ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

VIBRATION-BASED FAULT DETECTION FOR BALL BEARINGS

M.Sc. THESIS

Reza GOLAFSHAN

Department of Mechanical Engineering

Solid Mechanics Programme

MAY, 2015

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

VIBRATION-BASED FAULT DETECTION FOR BALL BEARINGS

M.Sc. THESIS

Reza GOLAFSHAN
(Student ID: 503121512)

Department of Mechanical Engineering

Solid Mechanics Programme

Thesis Advisor: Prof. Dr. Kenan YUCE SANLITURK

MAY, 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

Bilyalı Rulmanlarda Titreşim Verileri Kullanılarak Hasar Tespiti

YÜKSEK LİSANS TEZİ

Reza GOLAFSHAN
(Enstitü No: 503121512)

Makina Mühendisliği Anabilim Dalı

Katı Cisimlerin Mekaniği Programı

Tez Danışmanı: Prof. Dr. Kenan YÜCE ŞANLITÜRK

MAYIS, 2015

Reza-golafshan, a **M.Sc.** student of ITU, **Graduate School of Science Engineering and Technology** student ID **503121512**, successfully defended the **thesis** entitled "**VIBRATION-BASED FAULT DETECTION FOR BALL BEARINGS**", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor: Prof. Dr. Kenan Y. ŞANLITÜRK .....
Istanbul Technical University

Jury Members:

Prof. Dr. Ekrem TÜFEKÇİ .....
Istanbul Technical University

Dr. Egemen TINAR .....
Arçelik Company

Date of Submission : 04 May 2015
Date of Defense : 28 May 2015

To my mother & fiance
For their unconditional love

FOREWORD

It is a rare happening, if not impossible, for a graduate student like me to indulge himself in a set of unique opportunities and facilities in writing his thesis, for which I am sincerely thankful.
I mustn't forget that many people knowingly and unknowingly have contributed to this work, yet a few cannot be left unmentioned. First and foremost, I especially single out my dear Professor Sanliturk whom I am deeply indebted to. Not only did he allow me to utilize a fully equipped laboratory but also he gave me invaluable advice with meticulous care throughout the process of the formation of this thesis, which would not have come to its present stage without his unstinting support. It was his insightful and direct comments that helped me to rethink from the title to writing the last line. I am extremely grateful to him.
I owe a great deal to ORS Company in general and Dr. Merdane in particular in supplying me with the required equipments in the lab. My thanks are also due to Dr. Cakar from Firat Univesity for his help and support. It would be wrong of me to forget to mention Dr. Abdi from Sahand University of Technology who aroused my interest in and directed my attention to this area of mechanical engineering. I am also thankful to Mr. Terzioglu for his translation of the abstract into Turkish.
I wish to deeply thank the owner of my heart, my lovely fiance, Saba, for whom I barely spared time while working on my thesis.
Last but not least, my very special thanks to my mother who has always shielded me graciously from all the toughness during my entire life.

May 2015

Reza GOLAFSHAN
Mechanical Engineer

TABLE OF CONTENTS

Page

ABBREVIATIONS

CM : Condition Monitoring
RMS : Root Mean Square
TSA : Time Synchronous Average
AR : Autoregressive
EMD : Empirical Mode Decomposition
EEMD : Ensemble Empirical Mode Decomposition
IMF : Instrinic Mode Function
SVD : Singular Value Decomposition
SSA : Singular Spectrum Analysis
FFT : Fast Fourier Transform
BPFO : Ball Pass Frequency Outer
BPFI : Ball Pass Frequency Inner
BSF : Ball Spin Frequency
FTF : Cage (train) frequency
MED : Minimum Entropy Decomposition
SK : Spectral Kurtosis
STFT : Short Time Fourier Transform
WT : Wavelet Transform
DWT : Discrete Wavelet Transform
CWT : Continuous Wavelet Transfom
FWT : Fast Wavelet Transform
WPT : Wavelet Packet Transform
SGWT : Second Generation Wavelet Transform
RDWT : Redundant Discrete Wavelet Transform
DTCWT : Dual-Tree Complex Wavelet Transform
WVD : Wigner Ville Distribution
CWD : Choi-Williams Distribution
HHT : Hilbert-Huang Transform
PDF : Probability Density Function
FRF : Frequency Response Function
IRF : Impulse Response Function
AM : Amplitude Modulation
FM : Frequency Modulation
HFRT : High Frequency Resonance Technique
HT : Hilbert Transform
FIR : Finite Impulse Response
IIR : Infinite Impulse Response
OFM : Objective Function Method
MSE : Mean Square Error
SNR : Signal-to-Noise Ratio
MRA : Multi-Resolution Analysis
QMF : Quadrature Mirror Filters

LIST OF TABLES

Page

LIST OF FIGURES

Page

VIBRATION-BASED FAULT DETECTION FOR BALL BEARINGS

SUMMARY

The need for more advanced and well-organized maintenance methods is increasing day by day in many industries. Condition-based monitoring as one of the most effective, reliable, and economical maintenance strategies is widely employed to minimize operating and overhaul costs. Among a wide range of available Condition Monitoring (CM) methods, vibration-based methods are the most prevalent methods as they offer many advantages compared to other options. It is widely accepted that vibration measurement is able to sense most of the machinery faults by the change in vibration signals, which they produce and is reported to be the chosen method by most researchers to study machine defects. The most important advantage of using the vibration signals is that, any change in the condition of the machine is reflected immediately, and therefore it can be used for discrete as well as continuous monitoring methods. In addition, since each fault on the machine has specific symptom(s) on dynamic behavior of the system, vibration-based condition monitoring can also be considered as a strong tool, which allows detecting and diagnosis the type, size, location and even development of the faults on machines.

It is well known that ball bearings have very significant role on the global vibrations of rotating machineries. Ball bearings are one of the most crucial components in industrial machines and due to their critical role, it's of great importance to monitor their conditions under operation. However, the captured vibration signals from bearing housings can contain the effects of many vibration sources that may not have explicit diagnosis information. Ball bearing vibration signals can be corrupted by uninterested components as well as heavy background noise. Therefore, in most cases, routine procedures and techniques in the field of vibration signals processing are unable to detect the presence of probable faults on ball bearings. Consequently, the researchers are always trying to propose and develop advanced and novel methods in the field of signal processing for condition monitoring.

Realization of the need for novel methods that can provide more reliable results for diagnosis purposes has led to investigate and study the advanced signal processing methods. For this purpose, a wide range of available fault detection as well as noise elimination methods for ball bearings are investigated in the present study.

In the line with the purpose of this study, the first chapter is an introduction containing a very brief description of CM and a literature review for available research studies. Then, a chapter including some mathematical concepts of the available fault detection methods is provided to ensure comprehension of the full text. This section includes vital concepts of the investigated methods and their applications in ball bearing fault detection such as various de-noising and filtering methods, envelope analysis, Minimum Entropy Deconvolution (MED), and Wavelet Transform (include both discrete and continuous forms). In addition, a novel de-noising method based on Singular Value Decomposition (SVD) for ball bearing vibration signals is proposed in

the present study. The remainder of this study is dedicated to numerical and experimental results.

Generally, it can be said that the fault diagnosis algorithms may yield excellent results under some specific conditions. However, their performance can be very poor in other situations. Since there is not a universally superior method for all ball bearings fault detection purposes, it is necessary to study a wide range of fault detection methods for the purpose. Before carrying out experimental studies, numerically simulated ball bearing vibration signals are used to test various methods and approaches so as to complete understanding of the available methods. MATLAB® software is used to generate and analyze simulated vibration signals using advanced signal processing tools used in the field of fault detection. However, very significant and important part of the CM research is devoted to investigation of the diagnosis methods experimentally. The test rig system for this research work presented in this thesis is provided by Istanbul Technical University (ITU), Mechanical Engineering Department, Vibration&Acoustic Laboratory. Captured vibration signals are used for comparing the results of a wide range of diagnosis methods in time, frequency, and time-frequency domains. Empirical Mode Decomposition (EMD), Linear Prediction, Minimum Entropy Deconvolution (MED), Discrete Wavelet Transform (DWT) and its revised versions, and Wiener filtering was used for de-noising purpose. In addition, Continuous Wavelet Transform (CWT) as strong time-frequency tool is used to analyze various frequency sub-bands of the signals. After all, a very effective and reliable de-noising method based on SVD and Hankel matrix is proposed and successfully applied for ball bearing vibration signals.

This study inspires the usage of modern and more advanced methods in vibration signals analyses for ball bearing applications. Further studies are also possible such as the opportunities on modifying and improving the available fault detection algorithms to achieve a faster implementation, which is quite important for online signal conditioning.

Bilyalı Rulmanlarda Titreşim Verileri Kullanılarak Hasar Tespiti

ÖZET

Pek çok sanayi alanında daha gelişmiş ve iyi düzenlenmiş olan bakım metotlarına duyulan ihtiyaç, gün geçtikçe artmaktadır. En etkili, en güvenilir ve en ekonomik bakım stratejilerinden biri olan "Condition-Based Monitoring", bakım işlemini minimize etmek ve maliyeti azaltmak etmek için yaygın bir şekilde kullanılır. Ulaşılabilir pek çok Condition Monitoring (CM) metotları içerisinde, titreşim tabanli metotlar, diğer metotlarla karşılaştırıldığında pek çok avantaj sunmakta ve bu nedenle de yaygın olarak kullanılmaktadır. Titreşim sinyallerindeki değişimlerin makine arızalarının çoğunu tanımlayabildiği görülmüştür. Titreşim verilerinin kullanılmasının en önemli avantajlarından birisi de, makinenin durumundaki herhangi bir değişikliğin anında titreşim sinyaline yansımasıdır. Böylece, sürekli görüntüleme metotlarında kullanıldığı gibi parçalı durumlarda da kullanılabilir. Ayrıca, makinedeki her arıza, sistemin dinamik davranışında farklı belirtilere sahip olduğu için, titreşim analizi makinelerdeki hataların türünü, büyüklüğünü, yerini ve hatta nasıl geliştiğini tanımlayabilen güçlü bir araç olarak bilinmektedir.

Makinalarda çeşitli mekanizmalar ile üretilen dinamik yükler gövdeye çoğunlukla yataklardaki yuvarlanan elemanlar tarafından iletilirler. Bunun için, bilyalı rulmanların, dönen makinelerin genel titreşimleri üzerinde oldukça önemli bir rolü olduğu bilinmektedir. Bilyalı rulmanlar, endüstriyel makinelerin içerisindeki en önemli parçalardan biridir ve onların önemli bir role sahip olmalarından dolayı, işlem altındaki koşullarını gözlemlemek oldukça önemlidir. Ancak, yataklardan elde edilen titreşim sinyalleri, açık ve net hata tespit bilgisi olmayan pek çok titreşimleri de içerebilir. Bilyalı rulmanların titreşim sinyalleri, yüksek arka plan gürültüsü gibi, ilgilenilmeyen bileşenler tarafından bozulabilir. Bu sebeple çoğu durumda, titreşim sinyalleri alanındaki rutin yöntem ve teknikler, bilyalı rulmanların olası arızalarının varlığını belirleyemeyebilir. Dolayısıyla araştırmacılar sürekli olarak, sinyal işleme alanında ileri düzeyde yeni metotları geliştirip, çözüm geliştirme çabası içerisindedirler

Arıza tespiti amacıyla, daha güvenilir sonuçlar sağlayabilecek olan yeni metotlara duyulan bu ihtiyacın fark edilmesi, gelişmiş sinyal işleme metotlarının incelenmesini ve araştırılmasını sağlamıştır. Bu amaç doğrultusunda, gürültüyü minimize etme metotlarının yanı sıra, bilyalı rulmanlar için mevcut olan, pek çok arıza tespiti bu tez çalışması kapsamında incelenmiştir. Bu çalışmada bilyalı rulmanların titreşimlerini incelemek için bilgisayar destekli mühendislik yöntemleri kullanılmıştır.

Bu çalışmanın amacı doğrultusunda, ilk bölüm CM'nin çok kısa bir tanımını ve mevcut araştırma çalışmaları için kaynak taramasını kapsayan giriş bölümüdür. Bu bölümde, önceki çalışmaların kısa tanıtımları yapılıp ve avantajlar/dezavantajları hakkında bilgiler verilmektedir.

İkinci bölüm ise, mevcut hata tespit metotlarının matematiksel kavramlarını içeren teori bölümüdür. Bu bölüm, de-noising ve filtreleme metotları, Envelope Analizi,

Minimum Entropy Deconvolution (MED) ve Wavelet Trasform (ayrık ve sürekli durum için) gibi bilyali rulmanların arıza tespitinde incelenen metotların ve uygulamalarının önemli kavramlarını içermektedir. Ayrıca, bilyali rulmanların titreşim sinyalleri için, Singular Value Decomposition (SVD)' na dayalı ve yeni de-noising metodunun temel kapsamları bu tezde tanıtılmıştır. Çalışmanın geri kalanı, aşağıda kısaca bahsedildiği gibi nümerik ve deneysel sonuçları kapsamaktadır.

Genel olarak, bazı özel koşullar altında, arıza tespit algoritmalarının mükemmel sonuçlar sağladığı söylenebilir. Fakat diğer durumlarda ise performansları oldukça zayıf olabilir. Bilyalı rulmanların arıza tespiti amaçları doğrultusunda, dünya geneline bakıldığında, üstün başarılı sayılabilecek bir metot bulunmadığından dolayı, daha geniş kapsamlı bir şekilde arıza tespiti için çalışmalar yapmak gereklidir. Bu tezde, deneysel çalışmalara geçmeden önce, sayısal olarak simule edilen bilyali rulmanın titreşim sinyalleri kullanılmış ve farklı metotlar sayısal simulasyonlar ile test edilmiştir. Arıza tespiti alanında gelişmiş sinyal işleme araçlarını içerisinde barındıran, MATLAB® yazılımı, simule edilen titreşim sinyallerini analiz etmek için kullanılmıştır. Bu çerçevede, birçok farklı metotlar sayısal olarak incelenmiştir. Empirical Mode Decomposition (EMD) ve Ensemble Empirical Mode Decomposition (EEMD) simule edilmiş sinyaller üzerinde uygulanıp, bu metotların kullanım sınırları tespit edilmiştir. Wavelet Transform ile mevcut sinyal temizleme metotlar çoğunlukla thresholding tekniği ile yapılmaktadır. Bunun için, farklı Discrete Wavelet Transform (DWT)'un versiyonları, değişik thresholdlarla (Universal Threshold, Rigorous Sure Threshold ve Minimax Threshold) thresholding tekniklerle (Hard ve Soft) incelenmiştir. Second Generatin Wavelet Transform (SGWT), Redundant Discrete Wavelet Trasnform (RDWT) ve Dual-Tree Comlex Wavelet Transform (DTCWT) geleneksel DWT'unun iyileştirilmiş ve daha gelişmiş versiyonları olarak karşılaştırılmıştır. Sayısal testlerden çıkan sonuçlar, DTCWT'nun daha etkin olduğunu göstermektedir. Bu bölümde, ayrıca Continuous Wavelet Transform (CWT) ve Short Time Fourier Transform metotlar, en çok kullanılan zaman-frekans analiz teknikleri olarak sayısal sinyaller üzerinde test edilip her iki yöntemin sınırları ve performansları ortaya konulmuştur.

CM araştırmasının önemli bir kısmının, deneysel olarak yapılan tespit incelemesi olduğu dikkate alınmalıdır. Bu araştırma çalışmasında kullanılan test donanım sistemi, İstanbul Teknik Üniversitesi, Makine Mühendisliği Bölümü, Akustik&Titreşim Laboratuvarı tarafından sağlanmıştır. Dördüncü bölümde kullanılan mevcut test düzeneği, arızalı rulmanların teknik bilgileri ve titreşim sinyallerini almak için kullanılan sensörlerin tanıtımıları yaılmaktadır. Bu çalışmada değişik arızalara sahip olan rulmanlar kullanılıp deneysel testler farklı rulmanlar üzerinde yapılmıştır. Ayrıca ilerki çalışmalarda benzer testleri tekrarlamak isteyenler için gereken bilgiler verilmektedir.

Alınan titreşim sinyalleri, zaman, frekans ve zaman-frekans düzleminde, tanıtılan metotların geniş bir alanda sonuçlarının karşılaştırılması amacıyla kullanılmıştır. EEMD, ARMED filtering (AR ve MED filtering'lerin birleşmesi), DWT ve iyileştirilmiş versiyonları ile Wiener filtrelemesi, de-noising yöntemi amacıyla kullanılmıştır. EEMD, EMD metodunun iyileştirilmiş versiyonu olarak, bu bölüm'de sadece EEMD sinyal temizleme için kullanılmıştır. Yine üçüncü bölüm'de kullanılan farklı Wavelet Transform'lar ve thresholding teknikler deneysel olarak da test edilip karşılaştırılmıştır. Güçlü bir zaman-frekans aracı olarak CWT, sinyallerin çeşitli frekansın altbantlarını analiz etmek için kullanılmıştır. Ayrıca, bu bölüm'de SVD ve Hankel matrisine dayanan, oldukça etkin ve güvenilir bir sınyal temizleme yöntemi

önerilmiştir. Bu yöntem başarılı bir şekilde bilyali rulmanların titreşim sinyallerine uygulanmıştır. Küçük, orta ve büyük boy rulman hatarlar için test edilen yöntem, çok gürültülü ortamlardada çok iyi sonuç verdiği gösterilmiştir.

Bu çalışma, bilyali rulman uygulamaları için titreşim analizinde modern ve daha ileri düzeydeki metotların kullanımını teşvik etmiştir. Anlık sinyal iyileştirilmesi için oldukça önemli olan, daha hızlı bir uygulamayı başarmak adına, mevcut arıza tespiti algoritmalarının değişimi ve geliştirilmesi gibi daha ileri çalışmaların yapılması da mümkündür.

1. INTRODUCTION

1.1 Background

Industrial companies nowadays need well-organized maintenance strategies for optimum productivity. The maintenance strategies can be listed as 1) Run-to-Break, 2) Preventive Maintenance and 3) Condition-Based Maintenance [1]. In order to obtain the greatest effectiveness in the operation of complex machines and mechanical systems, condition-based maintenance methods are being widely employed to minimize operating and overhaul costs. Such modern strategies can provide many advantages over the traditional ones (e.g. Run-to-Break). Figure 1.1 illustrates the well known P-F curve, which shows the machine conditions before the failure. Detection of the probable faults when they are at early stages in their developments can prevent the machine failure and consequently, extra imposed costs. The operating conditions of the industrial machineries can be obtained via the advanced monitoring tools.

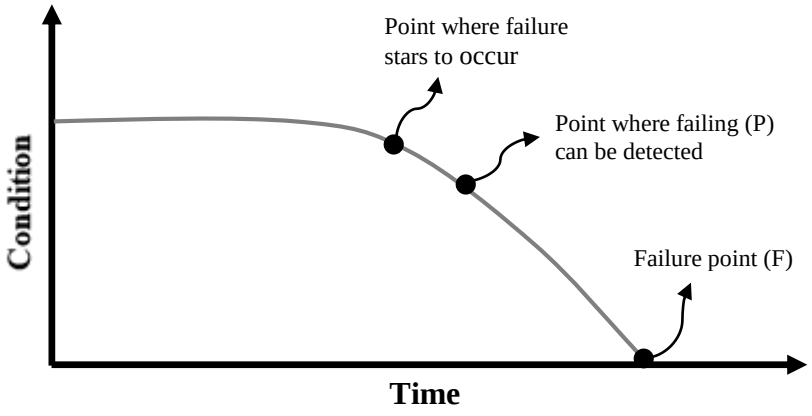

Figure 1.1: P-F curve related to the machine condition.

Condition Monitoring (CM) monitors the current condition and predicts the future condition of machines while in operation. This can be done in terms of machine internal conditions. Two most common and reliable techniques for obtaining information about internal conditions are vibration and lubricant analyses. Other CM methods such as temperature analysis (thermography), performance analysis, and acoustic emission analysis are also used in order to obtain diagnosis information.

Among a wide range of available CM methods, vibration analysis is the most widely used method for machine Condition Monitoring because of having more advantages compared to other methods (e.g. lubricant analysis) [2]. It is widely accepted that vibration measurement is able to identify 90% of machinery faults by the change in vibration signals, which they produce [1] and is reported to be the chosen method by most researchers to study machine defects [3, 4]. The most important advantage of the vibration analysis is that it reacts immediately to any change in the condition of the machine; therefore, it can be used for discrete as well as continuous monitoring methods [5].

According to Robert B. Randall (2011), "*Even in good condition, machines generate vibrations*". Consequently, some post processing techniques are strongly needed to describe the condition of the analyzed machine. Machines with moving parts give rise to vibrations and noise. The status of each machine and its environment yield to create a specific vibration signature. Therefore, a change in the vibration signature related to each machine due to a change in the machine state can be used to detect incipient defects before they become a critical issue. This is the goal of the Condition Monitoring based on vibration analysis. The captured machine vibration signal can be compared with a data bank of the healthy signals to detect abnormalities in the corresponding machine. This procedure does not require a shut down and can be used as an online diagnostic and trend monitoring tool. In addition, since each fault has the specific symptom(s) on the machine dynamic behavior, vibration analysis can also be used for detecting and diagnosis the type, size, location and even development of the faults on machines. Nevertheless, the captured vibration signals can be affected by different components and uninterested vibration sources. Therefore, the major challenge of CM is to specify that the signal contents are related to the state of the monitored component or machine (and not other sources) for further analyses (i.e. extract the diagnosis information). Consequently, a large number of signal processing techniques should be used in order to enhance the signal and to extract relevant information from a measured vibration signal.

1.2 Statement of the Problem

It is well known that the ball bearings have very significant role on the global vibrations of the rotating machineries. Ball bearings remain one of the most crucial

components in industrial machines and due to their critical role, it's of great importance to monitor their conditions under operation.

In most cases, routine procedures and techniques in the field of signal processing are unable to detect the faults on ball bearings. In particular, in very noisy environments, which is quite common in industrial cases, the captured vibration signal from bearing housings can be due to many vibration sources some of which may not have explicit diagnosis information. In addition, ball bearings are made by various components as shown in Figure 1.2 and contributes to vibration signal with specific frequency patterns or components. In many cases, these frequency components can be modulated and consequently, very complex frequency spectrums may result in from machines with ball bearings. Morever, in the case of multiple faults on the ball bearing elements, the mixture of symptoms associated to each fault, can create even more complex time and frequency domains.

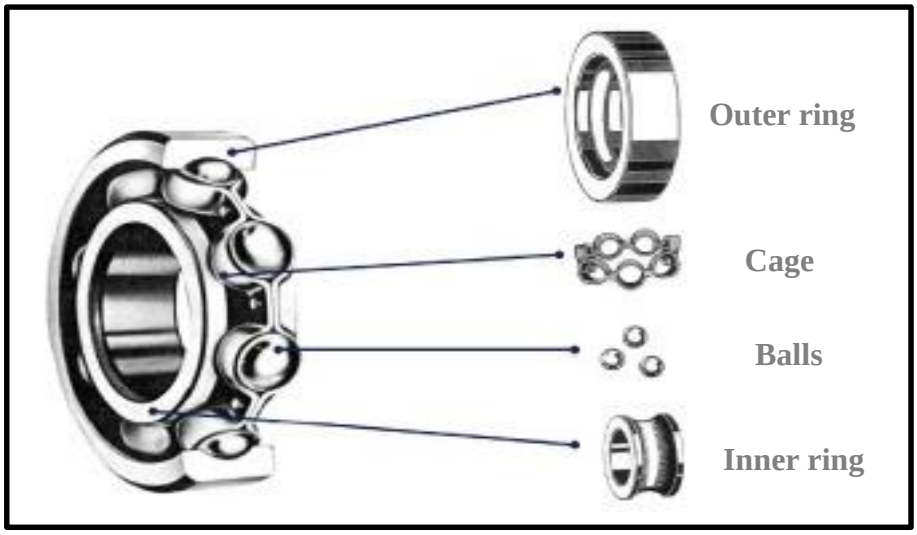

Figure 1.2: A typical ball bearing components.

Many novel and advanced vibration-based fault detection techniques for ball bearings have been developed in recent years; however, there is still a need for more suitable and effective methods for identifying the various faults in the presence of heavy background noise and/or at early stages of its development.

1.3 Literature Review

As stated, vibration analysis has been extensively used for evaluating the conditions of the mechanical rotating systems with ball bearings. Many studies in the literature

have been carried out in order to enhance the capability of signal processing tools used for vibration analyses. It is known that even if the ball bearings are healthy and operate under normal conditions, they produce vibrations and consequently noise, though at low levels. Ball bearing vibration sources include imperfection of the bearing surfaces (i.e. roughness and waviness) and external vibration transmitted via the bearings etc. These external sources along with the unavoidable errors in the measurement system create the background noise and unwanted components in the captured vibration signals. In the literature, there are several signal processing tools in order to extract useful diagnosis information. These methods are either in time domain, or in frequency domain or both.

1.3.1 Time domain techniques

Time domain methods for captured vibration signals try to analyze the amplitude and phase characteristics of the time domain vibration signals and/or eliminate the background noise to reveal the interested features. Depending on environmental conditions and properties of the localized fault (i.e. size and location), some faults may manifest themselves in time domain. However, it is generally agreed that time domain techniques are unable to describe the exact feature(s) of the fault. Time domain analyses are mostly employed as preprocessing steps such as de-noising and enhancing of the vibration signals. In addition, some statistical parameter estimations can be obtained in time domain, which can describe the properties of the analysed signal.

Simple statistical parameters evaluated over the measured time domain signal can give some interesting information about potential of defects. Root Mean Square (RMS), Crest Factor, and the fourth normalized moment of probability density function, so-called Kurtosis, are the most common parameters used in the field of online conditioning and fault detection [6, 7, 8]. RMS gives information about the amplitude of the signal and it can be considered as a parameter to determine the energy level of the signal. Crest factor indicates how extreme the peaks are in a waveform. Kurtosis is a parameter for determining the impulsiveness feature of the vibration signals and widely used in the literature as an indicator of faults [9].

One of the most fundamental signal enhancing techniques is Time Synchreonous Averaging (TSA). This method is considered as a powerfulvibration signal processing technique that can extract periodic components in noisy data, which contain the

diagnosis information. Although the main area of application of the TSA is for gear applications [10, 11], it can also be adapted for ball bearings fault detection [12]. However, it is reported in the literature that this method may be applicable only in case of inner race faults and cannot detect widely separated (wider than distance between two adjacent balls) faults [13].

Linear prediction is another classical way to obtain the periodic part of the signal. The prediction is done with respect to a predefined model using a specific number of data points to predict the next sample of a vibration signal. The major dificulties with this method are (i) choosing the model, (ii) the model coefficients computing algorithm, and (iii) the model order selection criteria (number of taken data points). Among the signal representation models, it is proven that the Autoregressive (AR) model has the most reliable de-noising properties [14]. Altough the Burg method described in [15] is considered as the most preferred and strong coefficient estimation technique [16], the Yule–Walker method offers simple, effective, and low prior requirements for its applications [17]. The choice of optimal values depends on the objectives of the problem as shown later in detail.

Empirical Mode Decomposition (EMD), proposed in [18] and its revised version, so-called Ensembel Empirical Mode Decomposition (EEMD) [19], are two self-adaptive decomposition techniques, which can be considered as the signal enhancing methods. Since captured vibration signals from ball bearings have various components due to its various constitutive parts, the decomposition of the vibration signals into its Instrinic Mode Functions (IMFs) can enable an opportunity to ignore and isolate unwanted components in the signal.

Among a wide range of de-noising methods available in the literature, noise elimination algorithms based on SVD are considerably faster and easier to implement. However, it appears that SVD based methods are mostly used in order to extract statistical parameters, which can be used for characterizing the vibration signals captured from ball bearings [20, 21, 22]. SVD based feature extraction method or Singular Spectrum Analysis (SSA) enables a suitable opportunity to monitor the singular values obtained from vibration signals at different operation conditions. In practice, the measured signals are mostly vectors rather than matrices. The available SVD based de-noising methods employ a trajectory matrix (i.e. Hankel matrix) using the captured data and try to find an appropriate threshold to reconstruct a lower rank

matrix for post processing purposes [23, 24, 25]. In the literature, some methods have been proposed for forming the input matrix in SVD process using measured data [26, 25, 23]. The main problem with the use of SVD in signal processing is the estimation of the noise threshold to separate the unwanted and uncontaminated parts of the signal.

1.3.2 Frequency domain techniques

Although the available time domain techniques can be employed as a preprocessing step in the field of fault detection, these methods are unable to describe the nature of the probable faults. This limitation can be overcomed by frequency domain analyses. Because each component of a ball bearing has its own characteristic frequency, any fault on a specific component excite its corresponding fault frequency. This is the basis of all frequency domain analyses. The FFT of the vibration signal is the most common method for converting the time domain to the frequency domain. As a fault develops in a ball bearing component, a train of peaks appear in the vibration signal with a certain repetition frequency associated with that component. As a result, the corresponding frequency can be detected in the frequency spectrum. When there is a local fault on one of the raceways of a ball bearing, passing roller elements (balls) over the local fault creates pulses at one of the fundamental fault frequencies (BPFO or BPFI, depending on whether the fault is on outer or inner raceways) and the spectrum of the vibration signal contains this fault frequency and its harmonics. On the other hand, if there is a local fault on the rolling elements (balls), the local fault creates impacts on both inner and outer raceways. Therefore, a frequency close to 2×BSF (Ball Spin Frequency) appears in the spectrum. In many cases, this frequency is modulated with BPFO or BPFI. Consequently, a very complex frequency spectrum may result in from this process. In addition, a captured raw ball bearing vibration signal certainly contains some background noise due to environmental conditions and it frequently makes interpretations of the spectrum quite difficult. Therefore, some further processes are strongly recommended in the literature in order to get a smoother spectrum, which can be used to extract diagnostic information [27, 28, 8].

The most fundamental and common technique used in frequency domain is the envelope analysis. Envelope analysis is for amplitude demodulating of a signal for extracting the periodic excitation [29, 30]. However, it is reported in [1] that the envelope analysis is applicable if the impulses are shorter than the spacings between

them, and this is not always the case for high speed machines. The Minimum Entropy Deconvolution (MED) method [31] is designed to reduce the spread of the impulses before the envelope analysis. In addition, MED can be merged with linear prediction filtering based on AR models and the total process is so-called ARMED filtering [17, 32].

Spectral kurtosis (SK) as a very useful tool, particularly for non-stationary signals has been proposed in [33] and successfully applied for ball bearing and gear fault detection in [34]. In very noisy environments, the impacts can be easily covered by background noise. Consequently, the kurtosis as a global indicator is not appropriate, so it is preferable to apply it locally in different frequency bands based on Short Time Fourier Transform (STFT). This is exactly what the SK does. It is also shown in [34] that the SK can be used to optimize the well known Wiener filter to filter out the noisy part of the signal. By combination of the various spectral kurtosis plots obtained by different window length in STFT process, the so-called Kurtogram can be obtained [35, 1]. A fast implemention of the Kurtogram is proposed in [36]. An enhanced version of Kurtogram is also developed and succefully adopted for ball bearings fault detection in [37].

Wavelet Transform (WT) as a powerfull method in time-frequency domain, is widely used for fault detection in literature [38, 39, 5, 40, 41]. WT enables an opportunity to analyze the signal in different frequency bands, as a result, any transient event can be observed in the time domain as well as in the frequency domain. However, the discrete form of WT can be considered as frequency domain tool, which can decompose the discrete vibration signals into different levels related to a specific frequency range [42, 28, 43]. Since the Discrete Wavelet Transform (DWT) decomposes the input signal into various frequency bands, it can be used as a de-noising technique. Wavelet Transform has a quite complex calculations and to overcome this complexity, a fast and simple implemention of WT has been proposed in [44]. The wavelet-based de-nosing process can be roughly summarized in three steps as: (i) decomposing, (ii) thresholding, and (iii) reconstructing [45]. Many improvements have been carried out on traditional DWT such as Second Generation Wavelet Transform (SGWT) [46], Redundant Discrete Wavelet Transform (RDWT) [47], and Dual-Tree Complex Wavelet Transform [48].

1.3.3 Time-Frequency domain techniques

Time-frequency analyses is designed to overcome the limitation of time and frequency domains. Traditional methods in time or frequency domains assume the captured signals are stationary over time. Time-frequency domain techniques, however, contain both time and frequency domain information, which enables investigation of transient features in non-stationary signals. To deal with non-stationary signals, several time-frequency analyses have been developed such as Short Time Fourier Transform (STFT) [1], Continuous Wavelet Transfom (CWT) [42], Wigner Ville Distribution (WVD) [49], Choi-Williams Distribution (CWD) [50], and Hilbert-Huang Transform (HHT) [51]. It should be noted that the most common time-frequency domain methods used for the diagnosis purposes are STFT and CWT.

1.4 Objectives and Scope of Thesis

As stated before, ball bearings have very critical role on the global vibrations of rotating machines. As a result, the aim of the present research is to investigate the available advanced and novel signal processing techniques used for the vibration-based ball bearing fault detection applications. For this purpose, vibration patterns of ball bearings for fault detection purposes investigated experimentally and numerically. A wide range of methods in time, frequency, and time-frequency domains are studied using both the simulated vibration signals and the actual captured ball bearing vibration signals in MATLAB® platform. The research in this thesis concentrates on some of the signal pre-processing steps in ball bearing fault detection procedures such as signal enhancing and de-noising. It is shown that some signal de-noising techniques as well as signal enhancing methods in time, frequency, and time-frequency domains have very significant role in the field of vibration-based ball bearings fault detection.

More specificlly, the main objectives of this research may be listed as:

1) Review of the vibration-based fault detection, signal de-noising and enhancing methods and their applications in ball bearings.
2) Modelling and generating the ball bearing vibration signals for numerical investigations.
3) Developing MATLAB® programs for advanced diagnosis.

4) Investigation of the signal enhancing and de-noising in time domain prior to post processing.
5) Analyzing the fault detection methods in frequency and time-frequency domains.
6) Investigation of various techniques used for ball bearing fault detections.
7) Determining amd comparing the performance of the fault detection methods for different ball bearing faults.

1.5 Outline of the Thesis

Following this introductory chapter, Chapter 2 reports more extensive literature review of the previous works as well as detailed description of the mathematical formulations of a wide range of signal processing methods and their applications in ball bearings fault detection, including some statistical parameters obtained in time domain time domain, signal de-noising and enhancing methods, signal decomposition methods, wavelet analysis, etc. Chapter 3 provides a comprehensive numerical investigations of the ball bearing diagnosis methods. The simulated vibration signals of single and multiple defective ball bearings are generated and analysed in this chapter. Prior to experimental investigations, Chapter 4 provides a describtaion of the test rig used in the experiments, including system design, vibration trasducers, data acquisition system and tested ball bearings. As a matter of fact, the fault detection methods need to be validated experimentally. In line with the purpose of this study, Chapter 5 presents the experimental studies, aiming to determine the performances of the fault detection methods when measured data are used. Finally, major conlusions and contibutions are summarized in Chapter 6, together with some recommedations for future works.

2. THE THEORY OF VIBRATION-BASED FAULT DETECTION

2.1 Introduction

Any abnormal change of modal properties of structures of machines leads to changes in their dynamic behavior. Consequently, these changes can be detected in vibration signals that have been captured from the corresponding machine. In this chapter, available faults detection and diagnosis methods for ball bearings based on signal processing approach are summarized in some detail.

A localized fault in ball bearings means a local small damage on raceways or balls. Hence, it acts like an impact to each moving part of defective ball bearing on each rotation cycle. Impacts increase vibration levels of the ball bearings with a certain repetition frequency depending on the location of the fault. Figure 2.1 from [52] illustrates a defective bearing and its ideal vibrations signal. As can be seen, there is a local fault on the outer race of the ball bearing, which creates impacts with a certain repetition frequency related to outer race, so-called Ball Passing Frequency Outer (BPFO). This fault can be easily detected in both time and frequency domains since it produces periodic vibrations.

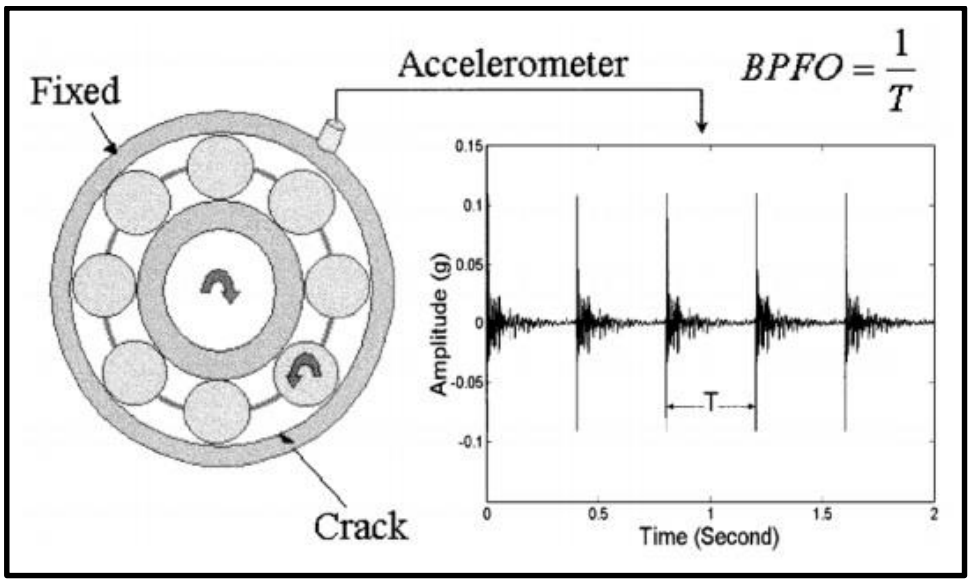

Figure 2.1: Illustration of localized defect on outer race and its vibration signal, from [52].

However, it is not always possible to identify faults and this could be the case in many ball bearing vibration analyses. In industrial environments, vibration signals can be covered with heavy background noise originating from measurement system. Also, the vibration signal might be quite complicated due to additional complications including the effects of misalignment, unbalance, crack(s) on the shaft, looseness, and distortions [53]. Morever, multiple faults on a ball bearing, which is also very common in industrial environments, generate various impulse trains in vibration signal and it may lead to wrong diagnostic information. This is the reason for carrying out some detailed signal processing to extract valuable information from contaminated signals.

Bearing defects may be categorized as distributed or local. Distributed defects include surface roughness, waviness, misaligned races, and off-size rolling elements [4]. In the present study, local faults include cracks, pits, and spalls on ball bearings have been investigated in details.

As briefly discussed in the previous chapter, there are a large number of signal processing techniques in order to extract interesting information (i.e. diagnosis) from measured vibration signals of the ball bearings. These techniques can be divided into three categories: Time Domain Analyses, Frequency Domain Analyses, and Time-Frequency Domain Analyses. In what follows, some of these techniques are described in some detail.

2.2 Time Domain Analyses

Localized faults on ball bearings generate sharp impacts in vibration signals and as a result increase the amplitudes of the corresponding signals. Time domain methods try to analyze the amplitude (and phase) information of the time domain vibration signals and/or eliminate the background noise to reveal the periodic part of the signal from measured data. Time domain techniques include Time Synchronous Average (TSA), correlation analyses, de-noising processes, statistical parameter estimations etc.

Depending on environmental conditions and properties of the localized fault (i.e. size and location), some faults may manifest themselves in time domain. However, it is generally agreed that time domain techniques are unable to describe the exact feature(s) of the fault.

2.2.1 Time synchreonous averaging (TSA) method

The classic way of separating periodic components from background noise for a given signal is by Time Synchronous Averaging (TSA). For example, for each rotation of a shaft, all vibrations data are collected and averaged in order to obtain periodic parts of the signals, which have valuable diagnosis information. It should be noted that this technique has a pivotal drawback that is related to the complexity of the measurement equipment. As a matter of fact, an additional sensor (tachometer) is mostly required in order to measure, for example, the rotational shaft speed for the synchronisation. The time synchronous averaging also requires many revolutions to be averaged in order to yield good results [54].

In practice, it is done by averaging together a series of signal segments each corresponding to one period of a synchronising signal. Thus,

$$\bar{x}(t) = \frac{1}{N}\sum_{n=0}^{N-1} x(t + nT) \qquad \textbf{(2.1)}$$

where $\bar{x}(t)$, $x(t)$, N, and T are output signal, captured (original) signal, number of averaging and the period of rotating, respectively.

The main area of application of the TSA is for gear applications. However, TSA can also be adapted for ball bearings fault detection [12]. From bearing fault detection point of view, this method may be used to obtain repetition frequencies of the impacts due to defects on the ball bearing components on each rotation cycle of ball bearing. Nevertheless, it is reported in the literature that this method may be applicable only in case of inner race faults and cannot detect widely separated (wider than distance between two adjacent balls) faults [13].

2.2.2 Estimation of statistical parameters

Simple statistical parameters evaluated over the measured time domain signal can give some interesting information about potential of defects. In particular, in on-line conditioning of a signal, which real time analyses is quite important, having some statistical information can assist to have at least an initial idea about the status of the signal. Since ball bearing faults make impulse trains, statistical parameters can specify the impulsiveness in the vibration signals. For example, the peak value and the root mean square (RMS) are referred to the overall vibration level of the signal. These

statistical parameters are simple to estimate, however, they are rather insensitive tools for fault detection.

RMS: One of the most common vibration level indicators is the Root Mean Square (RMS) value. According to Antoni and Randall [34], RMS can characterize the intensity of the signal. It is also one of the most relevant and simple measurement of the amplitude. Since RMS give information about the amplitude of the signal, it can be considered as a parameter to determine the energy level of the signal. In machinery vibrations, the RMS value can be used as a health indicator of the machine. This parameter for the vibration signal, $x(t)$, can be easily obtained as

$$\mathrm{RMS} = \sqrt{\frac{1}{T}\int_0^T x(t)^2 dt} \tag{2.2}$$

where T is the measured period.

Crest factor: Crest factor is a statistical parameter of signals, which gives information about the peakiness of the signal. Crest factor indicates how extreme the peaks are in a waveform. Crest factor can be obtained through

$$\text{Crest Factor} = \frac{|\text{Peak Value}|}{\text{RMS}} \tag{2.3}$$

It is worth mentioning that the Crest Factor is unable to detect impulsiveness feature of vibration signals in many cases.

Normalized moments of PDF: Since vibration signals captured from defective ball bearings have both periodic and random components, which can be expressed in terms of propability distribution. This can be expressed for time signals as [1]

$$p(x) = \Pr[x(t) \leq x] \tag{2.4}$$

Probability Distribution can lead to introducing a new function, so-called Probability Density Function (PDF). The probability density function of vibration signal of a normal bearing in good condition has a Gaussian distribution (Figure 2.2) or near-Gaussian distribution [55], whereas a damaged bearing results in non-Gaussian distribution with dominant tails because of existence of high amplitude peaks due to impacts originating from local faults.

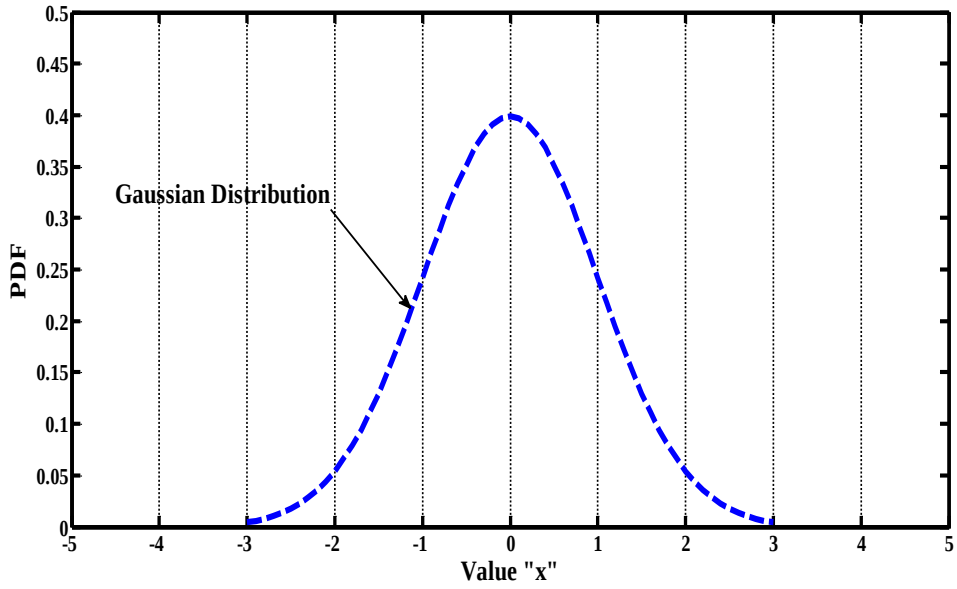

Figure 2.2: Probability density function for the Gaussian distribution.

Instead of studying the probability density curves, it is often more informative and simple to examine the statistical moments of the data points, defined as [4]

$$M_x = \int_{-\infty}^{+\infty} x^n p(x)dx \quad n = 1,2,3,\dots,m \tag{2.5}$$

where n denotes the number of central moments.

The first moment of PDF is the mean value, μ, and given by

$$\mu = \int_{-\infty}^{+\infty} xp(x)dx \tag{2.6}$$

Second, third, and fourth moments of PDF can be taken about the mean value. For example, the second moment can be expressed as

$$\sigma^2 = \int_{-\infty}^{+\infty} [x-\mu]^2 p(x)dx \tag{2.7}$$

where σ is the standard deviation and σ^2 is the variance of the signal.

The third and fourth moments are normalised with respect to the third and fourth power of standard deviation, respectively. The third normalized moment, so-called skewness, which can be calculated as

$$S = \frac{\int_{-\infty}^{+\infty} [x-\mu]^3 p(x)dx}{\sigma^3} \tag{2.8}$$

Since it can detect the impulses of the signal, the fourth moment, so-called kurtosis is quite useful in the field of ball bearing fault detection. Kurtosis is widely used in the literature as an indicator of faults. Kurtosis is given by

$$K = \frac{\int_{-\infty}^{+\infty}[x-\mu]^4 p(x)dx}{\sigma^4} \qquad \textbf{(2.9)}$$

The kurtosis value can immediately reacts to local faults and becomes greater for defective ball bearings. Although kurtosis is an effective tool for fault detection [6, 7], it could in noisy environments have high value due to the occurrence of random small impacts or low value due to covering impulses behind the heavy noise. To overcome the false alarms drawback, P. Tse and J. Leung in [8] have proposed the use of RMS×Kurtosis. When a vibration signal from a defective ball bearing is influenced by small noise, the overall vibration energy is still small, hence, the RMS value should be small. However, Kurtosis value for the same signal is large due to impacts. On the other hand, when a vibration signal from the same ball bearing is immersed in a heavy noisy environment, the overall vibration energy increases, resulting in increase in RMS value. However, due to the large energy, impacts are either averaged or overwhelmed by this large noise. Therefore, the kurtosis value becomes small. In Figure 2.3, two vibration signals from the same ball bearing with outer race defect but in different environments (noisy and ideal conditions) illustrates this point. As can be seen, kurtosis value for the signal covered by heavy noise is small. Nevertheless, the actual value of kurtosis for vibration signal of that ball bearing is high.

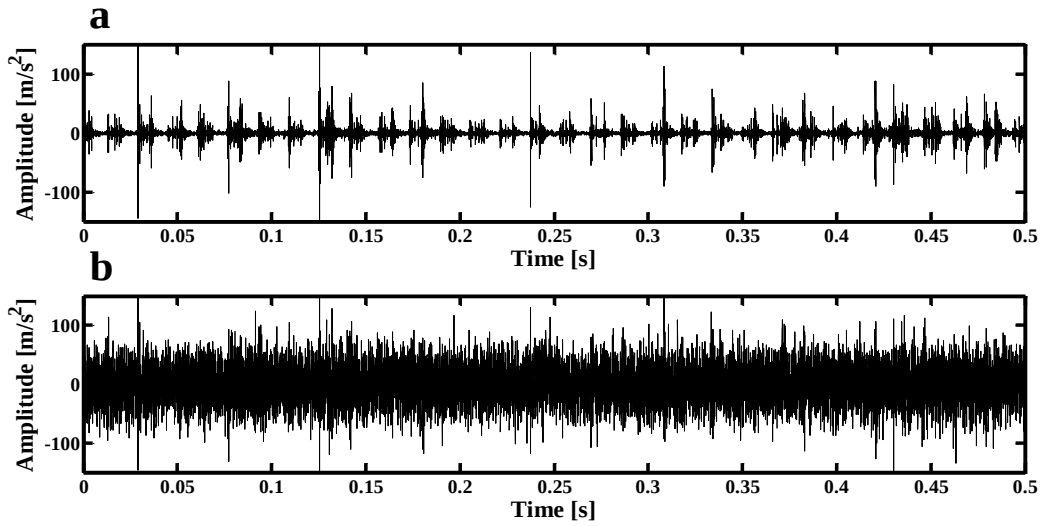

Figure 2.3: Vibration signal captured from an outer race defective ball bearing in a) ideal condition (RMS= 11.4, Kurtosis=44); b) noisy condition (RMS= 32, Kurtosis= 3.7).

2.2.3 De-noising methods

De-noising appears to be one of most important steps in field of Condition Monitoring and fault detection. When time domain methods are unable to describe the nature of the faults, de-noising of the vibration signals in time domain prior to future process (i.e. frequency domain analyses) is strongly recommended. Even though the captured vibration signal contains small background noise, try to improve the diagnostic feature of the signal by cancelling the unwanted components is always one of most interesting studies in the literature.

2.2.3.1 Singular value decomposition (SVD)

SVD is a numerical method which states that a matrix [A] of rank L can be decomposed into the product of three matrices, [U] (an orthogonal matrix), [S] (a diagonal matrix), and $[V]^T$ (the transpose of an orthogonal matrix [V]). This method is usually presented as

$$[A]_{m\times n} = [U]_{m\times m}[S]_{m\times n}[V]^T_{n\times n} \quad \textbf{(2.10)}$$

where $U^TU = I$ and $V^TV = I$; [S] is a diagonal matrix containing the square roots of eigenvalues of A^TA, which can be expressed as $S = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_L)$, where $L = \min(m,n)$. These σ_i ($i = 1, 2, \ldots, L$) are called the singular values of the matrix [A].

The method presented in [56] offers a possibility to find the best approximation of the original data points using fewer dimensions by SVD. Therefore, it can be used as a process for data reduction and this makes the SVD a useful tool for de-noising of vibration signals. There exists a m×n matrix $[\bar{A}]$ of rank $l \leq L$, which minimizes the sum of the squared error between the elements of the matrix [A] and the corresponding elements of the matrix $[\bar{A}]$ which can be formed as [56]

$$[\bar{A}] = [U_l][S_l][V_l]^T \quad \textbf{(2.11)}$$

where $[\bar{A}]$ is the reconstructed matrix using only the largest l number of singular values. The rest of singular values are replaced by zero such that. The mentioned theorem is Eckart-Young theorem and the proof of this theorem can be found in [57].

Some analytical investigation may be needed to have a better understanding about the Eckart-Young theorem. Matrix $[A]$ in equation (2.10) is said to be singular if one or more of the last singular values are zero or nearly singular when one or more of them

are below a threshold, ε [23]. Theoretically, the number of non-zero singular values, L, determines the rank of this matrix, i.e.

$$\begin{aligned} \sigma_i > \varepsilon, \quad i = 1, \dots, l \\ \sigma_i \leq \varepsilon, \quad i = l+1, \dots, L \end{aligned} \tag{2.12}$$

where ε is threshold. Note that, various methods can be implemented to determine a proper threshold value.

When the measured vibration data are noisy, matrix [A] comprises two parts; uncontaminated data which contains fault diagnosis information and unwanted part which represents noise. From Equations (2.10) and (2.11), the following equation can also be derived

$$[A] = [\bar{A}] + [N] = [U_l \quad U_0] \begin{bmatrix} S_l & 0 \\ 0 & S_0 \end{bmatrix} \begin{bmatrix} V_l^T \\ V_0^T \end{bmatrix} \tag{2.13}$$

where $[\bar{A}]$ and [N] represent uncontaminated data and noise matrices, respectively. Also S_l contains significant singular values σ_i (i = 1, 2, …, l) which are used to construct the uncontaminated data and S_0 contains small singular values σ_i (i = l +1, …, L).

In practice, the measured signals are vectors rather than matrices. As stated in previous chapter, various approaches have been proposed for forming the matrix [A] in literature. Here, the most effective two approaches are summarized below following.

Y. Chen and P. Zhang in [25] have proposed a simple yet powerfull de-noising method based on SVD. Authors have used a key parameter introduced in [58] to determine how many singular values are required to have an optimal de-noised vibration signal. According to this approach, $[A]_{m\times n}$ can be formed by m vibration signals and each signal has n data points.

According to [59], the following score can be assigned to the first k singular values

$$\lambda_k^2 = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{l} \sigma_i^2} \tag{2.14}$$

where λ_k^2 indicates the fraction of the captured overall variance. For this reason, the fixed value of λ_k^2 can be used for determining the value of k to obtain $[\bar{A}]$.

For extracting de-noised signal, the following steps are followed in [25]:

1) Signals form the the matrix [A] (rows represent captured signals and columns represent the legth of captured signals).
2) Matrix [A] is decomposed by SVD (matrices [U], [S], and [V] are calculated).
3) Assign λ_k^2 to algebraic value 0.5 (the parameter k is obtained).
4) Matrix $[\bar{A}]$ is recostructed by using the first k singular values.

Then, each row of the matrix $[\bar{A}]$ is a de-noised signal.

The major difficulty with this method is that if each signal has, say, 12000 data points, in case of a square or a nearly square matrix which is a key factor to yield convenient results, 10000 or a value near to 10000 vibration signals (experiments) are required for constructing the matrix [A] based on the corresponding approach.

The second de-noising approach based on SVD, which is investigated in the present work has been proposed by K. Y. Sanliturk and O. Cakar in [23] (originally a method for elimination of noise from Frequency Response Functions, FRFs, and Impulse Response Functions, IRFs) based on Hankel matrix for forming the matrix [A] [58]. Although the corresponding method do not suffer from the need of huge amount of experiments for the matrix [A], it needs some preliminary considerations compared to the previous method. From [58], the Hankel matrix for an input discrete vibration signal, $x(k) = x_k$ ($k = 1, 2, \ldots, N$), can be obtained by sliding a window having length of m over the corresponding vector and can be written as

$$[A] = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_2 & x_3 & \cdots & x_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_m & x_{m+1} & \cdots & x_N \end{bmatrix} \quad \textbf{(2.15)}$$

where $m + n - 1 = N$.

Authors in [23] have determined the Hankel matrix for niose cancelling on FRFs and its time domain counterpart, IRFs.

One way to determine the noise threshold and, hence the first *l* singular values in Equation (2.13) is to plot the normalized singular values and to choose the appropriate rank when the normalized singular values approach to an asymptote [26], which is also adopted in [22]. Figure 2.4 shows the first 800 normalized singular values in logarithmic scale for a case study (measured data on a fan casing) in [18].

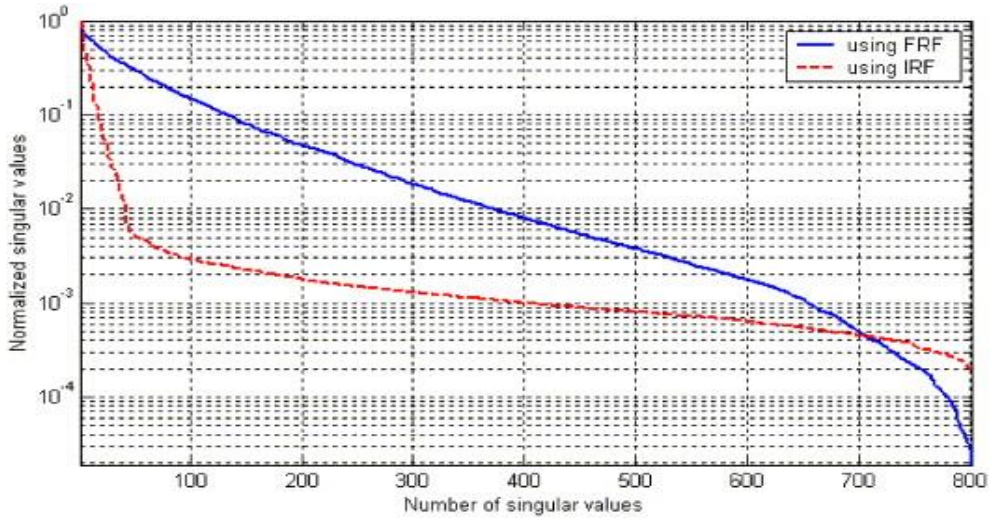

Figure 2.4: Comparison of normalised singular values of matrix [A] using FRF and IRF approaches from [23].

After the matrix $[\bar{A}]$ (using the first *l* singular values) is obtained, it must be converted to the vector form. The reconstruction procedure is done by arithmetic averaging along the anti-diagonals of the matrix $[\bar{A}]$ [23]. This method is also known as diagonal averaging method in [60, 21] and can be expressed as

$$\bar{x}_i = \frac{1}{\beta - \alpha + 1} \sum_{j=\alpha}^{\beta} \bar{A}_{i-j+1,j} \qquad \textbf{(2.16)}$$

where $\alpha = \max(1, i - m + 1)$ and $\beta = \min(n, i)$.

2.2.3.2 Empirical mode decomposition (EMD)

Vibration signals captured from ball bearings have various components (frequencies) due to its various constitutive parts and unbalances in the rotating system. These unwanted components can be modulated in terms of frequencies or amplitudes in the captured vibration signal. Empirical Mode Decomposition (EMD) is a method based on decomposition of the vibration signal into various frequency bands. Unlike the Discrete Wavelet Transform (see section 2.3.6), EMD is a self-adaptive method on each step of decomposition process [61], which is suitable for decomposing of a multi-component signal into its Intrinsic Mode Functions (IMFs). Since EMD does not need to define a basic function and transformation, it can also decompose non-linear and non-stationary signals into a sequence of Amplitude Modulation (AM)/Frequency Modulation (FM) components (IMFs) [62].

EMD originally have been proposed by N. E. Huang *et al*. in [18]. According to their study, each IMF must satisfy two constraints:

C1) In the whole data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one.

C2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

As noted above, the difficulty with EMD method is that there is no any analytical formulation or basic function and transformation. For obtaining an IMF from the captured signal, x(t), the sifting process is comprised the following steps [62]:

1) Detect all the local extrema and then connect all the local maxima by a cubic spline as the upper envelope and all local minima by another cubic spline as the lower envelope (The upper and lower envelopes should include all the data between them).

2) The mean of upper and lower envelope values is defined as m_1, and difference between the signal, x(t), and m_1 is the first component of IMFs, h_1 as

$$h_1(t) = x(t) - m_1(t) \quad \textbf{(2.17)}$$

3) If h_1 is not an IMF, it is treated as the input signal and repeat 1) – 2); then

$$h_{11}(t) = h_1(t) - m_{11}(t) \quad \textbf{(2.18)}$$

In which, m_{11} is the mean of upper and low envelope value of h_1. This process can be repeated up to k times, until h_{1k} becomes an IMF, that is

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \quad \textbf{(2.19)}$$

After each sifting processing, it is necessary to check whether the number of zero crossing equals the number of exterma. If the obtained $h_{1k}(t)$ meets the corresponding conditions, it is considered as the first IMF as follow

$$IMF_1(t) = h_{1k}(t) \quad \textbf{(2.20)}$$

It should be noted that, the first IMF component, IMF_1, should include the highest oscillation frequencies or the shortest period component of signal.

4) Separate IMF_1 from the original data x(t)

$$r_1(t) = x(t) - IMF_1(t) \quad \textbf{(2.21)}$$

where r_1 is called the first residue and contains information of longer period components. Then, r_1 is considered as the first data and repeat the above sifting process. This procedure can be repeated n times to generate n IMFs component. The result is

$$\begin{cases} r_1(t) - IMF_2(t) = r_2(t) \\ \vdots \\ r_{n-1}(t) - IMF_n(t) = r_n(t) \end{cases} \quad \textbf{(2.22)}$$

The sifting process can be stopped when meet below two criteria [62]:

a) The component, $IMF_n(t)$, or the residue, $r_n(t)$, becomes so small as to be considered inconsequential, or

b) The residue, $r_n(t)$, becomes a monotonic function for which no more IMFs can be extracted.

Finally, the below equation can be obtained

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \quad \textbf{(2.23)}$$

Thus, the original signal, $x(t)$, can be decomposed into n empirical modes and residue $r_n(t)$. Residue, $r_n(t)$, is the mean trend of the vibration signal, $x(t)$. The IMFs (IMF_1, IMF_2, IMF_3, … , IMF_n) represent different frequency bands ranking from high to low.

It should be noted that, although EMD is a powerfull tool to separate an input signal into various frequency bands (or scales), this method still suffers from so-called mode-mixing problem, which means that a single IMF can contain components of widely disparate frequencies, or share components with other IMF(s). Mode mixing problem not only leads to serious aliasing in the time-frequency distribution but also renders the physical meaning of individual IMF unclear [19].

Z. H. Wu and N. E. Huang in [19] have proposed a revised EMD method, named Ensemble Empirical Mode Decomposition (EEMD), which can relieve the mode-mixing problem. EEMD is based on the classical EMD method with added white noise – which can be easily generated in MATLAB® platform– at each loop in EMD process. Figure 2.5 shows the flowcharts of EMD and EEMD methods, respectively.

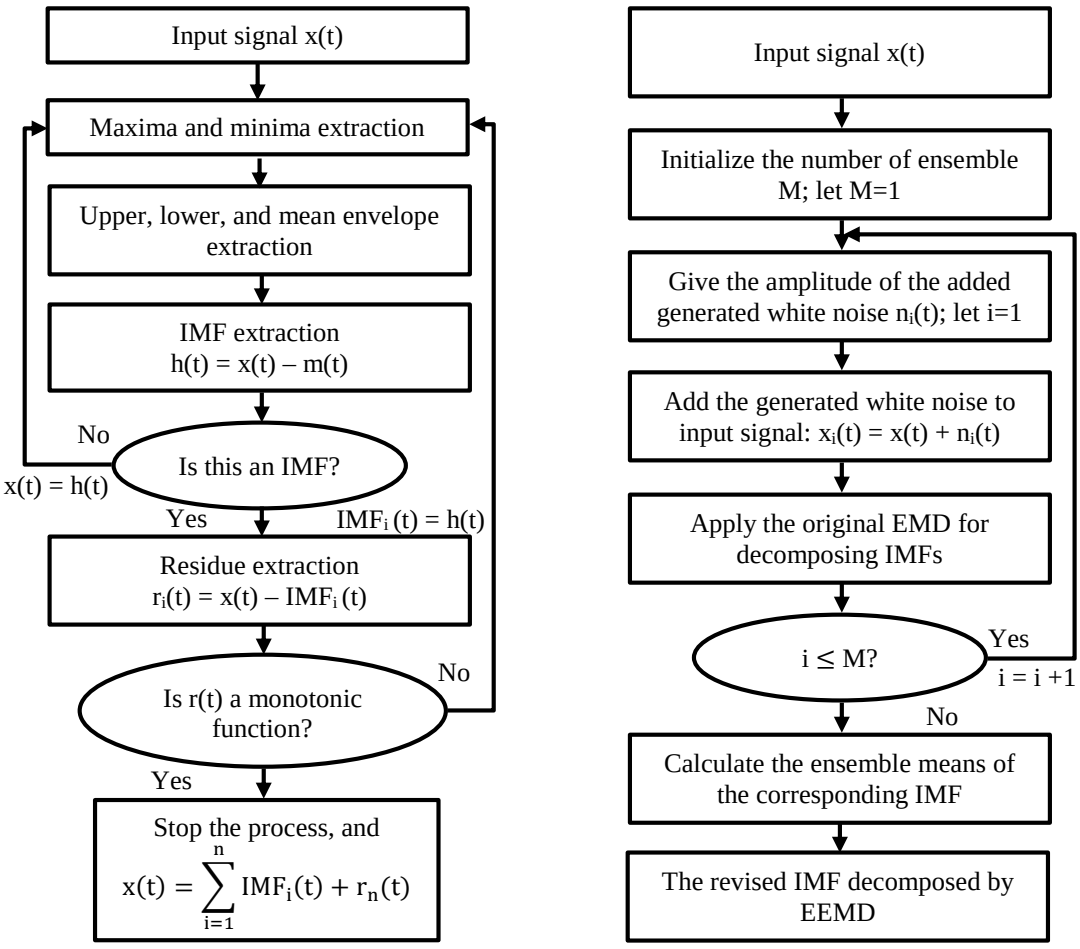

Figure 2.5: The flow charts of EMD (left), and EEMD (right).

As can be seen in Figure 2.5, the principle of EEMD is very simple and it is based on traditional EMD. The added white noise forms by components of various frequencies (or scales) and it would be uniform to inhabit the whole time-frequency representation of the input signal [63]. When a signal is added to the uniformly distributed white noise, the different scale components of the signal are automatically transformed onto proper scales with the reference established by the white noise in the background. Since the white noise is added in every loop, each IMF would be very noisy. As the noise level at each IMF is different, the white noise can be completely removed by the ensemble mean of entire IMFs [63] .

It is quite understandable that the effectiveness of EEMD method depends on the amplitude of added white noise and the ensemble number used in iteration process. Generally, an ensemble number of a few hundred produce reliable results, and the remaining noise would cause no more than a fraction of one percent of error if the added noise has the amplitude that is a fraction of the standard deviation of the

investigated signal [64]. Many studies have been carried out to determine the optimum amplitude of the added noise and the ensemble number [65, 66, 64, 63]. In the present study, however, it is seen that within a certain range of noise amplitude, the decomposition results of the EEMD have a small sensitivity to the noise amplitude. Generally, the added white noise with the standard deviation of 0.8 times the standard deviation of the analyzed signal might be enough for obtaining the reliable decomposition results.

In order to compare the application of EMD and EEMD methods in ball bearing fault detection, W. Guo *et al*. in [67] decompose a captured faulty ball bearing vibration signal into its IMFs using both EMD and EEMD methods. Figure 2.6 illustrates the first nine IMFs for the corresponding case study in [67]. As seen, the mode-mixing problem exists in IMF7 and IMF8 obtained using EMD, whereas the decomposition results in EEMD do not suffer from mode-mixing problem.

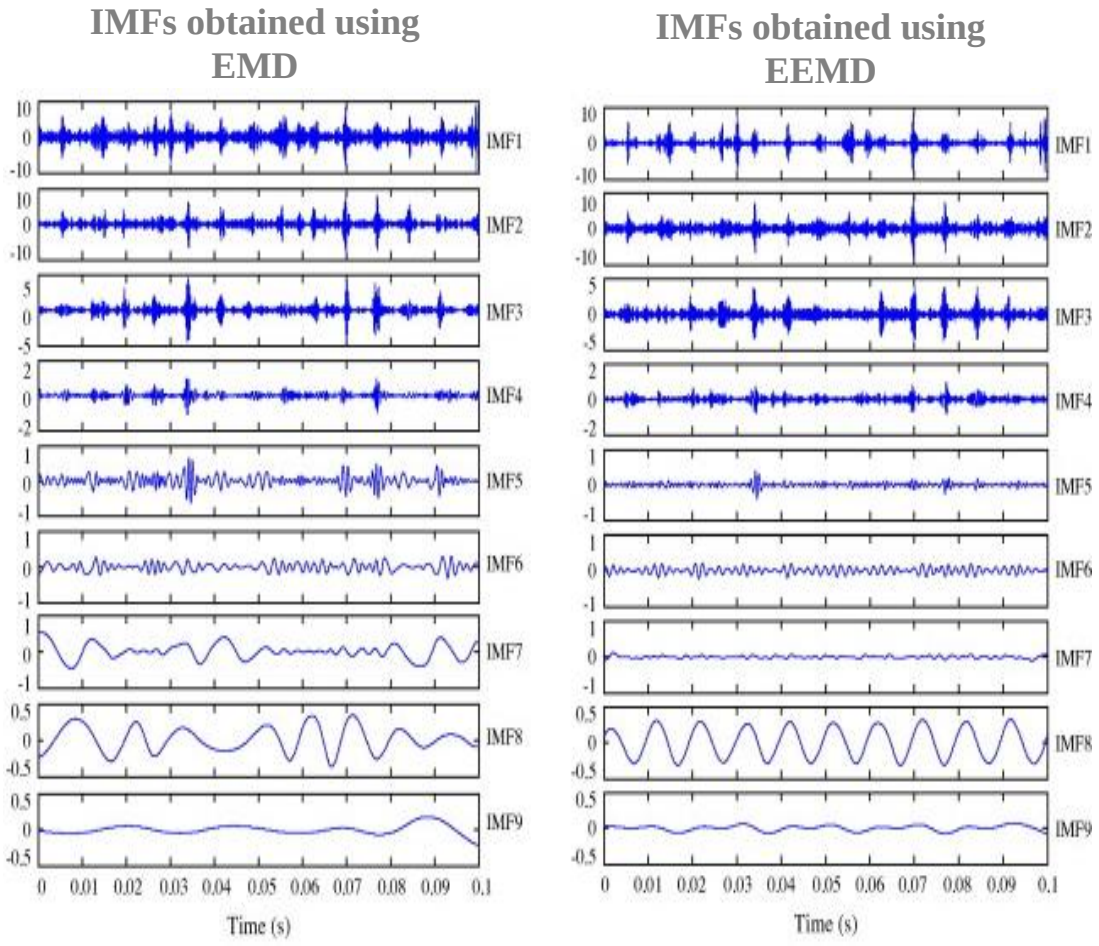

Figure 2.6: The first nine IMFs obtained using EMD (left) and EEMD (right) methods from [67].

2.2.3.3 Linear prediction

Linear prediction is basically a way of obtaining a model of the deterministic (periodic) part of a signal, based on a certain number of previous samples, and then using this model to predict the next value in the series [16]. The residual (unpredictable) part of the signal is then obtained by subtraction from the actual signal value.

The major problems with this method are (i) choosing the model, (ii) the model coefficients computing algorithm, and (iii) the model order selection criteria (number of taken data). Among the signal representation models, it is proven that the Autoregressive (AR) model has the most reliable de-noising results [14]. The AR model can be described by the following equation [16]

$$y(t) = -\sum_{k=1}^{p} a(k)x(t-k) \tag{2.24}$$

where the predicted current value $y(t)$ is obtained as a weighted sum of the p previous values. Here, the parameter, p, determine the order of the AR model. The actual current value is given by the sum of the predicted value and a noise term

$$x(t) = y(t) + e(t) \tag{2.25}$$

Altough the Burg method described in [15] is considered as the most preferred coefficient estimation technique which give better results [16], the Yule–Walker method offers simple, effective, and low prior requirements for its applications [17]. Consequently, the weighting coefficients, $a(k)$, in equation (2.24) can be obtained by Yule-Walker equations which can be represent in the matrix form as

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[-1] & \dots & r_{xx}[-p+1] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[-p+2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[p-1] & r_{xx}[p-2] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} a[1] \\ a[2] \\ \vdots \\ a[p] \end{bmatrix} = - \begin{bmatrix} r_{xx}[1] \\ r_{xx}[2] \\ \vdots \\ r_{xx}[p] \end{bmatrix} \tag{2.26}$$

where $r_{xx}(t)$ is the auto-correlation function of the time series $x(t)$, for which biased estimates can be obtained from

$$r_{xx}[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[t]x[t-k], \qquad 0 \le k \le p-1 \tag{2.27}$$

From ball bearing point of view, since a desired output signal should have high kurtosis value, a proper order, p, can be determined by some iterations between Equations (2.9)

and (2.24). The benefits of using an AR model arise from its simplicity, efficiency and the low prior requirements for its application [16].

2.3 Frequency Domain Analyses

Time representation of a vibration signal captured from the ball bearings shows how the corresponding signal changes over the time. However, it can not describe the constitutive components of the signal. Frequency domain can be achieved by using Fourier Transform. The advantage of frequency domain analysis over time domain analysis is its ability to easily identify and isolate a certain frequency band of interest [27]. Fast version of Fourier Transfor called Fast Fourier Transform (FFT), is mostly employed to obtain frequency spectrum of the vibration signal.

It is worth pointing out that, in the field of ball bearing fault detection, although the localized fault excite the resonances of the system, the resonance frequency is usually not of interest since it doesn't have explicit diagnostic information. Resonance frequency of the system (including the rotating shaft and the machine foundation) usually exceeds 1.5 kHz. In a vibration signal captured from faulty bearings, impulses are repeated with a certain time interval. Each local fault depending on its location and shaft rotating speed, produces impacts with a specific repetition frequency. The mentioned repetition frequencies are called fundamental fault frequencies and they can be predicted by some mathematical calculations.

2.3.1 Fundamental defect frequencies

There are five important motions in ball bearings; each produces a specific frequency in frequency domain. These frequencies are:

1) Shaft rotating frequency, f_r
2) Cage (train) frequency, F_c (FTF)
3) Ball passing frequency outer race, F_{BPO} (BPFO)
4) Ball passing frequency inner race, F_{BPI} (BPFI)
5) Ball spin frequency, F_b (BSF)

Ball bearing fundamental frequencies can be mathematically calculated from the bearing geometry and the rotating shaft speed. As can be seen in Figure 2.7, V_i, V_c,

and V_o are linear velocity of inner race, ball center, and outer race, respectively. Also D_b, D_p, and β are ball diameter, pitch diameter, and ball contact angle, respectively.

Shaft rotating frequency (f_r): The rotating shaft speed is very important in dynamics of the ball bearings. All other fundamental frequencies are dependent on f_r.

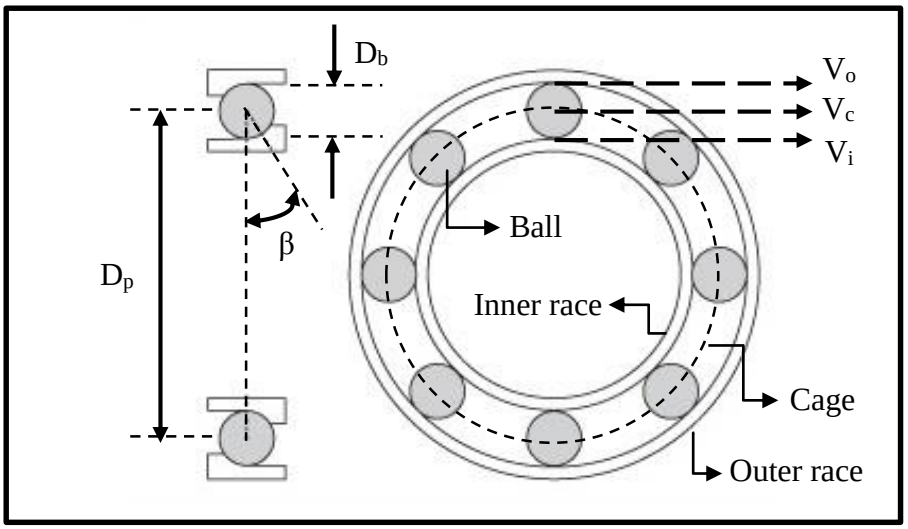

Figure 2.7: A simple ball bearing geometry.

Cage frequency (FTF): This frequency can be calculated from linear velocity, V_c, of the point on the cage. Linear velocity of the cage is the average of linear velocity of inner and outer races. Frequency of the cage can be calculated as follow.

$$\mathrm{FTF} = \frac{V_c}{r_c} = \frac{V_i + V_o}{D_p} \tag{2.28}$$

where r_p is $\frac{D_p}{2}$.

It can be expressed in terms of outer race spin frequency, F_o, inner race spin frequency, F_i, outer race radius, r_o, and inner race radius, r_i, as

$$\mathrm{FTF} = \frac{V_c}{r_p} = \frac{V_i + V_o}{D_p} = \frac{F_i r_i + F_o r_o}{D_p}$$

$$\mathrm{FTF} = \frac{1}{D_p}\left(F_i \frac{D_p - D_b \cos\beta}{2} + F_o \frac{D_p + D_b \cos\beta}{2}\right) \tag{2.29}$$

Ball passing frequency outer race (BPFO): Since this frequency illustrates ball passing frequency on a point on the outer race, it can be calculated by the product of number of balls, N_B, and difference of TFT and F_o as

$$BPFO = N_B|FTF - F_o| = N_B\left|\frac{F_i r_i + F_o r_o}{D_p} - F_o\right| \quad \textbf{(2.30)}$$

$$BPFO = N_B|TFT - F_o|$$
$$= N_B\left|\frac{F_i\left[r_p - \left({}^{(D_b\cos\beta)}/_{2}\right)\right] + F_o\left[r_p + \left({}^{(D_b\cos\beta)}/_{2}\right)\right]}{D_p} - F_o\right|$$

$$BPFO = \frac{N_B}{2}\left|(F_i - F_o)\left(1 - \left(\frac{D_b\cos\beta}{D_p}\right)\right)\right| \quad \textbf{(2.31)}$$

Ball passing frequency inner race (BPFI): Since this frequency illustrates ball passing frequency on a point on the inner race, it can be calculated by the product of number of balls, N_B, and difference of TFT and F_i as

$$BPFI = N_B|TFT - F_i| = N_B\left|\frac{F_i r_i + F_o r_o}{D_p} - F_i\right| \quad \textbf{(2.32)}$$

$$BPFO = N_B|TFT - F_i|$$
$$= N_B\left|\frac{F_i\left[r_p - \left({}^{(D_b\cos\beta)}/_{2}\right)\right] + F_o\left[r_p + \left({}^{(D_b\cos\beta)}/_{2}\right)\right]}{D_p} - F_i\right|$$

$$BPFI = \frac{N_B}{2}\left|(F_i - F_o)\left(1 + \left(\frac{D_b\cos\beta}{D_p}\right)\right)\right| \quad \textbf{(2.33)}$$

Ball spin frequency (BSF): BSF illustrates the spin of the ball about its axis and it can be calculated whether by BPFO or BPFI as

$$BSF = \left|(F_i - TFT)\frac{r_i}{r_b}\right| = \left|(F_o - TFT)\frac{r_o}{r_b}\right| \quad \textbf{(2.34)}$$

$$BSF = \frac{D_P}{2D_b}\left|(F_i - F_o)\left(1 - \left(\frac{D_b^2\cos\beta^2}{D_p^2}\right)\right)\right| \quad \textbf{(2.35)}$$

where r_b is $\frac{D_b}{2}$.

It should be noted that, when outer race is stationary and inner race is rotating with the shaft, equations (2.29), (2.31), (2.33), and (2.35) can be written, respectively, as

$$FTF = \frac{1}{2} f_r \left(1 - \frac{D_b \cos\beta}{D_p}\right) \text{ [Hz]} \quad \textbf{(2.36)}$$

$$BPFO = \frac{N_B}{2} f_r \left(1 - \frac{D_b \cos\beta}{D_p}\right) \text{ [Hz]} \quad \textbf{(2.37)}$$

$$BPFI = \frac{N_B}{2} f_r \left(1 + \frac{D_b \cos\beta}{D_p}\right) \text{ [Hz]} \quad \textbf{(2.38)}$$

$$BSF = \frac{D_p}{2D_b} f_r \left(1 - \frac{D_b^2 \cos\beta^2}{D_p^2}\right) \text{ [Hz]} \quad \textbf{(2.39)}$$

When there is a local fault on one of the raceways of a ball bearing, passing roller elements (balls) over the local fault creates pulses at one of the fundamental fault frequencies (BPFO or BPFI, depend on whether the fault is on outer or inner raceways) and the spectrum of the vibration signal contains this fault frequency and its harmonics. On the other hand, if there is a local fault on the rolling elements (balls), the local fault creates impacts on both inner and outer raceways. Therefore, a frequency close to 2×BSF (Ball Spin Frequency) appears in the spectrum. In many cases, this frequency is modulated with BPFO or BPFI. Consequently, a very complex frequency spectrum may result in from this process. In addition, a captured raw ball bearing vibration signal certainly contains some background noise due to environmental conditions and it frequently makes interpretations of the spectrum quite complex. Therefore, some further processes are strongly recommended in the literature in order to get a smoother spectrum which can be used to extract diagnostic information.

2.3.2 Envelope analysis

Envelope analysis is considered as a powerful technique for the detection of faults in rolling element bearing in frequency domain. The mentioned method has popularity compared to other signal processing based fault detection methods in the literature because of its simplicity and reliability [68]. Envelope analysis is for amplitude demodulating of a signal for extracting the periodic excitation. The output of the envelope analysis can be either the envelope signal or spectrum of the envelope signal. In many cases traditional amplitude spectrum (from now on such spectrums are referred to as amplitude spectrums) is unable to detect fault frequencies, whereas it is possible to distinguish impulsive patterns in ball bearing vibrations from other unwanted vibrations in envelope spectrum.

There are different ways to obtain the envelope signal; traditionally band pass filtering, rectifying, and low pass filtering can be used to carry out the demodulation. However, recent studies tend to use the Hilbert transform for the envelope detection [69].

As stated earlier, when a local fault occurs in ball bearings, it excites the resonance frequency of the system periodically. Consequently, impacts having resonance frequency are repeated with a certain repetition frequency. In other words, repetation frequency is modulated with the resonance frequency. So it makes sense to design a band-pass filter with central frequency equal to resonance frequency to pass the components having the resonance frequency and then demodulate the filtered signal to obtain its repetition frequency. This method is called envelope analysis or High Frequency Resonance Technique (HFRT) [70]. Figure 2.8 illustrates the envelope signals for different type of local faults on ball bearings.

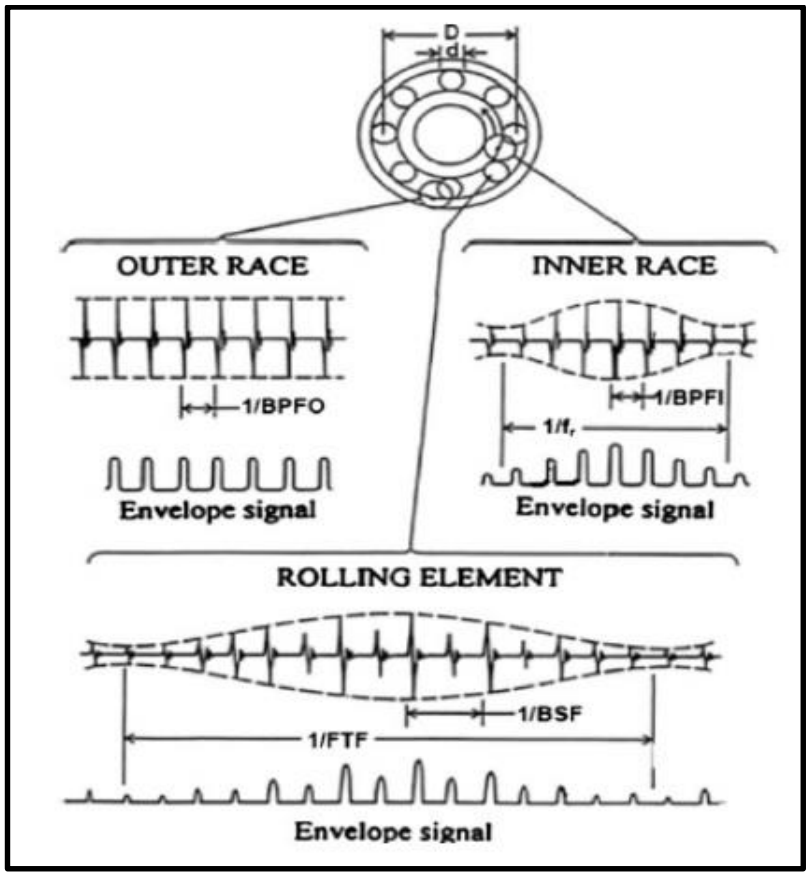

Figure 2.8: Illustration of Envelope Analysis from [71].

Envelope analysis can be summarized in four steps as [72]

1) Obtain the amplitude spectrum of the ball bearing signal for identification of the resonance frequency (carrier frequency)

2) Select the zone to employ the band-pass filtering around the center frequency.

3) Calculate the envelope of the filtered signal by Hilbert transform.

4) Calculate the spectrum of the envelope signal.

Of course, determining the resonance frequency of the ball bearing system is the most important step in envelope analysis. Usually the resonance frequency depends on geometry of the ball bearing varies between 1.5 kHz and 10 kHz [29].

2.3.2.1 Band-pass filtering

One of the main advantages of the frequency domain analyses is the ability of filtering out the unwanted components to reveal diagnosis features of the signal. The most common filter used in ball bearing fault detection is the band-pass filter. The response function of a typical band-pass is shown in Figure 2.9. The mentioned filter has two cut-off frequencies which block both frequencies lower than f_{c1} and higher than f_{c2}. A central frequency, f_0, is in middle of the two cut-off frequencies.

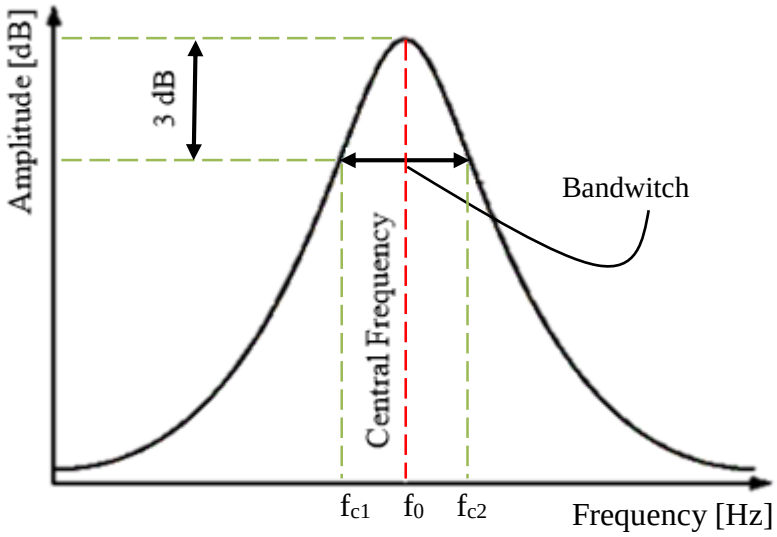

Figure 2.9: Band pass filter structure.

Since traditional amplitude spectrum in many cases can be very complicated, identification of resonance frequency may not be always feasible. For this purpose, taking FFT over a simplified short-time signal as a representative of the whole signal can overcome this problem in order to have a smoother amplitude spectrum. F. Cong *et al.* in [73] have proposed a method based on SVD to obtain a simplified short-time signal using the captured vibration signal to optimize the band-pass filter features. The more detailed information is available in [73].

It should be noted that, it is desired to use an optimum and advanced filtering method instead of a traditional band-pass filtering. A powerfull and effective filter which is succefully applied for the ball bearing fault detection is so-called Wiener filtering [74, 75, 76, 34] (see section 2.3.4). A digital Wiener filter is an optimum digital filter in the sense of producing the best estimate (depend on the application) of the signal in the presence of noise. This is compared to the classical filters such as a low-pass, high-pass, and band-pass filters, in which it is difficult to recover the signal that is corrupted and distorted by noise.

In addition, other de-noising algorithms such as wavelet-based de-noising, SVD based de-noising, MED, etc. can be also employed to achieve a desired impulsive vibration signal which contains diagnosis information.

2.3.2.2 Hilbert transform (HT)

Hilbert transform is a time domain convolution and can be defined by [77]

$$H[x(t)] = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{x(t)}{t-\tau} d\tau \qquad \textbf{(2.40)}$$

where t, $x(t)$, and $H[x(t)]$ are time, time domain signal, and the Hilbert transform of $x(t)$, respectively. According to [1], the demodulated signal can be obtained by the magnitude of the corresponding analytical signal. Hilbert Transform produces an orthogonal signal to the original input signal, $x(t)$. Consequently, the complex-valued time domain analytic signal, $A[x(t)]$ can be defined as

$$A[x(t)] = x(t) + iH[x(t)] = a(t)e^{i\emptyset(t)} \qquad \textbf{(2.41)}$$

where $\emptyset$ is phase angle.

The resulting analytic signal can be converted from the real/imaginary format to the magnitude/phase format as

$$a(t) = \sqrt{x^2(t) + H^2[x(t)]} \qquad \textbf{(2.42)}$$

$$\emptyset(t) = \arctan\frac{H[x(t)]}{x(t)} \qquad \textbf{(2.43)}$$

where $a(t)$ is the envelope signal. It is worth mentioning that, instead of using the integration in Equation (2.40) it is much simpler to obtain amplitude spectrum of the signal, multiply the positive frequency components by 2 (but not the zero frequency

or Nyquist frequency), set the negative part of the spectrum to zero, and finaly employ the IFFT (Inverse Fourier Transform) to get the analytic time signal [1].

2.3.3 Minimum entropy deconvolution (MED)

The efficiency of the envelope analysis and most of the other fault detection methods depend on the ability to identify series of impacts due to local faults. Nevertheless, identifying of impact train is not always possible. For example, when ball bearings rotate at very high speeds (usually more than 10,000 RPM [1]), when damping is low (resonance frequency is very close to impact repetition frequency), envelope analysis may be unable to find diagnosis features.

N. Sawalhi *et al*. in [32] show that the Minimum Entropy Deconvolution (MED), which is originally proposed by R.A. Wiggins in [31] can reduce the so-called smearing effect, which is a common issue in high speed machines. A desired output signal from any filtering process is assumed to be impulsive, and thus having high kurtosis value. So it makes sense to create an optimal filter with having an output signal as impulsiveness as possible. MED process uses the physical meaning of well-known thermodynamic term, entropy, to design the corresponding optimal filter. The following paragraphs explain the MED filtering process.

By increasing entropy of the signal, disorder value increases. Impulsive part of signals (with high kurtosis value) are very structured, whereas random components (with low kurtosis value) have maximum disorder (entropy). This explains the name minimum entropy method, because higher entropy corresponds to a tendency with Gaussian behaviour (i.e. less structured and more disordered). Therefore, minimizing the entropy of the output signal in the filtering process enhances the structured part in the signal and consequently increases the kurtosis value [17].

The optimal output filter can be obtained by Objective Function Method (OFM). The OFM is applied to design of the MED filter by Lee [78]. The aim is to find an optimal Finite Impulse Response (FIR) filter, $\mathrm{f}(l)$, to let the filtered discrete signal, $\mathrm{y(k)}$, much closer to the original discrete signal, $\mathrm{x(k)}$. A general expression for inverse filtering is given as [79]

$$\mathrm{y(k)} = \sum_{l=1}^{\mathrm{L}} \mathrm{f}(l)\mathrm{x}(\mathrm{k}-l) \qquad \textbf{(2.44)}$$

where L which is the length of the optimal FIR filter.

According to [32], the OFM is shown as the following

$$O_4\big(f(l)\big) = \sum_{k=1}^{N} y^4(k) \Big/ \left[\sum_{k=1}^{N} y^2(k)\right]^2 \qquad \textbf{(2.45)}$$

By combination of Equations (2.44) and (2.45) the following equation can be obtained

$$\underbrace{\left[\sum_{k=1}^{N} y^2(k) \Big/ \sum_{k=1}^{N} y^4(k)\right] \sum_{k=1}^{N} y^3(k)\, x(k-l)}_{b} \qquad \textbf{(2.46)}$$

$$= \underbrace{\sum_{P=1}^{L} f(p)}_{f} \underbrace{\sum_{k=1}^{N} x(k-l)x(k-p)}_{A}$$

Above equation can be written in matrix form as

$$\{b\} = [A]\{f\} \qquad \textbf{(2.47)}$$

The column vector b is computed from the cross correlation of input and output signals (y and x) of the FIR inverse filter [32]; A is the Toeplitz auto-correlation matrix of input signal, x(k), with dimension L×L and f is the column vector of the FIR inverse filter coefficients.

The solution for the optimum inverse filter needs some iterations between Equations (2.44) and (2.46). Iterative steps can be summarized as:

> *Step* 1: Compute the matrix [A] based on the original vibration signal, x(k), and initialize the optimal FIR filter by assuming a zero L×1 vector. (Note that, the matrix [A] is computed only once at the beginning of the iteration and is reused in the iterative loop)
>
> *Step* 2: Compute the output signal, y(k), using the input signal, x(k), and the initialized optimum filter, according to the Equation (2.44).
>
> *Step* 3: Compute the vector b via Equation (2.46).
>
> *Step* 4: Solve for the new optimal FIR filter coefficients via Equation (2.46).

Note that the mentioned iteration process can be repeated until the output signal have a proper kurtosis value. However, according to work by H. Endo and R.B. Randall in [17], although having a greater number of iterations can increase the kurtosis value of the output signal, the result did not gain any benefit for the purpose of fault diagnosis.

Figure 2.10 illustrates the kurtosis value for the output signals at each iteration step of the MED algorithm having different filter length, L (i.e. 250, 500, and 1000 samples). Altough the kurtosis reaches the highest value at the fifth iteration, but two iterations are sufficient to reveal the fault features of the signal for that corresponding case study.

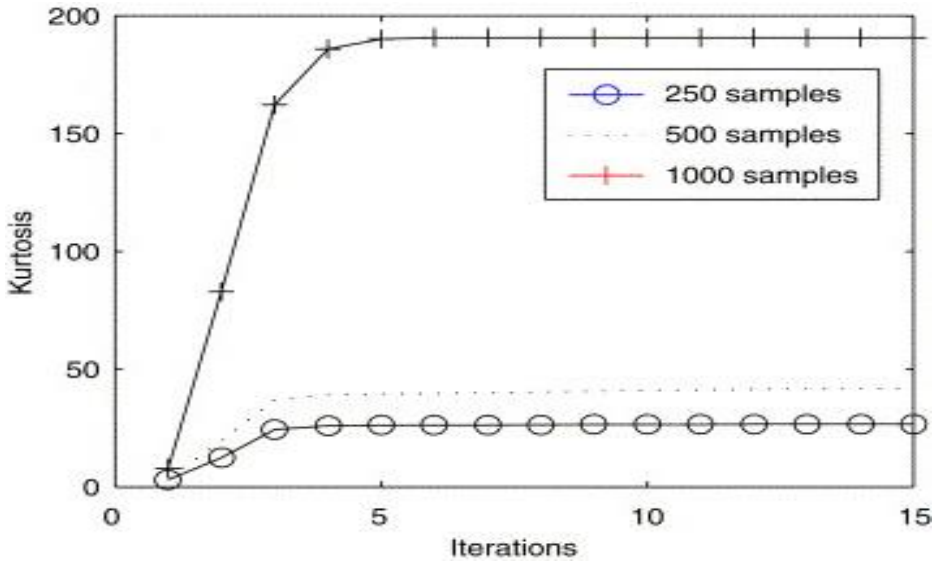

Figure 2.10: Convergence the kurtosis of the output signal in MED algorithm for gearbox fault detection from [17].

In some studies [17, 32], MED have been combined with AR linear prediction filtering (see section 2.2.3.3) to achieve excellent enhancement in diagnosis of the fault frequencies in the frequency domain analyses. The total process is so-called ARMED filtering. In addition, the so-called EEMED method proposed by R. Ricci and P. Pennacchi in [80], is defined as the combination of MED method and EMD algorithm (see section 2.2.3.2).

2.3.4 Wiener Filtering

Wiener is a filtering algorithm designed to obtain an estimation of the desired output signal [74]. A desired output signal can be different from an application to another. A Wiener filter can be an Infinite Impulse Response (IIR) filter (results in a set of non-linear equations) or a Finite Impulse Response (FIR) filter (results in a set of linear equations) [81]. The Wiener filter minimizes the Mean Square Error (MSE) between the estimated output signal and the desired output signal in filtering process. The traditional filter input and output relation for a discrete input signal, x(k), is given by

$$\bar{y}(k) = \sum_{l=1}^{p} w(l)x(k-l) \tag{2.48}$$

or in other form as

$$\bar{y}(k) = w^T x \tag{2.49}$$

where $\bar{y}(k)$, $w(l)$, and $x(k)$ are the desired output signal with minimum MSE, Wiener filter coefficients vector, and the input signal, respectively. In what follows, the corresponding procedure for obtaining the Wiener filter coefficients are shown.

The difference between the desired signal, $\bar{y}(k)$, and the filter output signal, $y(k)$, in Equation (2.48) can be written as [74]

$$\begin{aligned} e(k) &= y(k) - \bar{y}(k) \\ e(k) &= y(k) - w^T x \end{aligned} \tag{2.50}$$

As mentioned, Wiener filtering algorithm uses Mean Square Error (MSE) criterion. By minimizing the mean square of the difference function defined in Equation (2.50) with respect to the filter coefficient vector w, the desired output signal with the possible minimum error may be achieved. From Equations (2.50), the mean square estimation error is given by

$$\begin{aligned} E[e^2(k)] &= E[(y(k) - w^T x)^2] \\ &= E[y^2(k)] - 2w^T E[x(k)y(k)] + w^T E[x(k)x^T(k)]w \end{aligned} \tag{2.51}$$

As a result,

$$E[e^2(k)] = r_{yy}(0) - 2w^T r_{xy} + w^T R_{xx} w \tag{2.52}$$

where $R_{xx} = E[x(k)x^T(k)]$ is the auto-correlation matrix of the input signal and $r_{xy} = E[x(k)y(k)]$ is cross correlation vector of the input and the desired signals.

To minimize the MSE, the gradient of the function in Equation (2.51) with respect to the filter coefficient vector, w, must be set to zero as

$$\frac{\partial}{\partial w} E[e^2(k)] = 0 \tag{2.53}$$

Hence the following

$$\frac{\partial}{\partial w} E[e^2(k)] = -2E[y(k)x(k)] + 2w^T E[x(k)x^T(k)]w = -2r_{xy} + 2w^T R_{xx} = 0$$

The desired Wiener coefficients vector can be obtained as

$$w = R_{xx}^{-1} r_{xy} \tag{2.54}$$

In noisy captured ball bearing vibration signals, the background noise can be assumed as Gaussian white noise. Therefore, a typical ball bearing vibration signal, x(k), can be modelled as

$$x(k) = y(k) + n(k) \tag{2.55}$$

where y(k) and n(k) are faulty vibration signal and white noise, respectively.

Assuming that the signal and the noise are uncorrelated, it follows that the auto-correlation matrix of the noisy signal is the sum of the auto-correlation matrix of the signal x(k) and the noise n(k) as [81]

$$R_{xx} = R_{yy} + R_{nn} \tag{2.56}$$

and also

$$r_{xy} = r_{yy} \tag{2.57}$$

where R_{xx}, R_{yy} and R_{nn} are the auto-correlation matrices of the noisy signal, the de-noised signal and the noise, respectively, and r_{xy} is the cross correlation vector of the noisy signal and the noise free signal. By substitution of Equations (2.56) and (2.57) into the Wiener filter obtained in Equation (2.54), the following relation can be achieved

$$w = \frac{r_{yy}}{R_{yy} + R_{nn}} \tag{2.58}$$

Equation (2.58) is the optimal linear filter for the removal of additive white noise. This can also be written in term of power spectra as

$$w = \frac{P_{yy}(f)}{P_{yy}(f) + P_{nn}(f)} \tag{2.59}$$

where $P_{yy}(f)$ and $P_{NN}(f)$ are the noiseless signal and noise power spectra, respectively.

Figure 2.11 illustrates the principle of the Wiener filtering for a synthetic signal corrupted by gaussian white noise.

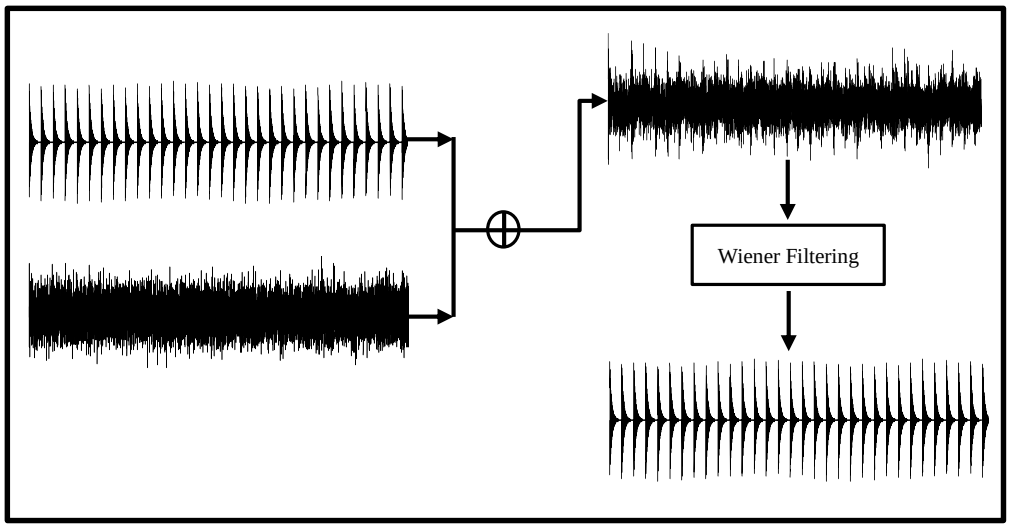

Figure 2.11: Principle of the Wiener filter.

Equation (2.59) can be also written in term of SNR (Signal-to-Noise ratio). By dividing the numerator and the denominator of Equation (2.59) by the noise power spectra $P_{NN}(f)$ and substituting the variable $SNR = {P_{xx}(f)}/{P_{NN}(f)}$ yields [81]

$$w(f) = \frac{SNR(f)}{SNR(f) + 1} \tag{2.60}$$

or

$$w(f) = \frac{1}{1 + \rho(f)} \tag{2.61}$$

where $\rho(f)$ is the Noise-to-Signal ratio. Note that the variable, $SNR(f)$ is expressed in terms of the power spectral ratio, and not in the more usual terms of logarithmic power ratio. Therefore, $SNR(f) = 0$ corresponds to $-\infty$ dB.

For the captured vibration signals with unknown noise level, an appropriate estimation of the noise (depend on environmental conditions) can be used to characterize the Wiener filter coefficients.

2.3.5 Spectral kurtosis and Kurtogram

As stated in section 2.2.2, vibration signals can be characterized in terms of some statistical indicators (e.g. kurtosis). Kurtosis value takes high value for impulsive signals which shows the presence of fault in the corresponding ball bearing, whereas it is low value for healthy vibration signals. In practice, however, the impulses can be covered by background noise. As a result, the kurtosis is unable to capture the

peakiness of the faulty vibration signal and hardly departs from the zero value [34]. In this situations, the kurtosis as a global indicator is not appropriate, and it is preferable to apply it locally in different frequency bands. This is exactly what the Spectral Kurtosis (SK) does. R. Dwyer in [82] originally proposed spectral kurtosis for stationary signals as the normalized fourth order moment of the real part of the Fourier transform. Recently, J. Antoni in [33] proposed the formalization of SK for both stationary and non-stationary signals as a strong signal processing tool to deal with non-stationary signals.

The estimation of the SK has been discussed in depth in [33]; therefore, only some main formulations of SK is stated here. For the non-stationary signal, $y(t)$, in Equation (2.55), the Word-Cramér's decomposition is as [33]

$$y(t) = \int_{-\infty}^{\infty} H(t,f)e^{j2\pi ft}dX(f) \qquad \textbf{(2.62)}$$

where $H(t,f)$ is the Fourier transform of time-varying impulse response, $h(t,f)$; $dX(f)$ is the spectral process associated with the white signal $X(t)$ which has a symmetric probability density function. In other words, the time-varying transfer function, $H(t,f)$, can also be interpreted as the complex envelope of $y(t)$ at frequency f.

Let $C_{4y}(f)$ and $S_{2ky}(f)$ be the fourth order spectral cumulants and 2k order spectral components of non-stationary signal, $x(t)$, respectively. Then, the spectral kurtosis of $y(t)$ can be defined as

$$SK_y(f) = \frac{C_{4y}(f)}{S_{2y}^2(f)} = \frac{S_{4y}(f) - 2S_{2y}^2(f)}{S_{2y}^2(f)} = \frac{S_{4y}(f)}{S_{2y}^2(f)} - 2 \qquad \textbf{(2.63)}$$

where

$$S_{2ky}(f) = E\{|H(t,f)dX(f)|^{2k}\}/f = E\{|H(t,f)|^{2k}\}.S_{2ky} \qquad \textbf{(2.64)}$$

when k=1, it gives the classical power spectral density of $y(t)$. The energy normalized fourth order spectral cumulant gives a measure of the peakiness of the probability density function of the signal [83]. J. Antoni in [33] shown that for a non-stationary signal, $y(t)$, and its Short Time Fourier Transfrom (see section 2.4.1) as $STFT(f,n)$, the SK estimator is defined as follows

$$SK_y(f) = \frac{\langle |STFT(f, np)|^4 \rangle_n}{\langle |STFT(f, np)|^4 \rangle_n^2} - 2, \quad |f - mod(1/2)| > 1/N \qquad \textbf{(2.65)}$$

where $STFT(f, k) = \sum_{k=-\infty}^{\infty} y(k)w(k - nP)e^{j2\pi kf}$, $w(k)$ is the analysis window with the length of N_w, p is a temporal step, and $\langle . , . \rangle$ stands for the time-average operator over index n. In addition, the optimal step, p, should satisfy the following condition

$$p \leq N/4$$

Figure 2.12 from [84] shows the STFT based SK for a non-stationary signal.

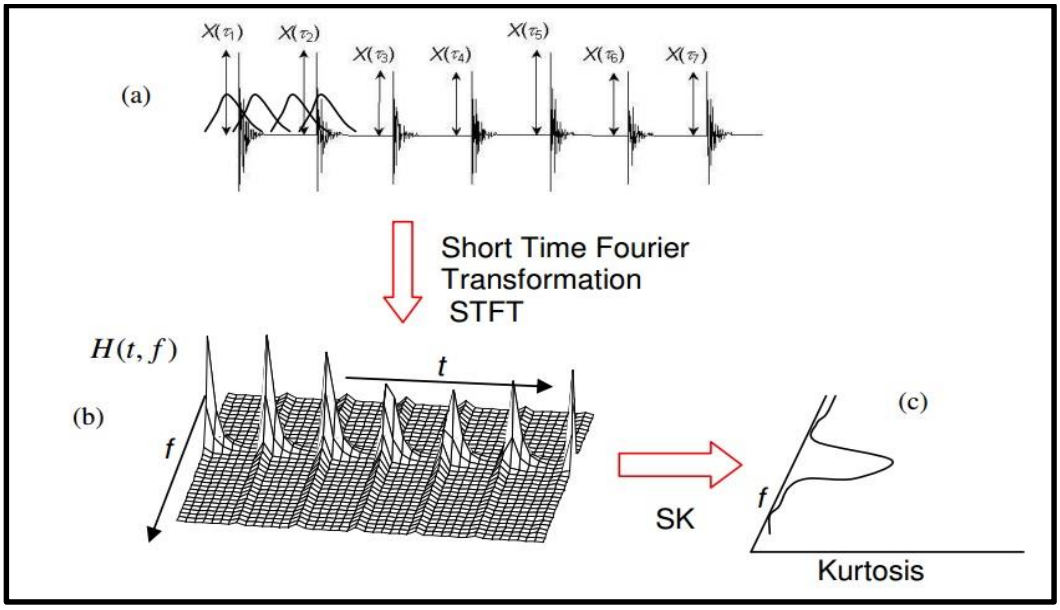

Figure 2.12: STFT based spectral kurtosis from [84].

J. Antoni in [33] has shown that the SK of the captured vibration signal, x(t), in Equation (2.55), can be obtained as

$$SK_x(f) = \frac{SK_y(f)}{[1 + \rho(f)]^2} \qquad \textbf{(2.66)}$$

where $\rho(f)$ is the Noise-to-Signal ratio. From Equation (2.71), it can be seen that the value of $SK_x(f)$ is close to $SK_y(f)$ at frequencies with low ρ value (high SNR), where the signal contains diagnois information. On the other hand, if the ρ is very high, the $SK_x(f)$ approaches to zero. Therefore, Equation (2.71) takes very important and useful property which can be employed on the optimum band-pass filter design.

By comparing Equations (2.61) and (2.66), it can be easily concluded that the optimum Wiener filter is proportional to to the square root of SK as

$$SK_x(f) = k.w(f)^2 \quad \textbf{(2.67)}$$

where k is the so-called scaling factor. This is exactly what J. Antoni and R. B. Randall in [34] have proven and employed on the ball bearing fault detection.

Figure 2.13 from [34] represents a noisy vibration signal captured from an outer race defective ball bearing and its de-noised version by optimum Wiener filtering. According to Figure 2.13, kurtosis value takes the highest value at 23 kHz where the impacts excite the resonances.

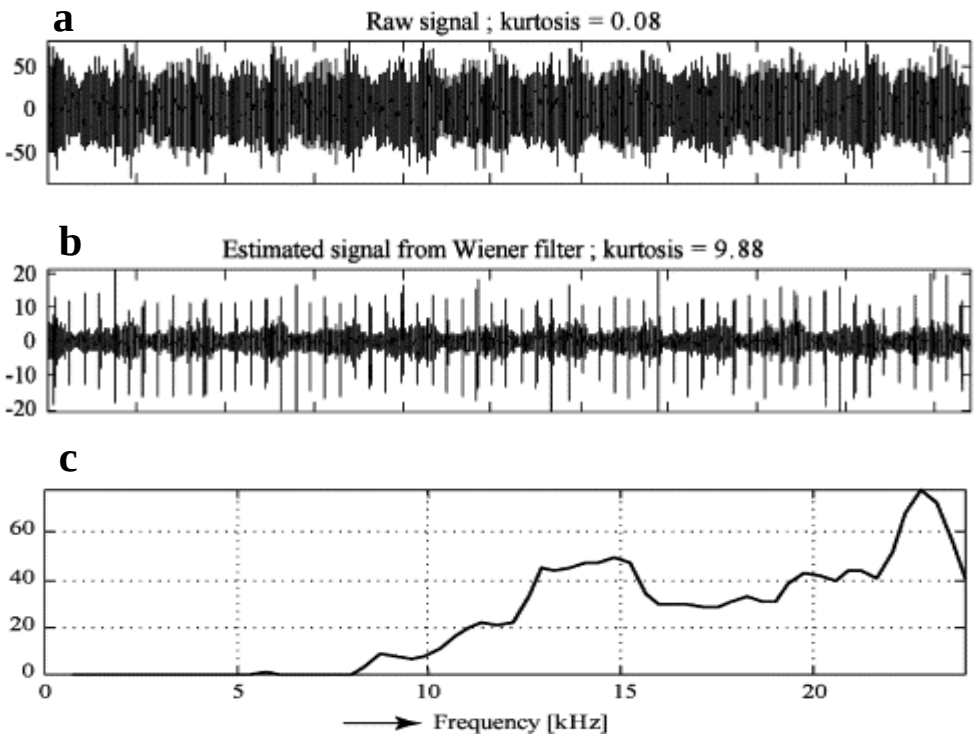

Figure 2.13: a) Noisy ball bearing vibration signal, b) De-noised vibration signal by optimum Wiener filter, c) Spectral kurtosis from [34].

Along with many benefits of the SK, it suffers from the lack of optimum choice of windows length, N_w, used in STFT process which can adversely affect the de-noising properties of the optimum Wiener filter. The windowing process in SK can lead to the concept of a (frequency/ frequency resolution) dyad. The question is then how to select the best frequency resolution of the SK of a given vibration signal, at a given frequency. J. Antoni and R. B. Randall in [34] responded to this question by introducing a novel concept derived from kurtosis, namely kurtogram which is a map formed using the SK as a function of frequency and window length, N_w. The idea behind the kurtogram is to find an optimal combination of a frequency, f, and a

frequency resolution, Δf which maximizes the SK. The kurtogram is obtained based on SK and can be considered as the generalized version of the SK. Figure 2.14 from [36] shows the kurtogram for a vibration signal captured from a defective ball bearing. The corresponding vibration signal has the carrier frequency−which is actually the resonance of the system−at 6750 Hz. According to this plot, the maximum kurtosis (K=12) can be obtained at 6750 Hz with a 1460 Hz bandwith. Consequently, with these two parameters, the optimum Wiener filter can be designed to filter out unwanted components.

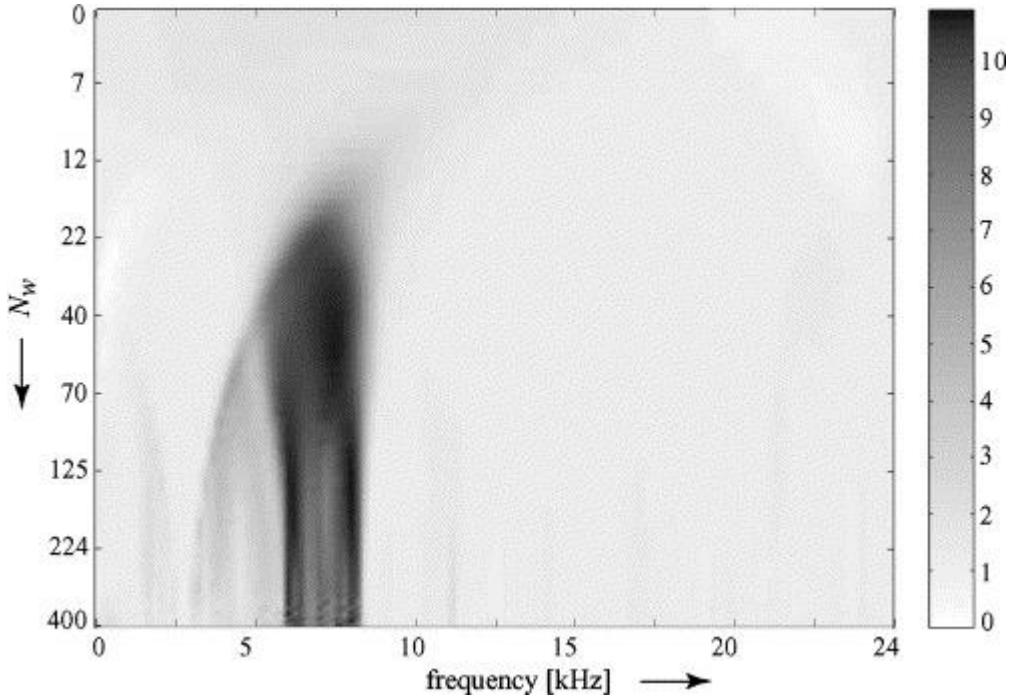

Figure 2.14: Kurtogram based on SK as a function of window length from [36].

The kurtogram still have a quite complex implementions and needs much CPU time for calculating. J. Antoni in [36] proposed a fast algorithm for kurtogram very similar to what S. G. Mallat did in [44] (see section 2.3.6.1). The fast implemention of the kurtogram proposed by J. Antoni is based on a series of band-pass filters (filter bank). By passing through the vibration signal, in a filter bank that has a low-pass and high-pass filters at each level of decomposition, the SK of the sub-signals at each level can be obtained. The details of this method is available in [36]. Figure 2.15 shows the fast version of the kurtogram in Figure 2.14. The maximum detected kurtosis is about 11 and the optimum central frequency and the bandwidth are 6750 Hz and 1500 Hz, respectively, which are very close to the parameters obtained in original kurtogram.

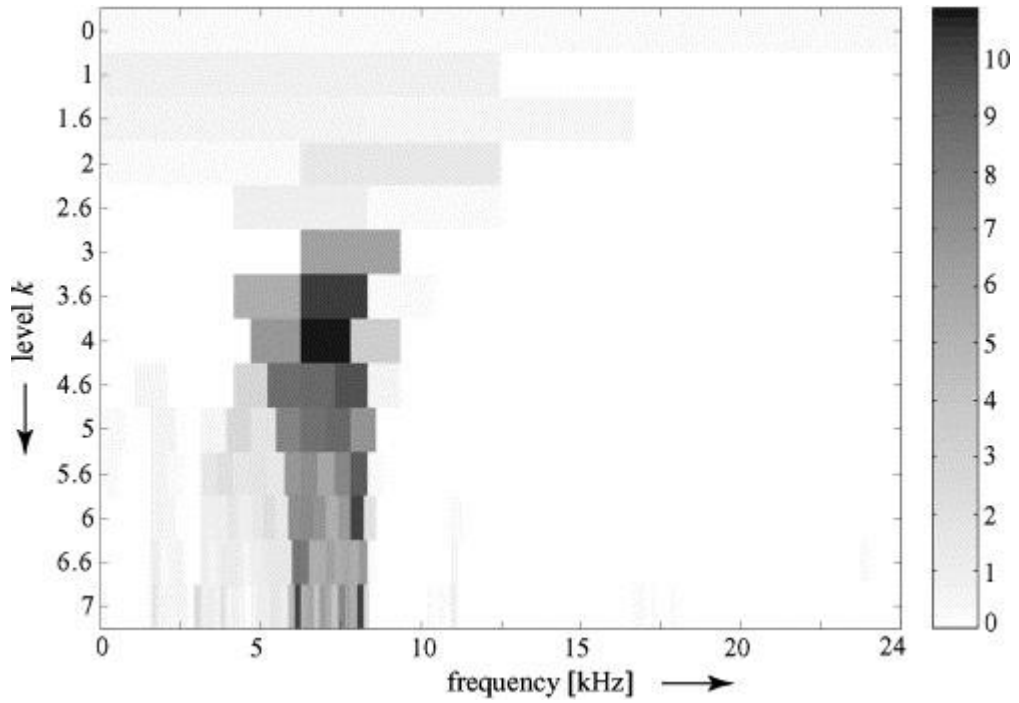

Figure 2.15: Fast version of the kurtogram in Figure 2.14 from [36].

2.3.6 Discrete Wavelet transform

Wavelet Transform (WT) is a method to decompose the input time singal into its wavelets components. WT analyzes the signal in different frequency bands so that any transient event can be observed in the time domain as well as in the frequency domain. The mentioned ability leads to introducing the concept so-called Multi-Resolution Analysis (MRA) in the field of signal processing. WT makes an effective tool for signal processing which can be employed in time-frequency domain (see section 2.4.2) in its continuous form or in separate components in its discrete form. The decomposition process is done in terms of a family of wavelets which can be shifted and dilated in time.

It should be emphasized that, the present study focuses on the application of Wavelet Transform in the field of ball bearing fault detection so that only some key concepts of Wavelet Transform are stated here. More rigorous mathematical treatment of the mentioned subject can be found in the signal processing and Wavelet Transform theory text books.

Mathematically, Continuous Wavelet Transform (CWT) of the vibration signal, x(t), can be obtained as [42]

$$CWT_{\Psi}(a,b) = \int_{-\infty}^{+\infty} x(t)\Psi_{a,b}{}^{*}(t)dt \qquad \textbf{(2.68)}$$

where $CWT_{\Psi}(a,b)$ are wavelet coefficients and $\Psi(t)$ denotes the mother wavelet and * stands for the complex conjugate. Wavelets are generated by scaling using the parameter a, and by shifting using the parameter b, from the single mother wavelet which can be defined as

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \qquad \textbf{(2.69)}$$

The parameter a represents the scale index which is a positive integer and can be considered as a reciprocal of frequency (pseudo-frequency) and the parameter b indicates the time shifting (translation).

A wide rang of mother wavelets such as Morlet, Meyer, Mexican hat, and Haar - each has a specific formulation - can be found in the literature. Mother wavelets can be classified as orthogonal or non-orthogonal [46]. The family of Daubechies wavelets (db1, db2, …) and the Gaussian or Morlet wavelet (morl) are the most known orthogonal and non-orthogonal wavelets, respectively.

Since parameters a and b are continuous over R (real numbers), CWT is a very redundant representation of the time signal and prodecues so many coefficiants to describe the properties of the signal in time-scale domain. As a result, redundancy property of CWT may make it impracticable and not suitable for some applications which seek a description with as few components as possible. To overcome this issue, Discrete Wavelet Transform (DWT) have been introduced. The main idea of DWT is the same as CWT, but it is considerably easier and faster to implement. In DWT, the scale and shift parameters (a and b) are evaluated on a discrete grid of time-scale plane leading to a discrete set of the continuous basis functions. The discretization is performed by setting $a = 2^{j}$ and $b = k2^{j}$ in equation (2.69), then below calculation can be done

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) = \frac{1}{\sqrt{2^{j}}}\Psi\left(\frac{t-k2^{j}}{2^{j}}\right) = 2^{-\frac{j}{2}}\Psi\left(2^{-j}t-k\right)$$

As a result, Equation (2.69) can be written in discrete form as follow,

$$\Psi_{j,k}(t) = 2^{-\frac{j}{2}}\Psi\left(2^{-j}t-k\right) \qquad \textbf{(2.70)}$$

From equations (2.68) and (2.70), the DWT of the vibration signal, x(t), is derived as [85]

$$\mathrm{DWT}_{\Psi}(j,k) = 2^{-\frac{j}{2}} \int_{-\infty}^{+\infty} x(t)\Psi^{*}(2^{-j}t - k)dt \qquad \textbf{(2.71)}$$

where $\mathrm{DWT}_{\Psi}(j,k)$ is called wavelet coefficients.

The most important property of the DWT is that the original signal can be reconstructed by a suitable integration over all the resulting components. Since in the field of signal de-noising based on Wavelet Transform, a perfect reconstruction (means unique reconstructed signal) is required, orthogonal wavelets are used in DWT. Orthogonality properties in orthogonal wavelets make the de-noised signal (output) as close as possible to the original input raw signal. The following sub-sections in Wavelet Transform part are dedicated to the various forms and implementations of the DWT.

2.3.6.1 Fast wavelet transform (FWT)

According to Equation (2.71), WT suffers from complexity of calculations. S. G. Mallat in [44] proposed an efficient way to implement DWT using a filter bank (a fast algorithm so-called Fast Wavelet Transform (FWT) and also known as Mallat algorithm) for orthogonal wavelets. According to studies by S. G. Mallat in [44], DWT can be implemented by means of a pair of low-pass and high-pass wavelet filters, denoted as h(k) and g(k), respectively. These wavelet filters are also known as the Quadrature Mirror Filters (QMF). For this purpose, S. G. Mallat proposed an orthogonal function to wavelet function called scaling function, φ(t), as [44]

$$\varphi(t) = 2^{-\frac{j}{2}}\varphi\left(2^{-j}t - k\right) \qquad \textbf{(2.72)}$$

The mentioned low-pass and high-pass filters can be derived via the following inner products, respectively, as [86]

$$h(k) = \frac{1}{\sqrt{2}}\langle \varphi(t)\,, \varphi(2t-k)\rangle \qquad \textbf{(2.73)}$$

$$g(k) = \frac{1}{\sqrt{2}}\langle \Psi(t)\,, \Psi(2t-k)\rangle \qquad \textbf{(2.74)}$$

where $g(k) = (-1)^{k}h(1-k)$ and $\langle \cdot\,, \cdot \rangle$ holds for the inner product operation.

The mentioned filters can be employed recursively to obtain wavelet coefficients of the vibration signal. The mentioned recursive algorithm is initialized by the discrete vibration signal, x(k), (i.e. $x(k) = cA_k^0$) and can be expressed as [87]

$$cA_k^j = \sum_n h(2k-n)cA_n^{j-1} \tag{2.75}$$

$$cD_k^j = \sum_n g(2k-n)cA_n^{j-1} \tag{2.76}$$

where cA_k^j and cD_k^j are two vectors containing approximation and detail coefficient elements, respectively. The elements of cA_k^j contains low frequency components (give the information about the overall shape of the signal) and the elements of cD_k^j contains high frequency components (describe the details in overall shape). It should be noted that, the wavelet coefficients, $DWT_\Psi(j,k)$ in Equation (2.71) are splitted into so-called approximation and detail parts in Mallat's algorithm in Equations (2.75) and (2.76).

Figure 2.16 illustrates the basic step (j=1) of the wavelet decomposition scheme. In the decomposition process, the discrete signal, x(k), is convolved with a low-pass filter and a high-pass filter, resulting in two vectors, cA^1 and cD^1, respectively. Here "↓2" illustrates down sampling (i.e. omitting the odd indexed elements of the filtered signal) so that the number of the total coefficients produced by the decomposition process is approximately the same as the number of elements of the discrete signal, x(k).

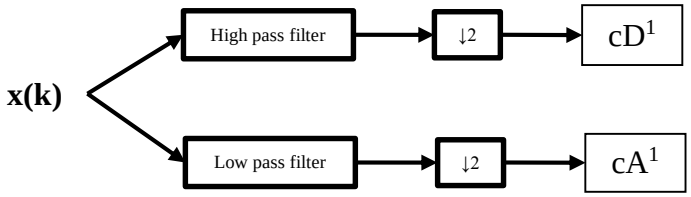

Figure 2.16: Basic step of decomposition process in the Wavelet Transform.

The reconstruction process for obtaining the reconstructed signal can be implemented by using Inverse Discrete Wavelet Transform (IDWT), given by [85]

$$x(t) = \sum_k \sum_j DWT_\Psi(j,k)\, \Psi_{j,k}(t) \tag{2.77}$$

The above reconstruction equation can be also implemented by the similar algorithm in FWT but in reverse. Note that, the low-pass and high-pass filters in the decomposition and the reconstruction processes are different.

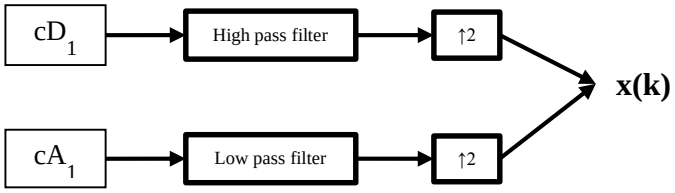

Figure 2.17: Basic step of reconstruction process in the Wavelet Transform.

Figure 2.17 illustrates the basic step of the reconstruction process. In the reconstruction process, a pair of low-pass and high-pass reconstruction filters are convolved with the vectors cA^1 and cD^1, respectively. For a single level decomposition, two signals are produced in the reconstruction process namely signal A^1 (approximation signal) and signal D^1 (detail signal). The symbol ↑2 denotes upsampling (i.e. inserting zeros between the elements of the vector cA^1 and cD^1). An important property of this step is: $x(k) = A^1 + D^1$. Figure 2.18 represents an example of waterfall decomposition for j = 3. As seen, the decomposition process is performed on output signals in low-pass filters at each step. As mentioned, the downsampling at each step is a technique to prevent the redundancy in the decomposition process. In this case, the reconstruction process can be applied to the corresponding coefficient vectors at any level up to third level. By applying the reconstruction process to the coefficient vectors – cA^3, cD^3, cD^2, and cD^1 –, the sub-signals– A^3, D^3, D^2, and D^1 –are obtained, where $x(k) = A^3 + D^3 + D^2 + D^1$.

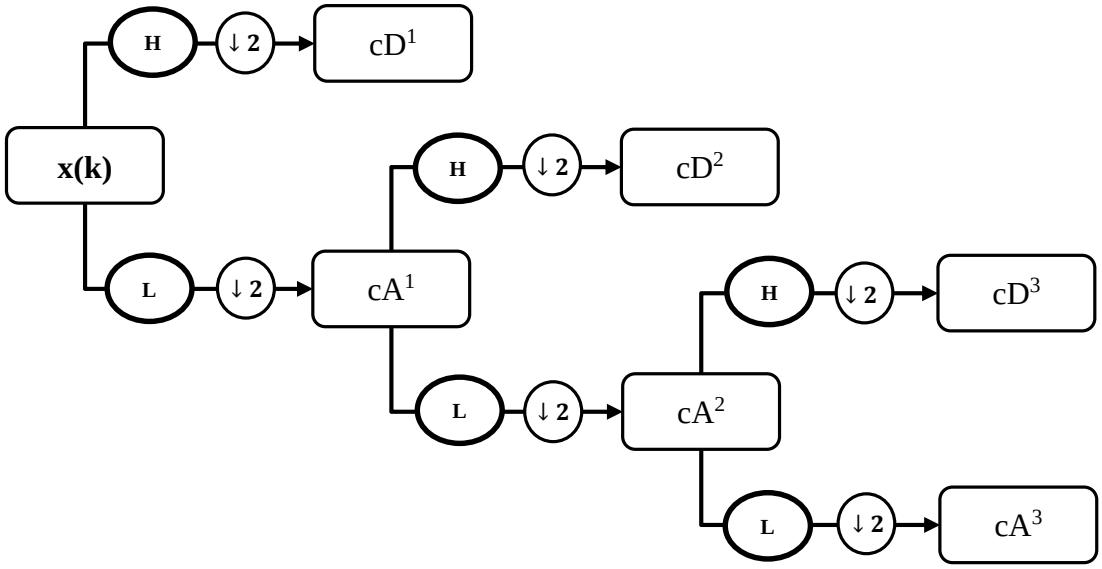

Figure 2.18: Waterfall diagram for DWT decomposition at 3 levels (j = 3).

By indicating that the maximum frequency of the input signal (F_{max}) is equal to the Nyquist frequency ($F_{Nyquist}$), the frequency bandwidth of each level is approximately

$$BW_j = {F_{Nyquist}}/{2^j} \quad \textbf{(2.78)}$$

where BW_j is the bandwidth of the j^{th} level.

Knowing that the maximum frequency of the level m is equal to the measured signal maximum frequency divided by 2^j, the maximum frequency of cA^j can be obtained as

$$F_{max}(A^j) = \frac{F_{Nyquist}}{2^j} \quad \textbf{(2.79)}$$

For example, in case of a signal with 20 kHz sampling frequency, the maximum frequency of the approximate coefficients at level 3 will be 1250 Hz.

According to the Equations (2.75) and (2.76), the decomposition process in DWT is only applied for the output signals in low-pass filters (approximation coefficients) at each step. Wavelet Packet Transform (WPT) is a simple generalization of the DWT that employe the decomposition process on the output signals in both high-pass and low-pass filters (approximation and detail coefficients). To perform the WPT on a vibration signal at a certain level (e.g. level 3), the wavelet and scaling functions, respectively, in Equations (2.70) and (2.72) are unified as [38]

$$w_{2p}(t) = \sqrt{2}\sum_k h(k) w_p(2t-k) \quad \textbf{(2.80)}$$

$$w_{2p+1}(t) = \sqrt{2}\sum_k g(k) w_p(2t-k) \quad \textbf{(2.81)}$$

where $w_0(t) = \varphi(t)$ and $w_1(t) = \Psi(t)$. Accordingly, the input vibration signal can be decomposed as [38]

$$cD_{2p}^{j+1} = \sum_n h(n-2k) cD_p^j \quad \textbf{(2.82)}$$

$$cD_{2p+1}^{j+1} = \sum_n g(n-2k) cD_p^j \quad \textbf{(2.83)}$$

where cD_p^j denotes the wavelet coefficients at the j^{th} level and p^{th} sub-band and cD_{2p}^{j+1} and cD_{2p+1}^{j+1} denotes the wavelet coefficients at the $(j+1)^{th}$ level, $(2p)^{th}$ and $(2p+1)^{th}$ sub-bands, respectively. Here, n is the number of the wavelet coefficients.

It is worth mentioning that, even if the complexity in implementation of the Mallat algorithm is ignored (because, it can be done in MATLAB® platform), the traditional DWT (also called first generation Wavelet Transform) has four major issues. Firstly, DWT suffers from time-invariant problem [38]. This means that the DWT of a temporal shifted version of a signal, x(k), is not the shifted version of the DWT of signal, x(k), and produces different wavelet coefficients. Secondly, DWT has the aliasing problem in the Wavelet Transform domain, particularly for lower scale range [39]. Although the IDWT cancels this aliasing, but only in case of the same coefficients in the reconstruction process which is not the case in the wavelet-based signal de-noising. In general, de-noising methods modify the wavelet coefficients so that the coefficients are not the same in the forward and inverse transforms. Thirdly, wavelet coefficients oscillate around the singularities as wavelets act like a band-pass function [88]. In a vibration signal, singularities or singular points are characterized by abrupt changes at confined regions (where the impacts occur).This complicates signal analysis based on Wavelet Transform [89]. Fourthly, DWT suffers from the lack of the directionally selective properties (a usefull property in 2D wavelet-based de-noising i.e. image processing) in wavelet filters in higher dimensions [90].

2.3.6.2 Second generation wavelet transform (SGWT)

The Wavelet Transform via lifting scheme is called Second Generation Wavelet Transform (SGWT) and has been introduced by W. Sweldens in [46] and [91] as an alternative implementation of the classical DWT. Figure 2.19 illustruates how the lifting scheme works, which is very similar to its counterpart in the classical DWT. Here, P is the prediction operator and U is the update operator.

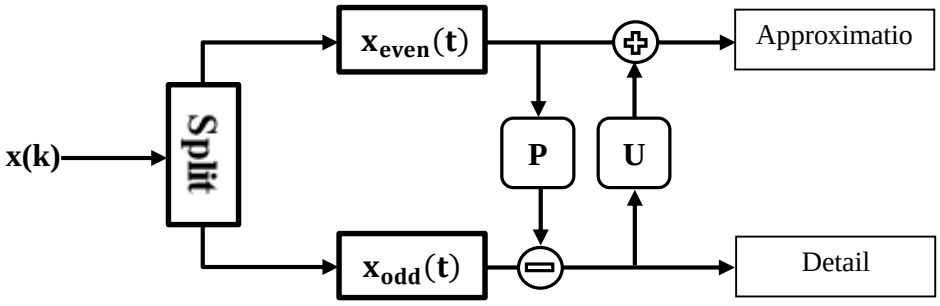

Figure 2.19: Illustration of signal decomposition using lifting scheme.

According to Figure 2.19, SGWT can be implemented in three main steps: 1) splitting, 2) prediction, and 3) updating. When the lifting scheme is performed, the discrete

signal is firstly splitted into two subsets; the odd samples, $x_{odd}(k)$, and the even samples, $x_{even}(k)$ by means of a sample operation. For example, for a given input signal, x(k), where k = {1, 2, 3, …, N} and assume N to be an even natural number, it split as

$$x_{odd} = \{x(1), x(3), x(5), \dots, x(N-1)\} \quad \textbf{(2.84)}$$

$$x_{even} = \{x(2), x(4), x(6), \dots, x(N)\} \quad \textbf{(2.85)}$$

After the splitting operation is completed, the odd and even sub-samples are obtained and the signal is sub-sampled by factor of two. Then the prediction operation is executed, which predicts the odd data sample according to the even data sample as [38]

$$\overline{x_{odd}} = P(x_{even}) \quad \textbf{(2.86)}$$

The difference between the predicted odd samples and the actual odd samples is then named as cD, and considered as the detail coefficients of the signal, x(k).

$$cD = x_{odd} - \overline{x_{odd}} = x_{odd} - P(x_{even}) \quad \textbf{(2.87)}$$

According to the even samples, x_{even}, and the detail coefficients, the approximation coefficients can be obtained via update operator U as

$$cA = x_{even} + U(cD) \quad \textbf{(2.88)}$$

Both prediction and update operations are independent of the input signal. These functions are similar to pair of h(k) and g(k) filters in the traditional Wavelet Transform (Mallat algorithm), and can be derived from the wavelet function $\Psi(t)$ and the scaling function $\varphi(t)$ in Equations (2.70) and (2.72), respectively, by the iteration algorithm [46]. It should also be noted that the prediction and update operators can be optimized using different algorithms, such as the Claypoole's optimization algorithm in [92]. After the mentioned steps, the input vibration signal is decomposed into approximation and detail segments and this process can be extended by taking the approximation parts at each level as the input signal.

The reconstruction of the SGWT, so-called Inverse Second Generation Wavelet Transform (ISGWT) can be obtained from Equations (2.87) and (2.88) but in reverse as follow [93]

$$a = x_{even} - U(d) \quad \textbf{(2.89)}$$

$$d = x_{odd} - \overline{x_{odd}} = x_{odd} + P(x_{even}) \quad \textbf{(2.90)}$$

The coefficients, a and d in above equations are then merged in order to obtain the original signal. Figure 2.20 illustrates the reconstruction process using lifting sheme in ISGWT.

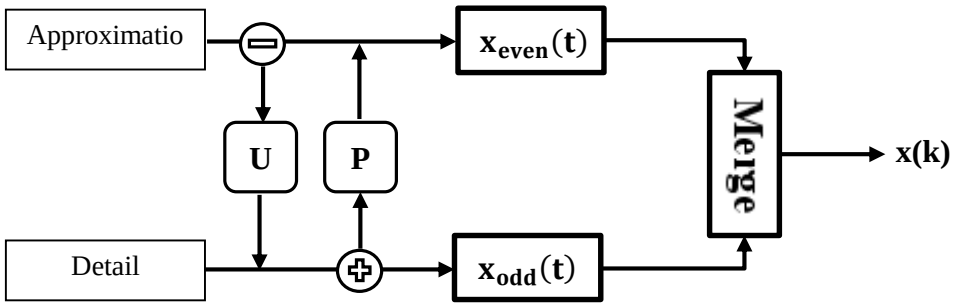

Figure 2.20: Illustration of signal reconstruction using lifting scheme.

In SGWT, only approximation coefficients have taken into acount at each step of lifting scheme. By decomposition of the detail and the approximation parts together via the lifting scheme, the so-called Second Generation Wavelet Packet Transform (SGWPT) is achieved. N. Li *et al*. have shown that the SGWPT has a strong capability in the fields of gearbox and ball bearing fault detection in [94, 88].

Although SGWT still suffers from the aliasing, which restricts its further application in fault feature extraction [93], it has a faster decomposition algorithm compared to traditional DWT by a factor of 2 [95].

2.3.6.3 Redundant discrete wavelet transform (RWT)

The Redundant Discrete Wavelet Transform (RDWT) [47] also known as algorithme à trous [96, 97], the Undecimated DWT (UDWT) [28], and the Stationary Wavelet Transform (SWT) [87] is another Wavelet Transform algorithm, which is designed to overcome the lack of time-invariant and oscillation of the wavelet coefficients around the singularity points in traditional discrete Wavelet Transform.

As stated earlier, to avoid the redundancy of the coefficients, DWT uses downsampling technique that prevents redundancy and allows using the same pair of filters in different levels. As a result, it causes the lack of shift-invariance property and may causes some error in reconstruction [98]. By eliminating of the downsampling operation in RDWT, the number of coefficients at each level is the same with original signal. As a result, in decomposition process through a filter bank, the high-pass and

low-pass filters at each level must be modified for the next level of decomposition. For this purpose, the filters at each level are subjected to the upsampling operator (by putting zero between each filter's coefficients of the previous level). Figure 2.21 shows the decomposition process by the RDWT at 3 levels. As can be seen, different filters are used at each step of the decomposition.

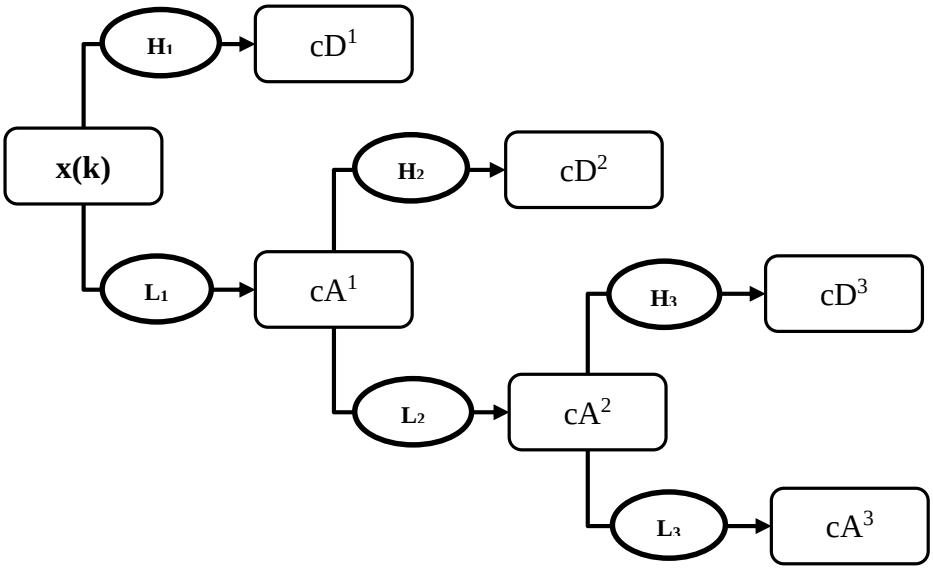

Figure 2.21: Waterfall diagram for RDWT decomposition at 3 levels (j = 3).

Suppose an input vibration signal, x(k), of length N, where $N = 2^a$ for some integer a. The first step of the decomposition process based on RDWT can be expressed as

$$cA_k^1 = \sum_j h_1(k-j)x(k) \quad \textbf{(2.91)}$$

$$cD_k^1 = \sum_j g_1(k-j)x(k) \quad \textbf{(2.92)}$$

where h_1 and g_1 are the orthogonal low-pass and high-pass wavelet filters, respectively. Since no downsampling is used in RDWT, cA_k^1 and cD_k^1 are the sequences of length N. For the further levels in RDWT, the decomposition can be written as [99]

$$cA_k^{j+1} = \sum_n h_{j+1}(k-n)cA_k^j \quad \textbf{(2.93)}$$

$$cD_k^{j+1} = \sum_n g_{j+1}(k-n)cA_k^j \quad \textbf{(2.94)}$$

where h_{j+1} and g_{j+1} are upsampled versions of the h_j and g_j, respectively.

The sub-signals in RDWT can be reconstructed via inverse RDWT in the reconstruction process. Since RDWT is an overcomplete data representation, the reconstructed signal is not unique if the coefficients are modified before the reconstruction process (i.e. de-noising applications).

RDWT decomposes the output coefficient vectors in low-pass filters at each step. The concept of wavelet packet can also be employed here to reach the so-called Redundant Discrete Wavelet Packet Transform (RDWPT) or Stationary Wavelet Packet Transform (SWPT). The corresponding formulations and application of RDWPT in the field of fault detection can be found in [100].

2.3.6.4 Dual-tree complex wavelet transform (DTWT)

The Dual-Tree Complex Wavelet Transform (DTCWT) is a relatively recent enhancement to the traditional DWT, which can overcome all mentioned major issues with DWT [48]. Although RWT solves the lack of shift-invariant problem, but it is a very high redundant transform. DTCTW is a nearly shift-invariant transform [39], which gives sufficient good results in wavelet-based decomposition process. In the dual-tree implementation of decomposition, two parallel DWTs with different low-pass and high-pass filters at each step are used. According to the nature of the Wavelet Transform, two different wavelet functions which must be orthogonal to each other are needed to employ two parallel DWT [48]. The mentioned wavelet functions can merged and accordingly made an analytic function. The analytic wavelet function $\Psi_c(t)$ can be composed of two wavelets as [101]

$$\Psi_c(t) = \Psi_h(t) + i\Psi_g(t) \quad \textbf{(2.95)}$$

where $\Psi_h(t)$ and $\Psi_g(t)$ are real even and real odd functions, respectively. Here, i stands for imaginary unit. As seen in Equation (2.95), the two real wavelets constitute a complex analytical wavelet $\Psi_c(t)$ which is only supported on the positive of the frequency axis [39]. In addition, as $\Psi_h(t)$ and $\Psi_g(t)$ must be orthogonal to each other, these functions must satisfy the following relationship [101]

$$\Psi_g = H[\Psi_h] \quad \textbf{(2.96)}$$

where H[.] denotes the Hilbert Transform operator.

Similarly to wavelet functions, according to the wavelet theory, the scaling functions can be obtained as

$$\varphi_c(t) = \varphi_h(t) + i\varphi_g(t) \quad \textbf{(2.97)}$$

where the same relation holds here i.e. orthogonality of the functions.

As can be seen in Figure 2.22, DTCWT uses two separate decomposition tree; one is so-called real tree (tree-a) and the other is so-called imaginary tree (tree-b). Accordingly, the wavelet and scaling coefficients can be obtained via inner product (i.e., Equation (2.71)). From Wavelet Transform theory, for the real tree the coefficients are as

$$cA_j^{Re}(k) = 2^{\frac{j}{2}} \int_{-\infty}^{+\infty} x(t)\Psi_h(2^j t - n)dt \quad \textbf{(2.98)}$$

$$cD_j^{Re}(k) = 2^{\frac{j}{2}} \int_{-\infty}^{+\infty} x(t)\varphi_h(2^j t - n)dt \quad \textbf{(2.99)}$$

Similarly for the imaginary tree the coefficients are as

$$cA_j^{Im}(k) = 2^{\frac{j}{2}} \int_{-\infty}^{+\infty} x(t)\Psi_g(2^j t - n)dt \quad \textbf{(2.100)}$$

$$cD_j^{Im}(k) = 2^{\frac{j}{2}} \int_{-\infty}^{+\infty} x(t)\varphi_g(2^j t - n)dt \quad \textbf{(2.101)}$$

DTCWT can be implemented using traditional method proposed by Mallat (fast algorithm) as

$$cD_{j+1}^{Re}(k) = \sum_m h_0(m - 2n)cD_j^{Re}(m) \quad \textbf{(2.102)}$$

$$cA_{j+1}^{Re}(k) = \sum_m h_1(m - 2n)cD_j^{Re}(m) \quad \textbf{(2.103)}$$

Similarly, for the imaginary tree coefficients can be obtained as

$$cD_{j+1}^{Im}(k) = \sum_n g_0(n - 2n)cD_j^{Re}(n) \quad \textbf{(2.104)}$$

$$cA_{j+1}^{Im}(k) = \sum_n g_1(n - 2n)cD_j^{Re}(n) \quad \textbf{(2.105)}$$

where the pairs of (h_0, h_1) and (g_0, g_1) are conjugate filters of real tree and imaginary tree, respectively, where the below relation holds between two low-pass filters in DTCWT [102]

$$g_0(n) = h_0(k - 0.5) \tag{2.106}$$

Note that, cA_j and cA_j are real and have equal length with the original signal, x(t), whereas cA_j^C and cD_j^C are complex and do not have equal length with the original signal. According to the nature of the DTCWT, the parallel filters in the frst level are different from the filters in upper decomposition levels [48].

The reconstruction algorithm of the DTCWT for real and imaginary trees can be expressed as [39]

$$cD_m^{Re}(k) = \sum_m \tilde{h}_0(k - 2m)cD_{j+1}^{Re}(m) + \sum_m \tilde{h}_1(k - 2m)cA_{j+1}^{Re}(m) \tag{2.107}$$

$$cD_m^{Im}(k) = \sum_n \tilde{g}_0(k - 2n)cD_{j+1}^{Im}(n) + \sum_n \tilde{g}_1(k - 2n)cA_{j+1}^{Im}(n) \tag{2.108}$$

The synthesis filter pairs $(\tilde{h}_0, \tilde{h}_1)$ and $(\tilde{g}_0, \tilde{g}_1)$ are form orthogonal or biorthogonal pairs with their respective counterpart filters (h_0, h_1) and (g_0, g_1), respectively.

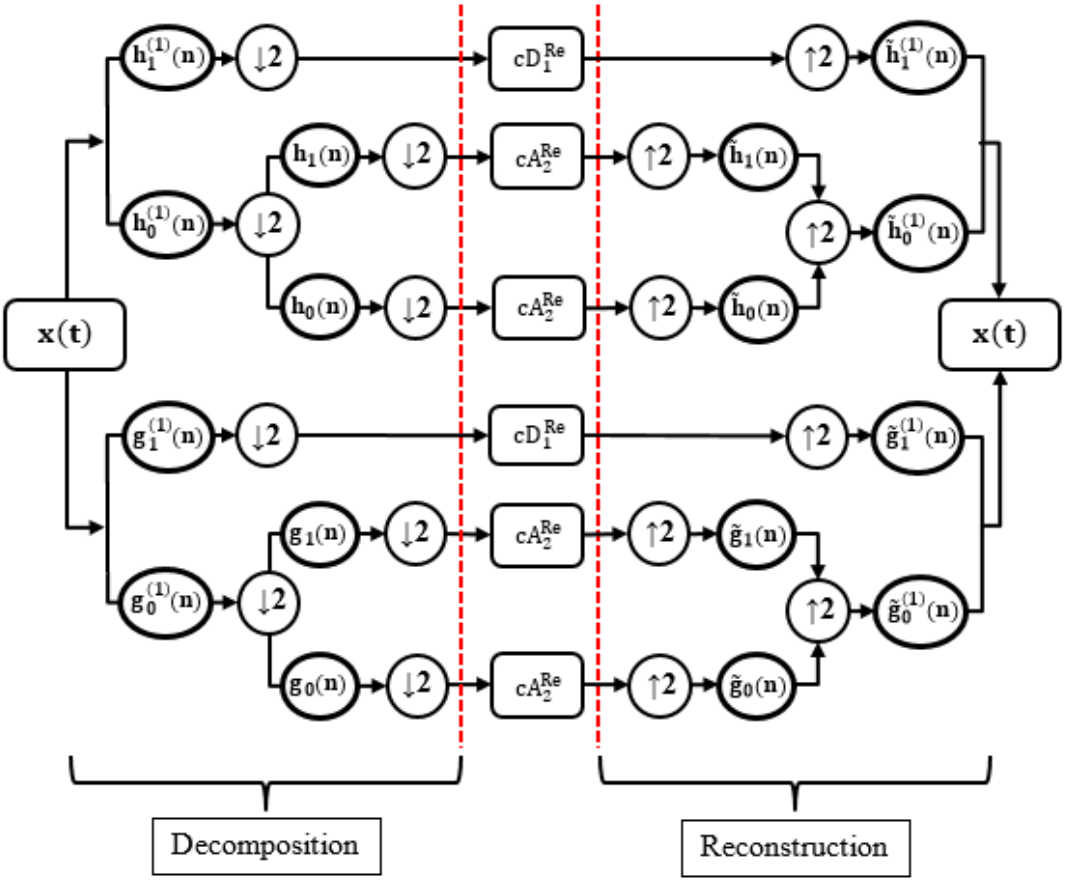

Figure 2.22: The decomposition and reconstruction process of Dual-Tree Complex Wavelet Transform (DTCWT).

The concept of wavelet packet transform can be also employed on the DTCWT which leade to a new algorithm called Dual-Tree Complex Wavelet Packet Transform (DTCWPT). More information about the mentioned method can be found in [103, 104].

2.3.6.5 De-noising based on wavelet transform

De-noising is one of the primary applications of WT. In decomposition process, it is possible to identify noisy components, modify, and then reconstruct the signal with modified coefficients. This is the basis of the all available wavelet-based de-noising methods.

To have better understanding of the wavelet-based de-noising procedure, assume the noisy raw vibration signal, x(t), in Equation (2.55). After applying the Wavelet Transform to the vibration signal, Equation (2.55) can be represented in terms of wavelet coefficients as

$$W_x(t) = W_y(t) + W_n(t) \quad \textbf{(2.109)}$$

where $W_x(t)$, $W_y(t)$, and $W_n(t)$ are wavelet coefficients related to the original signal, noiseless signal, and Gaussian white noise signal, respectively.

Owing to the properties of orthogonality of the mother wavelets used in all discrete Wavelet Transforms described in the previous sections, signal and noise energy remain the same after the decomposition process. Therefore, the modifications on wavelet coefficients can lead to reconstruct the signal with less noie components. The corresponding modifications can be done by thresholding the coefiicients according to an appropriate threshold. Most of the noise in rotating machineries occur in high frequencies where the detail coefficients show their greatest influence. Therefore, restriction (thresholding) of the detail coefficients makes the reconstructed signal with less background noise.

Many studies have been released to define a threshold and thresholding algorithm for wavelet-based de-noising [105, 106, 107, 86, 28]. Most of them are an extension of the study by Donoho and Johnstone in [108] who defined a proper threshold (i.e. Universal Threshold) and two types of thresholding methods, so-called Hard thresholding and Soft thresholding. The thresholded detail coefficients via Hard thresholding method can be written as [108]

$$c\bar{D}_n^m = \begin{cases} cD_n^m, & \text{for } |cD_n^m| \geq t \\ 0, & \text{otherwise} \end{cases} \quad \textbf{(2.110)}$$

and the thresholded detail coefficients via Soft thresholding method can be written as

$$c\bar{D}_n^m = \begin{cases} \text{sgn}(cD_n^m)(|cD_n^m| - t), & \text{for } |cD_n^m| \geq t \\ 0, & \text{otherwise} \end{cases} \quad \textbf{(2.111)}$$

In both Soft and Hard thresholding, the de-noised coefficients is achieved by setting the small absolute value coefficients (below the selected threshold) to zero. Figure 2.23 shows an example of Hard and Soft threshlding operations for t = 0.4.

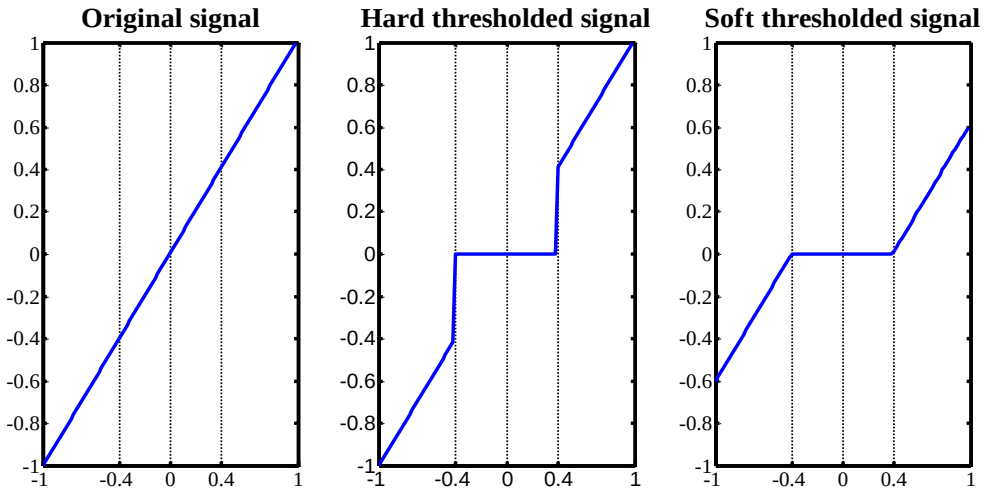

Figure 2.23: Illustration of the Hard and Soft thresholding methods.

According to the works by Donoho and Johnstone in [108], three different methods for determining a proper threshold are proposed. In what follow these methods are described in detail.

1) The Universal Threshold: It can be expressed as

$$t = \sigma\sqrt{2\log N} \quad \textbf{(2.112)}$$

where N is the signal length and σ is the estimate of the noise standard deviation. In practice, σ is taken as 1/4 or 1/3 of the standard deviation of the original signal.

2) Rigorous Sure Threshold (RST): It is a soft threshold evaluator of unbiased risk. Suppose $W = \{w_1, w_2, \ldots, w_n\}$ is a vector containing the square of wavelet coefficients for level m. Then the risk vetor can be written as

$$R_i = \frac{[n - 2i + (n - i)w_i + \sum_{k=1}^{n} w_k]}{n} \quad \textbf{(2.113)}$$

The selected threshold, t, is

$$t = \sigma\sqrt{w_i} \quad \textbf{(2.114)}$$

where w_i is the squared wavelet coefficients from the vector, W, related to the minimum value in risk vector. This threshold is always less than the Universal Threshold.

3) The Minimax Threshold: This method finds threshold, t, using min-max principle. It uses a fixed threshold to yield Minimax performance for mean square error against an ideal procedure. The related formulations can be found in [108].

Recent studies confirmed that although the Soft thresholding operation tends to smooth the signal slightly more than the Hard thresholding operation, it yields signal with better quality especially when the noise power is significant. Furthermore, with the chosen probability distribution, optimal Soft thresholding yields a lower risk than optimal Hard thresholding [109]. In the present study both thresholding methods have been applied to the vibration ball bearing signals.

The general wavelet-based de-nosing procedure can be summarized as follows

1) Select a discrete wavelet trasform method (e.g. SGWT).
2) Apply the decomposition based on the selected method to the input raw vibration signal.
3) Select the appropriate threshold (e.g. the universal threshold) for detail coefficient vectors at each level of decomposition.
4) Select the thresholding method (i.e. Hard or Soft) and apply the thresholding to the detail coefficients.
5) Reconstrcut the de-noised signal via inverse transform of the selected wavelet method in the first step using the modified coefficients.

Figure 2.24 from [39] compares the de-noised ball bearing vibration signals based on different approaches on Wavelet Transform in time domain. According to the obtained results the kurtosis values of the de-noised vibration signals by DWT, SGWT, and

DTCWT, respectively, are 16.94, 14.66, and 41.30, whereas the kurtosis value of the original noisy signal is 3.9. As seen, DTCWT shows superior performance in noise elimination of the ball bearing vibration signals compared to other Wavelet Transforms.

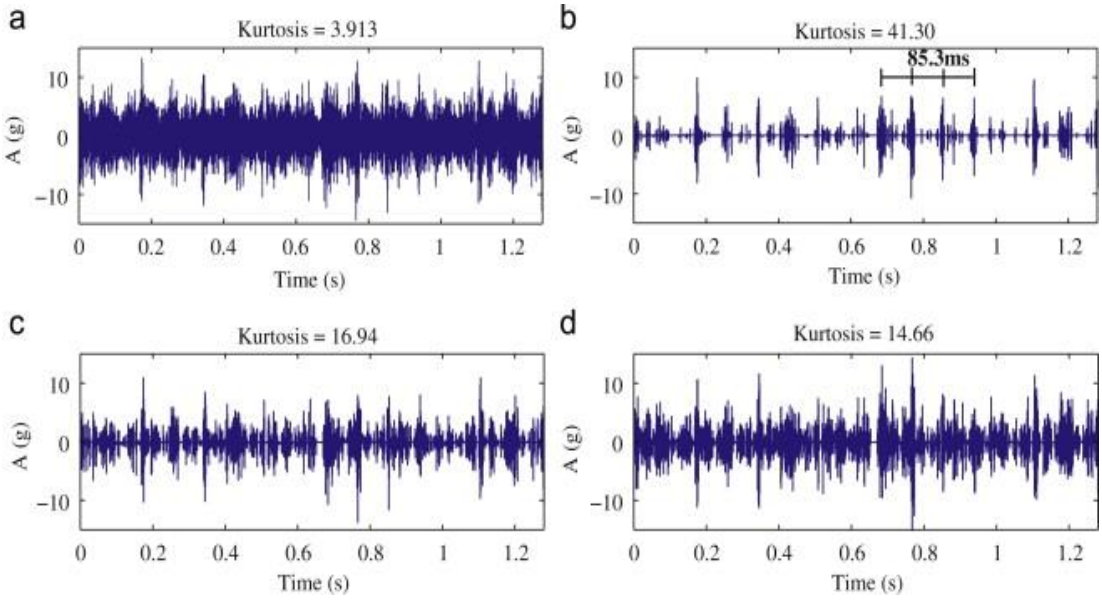

Figure 2.24: a) Original vibration signal; denoised vibration signal based on b) DTCWT c) DWT d) SGWT from [39].

2.4 Time-Frequency Domain Analyses

Up until now, the described fault detection methods assume the vibration signals to be stationary over the measurement time. However, the mentioned methods are unable to extract fault diagnosis information in non-stationary case. Varying the rotational speed or other environmental conditions (i.e., varying excitation forces) can make the signal to be non-stationary. During run-up where the rotating machine is not in its steady state condition all excitation forces may change both in amplitude and frequency. Therefore, some phenomenas which are usually not obvious at constant speed operation, may become more apparent under varying rotational speeds. To deal with non-stationary signals, several time-frequency analyses have been developed such as Short Time Fourier Transform (STFT), Continuous Wavelet Transfom (CWT), Wigner Ville Distribution (WVD), Choi-Williams Distribution (CWD), and Hilbert Huang Transform (HHT). In the present study the first two approaches which are more suitable for signal processing based fault detection purposes have been investigated in some detail.

2.4.1 Short time fourier transform (STFT)

The disadvantages of Fast Fourier Transform in processing of the non-stationary signals are overcomed by the introduction of Short Time Fourier Transform (STFT) which is localized in both time and frequency domains [110]. The mathematical model of STFT is given as [111]

$$\mathrm{STFT}(f,\tau) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j2\pi ft}dt \qquad \textbf{(2.115)}$$

where $w(t)$ is a window which is moved along the time axis. Normaly, the amplitude squared of STFT ($|\mathrm{STFT}(f,\tau)|^2$) is displayed on a time-frequency plot which is sometimes known as spectrogram or Campbell diagram. The window could be of finite length such as a Hanning window, or theoretically infinite such as a Gaussian window [1].

Figure 2.25 from [112] shows a typical application of STFT for gearbox fault diagnosis, where the fault frequency can be detected clearly around 250 Hz and 500 Hz (second harmonic) for this case.

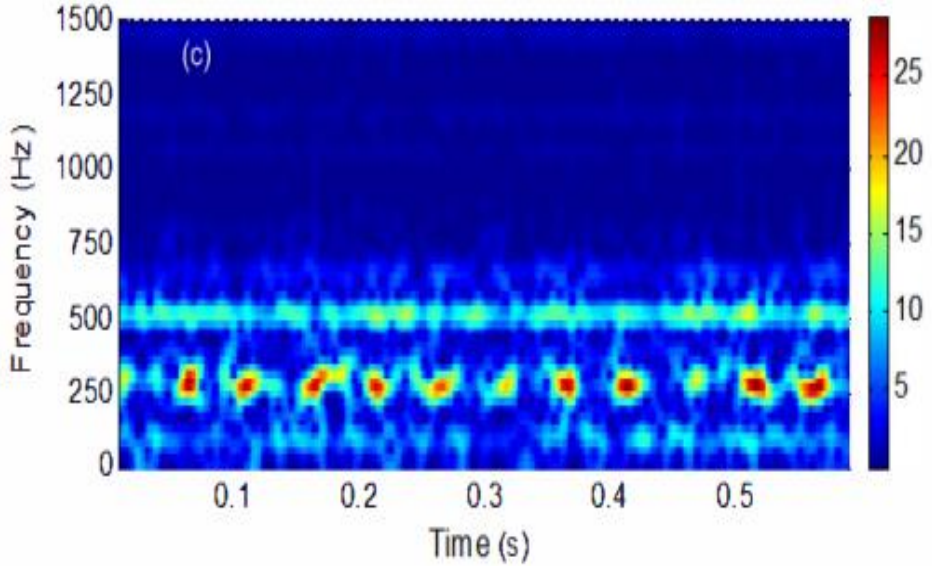

Figure 2.25: STFT based time-frequency representation of the vibration signal of a faulty gearbox, from [112].

According to research study in [113], some averaging in time-frequency domain on spectrograms may enhance the diagnosis capability of the STFT. Figure 2.26 illustrate the averaging effects on the readability of the spectrogram.

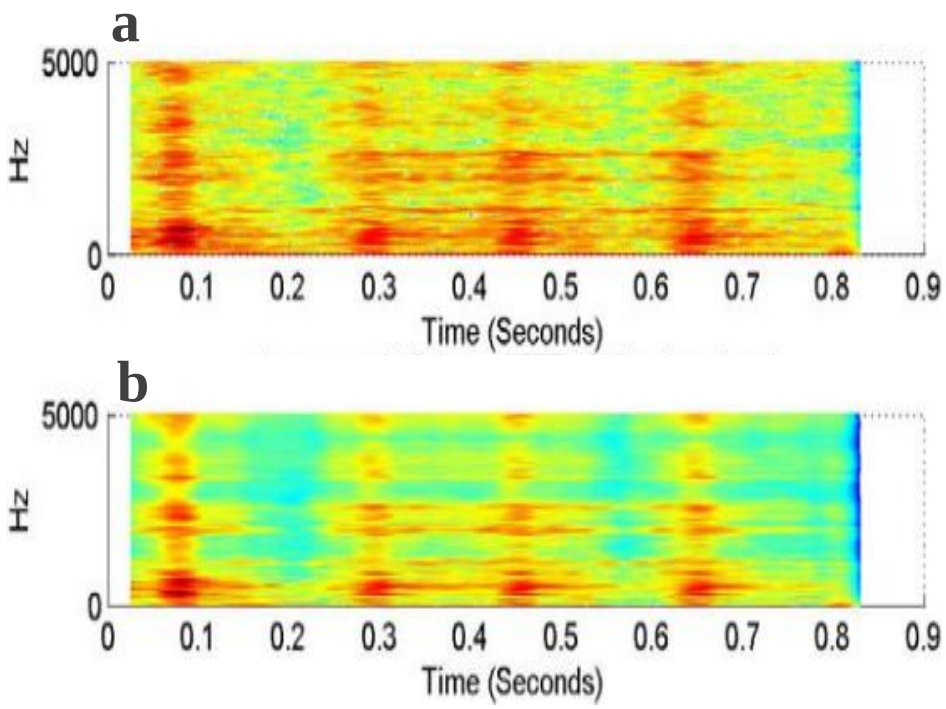

Figure 2.26: a) STFT, b) 48 times averaged STFT of the faulty ball bearing vibration signal from [113].

The STFT is computationally efficient, but has some drawbacks. Firstly, STFT uses a a sliding window having constant length in time and in some applications this may be a disadvantage. When a specific local area on time signal has a high frequency, the window will be shorter, while when the local area has a low frequency, the window will be longer [114]. Secondly, the choice of the window length simultaneously affects both frequency and time resolutions. For a good frequency resolution, a wider window length has to be chosen, but this choice adversely affects time resolution.

2.4.2 Continuous wavelet transform (CWT)

In this section wavelet trasfrom in its continuous form is summarized for time domain vibration signals to obtain time-scale and consequently time-frequency domain representation. The corresponding formulation for the Continuous Wavelet Transform (CWT) is given in Equation (2.68). According to the Equations (2.68) and (2.69), a 3D representation of the vibration signal in the time and frequency domains can be achieved.

A wavelet is a function that looks like a small wave, a ripple of the baseline, hence its name. As described earlier, there are many types of wavelets (as the mother wavelet)

in the literature; each is suitable for some specific applications. The analyzed signal properties and the corresponding application determine the type of the mother wavelet in Equation (2.68). Choosing an appropriate wavelet basis is usually very critical. From ball bearing fault detection point of view, the desired mother wavelet should have the ability to extract the impulsive features that are induced by localized bearing fault(s). By a series of tests and comparisons, it is found that the Morlet wavelet in its continuous form which has a similar shape to an impact, gives superior results compared to other wavelet basis [48], [49]. The Morlet wavelet can be expressed as

$$\Psi(t) = ce^{-\sigma^2 t^2} e^{i2\pi f_0 t} \qquad \textbf{(2.116)}$$

where σ, f_0, and c are the shape factor, wavelet center frequency, and a positive parameter which is typically chosen as $c = {}^{\sigma}/_{\sqrt{\pi}}$, respectively. Note that the Morlet wavelet is a non-orthogonal wavelet basis which is widely used in CWT. Figure 2.27 shows the Morlet wavelet with central frequency of 1 Hz at a ceratin scale level.

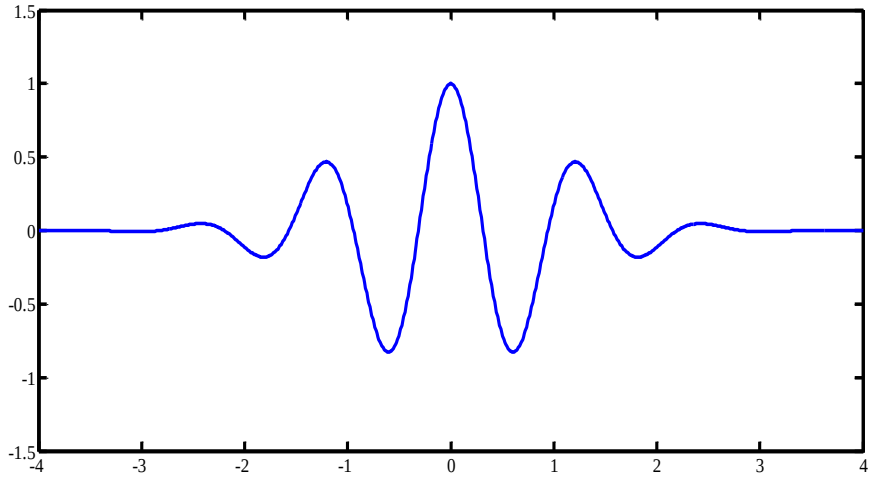

Figure 2.27: Illustration of the Morlet wavelet at a certain scale.

According to the Equations (2.68) and (2.69), as the parameters a and b vary, the output of the CWT is time-scale domain. The scale parameter, a, determine the dilatation or compression level of the mother wavelet. Therefore, it can nearly reflect the frequency of the wavelet at that scale. Consequently, the time-scale domain can be converted to time-(pseudo)frequency domain.

The variable window size of the CWT leads to yields good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies [115]. The basic difference between CWT and STFT is

that the CWT uses a size-adjustable window that is more advantageous than a fixed window used offered by STFT. Figure 2.28 from [116] compares the time-frequency representations of a case study based on STFT and CWT. As seen, STFT suffers from the energy leakage problem around the fault frequency.

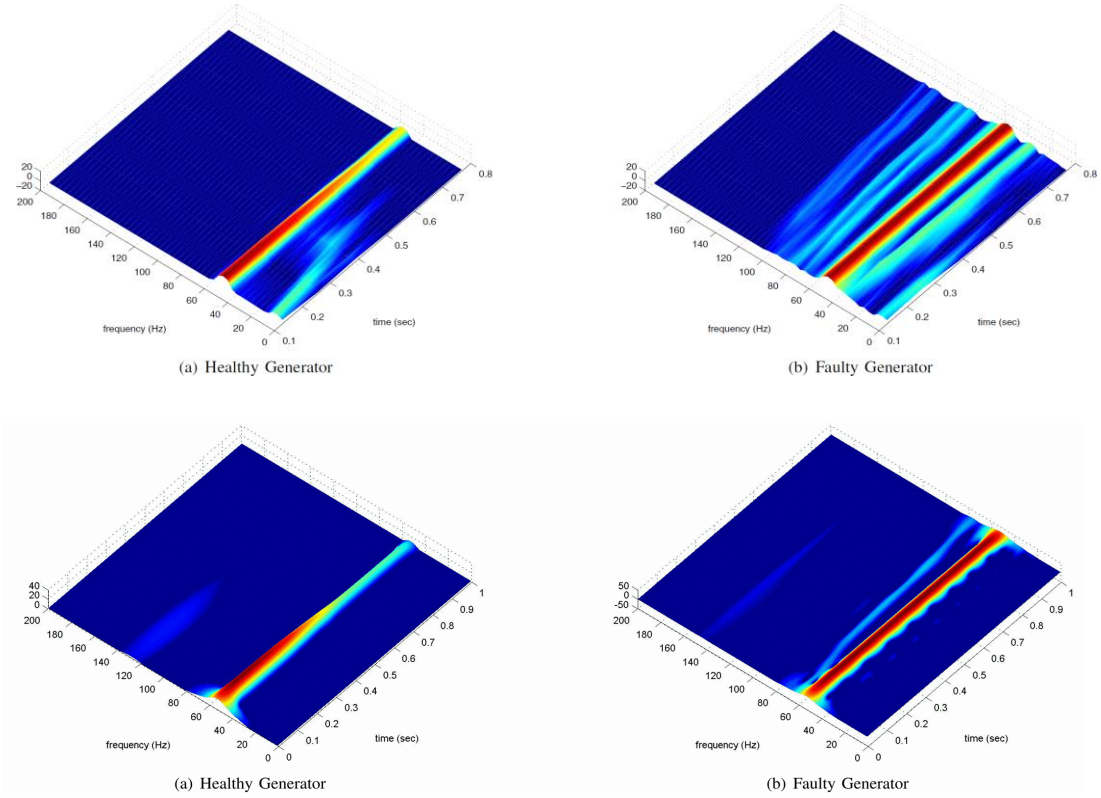

Figure 2.28: Time-frequency representation of the vibration signals captured from healthy (left) and faulty (right) wind turbines based on STFT (up) and the CWT (below) from [116].

Most of the time-frequency techniques for ball bearing fault detection are based on the detection of repetition frequency of the excited natural frequency. However, the captured signals may contain heavy background noise and this can adversely affect the ability of the diagnosis technique. According to the study by Y. Jiang *et al*. in [24], SVD based de-noising method which is described earlier in section 2.2.3.1 can be employed for the coefficient matrix obtained by CWT in order to achieve a smoother plot. The following steps can summarize the SVD based de-noising process in time-frequency domain:

1) Obtain the coefficient matrix by CWD
2) Consider the obtained coefficient matrix as the matrix [A] in Equation (2.10) and employ the SVD on coefficient matrix.

3) Plot the normalized singular values and determine a proper threshold value the singular values.

4) Replace all singular values below the threshold value with zero and obtain the modified coefficient matrix.

Figure 2.29 from [24] illustrates the original and de-noised time-frequency domains for the corresponding case study. As seen, the resonance frequency of the faulty ball bearing is excited with a certain repetition frequency. Comparing the results of the Wavelet Transforms with and without de-noising, it is seen that the de-noising process makes it possible to identity the fault period quite clearly.

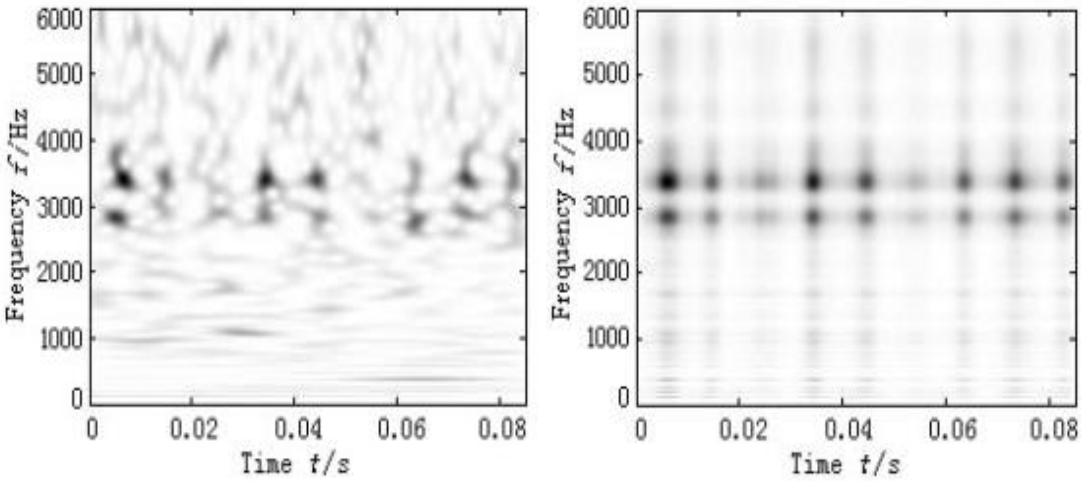

Figure 2.29: Time-frequency representation of the ball bearing vibration signal; original (left) and SVD based de-noised (right) time-frequency plots from [24].

Despite of the reliable results obtained by CWT in the field of ball bearing fault detection, it still has a major issue need to be studied and improved. When applying the Wavelet Transform method, the so-called energy leakage problem occurs in the time-frequency representation.

3. NUMERICAL INVESTIGATIONS

3.1 Introduction

In the preceding chapter, some of the available important techniques in the field of vibration-based ball bearing fault detection were investigated in some detail. Generally, it can be said that the fault diagnosis algorithms may yield excellent results under some specific conditions. However, their performance can be very poor in other situations. Since there is not a universally superior method for all ball bearings fault detection puprposes, it is necessary to study a wide range of fault detection methods for the purpose. Using synthetic data and simulated ball bearing vibration signals may give complete understanding of the analyzsed method. The present chapter starts with modelling of the ball bearing vibration signals both for healthy and defective types. The ability of the EMD and its revised version, named EEMD, as powerfull signal decomposition methods, are then investigated numerically. Next, the capability of the various type of discrete Wavelet Transforms for the signal de-noising purposes are studied. Finaly, the efficiency of STFT and CWT as the well-known time-frequency domain analyses methods in ball bearing fault detection are compared by perfoming numerical simulations using simulated faulty ball bearing vibration signals.

3.2 Simulation of Ball Bearing Signals

In general, modelling of the vibration signals of the real rotating machineries with ball bearings are complicated since the real machines have many vibration sources. Typical vibration sources in a rotating machinery can be devided into four categories: (i) synchronous excitations (i.e. shaft rotating frequency), (ii) non-synchronous excitations such as rubs and rotor whirl, (iii) rotor unbalances and misalignments, and (iv) impulsive periodic excitations (e.g. from defective supporting ball bearings) [87]. In addition, some background noise originating from the measurement system, environmental conditions, roughness of the bearing surfaces etc are unavoidable in captured vibration signals. From ball bearing fault detection point of view, all vibration sources, except the last part which is related to the ball bearings, are considered as the

uninterested components. In the present section, first, a general vibration signal model is defined for the rotating shaft as the superposition of the vibration responses due to all uninterested excitation sources. Then, the harmonic excitations from defective ball bearings are added to the shaft vibration signal in order to obtain faulty ball bearing vibration signal.

3.2.1 Rotating shaft harmonics

Most of the vibration sources in a rotating shaft are integer multiples of the rotating frequency, f_r. When the unbalance or misalignment occur, the response amplitude of the corresponding harmonic(s) increase in the vibration pattern as an indication of an abnormal condition. A general vibration signal of a rotating shaft having healthy supporting ball bearings in ideal conditions can be simply modeled as

$$h(t) = \sum_{n=1}^{p} A_n \cos(2\pi f_n t + \varphi_n) \qquad \textbf{(3.1)}$$

where A_n, f_n, p, and φ_n are amplitude, frequency, the number of harmonic components, and the corresponding initial phase, respectively. Note that, in above equation the first frequency, f_1, is considered as the shaft rotating frequency, f_r.

Since the real machineries are not in ideal conditions, the Gaussian white noise with zero mean is added to the signal in Equation (3.1). Consequently, the vibration signal captured from rotating shaft can be expressed as

$$x(t) = h(t) + n(t) \qquad \textbf{(3.2)}$$

where $n(t)$ represents the added noise.

3.2.2 Faulty bearing-induced harmonics

Defective ball bearings produce impulsive periodic signals that are at non-integer multiples of the shaft frequency as described in section 2.3.1. These impulses excite the resonances in the bearing and the whole rotating machine that can be approximated as an amplitude modulated of the signal in Equation (3.1). In what follows the vibration signals of two types of defective ball bearings are modelled.

3.2.2.1 Single faults

An impulsive vibration signal captured from a defective ball bearing having a single localized fault and single resonance frequency can be expressed as [117]

$$y_s(t) = d(t)\sum_{n=1}^{\infty}\delta\left(t-\frac{n}{f_b}-\tau\right)e^{-\alpha t}\cos(2\pi f_e t+\varphi) \qquad \textbf{(3.3)}$$

where $y_s(t)$, $d(t)$, α, f_e, f_b, φ and τ are multiple faulty ball bearing vibration signal, impact amplitude, damping factor, resonance frequency, fault repetition frequency, initial phase and initial delay, respectively. Here, δ stands for Dirac delta function.

Note that if the defect is on the outer race, the impact impulses have constant amplitude because the fault location does not change with respect to the bearing load distribution. However, for a bearing with an inner race fault, the amplitude of the defect impact impulses change with respect to angular position, and consequently, the impact impulses should be considered as time dependent variable as

$$d(t) = d_0|\cos 2\pi f_e t| \qquad \textbf{(3.4)}$$

where d_0 is amplitude of the impacts.

3.2.2.2 Multiple faults

In case of multiple faults, by asumming a linearly independent relation between the faults, the resultant vibration ideal signal is a summation of several impulse trains with different delays and repetition frequencies. As a result, Equation (3.3) can be generalized as [117]

$$y_m(t) = \sum_{p=1}^{m} d_p(t)\sum_{n=1}^{\infty}\delta\left(t-\frac{n}{f_{b,p}}-\tau_p\right)e^{-\alpha t}\cos(2\pi f_e t+\varphi) \qquad \textbf{(3.5)}$$

where p is the number of linearly independent localized faults on the ball bearing.

3.2.3 Simulation results

Based on Equation (3.3) and its generalized form in Equation (3.5), the numerical vibration signals can be generated in MATLAB® paltform. Depending on the ball bearing condition, synthetic signals of the normal ball bearing as well as defective ball bearings whether having a single fault or multiple faults can be constructed using specific desired properties. Accordingly three different vibration singnals are

presented and discussed: one for healthy ball bearings, one for the ball bearing having a single fault whether on inner or outer races, and one for the ball bearing having multiple faults simultaneously on outer and inner races.

Case Study 1: Consider a misalignment on a shaft rotates at 1200 RPM (20 Hz), which can be modelled as

$$h(t) = A_1 \cos(2\pi f_1 t + \varphi_1) + A_2 \cos(2\pi f_2 t + \varphi_2)$$

where $f_1 = 20$ Hz and $f_2 = 2 \times f_1$. Note that, as mentioned before, the gaussian white noise[1] must be added to the shaft vibration signal, $h(t)$. Figure 3.1 illustartes the construction steps of the healthy ball bearing simulated vibration signal with kurtosis = 2.9 for a time period of 0.5 second and the sampling frequency as 32 kHz. In addition, Figure 3.2 shows the amplitude spectrum of the signal in Figure 3.1c.

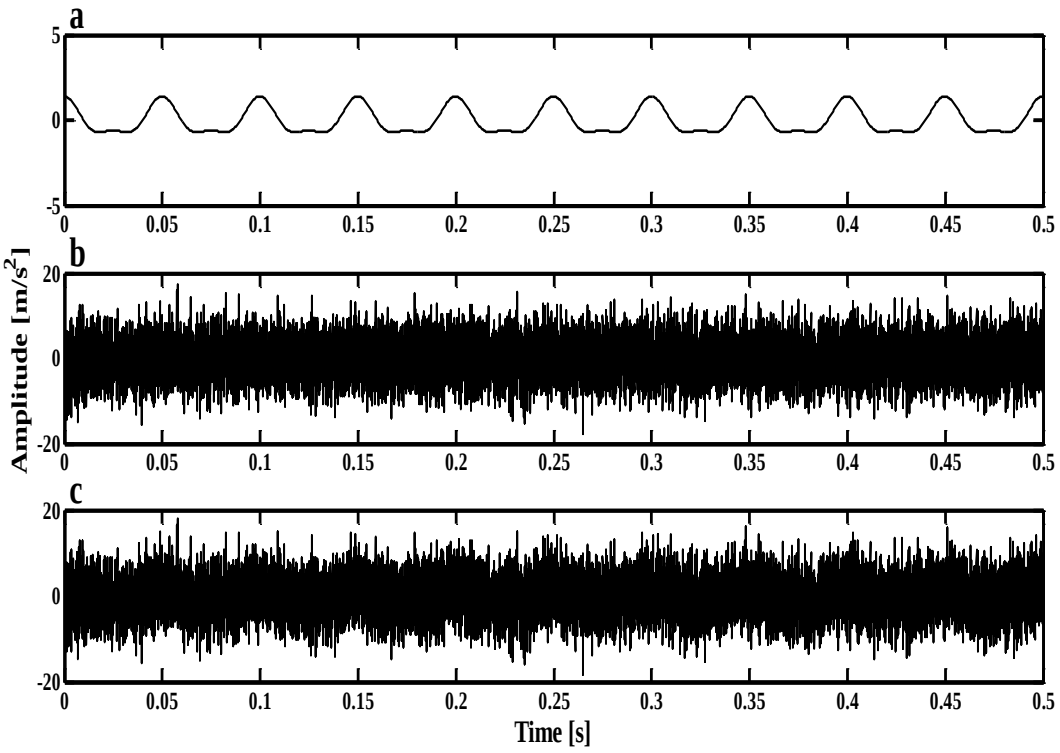

Figure 3.1: Construction of the numerical healthy ball bearing vibration signal; a) rotating shaft vibration signal b) white noise c) summation of the signals in a) and b).

[1] Added gaussian white noise having power as 13 dB.

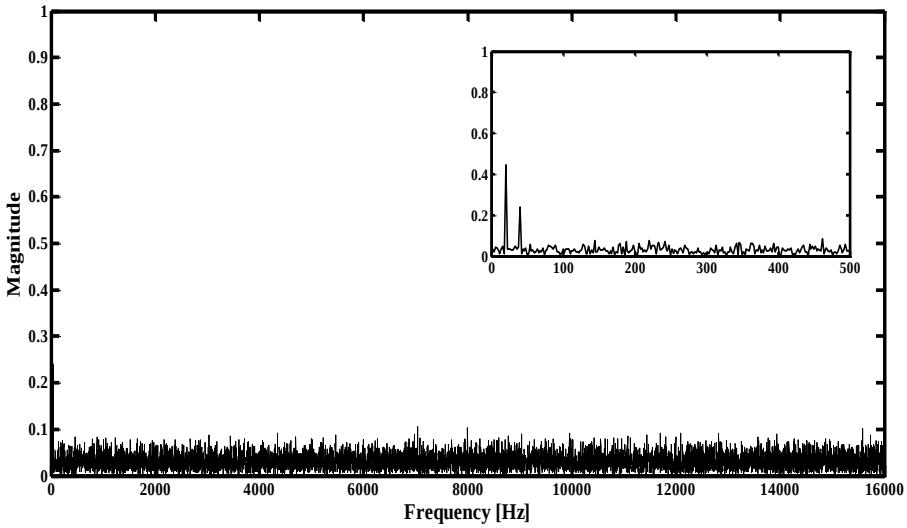

Figure 3.2: Amplitude spectrum of the vibration signal in Figure 3.1c.

Case Study 2: Assume the same rotating shaft (and having same background noise) in case study 1, but this time having the supporting ball bearing with a single localized fault on its outer race and a single resonance frequency at 4.5 kHz. Again, the vibration characteristics of the misalignment include an increase of vibration amplitude at 1× rotating frequency and higher than 2× harmonic amplitude. The corresponding fault repetitation frequency, BPFO, is assumed to be at 63 Hz. The mentioned faulty ball bearing vibration signal can be modelled by summation of three linearly independent signals; rotating shaft signal, white noise with zero mean, and impulse train as the representative of the periodic excitation due to faulty ball bearing. Figure 3.3 illustrates the constructing steps for the simulated faulty ball bearing vibration singnal with single localized fault with kurtosis = 3.3 for a time period of 0.5 second and the sampling frequency as 32 kHz. Figure 3.4 shows the amplitude spectrum of the signal in Figure 3.3d.

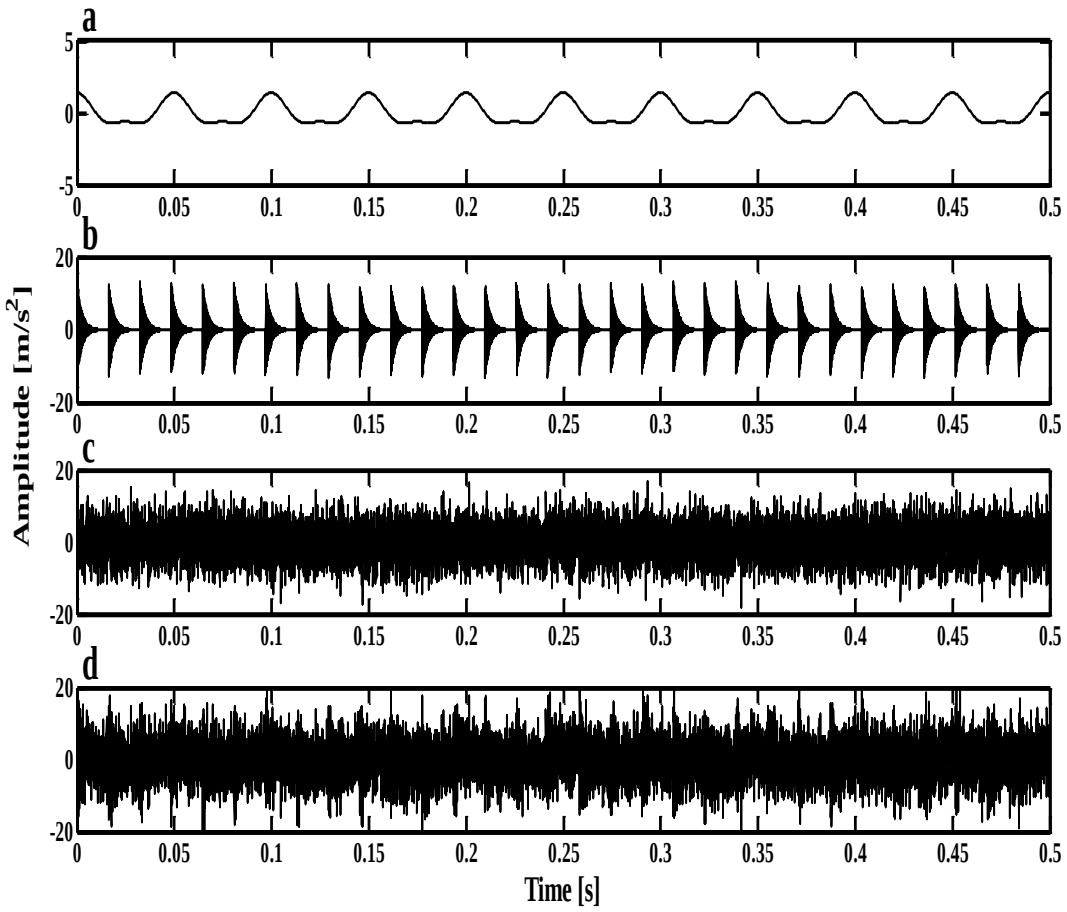

Figure 3.3: Construction of the faulty vibration signal with single defect on outer race; a) rotating shaft vibration singnal b) impulse train c) white noise d) summation of the signals in a), b), and c).

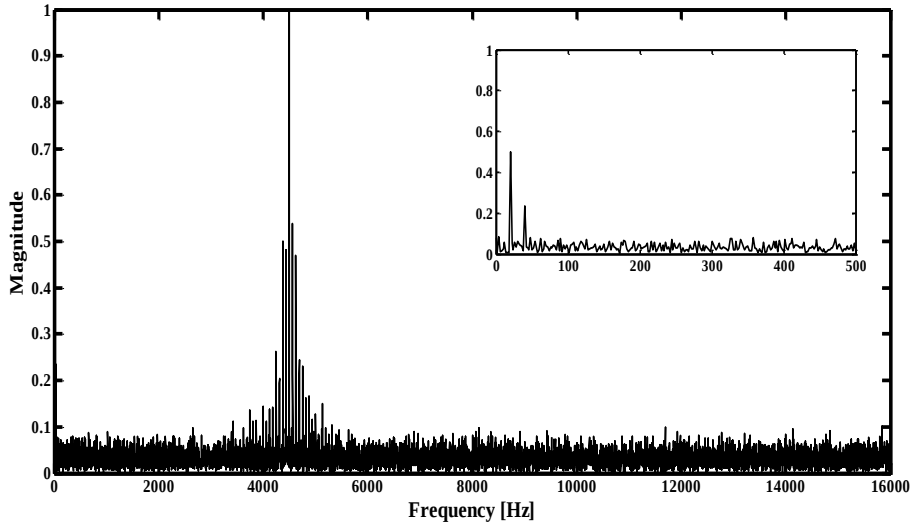

Figure 3.4: Amplitude spectrum of the vibration singal in Figure 3.3d.

As seen, the amplitude spectrum (traditional method) is unable to detect the fault symptoms of the faulty ball bearing simulated vibration signal due to modulated

components and added heavy background noise, whereas the fault frequency and its higher harmonics can be easily detected in the envelope spectrum in Figure 3.5.

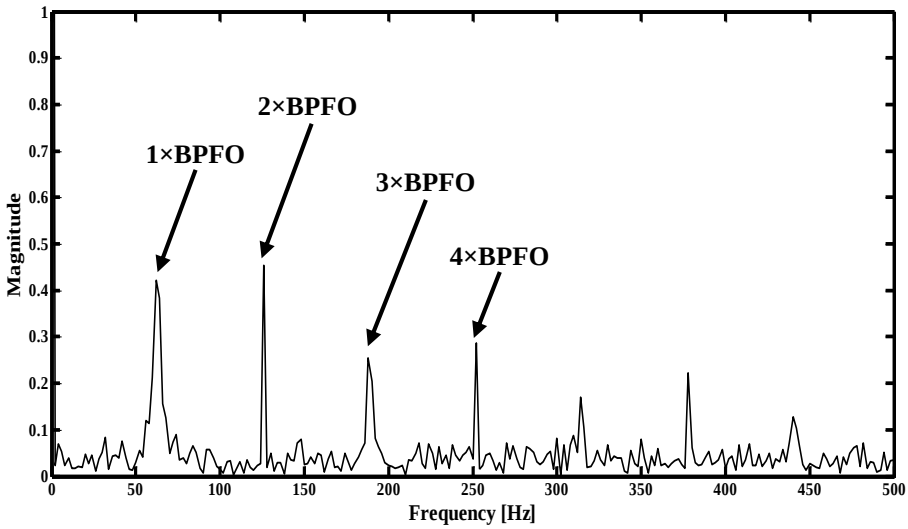

Figure 3.5: Zoomed envelope spectrum of the vibration signal in Figure 3.3d.

Case Study 3: Again, consider the same shaft with a misalignment rotating at 1200 RPM (20 Hz) with a supporting ball bearing having two localized faults on outer and inner races. The corresponding faulty vibration signal is generated using Equation (3.5) with 63 Hz and 96 Hz, respectively, as the outer and inner races fault repetition frequencies. It is assumed that both defects excite the same resonance frequency (4.5 kHz). In addition, the amplitudes of the simulated impulse trains depend on the assumed severty of the defects. In this case study, the defect on the outer race is assumed to be larger than the fault on the inner race. For the multiple fault case, Figure 3.6 illustrates the construction steps of a simulated ball bearing vibration signal with kurtosis = 3.5.

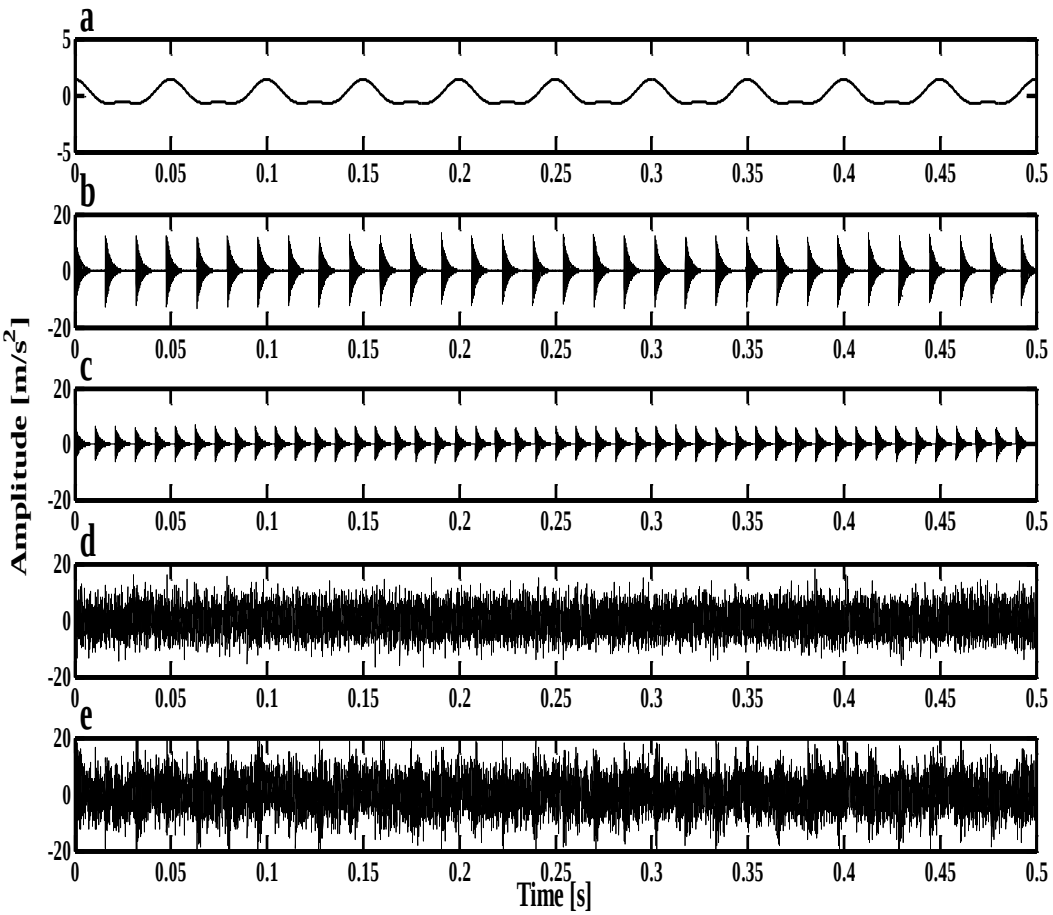

Figure 3.6: Construction of the faulty vibration signal with multiple defects on outer and inner races; a) rotating shaft vibration singnal b) impulse train due to outer race fault c) impulse train due to inner race fault d) white noise e) summation of the signals in a), b), c), and d).

Figure 3.7 shows the amplitude spectrum of the vibration signal in Figure 3.6e. Again, due to the same reasons mentioned before, amplitude spectrum cannot detect the fault frequencies. However, the envelope spectrum presented in Figure 3.8 reveals the corresponding fault frequencies. As expected, both fault frequencies related to outer and inner races and their higher harmonics can be detected in envelope spectrum. Figure 3.8 shows the advantageous of the envelope spectrum over the amplitude spectrum for the simulated vibration signals in the field of ball bearing fault detection.

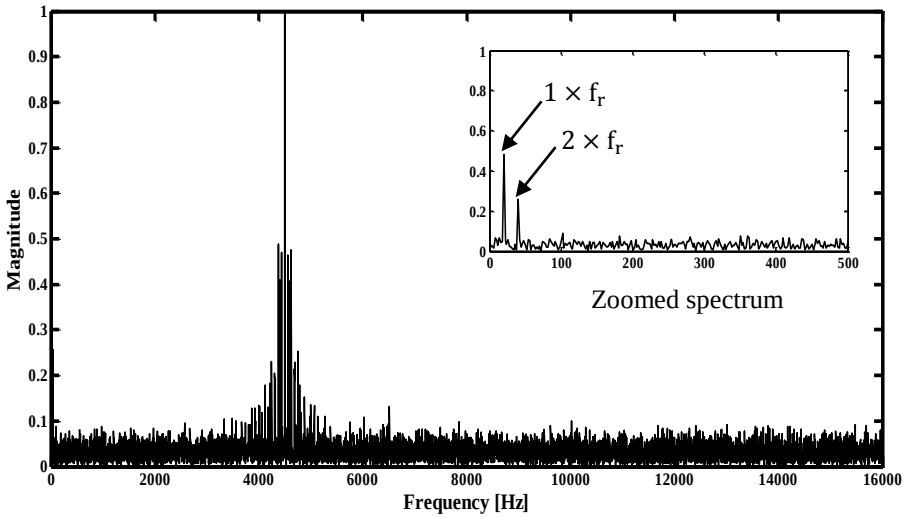

Figure 3.7: Amplitude spectrum of the vibration singal in Figure 3.6e.

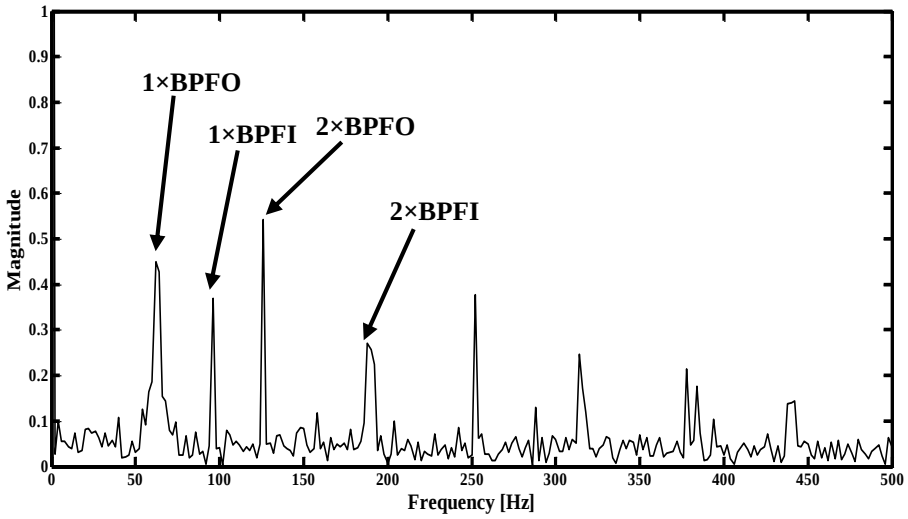

Figure 3.8: Zoomed envelope spectrum of the vibration signal in Figure 3.6e.

3.3 Comparison of the EMD and EEMD Methods

In this section, the powerful signal decomposition method so-called EMD and its revised version so-called EEMD are compared and investigated numerically for the purpose of ball bearing fault detection using the simulated vibration signal modelled in case study 2 (Figure 3.3d) in the previous section. Figure 3.9 and Figure 3.10 show the first 10 extracted Instrinic Mode Functions (IMFs) via EMD and EEMD methods, respectively. By considering the amplitude scales for the third and higher IMFs in both methods, it is pointed out that the only first and second IMFs contain the fault features for the corresponding case study. Accordingly, the first two IMFs are considered for further processes.

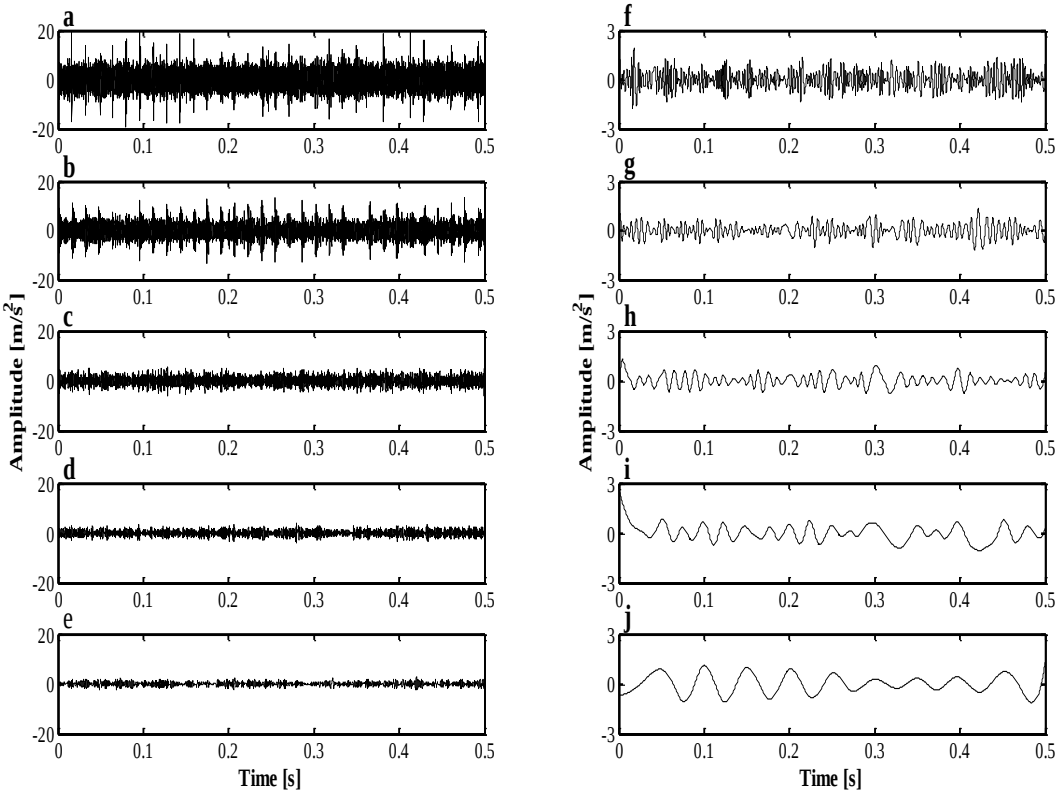

Figure 3.9: The first 10 IMFs of the vibration signal in Figure 3.3c via EMD, respectively, from a) to j).

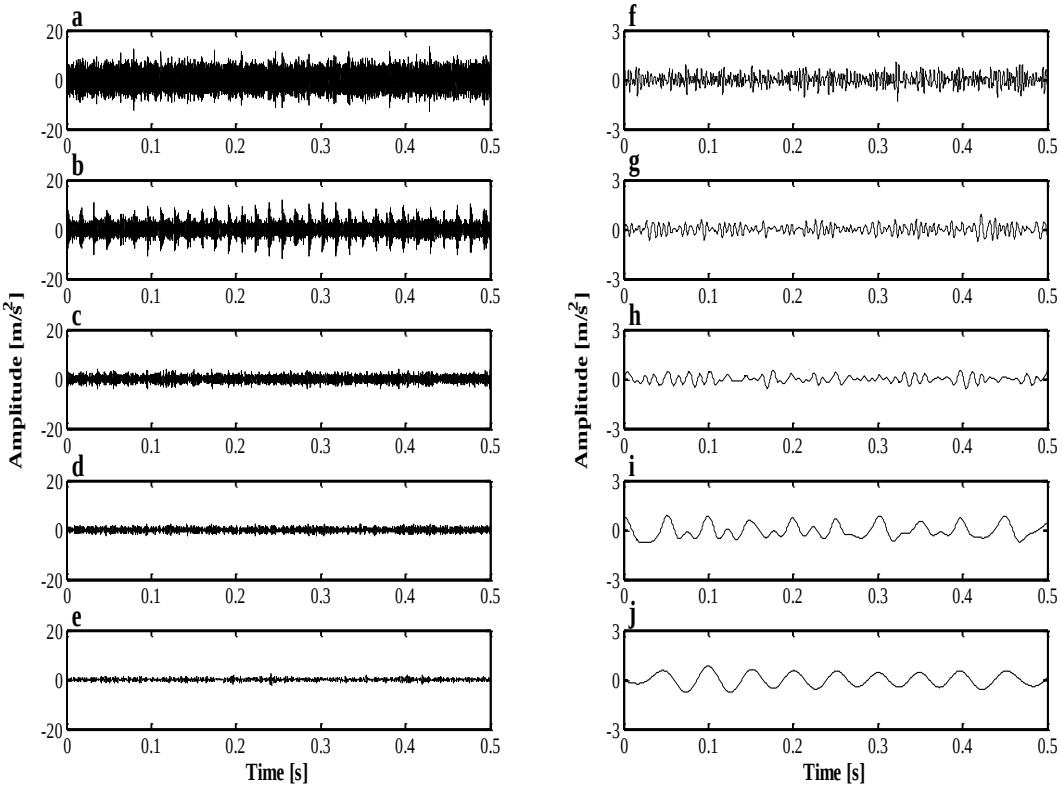

Figure 3.10: The first 10 IMFs of the vibration signal in Figure 3.3c via EEMD, respectively, from a) to j).

It should be noted that although the amplitude of the added white noise in EEMD method has significant role in signal enhancement, it is concluded in the present study that within a certain range of noise levels, the decomposition results have small sensitivity to the noise level. The ratio of standard deviations of the analyzed signal and the added white noise in EEMD is proposed to be about 0.5 ~ 0.8 for the vibration signals with the heavy background noise. It is also concluded that at least 100 ensemble number is needed to yield very reliable results in EEMD. Generaly, EEMD has more noise cancelation features than EMD, but for the analyzed simulated vibration signal, EMD performs better than EEMD. However, the so-called mode-mixing problem described in section 2.2.3.2, can be seen in IMFs (e.g. IMF7 and IMF8) obtained via EMD. Figure 3.11 and 3.12 show the zoomed envelope spectrums of the first, second, and third IMFs extracted using EMD and EEMD, respectively. As expected from time domain signals in Figure 3.9 and Figure 3.10, only first two IMFs have fault features of the corresponding simulated signal. The envelope spectrum of the first IMF related to the EMD method, however, have sharper fault amplitudes compared to EEMD method, hence yielding a slightly better detection of the related fault frequency.

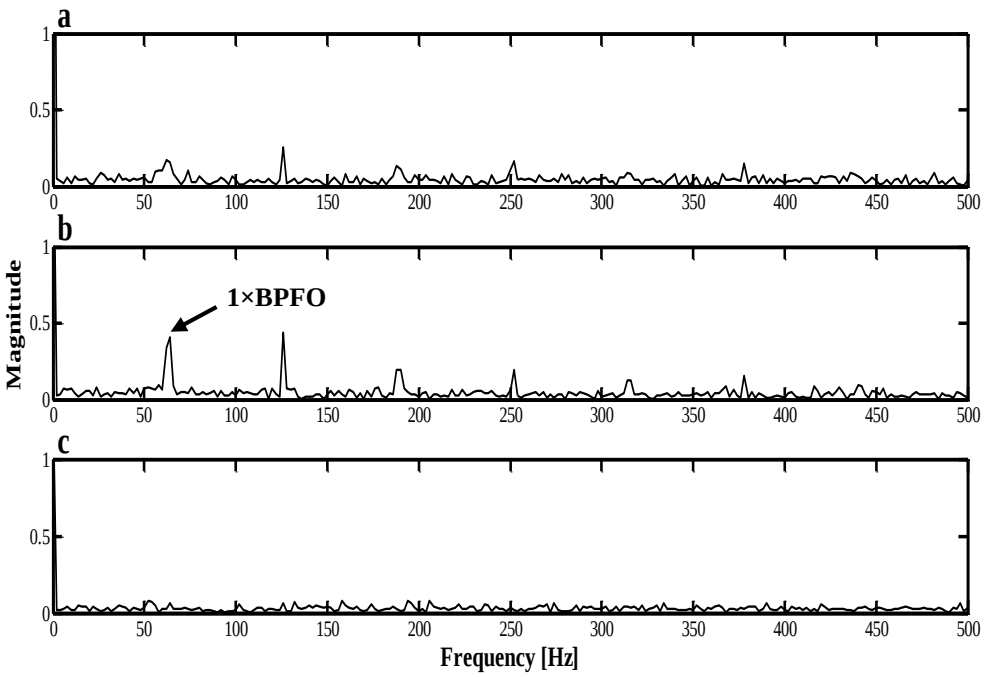

Figure 3.11: Zoomed envelope spectrums of the a) first b) second c) third IMFs extracted using EMD.

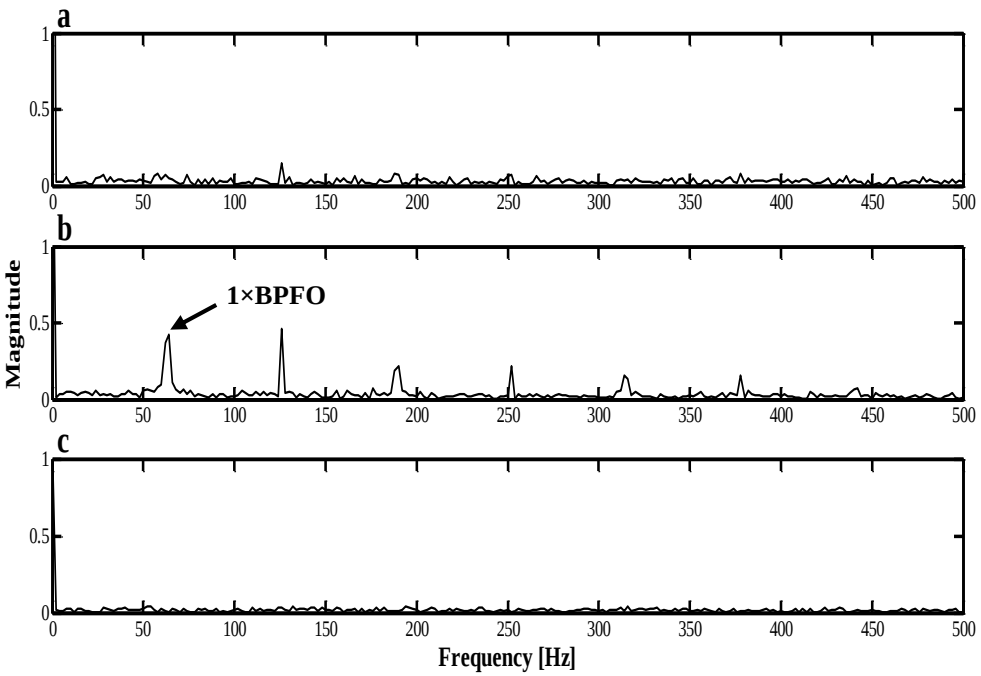

Figure 3.12: Zoomed envelope spectrums of the a) first b) second c) third IMFs extracted using EEMD.

3.4 Wavelet Based De-noising

In this section, one of the most crucial applications of the discrete Wavelet Transform are studied numerically; the case study 3 (i.e. ball bearing having two localized faults) is chosen. For this purpose, the Discrete Wavelet Transforms discussed in the previous chapter are employed for de-noising algorithm. Three thresholds (i.e. universal threshold, minimax threshold, and rigorous sure threshold) and two thresholding methods (i.e. hard and soft) are used and compared in de-niosing process using the simulated vibration signal in Figure 3.6e. This section shows the vibration signal enhancement based on noise elimination via Wavelet Transforms. For this purpose, db4 from family of Daubechies orthogonal wavelets is chosen as the wavelet basis and employed on all wavelet algorithms. The decomposition processes is done at 5 levels. Figure 3.13 to Figure 3.16 illustrate the de-noised signals based on DWT, SGWT, RDWT, and DTCWT, respectively, using various de-noising thresholds and thresholding methods. In addition, kurtosis values of the de-noised signals are listeded in Table 3.1 for each thresholding and Wavelet Transform methods. The highest kurtosis value for the de-noised signal is obtained by DTCWT with universal threshold and soft thresholding.

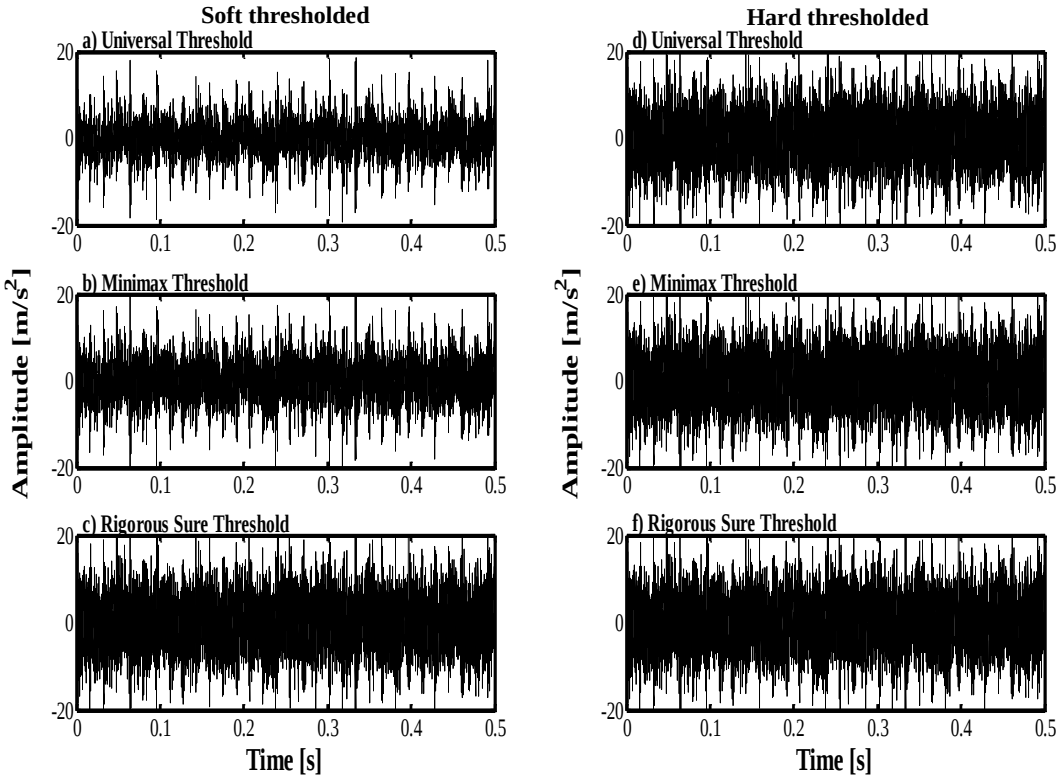

Figure 3.13: De-noised signals of the simulated vibration signal in Figure 3.6e based on DWT and soft (left) and hard (right) thresholding methods.

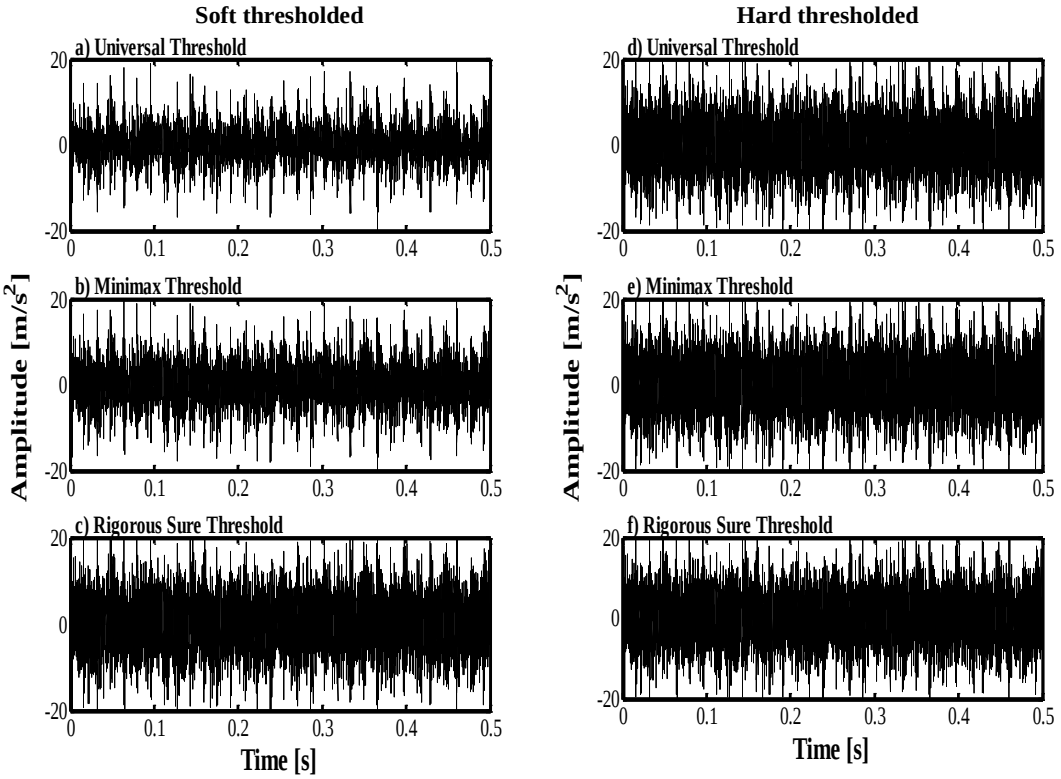

Figure 3.14: De-noised signals of the simulated vibration signal in Figure 3.6e based on SGWT and soft (left) and hard (right) thresholding methods.

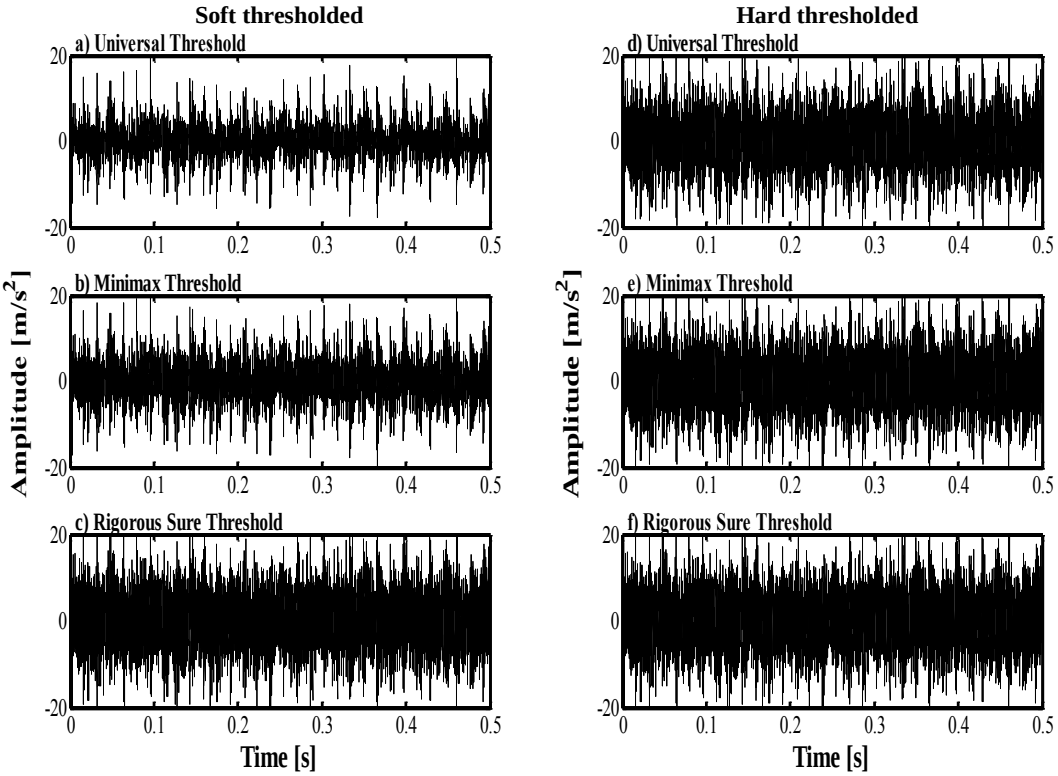

Figure 3.15: De-noised signals of the simulated vibration signal in Figure 3.6e based on RDWT and soft (left) and hard (right) thresholding methods.

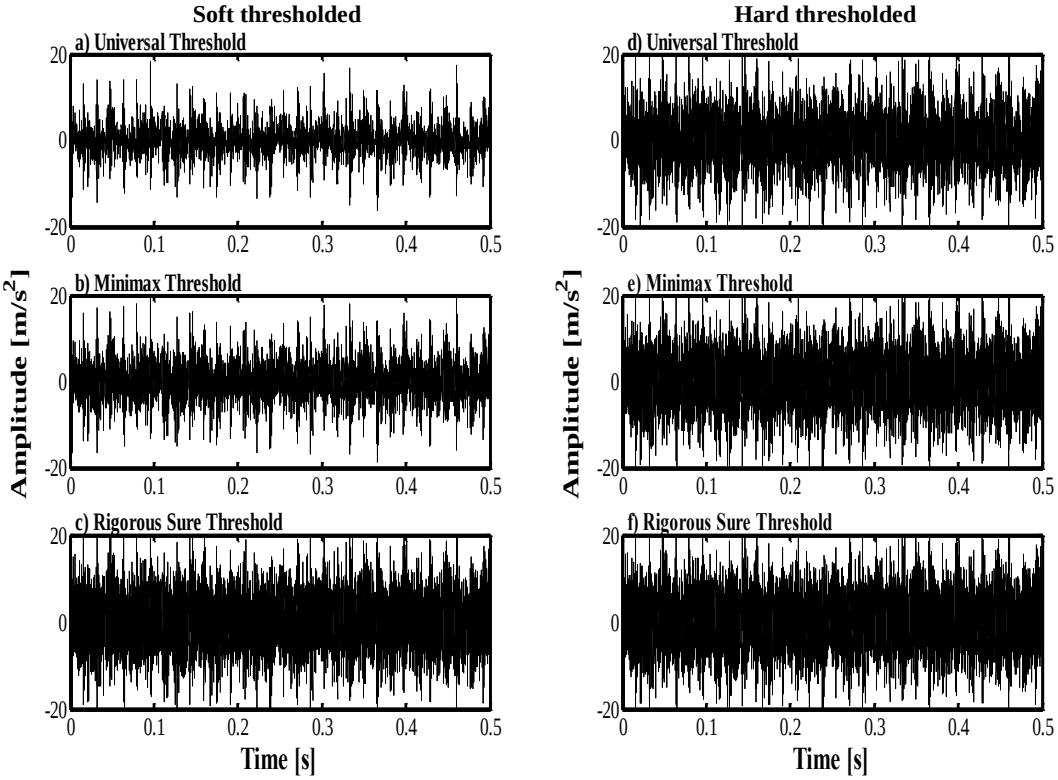

Figure 3.16: De-noised signals of the simulated vibration signal in Figure 3.6e based on DTCWT and soft (left) and hard (right) thresholding methods.

According to the obtained kurtosis values presented in Table 3.1, the soft thresholding method yields superior signal enhancement and noise reduction effects in time domain analyses compared to hard thresholding method. On the other hand, among the three studied thresholds, the universal threshold reveals the impulsiveness feature of the noisy simulated vibration signal better than the other thresholds. It is also worth mentioning that the DTCWT results in the best noise canceling effect compared to the other Wavelet Transforms.

Table 3.1: Kurtosis value of wavelet-based de-noised simulated vibration signals.

	Universal Threshold		Minimax Threshold		Rigorous Sure Threshold	
	Soft thresholding	Hard thresholding	Soft thresholding	Hard thresholding	Soft thresholding	Hard thresholding
DWT	5.57	3.77	4.67	3.58	3.56	3.50
SGWT	5.58	3.77	4.62	3.58	3.57	3.51
RDWT	6.12	4.01	4.98	3.68	3.56	3.51
DTCWT	6.50	4.48	5.26	3.82	3.61	3.52

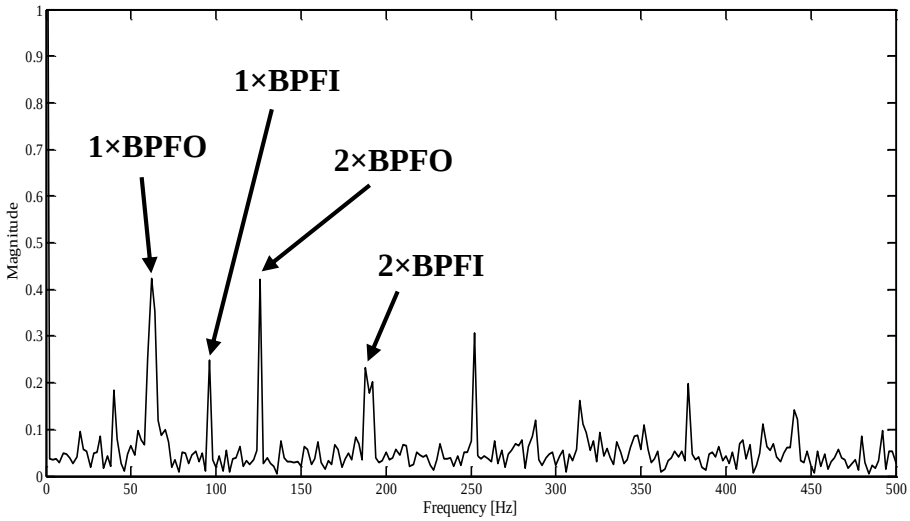

Figure 3.17: Envelope spectrum of the de-noised simulated vibration signal in Figure 3.16a.

Nevertheless, as stated earlier, the higher kurtosis value in time domain signals may not lead to the better and more readable spectrums in frequency domain analyses and the results may not provide valuable information for the purpose of fault diagnosis.

Figure 3.17 and Figure 3.18 show the envelope spectrums of the de-noised simulated vibration signals based on DTCWT both for soft and hard thresholding methods, respectively. As seen, both of the envelope spectrums can detect both of the fault frequencies (BPFI and BPFO) clearly, but the envelope spectrum of the hard thresholded signal is sharper than the soft thresholded one.

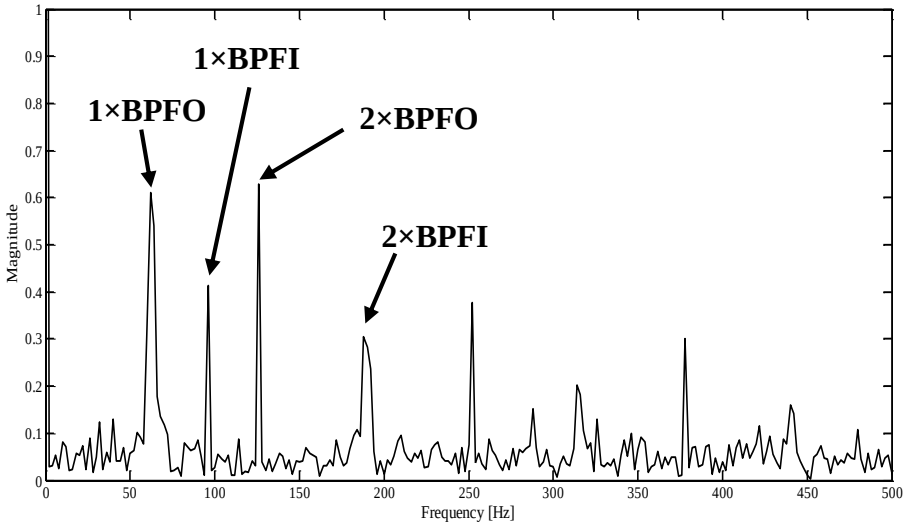

Figure 3.18: Envelope spectrum of the de-noised simulated vibration signal in Figure 3.16d.

3.5 Application of Kurtogram

Kurtogram is a non-stationary signal processing tool that can reveal the frequency band with highest kurtosis value in the analyzed signal. This frequency range contains the natural frequency excited by impulses with specific time intervals. As mentioned before, since kurtogram can detect the best frequency range of interest for the fault detection purpose, it can be used for optimizing the Wiener filtering. This optimum filter is used to separate the impulsive part of the signal from background noise and/or unwanted components. In this section, case study 3 (simulated ball bearing vibration signal having multiple faults) is investigated based on the kurtogram, first as a non-stationary tool, then as an optimizer for the Wiener filtering. It should be noted that, here, the kurtogram referes to its fast version so-called fast kurtogram. Fast version of the kurtogram uses the filter bank similar to mirror filters in DWT to decompose the signal into sub-bands, then obtain the kurtosis for each sub-band. Again, as seen in Figure 3.19, the optimum decomposition level is 4 for this case study. According to

the obtained results in Figure 3.19, the natural frequency of the analyzed signal–which was 4500 Hz–can be successfully detected in kurtogram.

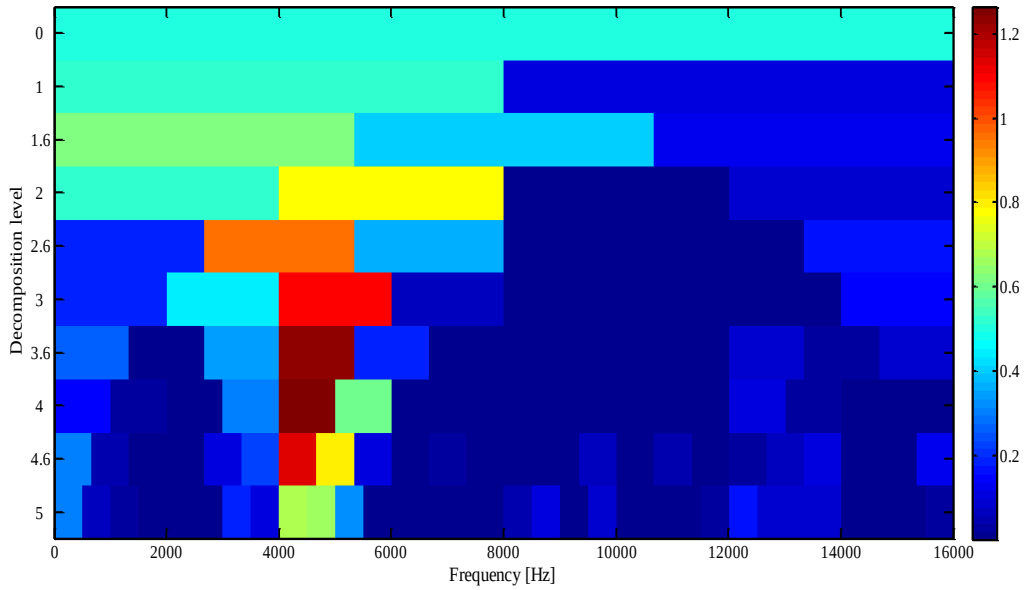

Figure 3.19: Kurtogram of the vibration signal in Figure 3.6e.

The kurtogram presented in Figure 3.19 decomposes the signal into two frequency bands in the first level, four frequency bands in the second level, eight frequency bands in the third level and so on. Since the power of added white noise in numerical signals is known, the best Wiener filter coefficients can be designed. On the other hand, the optimum Wiener filter obtained based on the kurtogram can also be achieved. Figure 3.20 shows the envelope spectrum obtained from de-noised simulated vibration signal related to the Wiener filter based on kurtogram.

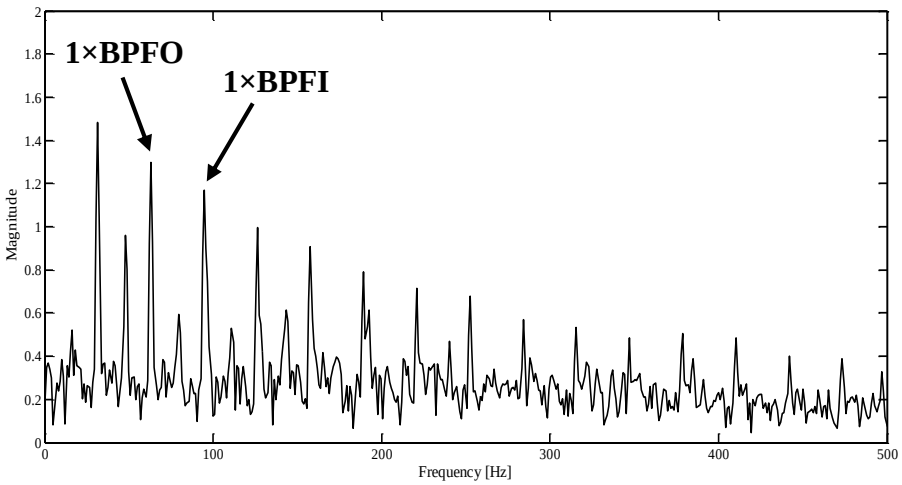

Figure 3.20: Envelope spectrum of the kurtogram based optimum Wiener filtered simulated vibration signal in Figure 3.6e.

Figure 3.21 shows the Wiener filtered signal, which designed based on the observed signal in Figure 3.6e and the added noise in Figure 3.6b. As can be seen, since the added noise is known, the magnitude of the noise components are close to zero in the filtered signal. Although the fault frequency magnitudes are higher in the envelope spectrum in Figure 3.20, the fault frequency magnitudes are clearer in Figure 3.21. To have better comparison between the two Wiener filter based de-noised envelope spectrums, the results in Figure 3.20 and Figure 3.21 are overlaid in Figure 3.22. The red line is related to the kurtogram based optimum Wiener, whereas the blue line is related to the best Wiener filter based on known added white noise.

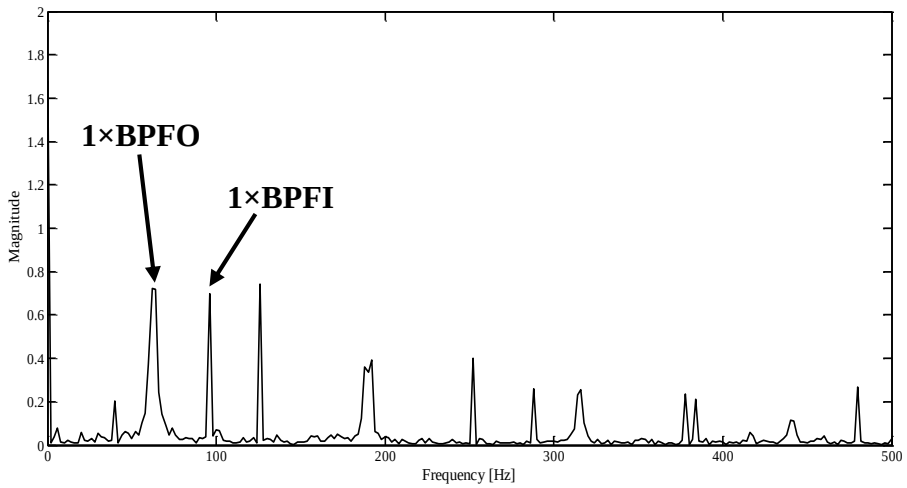

Figure 3.21: Envelope spectrum of the best Wiener filtered simulated vibration signal in Figure 3.6e.

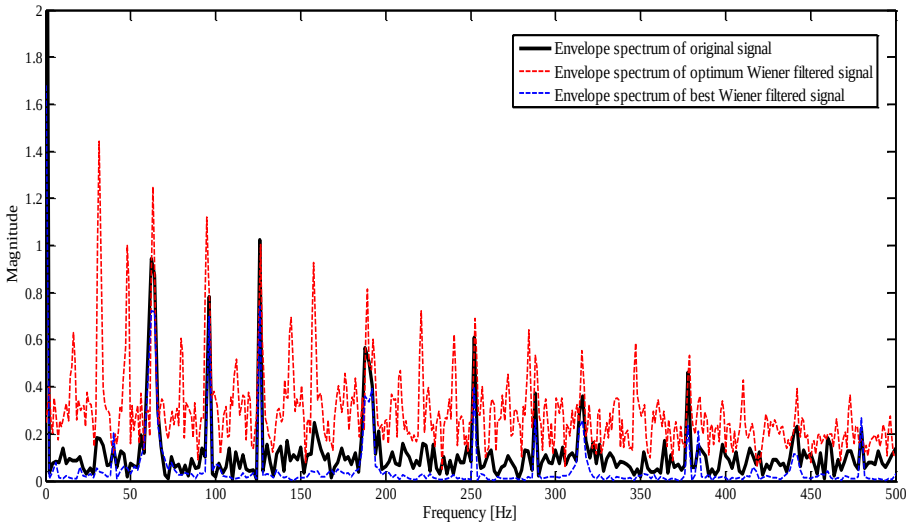

Figure 3.22: Envelope spectrums of de-noised simulated vibration signals related to the best and optimum Wiener filters.

3.6 Time-Frequency Analysis

Time-frequency analyses have played a very significant role in the fault detection of the rotating machines. The STFT and CWT methods, as the most common time-frequency analyses techniques in fault detection, are compared and studied numerically using simulated ball bearing vibration signal. The time-frequency techniques can be evaluated based on their ability to detect the repetition of the machine resonance excitations. Figure 3.23 illustrates the waterfall diagram obtained by STFT with hanning window and window lenght of 256 data for the simulated vibration signal in Figure 3.3d. According to this plot, the resonance frequency of the analysed simulated vibration signal occurs at about 4.5 kHz which can indicate the location of the impact train in frequency domain.

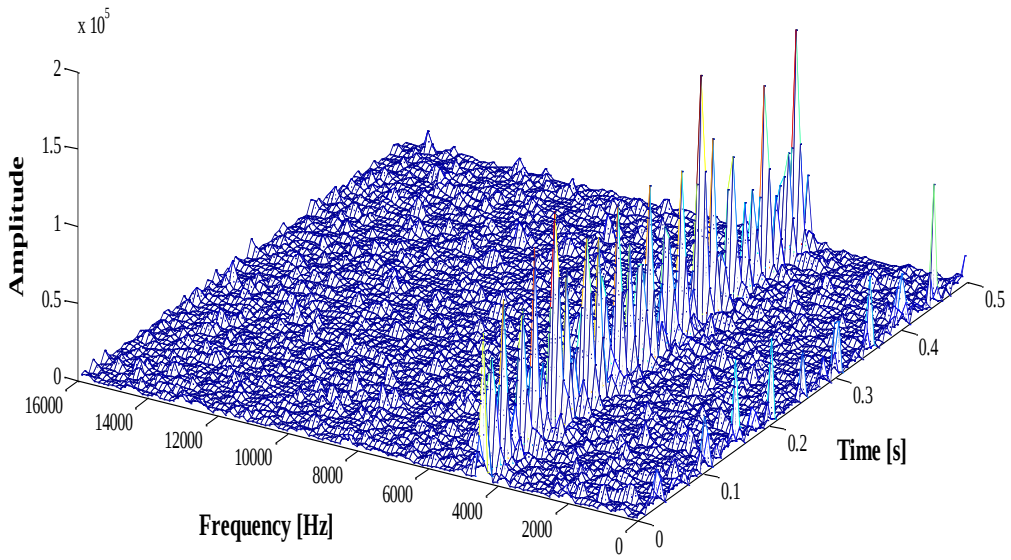

Figure 3.23: Waterfall diagram for the time-frequency representation of the simulated vibration signal in Figure 3.3d.

Figure 3.24, on the other hand, shows the time-frequency representation of the simulated vibration signal in case study 2 in Figure 3.3d obtained by CWT. The Morlet wavelet with a range of 1 to 256 scales is selected as the mother (main) wavelet basis, which looks a symmetric exponentially decaying coise type of sinusoidal function as shown in Figure 2.27. This exponentially decaying function is quite similar to the vibrations caused by impacts in faulty ball bearings. Consequently, the impacts can be magnified in each frequency scale by adjusting the exponential value (decay rate) in Morlet wavelet.

As seen in Figure 3.24, due to nature of the wavelets, the energy leakage at the location of impuls train from high to lower frequencies in frequency axis occurs. Nevertheless, CWT can successfully detect the occurrence time interval of the impacts, T_o, which is inverse of the fault repetition frequency, BPFO.

Figure 3.25 shows again the time-frequency representation, but time time obtained based on STFT sing hanning widow of size 256 data points for the simulated vibration signal in case study 2. After a series of tests, the optimum window size is obtained as 256 data points for the corresponding case study. The ball bearing resonance and its occurrence repetition can be detected via STFT in Figure 3.25. However, the so-called energy leakage problem in STFT is less than CWT.

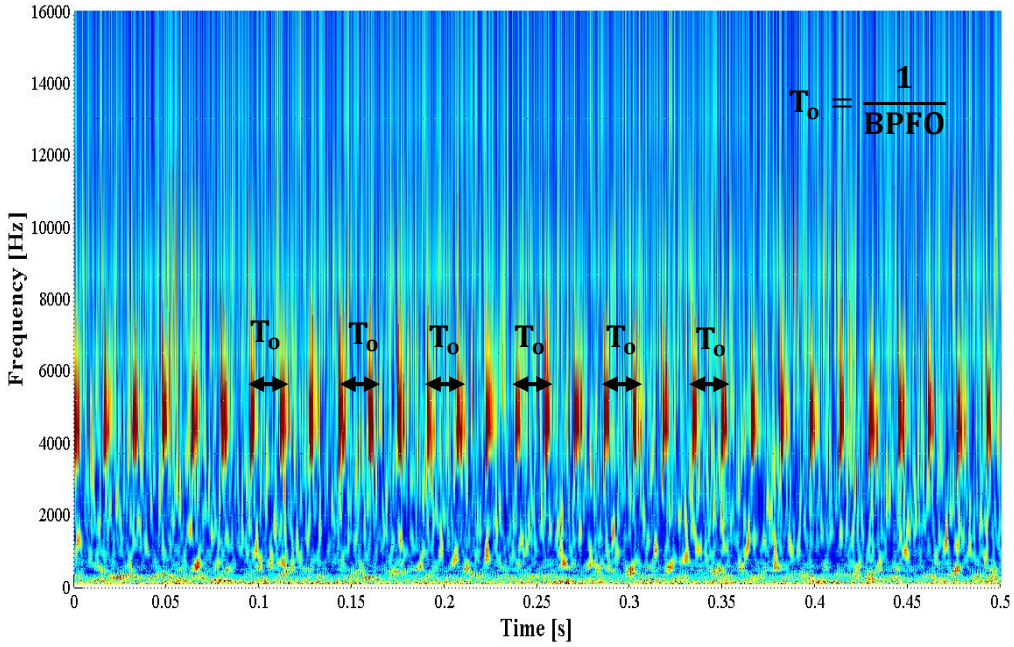

Figure 3.24: Time-frequency representation of the simulated vibration signal in Figure 3.3d based on CWT.

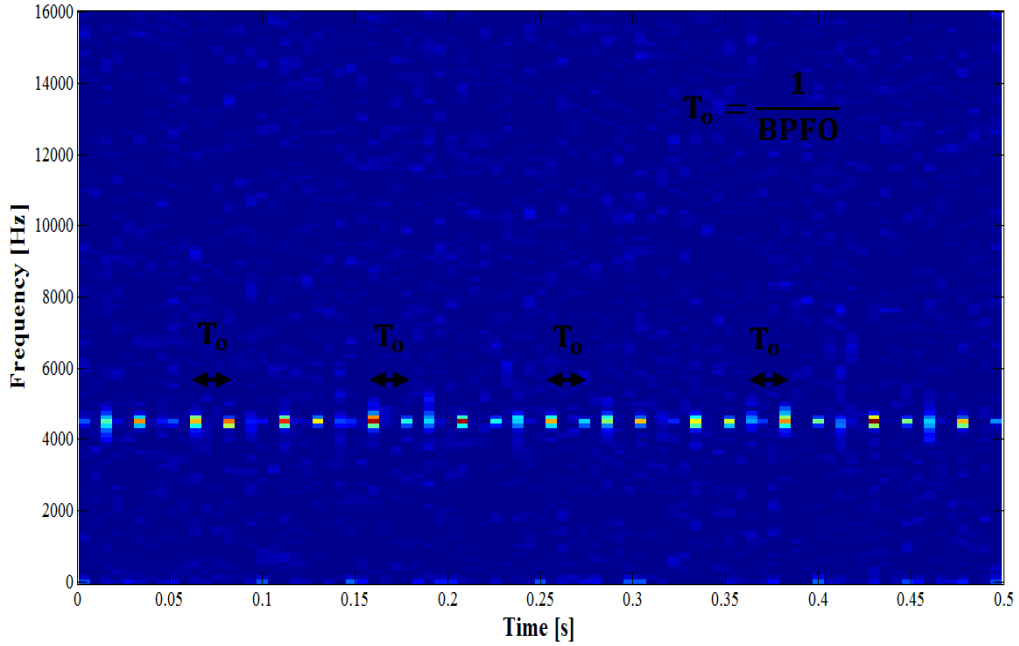

Figure 3.25: Time-frequency representation of the simulated vibration signal in Figure 3.3d based on STFT.

3.7 Summary

In this chapter, some advanced signal processing methods described in the previous chapter are studied numerically using the simulated vibration signals for various case studies. Three vibration signals for different ball bearing conditions are generated in MATLAB® platform. EMD and its revised version, EEMD are investigated for noise elimination applications. However, EMD shows superior noise canceling properties for the analyzed case study, but it suffers from mode-mixing problem, which can be adversely affected the further processes in ball bearing fault detections. Wavelet-based de-noisig algorithms with different threshold values and thresholding methods are also examined and the noise elimination results are presented and compared in term of kurtosis of the de-noised output signals. The application of the spectral kurtosis based on STFT as an optimizer of the Wiener filtering to reveal the impulsiveness properties of the analyzed vibration signal, are also shown. In addition, STFT and CWT as the time-frequency signal processing methods are employed on the simulated vibration signals to detect the repetition frequency of the impacts.

4. EXPERIMENTAL STUDY: DESCRIPTION OF THE TEST RIG

4.1 Introduction

Correct interpretation of the measured signals requires a good understanding of the test equipments, data acquisition system and professional experience. This chapter describes the used test rig and its components. In addition, the parameters to be measured for vibration-based ball bearing fault detection are discussed and details of the relevant transducer and data acquisition system are given. Also, the defects which are artificially created on the sample ball bearings are described in detail.

4.2 Machinery Diagnosis System PT500

An important part of a CM research is introducing of probable machinery faults in a controlled manner into a practical system, which can provide real data similar to ones might be recorded in an industrial situation. The environment for this research work was provided by Istanbul Technical University (ITU), Mechanical Engineering Department, Vibration&Acoustic Laboratory. The test setup named Machinery Diagnosis System, PT500, is made by GUNT Co., having a rotating shaft, two accelerometers, an electric motor, two loading disks, two supporting ball bearings, Pulley/V-belt mechanism, USB box, an amplifier, and an analyzer. Figure 4.1 and Figure 4.2 show the test machine, PT500. The speed of the electric motor is adjustable by a controller so the shaft can be run at different rotating speeds. There are two bearing housings to support the main rotating shaft and one bearing housing for secondary rotating shaft related to Pulley/V-belt mechanism. The housing nearer to the electric motor contains a single row and self-aligning healthy ball bearing type NU204-E-TVP2 and the other one contains a single row and self-aligning faulty ball bearing type ORS6204. Loading disks have a total weight of 2×1675 gr. The pre-tension of the V-belt can be adjusted using a sliding bearing located under the V-belt and the corresponding pre-tension value was set to 130 N for all the tests. Consequently, the variation of loading on the bearing was not studied in the experiments. It should be

noted that the loading disks and V-belt mechanism are employed in order to apply radial load.

Figure 4.1: Machinery Diagnosis System, PT500.

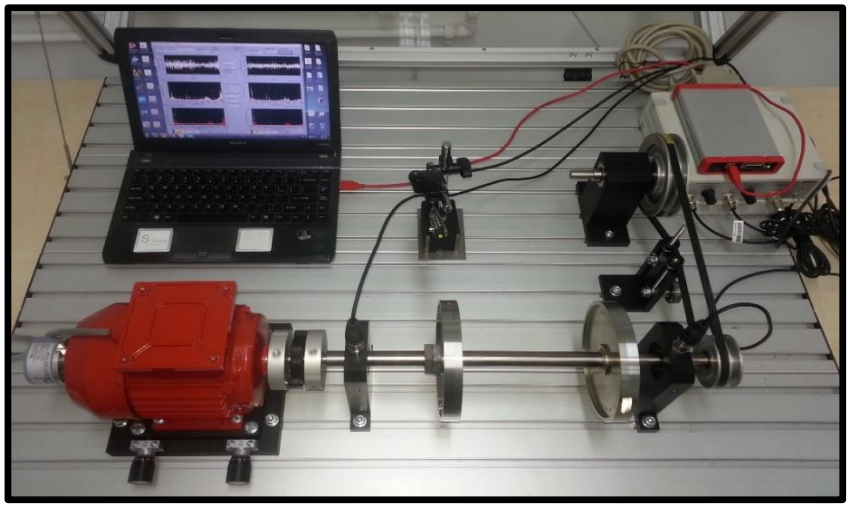

Figure 4.2: Ball bearing fault detection test rig.

4.3 Vibration Transducer

An Integrated Circuit Piezoelectric (ICP®) sensor is a device which can be designed to measure dynamic pressure, force, strain or acceleration. It contains a sensing element made of a piezoelectric material which converts mechanical strain into an electrical signal and an electronic circuit to amplify this signal and transmit it to an external device. The Machinery Diagnosis System, PT500, uses the ICP® type accelerometers, namely IMI603C01, mounted on both supporting bearing housing to collect the

vibration data. The basic construction of a piezoelectric accelerometer is shown schematically in Figure 4.3. When the accelerometer is subjected to vibration, the so-called seismic mass−depending on the severity of the oscillation−applies a variable force to the piezoelectric disks and consequently, a variable potential can be developed across the piezoelectric element. The potential is directly proportional to the acceleration over a particular frequency range. The piezoelectric material generates a charge when subjected to a stress.

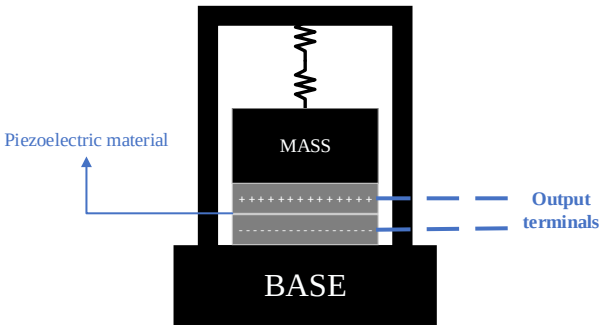

Figure 4.3: Schematic drawing of a piezoeletronic accelerometer.

Important properties of piezoelectric material of an accelerometer are piezoelectric constant, stiffness, dielectric constant, resistance and curie point. The sensitivity of an accelerometer is then defined as the ratio of its electrical output to the mechanical input. The natural frequency of an accelerometer depends on the its mass and stiffness properties as well as the stiffness of the mounting method. The corresponding natural frequency determines the maximum frequency range that an accelerometer can be used. The IMI603C01 accelerometer and its natural frequencies based on the various mounting methods are shown in Figure 4.4.

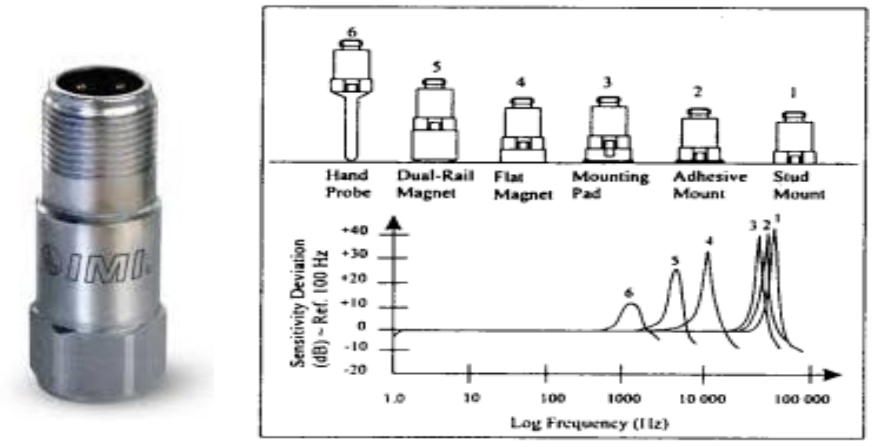

Figure 4.4: Illustration of the ICP type IMI603C01 accelerometer.

During the tests, a stud was used for mounting the accelerometer on bearing housing in order to have a wider frequency range. It should also be noted that the ICP® accelerometers are connected to the amplifier in Machinery Diagnosis System, PT500.

4.4 Tested Ball Bearings

As stated before, the ball bearing chosen for the present study was ORS6204 type ball bearing, which is a typical bearing widely used in industry and, using this type of bearing, the localized bearing faults can be created easily as required. Figure 4.5 shows a typical ball bearing used in the tests and Table 4.1 lists the geometric details of the ORS6204 type ball bearing.

Figure 4.5: A typical ORS6204 type ball bearing.

Table 4.1: Geometry details of ORS6204 ball bearing.

Outer Race Diameter	Inner Race Diameter	Number of Balls	Contact Angle	Bearing Width	Balls Diameter
47 [mm]	20 [mm]	8	0 °	14 [mm]	7 [mm]

For each tested ball bearing, a drill was used to produce the artificial fault (pit) on its raceways (outer and/or inner raceways). Starting from the undamaged condition, a transversal line on the contact area between outer race and balls was created and extended step by step to cover a square or nearly square zone, which can be considered as a typical localized fault on a ball bearing at its final stage. Seven ball bearings–six bearings with single localized fault and one with multiple faults–were damaged with a

drill to create different fault sizes, listed in Table 4.2. The localized faults on the outer and inner raceways of ball bearings are schematically shown in Figure 4.6.

Table 4.2: Defective ball bearings fault size details.

	Small	Medium	Large
Outer race	2[mm] × 2[mm]	4[mm] × 4[mm]	7[mm] × 6[mm]
Inner race	2[mm] × 1[mm]	3[mm] × 3[mm]	6[mm] × 5[mm]
Outer + Inner races	Outer race: 4[mm] × 4[mm]; inner races: 3[mm] × 3[mm]		

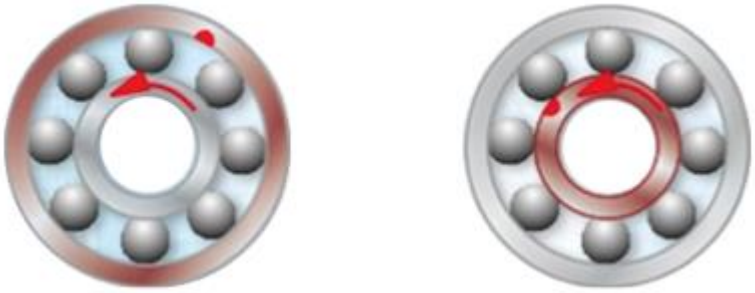

Figure 4.6: Outer race (left) and inner race (right) defective ball bearings.

4.5 Experimental Data Collection Procedure

All the experiments were conducted in a laboratory with climate control system; therefore, it can be assumed that all tests were carried out under the same condition. The ball bearing test rig was allowed to run for 2 minutes to warm up before starting the data collection. Seven faulty ball bearings having artificially added local faults and one healthy ball bearing at two different rotating speeds−900 and 1200 RPM−were used to study only one type of fault condition. Note that, all the ball bearings assumed to have no clearance and no slippage error. Experiments were conducted to study bearing fault detection and diagnosis so that the vibration signals were collected via the mounted accelerometer described earlier. In the measurements, the vibration transducer is positioned to measure the vertical accelerations of the bearing housing.

Figure 4.7 illustrates the fault detection mechanism used in PT500 system where a matched flexible coupling, couple the motor to the main rotating shaft which have two supporting ball bearings. The radial load from the loading disks and Pulley/V-belt mechanism are applied to simulate real conditions in industry and to be sure that the ball bearings are permissible by the rig tolerance. The healthy ball bearing located nearer the electric motor and the healthy supporting ball bearing on the secondary

rotating shaft were fixed for all the tests and were not subjected to the experiments. Each faulty ball bearing was employed on the specified location in Figure 4.7 in order to collect the vibration data.

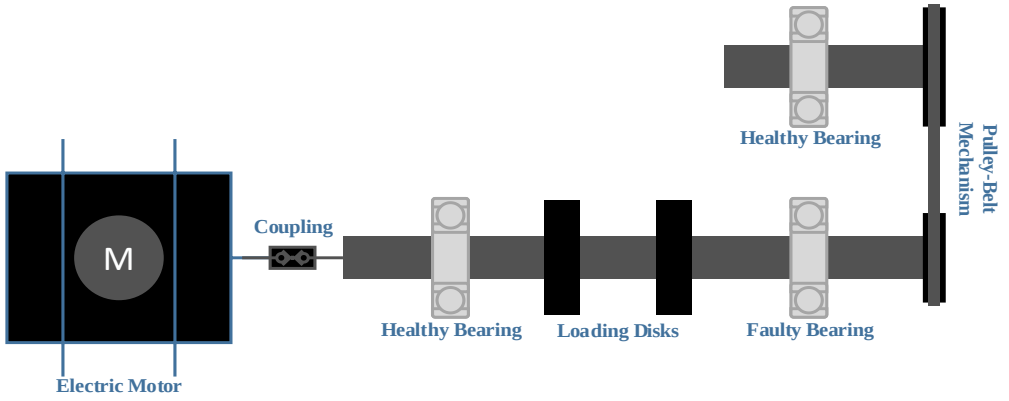

Figure 4.7: Ball bearing fault detection system used in PT500.

Two sets of data were collected for individual outer race defective ball bearings: One set of data was taken so that the position of the defective area on the outer race of the ball bearing was about 60° from the vertical axis along which the transducer was positioned and other set was collected when the defective area of the outer ring was located towards the top of the bearing housing, i.e. along the vertical direction.

The sampling rate for all the tests was 32 kHz and, hence 16000 data points were acquired during 0.5 seconds. The sampling rate of 32 kHz means the upper frequency limit of the spectrum was about over 12.5 kHz; this was about 130 times more than the expected highest fundamental fault frequencies, ensuring that all harmonics of interest would be included in the spectrum. Once the system is ready, making the measurements does not take long time; the time-consuming part of the experiments was replacing bearings in the system. Swopping a bearing requires the following actions: First, the Pulley/V-belt mechanism is needed to be dismantled and the shaft taken out. Then, the first tested bearing is taken out and the next bearing to be tested is assemble into the bearing housing. Following this, the Pulley/V-belt mechanism is re-assembled and then 130 N as the pre-tension is adjusted using the sliding bearing under the V-belt. This procedure has to be repeated every time when a ball bearing is to be tested. The order of the experimental procedure was as follows: (i) test the healthy bearing, (ii) test the bearings with outer race faults, (iii) test the bearing with inner race fault and finally (iv) test the bearing with multiple faults (outer and inner raceways).

5. CONDITION MONITORING OF BALL BEARINGS - EXERIMENTAL RESULTS

5.1 Introduction

After the successful simulations of some fault detection methods, as the most crucial part of a CM research work, the fault detection techniques need to be validated experimentally. Experimental investigations can reveal the practical limitations and benefits of the diagnosis methods. Accordingly, the present chapter presents the experimental results for the measured ball bearing vibration signals. In the previous chapter, a brief description of the test rig and its components used for the tests were provided. The mentioned ball bearing sets included a single and multiple faults and these faulty ball bearings are used in this chapter for experimental investigations. This chapter emphasizes further processes that can be applied to raw vibrations signals to have better understanding of ball bearing faults. In line with this purpose, various ball bearings are used to illustrate the importance of the signal enhancing, noise elimination, and sub-bands analyses for the proper fault detections.

5.2 Ensemble Empirical Mode Decomposition

In chapter 2, the theoretical foundations of EMD and its revised version, so-called EEMD were investigated in some detail. In addition, numerical analyses of EMD and EEMD and their comparisons were carried out in chapter 3 using simulated ball bearing vibration signals. In this section, EEMD as a better alternative to EMD is applied to real vibration signals captured from single (inner race) defective and multiple (both outer and inner races) defective ball bearings. Starting with the ball bearing having single fault and then multiple faults, EEMD method is investigated experimentally. Figure 5.1 shows the vibration signal with kurtosis = 5.38 and its amplitude spectrum, corresponding to inner race defective ball bearing with medium fault size, the rotational speed of the shaft being 1200 RPM. It should be noted that the corresponding fault sizes related to each ball bearing are available in Table 4.2. As can be seen, the resonances of the system are excited approximately around 1000 Hz and

7500 Hz. According to Equation (2.38), the fundamental fault frequency of the inner race defects for the tested ball bearing at 1200 RPM is estimated as 96 Hz.

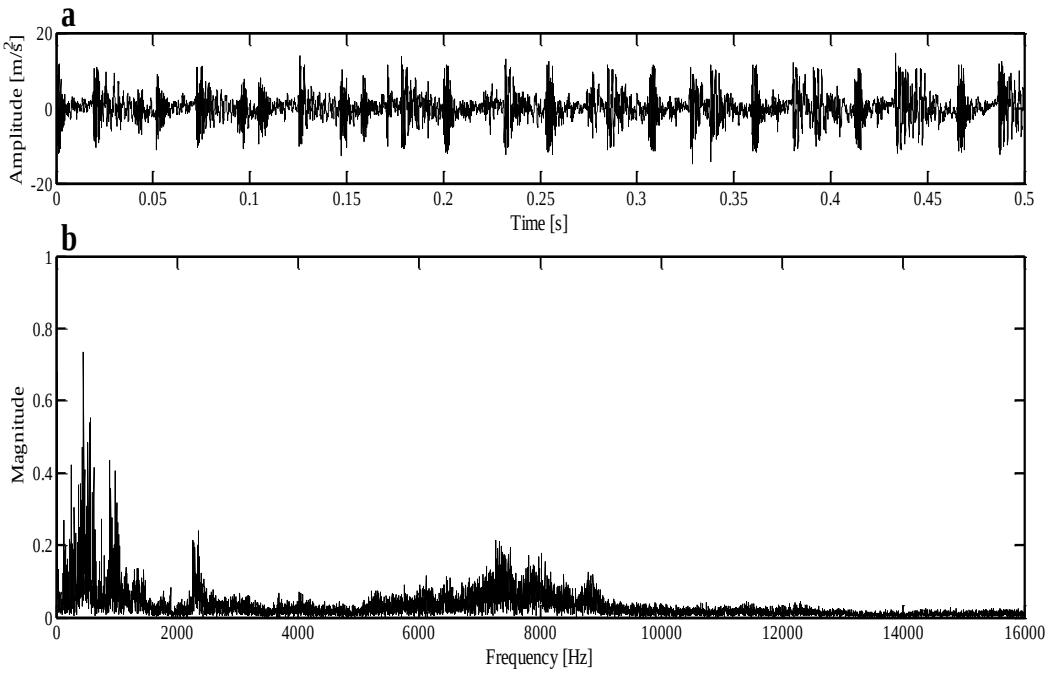

Figure 5.1: a) Captured vibration signal from an inner race defective ball bearing with medium fault rotates at 1200 RPM b) the corresponding amplitude spectrum.

The Gaussian white noise having standard deviation equal to half of the standard deviation of the analyzed signal is used in EEMD process. In chapter 3, it was pointed out that an ensemble number of about 100 is sufficient to yield reliable results. However, to be on the safe side, 200 ensembles are used to produce the IMFs based on EEMD. Figure 5.2 shows the decomposition results–first 10 IMFs–of the vibration signal presented in Figure 5.1a. According to the results in Figure 5.2, there is no so-called mode-mixing problem, which is a common issue in EMD. This means that each IMFs exclusively has only a specific range of frequencies. By considering the amplitudes of the fifth and the higher IMFs, it can be concluded that the only first four IMFs can be used for fault detection purposes and other IMFs may be considered as noise and/or unwanted components. In order to detect the fault frequencies, envelope analysis is applied to the first four IMFs, and the results are presented in Figure 5.3. As seen, all the envelope spectrums of the first to the fourth IMFs detect the corresponding fault frequency.

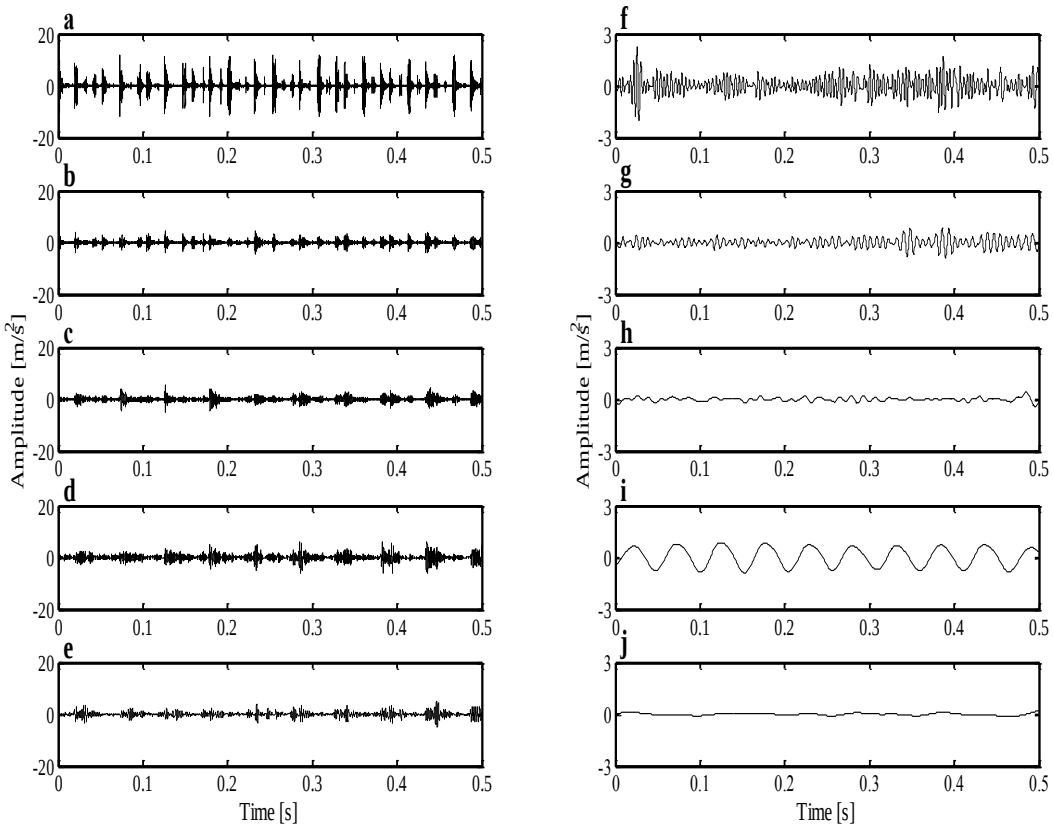

Figure 5.2: The first 10 IMFs of the vibration signal in Figure 5.1a based on EEMD, respectively, from a) to j).

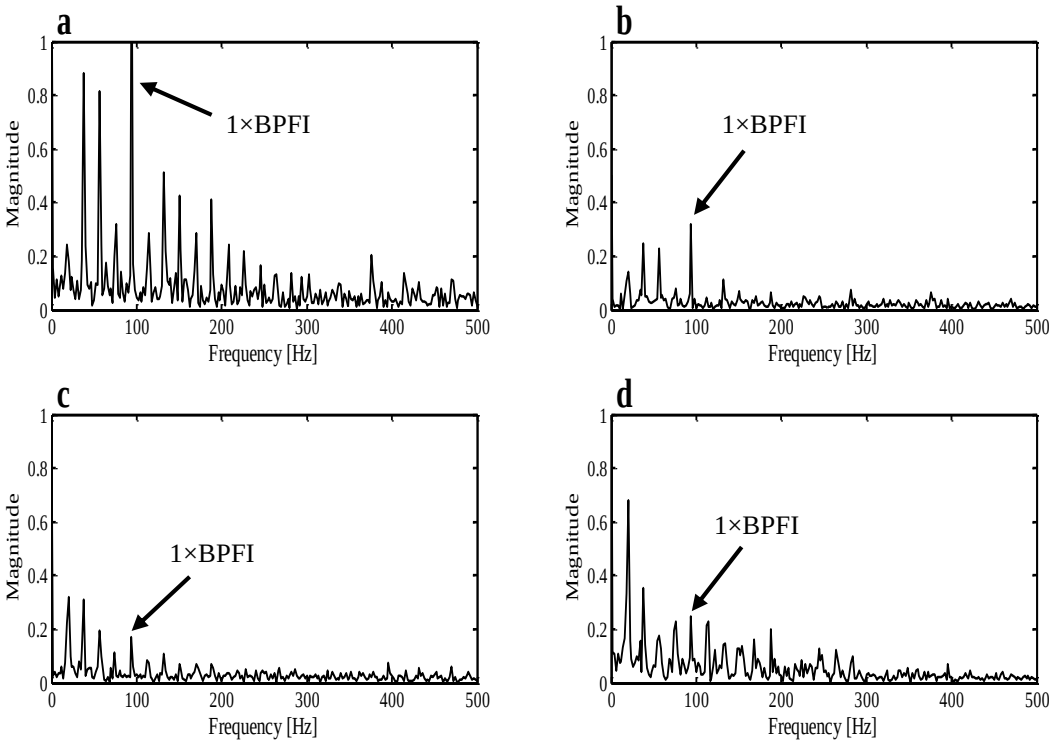

Figure 5.3: Envelope spectrum of a) first IMF b) second IMF c) third IMF d) fourth IMF in Figure 5.2.

It should be noted that, although the fault frequency in the envelope spectrum related to the first IMF has higher magnitude, the second IMF has the clearer envelope spectrum compared to other IMFs and the raw vibration signal.

Figure 5.4 shows the measured time domain vibration signal and its amplitude spectrum corresponding to a ball bearing with multiple faults (one fault on each raceway). In this case, the kurtosis vaue of the signal = 3.17, the rotaitonal speed was again 1200 RPM and the position of the defective area on the outer race was located along upward directon. The fault frequencies, BPFO and BPFI are determined, respectively, as 63 and 96 Hz. As expected, the resonances of the system excited by the faults have higher magnitudes compared to the single fault case study.

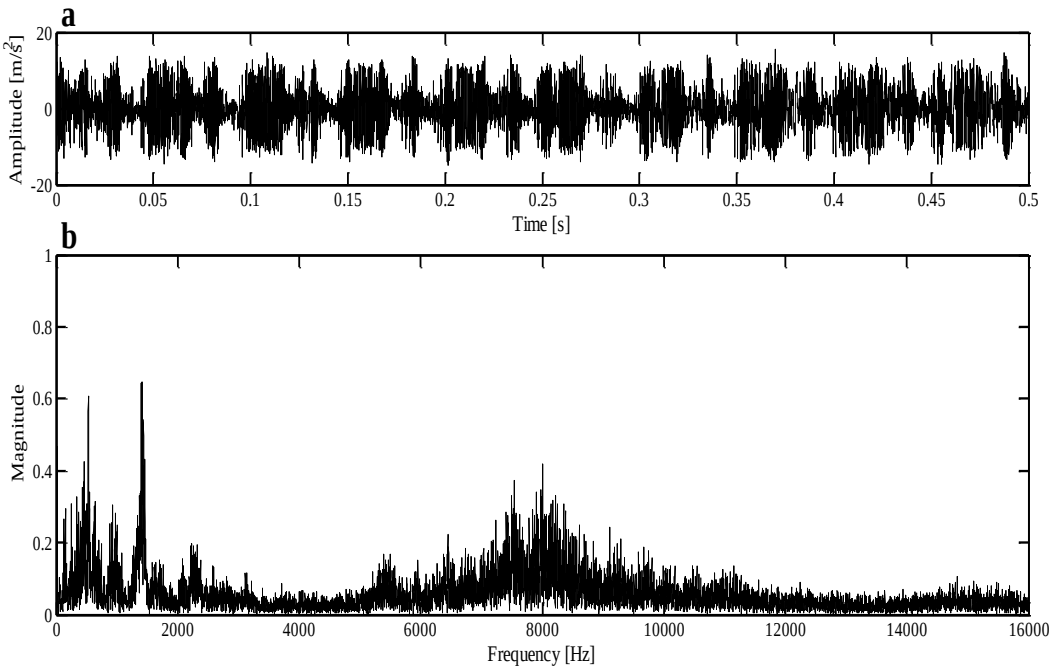

Figure 5.4: a) The captured vibratin signal from an outer and inner races defective ball bearing rotates at 1200 RPM b) the corresponding amplitude spectrum.

EEMD method is applied to the vibration signal in Figure 5.4a and the decomposition results obtained for the first 10 IMFs are presented in Figure 5.5. Again, no mode-mixing problem is seen in the IMFs. Similar to the previous case study, the fifth and higher IMFs are considered as the background noise and are not included in the post processings (i.e. envelope analysis). The envelope spectrums of the first four IMFs are presented in Figure 5.6. Generally, diagnosis of the inner race faults are more complicated than the outer race faults due to the rotation of the inner race. According to the envelope spectrums in Figure 5.6, the outer race fault manifests itself in the first envelope spectrum clearly.

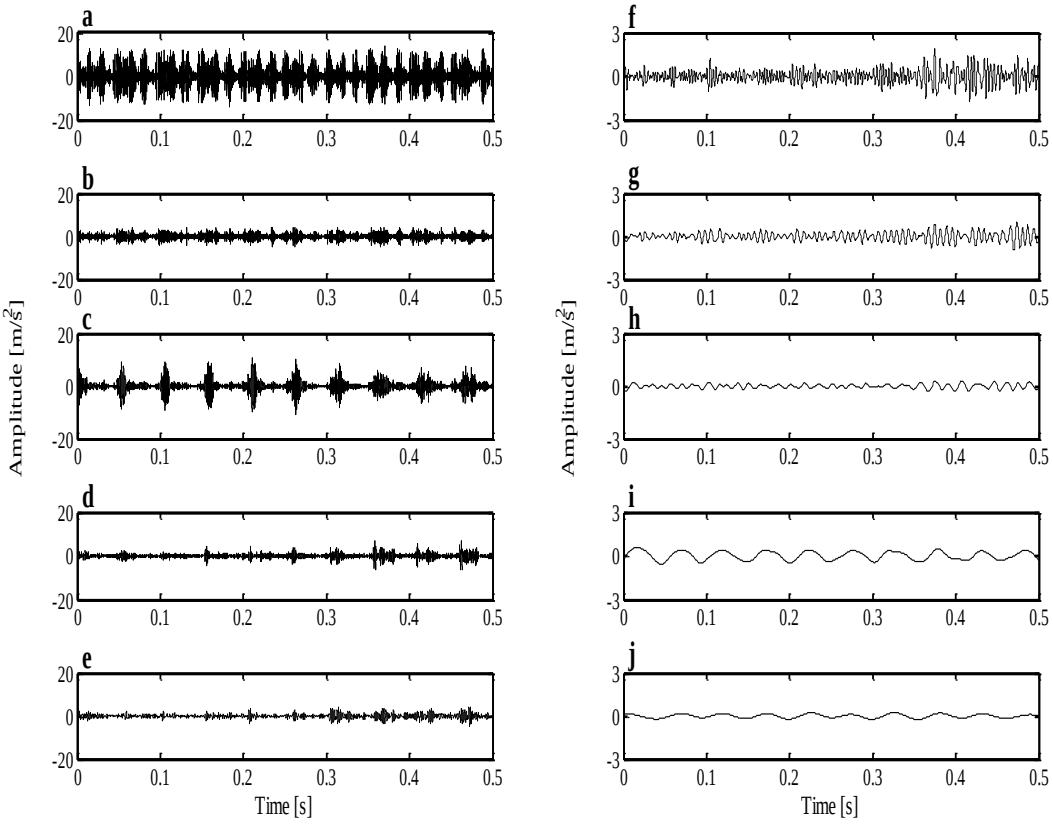

Figure 5.5: The first 10 IMFs of the vibration signal in Figure 5.4a based on EEMD, respectively, from a) to j).

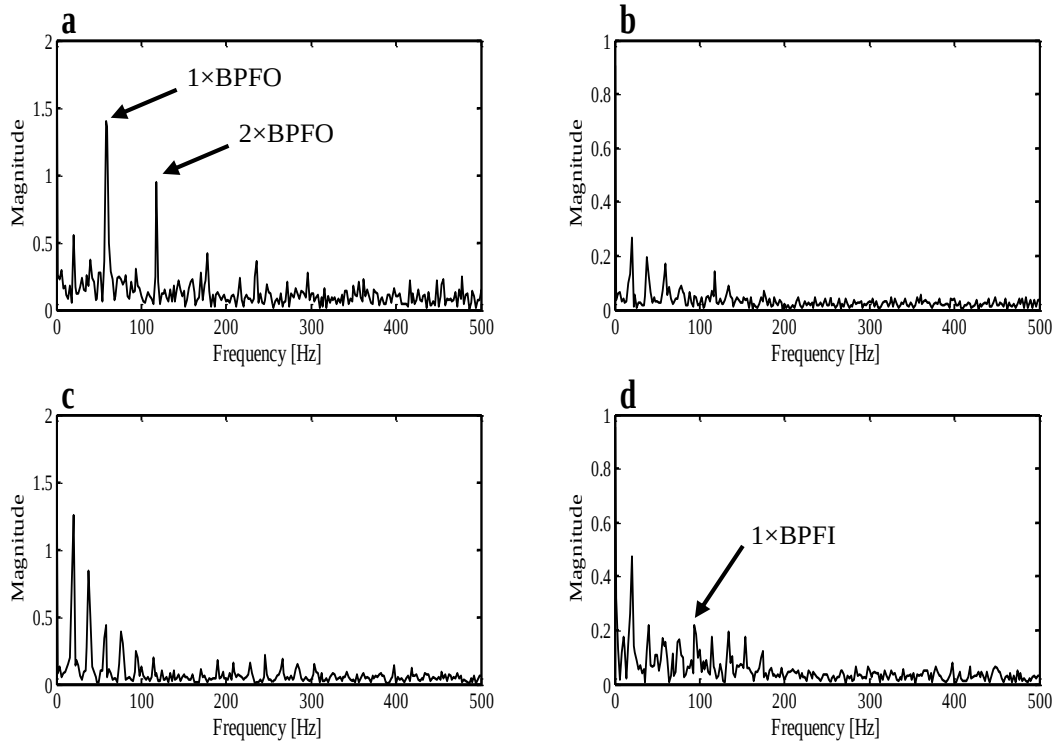

Figure 5.6: Envelope spectrum of a) first IMF b) second IMF c) third IMF d) fourth IMF in Figure 5.5.

However, only the envelope spectrum of the fourth IMF detected the fault frequency related to the inner race, BPFI, whereas almost all other envelope spectrums failed to identify the inner race fault. However, by considering the first and fourth envelope spectrums in Figure 5.6, it is detected that the tested ball bearing have both outer and inner races faults. The present case study demonstrates the importance of the signal sub-band analysis to have reliable fault detections.

5.3 ARMED Filtering

In this section, the so-called ARMED filtering which is a combination of AR model and the MED filterings is investigated experimentally based on captured vibration signals from two inner race defective ball bearings having small and medium fault sizes. The rotational speed of the rotor was 900 RPM. Figure 5.7 and Figure 5.8 show, respectively, the vibration signals of the ball bearing having small and medium faults and their amplitude spectrums up to Nyquist frequency.

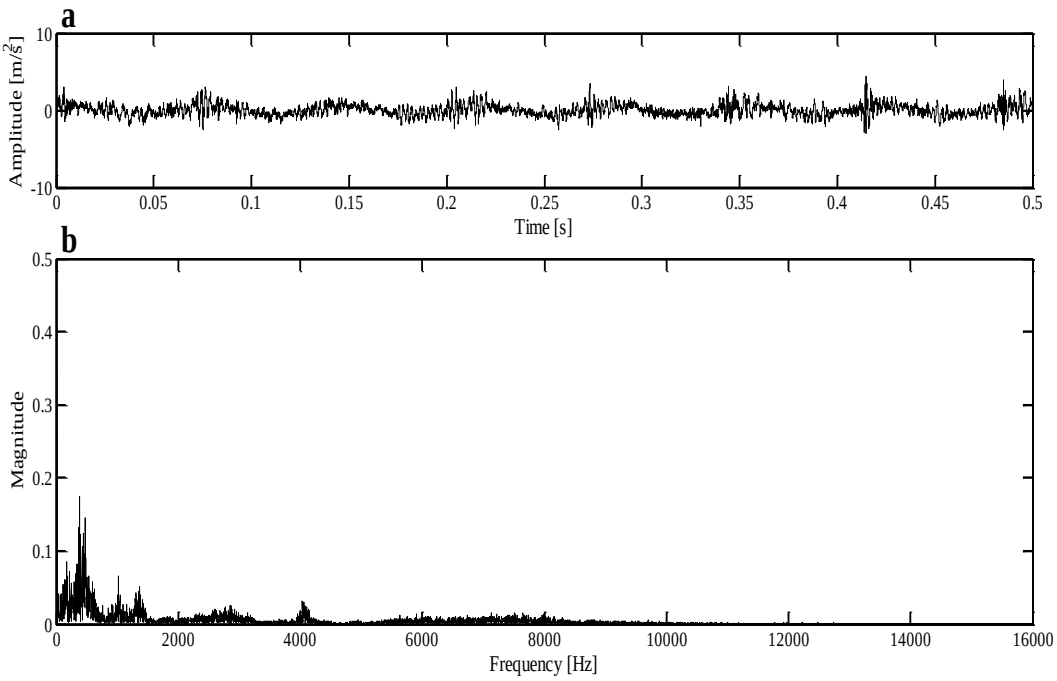

Figure 5.7: a) Captured vibratin signal from an inner race defective ball bearing with small fault rotates at 900 RPM b) the corresponding amplitude spectrum.

As seen, resonance amplitudes of the system with the ball bearing having medium fault have higher magnitudes than those of the bearing with small fault.

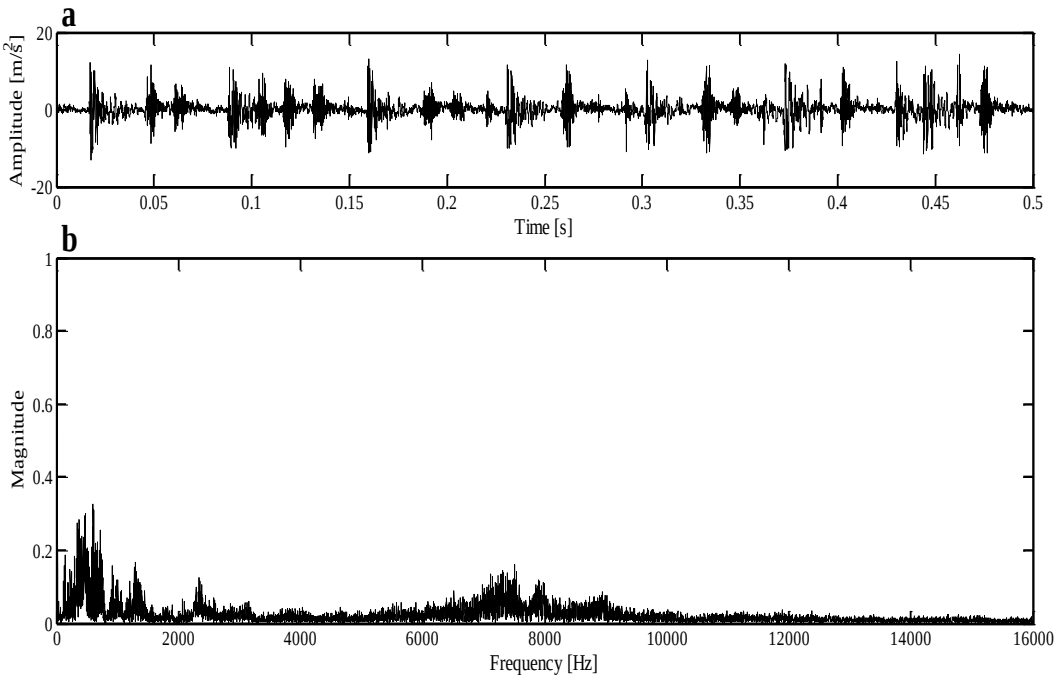

Figure 5.8: a) Captured vibratin signal from an inner race defective ball bearing with medium fault rotates at 900 RPM b) the corresponding amplitude spectrum.

Figure 5.9 show the envelope spectrums of the vibration signals in Figure 5.7a and Figure 5.8a, corresponding to small and medium faults, respectively. According to the obtained results, the envelope analysis is unable to detect the corresponding fault frequency for the ball bearing having small inner race fault. This is probably because of the small impacts, which can be covered by background noise.

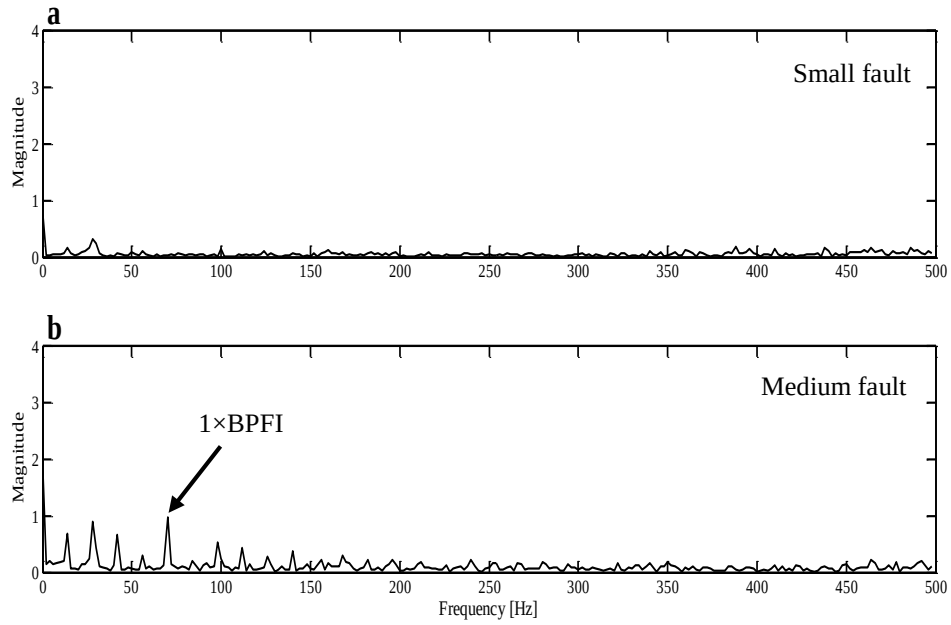

Figure 5.9: Envelope spectrums of the a) vibration signal in Figure 5.7a and b) Figure 5.8a.

First step in the ARMED filtering is AR model based linear prediction. The Yule-Walker method is employed for obtaininig the AR model coefficients. The optimum order in the linear prediction process is obtained by some iterations between Equations (2.9) and (2.24) and the results are presented in Figure 5.10, which shows the use of kurtosis as an indicator for selecting the optimum order, p, in Equation (2.24). The results indicate that the orders of p = 27 and p = 43, respectively, for the small and medium fault size case studies yield the highest kurtosis and consequently the best separation of the impulses.

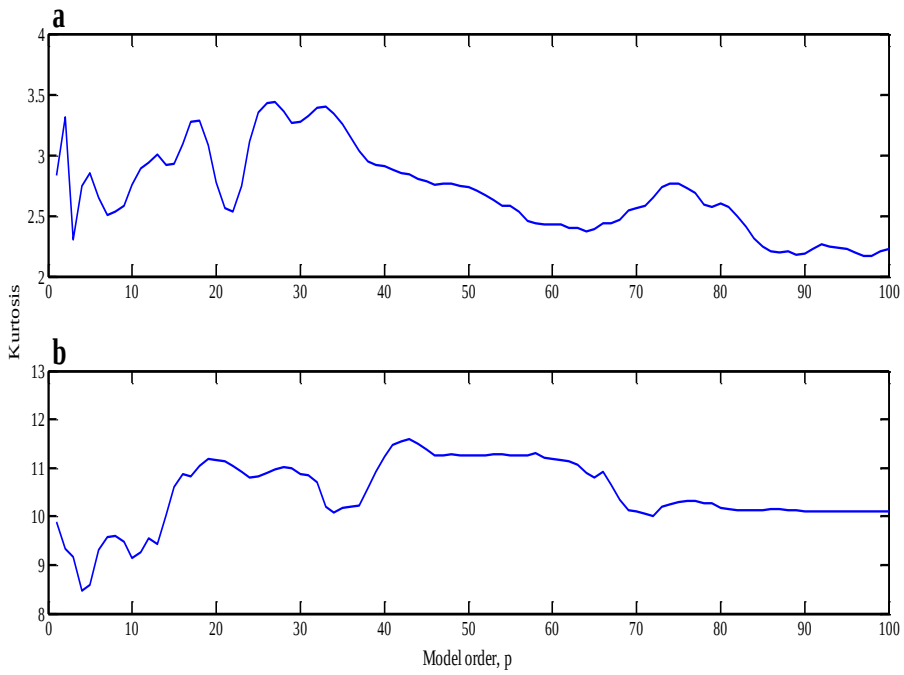

Figure 5.10: Kurtosis criterion for selecting the optimum order in AR model.

Envelope spectrums of the de-noised signals by AR linear prediction are presented in Figure 5.11. It is seen that, although the both spectrums are revealed the fault frequency, BPFI, the result for the small fault case is not clear enough.

The MED is the next step in ARMED filtering process to enhance the detection of fault frequencies and to increase the kurtosis values. For this, a series of filter lengths, L (50, 100, 200, 300, and 400), are investigated in order to obtain the best MED filtering performance. By comparing the envelope spectrums obtained by MED filter having different filter lengths, it is found that L = 100 offers the best de-noising performance for both case studies compared to other filter lengths.

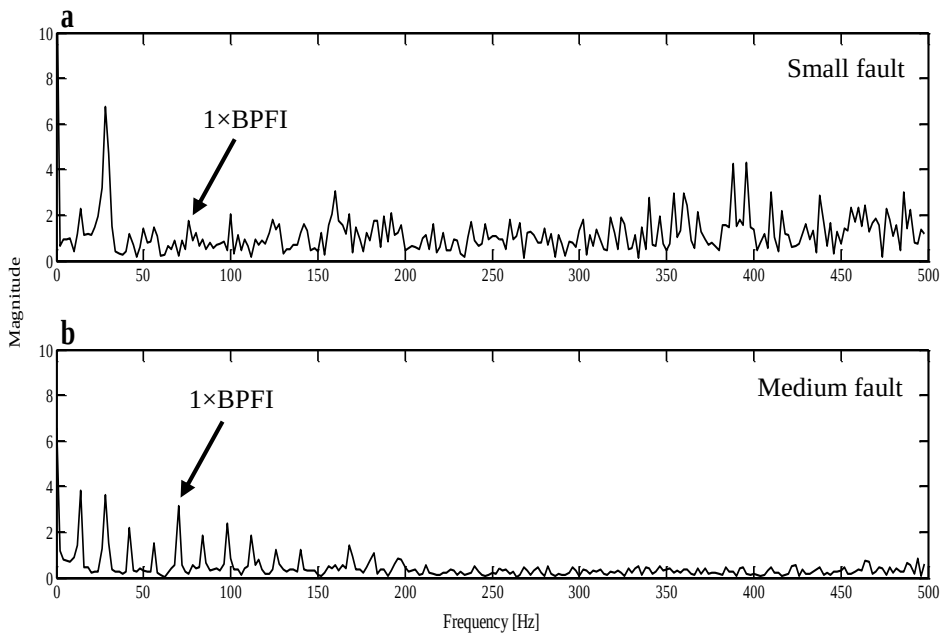

Figure 5.11: Envelope spectrums of AR linear prediction based de-noised vibration signals of a) small fault b) medium fault case studies.

Figure 5.12 shows the ratio of the kurtosis value of the output signals to the input signal for the different filter lengths and iterations used in MED filtering. Although more iterations increase the kurtosis of the output signals, the results do not provide additional benefits in terms of fault frequency detection. By doing a series of tests, it is concluded that for both case studies, three iterations provide sufficient noise elimination property.

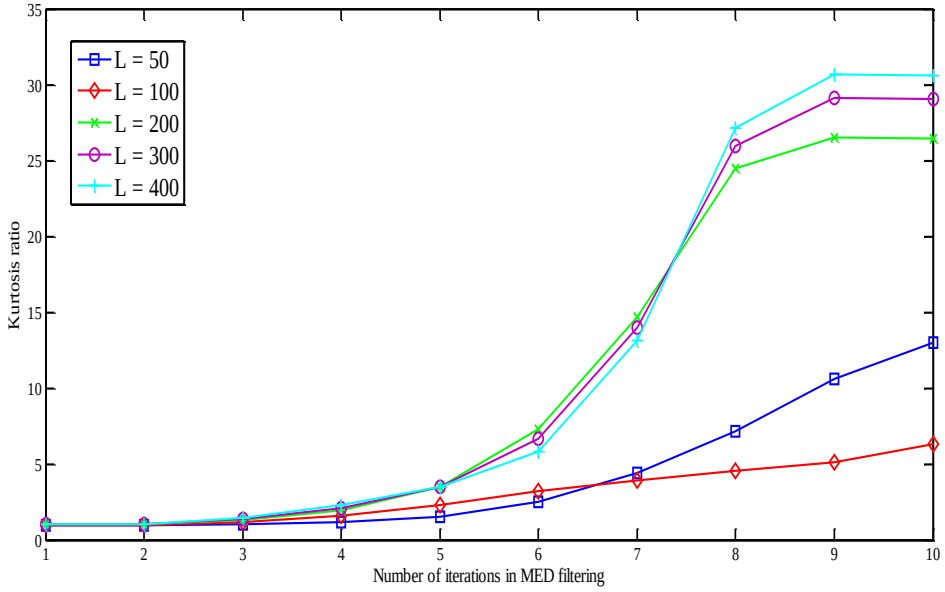

Figure 5.12: Convergence of the MED algorithm for the small fault case study.

Figure 5.13 shows the ARMED filtering based de-noised vibration signals of the ball bearings having small and medium faults and Figure 5.14 represents the envelope spectrums of the corresponding vibration signals.

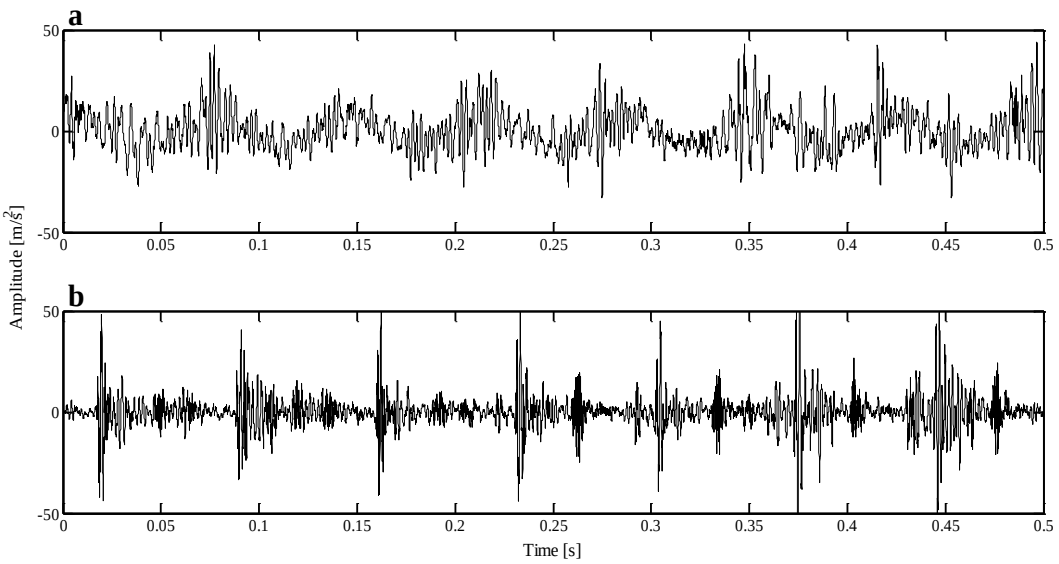

Figure 5.13: ARMED filtering based de-noised vibration signal of the a) small fault b) medium fault case studies.

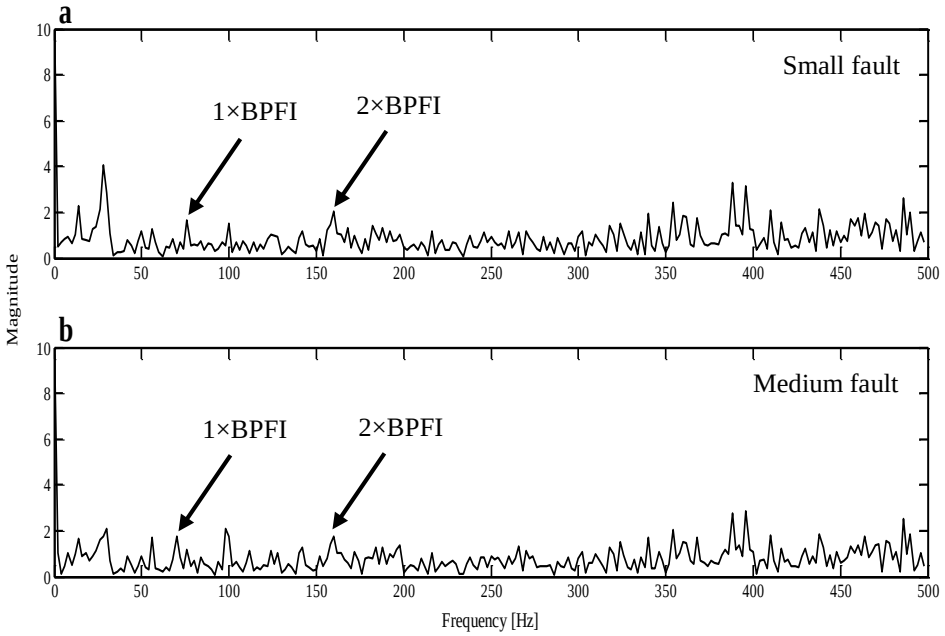

Figure 5.14: Envelope spectrums of the vibration signals in a) Figure 5.13a b) Figure 5.13b.

In spite of the fact that the overall amplitudes of the envelope spectrums of the ARMED based de-noised vibration signals are somewhat reduced in Figure 5.14, the envelope spectrums after de-noising are smoother than the previous envelope

spectrums related to the original and AR linear prediction based de-noised vibration signals, hence yielding a slightly better detection of the related fault frequency (BPFI).

5.4 SVD Based De-noising Method

In the present section, a very effective and useful signal de-noising and enhancing approach based on SVD is proposed and investigated using the measured ball bearing vibration signals. The outer race defective ball bearings with various fault are used in this section, the rotational speed of the rotor being 120 RPM. The position of the defective area on the outer race of a ball bearing was about 60° from the vertical axis along which the transducer is also positioned. The fundamental fault frequency of the corresponding ball bearings is determined to be 63 Hz.

The SVD based noise elimination method described in section 2.2.3.1 can be applied either to time domain or frequency domain signals for noise elimination purposes. The captured vibration signals and their amplitude spectrums are presented in Figure 5.15 and Figure 5.16.

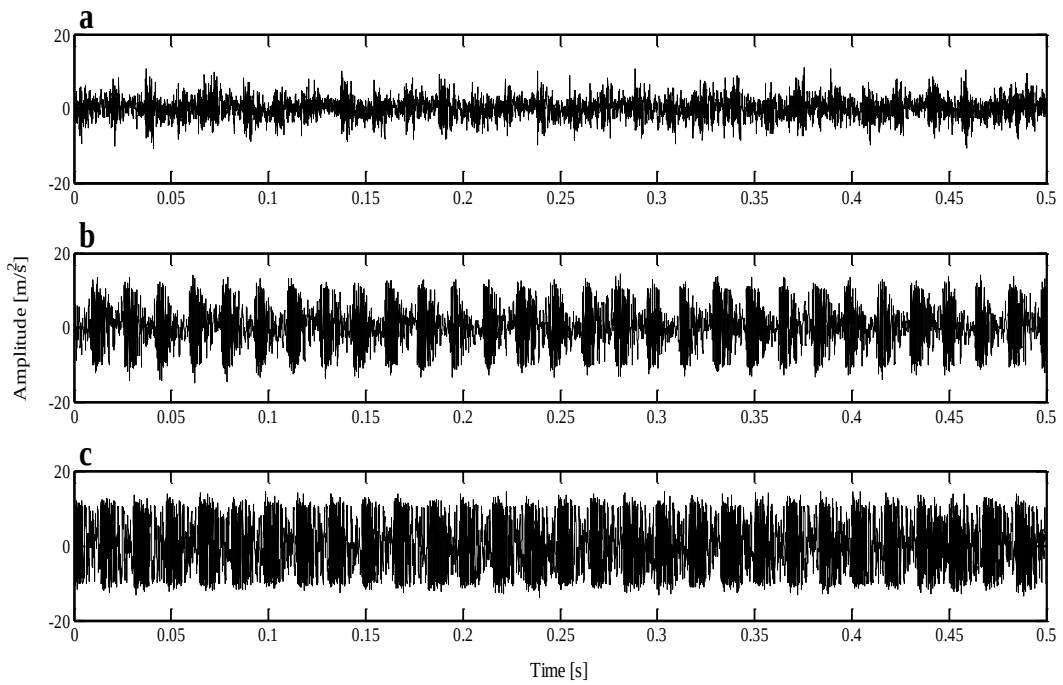

Figure 5.15: Vibration signals corresponding to outer race defective ball bearings with a) small fault b) medium fault c) large fault.

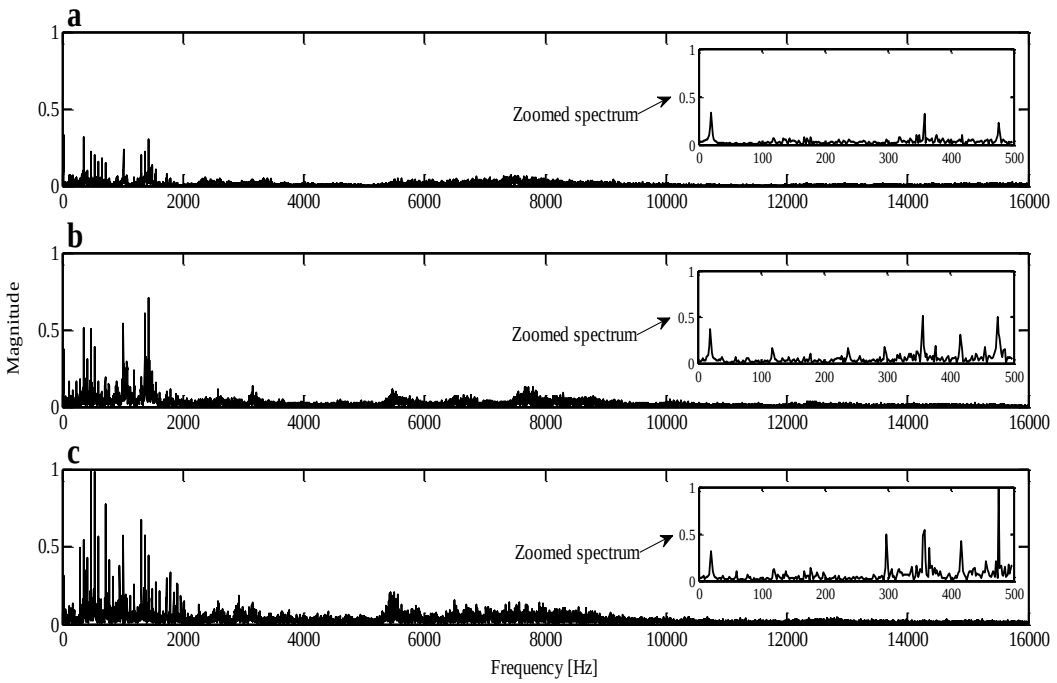

Figure 5.16: Amplitude spectrum of the vibration signals in Figure 5.15; a) small fault b) medium fault c) large fault.

Three Hankel matrices, each of which has 1000 rows, are formed for individual matrices corresponding to the test cases with ball bearings having small, medium and large local faults. The number of singular values to be used for noise elimination is determined systematically. For this purpose, the kurtosis ratio of the output de-noised signal to the input raw signal is obtained as function of the first *l* number of singular values for the reconstruction, up to 500 first singular values. Figure 5.17 presents the kurtosis results for measurements corresponding to different size of faults.

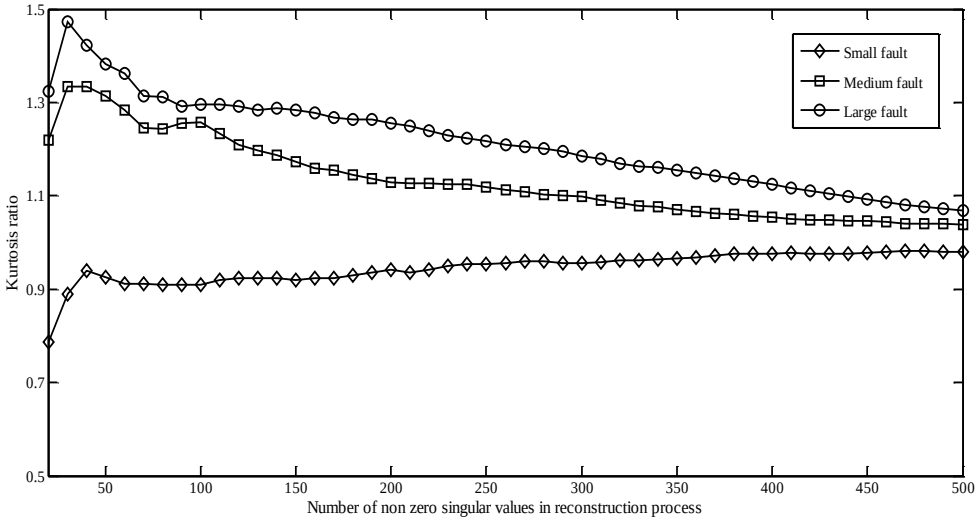

Figure 5.17: Illustration of the kurtosis ratio of the output to input signals in SVD based de-noising algorithm.

According to the results in Figure 5.17, the first 30 singular values for the ball bearing with small fault and the first 40 singular values for ball bearings with medium and large faults are considered to be sufficient for the elimination of the background noise and to maximize the kurtosis ratio. It is worth mentioning here that as the number of singular values included in the reconstruction process approaches to the maximum possible quantity (i.e. the number of all singular values), the ratio of the kurtosis, as expected, approaches to 1.

The de-noised vibration signals are plotted in Figure 5.18. In order to have a better understanding of de-noising effects in time domain, the envelope analysis is performed for the original and the de-noised signal captured from small fault case study up to 500 Hz and the results are presented in Figure 5.19.

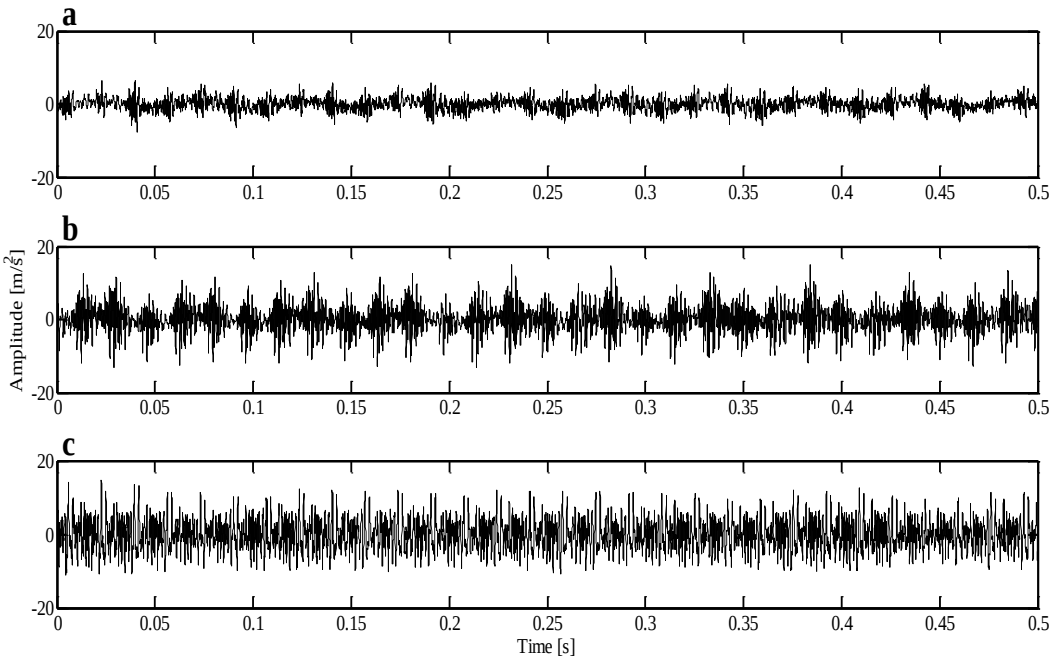

Figure 5.18: SVD based de-noised vibration signals correspond to the vibration signals in Figure 2; a) small fault b) medium fault c) large fault.

It is seen that, in spite of the fact that the overall amplitudes of envelope spectrum of the de-noised signal are somewhat reduced, the envelope spectrum of the de-noised signal has lower noise level than that of the original signal, hence yielding a slightly better detection of the related fault frequency (BPFO). It is also worth noting that while the original and de-noised signals in their envelope spectrums share more or less the same maxima at the first and second harmonics of BPFO, the third harmonic is visible only in the envelope spectrum of the de-noised signal.

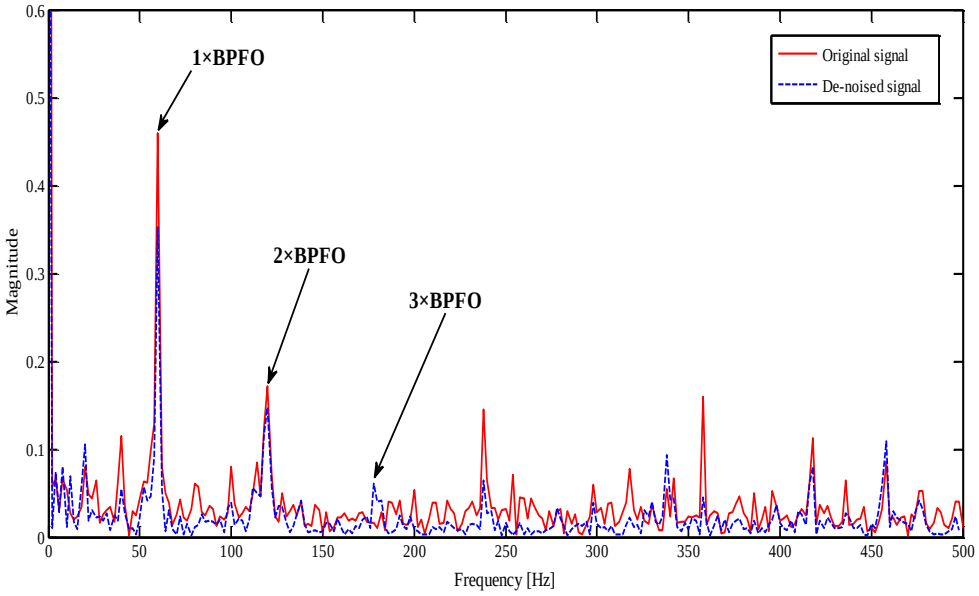

Figure 5.19: Envelope spectrum of the ball bearing with small fault for original and de-noised vibration signals.

In this section SVD of Hankel matrix based de-noising process is also applied to the amplitude spectrums of the ball bearing vibration signals. For this purpose, again, three Hankel matrices each with 1000 rows are formed using the spectrum vectors obtained via FFT. Figure 5.20 presents the normalized singular values obtained for the tested ball bearings in logarithmic scale. For de-noising purpose, a threshold value needs to be chosen based on the results in Figure 5.20.

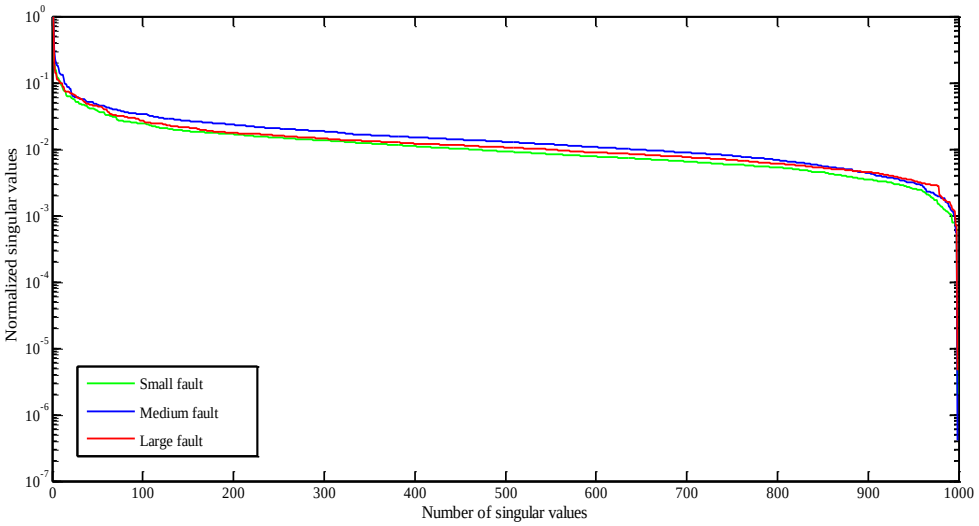

Figure 5.20: Normalized singular values of the spectrum in Figure 5.16.

Inspection of Figure 5.20 reveals that the singular values start to drop sharply up to about 100th singular value and then remain quite steady over a wide range. Therefore, in all cases, approximately, the first 100 singular values are used for the noise elimination and reconstruction process. Figure 5.21 and Figure 5.22 illustrate the original and the de-noised amplitude spectrums up to 350 Hz, respectively. As can be seen, the original amplitude spectrums in Figure 5.21, especially for the small fault case are unable to indicate the expected fault frequency clearly. However, inspection of the de-noised spectrum in Figure 5.22 reveals that noise removal from vibration spectrum makes the fault frequency and its harmonics very visible.

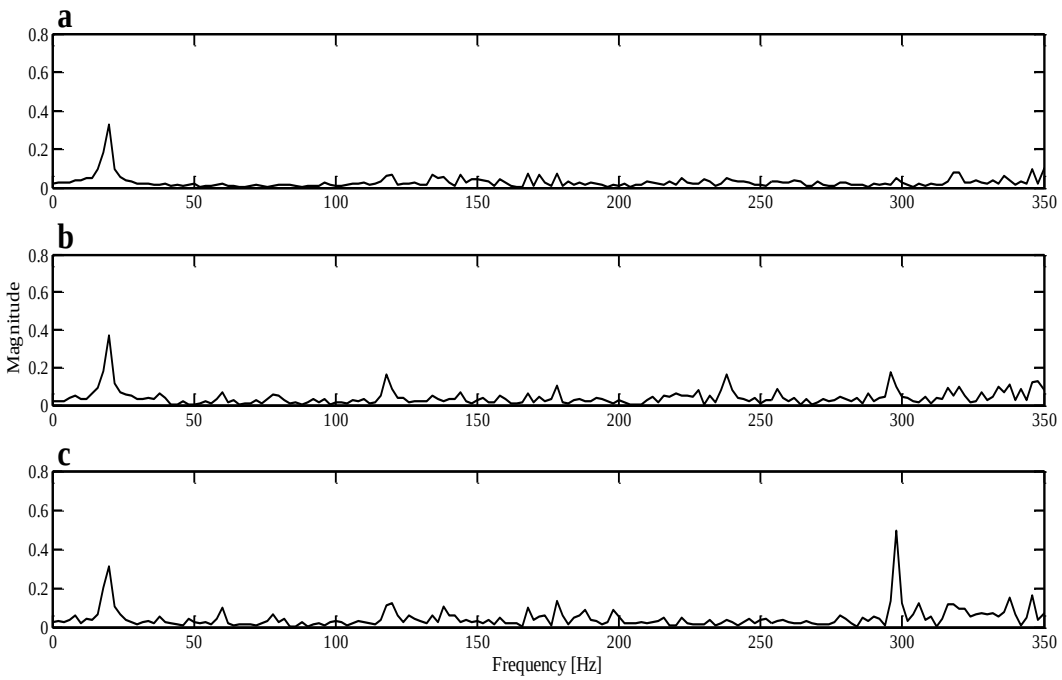

Figure 5.21: Zoomed amplitude spectrums of the original vibration signals in Figure 5.15; a) small fault b) medium fault c) large fault.

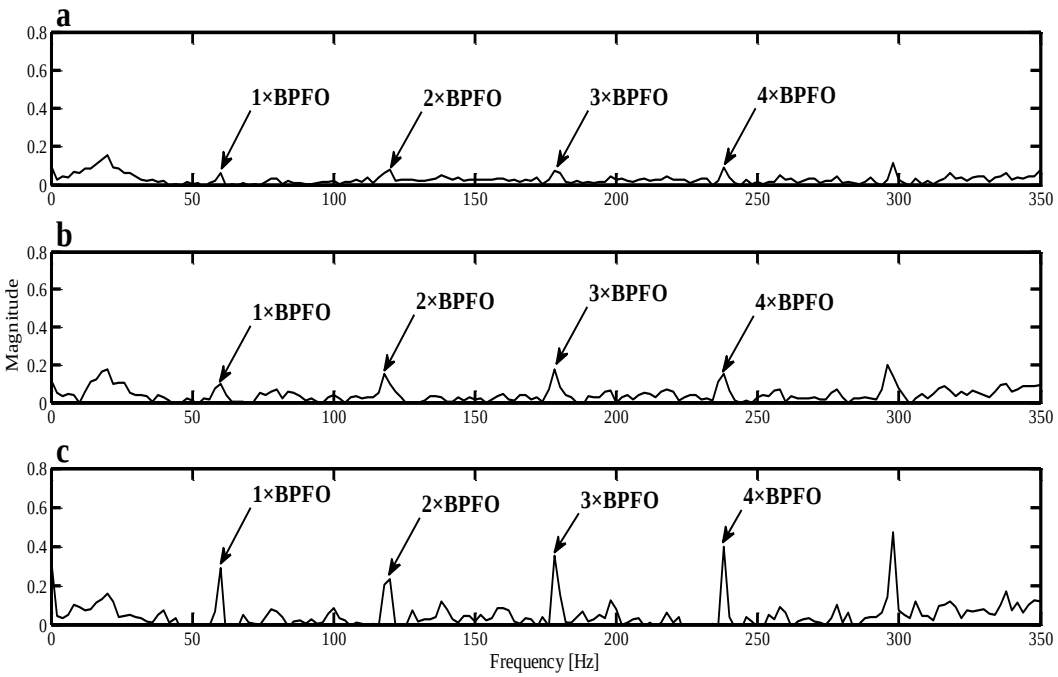

Figure 5.22: SVD based de-noised amplitude spectrums (using 100 singular values) of vibration signals in Figure 5.15; a) small fault b) medium fault c) large fault.

For better comparisons, amplitude spectrums presented in Figure 5.21c and Figure 5.22c are overlaid in Figure 5.23. There is no doubt that the de-noising the amplitude spectrum using the SVD based method makes the fault frequency much sharper and very visible indeed.

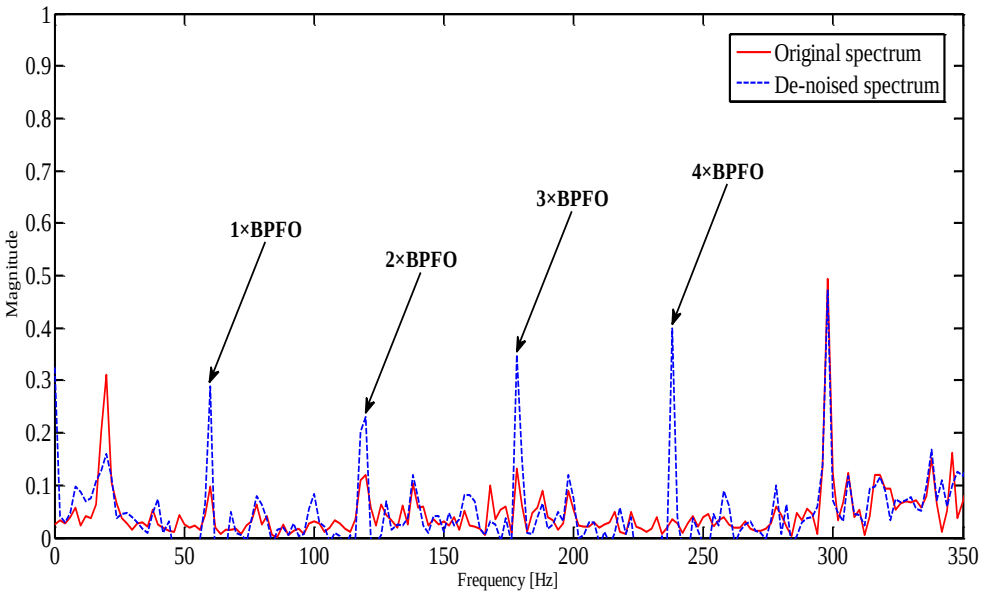

Figure 5.23: Original and de-noised amplitude spectrums of the vibration signal captured from ball bearing with large fault.

It is obvious that selecting the threshold, hence the number of most significant singular values to be included during the reconstruction process is important. Therefore, it is essential to assess the sensitivity of this threshold for fault detection purposes. The spectrums presented in Figure 5.24 are obtained by including different numbers of most significant singular values for noise elimination purpose. Comparing the results in Figure 5.24 with those in Figure 5.23 reveals that including the first most significant 50 or 100 or 150 singular values do not make much difference, all of them yielding much better detection of the fault frequency relative to the original spectrum. However, as expected, including higher number of singular values (e.g., 250 or 350) during the reconstruction process gradually reduces the benefit of noise removal. It can be said that the fault detection via de-noising is not very sensitive to the exact threshold value nor to the exact number of singular values included for de-noising purpose, provided that a sensible number of the most significant singular values are included during this process by considering the trend in Figure 5.20.

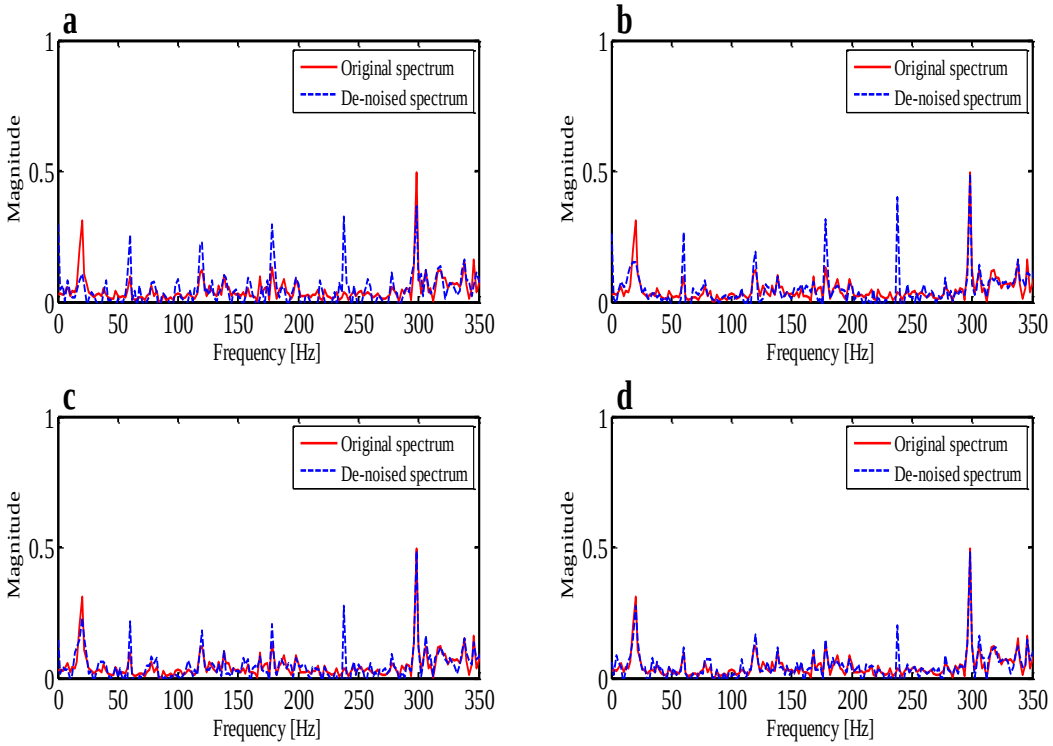

Figure 5.24: Original and de-noised amplitude spectrums using a) 50 b) 150 c) 250 d) 350 singular values of the vibration signal captured from ball bearing with large fault.

In this section, it is shown that de-noising the amplitude spectrum using the SVD based method makes the fault frequency much sharper and very visible. This suggests that

the fault identification using the de-noised amplitude spectrum can be a viable alternative to envelope analysis. It should be noted, however, that the effectiveness of the de-noising method depends on appropriate selection of the threshold level. Therefore, the number of most significant singular values, which needs to be included for the reconstruction process is an important parameter hence, a systematic identification of this threshold is recommended.

5.5 Wavelet Analysis

In this section, wavelet analysis as a time-frequency tool is studied experimentally using real ball bearings. Wavelet analysis decomposes the input vibration signal into its frequency bands as a function of time. Thus, it can be used as a non-stationary signal processing method or as a noise elimination method. Figure 5.25 shows a measured time domain vibration signal and its amplitude spectrum corresponding to a ball bearing with multiple faults (one fault on each raceway). In this case, the kurtosis vaue of the signal = 5.94, the rotaitonal speed was again 900 RPM and the position of the defective area on the outer race was located along upward directon. The fault frequencies, BPFO and BPFI are determined, respectively, as 46 and 73 Hz.

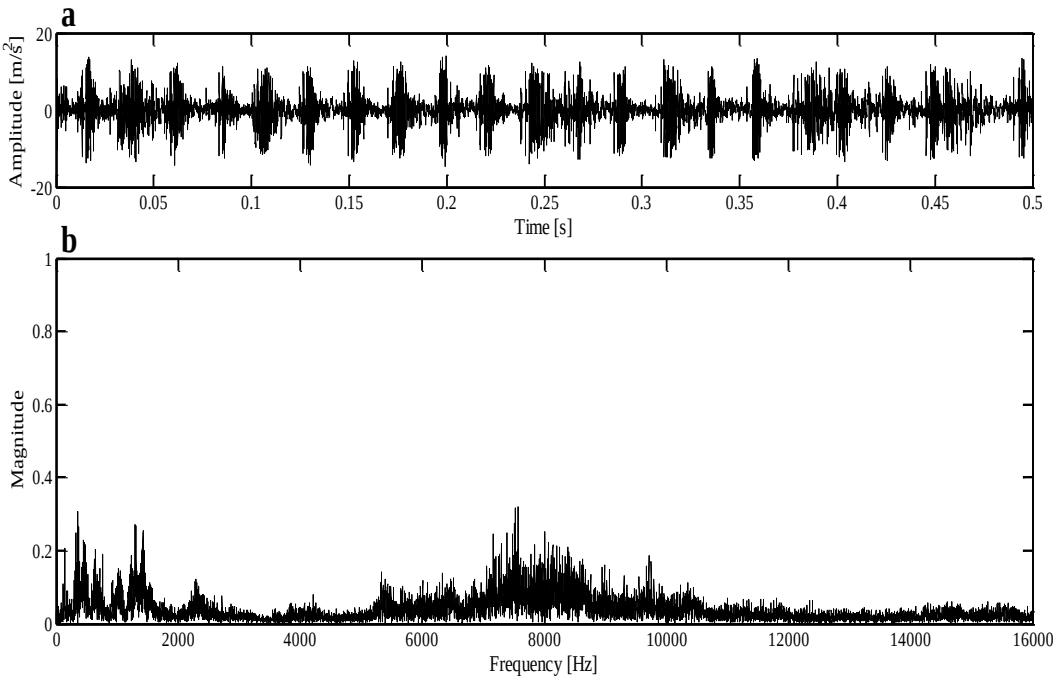

Figure 5.25: a) Captured vibratin signal from an outer and inner races defective ball bearing at 900 RPM b) the corresponding amplitude spectrum.

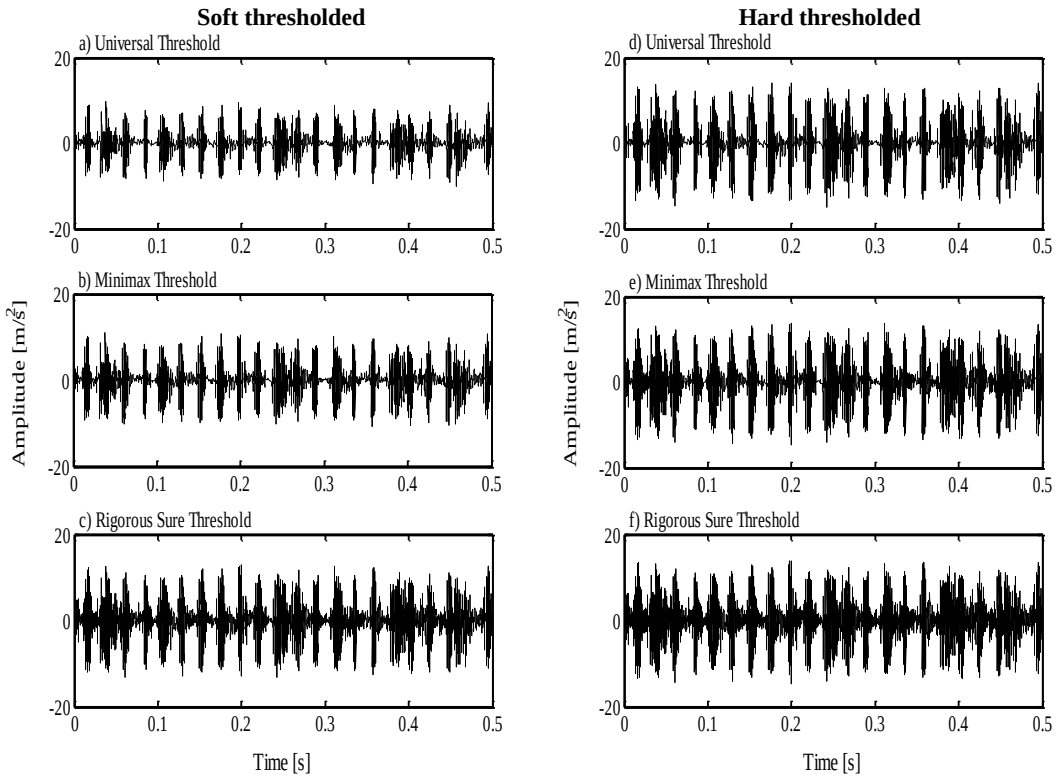

Figure 5.26: De-noised signals of the simulated vibration signal in Figure 5.25a based on DWT and soft (left) and hard (right) thresholding methods.

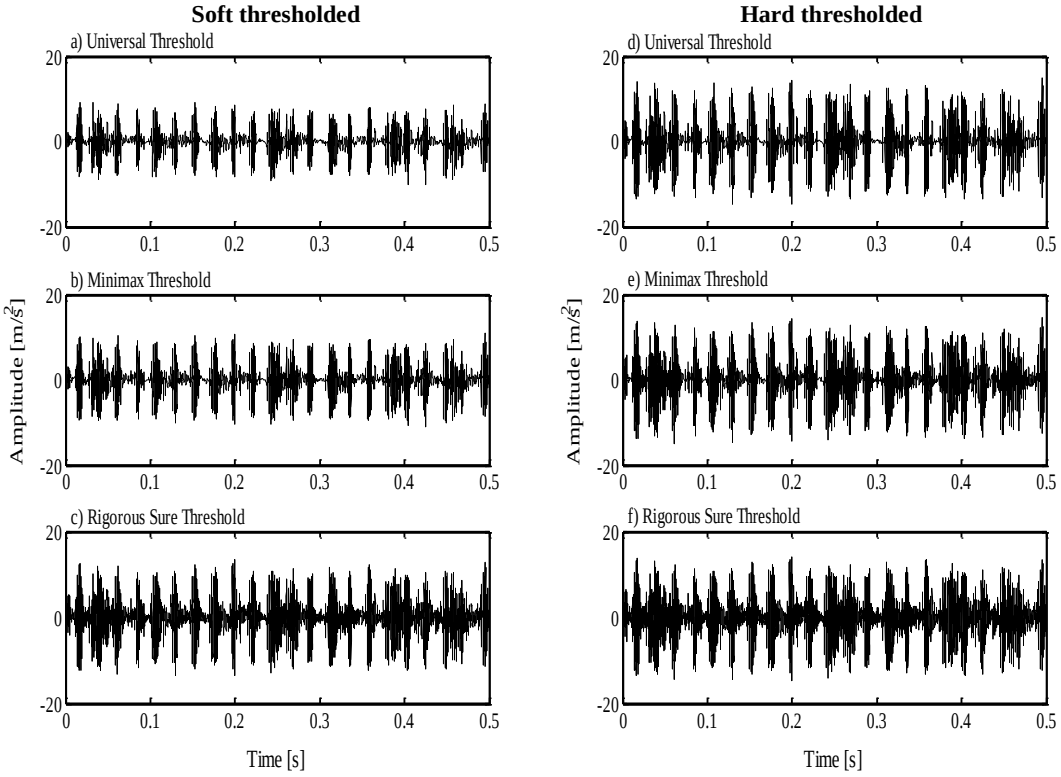

Figure 5.27: De-noised signals of the simulated vibration signal in Figure 5.25a based on SGWT and soft (left) and hard (right) thresholding methods.

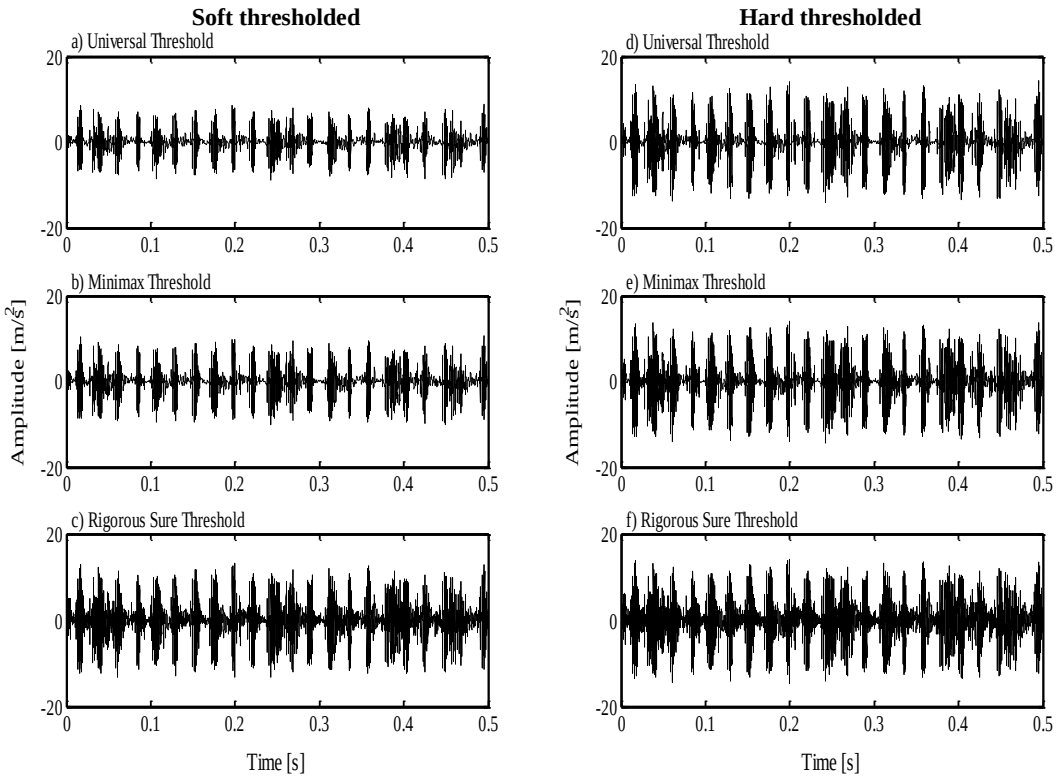

Figure 5.28: De-noised signals of the simulated vibration signal in Figure 5.25a based on RDWT and soft (left) and hard (right) thresholding methods.

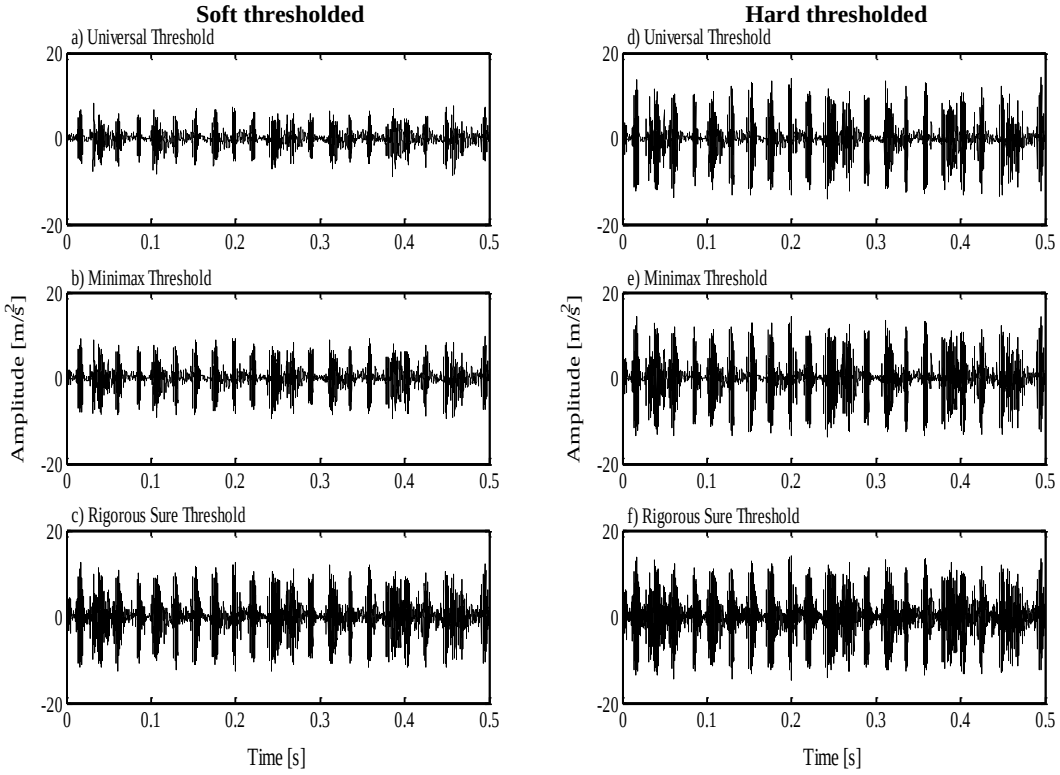

Figure 5.29: De-noised signals of the simulated vibration signal in Figure 5.25a based on DTCWT and soft (left) and hard (right) thresholding methods.

Figure 5.26 to Figure 5.29 illustrate de-noised signals based on DWT, SGWT, RDWT, and DTCWT, respectively, using various de-noising thresholds and thresholding methods. In addition, kurtosis values of the de-noised signals are listed in Table 5.1 for each thresholding and Wavelet Transform methods.

Table 5.1: Kurtosis values of the wavelet-based de-noised vibration signals.

	Universal Threshold		Minimax Threshold		Rigorous Sure Threshold	
	Soft thresholding	Hard thresholding	Soft thresholding	Hard thresholding	Soft thresholding	Hard thresholding
DWT	7.25	7.22	7.04	6.62	6.44	6.01
SGWT	7.32	7.21	7.08	6.59	6.39	6
RDWT	7.33	7.65	7.33	6.98	6.46	6.04
DTCWT	6.60	7.92	6.74	7.24	6.63	6.18

It is worth stating that all the de-noised vibration signals in Figure 5.26 to Figure 5.29 can detect the fault frequencies, BPFO and BPFI. According to the results in Table 5.1, the universal threshold yields the highest kurtosis values. Nevertheless, in spite of the numerical results in section 3.4, it is also concluded that the universal threshold gives the worst and Rigorous threshold gives the best fault feature extraction property (by analyzing the envelope spectrums of the de-noised vibration signals) compared to other thresholds. It should also noted that, RDWT yields the best noise cancelling property compared to other Wavelet Transforms for this case study.

After the investigation of the discrete wavelet-based noise elimination methods in frequency domain, Wavelet Transform in its continuous form, called CWT, is applied to two measured vibration signals (one for outer race and one for inner race defective ball bearings). The rotational shaft speed was 1200 RPM, consequently the fault frequencies are determined as 63 and 96 Hz, respectively. Figure 5.30 and Figure 5.31 present the time-frequency representaions of the vibrations signals. As can be seen, both CWT figures can detect the corresponding fault frequency. To have better understanding of time-frequency analyses, Figure 5.32 and Figure 5.33 show the time and frequency domains of the analyzed signals as well as their time-frequency representations. As seen, high energy peakes around the resonances of the signals can

be clearly seen in both plots. These peaks are repeated with certain time intervals, which characterize the corresponding fault frequencies.

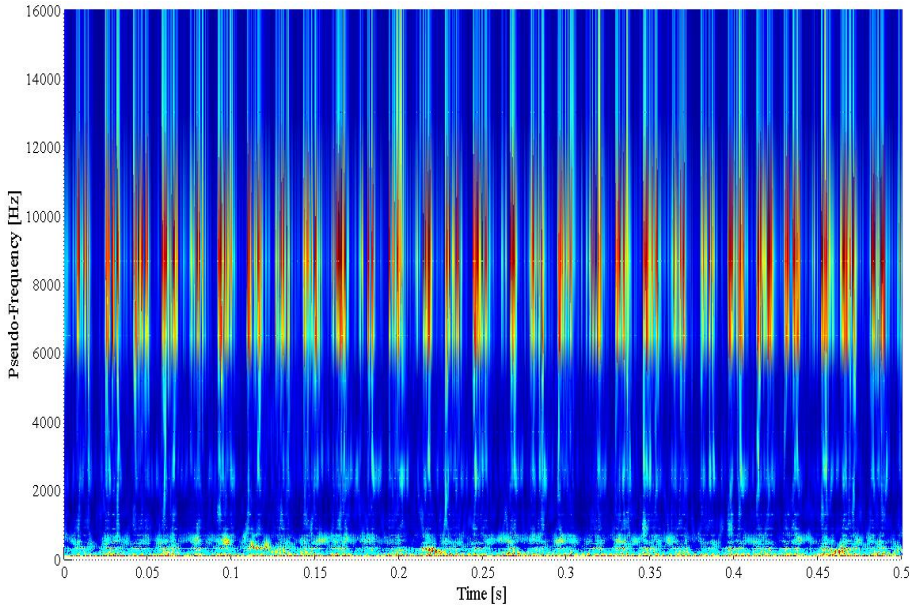

Figure 5.30: Time-frequency representation of the vibration signal captured from an outer race defective ball bearing.

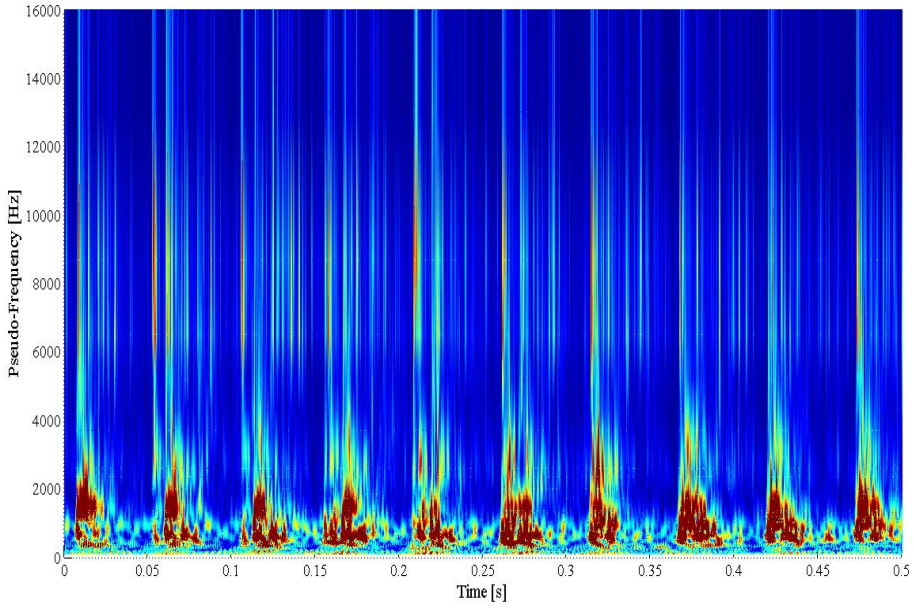

Figure 5.31: Time-frequency representation of the vibration signal captured from an inner race defective ball bearing.

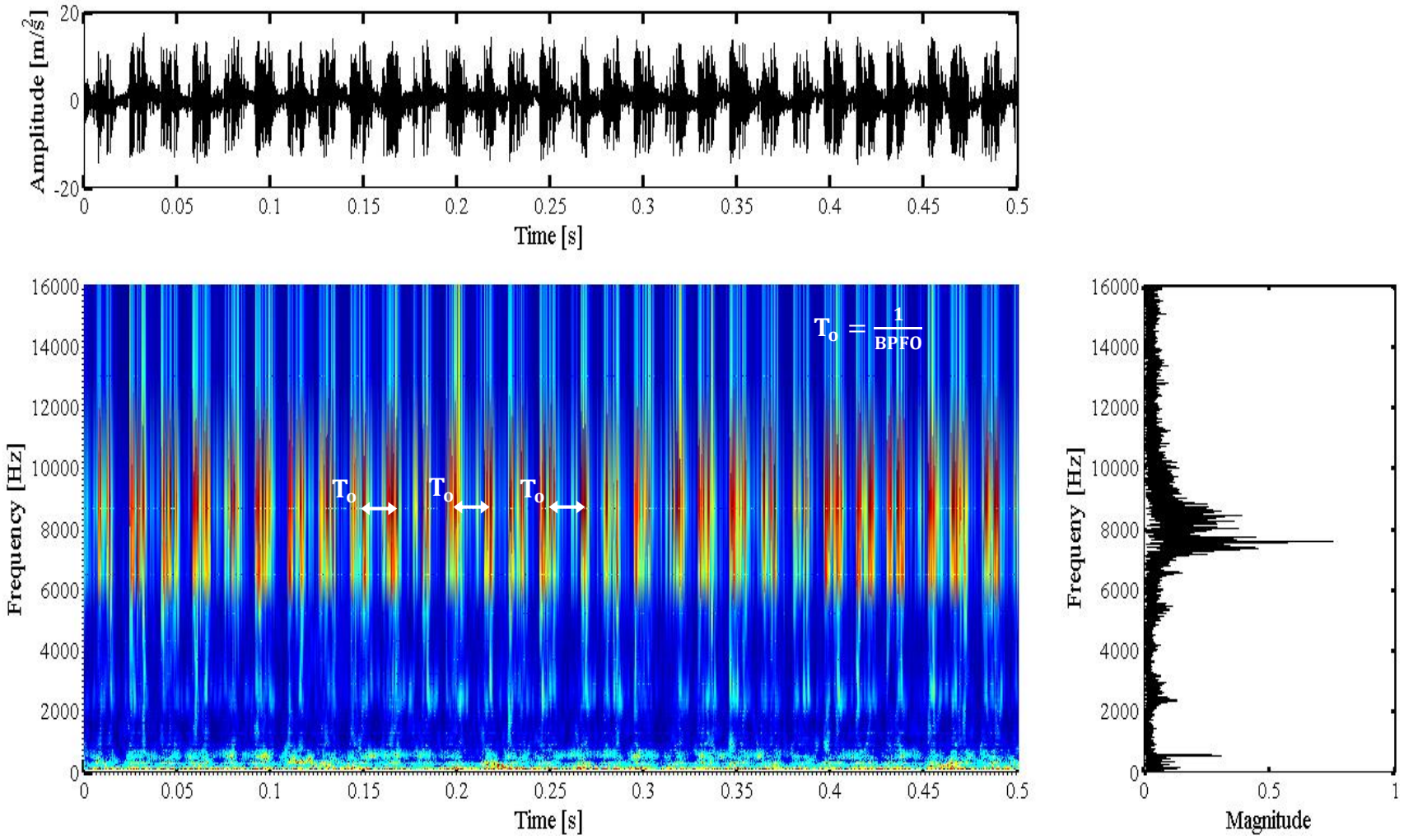

Figure 5.32: Time, frequency and time-frequency representations of the vibration signal captured from an outer race defective ball bearing.

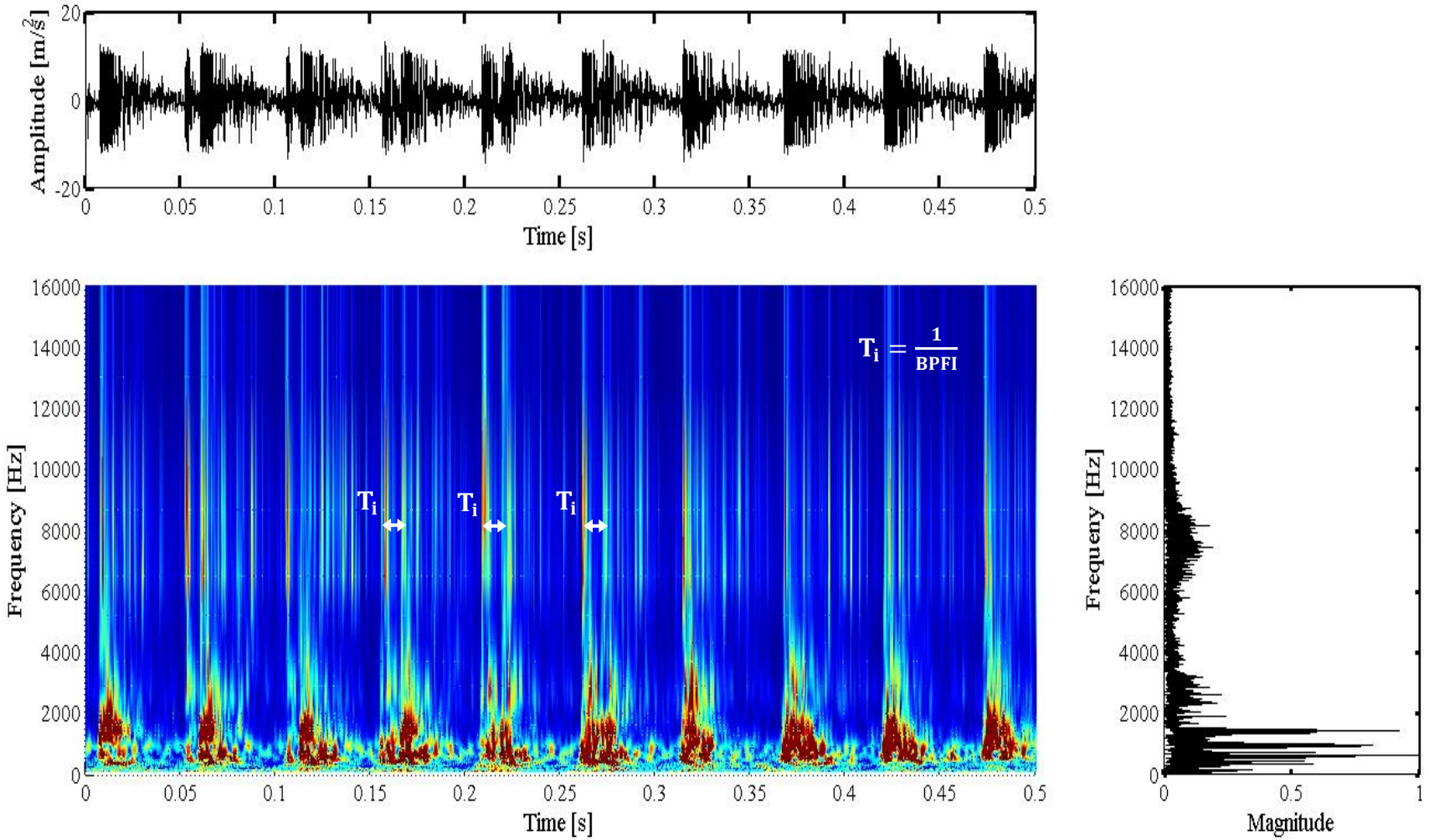

Figure 5.33: Time, frequency and time-frequency representations of the vibration signal captured from an inner race defective ball bearing.

5.6 Summary

This chapter presented contains the applications of some of the available advanced fault detection methods to measured vibration signals from bearing housings. Ball bearings having single and multiple faults and three various fault sizes are investigated using signal enhancing and de-noising techniques as the preprocessing step as well as the envelope analysis as a post processor and diagnosis information extraction. The present chapter demonstrates the importance of the signal sub-band analysis to have more reliable fault detections particularly for multiple ball bearing faults detection. EEMD is used to decompose the raw vibration signals into their frequency components so as to provide an opportunity for signal sub-band analyses. ARMED filter as a strong signal de-noising tools is also applied to measured ball bearing vibration signals. Wavelet analysis, both for signal de-noising and time-frequency analyses is also utilized in this chapter. Furthermore, a novel SVD based de-noising approach is proposed in this chapter and successfully applied to ball bearing localized fault detection in both time and frequency domains. It is also shown that de-noising the amplitude spectrum using the SVD based method makes the fault frequency much sharper and very visible. This suggests that the fault identification using the de-noised amplitude spectrum via the proposed approach can be a viable alternative to envelope analysis.

6. CONCLUSIONS AND SUGGESTIONS

6.1 Conclusions

This thesis addresses the applications of several signal processing techniques for diagnosis purposes. In particular, it is focused on the sensitivity and effectiveness of several signal enhancing and de-noising methods as well as non-stationary signal processing techniques applied for ball bearing fault detections.

After a brief introduction, mathematical foundations of several diagnosis signal processing tools are presented in second chapter. The third chapter contains numerical investigations based on generated vibration signals in MATLAB® platform. A brief description of the test rig used in the experiments is provided in the next chapter, followed by the experimental investigations presented in fifth chapter based on measured vibration signals from bearing housings.

Many vibration-based diagnosis methods have been proposed in the literature to monitor ball bearing conditions. Generally, a detection task based on the vibration analysis can be divided into signal enhancing (pre-processing step) and feature extraction (post processing step). Although the diagnostic feature extraction process is crucial for correct characterization of transient events (which contain valuable diagnosis information), signal de-noising and enhancing prior to any post processing remain perhaps the most important step in the whole fault detection process. In particular, in heavy noisy environments, when feature extraction methods are unable to describe the nature of the faults, de-noising of the vibration signals is strongly recommended. Even though the captured vibration signal contains small amount of background noise, it is worth improving the diagnostic feature of the analysed signal by cancelling the unwanted components and consequently enhancing the signal is desirable to obtain smoother and clearer spectrums. The signal enhancing and de-noising step(s) can be employed either in time or frequency domains, whereas the relevant diagnosis features are mostly associated with the frequency information and/or energy distribution in the vibration signal.

Several DWT theories with different mathematical foundations were investigated numerically and experimentally for wavelet-based de-noising purposes. It was shown that a specific Wavelet Transform may be very suitable in some conditions and may perform poorly in others. EMD and its revised version, EEMD, were also employed for single and multiple fault detections. In particular, the importance of the signal sub-band analyses to have more reliable fault detection of multiple faults was demonstrated. The combination of AR and MED filterings (called ARMED) was also applied for ball bearing faults detection. In spite of the fact that the overall amplitudes of the envelope spectrums of the ARMED based de-noised vibration signals were somewhat reduced, the envelope spectrums after de-noising were smoother than the previous envelope spectrums hence yielding a better detection of the related fault frequency. The optimum Wiener filter by kurtogram was another de-noising process that was applied to numerical vibration signals to remove the artificially added white noise. It was shown that the effectiveness of the Wiener filtering strongly depends on the correct estimation of the background noise level. Therefore, noise eliminations can be done either by signal decomposition such as EEMD and Wavelet Transforms or by filtering techniques such as ARMED and Wiener filterings.

In addition to available de-noising methods, a novel SVD based de-noising approach was proposed in this thesis and successfully applied for ball bearing localized fault detection in both time and frequency domains. For this purpose, the Hankel matrix was used for forming the input trajectory matrix in SVD process. The results showed that the proposed de-noising algorithm works very well in noise cancelling the vibration signals with heavy background noise.

In most cases, envelope analysis, the most commonly used diagnosis method, can reveal the modulated fault frequencies. Consequently, in this thesis, the envelope spectrum was used frequently as a post processor in the frequency domain. In addition, CWT as the time-frequency representation of the analysed signal was also used for fault features extractions. It was also shown that the de-noised amplitude spectrums using the proposed SVD based method makes the fault frequency much sharper and very visible and consequently can be considered as an alternative to the envelope spectrums.

6.2 Suggestions for Future Work

More detailed and further analyses in the field of vibration-based Condition Monitoring could yield more reliable results. For the purpose of fault detection, the signal processing methods in three major domains and noise elimination techniques investigated and applied in this thesis to the ball bearing vibration signals can be refined further to minimise the computation time. More specific recommendations are listed below:

1) Further works is needed to investigate the applicability of the signal processing tools in characterizing the fault signatures due to real ball bearing faults in industrial environments. Those tools include, but not limited to SVD based de-noised algorithm, Wiener filtering, EEMD, Wavelet Transforms etc.

2) The work in this thesis is concerned with the investigation of the overall diagnostics technique and has not particularly focussed on incipient faults. Further work may investigate the sensitivity of the techniques for incipient and very small faults.

3) The available ball bearing vibration signal models are unable to describe the non-linear relations in signal components as well as the dependency between the various faults in the same ball bearing. Future work is recommended to develop more representative mathematical models.

In addition to the above recommendations, further investigations can be carried out about the proposed SVD-based noise elimination method in this thesis. The investigations may focus on:

- Sensitivity, reliability, and boundaries of the proposed SVD-based de-noising method when it is applied to signals from other machines such as electric machines and gears.
- Combination of the proposed SVD-based de-noising method with other fault detection methods.
- Developing other reliable methods for determining the most significant singular values in SVD process.

-

REFERENCES

[1] **R. B. Randall.** (2011). Vibration-based Condition Moitoring. First ed., Illinois: John Wiley & Sons, Ltd, Publication.

[2] "*ISO 13373-1 Condition monitoring and diagnostics of machines - Vibration condition monitoring*" In International Standards Office, London: HMSO, 2002.

[3] **S. H. Ghaffari.** (2007). A Fault Diagnosis System for Rotary Machinery Supported by Rolling Element Bearings, PhD Thesis University of Waterloo.

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

[5] **W. He, Z.-N. Jiang and K. Feng.** (2009). Bearing fault detection based on optimal wavelet filter and sparse code shrinkage. *Measurement*, vol. 42, no. 7, pp. 1092-1102.

[6] **L. G. Martins and S. Gerges.** (1984). Comparison between signal analysis for detecting incipient bearing damage. In *Proceedings of the International Condition Monitoring Conference*, Swansea, UK, 10-13 April.

[7] **D. Dyer and R. M. Stewart.** (1978). Detecting of rolling element bearing damage by statistical vibration analysis. *Journal of Mechanical Design*, vol. 100, no. 2, pp. 229-235.

[8] **P. W. Tse and J. C. Leung.** (2010). Advanced System For Automatically Detecting Faults Occurring in Bearings. Newy York: Nova Science Publishers, Inc.

[9] **N. Sawalhi and R. B. Randall.** (2006). Helicopter Gearbox Bearing Blind Fault Identification Using a Range of Analysis Techniques. In *Proceedings of the 1st World Congress on Engineering Asset Management (WCEAM)*, London, UK, 11-14 July.

[10] **T. Heyns, P. Heyns and J. d. Villiers.** (2012). Combining synchronous averaging with a Gaussian mixture model novelty detection scheme for vibration-based condition monitoring of a gearbox. *Mechanical Systems and Signal Processing*, vol. 32, pp. 200-215.

[11] **F. Combet and L. Gelman.** (2007). An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor. *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2590-2606.

[12] **E. Sabini, J. Lorenc, O. Henyan and K. Hauenstein.** (2013). Bearing defect detection using time synchronous averaging (TSA) of an enveloped accelerometer signal. *United States*, Patent US 20030106375 A1, 12 June.

[13] P. D. MCFADDEN and M. M. TOOZHY. (2000). Application of Synchronous Averaging To Vibration Monitoring of Rolling Element Bearings. *Mechanical Systems and Signal Processing,* vol. 14, no. 6, pp. 891-906.

[14] J. Dron, L. Rasolofondraibe, C. Couet and A. Pavan. (2002). Fault Detection and Monitoring of A Ball Bearing Benchtest and A Production Machine Via Autoregressive Spectrum Analysis. *Journal of Sound and Vibration,* vol. 218, no. 3, pp. 501-525.

[15] P. Broersen. (2006). Automatic Autocorrelation and Spectral Analysis, Berlin: Springer.

[16] S. Thanagasundram, S. Spurgeon and F. S. Schlindwein. (2008). A fault detection tool using analysis from an autoregressive model pole trajectory. *Journal of Sound and Vibration,* vol. 317, no. 3-5, p. 975-993.

[17] H. Endo and R. Randall. (2007). Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. *Mechanical Systems and Signal Processing,* vol. 21, no. 2, pp. 906-919.

[18] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, *et al* (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceeings of Royal Society London, Series A,* vol. 454, pp. 903-995.

[19] Z. H. Wu and N. E. Huang. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis,* vol. 1, pp. 1-41.

[20] M. Kang and J.-M. Kim. (2013). Singular value decomposition based feature extraction approaches for classifying faults of induction motors. *Mechanical Systems and Signal Processing,* vol. 41, no. 1-2, pp. 348-356.

[21] B. Muruganatham, M. Sanjith, B. Krishnakumar and S. S. Murty. (2013). Roller element bearing fault diagnosis using singular spectrum analysis. *Mechanical Systems and Signal Processing,* vol. 35, no. 1-2, pp. 150-166, 2013.

[22] B. Kilundu, X. Chiementin and P. Dehombreux. (2011). Singular Spectrum Analysis for Bearing Defect Detection. *Journal of Vibration and Acoustics,* vol. 133, no. 5.

[23] K. Y. Sanliturk and O. Cakar. (2005). Noise elimination from measured frequency response functions. *Mechanical Systems and Signal Processing,* vol. 19, no. 3, pp. 615-631.

[24] Y. Jiang, J. Kan, Z. Zhang, G. Cheng and J. Wen. (2012). Machine Fault Diagnosis Based On Reassigned Wavelet Scalogram and SVD. *International Journal of Advancements in Computing Technology,* vol. 4, no. 20, pp. 10-21.

[25] Y. Chen and P. Zhang. (2012). Bearing Fault Detection Based on SVD And EMD. *Applied Mechanics and Materials,* Vols. 184-185, pp. 70-74.

[26] J. Jing, Y. Hu, X. Li and Z. Huang. (2007). Feature Extraction of Pulse Signal Based on Hilbert-Huang Transformation and Singular Value

Decomposition. In *The 1st International Conference on Bioinformatics and Biomedical Engineering*, Wuhan, China, 6-8 July.

[27] Andrew K.S. Jardine, Daming Lin, Dragan Banjevic. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510.

[28] M. Lang, H. Guo, J. Odegard and C. Burrus. (1996). Noise reduction using an undecimated discrete wavelet transform. *Signal Processing Letters, IEEE*, vol. 3, no. 1, pp. 10-12.

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

[30] I. Howard. (1994). A Review of rolling element bearing vibration "detection, diagnosis and prognosis". DSTO Aeronautical and Maritime Research Laboratory, Melbourne, Australia.

[31] R. Wiggins. (1978). Minimum entropy deconvolution. *Geoexploration*, vol. 16, no. 1-2, pp. 21-35.

[32] N. Sawalhi, R. Randall and H. Endo. (2007). The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2616-2633.

[33] J. Antoni. (2006). The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 282-307.

[34] J. Antoni and R. Randall. (2006). The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308-331.

[35] F. Tafinine and K. Mokrani. (2012). Real time automatic detection of bearing fault in induction machine using kurtogram analysis. *The Journal of the Acoustical Society of America*, vol. 132, no. 5, pp. EL 405-410.

[36] J. Antoni. (2007). Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 108-124.

[37] D. Wang, P. W. Tse and K. L. Tsui. (2013). An enhanced Kurtogram method for fault diagnosis of rolling element bearings. *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 176-199.

[38] R. Yan, R. X. Gao and X. Chen. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing*, vol. 96, no. Part A, pp. 1-15.

[39] Y. Wang, Z. He and Y. Zi. (2010). Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform. *Mechanical Systems and Signal Processing*, vol. 24, no. 1, pp. 119-137.

[40] D. E. Newland. (2007). Wavelet analysis of vibration signals. In *Handbook of Noise and Vibration Control*, M. Crocker, Hoboken HJ, John Wiley & Sons, Inc., Chapter 49.

[41] W.-X. Yang and X.-M. Ren. (2004). Detecting Impulses in Mechanical Signals by Wavelets. *EURASIP Journal on Advances in Signal Processing,* vol. 8, pp. 1156-1162.

[42] S. Prabhakar, A. R. Mohanty and A. S. Sekhar. (2002). Application of discrete wavelet transform for detection of ball Bearing Race Faults. *Tribology International,* vol. 35, no. 12, pp. 793-800.

[43] S. Khanam, N. Tandon and J. Dutt. (2014). Fault Size Estimation in the Outer Race of Ball Bearing Using Discrete Wavelet Transform of the Vibration Signal. In *2nd International Conference on Innovations in Automation and Mechatronics Engineering, ICIAME*, Tirgu Mures, Romania, 9-10 October.

[44] S. Mallat. (2002). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 11, no. 7, pp. 674-693.

[45] S. Abbasion, A. Rafsanjani. (2007). A. Farshidianfar and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing,* vol. 21, no. 7, pp. 2933-2945.

[46] W. Sweldens. (1996). The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets. *Applied and Computational Harmonic Analysis,* vol. 3, no. 2, p. 186–200.

[47] J. E. Fowler. (2005). The Redundant Discrete Wavelet Transform and Additive Noise. *IEEE Signal Processing Letters,* vol. 12, no. 9, pp. 629-632.

[48] I. Selesnick, R. Baraniuk and N. Kingsbury. (2005). The dual-tree complex wavelet transform. *Signal Processing Magazine, IEEE,* vol. 22, no. 6, pp. 123 - 151.

[49] C. Guanghua, M. Shiwei, L. Ming, Z. Jingming and Z. Weimin. (2008). Wigner-Ville distribution and cross Wigner-Ville distribution of noisy signals. *Journal of Systems Engineering and Electronics,* vol. 19, no. 5, pp. 1053-1057.

[50] S. U. Lee, D. Robb and C. Besant. (2001). The Directional Choi–Williams Distribution For The Analysis of Rotor-Vibration Signals. *Mechanical Systems and Signal Processing,* vol. 15, no. 4, pp. 789-811.

[51] Z. Yang, Z. Yu, C. Xie and Y. Huang. (2014). Application of Hilbert–Huang Transform to acoustic emission signal for burn feature extraction in surface grinding process. *Measurement,* vol. 47, pp. 14-21.

[52] S. A. McInerny and Y. Dai. (2003). Basic Vibration Signal Processing for Bearing Fault Detection. *IEEE Transactions On Education,* vol. 46, no. 1, pp. 149-156.

[53] M. P. Norton and D. G. Karczub. (2003). Fundamentals of noise and vibration analysis for engineers, 2nd ed., New York: Cambridge University Press.

[54] T. Barszcz and R. B. Randall. (2009). Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine. *Mechanical Systems and Signal Processing,* vol. 23, no. 4, pp. 1352-1365.

[55] J. Mathew and R. J. Alfredson. (1984). The condition monitoring of rolling element bearings using vibration analysis. *Journal of Vibration, Acoustics, Stress, and Reliability in Design,* vol. 106, pp. 447-453.

[56] D. W. Tufts, R. Kumerasan and I. Kirsteins. (1982). Data adaptive signal estimation by singular value decomposition of a data matrix. *Proceedings of the IEEE,* vol. 70, no. 6, pp. 684-685.

[57] G. Golub and C. Van Loan. (1989). Matrix Computations, 2nd ed., Baltimore, MD: John Hopkins University Press.

[58] S. Jensen, P. Hansen, S. J.A. and S. Hansen. (1995). Reduction of broad-band noise in speech by truncated QSVD. *IEEE Transactions on Speech and Audio Processing,* vol. 3, no. 6, pp. 439-448.

[59] A. D'addabbo, M. Papale and S. Di Paolo. (2008). SVD Based Feature Selection and Sample Classification of Proteomic Data. In *12th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems,* Zagreb, Croatia, September 3-5.

[60] B. Muruganatham, S. M.A., K. B.K., S. Murty and P. Swaminathan. (2010). Inner race bearing fault detection using Singular Spectrum Analysis. In *IEEE International Conference on Communication Control and Computing Technologies (ICCCCT),* Ramanathapuram, India, 7-9 October.

[61] Q. Miao, D. Wang and M. Pecht. (2011). Rolling Element Bearing Fault Feature Extraction Using EMD-Based Independent Component Analysis. In *IEEE International Conference on Prognostics and Health Management (PHM),* Montreal, QC, 20-23 June.

[62] W.-C. Tsao, Y.-F. Li and M.-C. Pan. (2010). Resonant-frequency band choice for bearing fault diagnosis based on EMD and envelope analysis. In *8th World Congress on Intelligent Control and Automation (WCICA),* Jinan, China, 7-9 July.

[63] X. Zhang and J. Zhou. (2013). Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mechanical Systems and Signal Processing,* vol. 41, no. 1-2, pp. 127-140.

[64] J. Lin and Q. Chen. (2010). Application of the EEMD Method to Multiple Faults Diagnosis of Gearbox. In *2nd International Conference on Advanced Computer Control (ICACC),* Shenyang, China, 27-29 March.

[65] Y. Lei, N. Li, J. Lin and S. Wang. (2013). Fault Diagnosis of Rotating Machinery Based on an Adaptive Ensemble Empirical Mode Decomposition. *Sensors,* vol. 13, no. 12, pp. 16950-16964.

[66] Y. Amirat, V. Choqueuse and M. Benbouzid. (2013). EEMD-based wind turbine bearing failure detection using the generator stator current homopolar component. *Mechanical Systems and Signal Processing,* vol. 41, no. 1-2, pp. 667-678.

[67] W. Guo, P. W. Tse and A. Djordjevich. (2012). Faulty Bearing Signal Recovery From Large Noise Using A Hybrid Method Based On Spectral Kurtosis And Ensemble Empirical Mode Decomposition. *Measurement,* vol. 45, no. 5, pp. 1308-1322.

[68] N. Sawahli and R. Randall. (2007). Simulation of vibrations produced by localized produced by localized. In *5th Australasian Congress on Applied Mechanics ACAM,* Brisbane, Australia, 10-12 December.

[69] J. Proakis and M. Salehi. (2000). Contemporary Communication Systems, BookWare.

[70] A. Al-khalidy. (2002). Health Monitoring of Dynamic Systems Using Wavelet Analysis.

[71] A. Tiwari and R. Jatola. (2013). Fault Detection In Bearing Using Envelope Analysis. *Indian Journal of Research,* vol. 3, no. 5, pp. 75-78.

[72] I. Moumene and N. Ouelaa. (2012). Gears and Bearings Combined Faults Detection Using Hilbert Transform and Wavelet Multiresolution Analysis. In *Proceedings of the Second International Conference on Condition Monitoring of Machinery in Non-Stationnary Operations (CMMNO), Wroclaw, Poland*.

[73] F. Cong, J. Chen, G. Dong and F. Zhao. (2013). Short-time matrix series based singular value decomposition for rolling bearing fault diagnosis. *Mechanical Systems and Signal Processing,* vol. 34, no. 1-2, pp. 218-230,.

[74] Y. Ming, J. Chen and G. Dong. (2011). Weak fault feature extraction of rolling bearing based on cyclic Wiener filter and envelope spectrum. *Mechanical Systems and Signal Processing,* vol. 25, no. 5, pp. 1773-1785.

[75] D. Koulocheris, A. Stathis, T. Costopoulos and A. Atsas. (2014). Wear and multiple fault diagnosis on rolling bearings using vibration signal analysis. *International Journal of Engineering Science Invention,* vol. 3, no. 4, pp. 2319-6726.

[76] L. Gelman, T. H. Patel, B. Murray and A. Thomson. (2013). Rolling Bearing Diagnosis Based on the Higher Order Spectra. *International Journal of Prognostics and Health Management,* vol. 4, no. 2, pp. 1-9.

[77] A. Carcaterra and A. Sestieri. (1997). Complex envelope analysis: a quasi-static approach to vibrations. *Journal of Sound and Vibration,* vol. 201, no. 2, p. 205–233.

[78] J.-Y. LEE and A. NANDI. (2000). Extraction of Impacting Signals Using Blind Deconvolution. *Journal of Sound and Vibration,* vol. 232, no. 5, pp. 945-962.

[79] R. Jiang, J. Chen, G. Dong, T. Liu and W. Xiao. (2013). The weak fault diagnosis and condition monitoring of rolling element bearing using minimum entropy deconvolution and envelop spectrum. *Journal of Mechanical Engineering Science,* vol. 227, no. 5, pp. 1116-1129.

[80] R. Ricci, P. Borghesani , S. Chatterton and P. Pennacc. (2012). The Combination of Empirical Mode Decomposition And Minimum. In *Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference,* Chicago, IL, USA, August 12-15.

[81] S. V. Vaseghi. (2000). Advanced Signal Processing and Noise Reduction, 2nd ed., Chichester, England: John Wiley & Sons Ltd.

[82] R. Dwyer. (1983). Detection of non-Gaussian signals by frequency domain Kurtosis estimation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP)*, Boston,MA, USA, 14-16 April.

[83] B. Chen, Z. Yan and W. Chen. (2014). Defect Detection for Wheel-Bearings with Time-Spectral Kurtosis and Entropy. *Entropy*, vol. 16, no. 1, pp. 607-626.

[84] R. Randall. (2005). Applications of Spectral Kurtosis in Machine Diagnostics and Prognostics. In *DAMAS Conference, Key Engineering Materials*.

[85] H.-T. Yang and C.-C. Liao. (2001). A De-Noising Scheme for Enhancing Wavelet-Based Power Quality Monitoring System. *IEEE Transactions On Power Delivery*, vol. 16, no. 3, pp. 353-360.

[86] N. G. Nikolaou and I. A. Antoniadis. (2002). Rolling element bearing fault diagnosis using wavelet packets. *NDT&E International*, vol. 35, no. 3, pp. 197-205.

[87] S. Lesecq, S. Gentil and S. Taleb. (2006). Fault Detection Based on Wavelets Transform: Application To a Roughing Mill. In *IFAC Fault Detection, Supervision and Safety of Technical Processes*, Beijing, China.

[88] N. Li and R. Zhou. (2001). Rolling Element Bearing Fault Detection Using Redundant Second Generation Wavelet Packet Transform. *Advanced Materials Research,* Vols. 199-200, pp. 931-935.

[89] H. Choi, J. K. Romberg, R. G. Baraniuk and N. G. Kingsbury. (2000). Hidden Markov tree modeling of complex wavelet transforms," in *Proc. IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP)*, Istanbul, Turkey, June 5–9.

[90] V. Kumar and S. K. Agarwal. (2012). Compression of Color Image by Using Dual Tree Complex Wavelet Transfrom. In *Proceedings of International Conference on Internet Computing and Information Communications*, India.

[91] W. Sweldens. (1997). The Lifting Scheme: A Construction of Second Generation Wavelets. *SIAM Journal on Mathematical Analysis*, vol. 29, no. 2, p. 511–546.

[92] R. Claypoole. (1999). Adaptive Wavelet Transform via Lifting.

[93] W. Bao, R. Zhou, J. Yang, D. Yu and N. Li. (2009). Anti-aliasing lifting scheme for mechanical vibration fault feature extraction. *Mechanical Systems and Signal Processing,* vol. 23, no. 5, p. 1458–1473.

[94] N. Li, R. Zhou, Q. Hu and X. Liu. (2012). Mechanical fault diagnosis based on redundant second generation wavelet packet transform, neighborhood rough set and support vector machine. *Mechanical Systems and Signal Processing,* vol. 28, p. 608–621.

[95] O. V. Vasilyev and C. Bowman. (2000). Second-Generation Wavelet Collocation Method for the Solution of Partial Differential Equations. *Journal of Computational Physics,* vol. 165, no. 2.

[96] M. Holschneider, R. Kronland-Martinet, J. Morlet and P. Tchamitchian. (1987). A real-time algorithm for signal analysis with the help of the wavelet transform. In *Proceedings of the International Conference on*

Time-Frequency Methods and Phase Space, Marseille, France, 14-18 December.

[97] P. Dutilleux. (1989). An Implementation of the "algorithme à trous" to Compute the Wavelet Transform. In *Wavelets: Time-Frequency Methods and Phase Space*, Springer Berlin Heidelberg, pp. 298-304.

[98] S. H. Mortazavi and S. M. Shahrtash. (2008). Comparing Denoising Performance of DWT, WPT, SWT and DT-CWT for Partial Discharge Signals. In *43rd International Conference on Universities Power Engineering Conference UPEC*, Padova, Italy, 1-4 September.

[99] R. Kumar and P. Patel. (2012). Signal Denoising with Interval Dependent Thresholding Using DWT and SWT. *International Journal of Innovative Technology and Exploring Engineering (IJITEE),* vol. 1, no. 6, pp. 47-50.

[100] L. Yi-Hua and Z. Guang-Zhou. (2008). Power System Fault Detection based on Stationary Wavelet Packet Transform and Hilbert Transform. In *International Congress on Image and Signal Processing CISP*, Sanya, China, 27-30 May.

[101] F. Wang and Z. Ji. (2014). Application of the Dual-tree Complex Wavelet Transform in Biomedical Signal Denoising. *Bio-Medical Materials and Engineering,* vol. 24, no. 1, p. 109–115.

[102] V. Kumar and A. K. Surendra. (2012). Compression of Color Image by Using Dual Tree Complex Wavelet Transfrom. In *Proceedings of sInternational Conference on Internet Computing and Information Communications ICICIC*, India.

[103] G. Serbes and N. Aydin. (2013). Directional Dual-Tree Complex Wavelet Packet Transform Implementation Using Band-Limited Synthetic Signals. In *IFMBE Proceedings of The International Conference on Health Informatics*, Osaka, Japan, 7-9 November.

[104] I. Bayram and I. Selesnick. (2008). On the Dual-Tree Complex Wavelet Packet and M -Band Transforms. In *IEEE Transactions on Signal Processing*, June.

[105] R. R. Coifman and D. L. Donoho. (1995). Translation-Invariant De-Noising," in *Lecture Notes in Statistics: Wavelet and Statistics*, vol. 103, Berlin, Springer, pp. 125-150.

[106] J. Lewis and C. Burrus. (1998). Approximate continous wavelet transform with an application to noise reduction. In *Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, Seattle, WA, USA, 12-15 May.

[107] L. Zhen, H. Zhengjia, Z. Yanyang and W. Yanxue. (2008). Customized wavelet denoising using intra-and inter-scale dependency for bearing fault detection. *Journal of Sound and Vibration,* vol. 313, no. 1-2, pp. 342-359.

[108] D. L. Donoho and I. M. Johnstone. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika,* vol. 81, no. 3, pp. 425-455.

[109] S. Tsung-Ying, L. Chan-Cheng, H. Sheng-Ta, *et al*. (2008). Optimal Determination of Wavelet Threshold and Decomposition Level via

Heuristic Learning for Noise Reduction. In *IEEE Conference on Soft Computing in Industrial Applications*, Muroran, Japan, 25-27 June.

[110] Y. Lei, J. Lin, Z. He and M. J. Zuo. (2013). A reviewon empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108-126.

[111] S. Qin and Y. Zhong. (2004). Research on the unified mathematical model for FT, STFT and WT and its applications. *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1335-1347.

[112] Q. He and X. Wang. (2012). Time-Frequency Manifold for Demodulation with Application to Gearbox Fault Detection, In *IEEE International Conference on Prognostics and System Health Management (PHM)*, Beijing, China, 23-25 May.

[113] M. Cocconcelli, R. Zimroz, R. Rubini and W. Bartelmus. (2012). STFT Based Approach for Ball Bearing Fault Detection in a Varying Speed Motor. In *Proceesings of the Second International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO)*, Hammamet, Tunisia, March 26-28.

[114] V. Giurgiutiu and L. Yu. (2003). Comparison of Short-time Fourier Transform and Wavelet Transform of Transient and Tone Burst Wave Propagation Signals For Structural Health Monitoring. Stanford University, Stanford, CA, September 15-17.

[115] T. Brotherton, T. Pollard and D. Jones. (1992). Applications of time-frequency and time-scale representations to fault detection and classification. In *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, BC, Canada, 4-6 Oct.

[116] E. Bouchikhi, V. Choqueuse, M. Benbouzid and J. Charpentier. (2011). A comparative study of time-frequency representations for fault detection in wind turbine. In *37th Conference on IEEE Industrial Electronics Society (IECON)*, Melbourne, VIC, 7-10 Nov.

[117] R. B. Randall, J. Antoni and S. Chobsaard. (2001). The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 945-962.

CURRICULUM VITAE

Name Surname: Reza Golafshan

Place and Date of Birth: URMIA - IRAN, 03.04.1988

E-Mail: reza.golafshan@gmail.com; reza.golafshan@itu.edu.tr

EDUCATION:

B.Sc., Mechanical Engineering, Islamic Azad University, IRAN, 2011

M.Sc., Mechanical Engineering, Istanbul Technical University, Turkey, 2015