating machines, of which ELMD can eliminate non-stationary and nonlinear interferences and highlight the components concerning faults. The feasibility and effectiveness are demonstrated by two experimental examples.

Furthermore, the application of SK in detecting faults in REBs can be improved and optimized by various approaches. For example, pre-whitening the power spectral density of signal, decomposing the pre-whitened signal using complex Morlet wavelets and so on. Details will not be covered in this thesis.

3.3 Cyclic modulation spectrum and Fast spectral correlation

The excellent behavior of cyclostationarity in describing the symptomatic modulations or repetition of transients has been related. Based on that, spectral correlation (SC) became a tool for cyclic spectral analysis that can express the whole structure of modulations and carriers of a signal into a bi-frequency map. The function of spectrum contains two variables namely spectral and cyclic. Suffering from the high computational cost, SC seems not a satisfactory tool in reality. Then The Cyclic Modulation Spectrum (CMS) was proposed by Antoni [61] 2009 as a faster alternative to return similar result as the SC. It tracks periodic flows of energy in frequency bands by evaluating the Fourier transform of the squared envelope at the output of a filter bank and it is interpreted as waterfall of envelope spectra. Due to the fact that the computational cost mainly centers on the calculation of STFT, the development of calculating devices helps CMS to be an efficient tool in real case. The application of CMS is mainly described in [62]. However, limited to detect relatively slow periodic modulations only, it cannot detect periodic patterns other than in the form of modulations whose frequencies are necessarily lower than the frequency resolution. Also the biased error will increase with the rise of cyclic frequency.

As a development of CMS, the Fast-SC was proposed by Antoni [63]. Benefiting from the property of cyclostationary signal, the STFT evidences periodic flows of energy in and across its frequency bands. The Fourier transform of the interactions of the STFT coefficients then returns a quantity which scans the SC along its cyclic frequency axis. Obviously, the same procedure can also be processed in other domains of application, for instance, angular domain. The results of simulation indicated that Fast-SC is an asymptotically convergent (unbiased and nil variance) estimator of the SC. And due to the algorithm structure, it can be computed in parallel, which tends to speed up further on. Certainly the essential advantage of Fast-SC is to save computational cost compared to CMS and other techniques, where the earning may be simply the ratio of the signal length to the STFT window length. The efforts for computation mainly lie on the fast Fourier transform (FFT) of STFT products, which makes it a competitive tool for analyzing long records over a wide cyclic frequency span. The use of the Fast-SC has been illustrated on several vibration signals in order to detect rolling bearing faults. The gain in computational time exerts considerable when the signal length increases, for example, compared to averaged cyclic periodogram (ACP) the central processing unit (CPU) time of implementing Fast-SC will be reduced significantly, especially when the length is larger than 2^{11}.

3.4 Autogram

To overcome the drawback of SK in detecting bearing damages with non-Gaussian noise, Ali Moshrefzadeh et al. [64] proposed a valuable tool, named Autogram. The auto-covariance function of a second order cyclostationary signal is periodic, and in virtue of this property it is considered to be sufficiently general for diagnosing faults in rotating machinery, for example, REBs and gears. It is also noted that the method is easy to handle and doesn't require any prior knowledge of signal processing.

The procedure to compute Autogram is presented as follows. First, the time domain data need to be divided in frequency bands and central frequencies. The Maximal Overlap Discrete Wavelet Packet Transform (MODWPT) is adopted as a filter to obtain a series of signals in each level of decomposition. Second, unbiased autocorrelation (AC) is computed on the squared envelope of each center frequency. As stated in Ref. [64], "The AC has the benefit of removing the uncorrelated components of the signal, i.e. noise and random impulsive contents, both unrelated to any specific bearing fault. Furthermore, the periodic part of the signal (directly related to the defects) is enhanced, showing an additional virtue of this process. This is even more advantageous since it is done for each node separately rather than on the complete raw signal, so that SNR for each demodulated band signal is increased.". Third, by computing the impulsiveness of each AC, Autogram displays the result in a color map, from which the node with maximum kurtosis can be addressed. Herein two types of kurtosis are given which are called Lower and Upper Autogram. This is a thresholding process to separate the two parts of the signal (noise and defect impulses) without losing any useful information for diagnosis of bearings, which directly affects the quality of the frequency analysis because it decides which coefficients will be retained and which will be rejected. The last step, with the node corresponding to fault characteristics selected in the previous step, computation of Fourier transform to the squared envelop is made to extract the diagnostic information, further defect frequencies.

Two cases related to bearing fault detection are also provided to demonstrate the effectiveness of Autogram. It is able to automatically select the carrier frequencies containing crucial information, even if the signal is processed in different ways. In brief, Autogram has strong performance in identifying damages in REBs. Deserved to be mentioned, Autogram is specifically designed to enhance the detection of periodic impulses, thereby not suitable to discover other damages such as pitting or corrosion.

3.5 Order tracking

To avoid the smearing effect of discrete frequency components owing to speed fluctuations, displaying the rotating machine signals in a different frequency axis is often expected, where "orders" of shaft speed replace the usual frequency axis. Concerning varying case, it is necessary to sample the signal from a tachometer or shaft encoder synchronous with shaft rotation for the sake of order analysis. In early years, a phase-locked loop was used to track the signal and a series of sampling pulses per tracked period could be obtained. Nevertheless, "angular resampling" may be a better way to sample along uniform increments in shaft angle, where digitally resampling on each corresponding period of tachometer signal is processed. It can be achieved by simply increasing the sampling rate by a large number and selecting the sample next to each theoretical interpolated position, which is available both in time domain and frequency domain. Alternatively, for the purpose of angular resampling, using phase demodulation of some shaft speed related component to deliver a map of shaft angle vs time is also valid, where the times corresponding to uniform phase increments can be addressed to interpolate between the original time samples.

Another issue worth of attention is to ensure that the signal is adequately low-pass filtered to prevent aliasing. It can be solved by oversampling, and thus the unexpected components can be isolated from the measurement range. In addition, order tracking (OT) has a positive effect on the spectrum of a signal with distinguishable speed related components. Without the OT, discrete frequency components are much more difficult to be detected in the spectrum.

The method OT still can be improved in some respects. An online tacholess OT method is given by Yi Wang [66] based on extraction of instantaneous tachometer information from the collected vibration signal itself. In this method, the problem of mono-component separation under changing rotational speed conditions and transient components extraction is solved involving the combination of generalized demodulation and resonance demodulation. In the mean while there are also some other techniques employed such as generalized Fourier transform, an adaptive ridge detection algorithm, band-pass filtering and FK. The results from both simulated and experimental bearing signal indicates the effectiveness of this online technique for detecting bearing faults under speed variation condition. Another approach is proposed by Tangfeng Yang [67], who combined the envelope order tracking and the constrained independent component analysis (cICA) to address the issue of OT's disability for multi-impulsive sources. In this scheme, OT collected sensor data from different positons and cICA extracted the interesting envelope independent components (ICs) by a reference signal from the prior-known feature frequency of the bearing. As a consequence, the features of bearing faults can be clearly expressed in the spectrum of the obtained interesting envelope.

3.6 Novel information extraction

When detecting the bearing faults, most of the analysis techniques are applied directly on signal acquired without considering historical information. From this point of view, Stephan Schmidt [68] proposed a methodology to enhance the performance of the analysis techniques by incorporating historical data in reference condition. Thereafter a novel signal, i.e. a signal that contains information that is not present in the historical reference data, can be consequently extracted. Generally speaking, it is a bit arduous and expensive to gain historical fault data of a machine, especially in conditions where the components are equipped in untouched region or built at a big cost. But the data of a healthy machine can be obtained easily as a reference, which is beneficial to the implementation of the fault diagnosis technique. Theoretically, after the filtering operation, the novel signal contains rich information corresponding to fault of the machine, which can be subsequently analyzed through conventional signal analysis techniques. The specific details regarding the algorithm of the methodology is elaborated in next few paragraphs.

Intuitively, novel signal expresses the deviation between the newly acquired vibration signal and the vibration signal from healthy machine, which explains the reason why a bulk of diagnostic information is included. There are two main phases inside the methodology, namely training phase and testing phase. The former extracts model parameters from reference data and the latter phase utilizes the parameters to design a filter for extracting novel signal accordingly. Based on the assumption that the damage locates in specific frequency range and changes in statistical properties are also resulted in relevant bands, the methodology can be developed step by step. And an

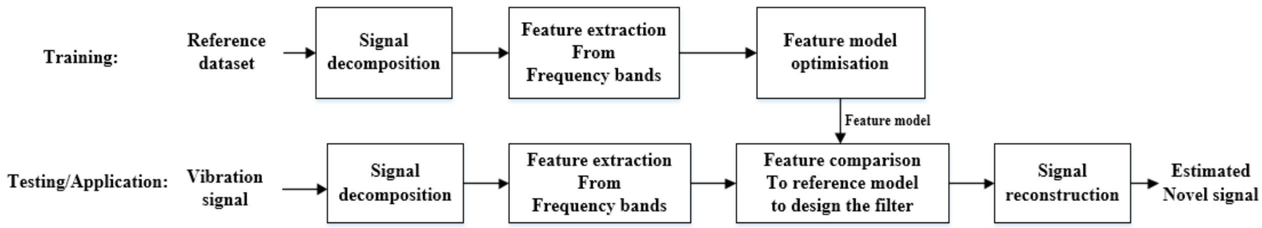

Fig. 7. Overview of the methodology

overview of the methodology is illustrated in Fig. 7.

First of all, the original signals have to be decomposed into time-frequency representation by using STFT, where a series of window lengths are involved and the overlapping rate is set as 75%. Then a feature extraction function is used to extract a diagnostic feature (as required) associated with each center frequency of STFT with a corresponding frequency resolution. In most cases, the feature extraction function has a definite function relationship with STFT. The same procedure can also be applied to reference data to deliver features in training phase which serve the purpose of modelling in next step. The second step mainly focuses on feature modelling. The features are parametrized by the Gaussian Probability Density Function (PDF), which can detect changes in features effectively. The author pointed out that "The naive assumption is made that the features are uncorrelated, i.e., the full covariance matrix reduces to a diagonal matrix with the variances of the respective features being on the diagonal of the covariance matrix. This assumption is made to reduce the number of parameters that need to be estimated from the reference datasets." After that a novelty detection scoring function compares the associated values in specific frequency bands with a threshold to detect novelties. As recommended, the squared Mahalanobis distance acts as an excellent novelty detection scoring function. In step three, the filter coefficient can be derived from the comparison between novelty detection score and threshold of "9" which covers 99.7% of the probability density function of a Gaussian distribution. Hence the coefficient equals unity if a novelty is detected and zero if not. Finally, novel signal can be estimated in two stages. Primarily, every frequency band of the time-frequency representations concerning different window lengths are scaled by the novelty filter. Ultimately, the inverse Fourier transform is applied on each scaled time-frequency representation to obtain the desired novel signal.

Afterwards, novel signal can be analyzed by conventional diagnostic techniques. In the thesis,

Items	Values	Units
Rotational speed	200	Hz
Sampling frequency	35*1024	Hz
Duration	1	Second
Random fluctuation	0	Percent
Noise amplitude	0.15	Percent
Teeth of wheel	23	Teeth
Resonance of structure	5600	Hz
Damping factor	0.05	-

Table 1. Specific quantities of simulated bearing signal

numerical gearbox simulation and practical experiment are conducted to validate the sensitivity of the novel vibration signal to presence of damage, which makes it a convincing and favorable pre-processing technique. In short, the benefits of the methodology are: historical reference data can be incorporated, and the filter is designed to attenuate frequency bands that do not contain any novel information.

The methodology can be extended to track the damage in varieties of rotating machines, of course, bearing included. It is noted in the conclusion that future work may focus on making the methodology more robust by optimizing each step in the procedure. The numerical modelling and novel information extraction for a rolling bearing will be carried out in next chapter.

4. Novel information extraction by implementing the methodology

Detailed introduction with respect to the pre-processing methodology has been presented in the previous chapter and herein the implementation of that based on simulated bearing signal will be given. In consideration of the convenience for carrying out the methodology, the whole procedure is going to be running in Matlab including the generation of simulated bearing signals and novel information extraction, where debugging and algorithm testing can be easily achieved. In addition, a mass of data analysis tools and functions can be employed readily. Section 4.1 is concerned with how the signals, both healthy and damaged, are produced on basis of a phenomenological model. As displayed in section 4.2 and 4.3, the final results are to be exported as diagrams thoroughly.

4.1 Generation of simulated bearing signals

Overall, the vibrational signal is induced by a simulated gearbox model convolving three components, i.e., periodic, impulse and noise, where the amplitudes and phase shifting can be set

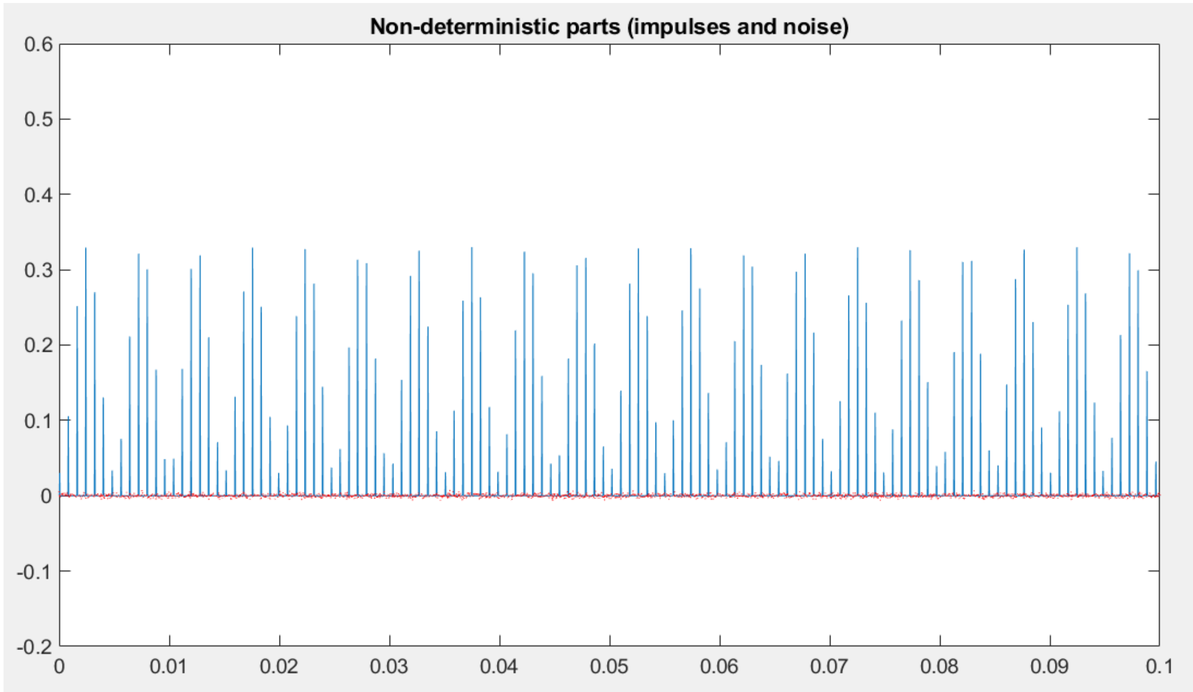

Fig. 8(a). Non-deterministic parts of the damaged signal in time domain

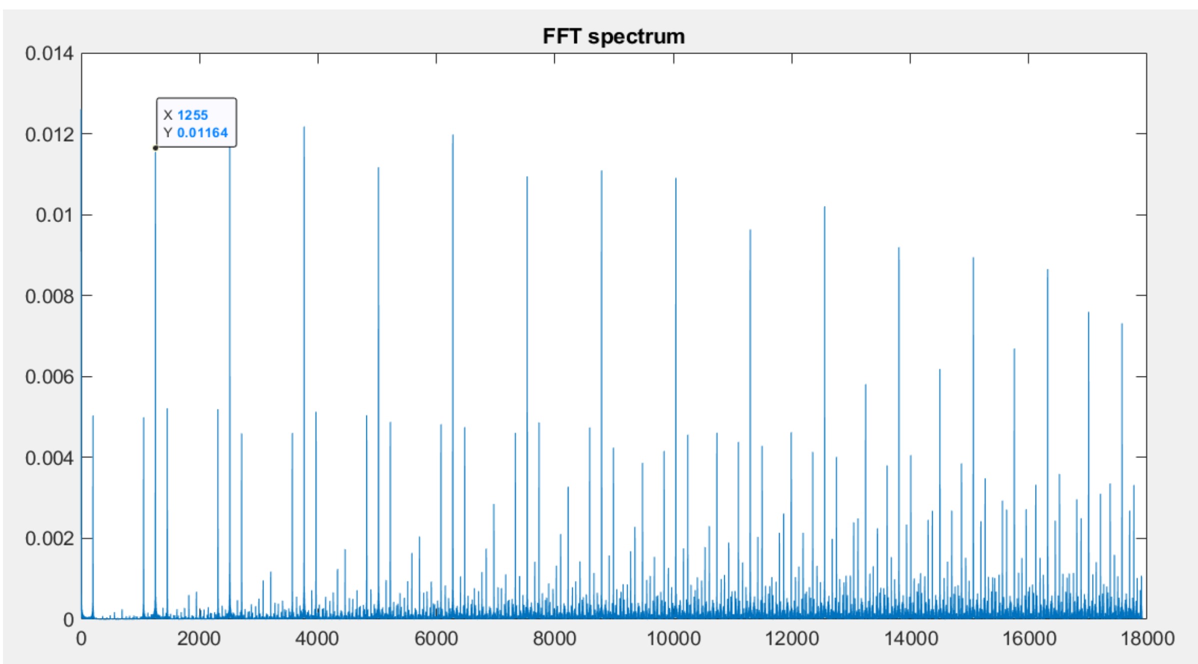

Fig. 8(b). FFT spectrum of non-deterministic part

as required. Then the signal is obtained by crossing a transmission path, whose resonance frequency and damping factor can be customized according to the need for monitoring conditions. Moreover, the model provides two types of conditions namely variation and constant speed, which is advantageous to reconstruct the real case situation.

To simplify the signal generation, sweeping in rotational speed is not considered thus the bearing is working in constant speed condition, namely 200 Hz (12000 rpm). Inner race damage is selected among local faults which exerts distinct envelope profile. The amplitudes of damage, noise and random fluctuation of bearing signal are shown in Table 1. In order to simulate an actual case, the dimensions of bearing are in accordance with SKF 6006-Z. With the program running twice for different damages (healthy and inner race fault), two groups of data are derived on behalf of signals from reference phase and application phase. Normally the training phase should be run for a long period to improve the precision of the reference condition. For convenience only one cluster of signals is generated as data in reference condition. Theoretically, the defect frequency computed from the formula equals to 1225 Hz, and its multiples are 2510 Hz, 3765 Hz and so on.

The impulse induced by inner race is demonstrated in Fig. 8 where the envelope is indicated in (a) and the characteristic frequency together with its multiples are arrayed in an axis of (b). Besides, the modulation on both sides are present respectively which simulates the actual condition. Fig. 9 clarifies the original and acquired signal in time domain and frequency domain where the latter is derived through a transmission path from the former. It can be easily spotted that the impulses have been buried, meanwhile, only periodic part exists distinctly (4600 Hz). In this way, the data have been prepared adequately for the realization of the pre-processing methodology to the extraction of novel signal.

4.2 Extraction procedure

Now that signals of healthy and damaged bearing have been obtained, novel information extraction is operated as follows in accordance with Section 3.6. For the first place, the signals in damaged case are decomposed by STFT with a suite of window lengths, viz., $[8,16,32,64,128,256,512,1024,2048]$. And then the feature extraction function, kurtogram, is used to collect features based on STFT for every window length. Fig. 10 illustrates the kurtogram values along window length axis. Window length 1024 conveys the kurtogram more clearly, so that this typical window length is also applied to training phase. In step two, features are parametrized by Gaussian PDF, such that a novelty detection scoring function, the squared Mahalanobis distance can be obtained by computation. Third, a filter is designed by comparing the novelty detecting score and an associated threshold which is 9 in this case, and coefficient of the filter only equals to unity or zero. In the end, each frequency band of the time-frequency

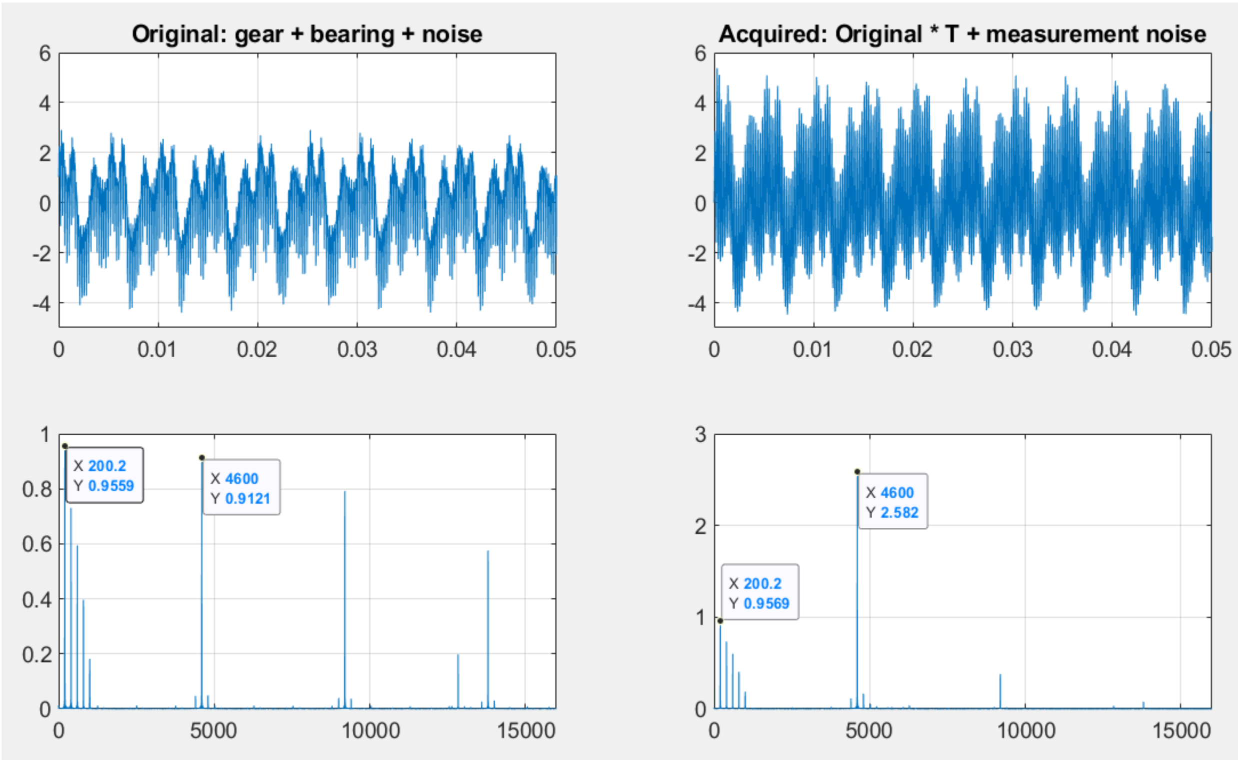

Fig. 9. Original together with acquired signal in time and frequency domain

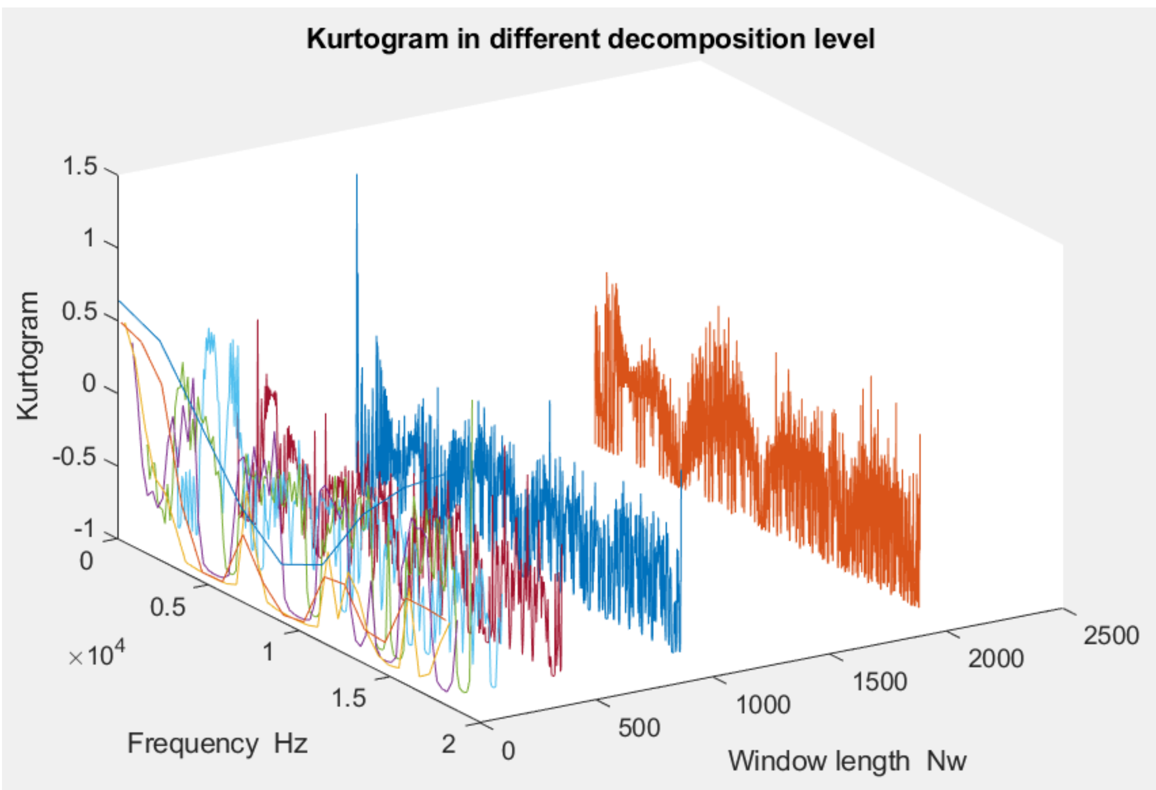

Fig. 10. Kurtogram values in different window lengths

representations for different window lengths are scaled by the novelty filter, where the center frequencies are tuned with those in training phase through a function relates the window lengths, centre frequencies to the index of filter coefficients. After performing inverse STFT for each group of scaled time-frequency representation, their average can be computed which represents the estimated novel signal of interest.

4.3 Results and diagrams

From the diagrams of the novel signal (depicted in Fig. 11), it can be distinguished that the frequencies concerning the damage have been retained and also the rotational speed in low frequency region. As indicated in Fig. 11(a), the amplitude of novel signal is always less than original signal, which is ascribed to the filter effect on the STFT during scaling procedure, where the signal in frequency domain only contains limited amount of information.

Nevertheless, the methodology is not always providing rosy results. Among the abundant trials, it is discovered that the novel information is sensitive to the magnitude of noise. Only generating signals under a certain amplitude, the rotational speed can be successfully kicked out. This is may be due to the influence of impulsivity in noise on the feature extraction. When the other damages are taken into consideration, the results are not sufficiently satisfactory as for the inner race case. Moreover, the implementation in variational speed is not practiced in this thesis, which requires further effort to validate the effectiveness of the methodology. As mentioned at the end of Ref. 68, "future work should focus on making the methodology more robust by optimising each step of the methodology." And a few future potential strategies are given to improve the performance of the method. Therefore, there are a great many paths to be investigated to anchorate the effectiveness of this technique.

Although the validity of the pre-processing methodology is partially proved in this thesis due to the imperfection of the modelling and feature extracting, it still can be believed that it appears an effective technique for rotating machine diagnosis where historical data are available.

5. **Experiments by different analyses**

5.1 Analysis by FK

Thanks to the availability of open access program provided by J. Antoni, the implementation of fast kurtogram for damaged bearing signal can be directly conducted in Matlab. Without pre-

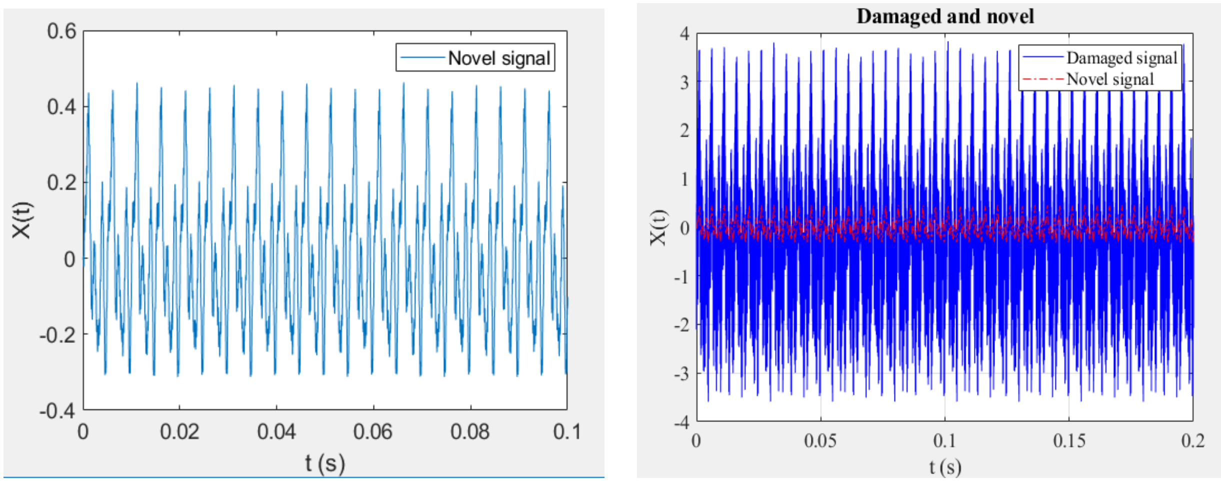

Fig. 11 (a). Derived novel signal in time domain.

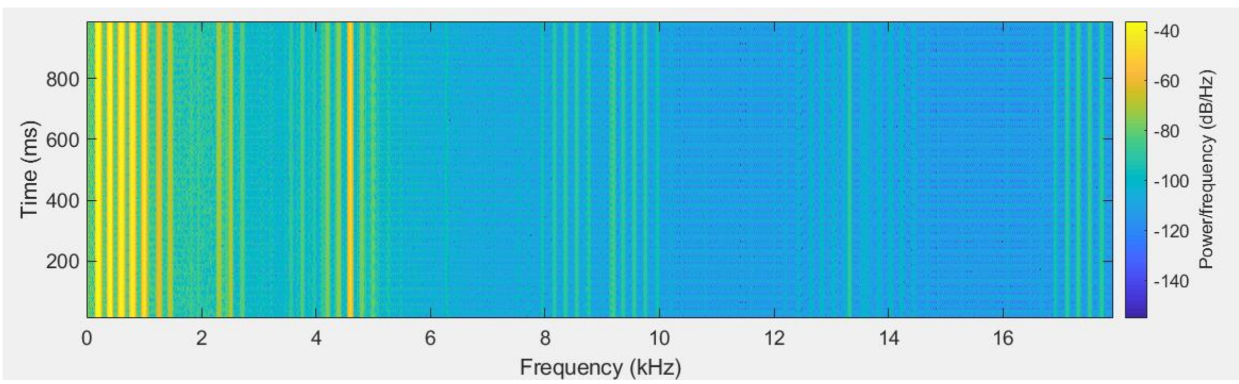

Fig. 11 (b). Spectrogram of novel signal.

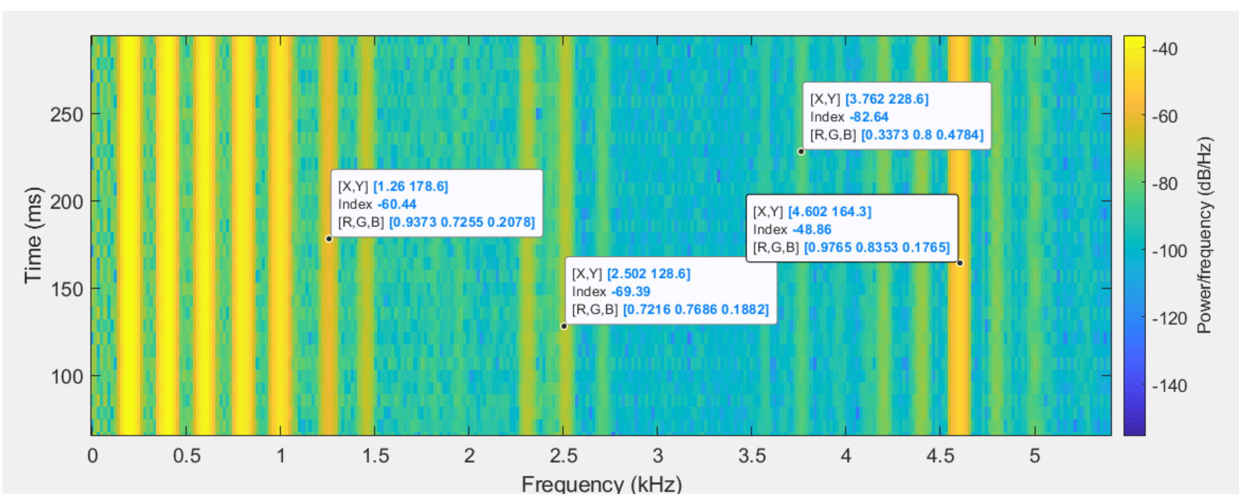

Fig. 11 (c). Data in zoom of spectrogram.

whitening the signal, the kurtogram can be displayed in two-dimension map containing several square segments, which is illustrated in Fig. 12(a). By selecting the area where the maximum kurtosis lies and inserting the corresponding carrier frequency and level, transient signals are able to be filtered out and also the envelope spectrum can be easily computed. Fig. 12(b) reveals the envelope and amplitude spectrum in frequency domain.

As shown in Fig. 12(c), the characteristic frequency can be easily captured with a spike whose value is in accordance with defect frequency for inner race. Besides, the other crests such as 200 and its multiples indicated existence of the rotational speed. By the way, if observing the plot in detail, it is possible to find the modulations on both sides of the defect frequency, which equal to plus or minus the rotational speed from defect frequency.

FK is proved to be an effective tool for detecting the impulses hidden in a signal. Its numerical efficiency is another advantage, which makes it suitable for a mass of trials in limited period. Moreover, through the verification test, a smaller sampling frequency had a bad impact on the final result, where a portion of information may be dropped out.

5.2 Analysis by Autogram

The algorithm for computing the Autogram has been shared by A. Moshrefzadeh [64] on Mathworks community, which offers great convenience for analyzing the diagnosis of the damaged bearing signal. With setting specific parameters, the Autogram can be obtained (Fig. 13(a)), and consequently the average combined square envelope spectrum (Fig. 13(b)) can be derived. Through repeated experiments, it manifested that upper Autogram delivered an excellent plot which contained defect frequency and its multiples more clearly.

As depicted in Fig. 13(b), besides the information concerning rotational speed in low frequency region, there are also clusters of spikes distributed in high frequency area which represent the defect frequencies and their modulations. As for the periodic part mixed in the original signal, it has been filtered out by the program. And noise generated in the simulated signal, though in presence along the envelope, plays a minor role in the bearing diagnosis.

5.3 Comparison

Now that the damage of the bearing has been detected by three different numerical analyses. On the whole, all of them have a certain effect on the diagnosis, which is worth to be discussed.

Speaking of detection capability, FK exports a clean result including rotation speed and the fundamental characteristic frequency; Autogram can detect rotational speed and multiple defect frequencies are displayed in vertical lines with gradient colors; novel information contains the

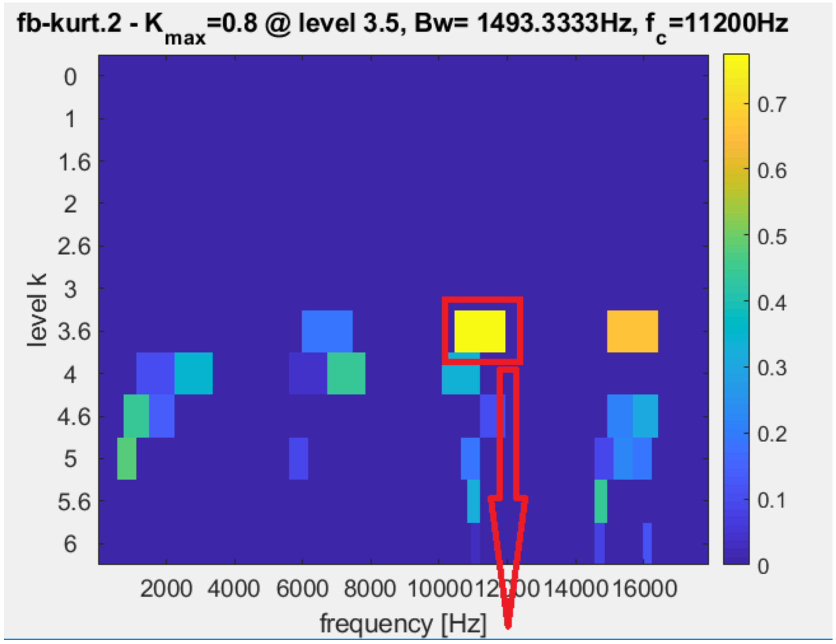

Fig. 12(a). Kurtogram by FK

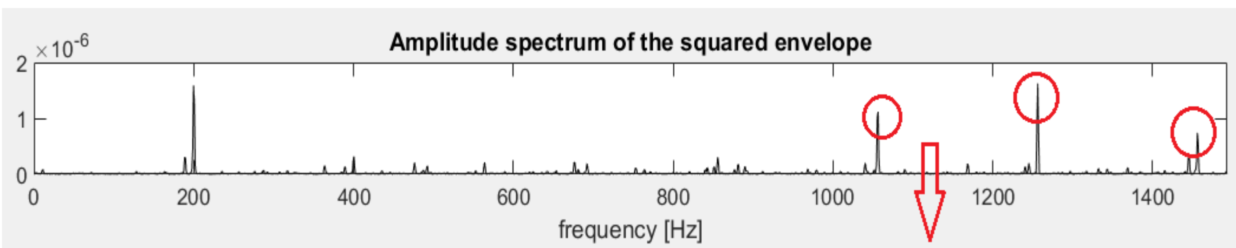

Fig. 12(b). Squared envelope of the selected segment.

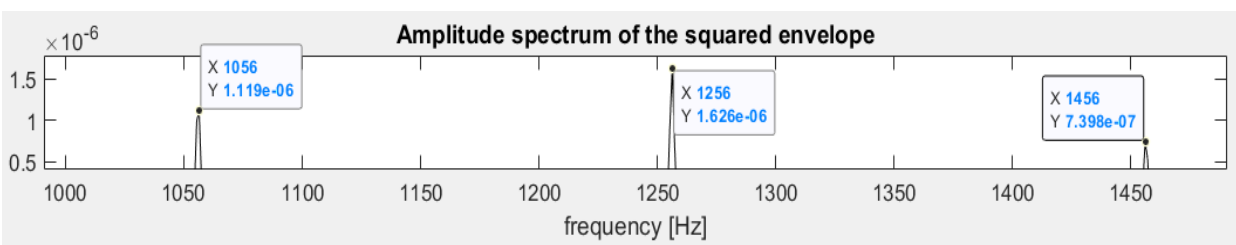

Fig. 12(c). Zoom of the squared envelope.

characteristic frequency and its multiples. Generally, they put up a good performance to discover the maximum impulsiveness regarding defects.

Referring to computational time, FK apparently outperforms the other two methods by taking advantage of multi-rate filtering; Autogram, developed from kurtogram, can automatically generate the final spectrum. Experiments manifest that it takes less than 10 seconds below level 7. Unfortunately, the novel signal extraction is only a pre-processing method to obtain the essential information, and further signal analysis needs to be conducted, which seems not a time-saving

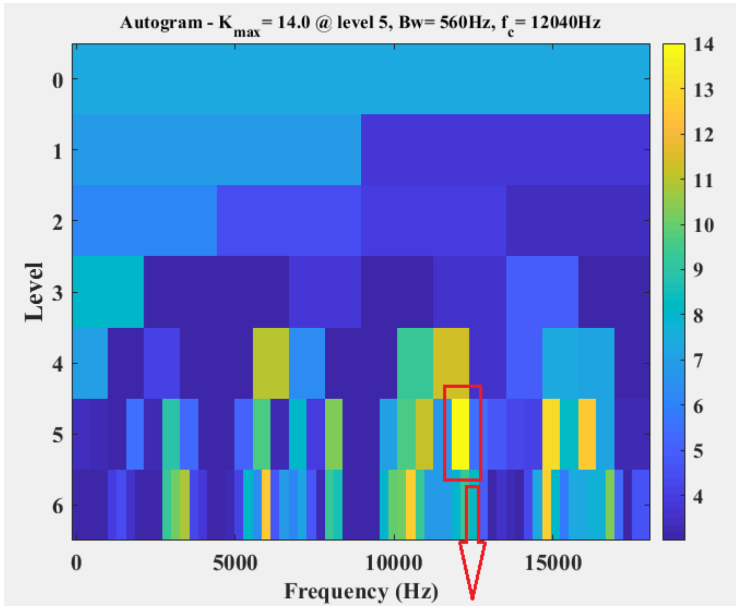

Fig. 13(a). Kurtogram by Autogram at level 6.

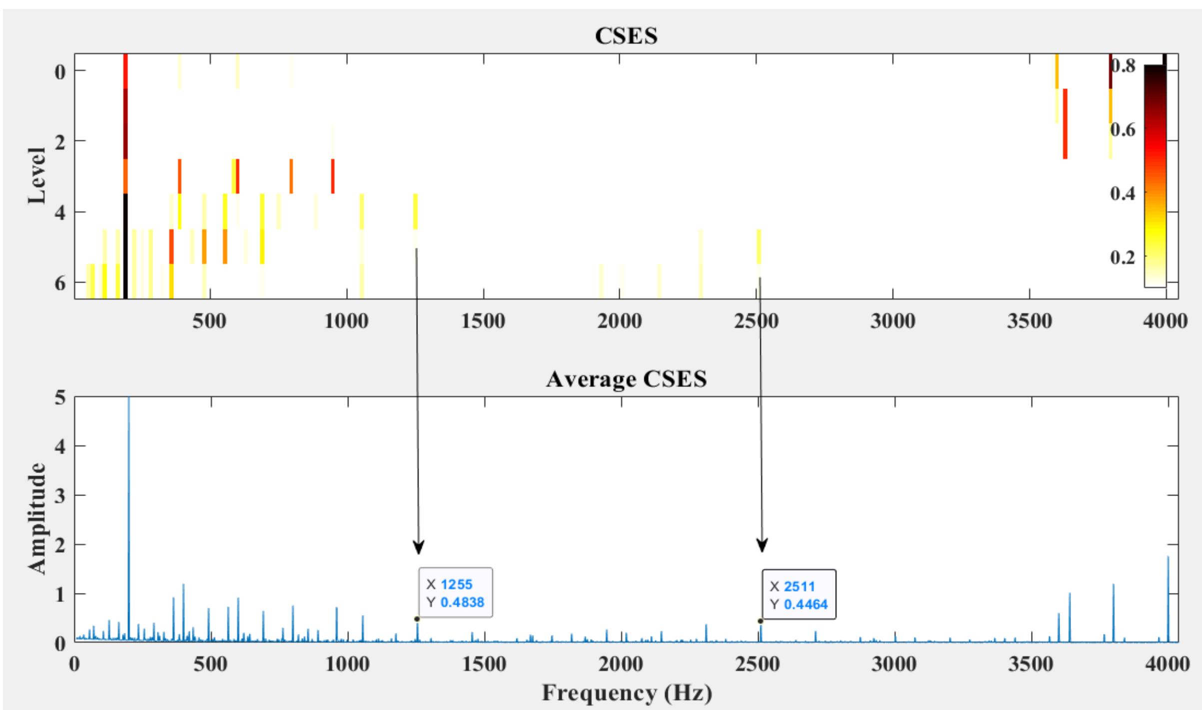

Fig. 13(b). CSES of the selected area.

technique.

As to operability, FK selects the area corresponding to the maximum kurtosis by hand and delivers the square envelope; on condition of specific parameters to be defined, the quality of Autogram's final result can be improved by alternating types of Autogram and enhancing the effect of upper portion; novel signal extraction requires a reference signal as a contrast, and the feature extraction function can be selected among many different types. Obviously, more investigations should be performed for adapting to sophisticated situations and improving the quality of results.

6. Conclusion

The thesis principally focuses on the detection of damage in REBs. With several common defects presented, an attempt to summarize the techniques for bearing fault diagnosis has been made. From the perspective of reliability and practicability, vibration analysis is the most commonly accepted technique due to its ease of application. Immediately following the introduction of signal analysis methods, three numerical analyses, specifically novel information extraction, FK and Autogram, are conducted successively to test the effectiveness for detecting damage to inner race.

Through comparison, FK and Autogram exhibit excellent performance in capturing the characteristic frequency of damaged bearing signal, as proved by many other researchers. And as for novel information extraction, a recently proposed pre-processing method, has its own merits from the view of algorithm. Although from the tests, it is limited by selection of feature extracting function and quantity of noise, a series of further investigations might target on different features, more sophisticating modelling procedure, alternative novelty filter, and experimental datasets to yield sufficiently accurate results than those previously found.

References

[1] Robert B. Randall, Jérôme Antoni, Rolling element bearing diagnostics—A tutorial, Mechanical Systems and Signal Processing, Volume 25, Issue 2, 2011, Pages 485-520, ISSN 0888-3270.

[2] Jeremy Wright, Noria Corporation, Grease Basics.

[3] 2015 The Timken Company, TIMKEN BEARING DAMAGE ANALYSIS WITH LUBRICATION REFERENCE GUIDE.

[4] Barden precision bearings, BEARING FAILURE: CAUSES AND CURES.

[5] Mobil Technical Topic, Grease High-Temperature Performance.

[6] Eugene Matzan, Detecting Premature Bearing Failure, Machinery Lubrication, May 2007

[7] Schaeffler group industrial, Rolling Bearing Damage - Recognition of damage and bearing inspection, Publ. No. WL 82 102/3 EA

[8] Tahsin Doguer, Jens Strackeljan, Vibration Analysis using Time Domain Methods for the Detection of small Roller Bearing Defects, SIRM 2009 - 8th International Conference on Vibrations in Rotating Machines, Vienna, Austria, 23 - 25 February 2009

[9] Kulkarni, Sham & Bewoor, Anand. (2016). Vibration based condition assessment of ball bearing with distributed defects. Journal of Measurements in Engineering. 4. 43-121.

[10] Bently, D. "Predictive maintenance through the monitoring and diagnostics of rolling element bearings." Bently Nevada Co., Applications Note, ANO44 (1989): 2-8.

[11] C.S. Sunnersjo, Varying compliance vibrations of rolling bearings, Journal of Sound and Vibration 58 (3) (1978) 363–373.

[12] Fukata, S, Gad, E.H., Kondou, T., Ayabe, T. and Tamura, H. (1985): On the Vibration of Ball Bearings. Bulletin of JSME, 28 (239), pp. 899-904.

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

[14] Y.T. Su, S.J. Lin, On initial fault detection of a tapered roller bearing: Frequency domain analysis, Journal of Sound and Vibration 155 (1) (1992) 75–84.

[15] N. Tandon, A. Choudhury, An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect, Journal of Sound and Vibration 205 (3) (1997) 275–292.

[16] Ho, D. and Randall, RB. (2000): Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. Mechanical Systems and Signal Processing, vol.14 (5), pp. 763-788

[17] M. Tadina, M. Boltezar, Improved model of a ball bearing for the simulation of vibration signals due to faults during run-up, Journal of Sound and Vibration 330 (17) (2011) 4287–4301.

[18] Gupta, P.K (April, 1975): Transient ball motion and skid in ball bearings. Transactions of the American Society of Mechanical Engineers, Journal of Lubrication technology, pp. 261-269

[19] Feng, N.S, Hahn, E. J and Randall, R. B. (2002): Using transient analysis software to simulate vibration signals due to rolling element bearing defects, Proceedings of the 3rd Australian Congress on Applied Mechanics, pp. 689-694, Sydney.

[20] Sawalhi, N., Randall R.B and Endo H. (2006 A): Simulating gear and bearing interactions in the presence of faults: paper to be presented at the 7th IFToMM-Conference on Rotor Dynamics, Vienna, Austria, 25-28 September 2006.

[21] Sawalhi, N., Randall R.B and Endo H. (2006 B): Gear and bearing fault simulation applied to diagnostics and prognostics, Proceedings of the 19th International COMADEM, Luleå, Sweden, pp.399-408.

[22] Tiwari, M., Gupta, K. and Prakash, O. (2000 A): Effect of radial internal clearance of a ball bearing on the dynamics of a balanced horizontal rotor. Journal of Sound and Vibration, vol. 238, pp.723–756.

[23] Tiwari, M., Gupta, K. and Prakash, O. (2000 B): Dynamic response of an unbalanced rotor supported on ball bearings. Journal of Sound and Vibration, vol. 238, pp. 757–779.

[24] Wijnant, Y & Venner, C & Larsson, Roland & Eriksson, P. (1999). Effects of Structural Vibrations on the Film Thickness in an EHL Circular Contact. Journal of Tribology-transactions of The Asme - J TRIBOL-TRANS ASME. 121. 10.1115/1.2833929.

[25] Sopanen, J, and Mikkola, A. (2003 A): Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 1: theory. Proc. Instn Mech. Engrs, Vol. 217 Part K: J. Multi-body Dynamics, pp. 201-211.

[26] Sopanen, J, and Mikkola, A. (2003 B): Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 2: implementation and results. Proc. Instn Mech. Engrs Vol. 217 Part K: J. Multi-body Dynamics, pp. 213-223.

[27] Gardner, W.A. (1986): Introduction to random processes with application to signals and systems, Macmillan, New York.

[28] Antoni, Jerome & Randall, R.B.. (2002). Differential Diagnosis of Gear and Bearing Faults. Journal of Vibration and Acoustics. 124. 10.1115/1.1456906.

[29] Sawalhi, Nader & Randall, R.B.. (2007). Simulation of vibrations produced by localized faults in rolling elements of bearings in gearboxes.

[30] Cui, L., Chen, X., & Chen, S. (2015). Dynamics Modeling and Analysis of Local Fault of Rolling Element Bearing. Advances in Mechanical Engineering.

[31] Feiyun Cong, Jin Chen, Guangming Dong, Michael Pecht, Vibration model of rolling element bearings in a rotor-bearing system for fault diagnosis, Journal of Sound and Vibration, Volume 332, Issue 8, 2013, Pages 2081-2097, ISSN 0022-460X.

[32] C. Mishra, A.K. Samantaray, G. Chakraborty, Ball bearing defect models: A study of simulated and experimental fault signatures, Journal of Sound and Vibration, Volume 400, 2017, Pages 86-112, ISSN 0022-460X.

[33] Peng Gao, Lei Hou, Rui Yang, Yushu Chen, Local defect modelling and nonlinear dynamic analysis for the inter-shaft bearing in a dual-rotor system, Applied Mathematical Modelling, Volume 68, 2019, Pages 29-47, ISSN 0307-904X.

[34] Dipen S. Shah, Vinod N. Patel, A Review of Dynamic Modeling and Fault Identifications Methods for Rolling Element Bearing, Procedia Technology, Volume 14, 2014, Pages 447-456, ISSN 2212-0173.

[35] Junning, Li & Xue, Jiafan & Ma, Zhitao. (2018). Study on the Thermal Distribution Characteristics of High-Speed and Light-Load Rolling Bearing Considering Skidding. Applied Sciences. 8. 1593. 10.3390/app8091593.

[36] Wang, Liqin & Chen, Guanci & Gu, Le & Z Zheng, D. (2007). Operating temperature in high-speed ball bearing. Proceedings of The Institution of Mechanical Engineers Part C-journal of Mechanical Engineering Science - PROC INST MECH ENG C-J MECH E. 221. 353-359. 10.1243/0954406JMES476.

[37] Jiang S, Mao H. Investigation of the High Speed Rolling Bearing Temperature Rise With Oil-Air Lubrication. ASME. J. Tribol. 2011;133(2):021101-021101-9. doi:10.1115/1.4003501.

[38] P. M. Cann, J. P. Doner, M. N. Webster & V. Wikstrom (2001) Grease Degradation in Rolling Element Bearings, Tribology Transactions, 44:3, 399-404, DOI: 10.1080/10402000108982473.

[39] P. M. Cann, M. N. Webster, J. P. Doner, V. Wikstrom & P. Lugt (2007) Grease Degradation in R0F Bearing Tests, Tribology Transactions, 50:2, 187-197, DOI: 10.1080/10402000701261003.

[40] Brüel & Kjær, 2850 Nærum, Denmark, Mearing vibration.

[41] Hansford Sensors, Acceleration enveloping to detect bearing damage, White Paper Nov 2016.

[42] Adeline Bourdon, Simon Chesné, Hugo André, Didier Rémond, Reconstruction of angular speed variations in the angular domain to diagnose and quantify taper roller bearing outer race fault, Mechanical Systems and Signal Processing, Volume 120, 2019, Pages 1-15, ISSN 0888-3270.

[43] Yi Wang, Guanghua Xu, Qing Zhang, Dan Liu, Kuosheng Jiang, Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions, Journal of Sound and Vibration, Volume 348, 2015, Pages 381-396, ISSN 0022-460X.

[44] Remond, Didier & Antoni, Jerome & Randall, R.B.. (2014). Editorial for the special issue on Instantaneous Angular Speed (IAS) processing and angular applications. Mechanical Systems and Signal Processing. 44. 1–4. 10.1016/j.ymssp.2013.11.006.

[45] Alan Hase, Hiroshi Mishina, Masaki Wada, Fundamental study on early detection of seizure in journal bearing by using acoustic emission technique, Wear, Volumes 346–347, 2016, Pages 132-139, ISSN 0043-1648.

[46] Md Junayed Hasan, M.M. Manjurul Islam, Jong-Myon Kim, Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions, Measurement, Volume 138, 2019, Pages 620-631, ISSN 0263-2241.

[47] Nabhan, Ahmed & Ghazaly, Nouby & Samy, Abdelhalim & M.O, Mousa. (2015). Bearing Fault Detection Techniques - A Review. Turkish Journal of Engineering, Sciences and Technology.

[48] Hasib Alian, Shlomi Konforty, Uri Ben-Simon, Renata Klein, Moshe Tur, Jacob Bortman, Bearing fault detection and fault size estimation using fiber-optic sensors, Mechanical Systems and Signal Processing, Volume 120, 2019, Pages 392-407, ISSN 0888-3270.

[49] Kuhnell BT, Stecki JS. Correlation of vibration, wear debris analysis and oil analysis in rolling element bearing condition monitoring. Maintenance Management Int., Vol. 5, PP.105–15, 1985.

[50] Duy-Tang Hoang, Hee-Jun Kang, Rolling element bearing fault diagnosis using convolutional neural network and vibration image, Cognitive Systems Research, Volume 53, 2019, Pages 42-50, ISSN 1389-0417.

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

[52] McFadden, P.D. and Smith, J.D. (1984 A): Vibration monitoring of rolling element bearings by the high frequency resonance technique – a review. Tribology International, vol. 117(1), pp.3-10.

[53] McFadden, P. D. and Smith, J.D., (1984 B): Model for the vibration produced by a single point defect. Journal of Sound and Vibration, vol. 96 (1), pp. 69-82.

[54] D. Abboud, J. Antoni, S. Sieg-Zieba, M. Eltabach, Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment, Mechanical Systems and Signal Processing, Volume 84, Part A, 2017, Pages 200-226, ISSN 0888-3270.

[55] Dyer, D. and Stewart, R.M. (1978): Detection of rolling element bearing damage by statistical vibration analysis. American Society of Mechanical Engineers, pp. 26 - 30.

[56] D.F. Dwyer, Detection of non-Gaussian signals by frequency domain kurtosis estimation, in: International Conference on Acoustic, Speech, and Signal Processing, Boston 1983, pp. 607–610.

[57] Antoni, J. and Randall, RB. (2006): The Spectral Kurtosis: application to the surveillance and diagnostics of rotating machines. Mechanical Systems and Signal Processing, vol. 20(2), pp.308-331.

[58] Jérôme Antoni, Fast computation of the kurtogram for the detection of transient faults, Mechanical Systems and Signal Processing, Volume 21, Issue 1, 2007, Pages 108-124, ISSN 0888-3270.

[59] Yongxiang Zhang, R.B. Randall, Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram, Mechanical Systems and Signal Processing, Volume 23, Issue 5, 2009, Pages 1509-1517, ISSN 0888-3270.

[60] Lei Wang, Zhiwen Liu, Qiang Miao, Xin Zhang, Time–frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis, Mechanical Systems and Signal Processing, Volume 103, 2018, Pages 60-75, ISSN 0888-3270

[61] J. Antoni, "Cyclostationarity by examples," Mechanical Systems and Signal Processing, vol. 23, no. 4, pp. 987–1036, 2009.

[62] J. Antoni, D. Hanson, Detection of surface ships from interception of cyclostationary signature with the cyclic modulation coherence, IEEE J. Oceanic Eng. 37 (3) (2012) 478–493.

[63] Jérôme Antoni, Ge Xin, Nacer Hamzaoui, Fast computation of the spectral correlation, Mechanical Systems and Signal Processing, Volume 92, 2017, Pages 248-277, ISSN 0888-3270.

[64] Ali Moshrefzadeh, Alessandro Fasana, The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis, Mechanical Systems and Signal Processing, Volume 105, 2018, Pages 294-318, ISSN 0888-3270.

[65] R. Potter, M. Gribler, Computed order tracking obsoletes older methods, SAE Paper 891131, 1989.

[66] Yi Wang, Guanghua Xu, Ailing Luo, Lin Liang, Kuosheng Jiang, An online tacholess order tracking technique based on generalized demodulation for rolling bearing fault detection, Journal of Sound and Vibration, Volume 367, 2016, Pages 233-249, ISSN 0022-460X.

[67] Tangfeng Yang, Yu Guo, Xing Wu, Jing Na, Rong-Fong Fung, Fault feature extraction based on combination of envelope order tracking and cICA for rolling element bearings, Mechanical Systems and Signal Processing, Volume 113, 2018, Pages 131-144, ISSN 0888-3270.

[68] Stephan Schmidt, P. Stephan Heyns, Konstantinos C. Gryllias, A pre-processing methodology to enhance novel information for rotating machine diagnostics, Mechanical Systems and Signal Processing, Volume 124, 2019, Pages 541-561, ISSN 0888-3270.