Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

School of Engineering and Sciences

Hilbert-Huang Transform based Methodology for Bearing Fault Detection

A thesis presented by

Rubén Campos García

Submitted to the

School of Engineering and Sciences
in partial fulfillment of the requirements for the degree of

Master of Science

In

Manufacturing Systems

Monterrey, Nuevo León, May 16th 2018

Dedication

Dedico este trabajo principalmente a mis padres, por su amor, ejemplo y todos los sacrificios realizados para brindarme una vida mejor, por enseñarme y motivarme a ser una persona mejor cada día.
A mis hermanos, con quienes he compartido momentos increíbles a lo largo de mi vida y que a pesar de la distancia me apoyan incondicionalmente.
A mis amigos, por permitirme formar parte de sus vidas y por todas las experiencias compartidas que hicieron placentero el tiempo durante mis estudios.
A Dios, por las grandes bendiciones a lo largo de mi vida y por permitirme alcanzar todas las metas que me he propuesto.

Acknowledgements

I would like to to express my deepest gratitude to Tecnológico de Monterrey for the tuition support and giving me the life changing opportunity to study in such a wonderful institution.

I thank the Automotive Consortium and Bocar SA de CV for the given experience to participate in the project *Monitoring and Diagnosis of High Speed Machining Spindles*.

I want to thank to my thesis advisor, Dr. Rubén Morales Menéndez, and co-advisor Dr. Antonio Jr. Vallejo Guevara for their patience, support and guidance that helped me to achieve the objectives of this research.

I would like to express my most sincere thanks to the committe member, Dr. David Isaac Ibarra Zárate for encouraging me during my studies, as well for the given recommendations and his valuable time in support of the development of my thesis.

Finally, thanks to CONACyT for the financial support during my studies.

Hilbert-Huang Transform based Methodology for Bearing Fault Detection

By
Rubén Campos García

Abstract

Rotating machinery is of great importance for manufacturing industry, and therefore huge investments for their acquisition are made every year. Machine preservation plays an important role in the exploitation of this resource. Rotating machines are more susceptible to certain types of faults, investigations report that at least 42 % of the root causes of failure in rotating machinery are related with bearings.

To detect the bearing condition many techniques have been developed. One of the most reliable is vibration analysis. The *Hilbert-Huang transform (HHT)* has been used for vibration analysis and has gained attention in recent years, a topic of controversy in this method is the selection of the *Intrinsic Mode Functions (IMFs)* with fault information.

Statistical parameters can be used to describe the characteristics of vibration signals, this attribute can be exploited to select the *IMFs*. There are many time domain features used for signal analysis. In this research, a study of 17 statistical parameters was made to determine which one is the best to represent *IMFs* with fault information. As a result of this analysis a new methodology based on *HHT* is proposed. This methodology deals with the *IMF* selection with the use of *KR (Kurtosis x RMS)* to detect the *IMFs* with fault information, and can be used to detect incipient bearing faults.

The proposed methodology was validated with 18 signals from the *Case Western Reserve University (CWRU)*, *Tian-Yau Wu*, and the *society for Machinery Failure Prevention Technology (MFPT Society)* databases. For the 18 analyzed signals, only one *IMF* was wrongly selected. The cause of this error was the end defect produced in the *EMD*, this caused the *KR* amplitude to increase even tough the *IMF* did not have fault information. The results on the *Envelope spectrum* from 14 signals were clear with fault components with large amplitude. For the remaining four signals the results on the *Envelope spectrum* was noisy, but the fault fault components were distinguishable.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Rotating machines are widely used in manufacturing industry. This type of machines includes pump systems, turbines, conveyor belts, rotating cutting machines, and so on. Some of these machines are key for product manufacturing, and usually, they are very expensive.

Companies make huge investments for the acquisition of these machinery and expect to recuperate the initial investment and to make profits, avoiding faults in the equipment is essential to achieve their goal. Unexpected faults can cause work accidents, downtimes in some processes or stop the production and result in financial losses for the companies, and even sometimes the machine can break down. For these reasons, the preservation of the machine has great importance.

1.1.1 Maintenance for machine preservation

Maintenance is the key to preserve the machine, there are three strategies: run-to-failure, preventive, and predictive maintenance [Scheffer and Girdha, 2004]. They are applied taking into consideration the cost and importance of the machine, and the maintenance cost.

Run-to-failure Maintenance: this strategy consists in repairing the machine when a fault is presented (*when it breaks*), and is focused to make the machine back to operation. This type of maintenance is applied to low cost general purpose machinery with low level of importance on production.

Preventive Maintenance: this strategy consists in scheduling a series of maintenance activities to preserve the machine, these activities are performed at predetermined time intervals that can be weeks or months or runtime hours of the machine. The goal of this strategy, contrary to the *run-to-failure*, is to avoid downtimes and keep the machine in normal operation. This is applied to machines with high costs and essential for the production.

Condition-based Maintenance (CBM)[1]: this strategy, also known as predictive maintenance, consists in the supervision of the most important components within the machine to detect how these components are deteriorating and when are close to fault. The goal is to anticipate the fault, programming and performing the maintenance activities when specific components of the machine requires it. This strategy implies an initial investment for specialized *condition monitoring systems* which often are expensive. *CBM* is applied to most expensive machines which are critical for production, often if this type of machine breaks the losses are much greater than the investment of the monitoring system.

Both, run-to-failure and preventive maintenance are the most common maintenance strategies in industry, the necessity of high skilled personnel makes difficult to apply *CBM*; nevertheless, it should be applied to the most important machines.

1.2 Problem Description

Rotating machinery is of great importance for manufacturing industry and by default preventive maintenance is applied to them, but in cases where rotating machines play a key role in production, the application *CBM* totally is justified. The first step to implement the *CBM* is to identify the most important components of the machine. One of the main components of motors from rotating machines is the shaft, which is a movile structure supported on bearings. These movile parts are vulnerable to faults during runtime, the friction, load forces, and the vibrations they experience can produce fatigue in their components, producing the machine to break. Rotating machinery are prone to some specific faults, their most frequent causes of fault are: *unbalance*, *misalignment*, *bent shaft*, and *bearing faults* [Randall, 2011].

The importance of rotating machinery has lead to investigations of the most frequent root causes of failures to pay an extra attention to these issues. Some examples of these investigations are: a study made by Motor Reliability Working Group in which 1,141 motors of 75 plants of 33 companies were analyzed, the survey revealed that around 43% faults in motors are bearing related, another study performed by General Electric of 4,797 motors of 132 plants of 56 companies, showed that almost 42% of faults are bearing related [Bell *et al.*, 1985]. Thorsen and Dalva study showed that 51% of motor faults are bearing related [Thorsen and Dalva, 1995], and Zhang et al. cites a fourth survey made by Allianz, which result showed that bearings were the second root cause of failures in motors [Zhang *et al.*, 2011], this is consistent with Immovilli et al. that mentions that depending of the type and size of the machine, bearing faults vary from about 40% to about 90% from large to small machines [Immovilli *et al.*, 2010].

The previous surveys lead to the conclusion that *CBM* is important for rotating machines and can be focused in bearing fault detection.

[1] All acronyms are defined in Appendix A

Condition monitoring systems use different techniques to detect the condition of components within the machine, they are based on the measurement of physical parameters. Some of them are vibration monitoring, acoustic emission, oil analysis, particle analysis, corrosion monitoring, and thermography.

According to Scheffer and Girdha vibration analysis is the most effective and the most applied technique to detect defects in rotating machinery since vibration analysis can identify problems while they are in early stages allowing to prevent unscheduled downtimes [Scheffer and Girdha, 2004]. For these reasons, *the research is focused in the vibration analysis techniques for bearing fault detection*.

1.3 Research Question

Vibration analysis has three main approaches: time domain, frequency domain and time-frequency domain. Standards are based on time domain, but these techniques are still limited in bearing fault detection, while frequency and time-frequency domain analysis are more powerful and have a wider variety of techniques, these techniques are called *transforms*, the most known are *Fourier transform*, *Envelope Spectrum*, *Wavelet transform*, *Wigner-Ville distribution*, *Hilbert-Huang transform*, and so on.

A technique that has caught the attention of scientific community is the *Hilbert-Huang Transform (HHT)* which introduces a new method called *Empirical Mode Decomposition (EMD)* that allows to decompose signals into *Intrinsic Mode Functions (IMFs)* which are functions with different frequency band information. The decomposition of signals is relevant because it helps to isolate target information from other components that can be seen as noise that affects the diagnosis. The *EMD* is based in the local properties of the signal allowing a decomposition without losing relevant information; this method has the advantage of decomposing signals without having a background mathematical knowledge of the components embedded in the signal.

HHT traditional methodology has two steps: first, the selection of the component (*IMF*) with the target information, and second, the interpretation of the final results. First step can be tedious and time consuming since the decomposition can produce up to 20 or more *IMFs* (determined by the frequency content of the signal), and they are analyzed one by one. The second step consists in the identification of fault components in the spectrum.

1.4 Solution Overview

HHT methodology for bearing fault detection requires trained personnel that can identify the proper *IMF* with relevant information and then calculate the spectrum. With the aim of reducing

human intervention and save time, the *IMF* selection was automated by using statistical parameters that can distinguish faulty signals.

An analysis of 17 time domain features were explored to determine the proper statistical parameter for the methodology. The selection was made based in the change of amplitude of these parameters from a normal condition signal to a faulty condition signals. First, the *EMD* was applied to the normal condition signals, and the faulty signals. Then, the statistical parameter were computed for each *IMF*. An *Envelope spectrum* analysis was made for each *IMF* to identify the component with faulty information. After the identification of the *IMF* was made, the largest change in amplitude was expected for that *IMF*. From the 17 statistical parameters only *Kurtosis* and *KR (Kurtosis x RMS)* were able to identify the *IMF* with fault information in all of the cases. From these two parameters, the *KR* showed a higher amplitude variation and was selected for the method.

1.5 Main Contribution

The main contribution of this thesis is the development of a methodology based on *HHT* to detect incipient bearing faults. This methodology deals with the tedious and time consuming process of *IMF* selection. This selection can be made in seconds with the exploitation of statistical parameters, and to determine the best statistical parameter, an exhaustive analysis was made for 17 time domain features exploring their capacity to represent faulty signals. It was determined that *KR* is the best parameter to identify *IMFs* with faulty information.

The methodology was validated with three different data sets. As a result, from the 18 signals used in the validation of the methodology only one *IMF* was wrongly selected (94.4% of effectiveness). Also, in 14 cases the spectrum result was clear and the faulty components were predominant. For the remaining four signals other peaks appeared in the spectrum (noise); but, the fault diagnosis was also possible.

1.6 Organization

This research work is organized as follows:

- Chapter 2 presents the state of the art of different approaches for signal analysis in bearing fault detection. Also, this chapter includes theoretical background of the *HHT* and the bearing faults.
- Chapter 3 describes the proposed methodology, and the bearing fault signal databases used to test the performance of the proposed methodology.

- Chapter 4 presents the results of the evaluation of the methodology.
- In Chapter 5 the contributions, conclusions, and future work of this research are presented.

Chapter 2

State of the Art

2.1 Introduction

This chapter is divided in two sections, the first section includes the most relevant investigations in different approaches for bearing fault detection.

The second section is focused on giving a stronger theoretical foundation of the key concepts that are used in the development of the proposed methodology for bearing fault detection.

2.2 Literature Review

Machines or any mechanism in spite of working in good conditions generate vibrations, they are generated by movile parts working within the machine. The constant movement and interaction between them produces the components wear and the vibrations tend to increase. Vibrations can be used to evaluate the machine condition, since they offer direct information of the structure and components of the machine.

The challenge in vibration analysis is how to extract signature information embedded in the vibration signals and associate it with specific faults in the machines. The answer to overcome this problem is the use of specialized techniques to process the signals in vibration analysis. Vibrations signals can be analyzed in three main approaches: time domain, frequency domain, and time-frequency domain. Figure 2.1 shows the classification of signal processing techniques and gives some of the most known and relevant techniques used, e.g. *Fourier Transform (FT)*, *Short-Time Fourier Transform (STFT)*, *Wigner-Ville Distribution (WVD)*, *Hilbert-Huang Transform (HHT)*, and *Wavelet Transform (WT)*.

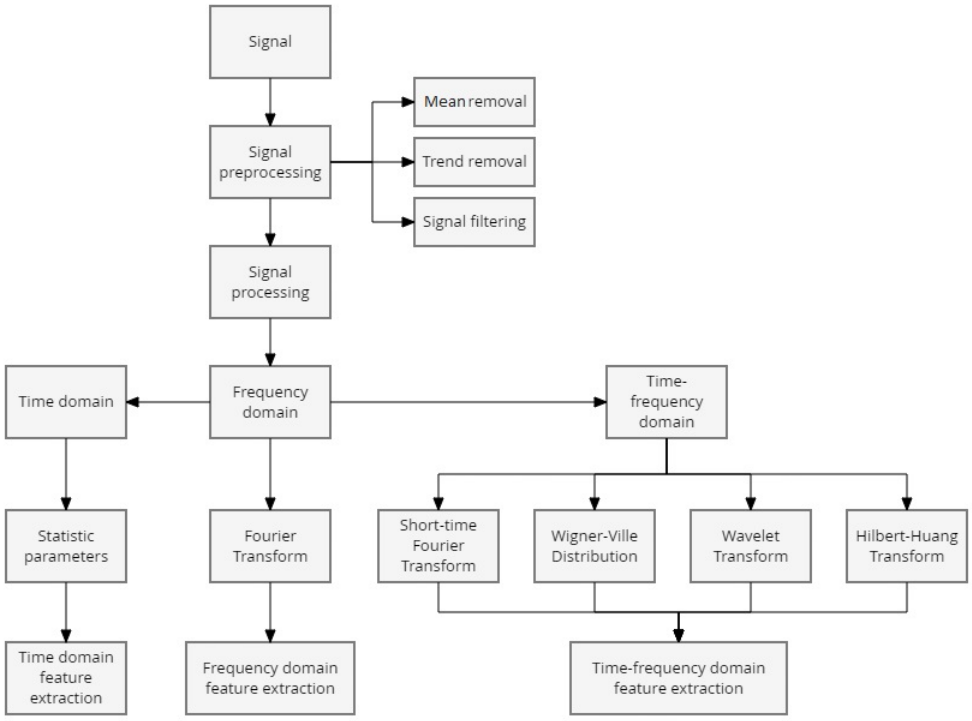


Figure 2.1: Signal processing and feature extraction [Lei, 2016]

2.2.1 Time domain analysis

Time domain analysis is the simplest and most common manner to analyze a signals. Signals from the sensors are captured in time domain and this type of analysis consists in the measurement of parameters (often statistical) directly related to the form of the signal. The existing guides for condition based maintenance are based on this approach.

The *International Organization for Standardization (ISO)*, is an organization that collaborates with specialized scientists and engineers in their areas to create detailed guides that help to improve the processes applied in industry by answering the question: *what is the best way of doing this?*. This organization has published some guides for rotating machinery.

ISO 17243-1: Evaluation of machine tool spindle vibrations by measurements on spindle housing. This standard provides information on how to assess the severity of machine tool spindle vibrations measured on the spindle housing. The part 1 of the standard corresponds to machine tools with nominal operating speeds between 600 and 30,000 RPM. The document describes the preliminary operations such as process load and spindle conditions besides of the measurement and the evaluation in the results. This norm emphasizes the use of *RMS* vibration velocity as assessment of long term spindle condition and the use of *RMS* vibration acceleration for short term spindle condition [ISO, 2014].

ISO 10816-1: Evaluation of machine vibration by measurements on non-rotating parts. The

standard establishes general conditions and procedures for the evaluation of vibration using measurements made on non-rotating parts. The part 1 presents the evaluation criteria related to the vibration and the part 3 gives criteria for assessing vibration levels when measurements are made in situ and applied to machine sets having a power above 15 kW and operating speeds between 120 and 15,000 RPM, the units for the vibration levels are in *RMS* [ISO, 1995] [ISO, 2009].

ISO 2954: Requirements for instruments for measuring vibration severity. This international standard specifies the measurement instrument requirements for vibration severity of machines, particularly when making repeated measurements for trend monitoring of a certain machine. The instruments covered by this international standard give direct indication of recording of *Root-Mean-Square (RMS)* vibration velocity that is defined as a general measurement unit, and *RMS vibration acceleration is recommended for high-speed machines and for rolling element bearings* [ISO, 2012].

2.2.2 Frequency domain analysis

Vibrations can be defined as *periodic* back-and-forth motions about an equilibrium point, produced as response after an external force disturbed the system steady conditions. The term periodic implies that these movements are repeated through time and have a corresponding frequency. Frequency domain analysis is focused on taking advantage of these characteristics of vibration signals and complement it with the force magnitude that is already used in time domain analysis. Frequency domain analysis is conducted with the assertion that individual components produce specific vibrations with their frequencies.

This new domain introduces frequency as a new parameter and requires a new form to display the information, in time domain the amplitude of the signal is plotted through the time, but in frequency domain data is visually represented by plotting the spectrum. *Spectrum* is defined as the arrangement of amplitude components of a set of waves of a vibration signal according their frequencies, for this reason the time axis is now replaced by the frequency axis.

For a simple periodic signal, the frequency can be easily calculated by plotting the signal in time domain, calculate the period (time that takes the system to move from the equilibrium point to an extreme, pass through the equilibrium point to the other extreme and return to equilibrium point), and then calculate the frequency $f = 1/\tau$, in the equation f is the frequency $cycles/second$ or Hz and τ is the period in $seconds$.

The simplicity of this equation is deceiving, because in real application machine vibration signals are the result of the combination of multiple signature signals, and it is hard to calculate the frequencies embedded using the method described before. To calculate the frequencies of real application signals, *integral transforms* are used to convert a signal from one domain to another, it can be from time to frequency domain or from frequency to time domain.

A characteristic of integral transforms is the *kernel function*, also known as mapping function

[Arfken and Weber, 2005], which is a bivariate (t for time and i for frequency) function that converts a continuous function from one domain to another [Courant and Hilbert, 1953], in this case, time and frequency, the difference between transforms is this kernel function, which has the waveform of the components to be converted from one domain to another. If the function that is being converted does not match the kernel, it will result in an incorrect transformation and the result will have no physical sense.

Basic frequency domain techniques for vibration analysis are *Fourier Transform (FT)*, *Power Spectral Density (PSD)*, *Cepstrum*, and *Envelope spectrum*. They perform well in simple applications, but they are limited when analyzing complex signals as in the case of bearing fault signals.

Fourier Transform in Bearing Fault Detection

FT is the most known and applied frequency domain tool, is used with the assumption that the signal only has periodic components. *FT* can be computed as:

$$F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(t) e^{-i\omega t} \, dt \tag{2.1}$$

where $F(t)$ is the *FT* of the signal, $X(t)$ is the signal, and $e^{-j\omega t}$ is the kernel function used for mapping sines and cosines from one domain to another; but, if a signal is a non-periodic function it will be interpreted as a periodic one, where the period tends to infinity and consequently the fundamental frequency tends to zero [James, 2011] producing elevated amplitude components that will be observed as peaks at low frequencies in the spectrum.

Equation (2.1) is used to convert signals expressed as mathematical functions, which is not the case when vibrations are recorded, to compute the *FT* using digital data there is an approach that is called *Fast Fourier Transform (FFT)* and introduces parameters as: the number of data points of the signal, window selection and sample frequency that affect the frequency resolution in the spectrum, further information can be found in [Sinha, 2014] and [Goldman, 1999]. The *FFT* has the same disadvantages as the *FT* of only mapping sines and cosines, and it adds the disadvantage of the selection of a proper signal data-length and sampling frequency to obtain good resolution in the spectrum and be able to make correct conclusions.

2.2.3 Envelope Spectrum

Vibration signals are the sum of multiple waves that vary in amplitude and frequency, a phenomena that occurs in many applications is that low-frequency and low-amplitude vibrations often excites the natural frequency of the system producing high-frequency and high-amplitude vibrations. The high frequency is called *carrier frequency* and when this occurs the low-frequency and low-amplitude vibrations are modulated by the carrier frequency causing the component being modulated to disappear from the spectrum when the *FT* is applied. This phenomena can cause a

misinterpretation of the spectrum and lead to wrong decisions, to determine the characteristics of the components being modulated the removal of such carrier frequency from the signal is essential [Sinha, 2014]. The technique to overcome this problem is called *Envelope spectrum*.

Envelope spectrum consists in two steps: first construct the envelope of the signal, this process is called *demodulation* and second apply the *FT* to the envelope.

An envelope is a new signal that covers the carrier signal and often has the form of the signal being modulated. It is known, that there are three methods to construct the envelope of signal: filter-demodulation, high-pass absolute demodulation, and the *Hilbert Transform (HT)* [Tsao *et al.*, 2010]. *HT* method is the most known and used.

2.2.4 Time-Frequency domain analysis

Time-frequency domain analysis allows to analyze the frequency components of the signal through time, for this type of analysis the representation of the data is in a 3D plot called *spectrogram* (*x-axis = time* , *y-axis = frequency* , and *z-axis = amplitude*). This representation allows to observe changes in frequency contents or changes in amplitude in a lapse of time to make inferences about the signals being analyzed. For such analysis techniques like *STFT*, *WVD*, *WT* and *HHT* are used.

Wavelets Transform (WT)

WT is a kernel based method, which implies that a previous knowledge of the waveform produced by the fault being investigated is needed, and the performance is directly affected by the kernel function used in the transform. *WT* can use more than one kernel or mapping function, in this technique kernels are called *mother wavelets* or *wavelet functions*. *WT* allows to decompose the signal into frequency bands called scales, where their characteristic information is represented as a series of coefficients. The mathematical form of *WT* is:

$$C_n = (X(t), \psi_n) = \int_{-\infty}^{\infty} X(t)\psi_n^*(t)\, dt \tag{2.2}$$

where $X(t)$ is the signal being decomposed, and $*$ represents the complex conjugate of the mapping function ψ_n, the more similar $X(t)$ to $\psi_n(t)$, the larger the value of the coefficient C_n will be [Gao and Yan, 2010]. Compared against *STFT* where the window size is fixed, *WT* uses variable window sizes to extract different frequency components, this can be achieved by scaling (dilation and contraction) and shift (translation on time axis) of the *mother wavelet*. Creating a new mathematical form:

$$wt(s,\tau) = (X(t), \psi_{s,\tau}) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} X(t)\psi^* \left(\frac{t-\tau}{s} \right) dt \tag{2.3}$$

where $s > 0$ represent the scaling parameter which determines the time and frequency resolution of the scaled *mother wavelet* $\psi^*(t-\tau/s)$, and τ represents the shifting parameter. Equation (2.3)

is known as *Continuous Wavelet Transform (CWT)*. Compared against the *FT* that is plotted in 2D plot, the scaling and shifting parameters enables the *WT* to be represented in a 3D plot that is called *scalogram* which represent time, scales or pseudo-frequency, and coefficients that represent the similarity between the signal and the *mother wavelet*.

A variation of the *CWT* is the *Discrete Wavelet Transform (DWT)* which reduces the computational time by implementing pairs of low-pass and high-pass wavelet filters, producing two main frequency bands, this process can applied to the resulting scales producing 2^n frequency bands where n is called the level of decomposition.

Other types of *WT* algorithms are the *Wavelet Packet Transform (WPT)* which focuses in provide detailed information from high frequencies, and the *Second Generation Wavelet Transform (SGWT)* that is an alternative implementation of the *DWT* [Yan *et al.*, 2014].

HHT in Bearing Fault Detection

HHT is an adaptive method that allows to decompose the signal into frequency bands similar to the *WT*, with the big difference of not being a kernel based method and eliminating the necessity of knowing the behaviour of the fault signal for a good performance in the decomposition and reducing the risk of a bad kernel selection. This is achieved by a method called *Empirical Mode Decomposition (EMD)*, which decomposes the signal based on local properties and not global properties (kernel based).

The research and contributions in *HHT* can be divided in two main areas: (a) advances in the performance of the *EMD* and (b) in hybrid methodologies, the advances in performance include pre-filtering techniques to enhance the *EMD* performance, Fig. 2.2. Norden E. Huang worked in the *Ensemble Empirical Mode Decomposition (EEMD)* and others researchers made similar numeric algorithms *EMD* as they are *Local Mean Decomposition (LMD)* [Smith, 2005], and *Hilbert Vibration Decomposition (HVD)* [Feldman, 2006]. In the area of hybrid methodologies some works to improve the *IMF* selection, which is an important step to obtain a good result in the analysis, combination of *EMD* with other techniques (transforms), and works focused on fault classification with help of *Artificial Neural Networks (ANN)* and other intelligent models.

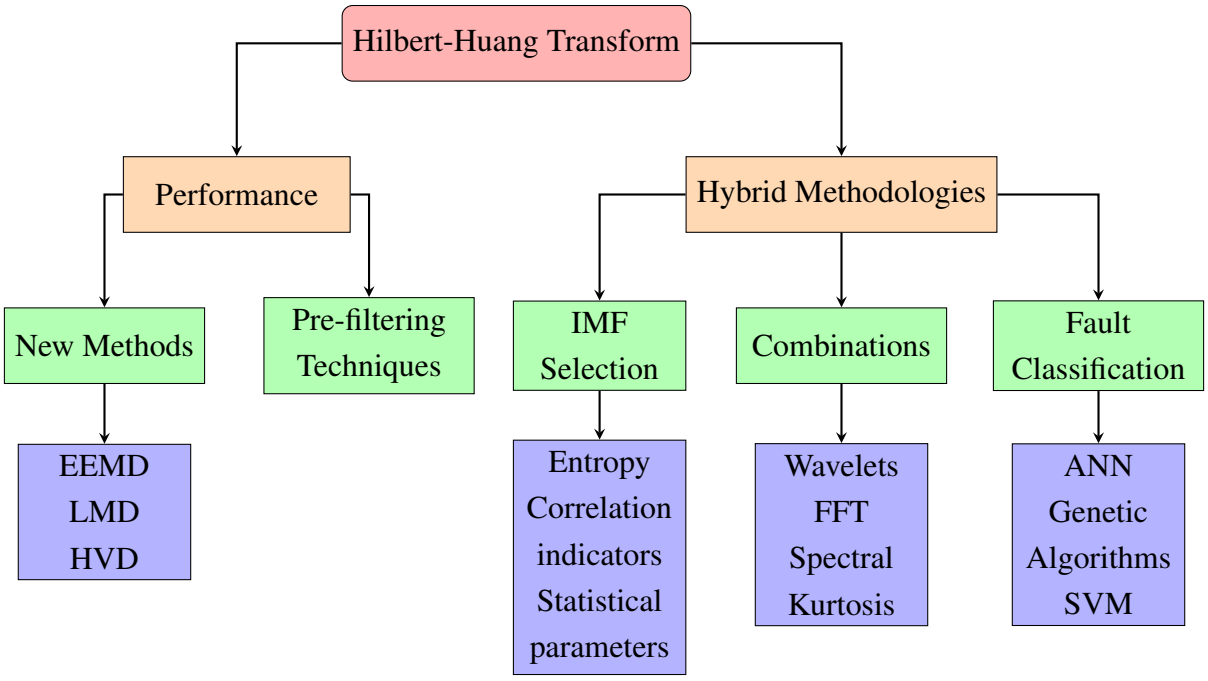


Figure 2.2: Research areas in HHT

An important research work that highlights the performance of the *HHT* over *STFT* and *WT* was made by Yan and Gao [Yan and Gao, 2006] where they applied *HHT* for machine health monitoring. In this research they made a comparative study of the performance of *STFT*, *WT*, and *HHT* on a synthetic signal with defect-induced transient components with time varying frequencies, these defect-induced were modelled as a series of exponentially decaying oscillations simulating a bearing fault. This comparative study showed that for having an acceptable performance the *STFT* depends of the window size selection while applying the technique. For *WT*, three different *mother wavelets* were used (*Db2*, *Meyer*, and *Morlet*) where *Morlet* gave the best result between them, but not as good as expected, and in the third case *HHT* got the best result over *STFT* and *WT* revealing that the constrain of window size selection for *STFT* and the proper selection of *mother wavelet* function for *WT* are overcomed by the *HHT*. After this analysis, they applied *HHT* for they seeded an outer race defect in a bearing mounted on an experimental test bed developed. They made a successful detection by plotting the *Hilbert Spectrum* and performing a visual inspection.

Li et al. used the *Hilbert Marginal Spectrum* [Li *et al.*, 2009]. The usual manner of identifying a fault on *HHT* is by plotting the *Hilbert Spectrum* which is a time-instantaneous frequency-amplitude distribution, instead they used the *Marginal Spectrum* proposed by Huang et al. [Huang *et al.*, 1998], which is the spectrum plot of every *IMF*, but the frequency is replaced by the instantaneous frequency of the *HHT* instead of using *FT*. This spectrum has a different meaning than the traditional spectrum, since the instantaneous frequency is based on the local properties of the signal it gives a measurement of probability of that frequency existing in the *IMF*, while the traditional

spectrum is based on the global properties of the signal and if a frequency appears means that is in the whole signal. They summarized their methodology as follows: (1) compute the envelope of the signal, (2) apply the *EMD*, (3) calculate the *Marginal Spectrum* of every *IMF* and (4) search for the *IMF* with the objective fault frequency, (5) analyze and (6) draw a conclusion.

Gao et al. [Gao *et al.*, 2008] made numerical experiments to the *EMD* to demonstrate that *EMD* performance is affected by the noise and therefore the resulting *IMFs* are also affected. They believe that the noise possibly introduces some *"false local extrema"* to the signal causing an over-decomposition of the signal while applying *EMD*. To reduce these effects, they propose the *Combined Mode Function (CMF)* that consists in combining two or more *IMFs*. There are two procedures to determine if *CMF* should be applied: the first is inspecting the waveform of every *IMF*, and if two *neighbor IMFs* waveforms vary abruptly simultaneously in a lapse of time the *CMF* should be applied. The second procedure is computing the instantaneous frequency of every *IMF*, if the instantaneous frequency of two neighbor *IMFs* change suddenly during a lapse of time, *CMF* should be applied. They applied the method to a power generator, with plain bearings, from a thermal-electric plant in China that had high levels of vibration, they applied *FT* and *WT* with *Db8* as *mother wavelet* and they could not identify the problem, when the generator was stopped they realized that the plain bearings were broken, they used the *EMD* with the *CMF* method and then applied Fourier to the resulting *IMFs* and they did identify that there were not oil circulation and it was not lubricating the plain bearing causing impacts and friction in the surface leading the bearing to break. The root cause of the problem was only identified by using *CMF*.

As well as Gao et al. [Gao *et al.*, 2008], Huang and Wu realized that the *EMD* performance was affected by the noise and worked in the *EEMD* a *Noise-Assisted Data Analysis (NADA)* method [Huang and Wu, 2009]. This new method is focused in solving the mode mixing caused by the noise, because a bad decomposition can lead to misinterpretations of the physical meaning of the *IMFs*; also, this method was designed to reduce the end effect of the *IMFs* as result of the decomposition.

This new method is inspired in the *dyadic filter bank* properties of the *EMD*, when *EMD* is applied to white noise it works as a dyadic filter bank. A dyadic filter bank is a collection of band-pass filters that have constant band pass shape (e.g. a *Gaussian* distribution), but with neighboring filters covering half or double of the frequency range of any single filter in the bank. The frequency filters can be overlapped. For example a dyadic filter bank can include filters covering 50-120 Hz, 100-240 Hz, 200-480Hz, and so on [Huang and Wu, 2009]. Intermittent signals can affect this properties, and to overcome this problem *EEMD* is performed with help of noise.

The method consists in applying series of *Gaussian* noise. the process of extraction of *IMFs* from the signal change and can be described as follows:

1. Add a white noise series to the signal.

2. Decompose the signal with added white noise into *IMFs* using the traditional *EMD*.

3. Return to step 1 and add a different white noise series. This is repeated a predefined number of iterations, but with different white noise series each time.

4. Compute the ensemble by calculating the mean value of every *IMF* among the different results of the decomposition.

As it can be observed *EEMD* defines the true *IMFs* components as the mean of an ensemble of trials, where each trial consists in the decomposition of the original signal plus a white noise of finite amplitude, by doing this, the different possible variations are considered increasing the physical sense of the *IMFs*. The problem of this method is that there are no equations to choose the white noise and its amplitude for every trial, and there is no rule to select the number of iterations to compute the ensembles, and the performance is affected by the selection of these values. The choice of the optimal parameters is still a research topic [Osman and Wang, 2016a], another disadvantage is that the time to perform the *EEMD* is *"n"* times longer than the *EMD*, since *EMD* is repeated in every trial, where *n* is the number of trials.

[Li *et al.*, 2010] worked in bearing fault diagnosis using *HHT* and replaced the traditional *EMD* by applying the *EEMD* instead. They detected *Inner Race (IR)* and *Outer Race (OR)* defects successfully in a visual manner proving that *EEMD* works for bearing fault detection; but, they did not compare the *HHT* with *EEMD* against the traditional method or any other method.

Another effort to improve the performance of *HHT* in bearing fault detection was made by Wu [Wu *et al.*, 2016], they worked on eliminate the effect of rotating speed fluctuations during the diagnostic process, which modifies the bearing fault frequencies, their solution was using instantaneous frequency normalization, this normalization is made by computing an envelope of the *IMF* by using a cubic spline curve (cubic interpolation) and dividing the original *IMF* over the envelope, this normalization of *IMFs* reduces frequency fluctuations within the *IMFs*; then, instead of applying the *HT* to construct the analytic signal and compute the instantaneous frequency, they used the *Generalized Zero-Crossing (GZC)* and *Direct Quadrature (DQ)* methods. *GZC* is computed by dividing the period in quarters and giving them different weighting factors to every one; and the *DQ* is calculated by differentiation of the normalized *IMFs*. They also normalized the fault frequencies to detect bearing faults to compute the spectrum, and used *Support Vector Machine (SVM)* to classify the faults.

In the area of pre-filtering techniques to improve the performance of *HHT*, [Osman and Wang, 2013] proposed an improved technique for bearing condition monitoring called *enhanced HHT (eHHT)*, this technique consists in two steps: (1) eliminate the noise of the signal to highlight the defect-related impulses for a better identification of the bearing faults, and (2) the *eHHT*. For the denoising of the signal they used a *Minimum Entropy Deconvolution (MED)* filter, which was

firstly proposed by Wiggins [Wiggins, 1978]; it helps to reduce the impedance effect of the transmission path of the signal acquisition, *Entropy* minimization is achieved by maximizing the signal *Kurtosis*. The selection of the parameters for the filter design are determined by trial and error for a specific level of noise, if the conditions of application change a new filter is needed to be designed complicating the application in different working conditions. The second part consists in applying the *EMD* and to obtain the most distinguished *IMF* that are related to a signature fault they propose a correlation measurement, *Normalized Correlation Measure (NCM)* and *Deficiency Mutual Information (DMI)*. *NCM* is an indicator for linear similarity between two distributions, while *DMI* provides a nonlinear similarity measure to characterize the uniqueness of the extracted *IMFs*. The *NCM/DMI* is calculated to every *IMF* by the sum of both measurements. The *IMF* with the highest *NCM/DMI* indicator value is selected, the selected *IMF* will have minimum *Entropy* and maximum correlation information, this procedure of *IMF* selection is called *eHHT* technique.

[Osman and Wang, 2016b] made another investigation and they proposed the method called *Normalized Hilbert-Huang Transform (NHHT)*, this method includes a *Maximum Kurtosis Deconvolution (MKD)* filter to demodulate the effect of the impedance in the signal transmission path by reducing the *Entropy* and maximizing *Kurtosis*, *MKD* filter is similar to the *MED* filter of [Osman and Wang, 2013] and it was also proposed by Wiggins. The second step is to apply the *EMD* and decompose the signal into *IMFs*. To choose the *IMFs* a D'Agostino-Pearson normality test is applied to determine their contribution to the information in the signal, then the *IMFs* with the highest values are selected and analyzed applying the *HT*. To validate this methodology they compared the results of the *Normalized HHT* method using *MKD* filter against the *Normalized HHT* without filter, and concluded that using *MKD* with *NHHT* gives a better result in the detection of faults in *RE*, *IR* and *OR* defects. They conclude that *MKD* is effective for denoising the signal, but the filter parameters must be selected by trials and errors so further research is being made for developing an adaptive method for the filter.

[Osman and Wang, 2016a] proposed the *Morphological HHT (MHHT)* technique , this method consists in the design of a mathematical morphological filter to reduce the noise of the signal produced by the impedance of the transmission path to highlight the impulses produced by the faults in the bearing, Morphological filters work to extract different types of information from the signal [Maragos and Schafer, 1987] by the convolution of the signal and a structuring element which is a function, the problem with the filter design is choosing a value of structuring element length, the proper value is calculated by trial and error and depends of the application. Once the filter is designed and applied to every *IMF* the *NHHT* is applied. This technique was applied and compared to the *WT*, *HHT*, and *EEMD* without filter and achieved a visual improvement for the fault detection.

In the area of hybrid methodologies, Yu et al. combined *WPT* with *HHT*, the *WPT* was used to decompose the signal into low and high frequency bands, then the high frequency signal is

reconstructed using the *Wavelet coefficients* and its envelope is computed by using the *HT*. This new signal was decomposed by the *EMD* and then the *Hilbert Marginal Spectrum* is applied to detect *IR*, and *OR* faults [Yu *et al.*, 2005]. This effort was made to improve the efficiency of the *Hilbert Marginal Spectrum*.

Djebala and Babouri proposed an hybrid method based on *EMD* and *Wavelet Multi-Resolution Analysis (WMRA)* also called *Discrete Wavelet* for bearing fault detection [Djebala *et al.*, 2015], they applied the *EMD* to the raw signal to decompose it into *IMFs*, then the *FFT spectrum* and the *Kurtosis* for every *IMF* was calculated, the selection of the *IMF* was made by searching the *IMF* with the *highest Kurtosis* and with important frequency components in the spectrum. The *Kurtosis* is used to guide the selection of the *IMF* with bearing fault because it is a sensitive shock indicator, once the *IMF* is selected an envelope is constructed by applying the *HT* and the *WMRA* is applied to decompose the signal, the frequency band with the highest *Kurtosis* is reconstructed with the wavelet coefficients and then the *Envelope spectrum* is performed. This method improved the visualization of the fault compared against the *EMD* or *WMRA* alone.

Another example of a hybrid methodology, but with different approach was made by Yu et al., they proposed the *EMD energy Entropy* combined with an *Artificial Neural Network (ANN)* for bearing fault diagnosis [Yu *et al.*, 2006], this method consists in the decomposition of the signal into *IMFs* using the *EMD*. Contrary to *Wavelet decomposition*, *EMD* does not suffer of energy leakage for this reason the sum of the energy of each *IMF* is equal to the energy of the original signal, then the *energy Entropy* is calculated to every *IMF* and normalized using the energy of all *IMFs*. The *energy Entropy* distribution between the *IMFs* of a healthy bearing is used as baseline, when a fault occurs components will appear in some *IMFs* end they energy will decrease changing the baseline distribution. The vector of *energy Entropy* values is used as input for an *ANN* that is trained by using Back Propagation algorithm. This method was compared against *WPT* using a *Db10 mother wavelet* obtaining a total of 8 frequency bands, the classification rate of the by using the *EMD* is superior to that *WPD*.

Yang et al. worked in a bearing fault diagnosis technique based on the *IMF Envelope spectrum* and *SVM* [Yang *et al.*, 2007]. It is known that bearing faults often are modulated in high frequencies and envelope analysis method provides an important and effective approach to analysis fault signals of high frequency impact vibrations, for this reason the signal was decomposed by using the *EMD* and the *Envelope spectrum* was computed for every *IMF*, then they made two characteristic amplitude ratios based on the bearing fault frequencies peak amplitudes, the first ratio was the *OR* over fault in *IR* frequency amplitudes and was a ratio of fault in *Rolling Element (RE)* over fault in *IR*, these two ratios were used as input for the *SVM* for a proper fault classification.

Dubey and Agrawal proposed a method for bearing fault classification using *ANN* in combination with *Hilbert footprint analysis* [Dubey and Agrawal, 2015]. They proposed to apply the *EMD* to decompose the vibration signals into *IMFs*, then the *HT* was applied to create the enve-

lope of the *two* first *IMFs* with the assumption that they contain the fundamental fault information of the signal, the next step is to create the footprint of the first and second *IMF* the $footprint_1$ and $footprint_2$ respectively, which are vectors with the *peak to peak-value* of every envelope, the *energy Entropy* and the *Kurtosis*. These vectors work as input for an *ANN* classifier. They used the data set from *Case Western Reserve University (CWRU)* and achieved 100% on accuracy classification, they compared using *Extreme Learning Machine (ELM)* and *SVM*, but the performance of the *ANN* was superior.

Junsheng et. al. proposed a fault extraction approach based on *EMD* and an *Auto-Regressive (AR)* model for roller bearings [Junsheng *et al.*, 2006]. *AR* models can not be applied to non-stationary signals like the produced by bearing faults, to overcome this problem they applied the *EMD* to decompose these non-stationary signals into *IMFs* which are stationary and simpler than the original signal, they applied the *AR* model to detect fault in *IR* and *OR* faults. This work was not based on the spectrum, but it highlights the advantages of the *EMD* to reduce the complexity of signals. This method does not need a mathematical knowledge of the bearing, but the signals with different bearing faults are needed to create the *AR*.

Tsao et al. worked on the methodology to select the proper *IMF* to perform the envelope analysis and is based on the resonant-frequency of the system. The first step is to calculate the spectrogram of the signal using *FFT* and identify the resonant frequency, then the *EMD* is applied to the original signal and the *FFT* is plotted for every *IMF*, then the resonant frequency is searched in the *IMFs* to select the one with those frequencies, after the *IMF* is selected the *Envelope spectrum* is plotted to search for bearing fault characteristic frequencies. This process is the most simple but it requires human intervention to perform that comparison analysis to select the proper *IMF* [Tsao *et al.*, 2010][Tsao *et al.*, 2012]. In Table 2.1 further information of the case of study and the type of faults detected with these methodologies can be found.

Table 2.1: Comparison of research projects using *HHT*.

References	*Defects*	*Case Study*	*Technique*	*Additional Analysis*	*Classifier and efficiency*
[Yu *et al.*, 2005]	*OR*	Test Rig 1,500 RPM	*WPT* and *EMD* + *HMS*	Does not apply	Visual
[Yan and Gao, 2006]	*OR*	Test Rig Not specified	*HTT*	Does not apply	Visual
[Yu *et al.*, 2006]	*IR, OR*	Test Rig 1,500 RPM	*EMD Energy Entropy*	Does not apply	ANN 93 %
[Junsheng *et al.*, 2006]	*IR, OR*	Test Rig 1,500 RPM	*EMD* and *AR Model*	Does not apply	AR Model 93.33 %
[Yang *et al.*, 2007]	*IR, OR*	Test Rig 1,500 RPM	*EMD* and *Envelope Spectrum*	Does not apply	SVM 100 %
[Li *et al.*, 2009]	*IR, OR*	Test Rig 1,500 RPM	*HHT Spectrum Marginal Hilbert*	Does not apply	Visual
[Li *et al.*, 2010]	*IR, OR*	Test Rig 1,500 RPM	*EEMD* and *HHT*	Does not apply	Visual
[Tsao *et al.*, 2010]	*IR, OR*	Test Rig 1,500 RPM	*EMD* and *Envelope Spectrum*	Does not apply	Visual
[Osman and Wang, 2013]	*IR, OR* and *RE*	Test Rig 1,800 RPM *CWRU* web data 1,735 RPM	*eHHT*	*Minimum Entropy Deconvolution (MED)*	Visual
[Djebala *et al.*, 2015]	*IR, OR* and *RE*	Test Rig 3,000 RPM and 1,800 RPM	*EMD* and *DWT*	*Kurtosis* and *Envelope Spectrum*	Visual
[Dubey and Agrawal, 2015]	*IR, OR* and *RE*	*CWRU* web data 1,736 RPM	*EMD* and *Hilbert Footprint Analysis*	*Entropy* *Kurtosis* *pk-pk value*	SVM 97.5 % ELM 98.5 % ANN 100 %

Table 2.1: Comparison of research projects using *HHT*. (Continued).

References	*Defects*	*Case Study*	*Technique*	*Additional Analysis*	*Classifier and efficiency*
[Osman and Wang, 2016b]	*IR, OR* and *RE*	Test Rig 900, 1200, 1500,..., 2100 RPM *CWRU* web data 1,735 RPM	*NHHT*	*MKD*	Visual
[Osman and Wang, 2016a]	*IR, OR* and *RE*	Test Rig 600, 1200, 1500,..., 2400 RPM	*Morphological HHT*	Does not apply	Visual
[Wu *et al.*, 2016]	*IR, OR* and *RE*	Test Rig Variable 300 - 750 RPM aprox.	*EMD* and *Frequency Normalization*	Does not apply	SVM 94.33 %

2.3 Theoretical Background

This section presents the concepts within *HHT*, which are needed for the development and implementation of the proposed methodology for bearing fault detection.

2.3.1 Hilbert-Huang Transform (HHT)

This *HHT* for signal processing consists in two steps:

1. The *Empirical Mode Decomposition (EMD)* (main contribution made by Huang).
2. The application of *the Hilbert Transform (HT)* to the result of step 1.

Empirical Mode Decomposition

EMD is a method that was designed to enhance and facilitate the decomposition and analysis of non-linear, and non-stationary oscillatory signals, this method compared against others methods as *WT* does not need a kernel that describes the characteristic behavior of the signals that want to be extracted, *EMD* is totally empirical and is based on the local characteristics of the signal instead of global characteristics (kernel based methods), allowing a better decomposition of the signal. A good signal decomposition gives the opportunity to select a signal with less noise and particularly including signature information of the phenomena being investigated.

The result of the *EMD* are signals or components called *IMFs* and a Residual (r). These *IMFs* have information that belong to many frequencies bands, and the residual represents the trend of the signal.

The *EMD* is based in the concept of trigonometric polynomials. This concept establish that a signal can be expressed as a combination of sines and cosines functions as in *Fourier* series. This can be expressed as:

$$X(t) = a_0 + \sum_{n=1}^{N} a_n \, cos(nt) + \sum_{n=1}^{N} b_n \, sin(nt) \tag{2.4}$$

The eqn. (2.4) can be adapted to the *EMD* as it follows:

$$X(t) = r + \sum_{n=1}^{N} C_n \tag{2.5}$$

where $X(t)$ is the original signal, res represents the residual from the *EMD* and C_n represents the n^{th} *IMF* obtained from the decomposition. The *IMFs* as they are extracted contain information

from the highest to the lowest frequencies.

To start the description of the *EMD* method it is necessary to introduce to key concepts, the concept of *IMF* and monotonic signal. An *Intrinsic Mode Function (IMF)* represents an oscillation mode embedded in the signal and it must satisfy two conditions:

1. The number of extremes and zero crossing in the whole signal must be equal or different at most by one.

2. At any point the mean envelope defined by the local maximum and the local minimum is zero.

These *IMFs* are often both amplitude and frequency modulated, and the two previous conditions allow to obtain a signal centered in zero with all the maximums in the positive side of the "x" axis and all minimums of the signal in the negative side of the "x" axis, and therefore we can say that the *EMD* works as a signal preprocessing method which removes the mean trend, and also filters the signal into frequency bands, all of this resulting in signals (*IMFs*) easier to analyze.

Monotonic Function: is a signal where the *sifting* process can not be applied because it does not have local maximums and minimums, in the *EMD* these signals are the residual of the decomposition and represent a *trend* of the original data.

The *EMD*, flow chart can be found in Fig. 2.3, it is an iterative method consisting in a series of steps:

1. Let $h_{j,k} = X(t)$, and $r = X(t)$, where $X(t)$ is the original signal, j is the number of *IMF* starting by 1, and k is the number of iteration starting by 0.

2. Localize all the local maxima of $h_{1,0}$ and construct an upper envelope E_{upper} interpolating all local maxima, for this process a cubic interpolation is used.

3. Localize all the local minima of $h_{1,0}$ and construct a lower envelope E_{lower} interpolating all local minima, for this process a cubic interpolation is used.

4. Construct a curve using the mean value of both envelopes (E_{upper} and E_{lower}) created in step 2 and 3, this curve will be named m.

$$m = E_{upper} - E_{lower}$$

5. Subtract the curve m to the signal $h_{1,k}$, this operation is represented as:

$$h_{1,k} = h_{1,k} - m$$

6. Verify if $h_{1,k}$ is an *IMF* evaluating the stopping criterion for *IMFs*. *SD* must be between $0.2 - 0.3$. *SD* is set as default to 1 when $k = 0$.

$$SD = \sum_{t=0}^{T} \frac{|h_{1,k-1} - h_{1,k}|^2}{h_{1,k-1}^2}$$

   If the stopping criterion of *SD* is not accomplished, it is necessary to return to step 2 using $h_{1,k+1}$, otherwise go to step 7.

7. Set $h_{1,k}$ as the first component (*IMF*) c_j, and subtract c_j from r.

$$c_j = h_{1,k}$$

$$r = r - c_j$$

8. Verify if r is monotonic.

   - If r is monotonic the process is finished.
   - If r is not monotonic, then $h_{j+1,0}(t)$ is set as r and then and go to step 2 using $h_{j+1,0}(t)$.

$$h_{j+1,0}(t) = r$$

The iterations of steps 2 to 6 is called *sifting* and it has two effects on the construction of the *IMF*:

- Eliminate riding waves in the signal
- To smooth uneven amplitudes from the *IMF* being constructed.

The *SD* criteria is used to guarantee that the *IMFs* components conserve enough physical sense of both amplitude and frequency modulations.

For the construction of the envelope besides cubic interpolation, many complex curve fitting functions were investigated and the improvement achieved does not compensate the increasing in computational load [Huang *et al.*, 1998] for this reason cubic interpolation is commonly used.

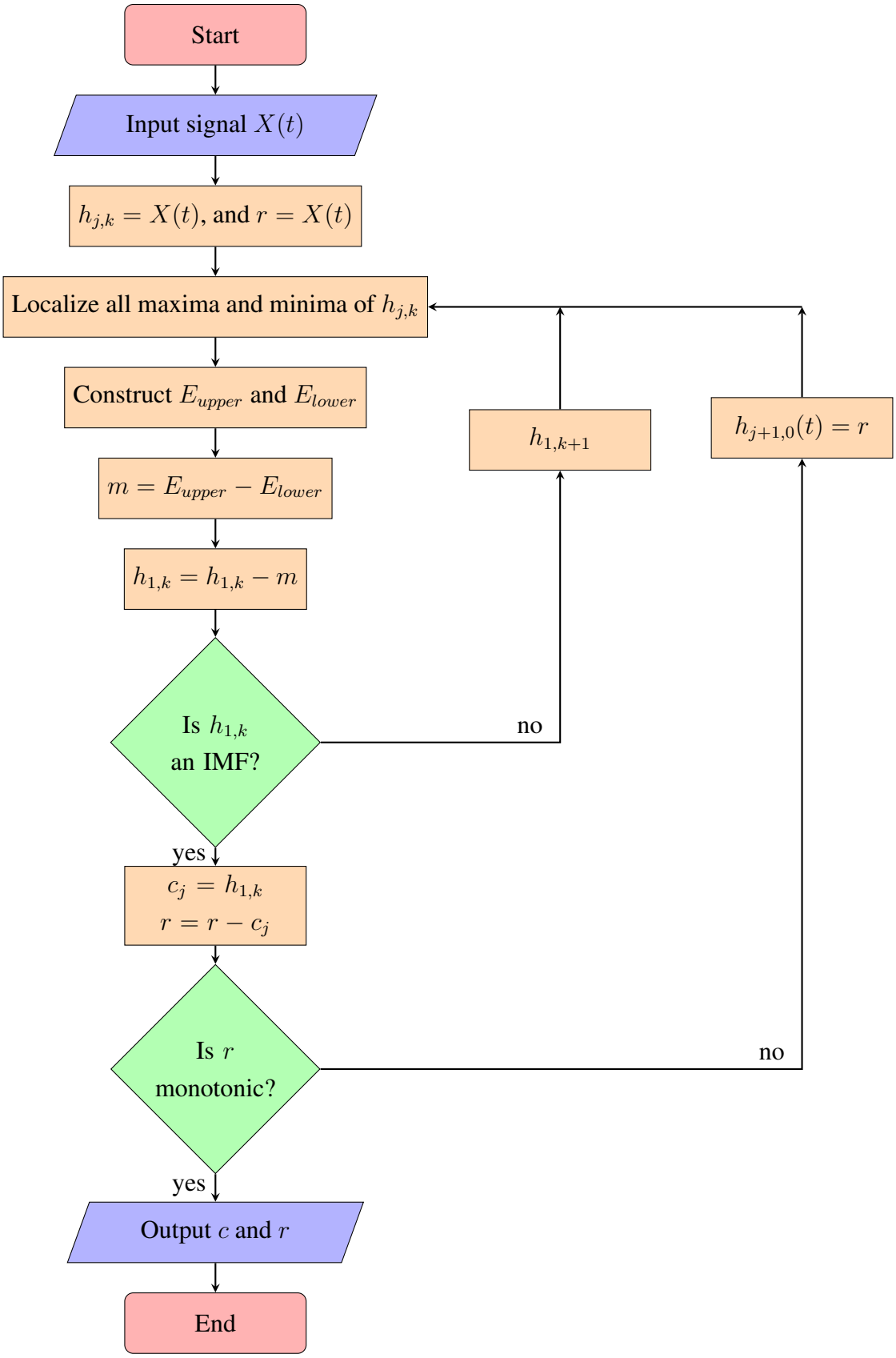


Figure 2.3: EMD algorithm

Hilbert Transform and the Analytic Signal

HT was developed by the mathematician David Hilbert in 1905 [Feldman, 2011a]. The *HT* has being used for *Single Side Band (SSB) Modulation* for data transfer in telecommunications and band pass systems [Briceno, 2012] and for the construction of the analytic signal, also known as quadrature of the signal, for the envelope of the signals. The *HHT* uses the second application (construction of the analytic signal), this transform is expressed mathematically as follows:

$$H[X(t)] = \tilde{x}(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{X(\tau)}{t-\tau}\, d\tau \tag{2.6}$$

where P is the *Cauchy Principal Value*, that has to be considered since the possibility of the singularity when $t = \tau$. According to [Feldman, 2011b], the mathematical integral does not give much information of the properties, anyway the *HT* is a lineal operator that works as a phase shift of $\pi/2$ to the signal being processed, this phase shift allows to create the conjugate of the signal and create the analytic signal that is represented as it follows:

$$X(t) = x(t) + i\,\tilde{x}(t) \tag{2.7}$$

where $\tilde{x}(t)$ represents the *HT* of $x(t)$. The amplitude of the analytic signal $X(t)$, also known as *envelope*, can be calculated using the next equation:

$$A(t) = \sqrt{x(t)^2 + \tilde{x}(t)^2} \tag{2.8}$$

The phase $\Phi(t)$ and therefore the frequency ω can both be calculated using these equations:

$$\Phi(t) = ArcTan\left(\frac{\tilde{x}(t)}{x(t)}\right) \tag{2.9}$$

$$\omega = \frac{d\Phi(t)}{dt} \tag{2.10}$$

where ω is also known as *instantaneous frequency* and represents the frequency in local characteristics ($\frac{d\phi(t)}{dt}$) and not global as in kernel functions. The computation of the amplitude and instantaneous frequency of the analytic signal allows to create the *Hilbert Spectrum*. The *Hilbert Spectrum* $H(\omega, t)$ is a distribution of time-frequency-energy that can be represented in a 3D plot for the signal analysis, and it is the equivalent to the *Fourier spectrogram*; but, because the *Hilbert Spectrum* is constructed after the *EMD* and the properties of the analytic signal that are based in local properties, there is no restriction in time resolution and frequency resolution, this resolution varies only with the sampling frequency used in the data acquisition.

2.3.2 Bearing Fault Frequencies

Bearings are one of the most important components in rotary machinery, in previous section techniques used to extract bearing fault features were introduced.

One of the oldest records of investigations to model the bearing fault frequencies was made by McFadden and Smith, they presented their work *"model for vibrations produced by single point defects in rolling element bearings"* where based on the geometry of the bearing and the rotation frequency of the shaft they developed and proposed the most known and accepted equations for bearing fault detection [McFadden and Smith, 1984].

As it can be observed in Fig. 2.4 bearings are composed of four main parts: outer race, inner race, rolling elements and the cage. The interaction forces between these parts can produce wear and defects in any of these parts, the four possible defects in bearings are *Inner Race (IR)* defect, *Outer Race (OR)* defect, *Rolling Element (RE)* or *Ball* defect, and *Cage* defect.

McFadden and Smith proposed equations to calculate bearing fault frequencies are:

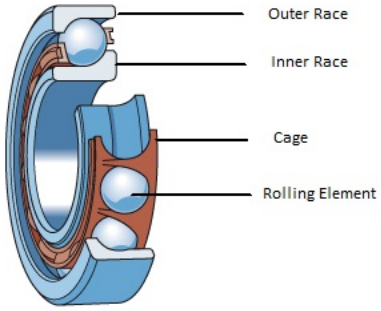


Figure 2.4: Bearing parts [SKF, 2017]

$$BPFO = \frac{nf_r}{2}\left(1 - \frac{d}{D}cos(\phi)\right) \tag{2.11}$$

$$BPFI = \frac{nf_r}{2}\left(1 + \frac{d}{D}cos(\phi)\right) \tag{2.12}$$

$$FTF = \frac{f_r}{2}\left(1 - \frac{d}{D}cos(\phi)\right) \tag{2.13}$$

$$BSF/RSF = \frac{D}{2d}\left[1 - \left(\frac{d}{D}cos(\phi)\right)^2\right] \tag{2.14}$$

where $BPFO$ is *Ball Pass Frequency of Outer Race*, $BPFI$ is *Ball Pass Frequency of Inner Race*, FTF is *Fundamental Train Frequency*, and BSF/RSF is *Ball/Roller Spin Frequency*, n is the number of rolling elements, f_r is the rotating frequency of the shaft, d is the ball diameter, D the pitch diameter and ϕ is the contact angle of the bearing, Fig. 2.5.

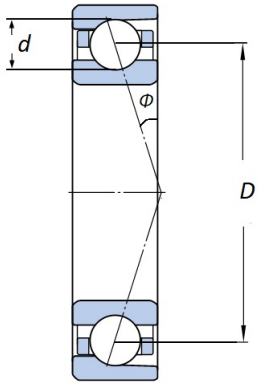


Figure 2.5: Pitch *(D)* and rolling element diameter *(d)*, and contact angle *(ϕ)* [SKF, 2017]

When bearing faults are being diagnosed, previous knowledge of the bearing characteristics as observed in the previous formulas is crucial, this can be a serious problem when trying to implement a bearing monitoring system in rotating machines, because sometimes there is not easy access to the bearings, and some manufacturing companies treat this information as confidential because they are part of the design of the machine.

Between the four type of bearing faults, the *cage* fault is not as common as *IR*, *OR* and *RE* faults, it can be observed that in Table 2.1 no one implemented the methodology for *cage* fault. In the remaining three type of faults, *RE* fault is the most difficult to detect when the rolling element is a ball because each ball has 6 degrees of freedom and factors like contact angle, friction and centrifugal forces lead the ball to have complex motions as rolling and sliding motions at the same time [Niu *et al.*, 2017]. This complex dynamic of the ball modifies the frequency of the ball defect, because when this effect occurs, produces that the defect may or may not make contact with both races of the bearing when rotating. Also, it is known that rotation speed can fluctuate, increasing the variation of the frequencies not only for *RE* fault, but for all types fault, this is an important consideration when performing a spectrum inspection is a variation of $1-2\,\%$ [Randall and Antoni, 2011].

Bearing deterioration is progressive, spectrum can be divided in four zones that help to identify the bearing condition:

- *Zone A*: in this zone rotating frequency and some harmonics can be found.
- *Zone B*: belongs to the frequencies calculated for bearing faults.
- *Zone C*: natural frequencies of the system.
- *Zone D*: *High Frequency Detection (HFD)* zone.

Bearing deterioration can be divided in four stages, the first stage is when a bearing begin to show the minimum wear, when this occurs in *zone D*, which correspond to ultrasonic frequencies (superior to $20\,kHz$), components begin to slightly increase in amplitude. When bearings are in stage 2 and begin to develop minimum pits, sidebands in *zone C* appear, and amplitudes in *zone D* may double form previous condition. When bearings are in stage 3 pits become bigger and components in *zone B* appear for the first time while components from previous stages increase, and finally in stage 4 the bearing pits combine creating rough tracks, and in the spectrum frequency components merge with the defect frequencies producing noise in that zone, and even in *zone A* rotation frequency component increases. Some studies indicate that by the end of the third stage the remaining bearing life is about 1 hour or $1\,\%$ of its average life [Scheffer and Girdha, 2004], this indicates that if a bearing component is found in *zone B* the machine must be stopped. For this reason it is recommended to detect bearing faults when they appear in *zones C* or *D*, while in *zone D* different phenomena can increase the amplitude of the components and complicates fault identifications, *zone C* allows to detect faults when they are incipient. In *zone C* defect signals are modulated in the natural frequency zone of the system and *Envelope spectrum* is essential to detect faults in this zone.

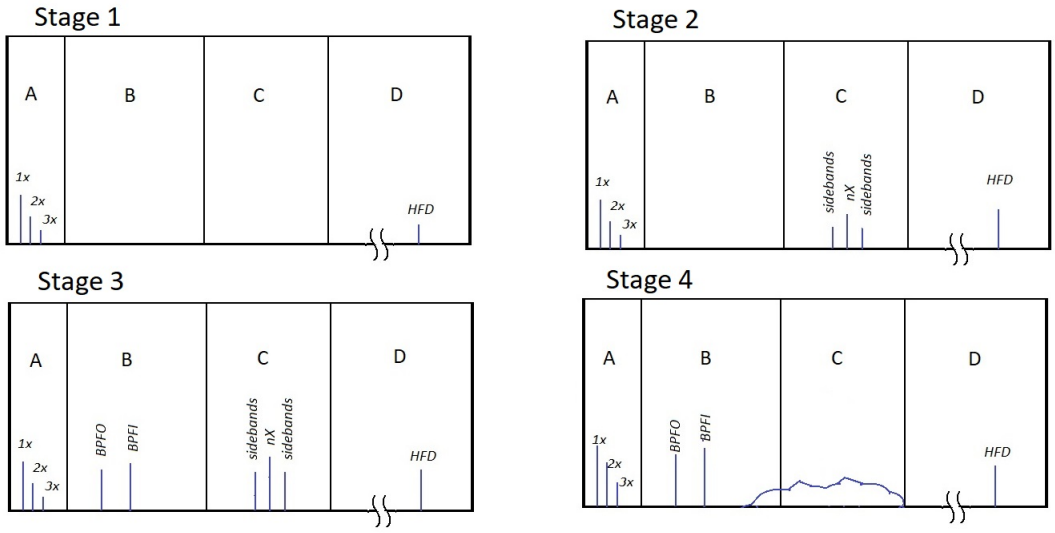


Figure 2.6: This figure shows the components that can be found in Stage 1: first wear indicator, Stage 2: incipient fault, Stage 3: fault condition, and Stage 4: broken condition [Scheffer and Girdha, 2004]

Chapter 3

Proposal

A methodology based on the *HHT* was developed to detect *IR, OR*, and *RE* faults in bearings. The *HHT* has been used for bearing fault detection in many occasions and there have always been a discussion of which *IMF* is the best to make a prognosis. The proposed methodology uses statistic techniques to select the best two *IMFs* to be analyzed.

3.1 Proposed Methodology

The methodology can be implemented by following the steps in the methodology flowchart, Fig. 3.1.

3.1. Signal Acquisition

The signal must be acquired following the recommendations of sensor specifications, sensor location, sampling frequency, and so on, *Appendix D*.

Two of the requirements for bearing fault monitoring systems are low computational time and good visualization to identify the bearing fault components. While increasing the window size of the analyzed signal increases the spectrum resolution, it also increases the computational time of the *EMD*. To achieve a balance in computational time and good spectrum resolution, an analysis for each problem was made in *Appendix B*. The analysis of computational time was made based on the mean computational time of the *EMD* applied to different window sizes of one, five, and ten revolutions. This analysis was made for 16 signals from *CWRU*, and 9 signals of *MFPT society* with different fault conditions (*Normal, OR, IR*, and *RE*). The mean computational time of five revolutions reached up to 10.6 times the mean computational of one revolution, and the mean computational of ten revolutions reached up to 30 times the mean computational of one revolution, where the expected was five and ten times the mean computational time of one revolution respectively. No direct relation was found between window size and computational time; however, the

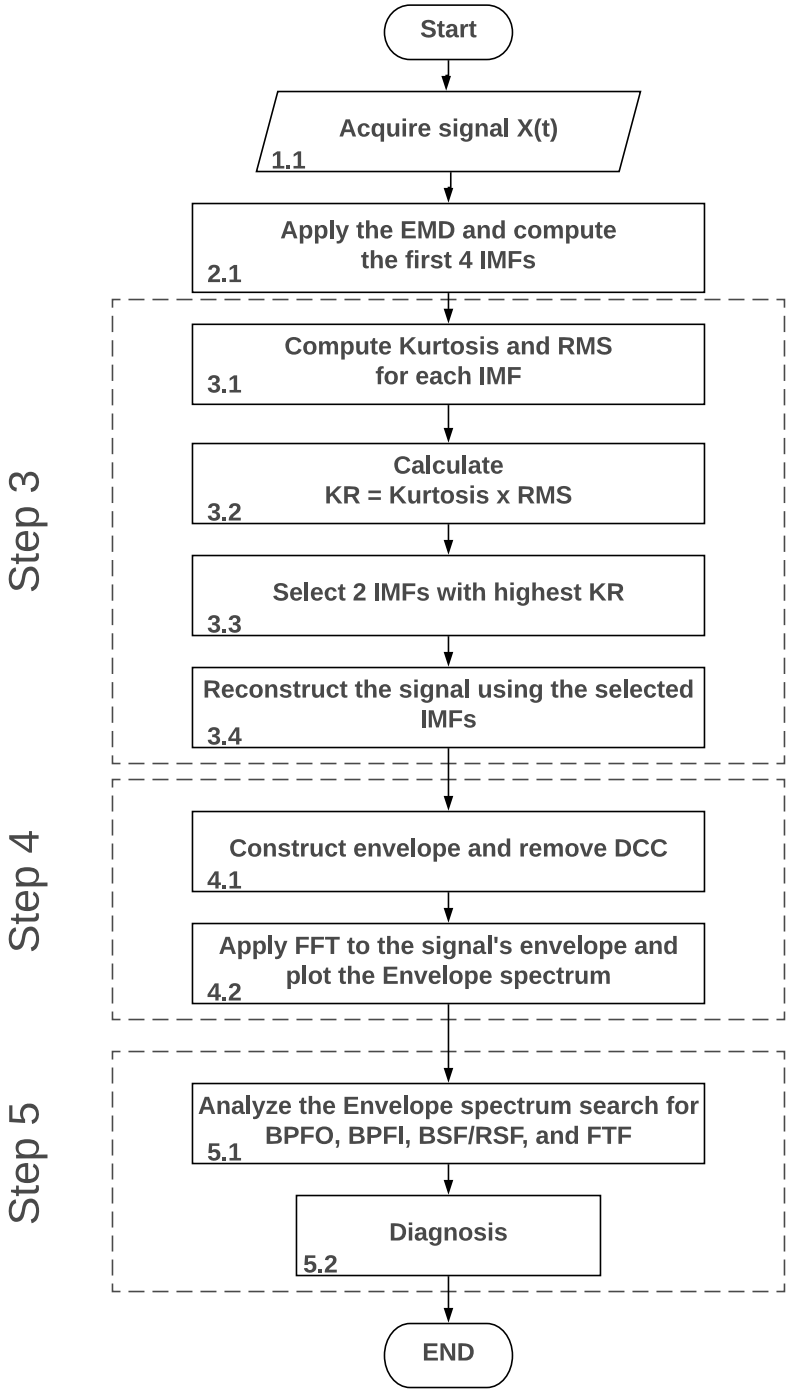


Figure 3.1: Proposed methodology for bearing faults detection

computational time increases more than the expected and suggest to reduce the window size as possible.

The analysis of resolution was made based on different window sizes, the frequency factors for the expected fault frequencies, and the sample frequency of the *CWRU* and *MFPT society* databases. The objective of this analysis was to avoid that one harmonic of the shaft speed and the frequency of any fault was represented for the same peak in the spectrum. It was determined that a minimum five revolutions of the signal produce good resolution in the spectrum.

Based on the analysis of computational time and spectrum resolution, it was determined that five revolutions of the shaft are the optimum for the analysis. However, if greater spectrum resolution is needed, the equations to calculate the optimum window size are provided in *Appendix B*.

3.2. Signal Decomposition

When vibration signals are recorded, noise from the transmission path and information produced by other components are embedded in the signal. The *EMD* uses the *sifting* process to compute the *IMFs* with different frequency bands information and isolate bearing fault information. At the same time, the sifting uses the *standard deviation* as a criterion to stop the process and guarantee that the *IMFs* retain enough physical sense of both amplitude and frequency modulations. However, in literature from the *EMD* the *end effect* phenomena produced by the spline (cubic interpolation) used in the sifting has been reported. This end effect introduces slight errors in the extremes of the *IMFs* that increase as the level of decomposition increases, causing loss of information and physical sense in the *IMFs* that is not detected with the *SD* criterion in the sifting process. This phenomena is deceiving for the method, because it can cause an increase in the computed statistics even though the *IMFs* do not have fault information.

Nevertheless, it was observed that most of the times the fault information can be found within the first three *IMFs*, and in the fourth in rare events. To avoid the components with end defects and still keep the most meaningful components, the *EMD* is stopped after the fourth *IMF* is obtained, *Step 2.1 in Fig. 3.1*.

3.3. Intrinsic Mode Function Selection and Signal Reconstruction

At this point the most reliable *IMFs* are saved; but, the *IMFs* with bearing fault information have not been identified. Two important characteristics of fault signals are impulsiveness and amplitude, these features can be measured using statistical parameters. For this reason, this research studied a total of 17 statistical parameters where 6 of them are amplitude measurements, 8 are distribution measurements and 3 are the combinations of amplitude and distribution measurements, *Appendix C*.

The objective of this analysis was to identify which parameters can represent the characteristics of faulty components to use them as an indicator for the selection of the *IMFs* with fault information. The analysis consisted in the comparison of the amplitude of the 17 statistical parameters for the *IMFs* obtained from the *EMD* from a *Normal* condition signal, against the amplitude of the statistical parameters for the *IMFs* produced for faulty signals, both with window size of five revolutions. A previous *Envelope spectrum* analysis was made for every component to determine which had the fault information, and changes in amplitude were expected for those *IMFs*. A second requirement for statistical parameters was their stability to window size changes, for this test the statistical parameters were computed for every *IMF* as well for the 90%, 80%, and 70% of the *IMFs*. To achieve this, symmetric cuts were made in the extremes, e.g. for 90% of the *IMF* 5% of each extreme was cut. After an extensive analysis it was determined that the *KR (Kurtosis x RMS)* is the best parameter to represent the amplitude and impulsiveness of faulty signals from bearings. Also, as a result of this analysis it was identified that the end effect can trigger the values of statistical parameters, and helped to limit the *EMD* to four *IMFs* to avoid errors.

The *KR* can be directly computed by the use of one equation; however, the *RMS* and *Kurtosis* can be useful on their own. In this methodology, first the *Kurtosis* and *RMS* are computed individually for each *IMF* using eqns. (3.1) and (3.2) respectively, (*Step 3.1 in Fig. 3.1*), and then the *KR* can be calculated with eqn. (3.3), (*Step 3.2*).

$$X(t)_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} y_n^2} \tag{3.1}$$

$$X(t)_{Kurtosis} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{y_n - \mu}{\sigma}\right)^4 \tag{3.2}$$

$$X(t)_{KR} = X(t)_{Kurtosis}\, X(t)_{RMS} \tag{3.3}$$

where N is the number of data points of the signal, y_n is the amplitude of the n^{th} data point, μ is the mean of the signal, σ is the standard deviation of the signal, $X(t)_{RMS}$ is the intensity of the signal, $X(t)_{Kurtosis}$ is the deviation of the signal's amplitude distribution from *Gaussian* ($Kurtosis = 3$ for *Gaussian* distribution), and $X(t)_{KR}$ represent the intensity and distribution of the signal.

The fault information might be spread in more than one component, the reconstruction of one signal using all the components with fault information is possible, but not recommendable. Some *IMFs* have fault information, but they also have high level of noise, a reconstruction using these components can introduce a lot of noise that can make difficult the diagnosis. For this reason in this methodology the reconstruction of the signal is limited to only two components.

To determine which components have the most useful information the two *IMFs* with the largest *KR* amplitude are selected and used to reconstruct the signal, (*Step 3.3*). This selection

is made automatically by programming the methodology. In *Appendix F* the codes for the implementation of the methodology are given.

The *EMD* is based on trigonometric polynomials, and the original signal can be reconstructed by summing all the *IMFs*. To reconstruct the signal without noise, the two selected components are summed, (*Step 3.4*).

3.4. Envelope spectrum

Faults can be identified in the different zones of the spectrum, 2.6. Incipient faults signals are modulated in resonance frequencies of the system, to demodulate the bearing faults the *Envelope spectrum* is used. For this reason, the envelope of the reconstructed signal is computed as described in *subsection 2.3.1*.

When the *IMFs* satisfy the *SD* condition in the sifting process they do not need any preprocessing, but when the signal is reconstructed and the envelope is computed a trend (*Mean trend*) or *Direct Current Component (DCC)* is introduced in the signal. This causes the *FT* to interpret it as a component with infinite period which is reflected as a peak in frequency of 0 Hz that make difficult a diagnosis. Since the *DCC* is present, the next step is to remove it, *Step 4.1*. The *DCC* can be removed with the eqn. (3.4) by removing the *Mean* value of the signal [Lei, 2016]:

$$X(t) = X(t) - X(t)_{Mean} \tag{3.4}$$

where $X(t)$ is the signal, and $X(t)_{Mean}$ is the *Mean* of the signal computed with equation C.7 (*Appendix C*). After the *DCC* is removed the *FFT* is computed and plotted to make the diagnosis, *Step 4.2*.

3.5. Diagnosis

In previous steps, noise and other useless frequencies were removed to facilitate the analysis. With the use of the *KR* the components more likely to present fault frequencies were selected. Now, based on the fault frequency values *OR (BPFO), IR (BPFI), Cage (FTF)*, and *RE (BSF/RSF)*, a visual inspection of the *Envelope spectrum* of the reconstructed signal is made. The analysis of the *Envelope spectrum* for bearing fault detection consists on the identification of bearing fault frequencies, their harmonics, and sometimes sidebands for some fault types (Table 3.1, and Fig. 3.2), *Step 5.1*. When searching for fault components, a possible variation of $1-2\%$ in the computed fault frequencies can be expected as product of speed fluctuations [Smith and Randall, 2015].

Table 3.1: Fault type and expected components in *Envelope spectrum*

Outer Race	*BPFO* and harmonics, no sidebands
Inner Race	*BPFI* and harmonics, sidebands spaced at f_r. Harmonics of f_r
Rolling Element	*BSF* and harmonics, sidebands spaced at *FTF*. Harmonics of *FTF*

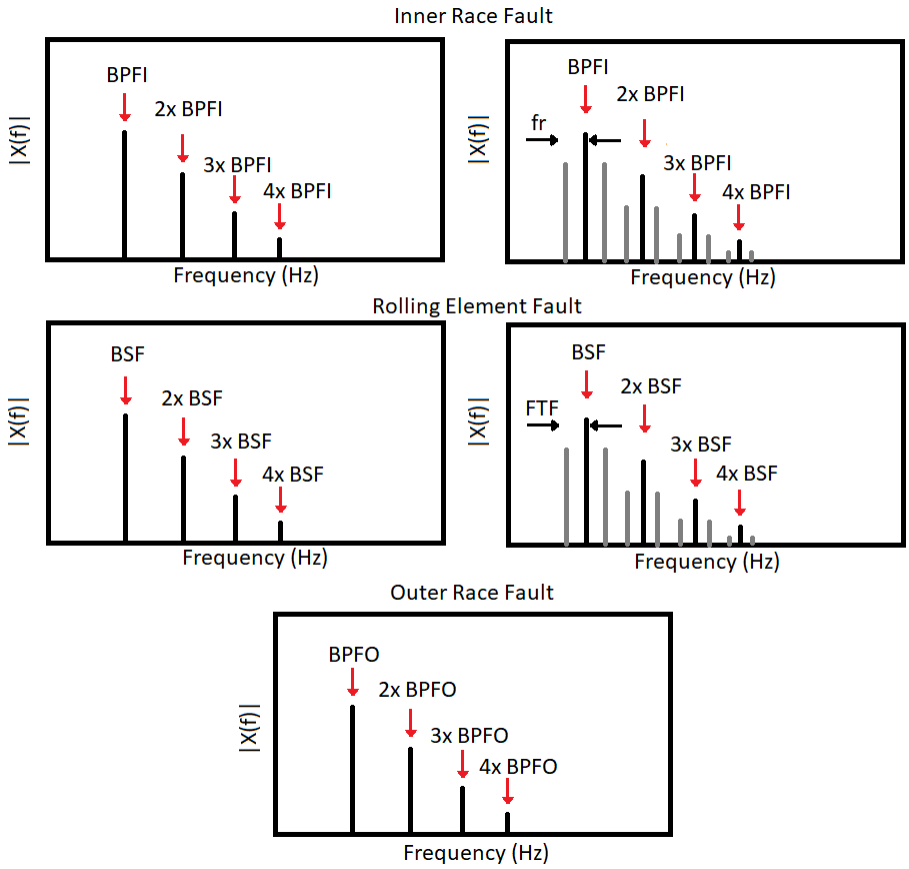


Figure 3.2: Expected components in *Envelope spectrum* for *IR, RE*, and *OR* faults

Finally, the diagnosis is made based on the presence of fault components in the spectrum, if fault components are visible, it can be determined that there is a bearing fault, *Step 5.2*. When a fault is identified it is recommended to confirm the result running at least other test because the machine disassembly to replace the bearings requires specialized personnel.

3.2 Experimental Systems

The description of three different experimental systems is presented. These cases of study are used to test the robustness of the proposed methodology as they present signals with localized

induced defects in bearings. The first experimental system is one of the most known database in the scientific community, it was obtained from *Case Western Reserve University (CWRU)*, and allows to compare new methodologies against previous investigations [CWRU, 1999] [Smith and Randall, 2015]. The second experimental system allows access to signals of bearing faults defects with different dimensions and speeds, the data was provided by T.Y. Wu from *National Chung Hsing University*, Taiwan [Wu *et al.*, 2016]. In the third database, provided by the *Society for Machinery Failure Prevention Technology (MFPT Society)* [Society, 2013], different axial loads were applied to the shaft, while in *CWRU* database they used torsional loads.

3.2.1 Case Western Reserve University (CWRU)

The experimental setup utilized by *CWRU* consists of a 2 HP electric motor, a torque transducer and an encoder. The electric motor was coupled to a dynamometer to apply loads of 0, 1, 2, and 3 HP, Fig. 3.3.

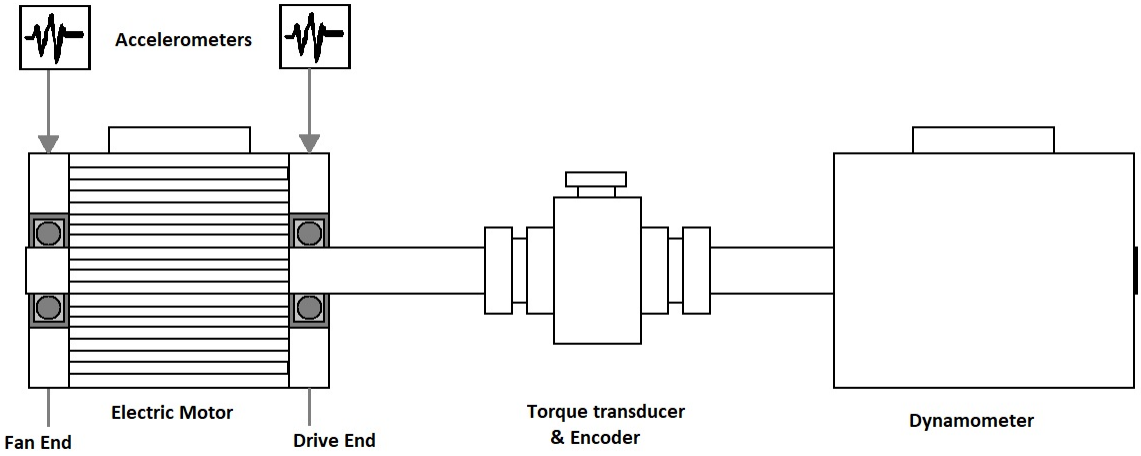


Figure 3.3: Test Rig used by CWRU

Faults of 0.007, 0.014, 0.021, and *0.028 inches* were seeded in *RE*, *IR*, and *OR* (one case per faulty bearing) by *Electro-Discharge Machining (EDM)* in *drive end* and *fan end* bearings. Model of bearings and their frequency factors can be found in Table 3.2.

Every test consisted on a faulty bearing (drive end or fan end), the motor running at speeds between 1,720 - 1,797 RPM with a load applied of 0-3 HP. Signals are recorded using accelerometers placed in vertical position on drive and fan end locations, Fig. 3.3, with sampling frequencies of 12,000 samples/second and at 48,000 samples/second in other cases.

Table 3.2: Frequency factors of *CWRU* bearings (multiple of shaft speed Hz)

Location on Test Rig	*Model*	*BPFI*	*BPFO*	*BSF/RSF*	*FTF*
Drive End	SKF6205-2RSJEM	5.4152	3.5848	4.7135	0.3983
Fan End	SKF6203-2RSJEM	4.9469	3.0530	3.9874	0.3817

3.2.2 Tian-Yau Wu Experimental Setup

Experimental setup consists of an electric motor with a motor driver, the motor is coupled to a shaft with two bearings and one encoder NA-MES-600 (800 indices/rev). The accelerometer is connected to *National Instruments Data Acquisition device (NI-DAQ)* NI 9234, the rotary encoder is connected to NI 9402, and the circuit for motor speed control is connected to NI 9401, and they communicate with the computer using NI 9172 module, Fig. 3.4.

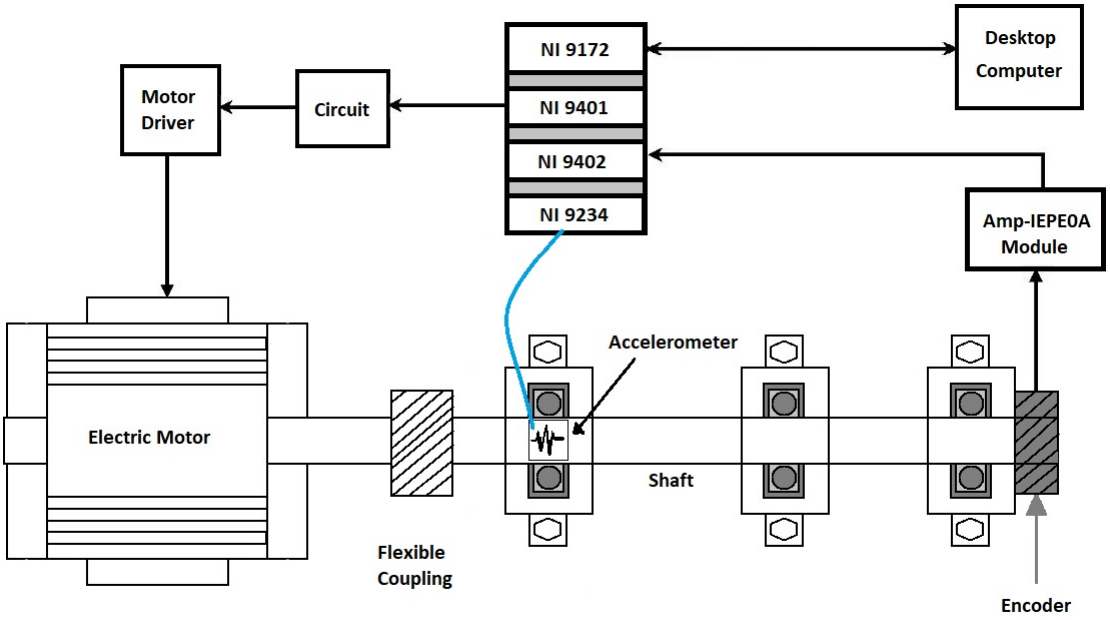


Figure 3.4: Connection diagram [Wu et al., 2016]

Bearing defects are produced by *EDM* in the *RE*, *IR*, and *OR*. Examples of defects are shown in Fig. 3.5. There are two levels of defects for *IR* and *OR*: 1) slight defect with dimensions of 0.4 mm x 0.3 mm, and 2) severe defect with dimensions of 0.8 mm x 0.3 mm. The three locations (RE, OR, IR), two levels for *IR* and *OR*, and normal condition create six classes of bearing signals. Experiments are recorded with motor running at four speed profiles (constant speed, speed-up, and

speed-down) between 300 - 750 RPM, and the vibration signals are recorded using a sampling frequency of 6,400 samples/second and a total length of 5 seconds.

Figure 3.5: *IR* (left), *RE* (center), and *OR* (right) seeded defects [Wu et al., 2016]

Table 3.3: Frequency factors of bearing (multiple of shaft speed in Hz)

BPFI	BPFO	BSF/RSF	FTF
6.14	3.86	4.17	0.39

3.2.3 MFPT Society Experimental Setup

The experimental system from the *society for Machinery Failure Prevention Technology (MFPT)* consists in of an electric motor of 1 HP coupled to a shaft with 3 bearings (2 of them in fixed position and other in one device that applies axial load), Fig. 3.6. The data acquisition systems uses an analog ADXL001 accelerometer and analog-to-digital converter ADS1271 ADC. This database includes 3 signals in normal condition and 3 *OR* fault condition signals with 270 lbs of load and sampling frequency of 97,656 samples/second and a total length of 6 seconds, there are also 7 *OR* and 7 *IR* fault condition signals with different loads applied (0 - 300 lbs) with sampling frequency of 48,828 samples/second and a total length of 3 seconds.

Bearing faults can be observed in Fig. 3.7, and the bearing frequency factors can be found in Table 3.4.

Table 3.4: Frequency factors of bearing (multiple of shaft speed in Hz)

BPFI	BPFO	BSF/RSF	FTF
4.755	3.245	2.555	0.406

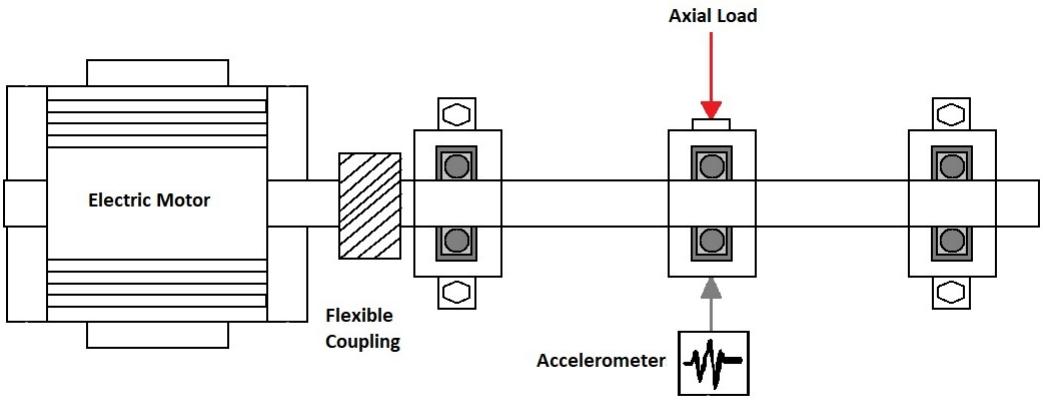


Figure 3.6: MFPT Society experimental setup (top view)

Figure 3.7: IR (left) and OR (right) bearing defects [Society, 2013]

3.3 Application of the Proposed Methodology

As demonstrative examples of the implementation of the methodology, one signal of each database is processed step by step. Finally, a comparison of the result of the proposed methodology against the traditional *Envelope spectrum*, and *FT* is presented.

3.3.1 CWRU database

For *CWRU* a signal with *RE* defect of *21 mils* and a load of 0 HP is analyzed. The signal is classified as partially diagnosable (P).

1. Signal

The methodology suggest a total of five revolutions for the analysis. As the signal was recorded at 1,796 RPM (29.93 Hz), the five revolutions correspond to 0.16705 seconds ($T = r/f_r$), where

T is the window size in seconds, r is the number of revolutions, and f_r is the shaft frequency, Fig. 3.8 .

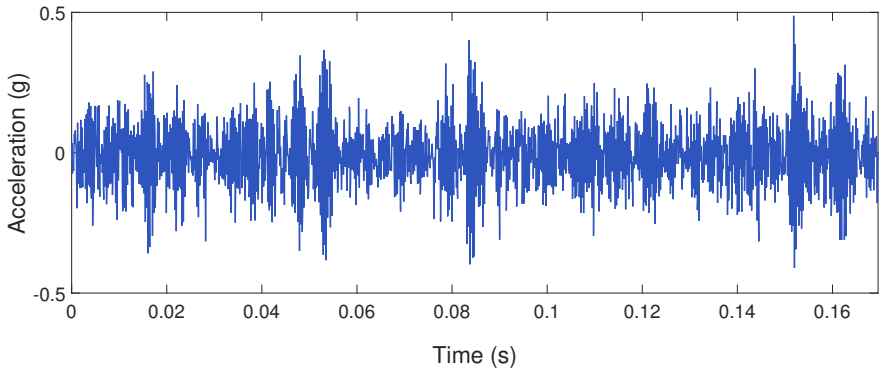


Figure 3.8: Analyzed signal with *RE* fault classified as partially diagnosable (*CWRU* database)

2. Signal Decomposition

The signal is decomposed using the *EMD* to separate the fault information from noise and other components. The *EMD* procedure is stopped when the fourth *IMF* is obtained. The corresponding components for the analyzed signal are shown in Fig. 3.9. It can be observed that the amplitude of the first *IMF* is greater than the others (around 4 times bigger for the second *IMF*, and up to 20 times bigger than the third *IMF*).

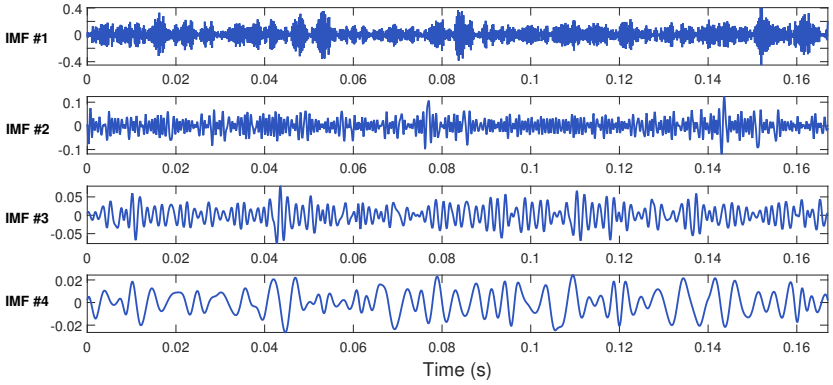


Figure 3.9: First four components (*IMFs*) of *RE-P* (*CWRU* database)

3. *IMF* Selection and Signal Reconstruction

The next step is to compute the *Kurtosis* and *RMS* for the *IMFS*, Fig. 3.10. It can be observed that the largest *RMS* value correspond to the first *IMF*, but the largest *Kurtosis* correspond to the second *IMF*. Then, the *KR* for each *IMF* computed, and based on the *KR* amplitudes, the two *IMFs* with largest values are selected to reconstruct the signal. For the analyzed signal, the 1^{st} and 2^{nd}

components have the largest *KR* with values of 0.3719 and 0.1076 respectively, Fig. 3.11. It is important to remind that the process of the selection of the two *IMFs* with largest *KR* is made automatically in real applications by programming the methodology.

The reconstructed signal is composed by the two *IMFs* with most useful information for the analysis and with less noise than the original signal, Fig. 3.12. The reconstruction is computed using eqn. (2.5) $X(t) = r + \sum_{n=1}^{N} C_n$, where r is the residual of the *EMD*, and C_n are the *IMFs*. For the reconstruction, only the two selected *IMFs* are used. At this point noise from low frequencies was removed to proceed with the *Envelope spectrum*.

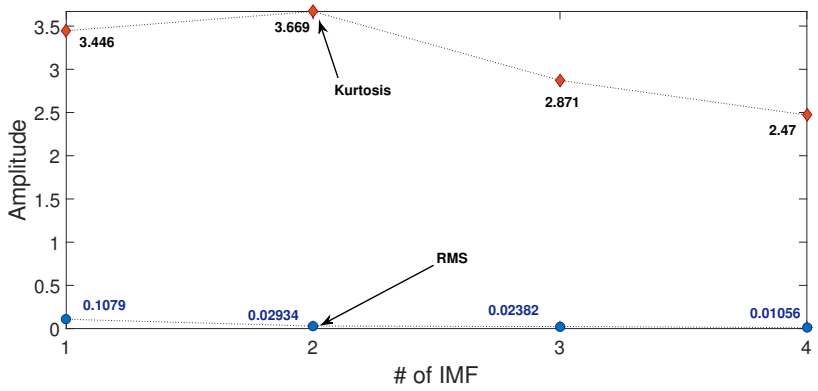


Figure 3.10: *Kurtosis* and *RMS* values, where the largest *RMS* and *Kurtosis* correspond to *IMF* 1, and *IMF* 2 respectively

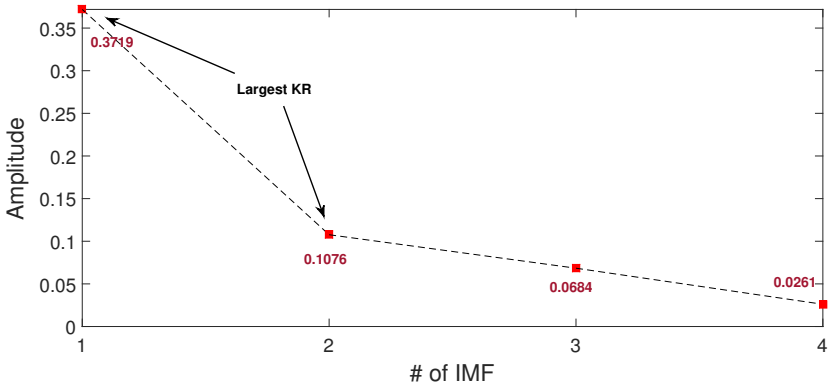


Figure 3.11: *KR* values, the largest values of *KR* correspond to the first two *IMFs*

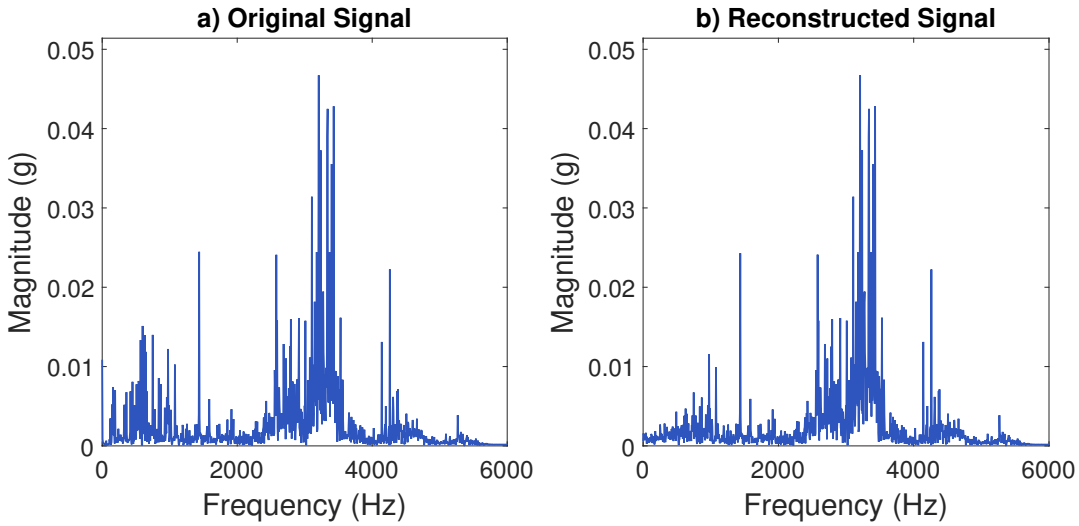


Figure 3.12: *FFT* spectrum of the (a) original signal and (b) the reconstructed signal with the *RE-P* fault (*CWRU* database). The reconstructed signal removed the noise from low frequencies to isolate the resonance frequency for the analysis

4. *Envelope spectrum*

The fault components are modulated in the resonance frequency, to visualize them in the spectrum, the *Envelope spectrum* is applied. First, the envelope is computed using the *HT*. Then, the mean trend or *DCC* is removed with eqn. (3.4) $X(t) = X(t) - X(t)_{Mean}$ to avoid the undesired peak in 0 Hz in the spectrum. The reconstructed signal, the envelope signal, and envelope signal with removed *DCC* are shown in Figs. 3.13 and 3.14.

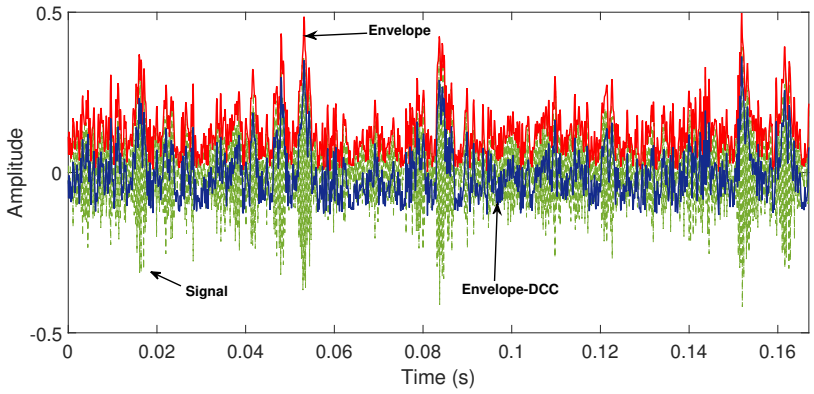


Figure 3.13: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*CWRU* database)

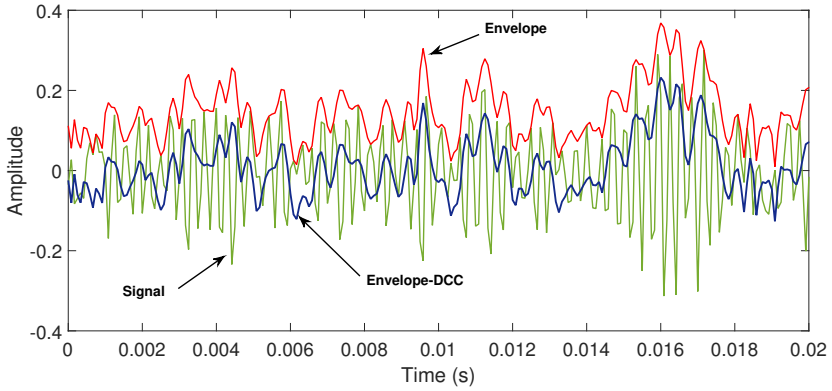


Figure 3.14: Envelope of the first 0.02 seconds from the reconstructed signal

After the envelope of the reconstructed signal is computed and the *DCC* is removed; The *FFT* is applied and the *Envelope spectrum* is plotted to make a diagnosis. To facilitate the diagnosis is recommended to include the fault frequencies and harmonics in the *Envelope spectrum*, because unlike controlled experiments, the expected fault is unpredictable.

Using the shaft speed of 29.93 Hz and the frequency factors, Table 3.2, the fault frequencies are obtained: *BPFO* = 107.3050 Hz, *BPFI* = 162.0950 Hz, *BSF/RSF* = 141.0908 Hz, and *FTF* = 11.9224 Hz, Fig. 3.15.

5. Diagnosis

The analysis of the *Envelope spectrum* for the diagnosis of bearing condition consist in the visual inspection of the spectrum and the identification of peaks in the computed fault frequencies, their harmonics, and in some cases sidebands, Fig. 3.2. For *RE* faults, peaks in the *BSF/RSF* and in the harmonics are expected, also sidebands spaced at *Cage* frequency can appear in the spectrum. In the *Envelope spectrum* of the reconstructed signal, Fig. 3.15, the *BSF/RSF* components can not be found; however, the *BPFI* and sidebands are clearly identified. The presence of these components confirm the presence of the *IR* fault.

In Fig. 3.16, the results for the comparison of the proposed methodology against the traditional *Envelope spectrum* and the *FFT* are shown. In this case, *BSF/RSF* components are not found in the spectrum, but there are components from *IR* fault, thus, the comparison is made in between these components. It can be observed that for the proposed methodology the *BPFI* and sidebands have higher magnitude than in the results form the traditional *Envelope spectrum* and the *FFT*. Another advantage of the methodology over the traditional *Envelope spectrum* and the *FFT* is the removal of the *DCC* that appears in 0 Hz for the other results.

The difference in amplitude and the presence of other components between each other are in terms of the effectiveness of the *EMD*, the use of statistical parameters (*KR*) for the selection of *IMFs* with fault information is intended to avoid the visual inspection of the *IMFs*, [Tsao *et al.*,

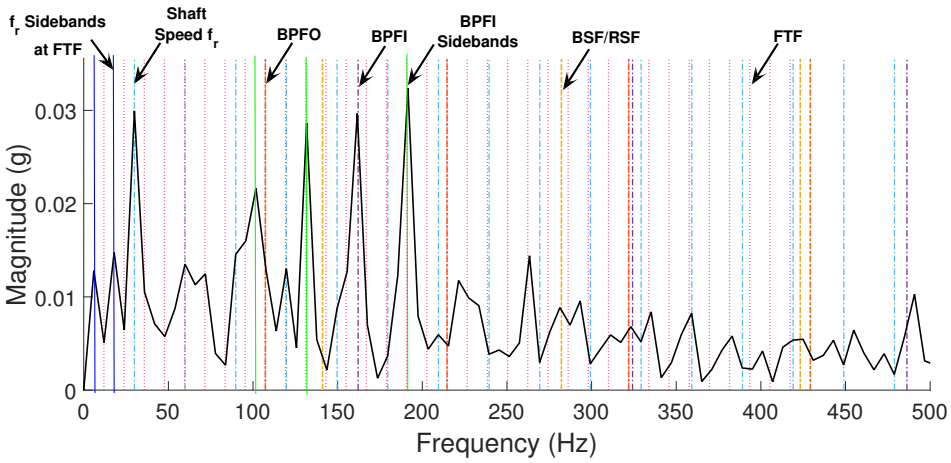


Figure 3.15: *Envelope spectrum* of reconstructed signal *RE-P* (*CWRU* database), the *BPFI* and sidebands are predominant in the spectrum

2010] [Tsao *et al.*, 2012]. The selection of the most significant *IMFs* is guarantee.

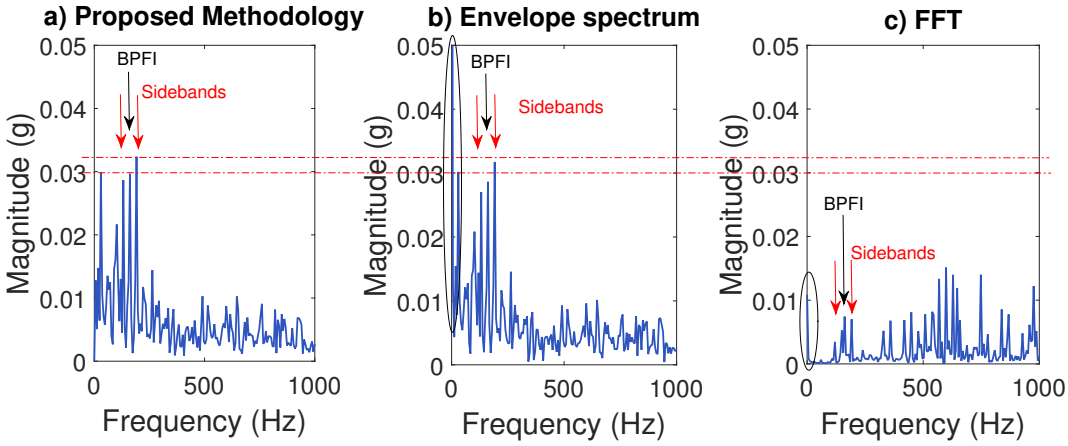


Figure 3.16: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *RE-P* (*CWRU* database), the proposed methodology removes the *DCC* (circled), and the magnitude of the *IR* fault components are greater for the proposed methodology

3.3.2 Tian-Yau Wu database

For *Tian-Yau Wu* database, a signal with *IR* slight defect (0.4 x 0.3 mm) is analyzed, Fig. 3.17.

1. Signal

First, the sample of five revolutions is made for the analysis. This signal was recorded at 7.45 Hz, the five revolutions correspond to 0.6711 seconds ($T = r/f_r$), where T is the window size in seconds, r is the number of revolutions, and f_r is the shaft frequency.

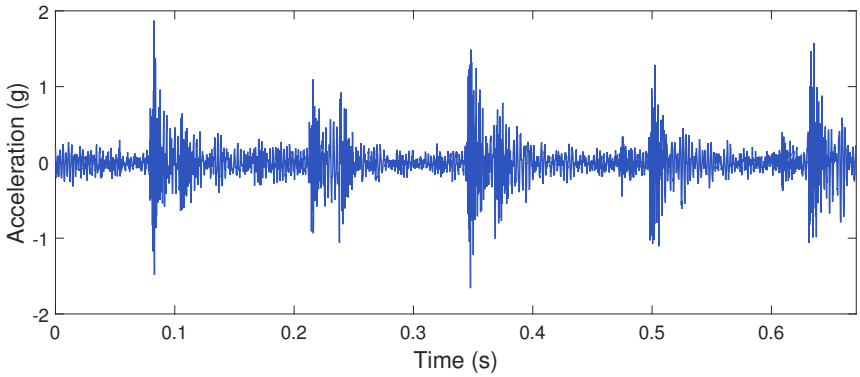


Figure 3.17: Analyzed signal with *IR* fault slight defect (*Tian-Yau Wu* database)

2. Signal Decomposition

The *EMD* is applied to decompose the signal in fault information and noise components. The *EMD* procedure is stopped when the fourth *IMF* is obtained to avoid *IMFs* with the end defect. The corresponding components for the analyzed signal are shown in Fig. 3.18. It can be observed that the amplitude of first *IMF* is twice the amplitude of the second and third *IMF*. In the first *IMF* the impulses are clearer, and there is less noise than in the others.

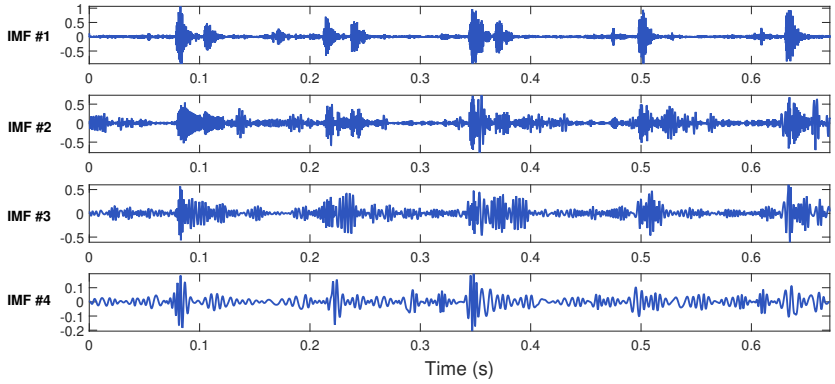


Figure 3.18: First four components (*IMFs*) of *IR* slight defect signal (*Tian-Yau Wu* database)

3. *IMF* Selection and Signal Reconstruction

The next step is to compute the *Kurtosis, RMS* for each *IMF*, Fig. 3.19. It can be observed that the first component produce the highest values of *RMS* and *Kurtosis*. The next step is to compute the *KR* for each *IMF*. The selection of the two components with faulty information for

the reconstruction of the signal is made based on the *KR* amplitudes, the 1^{st} and 2^{nd} components have the largest *KR* with values of 2.464 and 1.125 respectively and are selected, Fig. 3.20. This selection is made automatically by the code of the methodology, the presented images are only for illustrative purposes.

The reconstruction is computed using eqn. (2.5). The new signal is now with the most useful information from the original signal, Fig. 3.21. It can be observed that the noise from low frequencies was removed with the methodology. This new signal is now ready to apply the *Envelope spectrum*.

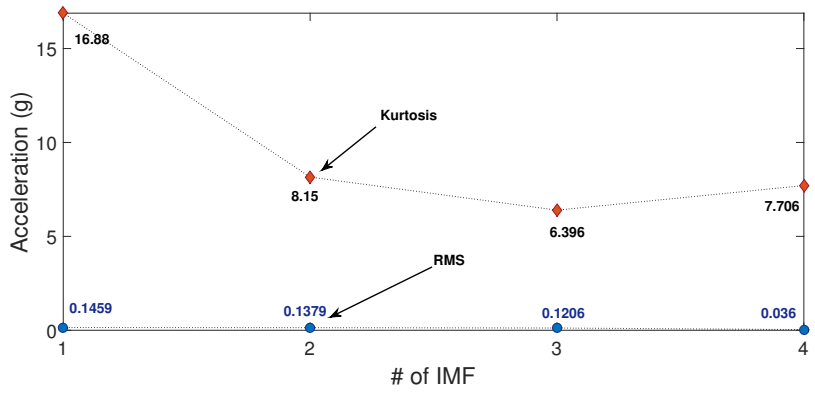


Figure 3.19: *Kurtosis* and *RMS* values of the *IMFs* from signal with *IR* slight defect (*Tian-Yau Wu* database), the largest *RMS* and *Kurtosis* correspond to *IMF* 1

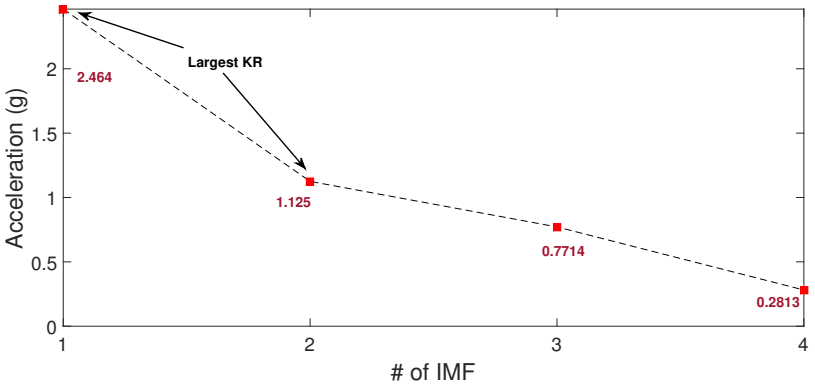


Figure 3.20: *KR* values of the *IMFs* from signal with *IR* slight defect (*Tian-Yau Wu* database), the largest values of *KR* correspond to the first two *IMFs*

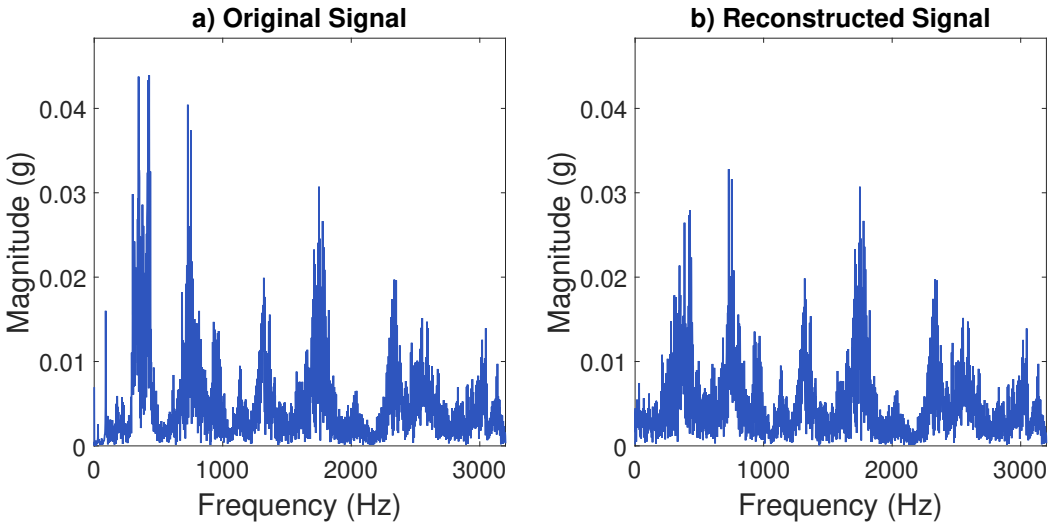


Figure 3.21: *FFT* spectrum of the (a) original signal and (b) the reconstructed signal with the *IR* slight defect (*Tian-Yau Wu* database). In the reconstructed signal, noise from low frequencies was removed to isolate the resonance frequency for the analysis

4. *Envelope spectrum*

First, the envelope is computed using the *HT* to demodulate the faults. Then, the mean trend or *DCC* is removed with eqn. (3.4) to avoid the undesired peak in 0 Hz in the spectrum. The reconstructed signal, the envelope signal, and envelope signal with removed *DCC* are shown in Figs. 3.22 and 3.23.

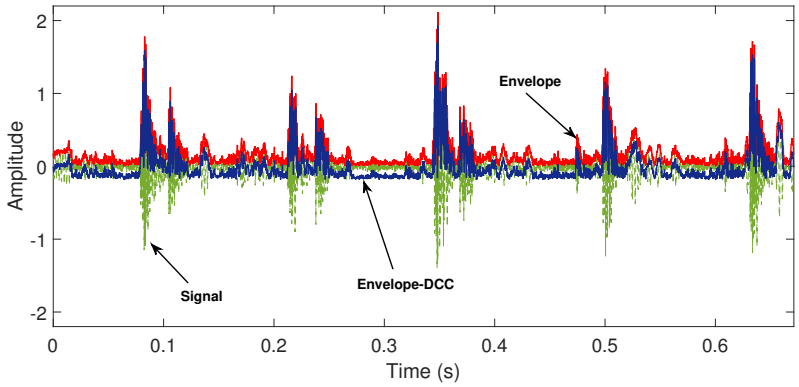


Figure 3.22: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*Tian-Yau Wu* database)

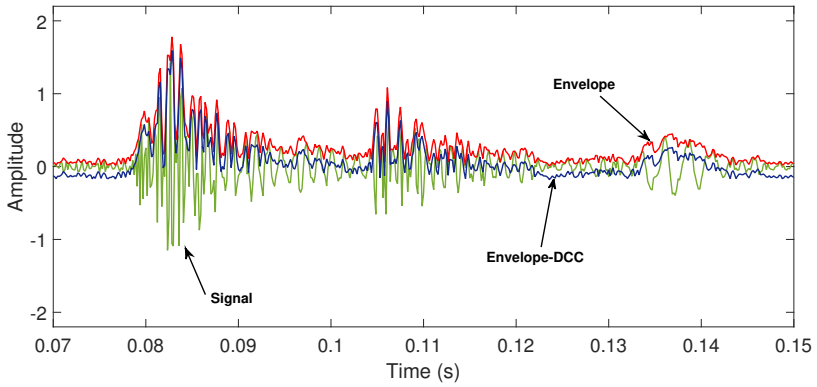


Figure 3.23: Envelope for seconds 0.07 to 0.15 of the *IR* slight defect reconstructed signal (*Tian-Yau Wu* database)

The next step is to apply *FFT* to the envelope signal with the removed *DCC* to make the diagnosis. Using the shaft speed of 7.45 Hz and the frequency factors, Table 3.3, the fault frequencies are obtained: *BPFO* = 28.757 Hz, *BPFI* = 45.743 Hz, *BSF/RSF* = 31.0665 Hz, and *FTF* = 2.9055 Hz, Fig. 3.24.

5. Diagnosis

The *Envelope spectrum* analysis consist in the visual inspection of the spectrum and the identification of peaks in the computed fault frequencies, their harmonics, and in some cases sidebands, Fig. 3.2. For *IR* faults, peaks in the *BPFI* and in the harmonics are expected, also some sidebands spaced at shaft frequency can appear in the spectrum. In Fig. 3.24, the *BPFI*, one harmonic and two sidebands are visible in the spectrum, Fig. 3.24. Also, the *BSF/RSF* with one sideband can be identified, which confirm that there are two faults in the bearing instead of one.

In Fig. 3.25, the comparison of the proposed methodology against the traditional *Envelope spectrum* and the *FFT* is shown. In this case, the *FFT* does not detect *BPFI* components; the proposed methodology and the traditional *Envelope spectrum* show the fault frequency and the two sidebands. In the comparison of these two methods, it can be observed that the proposed methodology deals correctly with the *DCC*. Other important peak is the shaft frequency, for the proposed methodology this peak reduces its amplitude. For the *BPFI* and the first sideband there are increases in amplitude of 12 % and 21 % respectively compared with the traditional *Envelope spectrum*.

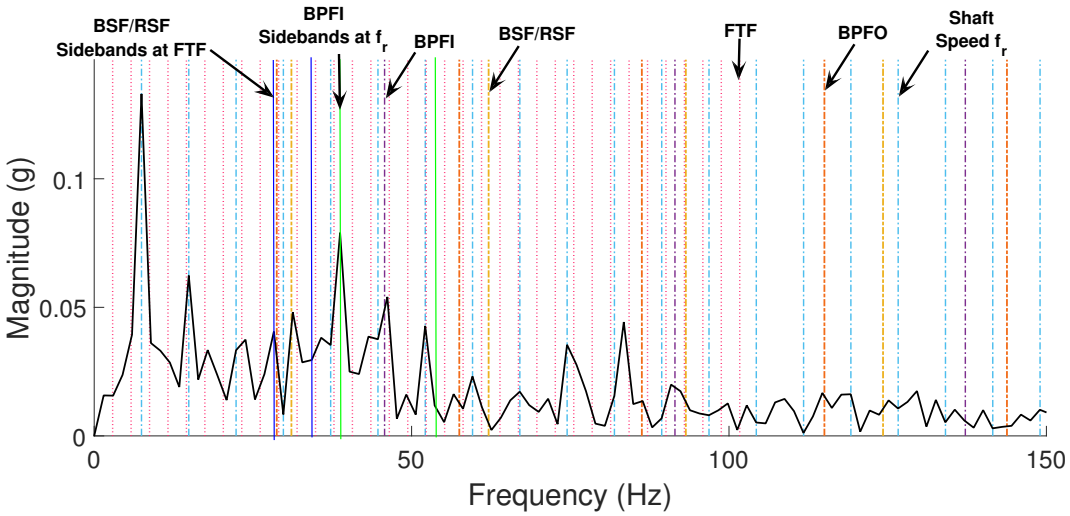


Figure 3.24: *Envelope spectrum* of reconstructed signal *IR* slight defect (*Tian-Yau Wu* database), the *BPFI, BSF/RSF* and sidebands are identified, this suggest multiple faults in the bearing

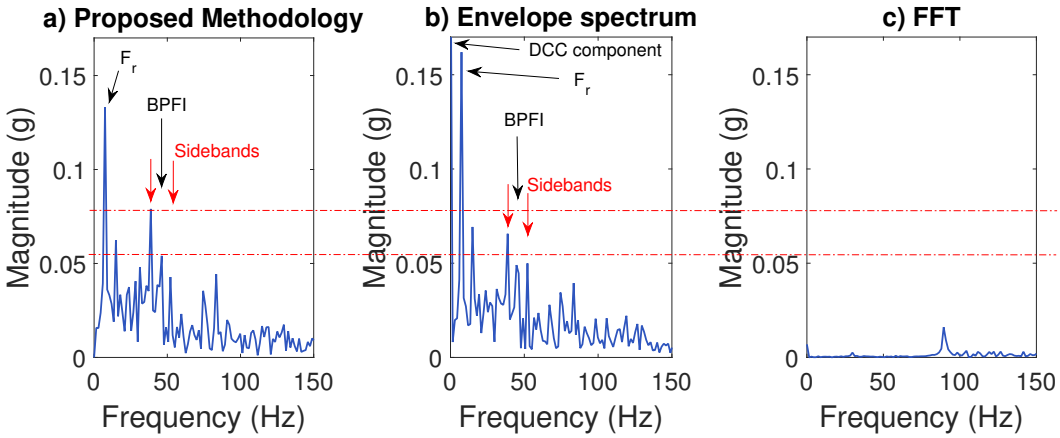


Figure 3.25: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *IR* slight defect (*Tian-Yau Wu* database), the proposed methodology increases the amplitude of the *BPFI* and the sidebands

3.3.3 MFPT database

For *MFPT* database a signal with *OR* defect with a load of 250 lbs is analyzed, Fig. 3.26.

1. Signal Decomposition

To compute the five revolutions for the analysis the shaft frequency is used. The signal was recorded at 1,500 RPM (25 Hz), the five revolutions correspond to 0.2 seconds ($T = r/f_r$), where

T is the window size in seconds, r is the number of revolutions, and f_r is the shaft frequency.

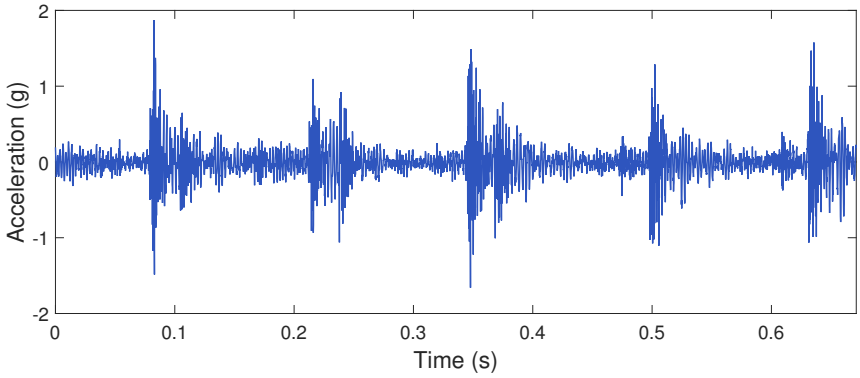


Figure 3.26: Analyzed signal with *OR* fault and load of 250 lbs (*MFPT* database)

2. Signal Decomposition

The next step is to apply the *EMD* to separate the fault information from noise, and the first four *IMFs* are obtained, Fig. 3.27. It can be observed that the amplitude of the first *IMF* is greater than the others, but the *IMF* with more impulses can not be identified.

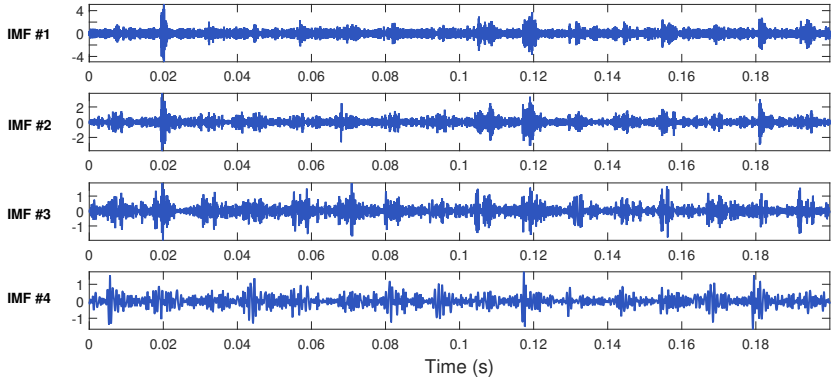


Figure 3.27: First four components (*IMFs*) of *OR* fault 250 lbs of load (*MFPT* database)

3. *IMF* Selection and Signal Reconstruction

The next step is to compute the *Kurtosis* and *RMS* for each *IMF*, Fig. 3.28. The highest *RMS* is 0.6363 and correspond to the first *IMF*, the highest *Kurtosis* is 8.599 and correspond to the second *IMF*, this suggest that the second *IMF* is the component with more impulses. The next step is to compute the *KR* for each *IMF*. Then components with largest *KR* are selected to reconstruct the signal, in this case the 1^{st} and 2^{nd} with *KR* values of 4.683 and 4.311 respectively are selected, Fig. 3.29. The identification of the two *IMFs* with highest *KR* is automated by programming the proposed methodology.

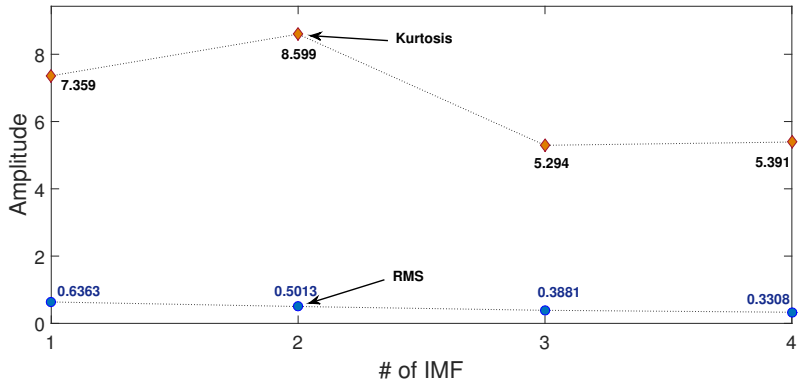


Figure 3.28: *Kurtosis* and *RMS* values of first four components (*IMFs*) of *OR* fault and 250 lbs of load (*MFPT* database), the largest values of *KR* correspond to the first two *IMFs*

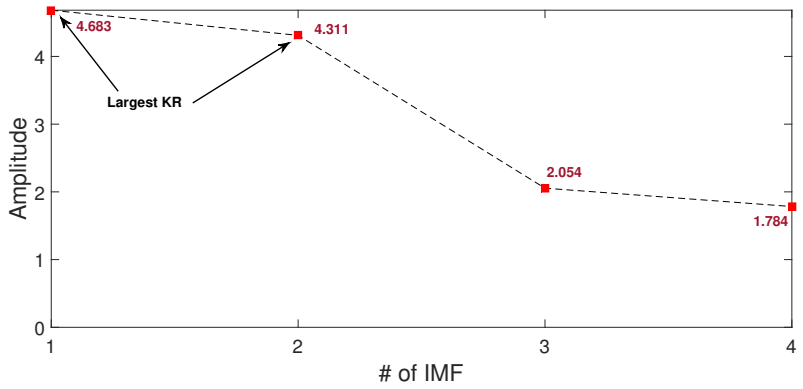


Figure 3.29: *KR* values of first four components (*IMFs*) of *OR* fault and 250 lbs of load (*MFPT* database), the largest values of *KR* correspond to the first two *IMFs*

The reconstructed signal is composed by the two *IMFs* with most useful information for the analysis and with less noise than the original signal, Fig. 3.30. The reconstruction is computed using eqn. (2.5). The next step is to proceed with the *Envelope spectrum*.

4. *Envelope spectrum*

To compute the *Envelope spectrum*, first, the envelope is computed using the *HT*. Then, the mean trend or *DCC* is removed with eqn. (3.4) to avoid the undesired peak in 0 Hz in the spectrum. This step with the reconstructed signal, the envelope signal, and envelope signal with *DCC* removed is shown in Figs. 3.31 and 3.32.

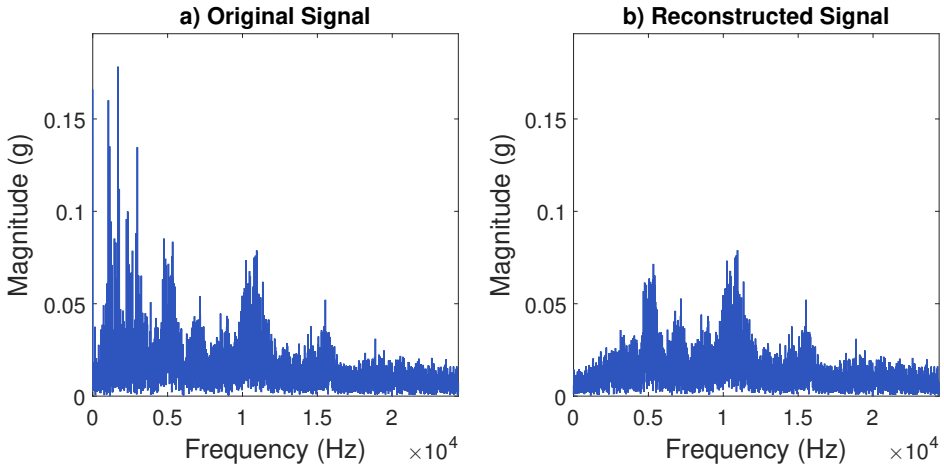


Figure 3.30: *FFT* spectrum of the (a) original signal and (b) the reconstructed signal with the *OR* fault and 250 lbs of load (*MFPT* database). The reconstructed signal removed the noise from low frequencies

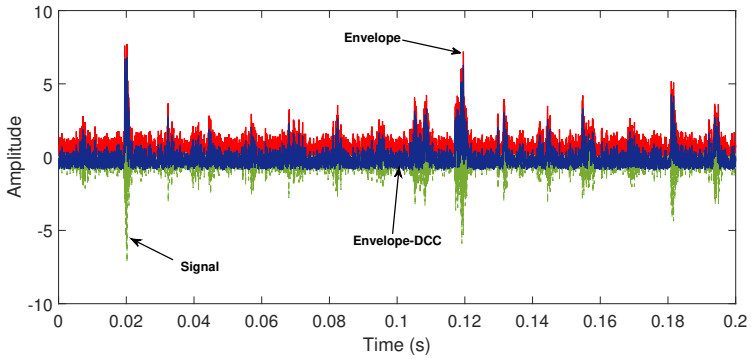


Figure 3.31: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*MFPT* database), the *DCC* removal does not affect the waveform and improves the diagnosis

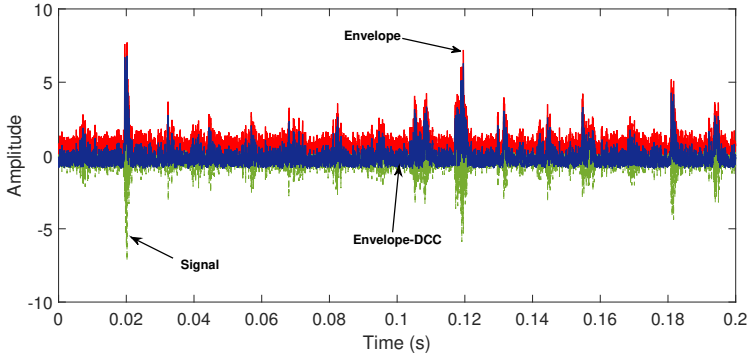


Figure 3.32: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*MFPT* database), the *DCC* removal does not affect the waveform and improves the diagnosis

After the envelope of the reconstructed signal is computed and the *DCC* is removed; the *FFT* is applied and the *Envelope spectrum* is plotted to make a diagnosis.

Using the shaft speed of 25 Hz and the frequency factors, Table 3.4, the fault frequencies are obtained: *BPFO* = 81.125 Hz, *BPFI* = 118.875 Hz, *BSF/RSF* = 63.875 Hz, and *FTF* = 10.15 Hz, Fig. 3.33.

5. Diagnosis

Finally, a visual inspection of the *Envelope spectrum* is made for the diagnosis. It consist in the identification of peaks in the computed fault frequencies, their harmonics, and in some cases sidebands, Table 3.1. For *OR* faults peaks in the *BPFO* and in the harmonics are expected. In Fig. E.13 peaks in the *BPFO* and first harmonic can be identified for the analyzed signal and confirm the presence of the *OR* fault in the signal.

In Fig. 3.34 the comparison of the proposed methodology against the traditional *Envelope spectrum* and the *FFT* is shown. In this case the amplitude in the *BPFO* is smaller than in the *Envelope spectrum*. This type of result is caused because of the *EMD*, when the fault information is more than two *IMFs* the reconstruction does not include it completely. The fault information in multiple in more than two *IMFs* can occur when there are multiple resonance frequencies of the system, e.g. in this case there are resonance frequencies around 2,500 Hz, 6,000 Hz, 12,000 Hz, and 15,000 Hz Fig.3.30 (a).

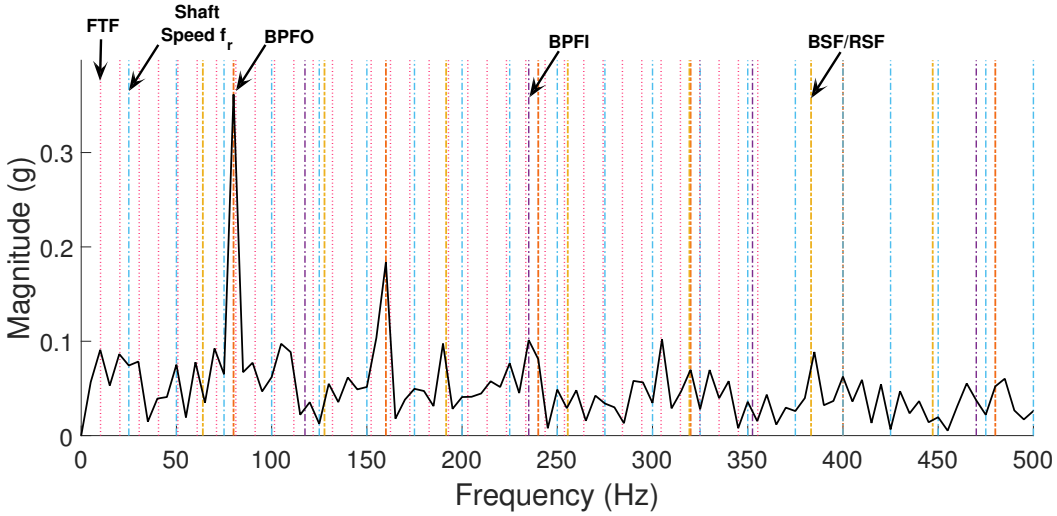


Figure 3.33: *Envelope spectrum* of reconstructed signal from *OR* fault and 250 lbs of load (*MFPT* database), the *BPFO* and one harmonic are predominant in the spectrum

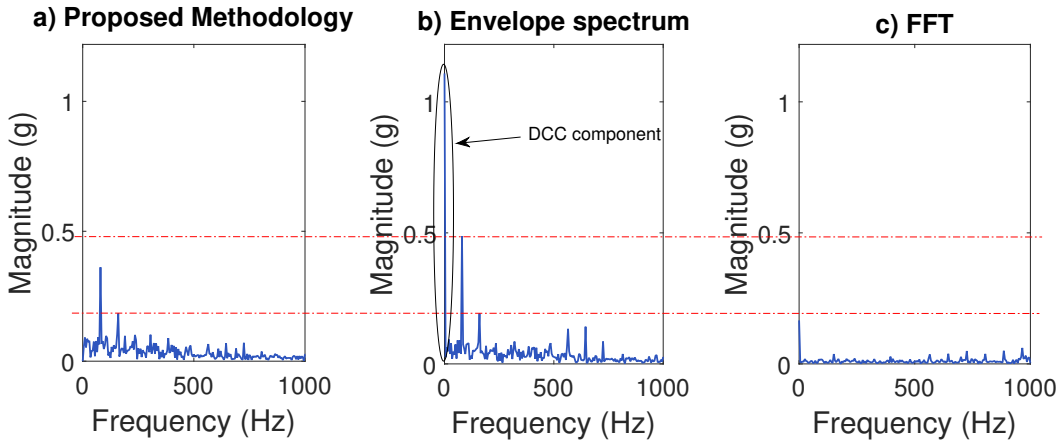


Figure 3.34: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *OR* fault and 250 lbs of load (*MFPT* database), the proposed methodology removes the *DCC*, but the *BPFO* is smaller than in the *Envelope spectrum*

Chapter 4

Results

4.1 Introduction

This chapter is divided in three sections for each database. In each section, first, the used signals for the validation of the methodology are presented. Finally, the methodology and results are discussed.

4.2 CWRU database

For the validation of the proposed methodology a total of 9 signals from *CWRU* database were analyzed covering the 3 types of fault, and 3 groups of signals based on the difficulty for their diagnosis: (Y) for diagnosable signals, (P) for partially diagnosable, and (N) for no diagnosable [Smith and Randall, 2015]. For the Y signals the fault waveform is clearly visible, Fig. 4.1 (a), (d) and (g), for P signals the components reduced their amplitude and the waveform is hard to distinguish, Fig. 4.1 (b), (e) and (h), and for N signal of *IR* the amplitude of the signal is similar to the Y classified but is overwhelmed by noise, Fig. 4.1 (c), and for *OR* and *RE* the fault components are as small in amplitude as the noise, Fig. 4.1 (f) and (i).

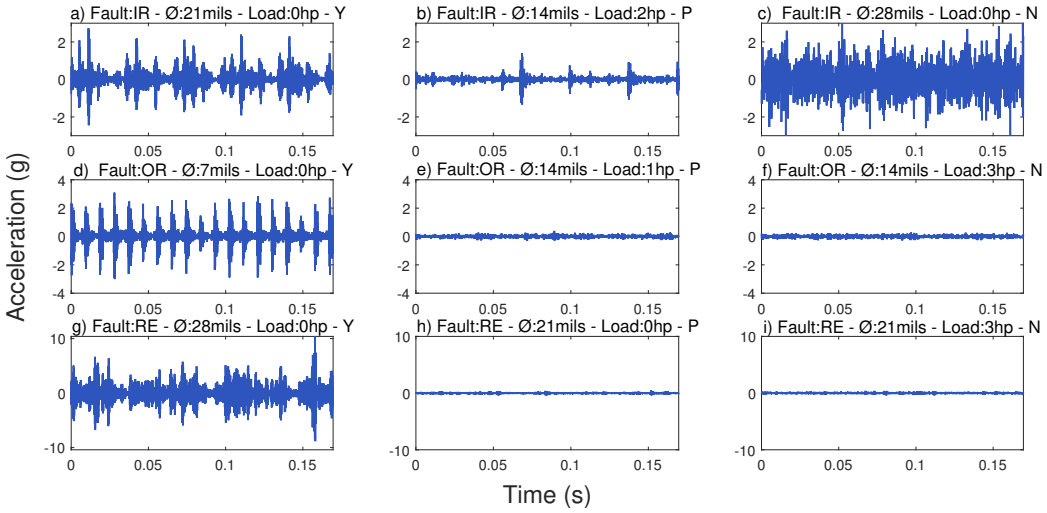


Figure 4.1: Analyzed signals to validate the proposed methodology (*CWRU* database)

4.2.1 CWRU Results

To validate the use of the *KR* as indicator of fault information, each *IMF* was analyzed individually to determine if the selection was correct. Also, the *Envelope spectrum* of the reconstructed signal was analyzed to make a diagnosis. In this section only some *Envelope spectrum* are presented, the remaining results can be found in *Appendix E*.

- **IR Signals.** First, for the diagnosable signal (Y) the two selected *IMFs* were the 1^{st} and 2^{nd} components, the analysis of the four *IMFs* indicate that the selection was correct, for the analysis of *Envelope spectrum* the fundamental frequency and harmonics are identified confirming that the fault is present. In the partially diagnosable signal the *BPFI* and harmonics were found, but also other peaks were identified in the spectrum, Fig. 4.2. These peaks can be interpreted as noise that makes difficult the diagnosis, but close analysis show that these peaks correspond to sidebands spaced at shaft frequency for the *BPFI*. In the non diagnosable signal the identification of the *IMFs* with fault information was successful, Table 4.1. Nevertheless, the peak in the *BPFI* was not identified nor the harmonics. On the other hand, the *BPFO* was clearly found instead and the *OR* fault can be confirmed, Fig. 4.3.

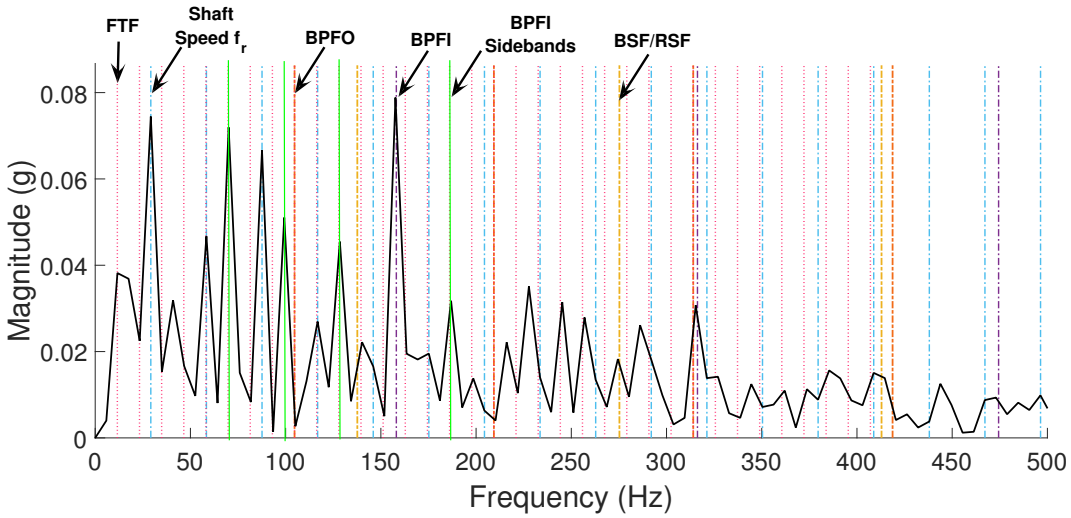


Figure 4.2: *Envelope spectrum* of reconstructed signal *IR-P* (*CWRU* database), sidebands of the *BPFI* appear in the spectrum

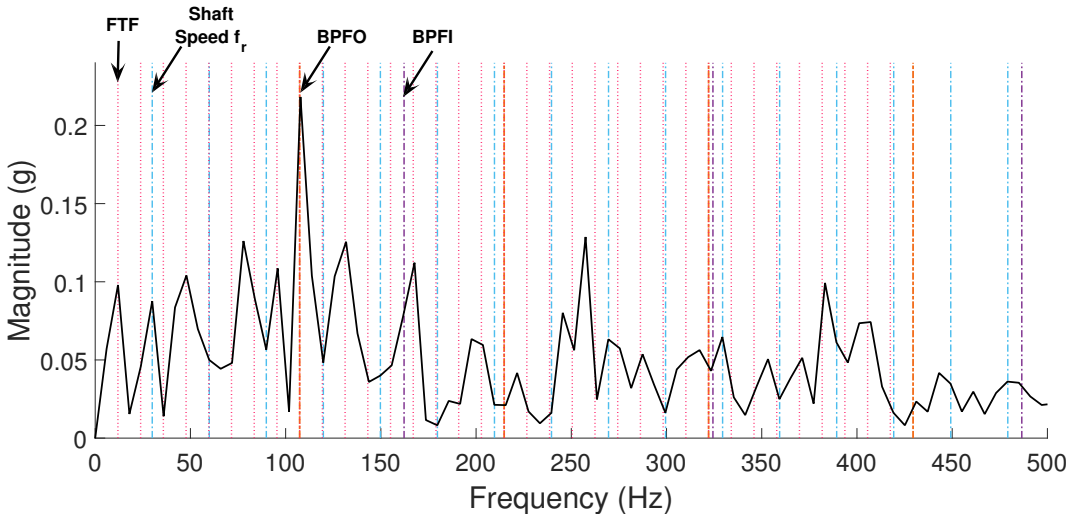


Figure 4.3: *Envelope spectrum* of reconstructed signal *IR-N* (*CWRU* database), the *BPFI* is not identified and the *BPFO* has the largest amplitude in the spectrum

- **OR Signals.** The result of the *OR* fault signal classified as diagnosable can be found in *Appendix E*. This signal was analyzed step by step, and the *BPFO* and harmonics can be clearly identified in the spectrum, which confirm the presence of fault. For the partially diagnosable case the *BPFO* was not identified neither the first harmonic, Fig. 4.4. In the *Envelope spectrum* of this signal, the *BPFI* and three sidebands can be noticed. Also, two sidebands spaced at +/- *FTF* are identified for the shaft speed frequency, the sidebands appear as an indicator of the modulation and in this case the shaft speed impulses are being modulated by

FTF impulses; The individual analysis of the *IMFs* showed that the selection of the first two components was made correctly. However, it was observed in this analysis that the *BPFO* can be identified in the second and third components. In Fig. 4.5 the *Envelope spectrum* of the third *IMF* is presented, the *BPFO, BPFI*, and *BSF/RSF* are identified which confirm multiple faults in the bearing. It is important to observe the amplitude of the peaks in this spectrum ($\times 10^{-3}$). These peaks can not be identified in the spectrum of the reconstructed signal because the amplitude of the first *IMF* is much greater than the amplitude of the second and third component.

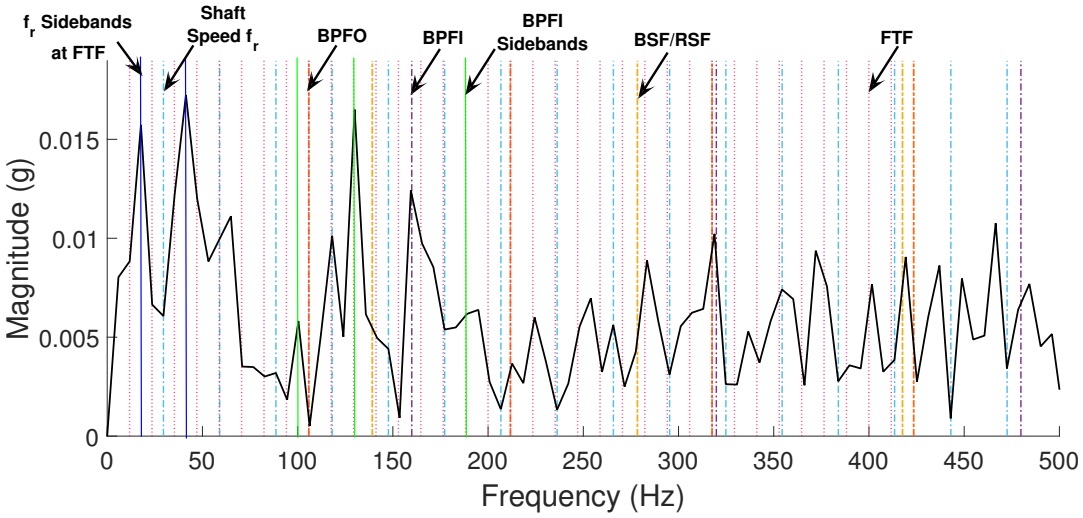


Figure 4.4: *Envelope spectrum* of reconstructed signal *OR-P* (*CWRU* database), the *BPFO* is not shown and the *BPFI* is identified instead, and two sidebands appear for the shaft speed

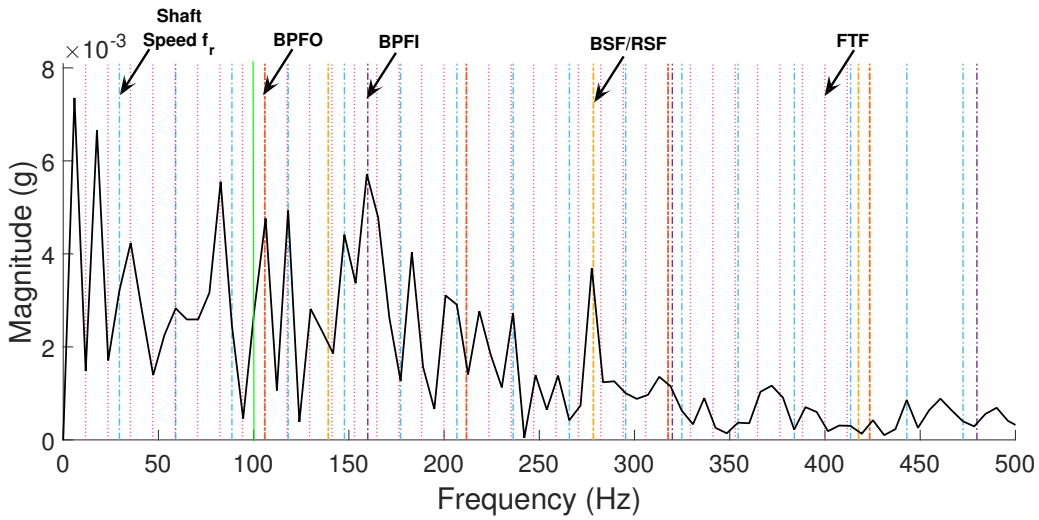


Figure 4.5: *Envelope spectrum* of third *IMF* of *OR-P* signal (*CWRU* database), the *BPFO*, *BPFI*, and *BSF/RSF* can be identified, but the magnitude is small

For the non diagnosable signal the *BPFO* can not be identified, individual analysis of the *Envelope spectrum* of each component produce the same result. In the *Envelope spectrum* of the reconstructed signal only one small peak of the *BPFI* and two harmonics of *IR* fault can be identified. The presence of these components confirm the *IR* fault in the bearing, but the diagnosis is not clear because there are other peaks with greater amplitude, Fig. 4.6.

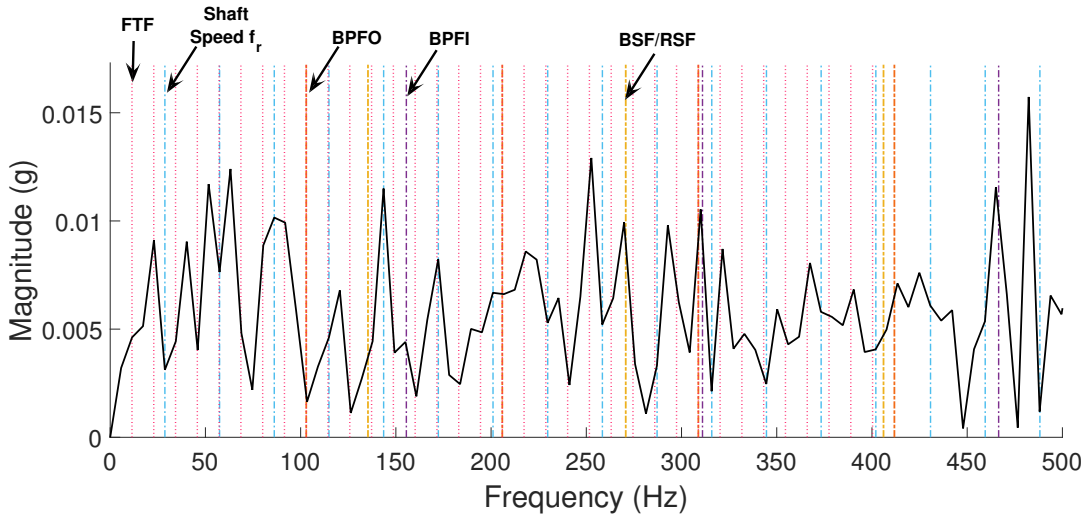


Figure 4.6: *Envelope spectrum* of reconstructed signal *OR-N* (*CWRU* database), the *BPFO* is not shown and the *BPFI* and *BSF/RSF* are identified instead

- **RE Signals.** The analysis of individual components for the diagnosable signal revealed that the *IMF* selection was not made correctly for this case, the largest *KR* values correspond to

the 1^{st} and 4^{th} *IMFs*. The first *IMF* is in fact the most reliable component, but the fourth *IMF* has the end defect, which made the *KR* to increase and cause the wrong selection, Fig. 4.7. For the *Envelope spectrum* of the reconstructed signal one peak can be found in the fault frequency, as well as the *FTF* and two harmonics that represent the largest components in the spectrum. Also, a small peak in the *BPFO* was found, which suggest multiple fault in the bearing, the result can be found in *Appendix E*. For the *RE* partially diagnosable signal, the result was presented in *Chapter 3*. Finally, for the non diagnosable signal the *BPFI* is identified again, but this time with more harmonics and sidebands, Fig. 4.8.

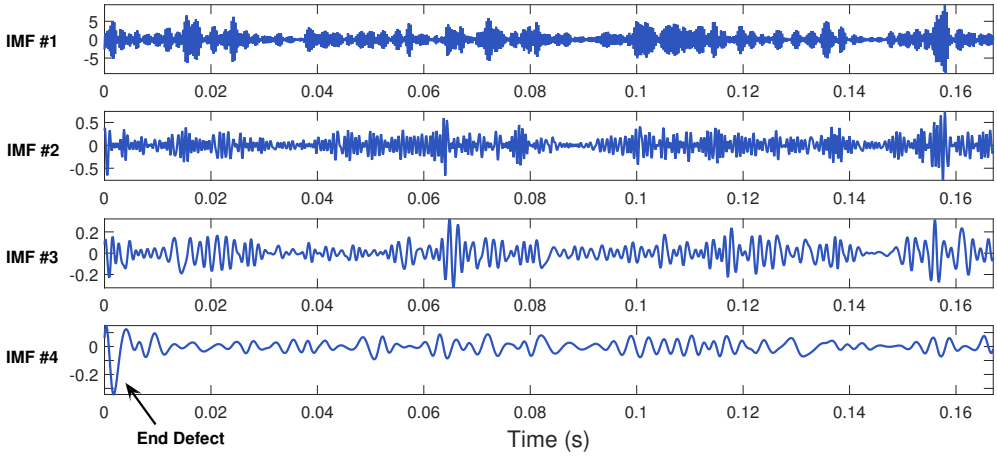


Figure 4.7: First four *IMFs RE-Y* signal (*CWRU* database), the fourth *IMF* has the end defect which caused a wrong selection

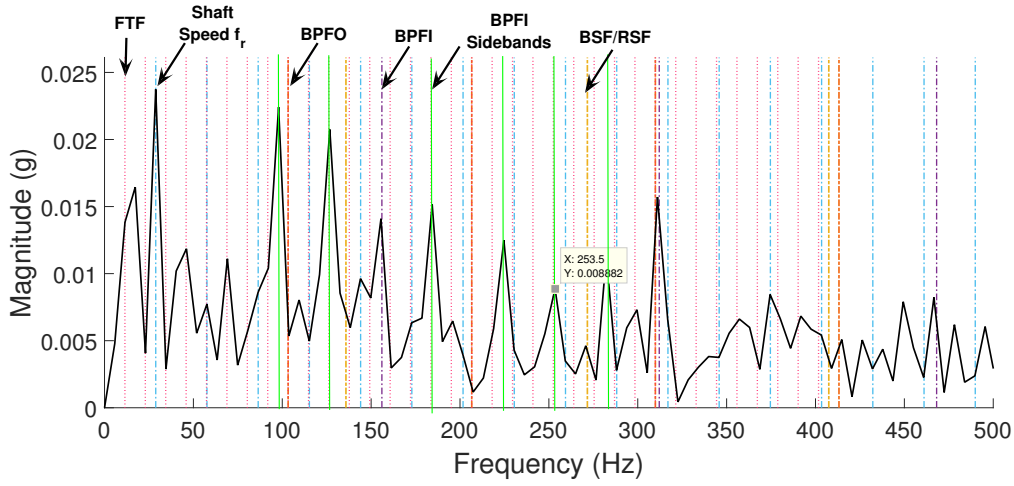


Figure 4.8: *Envelope spectrum* of reconstructed signal *RE-N* (*CWRU* database), the *BPFI*, harmonics and sidebands are found instead of the *RE* components

Contrary to expectations, in some signals labeled with one type of fault, the *Envelope spectrum* of the proposed methodology revealed the presence of other types of fault. The analysis performed to individual *IMFs* revealed the presence of other faults that were not visible in the spectrum of the reconstructed signal; however, this does not mean that the reconstruction of the signal was made incorrectly. The absence of this components is caused by the amplitude difference of the impulses produced by the types of fault. Only one error in the *IMF* selection was made, this was caused by the presence of the end defect.

4.3 Tian-Yau Wu database

For *Tian-Yau Wu* database 5 signals were studied considering slight and severe fault sizes for *IR* and *OR* faults and 1 signal for *RE* with slight defect, Fig. 4.9 . This database disposes of recorded signals with speed fluctuations, but as the assessment of rotating machines must be made at steady state condition (*ISO 10816-1* and *ISO 17243-1*) the profile of constant speed at 447 RPM was selected [ISO, 1995] [ISO, 2014]. Fig. 4.9 also shows the expected wave-forms produced by each type of the fault, for *IR* faults exponential decaying impulses are produced and are often modulated by the shaft speed because as the *IR* spins the force between the defect and the rolling element varies. The *OR* defect produces the same type of impulses as *IR* fault; but, they are not modulated by the shaft speed because the *OR* is in a fixed position and the force of contact with the *RE* does not vary, as a result more impulses are visible in *OR* fault signals; finally, the *RE* defect produces softer impulses caused by the combination of the forces and dynamic of the *RE*.

The methodology indicates that only 5 revolutions are needed for the analysis, in Fig. 4.9 (b) there are two unexpected peaks or sudden increases where there should be only five as in Fig. 4.9 (a) , these peaks are produced between seconds 0.38 and 0.43 and the cause will be discussed later. In Fig. 4.9 (e) there are some impulses between seconds 0.3 and 0.4 and not correspond to the period of the shaft speed, this phenomena also has a cause and the *Envelope spectrum* of the signals will help to explain it, Fig. 4.10 and 4.12.

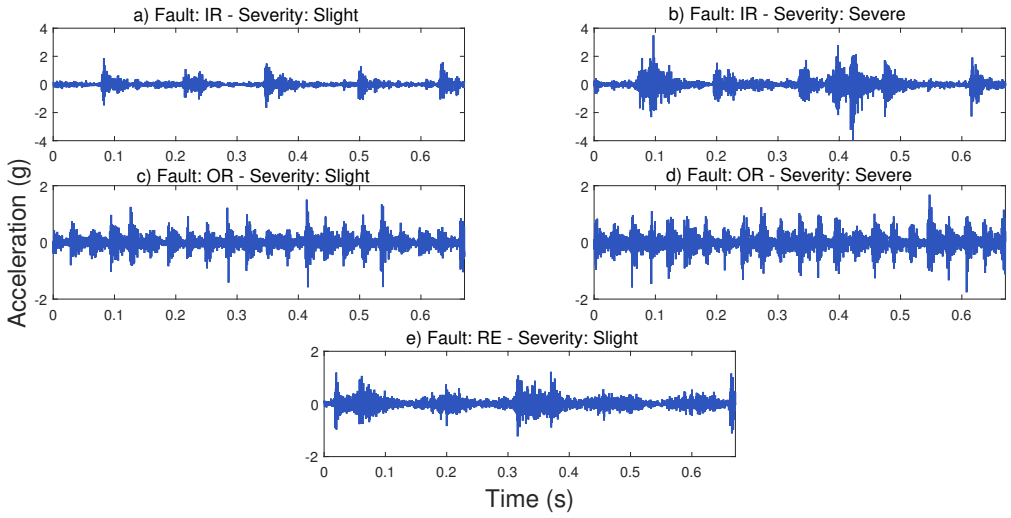


Figure 4.9: Analyzed signals to validate the proposed methodology (*Tian-Yau Wu* database)

4.3.1 Tian-Yau Wu Results

In this section, the findings of the analyzed signals are presented. In addition to the *Envelope spectrum* of the proposed methodology, an analysis of each individual *IMF* was made to confirm that the selection was made correctly.

- **IR Signals.**

  In the signal of severe fault on *IR* the *BPFI* and *BSF/RSF* have similar amplitude as in slight defect, but in this case the *FTF* and the harmonics have the largest amplitudes in the spectrum, Fig. 4.10. These amplitudes overwhelm other fault peaks in the spectrum, this suggest that there is a severe fault in the *Cage*. When the fault is introduced in the bearing by *EDM* the disassembly and reassembly of bearings is needed, this processes can introduce the defect in the *Cage*. The *Cage* defect explain the two unexpected impulses in the signal, Fig. 4.9 (b), as the *FTF* is a third of the shaft speed one impulse overlaps with the shaft impulses and the remaining two are visible.

- **OR Signals.** The methodology was applied to the *OR* signal with slight defect and the two first *IMFs* were selected. In the spectrum the *BPFO* does not match the peak because of the *Envelope spectrum* resolution; but, the difference is less than the 2% allowed. The magnitude of the *OR* peaks confirm the presence of the fault, Fig. 4.11. On the other hand, the result of the signal with *OR* severe can be found in *Appendix E* with the implementation of the methodology step by step.

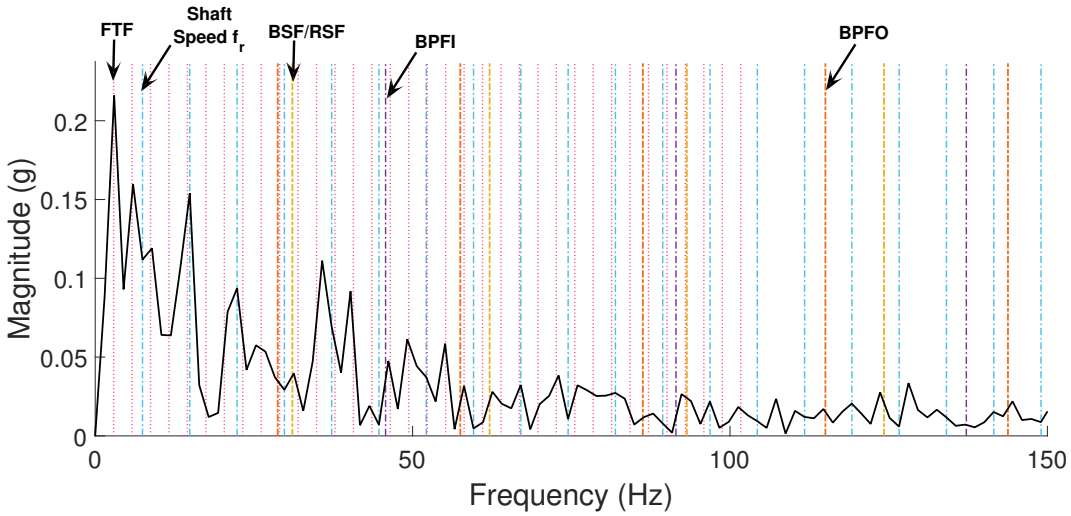


Figure 4.10: *Envelope spectrum* of reconstructed signal *IR* severe defect (*Tian-Yau Wu* database), the *BPFI* peak is small and *FTF* represent the largest peak in the spectrum, this suggest that there is a severe *Cage* defect

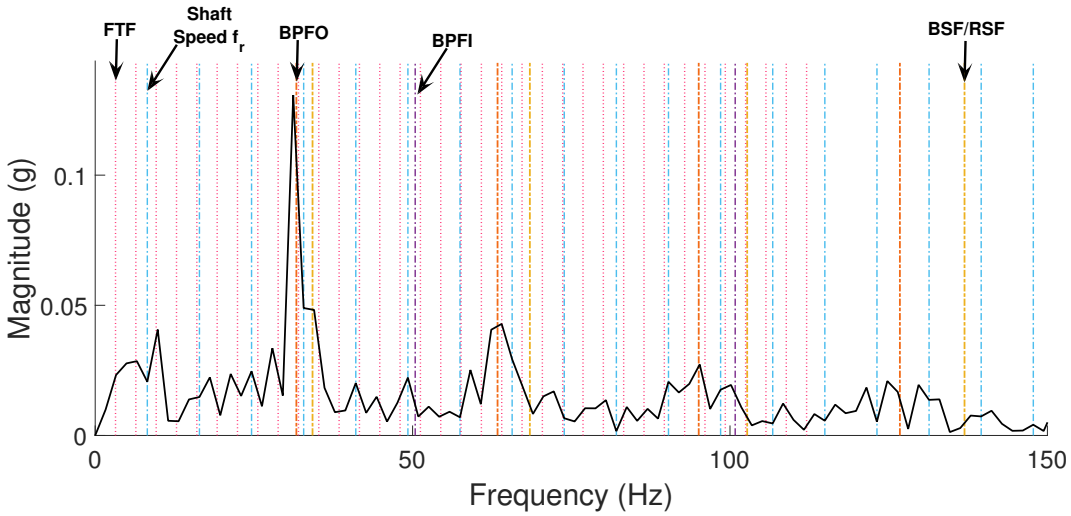


Figure 4.11: *Envelope spectrum* of reconstructed signal *OR* slight defect (*Tian-Yau Wu* database), the *BPFO* does not match the peak but is between the 2% allowed

- **RE Signal.** After the analysis of individual *IMFs* from this signal, it was determined that the selection of the first two *IMFs* is correct. In the *Envelope spectrum* the *BSF/RSF* appears with two sidebands spaced at *FTF* as can be expected, Fig. 4.12. In the spectrum the shaft frequency is the biggest peak, but the second is the *FTF*, this type of peaks on *Cage* frequency appeared in the signal of *IR* with severe defect. As in the case of the severe *IR* defect signal, the *Cage* impacts explain the sudden impulses in the *RE* signal, Fig. 4.9 (e).

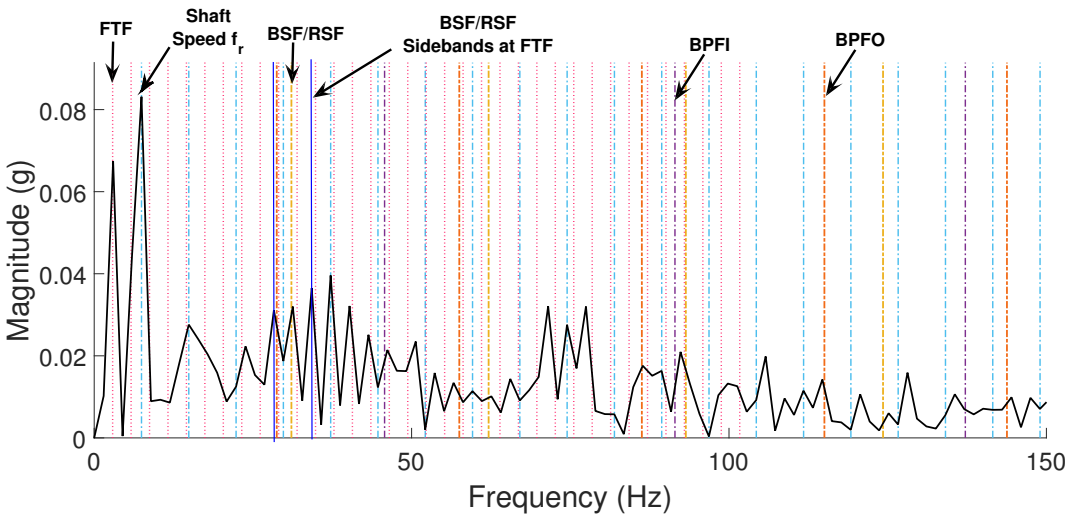


Figure 4.12: *Envelope spectrum* of reconstructed signal *RE* slight defect (*Tian-Yau Wu* database), the *BSF/RSF* is identified with two sidebands, and there is a large component in *FTF*

4.4 MFPT Society database

To validate the proposed methodology a total of 4 signals were selected from this database, two signals with *OR* defect and two signals with *IR* fault and loads of 250 and 300 lbs in each type of fault. Their respective waveforms can be observed in Fig. 4.13. Compared with *Tian-Yau Wu* signals, the waveforms of the *IR* fault change and there are some small impulses, this can be caused by the fact that there is a load applied to the shaft and perhaps the contact between the *RE* and the *IR* are stronger than without a load.

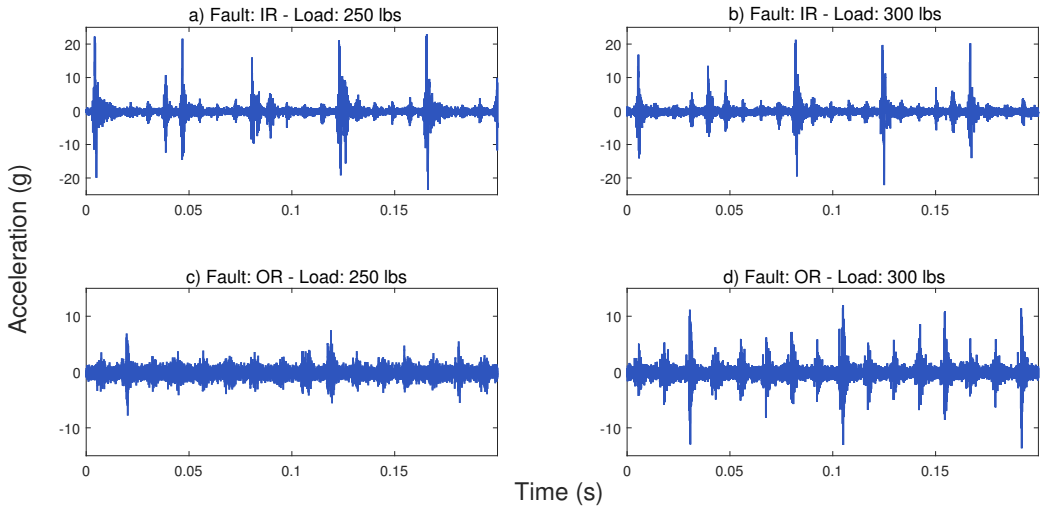


Figure 4.13: Analyzed signals to validate the proposed methodology (*MFPT* database)

4.4.1 MFPT Society Results

The results of the proposed methodology for the analyzed signals are presented. Also, the *Envelope spectrum* of the first four *IMFs* was made to determine if the selection of the *IMFs* to reconstruct the signal was correct.

- **IR Signals.** The analysis of individual *IMFs* for both *IR* defect signals revealed that there is fault information in all of the *IMFs*. However, the 2^{nd} and 4^{th} are the most optimum for the reconstruction of the signal, and the selection was made correctly. For both *Envelope spectrum*, the *BPFI* with sidebands and harmonics were identified, the amplitude of the *BPFI* does not present a substantial increase when the load is increased from 250 to 300 lbs. These results can be found in *Appedix E*. In Fig. 4.14, the comparison of the methodology against the traditional *Envelope spectrum*, and the *FFT* is shown. It can be observed that when the fault information is divided in more than two *IMFs* and the reconstruction of the signal is made only with two *IMFs* there is leakage of information.

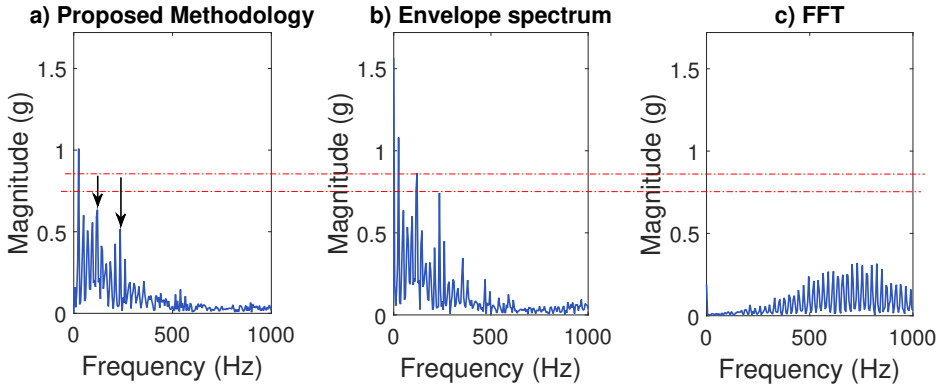


Figure 4.14: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *OR* 300 lbs of load (*MFPT* database), the amplitude in the proposed methodology decreased because the fault information is divided in more than two *IMFs*

- **OR Signals.** The result of the proposed methodology applied step by step to the *OR* defect signal with load of 250 lbs can be found in *Chapter 3*. For the *OR* defect signal with load of 300 lbs, the *Envelope spectrum* show the *BPFO* and one harmonic. Also, unusual sidebands spaced at f_r for the *BPFO* were identified, somehow the impulses of the *OR* fault are being modulated and therefore the sidebands appeared in the spectrum, Fig. 4.15. For this case the fault information was divided in the four *IMFs*; however, the most significant *IMFs* were

selected correctly. As in the *IR* fault case, there is leakage of information and the amplitude of the proposed methodology is smaller than in the traditional *Envelope spectrum*.

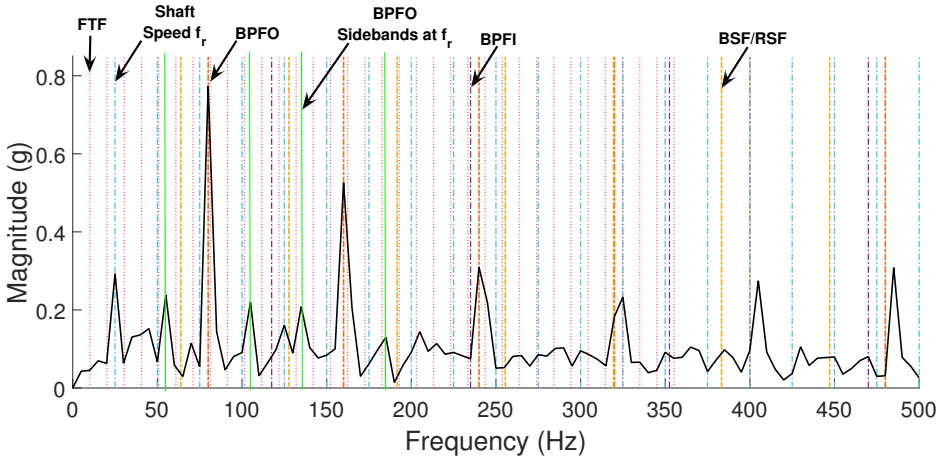


Figure 4.15: *Envelope spectrum* of reconstructed signal *OR-300 lbs* (*MFPT* database), small sidebands spaced at f_r are presents for *BPFO*

There are two important features that can be evaluated in the proposed methodology: (1) the effectiveness of the *KR* for the *IMF* selection, and (2) the identification of fault components in the analyzed signals. For the evaluation of the effectiveness of the *IMF* selection, a table was constructed with the information of the selected *IMFs*, the *KR* values, the percentage of increase of *KR* in the selected *IMF*, and if the *IMFsselected*correctly. were , Table 4.1

On the other hand, for the evaluation of the identification of fault components a table that summarizes the amplitude of the expected faults in the signal was constructed. Also, this table includes if fault components different to the expected were found. The analyzed signals were labelled according the difficulty of the diagnosis:

- Easy: The largest peaks coincide with the bearing fault frequencies. This does not consider the shaft speed frequency.
- Medium: High peaks coincide with the bearing fault frequencies, but there are too many components with similar amplitude that are associated to the shaft frequency and their harmonics, or there are other components with high amplitude considered as noise.
- Hard: No peaks coincide with the bearing fault frequencies.

In *Appendix* E, the remaining results are presented[1] as well as the information of the signals necessary to reproduce these results.

[1] All figures and relevant information can also be found in http://bit.ly/ThesisRCG

Table 4.1: *IMF* selection results using *KR*

Database	Signal	Selected *IMFs*	Largest *KR*	2^{nd} Largest *KR*	*KR* Baseline[a]				% of *KR* increase	Right *IMF* selection
					IMF 1	*IMF 2*	*IMF 3*	*IMF 4*		
CWRU	*IR* (Y)	1 and 2	3.3458	0.487	0.19	0.10	0.05	0.08	17.61	Yes
	IR (P)	1 and 2	3.2232	0.2381	0.19	0.10	0.05	0.08	16.96	Yes
	IR (N)	1 and 2	2.1971	1.0851	0.19	0.10	0.05	0.08	11.56	Yes
	OR (Y)	1 and 2	4.6961	0.7457	0.19	0.10	0.05	0.08	24.72	Yes
	OR (P)	1 and 2	0.2312	0.0851	0.19	0.10	0.05	0.08	1.22	Yes
	OR (N)	1 and 2	0.1951	0.0957	0.19	0.10	0.05	0.08	1.03	Yes
	RE (Y)	1 and 4	9.2034	0.7347	0.19	0.10	0.05	0.08	48.44	No
	RE (P)	1 and 2	0.3719	0.1076	0.19	0.10	0.05	0.08	1.96	Yes
	RE (N)	1 and 2	0.3136	0.1551	0.19	0.10	0.05	0.08	1.65	Yes
Tian-Yau Wu	*IR* SL	1 and 2	2.4681	1.053	0.13	0.16	0.14	0.08	18.99	Yes
	IR SE	1 and 2	3.6676	2.8158	0.13	0.16	0.14	0.08	28.21	Yes
	OR SL	1 and 2	1.2209	0.6578	0.13	0.16	0.14	0.08	9.39	Yes
	OR SE	1 and 2	1.2166	0.8275	0.13	0.16	0.14	0.08	9.36	Yes
	RE SL	1 and 2	1.1913	0.632	0.13	0.16	0.14	0.08	9.16	Yes
MFPT	*IR* 250 Lbs	2 and 4	46.1927	36.15	1.05	1.38	1.47	0.97	33.47	Yes
	IR 300 Lbs	2 and 4	34.2469	33.955	1.05	1.38	1.47	0.97	24.82	Yes
	OR 250 Lbs	1 and 2	4.6825	4.3109	1.05	1.38	1.47	0.97	4.46	Yes
	OR 300 Lbs	1 and 2	33.7207	9.04989	1.05	1.38	1.47	0.97	32.11	Yes

[a] *KR* values for normal condition signal

Table 4.2: Bearing fault results

Database	Signal	Difficulty	Fault amplitude		Maximum amplitude	Additional significant peaks [a]
			Fault freq.	Harmonic		
CWRU	*IR* (Y)	Easy	0.3061	0.1606	0.3107 Shaft	*BPFI* H & SB / Shaft
	IR (P)	Easy	0.07886	0.03071	0.07886 (fault)	*BPFI* H & SB / Shaft / *FTF*
	IR (N)	Easy	NA	NA	0.22182 *BPFO*	*BPFO* / *FTF*
	OR (Y)	Easy	0.6736	0.3811	0.6736 (fault)	Negligible
	OR (P)	Easy	NA	NA	0.01724 Shaft SB	*BPFI* with H & SB / Shaft SB
	OR (N)	Medium	NA	NA	0.015 noise	*BPFI* & H / *BSF/RSF* H
	RE (Y)	Easy	0.5461	NA	0.7393 3x*FTF*	*FTF* & H / *BPFO*
	RE (P)	Easy	NA	0.008	0.0323 *IR*SB	*BPFI* & SB / Shaft & Shaft SB
	RE (N)	Easy	NA	0.004	0.0237 Shaft	*BPFI* with SB & H / Shaft
Tian-Yau Wu	*IR* SL	Easy	0.05406	0.01995	0.133 Shaft	*BPFI* SB / *BSF/RSF* & SB
	IR SE	Medium	0.04758	NA	0.216 *FTF*	*FTF* / *BSF/RSF*,
	OR SL	Easy	0.1307	0.04285	0.1307 (fault)	Negligible
	OR SE	Easy	0.1371	0.0824	0.1371 (fault)	*FTF*
	RE SL	Easy	0.032	NA	0.082314 Shaft	*BSF/RSF* SB / *FTF*
MFPT Society	*IR*250 Lbs	Medium	0.6244	0.5279	1.236 Shaft	*BPFI* H & SB / Shaft & H
	IR 300 Lbs	Medium	0.6347	0.515	1.009 Shaft	*BPFI* H & SB / Shaft
	OR 250 Lbs	Easy	0.3616	0.1837	4.6825 (fault)	*BPFO* H / *BPFI*
	OR 300 Lbs	Easy	0.7733	0.5258	33.7207 (fault)	*BPFO* H & SB at Shaft freq.

[a] SB = Sidebands, H = Harmonics, SL = Slight, SE = Severe

4.5 Statistical Parameters for IMF selection

Statistical parameters can represent the characteristics of time series like vibrations. This characteristic can be exploited to identify *IMFs* with fault information. In this research, an analysis of 17 time domain features was made to determine which one can represent faulty signals the best.

The test consisted in analyzing the the change on amplitude of these statistical parameters from a normal condition signal to a faulty signal. First, the normal condition signal was decomposed using the *EMD*, then the 17 statistical parameters were computed for each *IMF* to make the comparison of these values against the obtained from the faulty signal. Then, the *EMD* was applied to the faulty signal, and the statistical features were computed for each *IMF*. The next step was to perform an *Envelope spectrum* analysis of the *IMFs* of the faulty signal to identify the component with faulty information. After the best *IMF* was identified, the comparison between the values of normal condition and fault condition was made for each statistical parameter. If the highest change of amplitude the statistical parameter corresponded to the *IMF* with fault information, this change was saved to summarize the results.

This test was applied to five signals from *CWRU* and *MFPT* databases. In Table 4.3, the result of this analysis is presented. In the first column the 17 statistical parameters can be found, the remaining columns correspond to the analyzed signals; for each signal the most significant *IMF* can be observed. If the highest change of amplitude of the statistical parameter corresponded to the most significant *IMF*, that value is included in the table. On the other hand, a NA was given if the largest change of amplitude for the statistical parameters that did not correspond to the *IMF* with faulty information.

From the 17 parameters, only the *Kurtosis*, and *KR* identified the most significant *IMF* in the five cases. It can be observed that the *KR* produces higher changes of amplitude in the comparison with normal condition, for this reason *KR* was selected to automate the *IMF* selection. Another reason to choose the *KR* over the *Kurtosis* is that *Kurtosis* only represents the impulsiveness from signals, while *KR* represents impulsiveness and the signal amplitude related to the resonance frequency of the system.

Table 4.3: Statistical parameters and their variation based on normal condition signals. Only *Kurtosis* and *RMS* identified the most significant *IMF* in each signal.

Parameter	CWRU		MFPT		
	IR 0 HP	*IR* 3 HP	*OR* 270 Lbs	*IR* 0 Lbs	*IR* 100 Lbs
	IMF 1 [a]	*IMF 1*	*IMF 4*	*IMF 1*	*IMF 1*
RMS	4.4	4.56	1.1	3.2	NA
Kurtosis	1.63	1.67	2.1	20.79	11.05
KR	7.16	7.6	2.31	66.57	22.33
Shannon Entropy	NA	NA	NA	-17	NA
Peak	7.1	6.82	NA	12.1	NA
Pk-Pk	6.73	6.55	NA	11.99	NA
CF	1.59	1.49	NA	3.78	2.69
Skewness	NA	NA	NA	NA	NA
Variance	19.29	20.84	1.2	10.25	NA
SD	4.39	4.57	1.1	3.2	NA
Mean	NA	NA	NA	4.6	NA
Root	4.27	4.34	NA	1.43	-11.46
Clearance	1.64	NA	NA	8.44	1.22
Shape	1.04	NA	1.13	1.75	NA
Impulse	1.65	NA	NA	6.6	NA
RMSxVariance	84.7	95.11	1.32	32.85	NA
RMSxSD	19.28	20.83	1.2	10.25	NA

[a] These are the most significant IMFs of the signal

Chapter 5

Conclusions

Vibration analysis is of great importance for *Condition Based Maintenance*, it allows to capture information from components that are not easy to access. The challenge in vibration analysis resides in the extraction of bearings signature information from the signal.

In this work, a methodology based on *HHT* for bearing fault detection is proposed, this methodology allows to detect faults in incipient stages. First, the vibration signal is decomposed using the *EMD*, after the first four *IMFs* are generated, the two *IMFs* with most fault information are used to reconstruct the signal without noise. To identify the most useful *IMFs* the *KR* is computed for the four *IMFs* and the two with highest amplitudes are used to reconstruct the signal. After the signal is reconstructed, the *Envelope spectrum* is used to extract the bearing fault information from the resonance frequency of the system and make a diagnosis. The *Envelope spectrum* allows to identify bearing faults in incipient stages since vibration produced by these types of faults are modulated in the resonance frequencies.

For the 18 analyzed signals good results were obtained, in 14 signals the identification of fault components can be easily made, for these signals the fault components were the largest in the spectrum. For the remaining four signals, the fault identification was possible; but, there were peaks with similar amplitude of the fault components. A disadvantage was found when the *EMD* produces more than two *IMFs* with fault information. As the signal reconstruction is limited to only two *IMFs*, the reconstructed signal produces a spectrum with smaller amplitude than the traditional *Envelope spectrum*.

It is important to mention that the effectiveness of the *Empirical Mode Decomposition* is not improved with this methodology, the key element of the proposed methodology is the automation of the *IMF* selection. For the 18 presented cases where two *IMFs* were selected for each signal, only one *IMF* was wrongly selected. This *IMF* had the end defect produced in the sifting process of the *EMD*. Unfortunately, statistical parameters are sensitive to these defects, causing their amplitudes to increase leading to a wrong selection.

5.1 Contributions

With the creation of the *Hilbert-Huang Transform* and the *EMD*, the controversy of the selection of useful *Intrinsic Mode Functions* emerged. Some authors have proposed to perform visual inspection of the spectrum of the *IMFs* to detect the component with information from the resonance frequency of the system [Tsao *et al.*, 2010] [Tsao *et al.*, 2010]; but, this approach can be time consuming, while other authors limit the analysis to one or two *IMFs* [Dubey and Agrawal, 2015], and they do not necessarily present the fault information.

In this research, a methodology is proposed to automate the *IMF* selection. The selection is based on the *KR* that can identify faulty *IMFs* from the resonance frequency of the system; therefore this methodology is useful to detect bearing faults in incipient stages. The *KR* values for the most useful *IMFs* are higher than for the others, and the identification of these *IMFs* can be programmed to avoid the visual inspection and obtain results in seconds.

5.2 Future work

This investigation made a contribution for the automation of bearing fault detection. In addition to this work some opportunities have been found:

- *EMD* has proven to be a reliable tool for signal analysis; however further research is necessary to determine a parameter to measure end defects and improve the methodology.
- This methodology can be used to create a hybrid methodology by its combination with *Wavelet Transform*. The *EMD* and statistical features can be used to filter the information from the resonance frequency band of the system and the *WT* can be used to eliminate noise from the reconstructed signal.
- The addition and implementation of intelligent methods as *Neural Networks, Support Vector Machine*, and *Deep Learning* can be made to classify the type of fault without a visual inspection.

Bibliography

[Arfken and Weber, 2005] G. B. Arfken and H. J. Weber. *Mathematical Methods for Physicists*. 6th Edition. Elsevier Academic Press Publications, 2005.

[Bell *et al.*, 1985] R. N. Bell, D. W. McWilliams, P. O'Donnel, C. Singh, and S. J. Wells. Report of Large Motor Reliability Survey of Industrial and Commercial Installations, Part I. *IEEE Trans on Industry Applications*, IA-21(4):853–864, 1985.

[Briceno, 2012] J. E. Briceno. *Principios de las Comunicaciones*. 3^{rd} Ed. Mérida: Universidad de los Andes, 2012.

[Courant and Hilbert, 1953] R. Courant and D. Hilbert. *Methods of Mathematical Physics*. 1^{st} Edition. Interscience Publishers, 1953.

[CWRU, 1999] CWRU. Case Western Reserve University Bearing Data Center Seeded Fault Test Data. http://csegroups.case.edu/bearingdatacenter/pages/apparatus-procedures, Accessed 01-12-2016. 1999.

[Djebala *et al.*, 2015] A. Djebala, M. K. Babouri, and N. Ouelaa. Rolling Bearing Fault Detection using a Hybrid Method based on Empirical Mode Decomposition and Optimized Wavelet Multi-Resolution Analysis. *Int. J. Adv. Manuf. Technol.*, 79:2093–2105, 2015.

[Dubey and Agrawal, 2015] R. Dubey and D. Agrawal. Bearing Fault Classification using ANN-based Hilbert Footprint Analysis. *IET Science, Measurement & Technology*, 9(8):1016–1022, 2015.

[Feldman, 2006] M. Feldman. Time-varying vibration decomposition and analysis based on the Hilbert transform. *J of Sound and Vibration*, 295(3-5):518–530, 2006.

[Feldman, 2011a] M. Feldman. Hilbert Transform in Vibration Analysis. *Mech. Syst. and Signal Process.*, 25(3):735–802, 2011.

[Feldman, 2011b] M. Feldman. *Hilbert Transform in Vibration Analysis*. 1^{st} Ed. J. Wiley & Sons, 2011.

[Gao and Yan, 2010] R. X. Gao and R. Yan. *Wavelets: Theory and Applications for Manufacturing*. 1^{st} Ed. Springer, 2010.

[Gao *et al.*, 2008] Q. Gao, C. Duan, H. Fan, and Q. Meng. Rotating Machine Fault Diagnosis using Empirical Mode Decomposition. *Mechanical Systems and Signal Processing*, 22(5):1072–1081, 2008.

[Goldman, 1999] S. Goldman. *Vibration Spectrum Analysis: a Practical Approach*. 2^{nd} Ed. Industrial Press Inc., 1999.

[Huang and Wu, 2009] N.E. Huang and Z. Wu. Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method. *Adv. in Adaptive Data Analysis*, 1(1):1–41, 2009.

[Huang *et al.*, 1998] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proc. R. Soc. Lond. A. Math. Phys. Sci.*, 454:903–995, 1998.

[Immovilli *et al.*, 2010] F. Immovilli, A. Bellini, C. Tassoni, and R. Rubini. Diagnosis of Bearing Faults in Induction Machines by Vibration or Current Signals: A Critical Comparison. *IEEE Trans on Industry Applications*, 46(4):1350–1359, 2010.

[ISO, 1995] ISO. ISO 10816 Part 1: General Guidelines. *In Mechanical vibration – Evaluation of Machine Vibrations by Measurements on Non-Rotating Parts*, pages 1–19, 1995.

[ISO, 2009] ISO. ISO 10816 Part 3: Industrial Machines with Nominal Power Above 15 kW and Nominal Speeds Between 120 r/min and 15 000 r/min when Measured in Situ. *In Mechanical vibration – Evaluation of Machine Vibrations by Measurements on Non-Rotating Parts*, pages 1–12, 2009.

[ISO, 2012] ISO. ISO 2954 Requirements for Instruments for Measuring Vibration Severity. *In Mechanical Vibration of Rotating and Reciprocating Machinery*, pages 1–13, 2012.

[ISO, 2014] ISO. ISO/TR 17243 Part 1: Spindles with Rolling Element Bearings and Integral Drives Operating at Speeds between 600 min^{-1} and 30 000 min^{-1} . *In Machine Tool Spindles – Evaluation of Machine Tool Spindle Vibrations by Measurements on Spindle Housing*, pages 1–18, 2014.

[James, 2011] J. F. James. *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering*. 3^{rd} Ed. Cambridge University Press, 2011.

[Junsheng *et al.*, 2006] C. Junsheng, Y. Deije, and Y. Yu. A Fault Diagnosis Approach for Roller Bearings based on EMD Method and AR Model. *Mech. Syst. and Signal Process.*, 20:350–362, 2006.

[Lei, 2016] Y. Lei. *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*. 1^{st} Ed. Butterworth-Heinemann, 2016.

[Li *et al.*, 2009] H. Li, Y. Zhang, and H. Zheng. Hilbert-Huang Transform and Marginal Spectrum for Detection and Diagnosis of Localized Defects in Roller Bearings. *J. of Mechanical Science and Technology*, 23(2):291–301, 2009.

[Li *et al.*, 2010] H. Li, Y. Wang, and Y. Ma. Ensemble Empirical Mode Decomposition and Hilbert-Huang Transform Applied to Bearing Fault Diagnosis. *2010* 3^{rd} *Int. Congress on Image and Signal Processing*, 7:3413–3417, 2010.

[Maragos and Schafer, 1987] P. Maragos and R. Schafer. Morphological filters–Part I: Their Set-Theoretic Analysis and Relations to Linear Shift-Invariant Filters. *IEEE Trans on Acoustics*, 35(8):1153–1169, 1987.

[McFadden and Smith, 1984] P. D. McFadden and J. D. Smith. Model for the Vibration Produced by a Single Point Defect in a Rolling Element Bearing. *J. of Sound and Vibration*, 96(1):69–82, 1984.

[Niu *et al.*, 2017] L. Niu, H. Cao, and X. Xiong. Dynamic Modeling and Vibration Response Simulations of Angular Contact Ball Bearings with Ball Defects Considering the Three-dimensional Motion of Balls. *Tribology Int*, 109:26 – 39, 2017.

[Osman and Wang, 2013] S. Osman and W. Wang. An Enhanced HilbertHuang Transform Technique for Bearing Condition Monitoring. *Measurement Science and Technology*, 24(8), 2013.

[Osman and Wang, 2016a] S. Osman and W. Wang. A Morphological Hilbert-Huang Transform Technique for Bearing Fault Detection. *IEEE Trans. on Instrumentation and Measurement*, 65(11):2646–2656, 2016.

[Osman and Wang, 2016b] S. Osman and W. Wang. A normalized Hilbert-Huang Transform Technique for Bearing Fault Detection. *J. of Vibration and Control*, 22(11):2771–2787, 2016.

[Randall and Antoni, 2011] R.B. Randall and J. Antoni. Rolling Element Bearing Diagnostics A Tutorial. *Mech. Syst. and Signal Processing*, 25:485–520, 2011.

[Randall, 2011] R.B. Randall. *Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications*. EBL-Schweitzer. J. Wiley & Sons, 2011.

[Scheffer and Girdha, 2004] C. Scheffer and P. Girdha. *Practical Machinery Vibration Analysis and Predictive Maintenance* . 1^{st} Ed. Newnes, 2004.

[Shannon, 1948] C. E. Shannon. A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, 27:379–423, 1948.

[Sinha, 2014] J. K. Sinha. *Vibration Analysis, Instruments, and Signal Processing*. 1^{st} Ed. CRC press, 2014.

[Smith and Randall, 2015] W. A. Smith and R. B. Randall. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100–131, 2015.

[Smith, 2005] J. S. Smith. The Local Mean Decomposition and Its Application to EEG Perception Data. *J of the Royal Society Interface*, 2(5):443–454, 2005.

[Society, 2013] MFPT Society. Condition Based Maintenance Fault Database for Testing of Diagnostic and Prognostics Algorithms Data Assembled and Prepared on behalf of MFPT. http://mfpt.org/fault-data-sets/, Accessed 03-16-2018. 2013.

[Thorsen and Dalva, 1995] O. V. Thorsen and M. Dalva. A Survey of Faults on Induction Motors in Offshore Oil Industry, Petrochemical Industry, Gas Terminals, and Oil Refineries. *IEEE Trans on Industrial Applications*, 31(5):1186–1196, 1995.

[Tsao *et al.*, 2010] W. Ch. Tsao, Y.F. Li, and M Ch. Pan. Resonant-Frequency Band Choice for Bearing Fault Diagnosis based on EMD and Envelope Analysis. volume 8. IEEE, 2010.

[Tsao *et al.*, 2012] W. Ch. Tsao, Y.F. Li, D. D. Le, and M Ch. Pan. An Insight Concept to Select Appropriate IMFs for Envelope Analysis of Bearing Fault Diagnosis. *Measurement*, 45:1489–1498, 2012.

[Tse and Leung, 2010] P.W. Tse and C.T. Leung. *Advanced System For Automatically Detecting Faults Occurring in Bearings*. Nova Science Publishers, 2010.

[Wiggins, 1978] R. A. Wiggins. Minimum Entropy Deconvolution. *Geoexploration*, 16(1-2):21 – 35, 1978.

[Wu *et al.*, 2016] T. Y. Wu, C. H. Lai, and D. C. Liu. Defect Diagnostics of Roller Bearing using Instantaneous Frequency Normalization Under Fluctuant Rotating Speed. *J. Mech. Sci. Technol.*, 30(3):1037–1048, 2016.

[Yan and Gao, 2006] R. Yan and R.X. Gao. Hilbert-Huang Transform-based Vibration Signal Analysis for Machine Health Monitoring. *IEEE Trans. on Instrumentation and Measurement*, 55(6):2320–2329, 2006.

[Yan *et al.*, 2014] R. Yan, R.X. Gao, and X. Chen. Wavelets for Fault Diagnosis of Rotary Machines: A Review with Applications. *Signal Processing*, 96:1–15, 2014.

[Yang *et al.*, 2007] Y. Yang, D. Yu, and J. Cheng. A Fault Diagnosis Approach for Roller Bearing based on IMF Envelope Spectrum and SVM. *Measurement*, 40:943–950, 2007.

[Yu *et al.*, 2005] D. Yu, J. Cheng, and Y. Yang. Application of EMD Method and Hilbert Spectrum to the Fault Diagnosis of Roller Bearings. *Mechanical Systems and Signal Processing*, 19(2):259–270, 2005.

[Yu *et al.*, 2006] Y. Yu, YuDejie, and C. Junsheng. A Roller Bearing Fault Diagnosis Method based on EMD Energy Entropy and ANN. *J. of Sound and Vibration*, 294(1–2):269–277, 2006.

[Zhang *et al.*, 2011] P. Zhang, Y. Du, B. Lu, and T. G. Habelter. A Survey of Condition Monitoring and Protection Methods for Medium-Voltage Induction Motors. *IEEE Trans on Industry Applications*, 47(1):34–46, 2011.

Appendix A

Acronyms and Variables Descriptions

Table A.1: Acronyms Definitions

Acronyms	*Description*	*Acronyms*	*Description*
ANN	Artificial Neural Network	BPFI	Ball Pass Frequency of Inner Race
BPFO	Ball Pass Frequency of Outer Race	BSF	Ball Spin Frequency
CF	Crest Factor	CBM	Condition Based Maintenance
CMF	Combined Mode Function	CWT	Continuous Wavelet Transform
CWRU	Case Western Reserve University	DCC	Direct Current Component
DQ	Direct Quadrature	DMI	Deficiency Mutual Information
DWT	Discrete Wavelet Transform	EEMD	Ensemble Empirical Mode Decomposition
eHHT	enhanced Hilbert-Huang Transform	EDM	Electro-discharge Machining
ELM	Extreme Learning Machine	EMD	Empirical Mode Decomposition
FFT	Fast Fourier Transform	FTF	Fundamental Train Frequency
HT	Hilbert Transform	HHT	Hilbert Huang Transform
IMF	Intrinsic Mode Function	IR	Inner Race
MKD	Maximum Kurtosis Deconvolution	MED	Minimum Entropy Deconvolution
MFPT	Machinery Fault Prevention Technology	NADA	Noise Assisted Data Analysis
NCM	Normalized Correlation Measure	NHHT	Normalized Hilbert-Huang Transform
OR	Outer Race	PSD	Power Spectral Density
RE	Rolling Element	RSF	Roller Spin Frequency
SD	Standard Deviation	SGWT	Second Generation Wavelet Transform
STFT	Short Time Fourier Transform	SVM	Support Vector Machine
WMRA	Wavelet Multi-Resolution Analysis	WPT	Wavelet Packet Transform

Table A.1: Acronyms Definitions (Continued)

Acronyms	*Description*	*Acronyms*	*Description*
WT	Wavelet Transform	WVD	Wigner-Ville Distribution

Table A.2: Variables Definitions

Variables	*Description*	*Variables*	*Description*
$A(t)$	Signal envelope	C_n	Scales of *WT*/ *IMFs* of *HHT*
d	*RE* diameter	D	Pitch diameter
F_q	Nyquist frequency	f_r	Shaft frequency
F_s	Sampling frequency	$F(t)$	*FT* of the signal
μ	Mean value	N	# of data points
P	*Cauchy Principal Value*	P_n	Probability of one event to occur
r	Revolutions / residual in *EMD*	s	Scaling parameter of *WT*
t	Time	T	Window size in seconds
τ	Period	$X(t)$	Signal
$\tilde{x}(t)$	*HT* of the signal	y_n	Amplitude of n^{th} data point
$e^{-i\omega t}$	Kernel function of *FT*	ω	frequency
ϕ	Angle of contact in bearings	$\Phi(t)$	Phase of the signal
ψ	*Mother Wavelet*	σ	Standard Deviation

Appendix B

Window Size Selection

Some of the monitoring system requirements for bearing fault detection are low computational time and good spectrum visualization (spectrum resolution) to identify bearing fault components. The proposed methodology includes two methods, the *Hilbert-Huang transform (HHT)* and the *Envelope spectrum*. The *HHT* is known by its adaptive method *Empirical Mode Decomposition (EMD)*, an iterative process that can demand elevated computational time if the analyzed *window size* or signal length is inappropriate (too large or too many data points) because of an over-sampling frequency selection, on the other hand, the *Envelope spectrum* resolution is affected by the signal length (the larger the better spectrum resolution), and the sampling frequency that defines the maximum frequency in the spectrum. The solution for one problem worsen the other problem, e.g. if a small sampling frequency is selected the computational time will be reduced, but the spectrum may not include fault frequencies or the resolution in the spectrum may not be appropriate to distinguish fault frequencies. This section is dedicated to analyze this problem and find an optimum window size and sampling frequency for the methodology.

Tests for *EMD* computational time and *Envelope spectrum* resolution are based on revolutions of the spindle, and the optimum window selection is given in revolutions.

B.1 CWRU Test

This data base includes signals with speed fluctuations from 1,720 - 1,797 RPM or 28.66 - 29.95 Hz. As tests are based on revolutions and the signals are recorded in time units, to guarantee that 1 revolution is covered in all signals, the period of the minimum speed was used as 1 revolution ($T = 1/f = 1/28.66 = 0.0349$ seconds). The signals length used in these tests is about 10 seconds with sampling frequency $Fs = 12,000$ samples/second, this makes $Fs/f_r = 12,000/28.66 = 418$ data points/revolution.

B.1.1 EMD Computational Time

In this analysis, a total of 16 signals corresponding to 4 bearing conditions (*Normal*, *IR*, *OR*, and *RE* Fault) under 4 conditions of load (0, 1, 2, 3 HP) are analyzed. For this test, the *mean computational time* of 20 samples of 1 revolution (418 data points) of each signal is compared against mean computational times of 20 samples of 5 revolutions (2,090 data points) and 20 samples of 10 revolutions (4,180 data points) to determine the proper window. The samples of the signal are selected every 5 seconds starting in second 0.001, i.e. sample 2 begins on second 0.501.

In Tables B.1, B.2, B.3, and B.4 can be found the name of the file, load for the specific test, the mean time of the *EMD* for the test Time_r (where r is the number of revolutions), and the SD of the mean times of the 20 samples.

It is observed in Table B.1 for *Normal* condition that computation time of the EMD increases as the number of revolutions (data points) increase, and not only the time but their variability. An interesting comparison is the mean time of 1 revolution against 5 and 10 revolutions, even though 5 revolutions means 5 times the window size of 1 revolution the time does not increase 5 times, it only increases around 2.5 times, and 4.8 times for 10 revolutions, which means that 418 data points are processed in approximately half of the time. The same behaviour can be observed for *IR* and *RE* faults condition times (Tables B.2 and B.4), but there is an exception in Table B.3 for *OR* fault condition times. This does not allow to determine a direct relation between window size and expected time.

Table B.1: Mean computational times (s) for 1, 5, and 10 revolutions of *CWRU Normal* condition signals

Normal Condition							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
X097	0	0.0562	0.0111	0.1722	0.0423	0.3508	0.1068
X098	1	0.0697	0.0157	0.165	0.0365	0.3153	0.0888
X099	2	0.0656	0.0137	0.1575	0.0435	0.2644	0.0585
X100	3	0.0563	0.0085	0.1315	0.0298	0.2652	0.0606
Mean		0.06195	0.012546	0.15655	0.038412	0.298925	0.081219
Increasing in computational time compared with 1 revolution				2.52704		4.82526	

Table B.2: Mean computational times (s) for 1, 5, and 10 revolutions of *CWRU IR* fault condition signals

Inner Race Fault 0.007 in							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
X105	0	0.0704	0.0138	0.1698	0.0489	0.4403	0.3048
X106	1	0.0718	0.0119	0.2078	0.1466	0.4342	0.1811
X107	2	0.0766	0.0148	0.2163	0.0995	0.3319	0.102
X108	3	0.0705	0.0136	0.1719	0.0268	0.3736	0.1135
Mean		0.072325	0.013565	0.19145	0.092873	0.395	0.192994
Increasing in computational time compared with 1 revolution				2.64708		5.46150	

Table B.3: Mean computational times (s) for 1, 5, and 10 revolutions of *CWRU OR* fault condition signals

Outer Race Fault 0.007 in							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
X130	0	0.0694	0.0152	0.3064	0.4384	1.5099	1.5832
X131	1	0.0683	0.0179	0.2645	0.1795	1.3393	1.7028
X132	2	0.0883	0.0566	0.2846	0.113	1.363	2.2271
X133	3	0.0711	0.0185	0.2291	0.1086	0.4721	0.18
Mean		0.074275	0.032005	0.27115	0.249488	1.171075	1.612329
Increasing in computational time compared with 1 revolution				3.65062		15.76675	

Table B.4: Mean computational times (s) for 1, 5, and 10 revolutions of *CWRU RE* fault condition signals

Rolling Element Fault 0.007 in							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
X048	0	0.0745	0.0231	0.1486	0.039	0.3317	0.0975
X049	1	0.0779	0.0241	0.1585	0.0463	0.3007	0.0748
X050	2	0.0682	0.0134	0.1632	0.038	0.3027	0.1206
X051	3	0.0759	0.0349	0.1788	0.0446	0.2826	0.0783
Mean		0.074125	0.02506	0.162275	0.042124	0.304425	0.094573
Increasing in computational time compared with 1 revolution				2.18921		4.10691	

B.1.2 Envelope Spectrum Resolution

Envelope spectrum resolution is crucial to identify fault frequency components. The expected frequency resolution is affected by the sample frequency and the length of the processed signal (Sinha p. 151) [Sinha, 2014]. Based on Sinha's example a new table (Table B.5) for *CWRU* data is constructed to explore the expected maximum frequency and the *Envelope spectrum* resolution when different window sizes are processed.

Table B.5: *CWRU Envelope spectrum* resolution based on shaft revolutions r and $Fs = 12,000$ samples/second

r	N	$dt = 1/Fs$	$T = Ndt$	$df = 1/T$	$Fq = Fs/2$	$nf = N/2$
1	418	8.33333E-05	0.03483333	28.708134	6,000	209
2	836	8.33333E-05	0.06966667	14.354067	6,000	418
3	1,254	8.33333E-05	0.1045	9.56937799	6,000	627
4	1,672	8.33333E-05	0.13933333	7.17703349	6,000	836
5	2,090	8.33333E-05	0.17416667	5.74162679	6,000	1045
6	2,508	8.33333E-05	0.209	4.784689	6,000	1254
7	2,926	8.33333E-05	0.24383333	4.101162	6,000	1463
8	3,344	8.33333E-05	0.27866667	3.58851675	6,000	1672
9	3,762	8.33333E-05	0.3135	3.18979266	6,000	1881
10	4,180	8.33333E-05	0.34833333	2.8708134	6,000	2090

Where r is the window size in revolutions, N is the number of data points, dt the time between two samples, T is the window size in seconds, df is the expected *Envelope spectrum* resolution, Fq is maximum frequency in the spectrum based on Nyquist frequency, and nf is the number of lines in the spectrum plot.

Table B.6: *CWRU* expected fault frequencies (Hz) using $f_r = 28.66$ Hz

FTF/CAGE	*BSF/RSF/RE*	*BPFO/OR*	*BPFI/IR*
11.415278	67.55162	102.7461	155.1939

Based on tables 3.2, 3.3, and 3.4 it is observed that frequency factors are not integer multiples of the shaft speed, and the identification of the fundamental fault frequencies in the spectrum should not be affected by the shaft frequency, however its first harmonic might be close to a harmonic of the speed. i.e. *BPFO* frequency (Table B.6) is 102.7461 Hz and its first harmonic is 205.4922 Hz, and the 6^{th} harmonic of the shaft speed is 28.66Hz x 7 = 200.62 Hz with a difference of 4.8722 Hz that could lead to one peak in the spectrum to represent both one harmonic of the shaft speed and one harmonic of the *BPFO* if the resolution is not adequate to avoid that problem the resolution should be at least the half (2.4361 Hz). Using the formulas in the frequency resolution table (Table B.5) based on the revolutions and data points of *CWRU* we can select a minimum of 12 revolutions ($df = 2.3923$ Hz) for the analysis. On the other hand, if the identification of the first harmonic is not relevant for the analysis a window size of *5 revolutions* is the optimum to reduce the *EMD* computational time and obtain good *Envelope spectrum* resolution.

It can be observed in figure B.1 that fault can not be detected and the fault component can be misinterpreted when the window size for the analysis is not chosen properly (1 revolution in this case), and in figure B.2 the *OR* fault frequency can be clearly identified in 107.3 Hz because the proper window selection.

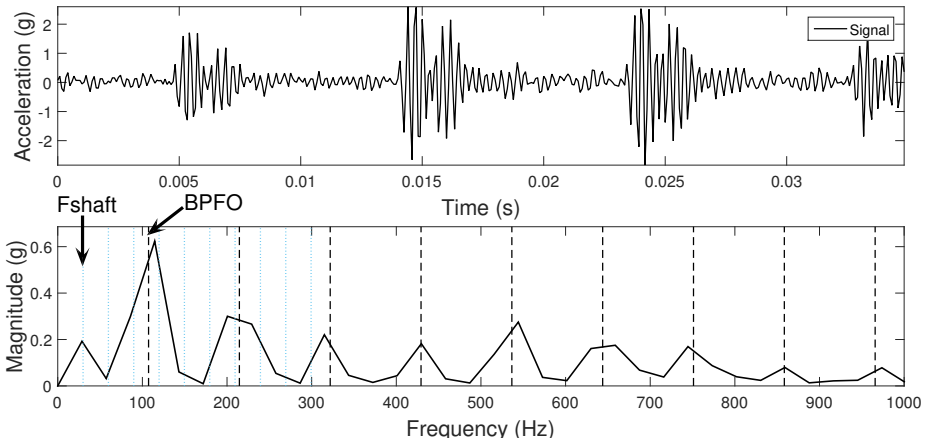


Figure B.1: The upper plot shows 1 revolution (0.0349 s) of the *OR* fault signal and the inferior plot shows the corresponding *Envelope spectrum* from *CWRU* data

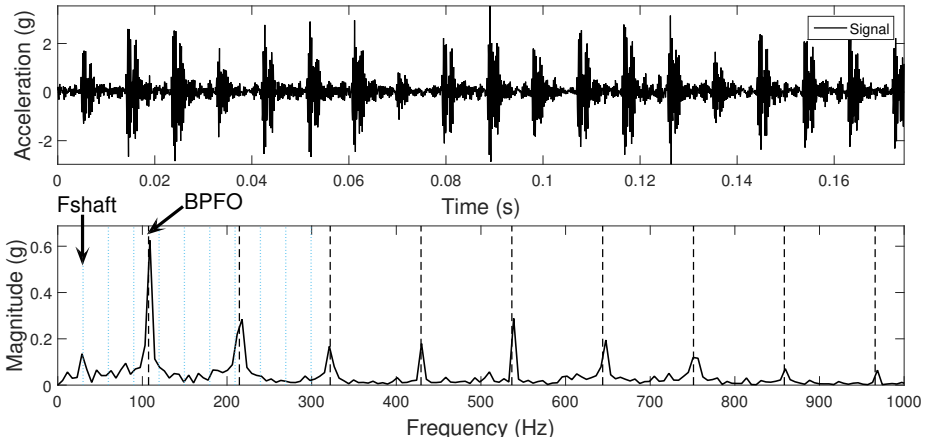


Figure B.2: The plot on the top shows 5 revolutions (0.1745 s) of the *OR* fault signal and the bottom plot shows the corresponding *Envelope spectrum* from *CWRU* data

Since *CWRU* signals were acquired on a controlled environment and are easy to analyze. The previous analysis is repeated for *MFPT* database, which has noisy signals as in real applications and to confirm the previous observations.

B.2 MFPT Test

This data base includes signals with speed of 1,500 RPM or 25 Hz. As tests are based on revolutions and the signals are recorded in time units, the period of 1 revolution ($T = 1/f = 1/25 = 0.04$ seconds). The signals length used in these tests is about 6 seconds or *Normal* and *OR* fault condition at $Fs = 97,656$ samples/second, and 3 seconds for *IR* fault condition at $Fs = 48,828$ samples/second. This makes $Fs/f_r = 97,656/25 = 3,906$ data points/revolution for *Normal* and *OR* fault condition, and $Fs/f_r = 48,828/25 = 1,953$ data points/revolution for *IR* fault condition.

B.2.1 EMD Computational Time

In this analysis, a total of 9 signals corresponding to 3 bearing conditions (*Normal*, *IR*, and *OR* Fault). For this test, the *mean computational time* of 10 samples of 1 revolution (3,906 data points) of each signal is compared against mean computational times of 10 samples of 5 revolutions (19,530 data points) and 10 samples of 10 revolutions (39,060 data points) for *Normal* and *OR* fault condition signals to determine the proper window. * *IR* fault signals are only 3 seconds, for this reason 6 samples are used in this case. The samples of the signal are selected every 5 seconds starting in second 0.001, i.e. sample 2 begins on second 0.501.

In tables B.7, B.8, and B.9 can be found the name of the file, load for the specific test, the mean time of the *EMD* for the test Time_r (where r is the number of revolutions), and the SD of the mean times of the 10 samples or 6 samples for *IR* signals.

It is observed in table B.7 for *Normal* condition that computation time of the *EMD* increases as the number of revolutions (data points) increase. The increase from the mean time of 1 revolution to 5 revolutions is between 6.5 to 10.6 times (Tables B.7, B.8, and B.9), and the increase from 1 to 10 revolutions is up to 30 times larger, this result differ from the results from *CWRU* data and rejects any direct relation between number of revolutions and expected processing time. This suggest that the window size must be as small as possible and meet the requirements for spectrum frequency.

Table B.7: Mean computational times (s) for 1, 5, and 10 revolutions of *MFPT Normal* condition signals

Normal Condition							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
Base_1	270 lb	0.4353	0.159	3.4707	0.5928	10.6683	2.5588
Base_2	270 lb	0.4351	0.1514	3.2161	0.9817	10.454	1.5718
Base_3	270 lb	0.3869	0.2164	3.1869	0.6659	11.22	2.0743
Mean		0.4191	0.1779812	3.29123333	0.76563036	10.7807667	2.10718874
Increasing in computational time compared with 1 revolution				7.85311		25.72361	

Table B.8: Mean computational times (s) for 1, 5, and 10 revolutions of *MFPT OR* fault condition signals

Outer Race Fault							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
OR_1	270 lb	0.345	0.0575	4.2508	3.2074	9.3577	1.1246
OR_2	270 lb	0.3622	0.0485	3.3075	0.7526	9.992	1.9971
OR_3	270 lb	0.3351	0.0581	3.5464	1.2015	12.3515	2.9393
Mean		0.34743333	0.05487595	3.70156667	2.02463361	10.5670667	2.15194626
Increasing in computational time compared with 1 revolution				10.65403		30.41466	

Table B.9: Mean computational times (s) for 1, 5, and 10 revolutions of *MFPT IR* fault condition signals

Inner Race Fault							
File	Load	Time_1	SD_1	Time_5	SD_5	Time_10	SD_10
IR_1	0 lb	0.181	0.0381	1.1627	0.2787	3.0197	0.5341
IR_2	50 lb	0.1965	0.043	1.1982	0.35	3.8365	0.7579
IR_3	100 lb	0.1734	0.0325	1.2657	0.2002	3.3479	0.5769
Mean		0.18363333	0.03810888	1.20886667	0.28299218	3.40136667	0.63047306
Increasing in computational time compared with 1 revolution				6.58305		18.52260	

B.2.2 Envelope Spectrum Resolution

The expected *Envelope spectrum* resolution for *MFPT* signals can be found in tables B.10 and B.11. Although the sampling frequency is reduced by half and the Nyquist frequency is reduced by half too, it is observed that the expected frequency resolution (df) does not change. This indicates that a proper selection of the sample frequency can reduce the number of data points on the selected window size (revolutions) and indeed reduce the computational time.

Table B.10: *MFPT Envelope spectrum* resolution based on shaft revolutions r and $Fs = 97,656$ samples/second

r	N	$dt = 1/Fs$	$T = Ndt$	$df = 1/T$	$Fq = Fs/2$	$nf = N/2$
1	3,906	1.024E-05	0.03999754	25.0015361	48,828	1953
2	7,812	1.024E-05	0.07999508	12.500768	48,828	3906
3	11,718	1.024E-05	0.11999263	8.33384537	48,828	5859
4	15,624	1.024E-05	0.15999017	6.25038402	48,828	7812
5	19,530	1.024E-05	0.19998771	5.00030722	48,828	9765
6	23,436	1.024E-05	0.23998525	4.16692268	48,828	11718
7	27,342	1.024E-05	0.2799828	3.57164801	48,828	13671
8	31,248	1.024E-05	0.31998034	3.12519201	48,828	15624
9	35,154	1.024E-05	0.35997788	2.77794846	48,828	17577
10	39,060	1.024E-05	0.39997542	2.50015361	48,828	19530

Table B.11: *MFPT Envelope spectrum* resolution based on shaft revolutions r and $Fs = 48,828$ samples/second

r	N	$dt = 1/Fs$	$T = Ndt$	$df = 1/T$	$Fq = Fs/2$	$nf = N/2$
1	1,953	2.04801E-05	0.03999754	25.0015361	24,414	976.5
2	3,906	2.04801E-05	0.07999508	12.500768	24,414	1953
3	5,859	2.04801E-05	0.11999263	8.33384537	24,414	2929.5
4	7,812	2.04801E-05	0.15999017	6.25038402	24,414	3906
5	9,765	2.04801E-05	0.19998771	5.00030722	24,414	4882.5
6	11,718	2.04801E-05	0.23998525	4.16692268	24,414	5859
7	13,671	2.04801E-05	0.2799828	3.57164801	24,414	6835.5
8	15,624	2.04801E-05	0.31998034	3.12519201	24,414	7812
9	17,577	2.04801E-05	0.35997788	2.77794846	24,414	8788.5
10	19,530	2.04801E-05	0.39997542	2.50015361	24,414	9765

Table B.12: *MFPT* expected fault frequencies (Hz) using $f_r = 25$ Hz

FTF/CAGE	*BSF/RSF/RE*	*BPFO/OR*	*BPFI/IR*
10.15	63.875	81.125	118.875

In agreement with the analysis of *CWRU Envelope spectrum* resolution, there is a consideration for the selection of the resolution and is based on the importance of the analysis of the first harmonic of the fault frequency. i.e. *BPFO* frequency (Table B.12 is 63.875 Hz and its first harmonic is 127.75 Hz, and the 5^{th} harmonic of the shaft speed is 25 Hz x 5 = 125 Hz with a difference of 2.75 Hz that could lead to one peak in the spectrum to represent both one harmonic of the shaft speed and the harmonic of the *BPFO* if the resolution is not adequate, to avoid that problem the resolution should be at least the half (1.375 Hz). Based on formulas and data of the tables B.10 and B.11, we should select a minimum of 19 revolutions ($df = 1.31587$ Hz) for the analysis. On the other hand, if the identification of the first harmonic is not relevant for the analysis a window size of *5 revolutions* is the optimum for the analysis.

It can be observed in figure B.3 that fault can not be detected and the fault component can be misinterpreted when the window size for the analysis is not chosen properly (1 revolution in this case), and in figure B.4 the *OR* fault frequency can be clearly identified in 81.125 Hz as table B.12 indicate because the proper window selection.

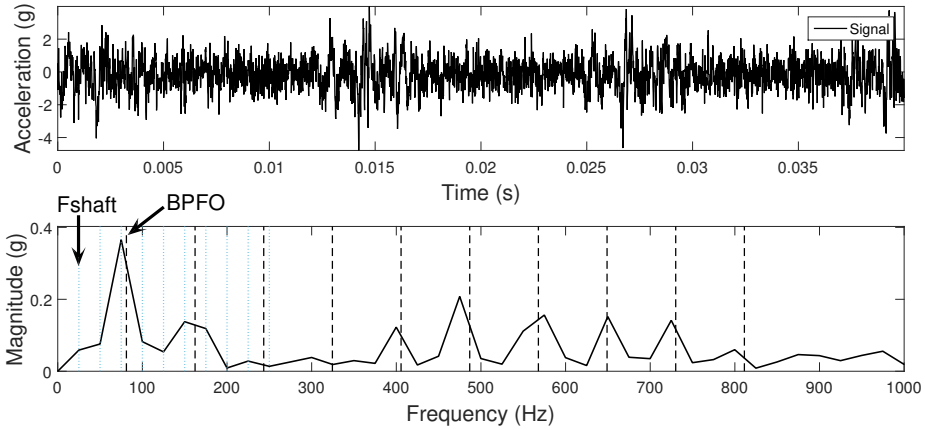


Figure B.3: The upper plot shows 1 revolution (0.04 s) of the *OR* fault signal and the inferior plot shows the corresponding *Envelope spectrum* from *MFPT* data

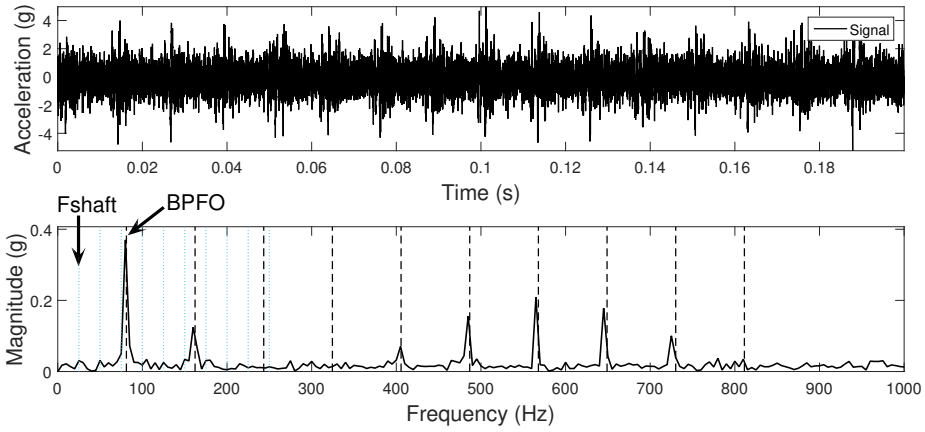


Figure B.4: The plot on the top shows 5 revolutions (0.2 s) of the *OR* fault signal and the bottom plot shows the corresponding *Envelope spectrum* from *MFPT* data

B.3 Window Size Selection Conclusions

The *EMD* computational time is important for online applications, nevertheless, the window selection should not be made based on computational times because these are not constant and there is not direct relation between window size (revolutions) and the expected time. i.e. if a window size is doubled the computational time is not necessarily doubled, it might take less of the expected (double) computational time or the triple or more. Despite there is no relation between number of revolutions and computational time, there is a trend of increasing of the computational time as the number of data points increase.

As consequence, the priority for window size selection should be based in the spectrum resolution required for the analysis, and as we observed in the *Envelope spectrum* resolution analysis, the number of revolutions of the window size is the key to guarantee the desired resolution. Another observation was that a proper sampling frequency allows to optimize the computational time.

The resolution in the spectrum can be computed by:

$$df = \frac{Fs}{Nr} = \frac{Fs}{\frac{Fs}{f_r}r} = \frac{Fsf_r}{Fsr} = \frac{f_r}{r} \tag{B.1}$$

where Fs is the sampling frequency, N are the data points in 1 revolution, r is the number of revolutions, and f_r is the shaft frequency. Equation B.1 allows to obtain the desired resolution, yet two conclusions can be made for the methodology:

1. Equation B.1 suggest that a reduction on the shaft frequency used in the diagnosis allows achieve good resolution using a smaller window size. To acquire reliable information the shaft frequency should not be excessively reduced, it is recommended to use normal working speeds for the analysis.

2. Sampling frequency Fs has an impact in the maximum frequency of the spectrum ($Fq = Fs/2$) and the number of data points that will be processed. The larger the number of data points the larger computation time. It can be concluded that Fs should be as small as possible and it must be large enough to include the information of the resonance frequency of the system.

Bearing frequency factors are often not integer multiples of the shaft speed and do not cause problems to misinterpret the fundamental fault frequencies with shaft harmonics. For *CWRU* and *MFPT* a window size of *5 revolutions* was an optimum selection for the analysis and therefore recommended. If the application requires to analyze the first harmonic of fault frequencies, the window must be adapted to the required resolution using equation B.1.

Appendix C

Statistical Parameters Selection

C.1 Statistical Parameters Selection

In Chapter 2 was shown that the *IMF* selection for fault diagnosis has been made in different ways. In this section, the selection of the *IMF* based on several statistical parameters is proposed.

One of the objectives of the proposed methodology is to automate the selection of the *Intrinsic Mode Function (IMF)* with bearing signature information. A premise is that the automation can be accomplished with the use of statistical parameters that can reveal fault information. In this section 15 statistical parameters are compared: *RMS, Kurtosis, KR, Shannon Entropy, peak value, peak to peak value, Crest Factor, Skewness, Mean, Variance, Standard Deviation, Root, Clearance Factor, Shape Factor*, and *Impulse Factor*. They can be computed using the next equations:

- **Root-Mean-Square (RMS)** is the most common vibration indicator used as a health indicator of the machines. *RMS* is a measurement of amplitude, it represents the intensity of the signal:

$$X(t)_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} y_n^2} \tag{C.1}$$

  where N is the number of data points of the signal, y_n is the amplitude of the n^{th} data point, and $X(t)_{RMS}$ is the intensity of the signal. *RMS* can be represented in units of acceleration, velocity or displacement, it depends of the units of the signal being analyzed. *RMS* values are associated with the condition of the machine. The standard recommends initiate the condition monitoring with the measurement of the machine working in good condition and then periodic measurements must be recorded and saved to create historic files. The vibration signals can be analyzed in a visual manner, identifying increasing trends that can be sudden or progressive. For each vibration signal, the *RMS* value must be computed and used as parameter for condition monitoring in agreement with the *ISO* standards, which include the description of the condition of the machine when taking measurements, sensors, machine condition zones, and so on.

- **Peak-value (pk) and peak to peak-value (pk-pk).** Also called *0 to peak-value*, is an alternative to quantify the amplitude of signals, *peak-value* represents the highest amplitude value in the whole signal and is calculated by finding the maximum. Another similar parameter is *peak to peak value (pk-pk)* that represents the total amplitude of the signal and is calculated measuring the distance from the minimum value to the maximum value.

  Compared to *RMS*, *pk* and *pk-pk value* allows to identify the maximum forces that are occurring in the machine. It is possible that the machine is exposed to random large amplitude vibrations; but, they are not being identified

by the *RMS* value because these vibrations are high force but short length and they are minimized. *Pk* and *pk-pk values* are perfect to identify collisions of the components within the machine.

- **Crest Factor (CF) or peak-to-RMS-ratio** is a non-dimensional parameter for any signal and it defines the ratio of the peak value of the signal to the *RMS* value.

$$CF = \frac{Y_{peak}}{Y_{RMS}} \tag{C.2}$$

where Y_{peak} represents the *peak value* of the signal and Y_{RMS} represents the *RMS* value of the signal. Crest factor is used to analyze the maximum vibration amplitude that the machine is being subject and compare it with *RMS*, since peaks values tend to be minimized in *RMS*.

- **Kurtosis** is called 4^{th} moment and it defines the deviation of a signal's amplitude distribution from Gaussian, this is also a non-dimensional value and can be calculated by using the equation:

$$Kurtosis = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{y_n - \mu}{\sigma}\right)^4 \tag{C.3}$$

where N is number of data points, y_n is the amplitude of the n^{th} data point, μ is the mean of the signal, and σ is the standard deviation of the signal. *Kurtosis* can be used to identify small impacts occurring in the machine, a value of *Kurtosis* = 3 represents a normal distribution, for a machine this would indicate a machine in good condition, when this value tend to increase means that something could be occurring.

- ***Skewness*** represents the symmetry of the data from normal distribution, *Skewness* con be positive or negative, if *Skewness* is negative the mass of the distribution is concentrated on the right of the mean, and if *Skewness* is positive the mass of the distribution is concentrated on the left of the mean. *Skewness* can be computed as:

$$X(t)_{Skewness} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} y_n^2} \tag{C.4}$$

where $X(t)_{Skewness}$ is the *Skewness* of the signal $X(t)$, N is the data length, and y_n is the amplitude of the n^{th} value of the signal.

- ***KR* (*Kurtosis x RMS*)** was proposed by Peter Tse [Tse and Leung, 2010]. This parameter was used in *Wavelet transform* to identify frequency bands with fault information, it combines *Kurtosis* as a measurement of impulsiveness and *RMS* as a measurement of amplitude. *KR* can be computed with the equation:

$$X(t)_{KR} = X(t)_{Kurtosis}\, X(t)_{RMS} \tag{C.5}$$

- ***Shannon Entropy or Entropy*** is a measurement of gain of information and a measurement of disorder and uncertainty. It was proposed by C. E. Shannon in 1948 [Shannon, 1948]. *Shannon Entropy* can be computed using the equation:

$$X(t)_{Shannon} = -\sum_{n=1}^{N} P_n \log(P_n) \tag{C.6}$$

where $X(t)_{Shannon}$ is the *Shannon Entropy* of the signal $X(t)$, and P_n is the probability of an event to occur. *Shannon Entropy* is based on the assumption that an event can take k values, and the probability for one event to occur is the same as any other ($P = 1/k$). E.g. Suppose a constant signal of 1, it means that amplitude only can take 1value ($k = 1$), if equation C.6 is applied the *Entropy* is 0 which means that the signal is simple. If the complexity of the signal increases, the *Entropy* does it too.

- ***Mean***

$$X(t)_{Mean} = \mu = \frac{1}{N}\sum_{n=1}^{N} y_n \tag{C.7}$$

where μ represents the mean value of the signal $X(t)$.

- ***Standard Deviation*** *(σ)* quantifies the amount of variation or dispersion of a set of data values.

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(y_n - \mu)^2} \tag{C.8}$$

- ***Variance*** *(σ^2)* is another measurement of dispersion of the data from its *Mean* value.

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N}(y_n - \mu)^2 \tag{C.9}$$

- ***Root*** is an alternative measurement of amplitude and energy of the signal.

$$X(t)_{Root} = \left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{|y_n|}\right)^2 \tag{C.10}$$

- ***Clearance Factor*** is a measurement of data distribution similar to *CF*.

$$X(t)_{Clearance} = \frac{X(t)_{peak}}{X(t)_{Root}} \tag{C.11}$$

- ***Shape Factor*** is also a measurement of data distribution similar to *CF*.

$$X(t)_{Shape} = \frac{X(t)_{RMS}}{\frac{1}{N}\sum_{n=1}^{N}|y_n|} \tag{C.12}$$

- ***Impulse Factor*** is a measurement of data distribution similar to *CF*.

$$X(t)_{Impulse} = \frac{X(t)_{peak}}{\frac{1}{N}\sum_{n=1}^{N}|y_n|} \tag{C.13}$$

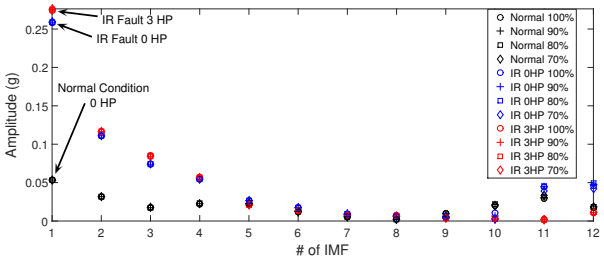
Clearance, Shape, and *Impulse Factors* are good indicators for spikiness of the sharp impulses generated in faulty bearings [Lei, 2016].

To determine which indicator can be used to detect the *IMF* with fault condition, all of the statistical parameters were applied to signals from the *CWRU* and *MFPT* data bases. The analysis consists in the computation of the 15 statistical parameters of the 12 first *IMFs* of the decomposition of one *Normal* condition signal (5 revolutions) and their comparison against the statistical values produced in the decomposition of two faulty signals (5 revolutions). A reported problem for the *Empirical Mode Decomposition* method is the *end effect phenomena* that introduces a slight error in the extreme of the *Intrinsic Mode Functions* that increases as the number of *IMFs* increase [Huang *et al.*, 1998], for his reason the statistical parameters were computed to 100%, 90%, 80%, and 70% of the *IMFs*, e.g. for 90% of the *IMF* 5% of each extreme was cut to avoid the end effect error.

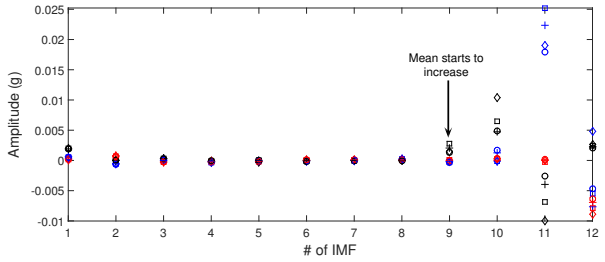
To determine if a parameter is useful, a change in their value is expected for faulty signals compared to *Normal* condition signals, and the result of these parameters must not be susceptible to changes produced by end effect errors (value produced by 100% of *IMF* must be equal or similar as the produced by 90% of the *IMF*).

The result of the first test for the *CWRU* data can be observed in Figures C.1 and C.2 [1]. After *Envelope Spectrum* analysis of the *IMFs*, it was determined that the fault information is found in the first *IMF*, as consequence the largest change in the statistical parameters must be in the first *IMF*. The statistical parameters that do not present difference in the first *IMF* are: *Mean, Shape Factor*, and *Skewness*, thus these parameters are discarded (Fig. C.1b, C.2c C.2h). The remaining parameters that present changes in the first *IMF* but major changes in the rest of the *IMFs* are: *Clearance Factor, Impulse Factor*, and *Crest Factor* and hence are discarded (Fig. C.2b, C.2d C.2g). Another discarded parameter is *Shannon Entropy* because despite a change between normal and fault condition, it is affected by the end effects and the distinction between faulty signals with different loads is not possible (Fig. C.1d). The *Peak* and *Peak to Peak* values allow to detect spikiness in the signal but does not allow to differentiate between a fault with different loads and are discarded. The remaining useful parameters are *RMS* and *Root* as amplitude measurements, *Kurtosis, Standard Deviation*, and *Variance* as measurements of signal distribution, and *KR* as combination of amplitude and spikiness measurement. The objective is determine which combination of parameters is the best to identify the *IMF* with fault information.

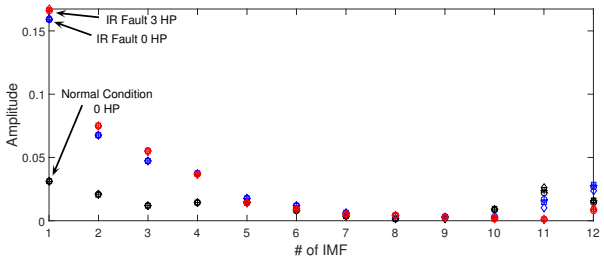
[1] All figures can be found in http://bit.ly/ThesisRCG



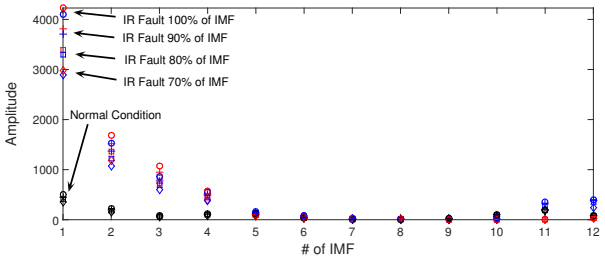
(a) *RMS*, the amplitude of the first four *IMF* increase in the presence of fault



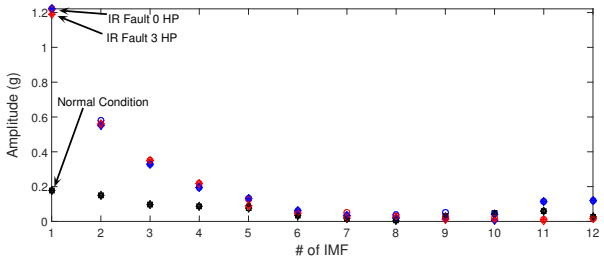
(b) *Mean*, there is no significant difference between the values of normal condition and fault condition



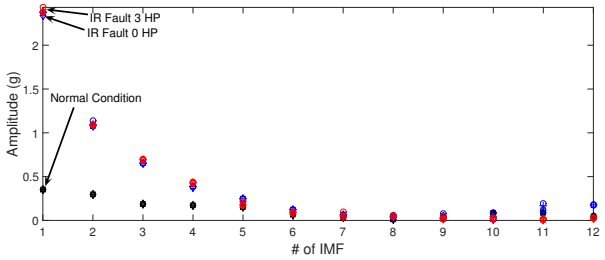
(c) *Root*, the amplitude of the first four *IMF* increase in the presence of fault



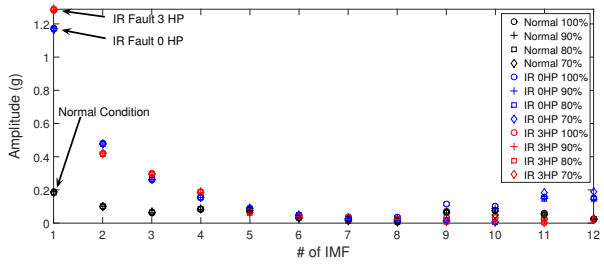
(d) *Shannon Entropy*, the amplitude varies with the length of the signal



(e) *Peak*, the amplitude of the first four *IMF* increase in the presence of fault

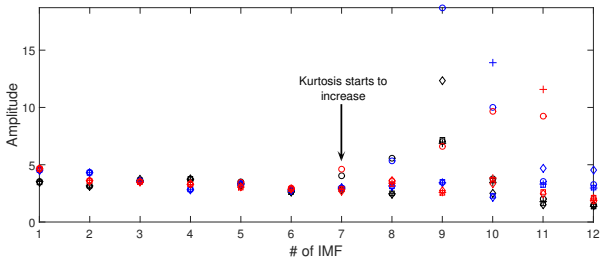


(f) *Peak to peak*, the amplitude of the first four *IMF* increase in the presence of fault

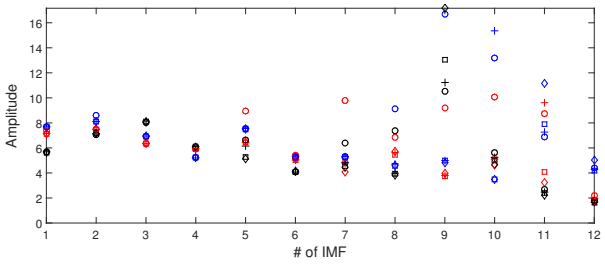


(g) *KR*, the amplitude of the first four *IMF* increase in the presence of fault

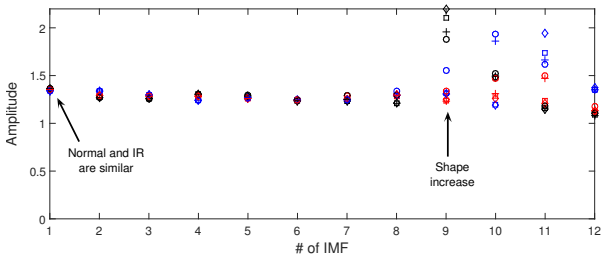
Figure C.1: Statistical parameters used to measure amplitude in signals, test of *CWRU* data



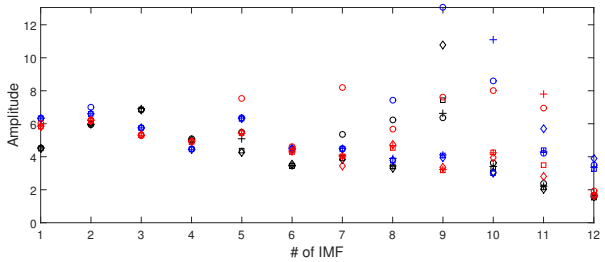
(a) *Kurtosis*, after 6th IMF the amplitude increases because of the end defect



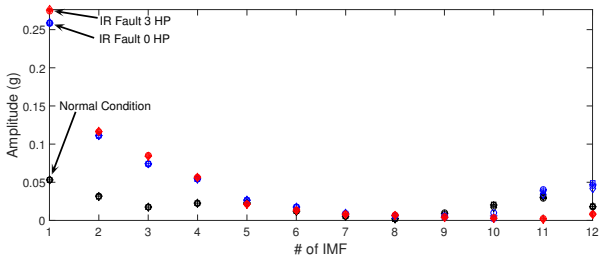
(b) *Clearance Factor*, there is no relation between a faulty and a normal condition signal



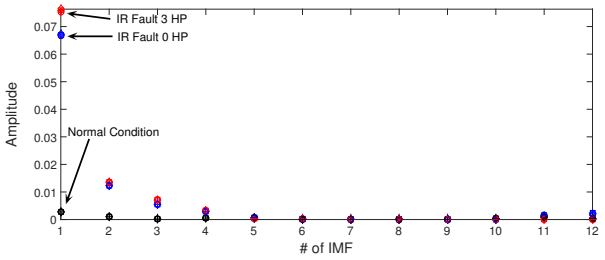
(c) *Shape Factor*, the amplitude does not increase in the first *IMF* as expected



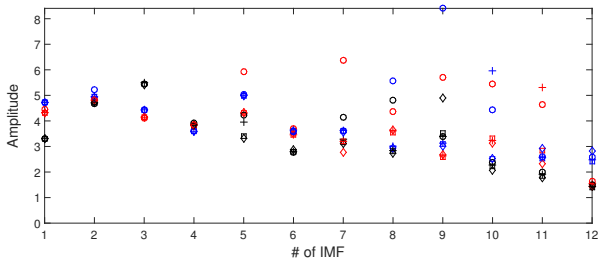
(d) *Impulse Factor*, there is a lot of variation in amplitude values



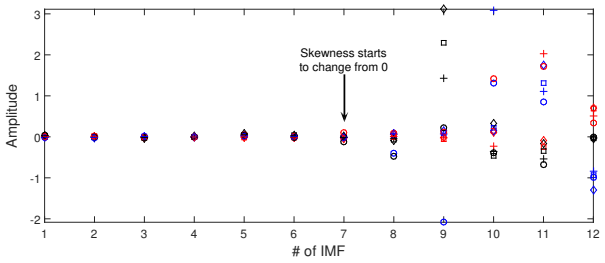
(e) *Standard Deviation*, the amplitude of the first four *IMF* increase in the presence of fault



(f) *Variance*, the amplitude of the first three *IMF* increase in the presence of fault



(g) *Crest Factor*, there is a lot of variation in amplitude values



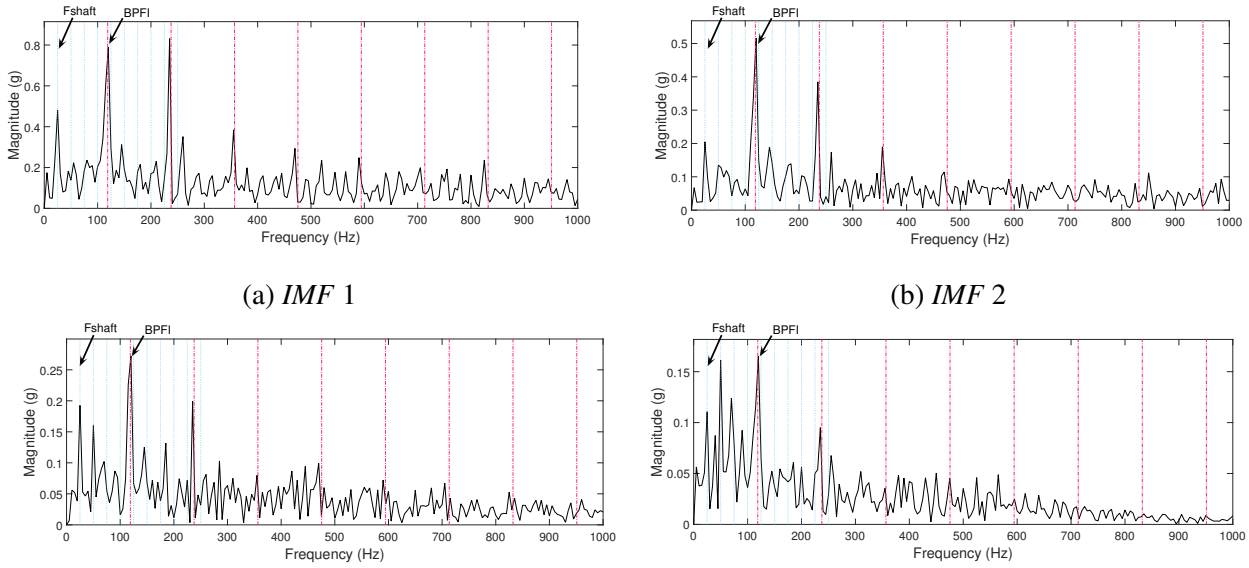
(h) *Skewness*, it produces values close to 0 for the first *IMFs*

Figure C.2: Statistical parameters used to detect spikiness produced by bearing faults, test of *CWRU* data

The *CWRU* data set is composed of low noise signals, to confirm if the 6 remaining parameters are reliable for the method, the same test is applied using more challenging signals from *MFPT* data base. For this test 5 revolutions from 1 *Normal* condition signal with load of 270 lbs, one *IR* fault condition signal with no load (0 lbs), and one *IR*

fault condition signal with load of 100 lbs were analyzed.

First, a visual inspection of the *Envelope spectrum* of the *IMFs* of the two faulty signals was made to detect which *IMFs* contains fault information. For the *IR* fault signal with 0 lbs of load, fault information can be found in the first 4 *IMFs* (see Fig. C.3). It can be observed that the first 2 *IMFs* present the largest amplitude of *BPFI* (Fig. C.3a and C.3b), as result the largest variation is expected in these *IMFs*.



(a) *IMF* 1 (b) *IMF* 2

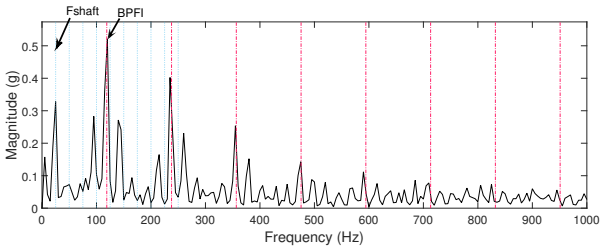
(c) *IMF* 3 (d) *IMF* 4

Figure C.3: *Envelope spectrum* of the first 4 *IMFs* of *IR* fault signal with load of 0 lbs from *MFPT* data, the 1st and 2nd *IMFs* have the most significant fault information

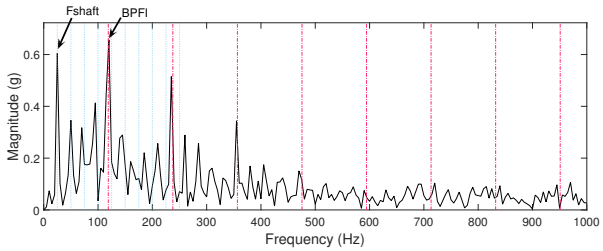
On the other hand, the analysis of the *IR* fault signal with 100 lbs of load presents *BPFI* information in the first 5 *IMFs* (see Fig. C.4), where the second *IMF* presents the largest amplitude, and therefore the largest change in the statistical parameters must be in the second *IMF*.

After the *Envelope spectrum* analysis, the computation and analysis of the 6 remaining statistical parameters was made. It is important to remind that two parameters are focused in amplitude measurement (*RMS* and *Root*), three are for distribution measurement (*Kurtosis, SD*, and *Variance*), and the last parameter combines both measurements (*Kurtosis* and *RMS*). First, in the analysis of amplitude *Root* fails to detect the *IMF* with *BPFI* information because it only presents changes from the *Normal* condition to fault condition in the first *IMF* (Fig. C.5b) , thus, *Root* is discarded for the method. On the contrary, *RMS* does represent the *Envelope spectrum* results. There are changes (in blue) in the first two *IMFs* for the *IR* with no load, and changes in the first 5 *IMFs* being the second the largest (Fig. C.5a). Another observation for the *RMS* values is that after the $8th$ *IMF* they start to increase and could lead to a wrong detection. As summary, *RMS* is the best parameter for amplitude measurement.

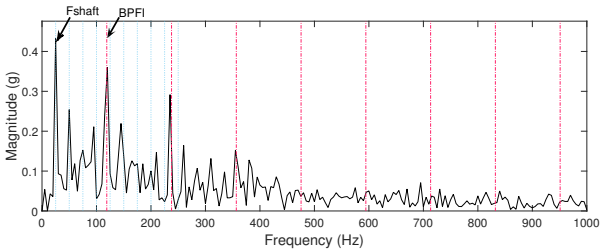
In the analysis of distribution measurement parameters or spikiness, *Kurtosis* increases in the first 4 *IMFs* for *IR* fault with no load, this in agreement with *Envelope spectrum* analysis of figure C.3. But there is variation in their values in *IMFs* 8 to 11 (Fig. C.5c). For *SD* is observed that the largest changes in values are in first 5 *IMFs* as in figure C.4. The largest changes in blue correspond to the first two *IMFs* for the signal with 0 lbs of load, and the changes in red for the first 5 *IMFs* of the *IR* fault signal with 100 lbs of load, being the second the largest as in the *Envelope spectrum* analysis. Finally, an increase in values of the last *IMFs* is found. The *Variance* describes the same pattern



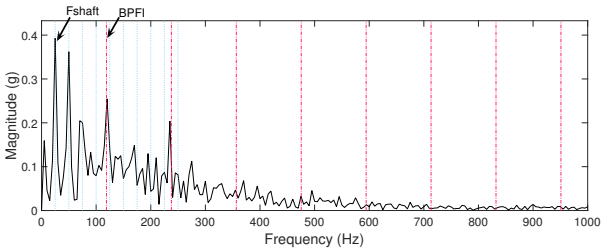
(a) *IMF* 1, the *BPFI* is the largest peak in the spectrum



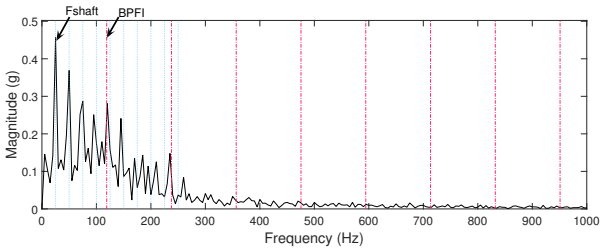
(b) *IMF* 2, the *BPFI* is the largest peak but the shaft speed component is similar



(c) *IMF* 3, the shaft speed component is the largest in the spectrum



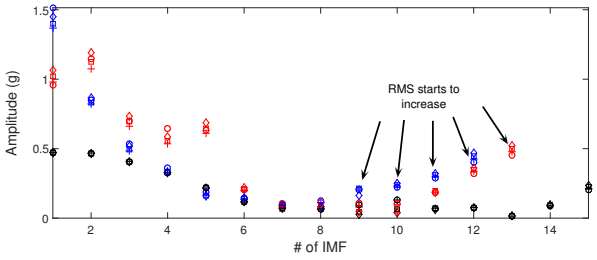
(d) *IMF* 4, the *BPFI* is half of the shaft speed component



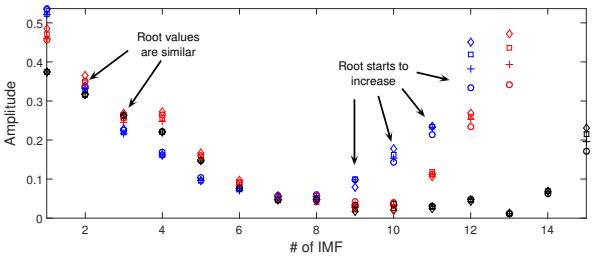
(e) *IMF* 5, the shaft speed component is the largest in the spectrum and overwhelms the *BPFI*

Figure C.4: *Envelope spectrum* of the first 5 *IMFs* of *IR* fault signal with load of 100 lbs from *MFPT* data

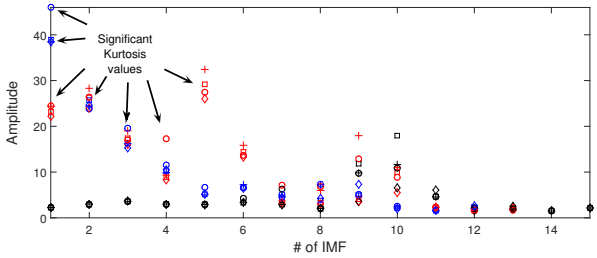
of change as in *SD* but the changes are larger, and the variation of the last *IMFs* is smaller than *Kurtosis* and *SD*. This results suggest that *Variance* represents better faulty signals, therefore a new parameter based on *Variance* and *RMS* could have better performance than the parameter *KR* (Fig. C.5f). To prove this hypothesis, two new parameters were included in the analysis: *RMSxSD* (Fig. C.5g) and *RMSxVariance* (Fig. C.5h). The comparison of the results of the *KR, RMSxSD*, and *RMSxVariance* demonstrate that in fact *RMSxVariance* performs better than *KR* and *RMSxSD* because the detection for *IR* fault with no load can be clearly detected in the first *IMF* (Fig. C.3a and for *IR* fault with 100 lbs of load can be clearly detected in the second *IMF* (Fig. C.3a, and these are the largest changes found in figure C.5h (*IMF* 1 for blue and *IMF* 2 for red respectively).



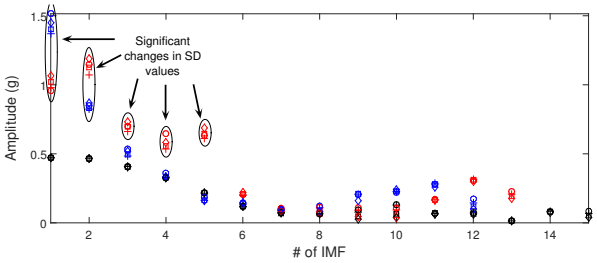
(a) *RMS*, there is an increase in the first five *IMFs* with fault information



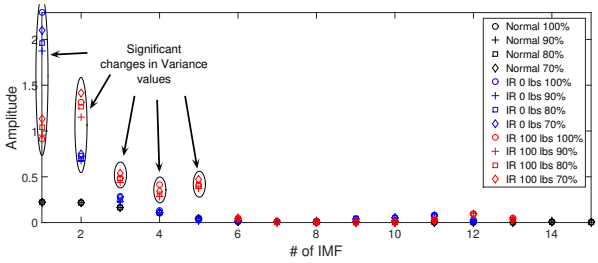
(b) *Root*, there is not an increase in the first five *IMFs* with fault information



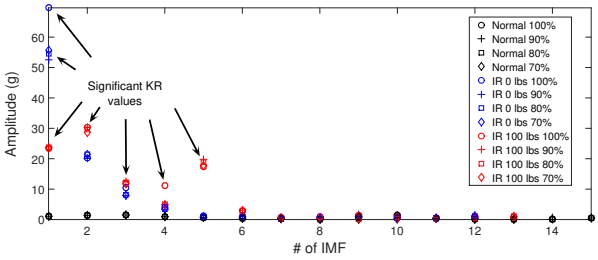
(c) *Kurtosis*, there is an increase in the first five *IMFs* with fault information



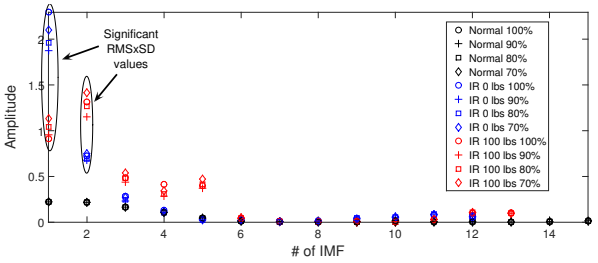
(d) *Standard Deviation*, there is an increase in the first five *IMFs* with fault information



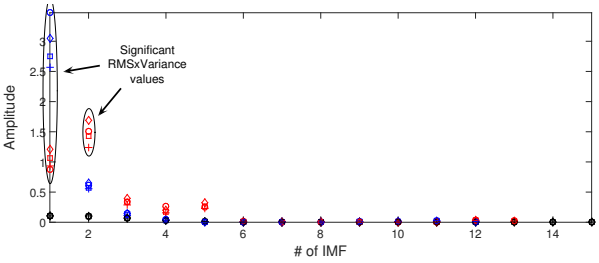
(e) *Variance*, there is an increase in the first five *IMFs* with fault information



(f) *KR*, there is an increase in the first five *IMFs* with fault information



(g) *RMS x Standard Deviation*, there is an increase in the first five *IMFs* with fault information

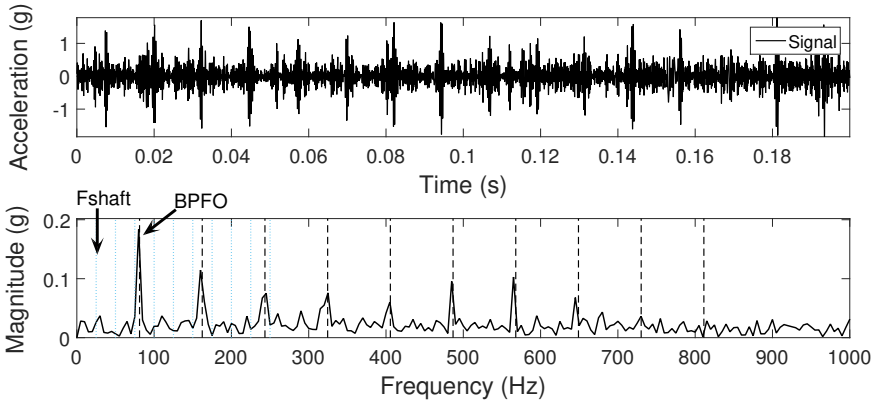


(h) *RMS x Variance*,there is an increase in the first five *IMFs* with fault information

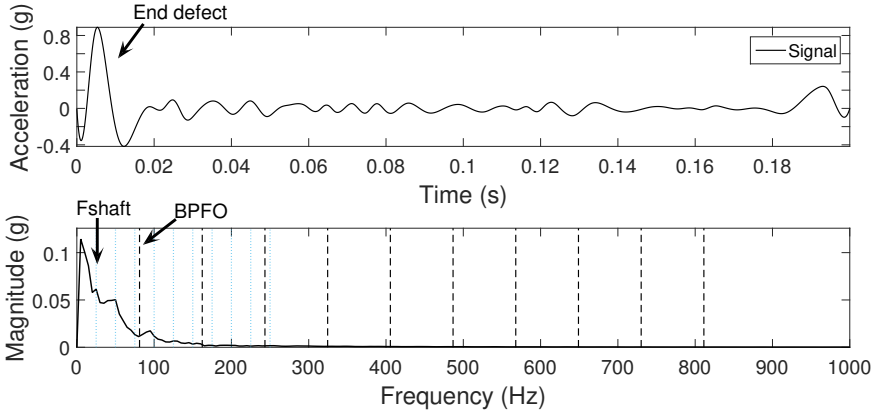
Figure C.5: Statistical parameters used in *IR* fault test of *MFPT* data, the most significant components are 1st and 2nd *IMFs*

An extra test was performed to *KR* and *RMSxVariance* to assure that *RMSxVariance* is the best parameter for the method. This time an *OR* fault signal with 270 lbs of load was chosen because is more challenging than *IR* fault signals. The *Envelope spectrum* analysis of the *IMFs* showed that the fault information is found in *IMF* 4, but the use

of both indicators failed in the correct detection, while *RMSxVariance* suggested that the correct *IMF* was the first, the *KR* chose the *IMF* 9. In figure C.6b the correct and the 9^{th} *IMF* can be observed,it is clear that the incorrect *IMF* has an end defect that might influenced the selection with *Kurtosis*.



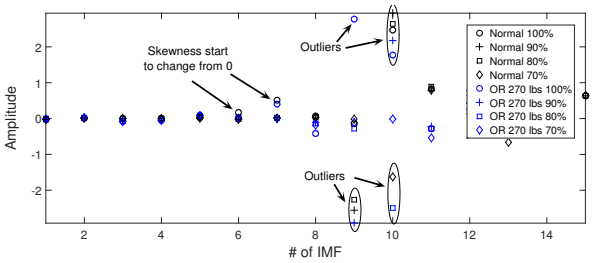
(a) *IMF* 4, the signal has no end defect



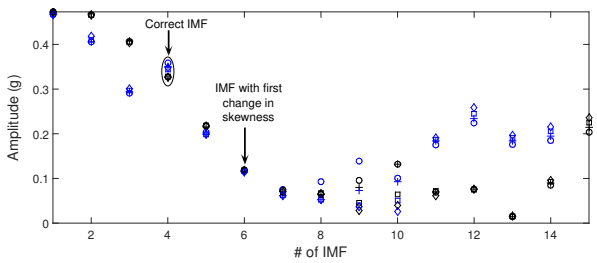
(b) *IMF* 9, the end defect can be observed

Figure C.6: *Envelope spectrum* of the 4^{th} (correct) and 9^{th} (selected) *IMFs* of the *OR* fault test from *MFPT* data

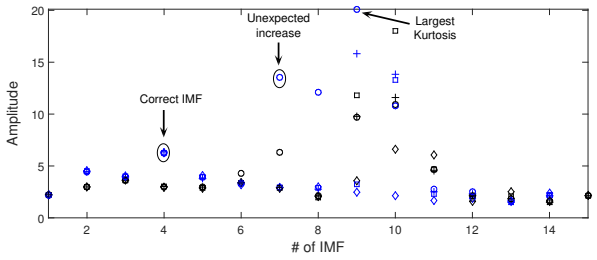
The resonance frequency of the system is found in the first *IMFs*, as a result the change in the parameter values is expected in the first *IMFs*. It was observed that at some point the end effect of the *EMD* plays an important role in the use of the statistical parameters to detect the *IMF* with fault information. This problem can be overcome by the stop of the decomposition when this error (End defect) becomes significant. Another advantage of a limit for the *EMD* is that computational resources can be saved, and the computational time can be enhanced. It was observed that *Skewness* can be used as parameter to select the grade of decomposition in the *EMD*, in the *CWRU* test (Fig. C.2h) the *Skewness* started to change from 0 in the 7^{th} *IMF*, the same *IMF* where *Kurtosis* starts to increase in unexplained manner (Fig. C.2a). In figure C.7a can be observed that the *Skewness* values start to change from 0 in the 6^{th} and 7^{th} *IMFs*, if we analyze the *Kurtosis* in these *IMFs* we observe that in the 7^{th} the *Kurtosis* has a sudden change, and in the next *IMFs* these values have large variation. For this case, if the *Skewness* criteria is applied the selection and the only the first 6 *IMFs* are used, the *KR* could successfully identify the correct *IMF* (Fig. C.7e). On the other hand, *RMSxVariance* would still choose the first (incorrect) *IMF* (Fig. C.7f).



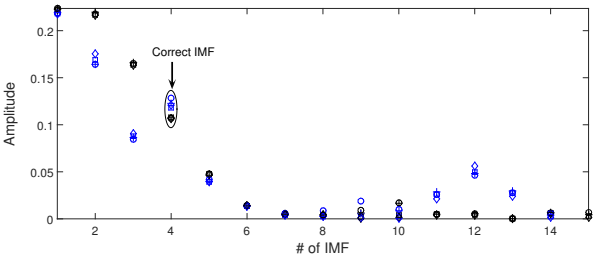
(a) *Skewness*, the values start to change from 0 in the 6th *IMF*



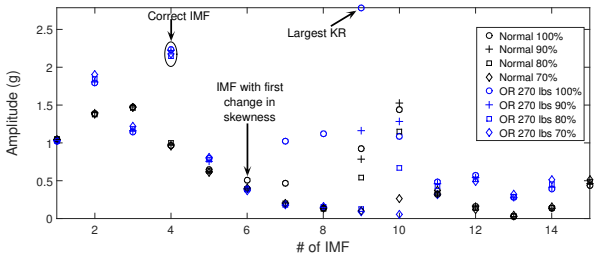
(b) *RMS*, in the 4th *IMF* the amplitude is larger than normal condition as expected



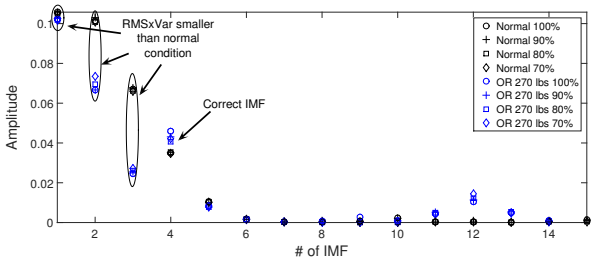
(c) *Kurtosis*, the largest amplitude in the first six *IMFs* correspond to the 4th as expected



(d) *Variance*, in the 4th *IMF* the amplitude is larger than normal condition as expected



(e) *KR*, the largest amplitude in the first six *IMFs* correspond to the 4th as expected



(f) *RMS x Variance*, the largest amplitude in the first six *IMFs* correspond to the 1st and not the 4th

Figure C.7: Statistical parameters used in *OR* fault test of *MFPT* data, the most significant *IMF* is the 4th component

C.1.1 Statistical Parameters Selection Conclusions

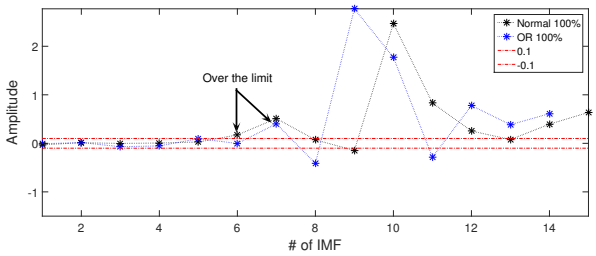
This analysis allowed to demonstrate that *Kurtosis x RMS* or *KR* is the best indicator for faulty signals, it combines measurements of both amplitude and signal distribution that allows to identify the resonance frequencies of the system and spikiness in the signal that reflects fault information. The higher is the *KR* the better the quality of information of the *IMF* to make a prognosis. It was found that the *end effect* phenomena is a relevant factor that can mislead the detection, unfortunately the statistical measurements are sensitive to the end defects and increase the difficulty of their application to the analysis, either way good results can be made by limiting the *Empirical Mode Decomposition* to four *IMFs*. The use of four *IMFs* is made based on the observed on the analysis of several signals, most of cases the *IMFs* with fault information are within the first three and in rare occasions in the fourth *IMFs*. Also, these tests demonstrated that the idea of the analysis of only the first two *IMFs* is wrong because they not necessarily have the fault information.

In next section the initial results for a possible solution to detect *IMFs* with end defects is presented.

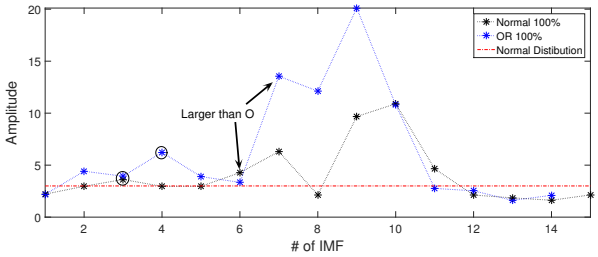
C.2 *Skewness* as *End Defect detector* (Future Work)

The end effect error produced in the *Empirical Mode Decomposition* can affect the diagnosis and the effectiveness of the statistical parameters to identify faulty signals (*IMFs*), in this section an analysis of the use of *Skewness* as a parameter to identify end defects is presented.

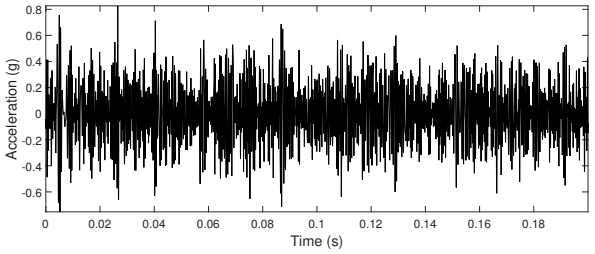
In figure C.8a we can observe the *Skewness* values of the *IMFs* from two signals (*Normal* and *OR* fault condition) and a zone from -0.1 to 0.1 that represents a region where the *IMF* has no end defect. On the right side (Fig. C.8b) the Kurtosis values are plotted and a line on 3 which represents a normal distribution. In *Skewness* plot can be identify the first *IMF* that exceeds the limit or both cases, being the 6^{th} and 7^{th} *IMF* for *Normal* and *OR* fault condition signals respectively. When we study the *Kurtosis* plot, it can be identified that the values in the 6^{th} and 7^{th} are larger than the produced in their respective previous *IMFs* (circled). To verify that this increase is caused by the end defect a plot of the *IMFs* was necessary. In figures C.8c and C.8d the 5^{th} and 6^{th} *IMFs* of *Normal* condition can be observed, in the fifth function no end defects are found in agreement with its *Skewness* value, but in the sixth function with *Skewness* value out of the proposed zone the end defect can be clearly identified. The same comparison can be made for the *OR* fault condition *IMFs* using figures C.8e and C.8f.



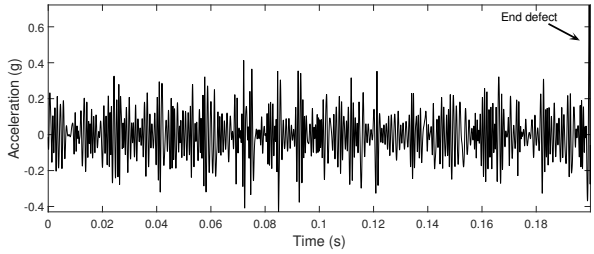
(a) *Skewness*, the amplitude of the 6th and 7th *IMFs* are out of the proposed zone



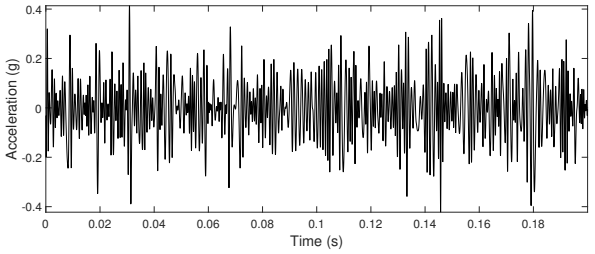
(b) *Kurtosis*, the amplitude for *IMFs* 6th and 7th are larger than in the previous *IMFs*



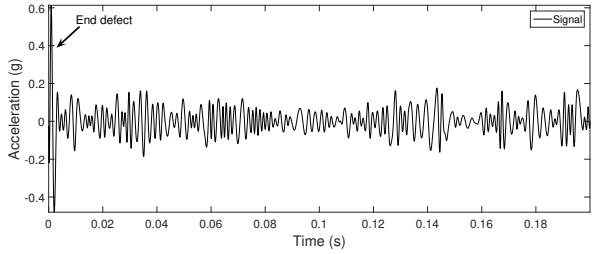
(c) *IMF* 5 *Normal* condition, there is no end defect



(d) *IMF* 6 *Normal* condition with end defect



(e) *IMF* 6 *OR* fault condition, there is no end defect



(f) *IMF* 7 *OR* fault condition with end defect

Figure C.8: *Skewness* as *End Defect* detection

Unfortunately the proposed limits for the *Skewness* does not work to detect the end defect in all signals. Further investigations will be made in the future to solve this problem, the goal is to detect which *IMFs* have end defects to exclude them from the reconstruction of the signal, or to use the *Skewness* as a stopping criteria for the *EMD* to save computational resources and improve the performance of the methodology.

Appendix D

Recommendations

- Tests performed for *Condition Based Monitoring (CBM)* requires that operating conditions during the machine evaluation can be recreated. The first step is to define the speed of the shaft that will be used for tests, it must be a speed within nominal speed and the spindle must reach steady-state operating conditions [ISO, 1995] [ISO, 2014].
- Sensors must satisfy the frequency range requirements of 10 - 10,000 Hz from *ISO 2954* [ISO, 2012]. For sensors with these characteristics a sample frequency of 20,000 samples/second is recommended based on Nyquist frequency; Nowadays there is access to sensors with larger frequency ranges, since methodology is based in demodulation, to guarantee good performance of the methodology it is necessary a sensor that covers the system resonance frequencies.

  To determine the resonance frequency of the system a spectrum plot from an over-sampled signal is recommended and based on that spectrum an optimum sample frequency can be selected.
- Measurements should be taken in the spindle housing (at front end and back end) in a minimum of two radial directions for rotating cutting machines *ISO 17243-1* [ISO, 2014].

Appendix E

Additional Results

E.1 CWRU additional results

For *CWRU* a signal with *OR* defect of *7 mils* and a load of 0 HP is analyzed step by step. The signal is classified as diagnosable (Y), Fig. 4.1 (d).

1. Signal

The methodology suggest a total of five revolutions for the analysis. As the signal was recorded at 1,796 RPM (29.93 Hz), the five revolutions correspond to 0.16705 seconds ($T = r/f_r$), where T is the window size in seconds, r is the number of revolutions, and f_r is the shaft frequency.

2. Signal Decomposition

The signal is decomposed using the *EMD* to separate the fault information from noise and other components. The *EMD* procedure is stopped when the fourth *IMF* is obtained. The corresponding components for the analyzed signal are shown in Fig. E.1. It can be observed that in the first and second *IMFs* there are more impulses than in the third and fourth, and the amplitude of the first *IMF* is greater than the others (around 10 times bigger).

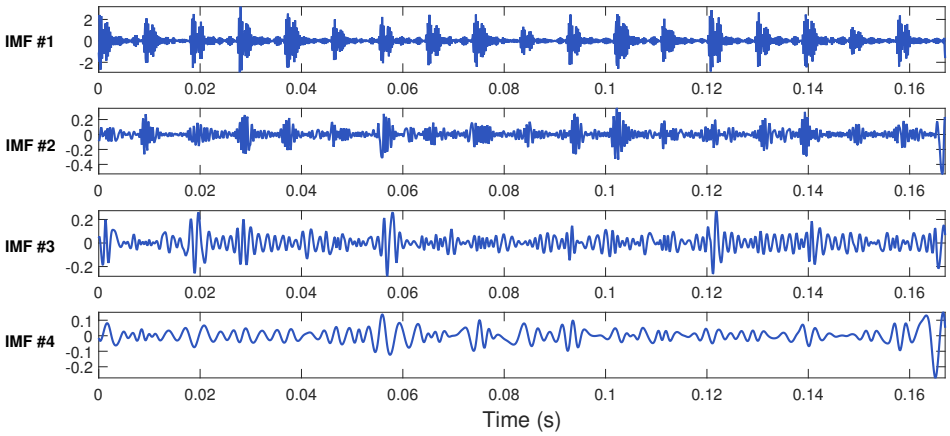


Figure E.1: First four components (*IMFs*) of *OR-Y* (*CWRU* database)

3. *IMF* Selection and Signal Reconstruction

The next step is to compute the *Kurtosis, RMS*, and *KR* for each *IMF*, and based on the *KR* amplitudes, the two *IMFs* with largest values are selected to reconstruct the signal. For the analyzed signal the 1^{st} and 2^{nd} components have the largest *KR* with values of 4.606 and 0.7457 respectively, Fig. E.2.

The reconstructed signal is composed by the two *IMFs* with most useful information for the analysis and with less noise than the original signal, Fig. E.3. The reconstruction is computed using Eq. 2.5. At this point the resonance frequency information is isolated to proceed with the *Envelope spectrum*.

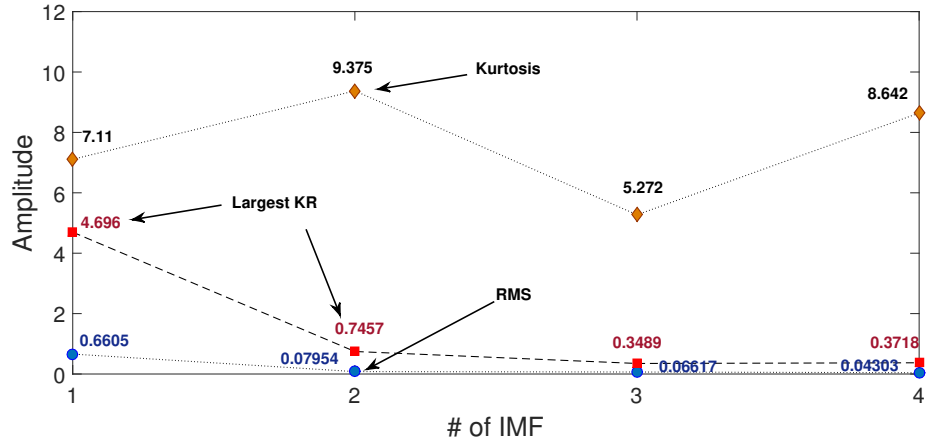


Figure E.2: Statistical values of first four components (*IMFs*) of *OR-Y* (*CWRU* database), the largest values of *KR* correspond to the first two *IMFs*

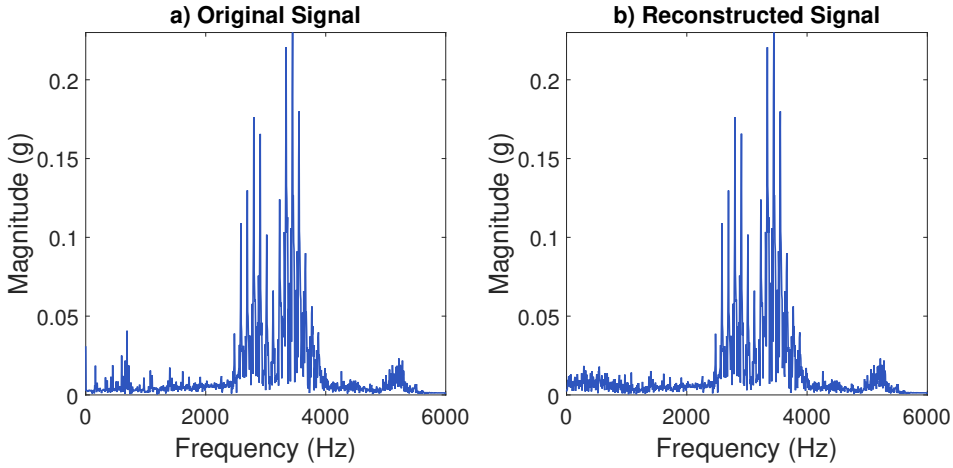


Figure E.3: *FFT* spectrum of the (a) original signal and (b) the reconstructed signal with the *OR-Y* fault (*CWRU* database). The reconstructed signal removed the noise from low frequencies to isolate the resonance frequency for the analysis

4. *Envelope spectrum*

The fault components are modulated in the resonance frequency, to visualize them in the spectrum, the *Envelope spectrum* is applied. First, the envelope is computed using the *HT*. Then, the mean trend or *DCC* is removed with eqn.

(3.4) to avoid the undesired peak in 0 Hz in the spectrum. The reconstructed signal, the envelope signal, and envelope signal with removed *DCC* are shown in Fig. E.4.

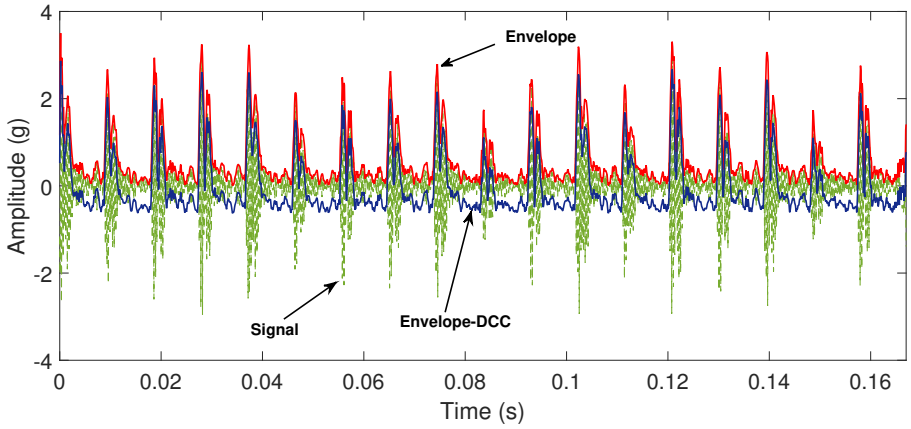


Figure E.4: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*CWRU* database), the *DCC* removal does not affect the waveform and improves the diagnosis

After the envelope of the reconstructed signal is computed and the *DCC* is removed; The *FFT* is applied and the *Envelope spectrum* is plotted to make a diagnosis. To facilitate the diagnosis is recommended to include the fault frequencies and harmonics in the *Envelope spectrum*, because unlike controlled experiments, the expected fault is unpredictable.

Using the shaft speed of 29.93 Hz and the frequency factors, Table 3.2, the fault frequencies are obtained: *BPFO* = 107.3050 Hz, *BPFI* = 162.0950 Hz, *BSF/RSF* = 141.0908 Hz, and *FTF* = 11.9224 Hz.

5. Diagnosis

The analysis of the *Envelope spectrum* for the diagnosis of bearing condition consist in the visual inspection of the spectrum and the identification of peaks in the computed fault frequencies, their harmonics, and in some cases sidebands, Table 3.1. For *OR* faults peaks in the *BPFO* and in the harmonics are expected. In Fig. E.5, peaks in the *BPFO* and three harmonics can be identified for the analyzed signal. As its classification is diagnosable these peaks can be clearly identified. The identification of these peaks confirm the presence of the *OR* fault in the signal.

In Fig. E.6 the comparison of the proposed methodology against the traditional *Envelope spectrum* and the *FFT* are shown. In this case, the analyzed signal is classified as diagnosable, there are no other components that make difficult the diagnosis, and the amplitude of the *Envelope spectrum* and the proposed methodology are similar. The difference in amplitude and the presence of other components between each other are in terms of the effectiveness of the *EMD*, the use of statistical parameters (*KR*) for the selection of *IMFs* with fault information is intended to avoid the visual inspection of the *IMFs*, [Tsao *et al.*, 2010] [Tsao *et al.*, 2012]. The selection of the most significant *IMFs* is guarantee.

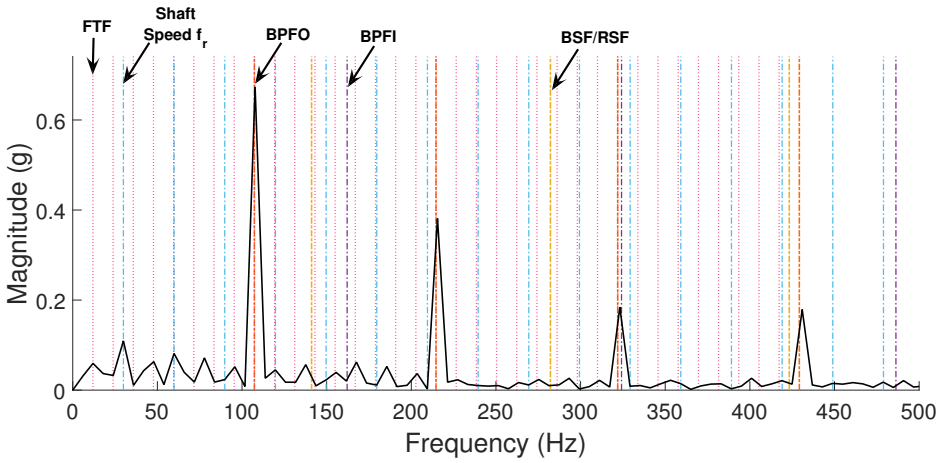


Figure E.5: *Envelope spectrum* of reconstructed signal *OR-Y* (*CWRU* database), the *BPFO* and harmonics are predominant in the spectrum

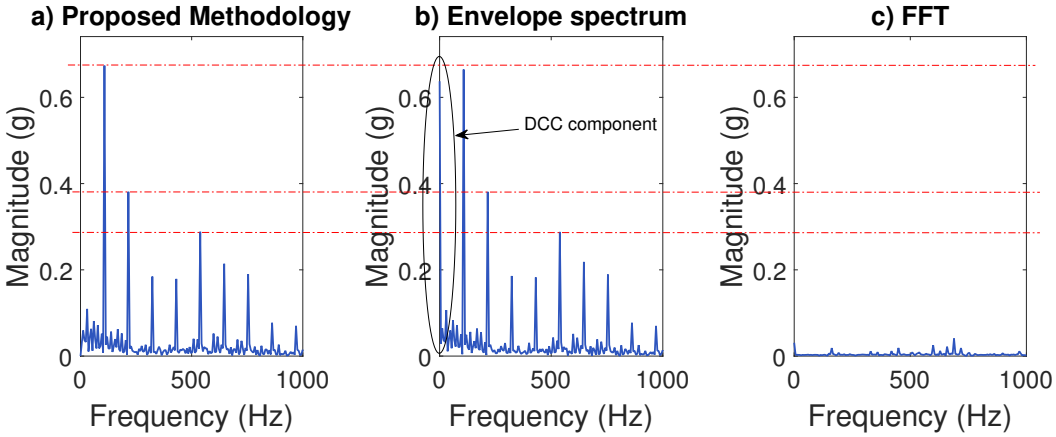


Figure E.6: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *OR-Y* (*CWRU* database), the proposed methodology removes the *DCC*, and the *FFT* does not detect the fault

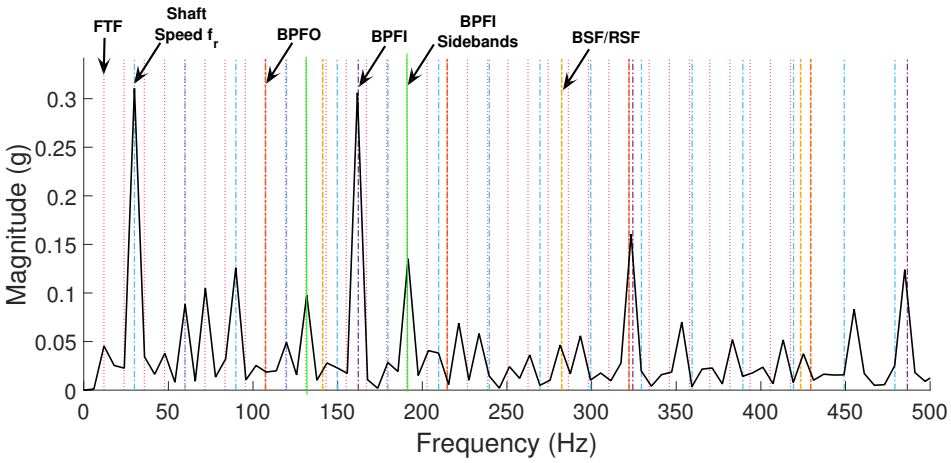


Figure E.7: *Envelope spectrum* of reconstructed signal *IR-Y* (*CWRU* database), the *BPFI*, their harmonics and sidebands can be observed

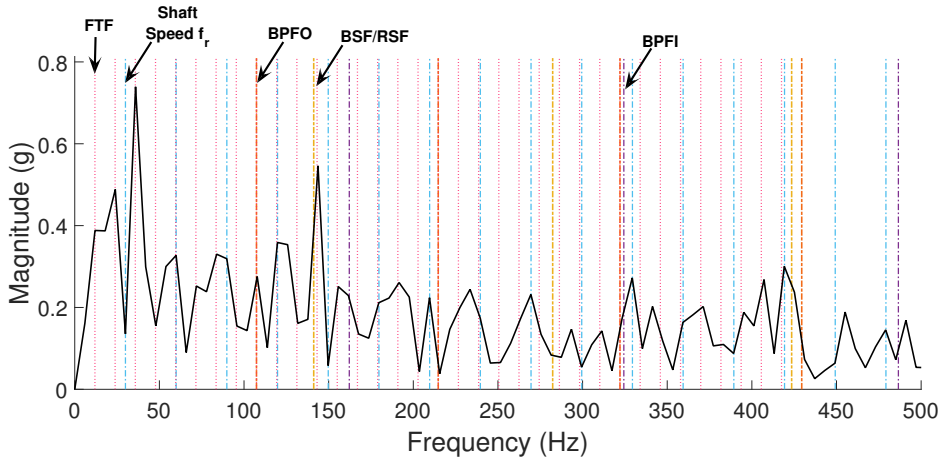


Figure E.8: *Envelope spectrum* of reconstructed signal *RE-Y* (*CWRU* database), the *BSF/RSF* can be clearly identified, and the *FTF* components are the largest in the spectrum

E.2 Tian-Yau Wu database

For *Tian-Yau Wu* database, a signal with *OR* and severe defect is analyzed step by step, Fig. 4.9 (d).

1. Signal

First, the sample of five revolutions is made for the analysis. This signal was recorded at 8.49 Hz, the five revolutions correspond to 0.5889 seconds ($T = r/f_r$), where T is the window size in seconds, r is the number of revolutions, and f_r is the shaft frequency.

2. Signal Decomposition

The *EMD* is applied to divide the signal in fault information and noise components. The *EMD* procedure is stopped when the fourth *IMF* is obtained. The corresponding components for the analyzed signal are shown in Fig. E.9. It can be observed that the first two *IMFs* are greater in amplitude than the others. In the first *IMF* the impulses are clearer.

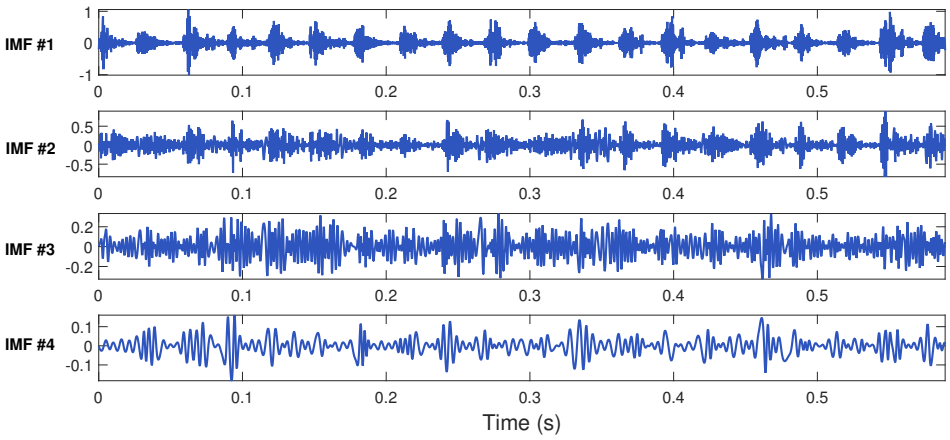


Figure E.9: First four components (*IMFs*) of *OR* severe fault (*Tian-Yau Wu* database)

3. *IMF* Selection and Signal Reconstruction

The next step is to compute the *Kurtosis, RMS*, and *KR* for each *IMF*. Based on the *KR* amplitudes, the 1^{st} and 2^{nd} components have the largest *KR* with values of 1.217 and 0.8175 respectively, Fig. E.10. These two components are used to reconstruct the signal with fault information.

The reconstructed signal is composed by the two *IMFs* with most useful information for the analysis and with less noise than the original signal, Fig. E.11. The reconstruction is computed using eqn. (2.5). At this point the resonance frequency information is isolated to proceed with the *Envelope spectrum*.

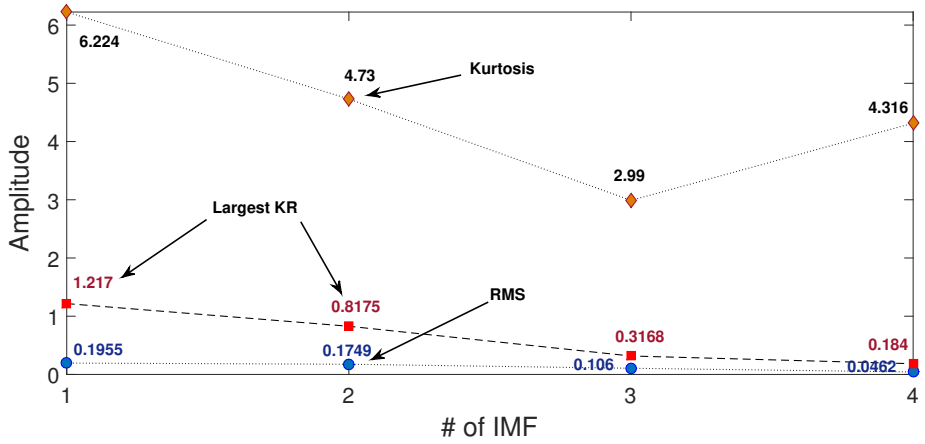


Figure E.10: Statistical values of first four components (*IMFs*) of (*IMFs*) of *OR* severe fault (*Tian-Yau Wu* database), the largest values of *KR* correspond to the first two *IMFs*

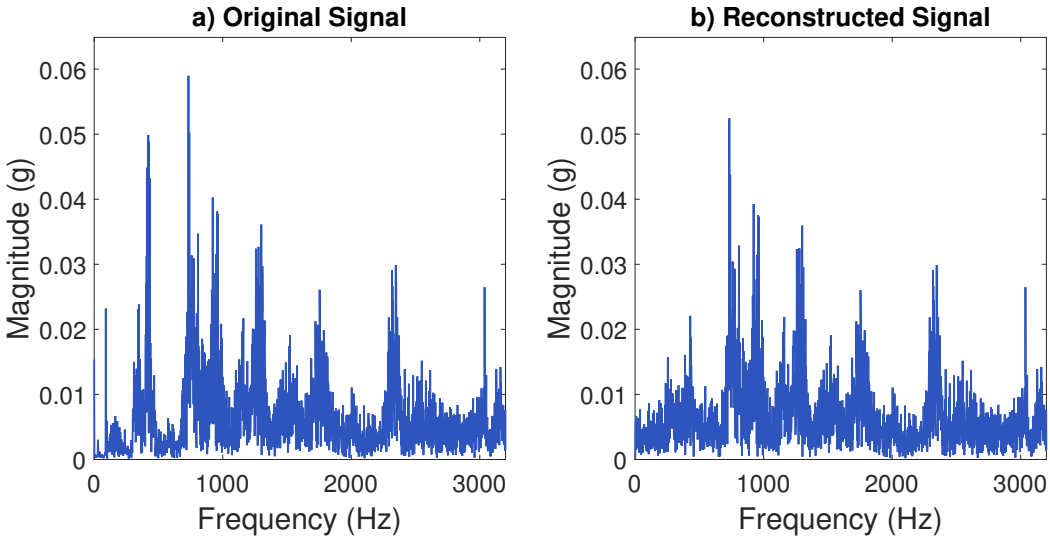


Figure E.11: *FFT* spectrum of the (a) original signal and (b) the reconstructed signal with the *OR* severe fault (*Tian-Yau Wu* database). The reconstructed signal removed the noise from low frequencies

4. *Envelope spectrum*

First, the envelope is computed using the *HT* to demodulate the faults. Then, the mean trend or *DCC* is removed with eqn. (3.4) to avoid the undesired peak in 0 Hz in the spectrum. The reconstructed signal, the envelope signal, and envelope signal with removed *DCC* are shown in Fig. E.12.

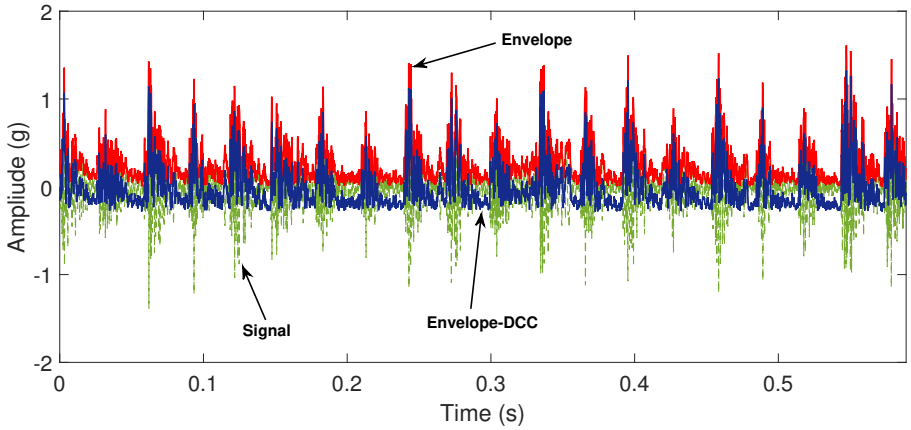


Figure E.12: Envelope of the reconstructed signal using *IMFs* 1 & 2 (*Tian-Yau Wu* database), the *DCC* removal does not affect the waveform and improves the diagnosis

The next step is to apply *FFT* to the envelope signal with the *DCC* removed to make the diagnosis. Using the shaft speed of 8.49 Hz and the frequency factors, Table 3.3, the fault frequencies are obtained: *BPFO* = 32.7714 Hz, *BPFI* = 52.1286 Hz, *BSF/RSF* = 35.4033 Hz, and *FTF* = 3.3111 Hz.

5. Diagnosis

The *Envelope spectrum* is analysis consist in the visual inspection of the spectrum and the identification of peaks in the computed fault frequencies, their harmonics, and in some cases sidebands, Table 3.1. For *OR* faults peaks in the *BPFO* and in the harmonics are expected. In Fig. E.13 peaks in the *BPFO* and two harmonics can be identified for the analyzed signal. These peaks can be clearly identified and confirm the presence of the *OR* fault in the signal.

In Fig. E.14 the comparison of the proposed methodology against the traditional *Envelope spectrum* and the *FFT* is shown. In this case there are no other components that make difficult the diagnosis, the amplitude of the fault component in the proposed methodology is greater than the amplitude of the fault in the *Envelope spectrum*. In the *Envelope spectrum* the largest pea correspond to the *DCC*.

The difference in amplitude and the presence of other components between each other are in terms of the effectiveness of the *EMD*, the use of statistical parameters (*KR*) for the selection of *IMFs* with fault information is intended to avoid the visual inspection of the *IMFs*, [Tsao *et al.*, 2010] [Tsao *et al.*, 2012]. The selection of the most significant *IMFs* is guarantee.

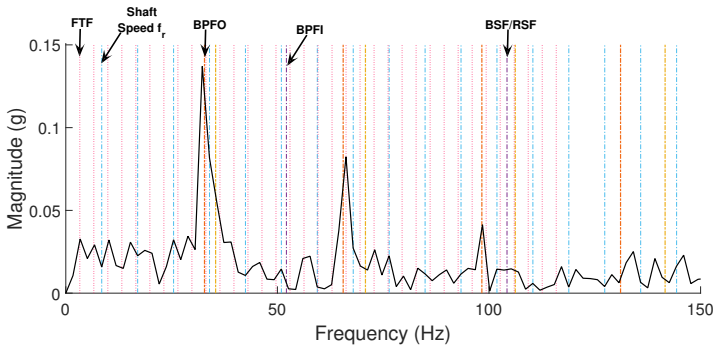


Figure E.13: *Envelope spectrum* of reconstructed signal from *OR* severe fault (*Tian-Yau Wu* database), the *BPFO* and harmonics are predominant in the spectrum

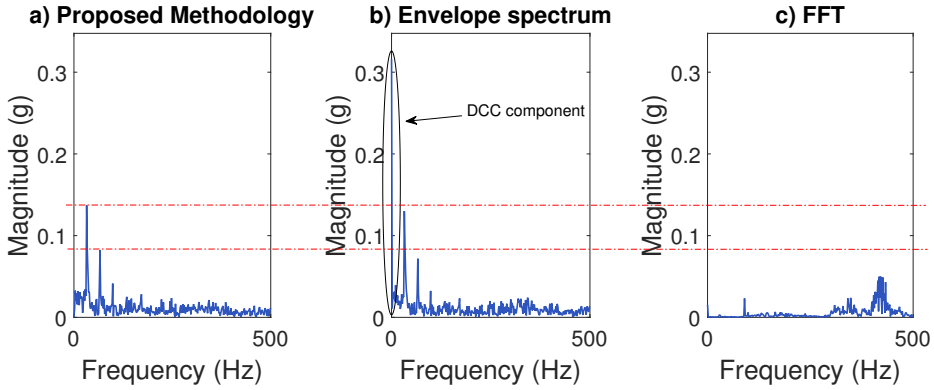


Figure E.14: (a) results of the proposed methodology, (b) *Envelope spectrum* of the original signal, and (c) the *FFT* of the original signal for *OR* severe fault (*Tian-Yau Wu* database), the proposed methodology removes the *DCC*, and the *FFT* does not detect the fault

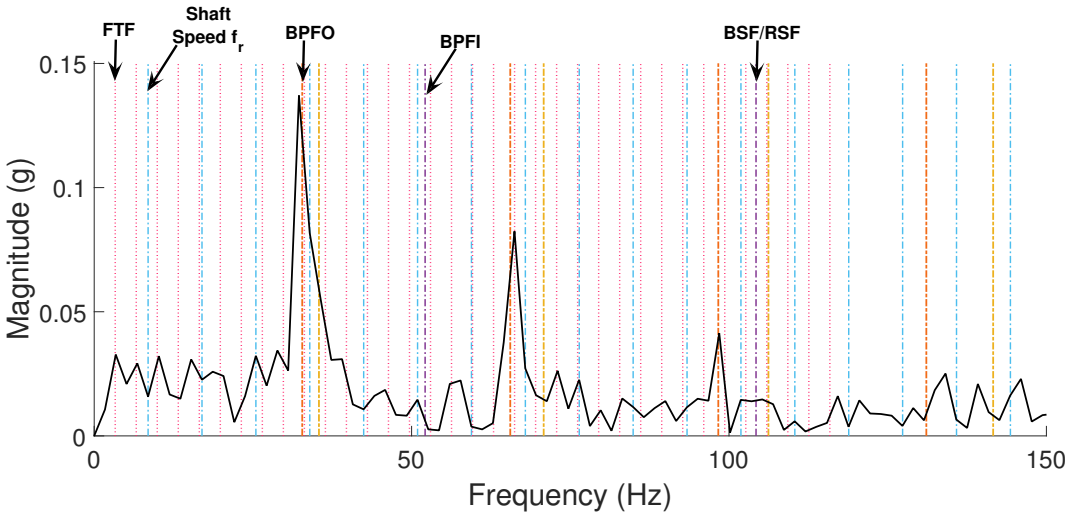


Figure E.15: *Envelope spectrum* of reconstructed signal *OR* severe defect (*Tian-Yau Wu* database), the *BPFO* and the two harmonics look sharper than in slight defect

E.3 *MFPT* Society additional results

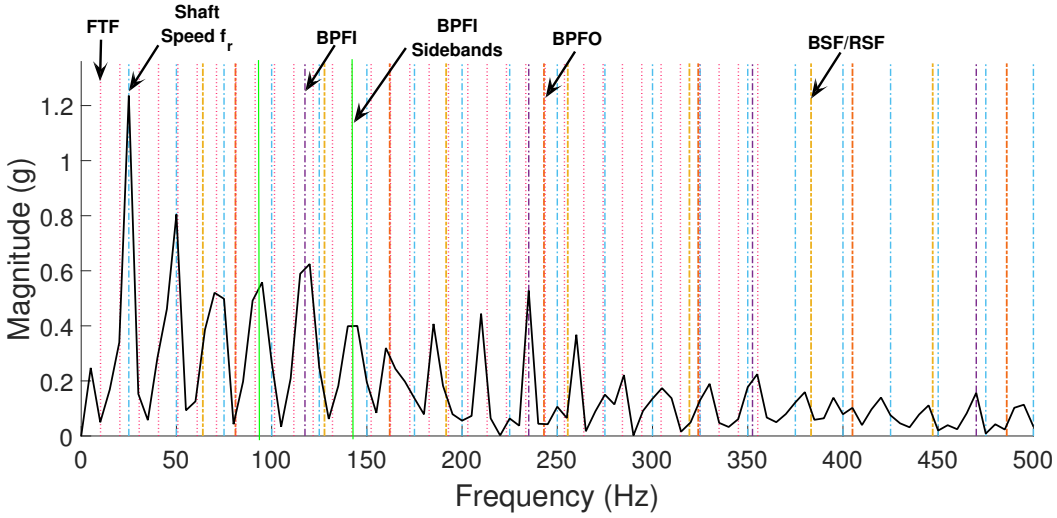


Figure E.16: *Envelope spectrum* of reconstructed signal *IR-250 lbs* (*MFPT* database), the *BPFI*, two harmonics and sidebands can be distinguished

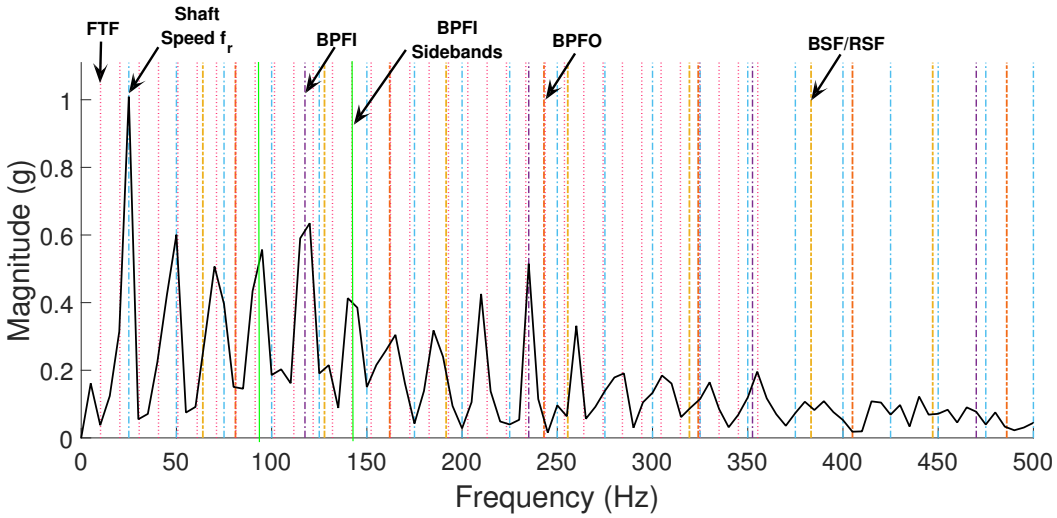


Figure E.17: *Envelope spectrum* of reconstructed signal *IR-300 lbs* (*MFPT* database), the same components as in *IR-250 lbs* are found but the amplitude is reduced

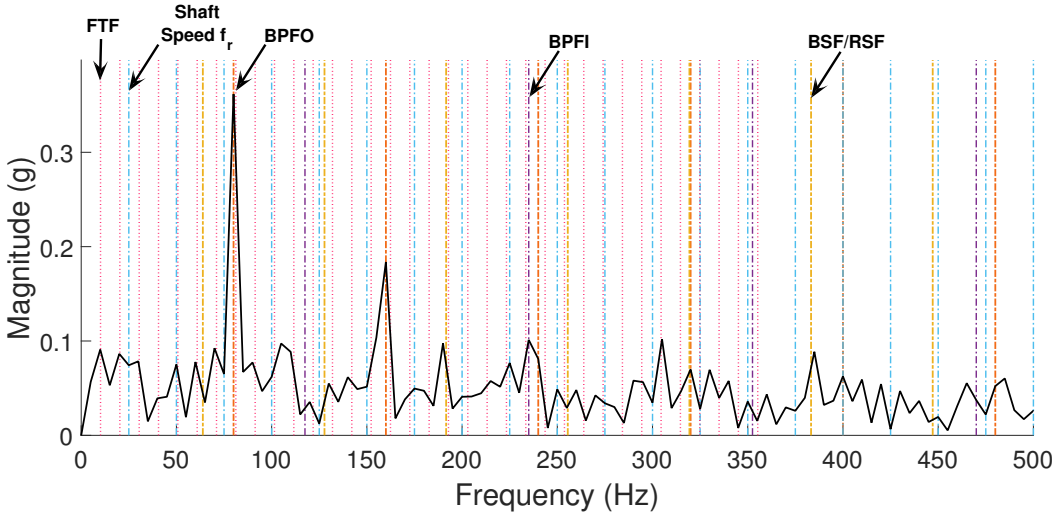


Figure E.18: *Envelope spectrum* of reconstructed signal *OR-250 lbs* (*MFPT* database), the *BPFO* and their harmonics are prominent, and a small peak is found at *BPFI*

Table E.1: Information of analyzed signals, the windows of the analyzed signals started from second 1

Database	File	Signal	Difficulty
CWRU	X209	*IR* (Y)	Clear
	X171	*IR* (P)	Clear
	X056	*IR* (N)	Clear
	X130	*OR* (Y)	Clear
	X198	*OR* (P)	Clear
	X200	*OR* (N)	Noisy
	X048	*RE* (Y)	Clear
	X222	*RE* (P)	Clear
	X225	*RE* (N)	Clear
Tian-Yau Wu	DFin04A18	*IR* SL	Clear
	DFin08A18	*IR* SE	Noisy
	DFout04A25	*OR* SL	Clear
	DFout08A25	*OR* SE	Clear
	DFroll04A24	*RE* SL	Clear
MFPT	InnerRaceFault_vload_6	*IR* 250 Lbs	Noisy
	InnerRaceFault_vload_7	*IR* 300 Lbs	Noisy
	OuterRaceFault_vload_6	*OR* 250 Lbs	Clear
	OuterRaceFault_vload_7	*OR* 300 Lbs	Clear

Appendix F

Developed Programs

For the developed methodology the *Empirical Mode Decomposition* function was obtained from Alan Tan[1]. To implement the methodology three basic Matlab functions were developed, one function to take a sample of the signal based on a given window size in (s), one function to compute the statistical features of the *IMFs*, and one function for plotting the *Envelope spectrum* and perform the diagnosis. The code used to analyze the faults in the three databases is:

```
%% Methodology
% Ruben Campos G.   rubencamposg@gmail.com
%
% This script is designed to analyze the signals from the 3 cases of study
%
%%
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Load the signal to analyze
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
load('filename')

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Select the language for the axis 1=Spanish  2=English
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
Languaje = 2;

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Select the type of fault to analyze 1=IR  2=OR 3=Ball 4=Cage
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
tFault = 1;

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Select the database of the signal  1=MPFT 2=CWRU 3=TY WU
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
```

[1] http://bit.ly/AlanTan

```
Database = 2;
switch(Database)
    case 1
        signal = bearing.gs;        %Vibration signal
        Fs=  48828;                 %Enter sampling fequency
        rpm = 1500;                 %Enter shaft speed in RPM
        RPM = rpm/60;               %Shaft speed in (Hz)
        OR = RPM * 3.2;             %BPFO
        IR = RPM * 4.7;             %BPFI
        BALL = RPM * 2.555;         %BSF/RSF
        CAGE = RPM * 0.406;         %FTF
        rev = 0.04;                 %Time of 1 revolution in (s)

    case 2
        signal = X198_DE_time';     %Enter vibration signal
        Fs = 12000;                 %Enter sampling fequency
        rpm = 1772;                 %Enter shaft speed in RPM
        RPM = rpm/60;               %Shaft speed in (Hz)
        OR = RPM * 3.5848;          %BPFO
        IR = RPM * 5.4152;          %BPFI
        BALL = RPM * 4.7135;        %BSF/RSF
        CAGE = RPM * 0.3983;        %FTF
        rev =  1/RPM;               %Time of 1 revolution in (s)

    case 3
        signal = Y;                 %Vibration signal
        Fs = 6400;                  %Sampling fequency
        RPM = 8.49;                 %Enter speed in (Hz)
        OR = RPM * 3.86;            %BPFO
        IR = RPM * 6.14;            %BPFI
        BALL = RPM * 4.17;          %BSF/RSF
        CAGE = RPM * 0.39;          %FTF
        rev =  1/RPM;               %Time of 1 revolution in (s)

end

%   Automatic selection of the fault frequency
switch tFault
    case 1
        AnalyzedF = IR;             %BPFI
    case 2
        AnalyzedF = OR;             %BPFO
    case 3
        AnalyzedF = BALL;           %BSF/RSF
    case 4
```

```
        AnalyzedF = CAGE;           %FTF
end

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% The signal is trimmed based on the desired # of revolutions
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
T1 = 1;                             %Initial time of the sample
T2 = T1+(rev *5);                   %Final timeof the sample
%                                   5 indicate the number of revolutions

% The function returns the trimmed signal (x) and time vector (t)
[x,t] = trimSignal(signal,T1,T2,Fs);

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Empirical Mode Decomposition
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

imf = emd(x);      % The EMD function, it returns the IMFs in a cell array

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Time domain features
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

%Conversion of the cell array of IMFs to a matrix to compute time domain
%features
N=size(imf,2);              % Determines the # of IMFs from the EMD
IMF=zeros(N,size(t,2));     % Variable to convert the IMFs to a matrix
for i=1:1:N
    IMF(i,:) = imf{1,i};    % Copy of the IMFs to the matrix
end

% Function to compute 17 time domain features for each IMF
[statVal] = statFeatures(IMF);


%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% IMF selection using the KR
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Identification of the largest KR, #4 limit the search to the first 4 IMFs
 HighestKR = max(statVal(1:4,3));
 % Find the # of IMF with largest KR
 IndexKR = find(statVal(:,3) == HighestKR);

 % Identification of the second largest KR
```

```
StatAux=statVal(1:4,3);                          % Auxiliar vector with the KR values of first 4 IMFs

StatAux(IndexKR)=0;                              % Replaces the largest KR to 0
HighestKR2 = max(StatAux);                       % Find the second largest KR
IndexKR2 = find(StatAux == HighestKR2);          % Find the # of IMF with largest KR

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Signal reconstruction with the selected IMFs
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
reconSignal = imf{1,IndexKR} +  imf{1,IndexKR2};       % Sum of the two IMFs

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Envelope spectrum of the reconstructed signal for the analysis
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

EnvSpect1(reconSignal,t,Fs,RPM,AnalyzedF,tFault,2);      % Function to plot the reconstructed signal and the envelope spectrum.
```

The functions used in the code are:

```
function [x, t] = trimSignal(Signal,iTime,fTime,Fs)
%   Rubn Campos G. A00821271  rubencamposg@gmail.com
%     This program trim the signal for a given window size (s)
%     t = Time vector
%     Signal = Original signal
%     iTime = Initial time (s)
%     fTime = Final time (s)
%     Fs = Sampling frequency of the original signal
%%

L = fTime-iTime;                    % Window size in (s)
ndatos =  abs(ceil(L*Fs));          % Window size in samples
datoInicial = ceil(iTime*(Fs));     % Initial data
x = Signal(datoInicial:(datoInicial+ndatos-1))';       %Trimms the signal
t = 0:(1/Fs):(ndatos-1)/Fs;         % Time vector of the new signal
```

```
function [statVal] = statFeatures(imf)
%   Rubn Campos G. A00821271  rubencamposg@gmail.com
%
%   This function computes time domain features for the IMFs
% imf = Matrix with the IMFs
% Rows = #IMF
% Columns: RMS, kurtosis, KR, Entropy, Peak value, PK-PK, Crest factor,
% skewness, variance, SD, promedio, root, clearance, shape, impulse, RMSxVar, RMSxSD.
```

```
9
10 %%
11 N=size(imf,1);            % # of IMFs
12 statVal= zeros(N,17);    % Creates a matrix to save the timedomain features
13
14
15 for i=1:1:N
16
17     statVal(i,1) = rms(imf(i,:));                    % RMS
18     statVal(i,2) = kurtosis(imf(i,:));               % Kurtosis
19     statVal(i,3) = statVal(i,1)*statVal(i,2);        % KR
20     statVal(i,4) = wentropy(imf(i,:),'shannon');     % Shannon Entropy
21     statVal(i,5) = max(abs(imf(i,:)));               % Peak
22     statVal(i,6) = peak2peak(imf(i,:));              % Peak-Peak
23     statVal(i,7) = peak2rms(imf(i,:));               % Crest Factor
24     statVal(i,8) = skewness(imf(i,:));               % Skewness
25     statVal(i,9) = var(imf(i,:));                    % Variance
26     statVal(i,10) = std(imf(i,:));                   % Standard Deviation
27     statVal(i,11) = mean(imf(i,:));                  % Mean
28     statVal(i,12) = (mean(sqrt(abs(imf(i,:)))))^2;     % Root
29     statVal(i,13) = statVal(i,5)/statVal(i,12);              % Clearance
30     statVal(i,14) = statVal(i,1)/mean(abs(imf(i,:)));        % Shape
31     statVal(i,15) = statVal(i,5)/mean(abs(imf(i,:)));        % Impulse
32     statVal(i,16) = statVal(i,1)*statVal(i,9);               % RMSxVar
33     statVal(i,17) = statVal(i,1)*statVal(i,10);              % RMSxSD
34
35 end
```

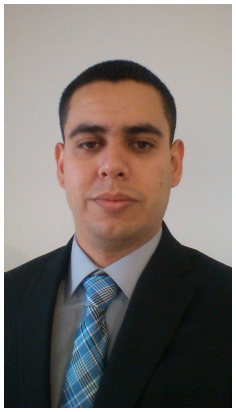
```
1 function  EnvSpect(Signal,t,Fs,RPM,AnalyzedF,tFault,Languaje)
2 %   Ruben Campos G. A00821271  rubencamposg@gmail.com
3 %
4 % Function to plot the reconstructed signal and the spectrum
5 %
6 % Signal = Signal to be plotted
7 % t = Time vector
8 % Fs = Sampling frequency
9 % RPM = Shaft speed in (Hz)
10 % AnalyzedF = Frequency of the fault investigated
11 % tFault = 1 for IR , 2 for OR, 3 for Ball , 4 for Cage
12 % Languaje = 1 for Spanish, Other for English
13 %%
14 %\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
15 % Label selection based on selected language
16 %\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
17 if Languaje==1
```

```
    xFreq = 'Frecuencia (Hz)';
    xTime = 'Tiempo (s)';
    yMag = 'Magnitud (g)';
    yAccel = 'Aceleracin (g)';
    titulo = 'SEAL Signal ';
else
    xFreq = 'Frequency (Hz)';
    xTime = 'Time (s)';
    yMag = 'Magnitude (g)';
    yAccel = 'Acceleration (g)';
    titulo = 'Signal SIGNAL ';
end

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Definee line colors for the plot
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
co=[...
    0.000 0.000 0.000; % 1 Black
    1.000 0.141 0.420; % 2 Pink
    0.180 0.333 0.745; % 3 Blue
    0.950 0.425 0.098; % 4 Orange
    0.466 0.674 0.188; % 5 Green
    0.929 0.694 0.125; % 6 Yellow
    0.301 0.745 0.933; % 7 Sky blue
    0.494 0.184 0.556; % 8 Purpure
    ];
set(groot,'defaultAxesColorOrder',co)

% Define line width
th1=1.2;      % Used for faults
th2=1.5;      % Used for signal

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Defines line type for faults
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
switch tFault
    case 1
        fault = 'BPFI';
        line = '-.';
        color = co(2,:);
    case 2
        fault = 'BPFO';
        line = '--';
        color = co(1,:);
    case 3
```

```
        fault = 'BSF';
        line = '-';
        color = co(6,:);
    case 4
        fault = 'FTF';
        line = '-';
        color = co(1,:);
end



%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Signal Envelope
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
recover = abs(hilbert(Signal));

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Trend Removal (DCC)
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
recover = recover - mean(recover);

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Envelope spectrum
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\

L=length(Signal);                % Length of the signal
freqFFT=Fs*(0:floor(L/2))/L;     % Create bins for Envelope spectrum

aux = abs(fft(recover)/L);       % computes the doubled sided FFT
FFT = aux(1:floor(L/2)+1);  FFT(2:end-1) = 2*FFT(2:end-1);    % Saves the one
    side of the spectrum
MAX = max(FFT);                  % Identifies the maximum magnitude

%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
% Signal plot
%\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
figure
subplot(2,1,1)
plot(t,Signal,'Color',co(1,:),'LineWidth',th1)
axis tight
set(gca, 'fontsize', 20)
xlabel(xTime,'FontSize',25)
ylabel(yAccel,'FontSize',25)
title(titulo)
legend('Signal')
```

```
107
108
109 %\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
110 % Envelope spectrum plot
111 %\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
112 subplot(2,1,2)
113 plot(freqFFT,FFT,'Color',co(1,:),'LineWidth',th2); hold on;
114 axis([0 1000 0 inf])
115 set(gca, 'fontsize', 20)
116
117 % Lines to identify the fault component and shaft speed component
118 for k=1:35
119     line([k*RPM k*RPM],[0 MAX*1.1],'Linestyle',':','Color',co(7,:),'LineWidth'
    ,th1)
120     line([k*AnalyzedF k*AnalyzedF],[0 MAX*1.1],'Linestyle',line,'Color',color,
    'LineWidth',th1)
121 end
122 xlabel(xFreq,'FontSize',25)
123 ylabel(yMag,'FontSize',25)
124 legend('Envelope spectrum Signal','SS-RPM',fault)
```

Curriculum Vitae



Rubén Campos García was born in Coahuila, México on April 29th, 1993. He received the degree of Mechatronics Engineer from the Tecnológico de la Región Carbonífera, Sabinas Coahuila, México in Dec 2015. He was part of the group of the Automotive Consortium, as member actively participated with the Bocar Group in the project named: MONITORING AND DIAGNOSIS OF HIGH SPEED MACHINING SPINDLES. He was accepted in the Master of Science in Manufacturing Systems program in the Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Campus Monterrey.

This document was typed in using LaTeX by Rubén Campos García.