USE OF THE CONTINUOUS WAVELET TRANFORM TO ENHANCE EARLY DIAGNOSIS OF INCIPIENT FAULTS IN ROTATING ELEMENT BEARINGS

A Thesis

by

SCOTT WEATHERWAX

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2008

Major Subject: Mechanical Engineering

USE OF THE CONTINUOUS WAVELET TRANFORM TO ENHANCE EARLY DIAGNOSIS OF INCIPIENT FAULTS IN ROTATING ELEMENT BEARINGS

A Thesis

by

SCOTT WEATHERWAX

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Alan Palazzolo
Committee Members,	Steve Suh
	Salih Yurttas
Head of Department,	Dennis O'Neal

August 2008

Major Subject: Mechanical Engineering

ABSTRACT

Use of the Continuous Wavelet Transform to Enhance Early Diagnosis of Incipient Faults in Rotating Element Bearings. (August 2008)

Scott Weatherwax, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Alan Palazzolo

This thesis focused on developing a new wavelet for use with the continuous wavelet transform, a new detection method and two de-noising algorithms for rolling element bearing fault signals. The work is based on the continuous wavelet transform and implements a unique Fourier Series estimation algorithm that allows for least squares estimation of arbitrary frequency components of a signal. The final results of the research also included use of the developed detection algorithm for a novel method of estimating the center frequency and bandwidth for use with the industry standard detection algorithm, envelope demodulation, based on actual fault data. Finally, the algorithms and wavelets developed in this paper were tested against seven other wavelet based de-noising algorithms and shown to be superior for the de-noising and detection of inner and outer rolling element race faults.

DEDICATION

I dedicate this thesis to my parents and family: The world can be a big, exciting, challenging, fun adventure. It can also be dangerous and daunting. I am very thankful to have a family I can count on and trust and will even remember me … even when my travels land me on an island in the middle of the Pacific Ocean. Thank you for believing in me.

ACKNOWLEDGEMENTS

I would like to thank Dr. Palazzolo for his patience and insight. Without his guidance, I would have been stumbling in the dark. I would also like to thank the other members of my committee, Dr. Salih Yurttas and Dr. Steve Suh, for their encouragement and help. Most importantly, beyond the guidance and encouragement these professors provided, were the always friendly smiles and open doors. Thank you.

Life would have been very dull indeed without my library buddy, Leah Silverman. Leah, thanks for sharing a table and trading stories. I will be one of the first in line to buy your published novel.

I would also like to thank my family for listening to me whine about writing my thesis and giving encouragement. More specifically, I would also like to thank my chief editor, my mom, Judy Weatherwax. While writing a document this big is a challenge, editing a document this big, written by me, is an even more daunting challenge. Thank you.

Finally, I want to thank Dr. Kenneth Laparo of Case Western Reserve University. He generously allowed me to use data that he collected from an experimental setup that he constructed for his own research.

NOMENCLATURE

CWT	Continuous Wavelet Transform
DRFS	Discrete Reduced Fourier Series
FT	Fourier Transform
STFT	Short Time Fourier Transform
WVD	Wigner-Ville Distribution
WT	Wavelet Transform
DWT	Discrete Wavelet Tranform
SVD	Singular Value Decomposition
MAD	Median of the Absolute Deviation
CPS	Cyclic Power Spectrum
PSD	Power Spectral Density
DSP	Digital Signal Processing

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

Page

Page

Page

Page

Page

Page

Page

LIST OF TABLES

1. INTRODUCTION

Rotating element bearings are prolific in today's machines and engines. They provide support for large loads and stiffness characteristics that are necessary in many applications [1]. Unfortunately, failure of these components often proves difficult to predict and diagnose. As a result bearing failure can often seem, without adequate diagnosis tools, very sudden and dramatic. Vibration resulting from failure can not only indicate a faulty bearing but also damage other components that are mechanically linked to the faulty bearing in question. Without accurate and early diagnosis of rolling element faults, machine operators suffer from large financial damages due to unscheduled maintenance and replacement of damaged components.

As a result, research has been diligent in this area for a number of decades [2,3]. Faulty bearing diagnosis has been predominantly detected and monitored using accelerometers attached either to or around bearing housings. With the advent of the digital computer and dramatic increase in both processing power and affordability, digital signal processing (DSP) techniques have become the prevalent tool used by engineers to detect, identify and give prognosis of faulty rolling element bearings. The current industry standard tool used by engineers for decades is Envelope Demodulation [2,4]. Using this technique, bearing vibration signals can be used to both detect and identify a rolling-element bearing fault. Limitations exist with this tool, however, in detecting incipient faults and detecting bearing fault severity. Obviously, the earlier a fault can be

---

This thesis follows the style of Mechanical Systems and Signal Processing.

detected and diagnosed the earlier engineers in charge of operating the equipment with faulty bearings can schedule maintenance with the minimum financial impact.

In recent decades a number of new signal processing techniques and methods have been developed in various engineering disciplines. A number of these techniques, including the use of the Wavelet Transform, have been modified, developed and adapted to detect incipient bearing faults [5-19]. This document will provide an introduction to the rolling element fault detection problem and a succinct literature review followed by an introduction to the wavelet transform. This research focused using a unique Fourier Series estimation tool, the Reduced Discrete Fourier Series, with the Continuous Wavelet Transform to develop a novel rolling element fault detection tool as well as two de-noising algorithms to enhance fault diagnosis. These two tools will be explained in detail as well as a new wavelet, the Frequency Domain Impulse Response Wavelet. Finally these new algorithms and the Frequency Domain Impulse Response wavelet will be compared to several current wavelet based de-noising algorithms and wavelets to evaluate their effectiveness.

1.1 The Rolling Element Bearing Fault Detection Problem

Bearing faults are detected using either vibration signals or, less commonly, high frequency acoustic emissions. Figure 1 shows a simple diagram of how a mechanical fault, appears with respect to the major components of the mechanical system. As the bearing rotates, the rolling elements in the bearing impact the fault located on the inner race. Faults can also exist on the rolling elements themselves as well as on the outer race.

These impacts cause vibrations that can be detected by sensing the vibrations set up in the mechanical structure surrounding the bearing.

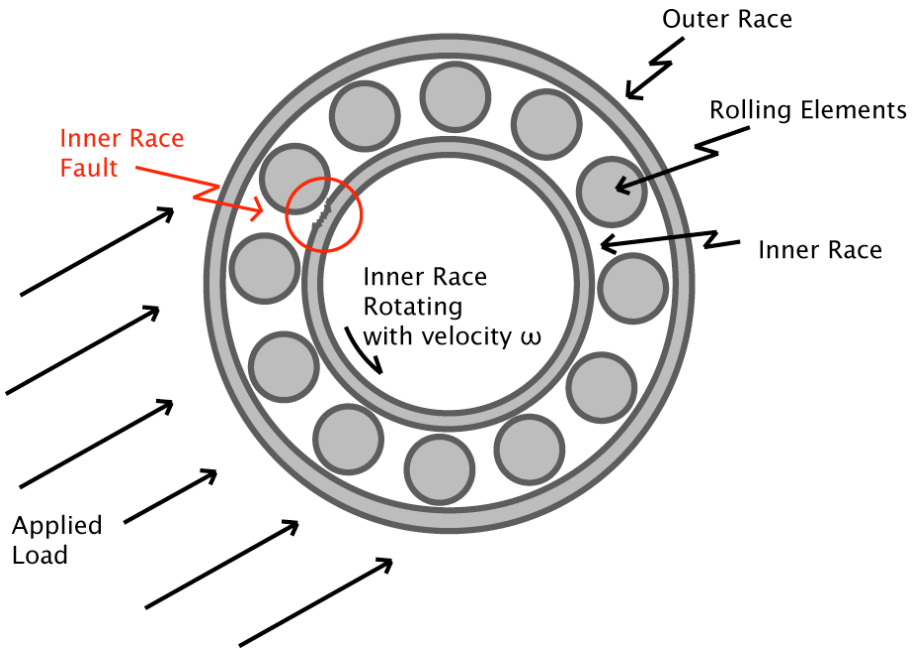

Figure 1: Graphical illustration of a rotating element bearing with an inner race fault.

The first successful work in characterizing the vibration signal produced faults in rotating element bearings was in the late 1970s and early 1980s by McFadden and Gupta [1, 4, 20]. More recently, random process theory has been successfully adopted to describe mechanical signals. The reader is referred to the sited documentation for a much more developed model of fault modeling, as this section will only give a high level overview of the faulty mechanical signal [21-23].

A very simple model of the signal from the mechanical fault can be gleaned from Figure 1. The rotation of the shaft causes periodic impulses as the rolling elements of the bearing strike a fault. These impacts excite one or more natural frequencies in the mechanical structure of the bearing and or vibration sensor. A very simple model of the signal produced by the mechanical system is an exponentially decaying sinusoid [20]:

$$s(t) = e^{-bt}\sin(\omega_o t) \tag{1}$$

Figure 2 shows a visual representation of the ideal fault signal.

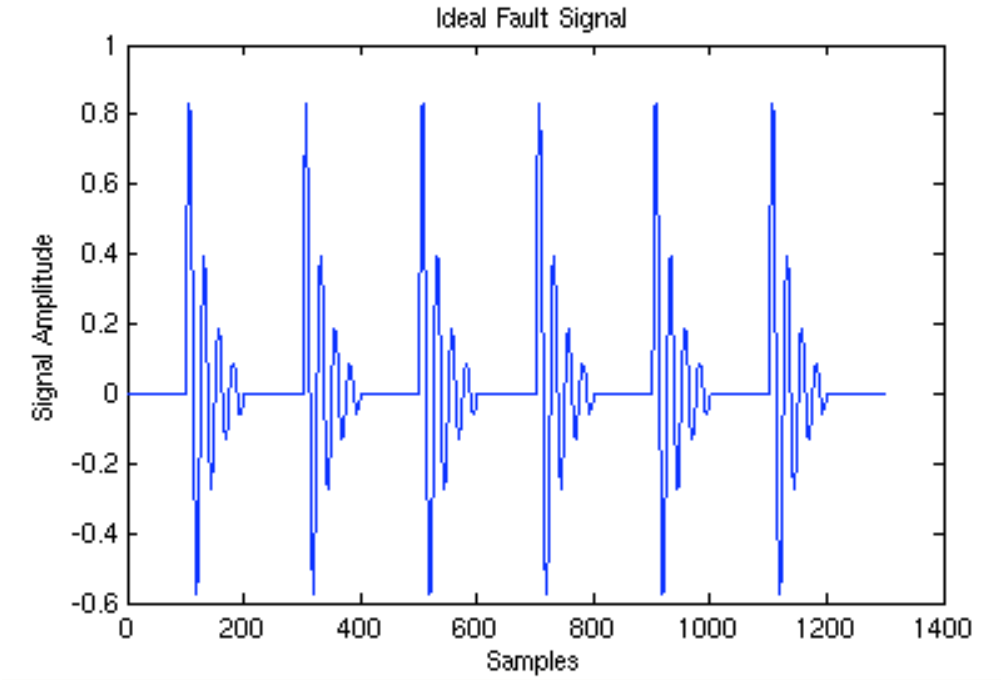

Figure 2: Visual representation of the ideal fault signal

As can be observed with this ideal signal, the faults themselves are periodic. Furthermore, the cyclic frequency of this impulse train is tied directly into the rotating speed of the shaft. Equations (2a) through (2d) and Figure 3, relate the frequency of impulses to the shaft speed assuming the shaft speed is constant.

Fundamental Train Frequency (FTF): $f_t = \frac{f_s}{2}\left[1 - \frac{B_d}{P_d}\cos(\phi)\right]$ (2a)

Ball Spin Frequency (BSF): $f_B = \frac{f_s}{2}\left(\frac{B_d}{P_d}\right)\left[1 - \left(\frac{B_d}{P_d}\right)^2 \cos(\phi)\right]$ (2b)

Outer Race Frequency (ORF): $f_O = Nf_t$ (2c)

Inner Race Frequency (IRF): $f_I = N(f_s - f_t)$ (2d)

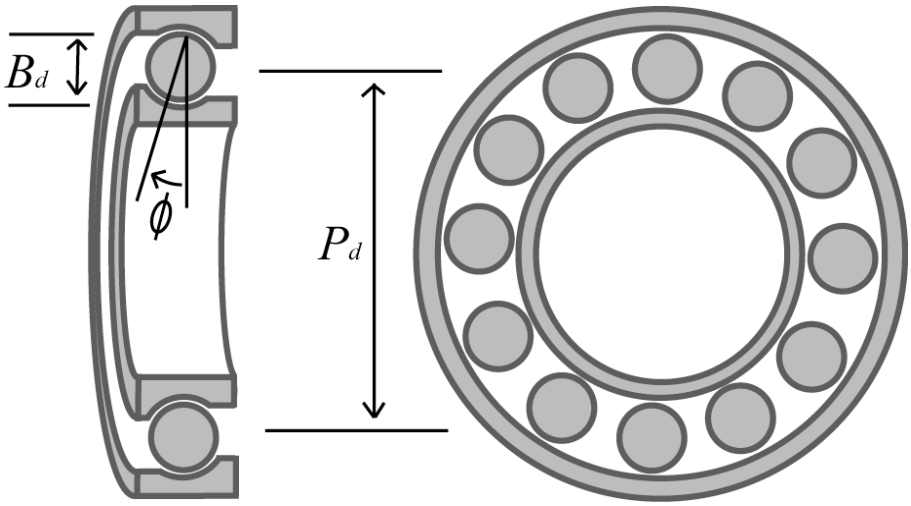

Figure 3: Ball geometry with important measurements for bearing fault frequencies highlighted. B_d is the rolling element diameter. P_d is the diameter of the bearing as measured to the center of the rolling elements. N is the number of rolling elements and ϕ is the contact angle of the ball with the outer race.

The frequency of the decaying sinusoid is generally, in practice, the natural frequency of the bearing's outer race or the natural frequency of the vibration sensor. However, in more complicated configurations multiple excitations can occur from different structures. Due to kinematic and load considerations there is variation in both the timing and amplitude of the fault signal. As a result, the rolling elements do not strike the bearing fault in an exactly periodic manner, and this leads to a degree of randomness in the spacing between impulses. The location of the inner race fault, for the case given in Figure 1, provides yet another source of variation. As the inner race rotates the periodic excitations produced by the fault on the inner race are modulated by the applied load on the bearing. Furthermore, there are natural amplitude variations dependent on how the rotating elements strike the developing inner race fault. All these variations lead to the definition of the fault signal using random process theory rather than the traditional deterministic model [4, 20, 21, 24-26].

A stochastic model of the bearing fault signature is [24]:

$$x(t) = \sum_i A_i s(t - iT - \tau_i) q(iT) + n(t) \tag{3}$$

where $n(t)$ is the additive noise to the signal, A_i is the amplitude of the ith impact and τ_i is an independent lag variable and $s(t)$ is from Eq. (1) with both A_i and τ_i as random variables. $q(iT)$ represents the modulation of the signal $s(t)$ due to the rotation over the applied load. Without the random time lag τ_i and amplitude fluctuations A_i this signal model falls out into a much more traditional signal model that has existed since the early to mid-1980s. The traditional signal model predicts, among other frequency components, a series of discrete spectral lines in the frequency domain corresponding to the fault

frequency [4, 20, 25, 26]. These discrete spectral lines theoretically extend across the entire frequency domain, and are used to detect and diagnose faults in the mechanical bearing. More recently, it has been shown that the effect of the random time lag τ_i, and amplitude fluctuations A_i, even if small, apply what can be considered a severe low pass filter on the discrete spectral lines [21]. This effect severely hampers traditional signal processing techniques in detecting fault signatures, especially in situations with noise.

Generally speaking, the natural frequency of the vibration sensor and/or mechanical housing, by modulating the impulse function from the rolling element striking the fault, preserve the spectral information that is repressed in the low frequency spectra. As a result, practical vibration analysis has consisted for many years of demodulating the bearing signal at the natural frequency observed in the frequency domain to attempt to recover spectral lines that have been preserved in the higher frequency band [2-4, 21, 23]. One of the central difficulties in the use of envelope demodulation is the determination of natural frequency that is modulating the bearing fault. Wavelet de-noising algorithms are an alternative means to recover spectral information preserved at the natural frequency of the fault signal that does not necessarily rely on the determination of the bearing fault natural frequency.

1.2 Literature Review

A large number of techniques are used to quantify bearing damage [2, 3, 14, 27]. These techniques exist both in the frequency and time domains, with the frequency domain methods being more prevalent. Several notable time domain methods are the

measure of Kurtosis, Crest factor and signal RMS value [7, 28]. Unfortunately, time domain analysis methods have shown that they do not provide consistent damage indication as each different indicator provides information about only specific stages of a growing bearing fault. For example, Kurtosis is a good indicator of incipient faults; unfortunately, at higher levels of damage Kurtosis will indicate a healthy bearing. However, it has been shown that a combination of time domain techniques when used in conjunction with a classifier can successfully diagnose a bearing fault [7].

As a result, frequency domain methods are much more commonly used to predict bearing damage and make up the bulk of current signal processing methods used by industry today. By transforming the signal into the frequency domain using the Discrete Fourier Transform (DFT), the spectral components in a faulty bearing can be observed depending on the strength of the underlying fault signature. The basic signal model is as follows: As rolling elements of the bearing pass over a fault an impulse function is generated. As this impulse function travels through the bearing to the vibration sensor, the signal is modulated by the resonate frequencies of the bearing and vibration sensor [2]. Many more advanced frequency domain methods rely on a number of more developed faulty bearing signal models [4, 25, 26, 29]. By using the developed models for the faulty bearing not only can the faulty bearing be identified as faulty, but the type of fault can also be diagnosed.

A number of frequency domain techniques are used for bearing analysis. The most established approach is Envelope Analysis or High Frequency Resonance. This technique demodulates the bearing vibration signal around a resonate frequency to pull

out salient damage frequency characteristics [3, 30-32]. Envelope Analysis has also been employed with adaptive filters to help eliminate errant noise and increase the SNR of the diagnostic signal [31, 33]. Other techniques, including the power spectral density (PSD), power spectral density of the expected signal squared and amplitude probability density (APD) have also been shown to be successful techniques [34].

More recently new methods in fault detection for bearings have been developed. During operation rotating bearings produce vibration signals that have non-stationary statistics. Notably, the bearing operation can be termed pseudo-cyclostationary due to the slip inherent in the rolling elements during operation [3]. Furthermore, changes in the environment also introduce non-stationary components into the vibration signal. These non-stationary components in the measured vibration often contain information about machine faults [27]. Standard Fourier Transform (FT) techniques implicitly assume the stationary of the signal that is being transformed. As a result, the non-stationary components of the signal cannot be measured and are detrimental and interfere with standard frequency domain techniques [6].

Methods to exploit the non-stationary characteristics of the signal have led to the development of the Short Time Fourier Transform (STFT) and the Wigner-Ville Distribution (WVD) as well as the Wavelet Transform (WT) [27]. All of these techniques lead to a time frequency distribution, however all have distinct disadvantages. The WVD has interference terms that can lead to possible misdiagnosis of information. The STFT has problems providing optimum frequency resolution and has no orthogonal basis leading to the inability to determine a fast implementation algorithm. The

development of the wavelet theory has led to a novel and fast growing signal processing tool that gives the frequency time distributions as attempted by the STFT and the WVD and similarly, not only can be inverted, but also modified in the time-frequency domain. This aspect of the WT is unique and lends itself to de-noising.

Wavelet analysis over the last ten years has been applied to machine condition monitoring and fault diagnostics in a wide variety of ways, such as fault feature extraction for classification techniques, singularity detection, de-noising of weak signals, vibration signal compression, and system and parameter identification techniques [27]. In the area of vibration analysis for fault, detection wavelet analysis has found to be a better indicator of fault damage when compared to the traditional frequency domain methods of fault detection [35-37]. Furthermore, it has been shown that Wavelet Analysis can diagnose faults in situations where more traditional bearing failure detection mechanisms fail to work, namely, in low speed situations where periodic excitations are more difficult to detect [12]. Wavelet analysis has been used to not only successfully detect faults in rotating machinery, but when used with a number of successful classification techniques also diagnose the type of defect present [18, 38, 39, 40]. WT techniques have also been used to successfully de-noise signals prior to fault detection [10, 15, 19, 41].

The wavelet transform has also been applied to remove from noise faulty bearing signals, however without the prolific success as seen in other fields. [42, 43] Because the bearing fault signal has a narrow bandwidth and is impulsive, displaying very little energy, it can easily get overwhelmed in broadband noise, as is often the case [22]. The

de-noising via wavelet thresholding using the discrete wavelet transform has been attempted, but the most research has attempted using a more non-traditional approach of de-noising the bearing fault signal using the continuous wavelet transform [19, 43-45]. Another approach has been to use the continuous wavelet transform as an optimal filter bank and select the scale that most represents the fault signal [10, 15, 18, 45, 46].

One of the main weaknesses that has kept wavelet analysis from being widely adopted for fault diagnosis is the variety of different processing parameters and interpretation techniques. For example, there are a large number of different wavelets that can be used with the same Wavelet Transform (WT) technique [47-50]. More recently with the development of the lifting scheme for wavelet construction, new wavelets tailored to particular signals can be constructed [39, 51]. Also, different WT techniques, such as Basis Pursuit and Matching Pursuit, can be employed and show better performance [52, 53]. Work has been done in an attempt to provide criteria to optimize wavelet parameters for signal detection; however, this is still a growing field [18, 52].

2. THE WAVELET TRANSFORM

Wavelet Theory has existed at various levels of development in a wide variety of different fields, from geophysics to quantum mechanics, dating back to prior to the 1970s. However, this mathematical tool has only really seen widespread development and acceptance since the late 1980s [54]. Wavelet analysis is a tool that has been used to overcome many of the problems with traditional signal processing by providing an optimum time-frequency relationship. Standard Fourier Transform (FT) theory states that all continuous-time signals can be decomposed into an infinite sum of sinusoids. Applications of FT are used in a variety of different applications from filter theory to vibration analysis. While a very useful transform, when the Fourier Transform is applied, information in the time domain, while not lost, cannot be observed. An example of this can be observed in Figure 4, where two signals show the identical signal shifted in time. The amplitude of the FT for both signals displays exactly the same information. The phase information, not shown, yielded from the FT of both the signals while different, gives no direct correlation to the difference in time between the actual signals.

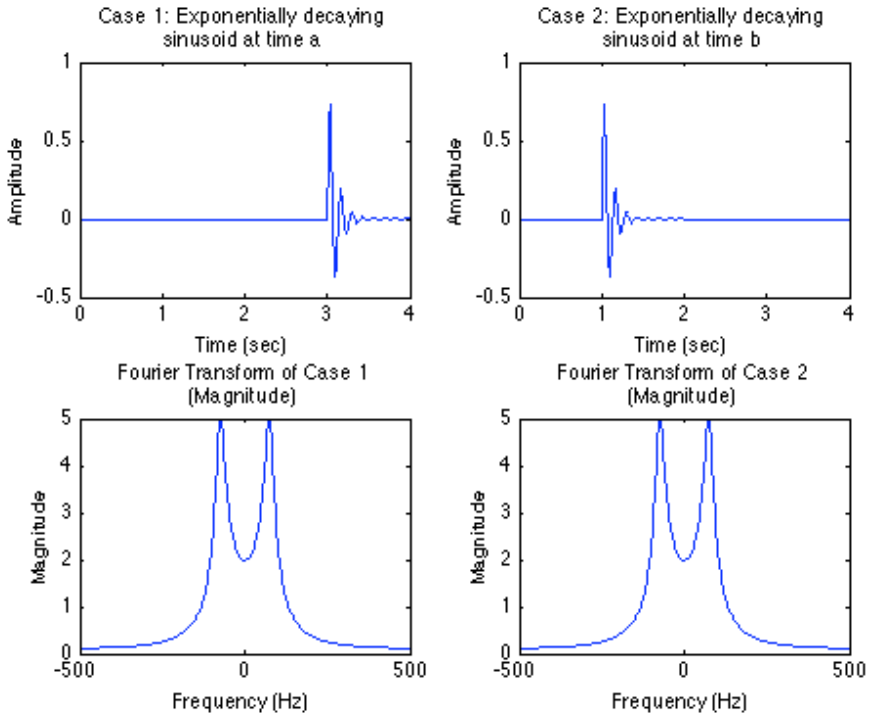

Figure 4: Application of the Fourier Transform of the same signal shifted to two different locations in time. The amplitude of the Fourier Transform is exactly the same for both signals. The phase information, not shown, is different, but gives no correlation to the time difference observed.

In dealing with transient signals, as with the exponentially decaying sinusoid seen in Figure 4, it is often important to not only determine the presence of a signal but also

localize the signal in time. As a result, while the FT is extremely useful for detecting a transient signal corrupted by noise based on spectral characteristics, it is impossible to localize the transient signal in time.

2.1 The Match Filter

A number of other tools have been developed to detect transient signals in time. One of the most popular, used often, in sonar and radar applications is a technique known as match filtering [55, 56]. Given a signal

$$x(t) = s(t) + n(t) \tag{4}$$

where $s(t)$ is a known transient signal occurring at an unknown time corrupted by noise $n(t)$. Locating the transient signal $s(t)$ in time can be solved by cross-correlating the signal of interest $x(t)$ with a detection filter $h(t)$. This cross-correlation product can be seen in (5).

$$y(t) = \int_{-\infty}^{\infty} h(\tau - t)x(\tau)d\tau \tag{5}$$

A graphical illustration of (5) is given in Figure 5.

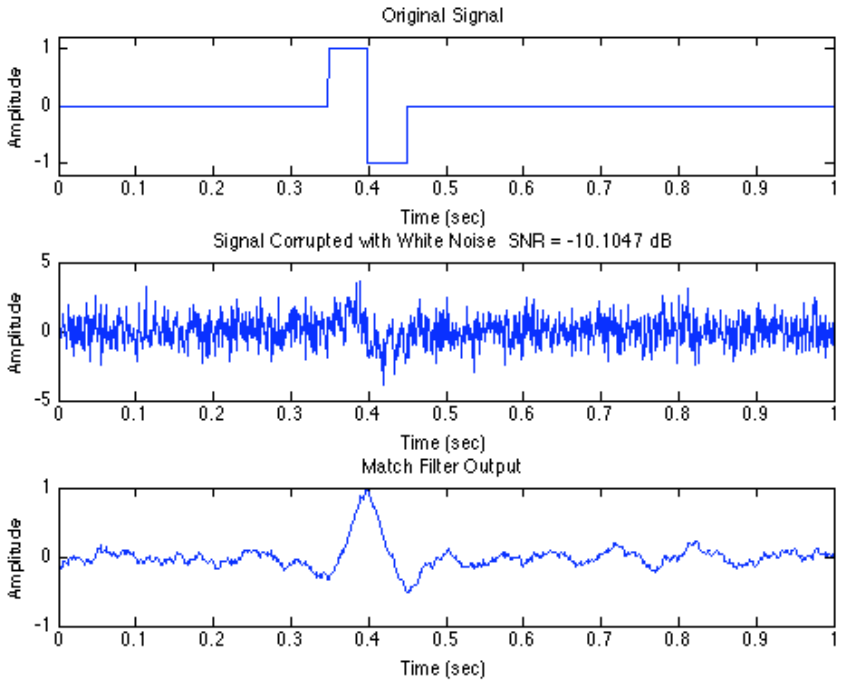

Figure 5: This figure demonstrates the effectiveness of a match filter. The original signal has been corrupted with white noise and using the original signal as a match filter the location of the square pulse has been located at 0.4 seconds.

An optimum filter $h(t)$, that maximizes the signal to noise ratio (SNR) of the output $y(t)$, can be built if both the estimate of the noise n(t) and the signal $s(t)$ is known. Furthermore, it can be shown that if the noise corrupting the signal $n(t)$ is white noise, that the detection filter, or match filter, $h(t)$ is optimum when it matches or is equal to $s(t)$ [64].

Match filtering can also be expressed in the frequency domain as (6) where $X(\omega)$, $H(\omega)$ and $Y(\omega)$ are the Fourier transforms of $y(t)$, $x(t)$ and $h(t)$ respectively with $H(\omega)^*$ representing the complex conjugate of $H(\omega)$.

$$Y(\omega) = X(\omega)H^*(\omega) \tag{6}$$

Equation (6) bears remarkable similarity to a convolution product seen in linear time invariant (LTI) theory. As a result, while the match filter concept is best understood in the time domain, it follows the time-frequency duality concepts that can be used with traditional LTI filters. Unfortunately, the creation of the optimum match filter $h(t)$, requires knowledge of a replica or estimate of the signal that is hidden or imbedded in the noise and an estimate of the noise. As a result, in a large number of cases, as with bearing fault identification, this is not realistic.

2.2 The Continuous Wavelet Transform

The Wavelet Transform, more specifically the Continuous Wavelet Transform (CWT) that is used in this research, is a tool that captures some of the best properties of both a traditional matched filter and is used to create an optimal time frequency distribution. The CWT is given in (7):

$$w_\psi(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \psi^*\left(\frac{t-b}{a}\right) x(t)dt \tag{7}$$

Readers will notice the immediate similarities between the CWT (7) and a traditional matched filter with the filter $h(t)$ replaced with a wavelet, $\psi(t)$. If the variable a is held to be one and the wavelet, ψ, is real, the CWT defaults to a simple match filter.

To understand the effect of the variable *a*, in the CWT, it is important to understand the definitions that define a wavelet [57]:

i.) A wavelet must have finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty \tag{8}$$

ii.) The wavelet must have a zero DC component and be band limited.

$$C_g = \int_{-\infty}^{\infty} \frac{|\psi(f)|^2}{f} \, df < \infty \tag{9}$$

iii.) For a complex wavelet the Fourier transform must be real and vanish for negative frequencies.

Focusing only on criteria *i* and *ii*, results in defining a wavelet to simply be a band pass filter. These are lenient restrictions. A multitude of different wavelets can exist with very different time and frequency domain properties. A wavelet can therefore be seen, simultaneously a replica that ideally matches the signal of interest, and a band pass filter that can be used to filter to separate the signal into discrete frequencies. Figure 6 displays several different common wavelets.

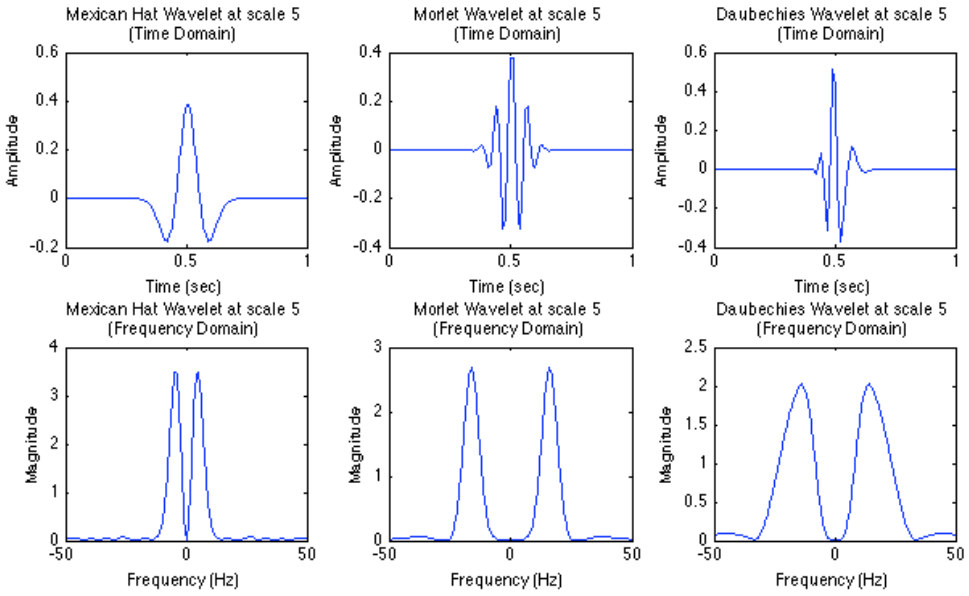

Figure 6: Three wavelets, the Mexican Hat, Morlet and Daubechies wavelets are shown for comparison in the time and frequency domains.

Going back to the CWT definition in (7), the variable a, is responsible for two effects. One, the variable a, denoted as the scale factor, dilates or contracts the wavelet in the time domain and second, simultaneously decreases or increases the center frequency of the wavelet. This can be seen by recalling the following Fourier Transform property,

$$\psi\left(\frac{t}{a}\right) \Leftrightarrow a\hat{\psi}(af) \tag{10}$$

where $\psi(t)$ represents a time domain signal and $\hat{\psi}(f)$ is its corresponding Fourier Transform [58]. A graphical impression of this concept is given in Figure 7 using the Morlet wavelet. As the wavelet scale is increased the wavelet energy also increases. To

counter this, ensuring that the energy of each wavelet is the same a normalization factor is included in front of the CWT integral in (7).

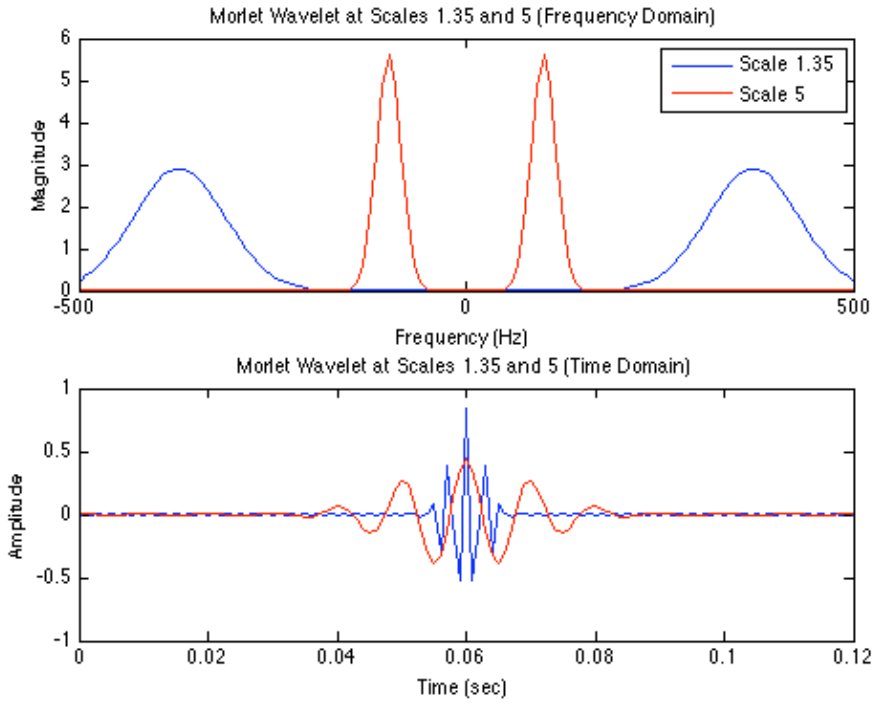

Figure 7: Morlet wavelet displayed for two different scales. As the scales increase the center frequency of the filter decreases and the effective width of the filter increases.

While there are a number of restrictions on what constitutes a wavelet, a correctly chosen wavelet can be used to detect a transient signal without detailed knowledge of exactly what the transient signal looks like by iteratively scaling the wavelet and performing the wavelet transform. If the wavelet ψ, at scale a, matches a transient in the signal, the CWT of the signal defaults to a match filter operation and any transients will be detected. An example of this concept is the use of the Morlet wavelet to capture a 300 Hz transient pulse. Figure 8 shows the time domain representation of the Morlet wavelet and a signal consisting of a 40 Hz sinusoid with an imbedded 300 Hz transient pulse. As the scale a of the Morlet wavelet increases the wavelet shrinks in time. At two scales, the Morelet wavelet will correlate to a high degree with different parts of the signal. The first correlation will occur when the Morlet wavelet matches the low frequency 40 Hz sinusoid. The second correlation will occur when the Morlet wavelet matches the high frequency transient pulse at 300 Hz.

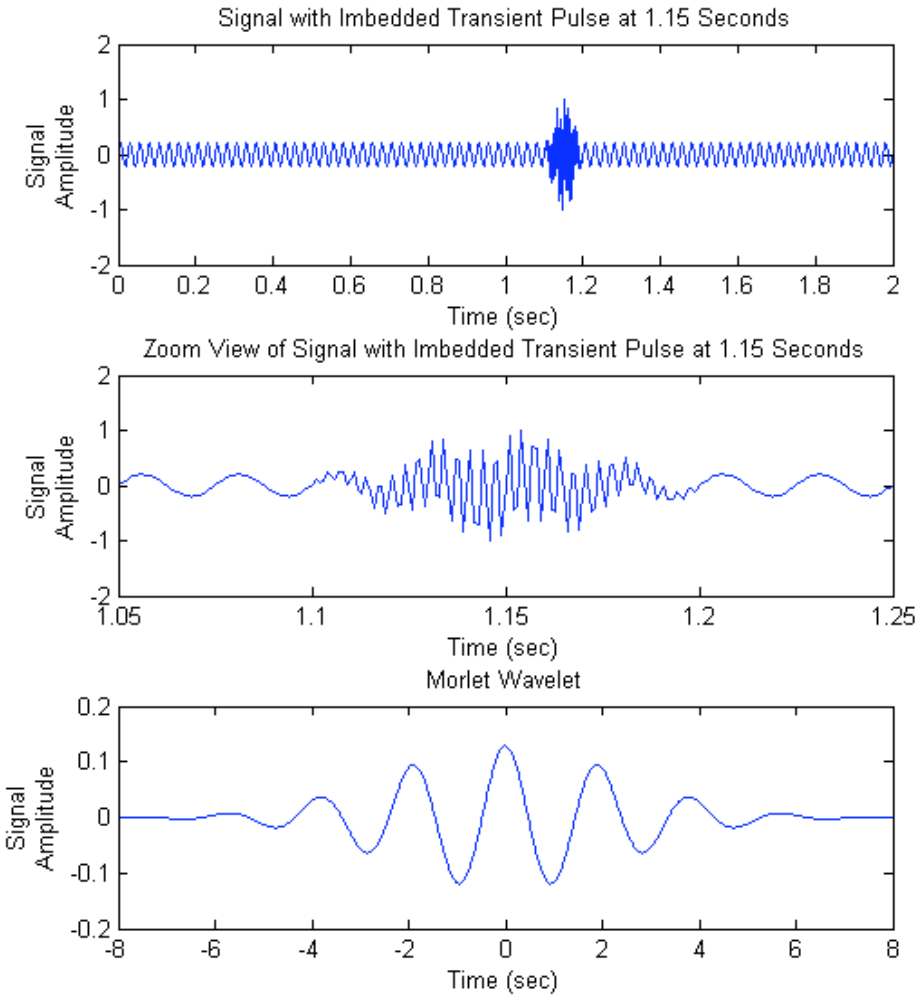

Figure 8: As the scale *a* increases the Morlet wavelet will shrink in time. As the wavelet shrinks it will correlate with the signal to a high degree at two scales. The first scale will be when the Morlet wavelet matches the 40 Hz sinusoid. The second place the Morlet wavelet will correlate to a high degree is when it matches the 300 Hz transient signal at 1.15 seconds.

Figure 9 displays the result of the correlation for 500 scales from 1 through 25. What is remarkable is because the Morlet wavelet can also be viewed as a band pass filter, a time frequency distribution is produced. Similarly, as the scale of the filter *a* increases, the temporal width of the filter shrinks to match the filter response of the wavelet. As a result, the width of the filter is directly related to the value of the scale *a*. For high frequency components of the signal, the width of the filter in the time domain is small. This allows for high frequency components of the signal to be localized in time with high accuracy. As the scale increases, the filter bandwidth shrinks and the filter grows in the time domain. This allows for high resolution of detection of low frequency components of the signal with high accuracy in the frequency domain.

As a result, it can be stated that for each successive filter iteration with the wavelet at an increased scale *a*, the filter that is band limiting the signal is at an optimum window size for center frequency of the wavelet.

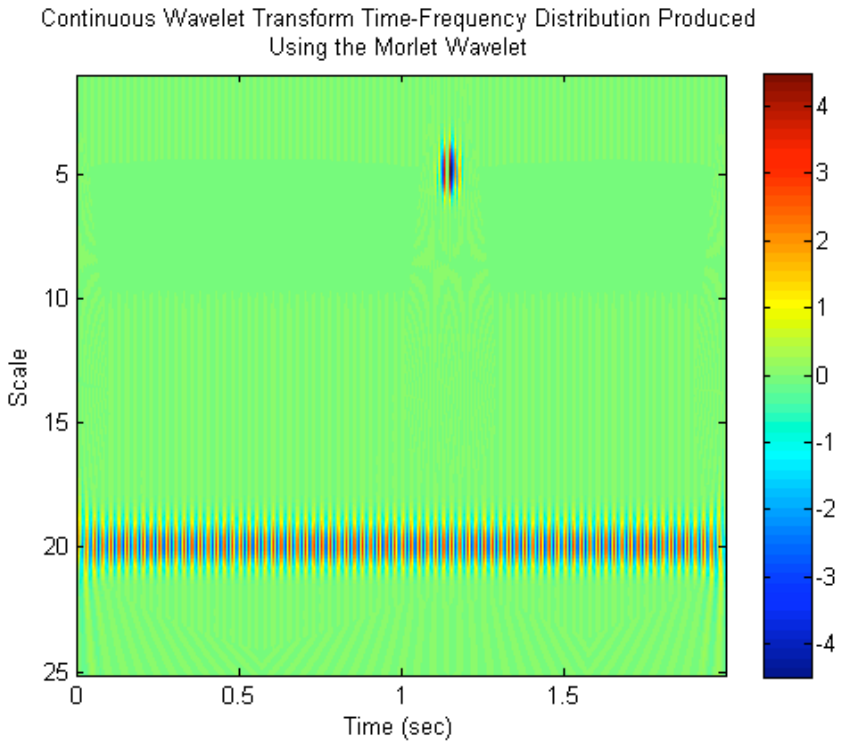

Figure 9: This is a time-scale distribution produced by implementing the Continuous Wavelet Transform using the Morlet wavelet. An alternate way of examining the continuous wavelet transform is as a match filter with continuously changing scale *a*.

2.3 Comparison to the Short Time Fourier Transform

Another tool often used in practice to both identify the presence of and locate transients in a signal is known as the Short Time Fourier Transform (STFT) or

spectrogram. A STFT calculates and plots the absolute value of the frequency components, via the Fast Fourier Transform (FFT) of a signal within a moving window. The results are arranged in a two dimensional image. As a result, the frequency components are approximated for individual time segments and the frequency distribution of the signal can be seen as it changes in time. This gives a time-frequency distribution that can be used to both identify transients based on changing frequency characteristics of the signal and the location where these transients occur.

Unfortunately, there are several distinct drawbacks to the STFT. The size of the window used to compute the STFT is critical to the computed results. In the case of a small window, high frequency components can be localized in time with high precision; however, due to the few number of data points used in the calculation of the FT, the signal resolution of the low frequency components in the signal are lost. Similarly, with large windows low frequency components of the signal can be localized in frequency very well. However, transient signals with high frequency components will be represented as existing across the entire length of the window size being implemented. An example can be seen in Figure 10.

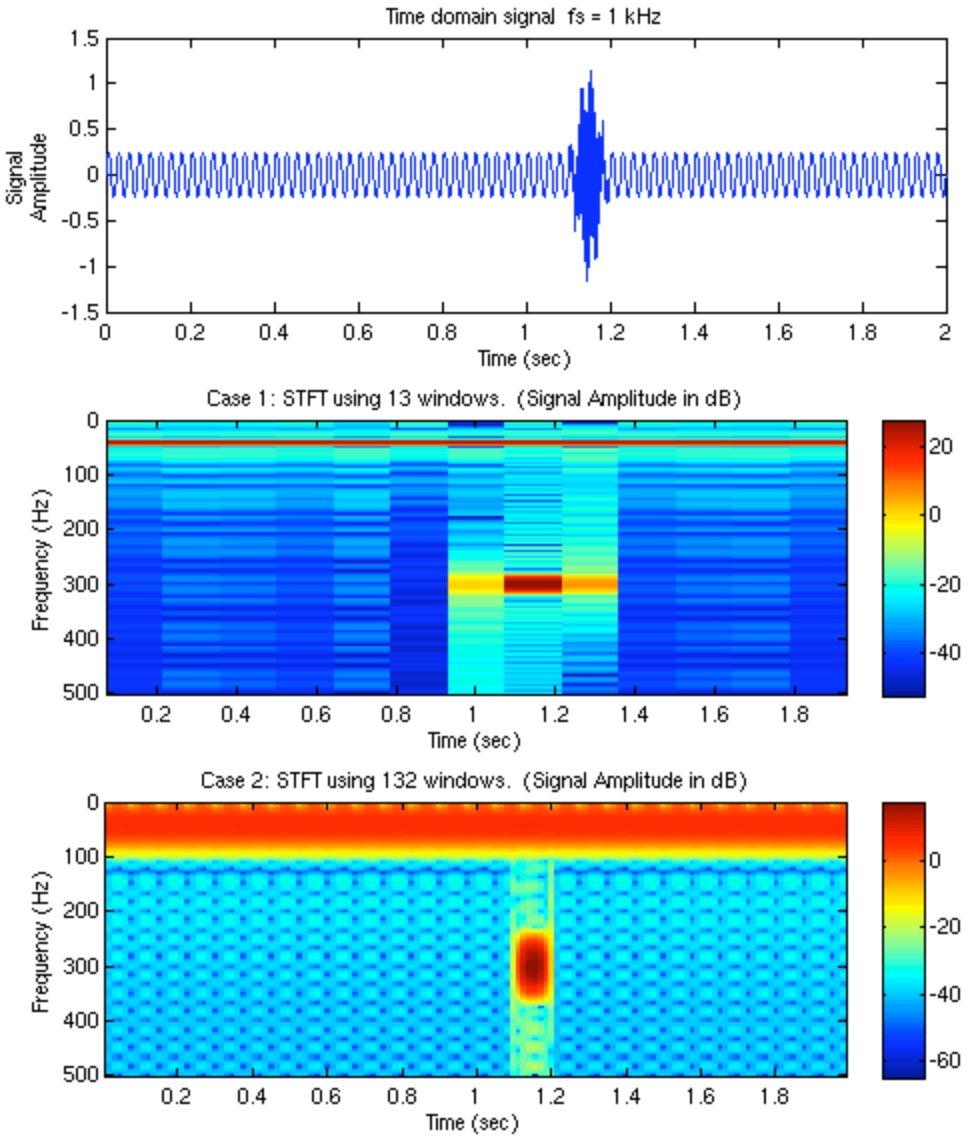

Figure 10: In this scenario a time domain signal, consisting of a 40 Hz sinusoid with an added transient 300 Hz pulse, has been transformed using the STFT. In Case 1, 13 windows over the length of the signal were used to compute the STFT. As can be seen there is low frequency resolution at the cost of time resolution. Case 2 demonstrates the opposite effect where a large number of windows were used to transform the signal. In case 2, it is much more clear where the transient is located at the cost of frequency resolution.

This relationship between localization in time and frequency estimation of the STFT has been compared to the Heisenberg Uncertainty Principle. By knowing exactly what frequency components are involved in the signal all time localization is lost. However, the closer frequency information is localized in time the less is known about the total content of the frequency information. The wavelet transform has the unique property of modifying its window size for each frequency band of interest. This is a phenomenal result, as it allows for a time-frequency distribution to be produced that does not have the drawback of attempting to select a window size, and the resulting time frequency tradeoffs as associated with the Short Time Fourier Transform.

2.4 Energy Distribution

One of the unique aspects of the Wavelet Transform is it's ability to conserve the energy in the original signal. In fact, for a given scale, *a*, and scale, *b*, the relative energy *E* of the signal can be determined as given in (11).

$$E(a,b) = \left|w(a,b)\right|^2 \tag{11}$$

$E(a,b)$ is known as a scalogram. This is an analogy to the STFT's denotation as a spectrogram. The energy of the scalogram can also be summed by scale, *a* as seen in (12) where C_g is called the admissibility constant and is given in (13).

$$E(a) = \frac{1}{C_g}\int_{-\infty}^{\infty}\left|w(a,b)\right|^2 db \tag{12}$$

$$C_g = \int_0^{\infty}\frac{\left|\hat{\psi}(f)\right|}{f}df \tag{13}$$

Finally the total energy of the signal can be determined from the scalogram by integrating the scalogram across scale as seen in (14).

$$E = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |w(a,b)|^2 \frac{da}{a^2} db \tag{14}$$

Because wavelets can be seen as band pass filters they also have defined center frequencies, as given in (15), defined as the second moment of the wavelets energy spectrum where f is frequency in hertz and $\hat{\psi}(f)$ is the Fourier transform of the wavelet.

$$f_c = \sqrt{\frac{\int_0^{\infty} f^2 \left|\hat{\psi}(f)\right|^2 df}{\int_0^{\infty} \left|\hat{\psi}(f)\right|^2 df}} \tag{15}$$

Using the Fourier Transform property given in (15) allows us to relate the effect of the scaling factor, *a*, to the center frequency of the wavelet. By effectively scaling the wavelet we are decreasing the wavelet's center frequency.

$$f(a) \propto \frac{f_c}{a} \tag{16}$$

This allows us to re-plot the scalogram in terms of frequency rather than scale and provide an analysis tool much more analogous to the STFT's spectrogram. An example is seen in Figure 11.

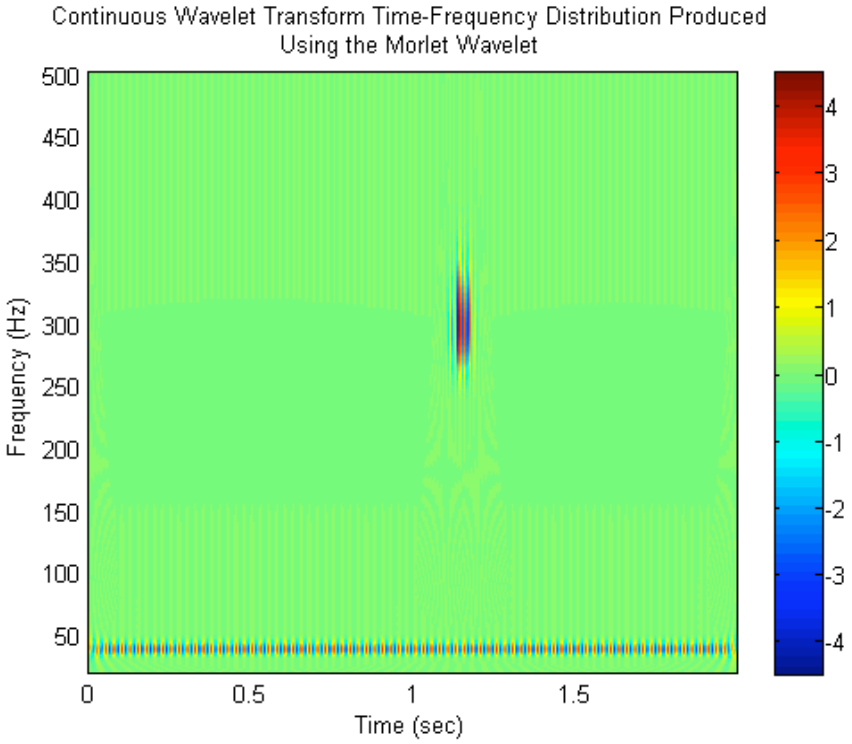

Figure 11: This is a time-frequency distribution produced by implementing the Continuous Wavelet Transform using the Morlet wavelet. The scale a can be recast in terms of the wavelet's center frequency f_c allowing for a time frequency distribution similar to that of the STFT's spectrogram.

2.5 Analytic Wavelets

There are also a number of nuances that can be implemented with the CWT that have novel properties. The examples so far in this document have used real wavelets that have acted as filters yielding real results. However by manipulating the real filter an imaginary filter can be created that splits the signal into amplitude and phase information. This is very analogous to the FT. Figure 12 shows the frequency domain representation

of the imaginary or analytical Morlet Wavelet compared to the Morlet Wavelet used in figure 10 above. As can be seen the analytic version of the Morlet Wavelet is a one-sided representation of the real wavelet.

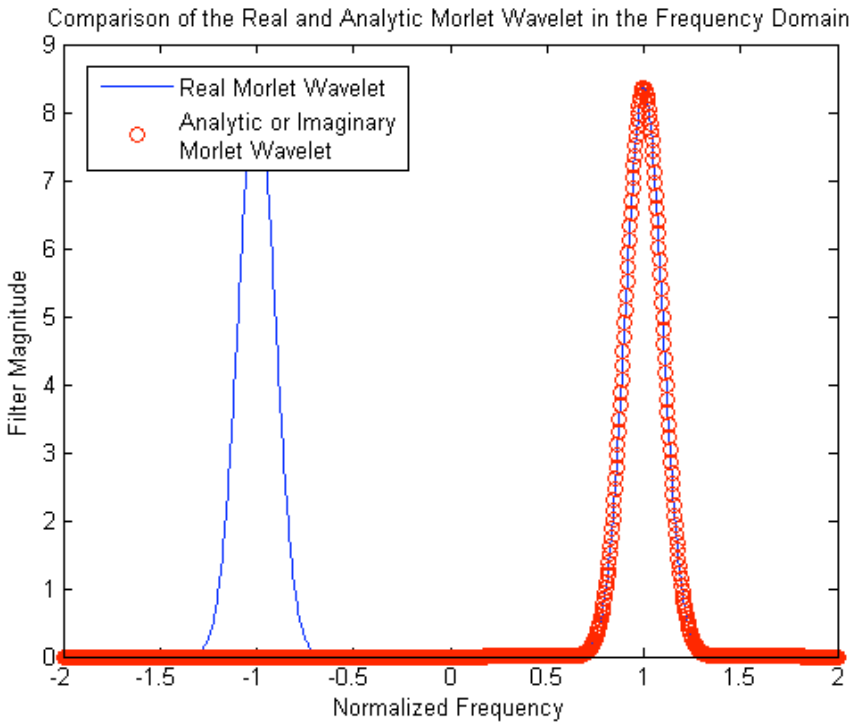

Figure 12: Magnitude of the frequency response of both the analytic and real Morlet wavelet for comparison.

The analytic wavelet provides another unique asset to the analyst. Figure 13 shows a 40 Hz sinusoid with a 300 Hz transient pulse. Scalograms from both the real and analytic Morlet wavelets computed for comparison can be seen in Figures 14, 15 and 16.

As can be seen, because the analytic wavelet is imaginary the results from the CWT using this wavelet be represented as magnitude and phase information as would be done with the results from the STFT.

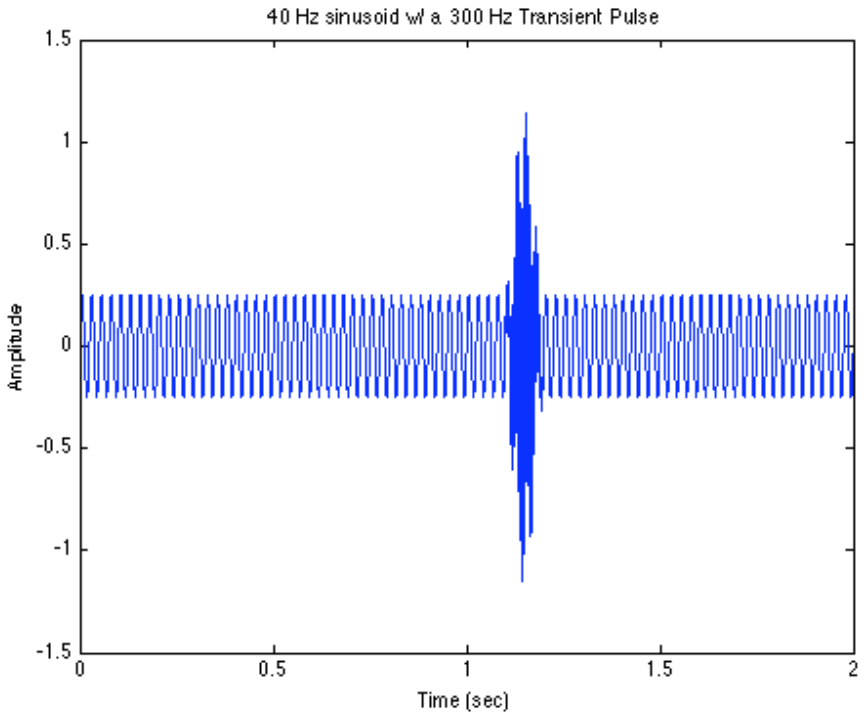

Figure 13: A 40 Hz sinusoid with an imbedded 300 Hz pulse.

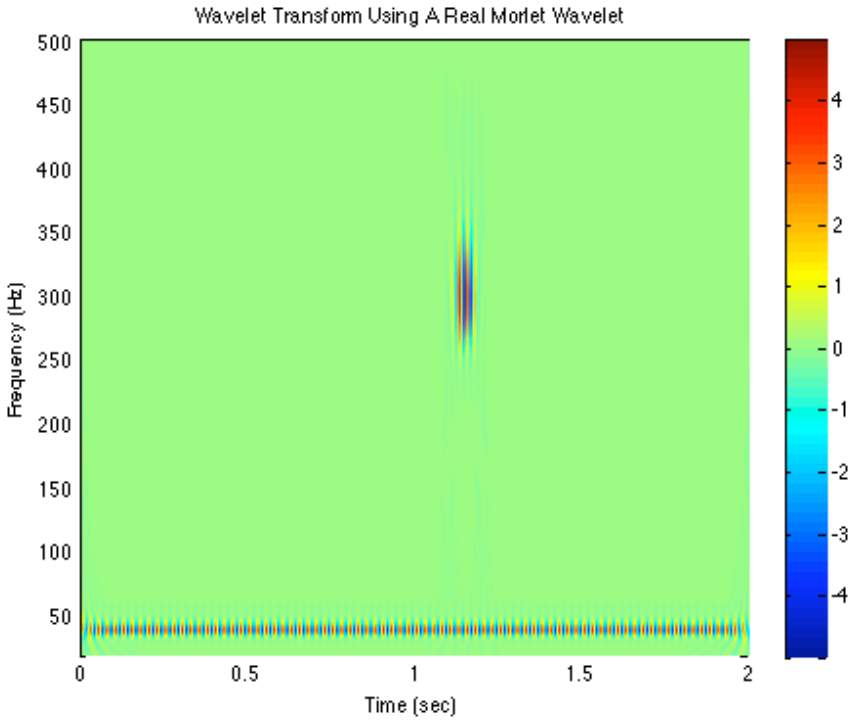

Figure 14: Wavelet scalogram using a real Morlet wavelet of the signal seen in Figure 13.

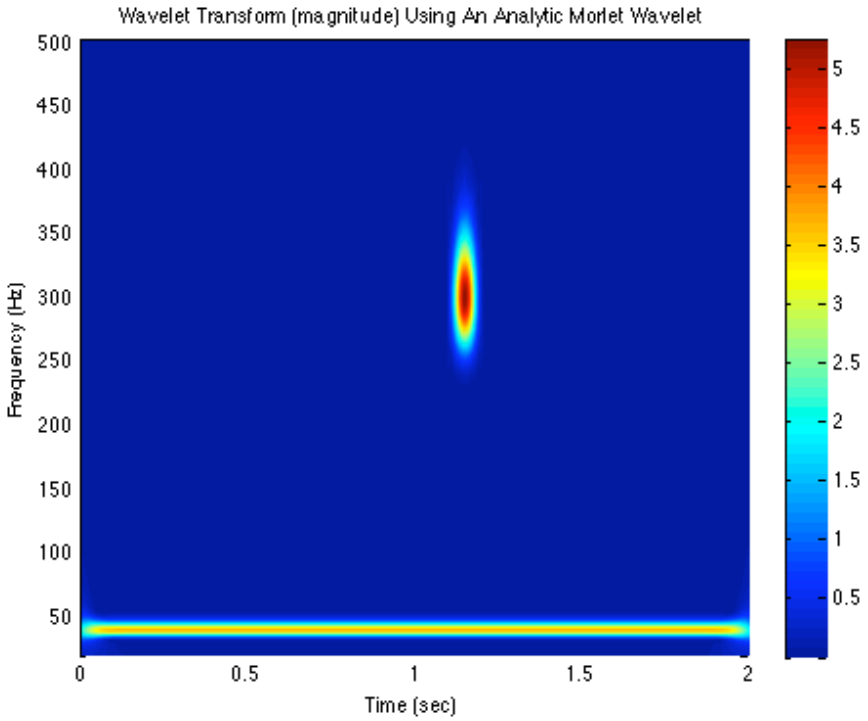

Figure 15: Magintude scalogram using an analytic Morlet wavelet of the signal seen in Figure 13.

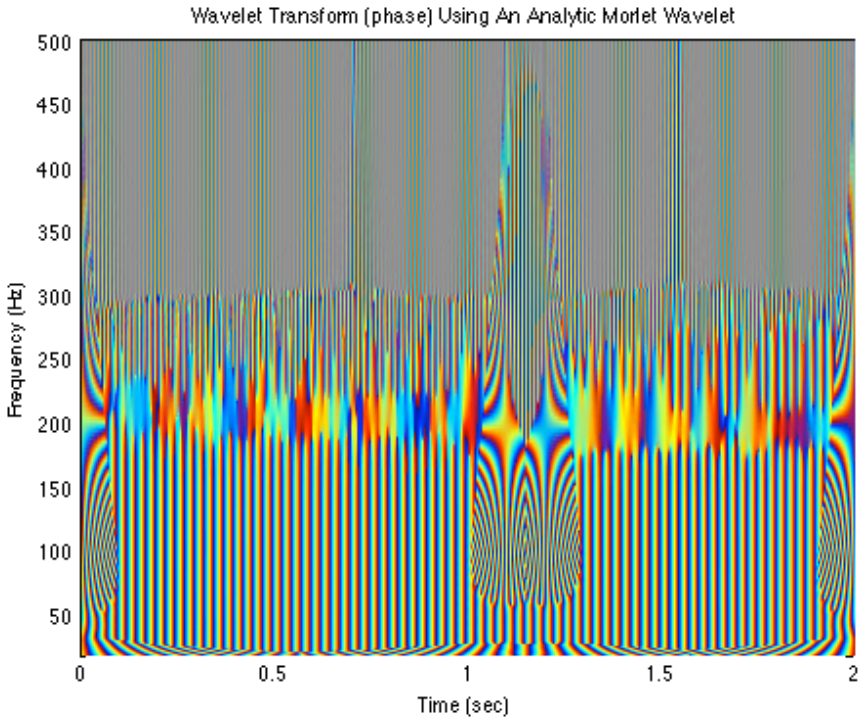

Figure 16: Phase scalogram using an analytic Morlet wavelet of the signal seen in Figure 13.

2.6 Inverse Continuous Wavelet Transform

The CWT, unlike the STFT, is also invertible and as a result not only can a signal be displayed in a time-frequency plot, but it can be manipulated in the time-frequency domain and reconstructed back into an approximation, $\hat{x}(t)$, of the original signal as given with (17) where C_g is the admissibility criterion given in (13).

$$\hat{x}(t) = \frac{1}{C_g}\int_{-\infty}^{\infty}\int_{0}^{\infty} w(a,b)\psi(t)\frac{da}{a^2}db \tag{17}$$

This inverse transform is only perfect reconstruction of the original signal when the integrals given in (17) can fully be evaluated. In practical implementations, for

example with a digital computer, the integrals in (17) are in effect summations and it is impossible to perform an exact reconstruction of the signal. Also, due to the bandpass nature of a wavelet it is impossible to capture the DC component of the signal as with a Fourier Transform. For some wavelets, not all, scaling functions can exist that capture the DC components of the signal. Wavelet transforms exist that, based on the CWT, provide orthogonal expansions of the original signal and ensure that the inverse transform is a perfect reconstruction. To do this the CWT gets modified to a discrete summation where the coefficients that define the transform, *a* and *b*, become discretely sampled values as seen in (18).

$$w(m,n) = \int_{-\infty}^{\infty} x(t)\psi\left(a_0^{-m}t - nb_0\right)\Big/ a_0^{m/2}dt \tag{18}$$

In the development of the CWT into an orthogonal transform, the Discrete Wavelet Transform (DWT), there are drawbacks. The CWT produces a very redundant transform. Redundancy is not always a drawback in some processing applications, as with detection of bearing fault signals. With redundancy, there is ambiguity involved in selecting both the maximum scale that the signal should be processed at as well as the fineness of divisions between successive applied scales. Furthermore if wavelet scales are not chosen without forethought, information can be lost from the signal. This would negatively impact any analysis as well as any inverse transformation. The majority of wavelets that exist for the CWT do not have a scaling function. As a result the DC component of the signal cannot be removed and added back to the signal approximation.

For applications where there is no DC component to the signal or it is negligible, as with acceleration or velocity signals, the inverse CWT is acceptable.

The CWT, depending on implementation, can also take significant time compared to calculating a standard Fourier transform or, as a better comparison, an STFT. At the cost of redundancy the DWT has been developed that is much faster and orthogonal, inherently providing a scaling filter. There also exists, using both the CWT and DWT, a plethora of different wavelets that can be chosen to process the signal. This is both a boon and a pitfall, as a good engineer attempting to perform signal processing with a wavelet must know something about the signal that will be processed to select a wavelet that will produce a good time-frequency distribution. This background has been a very brief introduction to wavelets. For a much more in depth review of the CWT, as well as the DWT, wavelet theory the reader is directed to [55, 57-59].

3. PROPOSED ALGORITHMS FROM LITERATURE

This section will highlight seven different algorithms from literature that will be used to study the effectiveness of the CWT in de-noising rolling element bearing faults. The first algorithm, envelope demodulation, is the industry adopted standard for bearing fault detection and was originally implemented using analog hardware [4]. More recently with the advent of the digital computer and low cost microprocessors this method has been implemented in digital computers. This thesis uses a digital implementation of envelope demodulation that models traditional analog hardware designs. This algorithm was implemented to provide a base line comparison for all other algorithms introduced in this thesis.

The second algorithm, Wavelet Based Weak Signature Detection, optimizes the wavelet transform to detect a bearing fault in a given signal and selects as single scale as an optimal filter [18]. The third and fourth de-noising methods, hard and soft thresholding were the first and most common forms of wavelet de-noising and have been proposed for use in bearing fault detection [10]. These forms of de-noising are fast and very effective means of removing noise from a signal. More complicated methods for de-noising bearing fault signals, using wavelets, must not only perform better than hard and soft thresholding, but also perform better than the standard envelope demodulation technique used by industry. The fifth and sixth algorithms, Hybrid thresholding and Mechanical thresholding introduced in this section have been suggested as better de-noising algorithms for bearing faults and are based of similar de-noising strategies as hard and soft thresholding [44, 50]. The seventh and final algorithm, Minimum Entropy

Deconvolution, is a new method of de-noising bearing fault signals that does not use the wavelet transform. [60] This last algorithm is being introduced to provide a comparison between CWT de-noising algorithms and alternate de-noising methods.

3.1 Envelope Demodulation

As previously mentioned, bearing fault impulses excite or resonate the natural frequency of one or more of the components of the mechanical structure of the bearing, bearing housing or the vibration sensor that is being used to detect the signal. As a result, in the frequency domain, most of the fault signal energy is located at this resonate frequency. Envelope Demodulation, the traditional detection method, takes advantage of this fact by applying a band pass filter at the resonance frequency. This is intended to remove all non-relevant information from the signal and, by application of demodulation, provide a base-banded representation of the fault.

This study will be using a digital implementation of envelope detection using guidelines developed by McFadden [4] for specific application to bearing fault detection. This detection method was developed prior to the explosion in the use of digital signal processing techniques and low cost microprocessors that transformed electronics. As a result, this detection method can be implemented using very low cost analog electronic components. In an effort to remain true to this commonly implemented method a more traditional digital signal processing approach was used that is comparable to analog detection methods. This is very similar to a more advanced, methodology was developed by Sheen [61].

An IIR Butterworth band pass filter was created, and the signal once filtered was demodulated using the Hilbert transform. There are a number of different parameters that go into the filter design. McFadden specifies that the filter bandwidth to be approximately four times the highest cyclic frequency however does not comment about the filter roll off or filter type except to mention that the most common filter used in practice is the Butterworth filter [4]. For this study a filter order of 2 was implemented to best simulate analog hardware that would be implemented in practice.

A plot of the band pass filter used in this study, assuming a 12 kHz sampling frequency, an arbitrary center frequency of 4 kHz and a bandwidth of 400 Hz is shown in Figure 17.

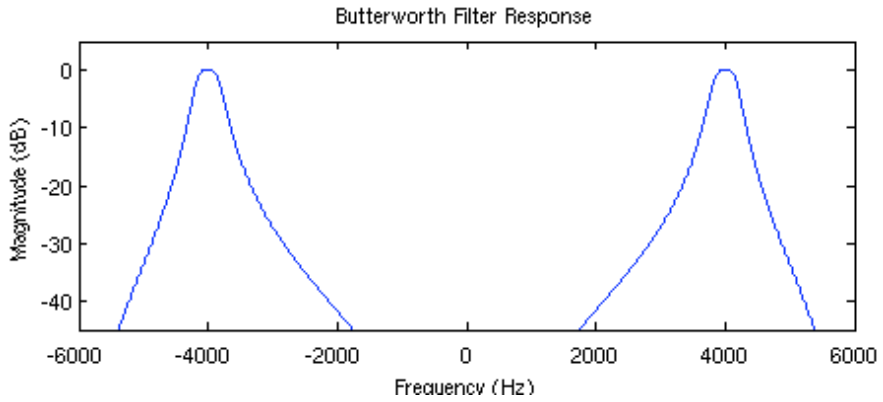

Figure 17: Frequency response of the digital IIR Butterworth band pass filter used in this study. The filter was created assuming a sampling frequency of 12 kHz, an arbitrary center frequency of 4 kHz and a bandwidth of 400 Hz.

A vibration signal with a hidden inner race fault, showing in Figure 18, was taken, filtered and demodulated. Periodic impulses from the bearing fault, with a cyclic frequency of approximately 147 Hz, are clearly identifiable after demodulation.

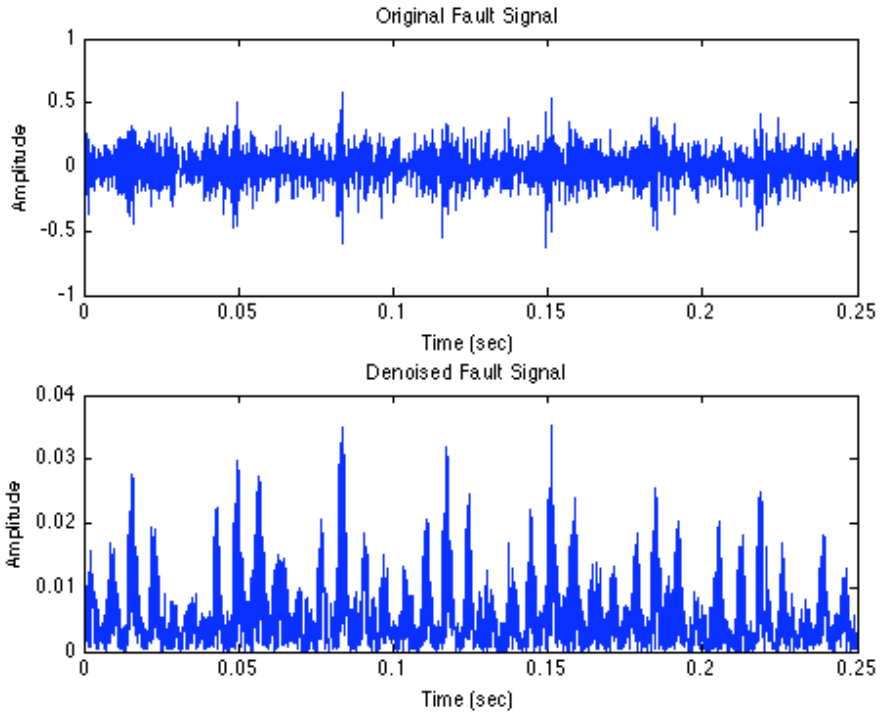

Figure 18: This figure shows the original signal prior to **(top)** and after **(bottom)** envelope demodulation. Note: Impulsive structures in the original signal are due to shaft vibrations and are harmonics of the running speed ~29.8 Hz.

For the example given in Figure 18, the filter bandwidth was set to be 500 Hz. This is approximately three and a half times the cyclic frequency of the bearing fault, 147.7 Hz. One of the central difficulties in using envelope demodulation is determining the natural frequency of the bearing. If the natural frequency of the bearing is known, this sets the center frequency of the Butterworth filter, however no reliable method exists to assign the filter bandwidth. Obviously, the filter must be large enough to detect the first harmonic of the fault and the more harmonics of the fault that are detected the better. However, above a limit and too much noise is being admitted, raising the noise floor, and reducing the amplitude of any detected harmonics.

For this research, the natural frequency of the bearing was identified using the cyclic frequency analysis tool, the CWT/DRFS algorithm, that will be introduced later in this thesis. The output of the CWT/DRFS algorithm can be seen in Figure 19.

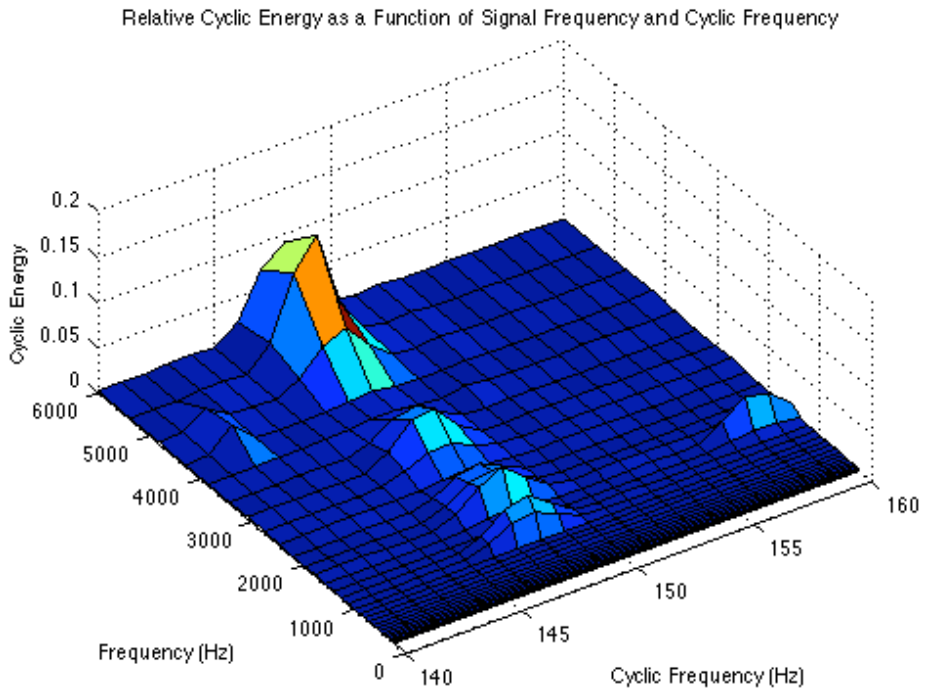

Figure 19: Output from the CWT/DRFS algorithm that can identify both the natural frequency of the bearing and the cyclic frequency of the fault.

As can be seen in Figure 19, the natural frequency of the bearing, ~5.5 kHz and the cyclic frequency of the fault, ~146 Hz are clearly identified. For the example seen in figure 18, the filter bandwidth was arbitrarily set at 3.5 times the detected cyclic frequency or ~500 Hz.

3.2 Wavelet Filter Based Weak Signature Detection Method

The Wavelet Filter Based Weak Signature Detection Method is a technique that uses the wavelet transform to design an optimal filter for bearing fault detection [18].

One of the central difficulties in envelope demodulation is not only detecting but filtering the fault signal at the resonate frequency at which the fault of interest is located. This technique uses the wavelet transform and the Morlet wavelet to isolates the fault signal.

The first step in the algorithm is to identify the optimal shape factor of the Morlet wavelet that captures the fault signal. The Morlet wavelet used in this study is given in (19).

$$\psi(t) = \frac{1}{\sqrt{2\pi}} e^{\left(-\beta^2 t^2/2\right)} \cos(\pi t) \tag{19}$$

This wavelet has a shape factor β, that can be used to balance the wavelet's shape both in the time and frequency domain. The shape factor can then be used to modify the Morlet wavelet to design an optimal filter shape for the fault signal. The shape factor is selected by using Shannon Entropy Measure (SEM) defined in (20) [57].

$$E_S(p) = -\sum_i p_i \log_2(p_i) \tag{20}$$

SEM was originally developed to characterize communication signals. SEM is a measure of the uncertainty inherent in each sample of a signal. The more random the signal the higher the value of entropy the SEM algorithm will yield. A good example of this can be seen in a perfect digital signal with values only at zero and one. In the case of a perfect digital signal equation (20) will yield zero entropy. As the signal becomes more random, elements of the signal start to deviate from zero or one and the SEM algorithm will begin to give a positive value of entropy indicating information in the signal is being lost. By repeatedly performing the CWT using the Morlet wavelet and different values of β, a shape factor can be selected that minimizes entropy over the resulting scalogram.

Once the shape factor has been selected the wavelet scale that best captures the fault signal can then be found. This is done using singular value decomposition (SVD). The SVD of a matrix D is defined by

$$D = UEV^T \tag{21}$$

where U is an m x m square matrix and V is an n x n square matrix such that

$U^TU = 1$ and $V^TV = 1$.

The matrix E is an m x n diagonal matrix with diagonal elements σ_p with p = min(m,n) arranged such that $\sigma_1 > \sigma_2 > \sigma_3 >$ etc. If D is of rank g, having only g independent column vectors, there will be only g values of σ. Remarkably enough this can be used to detect periodicity in a signal. Consider a periodic signal $X = [x_1\ x_2\ x_3\ \ldots\ x_n]$ with a period of length n. The signal can be partitioned into a matrix with each row representing one period of the signal as seen in (22).

$$X = \begin{bmatrix} x(1) & \ldots & x(n) \\ x(n+1) & \ldots & x(2n) \\ \vdots & \ddots & \vdots \\ x((m-1)n+1) & \ldots & x(mn) \end{bmatrix} \tag{22}$$

If the matrix X is periodic with period n, and there is no noise associated with the signal, the matrix is of rank 1 and consequently, after computing the SVD of matrix X, there is only one value of σ. If there is noise present in the signal, the signal is full rank, but dominated by the periodic structure of the signal. For this case, σ_1 would be very large compared to σ_2. An estimate of how periodic the signal is, assuming a period of n, can therefore be created using the ratio given by (23).

$$\delta = \left(\frac{\sigma_1}{\sigma_2} \right)^2 \tag{23}$$

Each wavelet scale can then be evaluated for all possible values of periodicity *n*. The maximum value given by (9) for all wavelet scales for all assumed periodicities will yield both the period of the fault and indicate the scale that best captures the fault signal. This wavelet scale is then taken as the de-noised signal.

This algorithm can best be understood by seeing an example. Figure 20 gives a fault signal covered by noise seen in the time domain. As can be seen the fault signal is not readily apparent in the time domain.

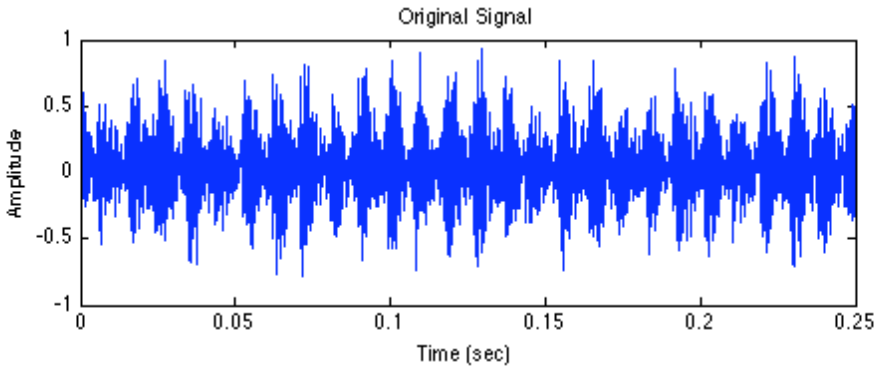

Figure 20: An outer race fault signal covered with noise.

The wavelet transform of this signal was taken for various values of β and entropy calculated for each transform. A plot of the entropy with respect to shape factor can be seen in Figure 21.

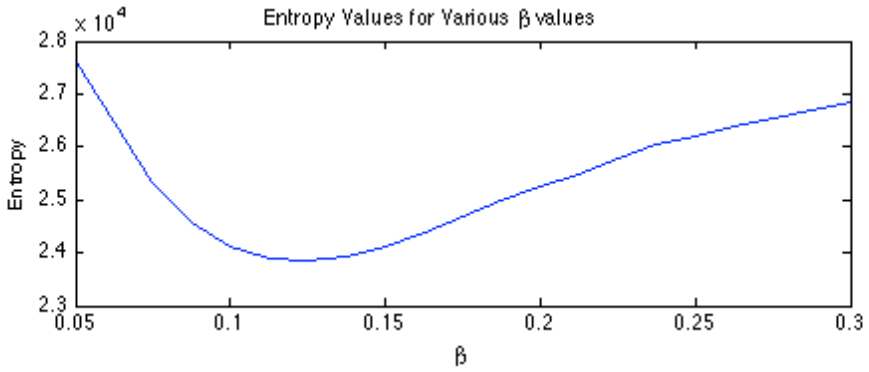

Figure 21: This plot gives calculated entropy values for various wavelet β values.

The wavelet transform of the fault signal was then generated using β equals ~0.27, the minimum shape factor. The periodicity ratio, δ, seen in (23) was then calculated for each scale for all possible values of period n. The result can be seen in Figure 22. The maximum value for δ, 59.4, occurs at scale 1.72 or 3483 Hz. This scale is used as the de-noised signal and can be seen in Figure 23.

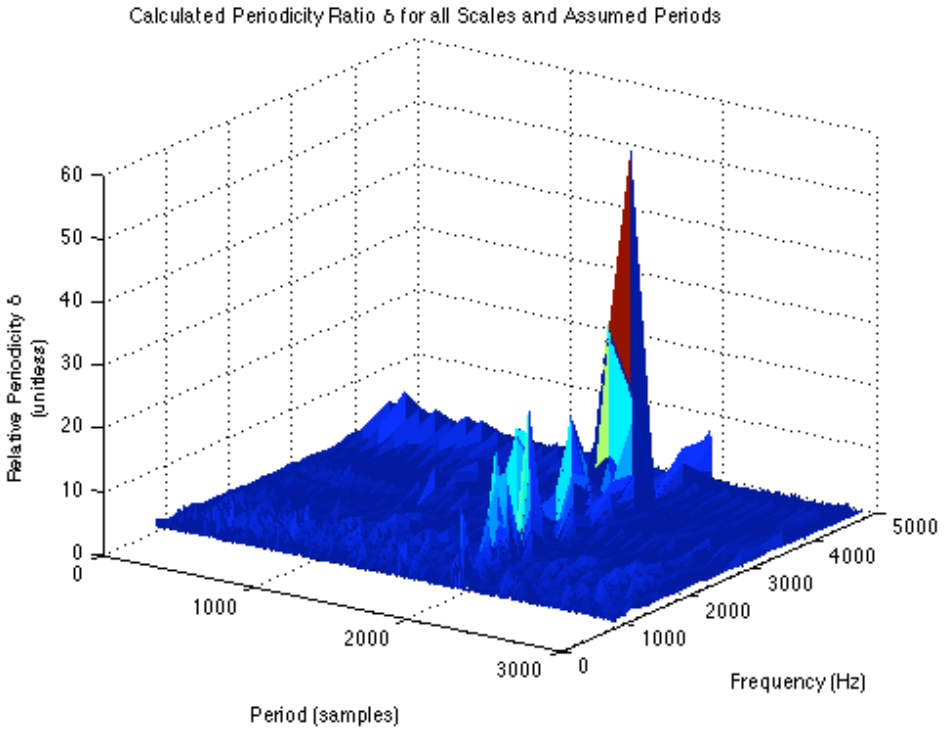

Figure 22: This image gives the value of δ for each scale and for various assumed periods. The maximum value of δ, 59.4, occurs at 3.5 kHz.

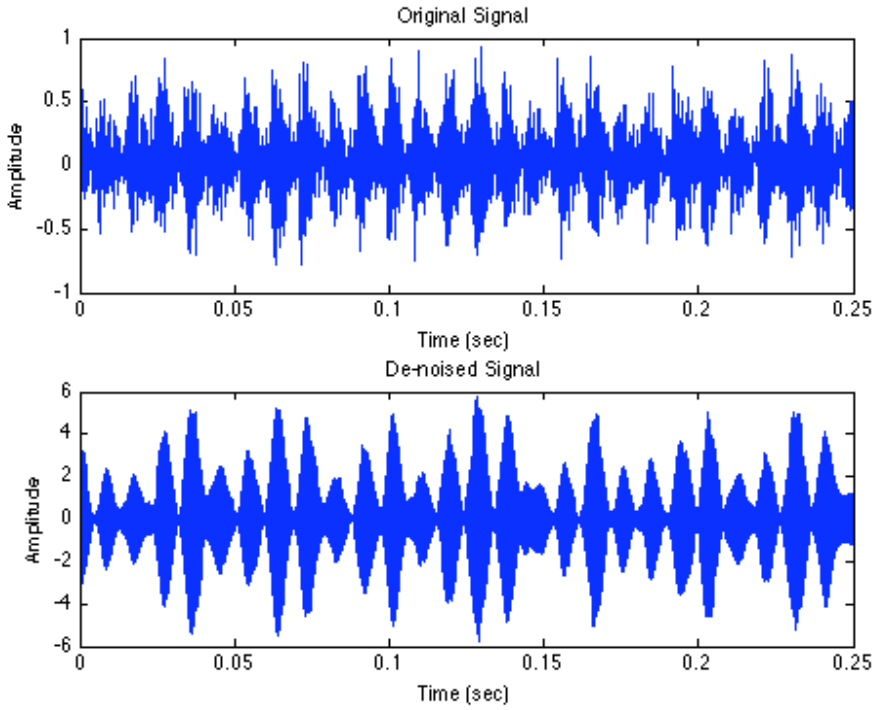

Figure 23: This figure gives the de-noised fault signal **(top)** using the Wavelet Filter Based Weak Signature Detection Method. Also for comparison the original fault signal is given **(bottom)**.

3.3 Hard and Soft Thresholding Using the Continuous Wavelet Transform

Hard and soft thresholding are a very simple procedures that uniquely filter a signal by modifying the signal in the wavelet domain and performing the inverse transform. Soft thresholding, in particular, using the Morlet wavelet has been shown to be effective in de-noising bearing faults [10]. The most prevalent example is that of a signal corrupted with white noise. In the time domain the signal might be filtered using a non-linear median filter or averaging filter. These filters would attempt to smooth the signal reducing the irregularities produced by the added white noise. A frequency domain

approach might be an attempt to design a filter that would extract the dominant portions of the signal. Both of these cases would not be optimum. For the time domain median filter, irregularities would still exist and due to the non-linear response of the filter the results would be highly dependent on the filter length and the characteristics of the signal of interest. Similarly, the frequency domain approach would, while extracting the dominant components of the signal, still keep significant portions of the noise that were centered about the dominant frequency components of the signal.

Hard and soft thresholding apply a filtering technique in a time-frequency distribution. When the wavelet transform is applied to the signal, white noise will be spread across all frequencies while portions of the signal that correlate with the wavelet, will be amplified. An example of this can be seen using the bearing fault signal seen in Figure 24 where a periodic bearing fault is dominating the signal however is corrupted by white noise.

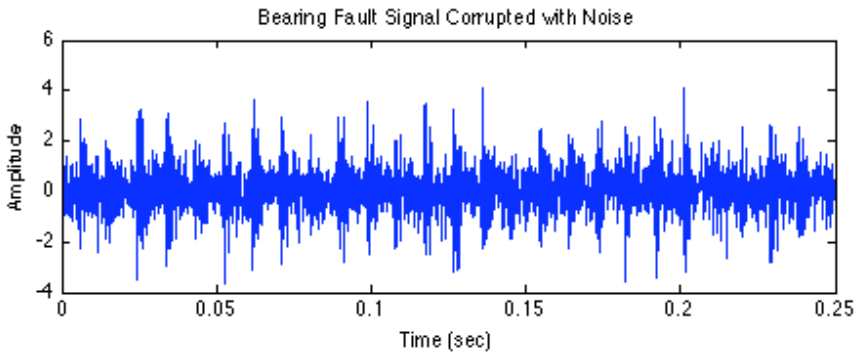

Figure 24: This is an outer race bearing fault signal corrupted by white noise.

The wavelet transform of the fault signal in Figure 24 can be seen in Figure 25.

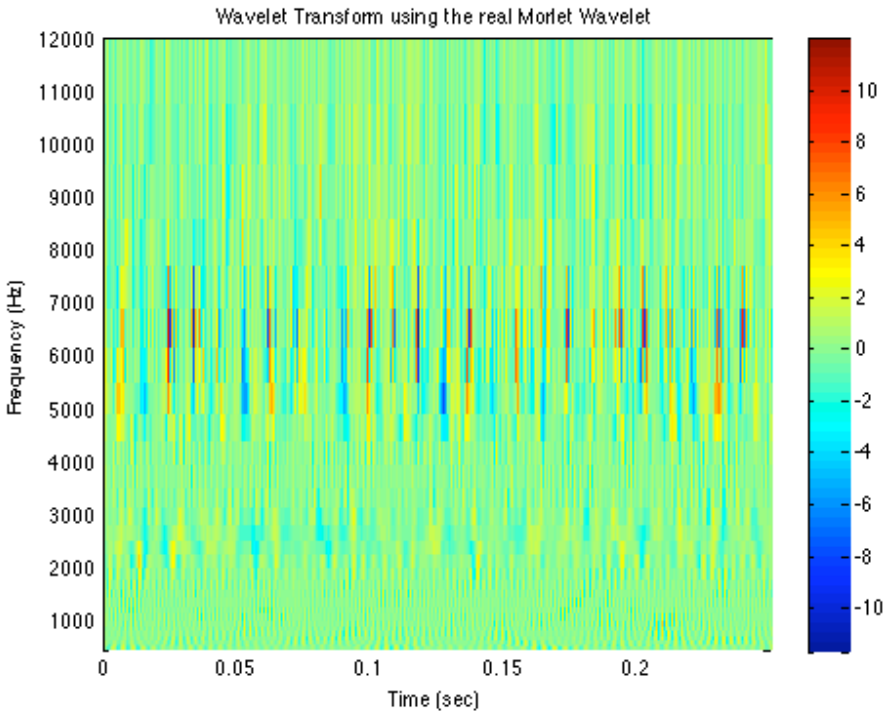

Figure 25: The image above is the wavelet transform of the fault signal seen in Figure 24.

By zeroing values of the wavelet transform that have an absolute value less than a given thresholding value, λ, noise can be removed from the signal with relatively little impact on the dominate components of the signal. This is known as hard thresholding and is given in (24).

$$W_{hard}(a,b) = \begin{cases} 0 & |W(a,b)| < \lambda \\ W(a,b) & |W(a,b)| \geq \lambda \end{cases} \tag{24}$$

where $W(a,b)$ are the wavelet transform coefficients at scale a and time b. A graphical illustration of hard the hard thresholding rule is given in Figure 26 with λ arbitrarily set to 0.1.

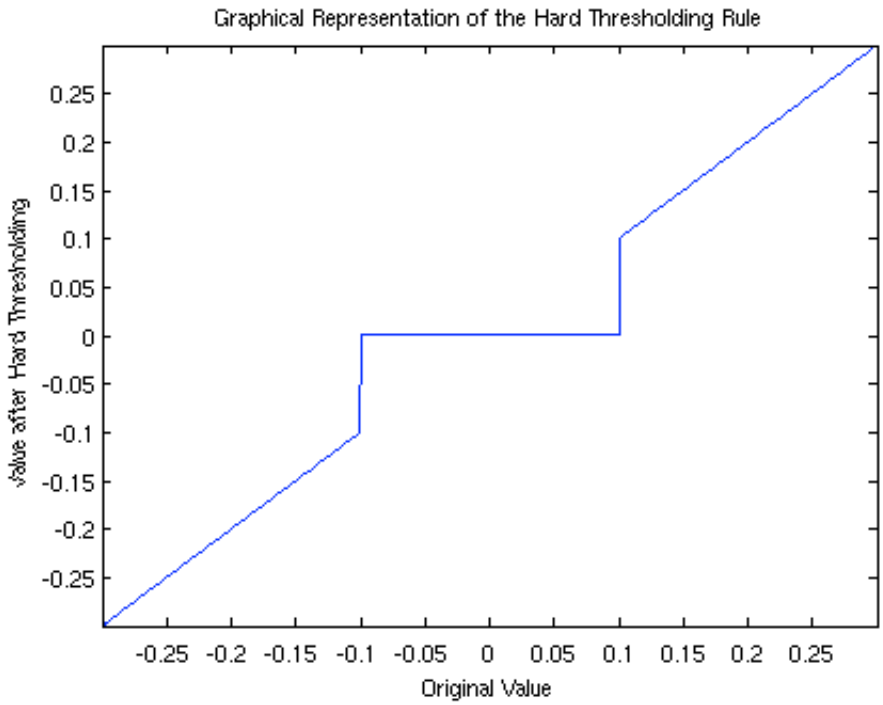

Figure 26: This is a graphical representation of the hard thresholding rule given in eq. (9). All values below the threshold value, for this example arbitrarily set to 0.1, map to 0, while all other values remain unchanged.

Figure 27 gives the de-noised wavelet transform using the hard threshold rule in (9) with a value of 2.91 for λ. Figure 28 gives the inverse wavelet transform of the de-noised wavelet scalogram seen in Figure 27.

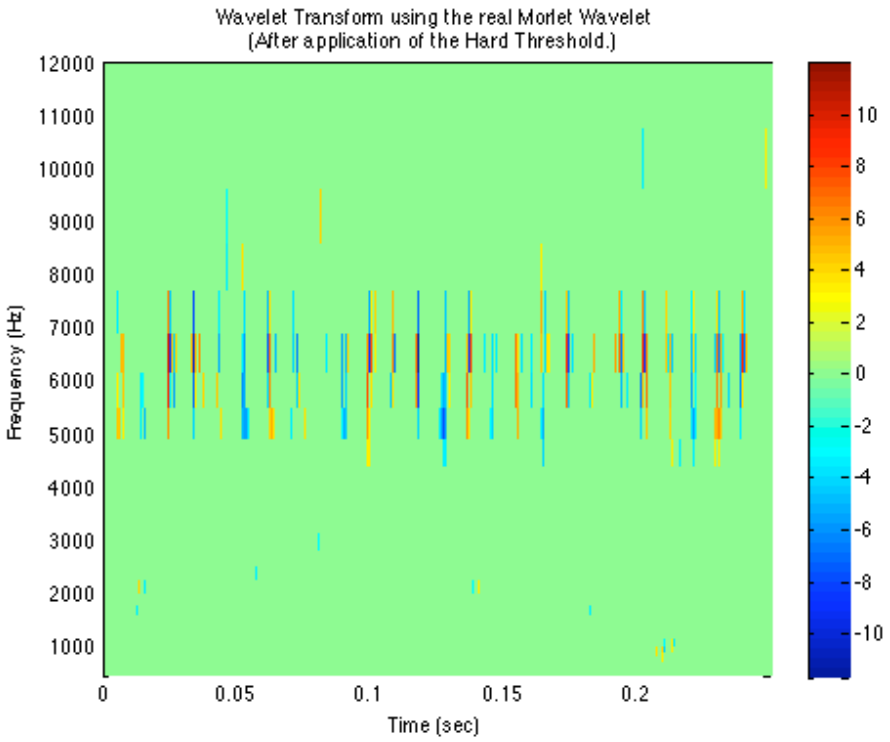

Figure 27: This is the wavelet transform of the fault signal in Figure 24 after hard thresholding has been applied.

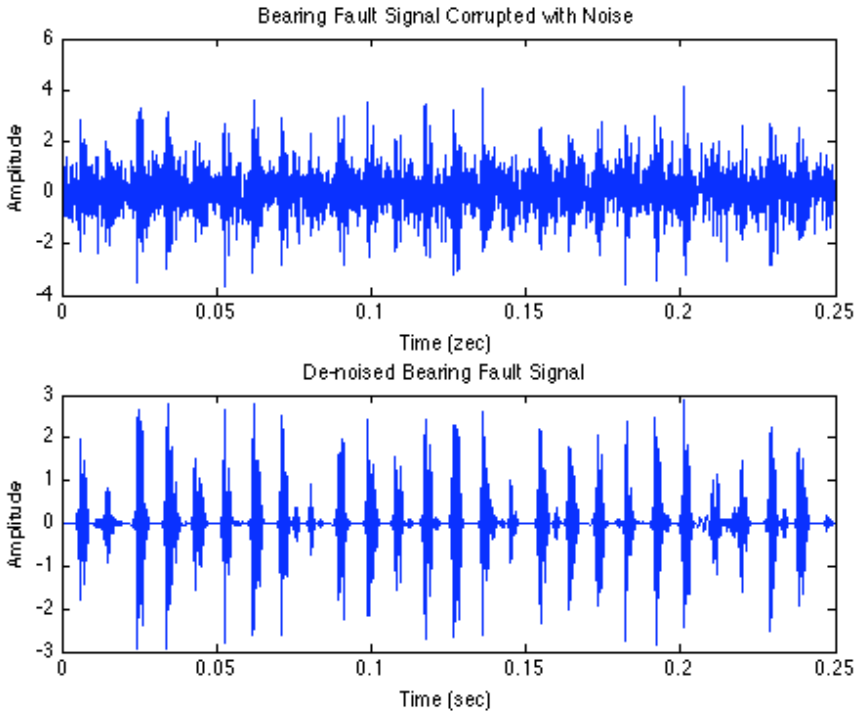

Figure 28: This figure shows the original fault signal **(top)** and the de-noised fault signal **(bottom)** using wavelet hard thresholding.

Soft thresholding is a very similar concept to hard thresholding [45]. Instead of zeroing all values below a threshold, however, values are selectively repressed as seen in (25).

$$W_{soft}(a,b) = \begin{cases} 0 & |W(a,b)| < \lambda \\ sign\big(W(a,b)\big)\big(|W(a,b)| - \lambda\big) & |W(a,b)| \geq \lambda \end{cases} \tag{25}$$

A graphical illustration of hard the hard thresholding rule is given in Figure 29.

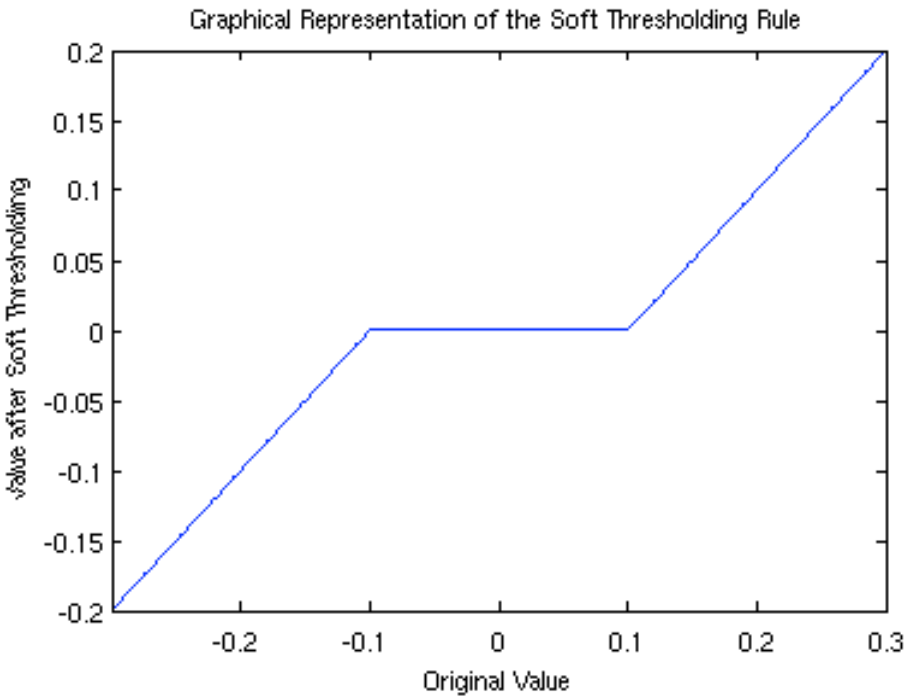

Figure 29: This is a graphical representation of the soft thresholding rule given in (10). All values below the threshold value, for this example arbitrarily set to 0.1, map to 0, while all other values are reduced by the threshold.

There are several limitations to hard and soft thresholding. First and foremost the signal that is being de-noised has to be the dominant characteristic component of the signal. If the signal is not strong enough, thresholding, can and will, remove components of the original signal. Similarly, if the noise added to the signal is not white and is concentrated in a bandwidth, it will not get removed from the signal. It is known that hard thresholding, due to the sharp transition that occurs at the threshold cut off can induce discontinuities in the de-noised signal. Soft thresholding on the other hand has the tendency to smooth the signal. This can lead to problems in the case where

discontinuities are present in the signal that is being de-noised. When the signal that is attempting to be de-noised is that of discontinuities, such as with bearing faults, both hard and soft thresholding produce undesirable effects [50].

It can also be difficult to select the level of the threshold. For this study the universal threshold, given in (26), was selected because of its popularity.

$$\lambda_U = \left(2\ln N\right)^{1/2}\sigma \tag{26}$$

where σ is the standard deviation of the noise in the signal. For this study a robust estimate of σ was used by setting σ to the median of the absolute deviation (MAD) of the wavelets smallest scale divided by 0.6745. The division by 0.6745 is a calibration term to normalize σ to that of a Gaussian distribution. The universal threshold then becomes (27) as seen below.

$$\lambda_U = \left(2\ln N\right)^{1/2}\frac{MAD}{0.6745} \tag{27}$$

It is important to note that there are a very wide variety of threholding techniques and criterion for threshold selection. It has been show that these de-noising techniques are not optimum for bearing fault detection [50]. For this study, hard and soft thresholding are simply being used as measures for comparison and as a result a much more in depth review of thresholding has not been presented.

Typically de-noising algorithms are employed with orthogonal and bi-orthogonal wavelet transforms and as a result the use of thresholding with the CWT is unusual. One of the advantages of the CWT compared to other wavelet transform methods its flexibility in wavelet selection. Wavelet choice with orthogonal and bi-orthogonal

wavelets is much more limited. Similarly, bearing faults typically are very faint signals and often because of their limited bandwidth and low energy lost in orthogonal and bi-orthogonal transforms. De-noising using orthogonal transforms helps smooth a signal. Signals with spikes, as seen with rolling element bearing faults, tend to be repressed as noise [44]. With that said, there are several different wavelet de-noising algorithms specifically designed using orthogonal and bi-orthogonal wavelets that have been used for bearing fault de-noising [57, 62].

3.4 The K-Hybrid Threshold Using the Continuous Wavelet Transform

The K-Hybrid Threshold is an alternative to hard or soft thresholding of the wavelet transform [50]. K-Hybrid Thresholding uses a combination of hard and soft thresholding with independent thresholds set for each wavelet scale based on kurtosis value. The CWT of the fault signal is computed and the kurtosis value of each scale is calculated using (28) where x is the signal of interest and $\bar{x}$ is the mean of x.

$$Kurt(x) = \frac{\sum_{n=1}^{i}(x-\bar{x})^4}{\left(\sum_{n=1}^{i}(x-\bar{x})^2\right)^2} \tag{28}$$

Kurtosis is a measure of 'spikiness' of a signal and it is defined as the fourth standardized moment. A signal with large but infrequent deviations from the mean will have higher kurtosis than that of a signal with the same variance but with more frequent deviations that are less extreme from the mean. For a more detailed review of kurtosis the reader is referred to [63]. For signal of perfect Gaussian distributed noise it has been

shown that the kurtosis calculation yields a value of three. (For simplicity of notation, W_a is defined as the scale a of the wavelet transform $W(a,b)$.) This allows each scale to be classified using the following three distinctions:

- super Gaussian $Kurt_a(W_a) > 3$: The noise shows characteristics of impulsiveness and the threshold is set to allow components of the scale to pass.
- Gaussian $Kurt_a(W_a) = 3$: The signal is dominated by Gaussian noise and as a result no portion of the wavelet scale is admitted.
- sub-Gaussian $Kurt_a(W_a) < 3$: The signal shows no signs of impulsiveness and no portion of the signal should be kept.

A threshold t_{2a} can then be defined to allow only impulsive components of the signal if the signal is super Guassian via (29).

$$t_{2a} = \frac{\max(W_a)}{Kurt(W_a)/3} \tag{29}$$

t_{2a} is defined as the noise floor for the wavelet scale a where everything above t_{2a} can be admitted as components of the fault signal and everything below is considered noise and removed. The choice therefore exists to apply either a hard or soft threshold. If the noise is constant across all scales, for all time and purely Gaussian, hard thresholding is be the optimum way to remove the noise. If the maximum value of kurtosis for all wavelet scales is approximately three, the signal is mostly white noise, and a hard threshold can be applied to the majority of the signal across all scales. However as the maximum value of kurtosis, evaluating all wavelet scales, increases, the value of the hard threshold needs to decrease and soft thresholding applied in its place. The value of t_{1a} representing the

hard threshold level at scale a, and the value of t_{3a}, representing the upper boundary of the soft threshold are given in (30), (31) and (32).

$$k_a = 1 + \max\left\{Kurt\left(W_a\right) - 3, 0\right\} \tag{30}$$

$$t_{1a} = \frac{t_{2a}}{k_a} \tag{31}$$

$$t_{3a} = \left(1 + \frac{1}{\sqrt{2}k_a}\right)t_{2a} \tag{32}$$

The threshold rule is given in (33), (34) and (35) where

$$R_{1a} = \left(1 - \frac{1}{k_a}\right)t_{2a} \tag{33}$$

$$R_{1a} = \left(\frac{1+\sqrt{2}}{k_a}\right)t_{2a} \tag{34}$$

$$W_{threshold} = \begin{cases} 0 & \left|W_a\right| < t_{1a} \\ sign\left(W_a\right)\left(R_{1a} - \sqrt{R_{1a}^2 - t_{1a}^2 + W_a\left(2t_{1a} - \left|W_a\right|\right)}\right) & t_{1a} \le \left|W_a\right| < t_{2a} \\ sign\left(W_a\right)\left(R_{1a} + \sqrt{2\left(R_{2a} + t_{2a}\right)W_a - \left|W_a\right|^2 - t_{2a}\left(2R_{2a} + t_{2a}\right)}\right) & t_{2a} \le \left|W_a\right| < t_{3a} \\ W_a & t_{3a} \le \left|W_a\right| \end{cases} \tag{35}$$

The threshold given in (35) is plotted for three values of kurtosis in Figure 30.

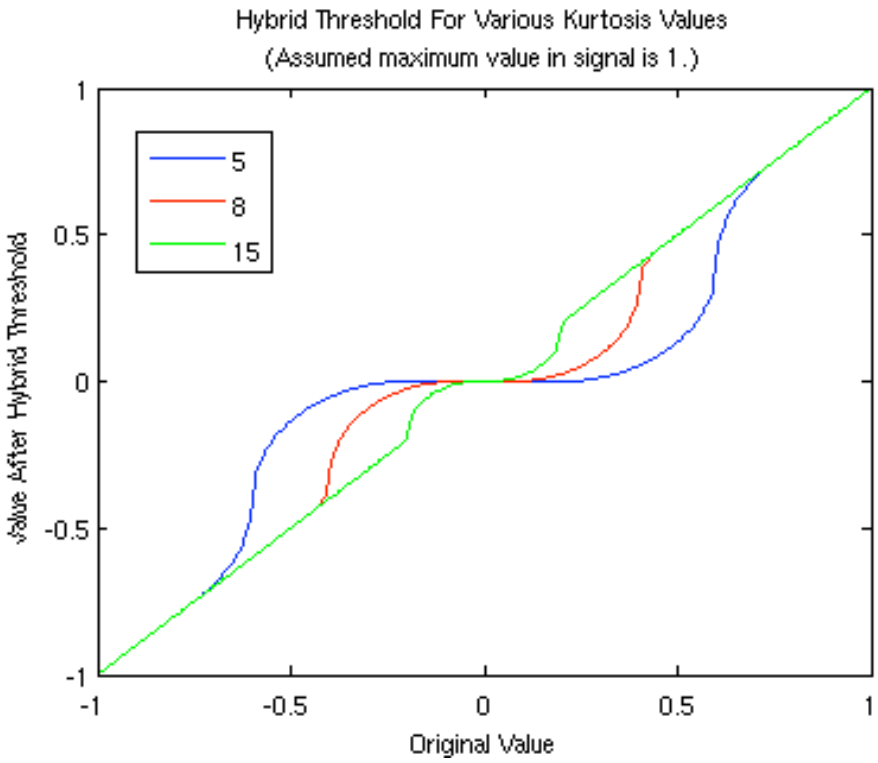

Figure 30: Three hybrid thresholds are graphically illustrated for different values of kurtosis.

As can be seen in Figure 30, as the value of kurtosis increases the relative thresholds are modified from soft to hard, and they decrease, allowing more of the signal to pass. As can be seen, the hybrid threshold is also continuous at the boundaries between the purely hard threshold and the pass region. Figure 31 gives a vibration signal computed with a visible inner race fault and Figure 32 gives the resulting wavelet transform. Figure 33 is the modified wavelet transform after the K-hybrid threshold has been applied. Figure 34 gives the inverse wavelet transform of Figure 33. Figure 33 is the denoised fault signal.

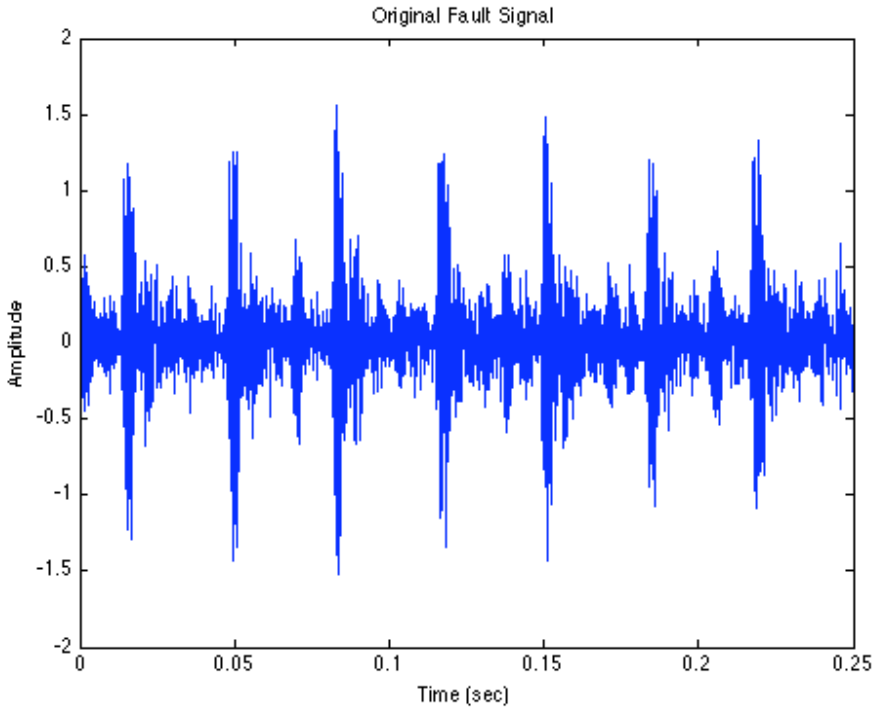

Figure 31: Vibration signal with an inner race fault.

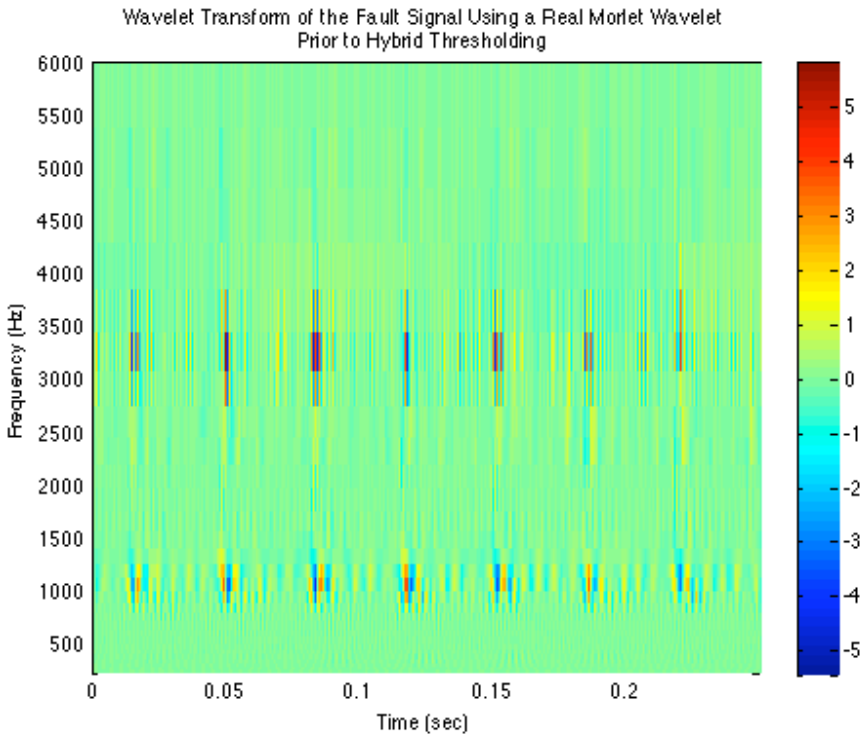

Figure 32: Wavelet transform of the vibration signal seen in Figure 31.

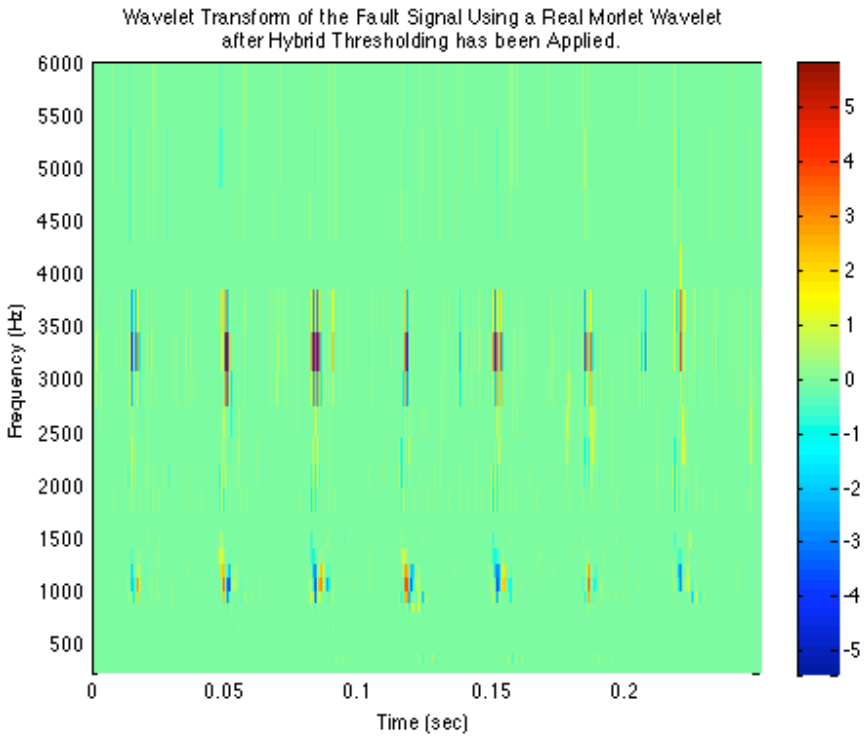

Figure 33: Wavelet transform seen in Figure 32 after the K-Hybrid threshold has been applied.

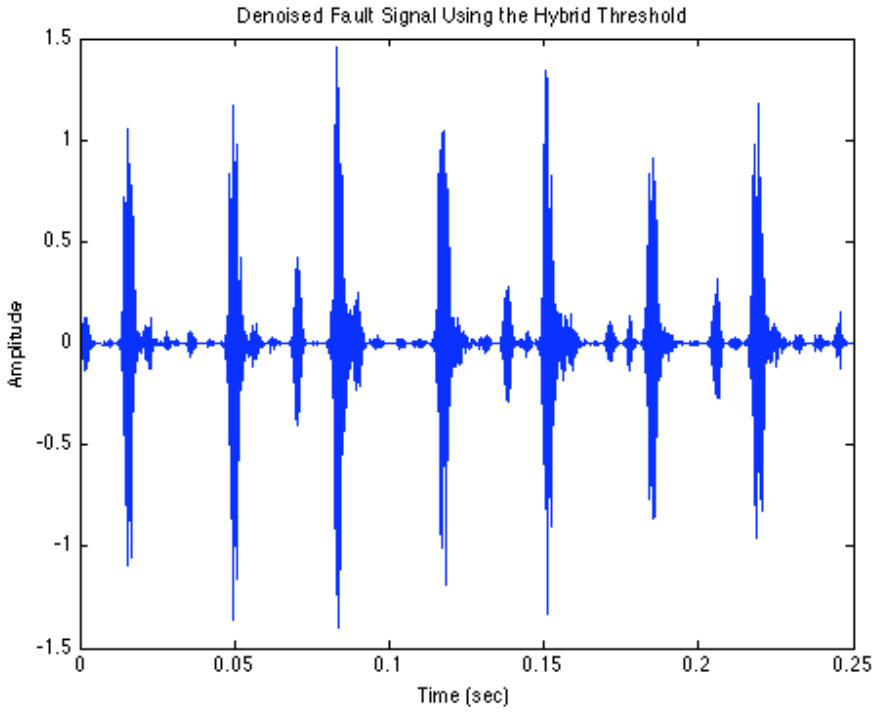

Figure 34: The de-noised signal, constructed from the inverse wavelet transform seen in Figure 33.

3.5 The Sparse Code Shrinkage Threshold Using the Continuous Wavelet Transform

The Sparse Code Shrinkage Threshold was originally proposed to de-noise a signal based on an orthogonal transform [64]. It has recently been adopted for use to de-noise mechanical systems using the non-orthogonal CWT and shown to work with bearing fault signals [51]. The original threshold was developed to estimate non-Gaussian data from Gaussian noise. The probability density function (pdf) of a sparse distribution is given in (36).

$$p(s) = \frac{1}{2d} \frac{(\alpha+2)\left[\alpha(\alpha+1)/2\right]^{(\alpha/2+1)}}{\left[\sqrt{\alpha(\alpha+1)/2} + |s/d|\right]^{(\alpha+3)}} \tag{36}$$

In [44], it was shown that the pdf, $p(s)$, given in (36) matches that of simulated impulses using $\alpha = 1$ and $d = 0.1529$. For signals represented by this pdf the following threshold rule given in (37) was proposed for orthogonal transforms based on the principle of Maximum Likelihood Estimation [64].

$$W_{threshold} = sign(W)\max\left(0, \frac{|W| - ad}{2} + \frac{1}{2}\sqrt{(|W| - ad)^2 - 4\sigma^2(\alpha+3)}\right) \tag{37}$$

where $a = \sqrt{\alpha(\alpha+1)/2}$ and σ is the standard deviation of the noise. For this study a robust estimate of σ was used by setting σ to the median of the absolute deviation (MAD) of the wavelets smallest scale divided by 0.6745 as seen in (38).

$$\sigma = \frac{MAD}{0.6745} \tag{38}$$

The division by 0.6745 is a calibration term to normalize σ to that of a Gaussian distribution. A graphical illustration of the threshold can be seen in Figure 35.

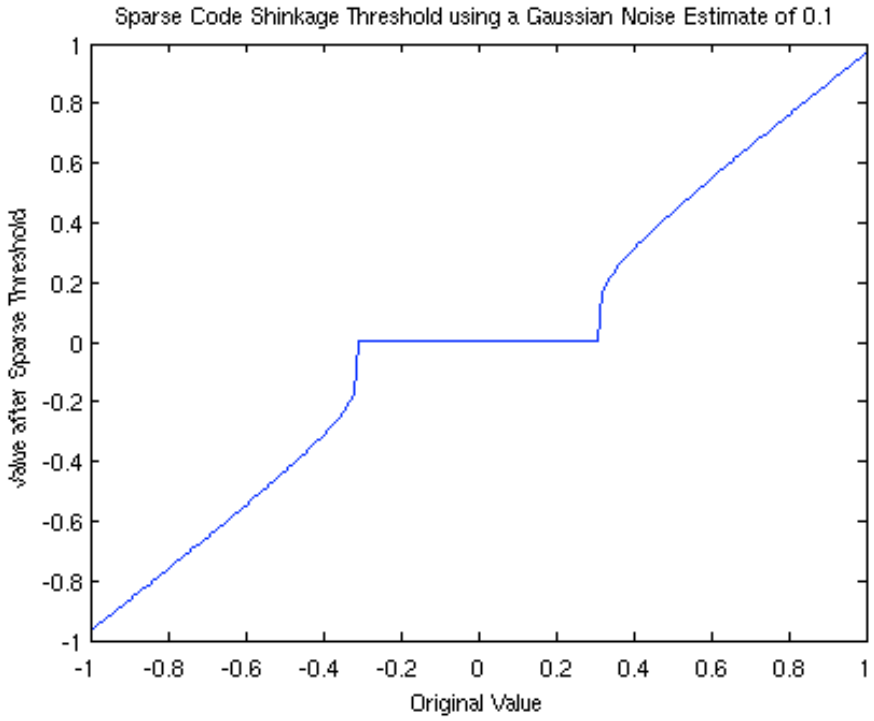

Figure 35: This is a graphical representation of the sparse code thresholding rule given in eq. (37).

Figures 36, 37, 38 and 39 give an example of the application of the sparse code shrinkage threshold applied to an actual fault signal for de-noising purposes. Figure 36 gives a vibration signal computed with a visible inner race fault and Figure 37 gives the resulting wavelet transform. Figure 38 is the modified wavelet transform after the sparse code shrinkage threshold has been applied. Figure 39 gives the inverse wavelet transform of Figure 38 and the de-noised fault signal.

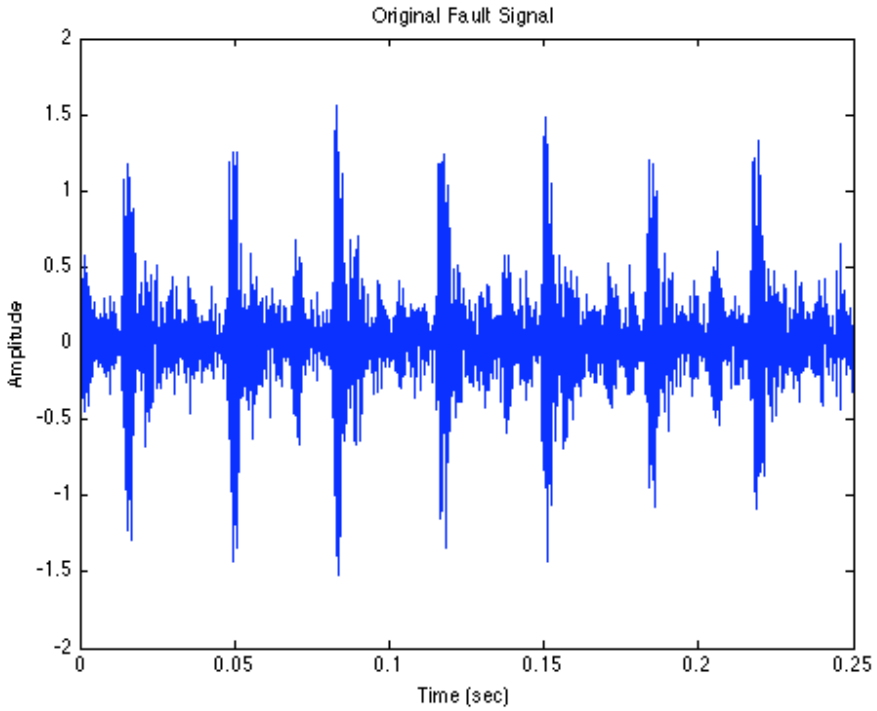

Figure 36: Vibration signal with an inner race fault.

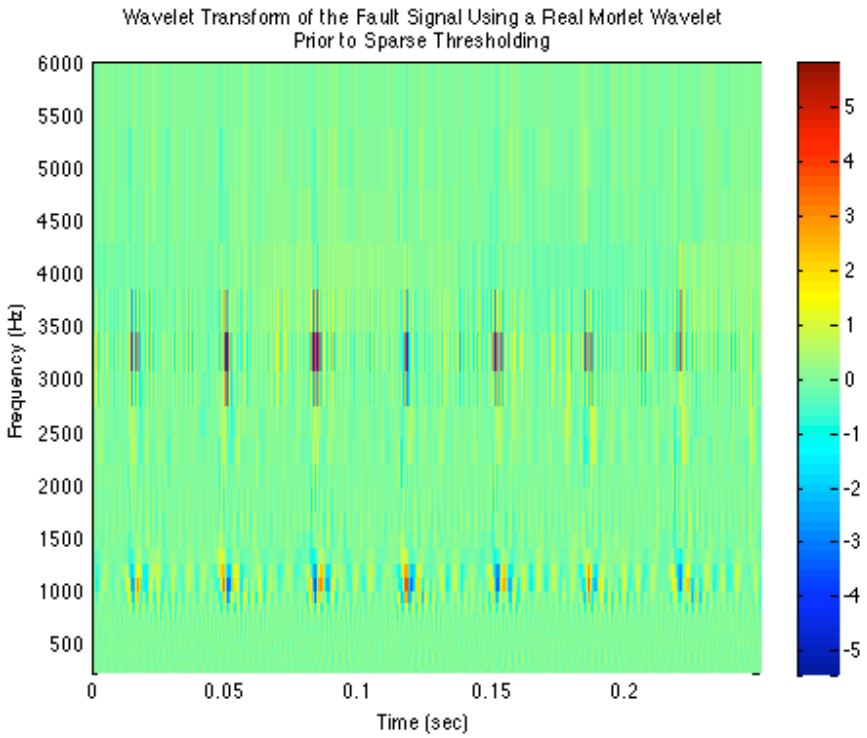

Figure 37: Wavelet transform of the vibration signal seen in Figure 36.

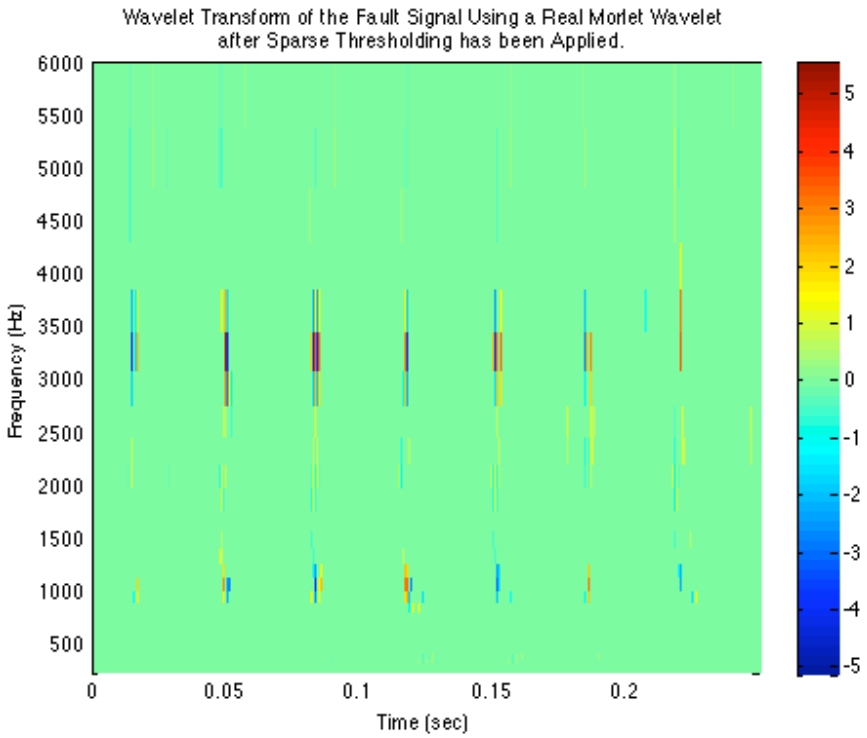

Figure 38: Wavelet transform seen in Figure 37 after the sparse coding threshold has been applied.

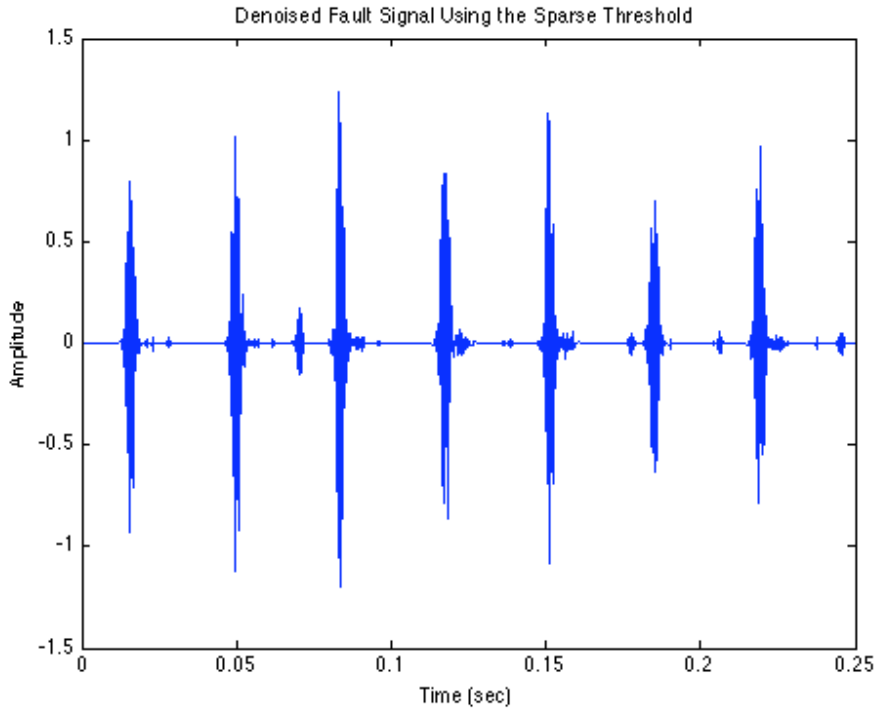

Figure 39: The de-noised signal, constructed from the inverse wavelet transform seen in Figure 38.

3.6 Minimum Entropy De-convolution

Minimum Entropy De-convolution (MED) with Autoregressive (AR) Modeling is a unique de-noising algorithm that uses a de-convolution filter originally developed for geophysical applications [60]. This algorithm does not use the CWT and, as a result, it is being introduced as an alternative to wavelet de-noising. This gives the comparison study that will be introduced in section 5 an opportunity not only to compare several de-noising algorithms that implement the CWT, but to see how well CWT de-noising

algorithms compare to a relatively new technique that de-noises bearing faults that does not use wavelets.

The first step in this algorithm attempts to remove dominant structures in the signal that are obscuring or repressing the fault signal. This technique models the stationary components of the signal using an AR filter. A wide-sense stationary random process can be represented as the output from a causal invertible linear system that is being excited by white noise [55]. In particular, an AR process, or all pole linear filter, can be represented as the difference equation given in (39), where $x(n)$ is the sequence that is being filtered, a_k are the filter coefficients and $\hat{x}(n)$ is the predicted output at data point n.

$$\hat{x}(n) = \sum_{k=1}^{p} a_k x(n-k) \tag{39}$$

The coefficients a_k, can be found by solving the famous Yule-Walker equations given in (40), where γ_{xx} is the autocorrelation sequence of $x(n)$.

$$\begin{bmatrix} \gamma_{xx}(0) & \gamma_{xx}(-1) & \gamma_{xx}(-2) & \cdots & \gamma_{xx}(-p) \\ \gamma_{xx}(1) & \gamma_{xx}(0) & \gamma_{xx}(-1) & \cdots & \gamma_{xx}(-p+1) \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma_{xx}(p) & \gamma_{xx}(p-1) & \gamma_{xx}(p-2) & \cdots & \gamma_{xx}(0) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \gamma_{xx}(0) \\ \gamma_{xx}(1) \\ \vdots \\ \gamma_{xx}(p) \end{bmatrix} \tag{40}$$

Once constructed for a given signal $x(n)$, this filter is then used to 'whiten' the signal and remove the stationary components as seen in (41) where $w(n)$ is white noise [62].

$$w(n) = x(n) - \sum_{k=1}^{p} a_k x(n-k) \tag{41}$$

The AR filter, if the correct length k of the filter is chosen, will only be able to model the dominant deterministic structures in the signal. This will leave only the weak impulsive components of the signal behind. This is a huge advantage as the mechanical noise that is usually corrupting the fault signal is usually not Gaussian nor white but structured in nature. Two examples of this type of noise would be vibrations from shaft imbalance or gear meshing in a transmission system. The order of the filter, k, can easily be selected by repetitively using the whitening filter given in (41) with steadily increasing filter order until the kurtosis, as given in (39), of the whitened signal is maximized [65]. An example of a bearing fault signal that has been whitened using an AR filter is given in Figure 40.

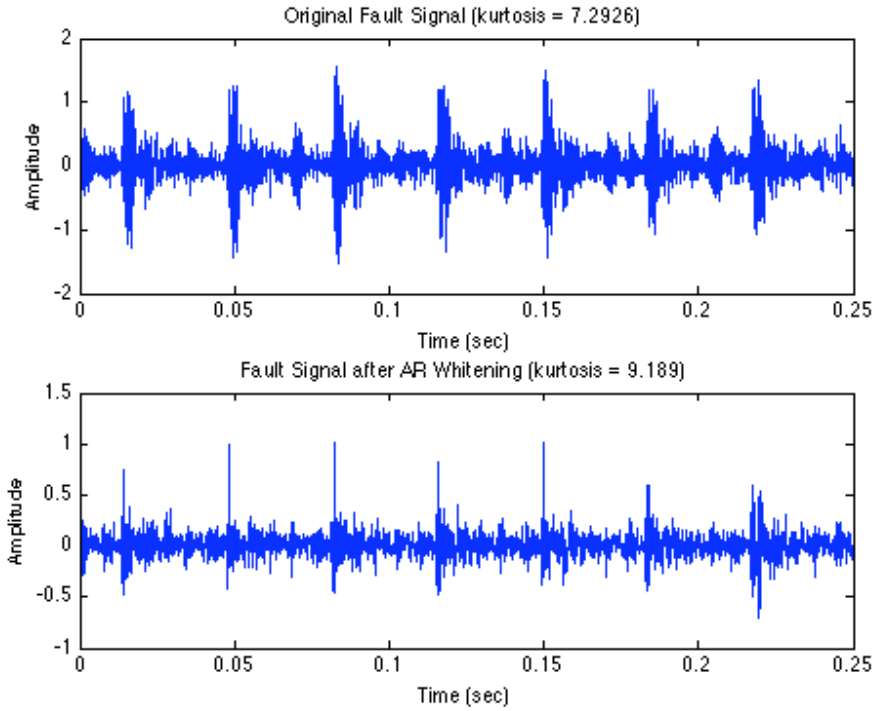

Figure 40: This figure shows the original signal **(top)** prior to using AR modeling to whiten the signal. The effect of the AR filter can also be observed **(bottom)**.

Once the signal is whitened, a blind de-convolution technique, known as Minimum Entropy De-convolution (MED), originally developed for geophysical application, is used to de-noise the signal. Using the MED technique, an optimum filter is constructed that maximizes the kurtosis value of the signal after de-convolution. The method used by Endo and Randall to perform MED is a repetitive procedure based on the Objective Function Method to maximize kurtosis [60, 65]. The objective function, kurtosis is given again in (42) where y is the signal of interest with a mean of zero.

$$O(k(l)) = \sum_{i=1}^{N} y^4(n) \Big/ \left[\sum_{i=1}^{N} y^2(n)\right]^2 \tag{42}$$

The general formula for inverse filtering using convolution is given in (43) where $h(n)$ is the inverse filter and $y(n)$ is the filter output.

$$y(n) = \sum_{n=1}^{L} k(l)x(n-l) \tag{43}$$

Optimizing (42) with respect to k(l) and using $\partial y(n)/\partial k(l) = x(n-l)$, from equation (42) yields (43).

$$\left[\sum_{n}^{N} y^2(n) \Big/ \sum_{n}^{N} y^4(n)\right] \sum_{n}^{N} y^3(n)x(n-l) = \sum_{p}^{L} k(p) \sum_{n}^{N} x(n-l)x(n-p) \tag{44}$$

Equation (44) simplifies to a simple matrix operation. The left hand side of (44) is a simple column vector b. The right hand side of (44) is the autocorrelation function of x, denoted as A, multiplied by a vector k, the unknown filter coefficients. This leads to redefining (44) as (45).

$$b = Ak \tag{45}$$

To find the optimum filter coefficients the following steps are performed:

1.) Compute the autocorrelation function A and assume an initial filter k^0. For this study a delayed impulse was used as the initial filter coefficients.

2.) Perform the inverse filter as seen in equation (44) to find y^0, the initial filter results.

3.) Compute the left hand side of equation (44) to find b^1.

4.) Solve for new filter coefficients, k^1, by solving equation (45) using b^1 and A.

5.) Compute the error criterion, *e*, given in (46) and (47). Restart at step 2 until the error reaches an acceptable level. *E* in equation (47) represents the expected value.

$$e = \left(k^1 - \mu k^0\right) \Big/ \mu k^0 \tag{46}$$

$$\mu = \frac{E\left(k^0\right)^2}{\left(E\left(k^1\right)^2\right)^{1/2}} \tag{47}$$

The MED algorithm was repeated using longer lengths for the filter *g* until the maximum kurtosis value was reached. The results of the MED algorithm can be seen in Figure 41, where the fault signal shown in Figure 40, after having applied AR modeling to whiten the signal, was subjected to MED. As can be seen in Figure 41, the faults are clearly defined as impulses.

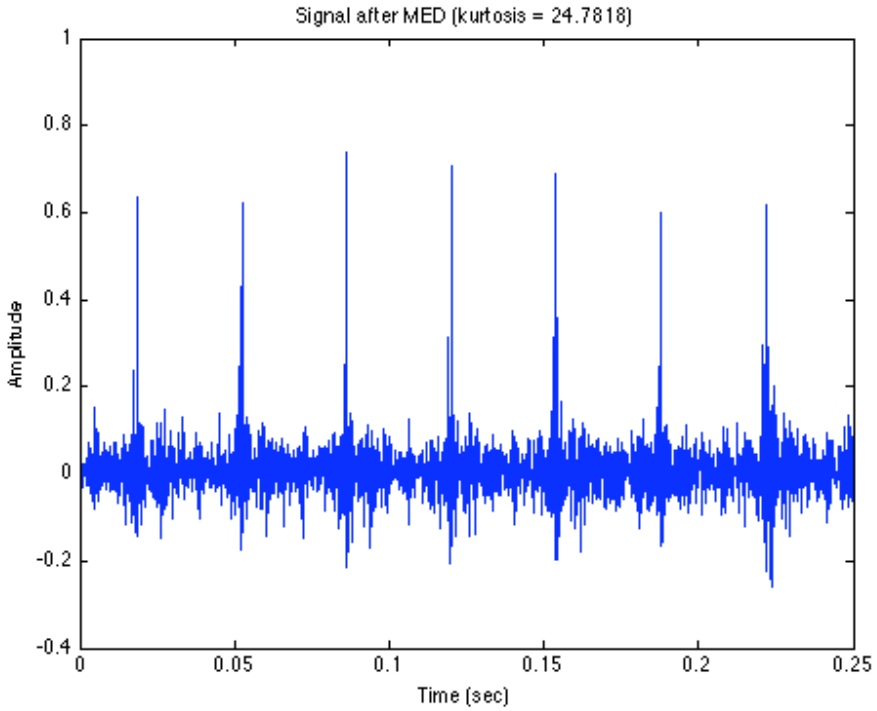

Figure 41: This figure gives the final de-noised signal after MED has been applied to the signal seen in Figure 40. The periodic impulses from the fault are clearly visible.

4. PROPOSED ALGORITHMS

In this section a new wavelet, a new detection algorithm as well as two new wavelet de-noising algorithms tailored to de-noise the bearing fault signals, are introduced. The wavelet, termed the Natural Frequency Impulse Response Wavelet, has been developed from the simple assumption of the exponentially decaying sinusoidal shape of the theoretical bearing fault impulse. The two de-noising algorithms and detection algorithm that will be introduced use the CWT along with a unique Fourier Series decomposition. One of the main drawbacks to Fourier analysis in the digital domain is the discretization of the frequency spectra of the signal. When this occurs weak spectra in the signal can, and often do become lumped with larger spectra in the signal rendering detection of weak signals, in the frequency domain, impossible. The Fourier Series technique introduced in this thesis allows for a set of arbitrary frequency spectra to be calculated for a signal providing a unique method of detection previously unused in rolling element bearing fault detection.

4.1 Natural Frequency Impulse Response Wavelet

For a wavelet to be effective in detecting transient signals it is optimal to have a wavelet that matches the transient signal of interest. In the case of bearing fault detection, a simple model of the transient signal of interest is an exponentially decaying sinusoid as given in (1).

The Fourier transform of the exponentially decaying sinusoid given in (1) is

$$X(\omega) = \frac{\omega_o}{(j\omega - b)^2 + \omega_o^2} \tag{48}$$

A proposed wavelet must meet the following criteria [57]:

i.) A wavelet must have finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty \tag{49}$$

ii.) The wavelet must not have a DC component.

$$C_g = \int_{-\infty}^{\infty} \frac{|\psi(f)|^2}{f} df < \infty \tag{50}$$

iii.) For a complex, or analytic, wavelets the Fourier transform must be real and vanish for negative frequencies.

An exponentially decaying sinusoid does not satisfy the requirements of a wavelet due to its non-zero DC component. As a result, to effectively turn the exponentially decaying sinusoid into a wavelet it is necessary to get rid of the DC component in the signal. By modifying (48) with a high pass filter, in the frequency domain, effectively makes the signal band pass and as a result a wavelet. The Frequency Domain Impulse Response Wavelet can be seen in (51).

$$X(\omega)=\left(\frac{j\omega}{j\omega+\omega_c}\right)\left(\frac{\omega_o}{(j\omega-b)^2+\omega_o^2}\right) \tag{51}$$

For a digitally sampled signal the signal is inherently band limited to π. Rewriting (51) for implementation as a discrete filter yields:

$$\psi(\omega)=\left(\frac{j\omega}{j\omega+\frac{\pi}{5}}\right)\left(\frac{\pi}{(j\omega-b)^2+\pi^2}\right) \tag{52}$$

The cut off frequency ω_c was determined experimentally and set to π/5. This cutoff frequency allowed most of the impulsive components of the exponentially decaying sinusoid to be unaffected by the addition of the high pass filter. For small scales the frequency domain is sparsely sampled. By setting the cut off frequency, ω_c, at π/5 this ensured that even for the smallest scale of one, no DC components of the exponentially decaying sinusoid exist and ensures that the admission criterion is met.

The Time domain and Frequency domain of the Frequency Domain Impulse Response Wavelet (51), as well as the time domain and frequency domain of the signal of interest (1) are given in Figure 42.

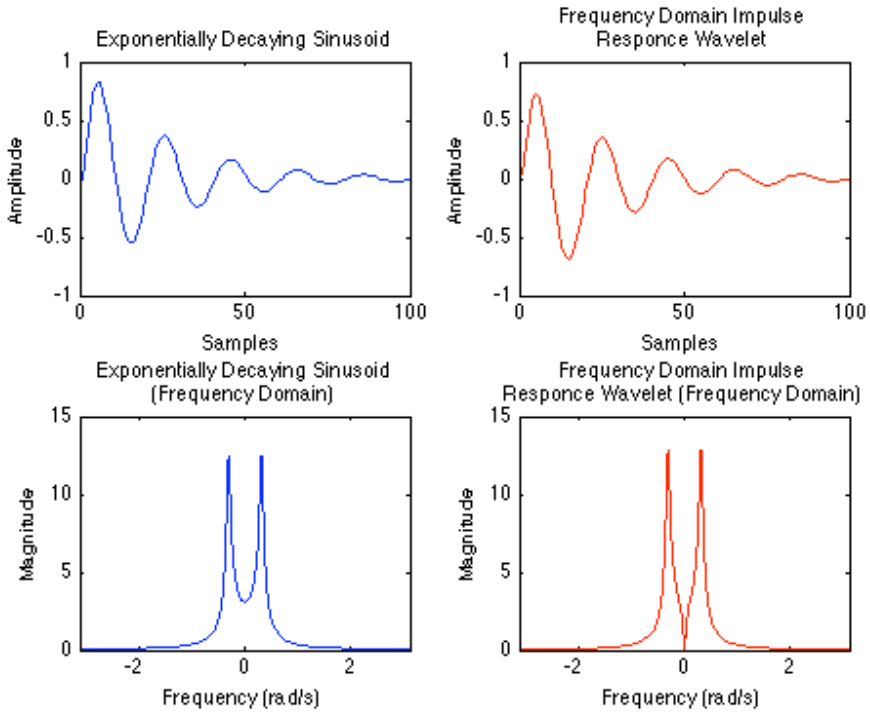

Figure 42: Shown are time domain **(top)** and frequency domain **(bottom)** representations of an exponentially decaying sinusoid **(left)** and the Frequency Domain Impulse Response Wavelet **(right)**.

The Frequency Domain Impulse Response wavelet matches the signal both in the time and frequency domain to a high degree. The simplicity of the equation representing the wavelet is lost when performing the inverse Fourier Transform of (51). As a result, this wavelet is best suited for implementation as a non-causal filter rather than in the time domain as a convolution product. Similarly, this wavelet, unlike many, is not even in the time domain, or equivalently real in the frequency domain. Consequently, no analytic version of this wavelet exists and thus limits its practicality.

4.2 The Discrete Reduced Fourier Series

Discrete Reduced Fourier Series (DRFS) was developed to de-noise low frequency vibration data [66]. By applying the DRFS a set of data can be approximated by arbitrary number of sinusoids, each with an arbitrary frequency. The development of the DRFS algorithm starts with the development of the traditional discrete Fourier Series analysis and synthesis equations. These equations can be seen in (53) and (54).

$$X_k = \frac{1}{N}\sum_{k=0}^{N-1} x_n W^{-kn} \tag{53}$$

$$x_n = \sum_{k=0}^{N-1} X_k W^{kn} \tag{54}$$

$$W^{kn} = e^{j2\pi kn/N}$$

The analysis equation (54) can be written as a matrix operation given in (9).

$$x_n = X_k W \tag{55}$$

Least squares estimation can then be applied to the solution of (55) allowing for the use of any set of arbitrary frequency components [63]. This representation is given below as (56).

$$X_k = \left(W^H W\right)^{-1} W^H x_n \tag{56}$$

The use of DRFS is remarkable as it allows for any data set to be approximated as a set of selective frequencies. What is most unique is that the frequencies that can be used do not necessarily need to be discrete and leakage, an issue with Fourier analysis is minimized [21]. This can easily be seen in Figure 43 where a low frequency signal has been corrupted by high frequency noise and a DRFS estimation of the signal has been

used to recover the corrupt signal by approximating the signal as the first 11 discrete Fourier Series.

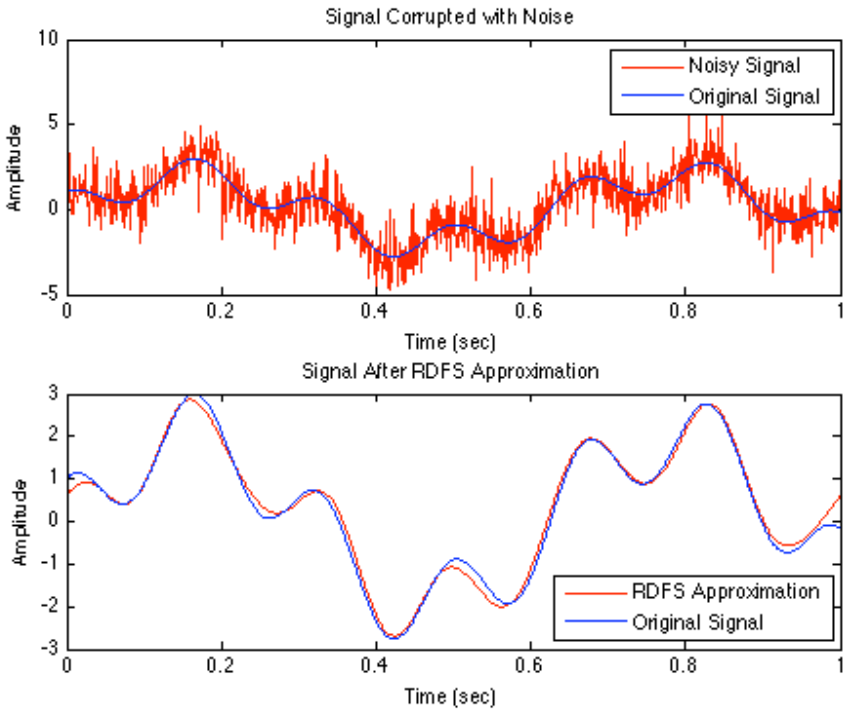

Figure 43: Top: A low frequency signal, labeled as 'original signal', can be seen covered with noise. **Bottom:** The 'original signal' covered with noise has been successfully recovered using a DRFS approximation with the first 11 discrete Fourier Series.

Assuming a signal has a periodic structure with period T_p it will admit to a traditional Fourier Series with discrete spectral lines. These discrete spectral lines are integer multiples of the cylcic frequency F_c where $F_c = 1/ T_p$ [55]. If the cyclc frequency

of a signal, F_c, is known, the signal can then be de-noised by using a DRFS approximation with discrete harmonic frequencies. An example of this can be seen in Figure 44 where a periodic signal with a cyclic frequency of F_c of 5.67 Hz is recovered after being corrupted with noise.

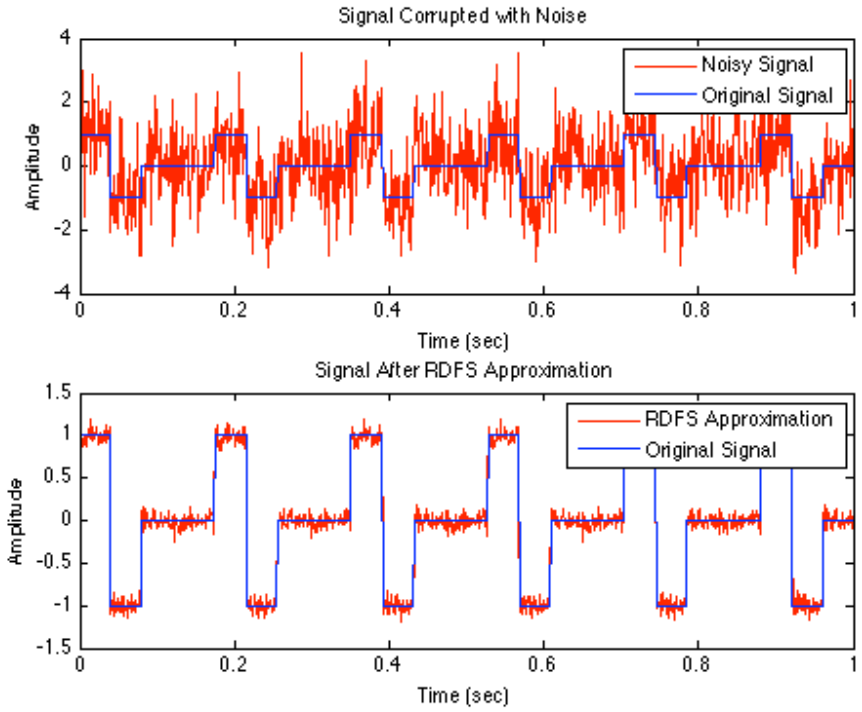

Figure 44: A periodic signal has been successfully recovered using a DRFS approximation using a non-discrete periodic Fourier Series representation.

There are four major weaknesses to the DRFS de-noising method on periodic signals. The fundamental period of the signal must be assumed prior to attempting the

algorithm. Second, to successfully de-noise a periodic sequence, the signal that will be de-noised must contain a significant number of periodic sequences. Third, periodic signals with a low cyclic frequency require a large number of discrete frequency spectral to accurately capture the signal. These last two facts can require, using the least squares approach, very large matrix inversions requiring lots of time and computational effort. Finally the stability of the DRFS algorithm is not guaranteed. For this research it was found that the DRFS algorithm, as long as the number of assumed harmonics was equal to or less than the number data points in the signal and the divisions between assumed frequencies were never less than those of a standard Fourier Series the algorithm always converged.

4.3 Cyclic Detection Using the DRFS with the Continuous Wavelet Transform

The relative periodicity, or inversely the cyclic frequency, of a signal can also be gauged by measuring how much energy was captured by the periodic DRFS approximation relative to the original signal. This can be used to find hidden periodic components of a signal. The noisy signal shown in Figure 17 has a cyclic frequency of 5.67 Hz. By trying successive approximations of steadily increasing and decreasing cyclic frequencies around 5.67 Hz and comparing the energy of the DRFS representation of the signal with the energy from the noisy signal, we can clearly identify 5.67 Hz as the cyclic frequency of the original signal. The results can be seen in Figure 45.

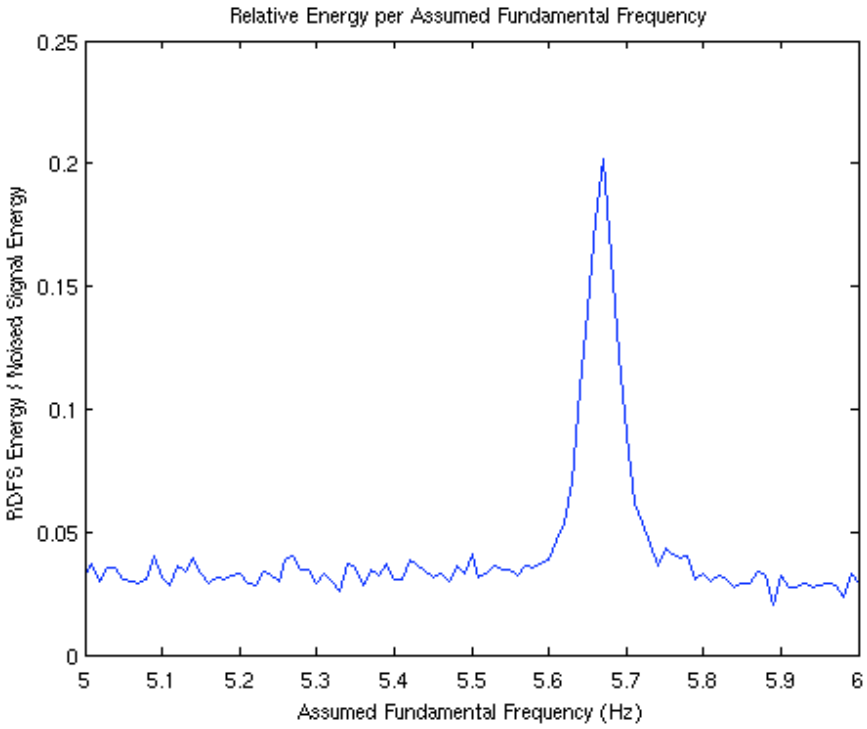

Figure 45: The DRFS approximation of the signal with an assumed fundamental frequency of 5.67 Hz captures the most energy from the noisy signal.

When combined with the CWT, this technique can be used to detect hidden amplitude modulated signals, such as rolling element bearing faults. The CWT can be viewed as a constant quality filter bank. By filtering the signal using the CWT and summing the energy from DRFS approximations at various cyclic frequencies, a signal can be analyzed to find hidden cyclic frequencies across its frequency spectrum.

The first step to perform this cyclic analysis is to perform the CWT using an analytic wavelet. Using an analytic wavelet both filters the signal into frequency bands

and, when the absolute value of the scalogram is computed, automatically demodulates the signal. For this research the analytic Morlet wavelet was used. Once the signal is in the wavelet domain each scale has its mean subtracted as seen in (57),

$$\hat{W}_a(b) = |W_a(b)| - \frac{1}{N}\sum_b |W_a(b)| \tag{57}$$

where W_a is the wavelet transform of a given signal of interest x using the analytic Morlet wavelet and N is the length of the arbitrary wavelet scale a. By performing this subtraction, the DC portion of the signal, constituting energy of the signal that is clearly not periodic, is removed. Second, the DRFS algorithm is applied assuming only harmonic components in the signal at an assumed cyclic frequency. The energy in the approximated scale at the assumed cyclic frequency is then calculated and divided by the original energy of the scale, given as,

$$E_{Cyclic}(a,cf) = \frac{\sum_{b=1}^{N} \hat{W}_{a,cf}(b)}{\sum_{b=1}^{N} \hat{W}_a(b)} \tag{58}$$

where E_{Cyclic} is the relative periodic energy of the signal at scale a, $\hat{W}_a$ is the demodulated wavelet scale given by (57) and $\hat{W}_{a,fc}$ is the DRFS approximation of $\hat{W}_a$ at the cyclic frequency fc.

This method of determining possibly hidden cyclic frequencies is very simple, yet powerful. The following example illustrates this concept. A periodic signal, x, sampled at 1000 Hz, with a cyclic frequency of 5.67 Hz is shown in Figure 46.

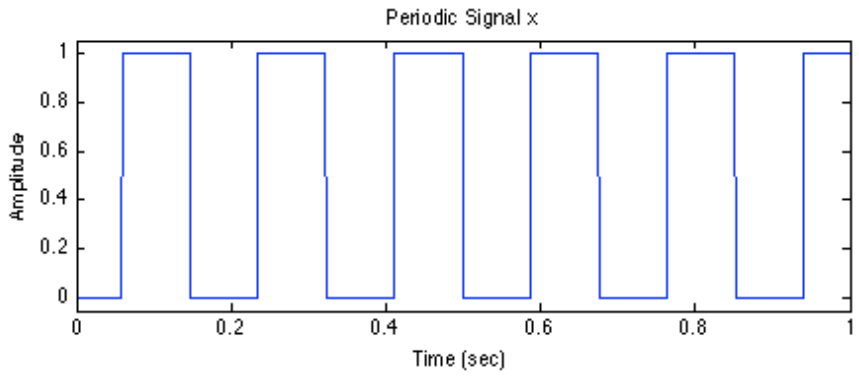

Figure 46: Periodic square wave signal with a period of 5.67 Hz sampled at 1000 Hz.

The signal *x* was modulated by a 200 Hz sinusoid and corrupted with white noise to give a signal to noise ratio (SNR) of approximately -10 dB. As can be seen in Figure 47, the modulating function, a simple 200 Hz sinusoid can be detected in the magnitude of the Fourier transform, the periodic nature of the modulated signal *x* can not easily be discerned.

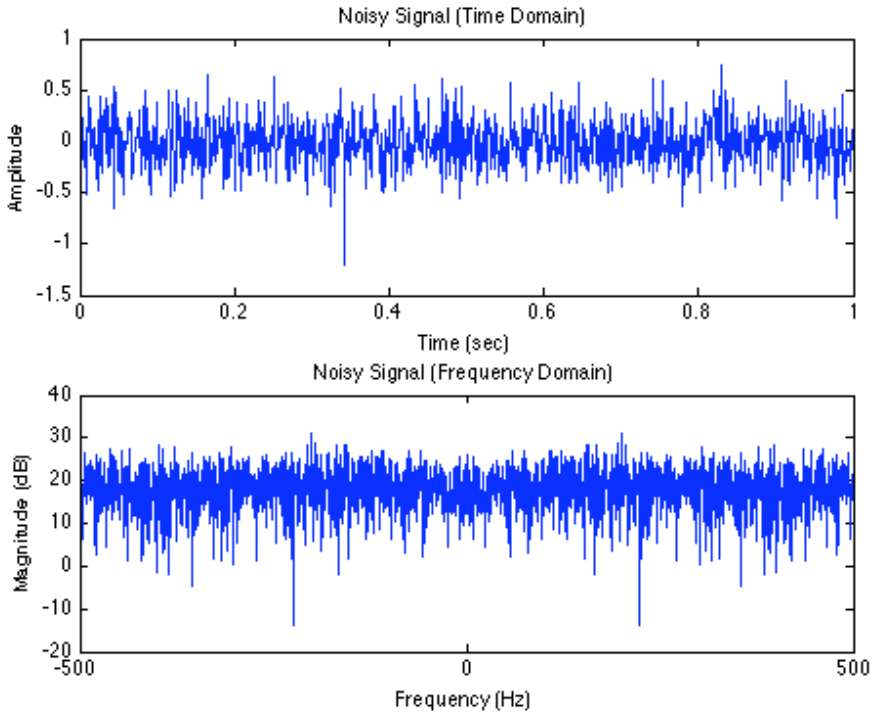

Figure 47: The periodic square wave signal *x*, seen in Figure 46, after modulation and added noise.

By performing the cyclical analysis procedure given above, the hidden cyclical frequency of the signal at 5.67 Hz can easily be seen, as given in Figure 48.

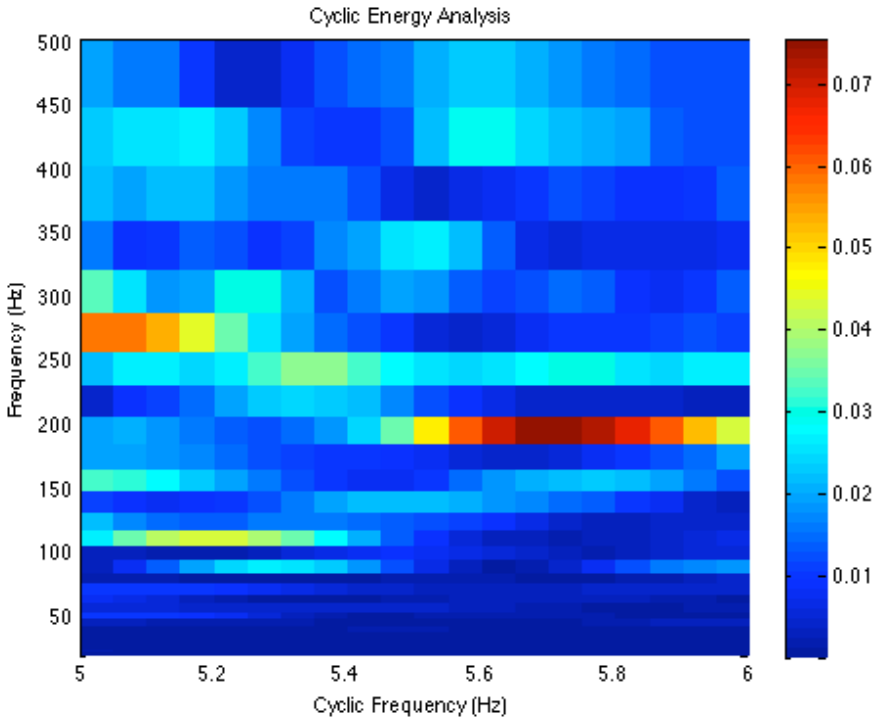

Figure 48: This figure shows the results of using DRFS in conjunction with the CWT to determine hidden cyclic signals. The periodicity and modulations frequencies are clearly identified at 5.67 Hz and 200 Hz respectively.

This method also works on bearing fault signals as seen in Figure 49 and 50 where both the natural frequency of the bearing fault response and the cyclic frequency of a bearing fault signal can be clearly identified even though the fault is not obvious in either the frequency or time domain. In the case of a bearing fault, the signal is not corrupted by white noise and the fault signature can be relatively isolated on one or more frequency bands. This can be observed in Figure 49 where the dominant noise is periodic in nature and not evenly distributed across the frequency domain. In fact there are

components of the signal that looks impulsive and periodic in the time domain and could be mistaken for the bearing fault. These periodic impulses are at the running speed of the shaft the bearing is supporting and not the much weaker fault signature.

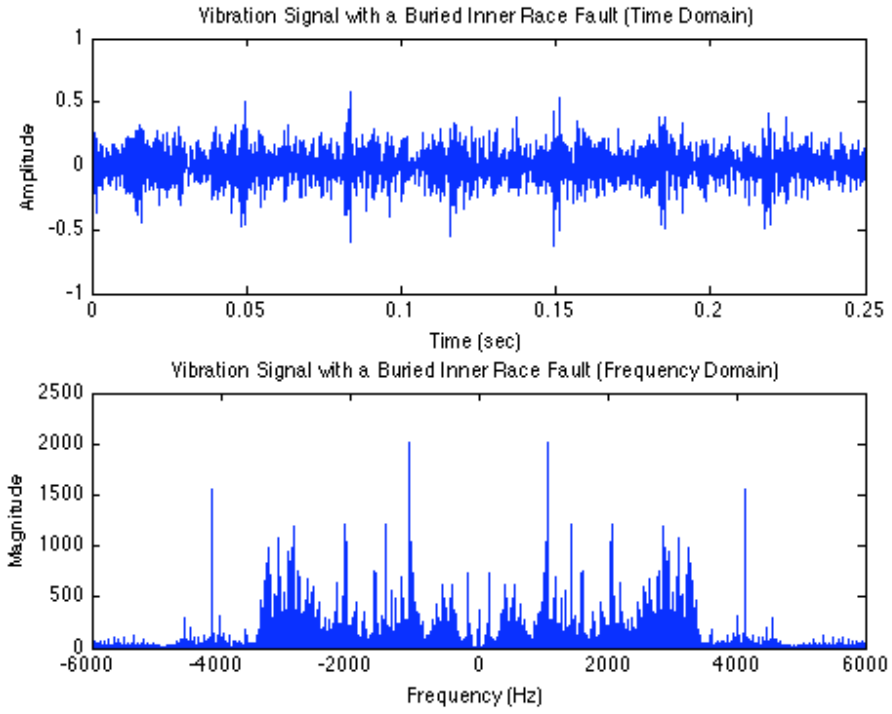

Figure 49: This figure shows the results a bearing fault signal overwhelmed with noise. The periodicity of the fault can not be identified in either the time or frequency domain.

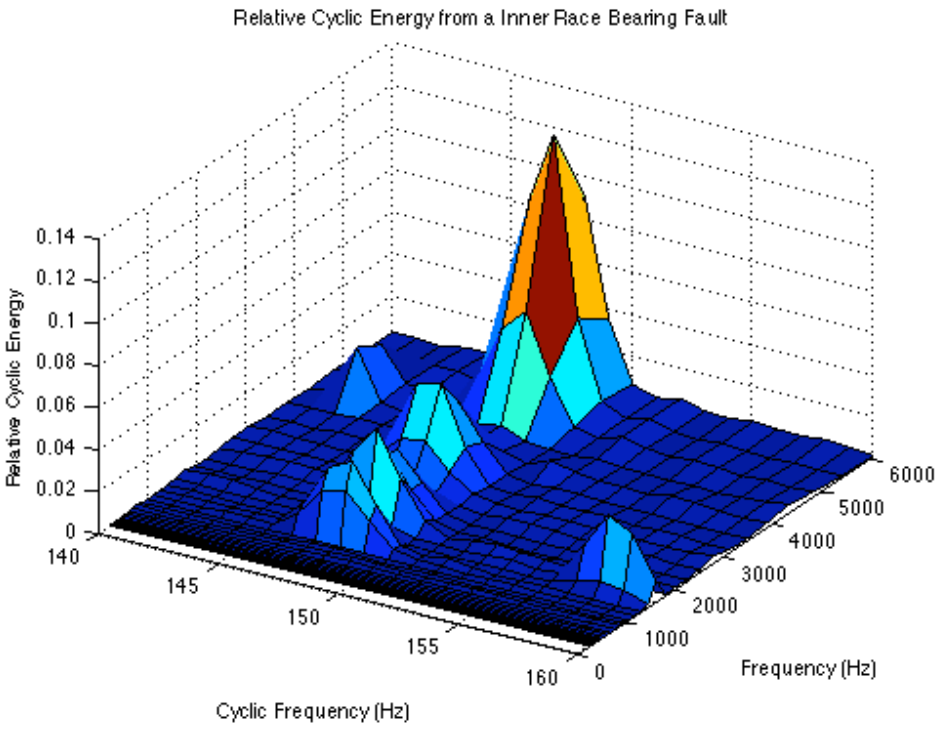

Figure 50: The cyclic frequency of the bearing fault can clearly be identified using the DRFS approximation as well as the natural frequency of the bearing fault.

As seen in Figure 50, the fault signal admits to a Fourier series at approximately 5.5 kHz with a cyclic frequency of ~147 Hz. The running speed of the shaft is approximately 30 Hz and the fourth harmonic of the running speed is approximately 150 Hz. This is very close to the fault cyclic frequency of 147 Hz. With traditional Fourier analysis the harmonics due to the running speed would be evident and, as in situations such as this where the cyclic frequency of the fault is very close to one of the harmonics from the shaft, it could be difficult to find the bearing fault.

One of the most powerful aspects of using the DRFS in conjunction with the CWT to detect hidden cyclic frequencies in specific frequency bands is the ability to zoom in on specific frequency bands and cyclic frequencies of interest. This can be seen in Figure 51 where a much better estimation of both the natural frequency of the bearing fault and cyclic frequency, found in Figure 50, can be identified by evaluating the CWT at scales localized around 5368.8 Hz and evaluating the DRFS at cyclic frequency about 147.65 Hz.

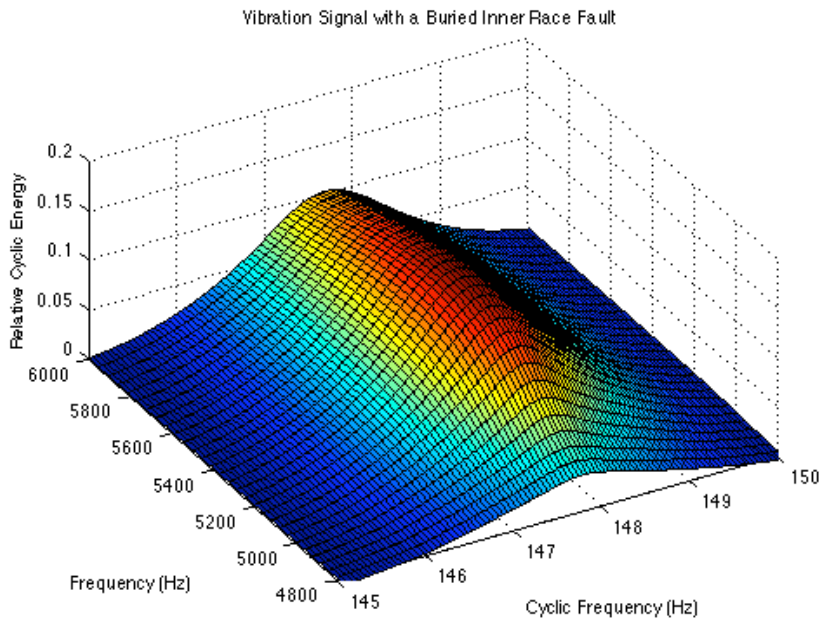

Figure 51: The cyclic frequency of the bearing fault can clearly be evaluated at different frequencies and cyclic frequencies to give a very accurate approximation of both the frequency band that shows cyclic characteristics as well as the cyclic frequency itself.

This technique is proposed as a detection tool for bearing fault signals and will be referred to, from this point, onward as the CWT/DRFS Cyclic Energy algorithm. There are a number of other techniques that exist to detect periodic or almost periodic signals by providing a frequency versus cyclic frequency distribution. This class of signal is also termed as cyclostationary. From cyclostationary theory, one of the most prominent methods to perform this analysis is known as the cyclic power spectrum (CPS),

$$S_x(f,\alpha) = \lim_{L\to\infty} \frac{1}{L} E\left\{ X_L\left(f + \frac{\alpha}{2}\right) X_L\left(f - \frac{\alpha}{2}\right)^* \right\} \tag{58}$$

where E denotes the expected value, X_L is the Fourier transform of a cyclostationary signal x of length L and α is the cyclic frequency of interest. The CPS produces a frequency versus cyclic frequency distribution similar to that given in Figure 41. The details of this tool are beyond the scope of this thesis, but can be found in [21]. What is unique about cyclic detection using the CWT/DRFS algorithm introduced in this paper is the ability of this algorithm to perform cyclic analysis without the use of the standard Fourier Transform. The Fourier Transform has issues with both frequency resolution and leakage. Furthermore the CWT/DRFS cyclic analysis algorithm allows for a particular defect or signal of interest to be highlighted when specifying a wavelet to be used with the CWT. The CWT/DRFS algorithm is also numerically efficient. Because each wavelet scale is demodulated prior to an DRFS estimation being applied the signal is basebanded and therefore only a limited number of harmonics need to be estimated. In this simulation only the first 10 harmonics of the signal were calculated for each cyclic

frequency leading to a relatively fast computation time compared to computing a harmonic set that spanned the width of the signals frequency domain.

Finally the CWT/DRFS cyclic analysis algorithm produces a de-noised signal. The scale that shows the most cyclic energy, E_{Cyclic}, at the cyclic frequency of interest is the original bearing fault frequency signal filtered and demodulated around the bearing fault frequency. This is exactly what envelope demodulation attempts to accomplish. However, by using the CWT/DRFS cyclic analysis the natural frequency of the bearing is easily found without the guesswork that is necessary when using envelope demodulation. Furthermore, the filter width is also inherently specified when performing the CWT while specifying the shape factor of the wavelet.

By determining the wavelet shape that minimizes SEM, the filter that best matches the bearing fault can be determined experimentally. While this filter width is not necessarily optimum in the sense of capturing the most harmonics of the signal it has been shown to work in practice. The de-noised version of the bearing fault seen in Figure 49 and analyzed in Figure 50 can be seen below in Figure 52. What is particularly pleasing about this algorithm is that the modulation effect of the amplitude of the impulses, predicted for inner race faults, due to the rotating load over the fault, is clearly visible. From this point forward this method of de-noising will be referred to as Automated CWT/DRFS de-noising.

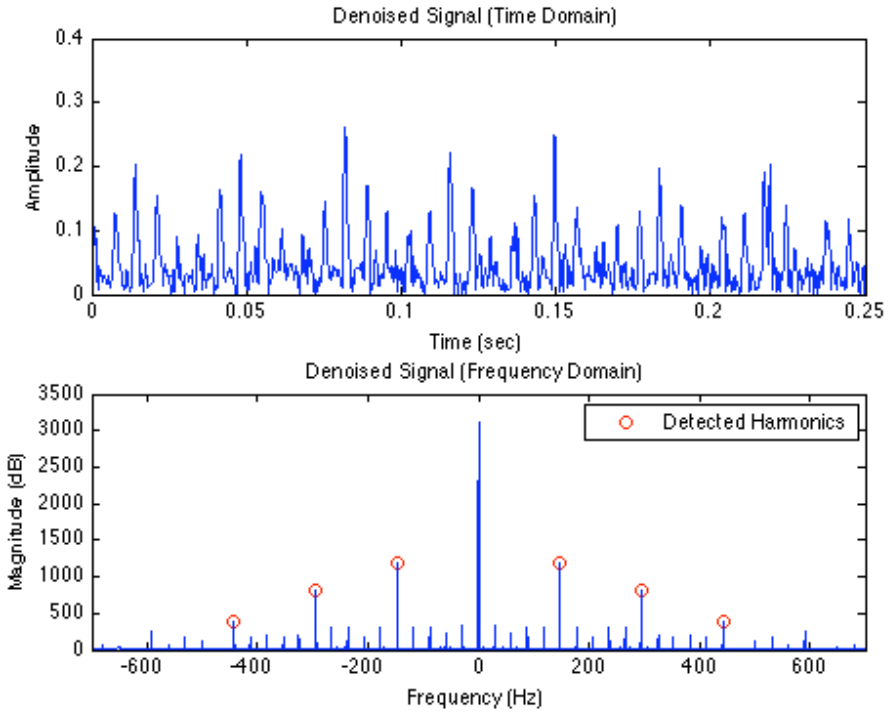

Figure 52: Demodulated Bearing Fault using the CWT/DRFS algorithm **(top)**. The frequency domain representation of the bearing fault **(bottom)** clearly shows, circled in red, four harmonics of the cyclic fault frequency of 140.76 Hz.

As an alternative de-noising method, the CWT/DRFS algorithm can easily be used to help specify not only the center frequency but also the bandwidth of an IIR Butterworth filter. The technique is used with the envelope demodulation algorithm used in this thesis. The center frequency of the filter can be located at the frequency that shows the maximum cyclic energy at the cyclic frequency of interest. Likewise, a rough estimate of the bandwidth of the filter can be estimated based on the amount of cyclic energy seen in the surrounding frequency bands.

4.4 The Discrete Reduced Fourier Series Threshold

This research also attempted to use the DRFS algorithm to threshold bearing fault signals using the CWT. The DRFS algorithm, in itself, is not an effective threshold for use with the wavelet transform. The first step of the algorithm is to perform a CWT of the noisy signal. Unlike the previous algorithm there is no restriction on the wavelet being analytic. Prior to application of a threshold the entropy of the wavelet scalogram is minimized for a set of shape factors. The scalogram with the least entropy, using SEM is chosen for the resulting thresholding.

To turn the DRFS into a threshold for use with wavelets the DRFS algorithm is modified further by assuming only a 'one-sided' periodic estimation. (i.e. using only positive harmonic frequencies) Performing a one-sided DRFS algorithm results in a complex signal, denoted here as x_n. This complex signal is still a valid estimate of the periodic signal, however, because it is one-sided and has separated both amplitude and phase information. By effectively scaling the amplitude information in x_n between 1 and 0 a periodic threshold T_{DRFS} can be created, given as,

$$T_{RDFS} = \frac{|x_n| - \min(|x_n|)}{\max(|x_n|)} \tag{59}$$

Figure 53 gives an example of a periodic signal, the associated developed threshold based off a one-sided DRFS estimate and the resulting signal after the threshold has been applied.

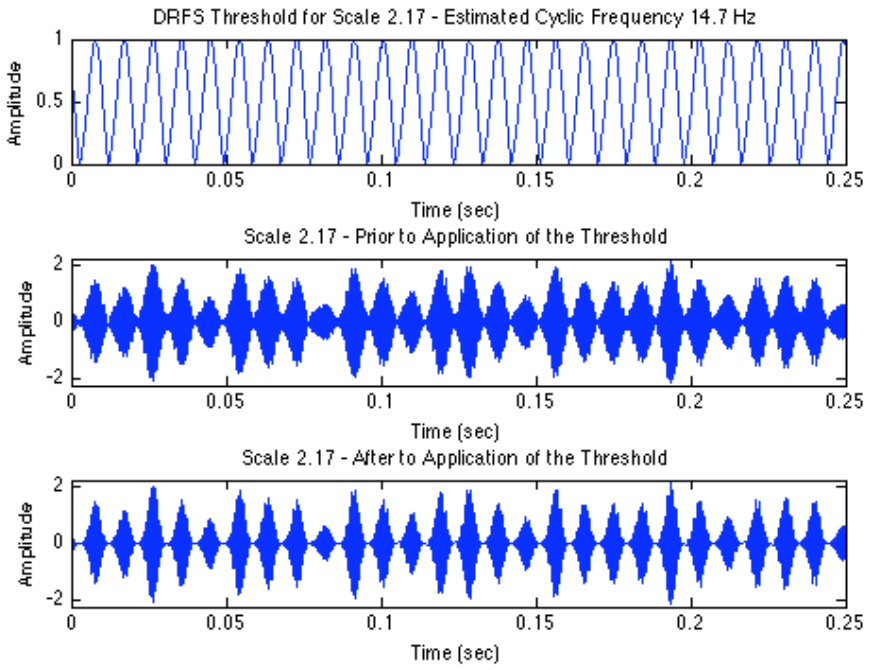

Figure 53: Example of a scale from a wavelet transform before and after a one-sided DRFS periodic threshold has been applied.

In Figure 53, the period of the signal was known while the one-sided DRFS estimate of the signal was created. It is not always the case that the scale to which the DRFS estimate will be applied will have components of the fault frequency present. It has been shown that kurtosis is a good indicator of the presence of a fault at a particular scale [65]. This threshold technique relies on kurtosis to select the wavelet scales that have fault characteristics. Prior to the inverse CWT of the wavelet transform, after the threshold has been applied, the kurtosis of each scale is calculated. Once inverse wavelet transform is performed the kurtosis of the final transformed signal is then calculated. An iterative procedure then begins to maximize the kurtosis of the final transformed signal. The scale with the lowest kurtosis value is set to zero effectively removing that scale from the scalogram. The inverse CWT is then applied to the new scalogram and a new kurtosis value of the de-noised is calculated. This procedure is repeated, selectively removing scales with low kurtosis, until the final de-noised signal has maximum kurtosis.

Figure 54 contains a flowchart of the proposed algorithm.

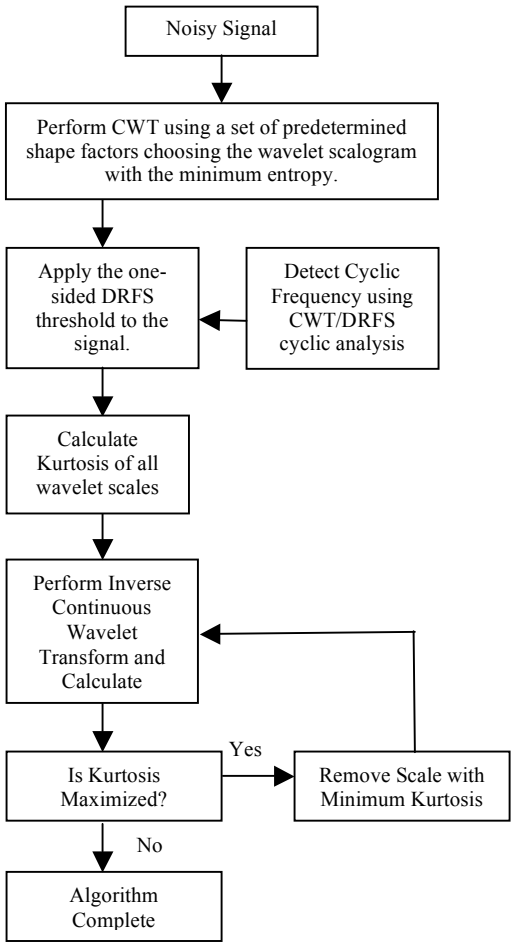

Figure 54: Flow chart of DRFS Signal Estimation algorithm.

The effectiveness of this algorithm can be seen in its operation. Actual bearing fault data, that will be reviewed in the section entitled Specific Objectives, was taken and used to verify that the DRFS Signal Estimation algorithm worked on actual fault signals with mechanical noise. Data, sampled at 12 kHz, was taken from a vibration sensor placed next to a healthy bearing and a bearing with an outer race fault that were

mechanically coupled. Time domain signal and the scalogram of the vibration signal from the healthy bearing can be seen in Figure 55 and 56 respectively. No obvious damage exists.

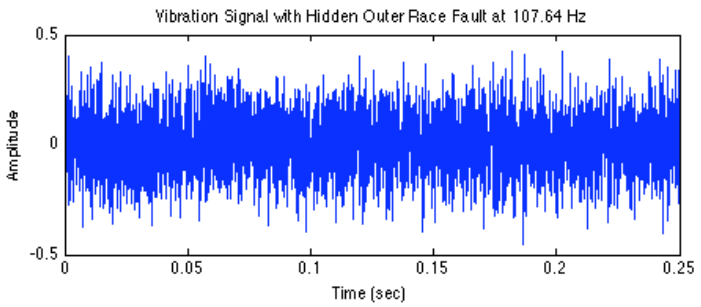

Figure 55: This is a time domain signal from a vibration sensor placed on a bearing with an outer race fault. No obvious fault signatures exist.

The outer race fault, as calculated, generates a periodic impulse at approximately 107.64 Hz, which is not evident in the scalogram of the healthy bearing seen in Figure 56. This scalogram was created after optimizing the shape factor, *b*, of the wavelet through the entropy method previously described.

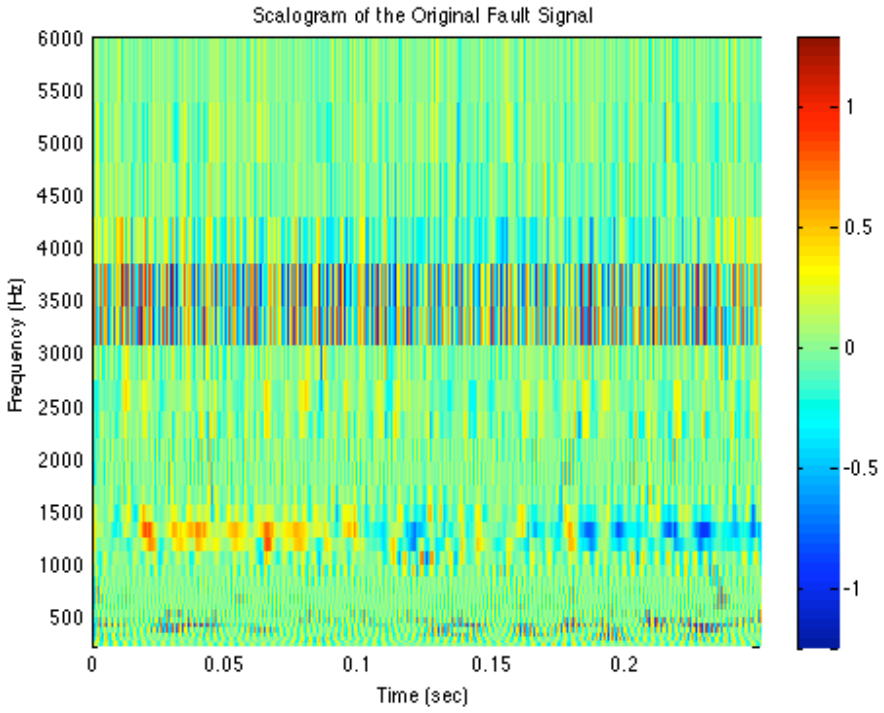

Figure 56: Scalogram from the vibration signal given in Figure 55.

The DRFS algorithm was implemented on each scale using the calculated fault frequency. The results of the DRFS threshold can be seen in Figure 57 for scale 2.17 (approximately 2764 Hz) where the scale is modified to allow only a 107.64 Hz cyclic frequency.

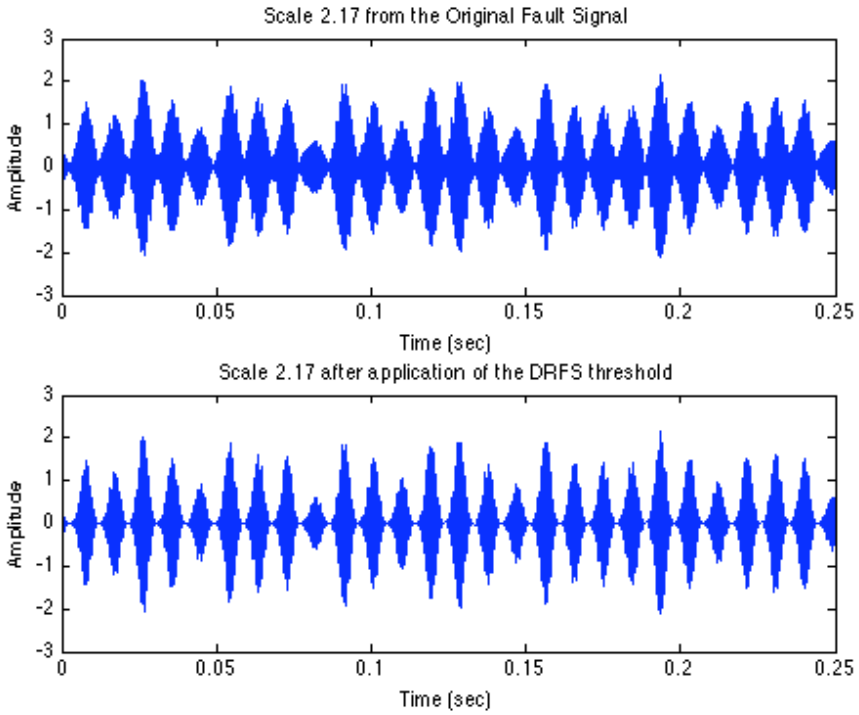

Figure 57: This figure shows scale 2.17 of the wavelet transform **(top)**. Using the DRFS as a threshold is it possible to produce an estimate of the given scale that is restrained to be periodic at the assumed fault frequency **(bottom)**.

Once the DRFS algorithm has been applied to each scale a new scalogram exists that has to be optimized to maximize the kurtosis of the inverse CWT. The kurtosis of all scales is then evaluated and can be seen in Figure 58.

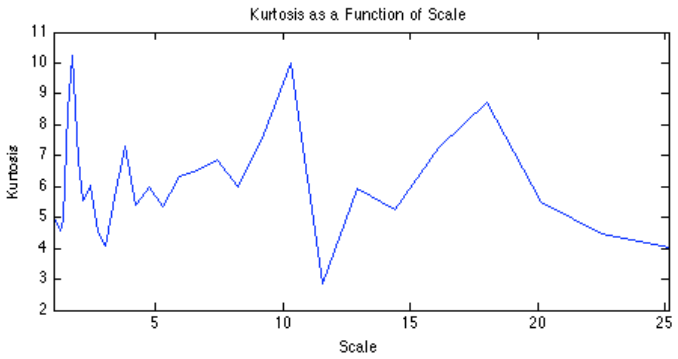

Figure 58: This figure gives kurtosis values for each scale of the wavelet transform. (The wavelet scales are logarithmic in increasing value.)

The kurtosis of the inverse CWT is repeatedly calculated, removing wavelet scales with the least kurtosis in sequential order from the scalogram. The final scalogram that maximizes kurtosis can be seen in Figure 59.

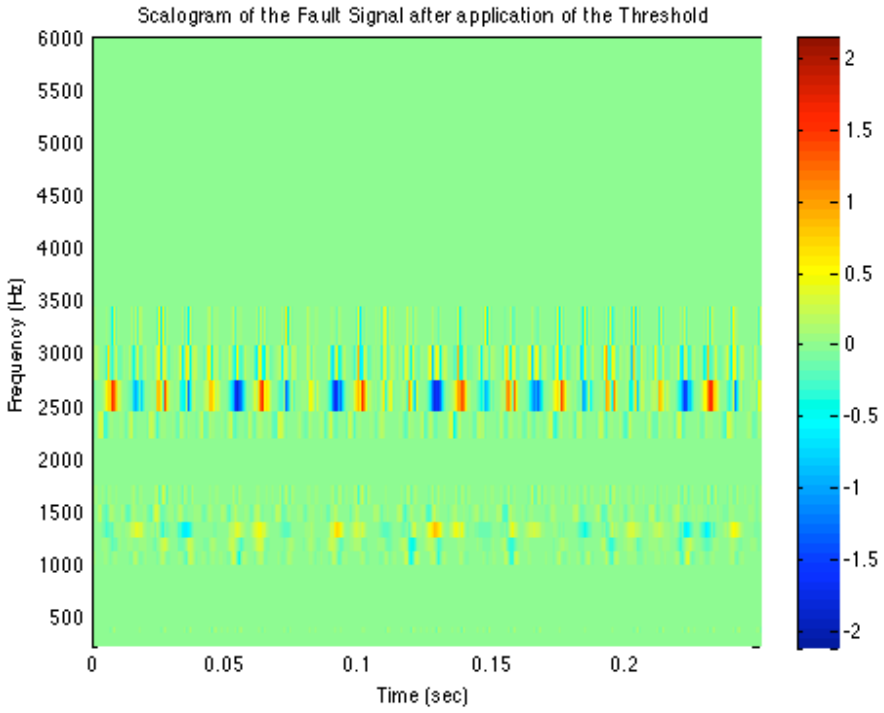

Figure 59: This figure gives final de-noised scalogram after kurtosis of the signal has been maximized.

The resulting de-noised signal can be seen in Figure 60. The periodic fault characteristic is clearly visible.

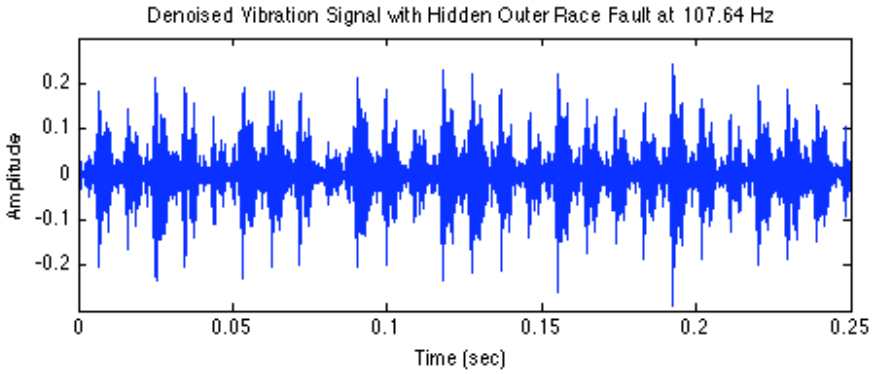

Figure 60: This figure gives final de-noised signal after the DRFS de-noising algorithm has been applied. The bearing fault at 107.64 Hz can clearly be identified.

The major drawback to this algorithm is computation time. The DRFS must be estimated across the entire frequency spectrum for each wavelet scale. This greatly increases computation time. Furthermore estimation of the wavelet scales that produce the maximum kurtosis of the inverse wavelet scalogram can also be very time intensive as this can require a large number of inverse CWTs to be computed.

5. BEARING FAULT DENOISING: CASE STUDY

Two major studies were undertaken in this thesis. The first study compared the three wavelets, the Morlet wavelet, the Impulse Response wavelet and the Frequency Domain Impulse Response Wavelet, that have been introduced for bearing fault detection using the CWT. To date no research was observed that gave an accurate comparison of the best wavelet that accurately captures a rolling element bearing fault. The second study took the algorithms introduced in this thesis and implemented them on experimentally collected bearing fault data to provide a clear estimate of how the algorithms performed with respect to each other and the traditional technique of envelope demodulation.

5.1 Wavelet Comparison

The results of the CWT are very dependent on the wavelet used for detection. The Morlet wavelet is the dominate wavelet chosen in bearing fault research using the CWT. In this research a new wavelet, the Frequency Domain Impulse Response Wavelet, was proposed. This case study compared the Frequency Domain Impulse Response Wavelet against two other wavelets that have been proposed in literature Morlet Wavelet and the Impulse Response Wavelet. [47, 57] These wavelets can be seen in Table 1. Figure 61 shows each of the wavelets in the time domain at scale 50.

Table 1: This is a list of the wavelets and their expressions that will be used in this study. *This wavelets expression is denoted in the frequency domain where ω denotes frequency in rad/s.

Morlet Wavelet [10]	$\psi(t) = e^{-bt/2}\cos(\pi t)$
Impulse Response Wavelet [47]	$\psi(t) = \begin{cases} e^{-bt}\sin(\pi t) & t > 0 \\ e^{bt}\sin(\pi t) & t \le 0 \end{cases}$
Frequency Domain Impulse Response Wavelet*	$\psi(\omega) = \left(\frac{j\omega}{j\omega + \frac{\pi}{5}}\right)\left(\frac{\pi}{(j\omega - a)^2 + \pi^2}\right)$

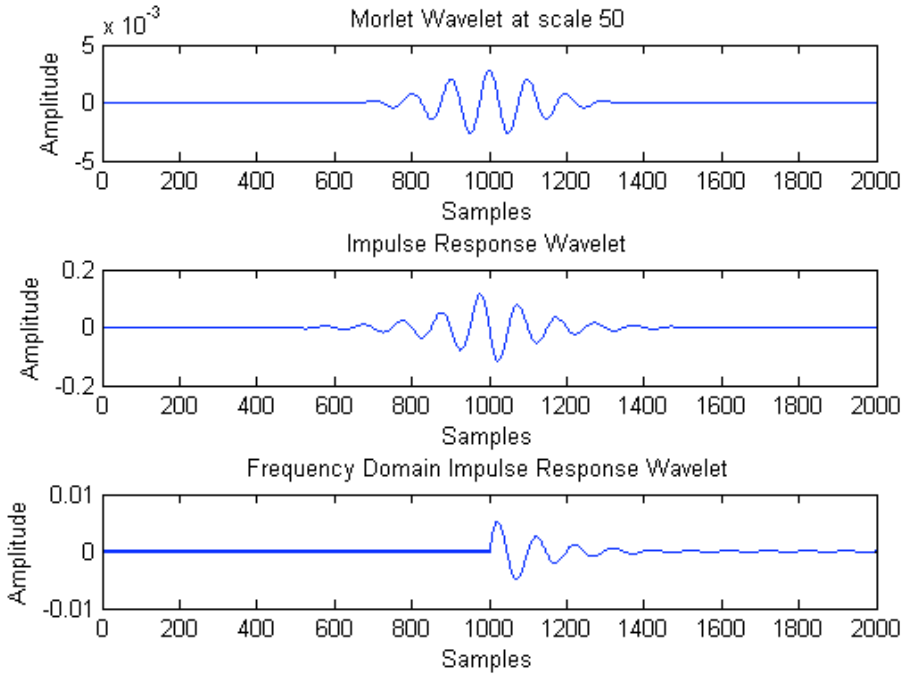

Figure 61: The Morlet, Impulse Response and Frequency Domain Impulse Response Wavelet at scale 50.

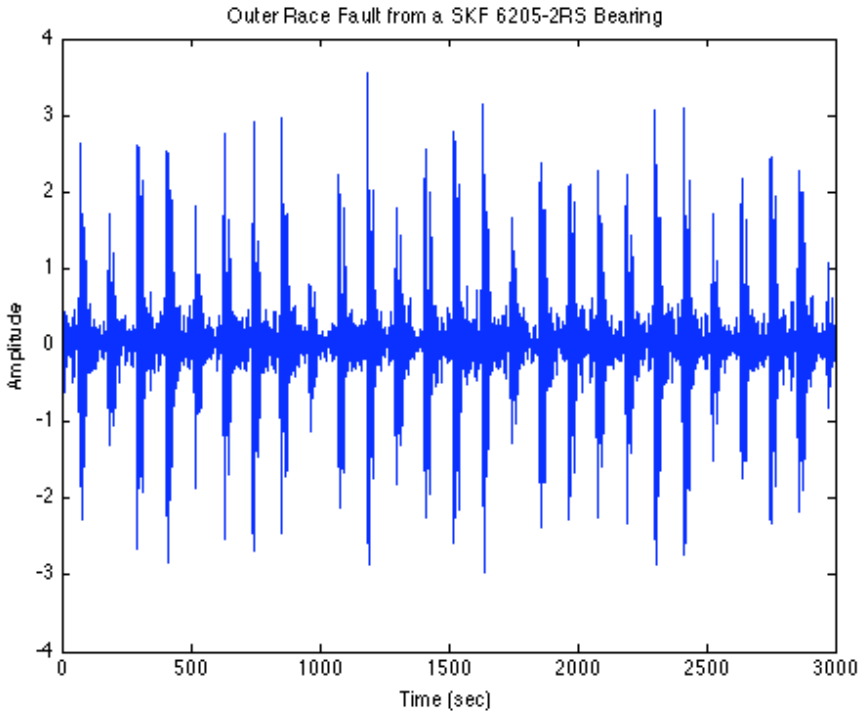

Figure 62: Outer race fault from a SKR 6205-2RS bearing that will be used for a wavelet comparison study.

The purpose of this study was to show which wavelet best typified bearing fault signals. Two different comparisons were attempted using data taken from the Case Western Reserve Bearing Fault Data Center run by Dr. Kenneth Laparo. [67] The first comparison attempted to compare how the wavelets performed based on how well they extracted a known fault signal from added levels of white noise. Increasing random levels of white noise were added to fault data taken from a SKR 6205-2RS bearing that clearly showed an incipient outer race fault. The fault data that was used can be seen in Figure 62 and, as can be seen, the actual fault is very clearly defined in the time domain.

The fault seeded on the bearing was 0.007 inches wide and 0.011 inches deep and the vibration data has been sampled at 12 kHz.

To perform this evaluation each of the three wavelets were used in a CWT decomposition of the bearing fault signal in Figure 62 with Shannon's Entropy Measure (SEM) being used to select the shape factor for each wavelet. SEM is a popular measure for determining how well a particular signal has been transformed into a scalogram. [57, 59] For convenience SEM is repeated from (60) below.

$$E_S(p) = -\sum_i p_i \log_2(p_i) \tag{60}$$

To effectively use SEM to evaluate a wavelet decomposition the CWT of the signal of interest was first performed using the wavelet in question. The scalogram was then normalized between zero and one. The entropy E_s was then calculated over the whole scalogram. If the wavelet matched dominate portions of the signal the resulting scalogram had very large peaks representing very high correlation values. When normalized this creates a very sparse distribution of values that either are close to zero or close to one. SEM can be seen as a non-linear transformation as seen in Figure 63. Values that are close to one and zero, map to low entropy values, and values that are between one and zero map to high entropy values.

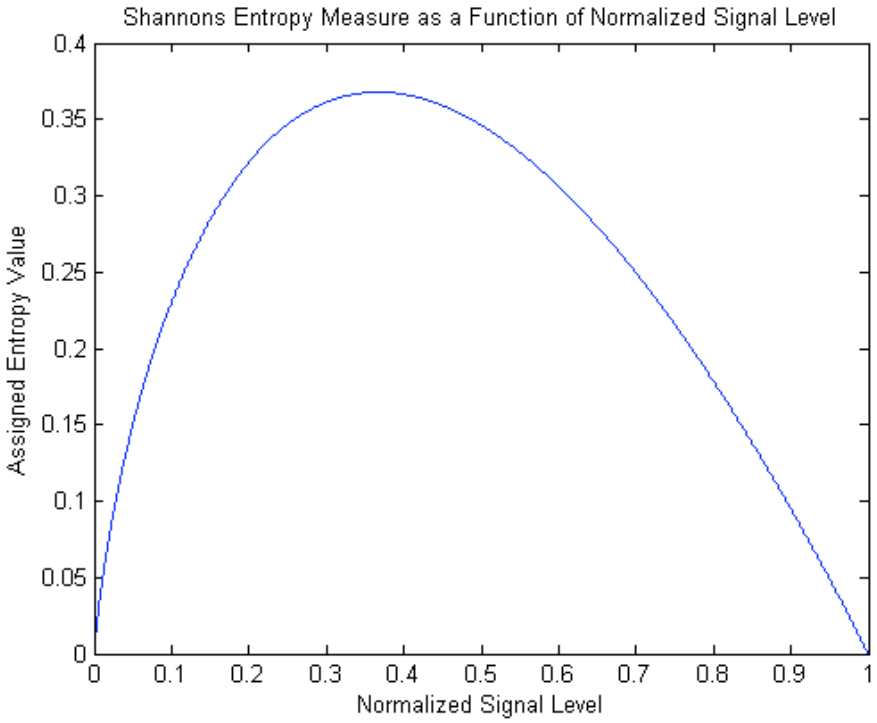

Figure 63: Shannon's Entropy Measure (SEM) as a function of normalized signal level.

The Morlet, Impulse Response and Frequency Domain Impulse Response wavelet were used to decompose the signal seen in Figure 62 with various levels of added white noise. After optimizing the CWT for each wavelet's shape factor based on SEM, the wavelets were then ranked using SEM (60) as the evaluation criterion.

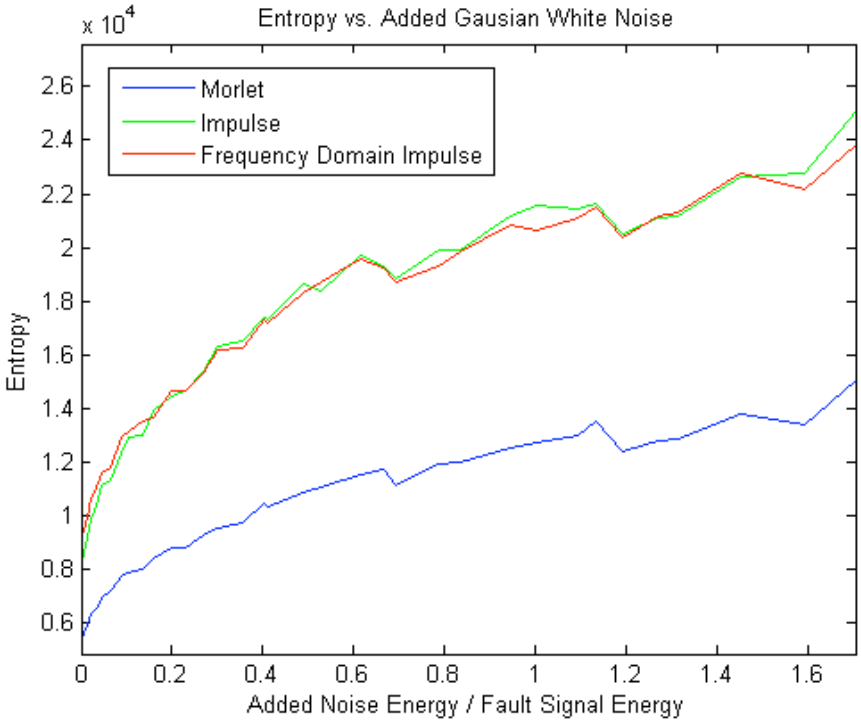

Figure 64: Entropy with respect to added noise entropy for CWT with 3 different wavelets.

The results of this comparison can be seen in Figure 64. The Morlet wavelet performs better than both the Impulse and Frequency Domain Impulse Response Wavelets. This can be seen by the lower entropy values given for the Morlet wavelet across all additive noise values. The effect of adding white noise to the bearing fault signal did not change the trend of the results. This could be predicted as white noise corrupts all frequency bands in the fault signal equally without adding dominant structures to the signal. Additive mechanical noise could easily provide dominant structures that might correlate to the Morlet wavelet stronger than the Impulse and Frequency Domain Impulse Response Wavelet.

The second comparison examined how the wavelets represented three different cases of bearing faults: inner race faults, outer race faults and rolling element faults that were corrupted with various amounts of mechanical noise. Fault signals from two bearings were evaluated, an SKF 6205-2RS and an SKF 6203-2RS. This gave a total of six different fault signals for the test. For each signal, each wavelet was optimized to find the shape factor that minimized SEM criterion producing the smallest amount of entropy for ten time segments consisting of 8192 samples or 0.68 seconds from each signal. Upon optimizing the shape factor and identifying the entropy the resulting wavelet scalogram was visually inspected to ensure that optimization procedure for the shape factor was optimized for the fault signal and no other dominate structure in the signal, as is possible. The results were then averaged and the variance calculated for each wavelet with respect to the six different signals. Samples of the six fault signals can be seen in Figure 65.

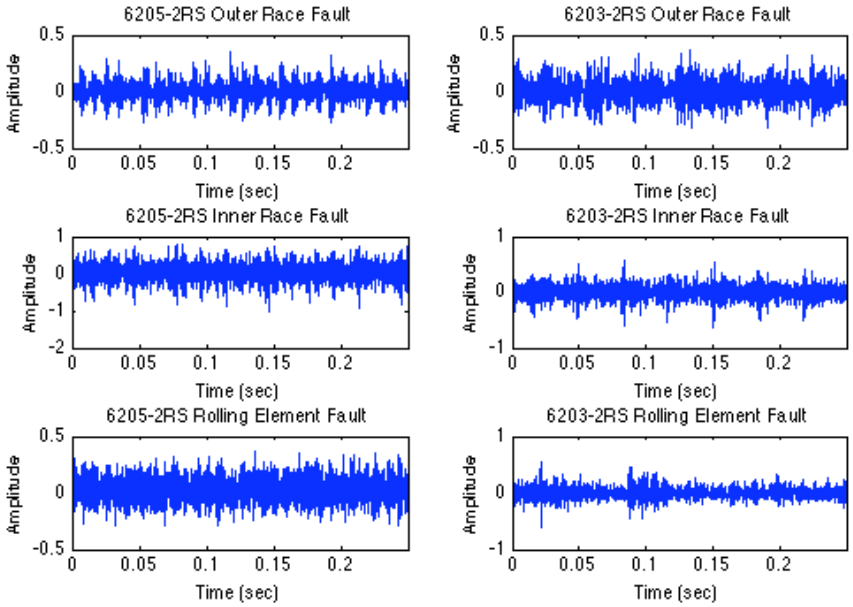

Figure 65: Samples from the six signals used to compare the Morlet, Impulse Response and Frequency Domain Impulse Response wavelets.

The results of the comparison are given in Figure 66 and Figure 67.

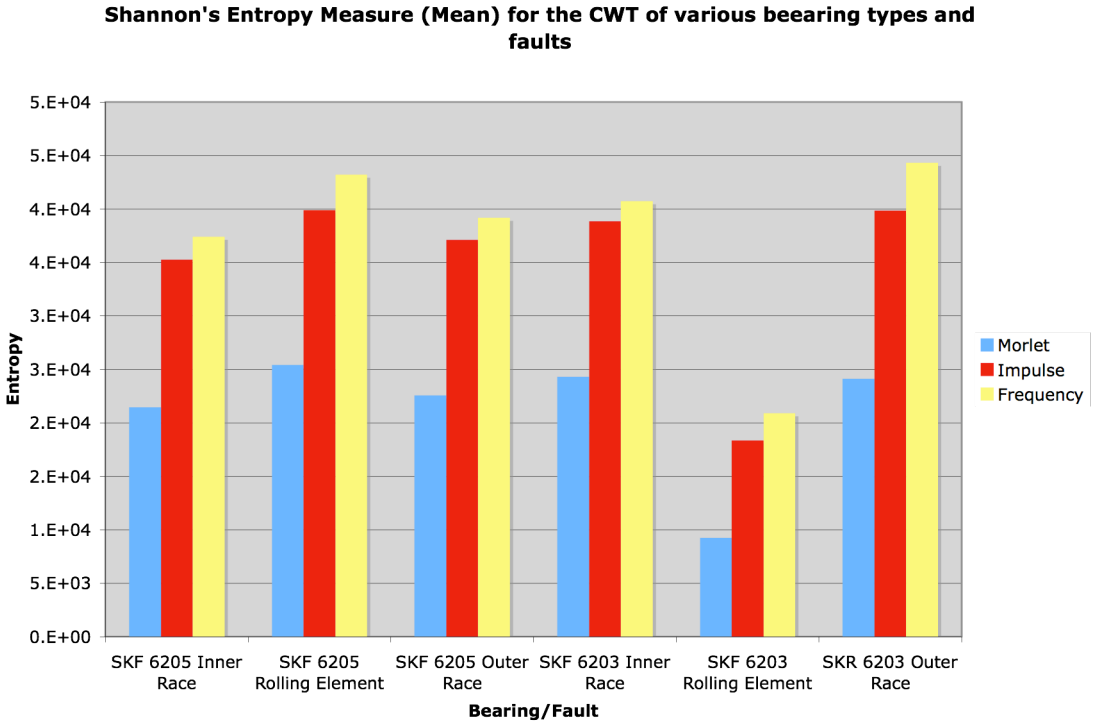

Figure 66: Mean calculated entropy for six fault signals of different types from different bearings.

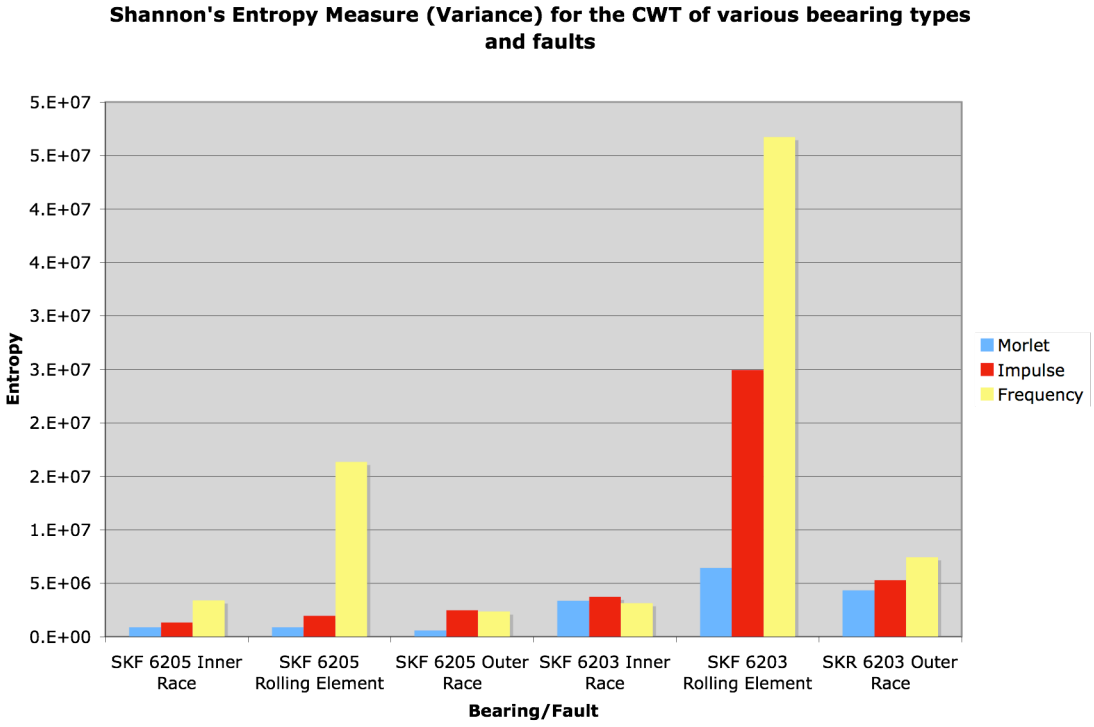

Figure 67: Variance of entropy measurements for six fault signals of different types from different bearings.

As can be seen in Figure 66, the Morlet wavelet for the three different cases examined consistently shows on average less entropy. Furthermore, for the Morlet wavelet, there is also much less variance between the ten trials for each signal. As a conclusion, it could be said that the Impulse Response and Frequency Domain Impulse Response wavelet show less ability to classify the six signals tested in this study. It is impossible to say that the wavelets actually correlated with the fault signal of interest. This renders this study inconclusive.

Both the Impulse Response and Frequency Domain Impulse Response wavelets were built off the assumption of the fault signal modeled by an exponentially decaying

sinusoid. These results indicate this assumption, assuming the wavelets correlated with the fault, is incorrect. As the rolling element impacts the developing fault the resulting vibration signal is very dependent on the actual fault structure. As the fault develops the path the rolling element takes as it rolls over the fault not well defined. Furthermore, each fault for all bearings has slightly different surface characteristics as it grows.

For very incipient faults where there is only a very slight nick or indentation in the surface of the outer race, inner race or the rolling element it is possible that the Impulse Response or Frequency Domain Impulse Response wavelet better matches the fault signal, however with the data set that was used in this study it is impossible to verify this assumption. Even if there were fault signals with ideal incipient faults available, the signal energy would be so weak as to be questionable if the fault signal could even be effectively detected, with or without any signal de-noising. A wavelet, developed for the Discrete Wavelet Transform (DWT), an orthogonal version of the CWT, using actual fault data from an incipient fault was developed by imbedding a vibration sensor on the outer race of a bearing with a very small outer race fault [50]. The resulting wavelet very closely matched an exponentially decaying sinusoid. This indicates that for other data sets the Impulse Response Wavelet or the Frequency Domain Impulse Response Wavelet might produce better results. With this said, this study found it impossible to verify that the developed wavelet, the Frequency Domain Impulse Response Wavelet, was in any way better than either the Morlet or Impulse Response Wavelet for bearing fault signals in general.

5.2 Algorithm Comparison

Three different possible bearing faults exist: inner race, outer race and rolling element fault. Each of these faults show similar fault characteristics, as all are modeled as periodic signals showing the characteristic exponentially decaying sinusoid, however each fault type has specific differences that make the possibility of detection slightly different. Outer race faults are the easiest to detect as they show the most consistency between rolling element impacts; however, they are dependant on the location of the load. The strength of inner race faults are often modulated due to the load applied on the bearing. Both inner and outer race faults show a high degree of periodicity especially at the incipient stage with periodicity gradually fading as the fault severity increases. Slipping between the rolling elements and the inner and outer race are intrinsic to the geometry of the rolling element bearing and, as a result, even in the best-case scenario, the inner and outer race fault signals are never completely periodic. Rolling element faults are the most difficult to detect as even at the incipient stage the impulses as the fault strikes a rolling surface are aperiodic and have been termed by some researchers as chaotic.

The impulse signature of the faults themselves can also be very erratic. The naïve assumption is that the fault signature is an exponentially decaying sinusoid. While this could possibly be true for very incipient bearing faults, realistic faults are generated from the rolling elements of the ball bearing traveling over rough and uneven surfaces. The true fault signature can therefore only be a function of the fault itself and how the rolling element travels across the fault. Even at the incipient stage it is improbable that the fault

itself could be considered simple enough to only generate an exponentially decaying sinusoid and that the path of each rolling element would be identical across the fault.

The additive noise that covers the fault signal is also very dependent on the bearing and the equipment to which the bearing is coupled. A large number of de-noising studies use white noise in an attempt to model mechanical noise. This is a facile assumption as white noise corrupts the entire frequency spectrum with an equivalent noise level. Mechanical noise, in a very general sense can be considered periodic. This creates a set of very strong discrete harmonics in the frequency spectrum. The actual fault signal is usually overwhelmed by the strength of the mechanical noise in the signal rather than the background noise that overwhelms all frequency spectra and can be represented by white noise.

All the factors mentioned play into creating a difficult signal de-noising problem and are one of the reasons why research is very prevalent in this field. With this said, not only is it difficult to detect the fault signal in real experimental data, but it is also very difficult to model the actual fault vibration signals. For this study the algorithms introduced in this thesis were compared on the basis of experimental data taken from Case Western Reserve University's Bearing Data Center run by Dr. Kenneth Laparo [67]. Separate cases of inner, outer and rolling element faults of steadily increasing width, from 0.007 inches to 0.04 inches and a depth of 0.011 inches, were seeded on SKF 6205-2RS and SKF 6205-2RS bearings using electro-discharge machining and vibration data collected at 12kHz. The vibration sensors used to collect this data were PCB Piezotronics accelerometers, model number 352C65. These accelerometers have a flat

spectral response from 1 to 10kHz within 5% tolerance. The resonate frequency of this sensor is greater than 35 kHz, well above the Nyquest frequency of 6kHz, used in this study. The in a number of cases the vibration data itself clearly shows the impulses associated with faulty bearings clearly evident in the time domain. Figure 68 has examples of the collected incipient 0.007 inch outer race fault, inner race fault and rolling element fault.

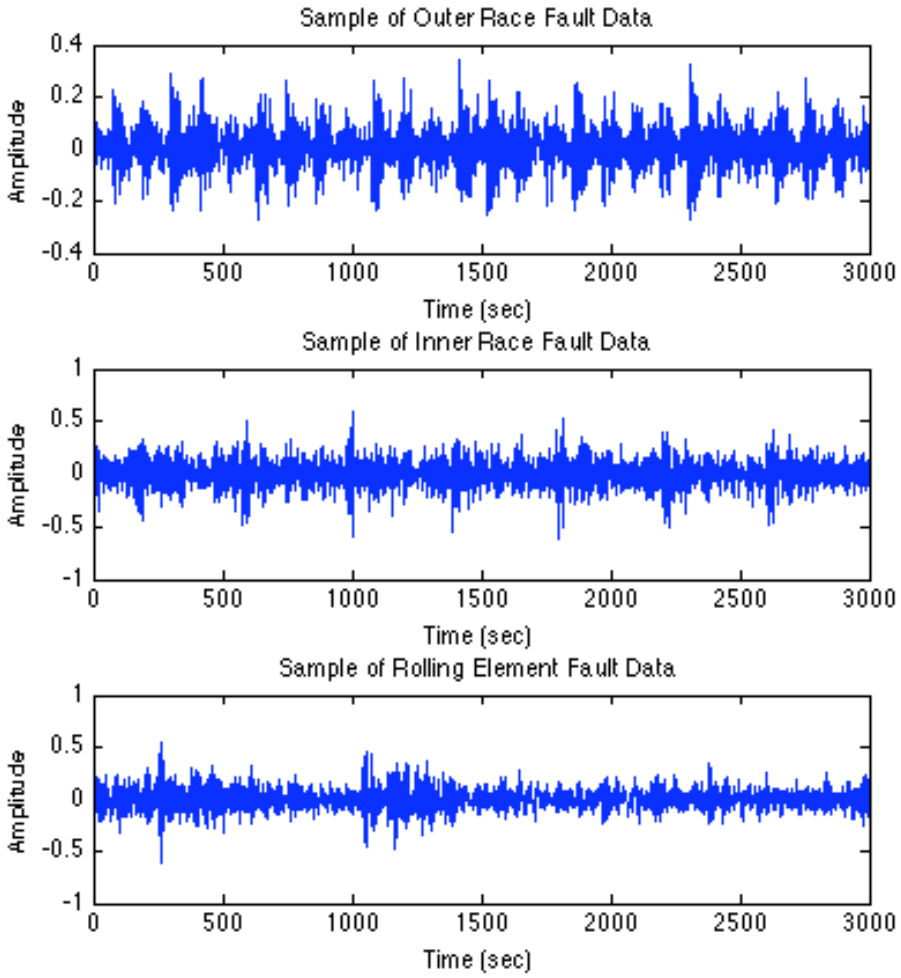

Figure 68: Example of the outer race, inner race and rolling element CWR bearing fault data.

What is unique about this data is that for a given faulty bearing there were three sets of vibration signals recorded. One vibration sensor was located directly on the faulty bearing. For the most part these vibration sensors showed the fault clearly visible. Two other vibration sensors were placed on healthy bearings that were mechanically coupled to the faulty bearing. These vibration signals showed varying degrees of noise with usually one of the vibration signals overwhelmed with noise. To test the given algorithms in this thesis, six vibration signals were taken from the aforementioned healthy bearings to detect the faults described and the algorithms introduced were used to enhance detection. To give a well rounded approach to the study, three vibration signals, from two different bearing types each, were taken from aforementioned healthy bearings that had fault signatures from inner race faults, outer race faults and rolling element faults. Table 2 gives details on the several sets of data. Data can be provided upon request or downloaded from the Case Western Reserve Bearing Data Center [67].

Table 2: Bearing fault data from the Case Western Reserve University's Bearing Fault Center. All faults were seeded using EDM and are 0.007 inches in width and 0.011 inches in depth.

Bearing	Data Set Name	Fault Type
6205-2RS JEM SKF	130.mat	Outer Race Fault
	118.mat	Rolling Element Fault
	105.mat	Inner Race Fault
6203-2RS JEM SKF	302.mat	Outer Race Fault
	282.mat	Rolling Element Fault
	278.mat	Inner Race Fault

From the results of the first study it was determined that the Morlet wavelet was the best fit for the data set introduced in this thesis. Consequently, only the Morlet wavelet was implemented for all algorithms. This allowed for a standard set of scales to be used. For all studies the CWT used 30 scales were computed logarithmically (base 10) space between 1 and 25. The magnitude of the frequency response of the CWT for l0 consecutive scales using the Morlet wavelet, can be seen in Figure 69.

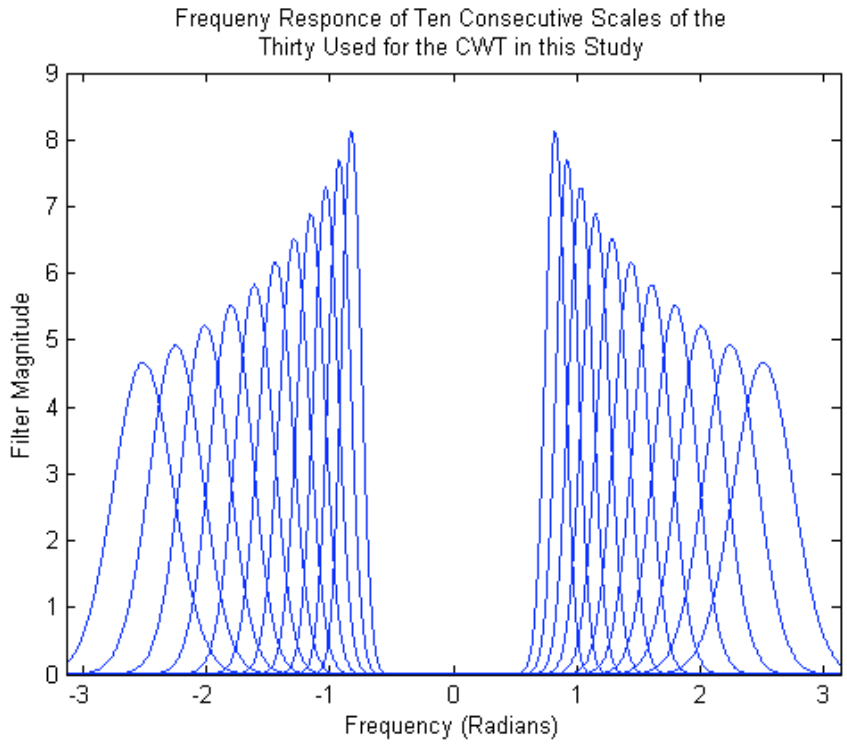

Figure 69: This figure shows 10 of the 30 wavelet scales used when computing the Morlet wavelet. As can be seen the scales provide clearly overlapping filter responses capturing the majority of the signal.

As can be seen, this created an overlapping filter bank that was sparse enough in scale to be computationally efficient but provide a low mean squared error (MSE) upon reconstruction using the inverse CWT. To demonstrate this ten trials, using the forward and inverses CWT, were performed on signals consisting of white noise with 65536 data points, and the mean squared error (MSE) was determined to be, on average, 0.0373 with a variance of 2.1813e-6. For each algorithm used in this study there were a number of input parameters that controlled the algorithms operation. For the automated CWT/DRFS algorithm the 5 sets of data each 4096 data points long were each used to estimate the CWT/DRFS Cyclic Energy distribution. The wavelet scale that then produced the highest amount of cyclic energy was then used to de-noise the entire 120000 signal. DRFS thresholding is a very computationally intensive algorithm. To minimize the computational overhead of the algorithm, the signal to be de-noised, 120000 samples long was broken into 29 individual data segements that were thresholded independently. The first set of 4096 points of the signal was used to train the algorithm and determine the optimum shape factor and scales that maximized kurtosis. For envelope demodulation, the two parameters that need to be set, the filter center frequency and filter bandwidth, were determined using the CWT/DRFS cyclic energy algorithm. The first 3000 data points of signal that were assigned to be de-noised were used to train the Autoregressive Filter/Minimum Entropy Deconvolution algorithm. Finally for the Weak Signature Detection algorithm the first 6000 data points were used to estimate the period of the fault and determine the scale to de-noise the fault signal.

As mentioned, all algorithms were implemented using 30 scales logarithmically spaced from 1 to 25 and the Morlet wavelet. Prior to implementation, the shape factor that minimized SEM was also determined using the first 4096 data points of the signal to be de-noised for the Weak Signature Detection and DRFS based algorithms. The other thresholding algorithms, hard, soft, hybrid and mechanical thresholding used a shape factor of 0.3. This guaranteed the overlapping filter effect seen in Figure 59. The input parameters for each algorithm can be seen in Table 3.

Table 3: Input arguments for each algorithm used in this study.

Automated CWT/DRFS De-noising	
Samples used to compute the CWT/DRFS cyclic energy distribution	4096 samples
Number of Averages	5
Cyclical Frequency Range	+/- 5 Hz of the estimated cyclic frequency
DRFS Thresholding	
Analysis Length	29 units of 4096 samples
Envelope Demodulation	
Center Frequency	User Defined
Bandwidth	User Defined
Autoregressive Filter / Minimum Entropy Deconvolution	
AR Training Length	50
MED Training Length	3000
Weak Signature Detection	
Period Range	From 5 samples / 6000 samples to 3000 samples / 6000 samples
Training Length	6000

5.3 Evaluation Criterion

To evaluate the ability of the algorithms to de-noise various signals, a standardized measurement system was implemented. Prior to comparison, all de-noising results, if needed, were demodulated. For this study, the Fourier Transform of the given demodulated de-noised fault signal was computed. This transform was then used as the basis for the evaluation criterion. The Power Spectral Density (PSD) of the signal was considered as a measurement tool; however, as suggested in literature, a very long Fourier Transform was more effective at extracting harmonics in the frequency domain of the periodic fault signals [4]. With only a limited amount of data, approximately 10 second for each fault signal, a PSD of the signal, with the length of the Fourier Transform suggested for fault identification, was not feasible. All signals were sampled at 12 kHz, leading to a Fourier transform of 120000 data points with a frequency resolution of approximately 0.0983 Hz per frequency bin.

The harmonics of the signal in the frequency domain were then located at the estimated cyclic frequency. To determine the effectiveness of the de-noising algorithm the number of detectible harmonics, 3dB above the surrounding noise, were counted, starting at the first fundamental harmonic and increasing in frequency. The count was stopped upon reaching a harmonic that had surrounding noise with amplitude greater than the magnitude of the harmonic being evaluated minus 3dB. The surrounding noise was defined as the area around the harmonic of interest bounded by the two closest harmonics. The Fourier Transform of a de-noised fault signal, in this case using envelope demodulation, can be seen in Figure 70, along with the detected harmonics highlighted.

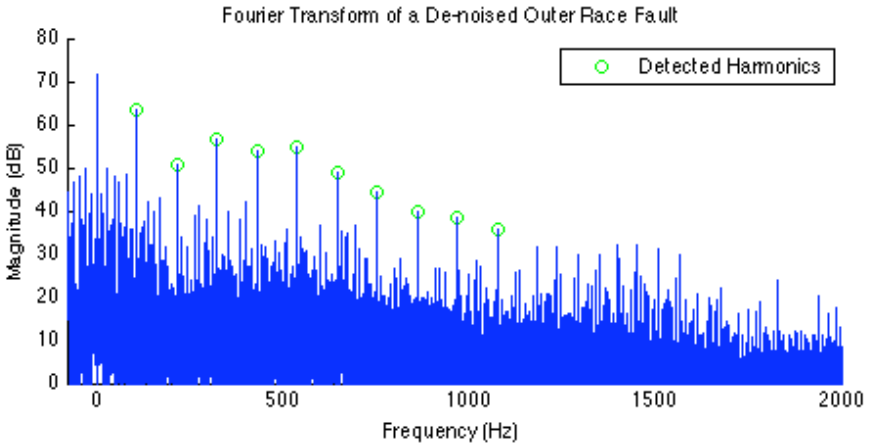

Figure 70: Detection method used to measure the effectiveness of bearing fault de-noising methods.

Using the detection method presented the following four criteria were used for evaluation of how well the fault signal was de-noised:

Number of harmonics detected. This gave a measure of how easy it was to visually detect the signal. With more harmonics present, the signal in the time domain becomes more prevalent, indicating sharper impulses present in the time domain.

SNR: Ratio of the energy of the detected harmonics with respect to the overall signal energy. For a truly periodic signal all of the energy of the signal would be located in discrete harmonics. By calculating the energy in the harmonics and dividing it by the total energy in the signal, an approximate signal to noise ratio can be determined.

Kurtosis of the de-noised signal. Kurtosis, historically, has been shown to be a good indicator of the 'impulsiveness' of a signal. This is a good indicator of the presence of a rolling element fault where, because the fault is not necessarily periodic, it cannot clearly be detected in the frequency domain.

Computation time of the algorithm. Due to the slowly developing nature of bearing fault signals, real time detection is not necessary. As a result, much more sophisticated algorithms can be implemented to detect faults and computational time is not a large issue. Computation is an issue, however, when the algorithm is implemented for fault detection in a setting where a technician might be left waiting for a result from an algorithm. Obviously, in cases such as this algorithms that take large amounts of time are costly. All algorithms were clocked using a 1.5 GHz Power PC G4 with 1.25 GB of ram running in Matlab R14 on the OSX operating system.

5.4 Outer Race Faults

Outer race faults are the easiest to detect. Generally speaking the amplitude of the impulses is fixed and the periodicity of the fault is much more constant. Two cases were examined for this type of fault. Case one, taken from a SKF 6205-2RS bearing, was a visible fault signature with relatively low amount of noise. The fault depth was 0.011 inches and the diameter was 0.007 inches with characteristic fault frequency being ~107 Hz. The de-noising algorithms were run on the signal to do a base comparison on how well each algorithm extracted the fault. For this case all algorithms showed success providing a base line comparison prior to attempted to extract a fully hidden signal from noise where not all algorithms were expected to work well. The fault signal can be seen in Figure 71.

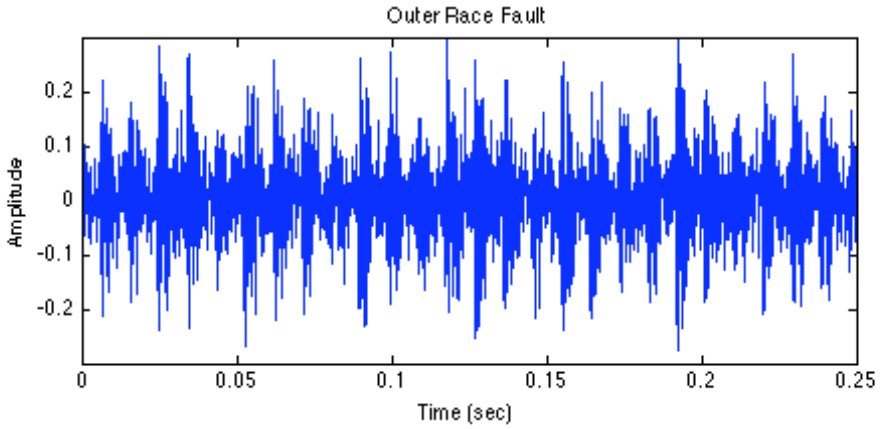

Figure 71: Outer race fault clearly visible in the time domain. The cyclic frequency of the fault is approximately 107 Hz.

Table 4 gives the results of de-noising the fault seen in Figure 71.

Table 4: Results from de-noising the outer race fault given in Figure 71.

Case 1	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS de-noising	2.0349	1	-2.9731	92.72
DRFS Threshold	6.9074	4	-7.7852	413.58
Envelope Demodulation	4.9356	10	-5.0367	NA
Hard Threshold	5.6399	7	-7.724	107.07
Hybrid Threshold	29.4773	8	-5.9774	131.11
Mechanical Threshold	18.5416	9	-5.9976	119.63
AR/MED	15.0807	9	-8.7975	65.78
Soft Threshold	5.8585	9	-7.2289	107.02
Weak Signature Detection	2.1148	0	-	144.83

The best de-noising signal for this set was obtained using Envelope Demodulation. It had the second lowest SNR of all signals and a large number of

harmonics that could be easily identified in the frequency domain. While the kurtosis of the de-noised signal was not the greatest, it is still very evident that this fault exists in the time domain. Both the frequency domain and time domain results of Envelope Demodulation can be seen in Figure 72.

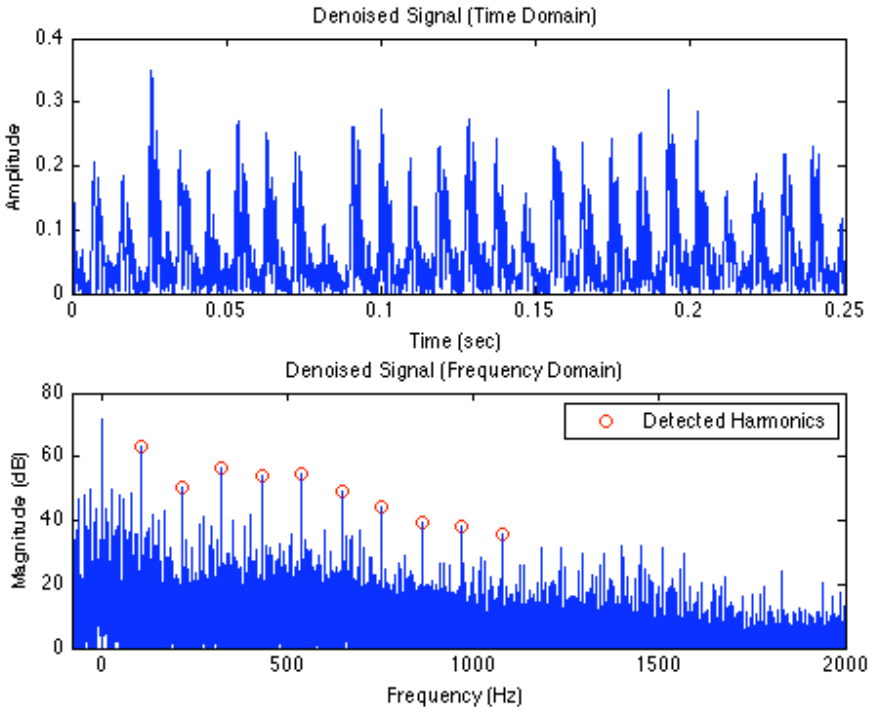

Figure 72: De-noised vibration signal from case one using Envelope Demodulation.

The results of Envelope Demodulation rely very heavily on identification of both the bandwidth and center frequency of the fault. The bandwidth, 1200 Hz, and center frequency, 3100 Hz, of the fault was located using the results of the CWT/DRFS cyclic

energy analysis as seen in Figures 73 and 74. Since this was not an automated procedure, the time that it took to develop this analysis was not quantifiable. As a result, there is no time given for calculation of the de-noising results given in Table 4.

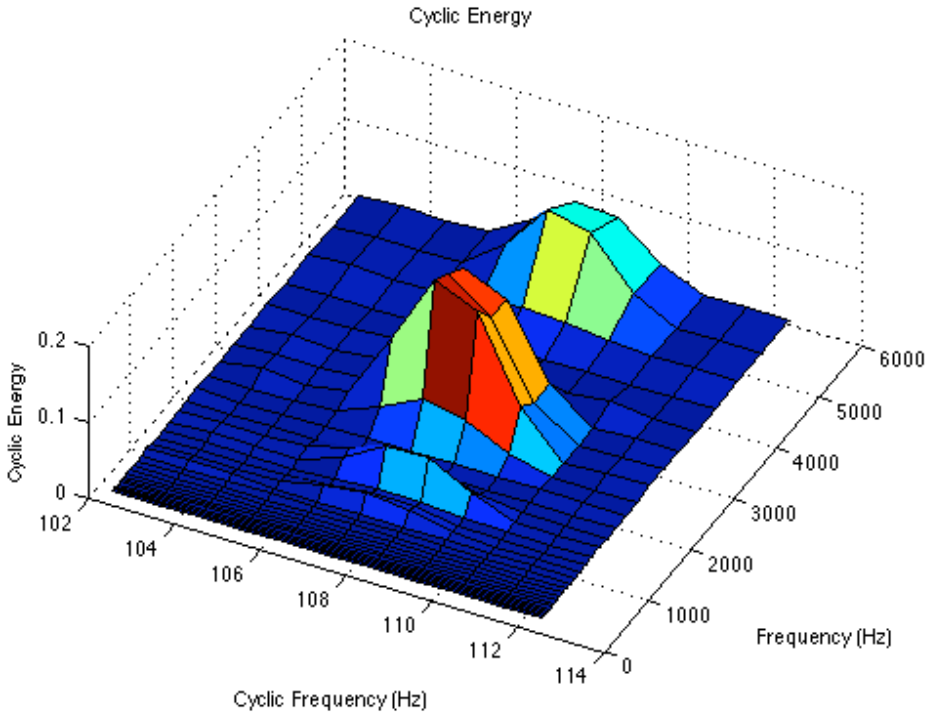

Figure 73: Cyclic Energy of the bearing fault given in Figure 71.

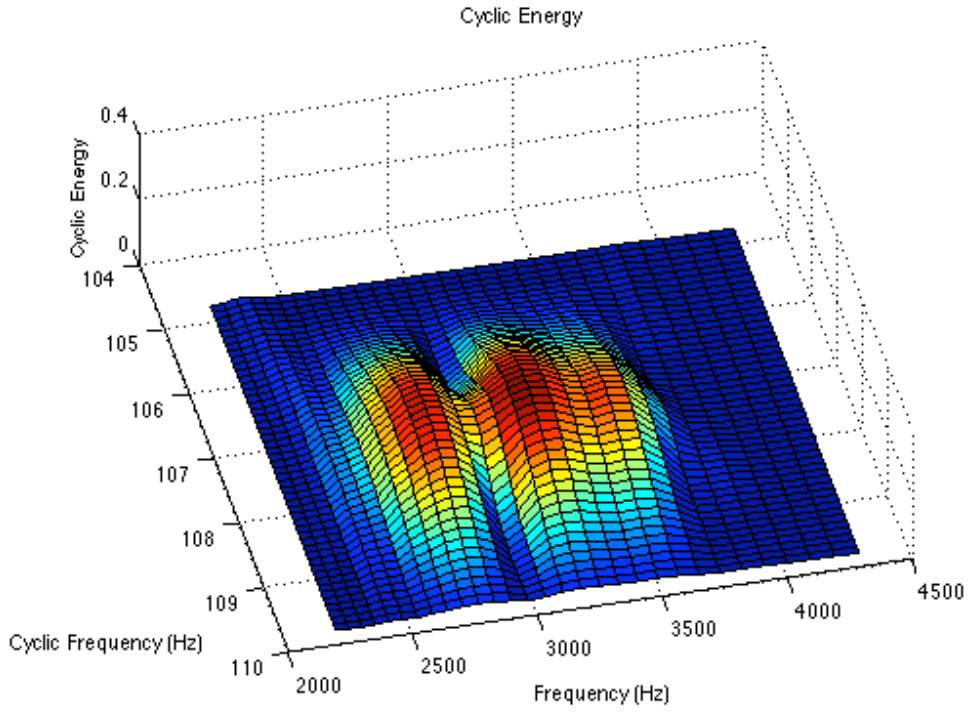

Figure 74: Recalculated the Cyclic Energy of the bearing fault given in Figure 71 over a frequency range and cyclic frequency range closer to the Cyclic Energy due to the fault.

The de-noised signal with the highest SNR was the automated result from the automated CWT/DRFS denoising algorithm and can be seen in Figure 75. The fault is clearly identifiable in both the time and frequency domain. Unfortunately the bandwidth of the wavelet scale limits the number of harmonics available for detection.

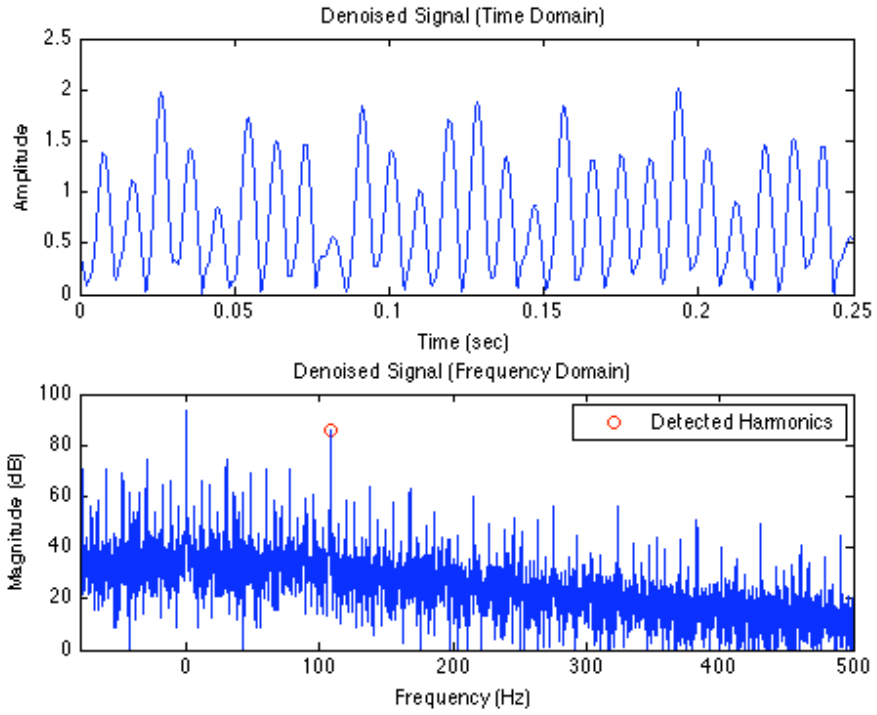

Figure 75: Results from CWT/DRFS wavelet de-noising.

Using Weak Signature Detection method to determine the most cyclic wavelet scale failed. The results of the de-noising attempt can be seen in Figure 76. Generally speaking, the noise covering bearing fault signals is from mechanical systems and, as with bearing faults, is periodic. What differentiates a bearing fault from general other vibrations is that it is modulated. Weak Signature Detection method does not take this into account. As seen in Figure 76, while the time domain signal shows periodic structure, it is at multiples of the running speed of the shaft speed and not the bearing fault frequency. As a result the algorithm totally missed identifying the fault finding the much stronger harmonics associated with shaft speed.

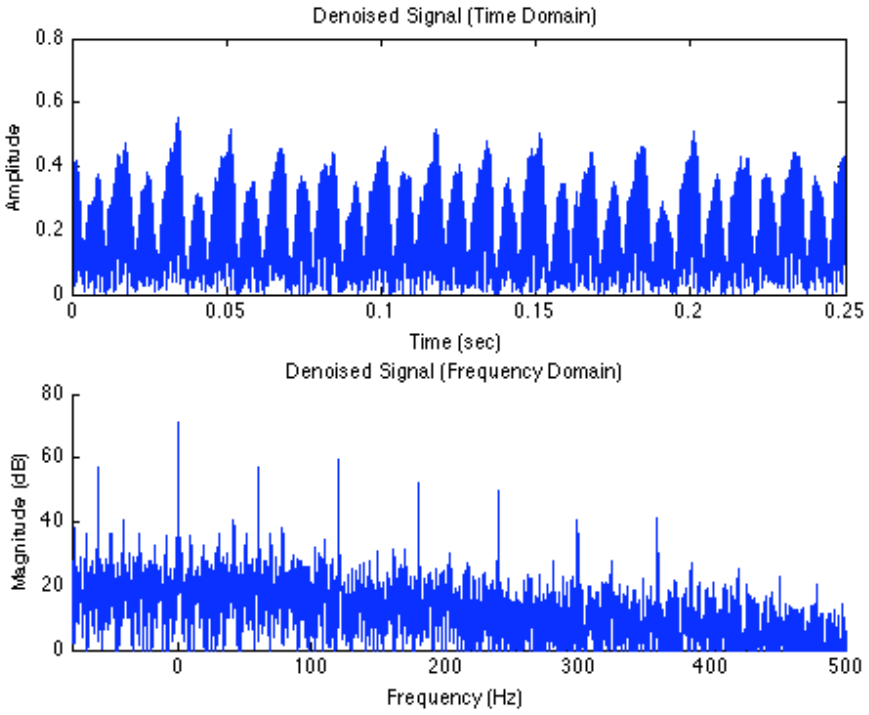

Figure 76: Results of using wavelet analysis with SVD.

At attempt was made to compare the results of these algorithms with the results from a standard Fourier Transform. A FFT was taken using 120000 points, or equivalently 10 seconds, of the original fault signal after a Hamming window was applied to reduce ripple effects. The results of the Fourier Transform as well as a sample of the original signal can be seen in Figure 77.

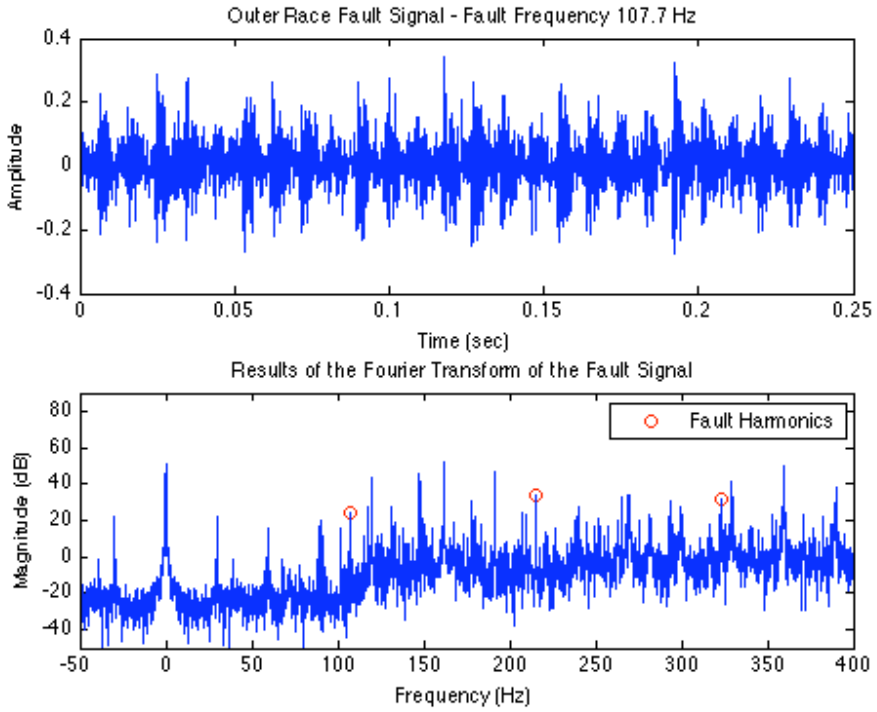

Figure 77: The results of the Fourier Transform of the original fault signal.

As can be seen in Figure 77, while the107.7 Hz cyclic frequency is visible in the time domain, in the frequency domain harmonics from the fault frequency is not clearly identifiable. This result matches with theoretical predictions. Due to the fact that the signal is not exactly periodic, harmonics associated with cyclic fault frequency are severely repressed rendering the standard Fourier Transform a poor tool to diagnose rolling element bearing faults. All other experimental faults seen in this thesis had much higher noise values and after the use of the standard Fourier Transform it was impossible to diagnose the fault.

The second case examined was an outer race fault hidden by severe noise on a SFK 6302-2RS bearing. The fault was 0.011 inches deep with a diameter of 0.007 inches. What is interesting about this set of data is the proximity of the harmonics of the shaft speed to the bearing fault frequency. The bearing fault frequency was calculated to be 91.43 Hz. This frequency is almost exactly 3 times the running speed of the shaft, 29.95 Hz. This makes detection difficult as traditional Fourier Transform methods will certainly show harmonics that are multiples of the shaft speed and any cyclical fault frequencies can easily be lost next to a dominate harmonic. As a result, for this case, not all algorithms were expected to work. The fault signal can be seen in Figure 78.

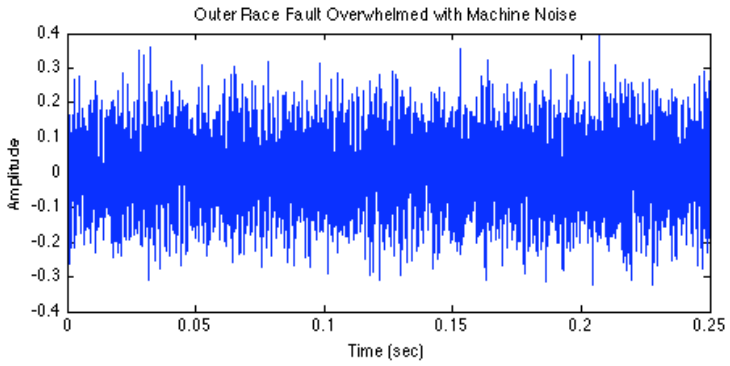

Figure 78: The time domain signal seen in this figure has a hidden outer race fault. The cyclic frequency of the fault is approximately 91.43 Hz.

Table 5 gives the results of de-noising the fault seen in Figure 78.

Table 5: Results from de-noising the outer race fault given in Figure 78.

Case 2	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS de-noising	3.2451	1	-6.1006	180.96
DRFS Threshold	4.7792	1	-12.9675	939.69
Envelope Demodulation	4.7193	1	-8.6621	NA
Hard Threshold	2.8357	0	-	234.49
Hybrid Threshold	78.0554	0	-	258.73
Mechanical Threshold	39.2061	0	-	222.38
AR/MED	3.9412	0	-	53.63
Soft Threshold	2.7574	0	-	242.52
Weak Signature Detection	2.1214	0	-	316.99

Surprisingly, only three of the nine algorithms were able to detect the bearing fault. Both envelope demodulation and the automated CWT/DRFS de-noising algorithm did the best in detecting the fault. The highlight of this comparison was the CWT/DRFS cyclic energy analysis. This analysis method was able to home in on the correct fault frequencies to make a detection, as seen in Figure 79 and 80. Figure 80 was used to set both the bandwidth at 220 Hz and the center frequency at 1340 Hz for envelope detection. The fault frequency as also identified to be closer 91.76 Hz rather than the calculated 91.4 Hz. As can be seen the limited bandwidth of the fault only allowed detection of the first harmonic.

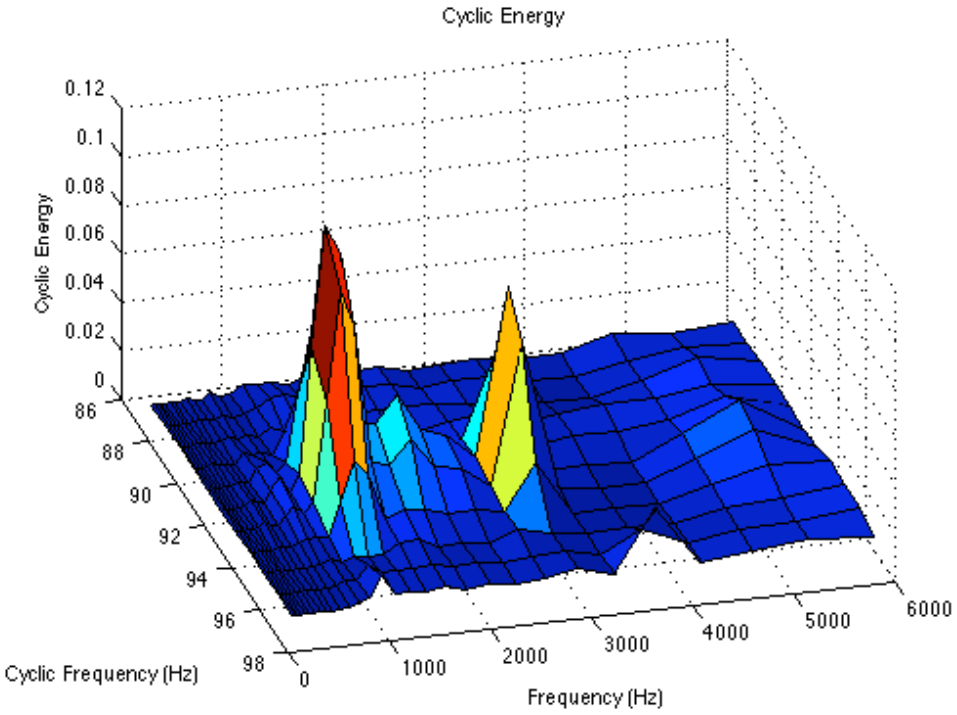

Figure 79: Results of applying Cyclic Energy Analysis to the outer race fault seen in Figure 78.

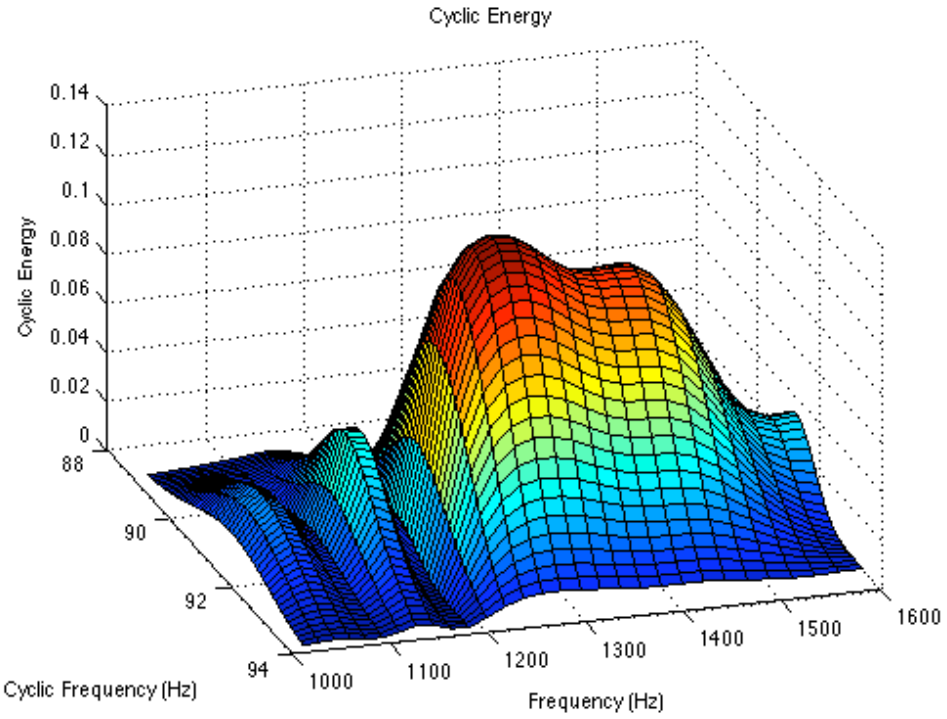

Figure 80: Results from applying Cyclic Energy Analysis to the outer race fault seen in Figure 78 over a selected region of frequencies and cyclical frequencies.

The actual fault signal itself is very weak, as seen in the results from the CWT/DRFS algorithm and envelope demodulation given in Figures 81 and 82.

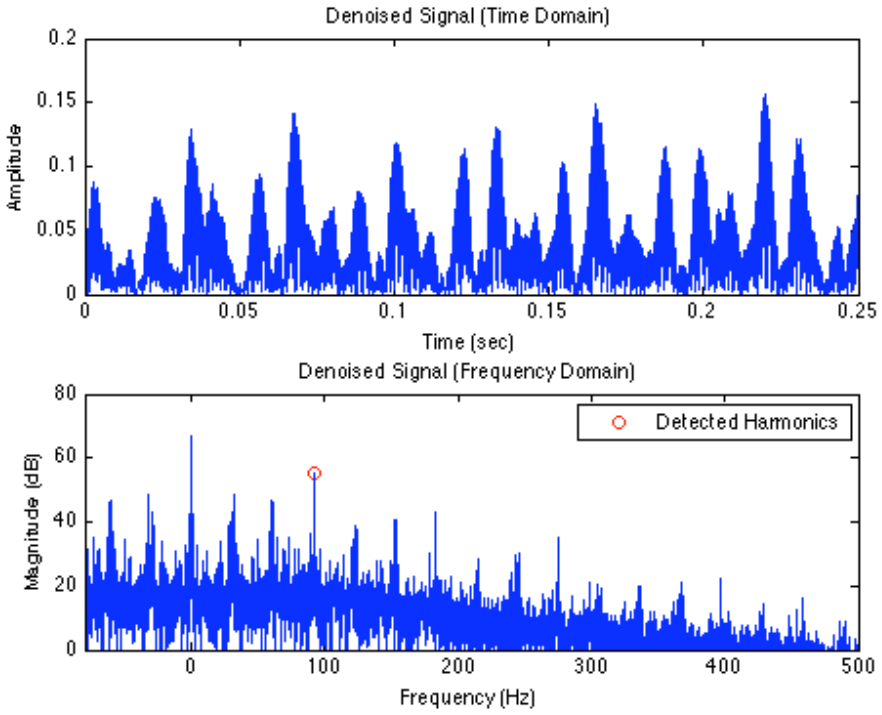

Figure 81: De-noised vibration signal from case two using Envelope Demodulation. The fault frequency of 91.7 can be seen.

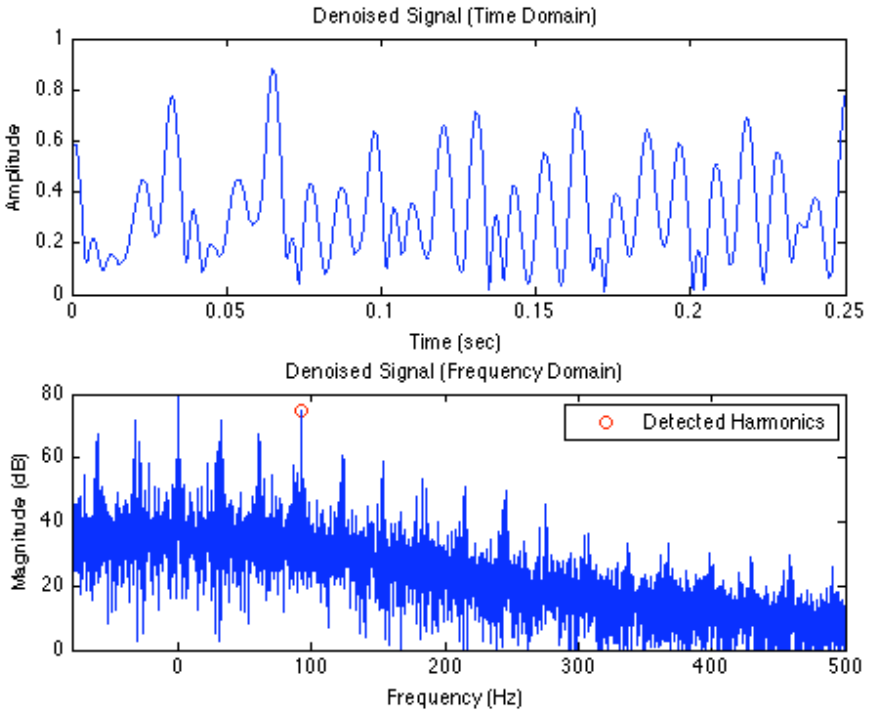

Figure 82: Results from applying CWT/DRFS de-noising. The fault frequency of 91.7 Hz can be seen.

What is remarkable is that although the fault frequency is very close to the third harmonic of the running shaft speed it is clearly defined upon close inspection. This can be seen in Figure 83 when closely examining the frequency domain results of envelope demodulation and the automated CWT/DRFS denoising algorithm.

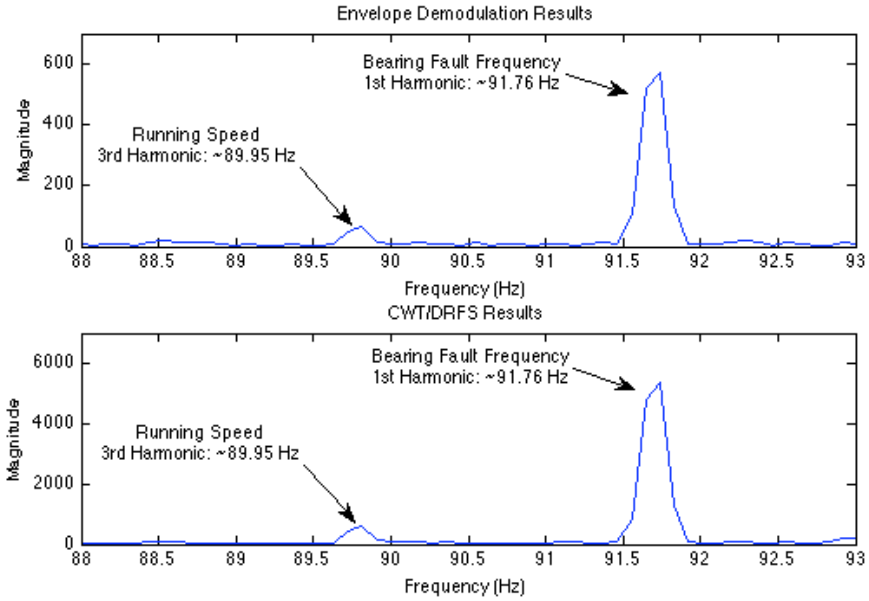

Figure 83: Upon close inspection in the frequency domain the bearing fault frequency can clearly be distinguished from the 3rd running speed harmonic using Envelope Demoduation **(top)** and automatic CWT/DRFS de-noising **(bottom)**.

The third algorithm that was able to successfully distinguish the bearing fault, DRFS thresholding, was also able to distinguish the bearing fault frequency. However the algorithm had a lower SNR as well as took almost 15 minutes to run compared to the CWT/DRFS cyclic analysis of 3 minutes.

5.5 Inner Race Faults

Inner race faults are generally more difficult to detect than outer race faults. Being located on the inner race the impulse of the rolling element striking the fault has a longer transmission path to the vibration sensor. Furthermore because the inner race is

rotating any radial load will have the effect of modulating the amplitude of the impulse, reducing the amount of periodicity in the fault signature. As the rolling elements travel around the bearing there is inherent slip in the rolling elements. The rolling slip has been shown to be more pronounced inner race faults. This further reduces the periodicity that can be expected when detecting this type of fault. Consequently inner race faults are generally characterized to be more difficult to detect.

To see the effect of the algorithms in de-noising inner race faults, the nine algorithms introduced in this thesis were applied on two experimentally gathered inner race fault signals. The first inner race fault signature was from a SKF 6302-2RS bearing with a fault depth of 0.011 inches and a diameter of 0.007 inches. The calculated inner race fault frequency was approximately 162 Hz. The fault signal can be seen in Figure 84.

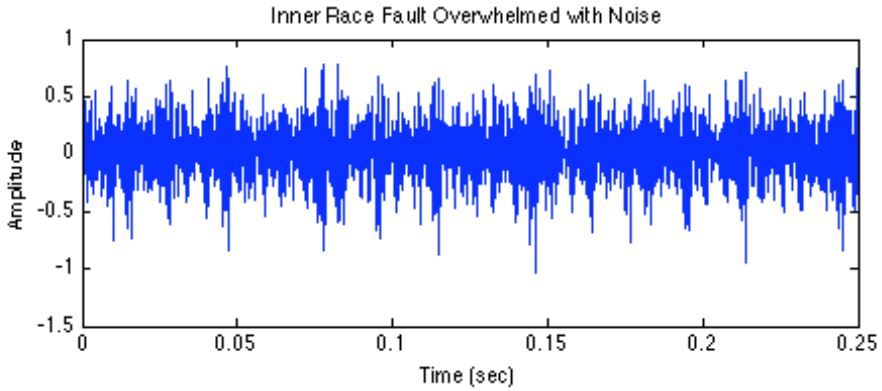

Figure 84: Inner race fault with a fault frequency of 162 Hz.

The results of the de-noising algorithms can be seen in Table 6.

Table 6: Results from de-noising the inner race fault given in Figure 84.

Case 1	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS de-noising	4.0059	0	-	140.78
DRFS Threshold	4.37	2	-9.6821	452.24
Envelope Demodulation	3.6364	3	-9.4498	NA
Hard Threshold	4.8628	0	-	148.9
Hybrid Threshold	27.6856	0	-	209.61
Mechanical Threshold	29.3701	0	-	194.1
AR/MED	93.5174	0	-	97.92
Soft Threshold	5.6157	0	-	139.2
Weak Signature Detection	1.9851	0	-	232.32

Only two of the nine algorithms attempted were able to clearly identify the fault. Again the clear winner was envelope demodulation with three harmonics above the noise floor and the highest SNR. Both the bandwidth and the center frequency again were estimated off the CWT/DRFS cyclic energy analysis. The results of this analysis can be seen in Figures 85 and 86.

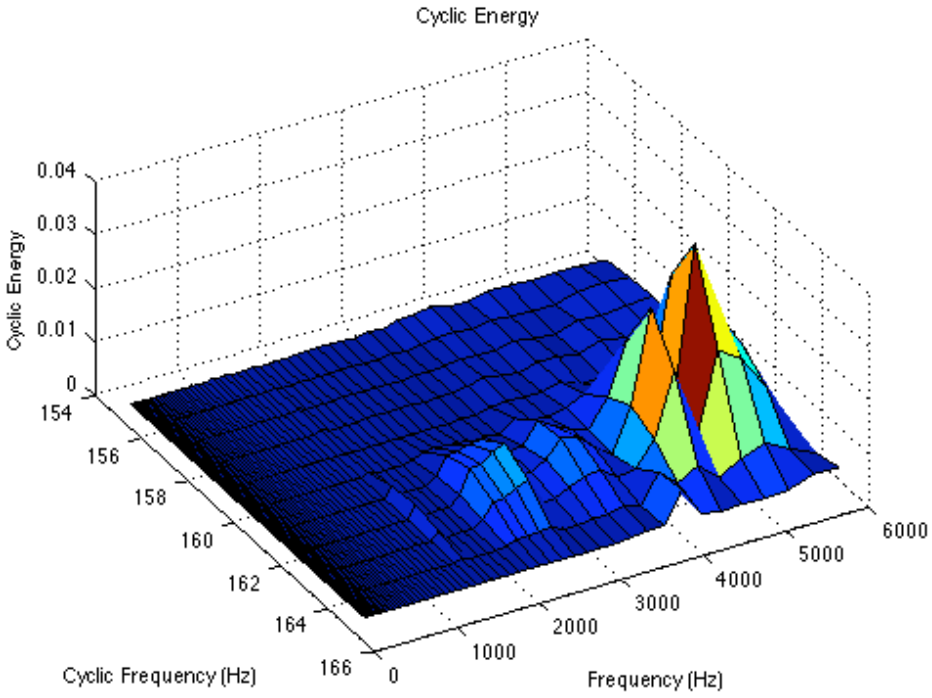

Figure 85: Results from Cyclic Energy Analysis of the outer race fault seen in Figure 84.

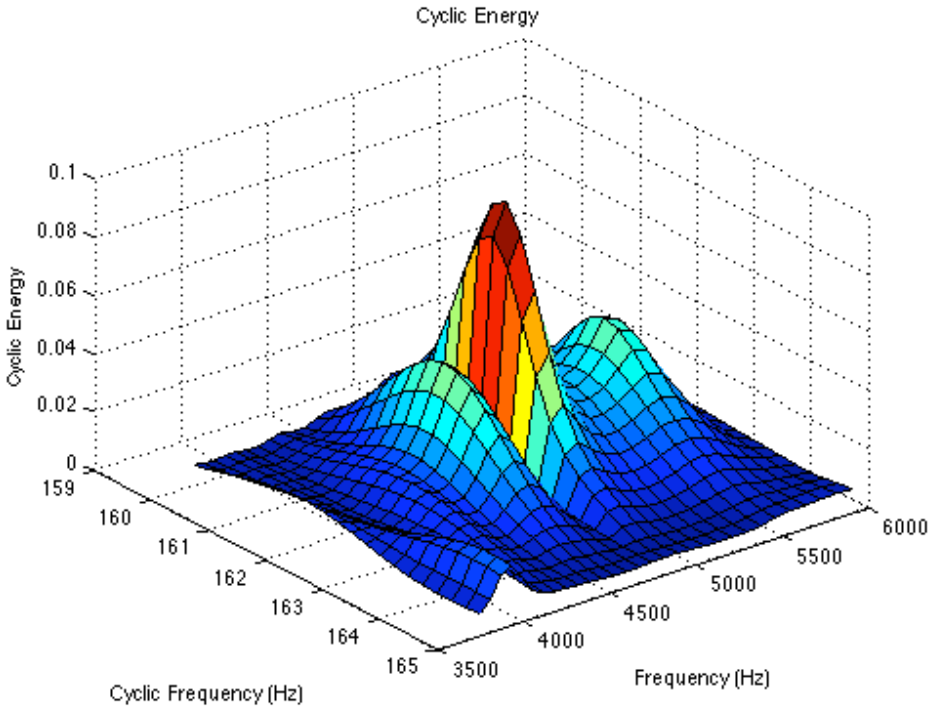

Figure 86: Results from Cyclic Energy Analysis of the outer race fault seen in Figure 84 over a selected region of frequencies and cyclical frequencies.

Using figures 85 and 86, the natural frequency of the fault, and therefore the center frequency of the filter, was located at 4969.8 Hz and the bandwidth of the filter was set at twice the bearing fault frequency or ~330 Hz. This allowed the capture of at least one harmonic in the frequency domain. As can be seen, setting this small of a window, only three harmonics were actually captured. The results from envelope demodulation can be seen in Figure 87.

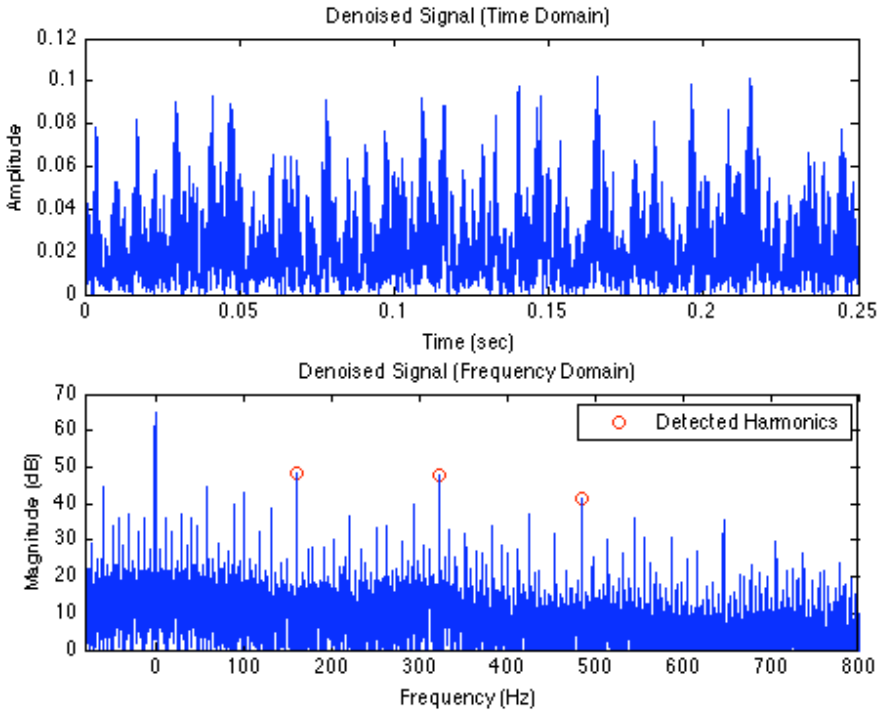

Figure 87: De-noised vibration signal from case three using Envelope Demodulation. Three harmonics can clearly be identified 3 dB above the surrounding noise.

The other algorithm that de-noised the signal effectively was the DRFS thresholding technique. The results from this algorithm are shown in Figure 88. Unfortunately, while the results were automatically computed with no user input, the algorithm took approximately 8 minutes to compute. The results from DRFS thresholding are also more visually appealing, resulting in a higher kurtosis value.

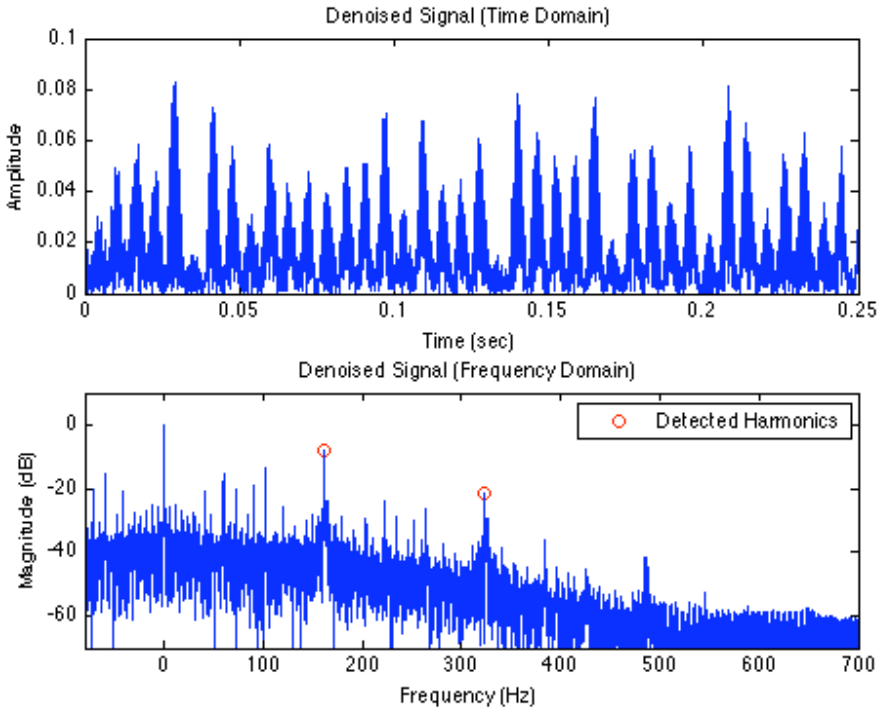

Figure 88: De-noising results from application of DRFS thresholding. As can be seen there are two actual clearly visible fault frequency harmonics at 161 Hz and 322 Hz.

The results from the automated CWT/DRFS de-noising algorithm were ineffective at finding the fault frequency. This is due to the limited number of scales, 30 used in computation. Are shown in Figure 86, the bandwidth of the fault was very small. It is interesting to note that the results of automated CWT/DRFS de-noising algorithm did include a very distinct harmonic at the bearing fault frequency, however low frequency spurious noise drowned out this detection. This can be seen in Figure 89. It is probable that an increase in the number of scales of the wavelet transform could rectify this failure.

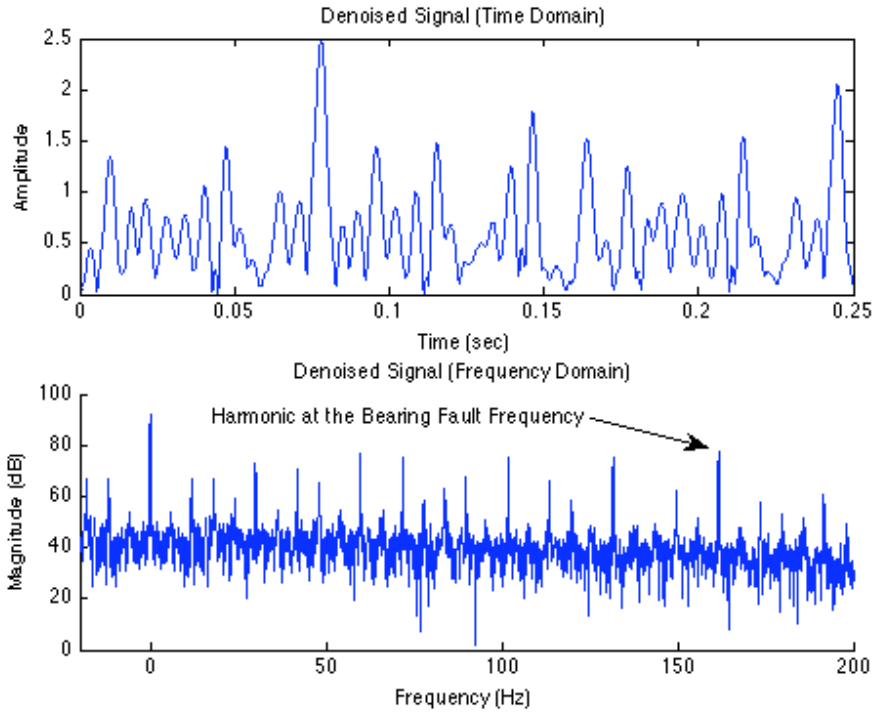

Figure 89: The automated results from the CWT/DRFS algorithm can be seen in this figure. This algorithm produced a noticeable harmonic at the bearing fault frequency, however had a number of low frequency noise components that rendered this algorithm a failure.

All the other algorithms failed in detecting the fault. Two of the algorithms, mechanical and hybrid thresholding found impulses from the fault signature, given in Figures 90 and 91. Unfortunately, due to the modulation effect seen with inner race faults the majority of the impulses were below the detection threshold of these algorithms resulting in limited detections.

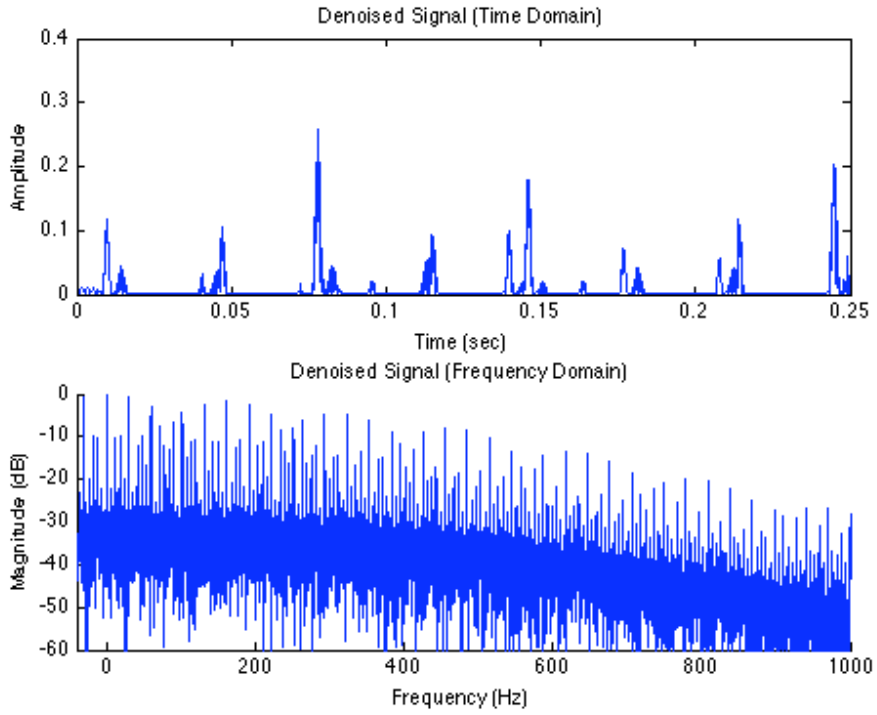

Figure 90: Results from application of Hybrid thresholding. Only impulses that were significantly large survived. This removed a number of impulses and resulted in not being able to identify the fault in the frequency domain.

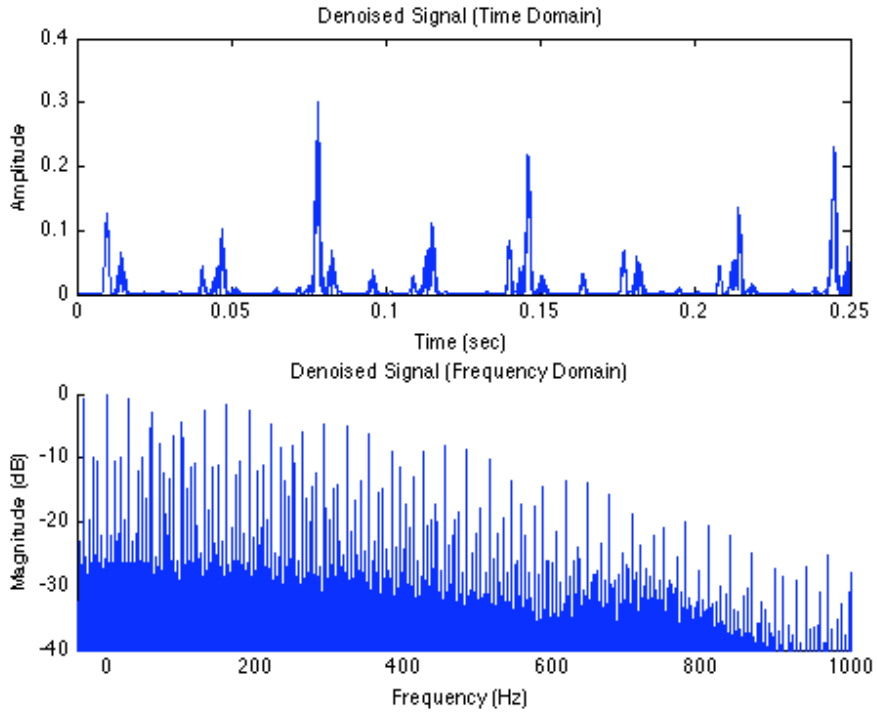

Figure 91: Results from application of Mechanical thresholding. Only impulses that were significantly large survived. This removed a number of impulses and resulted in not being able to identify the fault in the frequency domain.

Finally, two algorithms showed very visible signs of failing. Weak Signature Detection erroneously identified the running harmonics of the fault as the fault frequency. Minimum Entropy Deconvolution maximized the kurtosis of the signal based on a spurious impulse that was generated after applying the AR filter. This resulted maximizing the kurtosis of the signal, however also in completely obliterating any sign of the fault as can be seen in Figure 92.

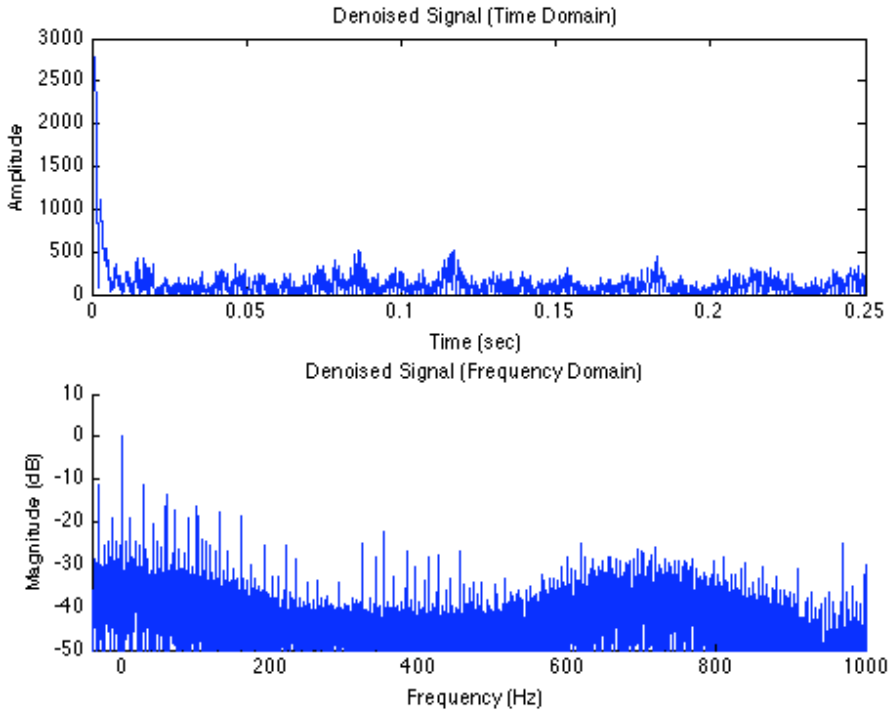

Figure 92: Results from application of the MED algorithm. The MED algorithm maximized kurtosis, unfortunately this resulted in the maximization of a spurious peak at the beginning of the signal.

The second inner race fault examined was a from a SKR 6205-2RS bearing. The fault was 0.011 inches in depth and 0.007 inches in diameter and corresponded to a fault frequency of 148 Hz. As with the second outer race fault examined, the inner race fault examined was almost exactly a multiple of the shaft running speed of 29.93 Hz making this fault more difficult to detect. The fault signal can be seen in Figure 93.

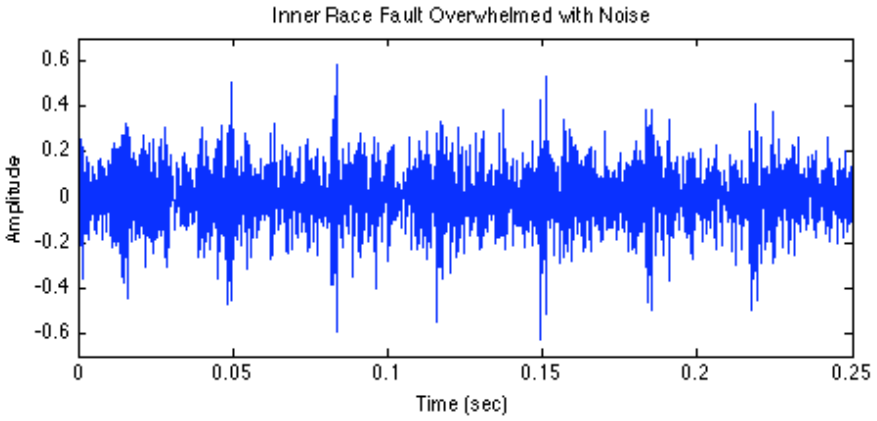

Figure 93: Inner race fault overwhelmed with noise. The impulses seen in the figure correspond to first harmonic of the running speed of the shaft or ~30 Hz.

The results of the de-noising algorithms can be seen in Table 7.

Table 7: Results from de-noising the inner race fault given in Figure 93.

Case 1	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS denoising	7.3543	4	-4.3106	82.23
DRFS Threshold	6.8787	2	-9.769	362.12
Envelope Demodulation	6.9258	3	-6.9729	NA
Hard Threshold	6.3387	0	-	103.84
Hybrid Threshold	68.7852	0	-	126.46
Mechanical Threshold	90.3228	0	-	120.06
AR/MED	4.3112	1	-	24.89
Soft Threshold	8.1944	0	-	103.61
Weak Signature Detection	7.3459	0	-	147.94

Only four of the nine algorithms were able to successfully detect this fault. The clear winner was the automated CWT/DRFS de-noising algorithm. These results can be

seen in Figure 94. What is remarkable with these results is that the modulation of the impulse amplitude is very clearly evident in the time domain.

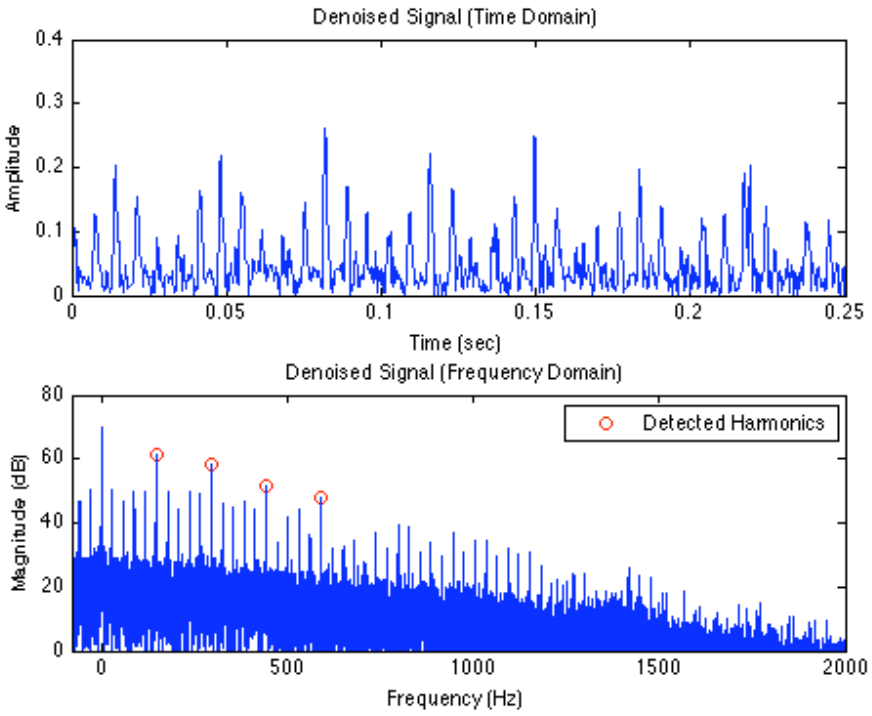

Figure 94: Results from applying the automated CWT/DRFS algorithm. Four harmonics were detected in the frequency domain and the impulse modulation associated with inner race faults is clearly visible.

Envelope demodulation was not as successful. Results from CWT/DRFS cyclic energy analysis can be seen in Figures 95 and 96. As can be seen, the fault is clearly evident over a wide frequency range, from 5 kHz to 6 kHz. While this made it easy to determine the center frequency of the filter used in envelope demodulation, several

different bandwidth values were attempted to maximize the number of detected harmonics. While the signal was detected with at least two harmonics for bandwidth values from 200 Hz to 900 Hz, the number of detected harmonics was maximized with a small bandwidth of 300 Hz.

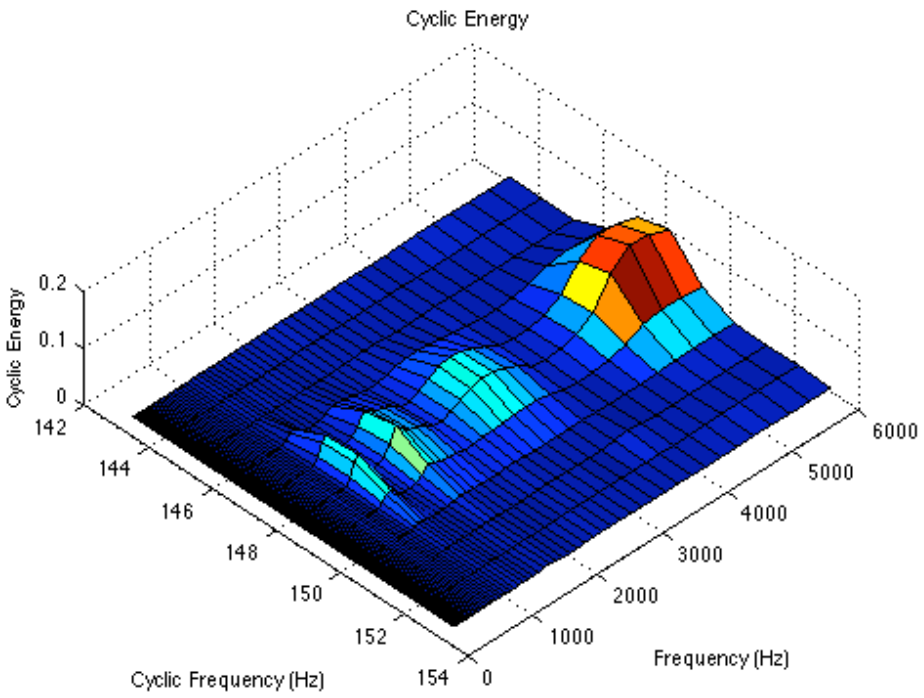

Figure 95: Results from applying Cyclic Energy Analysis to the outer race fault seen in Figure 92.

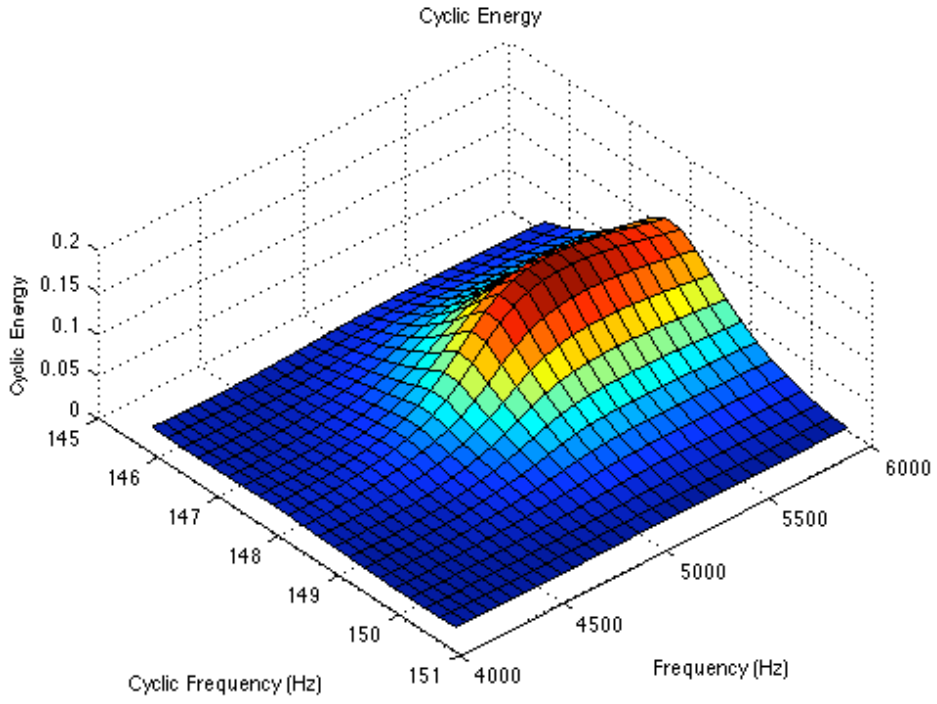

Figure 96: Cyclic Energy analysis of the outer race fault seen in Figure 92 over a selected region of frequencies and cyclical frequencies.

What is interesting to note, all the other de-noising algorithms that failed when attempting to de-noise the bearing fault frequency, successfully captured every 5th impulse. This can be seen in Figure 97 where the results of the results from Wavelet Analysis using SVD are plotted alongside the results from the automated CWT/DRFS de-noising algorithm and the original noisy fault signal.

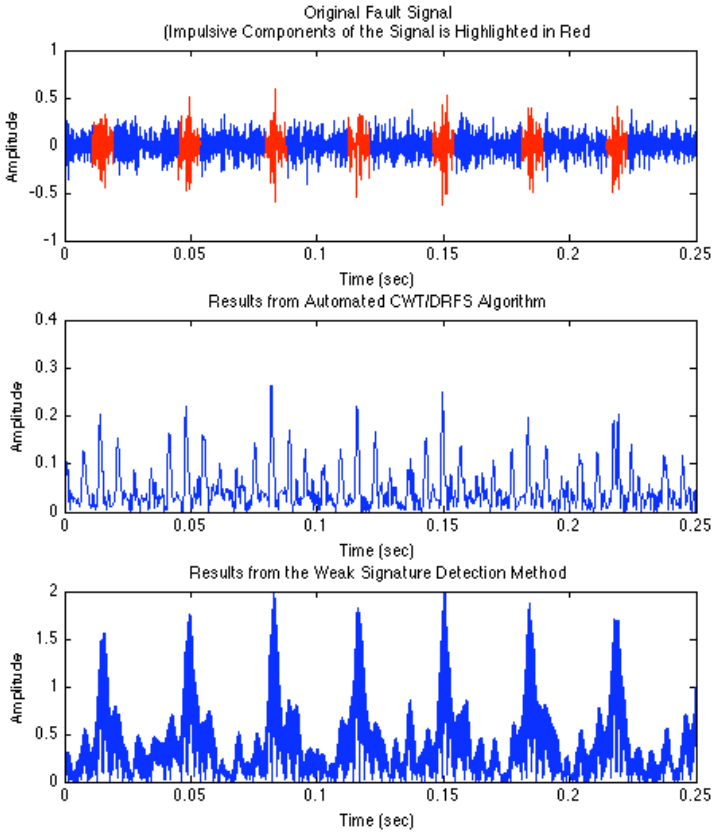

Figure 97: Results from the Automated CWT/DRFS algorithm **(middle)**, the Weak Signature Detection Method **(bottom)** and the original fault signal **(top)** are shown for comparison. The Weak Signature Detection Method was able to extract every fifth impulse from the fault however was unable to clearly define the weaker impulses that were modulated by the rotation of the load on the shaft. Remarkably, every fifth impulse can be seen in the original fault signal, highlighted in red.

The fifth harmonic of the running speed for this case is approximately 149.75 Hz. This is almost exactly the same speed as the first harmonic due to the bearing fault, 147.58 Hz. Consequently, fifth impulse resonated with the running speed harmonics overemphasizing the impulse and making every fifth impulse visible in original signal.

This makes diagnosing the fault difficult until envelope demodulation or the results from the automated CWT/DRFS de-noising algorithm are examined.

5.6 Rolling Element Faults

Rolling element faults are the most difficult faults to detect. As the rolling element travels round the bearing it is not guaranteed that the rolling element will take the same path every revolution. Rolling element faults are best characterized by detection of single impulses and periodicity, as measured with the FT, and used for detection with inner and outer race faults, does not detect the fault. Kurtosis has been a suggested means of automated detection of rolling element faults and will be used for this study.

Two inner race faults were studied. The first was from a SKF 6205-2RS bearing with a fault of 0.011 inches deep and 0.014 inches in diameter. The calculated fault frequency was 141.09 Hz. Two vibration signals were used for this analysis. One vibration signal was from a vibration sensor place very close to the faulty bearing and hence recorded a very clean vibration signal with little noise. The second vibration sensor was placed on a healthy bearing, and, as a result, the vibration from the fault was buried in noise. The signals seen from both vibration signals can be seen in Figure 98.

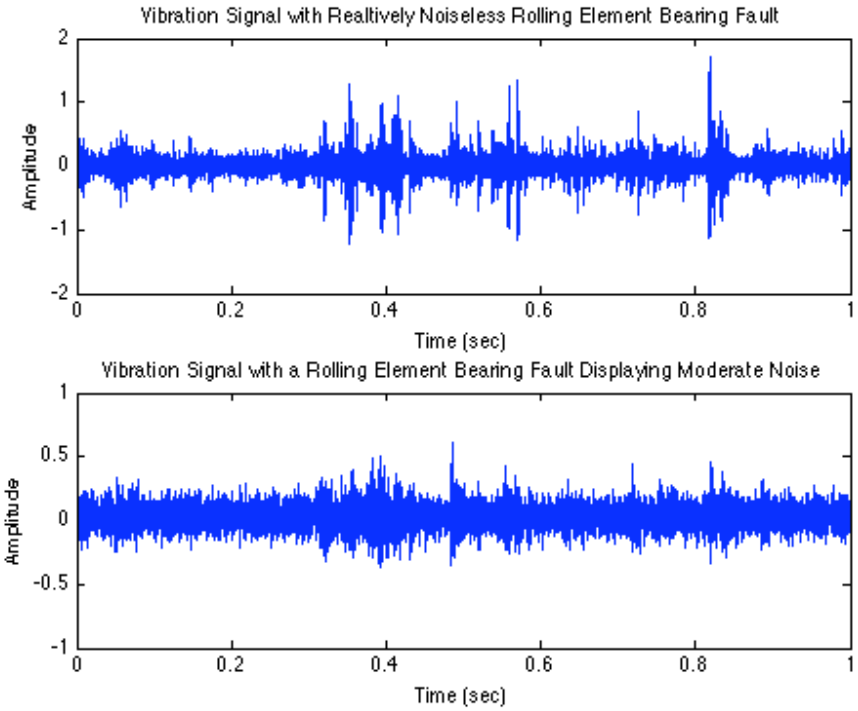

Figure 98: Two vibration signals showing the same rolling element fault. The first signal **(top)** is taken from a vibration sensor placed close to the faulty bearing. The second signal **(bottom)** is taken from a vibration sensor far away from the bearing fault.

The nine algorithms were applied to the noisy vibration signal in an attempt to de-noise it and the results can be seen in Table 8.

Table 8: Results from de-noising the rolling element fault given in Figure 98.

Case 1	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS de-noising	29.8329	0	-	78.51
DRFS Threshold	31.531	0	-	398.76
Envelope Demodulation	49.6762	0	-	NA
Hard Threshold	7.4509	0	-	107.61
Hybrid Threshold	56.3062	0	-	135.06
Mechanical Threshold	61.4654	0	-	119.93
AR/MED	58.2679	0	-	72.69
Soft Threshold	11.7294	0	-	110.86
Weak Signature Detection	9.2773	0	-	153.81

As can be seen in Table 8 no harmonics existed at the bearing fault frequency after the de-noising algorithms were applied. This is not surprising as the impulses from bearing faults have been termed chaotic by a number of researchers. The only measure of de-noising success is therefore kurtosis. In this case, Hybrid Thresholding, Mechanical Thresholding and Minimum Entropy Deconvolution all provided good indications of the presence of a fault. The results of these three algorithms can be seen with comparison to the noiseless vibration signal of the bearing fault in Figure 99. Mechanical Thresholding performs the best with the highest amount of impulsiveness with the lowest noise floor.

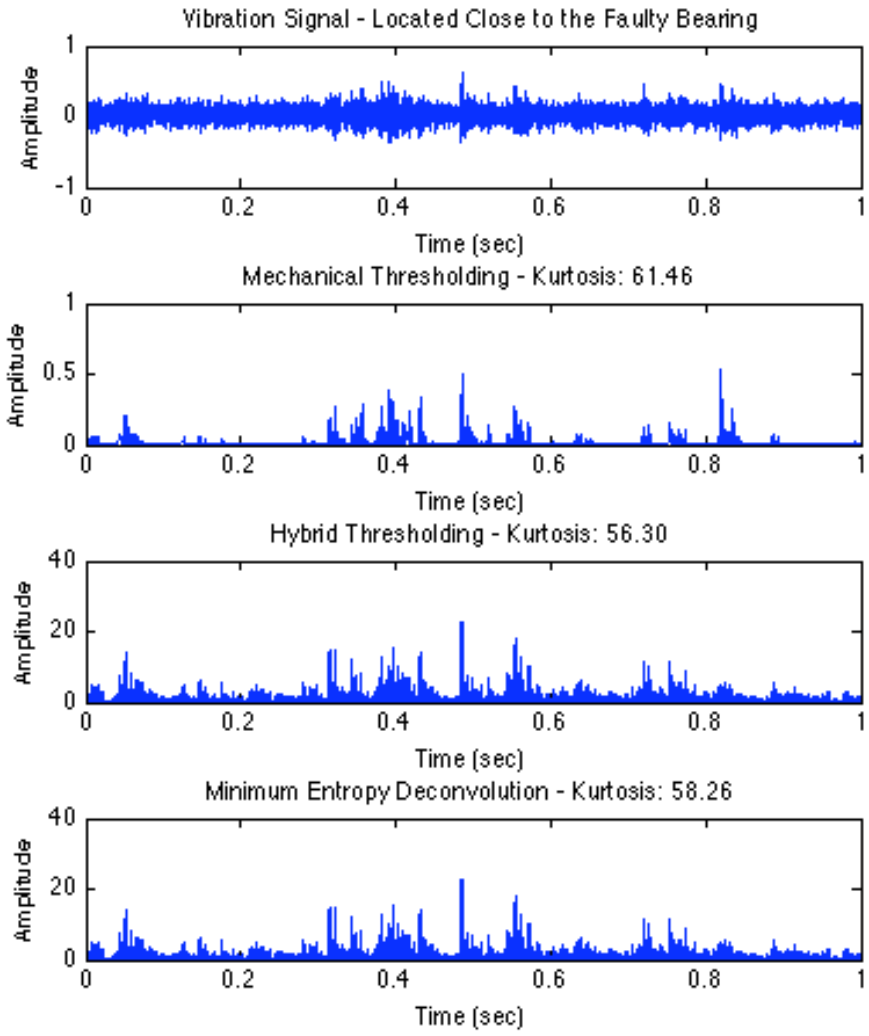

Figure 99: De-noised signals using Mechanical thresholding, Hybrid thresholding and Minimum Entropy Deconvolution.

Envelope demodulation was also relatively successful, as can be seen in Figure 100.

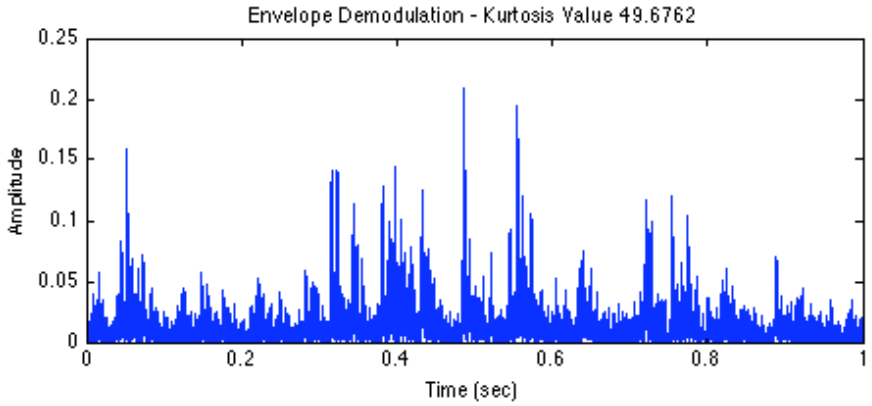

Figure 100: De-noised vibration signal from case five using Envelope Demodulation.

The results from the CWT/DRFS cyclical analysis, seen in Figures 101 and 102, were used to determine the filter bandwidth of 4923 Hz and the filter center frequency of 300 Hz.

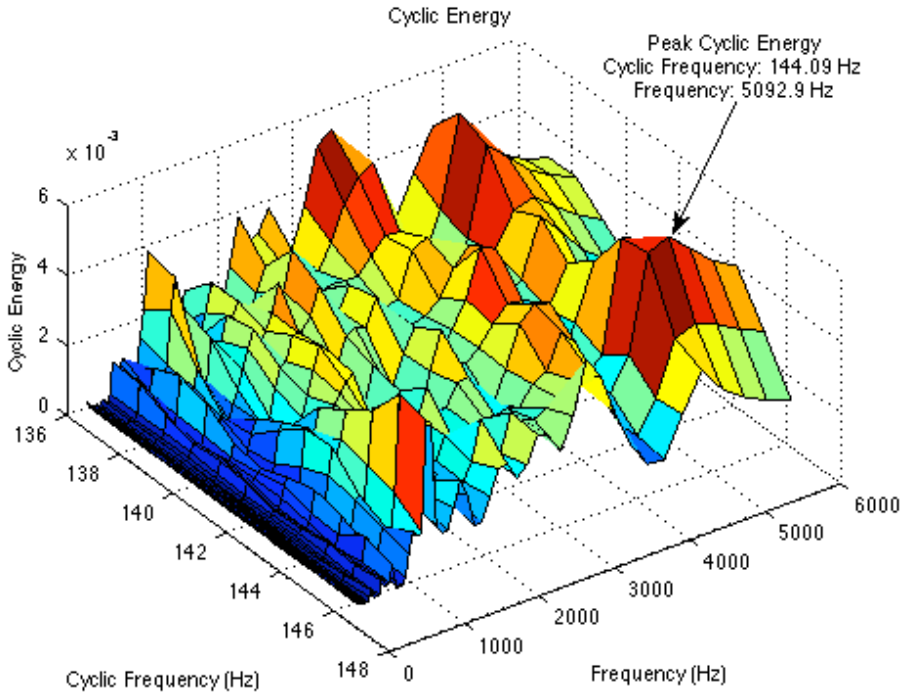

Figure 101: Results from Cyclic Energy analysis of the rolling element fault seen in Figure 98.

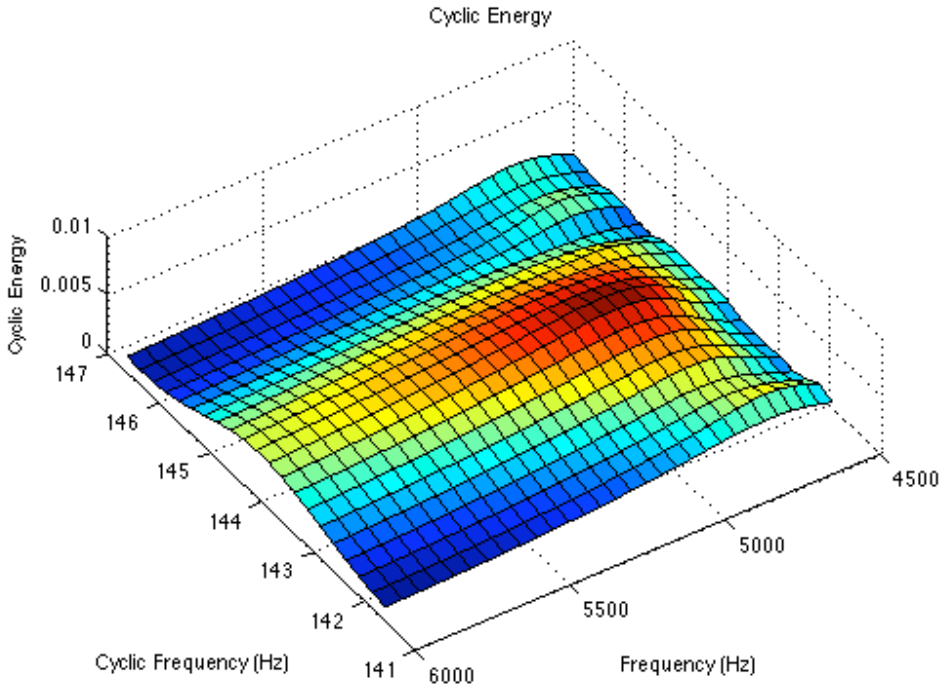

Figure 102: Results from Cyclic Energy Analysis of the rolling element fault seen in Figure 98 over a selected frequency range and cyclic frequency range.

As can be seen from Figures 101 and 102, CWT/DRFS cyclic energy analysis proved to be a very poor indicator of the presence of a fault. The fact that envelope demodulation worked for this case is more likely due to the fact that the filter was placed at a high frequency and low frequency noise, such as from the running speed of the shaft were removed from the signal. What is interesting is all de-noising methods that relied specifically on the fault being periodic or aperiodic failed providing very low kurtosis values.

The second case examined was an inner race fault from a SKF 6203-2RS bearing with fault 0.011 inches deep and 0.007 inches in diameter. The fault frequency was calculated to be 119.48 Hz. Again, as with the previous case two vibration signals were examined. The first signal was from a vibration sensor placed close to the faulty bearing. The second vibration sensor was placed further away and the vibration sensor recorded had more noise associated with the fault. Both vibration signals can be seen in Figure 103.

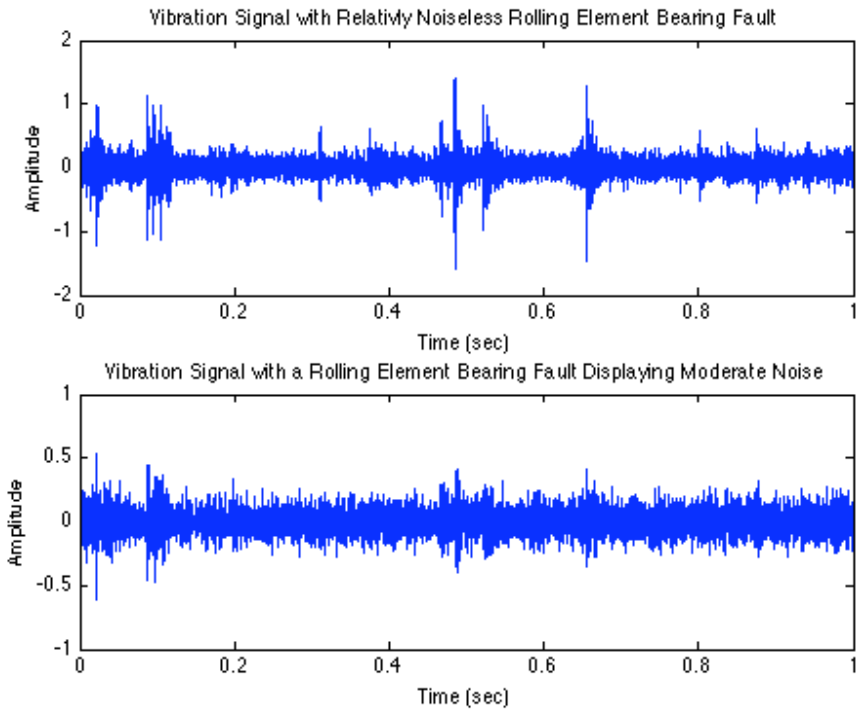

Figure 103: Two vibration signals showing the same rolling element fault. The first signal **(top)** is taken from a vibration sensor placed close to the faulty bearing. The second signal **(bottom)** is taken from a vibration sensor far away from the bearing fault.

The nine algorithms introduced in this paper were used to de-noised the faulty bearing signal and the results are given in Table 9.

Table 9: Results from de-noising the rolling element fault given in Figure 103.

Case 1	Kurtosis	Harmonics	SNR	Time (sec)
CWT/DRFS de-noising	4.8618	2	-5.1203	102.86
DRFS Threshold	10.8138	1	-22.7993	538.29
Envelope Demodulation	3.4301	1	-8.5084	NA
Hard Threshold	7.0382	0	-	155.09
Hybrid Threshold	590.02	0	-	153.02
Mechanical Threshold	321.03	0	-	138.26
AR/MED	3.8674	0	-	68.1837
Soft Threshold	9.7834	0	-	128.12
Weak Signature Detection	2.3164	0	-	236.20

Both Hybrid and Mechanical Thresholding were the clear winners in this case. The results from these algorithms compared to the low noise vibration signal can be seen in Figure 104.

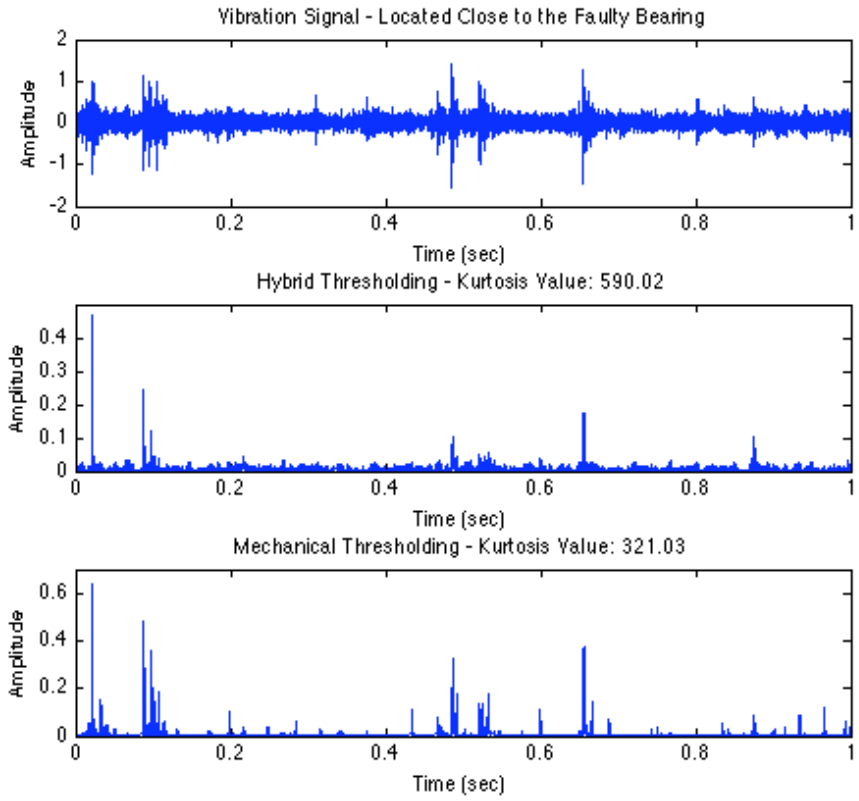

Figure 104: Results from Hybrid Thresholding **(middle)** and Mechanical Thresholding **(bottom)** compared to the noiseless vibration signal **(top)**.

As can be seen in Figure 104 the two signals with the highest Kurtosis values, Hybrid and Mechanical Thresholding de-noise the signal effectively. What is interesting is that Mechanical Thresholding, while having lower kurtosis, provides a much visually cleaner estimation of the fault locations in time. Kurtosis is a very sensitive measurement to signal outliers, and there is an outlier present at the very beginning of the time series produced by Hybrid Thresholding thus presenting Hybrid Thresholding as performing better than it should.

What is also interesting with this data set is the indication of a cyclical presence at the bearing fault frequency. This is due to other mechanical vibrations and is not surprising as fourth harmonic of the shaft frequency is 119.87 Hz, compared to the estimated fault frequency of 119.48 Hz. This can clearly be seen in the CWT/DRFS cyclical energy analysis provided in Figures 105 and 106.

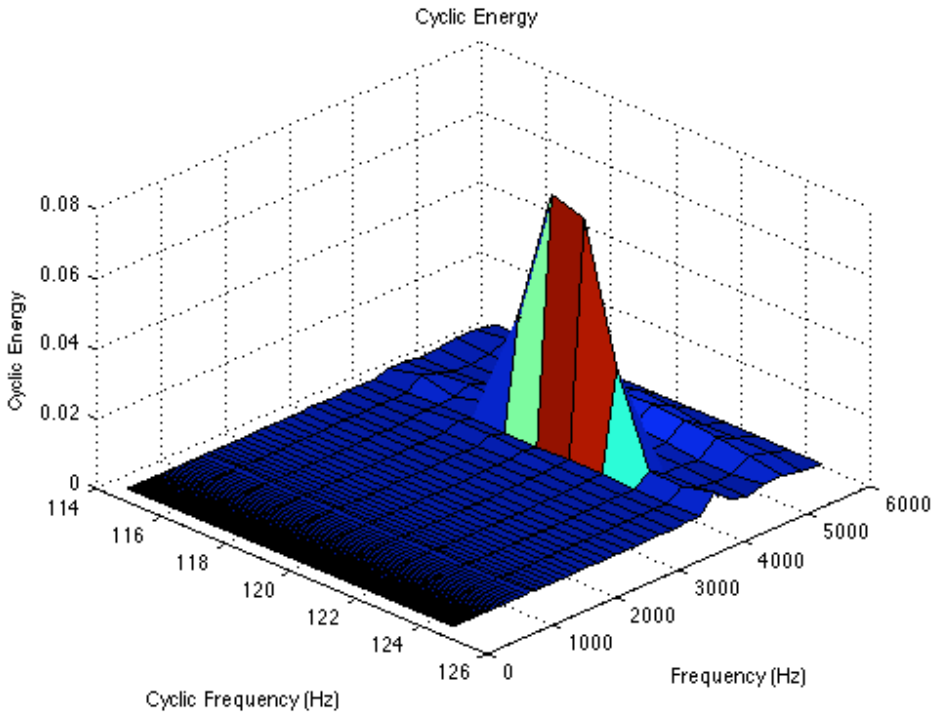

Figure 105: Results from applying Cyclic Energy Analysis to the rolling element fault seen in Figure 103.

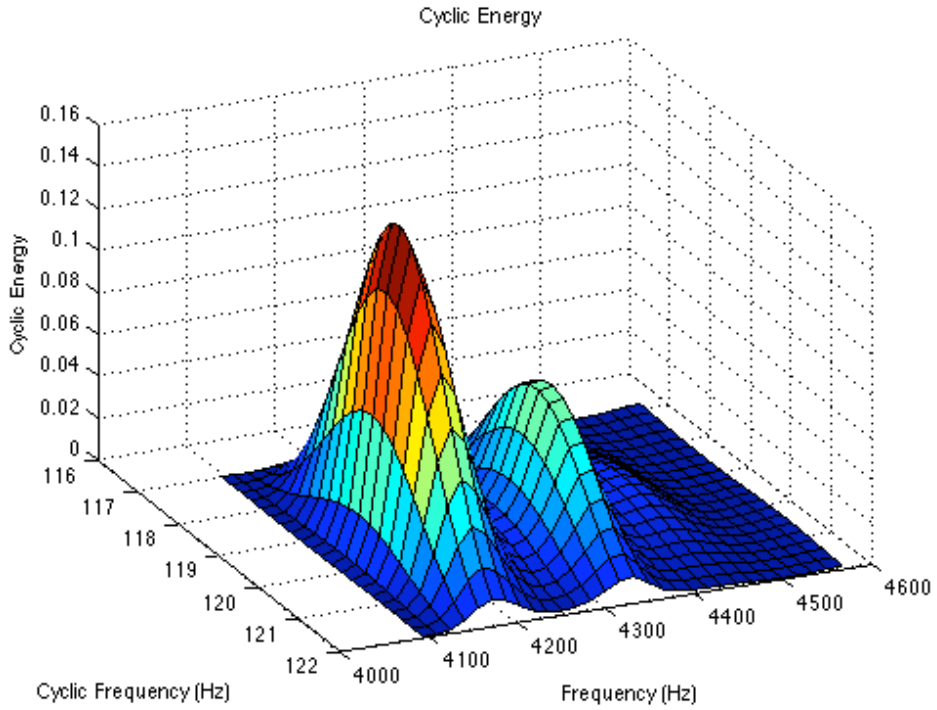

Figure 106: Results from applying Cyclic Energy Analysis to the rolling element fault seen in Figure 103 over a selected frequency and cyclic frequency range.

The CWT/DRFS cyclic energy distribution was developed to detect the presence of periodic amplitude-modulated signals. In this situation the CWT/DRFS analysis tool is picking up on other vibration signals that are present in the machinery. These types of signals are relatively common in mechanical systems, and it is not surprising that there are other components with hidden periodicities, especially at one of the harmonics of the running speed of the machine. Another interesting phenomenon, the Minimum Entropy Deconvolution algorithm, rather than amplifying the impulsive components of the signal also captured the hidden periodic signal that CWT/DRFS detected. This is unusual in

that there is no restraint on the Minimum Entropy Deconvolution algorithm to find periodic signals.

5.7 Results

The overwhelming results of this study favored envelope demodulation, the industry standard for bearing fault detection is better or equivalent to than the other eight algorithms for detecting inner and outer race faults, if the natural frequency of the bearing is known. For cases where the natural frequency of the bearing is not known, for inner and outer race faults, the two algorithms, particularly the automated CWT/DRFS algorithm produced detections with very little user input other than the bearing fault cyclic frequency.

For the rolling element faults, Hybrid and Mechanical Thresholding best extracted the fault signatures, maximizing kurtosis. Rolling element faults, while also impulses modulated at the natural frequency of the bearing are not periodic but chaotic. As a result it is possible to capture portions of a signal that, while the rolling element might be evident in the bearing, it is not producing vibration signatures. As a result it can be difficult to detect the natural frequency of the bearing based on the fault signature. These two algorithms also were effective on inner and outer race faults when the SNR ratio of the fault was significantly above the noise. This can be seen when examining data from the inner race faults where only the strongest of the periodic impulses were detected.

Minimum Entropy Deconvolution was also a successful algorithm in some cases; however, this algorithm had several faults. In one case the automated algorithm

produced a filter that completely destroyed the signal, maximizing kurtosis by amplifying a single data outlier at the beginning of the signal. Furthermore, it can be difficult to determine effective lengths for the AR filter and MED filter, and it is possible for the MED filter not to converge to a valid answer. As a result, for automated detection this algorithm is not robust enough to be considered valid. Finally, with that said, on this data set it did not perform as well as other algorithms for either inner, outer or rolling element faults.

The last three algorithms Weak Signature Detection, Hard Thresholding and Soft Thresholding did not provide good de-noising ability on the signals examined. Hard and Soft Thresholding were included to provide a baseline comparison for Hybrid and Mechanical Thresholding and were not expected to perform well. These algorithms are only really effective when the signal to be de-noised already has high SNR. For the cases where the SNR of the fault signal is weak, it will considered noise and removed.

Weak Signature Detection only worked in specific situations. In signals corrupted with harmonic noise the results of this algorithm to detect periodicity almost always found one of the signal harmonics and missed the bearing fault cyclic frequency. The algorithms developed in this paper used the unique ability of analytic wavelets to demodulate the signal therefore minimizing the harmonic content of the noise and automatically base-banding the harmonic content of the fault frequency. This proved invaluable in providing accurate estimation of the periodicity of the wavelet scale.

The dominant outcome of the study was the ability of the CWT/DRFS to detect inner and outer race bearing faults. What is remarkable, even in cases where only one

harmonic of the fault signal was clearly visible in the de-noised signal, the fault was clearly obvious when viewing the three dimensional CWT/DRFS energy analysis plot. One of the central difficulties for engineers attempting to detect bearing faults is locating the natural frequency of the bearing. Manufactures of bearings will often publish suggested center frequencies to examine when scanning the frequency domain looking for bearing faults. Using the CWT/DRFS energy plot, for this set of inner and outer race fault data examined, it was clearly evident where the natural frequencies of the bearing existed in the frequency domain.

To show the advantage of this tool more studies need to be conducted. Other tools currently exist to perform the same analysis as the CWT/DRFS cyclic energy algorithm given in this thesis and CWT/DRFS cyclic energy distribution needs to be compared against these other algorithms to judge it's relative effectiveness. Furthermore, it is very difficult to determine the minimum SNR that CWT/DRFS can effectively identify periodic amplitude modulated signals. Research would need to be conducted to see the effectiveness of this algorithm in detecting a truly incipient fault by running a bearing through its full life and determining the earliest point of fault detection using this algorithm compared to the standard envelope analysis as well as other bearing fault detection algorithms that were not necessarily wavelet based.

6. SUMMARY AND CONCLUSION

This thesis developed two wavelet based de-noising algorithms and one wavelet based detection algorithm for use with inner and outer race rolling element faults. A new wavelet, the Natural Frequency Impulse Response Wavelet was also introduced. All algorithms were compared in implementation with other wavelet de-noising algorithms to judge their effectiveness. It was shown that the fault detection and de-noising algorithms proposed in this paper were more effective than current algorithms that have been introduced in literature for detecting inner and outer race faults.

The success of the wavelet algorithms introduced in this paper rely on the properties of inner and outer race bearing fault signals. These signals are periodic modulated impulses. The goal in de-noising or detecting these faults is then to filter out modulation frequency and then detect the impulsive nature of the fault. The continuous wavelet transform is unique in that it provides a filter bank, an optimum time frequency distribution, and automatically demodulates the signal at each scale. The filter bank allows the extraction of the bearing fault into a small number of scales, separating it from the noise present in other frequency bands. With the signal already demodulated, the task is then to determine the scales that show periodicity close to or at the cyclic frequencies associated with the fault.

The Reduced Discrete Fourier Series is a unique tool that, in a least squared sense, fits a set of sinusoids of arbitrary frequency on a data set. It is the use of the DRFS algorithm that makes the proposed detection and de-noising algorithms in this thesis unique. The CWT/DRFS cyclic energy algorithm uses the DRFS algorithm to estimate

the periodicity of each demodulated scale at a given cyclic frequency. This allows the wavelet transform to be turned into a tool that can detect hidden amplitude modulated signals at discrete frequencies. The automatic CWT/DRFS de-noising algorithm uses the DRFS algorithm to determine the wavelet scale shows the most periodicity at the fault frequency of interest and returns that scale as the de-noised fault signal. Finally the DRFS Wavelet Thresholding algorithm uses the DRFS to create a periodic threshold for each scale at the cyclic frequency of the fault. It then selectively chooses the wavelet scales that maximize the inverse CWT to create a de-noised signal. The most unique algorithm provided by this research is the CWT/DRFS cyclic energy distribution. Viewed by an analyst this distribution is an easy tool to use to diagnose a bearing fault and can be used with envelope demodulation to center frequency and bandwidth of the filter.

The algorithms were then compared with the industry standard envelope demodulation and seven other algorithms. The proposed de-noising algorithms were more successful at de-noising inner and outer race faults than a number of other proposed continuous wavelet techniques as well as a novel filtering method. In several examples the methods proposed in this paper allowed automatic detection of very weak inner and outer race bearing fault signals that all other techniques were unable to detect.

In conclusion the initial results of this thesis are promising. The proposed algorithms worked well on the signals tested. Further verification needs to be performed to determine how well these algorithms work on other bearing faults as well as over the full lifetime of several rolling element bearings. This would allow these algorithms to be

judged on the basis of how early they could allow signal detection. The CWT/DRFS algorithm also shows promise in other areas of research as well. It is known that transmission faults show similar effects as inner and outer race faults with modulation of periodic impulses. In this case both the modulation frequency and the cyclic frequency can be estimated. It is proposed that the CWT/DRFS detection tool could also be used to detect and de-noise these types of faults as well.

REFERENCES

[1] P. Gupta, Dynamics of rolling element bearings, Part I, II, III and IV, Journal of Lubrication Technology 101 (1979) 293-326.

[2] R. Randall, State of the art in monitoring rotating machinery – Part 1, Sound & Vibration, March (2004) 14-21.

[3] R. Randall, State of the art in monitoring rotating machinery – Part 2, Sound & Vibration, May (2004) 10-17.

[4] P. McFadden, J. Smith, Vibration monitoring of rolling element bearings by the high-frequency resonance technique – a review, Tribology International 17 (1984) 3-10.

[5] C. Junsheng, Y. Dejie, Y. Yu, Application of an Impulse Response Wavelet to fault diagnosis of rolling bearings, Mechanical Systems and Signal Processing 21 (2007) 920-929.

[6] Z. Peng, F. Chu, Application of the Wavelet Transform in machine condition monitoring and fault diagnosics: a review with bibliography, Mechanical Systems and Signal Processing 18 (2004) 199-221.

[7] F. Xi, Q. Sun, G. Krishnappa, Bearing diagnostics based on pattern recognition of statistical parameters, Journal of Vibration and Control 6 (2000) 375-392.

[8] Z. Peng, P. Tse, F. Chu, A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing, Mechanical Systems and Signal Processing 19 (2005) 974-988.

[9] J. Lin, M. Zuo, Extraction of periodic components for gearbox diagnosis combining wavelet filtering and cyclostationary analysis 126 (2004) 449-451.

[10] J. Lin, L. Qu, Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis, Journal of Sound and Vibration 234 (2000) 135-148.

[11] L. Li, L. Qu, X Liao, Harr wavelet for machine fault diagnosis, Mechanical Systems and Signal Processing 21 (2007) 1773-1786.

[12] J. Altmann, J. Mathew, Multiple band-pass autoregressive demodulation for rolling-element bearing fault diagnosis, Mechanical Systems and Signal Processing 15 (2001) 963-977.

[13] H. Ocak, K. Loparo, F. Discenzo, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: a method for bearing prognostics, Journal of Sound and Vibration, 302 (2007) 951-961.

[14] A. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing 20 (2006) 1483-1510.

[15] Q. Sun, Y. Tang, Singularity analysis using Continuous Wavelet Transform for bearing fault diagnosis, Mechanical Systems and Signal Processing 16 (2002) 1025-1041.

[16] S. Ericsson, N. Grip, E. Johansson, L. Persson, R. Sjoberg, J. Stromberg, Towards automatic detection of local bearing defects in rotating machines, Mechanical Systems and Signal Processing 19 (2005) 509-535.

[17] P. Tse, Y. Peng, R. Yam, Wavelet analysis and envelope detection for rolling element bearing fault diagnosis – their effectiveness and flexibilities, Journal of Sound and Vibration 123 (2001) 303-310.

[18] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognosis, Journal of Sound and Vibration 289 (2006) 1066-1090.

[19] F. Li, J. Chen, G. C. Zhang, Z. He, Wavelet transform domain filter and its application in incipient fault prognosis, Key Engineering Materials 293-294 (2005) 127-134.

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

[21] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: facts and fictions, Journal of Sound and Vibration 304 (2007) 497-529.

[22] J. Antoni, F. Bonnardot, A. Radd, M. El Badaoui, Cyclostationary modeling of rotating machine vibration signals, Mechanical Systems and Signal Processing 18 (2004) 1285-1214.

[23] R. Randall, J. Antoni, S. Chobsaard, The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals, Mechanical Systems and Signal Processing 15 (2001) 945-962.

[24] J. Antoni, R. B. Randall, A stochastic model for simulation and diagnostics of rolling element bearings with localized faults, Mechanical Systems and Signal Processing 125 (2003) 282-289.

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

[26] P. McFadden, J. Smith, The vibration produced by multiple point defects in a rolling element bearing, Journal of Sound and Vibration 98 (1985) 263-273.

[27] M. Pan, P. Sas, Transient analysis on machinery condition monitoring, International Conference on Signal Processing Proceedings, ICSP 2 (1996) 1723-1726.

[28] S. Braun, The signature analysis of sonic bearing vibrations, IEEE Transactions of Sonics and Ultrasonics 27 (1980) 317-328.

[29] J. Antoni, R. Randall, A stochastic model for simulation and diagnostics of rolling element bearings with localized faults, Journal of Vibration and Acoustics 125 (2003) 282-289.

[30] J. Donelson, R. Dicus, Bearing defect detection using on-board accelerometer measurements, 2002 ASME/IEEE Joint Railroad Conference (2002) 95-102.

[31] Y. Wang, P. Kootsookos, Modeling of low shaft speed bearings faults for condition monitoring, Mechanical Systems and Signal Processing 12 (1998) 415-426.

[32] J. Shiroishi, Y. Li, S. Liang, T. Kurfess, S. Danyluk, Bearing condition diagnostics via vibration and acoustic emission measurements, Mechanical Systems and Signal Processing 11 (1997) 693-705.

[33] D. Ho, R. Randall, Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals, Mechanical Systems and Signal Processing 14 (2000) 763-788.

[34] G. Dalpiaz, A. Rivola, Conditions monitoring and diagnostics in automatic machines: comparison of vibration analysis techniques, Mechanical Systems and Signal Processing 11 (1997) 53-73.

[35] S. Ericsson, N. Grip, E. Johansson, L. Persson, R. Sjoberg, J. Stromberg, Towards automatic detection of local bearing defects in rotation machines, Mechanical Systems and Signal Processing 19 (2005) 509-535.

[36] P. Tse, Y. Peng, R. Yam, Wavelet analysis and envelope detection for rolling element bearing fault diagnosis – their effectiveness and flexibilities, Journal of Vibration and Acoustics 123 (2001) 303-310.

[37] J. Zarei, J. Poshtan, Bearing fault detection using wavelet packet transform of induction motor stator current, Tribology International 40 (2007) 763-769.

[38] P. Chen, M. Taniguchi, T. Toyota, Z. He, Fault diagnosis method for machinery in unsteady operating condition by instantaneous power spectrum and genetic programming, Mechanical Systems and Signal Processing 19 (2005) 175-194.

[39] Q. Hu, Z. He, Z. Zhang, Y. Zi, Fault diagnosis of rotating machinery based on improved wavelet package transform and SVM ensemble, Mechanical Systems and Signal Processing 21 (2007) 688-705.

[40] L. Zhang, L. Jack, A. Nandi, Fault detection using genetic programming, Mechanical Systems and Signal Processing 19 (2005) 271-289.

[41] D. Donoho, De-noising by soft-thresholding, IEEE Transactions on Information Theory 41 (1995) 613-621.

[42] F. Abramovich, T. Bailey, T. Sapatinas, Wavelet analysis and its statistical applications, The Statistician 49 (2000) 1-29.

[43] H. Hong, M. Liang, K-Hybrid: A kurtosis-based hybrid thresholding method for mechanical signal denoising, Journal of Vibration and Acoustics 129 (2007) 458-470.

[44] J. Lin, M. Zuo, K. Fyfe, Mechanical fault detection based on the wavelet de-noising technique, Journal of Vibration and Acoustics, 126 (2004) 9-16.

[45] I. Bozchalooi, M. Liang, A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection, Journal of Sound and Vibration 308 (2007) 246-267.

[46] N. Sawalhi, R. Randall, Spectral kurtosis optimization for rolling element bearings, Proceedings of the Eighth International Symposium on Signal processing and Its Applications. 2005, IEEE: Sydney, Australia.

[47] C. Junsheng, Y. Dejie, Y. Yu, Application of an Impulse Response Wavelet to fault diagnosis of rolling bearings, Mechanical Systems and Signal Processing 21 (2007) 920-929.

[48] L. Li, L. Qu, X. Liao, Haar wavelet for machine fault diagnosis, Mechanical Systems and Signal Processing 21 (2007) 1773-1786.

[49] K.F. Al-Raheem, A. Roy, K.P. Ramachandran, O.K. Harrison, S. Grainger, Detection of rolling element bearing faults using optimized – wavelet de-noising technique, The 8th International Conference on Signal Processing 2006, IEEE: Vienna, Austria.

[50] B.T. Holm-Hansen, R.X. Gao, L. Zhang, Customized wavelet for bearing defect detection, Journal of Dynamic Systems, Measurement and Control 126 (2004) 740-745.

[51] J. Hongkai, H. Zhengjia, D. Chendong, C. Peng, Gearbox fault diagnosis using adaptive redundant lifting scheme, Mechanical Systems and Signal Processing 20 (2006) 1992-2006.

[52] B. Lui, S.F. Ling, On the selection of informative wavelet for machinery diagnosis, Mechanical Systems and Signal Processing 13 (1999) 145-162.

[53] H. Yang, J. Mathew, L. Ma, Fault diagnosis of rolling element bearings using basis pursuit, Mechanical Systems and Signal Processing 19 (2005) 341-356.

[54] B. Hubbard, The World According to Wavelets: The Story of a Mathematical Technique in the Making, A K Peters, Wellesley, MA, 1998.

[55] J.G. Proakis, D.G. Manolakis, Digital Signal Processing, Principles, Algorithms and Applications, Prentice Hall, Upper Saddle River, NJ, 1996.

[56] L.G. Turin, An introduction to matched filters, IRE Transactions on Information Theory 6 (1960) 311-329.

[57] P.S. Addison, The Illustrated Wavelet Transform Handbook, Institute of Physics Publishing, Bristol, England, 2002.

[58] C.L. Phillips, J.M. Parr, Signals, Systems and Transforms, Prentice Hall, Upper Saddle River, NJ, 1999.

[59] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, London, 1998.

[60] H. Endo, R. Randall, Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter, Mechanical Systems and Signal Processing 21 (2007) 906-919.

[61] Y. Sheen, A compelx filter for vibration signal demodulation in bearing defect diagnosis, Journal of Sound and Vibration, 276 (2004) 105-119.

[62] L. Zhen, H. Zhengjia, Z. Zanyang, W. Yanxue, Customized wavelet denoising using intra- and inter-scale dependency for bearing fault detection, Journal of Sound and Vibration 313 (2008) 342-359.

[63] V.K. Madisetti, D.B. Williams, The Digital Signal Processing Handbook, CRC Press, New York, NY, 1998.

[64] A. Hyvarinen, Sparse code shrinkage: denoising of nongaussian data by maximum likelihood estimation. Neural Computation 11 (1999) 1739-1768.

[65] N. Sawalhi, R. Randall, H. Endo, The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis, Mechanical Systems and Signal Processing 21 (2007) 2616-2633.

[66] J.R.F. Arruda, Surface smoothing and partial spatial derivatives computation using a regressive discrete Fourier Series, Mechanical Systems and Signal Processing 6 (1992) 41-50.

[67] K. Laparo, Welcome/Overview, Bearing Data Center: Seeded Fault Test Data, 2008, Case Western Reserve University, 30 April, 2008. <http://www.eecs.case.edu/laboratory/bearing/welcome_overview.htm>

VITA

Name:	Scott Weatherwax
Address:	Texas A&M University Department of Mechanical Engineering 3123 TAMU College Station, TX 77843-3123
Email Address:	waxer@tamu.edu
Education:	B.S., Electrical Engineering, Texas A&M University, 2004 M.S., Mechanical Engineering, Texas A&M University, 2008