

University of Alberta

Vibration Signal-Based Fault Detection for Rotating Machines

by

Geoffrey Lyall McDonald

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
CONTROLS

Department of Electrical and Computer Engineering

©Geoffrey Lyall McDonald
Fall 2011
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

Fault detection in rotating machines from vibration data is a difficult task and is important for maintenance planning and preventing equipment damage or failure. The aim of this thesis is to improve upon existing vibration signal methods for detecting rotating machine faults in gears, bearings, and rotors. Faults manifesting in impulse-like vibration signals are focused on, which includes faults such as rotor-to-stator rubbing, bearing inner/outer race failures, and gear tooth faults. Towards this goal, two novel techniques for detecting these faults are proposed in this thesis and experimental data from gear tooth crack, gear tooth chip, and suspected turbine rotor-to-stator rubbing is analysed.

An adaptive sum-of-sinusoids model is presented and compared to the widely accepted autoregressive model approach. The results indicate that the proposed method performs better on experimental gear tooth crack data, requires no data fitting, and is of similar computational cost. This method is particularly suitable for equipment with changing rotational speed.

A deconvolution-based approach is presented as a periodic extension upon the minimum entropy deconvolution method. The proposed method aims to deconvolve periodic impulses, which is the vibration signal nature of many rotating machine faults, as opposed to the single impulse deconvolved by minimum entropy deconvolution. Performance of the deconvolution technique are shown to be strong on simulated and experimental gear tooth chip data, and an online implementation is proposed for the automated monitoring of equipment.

Both proposed methods are applied to turbine proximity sensor data, and a sensor fault plus two suspected rotor-to-stator rubbing faults are identified.

Acknowledgements

My supervisor, Dr. Qing Zhao, for her valuable support, contributions, and friendship.

Dr. Ming J. Zuo for his detailed review of and comments on the MCKD chapter.

David (Youliang Yang) for his work on data conversion of the power generator turbine data. Without him, the turbine data analysis chapter would not be possible.

X. Tian, Dr. Ming J. Zuo, J. Lin, M. Agelichaab, J. Zhang, M. Akhtar, and K. R. Fyfe for their work on the gearbox with tooth chip experiment. Their work is instrumental to the MCKD experimental results section.

Xianfeng Fan, Dr. Ming J. Zuo, Yao Wang, Wei Li, and Siyan Wu for their detailed experimental work on the gearbox with tooth fillet crack. Their results assisted me greatly with the presentation of the FE method.

Jian Qu and Tejas Patel from the Reliability Research Lab, Mechanical Engineering University of Alberta, for their help with access to the equipment and taking the time to navigate the data archives.

My parents for taking time to read, review, and provide feedback on my work.

NSERC and Syncrude Canada Ltd. for making this possible through funding support.

Contents

1	Introduction	1
1.1	Problem Statement	1
1.2	Scope of Research	1
1.3	Contribution	2
2	Literature Review	3
2.1	Rotating Machine Faults Overview	3
2.2	Spectral Analysis Methods	6
2.2.1	Wavelet Transform	6
2.2.2	Cyclostationary Analysis	7
2.2.3	Cepstrum Analysis	9
2.3	Model-Based Methods	10
2.3.1	Autoregressive Model	10
2.3.2	Observer Design	11
2.4	Filtering-Based Methods	11
2.4.1	Time-Synchronous Averaging	11
2.4.2	Filter Selected by Wavelet Transform	12
2.4.3	Filter Selected by Spectral Kurtosis	12
2.4.4	Minimum Entropy Deconvolution	14
2.5	Fault Indication	16
3	Frequency Estimator-Based Fault Detection	18
3.1	N-Component Frequency Estimator (FE)	19
3.2	Adaptive Identifiers	21
3.3	Least-Squares (LSQ) Parameter Selection	23
3.4	Computational Requirements	26
3.5	Simulation Results	26
3.6	Experimental Design and Results	27
3.7	Conclusion	35

4	Maximum Correlated Kurtosis Deconvolution Fault Detection	36
4.1	Correlated Kurtosis	37
4.2	Maximum Correlated Kurtosis Deconvolution	39
4.2.1	First-Shift Maximum Correlated Kurtosis Deconvolution	39
4.2.2	M-Shift Maximum Correlated Kurtosis Deconvolution	41
4.3	Simulation Results	43
4.4	Basic Concurrent Gear Fault Simulation	43
4.5	Experimental Results	47
4.6	Online Condition Monitoring Implementation	56
4.7	Conclusion	57
5	Steam Power Turbine Analysis	60
5.1	Turbine 1 Health Analysis	61
5.1.1	Sensor Fault	62
5.1.2	Rotor-Stator Rubbing	63
5.2	Turbine 2 Health Analysis	68
5.3	Conclusion	68
6	Conclusion and Future Work	71
6.1	Conclusion	71
6.2	Future Work	72
A	External Resources	73

List of Figures

2.1	Example gear faults under a) tooth chip, b) tooth fillet crack [11], c) tooth pitting [12], d) missing tooth [13], and e) tooth spall [14] cases.	4
2.2	General overview of gear and bearing fault detection.	5
2.3	W. J. Wang and P. D. McFadden [34] fault detection of a helicopter gearbox by TSA and continuous wavelet transform: a) under no-fault, and b) under fault.	6
2.4	Z.K. Peng, P.W. Tse, and F.L. Chu [37] performed a fault detection comparison between a wavelet-based scalogram and Improved Hilbert-Huang Transform-based scalogram. The vibration data was from an inner-race ball bearing fault. In the caption of the original figure the Hilbert-Huang spectrum is referred to as the Improved Hilbert-Huang spectrum. The Scalogram is the wavelet scalogram from the WPT.	8
2.5	A.C. McCormick and A.K. Nandi [45] analysed bearing faults by second-order cyclostationary analysis. The plot is of the the spectral correlation density function, which is the DFT amplitude spectrum of the time-varying cyclic autocorrelation. The sample autocorrelation is the second-order cyclostationary component.	9
2.6	AR prediction residual method. $y[n]$, $\hat{y}[n]$, and $r[n]$ are the input signal, predicted signal, and prediction residual respectively.	11
2.7	N.G. Nikolaou and I.A. Antoniadis [19] analysed inner-race bearing faults by shifted complex Morlet wavelet transform max-envelope technique. T_{BPFI} is the critical period of the bearing fault.	13
3.1	Continuous wavelet packet transform of vibration under gear chip using the complex Morlet mother wavelet.	19
3.2	N-component DT system model	20
3.3	LSQ minimization norm versus iteration for the parameter selection of the two and three-component FE.	24
3.4	Convergence of θ estimates for the three-component FE for the first four input signals each composed of three sinusoidal components at indicated frequencies.	24
3.5	Convergence of the frequency estimates for the first four input signals with the two-component FE model.	25
3.6	Simulated gear crack vibration signal and spectrum.	27
3.7	Prediction residual and corresponding kurtosis using the a) 2-component FE, b) 3-component FE, c) 9th order AR, and d) 146th order AR model. . . .	28

3.8	Kurtosis of model prediction residual plotted versus fault peak disturbance.	28
3.9	Experimental machine configuration.	29
3.10	Experimental gears with tooth cracks at levels of a) 25%, b) 50%, c) 75%, and d) 100% [56].	29
3.11	RMS of measured acceleration according to fault level.	30
3.12	Prediction error by method for the experimental setup at 3000 rpm, full load, and 100% fault level.	31
3.13	Prediction error by method for the experimental setup at 3000 rpm, full load, and 100% fault level.	31
3.14	$\hat{\theta}$ versus sample for the three-component frequency estimator for the machine at 3000 rpm, full load, and 100% fault level.	32
3.15	$\hat{\omega}$ versus sample for the two-component frequency estimator at 3000 rpm, full load, and 100% fault level.	32
3.16	Kurtosis of the signals averaged across all rpm settings under a) full load, b) half-load, and c) no load.	33
3.17	Kurtosis of the signals averaged across all rpm settings and plotted as a percentage of no-load fault kurtosis under a) full load, b) half-load, and c) no load.	34
4.1	Kurtosis and CK values for several signals. The kurtosis reaches a maximum with a single impulse, the CK_1 reaches a maximum with 3 impulses. The first signal is white noise.	38
4.2	CK_1 spectrum for an impulse train of fives impulses with spacing of T and strong additive Gaussian white noise corruption.	38
4.3	CK_1 values for a simple simulated repetitive fault with various amplitudes of the base harmonic vibration. All signals are normalized by their peak value for better illustration.	39
4.4	Deconvolution of a noisy periodic impulse train of period 100 with a) illustrating the fault detection versus impulse train amplitude, and b) normalized plot of the resulting deconvolved signals with impulse train amplitude of $f = 3$, some signals are polarity inverted for illustration purposes. A different white noise seed is used for each value of f.	44
4.5	Simulated resulting signals for concurrent fault detection after processing with indicated method. All signals are normalized and some signals are polarity flipped for better illustration.	46
4.6	Gearbox and experimental equipment layout.	47
4.7	Faulty gear 1 with seeded tooth chip.	48
4.8	Deconvolved signals versus deconvolution iteration for a) first-shift MCKD, b) MED, and c) ARMED for an AR model order of 100 and filter sizes of 200 samples. The vibration data is from the gearbox under gear chip fault at 40Hz shaft speed and 10% load and resulting signals are normalized by peak values with some signals being polarity swapped for better illustration.	49

4.9	a) MCKD and b) MED methods versus deconvolution filter size for the machine under gear chip fault at 40Hz shaft speed and 10% load. Each deconvolution method is iterated 50 times and the signals are normalized by peak values polarity swapped for clearer illustration.	49
4.10	Experimental acceleration data with fault detection processing with a) gear chip and shaft 1 speed of 35 Hz, and b) normal gear and shaft 1 speed of 35 Hz. All signals are normalized by peak values and some signals are reversed in polarity for better illustration.	51
4.11	Experimental acceleration data processed and normalized with a gear chip using a) <i>MCKD</i> ₁ method, b) <i>MCKD</i> ₃ method, c) MED method, and c) the ARMED method. All signals are normalized by peak values and some signals are polarity flipped for better illustration.	52
4.12	Experimental fault detection by method with varied deconvolution filter size for the machine operating at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz. Be careful to note the log-scale on the y-axis, meaning the <i>MCKD</i> ₃ and <i>MCKD</i> ₅ typically perform significantly better at fault detection.	54
4.13	Online implementation of MCKD-based fault detection for a two-shaft gearbox with a single accelerometer. Threshold and FIR filter plots are creative illustrations, not data.	55
4.14	Online results for first-shift MCKD fault detection for machine under 10% load at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz. The data is from looped the experimental data.	58
4.15	Online results for fifth-shift MCKD fault detection for machine under 10% load at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz.	58
5.1	The layout of the back pressure steam turbine generator. Source: [58]	60
5.2	RMS of turbine vibration for the X and Y axis on a) Turbine 1, and b) Turbine 2. Each dot represents the RMS of the dataset of approximately 1200 samples collected at the indicated time.	61
5.3	Sensor fault on the Y-Axis of Turbine 1. a) The long-term RMS vibration, b) vibration signal just before a sensor fault, c) vibration signal during a sensor fault, and d) vibration just after a sensor fault.	62
5.4	Fault indicators on the X-Axis of Turbine 1 vibration data by a) AR, b) two-component FE, c) three-component FE, d) MED, e) ARMED, f) third-shift MCKD, and g) fifth-shift MCKD. The Y-Axis generally agrees with the results, but is messy as a result of the sensor faults.	64
5.5	Detailed analysis of the MED result. The poor Kurtosis fault indicator performance is as a result of spurious deconvolved impulses, and MED agrees that there is a rotor rubbing fault in the highlighted region.	65
5.6	Downtime analysis of Turbine 1, with introduction and fix of rotor-rubbing indicated.	65
5.7	The original measured vibration and processed data for a suspected rotor rubbing vibration dataset.	66

5.8	The position of the indicated faults according to the orbit map. a) and c) illustrates the x and y-axis vibration with overlaid fifth-shift MCKD deconvolved signal, b) and d) illustrate the resulting deconvolution filters, and d) illustrates the x and y-axis fault locations. The vibration data is taken from Turbine 1 within the suspected rotor rubbing fault region.	67
5.9	The a) RMS vibration and fault indicators on the X-Axis of the Turbine 2 vibration data by b) AR, c) two-component FE, d) three-component FE, e) ARMED, f) third-shift MCKD, and g) fifth-shift MCKD. The Y-Axis generally agrees with the results.	69
5.10	The position of the indicated faults according to the orbit map. a) and c) illustrates the x and y-axis vibration with overlaid fifth-shift MCKD deconvolved signal, b) and d) illustrate the resulting deconvolution filters, and d) illustrates the x and y-axis fault locations. The vibration data is taken from Turbine 2 within Region 2 of Figure 5.9.	70

List of Tables

- 3.1 Final values for FE parameters γ and α by LSQ minimization 23
- 3.2 Floating point calculations required for each sample of predicted output as a function of model order, M , or number of frequency components, N 26

- 4.1 Fault indicator values by MCKD, MED, and ARMED. The fault indicators are final CK values for MCKD and the final Kurtosis values for MED and ARMED. All values are no-fault normalized. 47
- 4.2 Estimated shaft 1 period and corresponding deconvolution filter sizes for no-fault and fault cases. 50
- 4.3 Computation times for online *MCKD* gear chip fault detection on looped experimental data. 57

List of Symbols

ANN	Artificial Neural Network
AR	Autoregressive
CK	Correlated Kurtosis
DFT	Discrete Fourier Transform
DT	Discrete-Time
FE	Frequency Estimator
FIR	Finite Impulse Response
HMM	Hidden Markov Model
i.i.d.	Independent and Identically Distributed
LSQ	Least-Squares
MCKD	Maximum Correlated Kurtosis Deconvolution
MED	Minimum Entropy Deconvolution
RMS	Root Mean Square
SK	Spectral Kurtosis
SVM	Support Vector Machine
TSA	Time-Synchronous Average

1

Introduction

1.1 Problem Statement

Fault detection in rotating machinery has applications in fields such as wind turbines [1] and helicopter transmissions [2]. Detecting and diagnosing faults is important to maintenance planning, preventing equipment damage, and preventing failure. In some applications, such as helicopter transmissions, a gear fault can potentially result in a life-threatening situation [3]. Despite a large amount of research focus on this area, the detection of these faults can still be a very difficult problem.

For detecting gear and rolling bearing faults there has been a growing trend towards autoregressive (AR) model prediction residual fault detection [2, 4, 5, 6, 7], which has been shown to be effective in detecting the impulse-like behaviour associated with gear faults. Although this method is shown to be effective, the expected residual signal is the fault signal plus noise, and therefore requiring the fault amplitude to be significantly larger than the noise level to be detectable by using the AR model alone. In addition, the AR method requires a priori knowledge of the vibration under no-fault conditions for data-fitting.

1.2 Scope of Research

Detection of faults in rotating machines can be achieved through many different approaches including vibration, current, voltage, ultrasound, lubrication oil analysis, and many more approaches. The most popular approach is vibration due to the simplicity of measurement

(typically mounted on the machine casing), effectiveness in diagnosis, and a relatively low cost for the sensors. The scope of this thesis research is limited to vibration-based fault detection methods and is limited to impact faults; which typically share the same detection methodologies. These fault types include rotor rubbing, bearing inner/outer race faults, and various gear tooth faults. A detailed review of methods for detecting these types of faults is presented in Chapter 2.

1.3 Contribution

Two contributions to the field of rotating machine fault detection are presented in this thesis.

Firstly, an adaptive predictive model is proposed for the detection of gear tooth fillet cracks in a gearbox. The method is based upon a recently proposed N-component frequency estimator (FE) by M. Hou [8], and the method is adapted for the detection of rotating machine faults. Novel contributions include a new amplitude-invariant adaptive identifier, the discrete-time derivation, computational cost analysis, and LSQ-based parameter selection methodology. The method performs well in detecting the gear crack when compared to a commonly accepted alternative technique and includes some major advantages over this traditional approach. A large part of this work was presented as a regular paper at the American Controls Conference, June 28 to July 2nd 2011, San Francisco [9].

Secondly, a novel deconvolution technique called Maximum Correlated Kurtosis Deconvolution (MCKD) is proposed for the extraction of periodic impulse-like components from a one-dimensional signal. This proposed method is well-suited for the extraction of rotating machine faults which manifest in periodic impulse-like disturbances in the vibration. When compared to the standard Minimum Entropy Deconvolution (MED) technique, which aims to extract a single impulse from a signal, the proposed MCKD perform remarkably better on synthetic and experimental data. A large part of this work is currently under review for the journal Mechanical Systems and Signal Processing [10].

2

Literature Review

2.1 Rotating Machine Faults Overview

The focus of this literature review is on the detection of rotating machine faults from vibration data. Additionally, the main focus is on impact-faults such as rotor rubbing, rolling element bearing inner and outer-race faults, and gear tooth faults. These faults manifest in similar ways on the machine vibration data, and the detection methodologies typically work for all of these classes of faults.

Examples of gear tooth faults include tooth chips, fillet cracks, pitting, spalling, and missing tooth. Examples of these fault categories are illustrated in Figure 2.1.

These types of rotating machine faults are known to periodically excite the machine dynamics of the system with time-localized disturbances. I. Howard et al. [15] used a dynamic 16-degree of freedom model with friction to show that a gear tooth fillet crack causes a time localized disturbance at the meshing time, raises the overall whiteness of the spectrum, and the gear meshing and harmonics remain of similar amplitude. Vibration from an experimental setup under tooth fillet crack and spall were both found to be impulse-like excitations by H. Endo and R.B. Randall [5]. The experimental data from F. Combet and L. Gelman indicates that a gear pitting causes an impulse-like time localized disturbance [16]. Ball bearing faults manifest similarly as impulse-like disturbances in the vibration [17, 7, 18, 19, 20].

Fault detection methods are broken into several sections and are listed in Figure 2.2 along



Figure 2.1: Example gear faults under a) tooth chip, b) tooth fillet crack [11], c) tooth pitting [12], d) missing tooth [13], and e) tooth spall [14] cases.

with references. Vibration processing approaches utilize either a single method or a combination of the following:

- 1: Filtering methods. The goal of the filtering methods is to extract a time-domain signal where the fault is clearer than before processing. Fault indication is typically approached by threshold, artificial neural network (ANN), fuzzy logic, support vector machines (SVM), or hidden markov model (HMM) based on metrics of this resulting signal.
- 2: Spectral analysis methods. Spectral analysis methods aim to generate a spectral representation of the signal and is typically represented as three-dimensional plots, contour plots, or heat-map images. Fault detection based on these approaches are typically done by expert human analysis of the plots.
- 3: Model-based methods. Models are used for enhancing the contrast between the fault signal and the vibration data. The output of these approaches are typically a one-dimensional residual signal and faults are detected using signal metrics with a threshold, ANN, fuzzy logic, SVM, or HMM alarm.

Each category of processing has its own disadvantages and advantages. These general categories of vibration processing and fault indication techniques are reviewed in the following sections. A detailed background for the AR model and MED methods are provided since Chapters 3 and 4 are related to these methods.

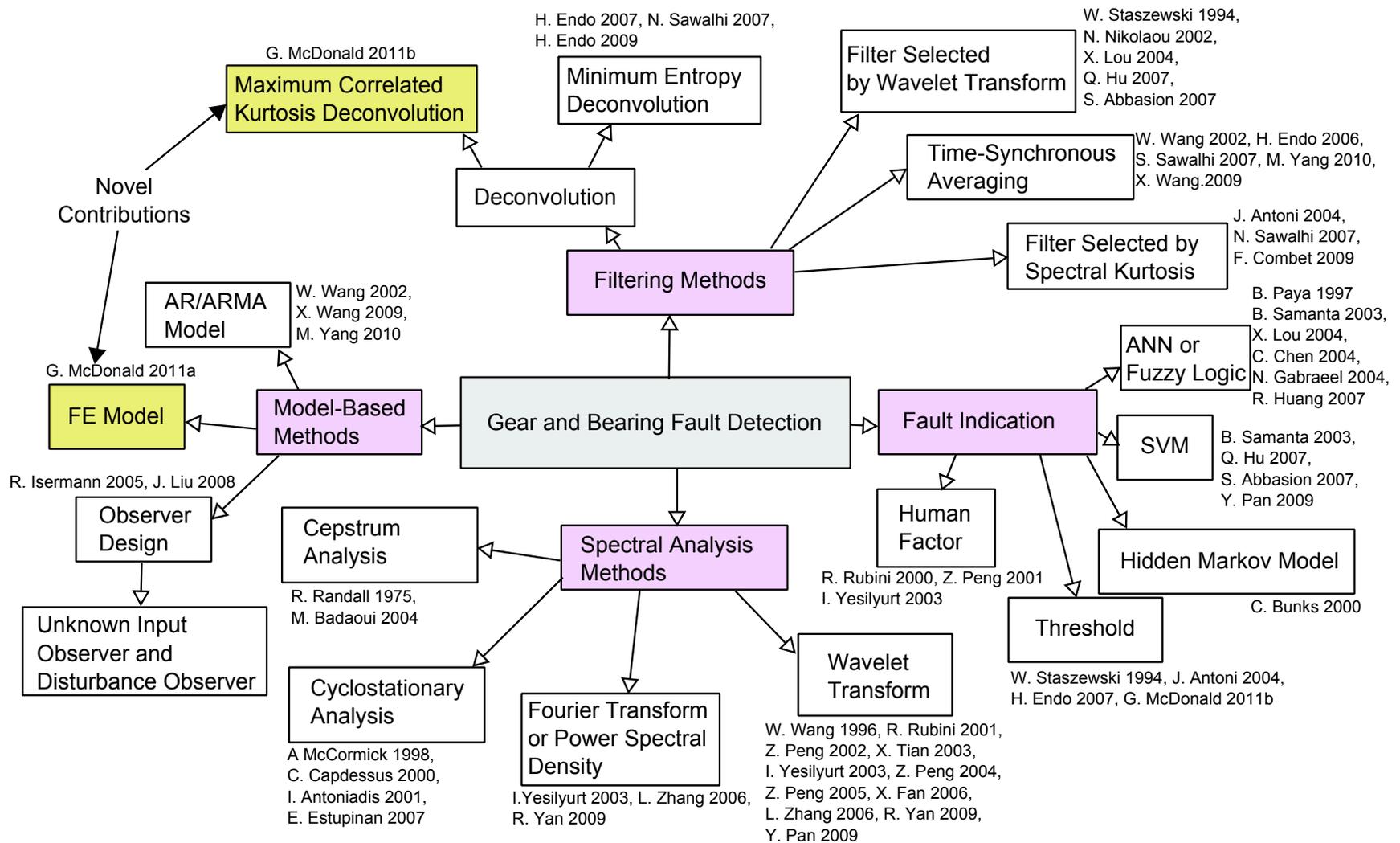


Figure 2.2: General overview of gear and bearing fault detection.

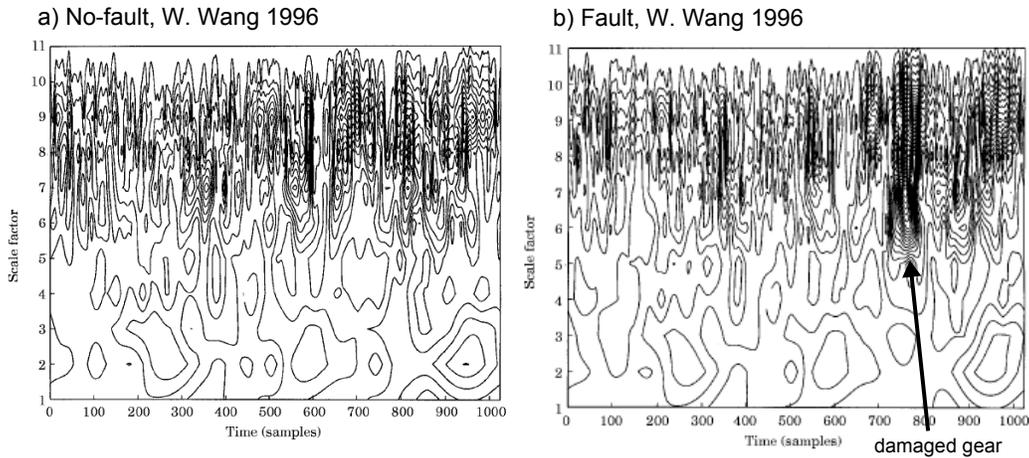


Figure 2.3: W. J. Wang and P. D. McFadden [34] fault detection of a helicopter gearbox by TSA and continuous wavelet transform: a) under no-fault, and b) under fault.

2.2 Spectral Analysis Methods

Spectral analysis methods have been used to analyse machine vibration data for decades and this section focuses on recent trends and literature. Most of the spectral analysis methods can be categorized into wavelet transform, cyclostationary analysis, and cepstrum-based methods. Presented in the following subsections is a brief review of some of these approaches with application of rotating machinery fault detection. Applications of the wavelet transform as a filtering method are included in Section 2.4.

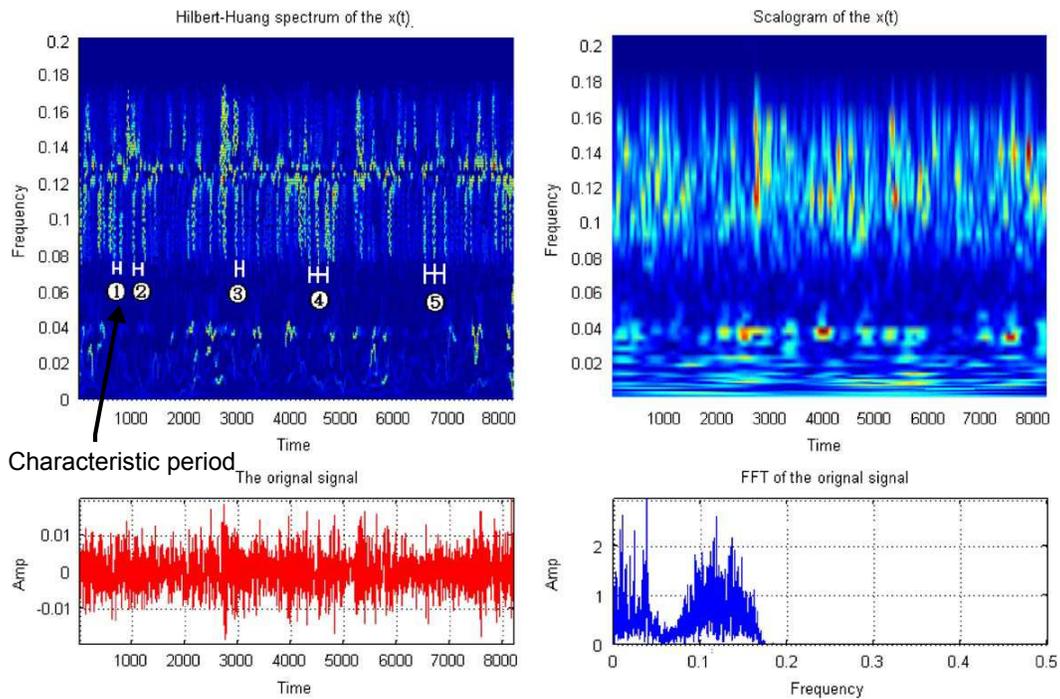
2.2.1 Wavelet Transform

The wavelet transform spectrum has been used for fault detection of gears and bearings in many papers dating back a few decades. The wavelet transform is the leading time-frequency signal analysis method, and is commonly applied in the analysis of non-stationary signals. It has been proven to be a very effective diagnostic tool and a brief summary of some of the papers is provided. In 1996, W. J. Wang and P. D. McFadden [34] applied time-synchronous averaging (TSA) and the continuous wavelet transform (CWT) to vibration data from a helicopter gearbox with a fatigue crack. The authors proposed a computation algorithm and demonstrated the first use incorporating a variable time-frequency resolution depending on the frequency for gear fault detection. Diagnosis was performed by analysis of the resulting wavelet scalogram contour maps, Figure 2.3, by an expert human. R. Rubini and U. Meneghetti [17] found that rolling bearing fatigue flaws resulted in increased spectrum peaks at the characteristic frequency. The authors proposed averaging the CWT sections across a frequency range, and taking the DFT of this averaged signal.

Fault indication was performed by human comparison of a characteristic frequency magnitude. Interestingly, after about 30 minutes of operation the fault was undetectable and the authors attributed this to the flattening-out of the fault. The reassigned wavelet scalogram application to vibration analysis was proposed by Z.K. Peng, F.L. Chu, and Y.He in 2002 [35]. The authors found that the discrete Fourier transform (DFT) spectrum failed to indicate faults at early stages, and that the reassigned wavelet scalogram reduced interference terms. Fault indication was performed through expert human analysis of these resulting wavelet scalograms. Another alternative scalogram representation was proposed by I. Yesilyurt in 2003 [13]. TSA was performed as an initial filtering stage, and the wavelet-based instantaneous power spectrum was applied and smoothed along the time axis to reduce oscillations. Results were strong in detecting a missing gear tooth fault. Z.K. Peng, P.W. Tse, and F.L. Chu in 2005 [37] proposed the application of an alternative to the wavelet spectrogram called the Improved Hilbert-Huang Transform. This transform is a combination of narrow-band filtering, Empirical Mode Decomposition, and Hilbert Transform. They compared the scalogram results on experimental data from bearing inner and outer-race faults. Their results, an excerpt in Figure 2.4, indicated a strong time localization of the signal. Fault identification was performed by expert human analysis of the scalograms. X. Fan and M.J. Zuo in 2006 [38] presented a method combining the WPT with the Hilbert Transform to detect faults in gear tooth chip vibration data. First the Hilbert Transform was performed to extract the signal envelope, then WPT was applied to this envelope to identify the modulating signals. Fault indication was performed by expert human analysis of the resulting scalogram. In general many methods of using the wavelet transform spectrum and variations upon the method have been investigated, and it is clearly still an active area of research.

2.2.2 Cyclostationary Analysis

Cyclostationary analysis is another spectrum method applied to fault detection of gears and bearings, and takes advantage of the stochastic process nature of the vibration signal. A signal is cyclostationary of degree N if its N th moment is periodic about any period. A.C. McCormick and A.K. Nandi applied cyclostationary analysis to machine vibrations in 1998 [45]. Their methodology analysed the DFT of the cyclostationary signal about a range of periods, and also used the sum of cyclostationarity order of vibration data under ball bearing faults around the expected fault period. Results for a second-order cyclostationary spectrum for bearing faults are illustrated in Figure 2.5. C. Capdessus, M. Sidahem, and J.L. Lacoume applied cyclostationary analysis towards gear faults in 2000 [46]. The presented vibration signal was shown to be second-order cyclostationarity and this parameter was



Characteristic period

Figure 2.4: Z.K. Peng, P.W. Tse, and F.L. Chu [37] performed a fault detection comparison between a wavelet-based scalogram and Improved Hilbert-Huang Transform-based scalogram. The vibration data was from an inner-race ball bearing fault. In the caption of the original figure the Hilbert-Huang spectrum is referred to as the Improved Hilbert-Huang spectrum. The Scalogram is the wavelet scalogram from the WPT.

source: A. McCormick 1998

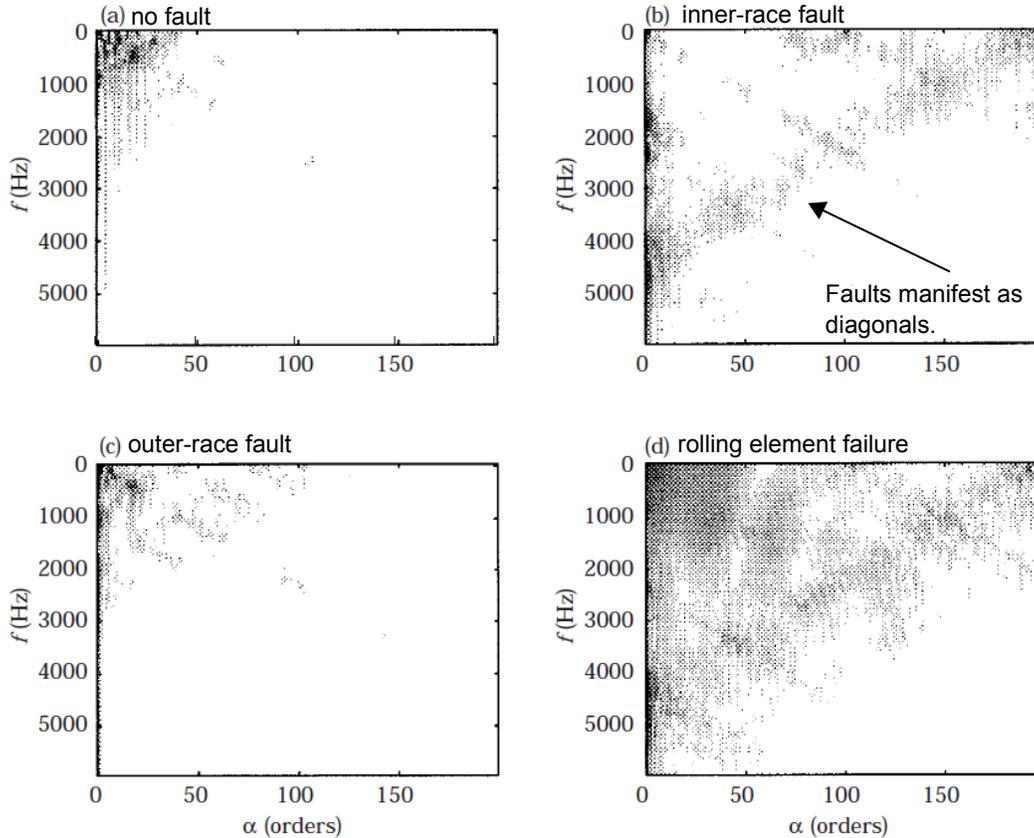


Figure 2.5: A.C. McCormick and A.K. Nandi [45] analysed bearing faults by second-order cyclostationary analysis. The plot is of the the spectral correlation density function, which is the DFT amplitude spectrum of the time-varying cyclic autocorrelation. The sample autocorrelation is the second-order cyclostationary component.

used as a fault indicator. The authors stated that TSA is in fact first-order cyclostationarity, and by performing TSA the randomness information in this second-order cyclostationarity is lost. Spectra correlation density is used to study the evolution of faults. Cyclostationary analysis has been applied more recently by E. Estupinan et al. in [42], 2007, with good success. This technique continues to be an area of moderate interest in machine fault detection. Application of this method typically results in either a one-dimensional fault indicator or a three-dimensional plot for human diagnostics.

2.2.3 Cepstrum Analysis

Cepstrum analysis is a topic of a relatively small amount of research activity in the field of machine fault detection. Cepstrum magnitude spectrum is the FT magnitude spectrum of the log-magnitude FT of the signal. This method is used to to detect the side-bands associated with time varying components or to quantify harmonics. Cepstrum was origi-

nally proposed for application of gear and bearing faults by R. Randall in 1975 [47]. The method has maintained moderate research interest over the years, and has recently been investigated by M. Badaoui, F. Guillet, and J. Danire [48] in 2004. The authors proposed using the cepstrum peaks at the critical frequencies as a fault indicator, and demonstrated the effectiveness in application to gear faults. In conclusion, the cepstrum proves a promising technique in monitoring gearbox conditions and provides a way for automated health monitoring, but does not indicate the type of problem with the machine. The lack of research papers in this field indicates that the method is likely not robust and does not work with many datasets.

2.3 Model-Based Methods

2.3.1 Autoregressive Model

Autoregressive (AR) models have been a growing trend in rotating machine vibration fault detection and have been shown to be effective in extracting gear faults with little a priori knowledge [2, 4, 5]. The AR system model with no input has structure

$$y_n = -a_1 y_{n-1} - a_2 y_{n-2} - \dots - a_N y_{n-N} + e_n,$$

where a_i are scalar model coefficients, N is the order of the AR model, e_n is white noise, and y_n is the signal being modelled. That is, the current sample is a linear combination of the N previous samples plus additive white noise. To solve for the scalar model coefficients, a_i , there are several approaches. For this thesis the Burg's lattice-based method [49] is applied due to the robustness of estimation. This method selects the parameters by minimizing the least-squares of both the forward and backward prediction errors.

Fault detection by AR model is approached through the following steps:

Step 1: Select AR model order N . This order is often selected by Akaike Information Criterion.

Step 2: Fit the AR model to the no-fault data by calculating a_i s by Burg's lattice-based method.

Step 3: Perform 1-step ahead prediction on the potentially faulty vibration data and calculate the prediction error, Fig. 2.6.

The prediction error consists of white noise, disturbances, and potentially some trended data from system dynamics changes. The impulse-like faults associated with gear cracks are expected to be more prominent in this residual. Several major drawbacks exist for this

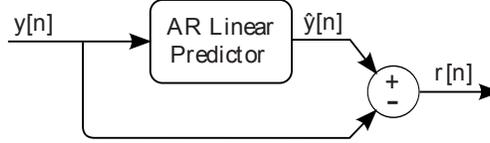


Figure 2.6: AR prediction residual method. $y[n]$, $\hat{y}[n]$, and $r[n]$ are the input signal, predicted signal, and prediction residual respectively.

proposed method. First of all and most importantly, the expected residual is white noise plus fault impulses; so the fault signal must be significantly larger in amplitude than the noise e_n for detection. And secondly, it requires knowledge of the system under no-fault conditions.

2.3.2 Observer Design

State observers, unknown input observers (or similarly disturbance observers) [50, 51], and fault-sensitive observers [44] have application in rotating machine vibration fault detection but no research has been applied to the investigation of these methods of rotating machine faults. The fault signal is expected to be more visible in this estimated disturbance, estimated input, or output residual signal. For the case of fault-sensitive observers, an observer can be designed to be noise tolerant while being sensitive to faults. The resulting prediction residual is expected to better emphasize the fault. Research should be dedicated to investigating the application of these methods towards rotating machine fault detection.

2.4 Filtering-Based Methods

Filtering-based methods have proven to be a very powerful method of fault detection in rotating machines. Presented here is a basic review of methods and literature regarding methods aiming to reduce noise, reduce non-fault signals, or enhance the visibility of the fault vibration. Several methods are presented in the following subsections including time-synchronous averaging (often just referred to as synchronous averaging), wavelet transform-based filtering, spectral Kurtosis-based filtering, and deconvolution.

2.4.1 Time-Synchronous Averaging

This technique is most commonly referred to as TSA in the gear and bearing fault detection field, but is sometimes referred to as synchronous signal averaging technique, synchronous averaging, or time domain averaging. As a denoising method, the TSA technique is very commonly applied [2, 5, 7, 6, 21]. To filter out asynchronous vibration and noise, speed

or tachometer information is combined with angular resampling to an integer number of samples per revolution of the gear or bearing critical period using an interpolation method, such as spline interpolation. This angular resampling is performed as a first step to deal with slight variances in the machine rotational speed, and to deal with the samples not being acquired at exactly the same angle positions. The method then averages the vibration over many revolutions of the gear or bearing critical period. This reduces asynchronous vibration components and zero-mean noise because the expected value of as the number of averaged revolutions gets very large approaches zero. A major drawback of TSA is the significant loss of information contained in the distribution or autocorrelation function, which contains important information related to fault detection [46]. Despite this loss of information, the method is widely applied as a preprocessing step and as a rough estimate approximately 40% of gear and bearing fault detection papers implement this method.

2.4.2 Filter Selected by Wavelet Transform

The wavelet transform is commonly applied as a filter to emphasize the fault signal. W.J. Staszewski and J.R. Tomlinson proposed the wavelet transform in detection of spur gear faults in 1994 [22]. One fault indicator shown to be effective by the authors is calculated as the Kurtosis of the envelope of the Morlet wavelet transform around a centre frequency of the fault critical frequency. In 2002, N.G. Nikolaou and I.A. Antoniadis proposed the application of the complex Morlet wavelet transform on fault detection of ball bearings [19]. The complex shifted Morlet wavelet transform was applied and the filtered signal was formed by selecting the maximum envelope of a frequency range at each time. The method was shown to be effective at extracting the time domain impulse-like structures of a bearing inner-race fault, Figure 2.7. X. Lou and K.A. Loparo proposed in 2004 [23] the application of the discrete wavelet transform decomposition followed by features of these resulting signals as input to an adaptive neural fuzzy inference system for fault identification. In Q. Hu et al. [25] 2007, the wavelet packet transform was used to decompose a vibration signal and generate features for bearing fault classification by support vector machine ensemble. In 2007, S. Abbasian et al. [26] implemented the discrete wavelet transform with Meyer wavelet and used a single decomposition signal as the output, hence just using the method as a basic filter design tool. Features of this filtered signal were used in fault classification of ball bearings with a support vector machine.

2.4.3 Filter Selected by Spectral Kurtosis

Spectral Kurtosis (SK) was proposed for fault detection on rotating machine vibration data by J. Antoni and R.B. Randall [18] in 2006. SK is a method in which a bandpass filter is

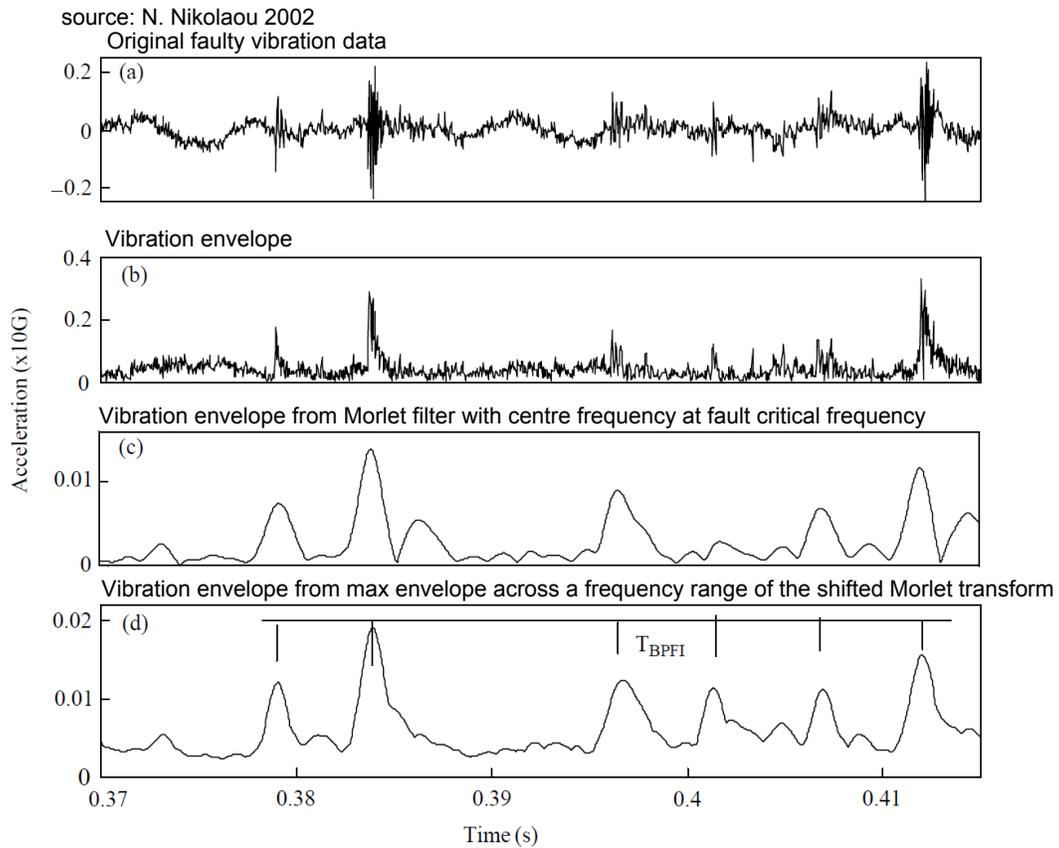


Figure 2.7: N.G. Nikolaou and I.A. Antoniadis [19] analysed inner-race bearing faults by shifted complex Morlet wavelet transform max-envelope technique. T_{BPFI} is the critical period of the bearing fault.

selected to maximize the Kurtosis of the resulting filtered signal. N. Sawalhi, R.B. Randall, and H. Endo [7] improved upon the method in 2007 by combining the method with MED. Results were analysed on the detection of ball bearing inner race fault vibration. The results were shown to be better than SK alone, but no investigation was performed for the implementation of the MED method alone. In 2008, F. Combet and L. Gelman presented a methodology involving TSA, windowed SK, enveloping, and smoothing [16]. Comparison of the performance on pitted and unpitted gear vibration was performed between the proposed method and ARMED. The results seem to indicate that both methods perform well, with the proposed method providing a slightly clearer distinction between the no-fault and fault cases. In general the SK method is a very strong method which has rightfully been receiving a lot of research attention within the last few years.

2.4.4 Minimum Entropy Deconvolution

MED was originally proposed for application on seismic recordings by R. Wiggins in 1978 [52] and recently applied to gear fault detection by H. Endo et. al. [5] in 2007. MED poses a deconvolution problem where a FIR filter is selected to minimize the entropy of the filtered signal.

Starting from a general linear time-invariant machine acceleration signal model:

$$x_n = (-a_1x_{n-1} - a_2x_{n-2} - \dots - a_Kx_{n-K}) + (b_1u_{n-1} + b_2u_{n-2} + \dots + b_Lu_{n-L}) \\ + (c_1d_{n-1} + c_2d_{n-2} + \dots + c_Md_{n-M}) + e_n$$

where x_n is the sampled acceleration signal, u_n is an unknown input sequence, d_n is the repeating impulse-like gear fault input sequence, and e_n is noise. a_k , b_k , and c_k are scalars representing the dependence of x_n on previous x , u , and d respectively. By taking the z -transform and solving for the system acceleration, we have:

$$X = \frac{B(z^{-1})}{A(z^{-1})}U + \frac{C(z^{-1})}{A(z^{-1})}D + \frac{1}{A(z^{-1})}E \\ A(z^{-1}) = 1 + a_1z^{-1} + \dots + a_Kz^{-K} \\ B(z^{-1}) = b_1z^{-1} + b_2z^{-2} + \dots + b_Lz^{-L} \\ C(z^{-1}) = c_1z^{-1} + c_2z^{-2} + \dots + c_Mz^{-M}$$

where X , U , and E are the z -transform of x , u , and e respectively.

Any stable transfer function $P(z^{-1})/Q(z^{-1})$ can be approximated as a FIR filter. Since these transfer functions are clearly stable or marginally stable in this case (otherwise the

machine would explode with infinitely growing acceleration), the resulting time domain approximation form in terms of convolution is

$$\vec{x} = \vec{h}_u * \vec{u} + \vec{h}_d * \vec{d} + \vec{h}_e * \vec{e},$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}, \quad \vec{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix},$$

where \vec{h} 's are the FIR filter approximations to their respective transfer functions under stable assumption, and N is the number of measured samples of vibration signal x_n . The deconvolution problem aims to reconstruct the fault signal \vec{d} by applying FIR filter \vec{f} with L samples to measured machine acceleration \vec{x} :

$$\vec{y} = \vec{f} * \vec{x} = \vec{f} * (\vec{h}_u * \vec{u}) + \vec{f} * (\vec{h}_d * \vec{d}) + \vec{f} * (\vec{h}_e * \vec{e}),$$

$$\vec{f} = [f_1 \ f_2 \ \dots \ f_L]^T$$

It is desired that the resulting filtered signal \vec{y} approximates fault signal \vec{d} and this is approached by selecting filter \vec{f} to minimize the noise effect $\vec{f} * (\vec{h}_e * \vec{e}) \rightarrow \vec{0}$, while closely cancelling the system $\vec{f} * (\vec{h}_u * \vec{u}) \rightarrow \vec{0}$, and extracting the fault signal $\vec{f} * (\vec{h}_d * \vec{d}) \approx \vec{d}$. Selection of the filter \vec{f} given only acceleration \vec{x} measurements may seem to be a difficult problem, but the fault signal \vec{d} is expected to be impulse-like (a signal of very high Kurtosis) while competing signals \vec{u} and \vec{e} are expected to be of much lower Kurtosis. As a result of this signal Kurtosis difference between the signals, the filter can be selected to reach a maximum in Kurtosis. To achieve this, R. Wiggins [52] proposed maximizing of a norm function called the Varimax Norm, which in the case of one-dimensional MED is equivalent to maximizing Kurtosis with assumed zero-mean:

$$\text{Kurtosis} = \frac{\frac{1}{n} \sum_{n=1}^N (y_n - \mu_y)^4}{\left(\frac{1}{n} \sum_{n=1}^N (y_n - \mu_y)^2\right)^2} - 3$$

$$\max_{\vec{f}} \text{Kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^N (y_n - \mu_y)^4}{\left(\sum_{n=1}^N (y_n - \mu_y)^2\right)^2}$$

Describing the deconvolution method in terms of Kurtosis is chosen because it is commonly used to quantify the impulse-like fault level of a vibration signal [2, 5, 18]. Assuming y_n is zero-mean, $\mu_y = 0$:

$$\max_{\vec{f}} \text{Kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^N y_n^4}{\left(\sum_{n=1}^N y_n^2\right)^2} \quad (2.1)$$

By taking the derivatives of Eq. 2.1 with respect to filter coefficients \vec{f} and solving it equal to zero, an iteratively converging local-maximum solution can be derived as:

$$\vec{f} = \frac{\sum_{n=1}^N y_n^2}{\sum_{n=1}^N y_n^4} (X_0 X_0^T)^{-1} X_0 [y_1^3 y_2^3 \dots y_N^3]^T \quad (2.2)$$

$$X_0 = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_N \\ 0 & x_1 & x_2 & \dots & x_{N-1} \\ 0 & 0 & x_1 & \dots & x_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x_{N-L+1} \end{bmatrix}_{L \text{ by } N}$$

where \vec{f} is iteratively selected. The iterative procedure is implemented with MATLAB and available in the External Resources Section, Appendix A. The general procedure is as follows:

Step 1: Assume initial filter as a centred impulse, $\vec{f} = [0 \ 0 \ \dots \ 1 \ \dots \ 0 \ 0]^T$.

Step 2: Calculate X_0 and $(X_0 X_0^T)^{-1}$ from input signal \vec{x} .

Step 3: Calculate \vec{y} as $\vec{y} = X_0^T \vec{f}$.

Step 4: Determine new filter coefficients by solving for \vec{f} in Eq. 2.2.

Step 5: Repeat from Step 3 for a specified number of iterations or until the change in Kurtosis between iterations is below a specified small value.

Step 6: The final deconvolved signal is calculated as $\vec{y} = X_0^T \vec{f}$.

H. Endo et. al. [5] proposed the fault detection method ARMED, which applied the AR fault detection method followed by MED on the resulting prediction residual. The expected results for MED is approximately the low entropy fault signal \vec{d} , unlike the AR method which expects the fault signal plus white noise. As a result, the ARMED resulted in improved performance over the traditional AR method. However, the author did not indicate a reason for inclusion of the AR method as a preprocessing stage and it should not be a required step; therefore both direct MED and ARMED techniques are applied for comparison purposes. Also important to note is that in the case of direct MED fault extraction, no a priori knowledge of no-fault machine vibrations is required for fault deconvolution.

2.5 Fault Indication

The goal of fault indication is to detect faults in the vibration data and depending on the method, indicate the type of fault as well. In general, most fault indication methods can be grouped into:

- Artificial neural network/fuzzy logic [27, 28, 23, 29, 30, 31], support vector machine [28, 25, 26, 32], or hidden markov model [33]. These approaches typically indicate the fault type and fault level but require extensive data knowledge of the machine under fault conditions and typically requires detailed training for each individual machine. Performance can be unreliable if a fault manifests in a different manner than expected, resulting in potentially missing faults. In general these approaches are typically not very applicable to industry machine condition monitoring or fault diagnosis.
- Human factor [17, 35, 13]. When authors provide no clear fault indication variable it is expected that fault detection and diagnosis is performed by human analysis of the results (eg. human analysis of a wavelet scalogram). This indication method is impractical in machine condition monitoring and only applicable towards diagnosis of faults.
- Threshold [22, 18, 5, 10]. Methods which perform fault indication by comparison of an indicator (typically Kurtosis or RMS of a filtered signal) are assumed to be implemented online by a threshold alarm. These approaches are very practical for industrial condition-monitoring applications, as the alarm threshold can be easily adjusted by on-site engineers. These approaches often do not indicate the type of fault, just that a fault is present.

These various fault indication methods have various pros and cons, and as a result have different applications.

3

Frequency Estimator-Based Fault Detection¹

Recently in vibration-based fault detection there has been a trend in the field towards detecting gear and bearing faults through vibration measurements using autoregressive (AR) matched model prediction error signals [2, 4, 5, 6, 7]. AR models are linear models matched to the vibration signal under no-fault conditions, and the model is applied to predict the next vibration sample and compared to the actually measured sample. This error in prediction, the residual, is well-suited to extract the impulse-like features associated with gear crack vibration. Major drawbacks of the AR-based approach is the non-linear nature of vibration data. Figure 3.1 illustrates a time-frequency plot of the vibration data by complex Morlet continuous wavelet packet transform. From this figure it is clear that the AR assumption of time-invariance is clearly incorrect, and it appears that the vibration consists of multiple time-variant sinusoids combined with noise and disturbance. Additionally, vibration is known to include non-linear and chaotic characteristics, and slight changes in motor speed change the model.

The nature of vibration is sinusoidal and robustness can be improved through prediction based on a sinusoid restriction [53]. Recently, adaptive system theory has been applied towards real-time estimation of multiple sinusoidal components of signals [8]. This frequency estimator (FE) is in the form of an adaptive state-space model and forms a non-linear system. This model is not only able to predict future samples based on a sum of

¹A version of this chapter has been published. G. L. McDonald and Q. Zhao, "Model-based adaptive frequency estimator for gear crack fault detection," in American Control Conference 2011, San Francisco, June 2011.

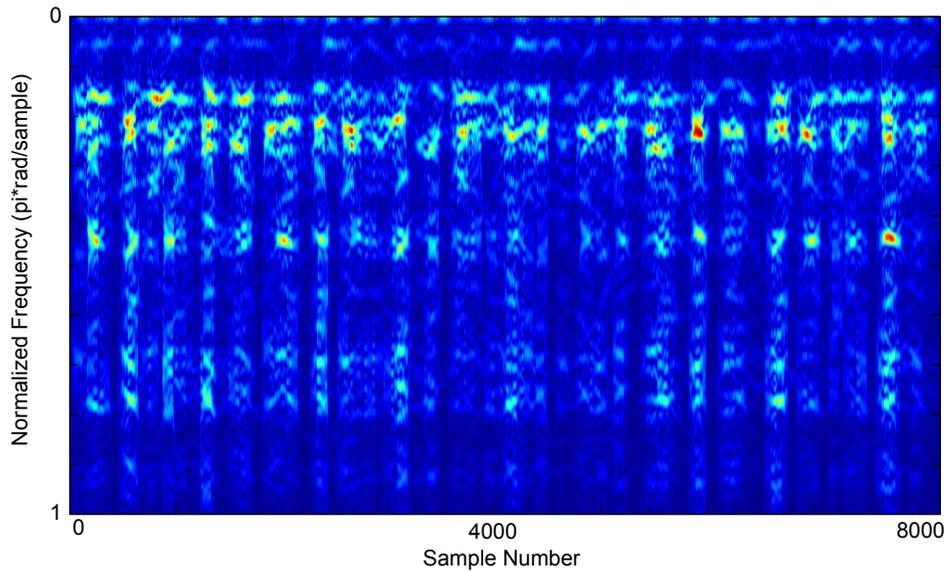


Figure 3.1: Continuous wavelet packet transform of vibration under gear chip using the complex Morlet mother wavelet.

sinusoids model, but also adapts the sinusoidal model during operation. This results in the ability to predict samples on-line even with time-variant sinusoid components and no a priori knowledge requirements such as the data fitting required in the AR approach. Despite the relatively complex operation of the FE method, this chapter demonstrates that it is of comparable calculations-per-sample to a low order AR algorithm. Clearly the FE model is well-suited to the nature of vibration signals and poses application in improving the field of gear-crack fault detection and other fields.

This section investigates the application of the adaptive system concept of FE for the detection of gear-crack faults in rotating machinery. Firstly, a discrete-time (DT) version of the continuous-time N-component FE proposed by M. Hou [8] is derived. Next a novel amplitude-invariant adaptive identifier is proposed. Finally, general model parameters are selected for a 2 and 3 component FE using a non-linear least-squares (LSQ) approach. For performance evaluation, the standard AR predictive model is compared to the FE predictive model in complexity, simulation performance, and experimental performance on a controlled gear-crack.

3.1 N-Component Frequency Estimator (FE)

Based on the N-component FE model by M. Hou [8] for continuous-time, a low complexity discrete-time sample predictor is first derived. The derivation starts from the zero-order

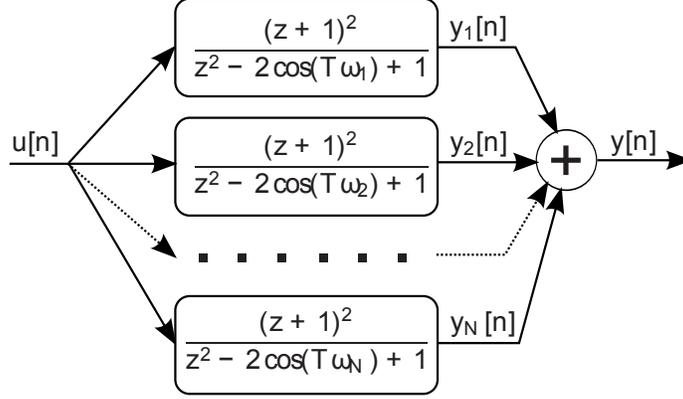


Figure 3.2: N-component DT system model

hold discretized model of the sum of sinusoid system for N-components, Figure 3.2.

Forming the transfer function for each sinusoidal component,

$$Y_i(z) = \frac{N(z)}{D_i(z)} U(z) = \frac{(z+1)^2}{z^2 - 2 \cos(T\omega_i)z + 1} U(z),$$

the complete transfer function follows as

$$Y(z) = U(z) \sum_{i=1}^N Y_i(z) = \frac{P(z)}{Q(z)} U(z),$$

where $P(z)$ and $Q(z)$ refer to the numerator and denominator polynomial of the complete transfer function,

$$\begin{aligned} Q(z) &= \prod_{i=1}^N (z^2 - 2 \cos(T\omega_i)z + 1) \\ &= z^{2N} + \theta_{N-1}z^{2N-1} + \theta_{N-2}z^{2N-2} + \dots \\ &\quad + \theta_0z^N + \dots + \theta_{N-2}z^2 + \theta_{N-1}z + 1, \end{aligned} \tag{3.1}$$

and θ_i 's refer to the resulting expanded polynomial coefficients. By introducing a Hurwitz polynomial,

$$\alpha(z) = z^{2N} + \alpha_{2N-1}z^{2N-1} + \dots + \alpha_1z + \alpha_0$$

we rewrite the transfer function as

$$Y(z) = \frac{P(z)}{\alpha(z) + (Q(z) - \alpha(z))} U(z).$$

Rearranging we have,

$$Y(z) = \frac{\alpha(z) - Q(z)}{\alpha(z)} Y(z) + \frac{P(z)}{\alpha(z)} U(z)$$

and taking $U(z) = 0$ we have

$$Y(z) = \frac{\alpha(z) - Q(z)}{\alpha(z)} Y(z)$$

with estimator state-space realization of

$$\begin{aligned}
x[n+1] &= Ax[n] + By[n] \\
\hat{y}[n] &= \hat{C}[n]x[n] \\
A &= \begin{bmatrix} -\alpha_{2N-1} & -\alpha_{2N-2} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \\
B &= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \hat{C}[n] = \alpha^T - \hat{\theta}[n]^T V \\
\alpha &= \begin{bmatrix} \alpha_{2N-1} \\ \alpha_{2N-2} \\ \vdots \\ \alpha_1 \\ \alpha_0 - 1 \end{bmatrix}, \hat{\theta} = \begin{bmatrix} \hat{\theta}_{N-1} \\ \hat{\theta}_{N-2} \\ \vdots \\ \hat{\theta}_0 \end{bmatrix} \\
V &= \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & 1 & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}
\end{aligned}$$

where the system output, $\hat{y}[n]$, is the signal estimation and $\hat{\theta}$ corresponds to the estimated frequency components $\hat{\omega}$ by Equation 3.1. The parameter α is tunable and controls the convergence behaviour of the system, and the selection of these parameters are discussed in Section 3.3. The $\hat{\theta}$ estimates are adapted according to the residual as described in the following section.

3.2 Adaptive Identifiers

Updating the frequency estimates, $\hat{\omega}$, is achieved through updating the related Hurwitz parameters, $\hat{\theta}$. The $\hat{\theta}$ parameter is updated according to the prediction residual, $r[n] = y[n] - \hat{y}[n]$, such that the system converges $\hat{\theta} \rightarrow \theta$. A simple adaptive identifier can be formed by simply adjusting $\hat{\theta}$ such that the residual decreases:

$$\hat{\theta}[n] = \hat{\theta}[n-1] - \gamma V x[n-1] r[n-1]$$

where γ is a tunable scalar parameter controlling the rate at which $\hat{\theta}$ is adapted. A smaller value will result in a longer convergence time but is more robust to noise, while a larger

value improves the convergence time but reduces noise tolerance. One drawback of this possible adaptive identifier is the amplitude sensitivity. By multiplying the input signal by a scalar, $y[n] = ks[n]$, we have

$$\begin{aligned}x[n+1] &= Ax[n] + Bks[n] \\ \frac{x[n+1]}{k} &= \frac{Ax[n]}{k} + Bs[n] \\ k\hat{s}[n] &= \hat{C}x[n]\end{aligned}$$

then redefining $\bar{x}[n] = x[n]/k$, we have the system

$$\begin{aligned}\bar{x}[n+1] &= A\bar{x}[n] + Bs[n] \\ \hat{s}[n] &= (\alpha^T - \hat{\theta}^T V)\bar{x}[n] \\ \hat{\theta}[n] &= \hat{\theta}[n-1] - \gamma V k \bar{x}[n-1](ks[n] - k\hat{s}[n])\end{aligned}$$

clearly resulting in a k^2 factor increase in the changes to $\hat{\theta}[n]$ each iteration and in turn resulting in significantly larger changes to the estimated frequencies $\hat{\omega}$. Ideally, the change to $\hat{\theta}[n]$ should be independent of the signal amplitude such that the parameter γ does not need to be tuned for each input signal. By introducing a state-normalization factor to the adaptive identifier, the amplitude sensitivity issue can be addressed,

$$\hat{\theta}[n] = \hat{\theta}[n-1] - \frac{\gamma V x[n-1](y[n-1] - \hat{y}[n-1])}{\|x[n-1]\|^2}$$

and following a similar procedure as above with input signal $y[n] = ks[n]$,

$$\begin{aligned}\hat{\theta}[n] &= \hat{\theta}[n-1] - \frac{\gamma V k \bar{x}[n-1](ks[n-1] - k\hat{s}[n-1])}{\|k\bar{x}[n-1]\|^2} \\ &= \hat{\theta}[n-1] - \frac{\gamma V \bar{x}[n-1](s[n-1] - \hat{s}[n-1])}{\|\bar{x}[n-1]\|^2}\end{aligned}$$

where the change to $\hat{\theta}[n]$ is now independent of the input signal amplitude. This proposed change introduces a singularity at $\|x[n]\| = 0$, and the adaptive identifier is slightly adjusted such that

$$\begin{aligned}\hat{\theta}[n] &= \hat{\theta}[n-1] - \Delta\hat{\theta}[n-1] \\ \Delta\hat{\theta}[n] &= \begin{cases} \frac{\gamma V x[n](y[n] - \hat{y}[n])}{\|x[n]\|^2} & \text{if } \|x[n]\|^2 > \epsilon \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

where ϵ is a small positive scalar. Although this simplifies the model by being amplitude invariant, the selection of α and γ is still a difficult process due to the non-linear characteristics of the system. For the rest of the chapter only this FE State-Normalized adaptive identifier is studied, as the results generalize to any input amplitude. The next section presents a least-squares minimization approach to selecting these parameters.

3.3 Least-Squares (LSQ) Parameter Selection

Parameter selection in a non-linear system is a difficult process. This problem is approached as a minimization problem where the input parameters, γ and α , are adjusted to minimize the error in the estimated $\hat{\theta}[n]$. Minimizing the estimated frequency error, $\hat{\omega}[n]$, is not considered due to the complexity involved with calculating $\hat{\omega}[n]$ from $\hat{\theta}[n]$. LSQ minimization is performed as

$$\min_{\gamma, \alpha} \|f_{\theta}\|_2^2$$

where f_{θ} is defined as the error in θ estimates in K -sample trajectories for L randomly generated input signals. Each input signal includes additive white Gaussian noise and N frequency components with uniformly-distributed frequency and initial phase. The L input signals are known and their corresponding correct θ is calculated by Equation 3.1. This forms the cost matrix,

$$f_{\theta} = [e_{\theta}[0] \quad e_{\theta}[1] \quad \cdots \quad e_{\theta}[K-1]]$$

$$e_{\theta}[n] = \begin{bmatrix} \theta_{y_0} - \hat{\theta}_{y_0}[n] \\ \theta_{y_1} - \hat{\theta}_{y_1}[n] \\ \vdots \\ \theta_{y_{L-1}} - \hat{\theta}_{y_{L-1}}[n] \end{bmatrix}$$

used for the LSQ minimization. Minimizing this least-squares non-linear minimization problem is approached using the Levenberg-Marquardt with line-searching algorithm [54, 55]. LSQ initial parameters of $\alpha = [0 \quad 0 \quad \cdots \quad 0 \quad -1]^T$ and $\gamma = 0$ are used for the adaption, and simulation initial conditions of $x[0] = 0$, and $\hat{\theta}[0] = 0$ are used. Table 3.1 indicates the final FE parameter values after 10 iterations with $K = 2000$, $M = 500$, and additive Gaussian white noise of zero mean and 0.1 standard deviation. These values for K and M were chosen experimentally through trying quite a few values and comparing the resulting FE models on additional simulation and vibration data.

Table 3.1: Final values for FE parameters γ and α by LSQ minimization

N	γ	α
2	0.522	$[-0.163 \quad 0.255 \quad -0.0840 \quad -0.911]^T$
3	1.06	$[0.227 \quad 0.834 \quad 0.135 \quad 0.357 \quad -0.0607 \quad -1.09]^T$

The convergence versus LSQ iteration is plotted in Figure 3.3 for the two and three-component FE models. Figure 3.4 plots the convergence behaviour of $\hat{\theta}$ for the first four input signals of the three-component FE after iterative parameter selection, while Figure 3.5 plots the convergence of the frequency estimates of the two-component FE. It can be seen that the estimates typically converge correctly in a relatively short number of samples.

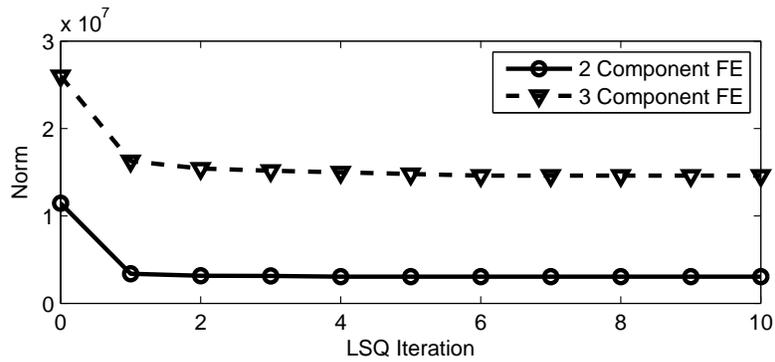


Figure 3.3: LSQ minimization norm versus iteration for the parameter selection of the two and three-component FE.

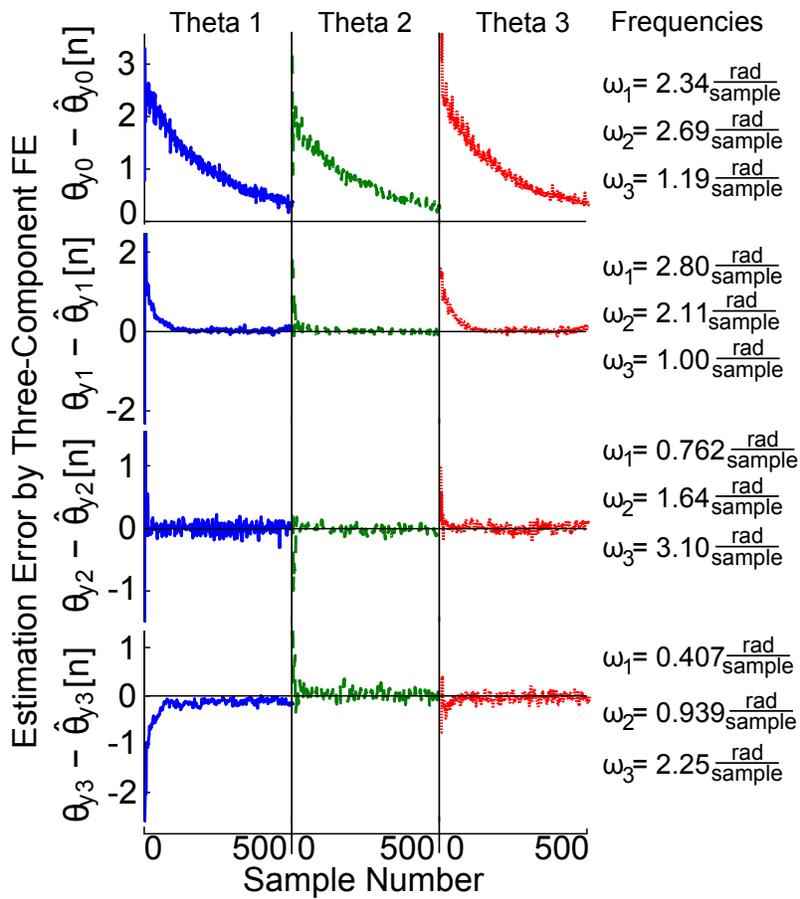


Figure 3.4: Convergence of θ estimates for the three-component FE for the first four input signals each composed of three sinusoidal components at indicated frequencies.

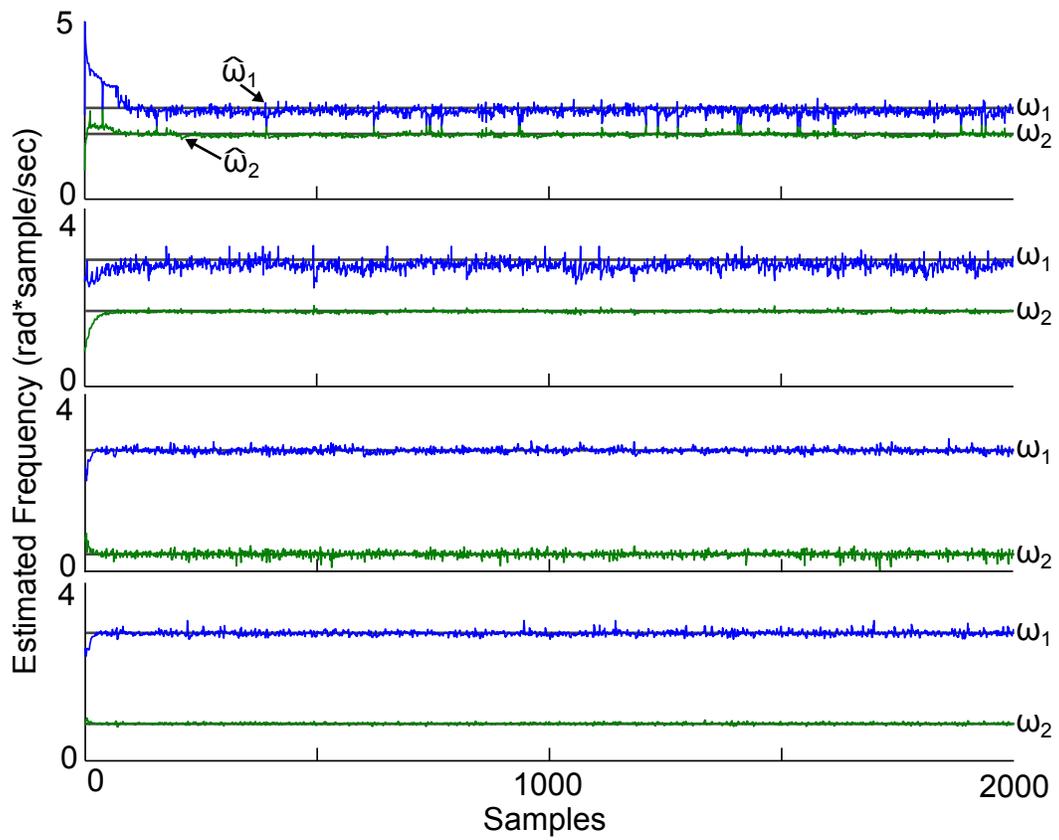


Figure 3.5: Convergence of the frequency estimates for the first four input signals with the two-component FE model.

3.4 Computational Requirements

The computational requirements of both the M-order AR model and the N-component FE is important for implementation. For the FE, the back-calculation of the $\hat{\omega}$ from $\hat{\theta}[n]$ is not included because this information is not required in the context of the proposed fault detection. Table 3.2 below indicates the computational requirements of the M-order AR model and N-component FE using both adaptive identifiers. All models were implemented in controllable canonical form and computations per sample were calculated.

Table 3.2: Floating point calculations required for each sample of predicted output as a function of model order, M, or number of frequency components, N.

Method	Additions per sample	Multiplications per sample
AR	$2M - 1$	$2M$
FE Classical	$5N - 1$	$3N + 2$
FE State-Normalized	$6N - 2$	$5N + 3$

For performance comparison purposes, an AR model of order $M = 9$ is included in the comparison as it is of similar complexity to the 3-component FE model.

3.5 Simulation Results

For validation, a simulated gear-crack vibration signal is generated with varied levels of fault. The vibration signal is formed as harmonics of the motor vibration of 60 Hz plus fault modelled as time-localized decaying exponential enveloped sinusoid vibrations repeating at the gear rotational period. The motor harmonics of 60 Hz, 120 Hz, ..., 360 Hz have amplitudes of 0.5, 0.2, 0.1, 0.02, 0.03, and 0.01 respectively and initial phase of 1.04, 1.82, 6.60, 2.83, 4.08, and 0.00 rad respectively. Additive white Gaussian noise with zero mean and 0.001 standard deviation is included. The disturbance gear crack vibration signal is added to the signal at a period of 83 milliseconds (12 Hz) and is composed of a decaying exponential envelope with time constant 2e-3 and carrier frequency of 360 Hz. Figure 3.6 plots the simulated vibration signal for a gear-crack peak disturbance of 0.362 and sampling rate of 10 kHz.

Figure 3.7 presents the residual results when model-matching the AR models to the no-fault vibration data, and applying the FE and resulting AR models to the fault vibration signal in Figure 3.6.

Kurtosis is defined as the fourth standardized moment,

$$k = \frac{E(x - \mu)^4}{\sigma^4}$$

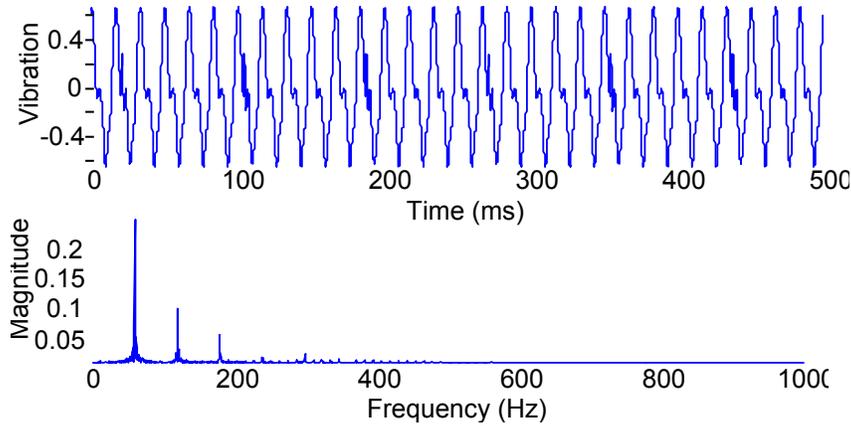


Figure 3.6: Simulated gear crack vibration signal and spectrum.

where $E()$ denotes the statistical expectation, x the signal, μ the signal mean, and σ the signal standard deviation. Kurtosis is a measure of how outlier-prone a signal is, and is commonly used as a fault indicator in gear-related faults due to the impulse-like vibration manifestation. The kurtosis values are analyzed to evaluate the FE and AR methods' ability to detect the gear crack fault. Comparing the fault peak disturbance versus the kurtosis, Figure 3.8, the 2-component FE performs the best for higher fault amplitudes while the 9th order AR model performs the best at small amplitudes. However, these results vary drastically according to the composition of the vibration signal and should not be interpreted as a general performance trend.

3.6 Experimental Design and Results

For validation and comparison, vibration data collected from a controlled gear tooth crack experiment under varying degrees of crack severity is analyzed. The machine configuration, Figure 3.9, is composed of a motor, gearbox, and brake.

A tooth crack is introduced in Gear 3 at varying levels using a cutting knife of width 0.4 mm controlling the depth and length of the cut. The gears for fault levels of 25%, 50%, 75%, and 100% can be seen in Figure 3.10.

The experiment consists of recording of the vibration along the axis of the impact force for three separate measurements during three separate machine start-ups under each operating condition. The motor operating conditions are varied between full-load, half-load, and no-load, with speeds ranging from 800 to 3000 rpm at 200 rpm intervals. In total, this results in 540 vibration measurements with approximately 8000 samples each. The sampling frequency is dependent on the rotation speed of the motor. Figure 3.11 illustrates the aver-

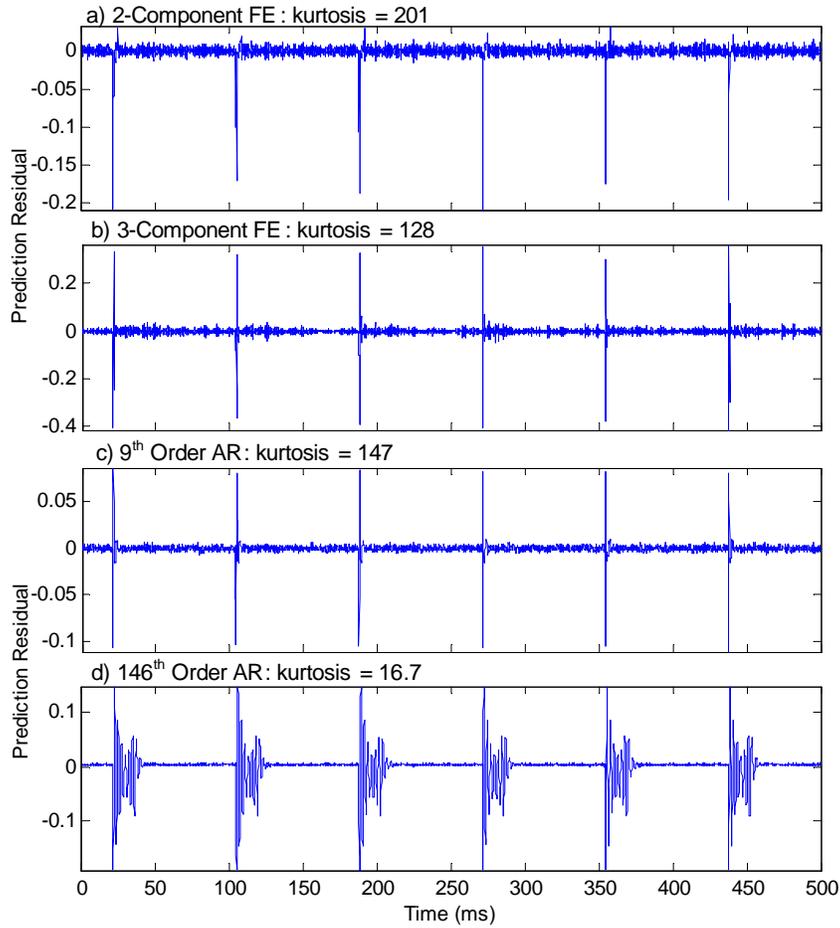


Figure 3.7: Prediction residual and corresponding kurtosis using the a) 2-component FE, b) 3-component FE, c) 9th order AR, and d) 146th order AR model.

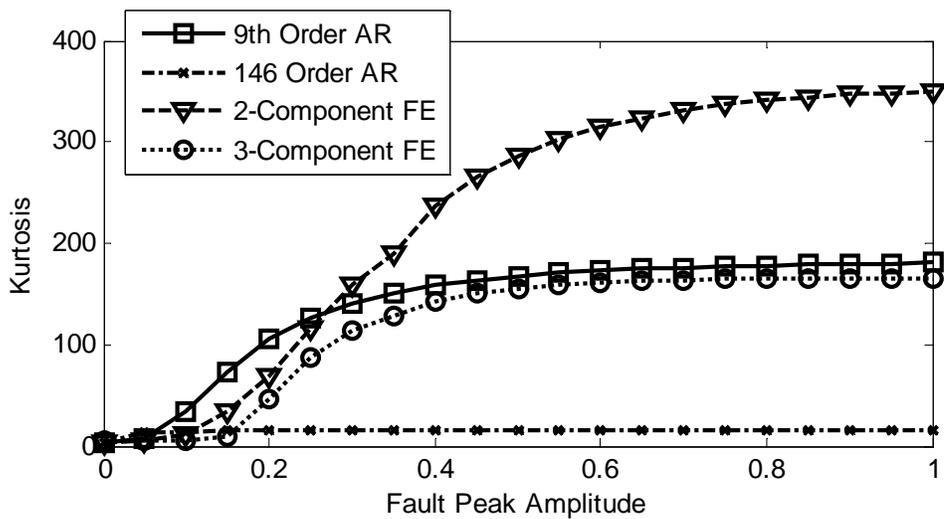


Figure 3.8: Kurtosis of model prediction residual plotted versus fault peak disturbance.

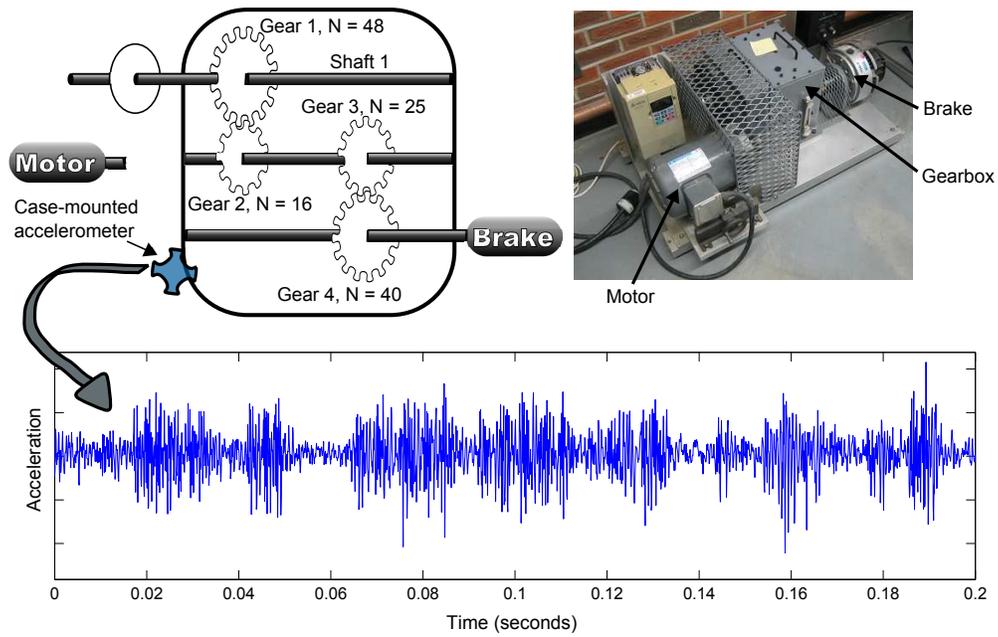


Figure 3.9: Experimental machine configuration.

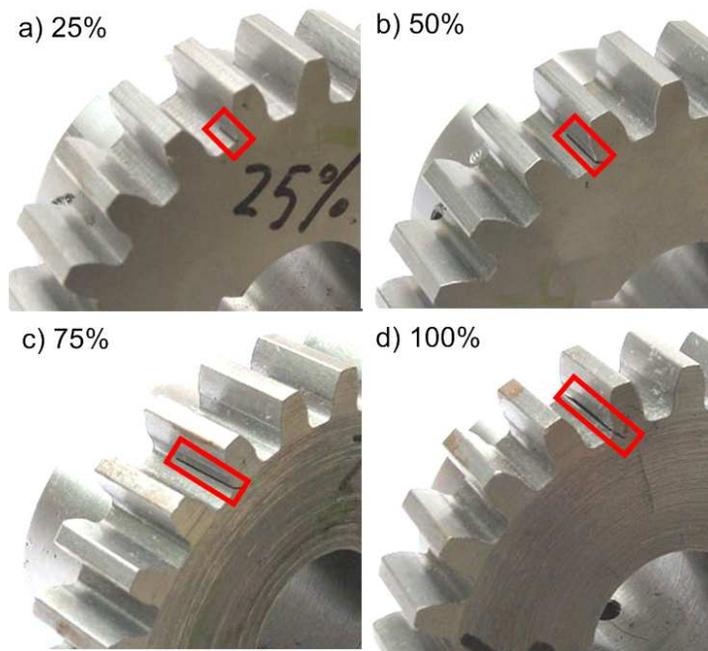


Figure 3.10: Experimental gears with tooth cracks at levels of a) 25%, b) 50%, c) 75%, and d) 100% [56].

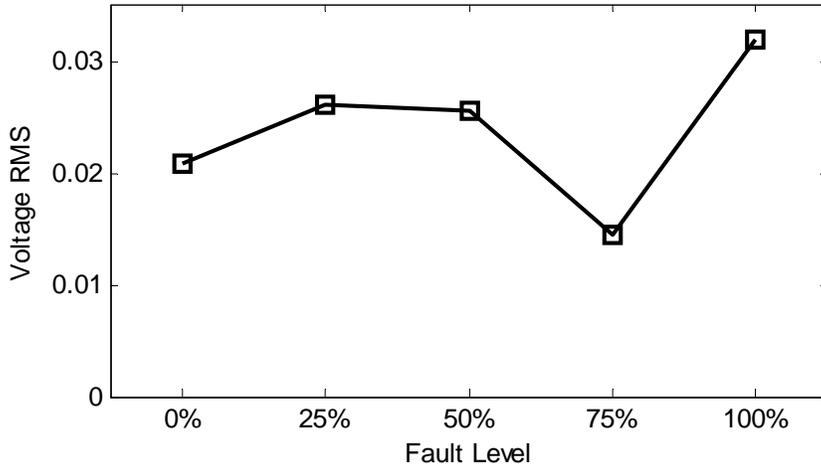


Figure 3.11: RMS of measured acceleration according to fault level.

aged rms vibration of the machine versus fault level. It can be seen that the vibration trend does not increase proportionally to the fault level; this is likely as a result of properties of the five different gears used or slight changes to the machine as the gear swapping is performed.

For the 540 datasets, the general data processing procedure is as follows:

1. Subtract the signal mean from each dataset.
2. Fit the AIC selected order and 9th order AR models to a dataset at no-fault condition at each rpm and load condition.
3. For each dataset, both AR models are used to predict the samples using the dataset's corresponding no-fault fit model. For example, a dataset measured at 3000 rpm, full-load, 25% fault level would use the AR models fit to the datasets at 3000 rpm, full-load, 0% fault level.
4. For each dataset, both FE models are simulated and the prediction residual calculated.
5. Calculate the kurtosis of each prediction residual signal while ignoring the first 200 samples to allow for the AR and FE models to converge.

From this procedure, it is clear that one big advantage of the FE-based method is the fact that it is a general model requiring no training to different datasets, different machines, different motor speeds, or even different applications. As a result the FE model will perform better under varied machine conditions. Figure 3.12 and 3.13 presents the prediction error for all four methods with the machine at 3000 rpm, full load, and 100% fault level.

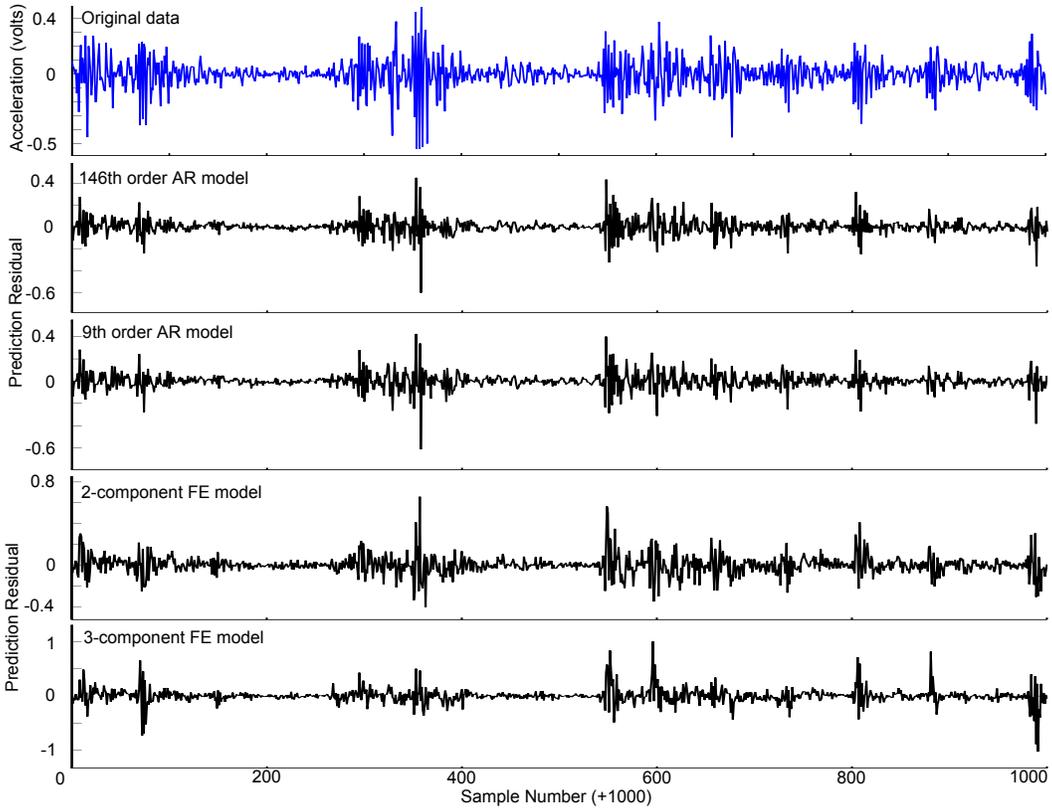


Figure 3.12: Prediction error by method for the experimental setup at 3000 rpm, full load, and 100% fault level.

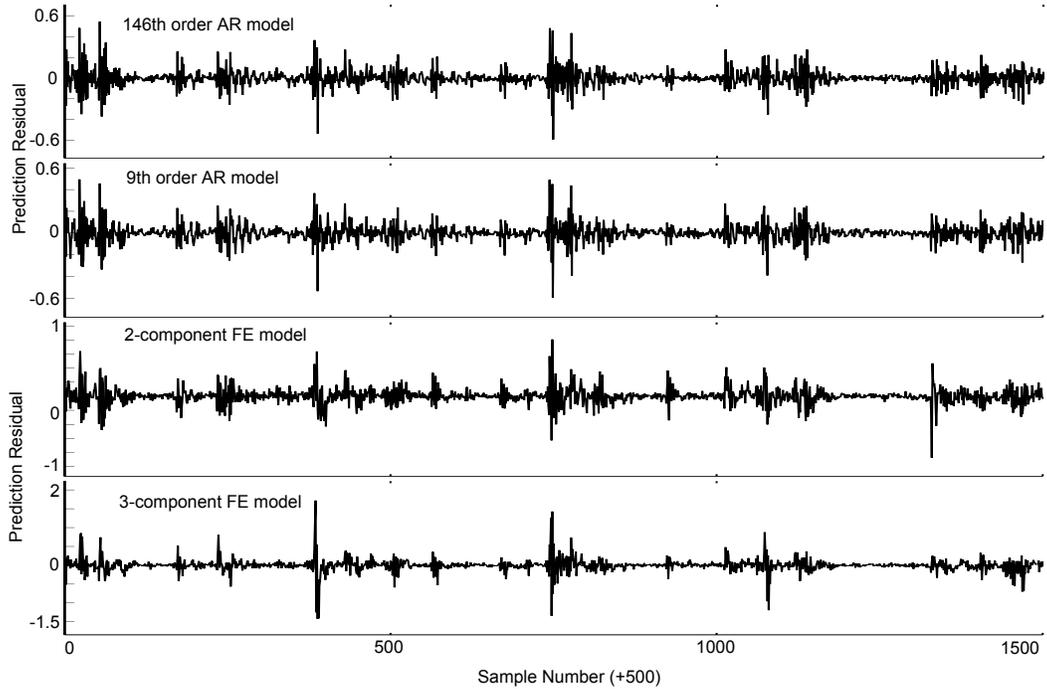


Figure 3.13: Prediction error by method for the experimental setup at 3000 rpm, full load, and 100% fault level.

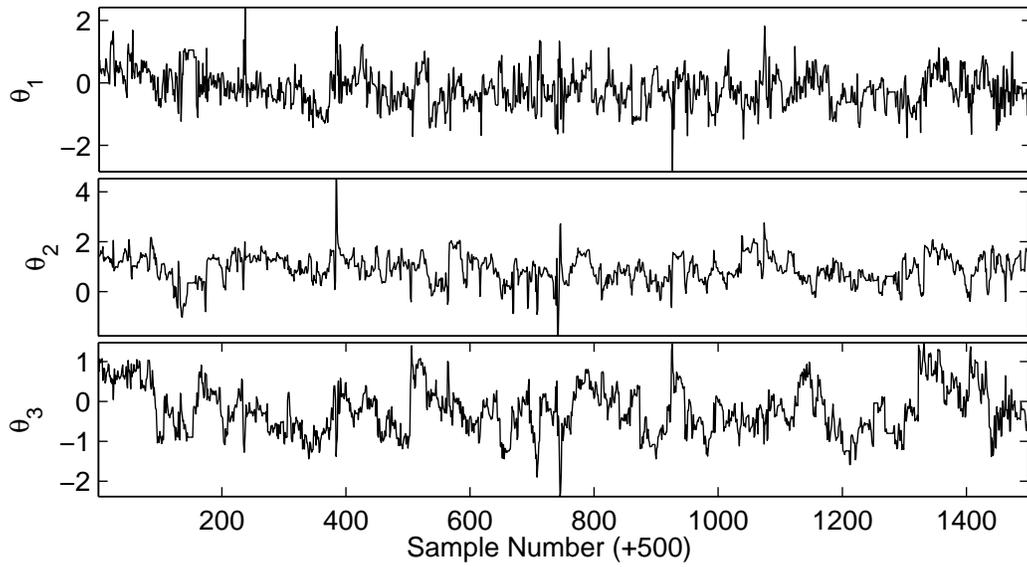


Figure 3.14: $\hat{\theta}$ versus sample for the three-component frequency estimator for the machine at 3000 rpm, full load, and 100% fault level.

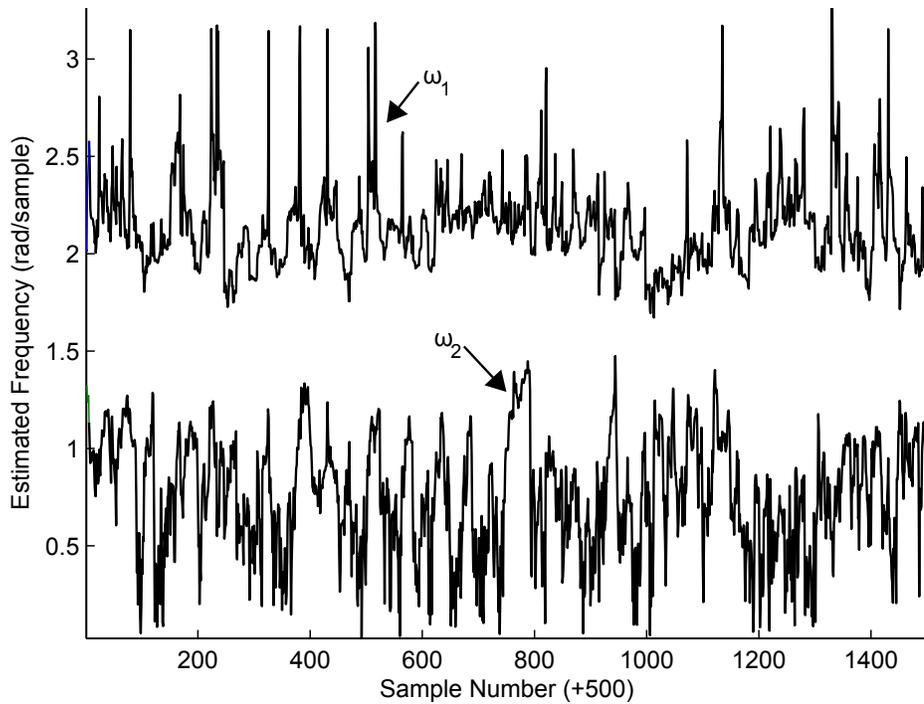


Figure 3.15: $\hat{\omega}$ versus sample for the two-component frequency estimator at 3000 rpm, full load, and 100% fault level.

Figure 3.14 and 3.15 illustrates the estimated θ and estimated frequency ω for the two and three-component FE respectively on the fault data. The frequency estimates for the three-component FE is not shown due to the complexity in calculated ω 's from θ 's. From the two-component frequency estimates, it can be seen that the FE method is modelling the system as a frequency-varying low frequency component plus high frequency component.

Figure 3.16 presents the kurtosis and Figure 3.17 presents the no-fault normalized kurtosis results for the four predictive models under full-load, half-load, and no-load. It can be seen that the 2 and 3 component FE provide a better correlation between increasing fault level and higher kurtosis values. It is clear that the FE results in a better gear fault indicator over both AR models.

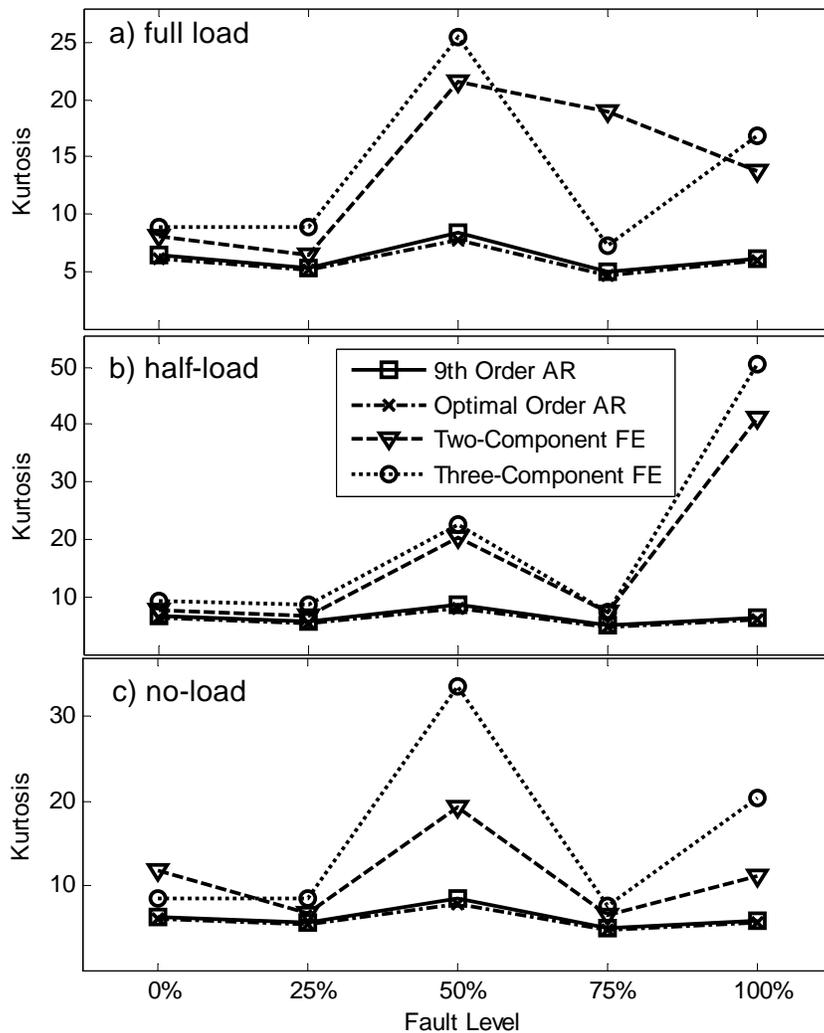


Figure 3.16: Kurtosis of the signals averaged across all rpm settings under a) full load, b) half-load, and c) no load.

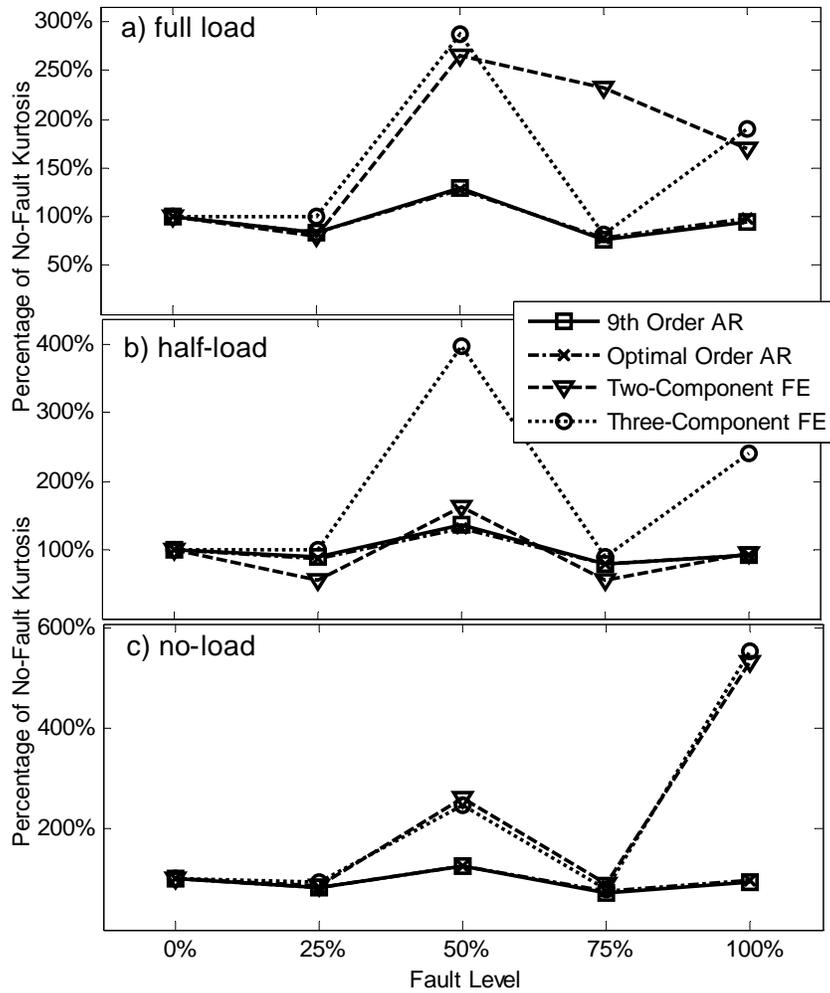


Figure 3.17: Kurtosis of the signals averaged across all rpm settings and plotted as a percentage of no-load fault kurtosis under a) full load, b) half-load, and c) no load.

3.7 Conclusion

In conclusion, the FE-based method not only detects gear crack faults better than the AR model based approach, but it requires no data fitting and is of similar computational requirements to a low order AR model. The experimental data indicates that the FE-based method outperforms the AR-based method in gear crack detection, while the simulation data suggests that the FE model only performs better at a higher fault level. Neither of these results are conclusive, but the FE model results proves a promising no a priori knowledge alternative to the AR model. Further work should investigate the FE-based method on additional experimental setups, and additional simulation models. Better parameter selection, convergence analysis, and higher order frequency estimators can be investigated for the FE model approach.

4

Maximum Correlated Kurtosis Deconvolution Fault Detection¹

Minimum Entropy Deconvolution (MED), originally proposed by R. Wiggins for application on seismic recordings in 1978 [52], iteratively selects a finite impulse response (FIR) filter to maximize the Kurtosis of the filtered signal and has had widespread applications across many fields. Unlike the AR method, the MED technique aims to extract the fault impulses while minimizing the noise and therefore results in clear detection results even under high noise. H. Endo and R. Randall [5] in 2007 proposed applying the AR method followed by MED, forming the method called ARMED and demonstrated the method to be very effective in detecting gear spalls and tooth cracks. N. Sawalhi et al. [7] demonstrated the effectiveness of the ARMED process in detecting faults in ball bearing elements. A limitation of the ARMED method is the preference of the MED algorithm to deconvolve only a single impulse or a selection of impulses, as opposed to the desired periodic impulses repeating at the period of the fault.

Inspired by the MED deconvolution technique, this chapter proposes an improved novel deconvolution norm, Correlated Kurtosis (CK), which takes advantage of the periodicity of the faults and requires no AR model stage prior to deconvolution. The deconvolution technique, Maximum Correlated Kurtosis Deconvolution (MCKD), is proposed to select a FIR filter to maximize the CK of the resulting signal which emphasizes high kurtosis while encouraging periodicity about a known period. An iteratively convergent solution to the

¹A version of this chapter has been submitted for publication to the journal Mechanical Systems and Signal Processing.

deconvolution is derived for first and M-shift MCKD, and the results are compared using simulation and experimental data from a controlled gear tooth chip experiment. Despite the faulty gear vibration data showing no visible indication of fault in the original or AR residual data, the deconvolution methods are able to successfully extract the fault clearly, with the proposed MCKD method performing the best. An online threshold alarm implementation of the MCKD method is presented, shown to be computationally achievable, and effective on looped experimental data.

The novel deconvolution norm CK is presented in Section 4.1, along with the CK values for some sample input signals. Next, an iteratively converging solution to the MCKD problem is derived for first and M-shift in Section 4.2. Simulated deconvolution results on an impulse train plus noise signal are presented for the standard MED and the proposed MCKD method in Section 4.3, from which the advantage of the MCKD method is clearly demonstrated. Next, Section 4.4 presents simulation results for repeating concurrent faults with different transmission paths. Here the MCKD method is shown to be able to extract the concurrent faults separately and indicate the faults clearly. Experimental validation, Section 4.5, is then performed on a controlled gear chip gearbox test, and results are compared among the AR, ARMED, MED and MCDK methods. The proposed MCKD method most clearly identifies the repeating fault in the time domain, and is able to indicate a fault significantly better when comparing fault indicators between fault and no-fault data. Finally a computationally simple online concurrent fault detection implementation of the MCKD fault detection method is presented in Section 4.6. Validation of the online implementation is performed on looped experimental data, and is shown to have strong fault detection results while being computationally achievable for online application.

4.1 Correlated Kurtosis

To improve upon the MED deconvolution technique, the periodicity of the fault can be taken advantage of through the definition of a new deconvolution norm. This proposed norm,

$$\begin{aligned} \text{Correlated Kurtosis of First-Shift} &= CK_1(T) = \frac{\sum_{n=1}^N (y_n y_{n-T})^2}{(\sum_{n=1}^N y_n^2)^2}, \\ \text{Correlated Kurtosis of M-Shift} &= CK_M(T) = \frac{\sum_{n=1}^N \left(\prod_{m=0}^M y_{n-mT} \right)^2}{(\sum_{n=1}^N y_n^2)^{M+1}}, \end{aligned} \quad (4.1)$$

$$y_n = \sum_{k=1}^L f_k x_{n-k+1}, \quad x_n = 0 \text{ and } y_n = 0 \text{ for } n \neq 1, 2, \dots, N, \quad ,$$

encourages filter output periodicity about a period T and high-kurtosis, hence the name Correlated Kurtosis. N is the number of samples in the input signal \vec{x} , L is the length of FIR

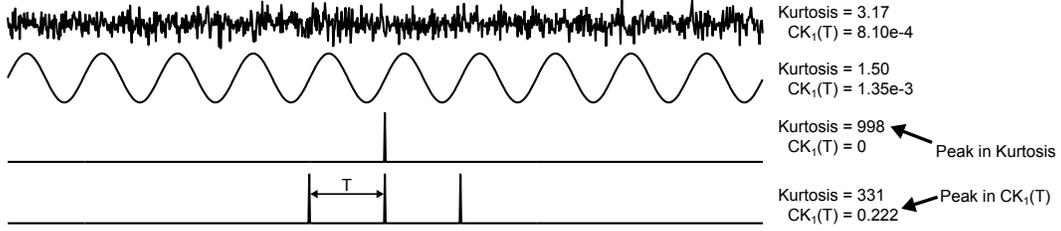


Figure 4.1: Kurtosis and CK values for several signals. The kurtosis reaches a maximum with a single impulse, the CK_1 reaches a maximum with 3 impulses. The first signal is white noise.

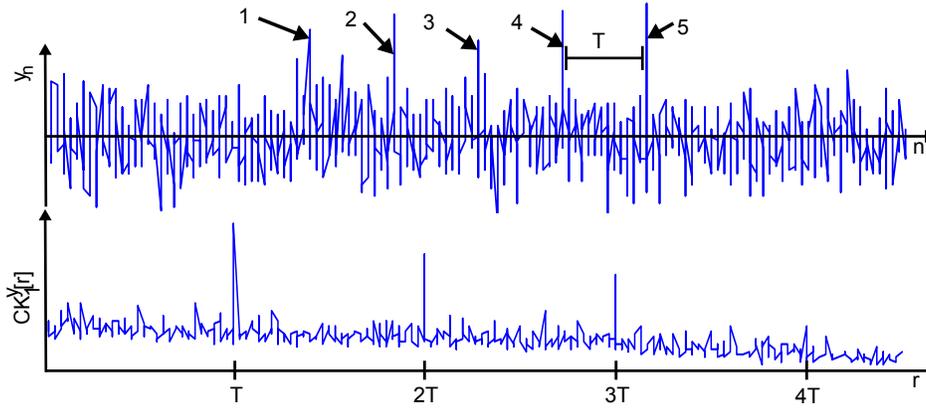


Figure 4.2: CK_1 spectrum for an impulse train of five impulses with spacing of T and strong additive Gaussian white noise corruption.

filter \vec{f} , and T is the period of interest. It should be noted that if $T = 0$ and $M = 1$, then CK is the Kurtosis norm used by MED in Eq. 2.1. Figure 4.1 illustrates the CK_1 versus Kurtosis for several simple signals. It can be seen that the proposed CK_1 approaches a maximum for a periodic impulse about the specified period as opposed to the Kurtosis which tends to a maximum with a single impulse. Higher shift CK emphasizes larger sequences of impulses in a row. Figure 4.2 illustrates the CK value versus period T . It can be seen that $CK(r)$ reaches a peak when $r = T, 2T, \dots$

To illustrate the usage of CK towards extracting fault signals, we compare the first-shift CK values for a simple simulated vibration example where we have a sinusoidal vibration with an impact once per rotation:

$$\text{Signal 1: } y_n = \sin(2\pi n/100) + 0.5\sin(4\pi n/100) + \text{noise}$$

$$\text{Signal 2: } y_n = 0.3 \sum_{k=0}^{\infty} \delta_{n-k100} + 1 [\sin(2\pi n/100) + 0.5 * \sin(4\pi n/100) + \text{noise}]$$

$$\text{Signal 3: } y_n = 0.3 \sum_{k=0}^{\infty} \delta_{n-k100} + 0.5 [(\sin(2\pi n/100) + 0.5\sin(4\pi n/100)) + \text{noise}]$$

$$\text{Signal 4: } y_n = 0.3 \sum_{k=0}^{\infty} \delta_{n-k100} + 0.2 [(\sin(2\pi n/100) + 0.5\sin(4\pi n/100)) + \text{noise}]$$

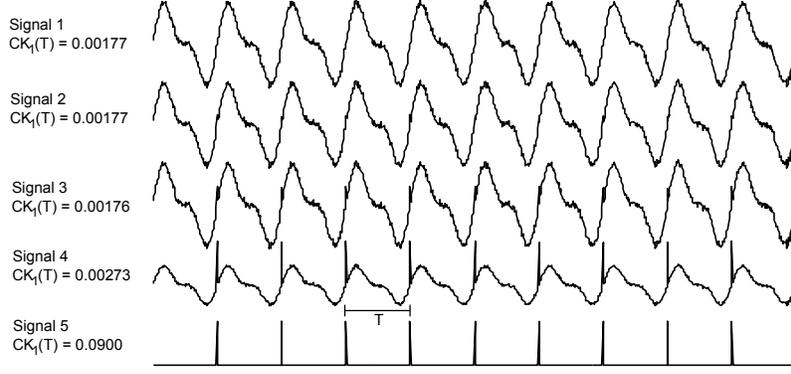


Figure 4.3: CK_1 values for a simple simulated repetitive fault with various amplitudes of the base harmonic vibration. All signals are normalized by their peak value for better illustration.

Signal 5: $y_n = 0.3 \sum_{k=0}^{\infty} \delta_{n-k100}$

where $\delta_k = 1$ for $k = 0$ and $\delta_k = 0$ otherwise, and $n = 1, 2, \dots, 2000$. The noise is zero-mean white noise with variance of 0.22. Fig. 4.3 illustrates $CK_1(100)$ for these signals and it is clear that the maximum of these signals is the fault signal by itself, Signal 5. The goal is to extract the fault impulses through maximizing CK , and this problem is approached through a deconvolution method presented in the following section. Special consideration is taken in the next section by initial conditions to prevent the local maximum solution achieved by Signal 1.

4.2 Maximum Correlated Kurtosis Deconvolution

4.2.1 First-Shift Maximum Correlated Kurtosis Deconvolution

The MCKD technique aims to maximize CK for input signal, \vec{x} , about the period, T , by selecting a FIR filter \vec{f} . First we only consider the first-shift MCKD algorithm, and then expand the method to M-shift in the next section. Starting from the maximization problem:

$$MCKD_1(T) = \max_{\vec{f}} CK_1(T) = \max_{\vec{f}} \frac{\sum_{n=1}^N (y_n y_{n-T})^2}{\left(\sum_{n=1}^N y_n^2\right)^2} \quad (4.2)$$

$$\vec{f} = [f_1 \ f_2 \ \dots \ f_L]^T$$

Towards solving for the filter coefficients corresponding to the maximum, we solve:

$$\frac{d}{df_k} CK_1(T) = 0, \quad k = 1, 2, \dots, L \quad (4.3)$$

First, the derivatives of the numerator and denominator portions are solved separately. Starting with the numerator,

$$\begin{aligned} \frac{d}{df_k} \text{CK}_1 \text{ Numerator} &= \frac{d}{df_k} \sum_{n=1}^N (y_n y_{n-T})^2 = \sum_{n=1}^N 2y_n y_{n-T} \frac{d}{df_k} y_n y_{n-T} \\ &= \left(\sum_{n=1}^N 2y_n y_{n-T}^2 \frac{d}{df_k} y_n \right) + \left(\sum_{n=1}^N 2y_n^2 y_{n-T} \frac{d}{df_k} y_{n-T} \right) \end{aligned}$$

and since,

$$\frac{d}{df_k} y_n = x_{n-k+1}$$

we have,

$$\frac{d}{df_k} \text{CK}_1 \text{ Numerator} = \sum_{n=1}^N 2x_{n-k+1} y_n y_{n-T}^2 + \sum_{n=1}^N 2x_{n-T-k+1} y_{n-T} y_n^2. \quad (4.4)$$

Similarly for the denominator,

$$\begin{aligned} \frac{d}{df_k} \text{CK}_1 \text{ Denominator} &= \frac{d}{df_k} \left(\sum_{n=1}^N y_n^2 \right)^2 = 2 \left(\sum_{n=1}^N y_n^2 \right) \frac{d}{df_k} \sum_{n=1}^N y_n^2 \\ &= 4 \left(\sum_{n=1}^N y_n^2 \right) \sum_{n=1}^N y_n x_{n-k+1}. \end{aligned} \quad (4.5)$$

From Eq. 4.3, Eq. 4.4, and Eq. 4.5 it follows:

$$\begin{aligned} \frac{d}{df_k} \text{CK}_1(T) &= 2\|\vec{y}\|^{-4} \left(\sum_{n=1}^N x_{n-k+1} y_n y_{n-T}^2 + \sum_{n=1}^N x_{n-T-k+1} y_{n-T} y_n^2 \right) \\ &\quad - 4\|\vec{y}\|^{-6} \sum_{n=1}^N (y_n y_{n-T})^2 \sum_{n=1}^N y_n x_{n-k+1} \end{aligned}$$

And rewriting in matrix form:

$$\frac{d}{df} \text{CK}_1(T) = \vec{0} = 2\|\vec{y}\|^{-4} (X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1) - 4\|\vec{y}\|^{-6} \|\vec{\beta}\|^2 X_0 \vec{y}, \quad (4.6)$$

$$X_r = \begin{bmatrix} x_{1-r} & x_{2-r} & x_{3-r} & \cdots & x_{N-r} \\ 0 & x_{1-r} & x_{2-r} & \cdots & x_{N-1-r} \\ 0 & 0 & x_{1-r} & \cdots & x_{N-2-r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_{N-L-r+1} \end{bmatrix}_{L \text{ by } N},$$

$$\vec{\alpha}_0 = [y_1 y_{1-T}^2 \quad y_2 y_{2-T}^2 \quad \cdots \quad y_N y_{N-T}^2]^T, \quad \vec{\alpha}_1 = [y_{1-T} y_1^2 \quad y_{2-T} y_2^2 \quad \cdots \quad y_{N-T} y_N^2]^T,$$

$$\vec{\beta} = [y_1 y_{1-T} \quad y_2 y_{2-T} \quad \cdots \quad y_N y_{N-T}]^T$$

Rearranging Eq. 4.6,

$$2\|\vec{\beta}\|^2 X_0 \vec{y} = \|\vec{y}\|^2 (X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1)$$

and substituting the relation,

$$\vec{y} = X_0^T \vec{f} \quad (4.7)$$

results in:

$$X_0 X_0^T \vec{f} = \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2} (X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1)$$

The matrix $X_0 X_0^T$ is the Toeplitz autocorrelation matrix of \vec{x} and the inverse $(X_0 X_0^T)^{-1}$ is assumed to exist:

$$\vec{f} = \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2} (X_0 X_0^T)^{-1} (X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1) \quad (4.8)$$

This resulting equation is nonlinear, but a local maximum solution for \vec{f} can be solved for iteratively. All datasets processed so far have been found to be monotonically convergent to a local maximum solution. The procedure for solving for \vec{f} iteratively is as follows, and a link to the MATLAB implementation can be found in the External Resources Section, Appendix A:

- Step 1: Select period of interest, T .
- Step 2: Calculate X_T , X_0^T , and $(X_0 X_0^T)^{-1}$ from your input signal \vec{x} .
- Step 3: Select filter size, L , and assume an initial filter of $\vec{f} = [0 \ 0 \ \dots \ 1 \ -1 \ \dots \ 0 \ 0]^T$. This is selected as a difference filter to prevent the algorithm from converging to the local solution of Signal 1 in Fig. 4.3 and the difference is in the centre because performance can be improved through not assuming a minimum phase filter.
- Step 4: Calculate the filtered output, \vec{y} , from Eq. 4.7.
- Step 5: Calculate $\vec{\alpha}_0$, $\vec{\alpha}_1$, and $\vec{\beta}$ from \vec{y} .
- Step 6: Calculate the new filter coefficient, \vec{f} , from Eq. 4.8.
- Step 7: Is $\Delta CK_1(T) > \epsilon$? Loop from Step 4 while true. ϵ is a small positive number controlling when the algorithm terminates and ΔCK_1 denotes the change for the iteration.
- Step 8: The final first-shift MCKD-filtered signal about period T is calculated from Eq. 4.7.

4.2.2 M-Shift Maximum Correlated Kurtosis Deconvolution

Results generally improve significantly on experimental data by using a higher shift MCKD method because increasing the shift increases the number of sequential impulses the algo-

rithm is looking to deconvolve. However, higher order shifts requires better estimates of the fault period T and increases the complexity of the calculation.

Similarly to the first-shift derivation, we start from the maximization problem

$$MCKD_M(T) = \max_{\vec{f}} CK_M(T) = \max_{\vec{f}} \frac{\sum_{n=1}^N \left(\prod_{m=0}^M y_{n-mT} \right)^2}{\left(\sum_{n=1}^N y_n^2 \right)^{M+1}} \quad (4.9)$$

and by solving the derivative of the numerator and denominator of $CK_M(T)$ with respect to filter coefficients f_k :

$$\frac{d}{df_k} CK_M \text{ Numerator} = 2 \sum_{n=1}^N \left[\left(\prod_{m=0}^M y_{n-mT} \right)^2 \left(\sum_{m=0}^M \frac{x_{n-mT-k+1}}{y_{n-mT}} \right) \right] \quad (4.10)$$

$$\frac{d}{df_k} CK_M \text{ Denominator} = 2(M+1) \|\vec{y}\|^{2M} \sum_{n=1}^N y_n x_{n-k+1} \quad (4.11)$$

Combining Eq. 4.10 and Eq. 4.11 results in,

$$\begin{aligned} \frac{d}{df_k} CK_M(T) = 0 &= 2 \|\vec{y}\|^{-2M-2} \sum_{n=1}^N \left[\left(\prod_{m=0}^M y_{n-mT} \right)^2 \left(\sum_{m=0}^M \frac{x_{n-mT-k+1}}{y_{n-mT}} \right) \right] \\ &\quad - 2(M+1) \|\vec{y}\|^{-2M-4} \sum_{n=1}^N \left(\left(\prod_{m=0}^M y_{n-mT} \right)^2 \right) \sum_{n=1}^N y_n x_{n-k+1}, \end{aligned}$$

and converting to matrix form with $k = 1, 2, \dots, L$ and rearranging results in the iterative solution:

$$\begin{aligned} \vec{f} &= \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2} (X_0 X_0^T)^{-1} \sum_{m=0}^M X_{mT} \vec{\alpha}_m \quad (4.12) \\ \vec{\alpha}_m &= \begin{bmatrix} y_{1-mT}^{-1} (y_1^2 y_{1-T}^2 \cdots y_{1-MT}^2) \\ y_{2-mT}^{-1} (y_2^2 y_{2-T}^2 \cdots y_{2-MT}^2) \\ \vdots \\ y_{N-mT}^{-1} (y_N^2 y_{N-T}^2 \cdots y_{N-MT}^2) \end{bmatrix}, \quad \vec{\beta} = \begin{bmatrix} y_1 y_{1-T} \cdots y_{1-MT} \\ y_2 y_{2-T} \cdots y_{2-MT} \\ \vdots \\ y_N y_{N-T} \cdots y_{N-MT} \end{bmatrix} \end{aligned}$$

Eq. 4.12 can be solved iteratively by following a similar procedure as to the first-shift MCKD presented above. Unlike the first-shift MCKD, the solution does not necessarily converge monotonically, and as a result the filter \vec{f} is chosen as the filter which results in the maximum in CK_M during the iterative process. For large M , from experience around 10 or more, the iterative method can result in loss of numerical precision because of exceeding the range of the floating point exponent.

Because the higher-shift MCKD requires better estimates of the fault period T and in application the period is a fractional number (eg. 170.21 samples per gear revolution), an

additional resampling stage is introduced as a preprocessing step. This additional step resamples the input signal \vec{x} at a ratio of:

$$\lfloor 20T + 0.5 \rfloor : \lceil 20\lceil T \rceil + 0.5 \rceil$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denotes the floor and ceiling operations respectively and the factor of 20 is chosen as a good balance between computation time and precision. This resamples the data so that the samples per revolution is approximately the nearest larger integer; for example 170.21 samples per gear revolution resamples at a ratio of 3404 : 3600, resulting in 180.01 samples per gear revolution. The resampling is performed by polyphase filter implementation.

The implementation of this method in MATLAB is available in the External Resources Section, Appendix A.

4.3 Simulation Results

To compare the MED and MCKD algorithms' ability to extract periodic repeating impulses, a simple deconvolution of impulses from white noise is analysed. The signal is formed as

$$x_n = e_n + f \left(\sum_{k=0}^{\infty} \delta_{n-k100} \right), \quad n = 0, 1, 3, \dots, 999$$

where e_n is zero-mean Gaussian white noise of 1 standard deviation and the fault amplitude, f , is varied between 0 and 5 at a step of 0.01. Fig. 4.4a illustrates the fault detection versus f for i.i.d. e_n for each f , while Fig. 4.4b illustrates the resulting deconvolved signals when $f = 3$. A deconvolution filter size of 200 and iteration count of 100 is used for each method. From these results it is clear that the MCKD method greatly outperforms the MED method when deconvolving periodic impulses of known period from white noise. This clearly illustrates the problem with the standard MED algorithm, which prefers to deconvolve a single impulse as the result. But not only that, the MED method fails to indicate any correlation at all between kurtosis and fault level.

4.4 Basic Concurrent Gear Fault Simulation

Consider the simple vibration model:

$$x_n = \sin(2\pi n/30) + 0.7\sin(3\pi n/30) + f_n + e_n$$

where f_n denotes the periodic fault signal caused by the faulty gear, e_n is zero-mean additive white Gaussian noise with a standard deviation of 0.1, and $n = 0, 2, \dots, 2999$. Then

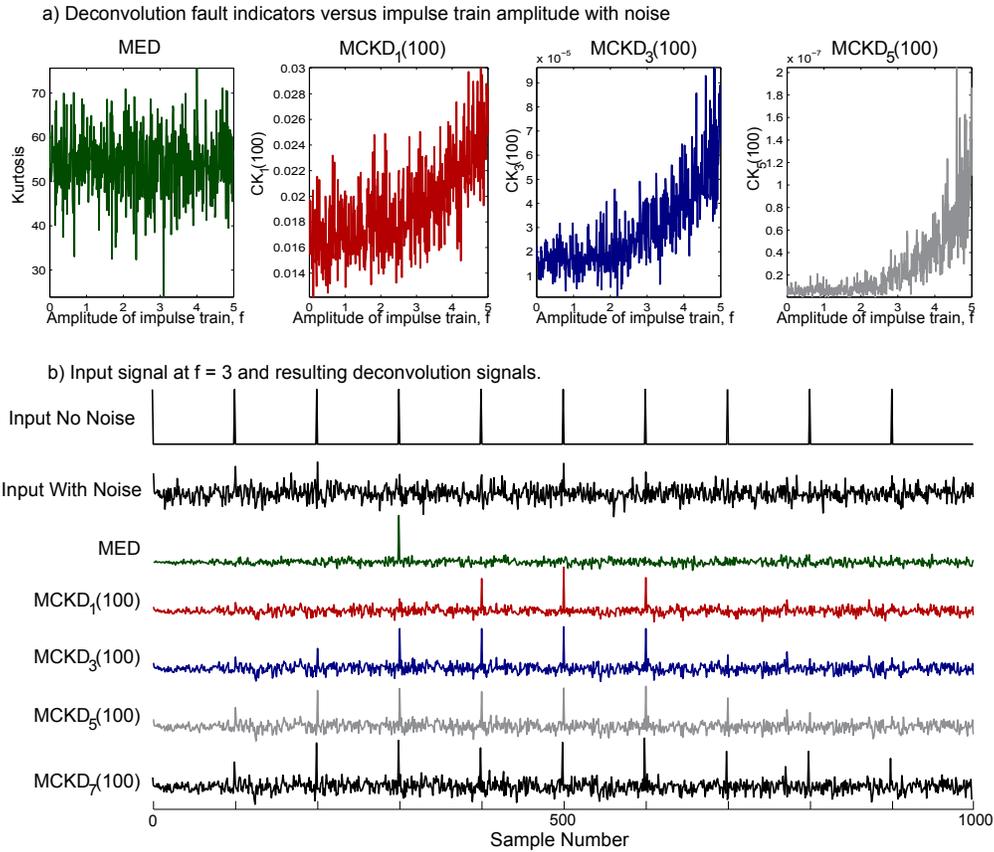


Figure 4.4: Deconvolution of a noisy periodic impulse train of period 100 with a) illustrating the fault detection versus impulse train amplitude, and b) normalized plot of the resulting deconvolved signals with impulse train amplitude of $f = 3$, some signals are polarity inverted for illustration purposes. A different white noise seed is used for each value of f .

consider two faulty elements in the system on with fault periods of 30 and 100 samples. These faulty elements result in impulse-like vibration signals convolved with two separate transmission paths:

$$f_n^a = h_n^a * \left(\sum_{k=0}^{\infty} \delta_{n-k30} \right)$$

$$f_n^b = h_n^b * \left(\sum_{k=0}^{\infty} \delta_{n-k100} \right)$$

Impulse responses h_n^a and h_n^b refer to two different the vibration transmission paths from the faulty gears to vibration sensor, combined with slightly different fault characteristics. We select these two transmission paths as:

$$\vec{h}^a = [0.9 \quad -0.3 \quad 0.1 \quad -0.05]^T$$

$$\vec{h}^b = [0.1 \quad 0.8 \quad -0.7 \quad -0.6 \quad 0.4 \quad -0.2 \quad 0.1 \quad -0.05]^T$$

and form four cases:

F1: No Fault: $f_n = 0$

F2: Fault on Period 30: $f_n = 0.4f_n^a$

F3: Fault on Period 100: $f_n = 0.4f_n^b$

F4: Concurrent Faults on Periods 30 and 100: $f_n = 0.4f_n^a + 0.4f_n^b$

The desired output signals include only the impulse fault signals before the transmission path. The resulting fault indicators, $CK_M(T)$ for MCKD and the Kurtosis for MED-based methods, are tabulated in Table 4.1. All algorithms were with filter size of 200, and the MED and MCKD algorithms were iterated exactly 100 times each. Shift-five MCKD was chosen as a good balance between a higher shift method and lower computation time. AR model orders of 20 were chosen because it is a fairly high model order while being smaller than the smallest fault period. Higher numbers indicate a larger detected fault. From the table, it is clear to see that only the MCKD algorithm detects the faults successfully. The MED technique performs so poorly as a result of incorrectly deconvolving a single impulse from the no-fault case. The MCKD method is not only able to strongly detect the faults, but is also able to indicate exactly which fault periods are present.

Figure 4.5 illustrates the resulting signals for the MCKD approach under concurrent fault condition, F4. It is clear that the faults with a period of 30 and 100 are successfully isolated separately during the deconvolution process.

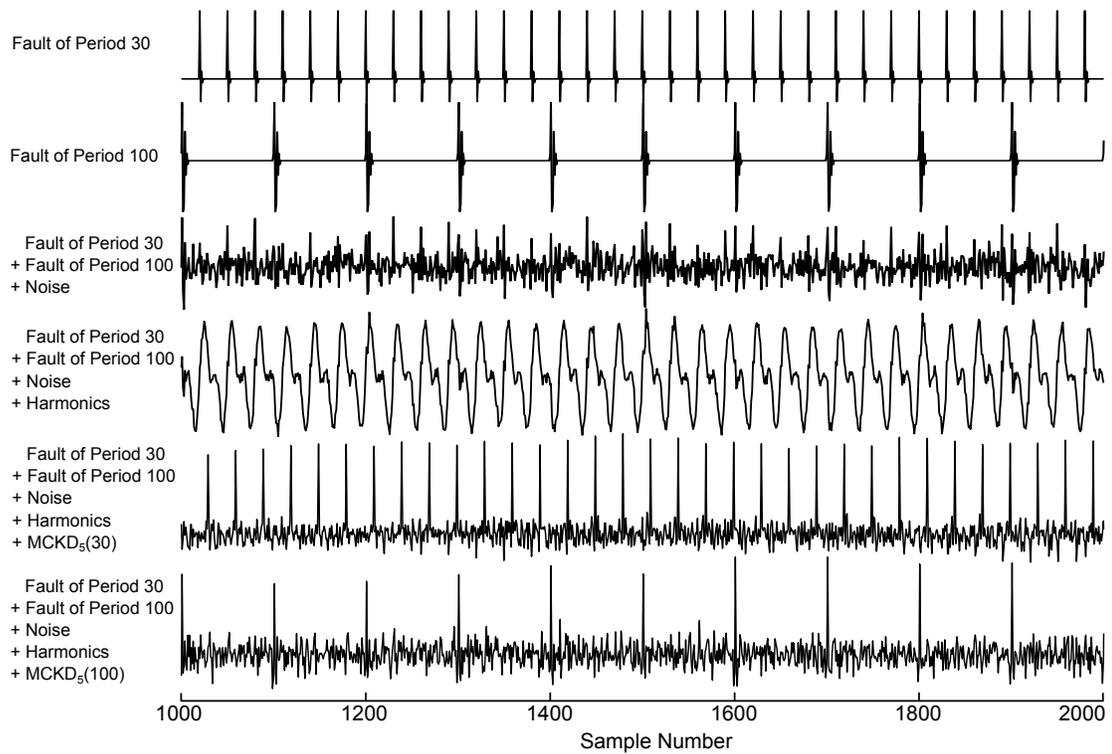


Figure 4.5: Simulated resulting signals for concurrent fault detection after processing with indicated method. All signals are normalized and some signals are polarity flipped for better illustration.

Table 4.1: Fault indicator values by MCKD, MED, and ARMED. The fault indicators are final CK values for MCKD and the final Kurtosis values for MED and ARMED. All values are no-fault normalized.

Method	$F1$, No Fault	$F2$, Fault on Period 30	$F3$, Fault on Period 100	$F4$, Concurrent Faults on Periods 30 and 100
$MCKD_5(30)$	1	9.75	0.929	8.55
$MCKD_5(100)$	1	2.62	79.8	77.4
MED	1	0.139	0.111	0.137
$ARMED$	1	1.01	1.09	0.995

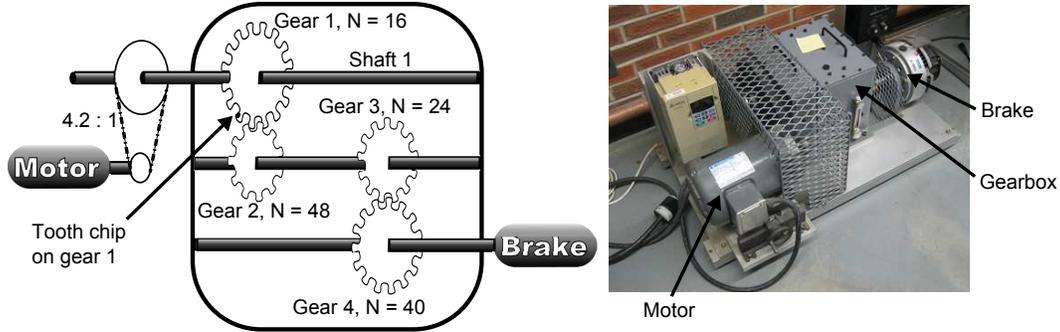


Figure 4.6: Gearbox and experimental equipment layout.

4.5 Experimental Results

For validation, vibration data is collected and compared from a gearbox with and without a gear chip. The machine configuration, Fig. 4.6, is composed of a motor, gearbox, and brake. Two gears are used for the Gear 1 position; one in healthy condition and one with a gear crack, Fig. 4.7. The machine is operated at 10% load and the rotational frequency of Shaft 1 is varied between 10, 15, 20, 25, 30, 35, and 40 Hz [57]. The accelerometer is of model PCB 352C67 and the data is acquired through a DSP Siglab 20-42 Signal Analyzer to a laptop [11]. The vibration accelerometer sampling frequency is varied according to the rotational speed between 1280 Hz and 5120 Hz, and 8192 samples are acquired for each measurement. At each frequency, two non-fault vibration measurements and a single fault vibration measurement is acquired and processed.

In the ARMED method proposed by H. Endo and R. Randall, the authors suggest that the MED stage should just be performed for only a few iterations to prevent erroneously extracting impulses [5]. Fig. 4.8 illustrates the results of MCKD, MED, and ARMED versus the iteration in the deconvolution stage for the experimental gear chip data. It can be seen that for both the MED and ARMED approach, the deconvolved signals indicate the fault clearest with a high iteration count, although the methods fail to extract a fault impulse for each revolution of the faulty gear. Reducing the iteration count does not appear to resolve



Figure 4.7: Faulty gear 1 with seeded tooth chip.

this problem in either case. As a result, for the duration of this thesis the MED algorithm is executed to convergence instead of being limited to just a few iterations. Fig. 4.9 illustrates the deconvolution results for the same fault vibration signal versus filter size. By close inspection, it is clear that for the MCKD method a higher filter size results in improved results. With the MED approach the repeating fault is roughly extracted at around a 50 filter size, and larger filter sizes cleans the signal up but at a cost of missing many of the fault impulses.

The general procedure for the data processing is as follows:

- Step 1: Shaft 1 speeds for each data measurement are assumed to be constant and close to the indicated frequency. To calculate the speed, the Discrete Fourier Transform is applied and peak detection is performed within the vicinity of the expected shaft speed. The period of Shaft 1 in number of samples is referred to as T , and can be a fractional number of samples (such as 172.32 samples per gear revolution). Direct speed measurement would likely result in better performance of the MCDK method.
- Step 2: Generate AR models at each operational frequency from the second no-fault data measurement of order $\lfloor 0.4T + 0.5 \rfloor$, i.e. 40% the number of data samples corresponding to a single Shaft 1 revolution and rounded to the nearest integer. The Akaike Information Criterion is not used to select the AR model orders because it erroneously overfits the data by suggesting model orders almost equal to the dataset sizes of around 8000 samples.
- Step 3: Calculate AR model residuals for the first no-fault data and the fault data measurements by 1-step ahead prediction.
- Step 4: Deconvolution filter sizes for both MED and MCKD are selected as

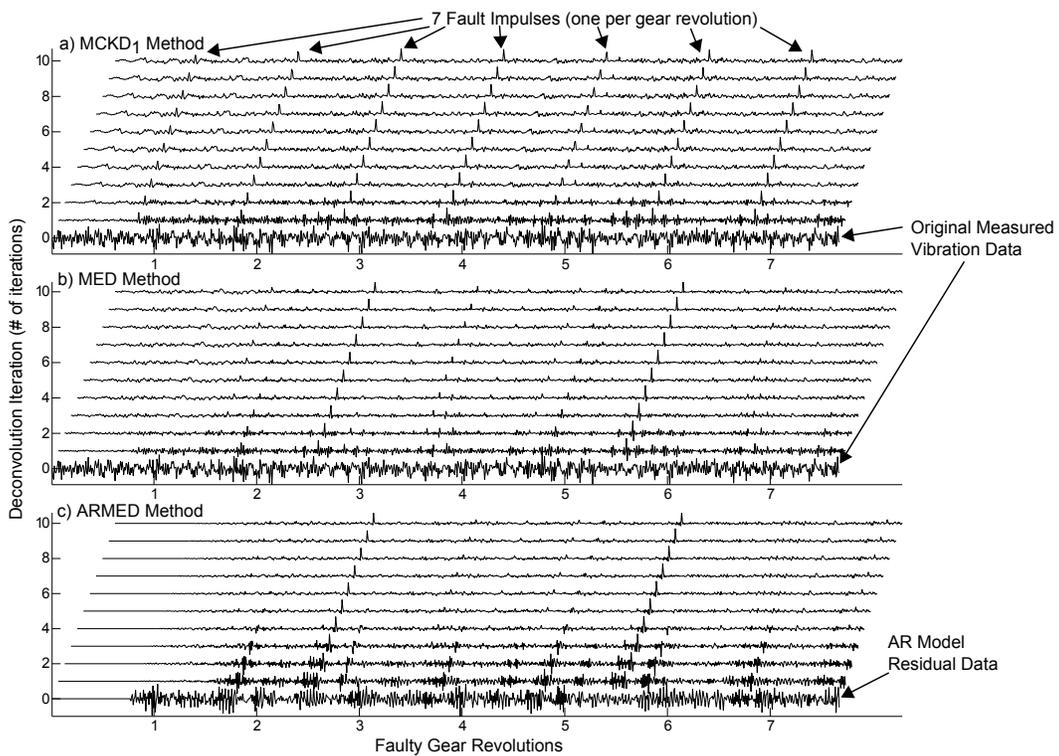


Figure 4.8: Deconvolved signals versus deconvolution iteration for a) first-shift MCKD, b) MED, and c) ARMED for an AR model order of 100 and filter sizes of 200 samples. The vibration data is from the gearbox under gear chip fault at 40Hz shaft speed and 10% load and resulting signals are normalized by peak values with some signals being polarity swapped for better illustration.

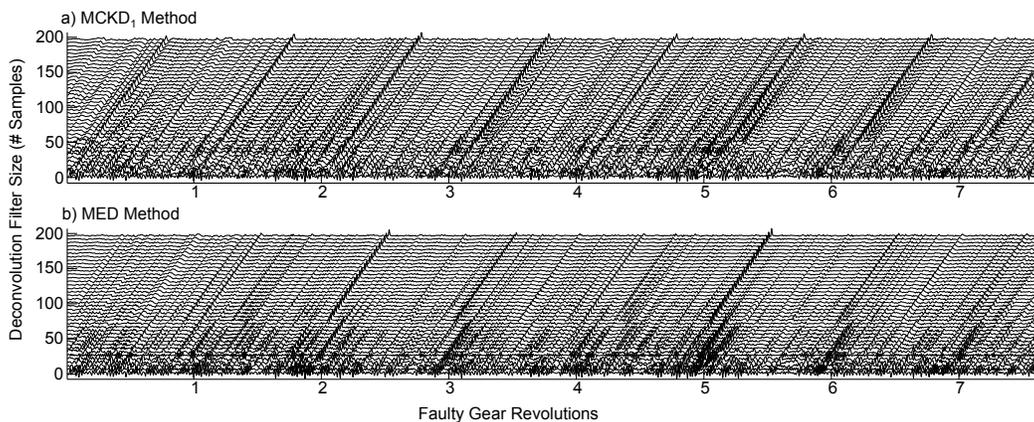


Figure 4.9: a) MCKD and b) MED methods versus deconvolution filter size for the machine under gear chip fault at 40Hz shaft speed and 10% load. Each deconvolution method is iterated 50 times and the signals are normalized by peak values polarity swapped for clearer illustration.

$\lfloor 0.8T + 0.5 \rfloor$, i.e. 80% the number of data samples in a single Shaft 1 revolution and rounded to the nearest integer.

Step 5: MED is performed on each AR residual, each first no-fault dataset, and each fault dataset; resulting in the signals for MED and ARMED methods. The kurtosis of each resulting output signal is calculated as a fault indicator.

Step 6: First, third, and fifth-shift MCKD is applied about period T for each first no-fault dataset, and each fault dataset. $CK_M(T)$ is calculated from the resulting deconvolved signals as a fault indicator where M corresponds to the shift of the MCKD algorithm (eg. $CK_3(T)$ is used as the fault indicator for third-shift MCKD).

Time synchronous averaging is not considered as a preprocessing step due the lack of tachometer or detailed speedometer information. The deconvolution filter sizes and AR model orders for the fault datasets and first no-fault datasets are tabulated in Table 4.2. These high-length deconvolution filters are selected because the time-domain plots more clearly illustrates the faults.

Table 4.2: Estimated shaft 1 period and corresponding deconvolution filter sizes for no-fault and fault cases.

Shaft 1 Frequency	No-Fault Dataset		Fault Dataset	
	Deconv. Filter Size L (# Samples)	AR Model Order	Deconv. Filter Size L (# Samples)	AR Model Order
10 Hz	208	104	222	111
15 Hz	139	69	142	71
20 Hz	207	104	211	105
25 Hz	166	83	168	84
30 Hz	138	69	139	70
35 Hz	118	59	119	60
40 Hz	103	52	104	52

The resulting processed signals for no-fault and fault conditions at 35 Hz are compared in Fig. 4.10. All processed signals under gear chip fault are plotted in Fig. 4.11 for all frequencies and all methods. From these results it is clear that the MCKD method extracts the expected single impulse per gear revolution, while the MED-based methods tend to miss most of the fault impulses. The detected fault impulses by different methods do not necessarily line up as a result of the different phase lags of the resulting filters. The exact placement of the faults in the time domain can be recovered by simply plotting the filter responses and typically the lag is clear in this response.

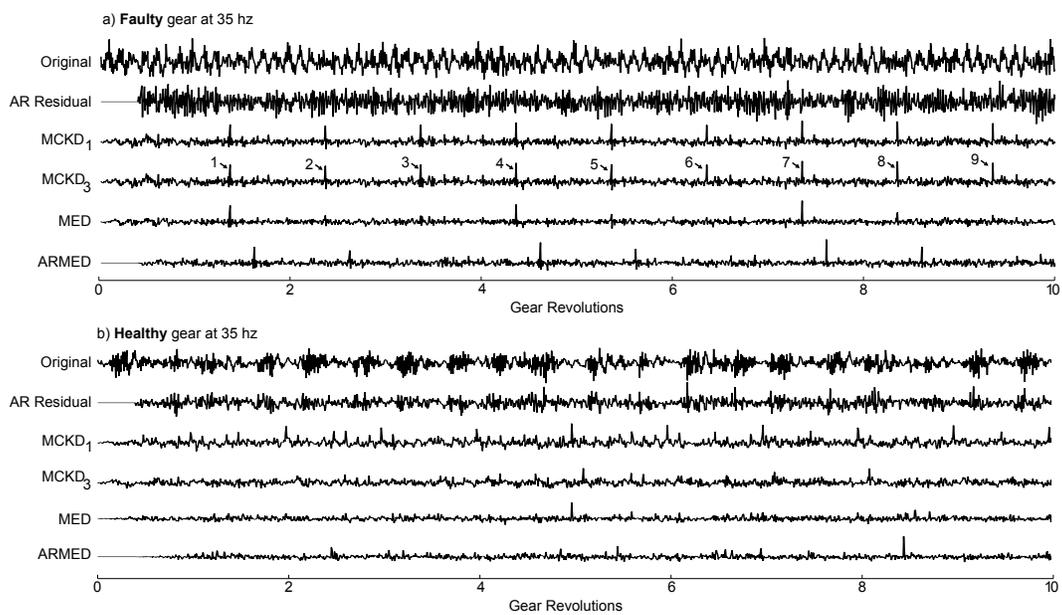


Figure 4.10: Experimental acceleration data with fault detection processing with a) gear chip and shaft 1 speed of 35 Hz, and b) normal gear and shaft 1 speed of 35 Hz. All signals are normalized by peak values and some signals are reversed in polarity for better illustration.

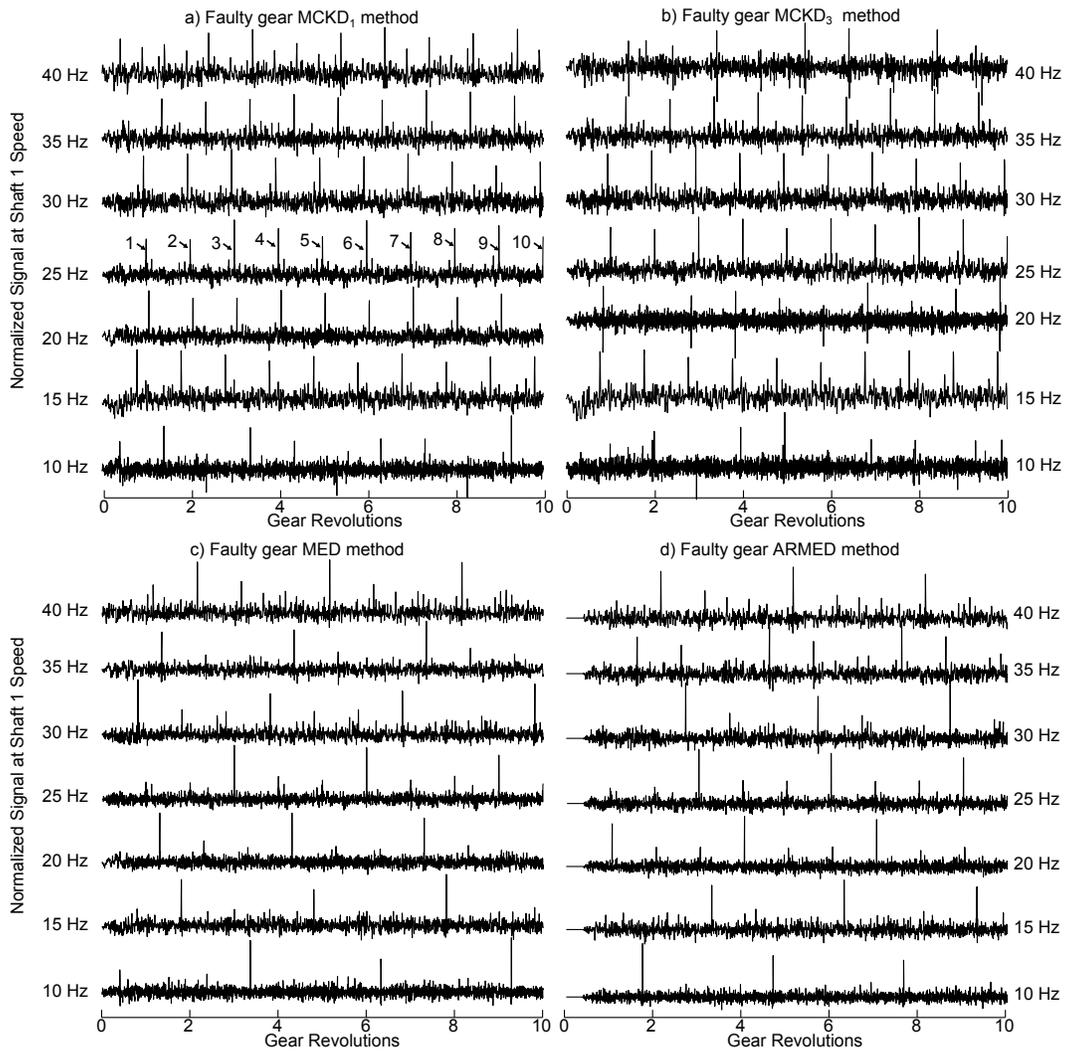


Figure 4.11: Experimental acceleration data processed and normalized with a gear chip using a) $MCKD_1$ method, b) $MCKD_3$ method, c) MED method, and c) the ARMED method. All signals are normalized by peak values and some signals are polarity flipped for better illustration.

Fault detection based on deconvolution filter sizes of 2 through 150 samples are analysed and compared, Fig. 4.12. From the plots it is clear that MCKD method significantly outperforms the MED method, often performing over 10 times better. The performance gain by increasing the shift from first to third and fifth shift is very significant. Also important to note is that the MED approach performs similarly to the ARMED approach, indicating that the AR stage is likely not required.

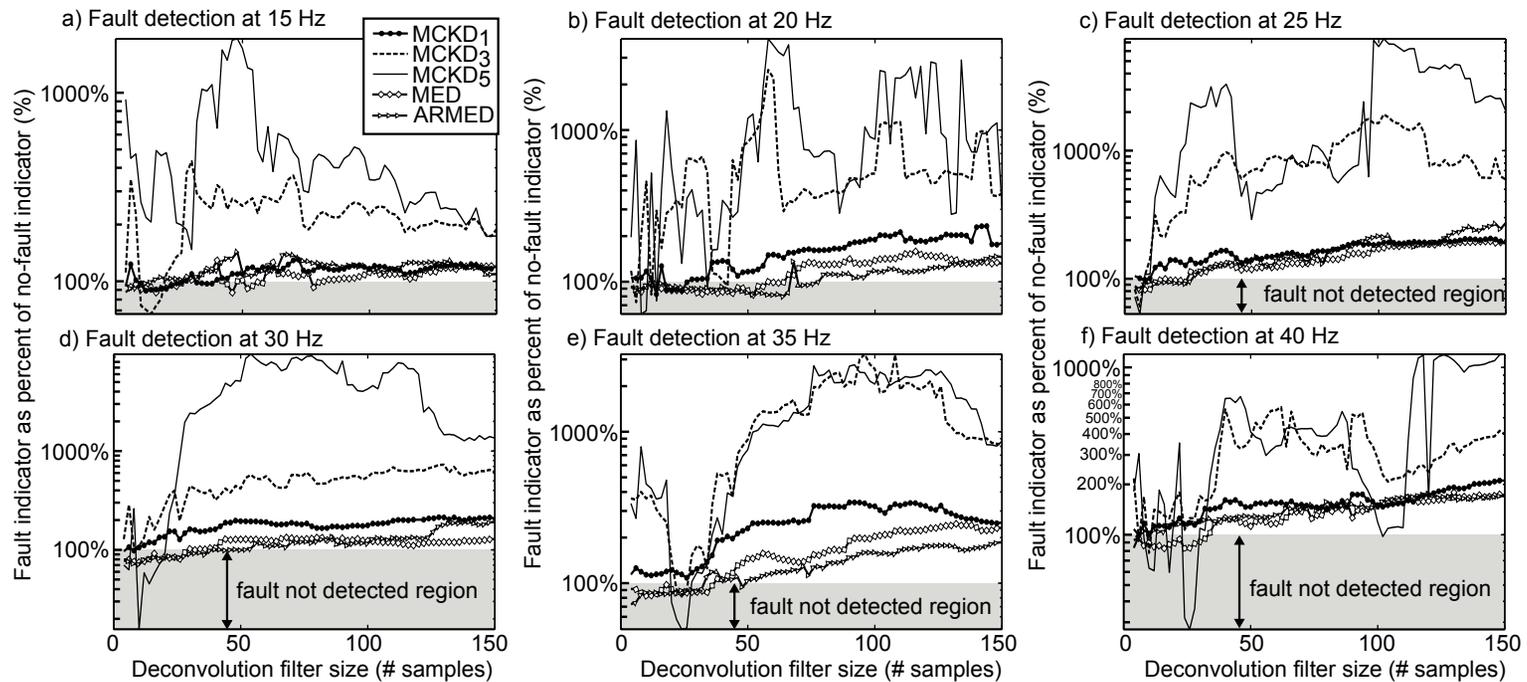


Figure 4.12: Experimental fault detection by method with varied deconvolution filter size for the machine operating at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz. Be careful to note the log-scale on the y-axis, meaning the $MCKD_3$ and $MCKD_5$ typically perform significantly better at fault detection.

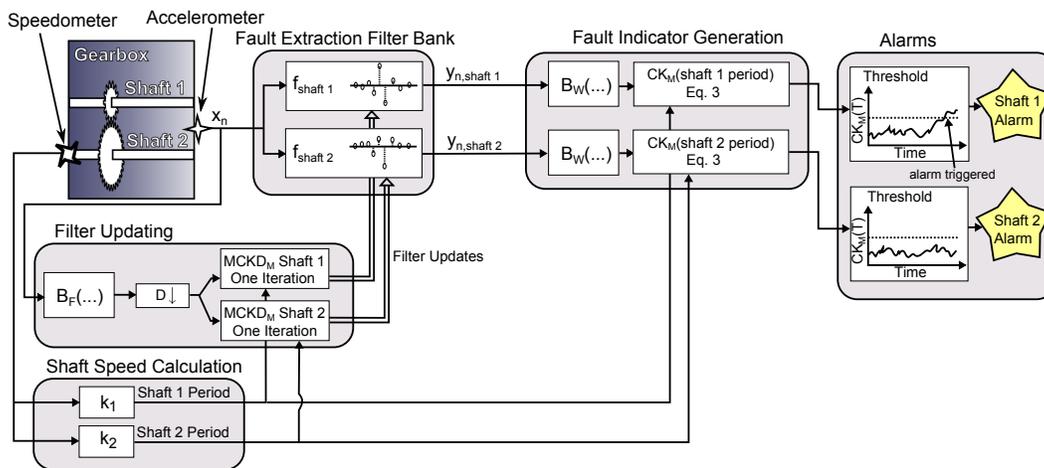


Figure 4.13: Online implementation of MCKD-based fault detection for a two-shaft gearbox with a single accelerometer. Threshold and FIR filter plots are creative illustrations, not data.

4.6 Online Condition Monitoring Implementation

In industry application, fault detection during machine operation is critical in preventing equipment damage or failure. Presented here is a modification of the MCKD technique to monitor the health of a machine online and generate fault alarms.

Two operations for buffering and downsampling are defined as:

Data buffering by factor R and no overlap, B_R :

$$\left[\begin{array}{cccc} x_1 & x_2 & x_3 & \dots \end{array} \right] \xrightarrow{B_R(\dots)} \left[\begin{array}{cccc} x_1 & x_{R+1} & x_{2R+1} & \dots \\ x_2 & x_{R+2} & x_{2R+2} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ x_R & x_{2R} & x_{3R} & \dots \end{array} \right]$$

Downsample by factor D , $D \downarrow$:

$$\left[\begin{array}{cccc} x_{0,0} & x_{0,1} & x_{0,2} & \dots \\ x_{1,0} & x_{1,1} & x_{1,2} & \dots \\ x_{2,0} & x_{2,1} & x_{2,2} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ x_{I,0} & x_{I,1} & x_{I,2} & \dots \end{array} \right] \xrightarrow{D\downarrow} \left[\begin{array}{cccc} x_{0,0} & x_{0,D} & x_{0,2D} & \dots \\ x_{1,0} & x_{1,D} & x_{1,2D} & \dots \\ x_{2,0} & x_{2,D} & x_{2,2D} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ x_{I,0} & x_{I,D} & x_{I,2D} & \dots \end{array} \right]$$

Fig. 4.13 illustrates the online implementation schematic. The presented implementation is for a single accelerometer near both Shaft 1 and Shaft 2, but performance can be improved by placing two different accelerometers close to each shaft. A speed measurement of one of the shafts is recommended, but if the system is expected to be operating at only a single operating speed then Discrete Fourier Transform peak detection on the acceleration signal can estimate the exact machine speed. To reduce complexity, the *MCKD* blocks performs only a single deconvolution filter update iteration for each column of input data and the filter update downsampling factor D controls how often a filter update iteration is performed. The buffer factor F controls the window size of MCKD iteration and a value resulting in at least 10 times the period of the shaft is recommended. The buffer factor W controls the window size for CK fault indicator generation, a larger factor will result in a more slowly changing fault indicator and slower time response to a fault while a smaller factor results in a more sporadic fault indicator but quicker time response to a fault. Constant gain values of k_1 and k_2 convert the measured speed to the periods of Shaft 1 and 2 respectively. The alarm thresholds should be selected based on indicator values under no-fault. $f_{\text{shaft } 1}$ and $f_{\text{shaft } 2}$ refer to the FIR deconvolution filters.

For testing the online implementation, basic input sequences are formed for each operating frequency by looping the experimental no-fault data ten times followed by looping fault data ten times. The data is looped by direct concatenation of the datasets reduced to the number of samples corresponding to the largest number of complete shaft revolutions in

the dataset. A deconvolution filter size of $L = 200$, no decimation factor $D = 1$, and filter update and fault indicator buffer sizes of 5000 is used $W = F = 5000$.

Table 4.3: Computation times for online *MCKD* gear chip fault detection on looped experimental data.

Shaft1 Fre- quency	Number of Samples (kSamples)	Duration of Data Acqui- sition (Seconds)	<i>MCKD</i> ₁ Data Pro- cessing Time (Seconds)	<i>MCKD</i> ₅ Data Pro- cessing Time (Seconds)
15 Hz	326	128	26	44
20 Hz	324	63	27	45
25 Hz	325	64	24	45
30 Hz	325	64	34	53
35 Hz	324	63	25	45
40 Hz	326	64	35	54

Table 4.3 illustrates the number of samples per dataset after looping, the time taken to measure the data, and the processing time for the proposed online fault detection implementation. The processing times are measured on Intel Core 2 Duo CPUs at 2.00GHz with implementation in MATLAB without parallel processing. From the processing times, it is clear that the proposed method is easily achievable in an online application. Fig. 4.14 illustrates the online fault detection results for first-shift *MCKD*. It is clear that the fault detection performance is good in online application, and the proposed threshold alarm on this signal is feasible. The trough in the fault indicator as the fault is introduced is as a result of both the dataset looping not being phase aligned during dataset concatenation between the no-fault and fault data, and the slight machine speed difference between the two datasets. Fig. 4.15 presents the fifth-shift *MCKD* results. Clearly the *MCKD* method is very strong at detecting the faults, with the fifth-shift *MCKD* method providing clearer results than the first-shift implementation but at a higher computational cost.

4.7 Conclusion

This chapter introduces a new deconvolution process, *MCKD*, which aims to deconvolve periodic impulse faults from a machine vibration signal. Simulation data and experimental gear chip vibration data were compared between *MCKD* and the established methods of AR residual, MED, and ARMED. The experimental results indicate that the AR residual method is inadequate, which is as a result of the fault impulses being of comparable amplitude to the system noise and the AR model residual method extracting noise plus disturbance. Additionally, the experimental results indicate that the ARMED method performs similarly to direct MED application; indicating that the AR stage of ARMED is unnecessary. Finally, the simulation and experimental results indicate the *MCKD* method is

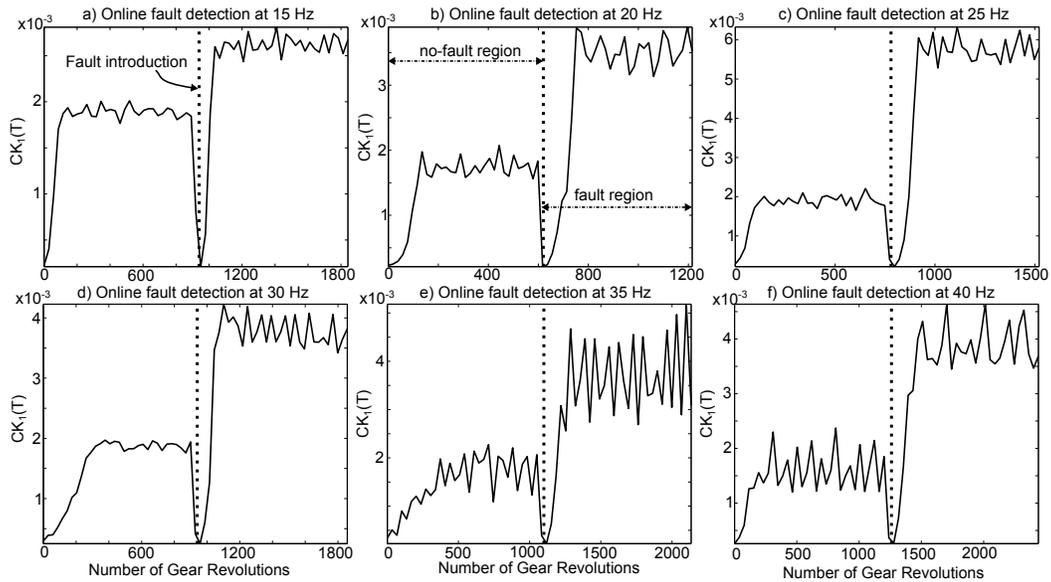


Figure 4.14: Online results for first-shift MCKD fault detection for machine under 10% load at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz. The data is from looped the experimental data.

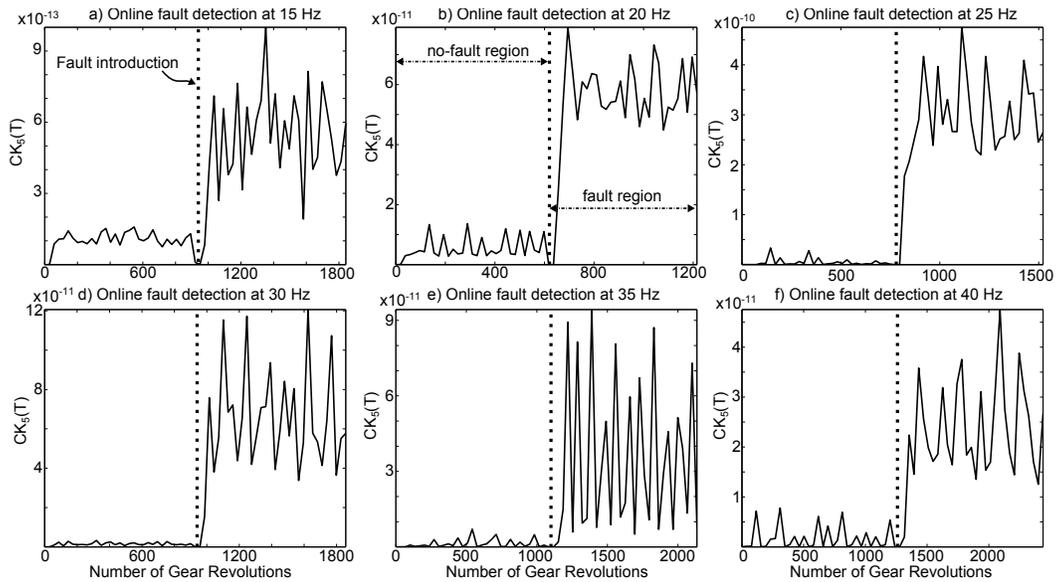


Figure 4.15: Online results for fifth-shift MCKD fault detection for machine under 10% load at a) 15 Hz, b) 20 Hz, c) 25 Hz, d) 30 Hz, e) 35 Hz, and f) 40 Hz.

the most successful in deconvolving the periodic faults and in has significantly better gear chip fault detection results, often on the order of 10 or more times better at fault detection. There is a significant performance advantage gain from increasing the shift-order from first to third, but fifth-shift MCKD did not significantly outperform third-shift. As a result the third-shift MCKD method is recommended in application. Clearly the MCKD method is a notable improvement upon the existing state-of-the-art methods and may include applications beyond rotating machine fault detection.

Online application of the MCKD method is shown to be computationally implementable on a personal computer, and results from fault detection on looped experimental data indicates it is a strong method in online application and that a simple threshold alarm is feasible.

Further work should investigate the application of MCKD towards detecting other types of gear faults and towards detecting bearing faults. Analysis of the proposed online implementation of the MCKD method should be performed through an accelerated gear or bearing degradation experimental setup.

5

Steam Power Turbine Analysis

Vibration data acquired from two 50MW back pressure steam turbine generators is collected by two proximity sensors on bearing 4, Figure 5.1. A data measurement of approximately 1200 samples is collected every two hours from 2003 to 2006 at a sampling rate of 9.7 kSamples/Second. Under normal operation, the turbines are rotating at approximately 60Hz, resulting in approximately 161 samples per revolution and over 7 rotor revolutions per dataset. Data is not consistently measured every 2 hours, but includes large gaps as a result of machine downtimes and includes more frequently measured data when, for example, the turbine health is being analysed by the on-site engineers.

The RMS of the vibration for both turbines, denoted Turbine 1 and 2, are presented in Figure 5.2. Faults often do not affect the vibration level and as a result it is not a good indication of the presence of faults. However, it is still important to keep this vibration level low to prevent faults such as rubbing, which are caused by the vibration exceeding available clearances.

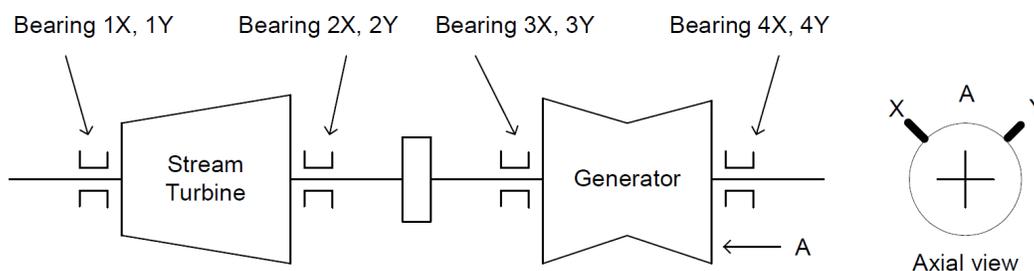


Figure 5.1: The layout of the back pressure steam turbine generator. Source: [58]

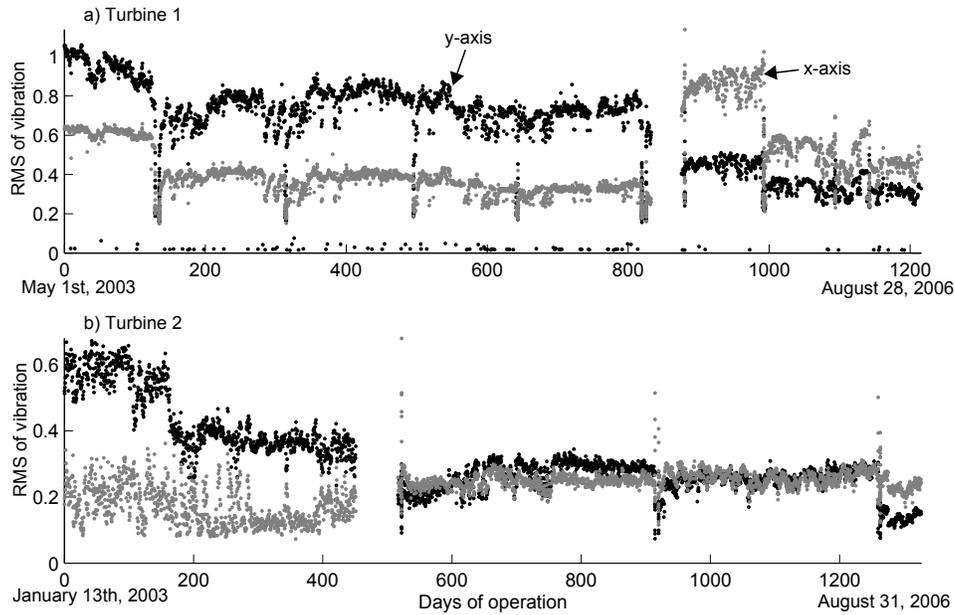


Figure 5.2: RMS of turbine vibration for the X and Y axis on a) Turbine 1, and b) Turbine 2. Each dot represents the RMS of the dataset of approximately 1200 samples collected at the indicated time.

The health of these two turbines, denoted Turbine 1 and Turbine 2, is analysed by the previously presented methods:

- AR prediction residual method, Section 2.3.1.
- Two and three-component frequency estimator prediction residual method, Chapter 3.
- MED filtering method, Section 2.4.4.
- MCKD filtering method, Chapter 4.

Because of a lack of knowledge regarding the bearing specifications, and consequently the inner and outer-race characteristic frequencies, the main focus of this investigation is on rotor faults such as rubbing.

5.1 Turbine 1 Health Analysis

Two problems have been identified in the Turbine 1 data; a sensor fault and a suspected rotor rubbing fault.

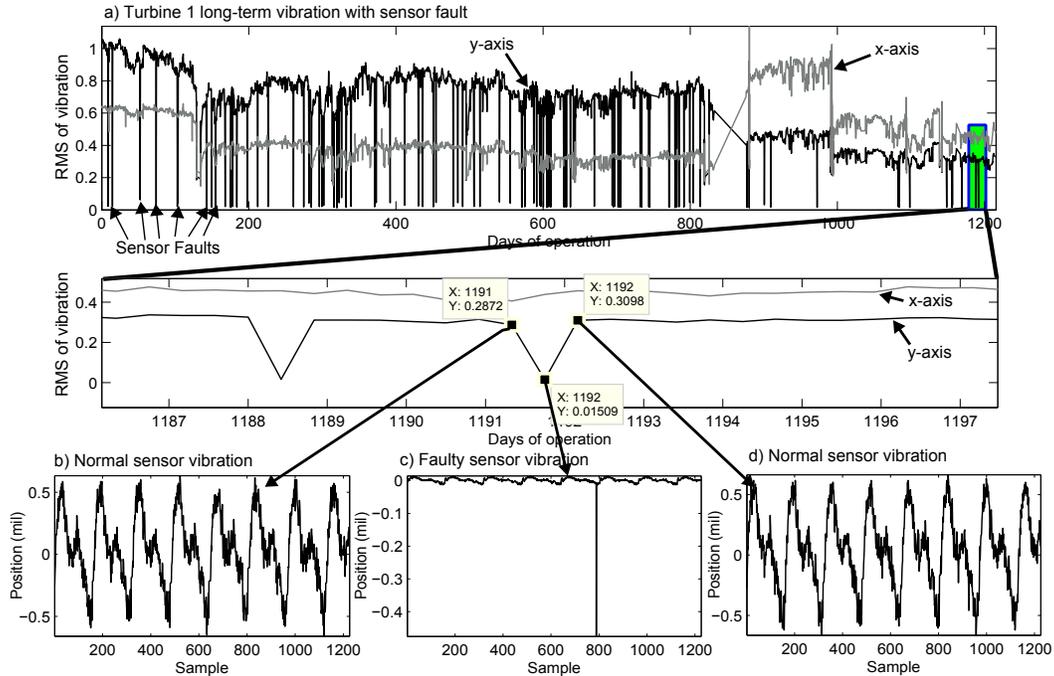


Figure 5.3: Sensor fault on the Y-Axis of Turbine 1. a) The long-term RMS vibration, b) vibration signal just before a sensor fault, c) vibration signal during a sensor fault, and d) vibration signal just after a sensor fault.

5.1.1 Sensor Fault

A clear sensor or data measurement fault is visible in the y-axis vibration sensor of Turbine 1. Figure 5.3a illustrates the RMS of the vibration with indicated sensor faults, and an expanded view of one sensor faults. Figure 5.3b through 5.3d indicate the recorded vibration data before, during, and after the indicated sensor fault respectively. It is clear that the sensor fault data, Figure 5.3c, is measuring the vibration incorrectly despite the x-axis vibration indicating a regular vibration signal. This fault is clearly present over the entire duration of the data, from May 1st 2003 through August 28 2006. The cause of the fault could be a loose wiring connection, internal proximity failure, or something like a data conversion failure. In general, the fault seems to be an occasional but can result in missing poor vibration behaviour of the machine.

5.1.2 Rotor-Stator Rubbing

Analysis of the vibration data by the fault detection methods, Figure 5.4, indicates a time region of particularly high fault level. All methods except MED indicate faulty behaviour in this region. The MED algorithm performance, Figure 5.5, does agree there is a rotor-rubbing fault in the region and the disagreement of the fault indicator is caused by the method spuriously deconvolving a single impulse.

The fault region starts after a month-long machine downtime on August 10 2005, and lasts until it is fixed on a short downtime on January 15th 2006. Figure 5.6 indicates the specific downtimes associated with the introduction and fix of the suspected rotor rubbing fault.

Within the fault region, the processed signals by the various methods are shown in Figure 5.7. All methods indicate a repeating fault once per shaft revolution, which is consistent with a rotor-to-stator rubbing fault. Figure 5.8 indicates the location of the faults with respect to the rotor orbit. It is not necessary that the fault correspond to a maximum in displacement and the rubbing damage will likely not match the angle indicated in the figure because this displacement measurement is not at the source of the impact; it is probable that the rotor tip is located at a different displacement position and angle at this time due to bending dynamics.

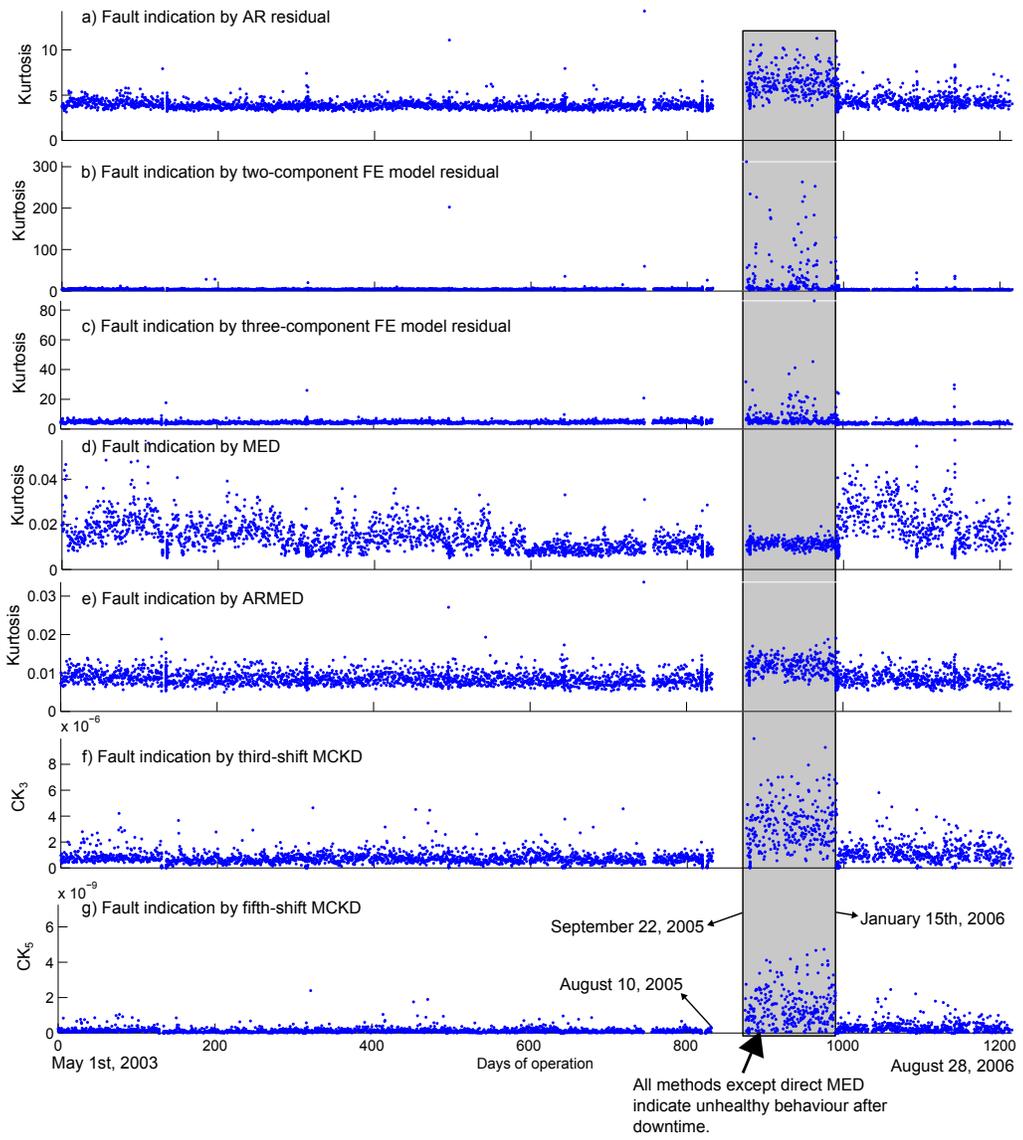


Figure 5.4: Fault indicators on the X-Axis of Turbine 1 vibration data by a) AR, b) two-component FE, c) three-component FE, d) MED, e) ARMED, f) third-shift MCKD, and g) fifth-shift MCKD. The Y-Axis generally agrees with the results, but is messy as a result of the sensor faults.

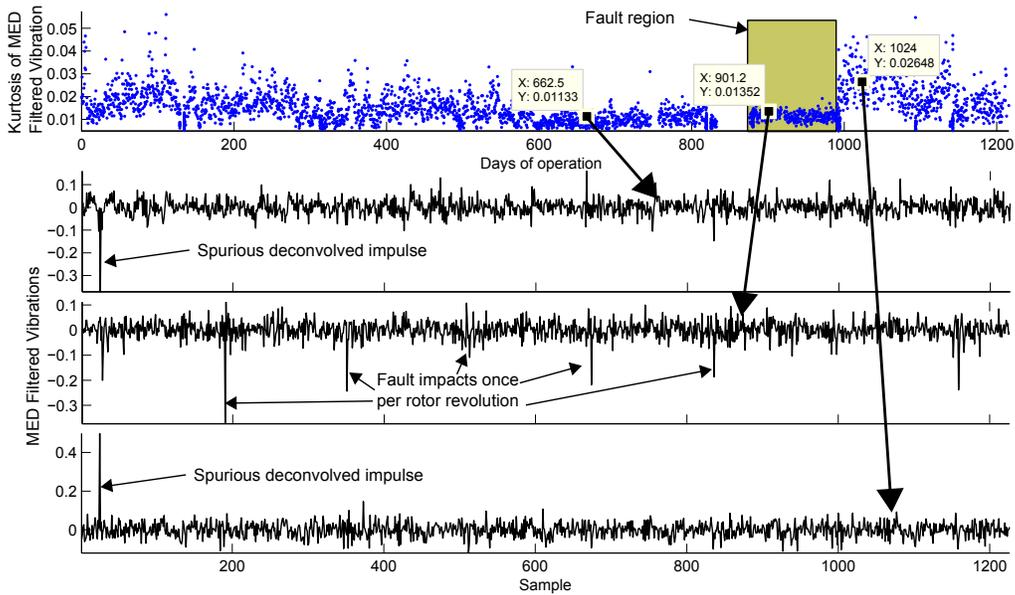


Figure 5.5: Detailed analysis of the MED result. The poor Kurtosis fault indicator performance is as a result of spurious deconvolved impulses, and MED agrees that there is a rotor rubbing fault in the highlighted region.

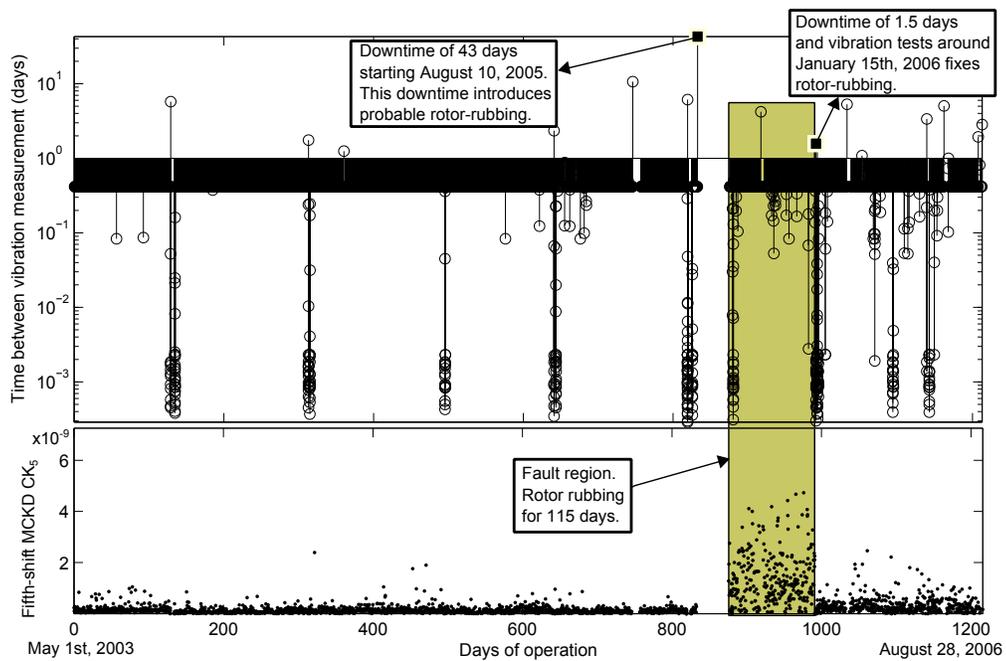


Figure 5.6: Downtime analysis of Turbine 1, with introduction and fix of rotor-rubbing indicated.

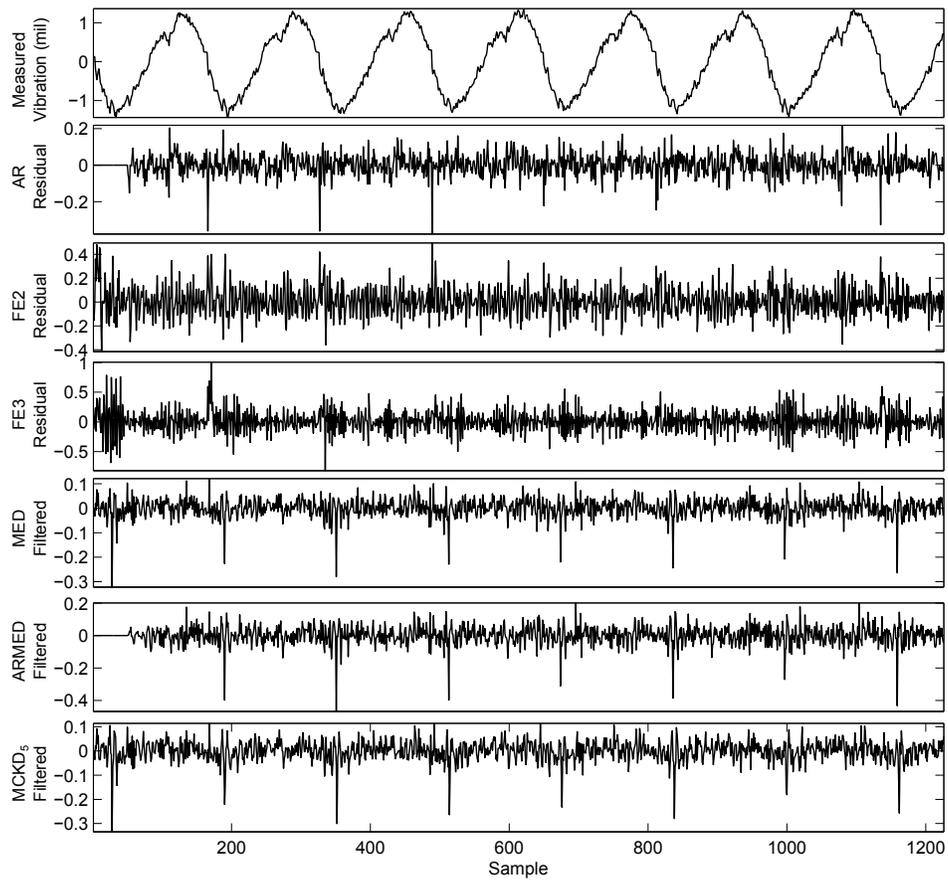


Figure 5.7: The original measured vibration and processed data for a suspected rotor rubbing vibration dataset.

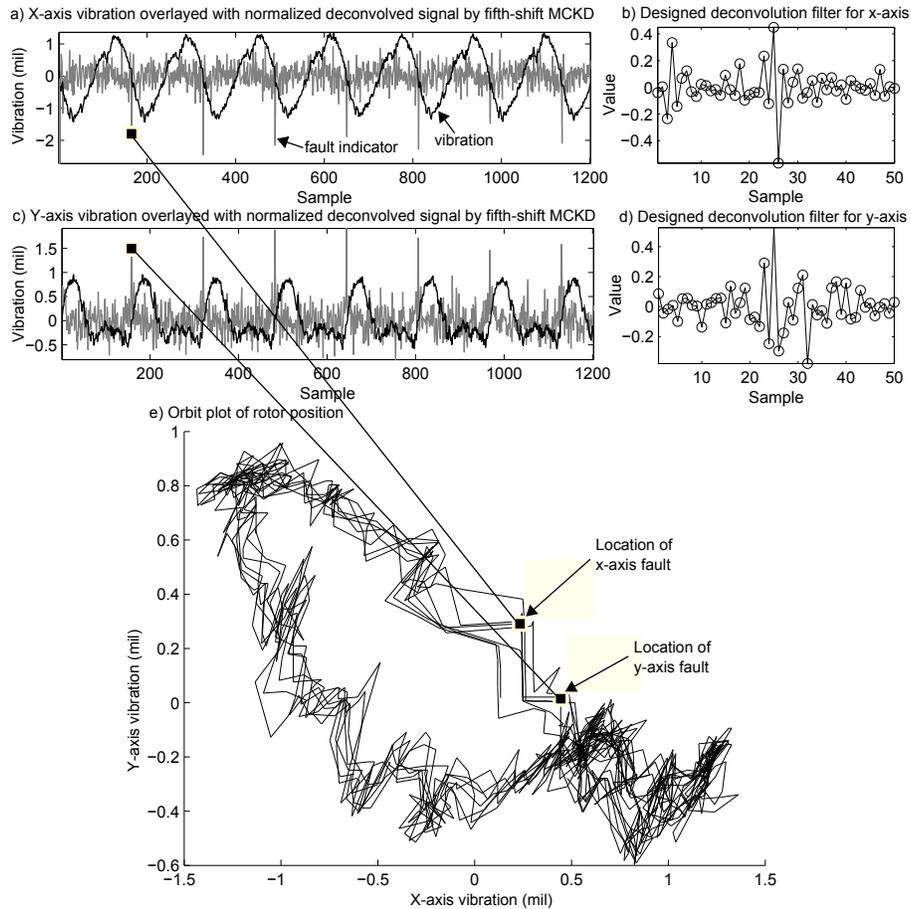


Figure 5.8: The position of the indicated faults according to the orbit map. a) and c) illustrates the x and y-axis vibration with overlaid fifth-shift MCKD deconvolved signal, b) and d) illustrate the resulting deconvolution filters, and e) illustrates the x and y-axis fault locations. The vibration data is taken from Turbine 1 within the suspected rotor rubbing fault region.

5.2 Turbine 2 Health Analysis

Vibration data from January 13th 2003 to August 31 2006 is analysed and it is suspected that there exists a rotor-stator rubbing event from before January 13th 2003 through April 8th, 2004. Figure 5.9 illustrates the fault indicators by method for the duration of the data. It is clear that the machine condition is poorer during the suspected rotor-stator rubbing region. When comparing the RMS vibration of this region, Figure 5.9a, it is clear that the orbit is imbalanced towards the y-axis. The orbit map, Figure 5.10, indicates the deconvolved fault indicator positions marked on the rotor orbit. The faults are indicated at the minimum and maximum of the x-direction displacement, but the actual rotor displacement position at further down the shaft at the rotor-stator position may differ because of the bending dynamics of the rotor.

5.3 Conclusion

Analysis of vibration data from Turbine 1 and 2 revealed important fault information.

Turbine 1 has an occasional sensor fault on the y-axis. This could be caused by a loose connection of the sensor, a problem with the sensor, a problem with the data acquisition, or a problem with the signal processing.

Both Turbine 1 and 2 underwent a long region of suspected rotor-stator rubbing. Rotor-stator rubbing causes damage to both the rotor and stator, which can be extremely costly. Verification of this fault could be performed through periodic visual inspection of rubbing damage, proximity sensor measurement at the stator, or performing temperature measurements. Better monitoring of vibration data is recommended to prevent such a long-running rubbing event from occurring in the future. Implementing monthly or weekly data inspections using the proposed methods addresses this problem and could be implemented in an automated fashion.

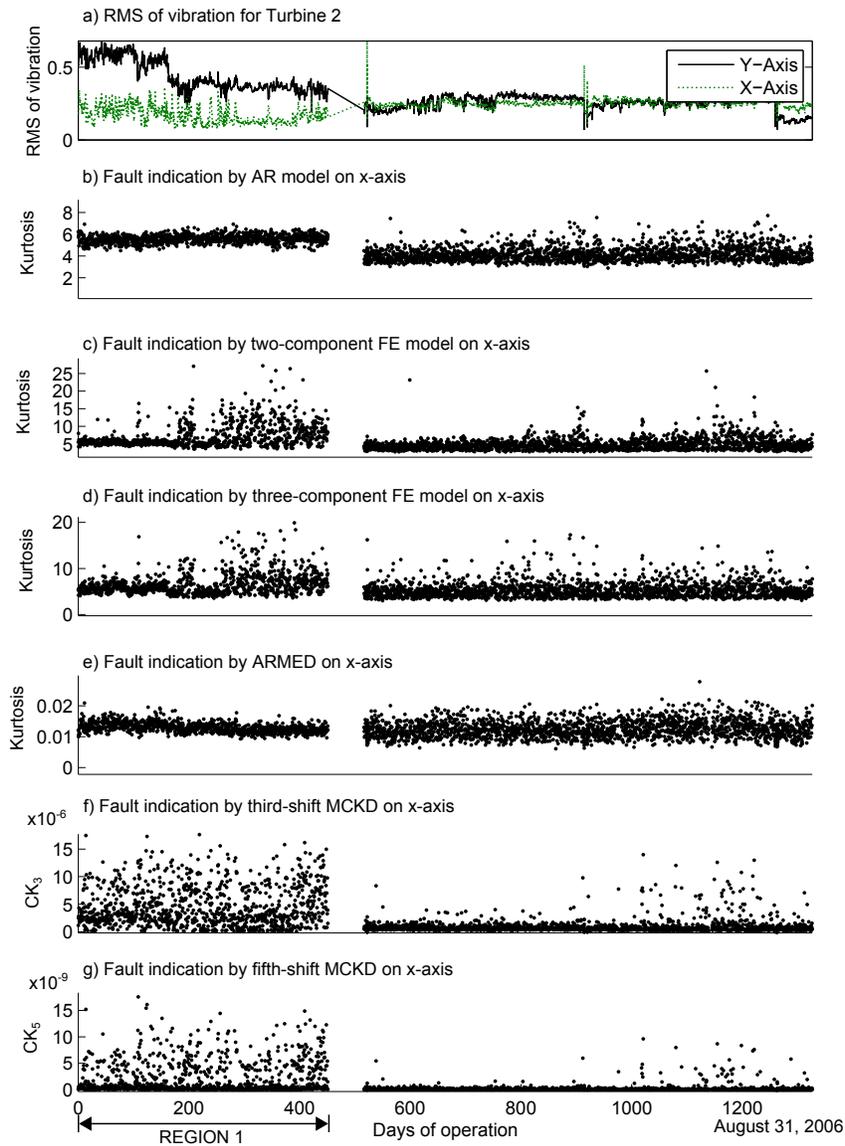


Figure 5.9: The a) RMS vibration and fault indicators on the X-Axis of the Turbine 2 vibration data by b) AR, c) two-component FE, d) three-component FE, e) ARMED, f) third-shift MCKD, and g) fifth-shift MCKD. The Y-Axis generally agrees with the results.

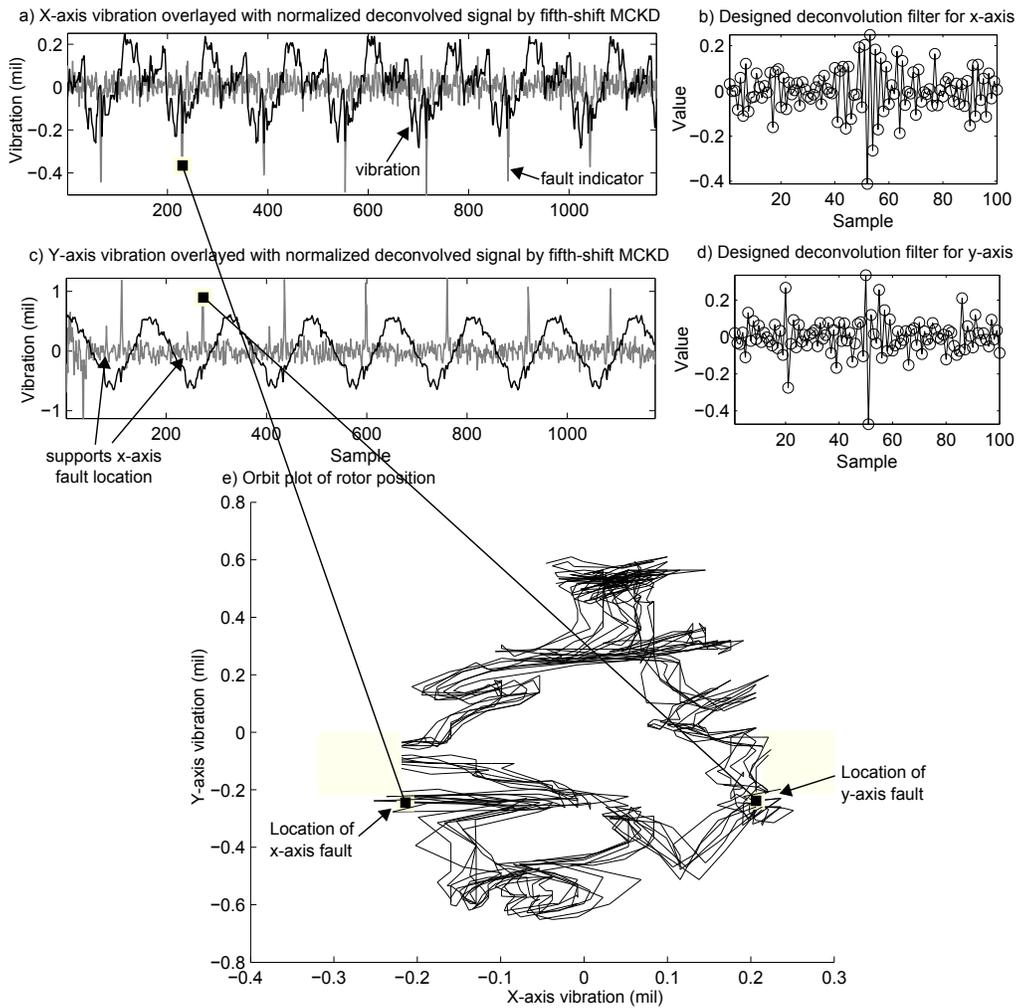


Figure 5.10: The position of the indicated faults according to the orbit map. a) and c) illustrates the x and y-axis vibration with overlaid fifth-shift MCKD deconvolved signal, b) and d) illustrate the resulting deconvolution filters, and d) illustrates the x and y-axis fault locations. The vibration data is taken from Turbine 2 within Region 2 of Figure 5.9.

6

Conclusion and Future Work

6.1 Conclusion

Firstly, in Chapter 3, an adaptive sum-of-sinusoids prediction residual fault detection method is presented. A novel discrete time derivation of the FE model is presented along with a new amplitude invariant adaptive identifier. General parameters for a two and three-component FE model are selected by a LSQ technique. Unlike the AR approach, the FE technique requires no data fitting and works with changing machine speed. The FE method is shown to have similar computational cost as a low order AR model. Based on simulation data, the two-component FE model outperforms the AR model approach in fault detection at higher fault amplitudes. Based on vibration data from a controlled gear tooth fillet crack experiment, it is shown that the proposed method outperforms the AR method. When applied to industrial turbine proximity sensor data, Chapter 5, the FE method agrees with the other techniques as to the location of the suspected rubbing faults. Clearly the proposed FE model is well-suited towards the application of gear faults and has significant advantages over the more accepted AR method.

Secondly, a novel deconvolution technique, MCKD, is proposed in Chapter 4 which expands upon the MED technique by taking advantage of the periodic nature of rotating machine faults. Unlike the MED technique which aims to deconvolve a single impulse from a signal, the proposed MCKD technique aims to deconvolve impulses periodic about a known period. An iterative approach is proposed to the proposed maximization deconvolution problem, and the method is shown to be more effective than MED on both

simulated and experimental data. Because the MCKD method takes in a known period input, concurrent fault detection of different periods is shown to be possible. As a result, this allows for separate fault indicators for different machine components (eg. bearing health, rotor health, and gear health) from a single vibration measurement. This allows for more detailed condition monitoring and identifies the cause of the fault. In Chapter 5, MCKD is shown to be effective in detecting a suspected rotor rubbing fault in turbine proximity sensor data, and indicating the position in the rotor orbit of the rubbing. MCKD is a logical improvement upon MED, which takes advantage of the periodic nature of the faults to provide better results.

On turbine proximity sensor data, two suspected long-running rotor-to-stator rubbing faults are identified. All of the investigated traditional and proposed novel methods agree with each other on these rubbing events. The rubbing location on the rotor orbit is indicated as well as the exact dates for the start and end of the time regions. Improved monitoring of this data is recommended to reduce rotor and stator damage in the future.

6.2 Future Work

For the proposed FE technique, several directions of further research are possible. Firstly, investigating a convergence proof for the model would be of great benefit. Alternative adaptive identifiers could be investigated to achieve better performance. Parameter selection could be investigated through more of an analytic approach, minimum prediction error, or minimum frequency estimation error. Finally, investigation and validation of the method's performance on addition datasets and fault types is important.

Future work on the MCKD method could include iterative convergence investigation, application on bearing inner/outer race fault detection, and fault detection of different datasets. To improve upon the iterative deconvolution technique, a non-linear local minimum solver can be applied to guarantee a local minimum solution at the cost of a longer computation time. The loss of numerical precision when applying higher-order shifts could be resolved through introducing a scaling operation.

Verification of the Turbine 1 and 2 rotor-rubbing should be performed through regular damage inspections if indicated by the fault indicators, and following confirmation the rotor should be rebalanced to reduce the vibration and consequently the rubbing. An automated indication of the rubbing should be developed to notify the engineers of the presence of a fault.



External Resources

Minimum Entropy Deconvolution MATLAB implementation:

<http://www.mathworks.com/matlabcentral/fileexchange/29151-minimum-entropy-deconvolution-med-1d-and-2d>

M-Shift Maximum Correlated Kurtosis Deconvolution MATLAB implementation:

<http://www.mathworks.com/matlabcentral/fileexchange/31326>

Bibliography

- [1] T. Barszcz and R. B. Randall, "Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1352 – 1365, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327008002239>
- [2] W. Wang and A. K. Wong, "Autoregressive model-based gear fault diagnosis," *Journal of Vibration and Acoustics*, vol. 124, no. 2, pp. 172–179, 2002. [Online]. Available: <http://link.aip.org/link/?VAJ/124/172/1>
- [3] F. Harris, E. Kasper, and L. Isler, "Us civil rotorcraft accidents 1963 through 1997," NASA, Ames Research Center, Tech. Rep. NASA/TM2000-209597, USAAMCOM-TR-00-A-006, December 2000.
- [4] W. Wang, "An evaluation of some emerging techniques for gear fault detection," *Structural Health Monitoring*, vol. 2, no. 3, pp. 225–242, 2003. [Online]. Available: <http://shm.sagepub.com/content/2/3/225.abstract>
- [5] H. Endo and 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*, vol. 21, no. 2, pp. 906 – 919, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327006000409>
- [6] M. Yang and V. Makis, "Arx model-based gearbox fault detection and localization under varying load conditions," *Journal of Sound and Vibration*, vol. 329, no. 24, pp. 5209 – 5221, 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X10004359>
- [7] N. Sawalhi, R. Randall, and 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*, vol. 21, no. 6, pp. 2616 – 2633, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327006002603>

- [8] M. Hou, "Estimation of sinusoidal frequencies and amplitudes using adaptive identifier and observer," *Automatic Control, IEEE Transactions on*, vol. 52, no. 3, pp. 493–499, march 2007.
- [9] G. L. McDonald and Q. Zhao, "Model-based adaptive frequency estimator for gear crack fault detection," in *American Control Conference, 2011*, june 2011.
- [10] G. L. McDonald, Q. Zhao, and M. J. Zuo, "Maximum correlated kurtosis deconvolution and application on gear tooth fault detection," in review process for *Mechanical Systems and Signal Processing* journal.
- [11] X. Fan and M. J. Zuo, "Gearbox fault detection using hilbert and wavelet packet transform," *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 966 – 982, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327005002037>
- [12] P. J. Dempsey, "A comparison of vibration and oil debris gear damage detection methods applied to pitting damage," NASA, Glenn Research Center, Cleveland, Ohio, Tech. Rep. NASA/TM2000-210371, September 2000.
- [13] I. Yesilyurt, "Fault detection and location in gears by the smoothed instantaneous power spectrum distribution," *NDT and E International*, vol. 36, no. 7, pp. 535 – 542, 2003. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0963869503000434>
- [14] H. Endo, R. Randall, and C. Gosselin, "Differential diagnosis of spall vs. cracks in the gear tooth fillet region: Experimental validation," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 636 – 651, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327008002203>
- [15] I. Howard, S. Jia, and J. Wang, "The dynamic modelling of a spur gear in mesh including friction and a crack," *Mechanical Systems and Signal Processing*, vol. 15, no. 5, pp. 831 – 853, 2001. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327001914141>
- [16] F. Combet and L. Gelman, "Optimal filtering of gear signals for early damage detection based on the spectral kurtosis," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 652 – 668, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327008002033>
- [17] R. Rubini and U. Meneghetti, "Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings," *Mechanical Systems*

- and Signal Processing*, vol. 15, no. 2, pp. 287 – 302, 2001. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S088832700091330X>
- [18] J. Antoni and R. Randall, “The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines,” *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308 – 331, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327004001529>
- [19] N. G. Nikolaou and I. A. Antoniadis, “Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets,” *Mechanical Systems and Signal Processing*, vol. 16, no. 4, pp. 677 – 694, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327001914591>
- [20] J. Lin and L. Qu, “Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis,” *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135 – 148, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X00928645>
- [21] X. Wang and V. Makis, “Autoregressive model-based gear shaft fault diagnosis using the kolmogorov-smirnov test,” *Journal of Sound and Vibration*, vol. 327, no. 3-5, pp. 413 – 423, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X09005550>
- [22] W. J. Staszewski and G. R. Tomlinson, “Application of the wavelet transform to fault detection in a spur gear,” *Mechanical Systems and Signal Processing*, vol. 8, no. 3, pp. 289 – 307, 1994. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327084710223>
- [23] X. Lou and K. A. Loparo, “Bearing fault diagnosis based on wavelet transform and fuzzy inference,” *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 1077 – 1095, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327003000773>
- [24] X. Tian, J. Lin, K. Fyfe, and M. Zuo, “Gearbox fault diagnosis using independent component analysis in the frequency domain and wavelet filtering,” in *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, vol. 2, april 2003, pp. II – 245–8 vol.2.
- [25] Q. Hu, Z. He, Z. Zhang, and Y. Zi, “Fault diagnosis of rotating machinery based on improved wavelet package transform and svms ensemble,” *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 688 – 705, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327006000306>

- [26] S. Abbasion, A. Rafsanjani, A. Farshidianfar, and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine," *Mechanical Systems and Signal Processing*, vol. 21, no. 7, pp. 2933 – 2945, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327007000349>
- [27] B. A. Paya, I. I. Esat, and M. N. M. Badi, "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor," *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 751 – 765, 1997. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S088832709790090X>
- [28] B. Samanta, K. R. Al-Balushi, and S. A. Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection," *Engineering Applications of Artificial Intelligence*, vol. 16, no. 7-8, pp. 657 – 665, 2003. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0952197603001143>
- [29] C. Chen and C. Mo, "A method for intelligent fault diagnosis of rotating machinery," *Digital Signal Processing*, vol. 14, no. 3, pp. 203 – 217, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1051200403000691>
- [30] N. Gebrael, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: a neural network approach," *Industrial Electronics, IEEE Transactions on*, vol. 51, no. 3, pp. 694 – 700, june 2004.
- [31] R. Huang, L. Xi, X. Li, C. R. Liu, H. Qiu, and J. Lee, "Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 193 – 207, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327005002165>
- [32] Y. Pan, J. Chen, and L. Guo, "Robust bearing performance degradation assessment method based on improved wavelet packet-support vector data description," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 669 – 681, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327008001404>
- [33] C. Bunks, D. McCarthy, and T. Al-Ani, "Condition-based maintenance of machines using hidden markov models," *Mechanical Systems and Signal Processing*, vol. 14, no. 4, pp. 597 – 612, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327000913098>
- [34] W. J. Wang and P. D. McFadden, "Application of wavelets to gearbox vibration signals for fault detection," *Journal of Sound and Vi-*

- bration*, vol. 192, no. 5, pp. 927 – 939, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X96902266>
- [35] Z. Peng, F. Chu, and Y. He, “Vibration signal analysis and feature extraction based on reassigned wavelet scalogram,” *Journal of Sound and Vibration*, vol. 253, no. 5, pp. 1087 – 1100, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X01940854>
- [36] Z. K. Peng and F. L. Chu, “Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography,” *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199 – 221, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S088832700300075X>
- [37] Z. Peng, P. W. Tse, and 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*, vol. 19, no. 5, pp. 974 – 988, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327004000111>
- [38] X. Fan and M. J. Zuo, “Gearbox fault detection using hilbert and wavelet packet transform,” *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 966 – 982, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327005002037>
- [39] L. Zhang, R. Gao, and K. Lee, “Spindle health diagnosis based on analytic wavelet enveloping,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 55, no. 5, pp. 1850–1858, oct. 2006.
- [40] R. Yan, R. X. Gao, and C. Wang, “Experimental evaluation of a unified time-scale-frequency technique for bearing defect feature extraction,” *Journal of Vibration and Acoustics*, vol. 131, no. 4, p. 041012, 2009. [Online]. Available: <http://link.aip.org/link/?VAJ/131/041012/1>
- [41] I. Antoniadis and G. Glossiotis, “Cyclostationary analysis of rolling-element bearing vibration signals,” *Journal of Sound and Vibration*, vol. 248, no. 5, pp. 829 – 845, 2001. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022460X01938155>
- [42] E. Estupian, P. White, and C. San Martin, “A cyclostationary analysis applied to detection and diagnosis of faults in helicopter gearboxes,” in *Progress in Pattern Recognition, Image Analysis and Applications*, ser. Lecture Notes in Computer Science, L. Rueda, D. Mery, and J. Kittler, Eds. Springer Berlin / Heidelberg, 2007, vol. 4756, pp. 61–70. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-76725-1_7

- [43] R. Isermann, "Model-based fault-detection and diagnosis - status and applications," *Annual Reviews in Control*, vol. 29, no. 1, pp. 71 – 85, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1367578805000052>
- [44] J. Liu, D. Xu, and X. Yang, "Sensor fault detection in variable speed wind turbine system using $h - \infty$ method," in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, june 2008, pp. 4265–4269.
- [45] A. C. McCormick and A. K. Nandi, "Cyclostationarity in rotating machine vibrations," *Mechanical Systems and Signal Processing*, vol. 12, no. 2, pp. 225 – 242, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327097901485>
- [46] C. Capdessus, M. Sidahmed, and J. L. Lacoume, "Cyclostationary processes: Application in gear faults early diagnosis," *Mechanical Systems and Signal Processing*, vol. 14, no. 3, pp. 371 – 385, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327099912608>
- [47] R. Randall, "Gearbox fault diagnosis using cepstrum analysis," in *T. of M. and M., Fourth World Congress on*, vol. 1, 1975, pp. 169–171.
- [48] M. E. Badaoui, F. Guillet, and J. Danière, "New applications of the real cepstrum to gear signals, including definition of a robust fault indicator," *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 1031 – 1046, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888327004000123>
- [49] S. L. Marple, Jr., *Digital spectral analysis with applications*, Marple, S. L., Jr., Ed., 1987.
- [50] D. Yu and D. N. Shields, "A bilinear fault detection observer," *Automatica*, vol. 32, no. 11, pp. 1597 – 1602, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0005109896001112>
- [51] W.-H. Chen, D. Ballance, P. Gawthrop, and J. O'Reilly, "A nonlinear disturbance observer for robotic manipulators," *Industrial Electronics, IEEE Transactions on*, vol. 47, no. 4, pp. 932–938, aug 2000.
- [52] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, no. 1-2, pp. 21 – 35, 1978. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0016714278900054>
- [53] A. Gerasimov, O. Morozov, E. Soldatov, and V. Fidelman, "Using the method of modified linear prediction for robust acoustic coding of speech," *Radiophysics and Quantum Electronics*, vol. 49, pp. 535–539, 2006, 10.1007/s11141-006-0086-1. [Online]. Available: <http://dx.doi.org/10.1007/s11141-006-0086-1>

- [54] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.
- [55] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. pp. 431–441, 1963. [Online]. Available: <http://www.jstor.org/stable/2098941>
- [56] X. Fan, Y. Wang, W. Li, S. Wu, and M. Zuo, "Experiment design of gear crack growth," Reliability Research Laboratory, Dept. Mech. Eng., University of Alberta, Tech. Rep., August 2005.
- [57] X. Tian, J. Lin, M. Agelinchaab, J. Zhang, M. Akhtar, M. Zuo, and K. Fyfe, "Technical report for experiment 02-05 on gearbox," Reliability Research Laboratory, Dept. Mech. Eng., University of Alberta, Tech. Rep., 2002.
- [58] Y. Yang and Q. Zhao, "Generator thermal sensitivity analysis with support vector regression," in *American Control Conference (ACC), 2010*, july 2010, pp. 944 –949.