Condition monitoring of rolling element bearings during low and variable speed conditions

Andreas Klausen

Condition monitoring of rolling element bearings during low and variable speed conditions

Doctoral Dissertation for the Degree *Philosophiae Doctor (PhD)* at
the Faculty of Engineering and Science, Specialisation in Mechatronics

University of Agder
Faculty of Engineering and Science
2019

Doctoral Dissertation at the University of Agder 227
ISSN: 1504-9272
ISBN: 978-82-7117-926-7

Printed by Media 07
Oslo

Acknowledgments

This project has been carried out in the period between October 2014 and January 2019 as part of an integrated PhD program that merged the last year of my master's education with the first year as a Research Fellow. The work towards this doctoral thesis has been carried out at the University of Agder, campus Grimstad, and the project is funded by the Ministry of Education in Norway. During this project I have learned a lot, mostly from moving to a new field of research which is not covered of my Mechatronics education. During this time, Prof. Kjell G. Robbersmyr has been my main supervisor. In addition, I have had the supervision of two co-supervisors: Prof. Hamid R. Karimi until December 2016 and Dr. Huynh V. Khang from January 2017.

I want to thank my main supervisor Prof. Kjell G. Robbersmyr for encouraging me in perusing this field of research and for his supervision. I also want to thank Prof. Hamid R. Karimi for his great help with writing papers and setting the theoretical framework of the plan. Finally, I would like to thank my co-supervisor Dr. Huynh V. Khang for his supervision and teaching me a great deal about publishing papers.

Incorporating both theoretical and experimental work has taught me that experiments doesn't always work as theorized, and problems occur. I am grateful for the help from the lab staff at the Mechatronics lab, especially Roy W. Folgerø for aiding in designing the test rig and building it afterwards.

The time spent at the office would not be as great if it weren't for my good colleagues. I have had the pleasure sharing an office with my friends Sondre Tørdal, Rune Husveg, Sondre Nordås and Philipp Pasolli who have always brought an enjoyable atmosphere for work and discussions. I want to thank Sondre Tørdal and Morten Ottestad for taking their time with discussions when I asked for help. I have also had great collaborations with my friends studying condition monitoring of rotating machinery, Surya, Martin and Jagath. Thank you for sharing your knowledge and working with me.

To my parents Håkon, Marianne, and my brother and sister Kristian and Solveig: Thank you for your unconditional love and support. Last, but not least, to my fiancée Veronica, thank you for mending my doubts, enduring my stress and late nights, and for your continuous love and support through my tough times.

Summary

Rotating machinery is common in industry, and the reliability is critical for continued operation and safety, as undetected component faults may force emergency shutdowns. With the increased complexity of machines, spare parts and maintenance crew are often not immediately available for a quick overhaul, which increases downtime. This is especially true when repairing an off-shore wind mill because suitable weather condition is required to board it. Rolling element bearings carry the shaft load using elements that rotate with shaft between two raceways. Despite being a mature technology, bearings are the most common component in rotating machinery to fail because of highly dynamic loading. The expected bearing lifetime is challenging to determine due to the statistical nature of manufacturing quality, and difficulties in modeling the actual operating conditions. Condition monitoring is therefore vital to increase reliability on critical machines. Vibration measurements have been used for several decades to estimate the bearing health state, and the currently available methods can solve most condition monitoring problems. However, low and variable speed conditions can present a problem for these existing techniques, because of low impact energy and non-stationary vibration signals. Such conditions are typical in wind mills which operate with low rotor speed and are subjected to variable wind speeds. A bearing is normally fully functional for a limited amount of time after the initial fault. Estimating this time frame is challenging as the degradation trend must be accurately predicted. However, an estimate of the remaining useful life is beneficial for maintenance planning. Despite being a mature research field, most of the up-to-date techniques require historic failure data to set parameters of prediction models. Such failure data is often not available when installing a condition monitoring system, which render a challenge for estimating the remaining useful life on new machines.

In this project, new techniques are proposed to alleviate the aforementioned challenges in fault diagnosis and prognosis. Two new algorithms are developed for fault detection of bearings operating under low and variable speed conditions. Additionally, a new algorithm is proposed for estimating the remaining useful life on new machines with no historic failure data. The developed methods are described in this thesis, while details are provided in the appended papers.

Publications

The following listed articles have been published or submitted for publication in peer reviewed conference proceedings and journals. The version presented in this thesis differs only in formatting and minor errata.

Paper A A. Klausen, R. W. Folgerø, K. G. Robbersmyr and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Journal of Modeling, Identification and Control*, 38(3):143-156, 2017. doi: 10.4173/mic.2017.3.4.

Paper B A. Klausen, K. G. Robbersmyr and H. R. Karimi. Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum. *IFAC-PapersOnLine*, 50(1):13378-13383, 2017. doi: 10.1016/j.ifacol.2017.08.2262.

Paper C A. Klausen and K. G. Robbersmyr. Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics. *Mechanical Systems and Signal Processing*, 118:226-244, 2019. doi: 10.1016/j.ymssp.2018.08.048.

Paper D A. Klausen, H. V. Khang and K. G. Robbersmyr. Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions. Under review at *Mechanical Systems and Signal Processing*.

Paper E A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings. *IEEE International Conference on Electrical Machines (ICEM)*, Greece, 1912-1918, 2018. doi: 10.1109/ICELMACH.2018.8507056

Paper F A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel RMS Based Health Indicator used for Remaining Useful Lifetime Estimation of Bearings. Under review at *Mechanical Systems and Signal Processing*.

The following articles were written and published during the time of this project but are not included in the thesis.

Paper G S. T. Kandukuri, A. Klausen, H. R. Karimi and K. G. Robbersmyr. A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management. *Renewable and Sustainable Energy Reviews*, 53:697-708. doi: 10.1016/j.rser.2015.08.061.

Paper H S. T. Kandukuri, A. Klausen and K. G. Robbersmyr. Fault diagnostics of wind turbine electric pitch systems using sensor fusion approach. *Journal of Physics: Conference Series*, 1037(3):032036. doi: 10.1088/1742-6596/1037/3/032036.

Paper I M. A. A. Ismail and A. Klausen. Multiple Defect Size Estimation of Rolling Bearings using Autonomous Diagnosis and Vibrational Jerk. *7th World Conference on Structural Control and Monitoring (7WCSCM)*, China, July 2018.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

ADC	analog-to-digital converter
AE	acoustic emission
BPFI	ball pass frequency inner race
BPFO	ball pass frequency outer race
BSF	ball spin frequency
CBM	condition based maintancne
CM	condition monitoring
CPW	cepstrum pre-whitening
CS2	second order cyclostationary
CWR	case western reserve
DE	drive-end
DFR	discrete Fourier transform
DRS	discrete/random separator
FE	fan-end
FIR	finite impulse response
FTF	fundamental train frequency
FT	failure threshold
HI	health indicator
hp	horse power
IMS	intelligent maintenance systems
LSC	low speed conditions
PCA	principal component analysis
PDF	probability density function
PF	particle filter
PHI	physical health indicator
RMS	root mean square
rpm	rounds per minute
RUL	remaining useful life
RVM	relevance vector machine
SOM	self-organizing map
STD	standard deviation
SVM	support vector machine
TSA	time synchronous average
VHI	virtual health indicator
VSC	variable speed condition

Variables

a	crack size
α	modified Paris-Law parameter
β	modified Paris-Law parameter
C	bearing dynamic capacity

D	bearing pitch diameter
d	roller diameter
Δf	frequency step
Δn_c	loading cycles
Δt	discrete time step
$\Delta\theta$	shaft position difference
E	signal energy
F	discrete Fourier transform
$f_{ot}(\theta)$	interpolated vibration signal
F_p	positive-sided frequency spectrum
f_r	shaft frequency
$\hat{R}_i$	FT of R_i
i, j, k	indexes
L_{10}	bearing nominal life
m	$\hat{R}_i$ FT scaling factor
μ_α	mean α
μ_i	mean of R_i during baseline
$\mathcal{N}$	normal distribution
N	vibration signal samples
N_b	number of frequency bands
n_b	samples in each frequency band
n_c	number of cycles
ν_α	process noise
ν_h	measurement noise
p	number of autoregressive model parameters
ϕ	radial plane load angle
P	bearing load
q_j	j'th autoregressive model parameter
R_a	vibration RMS
$\hat{R}_a$	acceleration-based RMS threshold
$\bar{R}$	Mean initial RMS value
ρ_L	lower Spearman threshold
R_i	RMS of frequency band i
R_v	velocity-based vibration RMS
$\hat{R}_v$	velocity-based RMS threshold
σ_a	STD of acceleration-based RMS
σ_α	process noise STD
σ_h	measurement noise STD
σ_i	STD of R_i during baseline
σ_v	STD of velocity-based RMS
$\theta(t)$	shaft position
W	PF weight for RUL estimation
w	particle weight
X	discrete frequncy spectrum
$x(t)$	vibration signal
$x_a(t)$	analytic vibration signal
$x_{env}(t)$	vibration envelope

$x_i(t)$	Hilbert transform of vibration signal	$x_w(t)$	whitened vibration signal
$x_{ot}(t)$	order-tracked vibration signal	$\boldsymbol{z}$	particle
$x_v(t)$	velocity-based vibration signal	Z_b	number of rollers

Chapter 1

Introduction

1.1 Background

Rotating machines have bearings which keep the shaft in place while providing as little friction and noise as possible. Various types of bearings are available, each with their own strengths and weaknesses. Rolling element bearings, plain bearings, magnetic bearings, and fluid film bearings are some commercially available types. Rolling element bearings are frequently used in gearboxes, motors and pumps, because of their ability to withstand high loads at high speeds. There are four main parts in a rolling element bearing. An inner race is fastened to the rotating shaft, while an outer race sits stationary inside a housing. Rolling elements move around the shaft between the raceways, and a cage keeps the elements at a near-fixed distance from each other to better distribute the shaft load and hinder them from crashing into each other.

Different rolling element types are available, and the optimal choice depends on the application. Roller elements are cylinder-shaped and provide a long surface line to distribute the load. Therefore, they can withstand high radial loads, but are not suitable for axial loads due to the regular cylindrical design. Tapered and spherical rollers are more suited for axial loads and misaligned systems but have a more complicated design and are therefore more expensive to manufacture. Additionally, more specialized designs, such as gear bearings and needle bearings, can be found in industry. In this thesis, the focus is on ball element bearings which can commonly be found in electrical machines. These bearings can be used at a high rotational speed, and the ball design permits both axial and radial loads. In addition, ball bearings are one of the cheapest to manufacture given the simplicity of a round element shape.

Oil or grease lubricant must be provided in the bearing to prevent metal-to-metal contact and fill gaps made by imperfections on the surface of rollers and raceways. Without adding a lubricant, the metal-to-metal friction is significant, which results in early

failure. The oil pressure between roller and raceway causes small elastic deformation of the elements. This cyclic deformation will eventually lead to fatigue, and the bearing may fail even if it is perfectly lubricated. The bearing life can be estimated by using the load, speed, temperature, lubricant viscosity, and load capacity of the bearing. The estimated life of a bearing is generally known as L_{10}, which denotes how many million revolutions 10 out of 100 bearings should last if installed correctly and operated under the specified conditions. If a higher confidence is desired, e.g. L_5, the estimated life is drastically reduced Similarly, L_{50} can be almost a magnitude greater than L_{10}. For these reasons, it is difficult to know accurately when a certain bearing will fail. It is similar to throwing a dice, as the outcome of a single throw is improbable to guess, while it is easier to estimate the probability distribution of each outcome if the dice is thrown thousands of times. This statistical nature of bearing life may arise from small differences in material and production quality between batches. If the actual load, temperature, and speed characteristics is dissimilar to the modeled conditions, the life estimation may get quite inaccurate. Lifetime calculators are often available via the bearing manufacturers web portal, and can be used to calculate L_{10}.

Without knowing the health indication of a machine, two maintenance practices can be employed. With a *reactive* maintenance strategy, the components are replaced after failure, which is ideal for non-critical appliances. A *preventive* maintenance strategy aims to replace components after a certain amount of time in use. Such a strategy brings increased maintenance cost, and high probabilities of early failure due to frequent machine overhaul and run-in of new components.

A condition-based maintenance (CBM) strategy aims at identifying the health of components, and replacing the ones that are worn out. This gives an economical advantage as fewer maintenance stops are necessary, and fewer parts are replaced before the end of their useful life. Additionally, knowledge of component health also improves system reliability. Although visual inspection at regular intervals is a normal procedure to assess the component health, this is difficult to perform on a rolling element bearing for a couple of reasons. First, the machine must be stopped and disassembled, and second, microscopic cracks have to be assessed on all rollers and raceways. Detecting tiny cracks can be difficult even in a laboratory setting, which makes this procedure infeasible. Instead, the bearing health can be estimated by measuring the machine states while it is running. Sensors are used to acquire physical states of the system, and signal processing algorithms can be used to detect faults or estimate the level of wear. Different types of sensors and algorithms can be used for this purpose, and the next section elaborates on the state-of-the-art within condition monitoring (CM) systems for bearings.

1.2 State of the Art

1.2.1 Fault diagnosis

Initial bearing faults are typically small surface defects on either the raceway or roller. When a roller passes the defect, an impulse of vibration is produced. This impulse excites the resonance modes of the bearing and supporting structure. To detect this phenomenon, a piezo-electric vibration accelerometer is fastened to the bearing housing, and the output voltage is digitized with an analog-to-digital converter (ADC). The measured voltage signal is proportional to the acceleration, e.g. with a rate of 100 mV/g. A bearing with an outer race fault is shown in Fig. 1.1, and a CM system monitors the vibration signal. The bearing has a surface fault on the outer race that generates vibration when rollers pass, and an accelerometer responds to this vibration. The signal is afterwards processed with a computer to generate results which an operator can analyze to determine the health state.

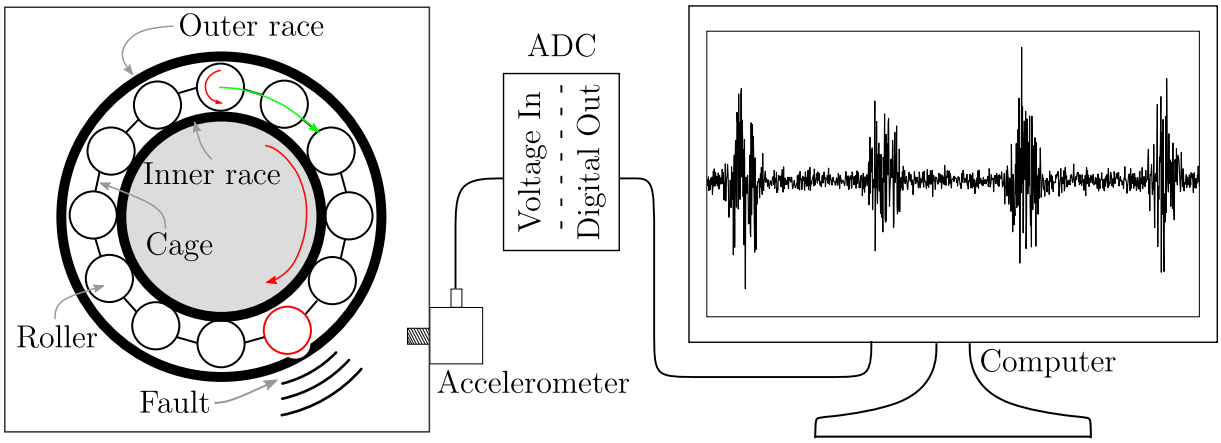

Figure 1.1: Simplified schematic of a bearing condition monitoring system.

The bearing resonance frequency is typically in the thousands of Hz, while the cyclic impact frequency is lower. Analyzing the kinematics of a bearing under the assumption of no slip reveals the cyclic impact frequency. Four characteristic fault frequencies are important for diagnosing a bearing:

$$\text{Fundamental train frequency, FTF} = \frac{f_r}{2}\left(1 - \frac{d}{D}\cos\phi\right) \tag{1.1}$$

$$\text{Ball pass frequency outer race, BPFO} = Z_b\text{FTF} \tag{1.2}$$

$$\text{Ball pass frequency inner race, BPFI} = \frac{Z_b f_r}{2}\left(1 + \frac{d}{D}\cos\phi\right) \tag{1.3}$$

$$\text{Ball spin frequency, BSF} = \frac{f_r D}{2d}\left(1 - \left(\frac{d}{D}\cos\phi\right)^2\right) \tag{1.4}$$

where f_r is the shaft frequency, d is the roller diameter, D is the pitch diameter, Z_b is the number of rollers, and ϕ is the radial plane load angle. Cyclic impacts at either of these frequencies indicate a damaged bearing. The bearing dimensions are shown in Fig. 1.2.

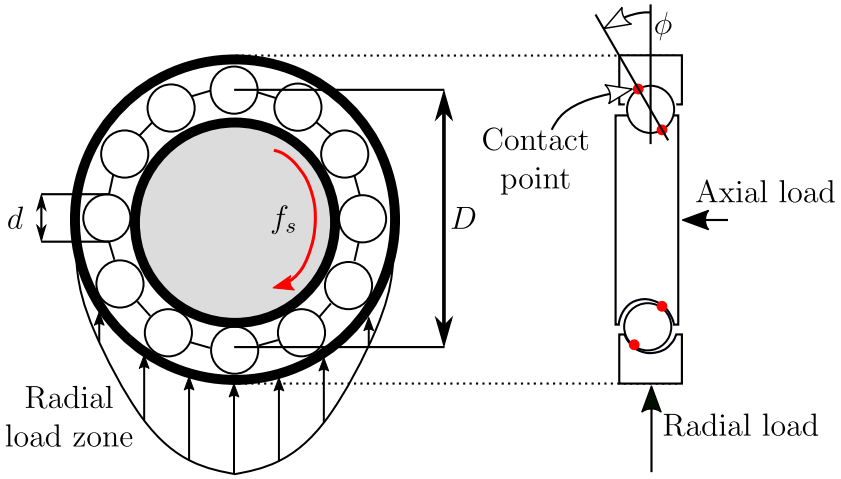

Figure 1.2: Bearing dimensions.

The radial load is spread un-evenly to the bottom half of the rollers, while the top half gets larger clearance in the raceway. Due to the radial load zone, the load angle for the rolling elements may differ, as the bottom half gets an angular contact point, while the top half may slide on the edges due to axial load. In a ball bearing, the roller speed (BSF) is dependent on the load angle, and thus the speed differs between rollers. While the cage is moving at the mean speed of all rollers, some of them must slip on the raceway if moving too fast or too slow. Therefore, there is slight variation in cyclic impact period, and the bearing impact vibration can be modeled as second order cyclostationary (CS2) [1]. On the other hand, deterministic vibration components originating from shafts and gearboxes are phase-locked to the shaft. The bearing vibration can be separated from the deterministic components by using a discrete/random separator (DRS). Time synchronous average (TSA) [2], linear prediction filter [3], and self-adaptive noise cancellation [4] are examples of DRSs. After separating the random components, the signal is denoted as whitened.

The vibration amplitude is often demodulated to extract the cyclic frequencies, and envelope analysis is one of the most successful methods for this purpose [5]. In the early days of bearing vibration analysis, analog rectifiers and low-pass filters were used to achieve the vibration envelope before digitizing the signal [6]. With faster digital technology, the Hilbert envelope has emerged as a new alternative with several advantages over analog rectification [4, 7]. To achieve the Hilbert envelope, the positive-sided frequency spectrum is first inversely transformed to the time domain, yielding the analytic signal. The analytic signal is the original signal plus an imaginary part, which is the Hilbert

transform of the signal. Then, the absolute value of the analytic signal is denoted the Hilbert envelope. Finally, the envelope spectrum is analyzed to detect prominent peaks at integer multiples of the bearing characteristic frequencies, i.e. harmonics.

The vibration signal should be band-pass filtered at the bearing resonance frequency before enveloping to separate the signal of interest from other sources. An inherit advantage of the Hilbert envelope is that the signal can be bandpass filtered in the positive-side of the spectrum, which gives the analytic signal directly. There are several ways of identifying the bearing resonance region. One approach is to track changes in the raw vibration spectrum over time and place the filter over emerging resonance modes [4], but that requires historic data of the system. More recent methods are based on the fourth statistical moment, kurtosis, which has a high value for impulsive signals such as bearing vibration. Spectral kurtosis is a tool for identifying frequency bands with high values of kurtosis [8, 9, 10], which is assumed the optimal frequency band for Hilbert envelope analysis. However, in case of non-Gaussian noise and large single impulses, spectral kurtosis may fail to identify the optimal frequency region. Hence, other methods such as Protrugram [11] and harmonic-to-noise ratio [12] have been proposed to mitigate those issues.

As stated in [4], the aforementioned methods are largely enough to solve most bearing diagnostics cases. However, there are two operating conditions that can pose difficulties when diagnosing a bearing, namely variable and low speed conditions. During variable speed conditions (VSCs), fault impacts no longer occur at a fixed time interval, therefore the cyclic frequencies are spread in the envelope spectrum. Computed order tracking [13] largely resolves the problem by transforming the vibration signal from the time domain to the shaft angle domain. This, however, requires knowledge of the shaft position acquired either by a tachometer or an encoder. Alternatively, the vibration signal can be used to estimate the shaft position by tracking frequency ridges in the time-frequency spectrum [14, 15, 16, 17]. Cepstrum pre-whitening (CPW) is proposed in [18] to whiten the vibration signal during VSC. Deterministic vibration components from shafts and gearboxes are periodic, but not sinusoidal, and therefore form multiple harmonics in the frequency spectrum. In the cepstrum, these harmonics form a single peak at the quefrency of the deterministic component, while cyclostationary signals form no significant peaks at all. By performing a series of liftering operations, the deterministic components are removed by setting the whole real cepstrum to zero, except for at the zeroth quefrency. Afterwards, the envelope spectrum of the order tracked whitened signal should reveal the bearing fault. However, CPW causes spectrum normalization that increases the noise floor, which could mask fault related harmonics. Methods for diagnosing bearings operating under VSC should be further investigated.

Low speed conditions (LSC) pose a second challenge for bearing diagnosis. The impact

energy is depending on the fault size, bearing load, and shaft frequency. With a slower shaft frequency, the impact energy is lower, which makes the impacts harder to detect when the signal is contaminated by noise. To obtain a high-resolution envelope spectrum, the signal must be measured for the duration of several shaft revolutions. At low speed, the data collection time increases drastically, which makes a high impact on memory footprint and required computational power. Some types of signal processing algorithms are not feasible to use if the signal contains too many samples, such as iterative algorithms.

To overcome the issues of low impact energy, several studies have successfully applied acoustic emission (AE) sensors, which are sensitive in the frequency range between 100 kHz and 1 MHz [19]. Such sensors require a high sampling rate ADC to properly sample the signal, which increases the price of the CM system. Multiple studies have shown that signal features such as AE energy [20], AE amplitude [21], AE counts [22], linear prediction filter coefficient values [23], root mean square (RMS) and kurtosis [24] are sensitive to the bearing fault at low rotational speed. Classification of faults using such features is demonstrated in [25] with support vector machine (SVM) and relevance vector machine (RVM). To reduce the amount of data required for fault diagnosis using envelope spectrum, a peak-hold down-sampling technique is proposed in [26], and heterodyne signal enveloping is demonstrated in [27]. A few studies are dedicated to using the vibration signal for diagnosing faults at low speed as well [28, 29]. Most of the referred papers conclude that AE is more suited to detect incipient bearing faults during LSC, compared to regular accelerometers. However, high cost and complexity make AE infeasible for some machines. Therefore, new methods for diagnosing a bearing under LSC with vibration signal should be developed.

Detecting initial faults in a bearing is an important first step towards predictive maintenance. However, in many cases the machine should not stop immediately after detection due to safety reasons or maintenance planning. Bearings can be in operation after the initial fault, and the time until a complete failure is commonly referred to as the remaining useful life (RUL). Prognostic techniques have been developed to estimate the RUL, and the state-of-the-art within prognostics is elaborated on in the next section.

1.2.2 Prognostics

The bearing prognostics procedure can be divided into four major steps: Health indicator (HI) estimation, initial fault detection, model prediction, and failure threshold (FT). A simplified example of this procedure is shown in Fig. 1.3, and the steps are detailed hereafter.

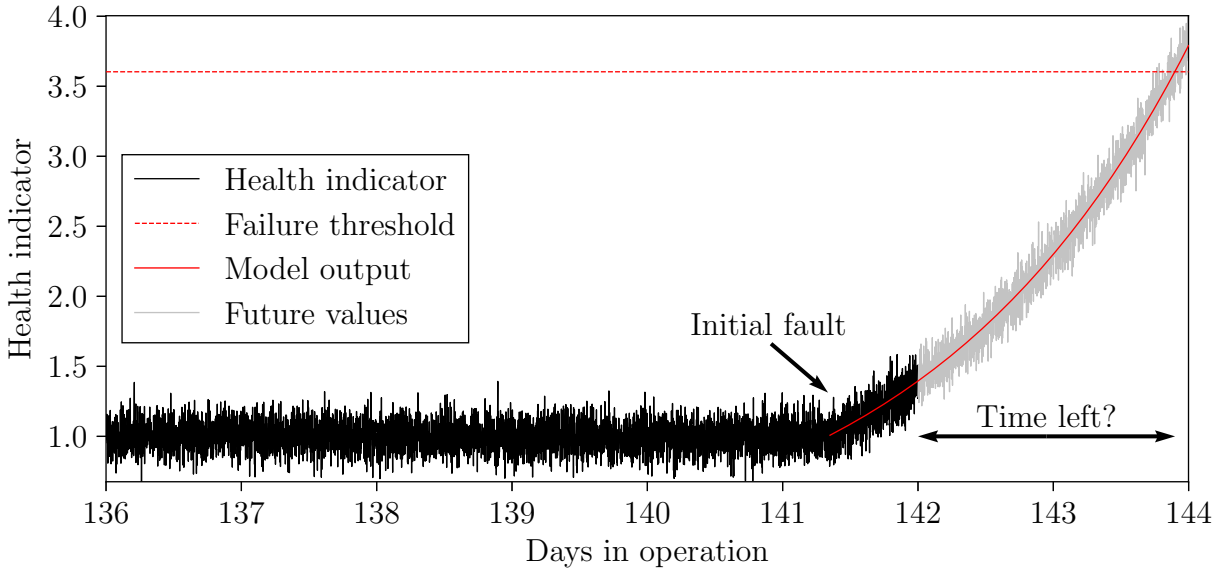

Figure 1.3: Simplified description of bearing prognostics procedure.

Health indicator estimation

The actual health state of a bearing can be assessed by quantifying the level of wear in the bearing. This would require disassembly of the machine and examination of the dismantled bearing, typically by a microscope. This is not feasible for most machines, therefore the actual state is estimated by an HI calculated from sensor data, such as vibration or AE. Generally, HIs can be categorized into physical HIs (PHIs) and virtual HIs (VHIs) [30]. PHIs are generated from signals closely related to the physics of failure, for example the vibration or AE signal. One of the most common PHIs is the RMS [31, 32, 33], as it is very closely linked to the vibration energy. Other PHIs are the kurtosis [34], and the energy ratio between the original signal and the linear prediction residual [35].

VHIs, on the other hand, have no direct correlation with physics of failure. For example, the combination of multiple PHIs using statistical methods or principle component analysis (PCA) is a VHI. In [36], the Mahalanobis distance is used to fuse 14 PHIs into a VHI, and in [37] a self-organizing map (SOM) is used to combine multiple PHIs. More examples of HIs are given in [38].

Initial fault detection

Estimating the bearing RUL can be done after the first fault has developed, and the HI starts to increase. In [31], a threshold on the kurtosis value is used to detect initial degradation. A statistical alarm threshold on the HI itself is also used in some research. For example, the mean and standard deviation (STD) of the HI is obtained with baseline

data, and the initial degradation starts when the HI passes its mean plus 3 times the standard deviation [36, 39].

Model prediction

After passing initial degradation, the future trend of the HI can be predicted to estimate the time until the FT is reached. Such algorithms can be divided into physics-based and data-based methods [40]. Physics-based methods use models that are closely linked to the failure mechanics of the component. Bearing vibration energy is often exponential as an increasing number of cracks results in a greater amount of impacts. Therefore, variants of the Paris-Erdogan crack law [37, 41, 42] have been used to track bearing degradation. Data-based or statistical methods predict the future trend based on historic data alone, without relying on physics of the fault. Some examples are the linear prediction model [43] and the wiener process model [36].

Failure threshold

The prognosis objective is to estimate the time until the chosen HI reaches the FT, and therefore the accuracy depends greatly on the FT value. Depending on the choice of HI, it may be difficult to determine a suitable FT, especially for the more abstract VHIs. In [40] it is stated that a challenge for rotating machinery prognosis is to define FTs for newly designed HIs. Two ISO standards can be used for two PHIs, RMS and peak-to-peak [40, 44], but are not applicable to other HIs. Consequently, most of the recently developed algorithms require historic failure data to determine a suitable FT. This is often not possible to achieve in practice due to lack of proper historic data when installing CM equipment on a new machine. Therefore, new methods for RUL estimation that does not involve historic failure data should be developed.

1.3 Motivation and Problem Statement

This project focuses on three topics under active research. The first topic is LSC bearing fault diagnosis. Most of the identified works have focused on using AE to detect low-energy impacts in the bearing. However, AE equipment is expensive, and accordingly used mostly on large, critical machines. Developing new methods using vibration signal to detect incipient faults in a bearing is desirable for LSC bearing diagnosis.

The second topic is bearing fault diagnosis under VSC in modern drive trains. While computed order tracking and cepstrum pre-whitening have proven to be effective for diagnosing bearings operating under VSC, there is considerable noise floor in the envelope

spectrum due to normalization of the entire vibration signal. Therefore, other methods that can isolate the bearing vibration better should to be investigated.

Finally, the last topic is remaining useful life (RUL) estimation of bearings. The need for historic failure data is common in all the referred literature. Once a CM system is installed on a new machine, it is impractical to require historical failure data before RUL can be estimated. Therefore, the last part of the thesis focuses on developing prognosis algorithms that do not require historic failure data.

1.4 Contributions

This thesis is based on 6 papers that have been published or submitted for publication during the project period.

1.4.1 Paper A: Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition

Summary: Experimental data is required to verify the performance of new diagnostics and prognostics algorithms. Some online repositories have vibration data captured during VSC or constant speed operation. However, vibration data captured during LSC was not available at the time of project initiation. Therefore, we designed and built an in-house test rig to generate our own vibration data. The test rig can apply both axial and radial loads to a small 40 mm bore-diameter 6008-type ball element bearing. The rig is structurally designed to handle long periods of operation to wear out the bearing from healthy to faulty stage. A permanent magnet synchronous motor with a planetary gearbox can run the bearing at low speed down to 20 rounds per minute (rpm) and variable speed up to 500 rpm. The test rig controller manages the entire test during the bearing life, and handles regular sensor data logging, motor control, and bearing health check. In other words, the operator does not have to be present and the accelerated life tests can run 24/7.

Contributions: The new test rig design allows for both axial and radial loads at the same time, while being able to gather sensor data at low and variable speed settings. The acquired data is used in all the appended papers in this thesis.

This paper has been published as:
A. Klausen, R. W. Folgerø, K. G. Robbersmyr and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Journal of Modeling, Identification and Control*, 38(3):143-156, 2017. doi: 10.4173/mic.2017.3.4.

1.4.2 Paper B: Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum

Summary: Analyzing vibration spectra for resonance regions and characteristic fault harmonics is time consuming, and false alarms or erroneous conclusions may be drawn. This paper presents an automatic method for bearing fault diagnosis. First, the vibration spectrum is split into all resonance regions by use of bandpass filters. For each resonance region, an automatic peak detection algorithm searches for harmonics of characteristic fault frequencies in the envelope spectrum. For every identified harmonic, a fault score level increases, and a higher score signifies greater confidence in fault detection. Experimental results show how all three major fault types are detected during a single accelerated lifetime test.

Contributions: In the direction of autonomous fault detection, the paper presents a method for automatic detection of faults in a bearing.

This paper has been published as:
A. Klausen, K. G. Robbersmyr and H. R. Karimi. Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum. *IFAC-PapersOnLine*, 50(1):13378-13383, 2017. doi: 10.1016/j.ifacol.2017.08.2262.

1.4.3 Paper C: Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics

Summary: Bearing fault diagnosis using vibration signal during LSC is a challenge due to low impact energy. This paper presents a new method for extracting more useful information of the vibration signal. First, the bearing vibration is isolated with a discrete/random separator. Afterwards, the envelope spectrum is correlated with the whitened vibration spectrum to combine useful information in both spectra. The reasoning is that fault related harmonics may be identifiable directly in the raw vibration spectrum, and the cross-correlation may improve strength of fault related harmonics. Experimental results at 20 rpm verifies that the presented method significantly improves the diagnosis capabilities compared to other methods reported in the literature.

Contributions: A new computationally efficient method for extracting more fault related harmonics of a vibration signal.

This paper has been published as:

A. Klausen and K. G. Robbersmyr. Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics. *Mechanical Systems and Signal Processing*, 118:226-244, 2019. doi: 10.1016/j.ymssp.2018.08.048.

1.4.4 Paper D: Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions

Summary: This paper presents a new method for identifying bearing resonance regions for bearings operating in VSC. It is known that order tracking makes shaft-speed dependent vibration components de-spread in the frequency spectrum. At the same time, bearing resonance is time dependent, and order tracking therefore causes a spread of resonance energy. The presented method exploits these phenomena by first normalizing the vibration spectrum after order tracking by means of CPW. Afterwards, the inverse order tracking is applied to de-spread and highlight resonance regions. Experimental results from three test rigs show that the proposed method can highlight multiple resonance bands, and that the envelope spectra contain numerous harmonics related to the fault. Compared to two other methods reported in literature, the fault diagnosis score is higher.

Contributions: A new method to identify all resonance regions in a vibration signal when operating under VSC.

This paper has been submitted as:
A. Klausen, H. V. Khang and K. G. Robbersmyr. Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions. Under review at *Mechanical Systems and Signal Processing*.

1.4.5 Paper E: Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings

Summary: ISO standard 10816-3 defines allowable RMS vibration levels in the position and velocity domains which may be used as an FT for RUL estimation. However, initial bearing faults cause high-frequency resonance vibration, and integration from acceleration to velocity reduces the effect of this component. Therefore, an early degradation trend is difficult to predict using the velocity- or position-based RMS. This paper presents a method for transforming the RMS threshold to the acceleration domain. Experimental results show that the acceleration-based RMS allows for longer prediction time, and that the transformed failure threshold is valid compared to final values.

Contributions: Two new failure thresholds for vibration signals that does not require any historic failure data.

This paper has been published as:
A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings. *IEEE International Conference on Electrical Machines (ICEM)*, Greece, 1912-1918, 2018. doi: 10.1109/ICELMACH.2018.8507056

1.4.6 Paper F: Novel RMS Based Health Indicator used for Remaining Useful Lifetime Estimation of Bearings

Summary: The vibration RMS is a good health indicator due to its connection to the signal energy. However, the trend is often non-stationary, which makes it difficult to predict future values. This paper presents a novel method for splitting the vibration signal into multiple frequency bands for RMS calculations. Without using digital filters, the proposed method splits the signal energy into multiple frequency bands using a single discrete Fourier transform. The RMS is calculated for each band, and suitable RMS trends for RUL estimation are identified online. Failure thresholds are calculated by extending the method presented in Paper E. A particle filter combined with the Paris-Erdogan law is applied to predict the RUL using suitable RMS trends. The method is verified with experimental results from two test rigs.

Contributions: A novel method for subdividing the vibration signal into multiple frequency bands for RMS calculations. A method for predicting remaining useful life without using historic failure data.

This paper has been submitted as:
A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel RMS Based Health Indicator used for Remaining Useful Lifetime Estimation of Bearings. Under review at *Mechanical Systems and Signal Processing*.

1.5 Outline

The rest of this thesis is divided into four main chapters. Chapter 2 contains brief descriptions of the datasets used in this thesis, which includes the in-house test rig and downloaded datasets from open repositories. Next, Chapter 3 is devoted to the contri-

butions within fault diagnosis during LSC and VSC. The chapter starts with a short background on existing techniques before elaborating on the contributions made from three papers written during the project. Chapter 4 involve the estimation of remaining useful life for bearings. In particular, contributions from the two last papers are presented. Finally, Chapter 5 concludes the thesis and presents the reader with limitations and possibilities for further work.

Chapter 2

Experimental test data

This chapter briefly describes the bearing vibration datasets used to generate experimental results. Section 2.1 presents an in-house test rig that was built during the project period, and the section is based on Paper A. Datasets from other sources have also been used to verify the performance of the developed methods, and these datasets are described in Section 2.2.

2.1 In-house test rig - Paper A

The test rig was designed and built at the University of Agder during the project period to conduct accelerated lifetime tests of bearings from healthy to faulty conditions. A brief description of the test rig is presented here, and more details are given in Paper A. Fig. 2.1 shows a schematic of the test rig, and Fig. 2.2 shows a picture of the finished test rig in the laboratory.

A permanent magnet motor drives the shaft with 4 bearings. The shaft can be operated with stable speeds between 20 and 500 rpm using a variable frequency drive and a planetary gearbox. Continuous variable speed operation is also permitted using this setup. A sealed 6008 type test bearing is installed in a custom housing on the left end of the shaft. Two electric linear actuators apply axial and radial loads to the test bearing, and Each actuator is connected to a lever arm to amplify the load, and load cells are used as rotational points to measure the applied loads. To the right in Fig. 2.1, the red lever arm is pushing on an axial bearing to generate up to 12 kN axial load. To the left, the second red lever pushes the test bearing housing upwards, producing up to 12 kN radial load. The two larger bearings are installed in the center to stabilize the shaft and bear the radial loads. A piezo-electric accelerometer is mounted on the test bearing housing to measure the vibrations. In addition, shaft movement is measured with an eddy current proximity sensor, and the bearing temperature is determined with a PT100 element. The

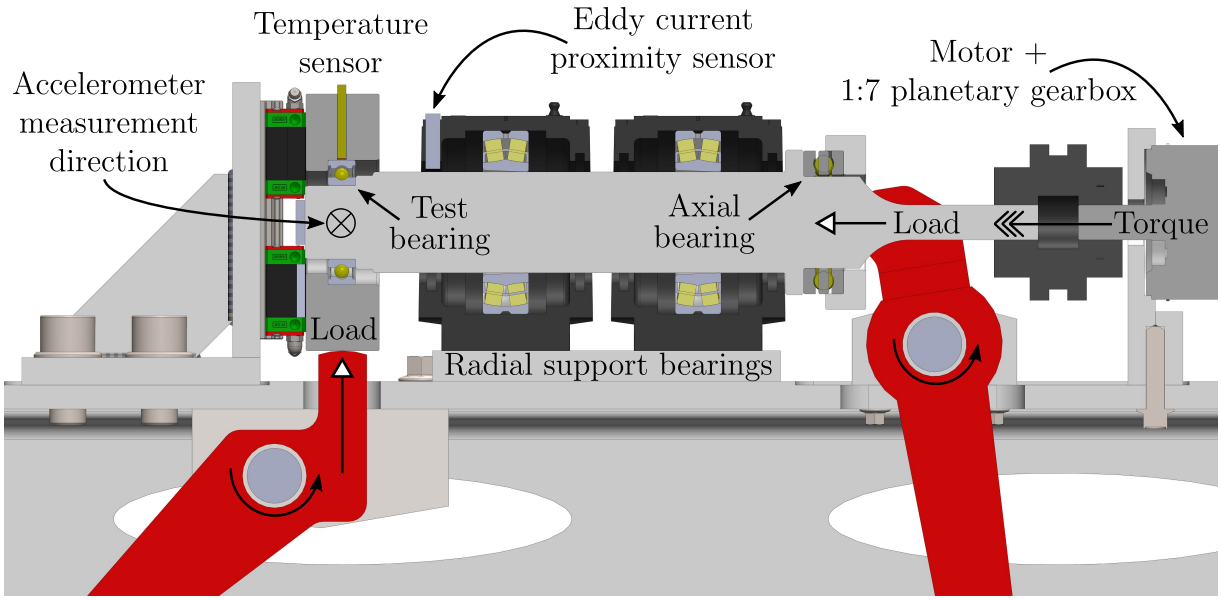

Figure 2.1: Schematic of the accelerated bearing life-time test rig.

proximity sensor data is acquired to verify the vibration analysis, as both sensors are recorded at the same time. In this project, datasets from three run to failure tests have been used. Table 2.1 shows the test conditions and which papers the datasets have been used in.

Table 2.1: Run to failure experiments using the in-house test rig.

Paper	Speed [rpm]	Radial load [kN]	Axial load [kN]	Lifetime [10^6 rounds]
B	250	9	4	42.7
A	50	9	7	6
C	20	9	7	6
E	100	9	7	6
D	50±35	9	5	35.7
F	100	9	5	35.7

Experimental results used in Paper B originates from the first successful attempt to damage the bearing, and this particular test was conducted before the planetary gearbox was installed. It was, however, observed that the measurements acquired during LSC below 100 rpm suffered from too large speed deviations due to variable friction and speed controller overshoots. Therefore, the planetary gearbox was installed for the remaining experiments to stabilize low speed operation. The second dataset used in Papers A, C and E included low speed measurements down to 20 rpm. Variable speed operation was also included in the last test, and datasets from this experiment were primarily used in

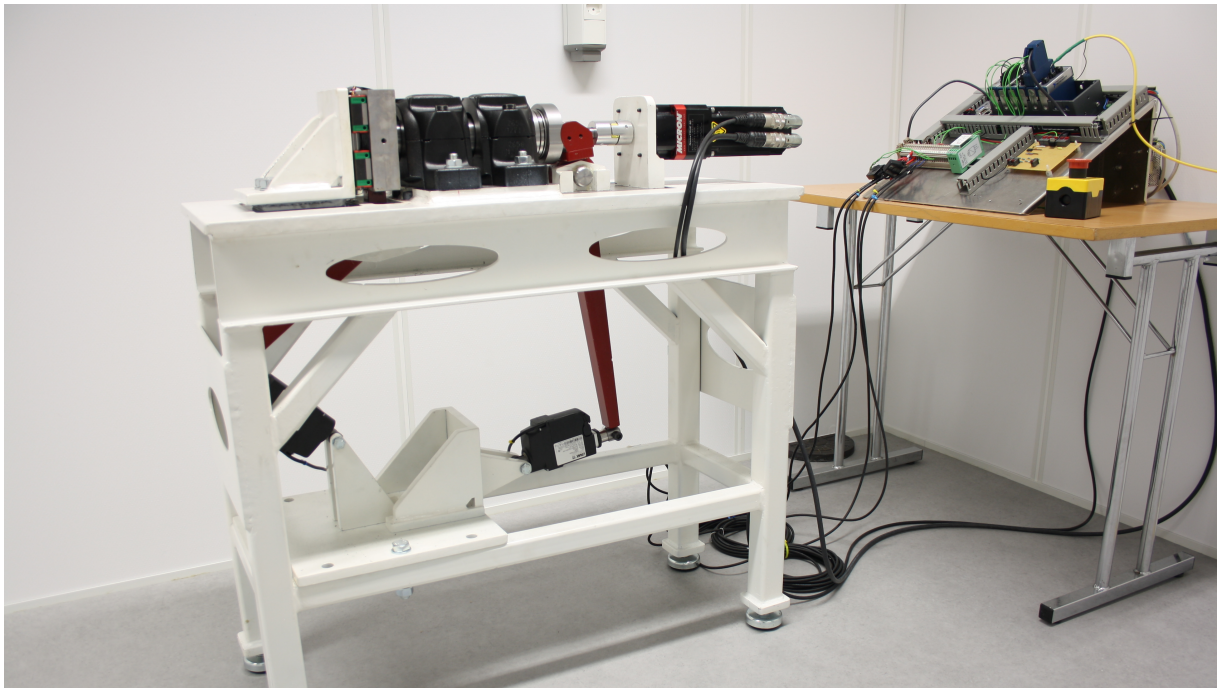

Figure 2.2: The finished test rig in the laboratory at the University of Agder.

Papers D and F. The surface faults on two of the disassembled bearings are shown in Fig. 2.3.

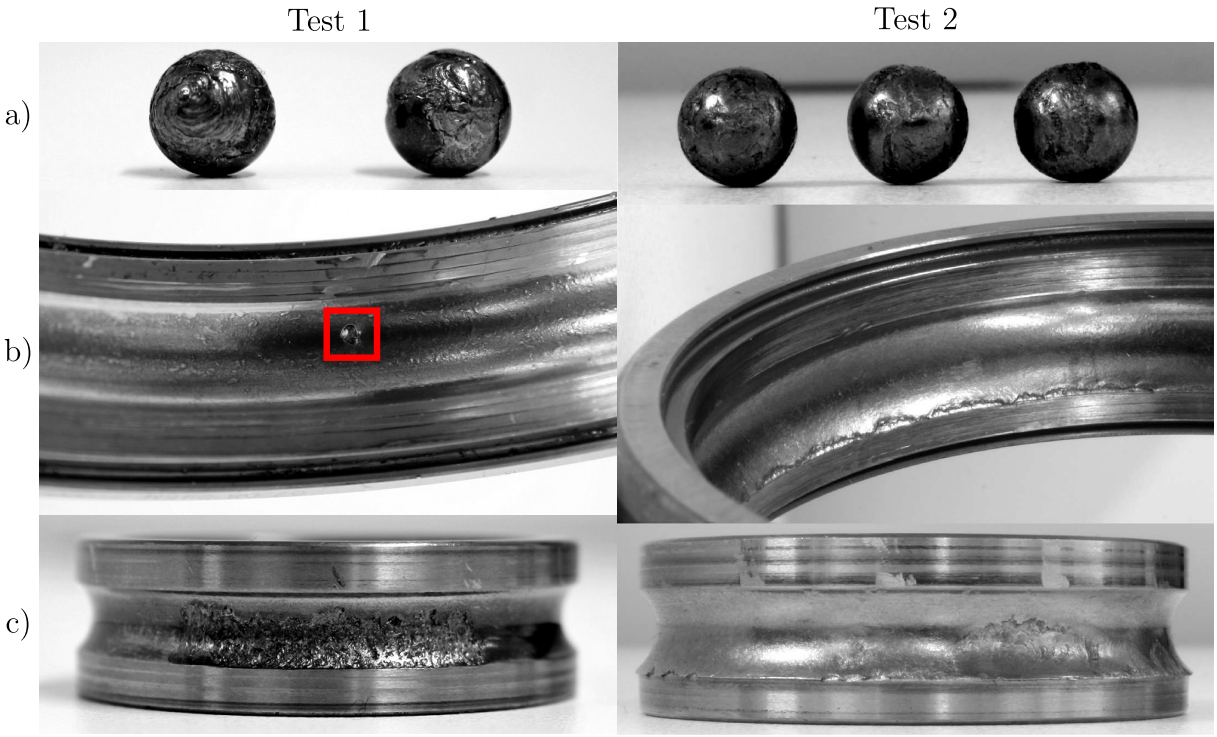

Figure 2.3: Damaged bearings after two accelerated life time tests. (a) Rollers; (b) Outer race; (c) Inner race.

2.2 Downloaded datasets

The datasets from three other sources have been used extensively in this project as a second comparison with in-house datasets.

2.2.1 Variable speed datasets

The datasets collected by Mishra et al. [45] have been applied for variable speed fault diagnosis. The test rig features a variable speed motor, bearings with artificially created faults, and a small radial load of 5 kg. A laser vibrometer measured the vibration at a sample rate of 50 kHz for $\approx$ 20 seconds for each dataset. The shaft encoder provided readings of the shaft position during the measurements. Three types of faults were tested, namely inner race, outer race and roller fault. During the measurements, the shaft speed was adjusted manually between 600 and 900 rpm. These datasets are extensively used in Paper D.

2.2.2 Accelerated lifetime datasets

The datasets generated by NSF I/UCR Center for Intelligent Maintenance Systems (IMS) have been used for testing prognostics algorithms. These datasets were first used in [46] and consists of three run-to-failure tests. The test rig has an induction motor running a shaft with 4 bearings at a constant speed of 2000 rpm. Radial load of $\approx$ 26700 N is applied to wear out the test bearings. For each test, the vibration at each bearing housing was recorded for 1 second at a rate of 20480 Hz, every 10 minutes. An oil circulation system lubricates the bearings, and a magnetic plug triggers a stop signal when enough metal debris is accumulated. These datasets are used extensively in Paper E and F.

2.2.3 Constant speed datasets

Case Western Reserve (CWR) University has an open database with bearing fault test data [47]. The test rig has an induction motor connected to a dynamometer as a load torque. Bearings with pre-seeded faults are installed on either the drive-end (DE) or fan-end (FE) for each test. For each fault and location combination, the vibration signal was measured at DE, FE and the motor base (BA). Vibration was sampled at a rate of 12 kHz for FE faults, while measured at both 12 and 48 kHz for DE faults. There was no radial load, but the torque load varied between 0, 1, 2 and 3 horse power (hp). While unloaded, the shaft speed was 1797 rpm, and 3 hp load caused the induction motor to run at a lower speed of 1730 rpm. These datasets are used extensively in Papers C and E.

Chapter 3

Fault diagnostics

This chapter presents three new methods for fault diagnosis developed during the project period. First, some relevant background used in the developed methods is given in Section 3.1. Afterwards, the three developed methods with results are described. Section 3.2 is based on Paper B and introduce a method for automatic bearing fault detection. A new method is developed in Paper C for extracting more fault-related information from the vibration signal during LSC, and the method is described in Section 3.3. Paper D demonstrates a new method for identifying resonance frequencies in the vibration signal while operating under VSC, and Section 3.4 covers this method briefly. The papers are appended to provide more detailed information about the proposed methods.

3.1 Theoretical background

The developed methods take advantage of background theory that is given in this section. Section 3.1.1 explain order tracking which is used to transform the vibration signal from time domain to shaft angle domain. In Section 3.1.2, three methods for separating random components from the vibration signal are disclosed. Next, Section 3.1.3 shows how the envelope spectrum is calculated. Finally, Section 3.1.4 explains how the envelope spectrum can be analyzed to detect and classify bearing faults.

3.1.1 Order tracking

The vibration signal is normally sampled at a constant rate, for example 51200 samples per second (Hz). A fixed number of samples between bearing fault impacts is recorded if the shaft is rotating at a constant speed. However, if there are changes in the shaft speed during the measurement, the impact frequency is no longer constant. Fourier techniques assume stationary frequency components in the signal, and this assumption fails under VSC. Therefore, the energy of cyclic components is spread over a larger frequency area,

which makes diagnosis difficult. Order tracking transforms the vibration signal from the time domain to the shaft angle domain [13], which results in a fixed number of samples between impacts. The shaft position or speed must be recorded along with the vibration measurements to perform order tracking. An encoder or tachometer is normally used to acquire the shaft position.

Let $x(t)$ be the vibration signal sampled in the time domain, and x_n is the same signal discretized at time index n. The shaft position signal is given by $\theta(t)$ and assumed to be sampled at the same rate as the vibration signal. A cubic spline is used to describe the vibration signal between the discrete samples. This is given as

$$f_{ot}(\theta) = \text{interpolate}(\theta(t), x(t)) \,, \tag{3.1}$$

where interpolate$(\theta(t), x(t))$ is the desired interpolation function (cubic spline) that is used to approximate the function $f_{ot}(\theta) = x(t)$. To order track the signal, vibration samples at a fixed $\Delta\theta$ period are calculated from $f_{ot}(\theta)$. The resulting order-tracked vibration signal is $x_{ot}(\theta)$. The x-axis of an order tracked vibration spectrum is given in orders, where the 1st order is the same as the shaft frequency. More details on this method are given in the appended Paper C, Section C.2.2.

3.1.2 Signal whitening

Bearing vibration is slightly random and is therefore modeled like a second order cyclostationary (CS2) signal as described in Section 1.2.1. Shaft and gearbox vibration components are phase-locked to the shaft and are therefore deterministic. Random and deterministic components can be separated using a DRS. For bearing fault diagnosis, the random signal is retained for further analysis. A whitened vibration signal contains mostly random, uncorrelated components such as white noise. Therefore, the random signal output of a DRS is a whitened signal. In this project, three types of whitening approaches are utilized.

The time synchronous average (TSA) is used to remove vibration components that are phase-locked to the shaft. The vibration signal is split into segments of equal lengths, each containing the vibration of one shaft rotation. Then, the mean of these segments is the average vibration occurring per shaft revolution, i.e. the TSA. By subtracting this average from of each shaft revolution of the original signal, the result is a slightly whitened vibration signal $x_w(t)$. If the signal is order tracked, the angle synchronous average is determined instead. More details are given in the appended Paper C, Section C.2.3.

An autoregressive model is trained to predict the next sample based on a linear

combination of p previous samples. This model is formulated as

$$x_{n+1} = x_{w,n+1} - \sum_{j=1}^{p} q_j x_{n-j-1} \,, \tag{3.2}$$

where $x_{w,n}$ is the n'th sample of the whitened vibration signal (model residual), and q_j is the j'th model parameter. Deterministic signal components can be predicted based on previous samples, while random components are left in the residual. The model parameters are identified using the Yule-Walker equations which is a least-squares approach involving the autocorrelation of the vibration signal [48, 49]. The model parameters are determined for each signal, and the whitened vibration signal is extracted using the trained model. More details are given in the appended Paper C, Section C.2.4.

Cepstrum pre-whitening [18] is an effective method that utilizes the cepstral domain to whiten the signal. Vibration components from shafts and gearboxes are periodic, but not sinusoidal. Therefore, the components make up multiple harmonics in the frequency spectrum. In the cepstral domain, these harmonics form a single peak at the quefrency equal to the period of the deterministic component. By performing a series of liftering operations, the deterministic components are mitigated by setting the whole real cepstrum to zero, except for the zeroth quefrency [50, 51]. This is simplified by [18]

$$x_w = F^{-1}\left(\frac{F(x(t))}{|F(x(t))|}\right) \,, \tag{3.3}$$

where F and F^{-1} are the forward and inverse discrete Fourier transform, respectively. Detailed information are given in the appended Paper D, Section D.6.2 and [18, 50, 51].

3.1.3 Envelope spectrum

The cyclic frequencies are extracted by demodulating the whitened vibration signal with the Hilbert envelope. The analytic signal is acquired by taking the inverse Fourier transform of the positive-sided frequency spectrum. If $F_p(x(t))$ is the positive-sided frequency spectrum, then the analytic signal is

$$x_a(t) = x(t) + ix_i(t) = F^{-1}(F_p(x(t))) \,, \tag{3.4}$$

where $x_i(t)$ is the Hilbert transform of $x(t)$. Afterwards, the envelope $x_{env}(t)$ is calculated with the absolute value as

$$x_{env}(t) = |x_a(t)| \,. \tag{3.5}$$

If, however, the signal is order-tracked, $x_{env}(\theta)$ is determined instead. Taking the Fourier transform of x_{env} returns the envelope spectrum.

3.1.4 Fault detection

The envelope spectrum can be analyzed to detect faults in the bearing as a faulty bearing will produce amplitude-modulated vibration signal. Fig. 3.1 shows the envelope in the time and frequency domain for three different fault types. The envelope of resonance

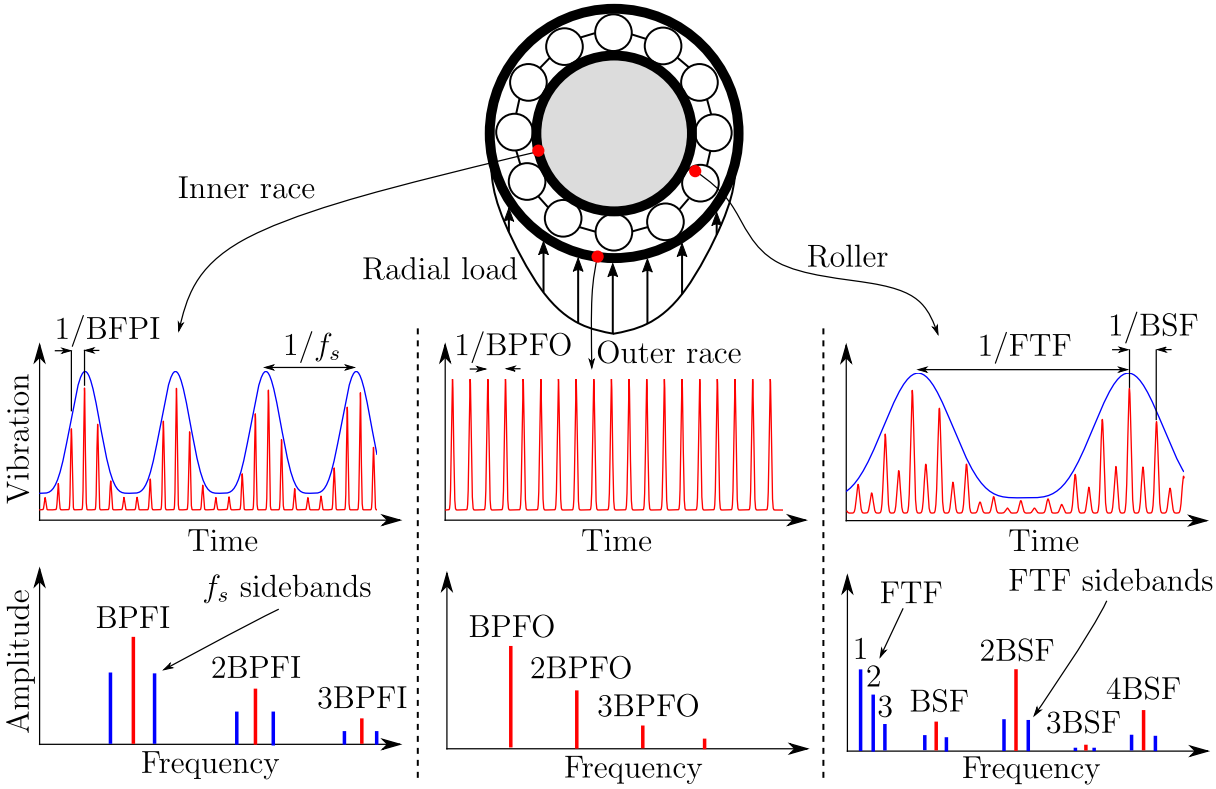

Figure 3.1: Envelope in time and frequency domain for the three fault types.

vibration impulses is periodic, but not sinusoidal, resulting in several harmonics in the spectrum. The side-bands appear due to amplitude modulation when the fault moves in and out of the radial load zone. The envelope spectrum is examined for the frequency domain characteristics in Fig. 3.1 to diagnose the bearing. If the spectrum does not contain any prominent harmonics or side-bands related to the fault, the bearing is either healthy, or the vibration signal contains too much noise.

3.2 Automatic fault detection - Paper B

Each fault impact causes resonance vibration in the bearing itself and the surrounding structure. It is beneficial to isolate the resonance frequency vibration before making the envelope spectrum to improve fault detection. However, identifying the resonance frequency requires a detailed model of the system, which is not suitable to make for all systems. Paper B presents a method for isolating resonance modes in the vibration

signal and make an envelope spectrum for each of them. Manual analysis of several envelope spectra is time consuming, therefore an automatic fault detection algorithm is also presented in this paper.

A dataset from the in-house test rig is used to demonstrate the automatic fault detection algorithm. During recording of this dataset, the bearing had an inner race fault, and the motor ran at 250 rpm. The vibration signal is first order-tracked, and the TSA is removed. The next step is to divide the vibration spectrum into resonance modes. A low-pass filter is applied to the vibration spectrum amplitude, and the local minima define points for dividing the spectrum into resonance bands. The vibration spectrum is shown in Fig. 3.2, where the y-axis is the vibration amplitude, and the x-axis is the frequency in Hz. In this example, 9 identified bands are investigated further.

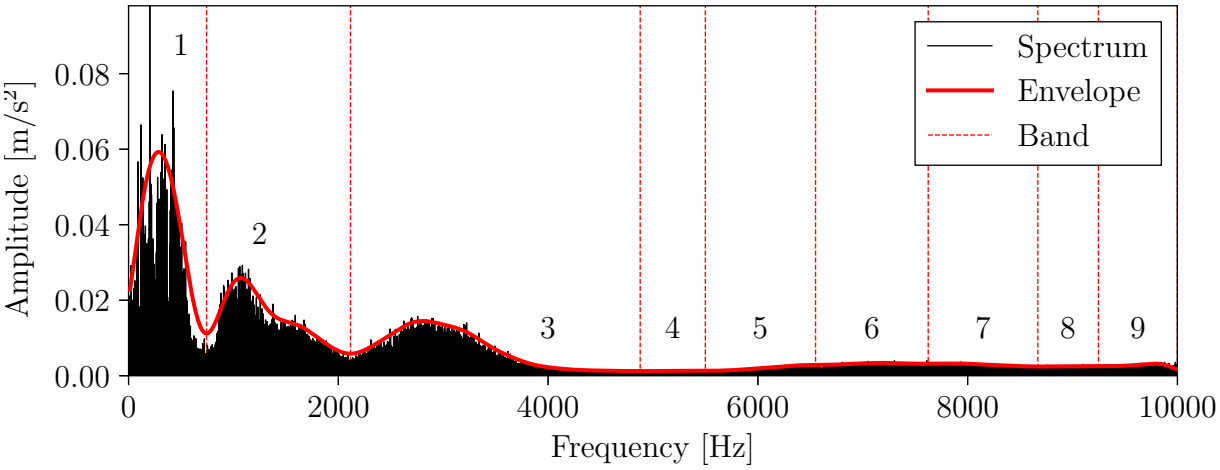

Figure 3.2: The vibration spectrum shown with the nine identified resonance mode bands.

The vibration signal is bandpass filtered at each identified band, and the envelope spectrum is afterwards calculated. For each envelope spectrum, a harmonic search algorithm is applied to identify prominent harmonics and side-bands related to bearing faults.

The search algorithm focuses a single fault case at a time, for example an inner race fault. Side-bands appear with 1 order spacing from the harmonic due to radial load modulation. Fig. 3.3 shows the search for the first harmonic and side-bands in the third frequency band.

The load angle ϕ normally makes up to 2 % difference in characteristic fault frequencies. Therefore, prominent harmonics are searched for within a band of ± 2 % width of the investigated characteristic fault frequency. The maximum values within these bands are chosen, and if the harmonic and one of the side-bands are greater than the threshold of three times the noise floor (3N), the harmonic is prominent. A score value is increasing for each prominent harmonic. The score is the ratio between the prominent harmonic and the

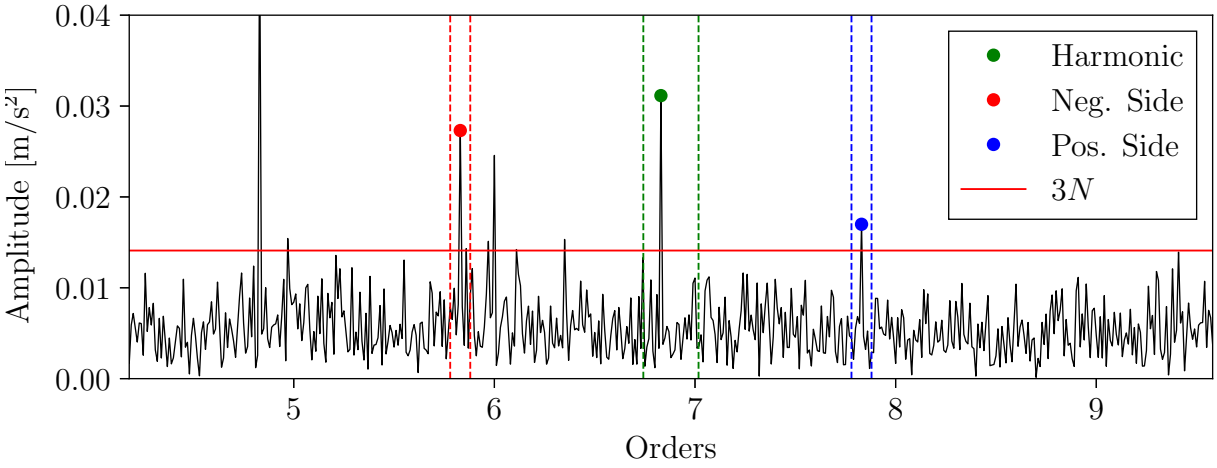

Figure 3.3: The first harmonic and its side-bands, and the prominence threshold given by 3 times the noise floor. The search width is given by the stapled lines.

threshold, multiplied with the harmonic number squared. The squaring causes the score to increase quickly if multiple harmonics are identified. This value can be interpreted as the probability of fault detection. In the third frequency band, the algorithm can identify three prominent harmonics as shown in Fig. 3.4. This procedure is continued with the 8

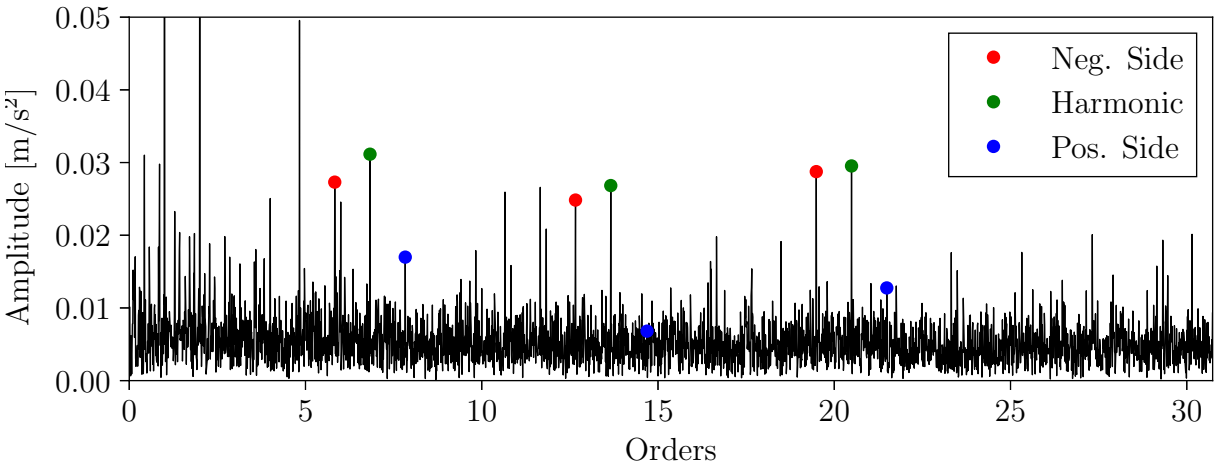

Figure 3.4: The three prominent harmonics and side-bands identified in the third frequency band.

other frequency bands shown in Fig. 3.2, and the score is summed to a single value for this dataset.

The scores for each fault type can be monitored over time for easier visualization of fault propagation. Fig. 3.5 shows the normalized fault scores (divided by their respective maximum value) for the last 100 datasets captured at 250 rpm using the in-house test rig. The normalized scores are small for the first 70 datasets, but the green trend increases to its maximum value at dataset 71. A fault score of less than 10 is typically observed for

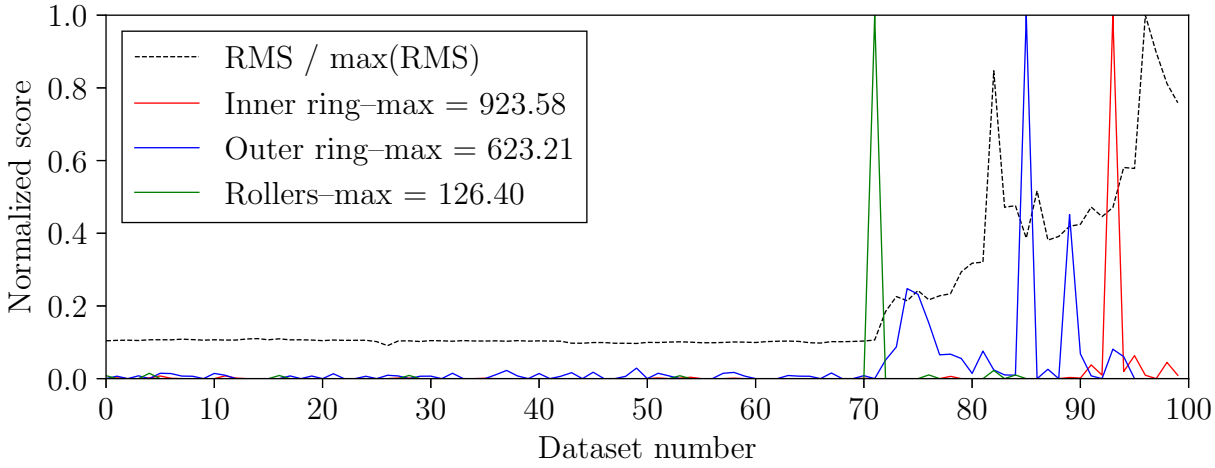

Figure 3.5: The fault scores over time for the last 100 datasets. The root mean square (RMS) is displayed to compare the fault score with the vibration energy.

noise, therefore 126.40 indicates a high probability of a roller fault. The outer ring is also damaged afterwards as seen by the increase of the blue trend during the next 20 datasets. Finally, the inner race was also damaged near the end of the lifetime.

The fault scores are compared with the RMS of the vibration signal, which is a measure of the mean vibration energy in the signal, given by

$$\text{RMS} = \sqrt{\frac{1}{T}\int_0^T x(t)^2 dt} \tag{3.6}$$

At dataset 71, the RMS value is unchanged from previous values, while the roller fault score increased to 126. Therefore, the proposed method can give earlier fault detection compared to the RMS, and the fault is automatically classified.

3.3 Whitened cross-correlation spectrum - Paper C

The automatic fault score method presented in the previous section can be used for fault diagnosis in many cases. During LSC, however, the bandpass filtered envelope spectrum may not provide correct diagnosis information. For this reason, AE sensors have been used in many applications to improve signal sensitivity and catch the low energy vibration.

Paper C presents a new method for extracting more information from the vibration signal for fault diagnosis during LSC. First, the vibration signal is order-tracked, and the TSA is removed. Afterwards, an autoregressive model with a limited number of coefficients is trained to further whiten the signal. The result is a vibration signal that is partially whitened as some deterministic components are still left in the signal. Bearing faults can in some cases be detected in the vibration spectrum directly, without using

the envelope. Before the envelope spectrum was invented, this was the common method for fault diagnosis. However, the bearing must be significantly damaged for prominent harmonics to appear in the vibration spectrum. Additionally, the resonance frequency needs to be low enough for identifying deterministic components [1].

The proposed whitened cross-correlation spectrum (WCCS) aims to combine the advantages of the vibration and envelope spectrum by correlating the useful information. The envelope spectrum is more likely to contain fault related information but may contain a high noise floor during low speed operation. The vibration spectrum may also show small signs of the bearing fault, only slightly higher than the noise floor. When combining these two signals, the information of both are fused in a single spectrum. To perform the fusing, the cross-correlation of the vibration and envelope signal is calculated and used for diagnosis analysis.

Vibration datasets from the in-house test rig captured during shaft speed of 20 rpm are used to verify the proposed method. The first fault to develop during the accelerated lifetime test is a roller fault, and Fig. 3.6 shows the result diagnosing the bearing using the proposed method.

Fig. 3.6 (a) shows the spectrum of the partially whitened vibration signal. The yellow lines show harmonic locations of 2 times BSF, and green stapled lines show the FTF side-bands. The first 5 harmonics are not prominent compared to the noise floor, but multiple harmonics of the side-bands are visible. On the contrary, the envelope spectrum in Fig. 3.6 (b) has a low noise floor and a few prominent harmonic amplitude values. The fault-related harmonics are only slightly higher than the noise floor, and therefore it is difficult to diagnose the bearing based on this spectrum alone. The proposed method combines these two spectra into the WCCS which is shown in Fig. 3.6 (c). The perceived noise floor is reduced compared to the spectrum of the whitened vibration signal, and the first five harmonics are now prominent compared to the noise floor. In addition, a prominent peak at FTF is showing, which further indicates a roller fault.

A limitation with the WCCS is that the deterministic bearing fault components are the result of a low-pass filter applied to the high-frequency resonance signal [1]. If the resonance frequency is too high, the whitened vibration spectrum may not show signs of the bearing fault at all. In this case, the envelope spectrum will be correlated with mostly white noise, which should neither improve or decrease the diagnosis capability of WCCS. Therefore, this method can be used even if the resonance frequency is high.

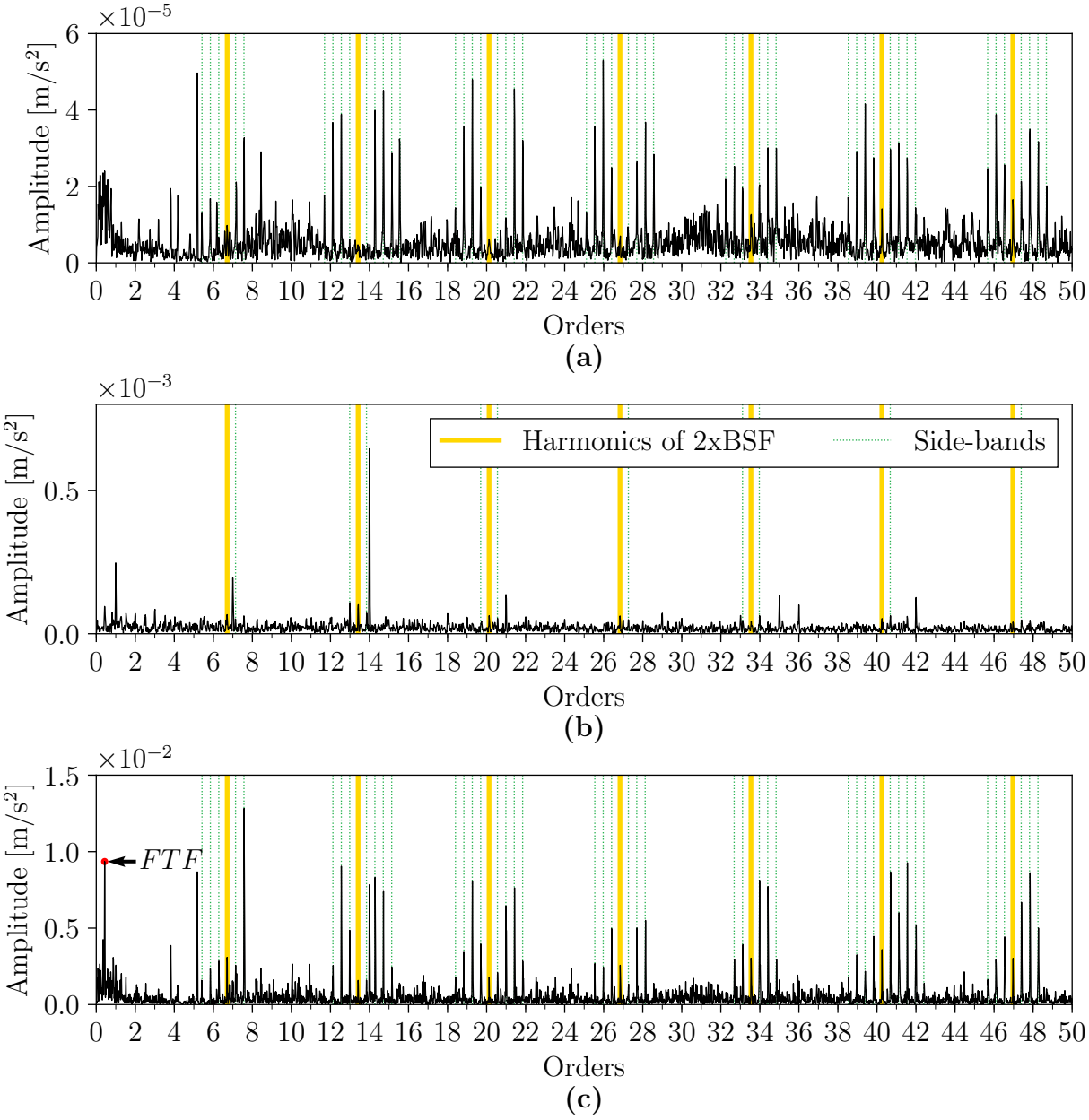

Figure 3.6: The WCCS used to diagnose a roller fault. Harmonics and side-bands linked to the roller fault are marked in all three subfigures. (a) Spectrum of the whitened vibration signal x_w; (b) Spectrum of the envelope x_{env}. (c) Proposed WCCS.

3.4 Resonance frequency identification - Paper D

Whitening methods can significantly improve the bearing diagnosis by isolating random signal components. However, white measurement noise is also random, and is therefore not removed in a whitening process. CPW is a whitening method proposed in [18] for bearing diagnosis under VSC. A brief description of CPW is given in Section 3.1.2. This method normalizes the vibration spectrum to reduce the power of deterministic components. The negative side effect is an increase in white noise energy, which could mask fault-related

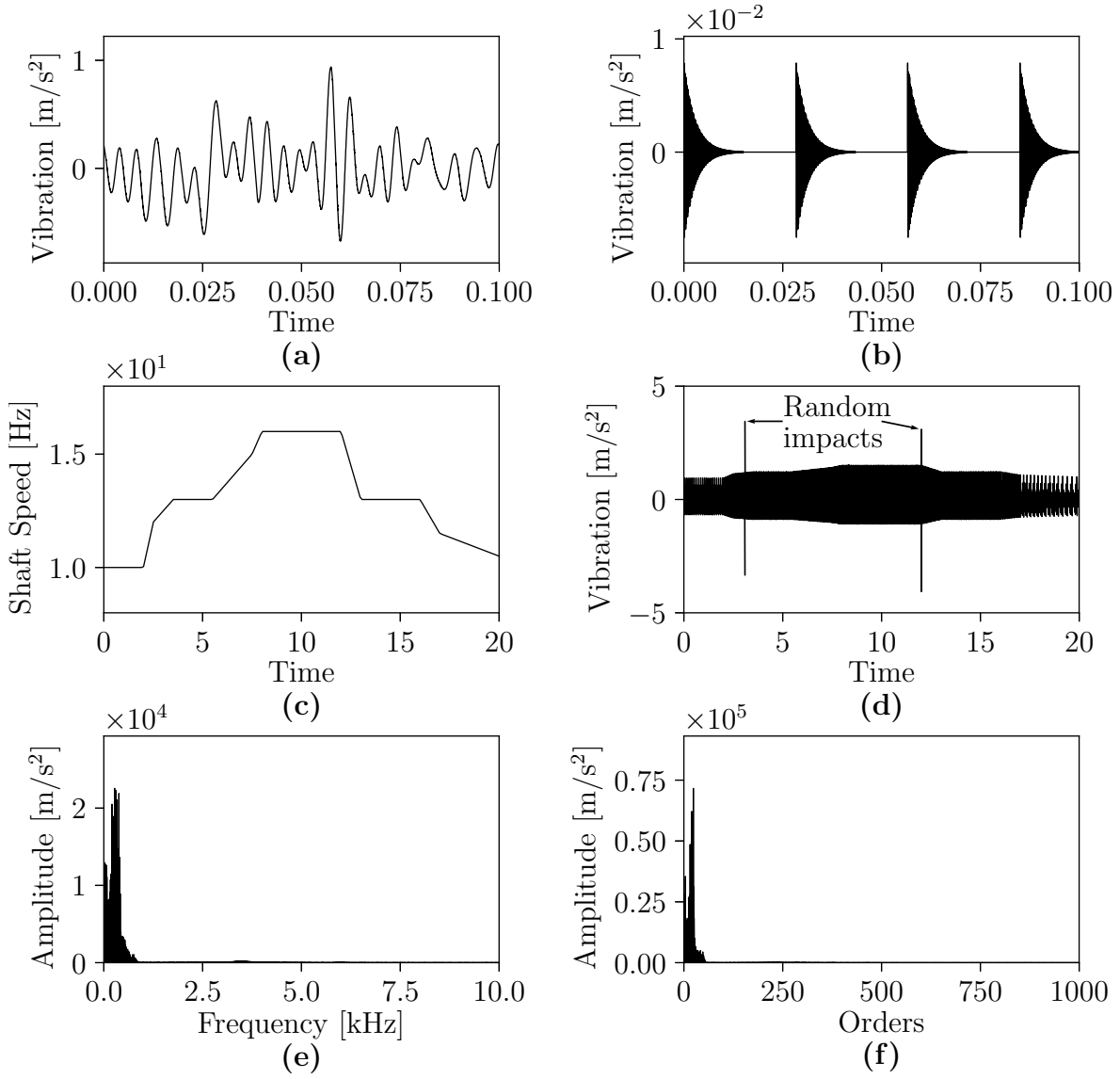

Figure 3.7: Resonance band identification using the proposed method on a simulated signal—Part 1. (a) shaft vibration; (b) bearing vibration; (c) shaft speed; (d) measured vibration signal; (e) frequency spectrum of raw signal; (f) order spectrum.

components. Although this whitening procedure is convenient, the diagnosis can be more accurate by isolating the resonance frequency instead.

Paper D presents a new method for identifying bearing resonance frequency when operating under VSC. The method utilizes CPW and order tracking to highlight shaft-speed invariant signal components. The procedure is shown as an example in Figs. 3.7 and 3.8 and explained hereafter. Deterministic shaft vibration harmonics are modeled as 27 sine wave components, and an underdamped second order model represents the bearing resonance vibration. The shaft and bearing vibration are shown in Figs. 3.7 (a) and (b), respectively. The shaft frequency ranges between 10 and 16 Hz as shown in Fig. 3.7 (c). The complete vibration signal is given in Fig. 3.7 (d), where two extra random impacts

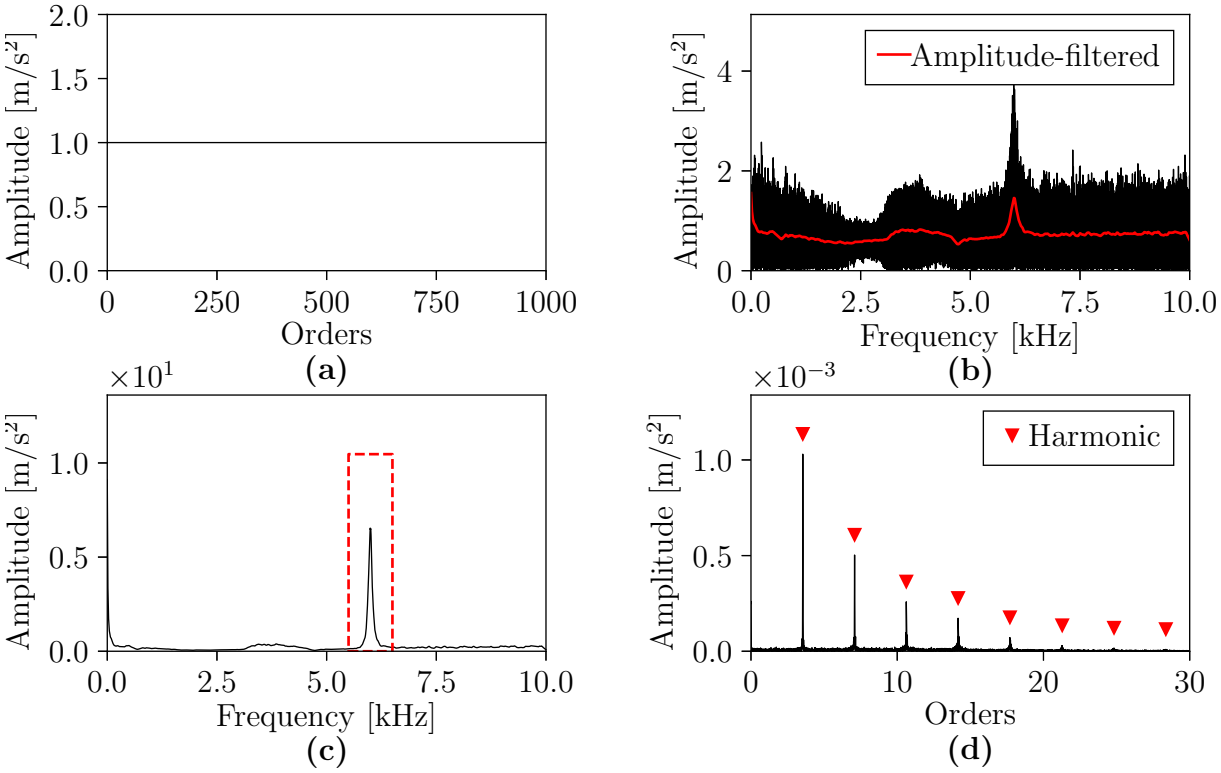

Figure 3.8: Resonance band identification using the proposed method on a simulated signal—Part 2. (a) order spectrum after CPW; (b) frequency spectrum after inverse order tracking with a red line showing the amplitude-filtered signal; (c) amplitude-filtered frequency spectrum raised to power of 5, with a suitable band-pass filter region marked in red; (d) envelope order spectrum after band-pass filtration of original signal.

are added to verify the proposed methods ability to ignore noise. First, the signal is order-tracked, which transforms the frequency spectrum as shown in Fig. 3.7 (e), into the order spectrum in Fig. 3.7 (f). The maximum value of the low-frequency peaks is amplified due to de-spreading of deterministic shaft-dependent components, and spreading of time-dependent vibration components. Afterwards, CPW is applied to normalize the signal as shown in Fig. 3.8 (a). Time-dependent signal components, such as bearing resonance vibration, are afterwards de-spread when inverse order tracking is applied, as shown in Fig. 3.8 (b).

A low-pass filter is applied on the spectrum amplitude to cluster single peaks into resonance bands. By raising the low-pass filter output to a power of 5, the resonance mode at 6000 Hz is isolated as shown in Fig 3.8 (c). The original signal is afterwards band-pass filtered at this frequency, given by the red-stapled square, and the envelope spectrum is shown in Fig 3.8 (d). In this spectrum, there are multiple harmonics related to the fault, which shows that the proposed method can be used to identify prominent resonance modes.

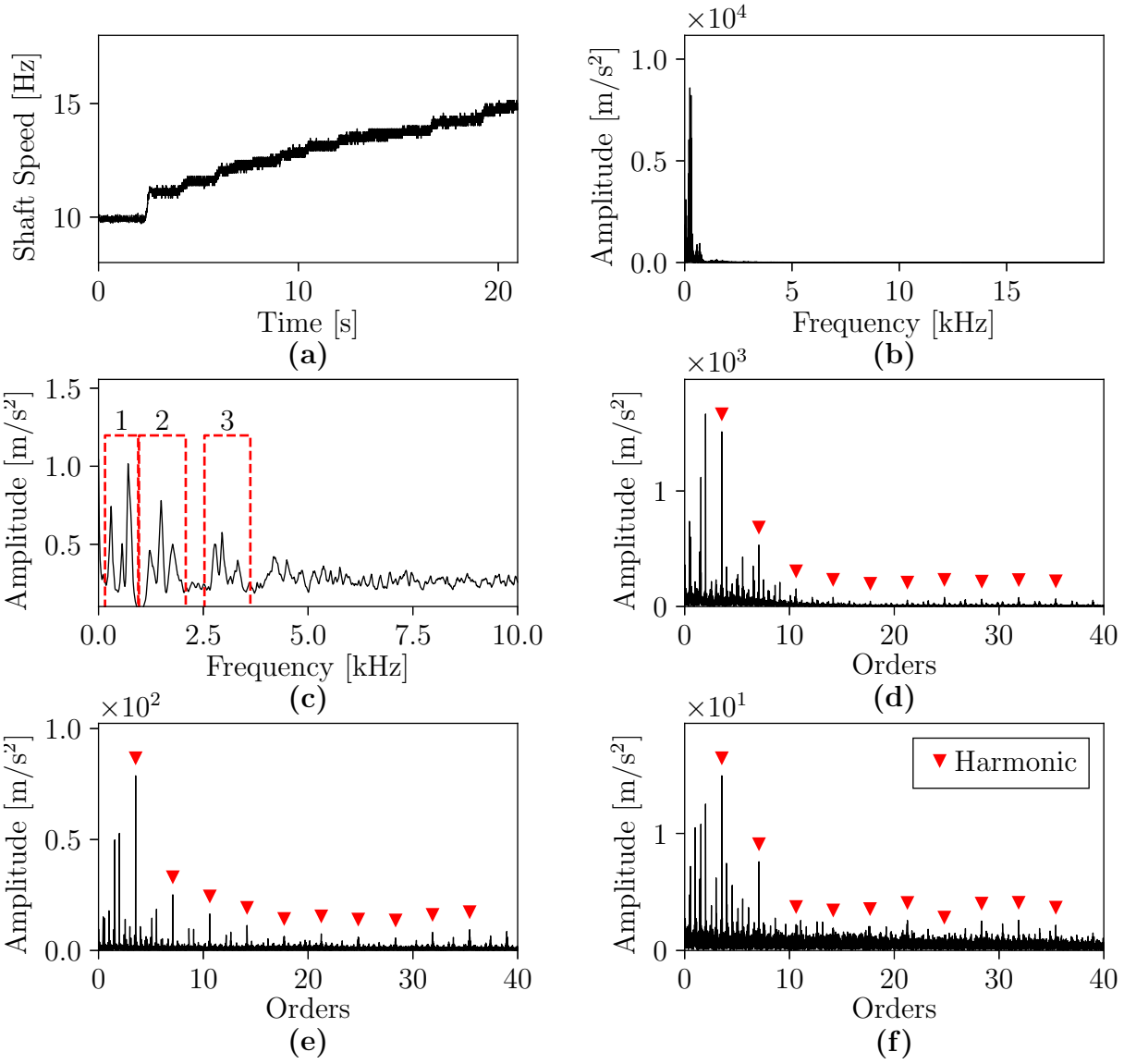

Figure 3.9: Results from diagnosing an outer race fault on a variable speed dataset. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where three suitable band-pass filter areas are marked; (d)–(f) envelope order spectra after band-pass filtration at the three marked areas, respectively. Red triangles show identified harmonics related to the fault.

As an example, a variable speed dataset is used to detect an outer race fault using the proposed method. A description of the dataset is given in Section 2.2.1. Fig. 3.9 (a) shows the shaft frequency, which ranges between 10 and 15 Hz. The vibration spectrum is shown in Fig. 3.9 (b), and it is dominated by low-frequency deterministic components. By applying the proposed method, resonance modes are highlighted as seen in Fig. 3.9 (c). Three suitable frequency bands are chosen manually to diagnose the bearing. The envelope spectrum is calculated for each band, and shown in Figs. 3.9 (d)–(f). In all

three envelope spectra, the automatic diagnosis method described in Paper B is used to detect prominent harmonics. The results make it easy to diagnose the bearing correctly, as there are multiple harmonics in each envelope spectrum. Selection of resonance bands can be automated using the automatic diagnosis algorithm presented in Section 3.2.

Chapter 4

Prognostics

This chapter describes a new method for estimating the RUL of a bearing without utilizing historical failure data. Section 4.1 is based on Paper E and briefly explains how the RMS vibration FT can be chosen based on ISO standard 10816-3 [44]. Next, Section 4.2, based on Paper F, describes how the vibration signal can be split into multiple frequency bands for RMS calculation to achieve more suitable trends for RUL estimation. Finally, the results of estimating RUL using a particle filter (PF) and the Paris-Erdogan law are shown in Section 4.3.

4.1 RMS health indicator - Paper E

Estimating the bearing RUL is challenging as the degradation trend must be predicted. More so, without historical failure data of the machine, it is difficult to determine a suitable HI and corresponding FT that correlates well with the actual degradation. To tackle this problem, the proposed methodology is to rely on industrial standards that can be generalized for many applications. The RMS is closely related to the vibration energy and can be used to assess the degradation of bearings because a larger number of defects results in higher vibration energy. ISO standard 10816-3 [44], "Mechanical vibration—Evaluation of machine vibration by measurements on non-rotating parts—Part 3", defines allowable vibration RMS values in the velocity domain, which can be used to set the FT. The RMS of a signal x can be calculated with

$$\mathrm{RMS}(x) = \sqrt{\frac{1}{N}\sum_{k=1}^{N} x_k^2}\,, \tag{4.1}$$

where N is the number of samples in the signal and x_k is the discrete value at time index k. Four classes are defined in ISO 10816-3: A–"new machine condition", B–"Unlimited long-term operation allowed", C–"Short-term operation allowed" and D–"vibration causes

damage". Vibration level D can be used as an FT because the vibration from continued operation can damage other components such as pump seals and gears. Vibration is commonly measured in the acceleration domain using accelerometers, because the piezoelectric crystal responds directly to pressure changes, and generates a voltage signal. To get velocity-based vibration, the signal must be integrated, either digitally or using analog circuits. The latter is preferred for accurate integration in standalone devices that checks only for high velocity-based RMS values. However, integrating the acceleration-based vibration signal to velocity domain presents drawbacks for bearing RUL estimation: high-frequency components are attenuated such that the effective bandwidth is reduced. Initial bearing wear generate high-frequency vibration above 1000 Hz, and therefore contribute insignificantly to the velocity-based RMS. The accelerated lifetime dataset introduced in Section 2.2.2 is used to show the difference between velocity- and acceleration based RMS. Let $R_a = \text{RMS}(x)$ be the acceleration-based RMS and $R_v = \text{RMS}(x_v)$ be the velocity-based RMS, where x_v is the vibration signal in the velocity domain. Fig. 4.1 (a) shows R_v and the mean initial value $\bar{R}_v$.

An initial degradation alarm is set as the mean value $\bar{R}_v$ plus 5 times the STD σ_v to achieve a low probability of false alarm. In Fig. 4.1 (a), the alarm is triggered when there is only a few hours left of the useful life, which gives little time to plan maintenance. R_a shown in Fig. 4.1 (b), is different, and the initial degradation alarm is triggered much sooner compared to R_v. In this case, there is approximately 70 hours left of the actual life. However, the FT must be defined for R_a to make it useful for estimating the RUL.

Paper E presents a new method for making an FT for R_a using ISO 10816-3 [44]. An analytic transformation that applies for any machine was not feasible to determine, because a detailed model of the vibration signal is required for that E. The transformation can instead be determined experimentally for each machine using baseline data. The ratio between R_a and R_v is determined to transform the velocity-based threshold $\hat{R}_v$ to the acceleration-based threshold $\hat{R}_a$. The ratio is given by

$$R_r = \bar{R}_a / \bar{R}_v \,, \tag{4.2}$$

where the mean RMS values during baseline measurements are used to determine the ratio. Afterwards, the ratio is used to transform the threshold from velocity-domain to acceleration-domain using

$$\hat{R}_a = R_r \hat{R}_v \,. \tag{4.3}$$

This transformation basically dictates that R_a must increase by a scale of $\bar{R}_v / R_v$ to reach the FT.

ISO standard 10816-3 [44] defines the vibration levels for machines that are 15 kW or larger. For 15 kW, the standard defines that vibration causes damage when $R_v = 4.5$

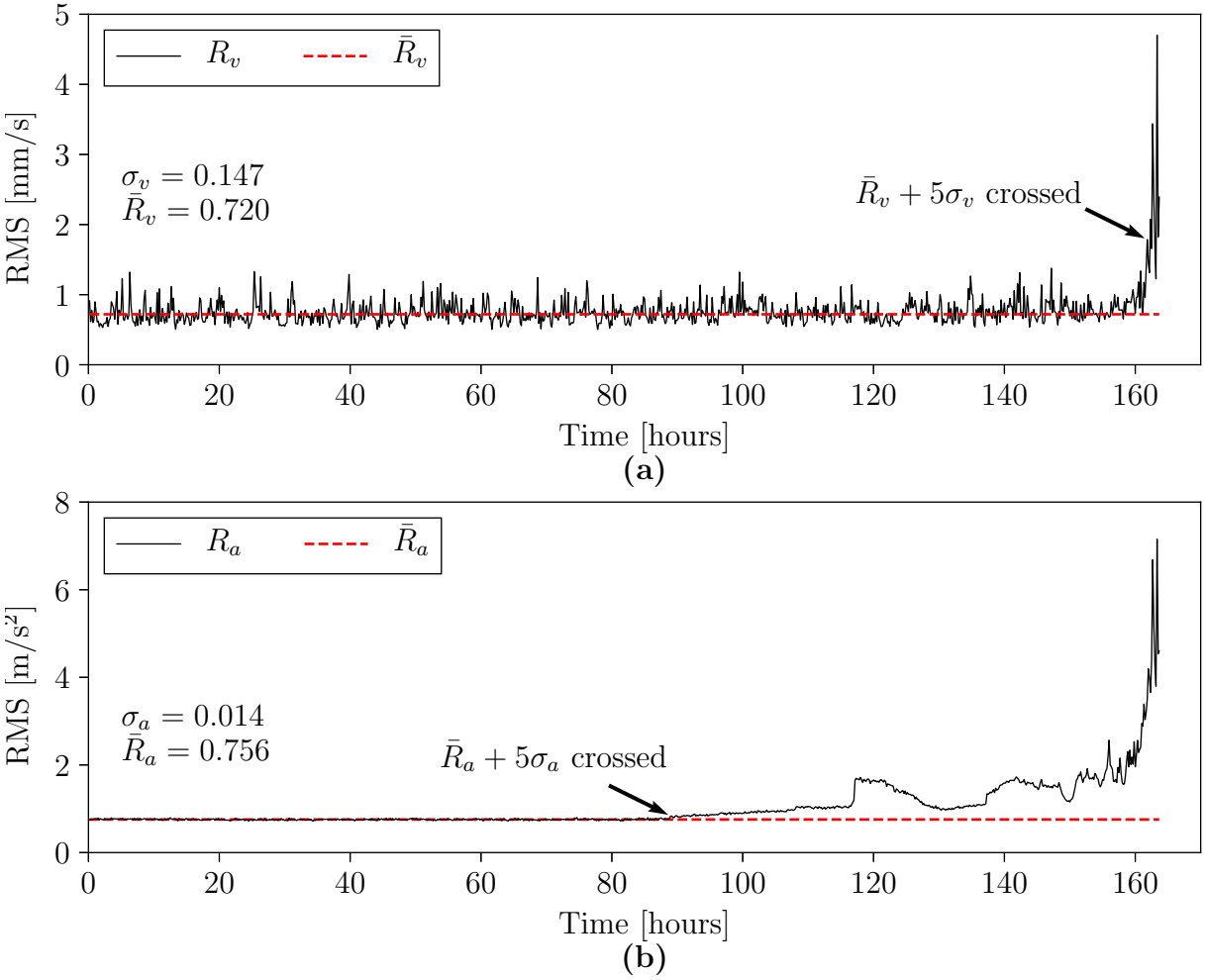

Figure 4.1: Difference between RMS values calculated using velocity and acceleration units. (a) RMS of the velocity signal; (b) RMS of the acceleration signal. The mean $\bar{R}$ and standard deviation σ are estimated using the first 50 hours of data files.

mm/s. As the motor used in the test rig may not be as powerful, the threshold is slightly reduced to $\hat{R}_v = 4$ mm/s. Using the baseline data in Fig. 4.1, the threshold is determined using (4.3) as $\hat{R}_a = 4.34$ m/s^2. The RMS trends are re-drawn in Fig. 4.2 together with the respective thresholds. $\hat{R}_v$ is reached by the velocity-based RMS as seen in Fig. 4.2 (a), showing that the ISO standard threshold is useful for stopping the machine. The acceleration-based threshold $\hat{R}_a$ shown in Fig. 4.2 (b) is also reached near the end of the useful life, which validates the transformation for this dataset. Paper E gives more details on the method and also more practical examples using other datasets.

While R_a can be used as an HI for RUL estimation, the trend is not ideal for predicting the future degradation level. The main reason is the oscillating behavior, which makes R_a a non-monotonic trend. R_a is cropped in Fig. 4.3 to better show the oscillations. The red-stapled line shows the mean degradation path, which is ideal to extract. Non-monotonic trends render a challenge for determining optimal model parameters to predict future

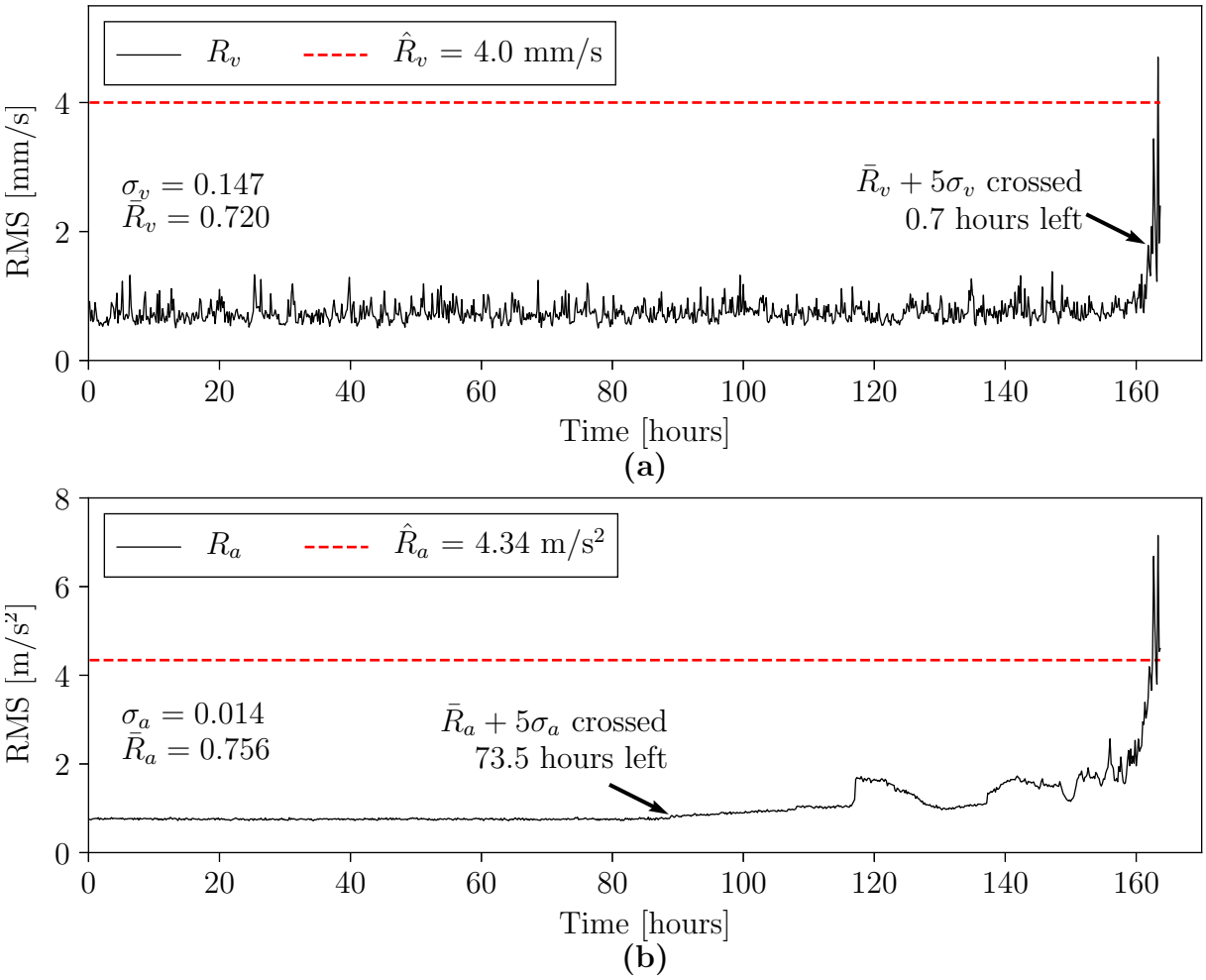

Figure 4.2: Comparison of the RMS threshold using velocity and acceleration units. (a) Velocity-based RMS; (b) Acceleration-based RMS.

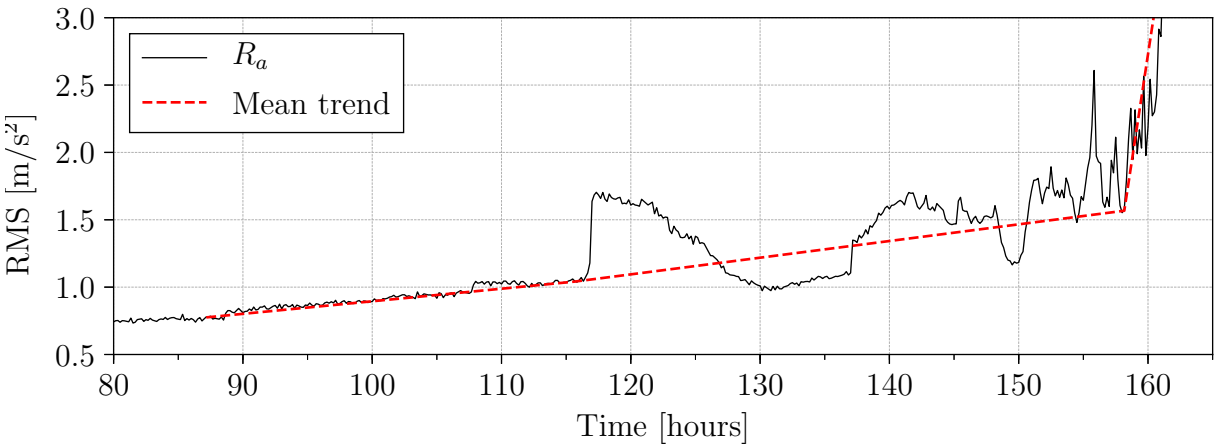

Figure 4.3: R_a compared with the mean degradation trend to highlight oscillations.

values. Therefore, the mean degradation path should be extracted from the vibration signal. The next section describes a new method for extracting more useful information

from the vibration signal by splitting the signal into multiple frequency bands.

4.2 Filter bank RMS - Paper F

The non-stationary oscillations of R_a can be mitigated by considering the RMS of the components within a certain frequency band rather than the entire signal. The vibration signal can be split into multiple components using a filter bank, where each component contains a frequency sub-band of the input signal. This is achieved by iteratively passing the signal through low- and high-pass finite impulse response (FIR) filters and decimating each output signal with a factor of 2. The number of levels in the filter bank decides how narrow each frequency sub band is. The frequency content for each band can afterwards be used to calculate RMS. Digital FIR filters use convolution to filter the signal, and only the overlapping part between the filter kernel and the input signal should be preserved to avoid adding artifacts. Therefore, there is information loss in the filtering. A larger filter kernel yields sharper frequency cutoff, but the information loss is greater per level of filtering compared to a shorter filter.

However, the time domain filter bank not necessary to acquire to calculate the RMS within different frequency bands. Paper F presents a novel method for splitting the frequency spectrum into suitable bands for RMS calculations, without losing signal information/energy in the process. This is achieved with a single discrete Fourier transform (DFT), and the method is described in the following.

The RMS can be calculated based on the energy E of the signal, such as

$$\mathrm{RMS}(x) = \sqrt{\frac{\mathrm{E}(x)}{T}}\,. \tag{4.4}$$

The signal energy can be calculated in both time and frequency domain with

$$\mathrm{E}(x) = \sum_{i=1}^{n} |x_i|^2 \Delta t = \sum_{i=1}^{n} |X_i|^2 \Delta f\,, \tag{4.5}$$

where Δt is the time step between samples, X_i is the i'th DFT bin, and Δf is the frequency step. Let X_L be the frequency bins between negative and positive 50% Nyquist frequency, with the rest set to zero. Similarly, let X_H contain the other high-frequency bins, and the rest are zero. This arrangement represents a single level filter bank where the frequency content is split in half. It is demonstrated in Section F.3.1 of Paper F that the energy of the entire signal can afterwards be calculated as

$$\mathrm{E}(X) = \mathrm{E}(X_L) + \mathrm{E}(X_H)\,. \tag{4.6}$$

Eq. (4.6) shows that the total signal energy can be calculated from the spectrum bins directly. Further, if the spectrum is split evenly into N_b bands, the signal energy and

RMS are calculated with

$$\mathrm{E}(X) = \sum_{i=1}^{N_b} \mathrm{E}(X_{(i-1)n_b+1:in_b}) \tag{4.7}$$

$$\mathrm{RMS}(X) = \sqrt{\sum_{i=1}^{N_b} \mathrm{RMS}(X_{(i-1)n_b+1:in_b})^2}\,, \tag{4.8}$$

where $n_b = n/N_b$ is the number of samples in the frequency band. For brevity, the RMS of a frequency band i is calculated with

$$R_i = \mathrm{RMS}(X_{(i-1)n_b+1:in_b})\,. \tag{4.9}$$

To acquire an FT for each R_i, the conservation of RMS must be conserved as indicated in (4.8). By substituting the RMS values for the respective thresholds, (4.8) becomes

$$\hat{R}_a = \sqrt{\sum_{i=1}^{N_b} \hat{R}_i^2}\,, \tag{4.10}$$

where $\hat{R}_i$ is the FT for R_i. A simple approach for solving (4.10) is to let $\hat{R}_i \; \forall \; i \in [1, N_b]$ be equal. However, some energy bands may have larger mean values than others, and therefore this FT may not serve the purpose very well for estimating RUL. Instead, the assumption for solving (4.10) is that white noise is spread across the entire DFT, and that the final value of each R_i is proportional to the noise standard deviation. In this case, the FT is set as

$$\hat{R}_i = \mu_i + m\sigma_i\,, \tag{4.11}$$

where μ_i and σ_i are the mean and standard deviation of R_i during baseline measurements. After combining (4.10) and (4.11), and squaring the equation, m can be determined by solving the quadratic formula

$$\hat{R}_a^2 = m^2 \sum_{i=0}^{N_b} (\sigma_i)^2 + m \sum_{i=0}^{N_b} 2\mu_i\sigma_i + \sum_{i=0}^{N_b} (\mu_i)^2\,, \tag{4.12}$$

for the maximum valued m. With m and baseline vibration measurements, the thresholds for all R_i can be determined with (4.11)

Fig. 4.4 shows four RMS trends generated by frequency sub-bands of the previously described dataset. R_{15} in Fig. 4.4 (a) is based on frequency content within [4480, 4800] Hz. The trend is not monotonously increasing and is therefore not suitable for RUL estimation. On the other hand, R_3 in Fig. 4.4 (d), is very suitable for RUL estimation since the trend is monotonic. Another difference between R_{15} and R_3 is the time of triggering the initial alarm, as R_{15} triggers it almost 25 hours before R_3. These observations suggest that

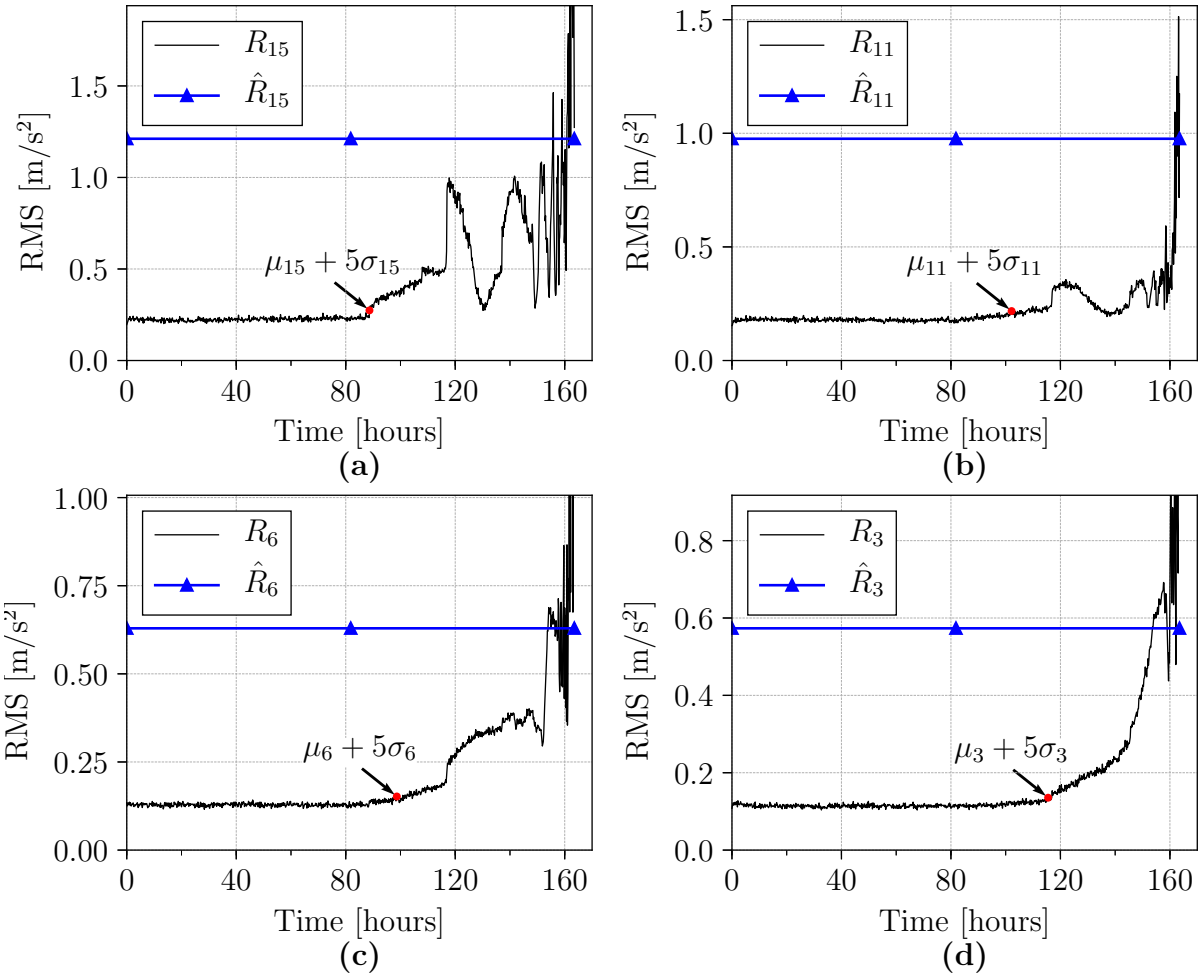

Figure 4.4: A collection of RMS trends with failure FTs for the vibration dataset. (a) R_{15}; (b) R_{11}; (c) R_6; (d) R_3.

the energy of high-frequency components increase first, but are more volatile, while the energy in low-frequency components changes last but rises more steadily. R_{11} and R_6 in Figs. 4.4 (b) and (c) are calculated from frequency bands between the two others. As is observed here, R_6 is more monotonic compared to R_{11}. In addition to these observations, the FT is reached for all the trends in this dataset. The next section shows how suitable RMS trends can be identified and used to estimate the bearing RUL.

4.3 Remaining useful life estimation - Paper F

Since it is unknown which of the RMS trends are suitable for RUL estimation, the most monotonously increasing ones needs to be identified online. The Spearman rank correlation coefficient ρ [52] assesses how well two signals are correlated in terms of their rank. Here, the RMS trend and time are used as input, which returns a measure of how monotonic the RMS signal is. By iteratively calculating the Spearman coefficient for all trends

with each new measurement, the most suitable trends for RUL estimation are determined online. Bearing degradation is estimated using the modified Paris-Erdogan law [41] given by [37]

$$\frac{da}{dn_c} = \alpha a^{\beta}\,, \tag{4.13}$$

where a is the crack size, n_c is the number of cycles, and α and β are the model parameters. This model is utilized to replicate the exponential degradation of a bearing [53].

A particle filter (PF) is applied to filter noise and predict the future R_i trend using the modified Paris-Erdogan law. The model is re-formulated to a state-space representation given by

$$\begin{cases} a_k &= a_{k-1} + \alpha_{k-1} a_{k-1}^{\beta} \Delta n_c \\ \alpha_k &= \alpha_{k-1} \\ h_k &= a_k + \nu_h\,, \end{cases} \tag{4.14}$$

where $\alpha_{k-1} \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$ is a random variable given by a normal distribution $\mathcal{N}$, β is constant, $\nu_h \sim \mathcal{N}(0, \sigma_h^2)$ is measurement noise, and Δn_c is the number of cycles since last update. Initial states of N_p particles $\boldsymbol{z}_k^j \,\forall\, j \in [1, N_p]$ at time index k are given with

$$\boldsymbol{z}_k^j \sim \mathcal{N}\left(\begin{bmatrix} R_{i,k} \\ \mu_\alpha \end{bmatrix}, \begin{bmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \right)\,, \tag{4.15}$$

where $R_{i,k}$ is the measured R_i value at time index k, and the weight of each particle is set to $w_k^j = 1/N_p$. Afterwards, the PF predicts new values for its states using the state-space model, and particles with high probabilities of matching new measurements are given a high weight. The weights describe a probability density function (PDF) that dictate which parameter values make the state-space model match the measured samples best. The RUL estimation of each particle is achieved by iteratively running the state-space model in (4.14) until the FT is reached. Combining the particle RUL with their weight creates the PDF of the RUL estimation from trend R_i.

When initiating a new PF, the parameters are first initialized to suitable values. Non-linear least squares (NLS) is applied to determine the initial mean and constant parameters, i.e. $\Theta_1 = (a_1, \mu_\alpha, \beta)$, where a_1 is the initial crack size. The cost function is the squared error between simulated values using (4.14) and R_i measurement samples. This cost function is minimized by optimizing values in Θ_1. The measurement noise variance is afterwards determined by calculating the variance of the baseline data, i.e. $\sigma_h^2 = \sigma_i^2$. There is also a process noise on α given by $\nu_\alpha \sim \mathcal{N}(0, \sigma_\alpha^2)$, which is the uncertainty of the model. α is a function of the cyclic load on the bearing, which may change during the degradation period, and therefore this process noise is added when predicting new particle states. Therefore, the performance of the PF is dependent on the process and

measurement noise. The process noise is set to $\sigma_\alpha^2 = \mu_\alpha^2$ to give the PF some margins when filtering the RMS trend.

A new PF (PF_i) is initiated for all R_i trends that achieve a Spearman coefficient greater than $\hat{\rho}$. In theory, all N_b RMS trends can therefore be utilized to estimate the RUL, and the available information is weighed to make a single RUL decision. The weights are based on the current Spearman coefficient value for the trend, and a higher value gives a higher weight. The weight for PF_i at time index k is given by

$$W_{i,k} = \left(\rho_{i,k}^3 - \rho_L^3\right) / \left(1 - \rho_L^3\right) \ , \tag{4.16}$$

where $\rho_{i,k}$ is the Spearman coefficient for trend R_i at time index k, and ρ_L is the lower Spearman threshold. Non-monotonic trends get a zero weight as they cannot be predicted by the Paris-Erdogan law. These weights are multiplied to the corresponding PFs particle weights, and afterwards, the RUL PDF of all PFs are combined to a weighted PDF. The weighted mean and 95% CI of this PDF is the estimate of the bearing RUL. More details on this implementation is given in Section F.4 in Paper F.

The IMS dataset introduced in Section 2.2.2 is used to validate the performance of this proposed method. The dataset is split into $N_b = 32$ bands, giving a bandwidth of 400 Hz. In addition, 30 measurement samples are used to optimize the initial PF values, and the Spearman coefficient must be greater than $\hat{\rho} = 0.9$ to initiate a new PF. The lower Spearman threshold is set to $\rho_L = 0.7$ to neglect non-monotonic trends. Fig. 4.5 shows the identified trends and predicted output of the initialized PFs. Each subplot row is allocated for a single RMS trend number, given by the index i in the upper left corner. The columns of subplots contain the following. Column 1 shows the identified R_i with its FT $\hat{R}_i$, in addition to the median and 95% confidence interval (CI) prediction of the initialized PF. Column 2 shows the median and 95% CI of the μ_α parameter in the model. In addition, the median and 95 % CI of the predicted PF output are given in Fig. 4.6 when $t = 130$ to show convergence of PFs.

At $t = 101$ hours, the first monotonic RMS trend is identified, which is R_{15} shown in Fig. 4.5 (a). The trend starts to increase linearly at the beginning, but at $t \approx 115$ hours, the value increases quickly before gradually decreasing. The Spearman coefficient ρ_{15} gets below 0.7 when $t \approx 130$ hours due to the fluctuations in R_{15}, and PF_{15} will not contribute to the weighted RUL after this point. μ_α is shown in Fig. 4.5 (b), and the parameter eventually converges to the median value, signifying that the predicted trend is not updated on new samples. The predicted trend is shown in Fig. 4.6 (a) for $t = 130$ hours. The 95% CI is much smaller compared to the initial PF prediction due to the convergence of μ_α.

R_{14} and R_{12} are later identified as monotonic trends at $t \approx 108$ hours, and the initial PF outputs are shown in Figs. 4.5 (c) and (e) for the two trends, respectively. R_{14} and R_{12}

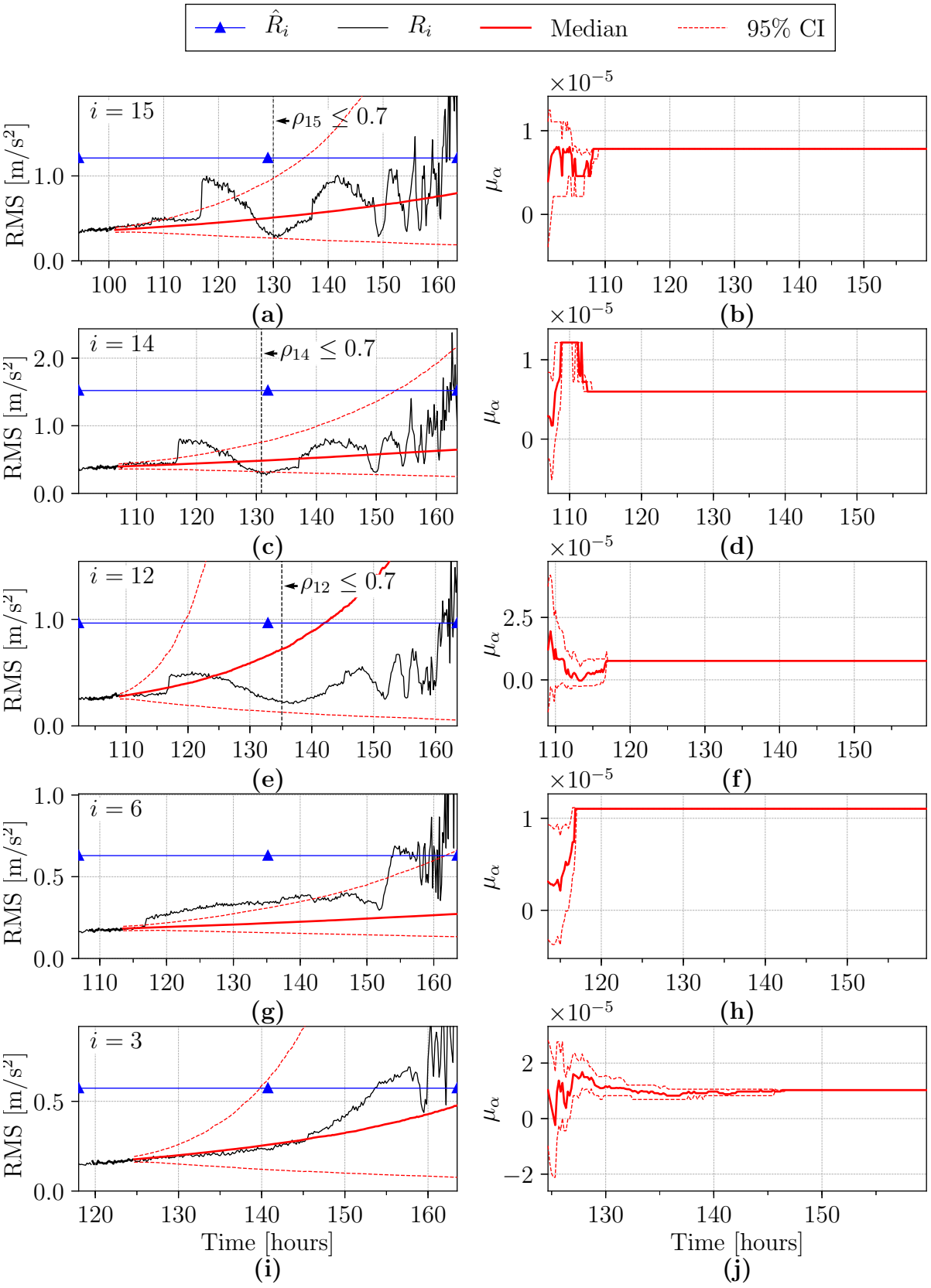

Figure 4.5: Identified RMS trends with high Spearman coefficient, and output of the corresponding PFs. Rows 1-5 indicate $i = [15, 14, 12, 6, 3]$. (column 1) R_i, FT $\hat{R}_i$, and median and 95% CI of initial PF output; (column 2) μ_α over time for the initiated PF.

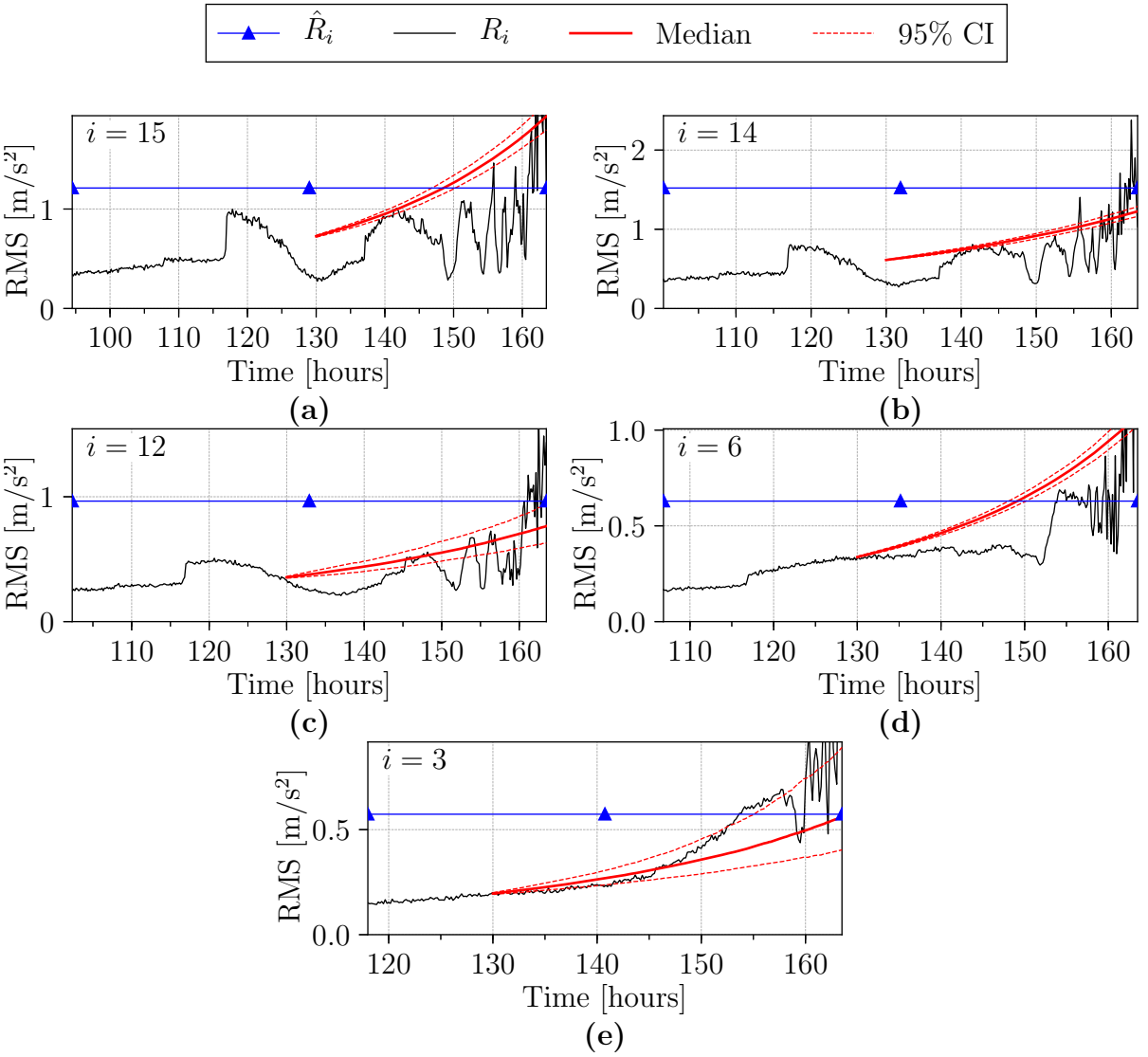

Figure 4.6: PF prediction of each identified RMS trend at $t = 130$ hours.

are similar to R_{15}, due to the RMS fluctuations. Additionally, the Spearman coefficient for these trends go below 0.7 at the points shown by the vertical stapled lines. The future PF predictions at $t = 130$ hours are shown in Figs. 4.6 (b), (c) for these two trends. The median of both PFs pass the FT near end of life due to good parameter estimation at the start. However, since the Spearman coefficient is lowered towards 0.7, the RUL estimation won't count in the weighted RUL decision.

A more promising trend R_6 is later identified at $t = 113$ hours, as shown in Fig. 4.5 (g). The trend is more monotonic compared to the previous three ones, and the Spearman coefficient never gets below 0.7. However, the trend is not increasing at a similar rate all the time. When μ_α converges, as seen in Fig. 4.6 (h), the predicted trend increases faster, as μ_α has increased from initial median value. The predicted PF output at $t = 130$ hours shown in Fig. 4.6 (d) increases faster than the new measured samples. Therefore, the

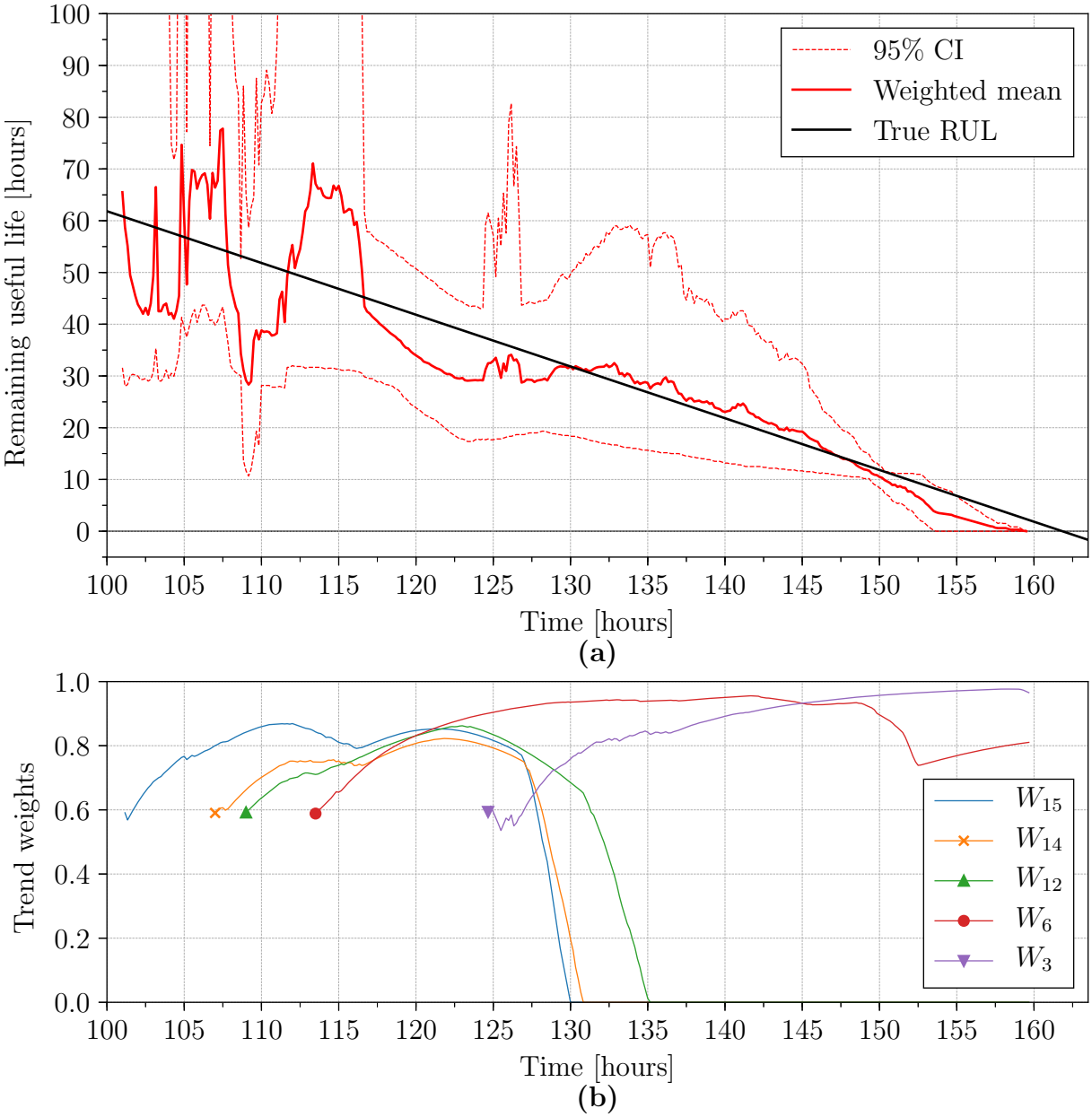

Figure 4.7: Weighted RUL decision. (a) weighted mean, 95% CI and true RUL of the dataset; (b) weights for each PF output.

RUL is underestimated by this trend.

The last identified RMS trend is R_3 shown in Fig. 4.6 (i). This trend is monotonic and is very useful for RUL estimation. The initial PF prediction matches the future samples well, which signifies a good RUL estimation by this trend. μ_α slowly converges to a median value in Fig. 4.5 (j) which imply that the initial value was chosen close to the final converged value. The predicted PF output at $t = 130$ is directed towards the true RUL, and therefore the RUL estimation is accurate at this point.

The weighted mean and true RUL are shown in Fig. 4.7 (a), while the weights for

each PF are shown in Fig. 4.7 (b). The weighted RUL is fluctuating around the true RUL throughout the estimation period, and changes for each new PF that is initialized. At $t = 130$ hours, the estimated RUL is the most accurate compared to the true RUL. Here, only W_{12}, W_6 and W_3 have significant weights. In Fig. 4.6, the predictions of PF_{12} overestimates the RUL, PF_6 underestimates the RUL, and PF_3 matches the RUL. Therefore, the weighted mean value falls between these three predictions, i.e. matching the true RUL.

Afterwards, R_{12} get a low Spearman coefficient within the next 5 hours. Therefore, the weights W_{15}, W_{14} and W_{12} are all turned to zero as seen in Fig. 4.7 (b) by $t = 135$ hours. After that, the RUL is estimated by PF_6 and PF_3 only, and the estimation matches the true RUL well.

A 3D plot in Fig. 4.8 shows how the weighed RUL PDF is changing over time. This plot makes it easier to visualize how the weighted mean is determined. At the beginning, the weighted mean RUL fluctuates due to several PFs being active, each with a high variance in μ_α. After $t = 135$ hours, only PF_6 and PF_3 are left, and the weighted mean RUL stabilizes between two peaks.

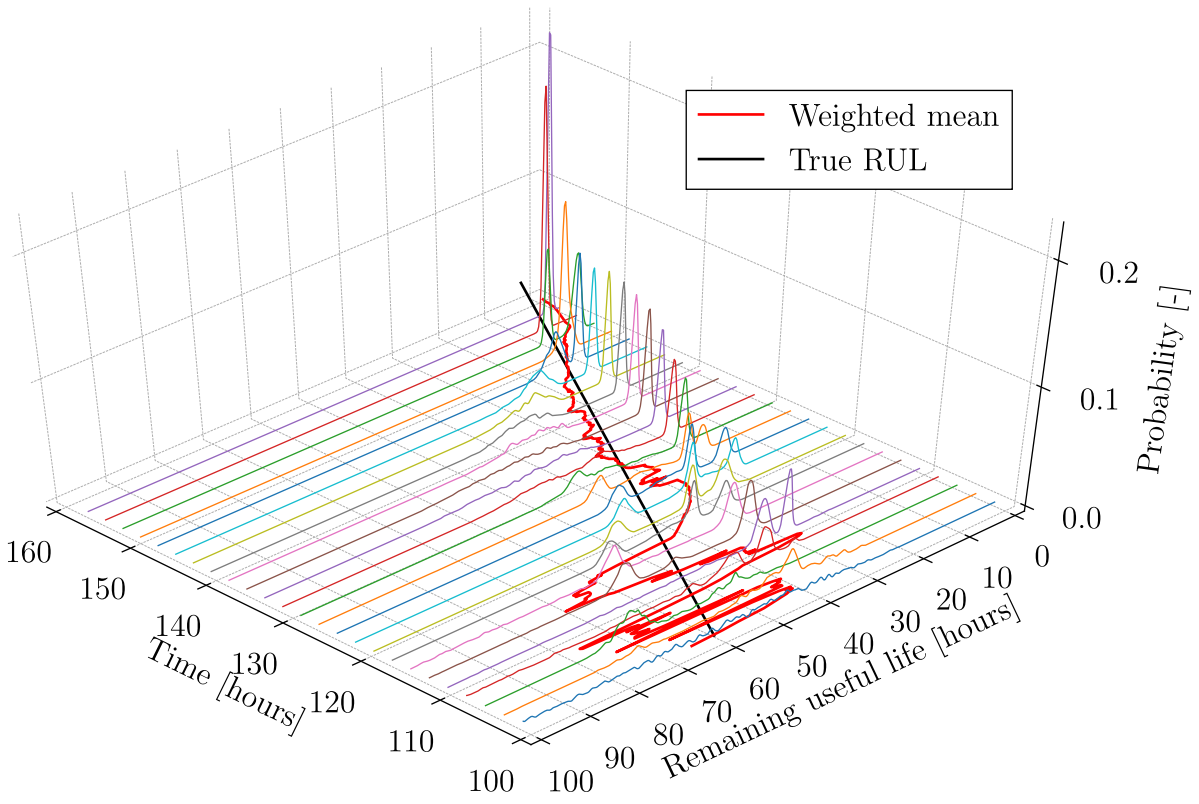

Figure 4.8: Weighted RUL PDF over time shown with the weighted mean RUL and true RUL.

Chapter 5

Concluding Remarks

5.1 Conclusions

The focus area of this project is fault detection and remaining useful lifetime (RUL) estimation of rolling element bearings. Several fault diagnosis and prognosis algorithms are proposed in this thesis to tackle different challenges in bearing health monitoring: automatic fault classification, low speed fault detection, variable speed fault detection, and remaining useful lifetime estimation on new machines. A test rig was first designed and built to make bearing failure sensor data at different operating conditions, such as low and variable speed. With the in-house datasets and downloaded datasets from open repositories, five algorithms have been proposed to tackle the aforementioned challenges.

With many machines in a plant to monitor, manual data analysis is cumbersome due to a large amount of data. As a solution, an automatic diagnosis algorithm has been developed to detect the most common bearing faults using the vibration signal. Vibration data captured at 250 rpm was successfully diagnosed for three common fault types, namely the rollers and two raceways.

During low speed conditions, a more robust algorithm is necessary compared to the envelope spectrum. A new algorithm has been developed for extracting more useful information from the vibration signal, by utilizing the deterministic components as well as the envelope. This made it possible to detect bearing faults during low-speed operation using a vibration accelerometer. With dataset from the in-house test rig, a roller fault was diagnosed at 20 rpm using this new method. Compared to three other methods reported in the literature, the proposed method resulted in stronger harmonics and more side-bands related to the fault.

For variable speed conditions, a new method was developed for estimating the bearing resonance frequency regions. The signal is bandpass filtered around each region, so the envelope spectrum is easier to analyze for faults as there is less noise in the signal. In

most of the investigated vibration datasets, several resonance modes were identified using this method. The envelope spectrum of each bandpass filtered resonance mode contained multiple harmonics and side-bands related to the fault. Compared to two other methods reported in the literature, the proposed method yielded the highest fault score using the automatic diagnosis algorithm.

Estimating the remaining useful lifetime of a bearing on new machines without historic failure data is the last challenge tackled in this thesis. Most prognosis algorithms reported in the literature require historic failure data for parameter tuning and setting the failure threshold. Such data may not be available in the industry, and therefore new algorithms needs to be developed for these situations. A new method has been developed for producing a general failure threshold for vibration root mean square (RMS) based on ISO 10816-3. This method only requires knowledge of the nominal machine power output, which is normally available in most plants. Experimental results from three test rigs show that the proposed method generated failure thresholds that are surpassed close to the actual time of failure.

The vibration RMS is, however, mostly non-stationary, and therefore fluctuates around a mean trend. These fluctuations make it difficult to predict the RUL using a mathematical model. A new approach for splitting the vibration signal into multiple frequency bands is proposed to remove signal non-stationarity. The RMS is calculated for each band, and monotonic RMS trends are identified using the Spearman coefficient. A particle filter algorithm is applied on each trend to predict future states by assuming the degradation follows the Paris-Erdogan law. Experimental results from the in-house setup and an online repository validate the performance of the proposed method.

In addition to the results presented in Chapters 3 and 4, several other datasets are tested with the proposed methods in the appended papers. In particular, Papers D, E and F include results using datasets acquired on the in-house test rig during low speed operation. The presented results show that the proposed methods are also applicable during low speed conditions.

5.2 Limitations

The automatic failure diagnosis algorithm proposed in Paper B requires a minimum vibration signal length to acquire a fine frequency resolution. If the resolution is poor, the algorithm has a trouble in distinguishing fault related peaks from surrounding noise. Measuring the vibration signal for at least 30 rounds should be enough for the algorithm to function properly.

The whitened cross-correlation spectrum (WCCS) proposed in paper C for low speed

bearing diagnosis is limited by machine resonance frequencies. An analytic derivation of the vibration spectrum shows that the deterministic components related to bearing faults are the result of a low-pass filter applied to the resonance vibration. If the resonance frequency is too high, the deterministic components may not be distinguishable from noise, and the proposed method may not improve the fault diagnosis.

In Paper D, a new approach for identifying resonance frequency bands is proposed. A clear requirement is the vibration signal must be recorded during variable speed operation, and that the shaft position must be acquired using an encoder. In addition, some of the resonance bands may contain components from other sources than the bearing, and therefore, several bands should be analyzed for bearing faults.

The failure threshold created in Paper E is based on the vibration energy, and therefore other failure criteria, such as maximum vibration, will require a different failure threshold. Further, the transformation from velocity-based to acceleration-based RMS failure threshold is based on experimental data acquired on each machine and may therefore not be optimal in all cases.

The particle filter implementation for remaining useful lifetime estimation in Paper F is computationally heavy, and may not be usable on small, embedded computers.

5.3 Further Work

Vibration data from real machines in the industry should be used to verify the performance of the proposed methods. In this thesis, vibration data has been acquired from either accelerated life-time tests or bearings with artificial faults. Bearing damages in real machines may have different characteristics compared to laboratory tests. Accelerated lifetime tests make somewhat comparable vibration data, but the high load may make the bearing break in different way than through normal fatigue over several years.

A vibration sensor has been used in this thesis as a cheaper alternative compared to acoustic emission sensors. To further reduce the cost of the diagnosis system, the shaft frequency may be estimated based on the vibration signal directly rather than using an encoder. Algorithms for estimating this has been reported in the literature, but more tests should be performed to verify the performance at low speed operation.

The method developed for low speed fault diagnosis should be validated on low speed vibration data from another test rig, as a single source of vibration data is not enough to validate the algorithm. In addition, it would be interesting to see if there is a limit to how low the shaft speed can be before the fault is indistinguishable from noise.

The methods developed in Paper F for splitting the vibration signal into frequency bands for RMS calculations can be used for estimating the RUL on new machines. How-

ever, the presented RUL estimation algorithm is computationally heavy, and more efficient methods for utilizing the RMS bands should be investigated.

References

[1] R. B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5):945–962, 2001. doi:10.1006/mssp.2001.1415.

[2] S. Braun. The extraction of periodic waveforms by time domain averaging. *Acta Acustica United with Acustica*, 32(2):69–77, 1975.

[3] W. Wang and A. K. Wong. Autoregressive Model-Based Gear Fault Diagnosis. *Journal of Vibration and Acoustics*, 124(2):172, 2002. doi:10.1115/1.1456905.

[4] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

[5] P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1):69–82, 1984. doi:10.1016/0022-460X(84)90595-9.

[6] M. S. Darlow, R. H. Badgley, and G. W. Hogg. Application of High-Frequency Resonance Techniques for Bearing Diagnostics in Helicopter Gearboxes. Technical report, US Army Air Mobility Research and Development Laboratory, 1974.

[7] M. Feldman. Hilbert transform in vibration analysis. *Mechanical Systems and Signal Processing*, 25(3):735–802, 2011. doi:10.1016/j.ymssp.2010.07.018.

[8] R. Dwyer. Detection of non-Gaussian signals by frequency domain kurtosis estimation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, volume 8, pages 607–610. IEEE, 1983.

[9] J. Antoni. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2):282–307, 2006. doi:10.1016/j.ymssp.2004.09.001.

[10] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007. doi:10.1016/j.ymssp.2005.12.002.

[11] T. Barszcz and A. JabŁoński. A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram. *Mechanical Systems and Signal Processing*, 25(1):431–451, 2011. doi:10.1016/j.ymssp.2010.05.018.

[12] X. Xu, M. Zhao, J. Lin, and Y. Lei. Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis. *Measurement*, 91:385–397, 2016. doi:10.1016/j.measurement.2016.05.073.

[13] K. R. Fyfe and E. D. S. Munck. Analysis of computed order tracking. *Mechanical Systems and Signal Processing*, 11(2):187–205, 1997. doi:10.1006/mssp.1996.0056.

[14] J. Urbanek, T. Barszcz, and J. Antoni. A two-step procedure for estimation of instantaneous rotational speed with large fluctuations. *Mechanical Systems and Signal Processing*, 38(1):96–102, 2013. doi:10.1016/j.ymssp.2012.05.009.

[15] T. Wang, M. Liang, J. Li, and W. Cheng. Rolling element bearing fault diagnosis via fault characteristic order (FCO) analysis. *Mechanical Systems and Signal Processing*, 45(1):139–153, 2014. doi:10.1016/j.ymssp.2013.11.011.

[16] Y. Wang, G. Xu, Q. Zhang, D. Liu, and K. Jiang. Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions. *Journal of Sound and Vibration*, 348:381–396, 2015. doi:10.1016/j.jsv.2015.03.018.

[17] J. Shi, M. Liang, and Y. Guan. Bearing fault diagnosis under variable rotational speed via the joint application of windowed fractal dimension transform and generalized demodulation: a method free from prefiltering and resampling. *Mechanical Systems and Signal Processing*, 68:15–33, 2016. doi:10.1016/j.ymssp.2015.08.019.

[18] P. Borghesani, P. Pennacchi, R. B. Randall, N. Sawalhi, and R. Ricci. Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions. *Mechanical Systems and Signal Processing*, 36(2):370–384, 2013. doi:10.1016/j.ymssp.2012.11.001.

[19] (ISO). 22096 condition monitoring and diagnostics of machines - acoustic emission, 2007.

[20] M. Elforjani and D. Mba. Accelerated natural fault diagnosis in slow speed bearings with acoustic emission. *Engineering Fracture Mechanics*, 77(1):112–127, 2010. doi:10.1016/j.engfracmech.2009.09.016.

[21] N. Jamaludin, D. Mba, and R. H. Bannister. Condition monitoring of slow-speed rolling element bearings using stress waves. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 215(4):245–271, 2001. doi:10.1177/095440890121500401.

[22] M. Elforjani and D. Mba. Detecting natural crack initiation and growth in slow speed shafts with the Acoustic Emission technology. *Engineering Failure Analysis*, 16(7):2121–2129, 2009. doi:10.1016/j.engfailanal.2009.02.005.

[23] N. Jamaludin and D. Mba. Monitoring extremely slow rolling element bearings: Part I. *NDT and E International*, 35(6):359–366, 2002. doi:10.1016/S0963-8695(02)00006-3.

[24] E. Y. Kim, A. C. C. Tan, J. Mathew, and B. S. Yang. Condition monitoring of low speed bearings: A comparative study of the ultrasound technique versus vibration measurements. *Australian Journal of Mechanical Engineering*, 5(2):177–189, 2008. doi:10.1007/978-1-84628-814-2_21.

[25] A. Widodo, E. Y. Kim, J.-D. Son, B.-S. Yang, A. C. C. Tan, D.-S. Gu, B.-K. Choi, and J. Mathew. Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. *Expert Systems with Applications*, 36(3):7252–7261, 2009. doi:10.1016/j.eswa.2008.09.033.

[26] T. R. Lin, E. Kim, and A. C. C. Tan. A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm. *Mechanical Systems and Signal Processing*, 36(2):256–270, 2013. doi:10.1016/j.ymssp.2012.11.003.

[27] B. Van Hecke, J. Yoon, and D. He. Low speed bearing fault diagnosis using acoustic emission sensors. *Applied Acoustics*, 105:35–44, 2016. doi:10.1016/j.apacoust.2015.10.028.

[28] Q. Xiong, Y. Xu, Y. Peng, W. Zhang, Y. Li, and L. Tang. Low-speed rolling bearing fault diagnosis based on EMD denoising and parameter estimate with alpha stable distribution. *Journal of Mechanical Science and Technology*, 31(4):1587–1601, 2017. doi:10.1007/s12206-017-0306-y.

[29] C. Mishra, A. K. Samantaray, and G. Chakraborty. Rolling element bearing fault diagnosis under slow speed operation using wavelet de-noising. *Measurement*, 103:77–86, 2017. doi:10.1016/j.measurement.2017.02.033.

[30] C. Hu, B. D. Youn, P. Wang, and J. T. Yoon. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, 103:120–135, 2012. doi:10.1016/j.ress.2012.03.008.

[31] N. Li, Y. Lei, J. Lin, and S. X. Ding. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Transactions on Industrial Electronics*, 62(12):7762–7773, 2015. doi:10.1109/TIE.2015.2455055.

[32] Y. Lei, N. Li, and J. Lin. A new method based on stochastic process models for machine remaining useful life prediction. *IEEE Transactions on Instrumentation and Measurement*, 65(12):2671–2684, 2016. doi:10.1109/TIM.2016.2601004.

[33] Z. Huang, Z. Xu, X. Ke, W. Wang, and Y. Sun. Remaining useful life prediction for an adaptive skew-Wiener process model. *Mechanical Systems and Signal Processing*, 87:294–306, 2017. doi:10.1016/j.ymssp.2016.10.027.

[34] Z.-X. Zhang, X.-S. Si, and C.-H. Hu. An Age-and State-Dependent Nonlinear Prognostic Model for Degrading Systems. *IEEE Transactions on Reliability*, 64(4):1214–1228, 2015. doi:10.1109/tr.2015.2419220.

[35] R. Li, P. Sopon, and D. He. Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*, 23(2):313–321, 2012. doi:10.1007/s10845-009-0353-z.

[36] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K.-L. Tsui. A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Transactions on Industrial Informatics*, 12(3):924–932, 2016. doi:10.1109/TII.2016.2535368.

[37] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, and J. Dybala. A model-based method for remaining useful life prediction of machinery. *IEEE Transactions on Reliability*, 65(3):1314–1326, 2016. doi:10.1109/TR.2016.2570568.

[38] J. Zhu, T. Nostrand, C. Spiegel, and B. Morton. Survey of condition indicators for condition monitoring systems. In *Annual conference of the prognostics and health management society*, volume 5, pages 1–13, 2014.

[39] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. S. Chow. Anomaly detection and fault prognosis for bearings. *IEEE Transactions on Instrumentation and Measurement*, 65(9):2046–2054, 2016. doi:10.1109/TIM.2016.2570398.

[40] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 104:799–834, 2018. doi:10.1016/j.ymssp.2017.11.016.

[41] P. Paris and F. Erdogan. A critical analysis of crack propagation laws. *Journal of basic engineering*, 85(4):528–533, 1963. doi:10.1115/1.3656900.

[42] D. Xu, Q. Zhu, X. Chen, Y. Xu, S. Wang, et al. Residual fatigue life prediction of ball bearings based on paris law and rms. *Chinese Journal of Mechanical Engineering*, 25(2):320–327, 2012. doi:10.3901/cjme.2012.02.320.

[43] Y. Qian, R. Yan, and S. Hu. Bearing degradation evaluation using recurrence quantification analysis and kalman filter. *IEEE Transactions on Instrumentation and Measurement*, 63(11):2599–2610, 2014. doi:10.1109/tim.2014.2313034.

[44] ISO. Mechanical vibration—Evaluation of machine vibration by measurements on non-rotating parts—Part 3 (Standard No. 10816-3). *ISO*, 1998.

[45] C. Mishra, A. K. Samantaray, and G. Chakraborty. Rolling element bearing defect diagnosis under variable speed operation through angle synchronous averaging of wavelet de-noised estimate. *Mechanical Systems and Signal Processing*, 72:206–222, 2016. doi:10.1016/j.ymssp.2015.10.019.

[46] H. Qiu, J. Lee, J. Lin, and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006. doi:10.1016/j.jsv.2005.03.007.

[47] Case Western Reserve University Bearing Data Center Website. 2015. URL: http://csegroups.case.edu/bearingdatacenter/home.

[48] G. U. Yule. On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London*, 226:267–298, 1927. doi:10.1098/rsta.1927.0007.

[49] G. Walker. On periodicity in series of related terms. *Proceedings of the Royal Society of London*, 131(818):518–532, 1931. doi:10.1175/1520-0493(1931)59<277:OPISOR>2.0.CO;2.

[50] N. Sawalhi and R. B. Randall. Signal pre-whitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings. In *COMADEM, Stavanger, Norway*, pages 330–336, May 2011.

[51] R. B. Randall, N. Sawalhi, and M. Coats. A comparison of methods for separation of deterministic and random signals. *International Journal of Condition Monitoring*, 1(1):11–19, 2011. doi:10.1784/204764211798089048.

[52] C. Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101, 1904. doi:10.1093/ije/dyq191.

[53] P. Rycerz, A. Olver, and A. Kadiric. Propagation of surface initiated rolling contact fatigue cracks in bearing steel. *International Journal of Fatigue*, 97:29–38, 2017. doi:10.1016/j.ijfatigue.2016.12.004.

Appendices

Paper A

Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition

Andreas Klausen, Roy Werner Folgerø, Kjell G. Robbersmyr and Hamid Reza Karimi

This paper has been published as:

A. Klausen, R. W. Folgerø, K. G. Robbersmyr and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Journal of Modeling, Identification and Control*, 38(3):143-156, 2017. doi: 10.4173/mic.2017.3.4.

Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition

Andreas Klausen*, Roy Werner Folgerø*, Kjell G. Robbersmyr*, Hamid Reza Karimi**

*University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

**Politecnico di Milano
Department of Mechanical Engineering
Via La Masa 1, 20156 Milan, Italy

Abstract **– Condition monitoring plays an important role in rotating machinery to ensure reliability of the equipment, and to detect fault conditions at an early stage. Although health monitoring methodologies have been thoroughly developed for rotating machinery, low-speed conditions often pose a challenge due to the low signal-to-noise ratio. To this aim, sophisticated algorithms that reduce noise and highlight the bearing faults are necessary to accurately diagnose machines undergoing this condition. In the development phase, sensor data from a healthy and damaged bearing rotating at low-speed is required to verify the performance of such algorithms. A test rig for performing accelerated life-time testing of small rolling element bearings is designed to collect necessary sensor data. Heavy loads at high-speed conditions are applied to the test bearing to wear it out fast. Sensor data is collected in intervals during the test to capture the degeneration features. The main objective of this paper is to provide a detailed overview for the development and analysis of this test rig. A case study with experimental vibration data is also presented to illustrate the efficacy of the developed test rig.**

A.1 Introduction

Rolling element bearings, or bearings for short, are necessary in rotating machinery to reduce the degree of freedom of moving parts. A typical bearing is made of an inner-race fastened to the shaft, a stationary outer-race, and rollers in between that transfers the shaft load. The relative distance between each roller is kept constant by a cage. Bearings

are precisely manufactured to withstand the dynamic loads acting on the shaft. Metal-to-metal contact is reduced to a minimum by lubricating the bearing with either oil or grease. However, wear will always be present, even in perfectly lubricated bearings. As a roller moves in and out of the radial load zone, the local lubrication is pressurized and causes stress to the rollers and the raceways. After millions of rotations, this cyclic load wears out the bearing components, and cause single or multiple faults which can occur in four different locations: the rollers, the inner-race, the outer-race, and the cage; although cage failures are uncommon. A worn bearing has increased friction, which in turn increases the machine temperature, noise, and vibration levels. A completely worn bearing could cause total system breakdown, injuries, and costly downtime.

It is important to monitor the condition of critical components, such as bearings, to schedule maintenance when necessary and avoid breakdowns. This is especially important for machines operating offshore, like windmill farms [1], as they are difficult to access and typically has a low profit margin. For bearing condition monitoring, the vibration signal is broadly accepted as a reliable data source. When a roller hits a fault, the impact energy causes the bearing to vibrate at its resonance frequency. This vibration is measured using an accelerometer, and is analyzed to determine the bearing condition. However, bearings faults of low-speed machines are difficult to diagnose as the fault impact energy is lowered, resulting in a low signal-to-noise ratio (SNR). Algorithms that reduce noise and highlight low-energy impacts are important to correctly diagnose a machine operating at low speed. Noise reduction examples in literature are autoregressive model filters [2], adaptive filters [3], and neural networks filters [4], for instance. Highlighting low-energy impacts can also be accomplished with a bandpass filter at the estimated bearing resonance frequency using the fast Kurtogram [5], or blindly deconvolve the raw signal to preserve parts with high kurtosis [6]. Analyzing proximity data to diagnose a faulty bearing has also been explored in [7] where the authors state that a proximity sensor is more sensitive than accelerometers during low-speed conditions. Acoustic emission sensors have also gained prominence in the past decade, promising better sensitivity compared to accelerometers during low-speed conditions. Apart from the sensor type selection, low-speed bearing sensor data is necessary to assess the diagnostic accuracy of a newly developed algorithm. Sensors can be used to acquire test data, and they can be mounted on industrial machinery in use, or on specialized test rigs in research facilities. The latter setup type is the most ideal for research as it grants full control of the test and the environment of the bearings. However, for verification purposes, it may be advantageous to test algorithms on industrial machinery as well.

In the literature, different bearing test rigs have been used to generate signal data. Some designs [8, 9, 10, 11, 12] combine a torque source and a test bearing that is monitored

for faults. These rigs are unable to apply heavy loads to the bearing, and are thus unable to naturally wear out the bearing within a feasible amount of time. Other test benches also include heavy load capabilities [13, 14], but due to a low rotating speed, the bearing did not wear out naturally. In these cases, the test bearing is installed with an artificially introduced, pre-seed fault, often shaped like a hole or a line. Comparing sensor data from a healthy and a damaged bearing is a typical scenario for testing a diagnostic algorithm.

One missing feature is the ability to follow the development of faults as it would happen in a real scenario. Remaining useful life prognostics of a machine is often dependent on the changes in sensor data as the amount of wear increases. To capture this trend, the bearing may be run to failure on an accelerated life-time test rig. Their designs include equipment for applying loads targeting the test bearing to accelerate its lifetime. The test rigs in [15, 16, 17, 18, 19, 20, 21, 22] combine high speed and heavy load to achieve faults in a reasonable amount of time. Unfortunately, their equipment does not handle low-speed scenarios for logging data. Some other designs have the loading capacity to wear out the bearing even during low-speed scenarios [23, 24], using only axial load.

The presented research gives a detailed overview of the development of a new test rig used to acquire low-speed sensor data. From the literature review, the adopted methodology for this rig is to accelerate the bearing life-time using heavy axial and radial loads at medium-to-high shaft speed. Sensor data is acquired in intervals during the test at lower speeds, reaching down to 20 revolutions per minute (rpm). Required hardware and software design for performing the test is also presented.

The rest of this paper is organized as follows. Section A.2 describes all specifications for the test rig, and the different solutions that fulfill these requirements. A combination of solutions are chosen and merged into a final design in Section A.3. Additionally, a case study from an accelerated life-time test is also presented. Finally, conclusions are drawn in Section A.4.

A.2 Methods

In this section, a short description of accelerated life-time testing is first given. Then, specifications for the test rig are provided, and possible solutions to fulfill them are discussed.

A.2.1 Accelerated life-time testing

Accelerated life-time testing (ALT) is the process of speeding up the degradation of a component by overloading it to uncover faults in a short amount of time. The test rig described in this paper is designed to drastically reduce the lifetime of a bearing from

several years to a few weeks. Subjecting the test bearing with large forces reduces the number of revolutions a bearing will hold. This comes from the following:

$$L_{10} = \left(\frac{C}{P}\right)^3, \tag{A.1}$$

where L_{10} denotes the bearing life-time in million revolutions with a 90% confidence (hence the lower case 10), C is the dynamic capacity of the bearing, and P is the weighted sum of radial and axial loads in Newtons. By running the bearing with a heavy load at a high speed, the life-time is accelerated to uncover faults relatively fast.

A.2.2 Specifications

Before designing the test rig, it is important to set up specifications describing the main features. Further development of concepts is based on these specifications:

1. The driving motor must be able to rotate the shaft at medium-to-high speed to accelerate the life-time of the bearing, and at a low speed (about 20 rpm).

2. The accessible radial and axial loads must be at least 10 kN. The load magnitude must also be easy to change.

3. It must be possible to mount sensors measuring bearing data including: temperature, vibration, shock pulses, acoustic emission, and shaft radial movement.

4. The test bearing housing must be big enough to house a 6008 size rolling element bearing with a 68mm outer diameter size.

5. The test rig equipment, software, and controller hardware must be designed for unsupervised 24/7 accelerated life-time testing.

The reason for designing a test rig around a small 6008 size bearing is to minimize the cost, space and complexity, as a larger bearing require heavier loads in addition to stronger support structure. In the next subsection, solutions that fulfill the specifications are presented.

A.2.3 Concept Development

Based on the specifications there are various solutions capable of achieving the desired functions. A function tree, illustrated in Figure A.1, visualizes feasible options to achieve the main function which is to accelerate the life-time of a bearing. The main function is split into four sub functions that must be achieved, and for each there are several

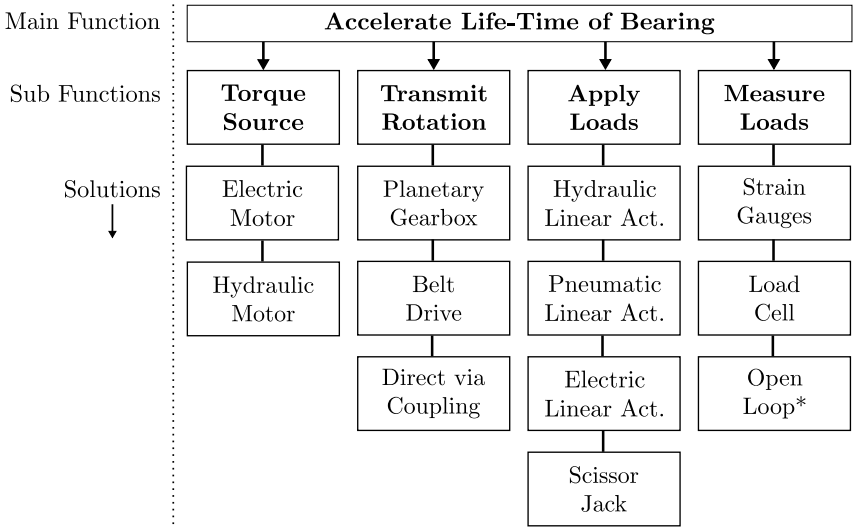

Figure A.1: Test rig pre-design function tree. *An Open-Loop force feedback is a feed-forward calculation of the assumed forces based on a particular input.

feasible solutions. The viability of each are discussed, with respect to the specifications, to determine which should be kept in the design phase:

Torque source: It is important that the torque source can control the test bearing shaft at low speed. If the shaft experiences too much unintentional speed fluctuations around the setpoint, the sensor measurements may be hard to analyze. A high-performance motor and frequency drive combination is necessary to fulfill this requirement. In addition, variable speed must also be realizable to simulate some machines undergoing variable conditions, such as winches or windmills. Most motor types have the possibility to be driven at both high and low speed. A hydraulic motor requires an external pressure source to be operational during the entire accelerated life-time test. Using a centralized hydraulic power unit (HPU) in a laboratory for 24/7 operation is not ideal due to power consumption and equipment safety. Installing an HPU for the test rig increase the complexity and cost of parts. Using an electric motor is more ideal since electricity is almost always readily available. The different electric motor types are: brushed DC motor, brush-less DC motor, induction motor, and permanent magnet motor. Brushed DC motors may require more maintenance due to wear of the brushes, and therefore these are not the most suitable. Brush-less DC motors and permanent magnet motors are quite similar in function, as they are synchronous machines. Low-speed operation is necessary, which require an encoder or resolver in the motor. Most induction motors are controlled using sensor-less speed control to reduce cost, but a position-feedback is important for low-speed scenarios to properly control the motor, and log the shaft position during measurements.

Transmit Rotation: The lowest speed requirement is 20 rpm, and it is improbable

that a motor coupled to the shaft directly will be able to rotate steadily at that low speed. This is because the shaft may be slightly unbalanced or bent, resulting in a variable load torque. Without a reasonably high inertia, the motor controller will most likely not react fast enough to the changes in the load torque. Therefore, a gearing ratio of at least 5 is probably necessary to increase the load inertia, and allow the motor to operate at a higher speed. This can be achieved with a belt drive, but a planetary gearbox is more compact to install.

Apply Loads: Using a hydraulic linear actuator (Act.) to load the bearing allows for a high amount of variable force, but require an HPU stationed nearby. Pneumatic linear actuators have some of the same disadvantages, but cannot subject the bearing with high forces due to low permissible air pressure. Hydraulic and pneumatic valves often leak over time, which require the control system to reapply pressure consistently. A scissor jack may create a high force, but varying the load requires manual operation and is not convenient. Electric linear actuators have the advantage that they can self-lock after the desired force is applied, therefore power is only necessary when the load is changing. However, they are often limited by their maximum force. Considering this, a mechanical lever can be used to amplify the load if necessary. It is important to allow for 24/7 operation, and as such it is inconvenient to keep a hydraulic or pneumatic power unit continuously powered on.

Measure Loads: Actuators have losses from friction and power conversions, and these losses may be non-linear, which makes them hard to estimate. Therefore open-loop force calculations will include errors and are thus not ideal to use. Strain gauges can be added to the structure to measure loads with good accuracy. However, such measurements require high-resolution analog-to-digital converters, a model of the system, and a way to calibrate the strain gauge signal based on a known reference. In addition, it may be difficult to suppress unwanted noise due to temperature changes, or to decouple load sources. Off the shelf load cells are pre-calibrated and temperature compensated by the manufacturer, and are thus much easier to use and less error prone.

A.3 Results and discussion

In this section, solutions for the sub functions are chosen and designed. Later, all parts are assembled on a suitable steel rig. The software and electronic hardware designs for performing the accelerated life-time test are elaborated afterwards. Finally, a case-study from an accelerated life-time test using the assembled test rig is presented.

A.3.1 The chosen solution

Based on the discussion of available components in Section A.2.3, the following choices are made:

The available electric motor types should all be capable of controlling the shaft at low-speed, assuming that some gearbox is installed, and the controller utilizes a motor position feedback sensor. A permanent magnet motor (PMM) with a high resolution built-in encoder, and a variable frequency drive with a high-performance controller was chosen to drive the shaft. This combination can drive the motor up to 4000 rpm, and due to the high efficiency of PMM, it is simples to keep cool during low-speed conditions as squirrel cage induction motor fans typically produce little airflow during this condition.

A planetary gearbox is installed to transmit the rotation torque from the motor to the shaft. It has a gearing ratio of 1:7, and is designed to be attached directly to the chosen PMM.

Two electric linear actuators were chosen to produce the radial and axial loads. Each have a loading capacity of 2.5 kN, which is too low considering the specifications. To amplify the low force, two levers are designed to amplify them.

Pre-calibrated load cells are used to measure the loads due to difficulties of applying strain-gauges. The chosen cells are shaped as bolts, and can measure the reaction force in a rotational joint up to 20 kN. These will be used as hinges for the two levers, and measure the reaction forces at the same time.

A.3.2 Test rig design

The design of the test rig starts with the shaft. The test bearing is installed on one end of the shaft, and the gearbox is connected via a coupling on the other end. Radial and axial loads are applied using linear actuators and levers. Aside from the test bearing, a minimum of two support bearings are also needed to counteract the added radial forces. To avoid wearing out the support bearings, durable ones with higher dynamic capacities compared to the test bearing are necessary. Also, it is important that the support bearings ignore the axial load on the shaft since it must propagate to the test bearing. Radial bearings within split plummer block housings are suitable for these specifications, as this combination give no axial support without installing extra equipment. The axial load must originate from a stationary source, hence a thrust (axial) bearing is installed on the shaft that allow for this load transfer. Four bearings are therefore installed on the shaft: a test bearing, two radial support bearings, and a thrust bearing. Figure A.2 (a) shows the acting forces on the shaft, and Figure A.2 (b) shows the bearing positions.

Stationary houses are required to transfer load to the rig structure, and to mount

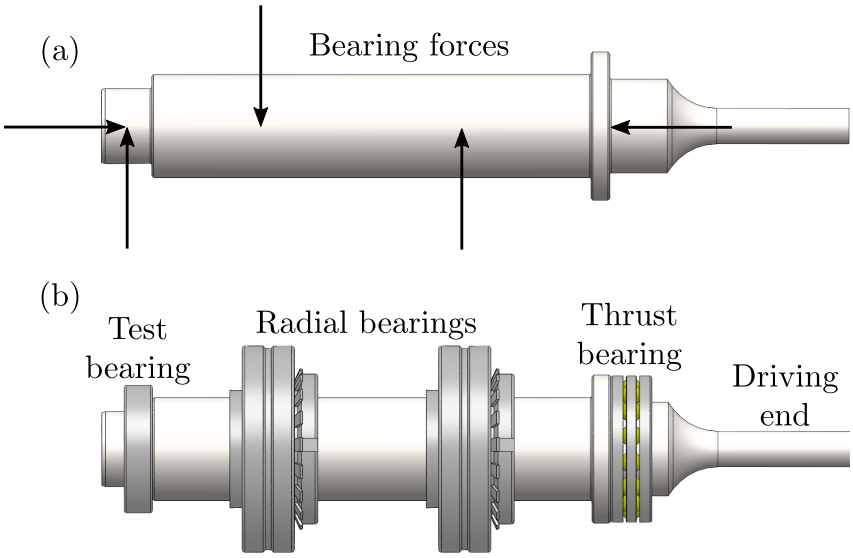

Figure A.2: Chosen shaft design and bearing placements. (a) Counteracting forces on all the bearings. (b) Position of the required bearings.

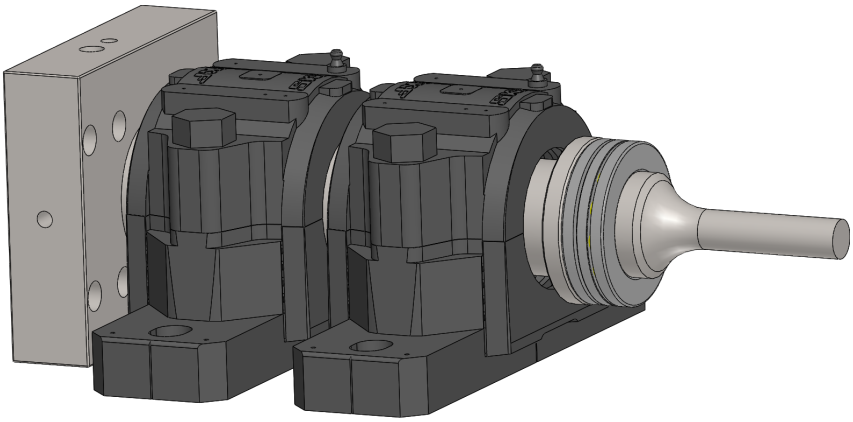

Figure A.3: Bearing housings as placed on the shaft.

sensors. Two suitable split plummer bearing housings are chosen to accommodate the two larger radial bearings. Inside these housings, the two bearings are free to slide in the axial direction. This is critical to transfer the axial load to the test bearing. The test bearing housing is processed in a CNC machine from a steel block. Figure A.3 shows the bearing houses on the shaft. The outer area of the test bearing housing can accommodate up to three vibration accelerometers, or shock pulse sensors, via stud mounts. A temperature sensor reaches the outer ring of the bearing via a hole in the housing. An eddy current proximity sensor is stud mounted in a drilled hole in the split plummer bearing housing close to the test bearing. Acoustic emission sensors can be placed using glue on any flat surface on the housing. Figure A.4 shows the possible sensor layout on the test bearing housing.

The axial load is added to the shaft via a concentric unit that is placed on the thrust

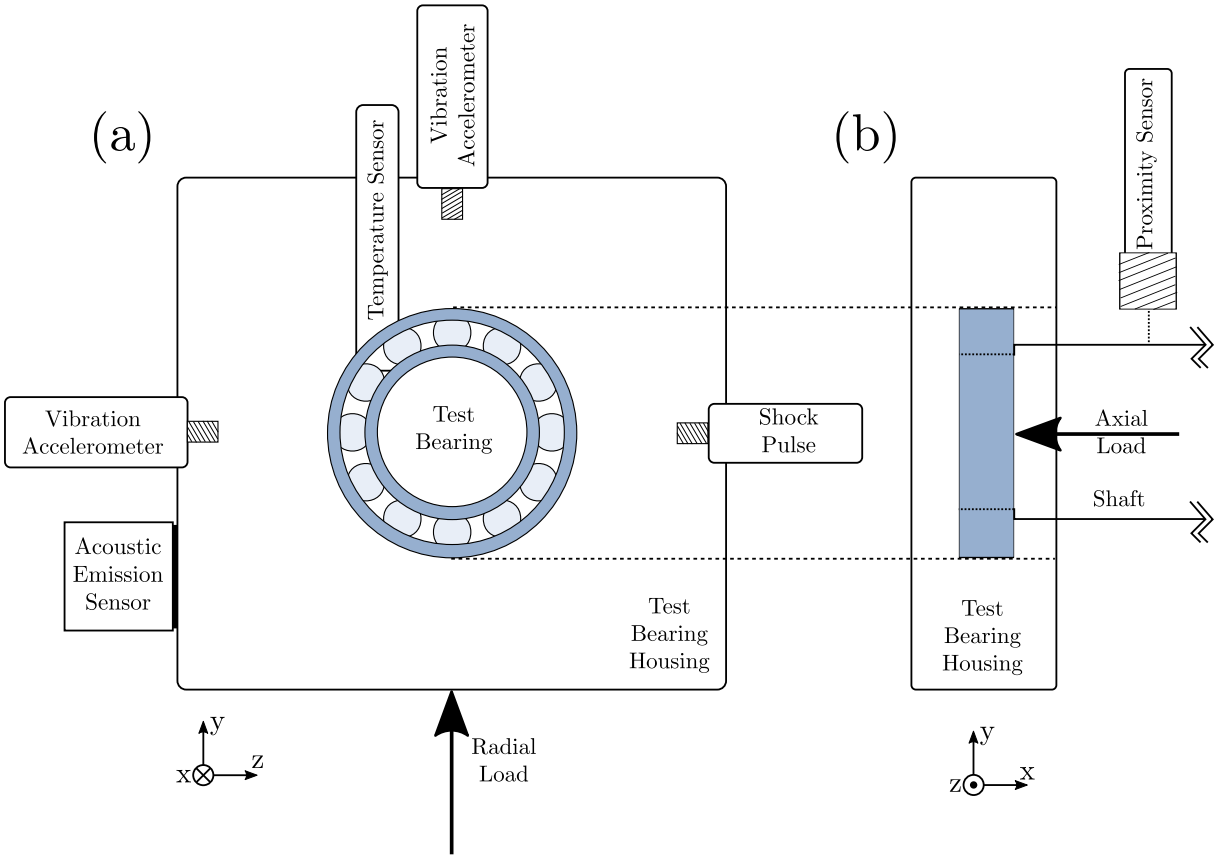

Figure A.4: The sensor layout on the test bearing housing: (a) Front of the housing with the bearing in the middle and sensors around it. (b) Side of the test bearing housing with the proximity sensor location.

bearing. Its curved area makes the load transfer from the lever easier. Figure A.5 shows the concentric unit combined with the thrust bearing. The concentric unit is stationary and acts as a bridge between the lever and the shaft.

From the specifications, at least 10 kN radial and axial loads must be available. Two electric linear actuators, each with a maximum force output of 2.5 kN, are used to apply the loads. Additionally, two levers are designed to amplify each linear actuator force to 17 kN. The high amplification was chosen to avoid loading each actuator 100% to achieve the required force, and to establish some flexibility for testing stronger bearings in the future. Figure A.6 shows the lever used to transfer axial load to the shaft. The two lengths L_1 and L_2 are determined in such a way that $L_1/L_2 = 17/2.5 = 6.8$. The concentric unit is located at the upper end of the lever, and the linear actuator at the bottom. The lever features a curved surface that ensures proper transfer of the force to the concentric unit. The revolution point A is a load cell bolt that can measure the reaction force in the hinge. The axial force is calculated based on the geometry of the lever and the reaction force feedback from the load cell. Section A.3.3 includes more details. The load cell is

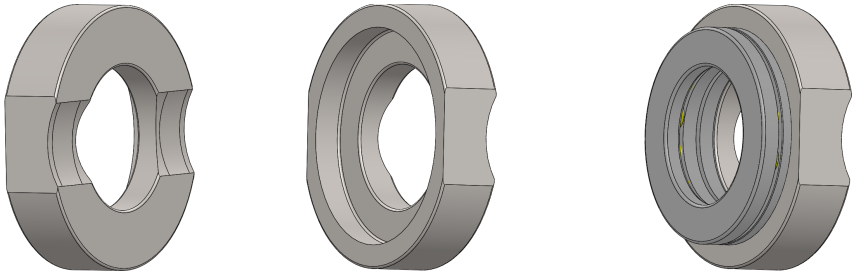

Figure A.5: Concentric unit housing the thrust bearing.

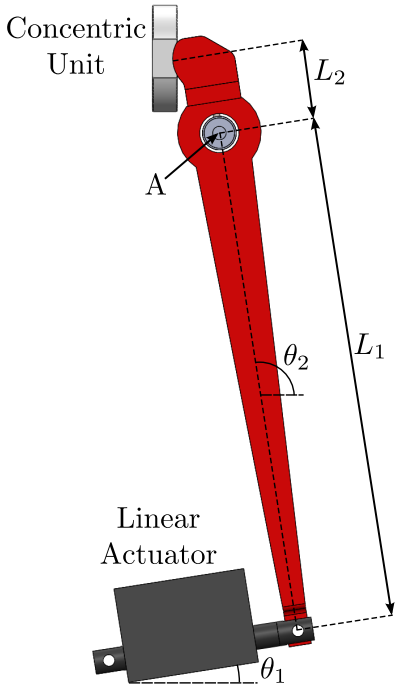

Figure A.6: Axial load setup.

supported by brackets bolted to the main steel structure. Figure A.7 shows the load cell connected to the lever on the axial load setup.

A similar setup is used to produce the radial load which is applied to the test bearing housing from underneath. A second linear actuator is connected to a lever which amplifies the radial load. Figure A.8 shows the radial load setup. The lever follows the aforementioned geometry ratio to amplify the linear actuator force to 17 kN by making $L_3/L_4 = 6.8$. A load cell bolt is used as the revolute joint to measure the reaction force in the hinge. Point B on Figure A.8 pushes the bottom of the test bearing housing when the actuator is retracting. The load cell is supported by two brackets as shown in Figure A.9. As shown, the lever pushes directly on the bottom of the housing, and the force propagates through the test bearing, to the shaft.

The added axial force propagates through the test bearing, to the test rig structure. The radial force is transferred through the test bearing, to the support bearing housings.

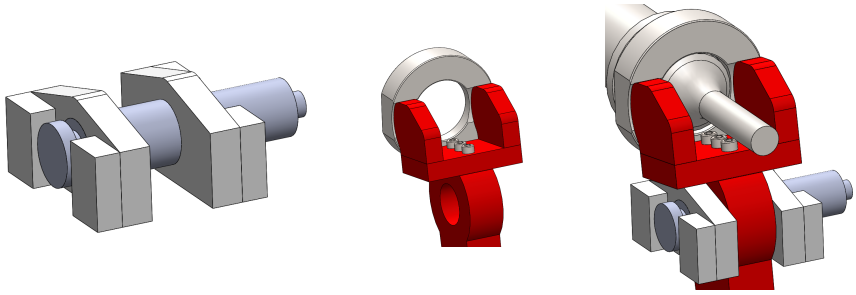

Figure A.7: Closeup details of the axial load setup.

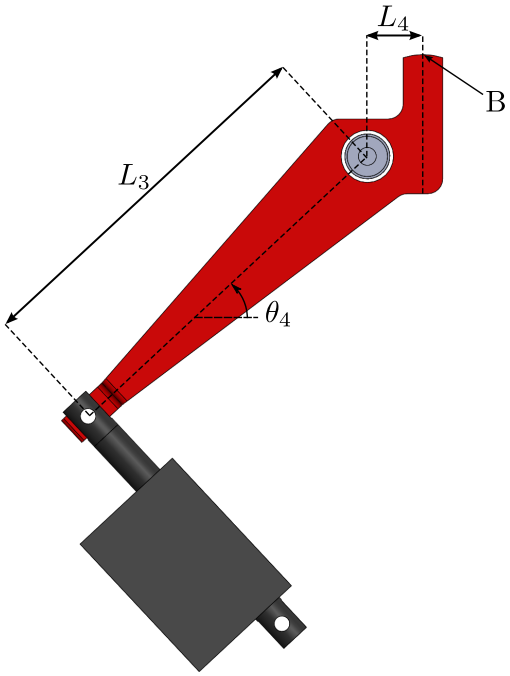

Figure A.8: Radial load setup.

Considering this, the test bearing housing require structural support in the axial direction, without interfering in the radial direction. Linear rolling element bearings on rails take care of this behavior. Using four suitable linear bearings between the housing and a support bracket, it is free to move in the vertical direction while supported in the axial direction. Figure A.10 shows the housing connected to the linear bearings.

A.3.3 Axial and Radial Loads

In the previous subsection, the linear actuator and lever combinations producing axial and radial loads were presented. Load cell bolts are used as lever hinges, and they measure the reaction force in a single direction, as indicated by an arrow on the bolt. The optimal measurement direction must be determined to measure the complete reaction force. During installation, it is important that the load cell orientations are correct. To this aim, the force equilibrium equations are established from a Free Body Diagram

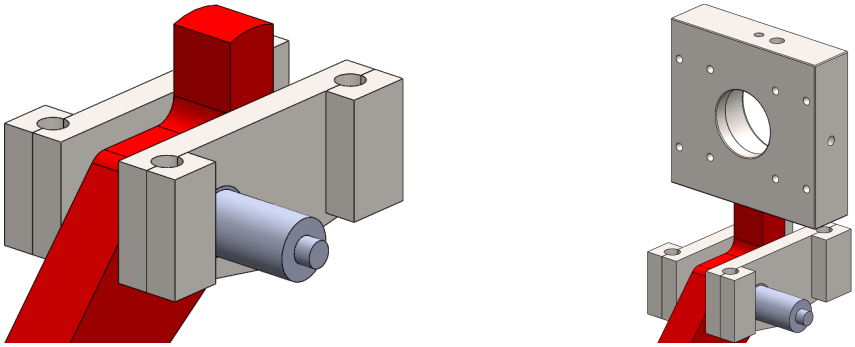

Figure A.9: Closeup details of the radial load setup.

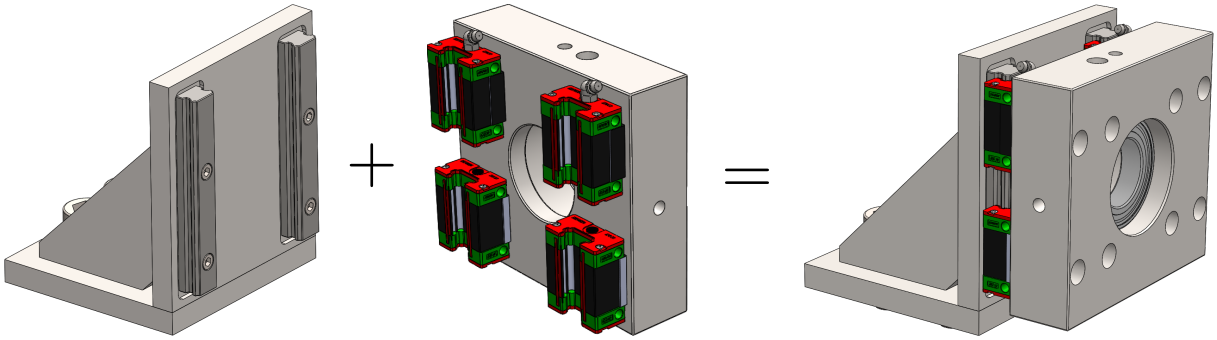

Figure A.10: Bracket with linear rolling element bearings connected to the test bearing housing.

(FBD). The FBD of the axial load setup is shown in Figure A.11. Here, three forces are acting on the lever, disregarding gravity: F_A is the axial force, R_1 is the load cell reaction force, and F_1 is the linear actuator force. θ_3 is the angular orientation of the load cell necessary to measure the entire reaction force R_1, i.e. the direction of this force vector. The force and moment equilibrium are given in (A.2) and (A.3) respectively.

$$\sum F = F_1 \begin{bmatrix} \cos\theta_1 \\ \sin\theta_1 \\ 0 \end{bmatrix} + R_1 \begin{bmatrix} \cos\theta_3 \\ \sin\theta_3 \\ 0 \end{bmatrix} + \begin{bmatrix} F_A \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{A.2}$$

$$\begin{aligned} \sum M = L_1 \begin{bmatrix} -\cos\theta_2 \\ -\sin\theta_2 \\ 0 \end{bmatrix} \times F_1 \begin{bmatrix} \cos\theta_1 \\ \sin\theta_1 \\ 0 \end{bmatrix} \\ + L_2 \begin{bmatrix} \cos\theta_2 \\ \sin\theta_2 \\ 0 \end{bmatrix} \times \begin{bmatrix} F_A \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \end{aligned} \tag{A.3}$$

where $\times$ represent the cross product in a right-hand coordinate system, $\sum M$ is the sum of moments, and $\sum F$ is the sum of forces, and the three dimensions are x, y, and z

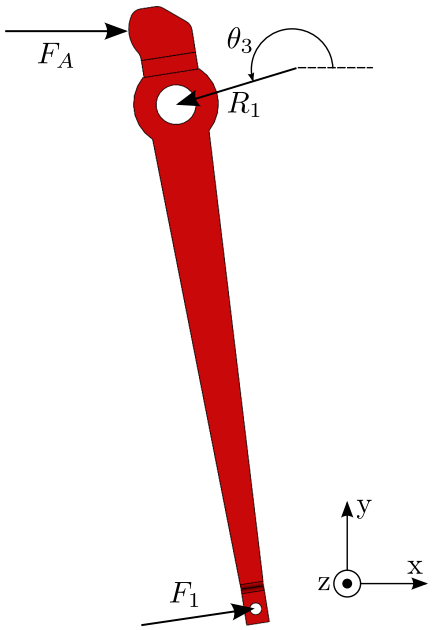

Figure A.11: Free Body Diagram of the axial load setup.

respectively. Evaluating (A.3) yields the following equilibrium in the z-direction

$$F_1 L_1(\sin\theta_2\cos\theta_1 - \cos\theta_2\sin\theta_1) - L_2 F_A \sin\theta_2 = 0\,. \tag{A.4}$$

Afterwards, (A.4) is rearranged to isolate the axial force

$$F_A = \frac{F_1 L_1}{L_2}\left(\cos\theta_1 - \frac{\sin\theta_1}{\tan\theta_2}\right)\,. \tag{A.5}$$

The linear actuator force is isolated from the y-dimension of (A.2)

$$F_1 = -R_1\frac{\sin\theta_3}{\sin\theta_1}\,. \tag{A.6}$$

Inserting (A.6) into (A.5) yields

$$F_A = \frac{R_1 L_1 \sin\theta_3}{L_2}\left(\frac{1}{\tan\theta_2} - \frac{1}{\tan\theta_1}\right)\,. \tag{A.7}$$

The x-dimension in (A.2) is used to determine the hinge reaction force angle

$$F_1\cos\theta_1 + R_1\cos\theta_3 + F_A = 0\,. \tag{A.8}$$

Inserting (A.6) and (A.7) into (A.8) yields

$$\begin{aligned} &- R_1\sin\theta_3\frac{\cos\theta_1}{\sin\theta_1} + R_1\cos\theta_3 \\ &+ R_1\sin\theta_3\frac{L_1}{L_2}\left(\frac{1}{\tan\theta_2} - \frac{1}{\tan\theta_1}\right) = 0\,. \end{aligned} \tag{A.9}$$

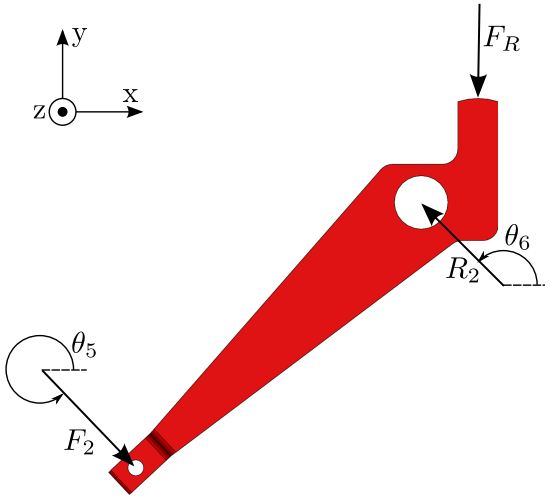

Figure A.12: Free Body Diagram of radial load setup.

Afterwards, (A.9) is rearranged by dividing all terms by $R_1 \cos\theta_3$, and isolating θ_3

$$\theta_3 = \arctan\left(\frac{1}{\frac{1}{\tan\theta_1} + \frac{L_1}{L_2}\left(\frac{1}{\tan\theta_1} - \frac{1}{\tan\theta_2}\right)}\right) . \tag{A.10}$$

θ_3 is a constant angle independent of the axial load magnitude, as it is only dependent on the geometry and orientation of the lever. During installation on the test rig, it is important that the rotational orientation of the load cell matches the angle θ_3, otherwise the measured reaction force is smaller than in reality. With the load cell in place, R_1 can be measured and recalculated into F_A using (A.7).

The FBD of the radial load setup is shown in Figure A.12. Here, F_R is the radial load, R_2 is the load cell reaction force, F_2 is the linear actuator force, and θ_6 is the angular orientation of the load cell. The force and moment equilibrium for the FBD are given in (A.11) and (A.12) respectively.

$$\sum F = F_2 \begin{bmatrix} \cos\theta_5 \\ \sin\theta_5 \\ 0 \end{bmatrix} + R_2 \begin{bmatrix} \cos\theta_6 \\ \sin\theta_6 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -F_R \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{A.11}$$

$$\sum M = L_3 \begin{bmatrix} -\cos\theta_4 \\ -\sin\theta_4 \\ 0 \end{bmatrix} \times F_2 \begin{bmatrix} \cos\theta_5 \\ \sin\theta_5 \\ 0 \end{bmatrix} + \begin{bmatrix} L_4 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ -F_R \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{A.12}$$

Evaluating (A.12) yields the following equilibrium in the z-direction,

$$L_3 F_2(\sin\theta_4\cos\theta_5 - \cos\theta_4\sin\theta_5) - L_4 F_R = 0\,. \tag{A.13}$$

Afterwards, (A.13) is rearranged to isolate F_R:

$$F_R = \frac{F_2 L_3}{L_4}(\sin\theta_4\cos\theta_5 - \cos\theta_4\sin\theta_5)\,. \tag{A.14}$$

The linear actuator force is isolated from the x-dimension in (A.11) as

$$F_2 = -R_2\frac{\cos\theta_6}{\cos\theta_5}\,. \tag{A.15}$$

Inserting (A.15) into (A.14) yields

$$F_R = \frac{R_2\cos\theta_6 L_3}{L_4}(\cos\theta_4\tan\theta_5 - \sin\theta_4)\,. \tag{A.16}$$

The y-dimension in (A.11),

$$F_2\sin\theta_5 + R_2\sin\theta_6 - F_R = 0\,, \tag{A.17}$$

is used to determine the hinge reaction force angle. Inserting (A.15) and (A.16) into (A.17) yields

$$\begin{aligned} &-\frac{R_2\cos\theta_6}{\cos\theta_5}\sin\theta_5 + R_2\sin\theta_6 \\ &+\frac{R_2\cos\theta_6 L_3}{L_4}(\sin\theta_4 - \cos\theta_4\tan\theta_5) = 0\,. \end{aligned} \tag{A.18}$$

θ_6 is isolated in (A.18) by dividing all terms by $R_2\cos(\theta_6)$, and rearranging the result as

$$\theta_6 = \arctan\left(\frac{L_3}{L_4}(\cos\theta_4\tan\theta_5 - \sin\theta_4) - \tan\theta_5\right)\,. \tag{A.19}$$

Here, θ_6 is a constant angle independent of the radial load, and is only dependent on the geometry and orientation of the lever. After installing the load cell bolt with the orientation given by θ_6, the radial force F_R is calculated using the measured hinge reaction force in (A.16).

A.3.4 Complete test rig

Based on the sub function solutions in the previous sections, a complete design of the test rig is created. The chosen solutions achieve the different sub functions; however, the final design shows how the components are placed on a suitable steel structure. The complete test rig is shown in Figure A.13. Details of all numbered components are given in the following: To drive the shaft, a permanent magnet synchronous AC motor (1) is

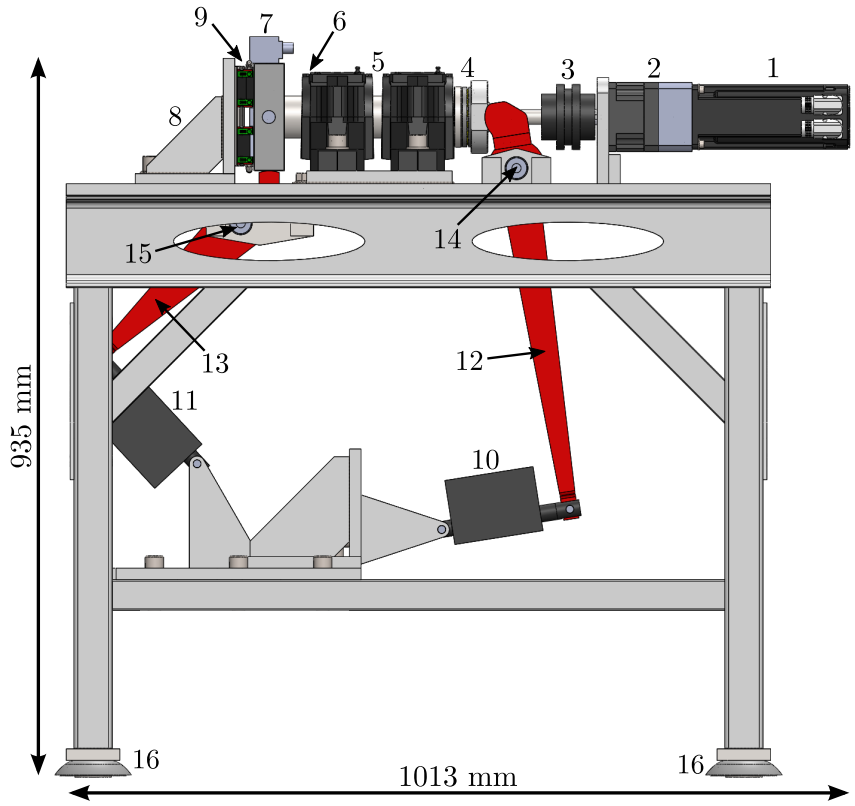

Figure A.13: Complete 3D model with numbered components. The width of the test rig is 350 mm.

chosen. It is powered by a variable frequency drive to realize variable-speed conditions. A 1:7 ratio planetary gearbox (2) is installed to aid the motor during low-speed conditions. This combination enables the shaft to reach speeds up to 500 rpm and down to 20 rpm.

A flexible claw coupling (3) that does not transfer axial load is installed to provide shaft rotation. A thrust bearing (4) is installed to transfer axial load to the shaft.

Two bearings (5) are used to counteract the radial load subjected to the test bearing, and they are free to slide in the axial direction inside their housings. On the left end of the shaft at position (6), a proximity sensor is mounted to measure the radial movement of the shaft close to the test bearing.

The test bearing housing (7) is made of steel and is currently manufactured to house a 6008-type ball bearing. Other similarly sized bearings with the same bore diameter may be used, provided that a new housing is created if the outer diameter is different. An accelerometer is placed on the side of the test bearing housing, and the outer ring temperature is measured with a tip sensitive PT100 resistance temperature detector. Other sensors are also available for installation as illustrated in Figure A.4.

The support bracket (8) is attached to the test bearing housing via linear ball element bearings (9). These bearings allow for vertical load transfer to the test bearing, while

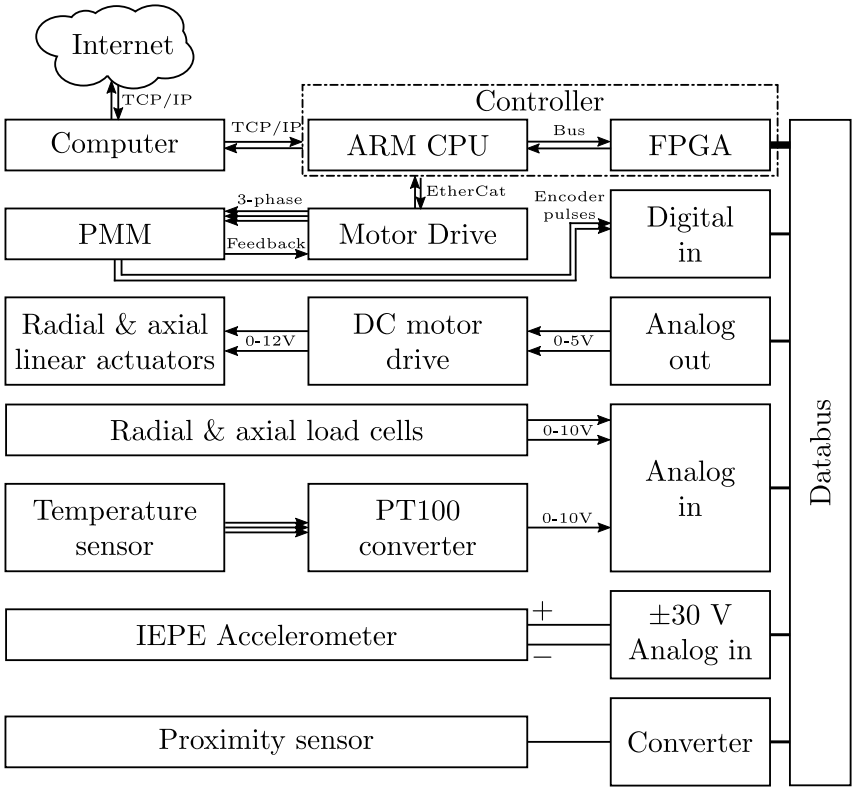

Figure A.14: Electronic hardware connections.

they support the shaft in the axial direction. The bracket itself is bolted to the structure.

(10) and (11) are electric linear actuators. They can produce loads up to 2.5 kN in their axial direction. Coupled with the levers (12) and (13), the system can theoretically subject the bearing with 17 kN of load in both axial and radial direction.

Two load cells, (14) and (15), act as hinges for the levers while measuring the respective reaction force up to 20 kN.

Machine damping feet (16) are used to reduce the vibration coming from external sources, and their height can also be adjusted to level out the rig on slightly uneven surfaces.

A.3.5 Controller hardware

Electronic hardware is used to control the active components, convert between analog and digital signals, and store sensor data for further processing. A detailed overview of the hardware is shown in Figure A.14. A computer with internet access is connected to a dedicated controller with a built-in ARM CPU and Field Programmable Gate Array (FPGA). The main task of the computer is to provide a local and remote interface to the test setup software, and to store data during measurement intervals. A program on the ARM CPU controls all the active components based on the feedback signals and test

configuration. The FPGA is responsible for acquiring and sending the analog signals provided by the various IO via a databus.

The permanent magnet motor (PMM) is controlled by a motor drive with a built-in regulator, and requires only a speed reference from the controller via a real-time EtherCat connection. The shaft angular position is acquired from a quadrature encoder in the motor. The electric linear actuators are powered by a 12 V DC motor drive, and the reference signal to the drive is between 0-5 V. An input of 0 V corresponds to full retraction force, 5 V to full push force, and 2.5 V for no force. The test-bearing temperature and load cell reaction forces are acquired via an analog-in module. A 24-bit, high-frequency ±30 V IEPE-compatible analog input module measures the vibration in the accelerometer attached to the test-bearing housing. The accelerometer produces a signal of 100 mV/g in the linear range up to 10 kHz. In addition, a proximity sensor measures the radial shaft movement in one direction. It has a measurement range of 1.1 mm and a resolution of 18.5 nm, which makes it capable of detecting small changes in shaft position when the bearing is damaged.

A.3.6 Load controller design

The electric linear actuators are controlled to apply the correct radial and axial loads to the test bearing. The choice of control system is not trivial due to several factors: 1) The linear actuator consists of a DC motor connected to a screw joint which moves in and out when the motor rotates. The friction between the screw and DC motor is considerable, and the actuator will therefore self-lock when not active. 2) The recommended duty cycle is less than 10% to avoid overheating. Specifically, the manufacturer recommends a maximum of 2 minute operation per 20 minutes. 3) Experimental testing on the finished test rig revealed that the measured axial and radial loads varies with the shaft angular position. The variation resembles a sine wave with a certain amplitude if the shaft speed is constant. This variation is most likely caused by asymmetries on the shaft. It may be slightly bent from the manufacturing process, or perhaps not perfectly circular. Due to this disturbance, the actuators are difficult to control in a closed loop.

In any case, care must be taken not to overload the bearing or overheat the actuators. A PID controller would attempt to cancel out the cyclic disturbance from the shaft, risking overshoots and overheating. Using small gains would be a possibility, however that will result in a poor bandwidth.

The self-locking ability is instead exploited to create a simple control system similar to *Sliding Mode Control*. For each actuator setup, the error e between the reference load $F^{(ref)}$ and the actual load F is continuously checked. If this error is outside a threshold T, the actuator is activated with maximum power ($u = \pm 1$) to reduce the error. Once

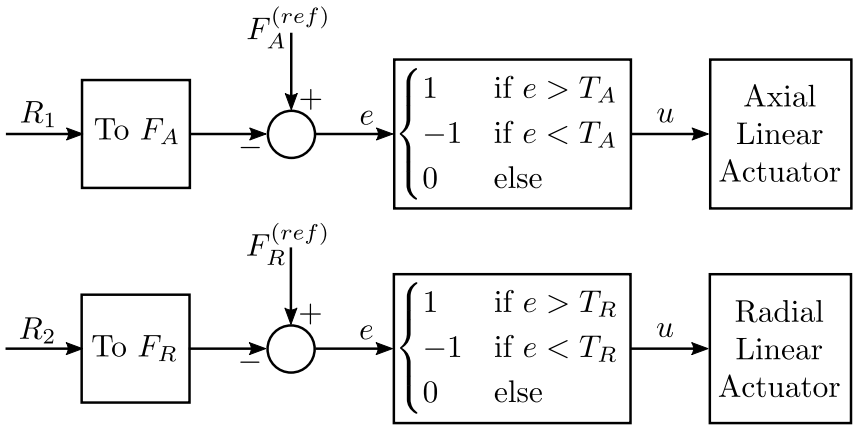

Figure A.15: Control scheme for both linear actuators.

the error is within the threshold, the actuator is deactivated ($u = 0$). By configuring this threshold slightly larger than the amplitude of the shaft disturbance, the actuator is only powered in a short period after load reference change. The control scheme for each actuator load setup is shown in Figure A.15. Here, R_1 and R_2 denotes the measured reaction force in the axial and radial load cell, respectively. The block "To F" determines the axial or radial load using Equation (A.7) or (A.16), respectively.

It should be noted that this control scheme is not suitable for following a continuously changing reference signal, i.e. a sine wave. For such an input, the controller will actuate the system in steps when the error exceeds the threshold, resulting in a step-wise load change.

A.3.7 Accelerated life-time test software design

An accelerated bearing life-time test may last for several weeks, or even months. Therefore, it is impractical to manually make measurements at a defined interval, or to monitor the test in person. To overcome this, an automated lab test program is developed. This program controls all the active components such as the motor and the linear actuators, and initiates sensor data logging at pre-defined intervals. The flowchart of the test program is shown in Figure A.16. Initially, the operator starts by preparing n_c configurations (configs for short) that the program runs through. Each config contains specifications for: the shaft speed, bearing load, and the duration. The duration is specified in either number of revolutions (revs) if sensor data is logged, or number of minutes if not. Afterwards, the test may start from the first config, and after it is finished, the next config is loaded and performed. Once all configs are finished, the program resets to the first one.

The test is stopped automatically if the vibration root mean square (RMS) exceeds a

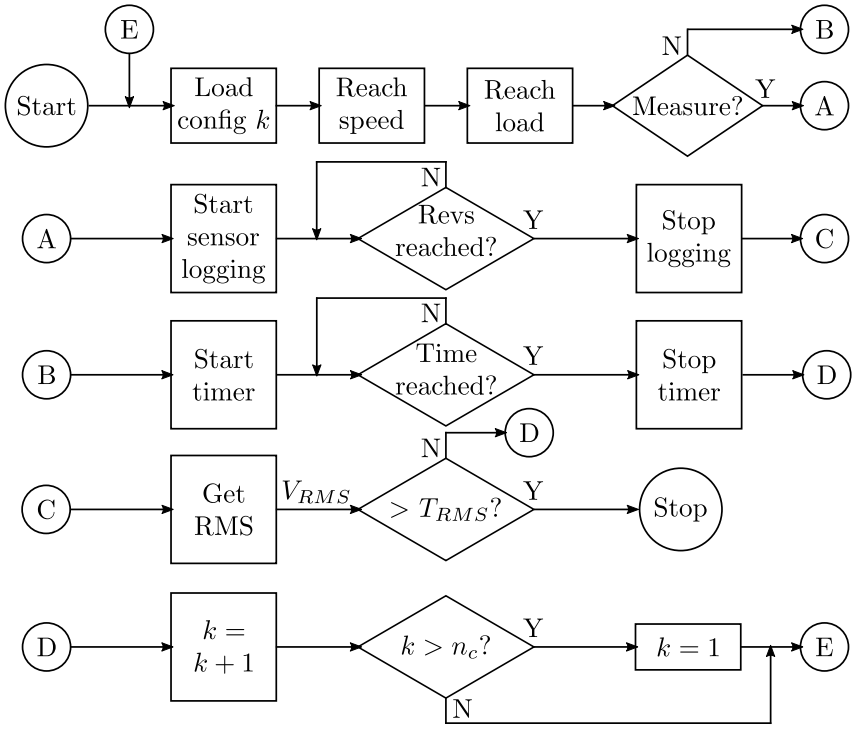

Figure A.16: Control scheme for the accelerated life-test.

preset threshold T_{RMS}. The RMS is calculated using

$$V_{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} V_i^2}\,, \tag{A.20}$$

where n is the number of samples in the dataset, and V_i denotes the i'th sample in the dataset V. Once the RMS reaches the threshold, the bearing is considered completely worn out, and the test is automatically stopped. It is also possible to use different sources for stopping the test such as motor torque or bearing temperature. Additionally, trends of the RMS, motor torque, and bearing temperature is automatically uploaded to a secure folder on the internet, and may be observed remotely by the operator. Further, the test computer may also be accessed via a secure remote desktop connection in case the test must be stopped/changed outside of working hours. Using this setup, the test rig is safe to operate 24/7 without supervision.

A.3.8 Case study

As described in Section A.2.1 the test rig is used to perform accelerated life-time tests on bearings. In addition to being able to run at high-speed to degenerate the bearing faster, it also operates at slower speeds to obtain low-speed vibration data. For this accelerated life-time test, the following options are set:

- The shaft is set to run at 500 rpm to wear out the test bearing.

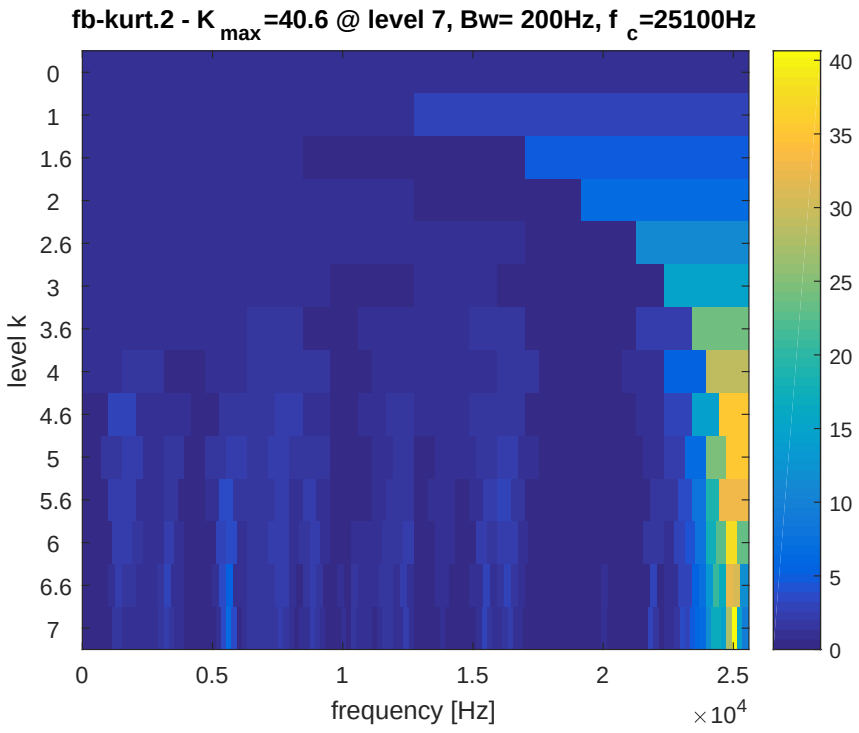

Figure A.17: Resulting kurtogram.

- The radial load is set to 9 kN, and the axial load to 7 kN. This results in a nominal lifetime of $L_{10} = 6$ million revolutions according to the bearing manufacturer's online calculator [25].

- The bearing vibration is recorded at a sample rate of 51.2 kHz, while the bearing temperature, motor torque, and shaft angular position are recorded at a sample rate of 512 Hz.

- Sensor datasets are obtained every 10 minutes at 500, 250, 100, 50, and 20 rpm, successively. The measurement duration is set to 100 revolutions.

- The automatic stop was set to an RMS threshold of 3 g (1 g = 9.81 m/s^2).

A summary of the configuration loaded into the test software is given in Table A.1.

Two weeks and approximately 6 million revolutions later, the bearing is sufficiently damaged that the RMS threshold is triggered. To diagnose the bearing for faults, the datasets acquired at 50 rpm are used on this occasion. To determine the fault condition in the bearing, a state-of-the-art method named the Fast Kurtogram [5] is employed. The method decomposes the vibration data into frequency narrow-bands at different central frequencies and widths. The kurtosis of each decomposed signal is calculated using

$$\text{Kurt}\{x\} = \frac{\mu_4}{\sigma^4}\,, \tag{A.21}$$

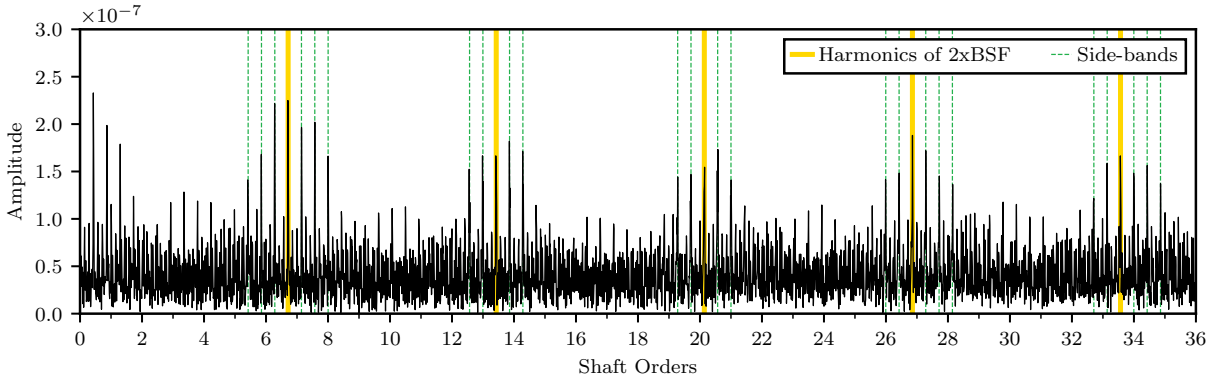

Figure A.18: Squared Envelope Spectrum after bandpass filtering.

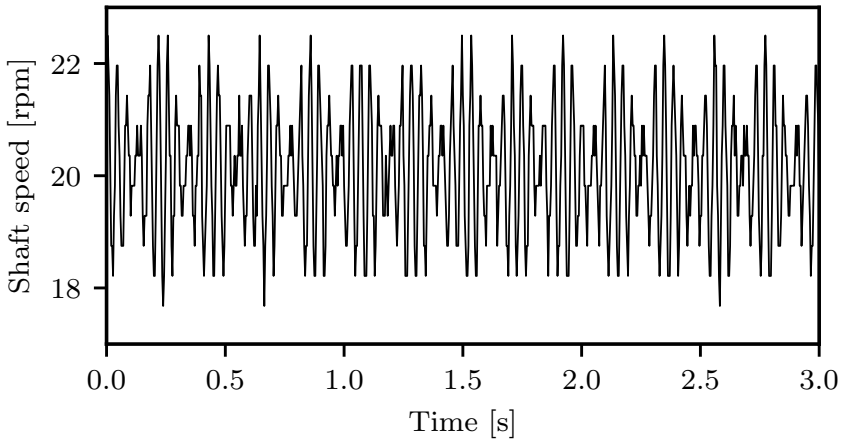

Figure A.19: Shaft rotational speed at a reference speed of 20 rpm.

where μ_4 is the fourth central moment, and σ is the standard deviation. A high kurtosis value imply an impulsive signal which resembles bearing fault impact vibration. The frequency narrow-band with the highest kurtosis is therefore chosen as the optimal filter, and this band should contain the bearing fault impact vibration. Using this method, the kurtogram is employed at every dataset acquired at 50 rpm shaft speed from the accelerated life-time test. After the signal is filtered using the optimal filter, it should contain the high-frequency resonance vibration from the bearing. To identify the cyclic frequencies in the signal, the signal is demodulated. Using the Hilbert transform, the complex-valued analytic signal is obtained. By computing the absolute value of this analytic signal, the envelope is obtained. The envelope spectrum (ES) is the Fourier transform of this envelope, which contain all cyclic impact frequencies. It has been shown that the envelope should be squared to remove extra peaks in the spectrum [26], hence resulting in the squared ES (SES). The resulting SES after bandpass filtration is manually analyzed to identify prominent peaks at the characteristic fault frequencies for the present bearing, which are given in Table A.2. In this table, BPFI is the ball pass frequency inner

Table A.1: Operator settings for the test ($n_c = 6$).

Conf.	Speed	F_A	F_R	Rec.?	Duration
1	500 rpm	7 kN	9 kN	No	10 min
2	500 rpm	7 kN	9 kN	Yes	100 revs
3	250 rpm	7 kN	9 kN	Yes	100 revs
4	100 rpm	7 kN	9 kN	Yes	100 revs
5	50 rpm	7 kN	9 kN	Yes	100 revs
6	20 rpm	7 kN	9 kN	Yes	100 revs

Table A.2: Characteristic bearing fault frequencies in shaft orders.

BPFI	BPFO	FTF	BSF
6.88	5.12	0.43	3.33

race, BPFO is the ball pass frequency outer race, FTF is the fundamental train frequency, and BSF is the ball spin frequency. One shaft order is defined as the shaft speed, and ball faults are shown at 2xBSF as there is an impact at the inner and outer race successively during one spin. Equations for calculating these characteristic frequencies are given in [27]. These characteristic bearing frequencies are of interest because they describe how many times an incipient fault on a certain location is passed on each shaft revolution. The earliest sign of bearing fault is identified after 5.37 million revolutions, and in this case, a ball fault is progressing. The kurtogram is shown in Figure A.17, where the width of each frequency band is given as levels in the y-axis, and the central frequency is given on the x-axis. The kurtosis values are displayed as colors, and the maximum kurtosis of 40.6 is identified at the central frequency 25,100 Hz with a bandwidth of 200 Hz (level 7). The resulting SES after band-pass filtering is shown in Figure A.18. Here, integer multiples of 2xBSF (harmonics) are marked as yellow lines together with side-bands marked as green stapled lines. The side-bands spaced apart by the FTF are shown due to the amplitude modulation from the non-homogeneous radial load [27]. As there are multiple harmonics with accompanying side-bands close to the theoretical 2xBSF, it is likely that the vibration is caused by a ball fault.

The fast Kurtogram filtered SES from a dataset recorded at 20 rom was also analyzed, however it did not show prominent peaks at this low speed. More advanced processing methods are required for diagnosing the bearing during this low-speed working condition. To check if the problem of diagnosing the bearing at 20 rpm is due to speed fluctuations, the shaft speed for the 10 first revolutions is derived from encoder data, and shown in Figure A.19. Accordingly, the speed fluctuations are within ±10%, which may deteriorate

the diagnostic capabilities of the fast Kurtogram slightly. Angular re-sampling should be applied to reduce the blurring effect of the speed fluctuation [27].

A.3.9 Future work

In the case study in Section A.3.8, it was shown that early fault detection is possible using the fast Kurtogram [5] at 50 rpm, however not at 20 rpm. The main purpose of this case study, however, is to demonstrate that the test rig is capable of accelerating the life-time of the bearing, and that the acquired signals may be used for fault diagnosis. Future research will focus on new methods and algorithms for fault detection at the low speed of 20 rpm.

After determining the initial fault, it is advisable to predict the remaining useful life-time (RUL) using prognostics methods. These models are important to help schedule maintenance and avoid unexpected breakdowns. Typically, extracted features from sensor data is used to create a prognostic model, and a variety of models and methods have already been suggested in the literature [28, 29]. However, there are still challenges to overcome in designing a standardized scheme that can be applied on any machine [30]. Additionally, one apparent challenge in rotating machinery prognostics is to make accurate estimations during variable speed or load conditions [31].

On the presented test rig, the bearing temperature and vibration, motor torque, and shaft movement is measured for the whole life-span of the bearing during the accelerated life-time test. Further, the working speed of the bearing is also changed between each measurement, and may even be variable during the measurement period. In comparison to bearing test rigs relying on pre-seeded faults, accelerated life-time tests allows for tracking the changes in sensor data features as time/wear progresses, making RUL estimation possible. The acquired sensor data will be used in future research to develop algorithms which estimate RUL for machines operating under variable conditions.

A.4 Conclusion

An accelerated life-time test rig for rolling element bearings have been developed in this paper. Using electric actuators and levers, the test bearing can be subjected with heavy radial and axial loads. Additionally, utilizing a variable frequency drive, a motor, and a planetary gearbox, both high and low speed working conditions are realizable. Sensors are installed to measure physical data on the test bearing including vibration, shaft radial movement, and temperature. Tests show that the test rig can be used to generate low-speed vibration data during all bearing states ranging from healthy to completely worn. Therefore, all the design criteria from Section A.2.2 are fulfilled. The list below shows

conclusive remarks of the design specifications, where ($\surd$) means fulfilled, ($\approx \surd$) means partially fulfilled, and ($\times$) means not fulfilled:

1. ($\surd$) - Results show that the low speed of 20 rpm is achievable using a permanent magnet motor and a planetary gearbox. There is a small fluctuation of ± 2 rpm, which may be corrected using digital angular re-sampling.

2. ($\surd$) - Theoretically, the load setups can apply up to 17 kN of load in the radial and axial direction.

3. ($\surd$) - The flat edges of the test bearing housing allow for using all the requested sensor types.

4. ($\surd$) - The test bearing housing can house the 6008-type bearing with 68mm outer diameter.

5. ($\surd$) - Using suitable hardware and software, the test rig may be operated safely 24/7 without supervision.

REFERENCES

[1] S. T. Kandukuri, A. Klausen, H. R. Karimi, and K. G. Robbersmyr. A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management. *Renewable and Sustainable Energy Reviews*, 53:697–708, 2016. doi:10.1016/j.rser.2015.08.061.

[2] C. Junsheng, Y. Dejie, and Y. Yu. A fault diagnosis approach for roller bearings based on EMD method and AR model. *Mechanical Systems and Signal Processing*, 20(2):350–362, 2006. doi:10.1016/j.ymssp.2004.11.002.

[3] J. Antoni and R. B. Randall. Unsupervised noise cancellation for vibration signals: part I-evaluation of adaptive algorithms. *Mechanical Systems and Signal Processing*, 18(1):89–101, 2004. doi:10.1016/S0888-3270(03)00012-8.

[4] J. Zarei, M. A. Tajeddini, and H. R. Karimi. Vibration analysis for bearing fault detection and classification using an intelligent filter. *Mechatronics*, 24(2):151–157, 2014. doi:10.1016/j.mechatronics.2014.01.003.

[5] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007. doi:10.1016/j.ymssp.2005.12.002.

[6] R. Peled, S. Braun, and M. Zacksenhouse. A blind deconvolution separation of multiple sources, with application to bearing diagnostics. *Mechanical Systems and Signal Processing*, 19(6):1181–1195, 2005. doi:10.1016/j.ymssp.2005.08.019.

[7] P. Shakya, A. K. Darpe, and M. S. Kulkarni. Bearing diagnosis using proximity probe and accelerometer. *Measurement*, 80:190–200, 2016. doi:10.1016/j.measurement.2015.11.029.

[8] N. Sawalhi and R. B. Randall. Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults. *Mechanical Systems and Signal Processing*, 22(8):1924–1951, 2008. doi:10.1016/j.ymssp.2007.12.001.

[9] R. Kumar and M. Singh. Outer race defect width measurement in taper roller bearing using discrete wavelet transform of vibration signal. *Measurement*, 46(1):537–545, 2013. doi:10.1016/j.measurement.2012.08.012.

[10] D. Siegel, H. Al-Atat, V. Shauche, L. Liao, J. Snyder, and J. Lee. Novel method for rolling element bearing health assessment—A tachometer-less synchronously averaged envelope feature extraction technique. *Mechanical Systems and Signal Processing*, 29:362–376, 2012. doi:10.1016/j.ymssp.2012.01.003.

[11] N. Sawalhi and R. B. Randall. Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size. *Mechanical Systems and Signal Processing*, 25(3):846–870, 2011. doi:10.1016/j.ymssp.2010.09.009.

[12] S. A. Niknam, V. Songmene, and Y. H. J. Au. The use of acoustic emission information to distinguish between dry and lubricated rolling element bearings in low-speed rotating machines. *International Journal of Advanced Manufacturing Technology*, 69(9-12):2679–2689, 2013. doi:10.1007/s00170-013-5222-4.

[13] T. R. Lin, E. Kim, and A. C. C. Tan. A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm. *Mechanical Systems and Signal Processing*, 36(2):256–270, 2013. doi:10.1016/j.ymssp.2012.11.003.

[14] N. Jamaludin, D. Mba, and R. H. Bannister. Condition monitoring of slow-speed rolling element bearings using stress waves. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 215(4):245–271, 2001. doi:10.1177/095440890121500401.

[15] B.-Q. Fan, K.-M. Lee, X.-P. Ouyang, and H.-Y. Yang. Soft-switchable dual-pi controlled axial loading system for high-speed emu axle-box bearing test rig. *IEEE Transactions on Industrial Electronics*, 62(12):7370–7381, 2015. doi:10.1109/TIE.2015.2458303.

[16] B. Muruganatham, M. A. Sanjith, B. Krishnakumar, and S. A. V. S. Murty. Roller element bearing fault diagnosis using singular spectrum analysis. *Mechanical Systems and Signal Processing*, 35(1):150–166, 2013. doi:10.1016/j.ymssp.2012.08.019.

[17] J. B. Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics*, 89:16–27, 2015. doi:10.1016/j.apacoust.2014.08.016.

[18] B. Zhang, C. Sconyers, C. Byington, R. Patrick, M. E. Orchard, and G. Vachtsevanos. A probabilistic fault detection approach: Application to bearing fault detection. *IEEE Transactions on Industrial Electronics*, 58(5), 2011. doi:10.1109/TIE.2010.2058072.

[19] J. Yu. Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. *Mechanical Systems and Signal Processing*, 25(7):2573–2588, 2011. doi:10.1016/j.ymssp.2011.02.006.

[20] R. Li, P. Sopon, and D. He. Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*, 23(2):313–321, 2012. doi:10.1007/s10845-009-0353-z.

[21] H. K. Danielsen, F. G. Guzmán, K. V. Dahl, Y. J. Li, J. Wu, G. Jacobs, G. Burghardt, S. Fæster, H. Alimadadi, S. Goto, et al. Multiscale characterization of White Etching Cracks (WEC) in a 100Cr6 bearing from a thrust bearing test rig. *Wear*, 370:73–82, 2017. doi:10.1016/j.wear.2016.11.016.

[22] H. Qiu, J. Lee, J. Lin, and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006. doi:10.1016/j.jsv.2005.03.007.

[23] W. Caesarendra, P. B. Kosasih, A. K. Tieu, C. A. S. Moodie, and B.-K. Choi. Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition. *Journal of Mechanical Science and Technology*, 27(8):2253–2262, 2013. doi:10.1007/s12206-013-0608-7.

[24] M. Elforjani and D. Mba. Accelerated natural fault diagnosis in slow speed bearings with acoustic emission. *Engineering Fracture Mechanics*, 77(1):112–127, 2010. doi:10.1016/j.engfracmech.2009.09.016.

[25] SKF. Bearing lifetime calculator, (Link accessible October 2017). URL: http://webtools3.skf.com/BearingCalc/.

[26] D. Ho and R. B. Randall. Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*, 14(5):763–788, 2000. doi:10.1006/mssp.2000.1304.

[27] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

[28] Y. Peng, M. Dong, and M. J. Zuo. Current status of machine prognostics in condition-based maintenance: a review. *International Journal of Advanced Manufacturing Technology*, 50(1):297–313, 2010. doi:10.1007/s00170-009-2482-0.

[29] F. Ahmadzadeh and J. Lundberg. Remaining useful life estimation: review. *International Journal of System Assurance Engineering and Management*, 5(4):461–474, 2014. doi:10.1007/s13198-013-0195-0.

[30] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel. Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42(1):314–334, 2014. doi:10.1016/j.ymssp.2013.06.004.

[31] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew. Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23(3):724–739, 2009. doi:10.1016/j.ymssp.2008.06.009.

Paper B

Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum

Andreas Klausen, Kjell G. Robbersmyr and Hamid Reza Karimi

This paper has been published as:

A. Klausen, K. G. Robbersmyr and H. R. Karimi. Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum. *IFAC-PapersOnLine*, 50(1):13378-13383, 2017. doi: 10.1016/j.ifacol.2017.08.2262.

Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum

Andreas Klausen*, Kjell G. Robbersmyr* and Hamid Reza Karimi**

*University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

**Politecnico di Milano
Department of Mechanical Engineering
Via La Masa 1, 20156 Milan, Italy

Abstract **– Rolling element bearings are one of the fundamental components of a machine, and their failure is the most frequent cause of machine breakdown. Monitoring the bearing condition is vital to preventing unexpected shutdowns and improving their maintenance planning. Specifically, the bearing vibration can be measured and analyzed to diagnose bearing faults. Accurate fault diagnosis can be achieved by analyzing the envelope spectrum of a narrow-band filtered vibration signal. The optimal narrow-band is centered at the resonance frequency of the bearing. However, how to determine the optimal narrow-band is a challenge. Several methods aim to identify the optimal narrow-band, but they are not always precise. The bearing fault vibration components are lost if the narrow-band is incorrectly chosen, thus leading to an incorrect fault diagnosis. For on-line systems, it is critical that bearing faults are diagnosed with a high degree of confidence. In this article, a method for analyzing multiple narrow bands is presented. Bearing faults are detected autonomously by a narrow-band envelope spectrum-based algorithm. This algorithm removes the need for manual spectrum analysis, allowing operators to focus on more important tasks. Bearing fault vibration data from an accelerated life-test is used to verify the performance of the proposed method. The proposed method accurately diagnoses the worn-out bearing for three characteristic defect types and shows when one fault propagates to a second one.**

B.1 Introduction

Rolling Element Bearings (REBs, or bearings for short) are used in all kinds of rotating machinery. They are designed to restrict the shaft rotational motion for transferring loads to stationary housings. Bearings are worn out after exceeding a certain life-time based on the rotational speed, the combined loads, and the bearing design. When a bearing suddenly fails, the rapidly increasing vibration may damage other components, causing a complete breakdown of the machine. Replacing a bearing before it fails is of great importance to avoid breakdowns and expensive overhauls. The exact life-time of a bearing cannot be perfectly predicted because the failure rate of a particular bearing is based on statistics. In addition, changes in bearing vibration is only detectable during the final stage of the life-time. Therefore it is useful to continually monitor the condition of critical components such as bearings to prevent sudden machine breakdowns. Sensors can be placed on the bearing housing to measure a physical quantity containing information about the condition. The bearing vibration signal is strongly linked to the amount of wear and is a suitable signal to use for bearing condition monitoring.

Roller impacts in a damaged bearing produce amplitude modulated ringing with the carrier frequency equal to the resonance frequency of the bearing. The frequency of the modulation waveform reveals the fault location in the bearing. As such, the vibration signal must be demodulated to diagnose the fault location and severity. After demodulation, the frequency spectrum can be analyzed to diagnose bearing faults. [1] used the Hilbert-Huang transform to demodulate the vibration signal and diagnosed bearing faults using the frequency spectrum. The bearing characteristic frequencies are derived from the bearing design and kinematics under the no-slip assumption. [2] calculated these characteristic frequencies and diagnosed the damaged bearing by observing the envelope spectrum for peaks at these frequencies. The vibration signal contains components that are not linked to impacts in the bearing, but from other sources such as mass unbalance, misalignment, and others. These noisy components can be removed by applying a band-pass filter on a narrow-band around the resonance frequency of the bearing. The result is a much clearer frequency diagram to analyze. However, it is difficult to determine the optimal narrow-band analytically, but several attempts in the literature have been made to determine it experimentally. [3] introduced the Fast Kurtogram which uses the kurtosis to determine the optimal narrow-band. The vibration signal is band-pass filtered at various narrow-bands, and the kurtosis value determines the extent of impulsiveness in the filtered data. The filter specification returning the highest kurtosis can be selected as the optimal narrow-band for further envelope spectrum analysis. However, while there are cases when the kurtogram detects the optimal narrow-band, there are also cases when it fails. [4] presented a new method inspired by the kurtogram that attempts to overcome

some of the drawbacks of the fast kurtogram. It is named the Protrugram and it calculates the kurtosis of the envelope spectrum of the band-pass filtered signal rather than the time-signal. The advantage is the ability to detect transients with a small signal-to-noise ratio. [5] presented a new feature named envelope harmonic-to-noise ratio (EHNR) that can replace the kurtosis calculation in the fast kurtogram. Its advantage over the kurtosis is that it is not sensitive to random single impulses and is therefore more robust.

The methods existing in literature aim at identifying the optimal narrow-band. However, incorrect selection leads to filtering out components related to the bearing fault. Also, the bearing fault impacts may excite more than one resonance frequency in the system, and thus multiple narrow-bands should be investigated on a single vibration dataset. In this article, we present a new method for diagnosing faults in a bearing, based on narrow-band envelope spectrum analysis. Initially, the entire vibration frequency spectrum is divided into multiple narrow-bands and band-pass filtered. Next, the envelope spectrum of each filtered signal is realized by applying the Hilbert-Huang transform and the Fast Fourier transform. Each spectrum is afterwards analyzed to diagnose the bearing for faults. Since there are multiple spectra to analyze, the task becomes impractical to perform manually. Therefore, a method for autonomous bearing fault diagnosis is presented in this article. The method is based on automatic envelope spectrum analysis to diagnose the bearing for faults. Only the shaft speed and the four characteristic bearing fault frequencies derived from the bearing design, are required prior knowledge. An algorithm searches the envelope spectra for harmonics linked to a characteristic fault frequency and scores the dataset based on the number of harmonics identified and their prominence compared to surrounding noise. This score is monitored over time to detect condition changes in the bearing and determine the fault location. In this article, the proposed method is presented along with results from diagnosing faults in a damaged bearing. The test bearing is worn naturally over time during an accelerated life-test, and three characteristic faults are identified in the de-assembled bearing. The presented method accurately identifies all three fault types, and score trends also show when one fault progresses to a second one.

B.2 Methods

B.2.1 Experimental Test Setup

The vibration signal is collected on a bearing test bench. The test bearing is a 6008 type with a dynamic and static load rating of $17.8kN$ and $11kN$, respectively. The test bearing is naturally worn over time during an accelerated life-time test. To reduce the

bearing life-time, radial and axial loads of, respectively, $9kN$ and $4kN$ are applied to the test bearing. During the test, the vibration signal $V(t)$ and encoder position signal $\theta(t)$ is measured every 30 minutes at a reference speed of $\dot{\theta}^{(ref)} = 250$ revolutions per minute (rpm). The sampling frequency is set to $51.2kHz$ and the measurement duration is $24s$ (100 revolutions). The test bearing was worn out after surpassing 42 million revolutions.

B.2.2 Pre-process Vibration Signal

The measured vibration signal contains frequency components that are smeared across the frequency spectrum due to small variations in motor speed. The vibration signal is re-sampled to obtain an equal number of samples per round and spaced at a fixed angle interval to reduce frequency smearing. The number of samples retained per round is calculated using the reference shaft speed $\dot{\theta}^{(ref)}$ and the sampling rate. The encoder signal $\theta(t)$ is used as the reference angular shaft position, and the vibration signal is re-sampled using linear interpolation. This method is named *Order-Tracking*.

After applying the order-tracking algorithm, the re-sampled vibration signal $V_{ot}(t)$ is obtained. The next step is to filter the stationary vibration components not linked to bearing fault impacts. Most of these components are synchronous with the shaft position and are generated by e.g. misalignments or mass unbalance. A Time Synchronous Average (TSA) algorithm is used to remove these components. The vibration signal is first split up into multiple parts each with a time length of 1 round. The average of these parts is the synchronous vibration signal $V_s(t)$. $V_s(t)$ is afterwards subtracted from the vibration signal $V_{ot}(t)$ to remove the synchronous components. This subtraction results in the asynchronous vibration signal $V_{as}(t)$. More information on order tracking and TSA can be found in [6].

B.2.3 Narrow-band Identifier Algorithm

The narrow-bands are identified in the vibration signal frequency spectrum. A Fast Fourier Transform (FFT) algorithm is used to obtain the frequency spectrum. It returns the absolute amplitude Y and the frequency f of each sine wave component. The frequency spectrum is obtained using (B.1).

$$[Y, f] = \text{FFT}(V(t)) \tag{B.1}$$

$Y(i)$ denotes the amplitude at the i'th frequency bin, and similarly $f(i)$ denotes the frequency at the i'th bin. After obtaining the spectrum, an envelope is determined and placed on top of it. The envelope is obtained by modifying the amplitude of the spectrum to the maximum value within $\pm 15Hz$ from each frequency bin. Afterwards, a low-pass

filter is applied to smoothen the transitions. The resulting envelope is seen on top of the spectrum in Fig. B.1.

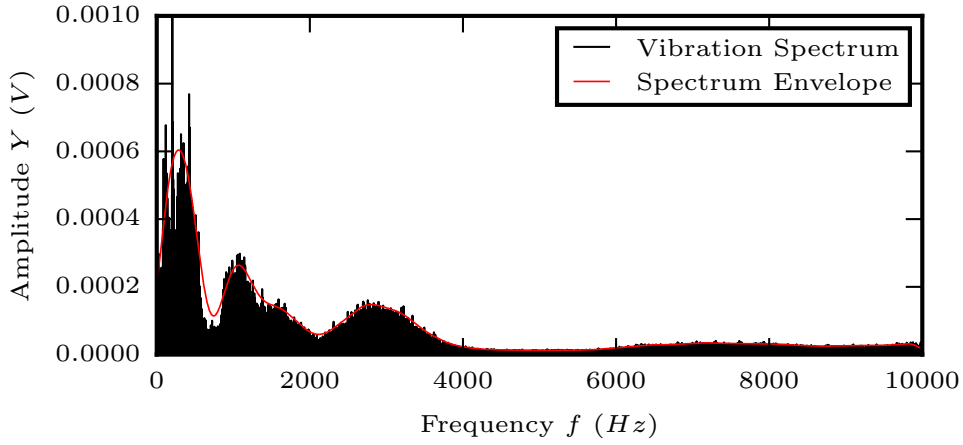

Figure B.1: The frequency spectrum of the vibration signal in addition to its envelope.

To determine each narrow-band in the spectrum, each local minimum of the envelope determines the separation frequency. These separation frequencies are shown in Fig. B.2 and each narrow-band is given a number ranging from 1 to 9 in this example.

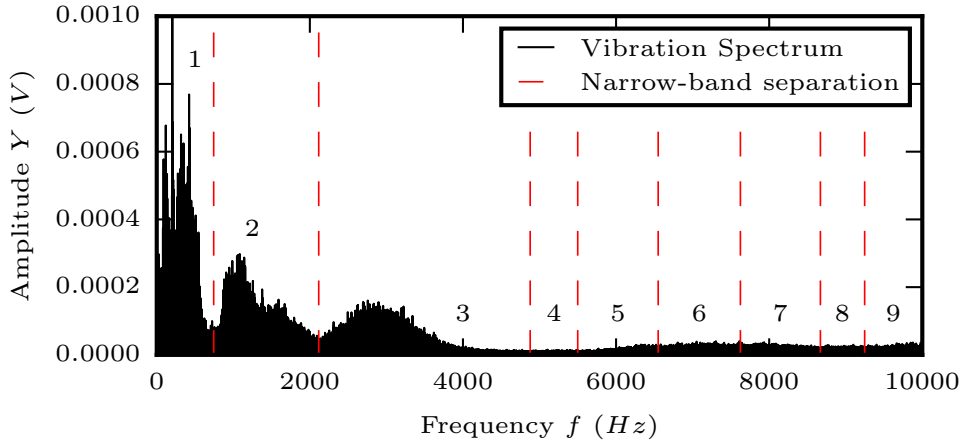

Figure B.2: The frequency spectrum of the vibration signal combined with the nine narrow-bands that are identified for this dataset.

The first three narrow-bands contain the most energy and should be analyzed for bearing fault impacts. However, the remaining six are also investigated in case the bearing natural frequency is within these narrow-bands.

B.2.4 Envelope Spectrum

Impacts in the bearing cause it to ring (vibrate) at its resonance frequency. The narrow-band may contain frequency components related to this ringing in the bearing. The

resonance frequency of the bearing is not directly linked to any characteristic bearing fault frequencies, but acts as an information carrier in an amplitude modulated signal. The information we seek lies in the frequency of the modulation waveform. Therefore it is necessary to de-modulate the signal to diagnose faults in the bearing. Each narrow-band is analyzed by performing three tasks. First, the vibration signal is band-pass filtered around the narrow-band frequency. Second, a Hilbert-Huang transform is applied to the bandpass-filtered signal to demodulate the signal. The result is the envelope of the signal, i.e. the modulation waveform. Finally, the envelope is transformed to the frequency domain via the FFT algorithm in (B.1). The resulting spectrum is named the *Envelope Spectrum*. It can be analyzed directly to find prominent peaks at the characteristic fault frequencies derived from the bearing kinematics. However, recall that the vibration signal is split up into multiple narrow-bands that should be analyzed using the Envelope Spectrum method. Analyzing multiple spectra manually is impractical due to time limitations. Therefore, there is a need for an autonomous bearing fault diagnostic system that can analyze the spectra. The next subsection presents such an autonomous bearing fault diagnosis system.

B.2.5 Autonomous Fault Diagnosis

The following method aims to mimic the way a diagnostician would analyze an envelope spectrum. Typically, one can observe the spectrum for prominent peaks corresponding to the characteristic bearing fault frequencies. Also, there may be several harmonics and also sub-bands if roller or inner-ring faults are present. Due to slip or a non-zero contact angle in the bearing, the spectrum peaks may not coincide perfectly with the characteristic fault frequencies. The diagnostician also has to account for this possible error. These tasks are built into the autonomous bearing fault diagnosis algorithm.

To illustrate the method, the envelope spectrum of the third narrow-band ($M = 3$) shown in Figure B.2 is used as an example. The test bearing has an inner-race fault that will be characterized by harmonics at the inner-race ball pass frequency, and sub-bands equal to the shaft frequency. The characteristic fault frequency for the inner race is $f_c = 6.88$ orders, and the expected sub-band frequency is $f_{sb} = 1$ order. One order is the same as the shaft speed as shown in (B.2)

$$1 \text{ order} \equiv f_s \tag{B.2}$$

where f_s is the shaft speed. Converting the envelope spectrum frequency to orders removes the need for scaling the fault frequencies based on the shaft speed. The steps in the autonomous fault diagnosis algorithm are detailed in the following. Suppose the

envelope spectrum is given by the amplitude Y and the frequency f. Let the harmonic number $H = 1$.

Step 1: Identify the maximum amplitude within a small frequency band around a characteristic fault frequency. This maximum value is the harmonic value and is calculated as follows:

$$M_h, I = \max(Y(i)) : f_{h1} \leq f(i) \leq f_{h2} \tag{B.3}$$

where $f_{h1} = c_\alpha f_c \cdot (H - w)$, $f_{h2} = c_\alpha f_c \cdot (H + w)$, and I is the index i where $Y(i)$ is maximum. The frequency band is determined by the slip tolerance w and the fault frequency correction factor c_α. These values are calculated using (B.4) and (B.5) respectively.

$$w = \begin{cases} 0.02, & \text{for } H = 1 \\ 0.01, & \text{else} \end{cases} \tag{B.4}$$

$$c_\alpha = \begin{cases} 1, & \text{for } H = 1 \\ \alpha, & \text{else} \end{cases} \tag{B.5}$$

The correction factor α is calculated in (B.6) to adjust the characteristic fault frequency to account for roller slip and contact angle.

$$\alpha = f(I)/(f_c \cdot H) \tag{B.6}$$

where $f(I)$ is the frequency where the most prominent harmonic peak is observed.

Step 2: A "noise level" within a small band around the corrected fault frequency is calculated using (B.7) to determine whether the maximum peak is prominent enough.

$$N = \frac{1}{n-1}\left(\left[\sum Y(i) : f_{n1} \leq f(i) \leq f_{n2}\right] - Y(I)\right) \tag{B.7}$$

where $f_{n1} = \alpha f_c \cdot (H - 0.02)$, $f_{n2} = \alpha f_c \cdot (H + 0.02)$, and n is the number of elements in the summation. N is the mean of all spectrum values within the frequency band, except for the value of the harmonic. The frequency bands used to determine the maximum harmonic value and the noise are shown as an example for $H = 1$ in Fig. B.3.

Step 3: If the fault frequency is amplitude modulated by a sub-band frequency f_{sb}, there should be at least one prominent peak $\pm f_{sb}$ away from αf_c. If it is not amplitude modulated, skip this step. Two prominent sub-bands are searched for: one negative sub-band and one positive sub-band. The negative sub-band is calculated using (B.8).

$$M_{ns} = \max(Y(i)) : f_{ns1} \leq f(i) \leq f_{ns2} \tag{B.8}$$

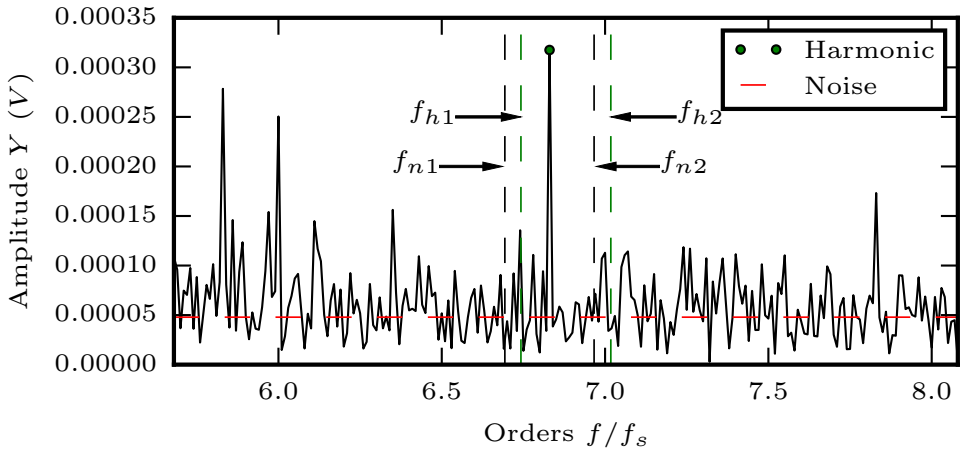

Figure B.3: For $H = 1$, the frequency bands for identifying the harmonic value and the noise are shown. In addition, the harmonic is marked as a green dot and the noise level is the red-stapled line.

where $f_{ns1} = H\alpha f_c - f_{sb} - 0.05f_s$, and $f_{ns2} = H\alpha f_c - f_{sb} + 0.05f_s$. Here a 5% tolerance away from the shaft speed is used to generate the frequency band where the negative sub-band should be. The positive sub-band is calculated in a similar manner in (B.9).

$$M_{ps} = \max(Y(i)) : f_{ps1} \leq f(i) \leq f_{ps2} \tag{B.9}$$

where $f_{ps1} = H\alpha f_c + f_{sb} - 0.05f_s$, and $f_{ps2} = H\alpha f_c + f_{sb} + 0.05f_s$. The frequency bands are shown together with the maximum sub-band values for $H = 1$ in Fig. B.4.

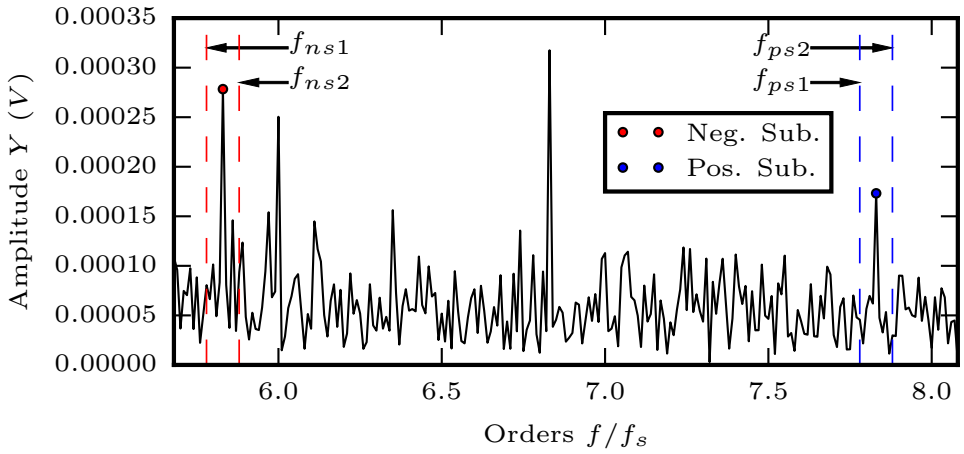

Figure B.4: For H = 1, the frequency bands for identifying the sub-bands are shown. In addition the identified sub-bands are marked with a red and blue dot.

Step 4: If the harmonic and possible sub-bands are prominent enough, a score is calculated based on the harmonic value. However, the harmonic value must pass a threshold if a score should be computed. The threshold is calculated using (B.10).

$$T = 3N \tag{B.10}$$

If there should be sub-bands present around the fault frequency, the following relation must be true to calculate a score:

$$M_h > T \text{ and } (M_{ns} > T \text{ or } M_{ps} > T) \tag{B.11}$$

Similarly if $f_{sb} = 0$, the following relation must be true to calculate a score:

$$M_h > T \tag{B.12}$$

If the relation above is true, the harmonic score is calculated in Step 5. Otherwise, skip to Step 6.

Step 5: A scoring system is developed to quantify the prominence of the harmonic peaks. The score for each identified harmonic is calculated using (B.13).

$$S(H) = (M_h/3N) \cdot H^2 \tag{B.13}$$

where $S(H)$ is the score for the H'th harmonic. The score is unit-less and rewards narrow-bands with multiple harmonics. A large score implies a high probability of damage being present. Afterward, more harmonics are searched for by increasing the harmonic number $H = H + 1$ and returning to Step 1. Figure B.5 shows an example where the second harmonic and sub-bands are identified in mode 3. In this example, the negative sub-band and the harmonic is greater than the threshold, and therefore the score of 10.21 is calculated using (B.13).

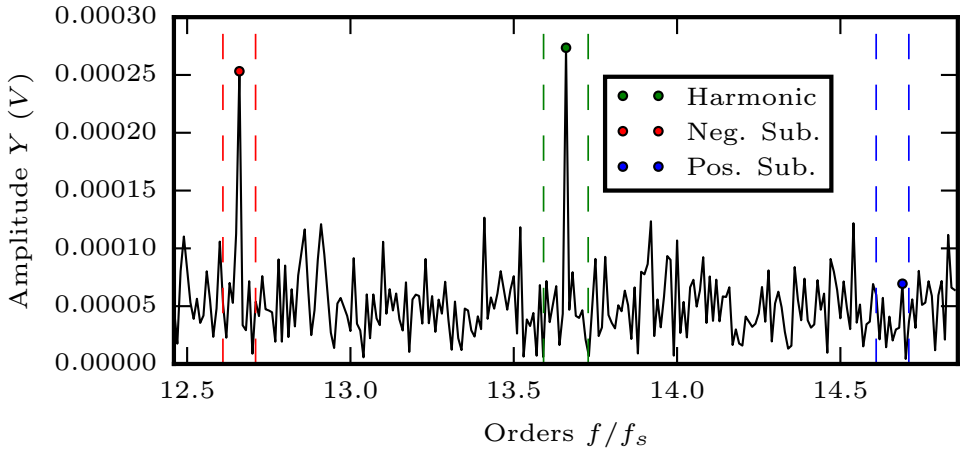

Figure B.5: For H = 2, the harmonic and two sub-bands are identified.

Step 6: Once all the prominent harmonics are identified, a combined score for the narrow-band can be calculated using (B.14).

$$S_m(M) = \sum_{i=1}^{n_H} S(i) \tag{B.14}$$

where n_H is the number of harmonics identified, and M is the narrow-band number. The score summation can be extended by also summing over all the narrow-bands in the vibration dataset. This combined score is calculated using (B.15).

$$S_{tr} = \sum_{i=1}^{n_M} S_m(i) \tag{B.15}$$

where n_M is the number of narrow-bands in the vibration dataset. Since S_{tr} is a single scalar value for a given dataset and a given fault type, it can be monitored over time like RMS and Kurtosis. This provides an excellent opportunity to visualize the changes in the bearing condition and to set up automatic warnings or alarms for the operator. The complete flow diagram for the algorithm described in this section is shown in Fig. B.6.

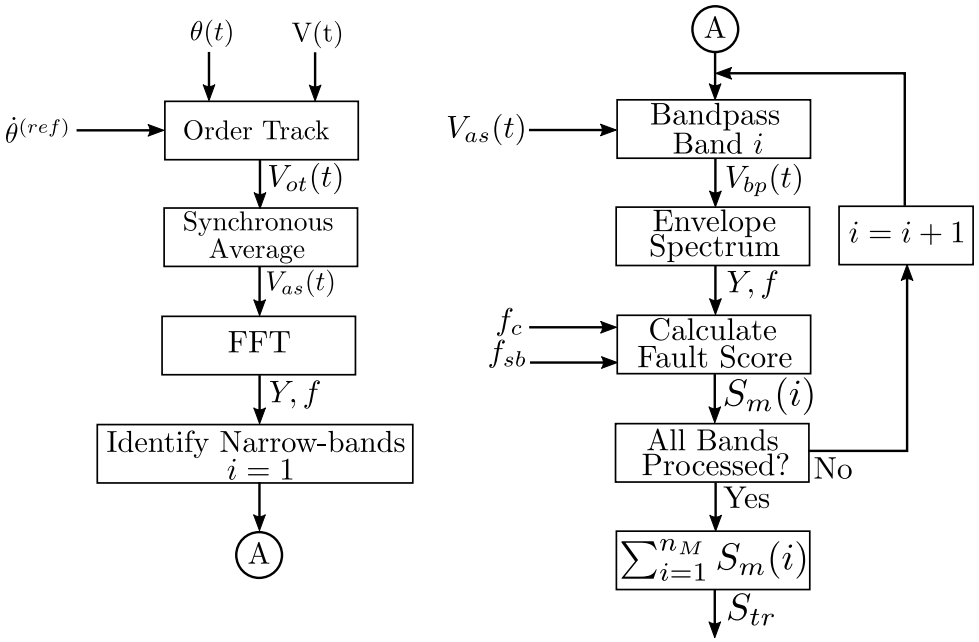

Figure B.6: Complete algorithm flowchart.

B.3 Results and Discussion

B.3.1 Vibration Mode Score

The vibration dataset whose frequency spectrum is shown in Fig. B.2, contains at least two narrow-bands that include the characteristic fault components for an inner-ring fault. The first one is narrow-band three, $M = 3$. The harmonics observed in the envelope spectrum of this narrow-band are shown in Fig. B.7.

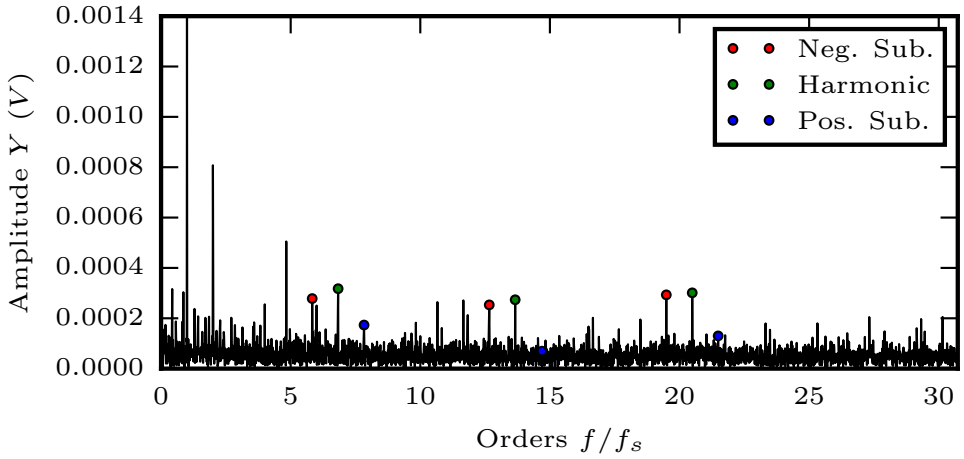

Figure B.7: For the third narrow-band, $M = 3$, a total of 3 harmonics related to the inner-race fault are identified.

The score calculated using (B.14) for the third narrow-band is $S_m(3) = 39.33$. Whether this value is large or small is difficult to determine. But normally, analyzing an undamaged bearing should return a score quite close to zero. In addition, analysis of the sixth narrow-band ($M = 6$) also reveals signs of a fault on the inner-ring. The harmonics observed in the analysis of this narrow-band are shown in Fig. B.8.

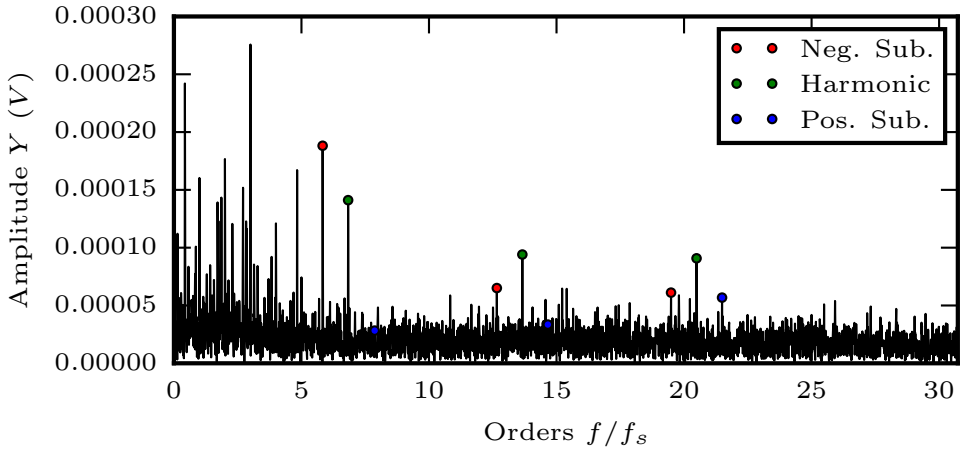

Figure B.8: For the sixth narrow-band, $M = 6$, a total of 3 harmonics related to the inner-race fault is identified.

The score calculated using (B.14) for the sixth narrow-band is $S_m(6) = 36.14$, which is almost as much as the third narrow-band score. Therefore, there are reasons to analyze more than one narrow-band on the vibration signal. For reference and comparison purposes, the score for the nine narrow-bands are given in Table B.1.

Table B.1: Inner ring diagnostic score for the nine narrow-bands in the example data.

$\boldsymbol{M}$	1	2	3	4	5	6	7	8	9
$\boldsymbol{S_m(M)}$	3.35	0	39.33	0	0	36.14	1.36	1.75	0

B.3.2 Vibration Trend Score

A great usage scenario for the proposed method is to monitor the score for the three characteristic bearing fault types over time. The algorithm in Fig B.6 is repeated for each characteristic fault by altering f_c and f_{sb} to the combinations given in Table B.2. The combined vibration dataset score is calculated using (B.15) for the three characteristic bearing faults and plotted in a graph. The proposed method will be compared to monitoring the RMS value of the vibration data. The RMS is calculated using (B.16).

$$V_{RMS} = \sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t} V(i)^2} \tag{B.16}$$

where n_t is the number of samples in the vibration dataset. The last 100 datasets acquired from the accelerated life test detailed in Section B.2.1 are used to monitor the progressing bearing faults. The scores calculated using (B.15) for the three fault types are shown in Fig. B.9 a), and, for comparison, the RMS value as calculated using (B.16) is shown in Fig. B.9 b).

Table B.2: Characteristic fault frequencies for the test bearing in orders.

	Inner Race	**Outer Race**	**Roller**
$\boldsymbol{f_c}$	6.88	5.12	6.66
$\boldsymbol{f_{sb}}$	1	0	0.43

In Fig. B.9 a) the scores are normalized to fit all three graphs in a single figure, but the respective maximum values are given in the legend. Analyzing Fig. B.9 a) it is seen that until the 71st dataset the condition of the bearing is unchanged and in a healthy state. At this point, however, the proposed algorithm returns a high score for the roller fault, while it has previously been close to zero. This finding corresponds well with the increase in RMS value as shown in Fig. B.9 b). However, the roller fault does not appear in the vibration datasets for long. Instead, an outer-ring damage propagates during the next 20 datasets. Finally, the inner-ring is also damaged, as is seen during the final 10 datasets in Fig. B.9 a). The score values calculated using the proposed method correspond well with the findings in the de-assembled bearing. The internal damage in the bearing is shown in Fig. B.10.

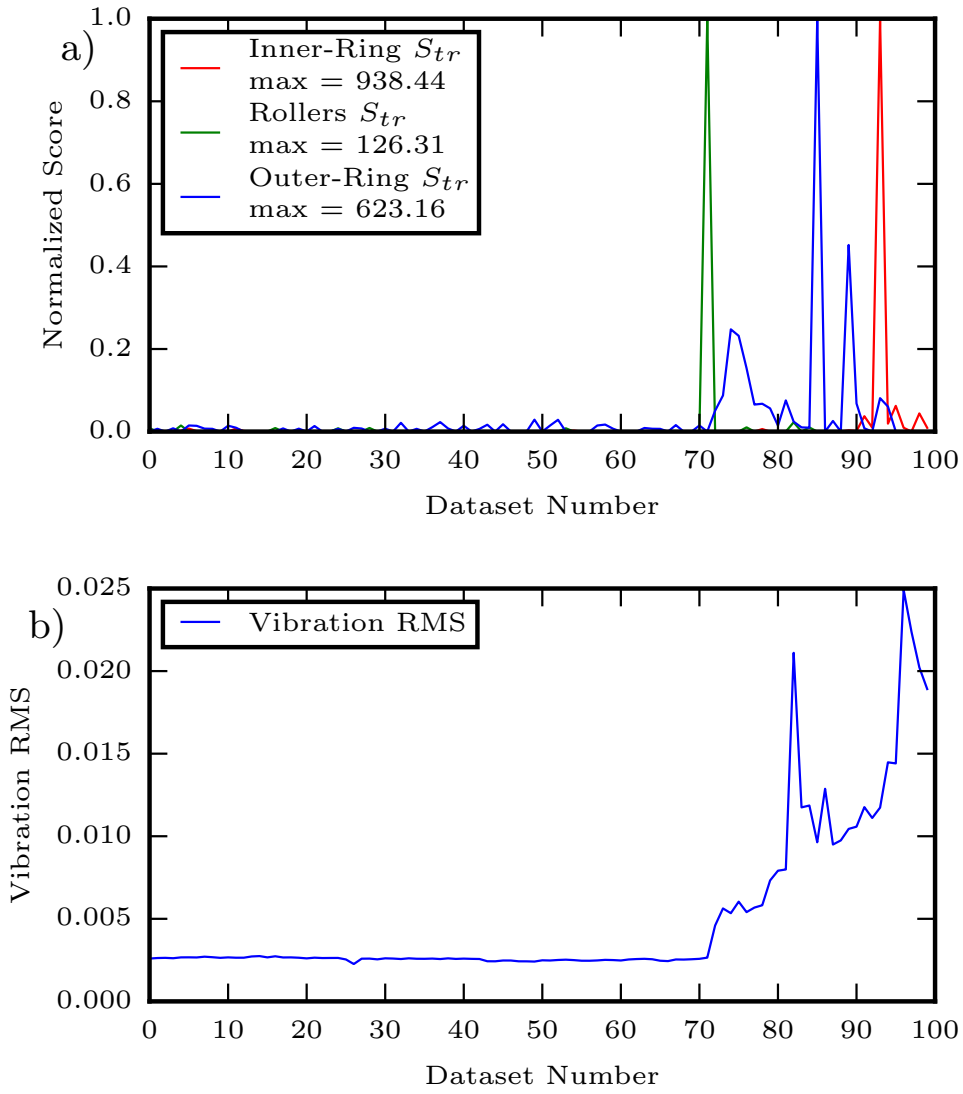

Figure B.9: For the last 100 datasets, the characteristic fault scores are calculated using (B.15) and shown in a). The RMS trend calculated using (B.16) is shown in b) for comparison.

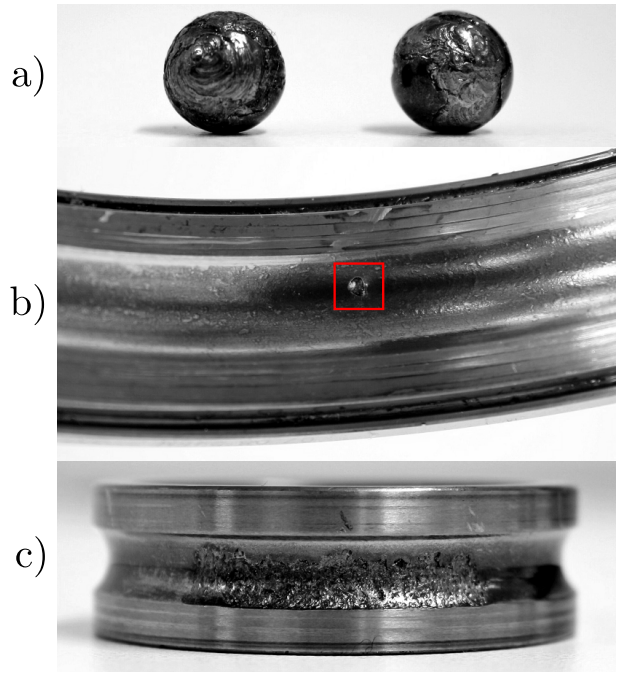

Figure B.10: The de-assembled test bearing after the accelerated life test. a) shows two damaged balls, b) shows a small pit in the outer-ring, and c) shows a large area of pitting on the inner-ring.

Fig. B.10 shows signs of damage in two rollers, the outer-ring, and the inner-ring. The locations of the faults correspond very well with the results from the proposed method. This comparison shows that the proposed algorithm is very efficient in diagnosing the bearing for the characteristic faults.

B.4 Conclusion

Raw vibration data is often bandpass-filtered around the bearing fundamental frequency only to preserve frequency components related to fault impacts. However, determining the resonance frequency of the bearing is a challenging task. Instead of identifying the optimal narrow-band using Kurtogram, the vibration data is divided into multiple narrow-bands and each band is analyzed to diagnose the bearing. However, this is a tremendous task to perform manually and thus an automatic bearing fault diagnosis algorithm is proposed in this article. It is capable of identifying the three most common faults in a bearing with good accuracy. The proposed algorithm has been explained in detail and validated using experimental data.

REFERENCES

[1] L. Guo, J. Chen, and X. Li. Rolling bearing fault classification based on envelope spectrum and support vector machine. *Journal of Vibration and Control*, 15(9):1349–1363, 2009. doi:10.1177/1077546308095224.

[2] C. Mishra, A. K. Samantaray, and G. Chakraborty. Rolling element bearing defect diagnosis under variable speed operation through angle synchronous averaging of wavelet de-noised estimate. *Mechanical Systems and Signal Processing*, 72:206–222, 2016. doi:10.1016/j.ymssp.2015.10.019.

[3] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007. doi:10.1016/j.ymssp.2005.12.002.

[4] T. Barszcz and A. JabŁoński. A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram. *Mechanical Systems and Signal Processing*, 25(1):431–451, 2011. doi:10.1016/j.ymssp.2010.05.018.

[5] X. Xu, M. Zhao, J. Lin, and Y. Lei. Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis. *Measurement*, 91:385–397, 2016. doi:10.1016/j.measurement.2016.05.073.

[6] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

Paper C

Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics

Andreas Klausen and Kjell G. Robbersmyr

This paper has been published as:

A. Klausen and K. G. Robbersmyr. Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics. *Mechanical Systems and Signal Processing*, 118:226-244, 2019. doi: 10.1016/j.ymssp.2018.08.048.

Cross-correlation of Whitened Vibration Signals for Low-Speed Bearing Diagnostics

Andreas Klausen and Kjell G. Robbersmyr

University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

Abstract **– Rolling-element bearings are crucial components in all rotating machinery, and their failure will initially degrade the machine performance, and later cause complete shutdown. The period between an initial crack and complete failure is short due to crack propagation. Therefore, early fault detection is important to avoid unexpected machine shutdown and to aid in maintenance scheduling. Bearing condition monitoring has been applied for several decades to detect incipient faults at an early stage. However, low-speed conditions pose a challenge for bearing fault diagnosis due to low fault impact energy. To reliably detect bearing faults at an early stage, a new method termed Whitened Cross-correlation Spectrum (WCCS) is proposed. The method computes the cross-correlation between the whitened vibration signal and its envelope. In this paper, it is detailed how this correlation can improve the fault diagnosis compared to analyzing the envelope spectrum alone. Compared to other methods reported in the literature, the WCCS provides accurate fault detection without involving experimentally tuned settings or bandpass-filtering. Vibration data at 20 revolutions per minute rotational speed from an accelerated life-time test of a 40 mm bore size bearing is used to verify the performance of the proposed method. An additional case study using the WCCS on a difficult dataset from the Case Western Reserve University database is also presented to verify the performance.**

C.1 Introduction

Rolling-element bearings, or bearings for short, are crucial components in all rotating machinery. Their failure is one of the most common cause of machine breakdown. A worn bearing is characterized by increased vibration levels, internal looseness, and higher friction. The increase of vibration can damage nearby components, and lead to a full stop of the machine. If worn bearings are not replaced in time, costly downtime or personnel injuries may occur. Condition monitoring techniques can be applied to estimate

the bearing health and remaining useful life-time. Data from sensors that measure a physical quantity, like the vibration, are used as input to such a system. The data is further analyzed using signal processing algorithms, before the results are presented to an operator. Based on the results, the operator can decide whether the bearing is in a healthy state, or if it is worn and should be replaced. Such condition monitoring systems have been used for several decades to monitor the health of all kinds of rotating machinery components. The most common sensor type to use is vibration accelerometers, as the bearing vibration is closely linked to the amount of internal wear. An incipient fault, on either bearing race-way or a roller, causes an impulse of vibration every time it is struck. The spectral frequency of the resulting vibration is based on the resonance frequency of the system and is normally in the thousands of Hertz. Further, the resonance frequency of a bearing system is normally not known as it is difficult to determine analytically or experimentally. However, the spectral frequency is not directly of interest when diagnosing a bearing. The cyclic frequency between each impact impulse may reveal its fault. By analyzing the kinematics of a bearing under no-slip conditions, the characteristic cyclic frequencies for the different fault types are determined. If the cyclic vibration frequency match any of the characteristic frequencies, the bearing is likely to be damaged at that certain location. To determine the cyclic impact frequencies of the measured vibration signal, the Fourier transform of its demodulated envelope may be analyzed to identify the bearing cyclic vibration. During low-speed conditions, there may be some challenges in diagnosing bearing faults using a vibration accelerometer. The fault impact energy is dependent on the shaft speed and is decreased if the shaft speed decreases [1]. Background noise, however, is the same no matter what the shaft speed is, and therefore the signal-to-noise-ratio (SNR) decreases with decreasing shaft speed. The envelope spectrum may not reveal the fault if the characteristic cyclic frequencies are masked in background noise during low-speed conditions. In this case, more advanced signal processing methods are necessary to highlight the low energy impacts during low-speed conditions.

A new method was proposed in [2] for diagnosing a low-speed bearing using extracted discrete wavelet packets that contains the bearing fault vibrations. The results show that the multiple band-pass filtered autoregressive envelope spectrum provided clear indication of faults at 60 revolutions per minute (rpm) rotational speed. However, one disadvantage is that the Adaptive Network-based Fuzzy Inference System trained to choose the bands, requires pre-labeled training data, and such historic data may not be available for every system. In recent years, research on acoustic emission (AE) technology has shown that it is sensitive to early sub-surface cracks in the bearing. While vibration sensors are only sensitive to impacts on the bearing surface, AE sensors may detect changes in the bearing sub-surface, which should aid in early fault diagnosis. The biggest disadvantage is the cost

of the equipment, and the high required sample rate, as the signal of interest is typically in the range of 100 kHz and 1 MHz [3]. The disadvantages of the high sampling rate is however reduced significantly by an efficient down-sampling technique that does not affect signal quality [4]. The results indicate that even when down-sampling by a ratio of 500, the bearing fault signature could be captured by the AE sensor. In [5], a thrust bearing is run to failure at 72 rpm, and the resulting AE was measured at four locations on the outer ring. The results indicate that the AE energy increased with increasing fault size, and that the fault type is detectable in certain pre-processed spectra. In [6], a bearing with a pre-seeded fault at an extremely low speed of 1.12 rpm is diagnosed using AE. It was concluded that parameters such as AE amplitude and energy provided valuable information on the condition of a low-speed rotating bearing. In addition, a method for detecting a fault in the bearing using AE signal was also presented. The method consists of grouping multiple stress-wave signals by the centroid of autoregressive model coefficients. If two distinct groups are formed, the bearing is considered faulty. If no distinct groups were formed, the stress wave signals are considered to only contain noise, and the bearing is assumed healthy. In [7], a low-speed bearing rotating at 10 rpm was diagnosed using Support Vector Machine and Relevance Vector Machine. The input features were generated from both vibration and AE sources, and it was concluded that the classifier trained on AE data had the best accuracy. In [8], the performance of AE and vibration was compared on a large slew bearing rotating at 8 rpm. Both signals were pre-processed using a combination of multivariate Principal Component Analysis and Ensemble Empirical Mode Decomposition, which adaptively decomposed the signal into different time scales. It was shown that both the vibration and AE signal contained enough information to diagnose the bearing for a seeded inner ring fault after proper pre-processing. Other research has also shown success in using AE for low-speed bearing fault diagnosis [9, 10, 11, 12].

Most of the referenced works state that AE is superior to vibration signals for bearing fault diagnosis under low-speed conditions. While a vibration sensor can capture the fault impact vibration, the signal is often masked by vibration from other machine components, and background noise. In addition, the bearing fault impact energy gets lower as the shaft speed is reduced. Therefore, a solid algorithm is necessary to extract the weak impulses generated in the bearing during low-speed conditions to properly diagnose its condition. In this article, vibration signals are used to diagnose a 6008 ball-type bearing for faults. The bearing is run to failure on an accelerated life-time test bench at 500 rpm. The vibration signal is measured every 10 minutes at 20 rpm for the duration of 50 revolutions to acquire low-speed vibration signals. Initially, the vibration signal is re-sampled using encoder data to achieve a constant angle increment rather than time. Next,

shaft-synchronous vibration components are removed by subtracting the average vibration per revolution. To highlight the slightly random vibration impulses in the bearing, an Autoregressive Model (ARM) is afterwards trained to predict deterministic components that are not related to the bearing fault. The components predicted by the ARM are removed to retain random components such as the bearing vibration. After this process, the vibration signal is whitened as it contains mostly random components. Finally, the cross-correlation between the whitened vibration signal and its envelope is computed. This cross-correlation results in an element-wise multiplication of the frequency spectrum of the whitened vibration signal and its envelope. The justification of computing the cross-correlation is that the bearing characteristic fault vibration may be visible directly in the spectrum of the raw vibration signal [13], as well in the spectrum of the demodulated high-frequency vibration signal. Therefore, to exploit both signals, the cross-correlation is calculated to fuse the information in both spectra. Frequency amplitude peaks that are present in both spectra will be amplified, while peaks that are only present in either of the two are attenuated. Using the Fourier transform, the resulting spectrum of the cross-correlation signal, termed the Whitened Cross-correlation Spectrum (WCCS), is analyzed for faults related to the bearing. The results indicate that, even at the low-speed of 20 rpm, early identification of fault is possible. For verification purposes, a difficult dataset from the Case Western Reserve University bearing fault database is also successfully diagnosed using the WCCS. The rest of the paper is organized as follows: the experimental setup and the algorithms used in the proposed method and comparison methods are presented in Section C.2. Afterwards, results of using the proposed method are presented, and the performance is compared to other methods reported in the literature, like the Envelope Spectrum, in Section C.3. Finally, conclusions are drawn in Section C.4.

C.2 Methods

C.2.1 Experimental test setup

The vibration data is collected on an accelerated bearing life-time test rig that is shown in Figure C.1. A permanent magnet motor combined with a 1:7 reduction planetary gearbox is used to drive the test bearing shaft, and a variable-frequency drive allows for variable speed conditions. The test bearing at the left end of the shaft is a 6008 type bearing with 40 mm bore size, and its characteristic fault orders are given in Table C.1. These frequencies are given in magnitudes of shaft orders, where the shaft speed equals 1 order, i.e. $O_S = 1$. The rated dynamic load of the bearing is $C = 17.8$ kN, and the static load rating is $C_0 = 11.6$ kN. Radial and axial loads are applied by two electric

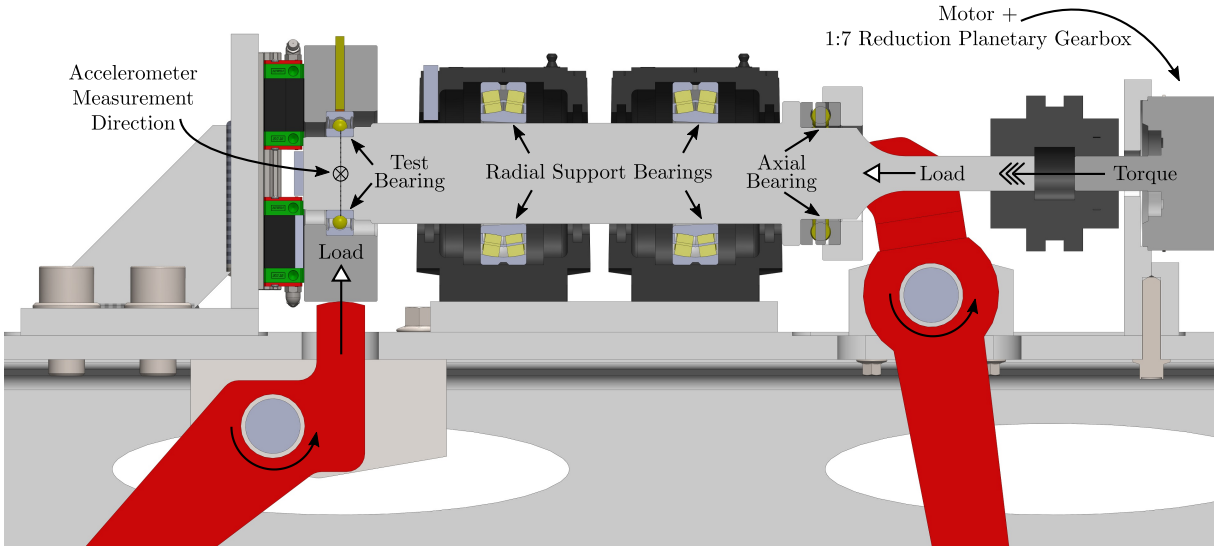

Figure C.1: Simplified view of the accelerated bearing life-time test bench.

linear actuators geared with mechanical levers. The two support bearings at the middle of the shaft aid to counteract the radial force on the test bearing. The axial bearing is installed to transfer the stationary axial load to the rotating shaft. A unidirectional vibration accelerometer is stud-mounted to the side of the test bearing housing with the direction and placement as shown in Figure C.1. Its linear range is between 2 Hz and 10 kHz, and the nominal sensitivity is 100 mV/g with a maximum peak acceleration of 60 g. The vibration is sampled at 51.2 kHz using a 24-bit ±30 V A/D converter. A quadrature incremental encoder is located inside the motor. With 1024 pulses per channel, a total of 4096 pulse edges can be identified per motor shaft revolution. Additionally, the motor rotates 7 times faster than the test bearing shaft due to the gearbox, resulting in a resolution of 28,672 pulse edges per revolution on the test bearing shaft. To record the shaft position, the number of pulse edges passed is sampled at a frequency of 512 Hz. The test bearing is subjected to a radial and axial load of 9 kN and 7 kN respectively to reduce the rotational life-time. The test procedure is as follows: The test bearing is generally driven at a speed of 500 rpm. Every 10 minutes, a measurement cycle is initiated. During this cycle, the bearing vibration and motor encoder position is measured for the duration of 50 revolutions at shaft speeds of 500, 250, 100, 50, and 20 rpm successively. After each cycle, the shaft speed returns to 500 rpm. In this article, the vibration data collected at 20 rpm will be used since low-speed bearing fault diagnostics is the primary objective. The bearing life-time was roughly 6 million revolutions (14 days of continuous operation) before complete stop. The bearing was afterwards disassembled to identify what components were damaged. Faults were identified on three of the rollers, the outer-race, and the inner-race. Pictures of the disassembled bearing are shown in Figure C.2.

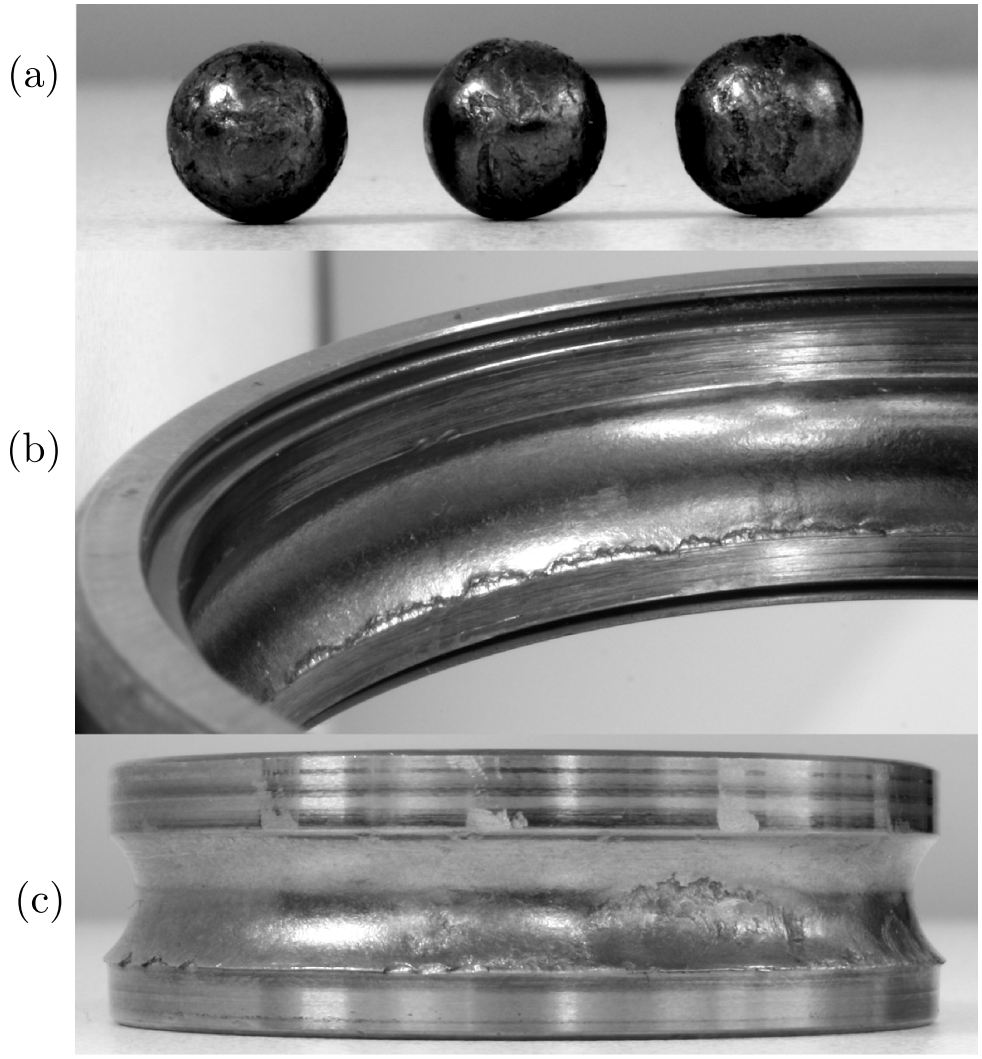

Figure C.2: The damaged components in the disassembled bearing. (a) Three damaged rollers. (b) Damaged area on the outer-race. (c) Large damaged area on the inner-race.

Table C.1: Characteristic fault orders for the test bearing given in shaft orders assuming zero contact angle and no roller slip. The expected harmonic order and the accompanying side-band order (due to non-homogeneous radial load) for the given fault type are given.

Characteristic fault orders	Acronym	Value	Side-band
Fundamental train	O_{FT}	0.43	-
Ball pass outer race	O_{BPO}	5.12	-
Ball pass inner race	O_{BPI}	6.88	O_S
2x Ball spin	O_{2BS}	6.66	O_{FT}

C.2.2 Order tracking

The motor speed controller is incapable of rotating the shaft at a completely constant speed. Consequently, the time period between impacts in a bearing will vary based on the instantaneous speed. The A/D converter that measures the vibration signal is storing data at a fixed time interval given by the sampling rate. Because of the instantaneous speed changes, this will result in a blurred frequency spectrum as the Fourier transform assumes a stationary process. To mitigate the symptoms, the vibration signal is digitally re-sampled with respect to the shaft encoder pulses. The re-sampled signal has a fixed shaft angle interval rather than a time interval. This makes sure that there is, theoretically, a fixed number of samples between each bearing fault impact. The following algorithm is inspired by the work done by Fyfe and Munck in [14].

On the presented test rig, the rotor angular position is measured at a sampling frequency of 512 Hz. Due to the fast sampling, it is assumed that the motor speed is constant between each measurement. Under this assumption, the shaft angle at any time t can be obtained using linear interpolation techniques. The shaft position is initially interpolated to match the digital sampling times of the vibration signal $V(t)$. This results in a vibration signal that is given with respect to the shaft angle, but the shaft angle interval is not fixed. To fix the interval, the vibration signal is interpolated to achieve a fixed angle interval. Here, a second-order cubic spline interpolation is chosen as it provides good interpolation performance at a low computational cost. The spline interpolation is initialized as:

$$f_{ot}(\theta) = \text{interp}\{\theta, V\} \tag{C.1}$$

where interp$\{\cdot\}$ is the cubic spline interpolation function, and $f_{ot}(\theta)$ is a continuous approximation of V describing points between the discrete samples of V using cubic spline interpolation. Using $f_{ot}(\theta)$, the vibration data is re-sampled to achieve a fixed shaft angle interval. The size of this angular interval can be arbitrarily chosen, but some guideline should be followed to avoid over- or under-sampling. Here, the angular interval is chosen to preserve the original sampling rate and total number of samples as close as possible using:

$$\Delta\theta_d = \left(\text{round}\left\{\frac{F_s}{\dot{\theta}^{(\text{ref})}}\right\}\right)^{-1}, \tag{C.2}$$

where $\dot{\theta}^{(\text{ref})}$ is the reference shaft speed in Hz, F_s is the vibration data sampling rate in Hz, and round$\{\cdot\}$ rounds the number to the nearest integer. The reason for using the round$\{\cdot\}$ function and inverse, is to make sure that an integer number of samples contain a complete shaft revolution. This is crucial when performing Angle Synchronous Averaging to avoid overlap between revolutions. Finally, the vibration signal is re-sampled to achieve the

constant shaft angle interval using:

$$V_{ot}(j \cdot \Delta\theta_d) = f_{ot}(j \cdot \Delta\theta_d). \tag{C.3}$$

where V_{ot} is the order tracked vibration signal, and j is a positive integer. In the next Subsection, the average vibration per shaft revolution is calculated to remove the shaft-synchronous vibration components.

C.2.3 Angle synchronous average

The vibration signal includes components from sources that are synchronous to the shaft rotation. Typical components include vibration from shaft misalignment, shaft unbalance, and motor vibration. In addition, the meshing vibration from fixed-axis gearboxes are also integer multiples of the shaft speed. It should be noted, however, that the test rig in Section C.2.1 does not include a fixed-axis gearbox. These shaft-synchronous vibration components may mask the bearing fault and should therefore be removed before further processing of the signal. The bearing characteristic fault orders in Table C.1 are non-integer orders, and are therefore not shaft-synchronous. Therefore, by estimating the shaft-synchronous vibration components and removing that from the vibration signal, the bearing fault vibration is not affected. Using a deterministic/random separator (DRS) such as the time synchronous average in [15], the shaft-synchronous components are estimated and removed from the signal. In this research, the signal was first order-tracked by means of computed order tracking, and therefore a modified Angle Synchronous Average (ASA) is performed instead. An estimation of the shaft-synchronous components is found by calculating the mean value of the vibration signal per revolution using:

$$V_s = \frac{1}{n_r} \sum_{j=0}^{n_r-1} V_{ot}\left[j \cdot N_r : (j+1) \cdot N_r\right], \tag{C.4}$$

where n_r is the number of revolutions the shaft has turned during the measurement, $N_r = 1/\Delta\theta_d$ is the number of samples measured per revolution, $V_{ot}[a:b]$ imply all discrete samples of V_{ot} including $V_{ot}[a]$ up to and including $V_{ot}[b-1]$, and the first sample is at index 0. This results in the ASA, which should contain the shaft-synchronous vibration components. The ASA is afterwards removed from each revolution in the order tracked vibration signal using:

$$V_{as}\left[j \cdot N_r : (j+1) \cdot N_r\right] = V_{ot}\left[j \cdot N_r : (j+1) \cdot N_r\right] - V_s : j \in [0, n_r - 1], \tag{C.5}$$

where V_{as} is the asynchronous vibration data containing the bearing fault. More information on the general time synchronous average is found in [15]. In the next subsection, other deterministic signal components are separated and removed using an autoregressive whitening filter.

C.2.4 Autoregressive whitening filter

The ASA algorithm attenuated most of the shaft-synchronous vibration originating from the shaft and the motor. However, there are still deterministic vibration components not related to the bearing fault present in the dataset. One such source is the planetary gearbox used in the test rig. Planetary gearboxes emit a complex vibration pattern due to multiple moving components and non-integer orders forcing frequencies [16]. These components are, however, deterministic and therefore predictable. The characteristic bearing orders from Table C.1 assumes a zero contact angle and no slip. However, as the local radial and axial load relationship for each roller may be different from one another, each roller moves at slightly different speeds. The bearing cage moves at the mean speed of all rollers, forcing some of them to slip on the raceway. Consequently, angular deviations of up to 2% between impacts may occur [15], and the bearing vibration can therefore be modeled as a cyclostationary process. This allows the bearing vibration to be separated from the planetary gearbox vibration using a DRS. A linear prediction model as described in [15] is used in this research to model the deterministic components. To this end, an Autoregressive Model (ARM) is estimated to predict the future values of the vibration signal based on the p previous values. The ARM acts as a model of the deterministic part of the signal. The ARM is afterwards subtracted from the original signal, separating the random components. Separating the deterministic components whitens the signal, and hence this process is described as a whitening filter. The ARM is given by:

$$V_{as}[k] = -\sum_{j=0}^{p-1} a[j] \cdot V_{as}[k-j-1] + V_w[k], \tag{C.6}$$

where $a[j]$ is the j'th parameter of the model, V_w is the whitened vibration signal (or ARM residual), and p is the ARM order. Random signal components, including the bearing vibration, are not perfectly predictable and will therefore be a part of the residual V_w. The unknown parameters are identified by solving the Yule-Walker equations [17, 18] for ARM training. Wang and Wong [19] describes a method for efficiently training the ARM using Levinson-Durbin Recursion (LDR). This training method allows for the identification of all ARM up to an order of p_{max}. Theoretically, the ARM can have as many parameters as the input signal length. C.7red However, to avoid overfitting, i.e. the ARM starts to predict random data, the maximum order must be less than the number of samples between bearing fault impulses [20]. This limit also imposes a selective whitening as the deterministic components related to the bearing fault are preserved in the residual. The maximum number of samples is calculated using:

$$p_{max} = \text{floor}\left\{\frac{1}{\max\{O_{bcf}\}\Delta\theta_d}\right\}, \tag{C.7}$$

where O_{bcf} is a list containing the characteristic bearing fault orders for the present bearing, floor$\{\cdot\}$ returns the nearest integer rounding in the negative direction, and max$\{\cdot\}$ returns the maximum number of a list. Since the type of fault is unknown during analysis, the maximum characteristic fault order is used.

The training method in [19] uses the Akaike Information Criterion (AIC) to determine the optimal model order up to p_{max}. One modification is however implemented in this research by using the Corrected Akaike Information Criterion (AIC_C) [21] rather than the standard AIC. This change is made for its advantage on finite size datasets. The optimal ARM order p_{opt} is identified by the minimal value of AIC_C. Using the optimal ARM, the residual of the ARM is calculated using:

$$V_w[k - p_{opt}] = V_{as}[k] + \sum_{j=0}^{p_{opt}-1} a_{opt}[j] \cdot V_{as}[k - j - 1], \tag{C.8}$$

where a_{opt} is the optimal ARM. The rest of the subsection is devoted to an example using the autoregressive whitening filter. Let y be a simulated signal containing a sum of sines, impulses, and white Gaussian noise given by:

$$y = I_t + W + \sum_{j=1}^{10} \sin\left\{j \cdot 0.5 \cdot 2\pi \cdot t + \phi[j]\right\}, \tag{C.9}$$

where I_t is an impulse train of strength 1.5 with frequency 1 Hz, W is white Gaussian noise with a standard deviation $\sigma = 1/5$, and $\phi[j]$ is a random phase $\in [0, 2\pi]$. The signal is simulated at a sample rate of 100Hz for 10 seconds, and is shown in Figure C.3 (a). Here, it is difficult to spot the impulses occurring every second. An ARM is trained to predict the deterministic components to separate the random components. Since the impulses occur at a rate of 1 Hz, the maximum ARM order is set to $p_{max} = 99$ to avoid prediction of the impulses. After model training, the residual is given by (C.8) and shown in Figure C.3 (b). The residual only contains the impulses and Gaussian noise and is therefore suitable for further processing. In addition, the AIC_C for each ARM order up to p_{max} is shown in Figure C.3 (c), and the residual power σ^2 is shown in Figure C.3 (d). The residual power decreases continuously with increasing filter order, but the AIC_C prevents overfitting the model. The optimal filter order is $p_{opt} = 19$ for this example.

C.2.5 Hilbert transform

The ARM residual retains the high-frequency amplitude-modulated bearing fault vibration impacts as well as low-frequency deterministic components related to the fault. The bearing characteristic frequency for a specific fault is the frequency of the modulation waveform. This modulation waveform is obtained by determining the envelope of the

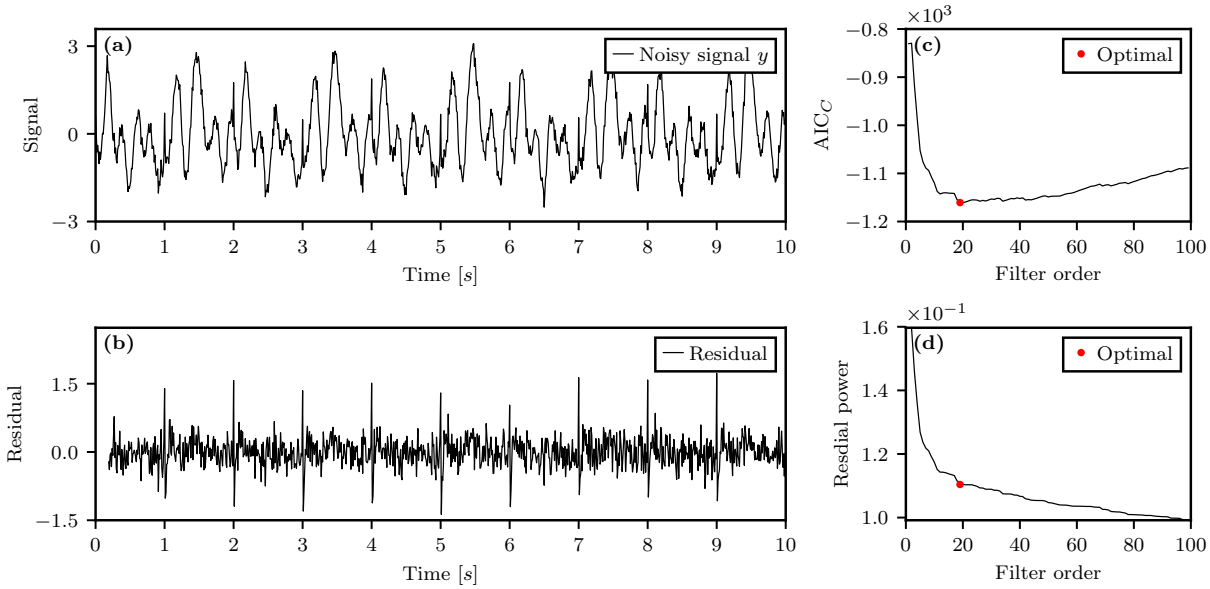

Figure C.3: Example scenario where an impulse train is revealed after applying the autoregressive whitening filter. **(a)** The example input signal. **(b)** The residual after applying the whitening filter. **(c)** The $\mathrm{AIC_C}$ criteria as a function of the filter order together with the optimal value. **(d)** The residual power as a function of the filter order.

signal. The envelope is obtained by computing the absolute value of the analytic signal. The analytic signal is acquired using:

$$V_a(\theta) = V_{as} + H\{V_{as}\} = V_{as} + iV_i, \tag{C.10}$$

where $H\{\cdot\}$ is the Hilbert transform, i is the imaginary unit, and V_i is the complex part of the analytic signal. The zero-mean envelope is obtained by computing the Pythagorean distance in the complex plane and removing the mean value using:

$$V_{env} = |V_a| = \sqrt{V_{as}^2 + V_i^2} - \mathrm{mean}\{\sqrt{V_{as}^2 + V_i^2}\}. \tag{C.11}$$

where $\mathrm{mean}\{\cdot\}$ returns the mean value of an array. In the next subsection, bearing fault vibration components are amplified by computing the cross-correlation between the whitened vibration signal and its envelope.

C.2.6 Cross-correlation

The enveloped vibration signal contains the demodulated cyclic frequency of the bearing fault impacts. In addition, the selectively whitened vibration signal also contains deterministic components at the bearing cyclic frequency. Instead of analyzing either of the two signals, the signals are fused by means of the cross-correlation operation. The resulting signal contain amplified frequency components related to the fault. A more

in-depth discussion on the effects of this cross-correlation is given in Section C.2.8. The cross-correlation is calculated using:

$$V_{cc}(\alpha) = V_w(\theta) \star V_{env}(\theta) = V_w^*(-\theta) * V_{env}(\theta), \tag{C.12}$$

where α is the angular lag, $\star$ is the notation for cross-correlation, superscript * indicates the complex conjugate, and $*$ is convolution. V_w^* is reversed to match the definition of cross-correlation, hence noting the input with $(-\theta)$. The cross-correlated signal is transformed to the order domain in the next subsection to be analyzed for bearing faults.

C.2.7 Order spectrum

The cross-correlated signal V_{cc} contains frequency components directly related to the bearing fault. Harmonics and side-bands of the bearing characteristic frequencies can be observed in the order spectrum of V_{cc}. To avoid spectral leakage, a Hann window is used in the Fourier transform. The amplitude of the windowed Fourier transform is calculated using:

$$Y = \left| F\{w \cdot V_{cc}\} \right|, \tag{C.13}$$

where w is the Hann-window, and $w \cdot V_{cc}$ must be treated as an element-wise multiplication. The absolute value of the spectrum is used because the phase of each component is not of interest, only the amplitude. The frequency bins of the resulting spectrum are given in orders using:

$$O[j] = \frac{j}{n_w \cdot \Delta\theta_d}, \tag{C.14}$$

where n_w is the size of the Hann window. The final order spectrum $Y(O)$ is termed the Whitened Cross-correlation Spectrum (WCCS) and represents the last step in the proposed algorithm.

C.2.8 Theoretical Discussion on WCCS

The effects of the cross-correlation in (C.12) on the WCCS is not readily apparent in the time domain. The reason for why this cross-correlation amplifies the bearing vibration is difficult to visualize in the time domain. Advantageously, the time domain convolution may be represented by an element-wise multiplication in the frequency domain thanks to the convolution theorem. Considering this, a different way of calculating the cross-correlated vibration is

$$V_{cc}(\theta) = F^{-1}\{F\{V_w^*(-\theta)\} \cdot F\{V_{env}(\theta)\}\}, \tag{C.15}$$

where F^{-1} is the inverse Fourier transform, and V_w^* and V_{env} must be padded with zeros to match the length of the original V_{cc} in (C.12). Inserting (C.15) into (C.13) shows that

the WCCS can be calculated with

$$Y = |F\{w \cdot F^{-1}\{F\{V_w^*(-\theta)\} \cdot F\{V_{env}(\theta)\}\}\}|, \tag{C.16}$$

which contain numerous Fourier transforms. However, if the Hann window w is replaced by a rectangular window (i.e. no window), (C.16) may be written as

$$Y = |F\{V_w^*(-\theta)\} \cdot F\{V_{env}(\theta)\}|. \tag{C.17}$$

(C.17) shows that the WCCS is actually a similarity measure between the spectrum of the whitened vibration signal, and the spectrum of the envelope. If both spectra have high values at certain frequency bins, the WCCS results in a magnified amplitude at these bins. At bins with a high amplitude on only one of the spectra, its amplitude is attenuated. The bearing characteristic fault frequency can be observed directly in the raw/deterministic vibration spectrum [13]. Additionally, it is determined in [22] that the demodulated vibration signal (envelope) also contains components at the bearing characteristic frequencies. As the characteristic bearing fault vibration may be observed in both the selectively whitened vibration signal and its envelope, the cross-correlation may therefore amplify the bearing vibration and attenuate unwanted noise and uncorrelated peaks. This also shows that it is only the low-frequency part of the two signals that are used. Therefore to rapidly calculate the cross-correlation in (C.12), the two signals should be low-pass filtered and decimated to reach a suitable sampling frequency. For example, as most harmonics of a bearing fault are visible up to 50 orders, the signals should be decimated to this amount before applying the cross-correlation to reduce the computational load.

However, there are some possible limitations to the presented WCCS that must be addressed. The presented method performs best when low-frequency bearing fault components are discernible from background noise in both the selectively whitened deterministic signal and its envelope. A theoretical model of the bearing vibration signal is thoroughly analyzed in [23]. In this reference, it is shown that the deterministic bearing signal components are the result of a low-pass filter applied to the high-frequency bearing resonance vibration signal. Consequently, the deterministic components related to the bearing fault are only discernible from background noise if the bearing resonance frequency overlaps this low-pass filter. Typically, small bearings have a high resonance frequency, and therefore this condition may not be met in all scenarios. A recent publication [24] demonstrate the limitation of using the deterministic torque vibration signal for bearing fault diagnosis. It is also shown that the demodulated high-frequency vibration signal is more suited for bearing fault diagnosis. However, it should be noted that even if the low-frequency bearing fault components in the deterministic signal are only marginally stronger than

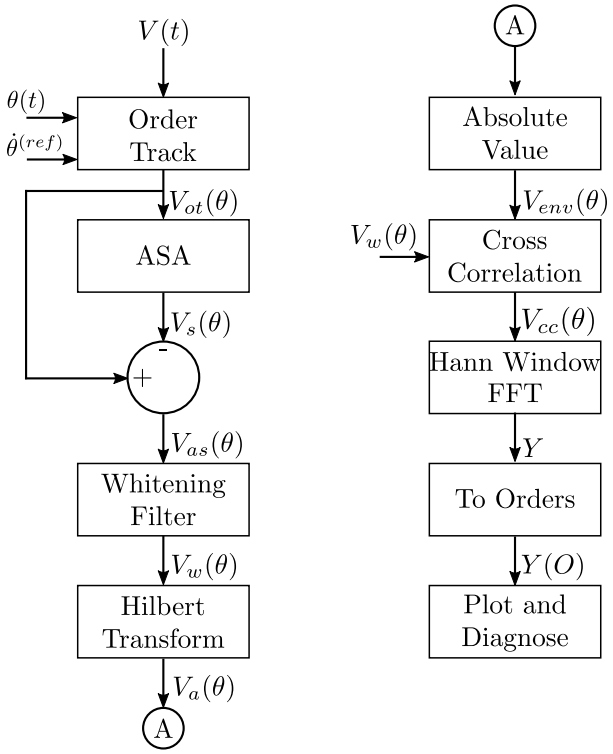

Figure C.4: Flowchart of WCCS used to diagnose the low-speed bearing for faults.

the background noise, the high-frequency signal envelope will still be amplified by the cross-correlation step in (C.12), although not by much. In the case that the deterministic signal peaks are completely drowned in noise, the amplification will be almost unity, i.e. no improvement. Such a scenario is shown and discussed in Section C.3.2.

C.2.9 Complete algorithms

In this subsection, the proposed method is summarized with an accompanying flowchart illustrating the implementation. In addition, other established methods in literature are presented to establish a comparison with the proposed method.

C.2.9.1 Whitened Cross-correlation Spectrum

The flowchart for producing the WCCS is shown in Figure C.4. Initially, experimental data consisting of the speed reference $\dot{\theta}^{(ref)}$, vibration signal $V(t)$, and encoder data $\theta(t)$, are obtained in an accelerated life-test as described in Section C.2.1. This data is used as input to the Order Track algorithm elaborated in Section C.2.2. After this re-sampling process, a constant angle increment between each sample is realized. Next, the synchronous components are estimated with the Angle Synchronous Average (ASA) algorithm from Section C.2.3 and subtracted from the order tracked signal V_{ot} to retain asynchronous components, V_{as}. Afterwards, a linear Autoregressive Model (ARM)

is used as a DRS in Section C.2.4 to separate the bearing fault vibration from other deterministic components. This process acts as a whitening filter and returns the whitened vibration signal. The envelope is obtained by using the Hilbert transform and calculating the absolute value of the analytic signal in Section C.2.5. This envelope contains the demodulated high-frequency bearing fault vibration, random peaks, and white noise. The cross-correlation between the whitened vibration signal and its envelope is calculated in Section C.2.6 to amplify the bearing fault vibration, and attenuate white noise and random peaks. V_{cc} is finally processed through a Hann-windowed Fourier transform as elaborated in Section C.2.7, which results in the WCCS that is plotted for analysis. The performance of the proposed method is compared against three other methods reported in the literature, and they are explained in the next Subsection.

C.2.9.2 Comparison methods

To showcase the performance of the WCCS, the method is compared to three other methods reported in the literature. These three are the Envelope Spectrum [22], Fast Kurtogram [25], and the Fast Spectral Correlation [26]. These methods are similarly computationally heavy, and do not involve experimentally tuned settings. To generate a fair comparison, the whitened vibration signal $V_w(\theta)$ is used as input to all three methods.

Method A: Envelope Spectrum

The Envelope Spectrum [22] is a well-known method used for bearing fault diagnosis, and its flowchart is shown in Figure C.5 (a). The method involves demodulating the vibration signal via an envelope function to identify the characteristic bearing fault frequencies in the frequency spectrum. In this application, the absolute valued analytic signal is applied to get the envelope.

Method B: Kurtogram

To further improve the envelope spectrum, a narrow-band filter centered on the resonance frequency of the bearing vibration can be applied beforehand to increase the signal-to-noise ratio. However, the resonance frequency is normally unknown, and is difficult to obtain analytically. To estimate this resonance frequency, the Fast Kurtogram [25] is employed. The flow chart of this method is shown in Figure C.5 b). In summary, the Fast Kurtogram passes the whitened vibration signal through a complex, narrow-band filter-bank with varying central frequencies and spectral width. The kurtosis is calculated for each filtered signal, and the narrow-band that maximizes the kurtosis is assumed to

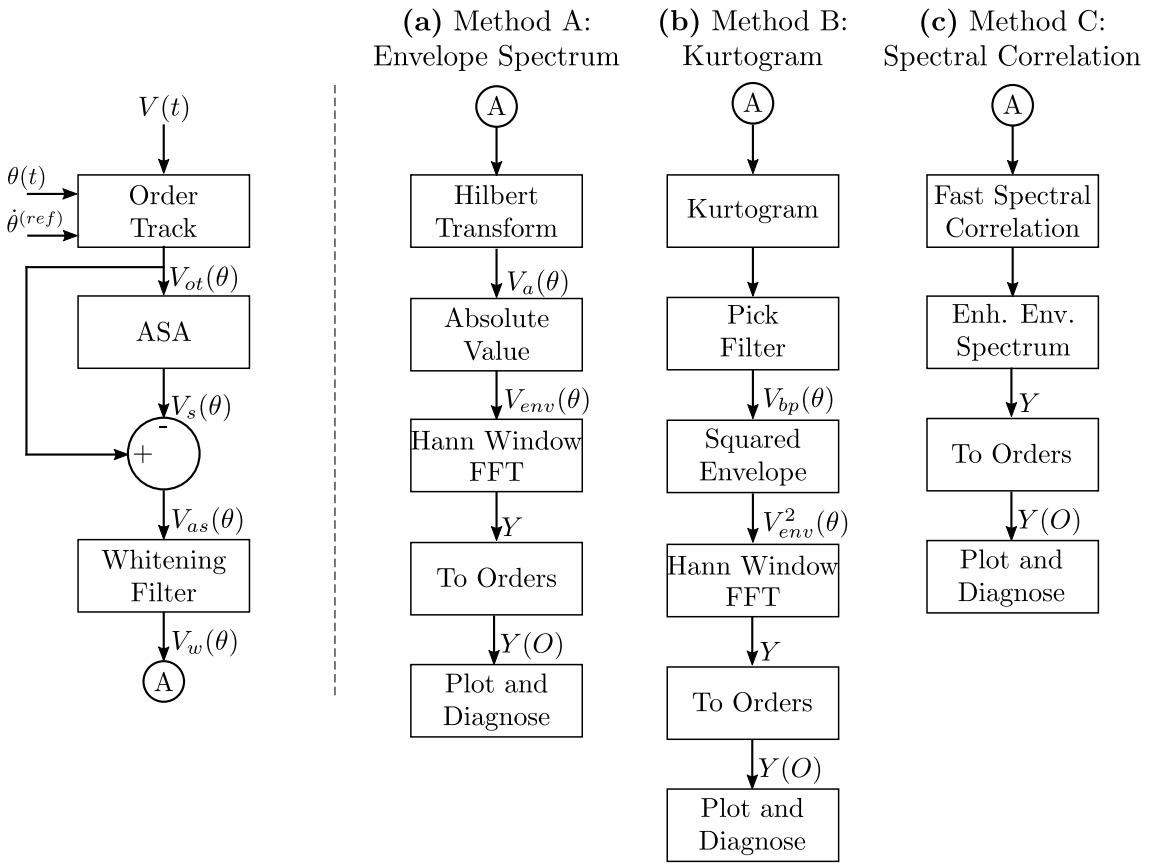

Figure C.5: Three additional methods used to compare performance with the WCCS. **(a)** Method A involves diagnosing the bearing using the Envelope Spectrum. **(b)** Method B improves on the Envelope Spectrum by bandpass filtering the whitened vibration signal around the narrow-band yielding maximum kurtosis. **(c)** Method C involves the fast calculation of the Spectral Correlation which is a tool for analyzing cyclostationary signals.

contain the bearing resonance frequency. The squared envelope spectrum of the narrow-band filtered signal $V_{bp}(\theta)$ is analyzed for the bearing characteristic fault frequencies. In obtaining the Kurtogram, the excess kurtosis is calculated using:

$$\text{Kurt}\{x\} - E = \frac{\mu_4}{\sigma^4} - E, \tag{C.18}$$

where μ_4 is the fourth central moment, σ is the standard deviation, and E is the kurtosis value for Gaussian white noise. For a real signal $E = 3$, and for a complex signal $E = 2$. Henceforth, the excessive kurtosis is named kurtosis to avoid repeating the word "excessive".

Method C: Fast Spectral Correlation

The Spectral Correlation (SC) is a tool for analyzing cyclostationary signals (signals with hidden periodicities or repetitive patterns), like the high-frequency vibration signal

exhibited from bearing fault impacts. The resonance vibration is amplitude-modulated with a comparably lower-frequency modulation waveform, making it ideal to analyze using such cyclostationary tools. The SC is a two-dimensional spectrum showing the cyclic-spectral frequency relationship. The usage of the SC has been limited in condition monitoring applications due to its high computational cost. A fast version of the SC, Fast SC, was developed in [26] which computes an estimate of the SC several magnitudes faster than the original algorithm. From the estimated SC, the Enhanced Envelope Spectrum [26] is identified by calculating the mean, absolute value of the SC in the direction of the spectral frequency. The Enhanced Envelope Spectrum is used in this method, and the flowchart is shown in Figure C.5 (c).

C.2.9.3 Performance metrics

Two performance metrics are created to quantify the performance of the WCCS and the three other methods. The first metric evaluates the method's ability to discern the bearing characteristic fault frequency harmonics from the noise floor (mean value of the spectrum), and is calculated using:

$$P1 = \left(\frac{1}{n_h} \sum_{j=1}^{n_h} Y_{h,j} \right) / \text{mean}\{Y\} \, , \tag{C.19}$$

where n_h is the number of harmonics in the spectrum, and $Y_{h,j}$ is the amplitude of harmonic j. $P1$ represents the mean ratio of the harmonic values to the noise floor. This dimensionless ratio is comparable between each method, as the amplitude-scale of each spectrum is widely different from each other.

The second metric evaluates the method's ability to discern the side-bands from the noise floor, and is calculated using:

$$P2 = \frac{1}{n_h} \sum_{j=1}^{n_h} n_{s,j}, \tag{C.20}$$

where $n_{s,j}$ is the number of side-bands linked to harmonic j in the spectrum. Visible side-bands are marked in the spectra and are expected to be prominent compared to the noise floor. $P2$ represents the mean number of visible side-bands per harmonic in the spectrum.

C.3 Results

Vibration data is collected from an accelerated life-time test as described in Section C.2.1. In this section, the measurements recorded at 20 rpm will be used to show the performance of the WCCS, and how it compares to the three other methods.

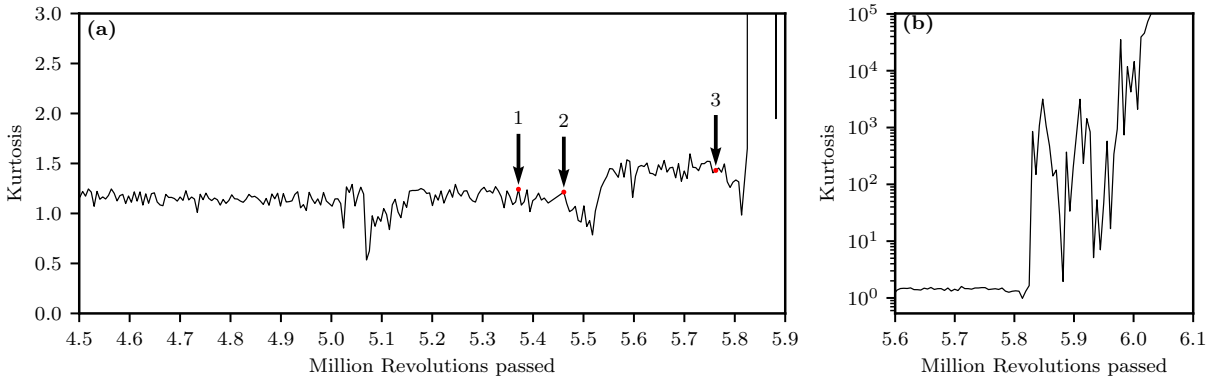

Figure C.6: Kurtosis of the vibration datasets measured at 20 rpm shaft speed. **(a)** The kurtosis at an early stage of damage. **(b)** The kurtosis at an advanced stage of damage.

C.3.1 Kurtosis trend

The kurtosis is a time domain feature that can be used to indicate damage in the bearing. This is because a high kurtosis value implies an impulsive signal containing shocks from, i.e., bearing fault impacts. To show the damage propagation of the bearing during the accelerated life-time test, the kurtosis is calculated using (C.18) for the 20 rpm datasets, and shown in Figure C.6. In this figure, there are two graphs. (a) and (b) both shows the kurtosis of the vibration data, but at different y- and x-axis scales. The x-axis indicates how many million revolutions the bearing has rotated before the data is recorded. Initially, the kurtosis value is low, but it rises to very high values when the bearing is sufficiently damaged. In Figure C.6 (a) there are three points that are marked. These will be referred to as kurtosis points in the following, and they indicate changes in the bearing health. Specifications for the datasets linked to these points are given in Table C.2. In the next Subsection, the WCCS is used to diagnose the bearing for faults at these three points.

C.3.2 Whitened Cross-correlation Spectrum

The proposed method from Section C.2.9.1 generates the WCCS which is used to analyze the vibration data at the various kurtosis points. At kurtosis point 1, there is no notable change in the kurtosis value. However, the WCCS generated from the proposed method shows a significant change compared to earlier measurements. The spectrum of the whitened vibration signal, the envelope spectrum, and the WCCS are shown in Figures C.7 (a)-(c), respectively. Here, the order spectra up to 50 orders are displayed. In addition, lines that indicate harmonics of O_{2BS} and side-bands are also shown. From Table C.1, the characteristic fault frequency for a roller fault is 6.66 orders, and the side-bands are spaced apart by the fundamental train at 0.43 orders (O_{FT}). Due to roller slip and non-zero contact angle, the actual frequency is observed at 6.71 orders (deviation

Table C.2: Specifications of the datasets at the three kurtosis points.

Kurtosis point →	Point 1	Point 2	Point 3
Sampling frequency F_s (kHz)	51.2	51.2	51.2
Duration (s)	153	153	153
Million revolutions passed	5.37	5.46	5.76
Number of samples	7,833,600	7,833,600	7,833,600
Kurtosis of raw data kurt$\{V\}$	1.14	1.10	1.41
Shaft speed $\dot{\theta}^{(ref)}$ (Hz)	0.33	0.33	0.33
Fault type	Roller	Outer-race	Inner-race
Max ARM order p_{max}	22,391	22,391	22,391
Optimal ARM order p_{opt}	22,325	22,160	22,094
Kurtosis of whitened data kurt$\{V_w\}$	0.69	0.65	0.61

of +0.75%). Examining Figure C.7 (a) reveals that the low-frequency vibration signal contains multiple side-bands around harmonics of the characteristic frequency. However, the first five harmonics are not readily visible. On the contrary, the envelope spectrum in Figure C.7 (b) shows small signs of the harmonics, and a few side-bands. Incorporating the cross-correlation fuses these two spectra, and amplifies the vibration components that are visible in both. Figure C.7 (c) show the WCCS. Here, the perceived noise-floor is smaller than in Figure C.7 (a), and the harmonics from Figure C.7 (b) are shown as discrete peaks. In addition, a prominent peak at the O_{FT} is also visible due to the strong vibration occurring when the damaged roller is in the radial load zone. From these results, it is shown that even if a characteristic peak is only readily visible on one of the two spectra, it is preserved after the cross-correlation step. This phenomenon is best observed at the second harmonic (2x 6.71 orders), where the envelope spectrum in Figure C.7 (b) shows a small peak, while the whitened vibration spectrum in Figure C.7 (a) does not indicate a significant peak. Even with these signals, the WCCS in Figure C.7 (c) still preserves the second harmonic peak from the envelope spectrum. Therefore, it should be unproblematic to use the proposed method even if the deterministic signal shows small to no signs of the bearing fault.

By visual inspection between the three spectra in Figure C.7, it is evident that the WCCS is easier to analyze than either of the two other spectra. From these results, it is clear that the bearing suffers from a roller fault at this stage. For comparison purposes, the WCCS from a dataset captured ≈10,000 revolutions earlier is obtained, and the spectrum

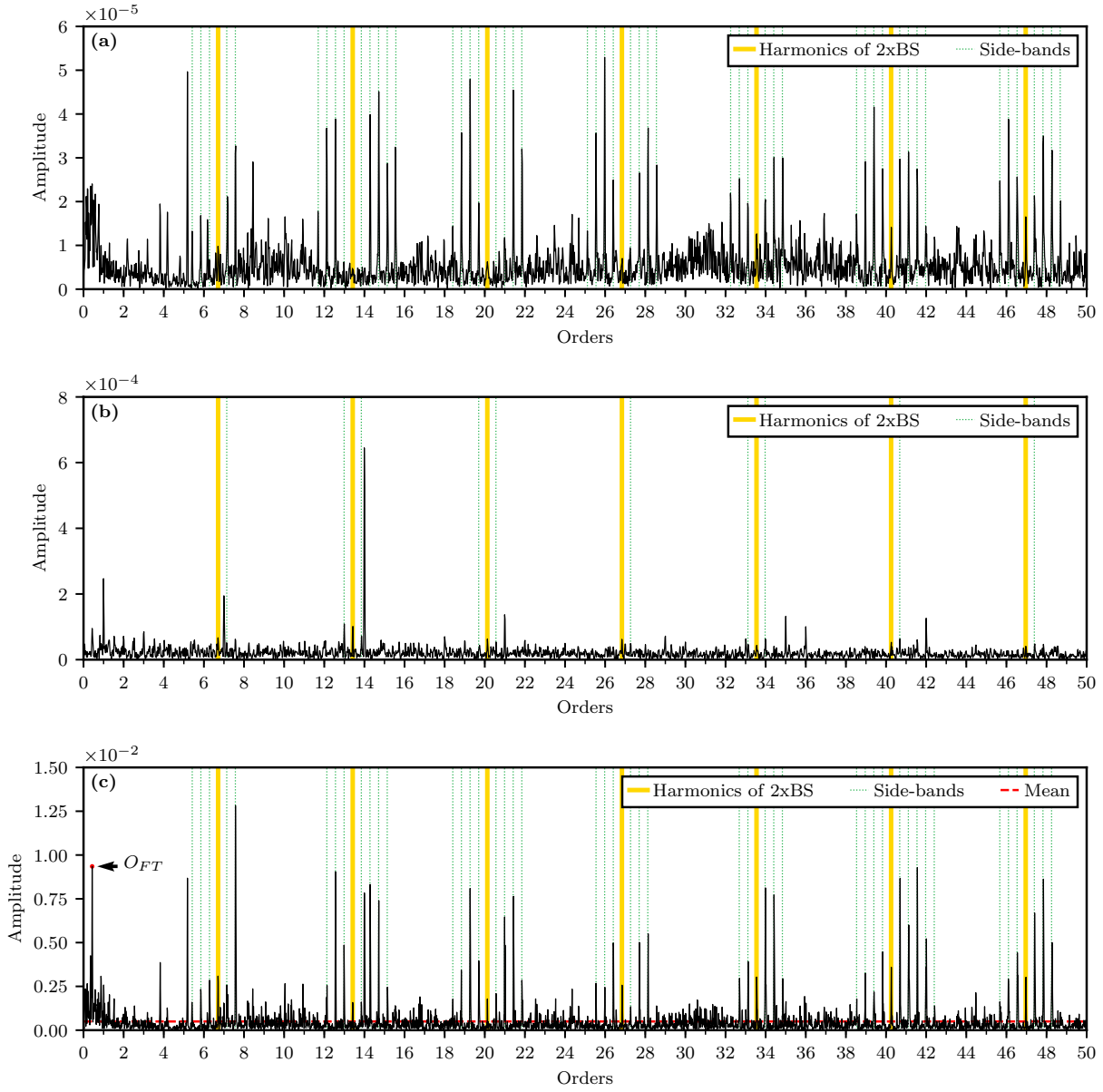

Figure C.7: The WCCS used to diagnose a roller fault at kurtosis point 1 in Figure C.6. Harmonics and side-bands linked to the roller fault are marked in all three subfigures. **(a)** The spectrum of the whitened vibration signal V_w. **(b)** The spectrum of the envelope V_{env}. **(c)** The proposed WCCS and a red stapled line which corresponds to the mean spectrum value.

is shown in Figure C.8. Here, there are no prominent amplitudes related to a roller fault, and the bearing seems undamaged. There are, however, four other prominent peaks in the spectrum, and these are marked. $14O_S$ is equal to two times the motor speed due to the gearbox. This vibration component was not attenuated in the whitening process and is hence noise in the spectrum. The second peak is marked as O_{BPO} which is the characteristic frequency for an outer race fault at roughly 5.12 orders. It should normally

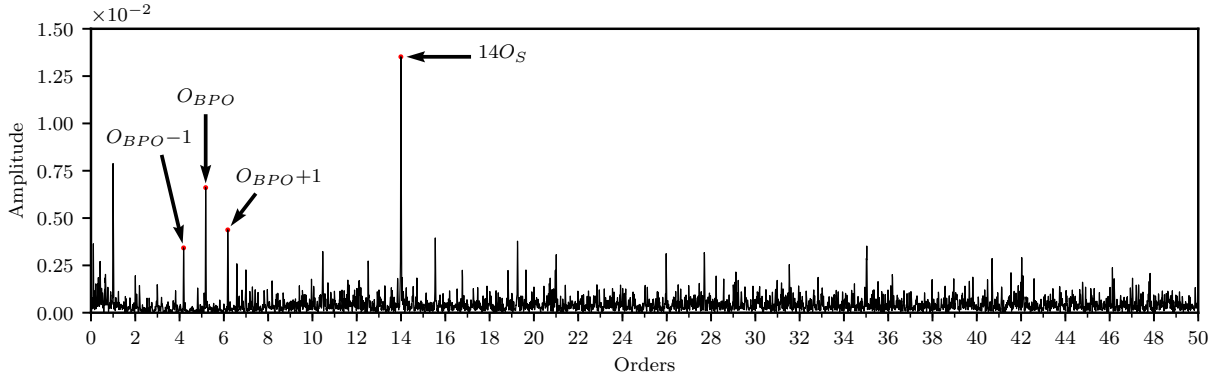

Figure C.8: The WCCS computed 10,000 revolutions prior to kurtosis point 1 in Figure C.6. The roller damage is not visible.

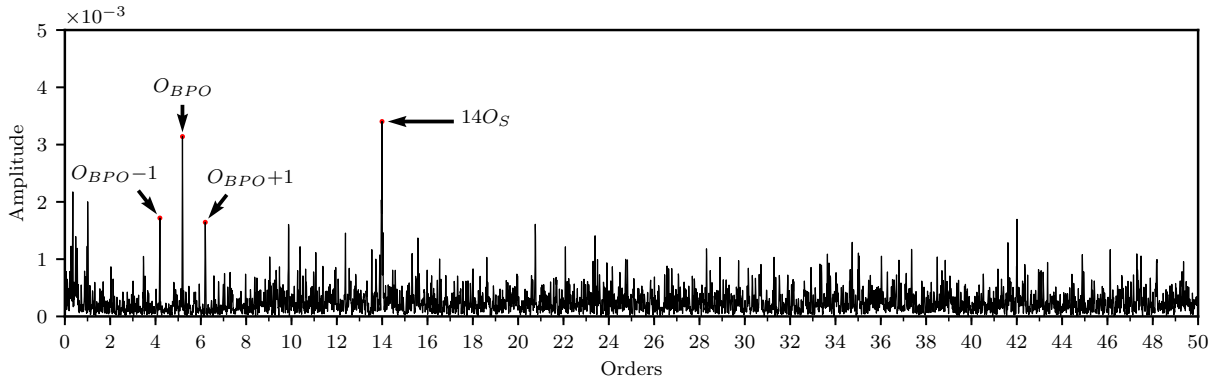

Figure C.9: Baseline WCCS at the start of the accelerated life-test. A peak at the characteristic frequency for an outer-race fault is visible, however the bearing is undamaged at this stage.

be visible if there is an outer-race fault. However, the authors believe that the outer race is undamaged at this point. That is because vibration at this cyclic frequency can also be spotted at the beginning of the accelerated life-time test. A baseline WCCS after the initial run-in of the healthy bearing is shown in Figure C.9. In this spectrum, a peak at the characteristic frequency for an outer-race fault is still visible. As the bearing isn't damaged at this point, the authors conclude that there is vibration at the O_{BPO} frequency regardless of the bearing condition. The authors suspect that rollers passing the radial load zone may cause the vibration at the O_{BPO} frequency. This passing may cause the shaft to move slightly up and down, which would explain the vibration. It could also be from stressing and de-stressing a roller once it passes the radial load zone. Additionally, the 1 order side-bands in Figures C.8 and C.9 are effects of amplitude-modulation of the O_{BPO} vibration. This could be caused by shaft mass-unbalance, or if the shaft is slightly bent. Nevertheless, to accurately diagnose an outer-race fault using the proposed method, there must be several stronger harmonics in the spectrum, as a single peak appears to

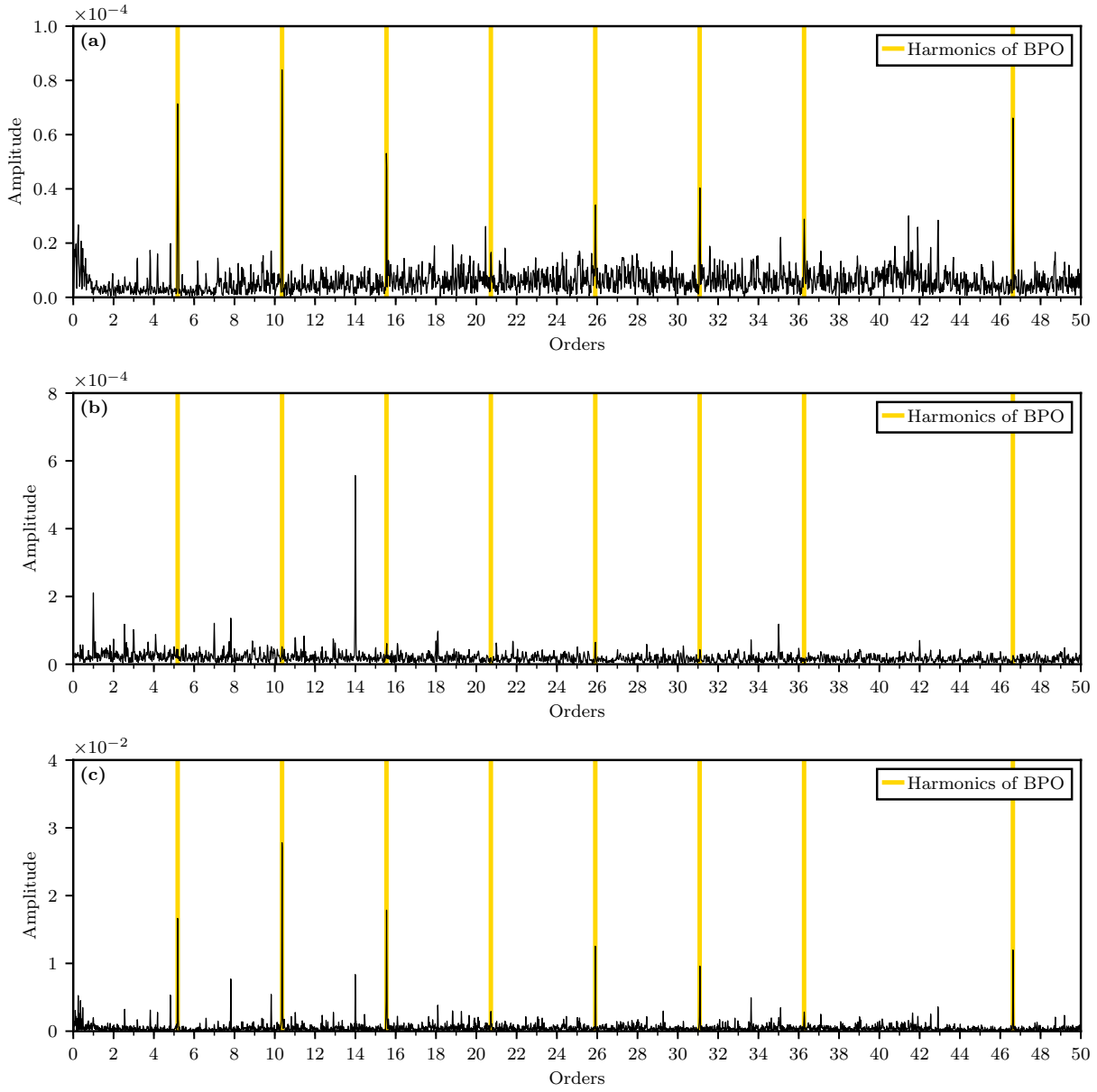

Figure C.10: The WCCS used to diagnose an outer-race fault at kurtosis point 2 in Figure C.6. Harmonics linked to the outer-race fault are marked in all three subfigures. **(a)** The spectrum of the whitened vibration signal V_w. **(b)** The spectrum of the envelope V_{env}. **(c)** The proposed WCCS.

always be visible.

Moving on, kurtosis point 2 from Figure C.6 marks a new health state for the bearing. The spectrum of the whitened vibration signal, the envelope spectrum, and the WCCS are shown in Figures C.10 (a)-(c), respectively. The roller fault is no longer visible in the WCCS in Figure C.10 (c). Instead, harmonics linked to an outer race fault are very prominent in the spectrum. The characteristic fault frequency for outer-race faults is 5.12 orders with no side-bands according to Table C.1. The actual observed harmonic

frequency is 5.18 orders (deviation of +1.17%). Compared to a healthy state, there are now multiple strong harmonics in the spectrum, which strengthen the probability that there is an outer-race fault. Examining Figure C.10 (a) and (b) reveals that the outer-race vibration peaks in the WCCS mostly arrive from the spectrum of the whitened vibration signal. This is a second case showing that even if only one of the two input spectra shows significant signs of bearing vibration, the resulting WCCS does not suffer significantly from this. The envelope spectrum in Figure C.10 doesn't show any prominent harmonics at the outer-race characteristic frequency. The outer-race fault stays visible for quite some time until kurtosis point 3 which marks yet a change of state in the bearing. The spectrum of the whitened vibration signal, the envelope spectrum, and the WCCS are shown in Figures C.11 (a)-(c), respectively. The WCCS in Figure C.11 (c) shows multiple harmonics and side-bands linked to an inner-race fault. The fault is characterized by strong harmonics at multiples of 6.88 orders accompanied by side-bands spaced 1 order away from the harmonic according to Table C.1. In the observed spectrum, the actual harmonic frequency is 6.82 (deviation of -0.87%). The peaks are very prominent, and there are multiple side-bands, making it easy to conclude the type of fault. Harmonics from the outer-race fault may also be observed in the spectrum, however they are not marked to avoid clutter in the plot. Examining the whitened vibration spectrum in Figure C.11 (a) shows prominent peaks at the third and higher harmonics. Meanwhile the first and second harmonic are visible in the envelope spectrum in Figure C.11 (b). The cross-correlation of the two signals make a clearer spectrum as shown in Figure C.11. Shortly after kurtosis point 3, the kurtosis value increases drastically as seen on the chart in Figure C.6 (b), and the bearing is approaching the end of its useful life. Luckily, the WCCS highlighted roller faults nearly 500,000 rounds prior to kurtosis increasing drastically. With an expected life-time of ≈ 6 million revolutions, this corresponds to $0.5/6 = 8.33\%$ of remaining useful life after the first fault was diagnosed. In the next subsection, the performance of the WCCS is compared to the three other methods.

C.3.3 Comparison to other methods

For the comparisons shown in this subsection, the dataset resulting in kurtosis point 1 from Figure C.6 is used, because early fault detection is the most critical. The WCCS shown in Figure C.7 (c) indicates that there is a roller fault in the bearing. This spectrum is compared to the three other processing methods (A, B, and C), whose algorithms are elaborated in Section C.2.9.2.

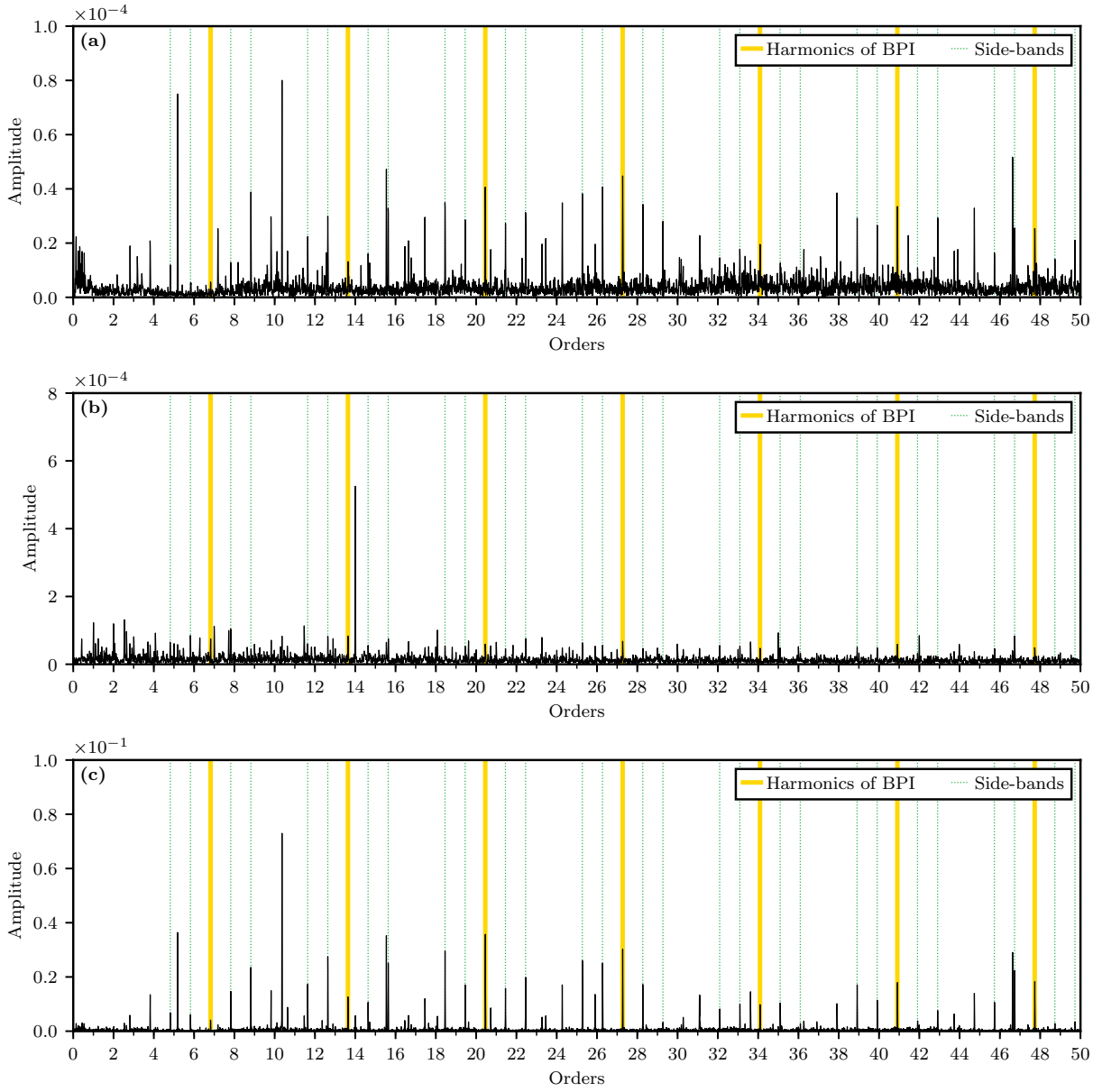

Figure C.11: The WCCS used to diagnose an inner-race fault at kurtosis point 3 in Figure C.6. Harmonics and side-bands linked to an inner-race fault are marked in all three subfigures. **(a)** The spectrum of the whitened vibration signal V_w. **(b)** The spectrum of the envelope V_{env}. **(c)** The proposed WCCS.

Method A: Envelope Spectrum

The resulting envelope spectrum from method A is shown in Figure C.12. The spectrum barely highlights the roller fault harmonics and its side-bands. They could easily be ignored due to the first harmonic's low amplitude. Shaft and motor vibrations, marked as O_S and $14O_S$, are the most prominent peaks in the spectrum. Therefore, method A is not suitable for diagnosing faults on the presented dataset.

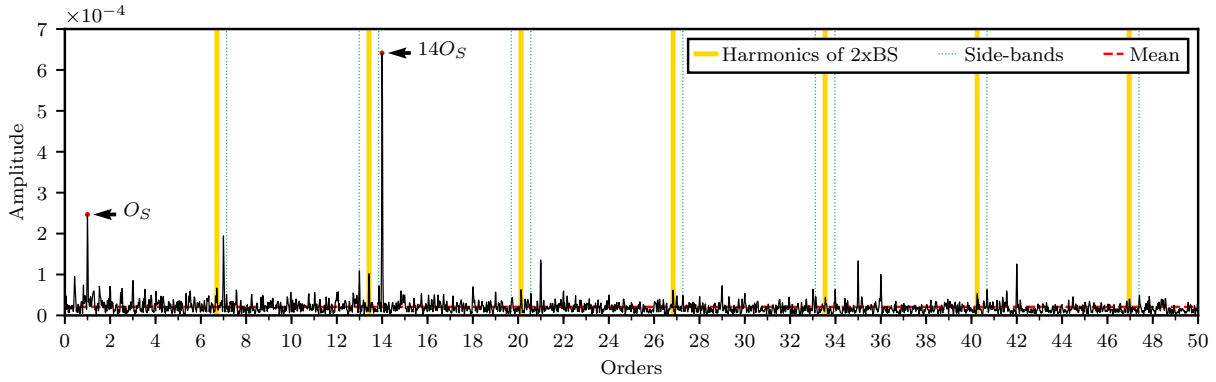

Figure C.12: The resulting Envelope Spectrum from using method A. The roller fault harmonics are barely visible, while shaft and motor vibration dominate the spectrum. The mean value of the spectrum is marked for comparison purposes.

Method B: Kurtogram

The resulting Kurtogram from method B is shown in Figure C.13. From this result, the chosen narrow-band is centered at 24,800 Hz with a bandwidth of 1600 Hz, because it represents the highest kurtosis. The squared envelope spectrum of the band-pass filtered signal is returned from this method and shown in Figure C.14. In this spectrum, it is possible to discern seven harmonics related to a roller fault. However, motor-synchronous vibration at $14O_S$ and $35O_S$ dominates the spectrum. The harmonics and side-bands of the roller fault frequency are more prominent compared to the non-filtered envelope spectrum. However, compared to the WCCS shown in Figure C.7 (c), this squared envelope spectrum has a higher noise-floor.

Method C: Fast Spectral Correlation

The resulting Enhanced Envelope Spectrum (EES) from method C is shown in Figure C.15. In the EES, the harmonics linked to the roller fault are small compared to the vibration from the shaft and motor. The spectrum is dominated by peaks at $1O_S$, $14O_S$, and $35O_S$. The harmonics linked to the roller fault are partially visible, however they are hard to discern from the noise-floor. Compared to the WCCS, it is harder to detect the harmonic peaks, and there are fewer side-bands.

C.3.3.1 Performance metrics

The performance of each method is quantified using the performance metrics from Section C.2.9.3. The spectra indicating a roller fault at kurtosis point 1 in Figure C.6 is used to evaluate the performance, and the values are given in Table C.3. The metrics show that the WCCS scores the best among the four different spectra. The Fast Kurtogram (B) is

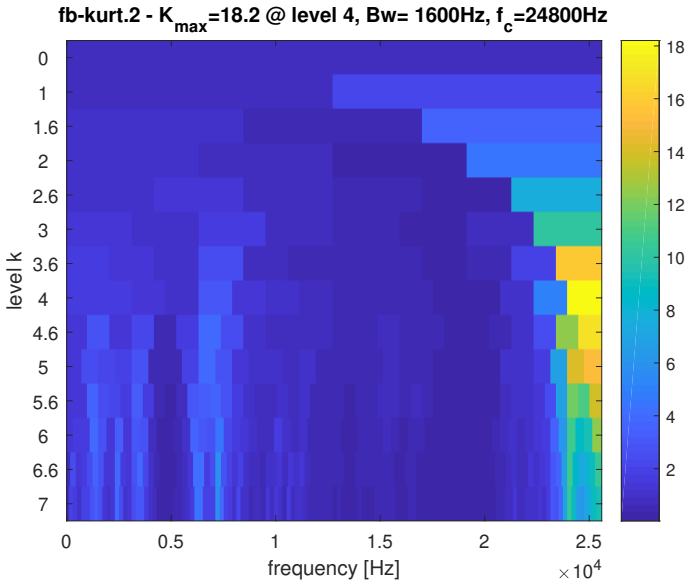

Figure C.13: Resulting Kurtogram in method B indicating a maximum of kurtosis at a central frequency of 24,800 Hz and a bandwidth of 1600 Hz. The kurtosis value at that narrow-band is 18.2.

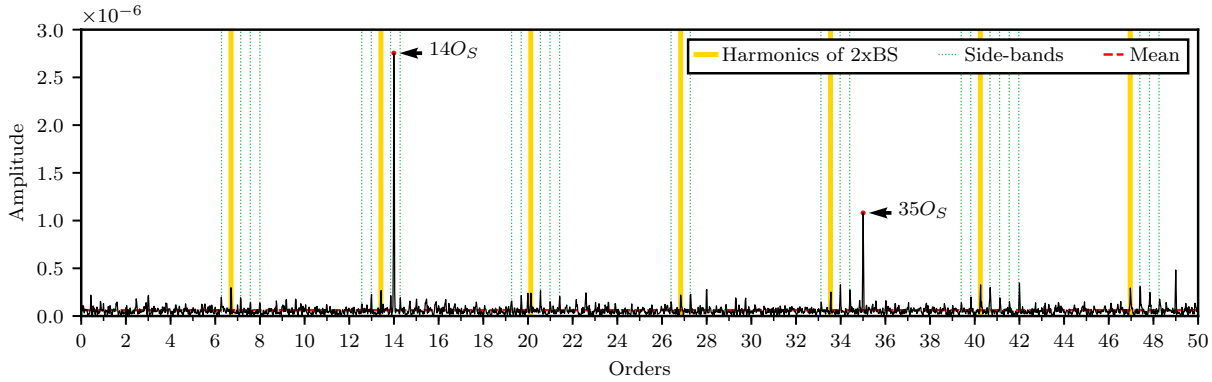

Figure C.14: Squared envelope spectrum of the band-pass filtered vibration signal in Method B. The roller fault harmonics and side-bands are visible, and the mean value of the spectrum is marked for comparison purposes.

second best, and the Envelope Spectrum (A) and the Fast Spectral Correlation (C) are tied for the third place. Further, the resulting spectrum from methods A, B, and C also contained strong peaks from the shaft and motor vibration, which reduces the visibility of the peaks linked to bearing faults. Therefore, based on these criteria, the WCCS is the best spectrum to perform fault diagnosis on among the four. It should be noted that the minimum value of the Enhanced Envelope spectrum was subtracted before the criteria was calculated, as the spectrum shown in Figure C.15 appears to have an offset in amplitude.

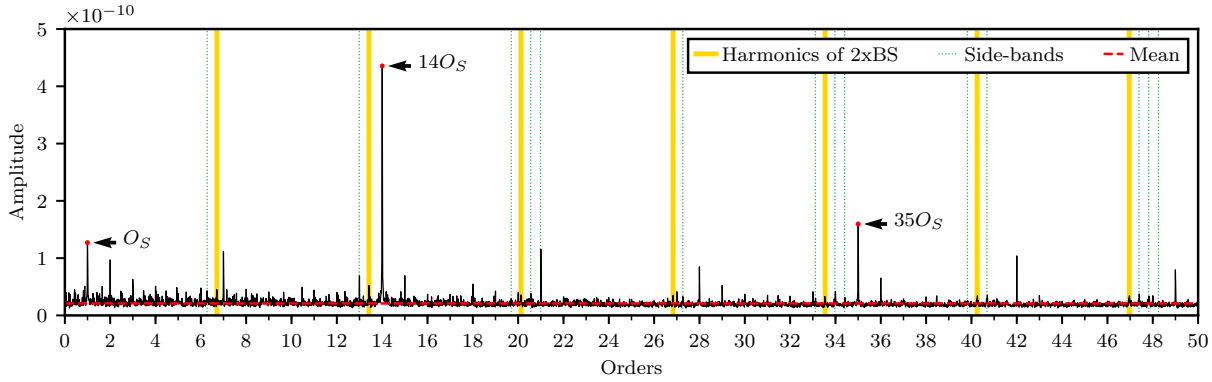

Figure C.15: Resulting Enhanced Envelope Spectrum from the Fast Spectral Correlation (Method C). The roller fault harmonics and side-bands are barely visible, and the mean value of the spectrum is marked for comparison purposes.

Table C.3: Performance metric for each method. The WCCS scores the highest in both categories.

Method	P1	P2
WCCS	**5.49**	**6.57**
A	2.95	1.42
B	4.42	3.85
C	2.65	2.00

C.3.4 Supplementary benchmark case study

The bearing fault database from the Case Western Reserve University (CWRU) [27] is a standard benchmark for diagnostics algorithm performance. Smith and Randall [28] published an extensive benchmark study where the performance of three established diagnostic techniques were evaluated on all available datasets. The three methods involved the squared envelope spectrum (SES); pre-whitening before SES; and pre-whitened Spectral kurtosis [25] and SES.

A dataset recorded during healthy conditions is examined first. This is to investigate the phenomenon of a visible vibration at the frequency of the outer-race O_{BPO}. The dataset used is named 100DE in which the vibration is measured at the drive-end, and there is no damage on either fan-end or drive-end bearing. Relevant specifications are given in Table C.4. The WCCS of this dataset is shown in Figure C.16.

The resulting WCCS show multiple peaks at integer multiples of the shaft speed, in addition to other non-integer peaks. The most interesting part is the inclusion of a peak at O_{BPO} and two side-bands spaced 1 order away from it. This is the same phenomenon

Table C.4: Specifications for the tested CWRU datasets. DE=Drive-end, FE=Fan-end

Dataset number →	100DE	275DE
Sampling frequency F_s (kHz)	48	12
Duration (s)	10.11	10.11
Accelerometer location	DE	DE
Number of samples N	485,335	121,351
Kurtosis of raw data	-0.0428	1.25
Shaft speed $\dot{\theta}^{(ref)}$ (Hz)	28.83	29.58
Fault type	Healthy	FE Inner-race
DE ball pass order outer-race O_{BPO}	3.5961	-
FE ball pass order inner-race O_{BPI}	-	4.974
Max ARM order p_{max}	308	82
Optimal ARM order p_{opt}	308	82
Kurtosis of whitened data	0.0708	1.22

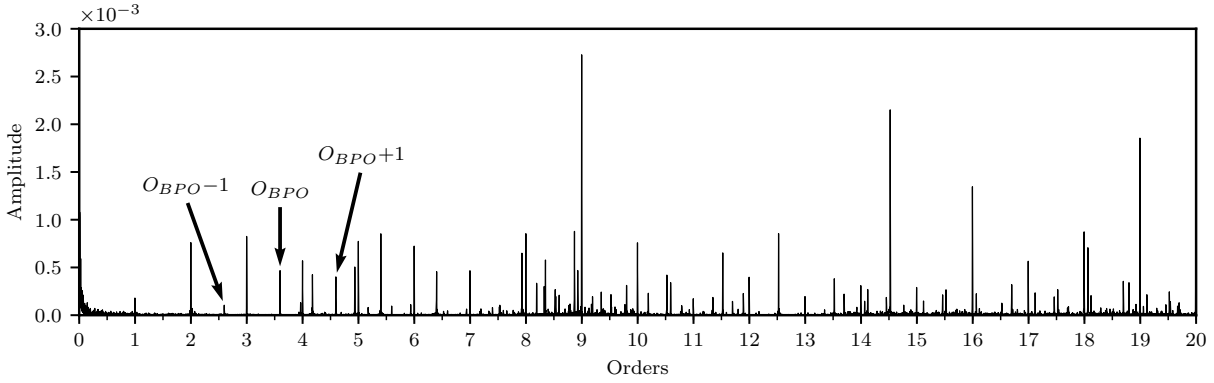

Figure C.16: The WCCS on the CWRU healthy dataset 100DE recorded on the drive-end.

shown in the healthy case of the first database in Figure C.8. It was suggested that the vibration at O_{BPO} comes from the passing of rollers over the radial load zone. A bent shaft would cause this O_{BPO} vibration to be modulated by the shaft frequency, which would result in the side-bands spaced apart by 1 order.

The performance of the WCCS is also tested on a dataset recorded with a known fault. A particular dataset (275DE) with an inner-race fault located at the fan-end (FE), and the vibration recorded at the drive-end (DE), proved to be difficult to diagnose using the aforementioned benchmark methods ([28], Table B4). The results are indicated as "partially successful" for the first two methods, and "not successful" for the last one as

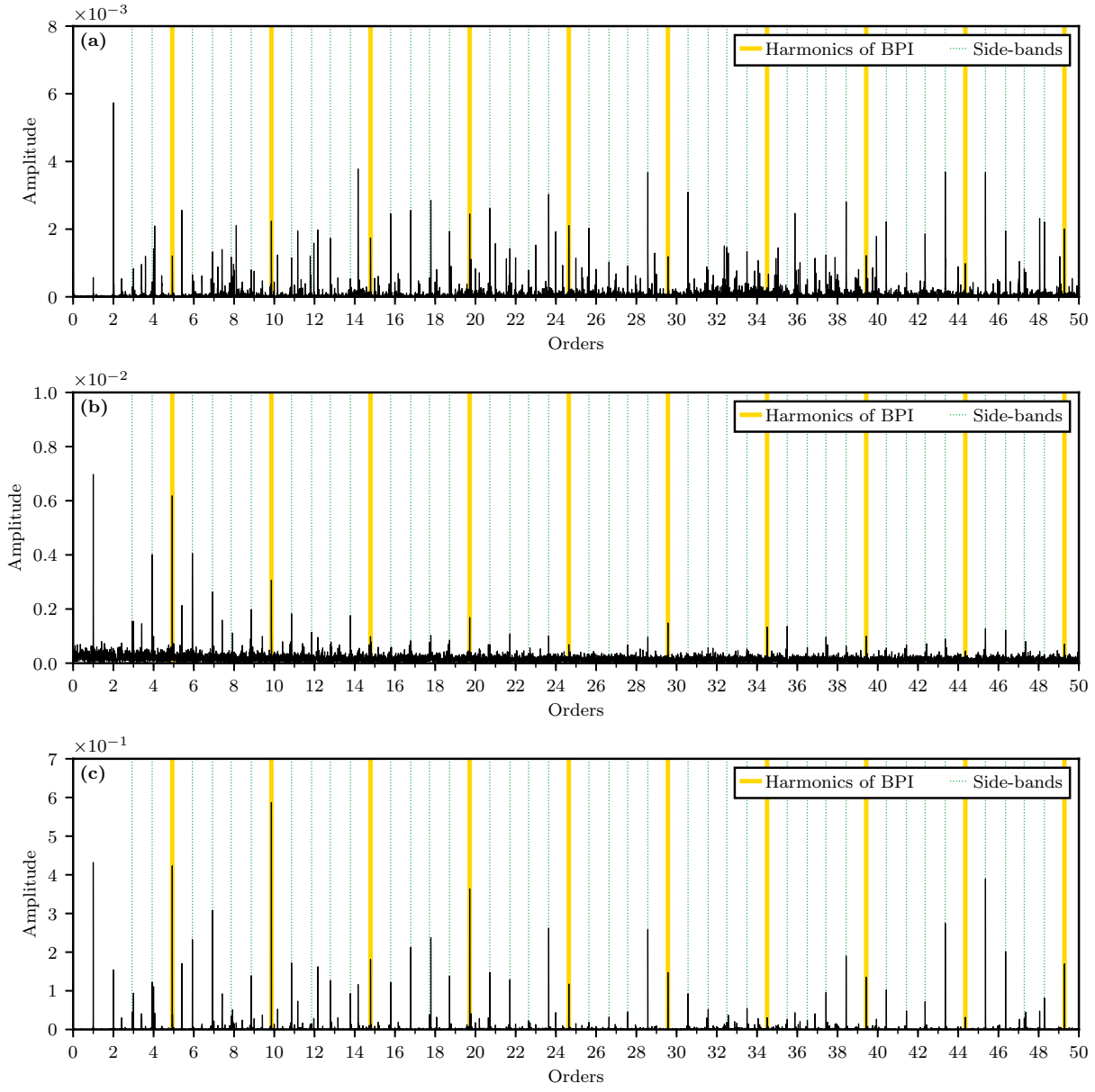

Figure C.17: The WCCS used to diagnose an inner-race fault located at the fan-end of the CWRU dataset. The vibration was measured at the drive-end. Harmonics and side-bands linked to an inner-race fault are marked in all three subfigures. **(a)** The spectrum of the whitened vibration signal V_w. **(b)** The spectrum of the envelope V_{env}. **(c)** The proposed WCCS.

shown in Table B4 in [28]. Performing diagnosis on this dataset is difficult as the bearing vibration must pass through the noisy induction motor. Relevant specifications for this dataset are given in Table C.4. Note that the ball-pass inner race order is very close to 5 shaft orders, which may cause problems during spectrum analysis.

The spectrum of the whitened vibration signal, the spectrum of the envelope, and the WCCS are shown in Figures C.17 (a)-(c), respectively. The vibration spectrum in Figure

C.17 (a) shows 10 harmonics of the O_{BPI}, and multiple side-bands spaced one order away from the harmonic. However, there also multiple other peaks present which makes the analysis more difficult. The envelope spectrum in Figure C.17 (b) doesn't contain as many peaks, but most are related to the bearing fault. However, from the third harmonic, the amplitude of the O_{BPI} harmonics diminish in value. The WCCS in Figure C.17 (c) is capable of combining the harmonics and side-bands of the whitened vibration spectrum and the envelope spectrum. This spectrum has a lower perceived noise-floor, and fewer peaks that are not related to the bearing fault. From these results it is clear that the WCCS is useful to successfully diagnose the fan-end bearing, and that it is easier to analyze than either the whitened vibration spectrum or the envelope spectrum.

C.4 Conclusions

The Whitened Cross-correlation Spectrum (WCCS) is proposed in this paper for diagnosing bearing faults. The method consists of whitening the vibration signal using synchronous average and an autoregressive model. The key feature of the method is the cross-correlation between the selectively whitened vibration signal and its envelope. This correlation results in an element-wise multiplication of the frequency spectra of the two signals. It has been shown in the literature that both the raw deterministic vibration signal and the envelope may contain low-frequency signal components related to the bearing fault. Therefore, this correlation causes frequency components that are visible in both to be amplified, while components that do not share correlation are attenuated. The computational complexity of the time-domain convolution can also be minimized by down-sampling the whitened vibration signal and the envelope prior to computing the cross-correlation. Vibration data from a bearing accelerated life-time test has been used to verify the diagnostic performance of WCCS. The method aids to accurately diagnose the bearing for all three fault types before the bearing is seriously worn out. The early fault was identified at $\approx 8.33\%$ of remaining useful life, which should be long enough to schedule a machine overhaul. Compared to other methods in the literature that are similarly computationally heavy, the WCCS provided more prominent harmonics and side-bands compared to the other methods. Additionally, a difficult dataset from the Case Western Reserve University (CWRU) has been successfully diagnosed using the WCCS.

Funding

This work has been funded by the Ministry of Education and Research in Norway.

REFERENCES

[1] C. Mishra, A. K. Samantaray, and G. Chakraborty. Rolling element bearing defect diagnosis under variable speed operation through angle synchronous averaging of wavelet de-noised estimate. *Mechanical Systems and Signal Processing*, 72:206–222, 2016. doi:10.1016/j.ymssp.2015.10.019.

[2] J. Altmann and J. Mathew. Multiple Band-Pass Autoregressive Demodulation for Rolling-Element Bearing Fault Diagnosis. *Mechanical Systems and Signal Processing*, 15(5):963–977, 2001. doi:10.1006/mssp.2001.1410.

[3] (ISO). 22096 condition monitoring and diagnostics of machines - acoustic emission, 2007.

[4] T. R. Lin, E. Kim, and A. C. C. Tan. A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm. *Mechanical Systems and Signal Processing*, 36(2):256–270, 2013. doi:10.1016/j.ymssp.2012.11.003.

[5] M. Elforjani and D. Mba. Accelerated natural fault diagnosis in slow speed bearings with acoustic emission. *Engineering Fracture Mechanics*, 77(1):112–127, 2010. doi:10.1016/j.engfracmech.2009.09.016.

[6] N. Jamaludin, D. Mba, and R. H. Bannister. Condition monitoring of slow-speed rolling element bearings using stress waves. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 215(4):245–271, 2001. doi:10.1177/095440890121500401.

[7] A. Widodo, E. Y. Kim, J.-D. Son, B.-S. Yang, A. C. C. Tan, D.-S. Gu, B.-K. Choi, and J. Mathew. Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. *Expert Systems with Applications*, 36(3):7252–7261, 2009. doi:10.1016/j.eswa.2008.09.033.

[8] M. Žvokelj, S. Zupan, and I. Prebil. Multivariate and multiscale monitoring of large-size low-speed bearings using Ensemble Empirical Mode Decomposition method com-

bined with Principal Component Analysis. *Mech. Syst. Signal Pr.*, 24(4):1049–1067, 2010. doi:10.1016/j.ymssp.2009.09.002.

[9] N. Jamaludin and D. Mba. Monitoring extremely slow rolling element bearings: Part I. *NDT and E International*, 35(6):359–366, 2002. doi:10.1016/S0963-8695(02)00006-3.

[10] E. Y. Kim, A. C. C. Tan, J. Mathew, and B. S. Yang. Condition monitoring of low speed bearings: A comparative study of the ultrasound technique versus vibration measurements. *Australian Journal of Mechanical Engineering*, 5(2):177–189, 2008. doi:10.1007/978-1-84628-814-2_21.

[11] M. Elforjani and D. Mba. Detecting natural crack initiation and growth in slow speed shafts with the Acoustic Emission technology. *Engineering Failure Analysis*, 16(7):2121–2129, 2009. doi:10.1016/j.engfailanal.2009.02.005.

[12] M. Elforjani and D. Mba. Condition Monitoring of Slow-Speed Shafts and Bearings with Acoustic Emission. *Strain*, 47(SUPPL. 2):350–363, 2011. doi:10.1111/j.1475-1305.2010.00776.x.

[13] N. Tandon and A. Choudhury. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology international*, 32(8):469–480, 1999. doi:10.1016/S0301-679X(99)00077-8.

[14] K. R. Fyfe and E. D. S. Munck. Analysis of computed order tracking. *Mechanical Systems and Signal Processing*, 11(2):187–205, 1997. doi:10.1006/mssp.1996.0056.

[15] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

[16] Z. Feng and M. J. Zuo. Vibration signal models for fault diagnosis of planetary gearboxes. *Journal of Sound and Vibration*, 331(22):4919–4939, 2012. doi:10.1016/j.jsv.2012.05.039.

[17] G. U. Yule. On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London*, 226:267–298, 1927. doi:10.1098/rsta.1927.0007.

[18] G. Walker. On periodicity in series of related terms. *Proceedings of the Royal Society of London*, 131(818):518–532, 1931. doi:10.1175/1520-0493(1931)59<277:OPISOR>2.0.CO;2.

[19] W. Wang and A. K. Wong. Autoregressive Model-Based Gear Fault Diagnosis. *Journal of Vibration and Acoustics*, 124(2):172, 2002. doi:10.1115/1.1456905.

[20] V. C. M. N. Leite, J. G. B. da Silva, G. F. C. Veloso, L. E. B. da Silva, G. Lambert-Torres, E. L. Bonaldi, and L. E. d. L. de Oliveira. Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Transactions on Industrial Electronics*, 62(3):1855–1865, 2015. doi:10.1109/TIE.2014.2345330.

[21] N. Sugiura. Further analysts of the data by akaike's information criterion and the finite corrections: Further analysts of the data by akaike's. *Communications in Statistics-Theory and Methods*, 7(1):13–26, 1978. doi:10.1080/03610927808827599.

[22] P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1):69–82, 1984. doi:10.1016/0022-460X(84)90595-9.

[23] R. B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5):945–962, 2001. doi:10.1006/mssp.2001.1415.

[24] M. R. Shahriar, P. Borghesani, and A. C. C. Tan. Electrical Signature Analysis-Based Detection of External Bearing Faults in Electromechanical Drivetrains. *IEEE Transactions on Industrial Electronics*, 65(7):5941–5950, 2018. doi:10.1109/tie.2017.2782240.

[25] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007. doi:10.1016/j.ymssp.2005.12.002.

[26] J. Antoni, G. Xin, and N. Hamzaoui. Fast computation of the spectral correlation. *Mechanical Systems and Signal Processing*, 92:248–277, 2017. doi:10.1016/j.ymssp.2017.01.011.

[27] Case Western Reserve University Bearing Data Center Website. 2015. URL: http://csegroups.case.edu/bearingdatacenter/home.

[28] W. A. Smith and R. B. Randall. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100–131, 2015. doi:10.1016/j.ymssp.2015.04.021.

Paper D

Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

This paper has been submitted as:

A. Klausen, H. V. Khang and K. G. Robbersmyr. Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions. Under review at *Mechanical Systems and Signal Processing*.

Multi-band Identification for Enhancing Bearing Fault Detection in Variable Speed Conditions

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

Abstract **– Rolling element bearings are crucial components in rotating machinery, and avoiding unexpected breakdowns using fault detection methods is an increased demand in industry today. Variable speed conditions render a challenge for vibration-based fault diagnosis due to the non-stationary impact frequency. Computed order tracking transforms the vibration signal from time domain to the shaft-angle domain, allowing order analysis with the envelope spectrum. To enhance fault detection, the bearing resonance frequency region is isolated in the raw signal prior to order tracking. Identification of this region is not trivial but may be estimated using kurtosis-based methods reported in the literature. However, such methods may fail in the presence of relatively strong non-Gaussian noise. Cepstrum pre-whitening has also been proposed for this diagnosis challenge, however the noise floor may increase significantly from the normalization of the entire spectrum. In this paper, a new approach for identifying multiple resonance regions is proposed. The proposed method highlights all resonance frequencies in the signal by combining computed order tracking and cepstrum pre-whitening in a new way. Simulations and experimental results prove the validity of the method, and comparisons with two existing methods show the increase in effectiveness of the proposed method.**

D.1 Introduction

Rolling-element bearings, or bearings for short, are critical components in rotating machinery. Unexpected bearing failure may cause machine breakdown and unplanned stops, followed by human safety risks and economic loss [1, 2, 3]. To avoid this scenario, critical machine components, such as bearings, should be monitored to detect irregularities early. Condition monitoring systems with sensor data input have been used for decades to diagnose rotating machinery for various faults. Commonly, the vibration measured

with an accelerometer is analyzed to detect bearing faults [4, 5, 6]. Electric motor stator current [7, 8, 9] and acoustic emission [10, 11] are also reported as viable sensor sources for bearing fault diagnosis. However, in this research, vibration signals are used towards drive-train applications, where accelerometers are always in place.

Bearing fault diagnosis methods based on the theory of cyclostationary (CS) signals have received substantial attention in both industry and academia [12].

Bearing vibration signals are weakly categorized as second order CS (CS2) signals due to internal slip and varying radial/axial load ratios for each roller [13].

The envelope spectrum (ES) [14] is a high-frequency demodulation technique that is often used for bearing diagnosis owing to its simplicity for identifying the cyclic frequency of CS2 signals.

Such existing techniques are commonly used in industrial production, in which the machinery mainly operates in a steady state, i.e. the motor shaft rotates at a near-fixed speed. This is not the case for electric drive-trains that often operate under variable speed conditions (VSC) based on driver's command.

Under VSC, bearing fault impacts no longer occur at a fixed time interval as the relation between time and shaft angle is not linear. The time domain vibration signal can be transformed to the shaft-angle domain by applying computed order tracking [15]. This transformation requires knowledge of the shaft position acquired from a tachometer or an encoder. Alternatively, the shaft phase can be extracted from high-energy harmonics in the vibration signal [16], or the electric motor current [17]. The bearing resonance vibration should be isolated using a band-pass filter prior to order tracking to detect bearing faults effectively. Determining the resonance frequency may be achieved by a hammer tap test [18], or from a thin-shell vibration computer program [19]. However, these operations are complex and hard to generalize for multiple different bearings and machines. An alternative is to estimate an optimal band for band-pass filtration using the vibration signal itself. Kurtosis-based methods, such as spectral kurtosis [20], may be used to identify optimal frequency regions. A high kurtosis value implies that the signal is impulsive, such as the quickly damped impact vibration from bearing fault impacts. Consequently, a frequency band with a high kurtosis value should contain bearing vibration. Methods such as the fast Kurtogram [21] and the Protrugram [22] both use the kurtosis to determine optimal frequency bands for filtering. Applications of these methods [23, 24] show their applicability in determining the optimal filter band. However, in the case of strong non-Gaussian noise or single impacts, kurtosis-based methods may fail to identify the optimal frequency band. Cepstrum pre-whitening (CPW) has also been proposed for fault diagnosis during VSC [25].

However, the spectrum normalization of CPW increases the noise floor which makes

it more difficult to detect fault-related peaks in the ES.

To address the mentioned challenges, a new method for identifying multiple resonance frequency regions is proposed in this paper. Instead of relying on kurtosis, the proposed method exploits the effect of resonance frequency spread due to order tracking, and the spectrum normalization effect of CPW [26, 27]. Advantageously, the method is robust against non-Gaussian noise and random impacts, and the noise floor is kept to a minimum compared to CPW as the signal is band-pass filtered at the bearing resonance region(s). After applying order tracking and CPW, the normalized spectrum is transformed back to the time domain, and time variant resonance modes rises from normalized spectrum. Raising this spectrum to a higher power allows separation resonance frequency modes, which can be chosen for bandpass filtration. Simulations and experimental results from three test rigs validate the proposed method. The rest of the paper is organized as follows. An introduction to bearing fault diagnosis using a vibration accelerometer is given in Section D.2. The algorithms used in the proposed method are elaborated in Section D.3. Methods used for comparison purposes are detailed in Section D.4. Experimental results from simulations and vibration datasets from three test rigs are shown in Section D.5. Comparisons between the proposed method versus spectral kurtosis and CPW are given in Sections D.6. Finally, conclusions are drawn in Section D.7.

D.2 Bearing fault diagnosis

A roller element bearing is made of four main components: an inner race fastened to the rotating shaft; an outer race stationary inside a housing; rollers/balls rotating in between the raceways; and a cage which keeps the distance between rollers constant. An exemplary bearing with an outer race fault is shown to the left in Fig. D.1. Fault impacts of passing rollers cause quickly damped vibration pulses at the resonance frequency of the bearing as seen in Fig. D.1. The impact frequency, referred as a characteristic bearing frequency, is determined by the kinematics of the bearing. The expected characteristic frequencies associated with each fault type (inner race, outer race and roller) can be determined with bearing dimensions, such as the number of rollers, the roller diameter, and the pitch diameter [18]. A resonance frequency mode should be isolated using a band-pass filter to remove frequency components, which are not related to the bearing impacts. Usually, multiple frequency bands contain bearing resonance vibration as indicated in [19, 28]. The proposed method allows for identifying several resonance frequency bands that can be used for band-pass filtration. To perform diagnosis, the Hilbert ES is analyzed for prominent peaks at the characteristic frequencies of the bearing.

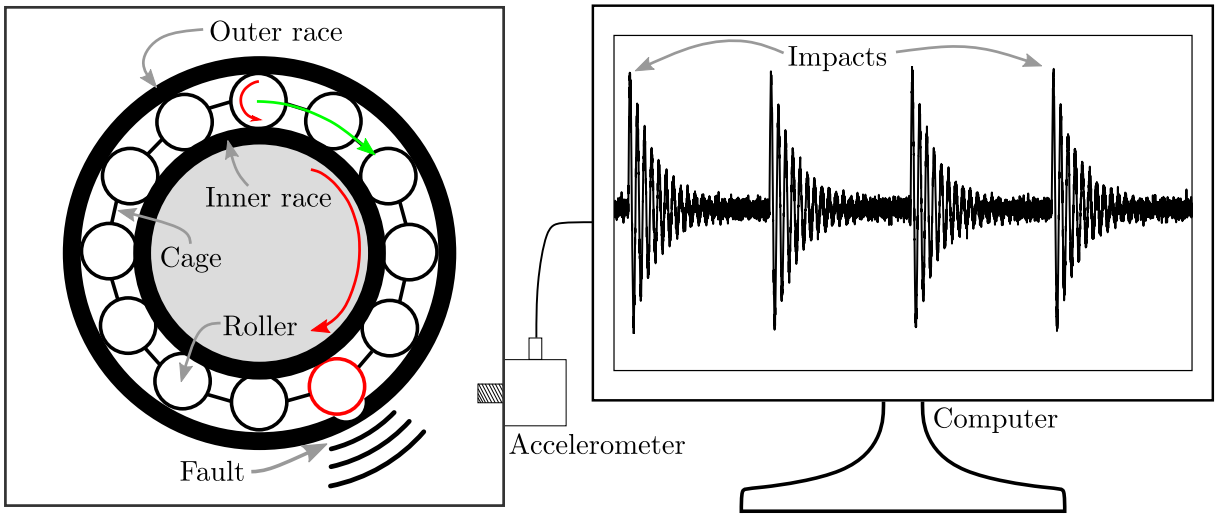

Figure D.1: Simplified bearing condition monitoring setup.

D.3 Proposed method

D.3.1 Complete algorithm

The flow diagram of the proposed method is shown in Fig. D.2. At first, the vibration signal $x(t)$ is acquired using an accelerometer, and the shaft position signal $\theta(t)$ is captured using an encoder.

A suitable frequency band $[f_c, f_{bw}]$ is first identified using the proposed method before computing a band-pass filtered Hilbert envelope order spectrum [14].

D.3.2 Order tracking

The vibration signal is usually measured at a fixed sample rate in the time domain. During constant speed operation, this corresponds to a near-fixed number of vibration samples between bearing fault impacts. Under VSC, however, this correspondence is no longer true. To acquire a fixed number of samples between impacts, the vibration signal is re-sampled from the time domain to the shaft-angle domain (also known as the order domain). The vibration signal is re-sampled using any interpolation-method of choice with

$$f_{ot}(\theta) = \text{Interpolate}\{x = \theta(t), y = x(t)\}\,, \tag{D.1}$$

where $f_{ot}(\theta)$ is a continuous description of the vibration signal in the order domain, where values between the discrete samples are described using the desired interpolation method.

In this research, cubic spline interpolation is chosen to compromise between computational burden and accuracy [15]. Using $f_{ot}(\theta)$, the vibration signal is re-sampled at

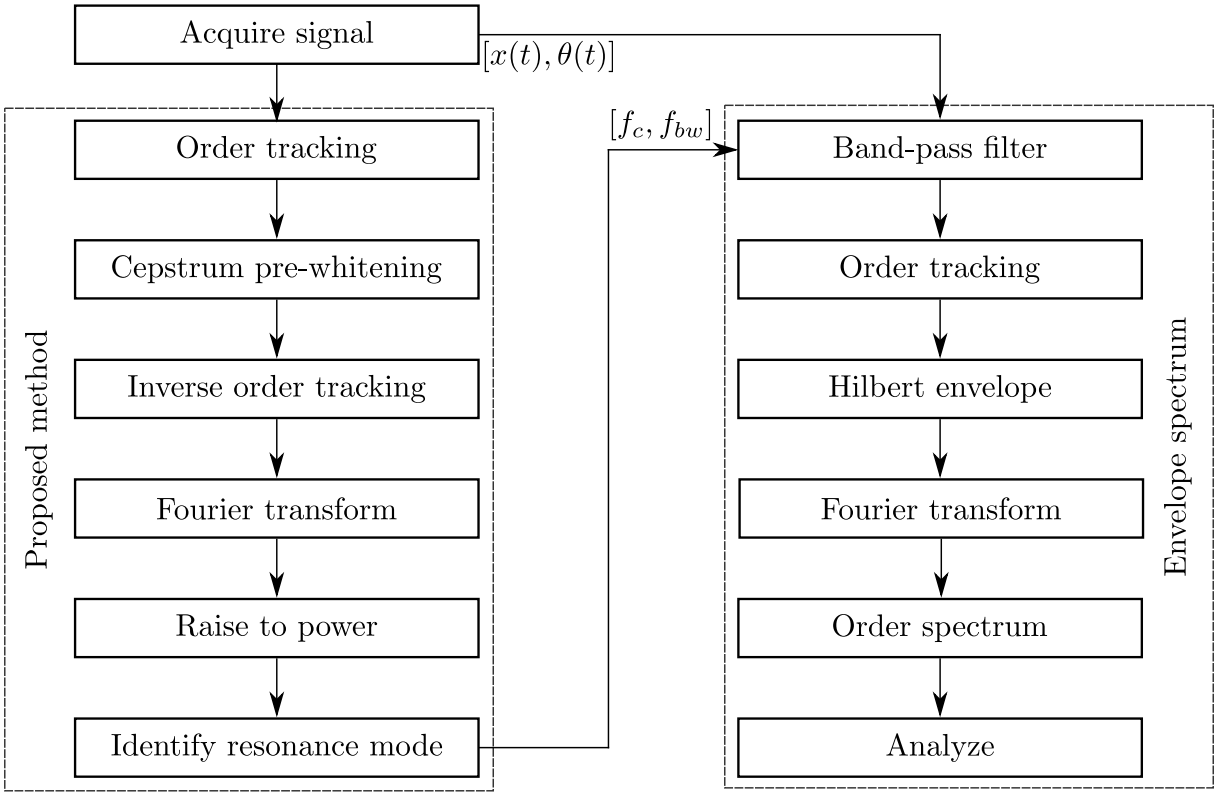

Figure D.2: Flow diagram of the proposed method. $x(t)$ and $\theta(t)$ are the vibration signal and encoder signal, respectively. f_c and f_{bw} are the central frequency and bandwidth of the identified resonance mode.

intervals of $\Delta\theta$ with

$$x_{ot}[j] = f_{ot}(j \cdot \Delta\theta)\,, \tag{D.2}$$

where $j \geq 0$ is an integer, $x_{ot}[j]$ is the j'th sample of the order tracked vibration signal, and $(j \cdot \Delta\theta)$ must be within limits of the measured $\theta(t)$.

It should be noted that the frequency spectrum of an order tracked signal is shown as a function of orders, where the 1st order is equal the shaft speed. In addition to order tracking, it is also possible to perform the inverse, i.e. time tracking, which transforms the signal back to the time domain. Deterministic vibration components from shafts are de-spread in the order spectrum, while the bearing resonance vibration is spread across a larger area as the resonance vibration is a function of time. Numerical examples of order tracking shaft and bearing vibration are shown in Fig. D.3. Figs. D.3 (a) and (b) show the simulated shaft vibration and bearing resonance vibration, respectively. The shaft vibration is the sum of 27 sine waves described in the shaft angle domain. A bearing with an outer race fault is included in the simulation. Each roller impact causes a response given by an underdamped second order system with a resonance frequency of 6000 Hz. Each signal is simulated for 20 seconds with a shaft speed that ranges between 10 and 16 Hz. The frequency spectra of the shaft and bearing vibration are shown in Figs. D.3

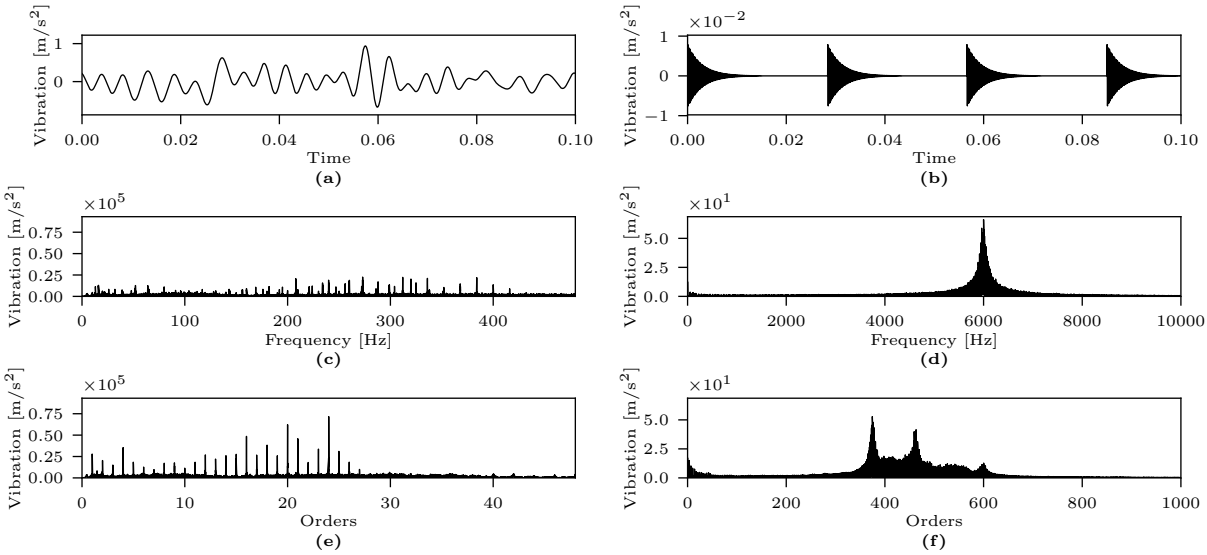

Figure D.3: Simulation showing the resonance frequency spreading effect, and deterministic component de-spreading. (a) shaft vibration; (b) bearing vibration; (c) shaft vibration frequency spectrum; (d) bearing vibration frequency spectrum; (e) shaft vibration order spectrum; (f) bearing vibration order spectrum. One order is equal 10 Hz in this example.

(c) and (d), respectively. As seen in Fig. D.3 (c) shaft vibration components are spread across the low-frequency region, while the bearing resonance is centered at 6000 Hz in Fig. D.3 (d). After applying order tracking, the order spectrum of the shaft and bearing vibration are shown in Figs. D.3 (e) and (f), respectively. In D.3 (e), the shaft vibration components are sharpened, while the bearing vibration order spectrum in Fig. D.3 (f) shows that the resonance vibration energy is spread across a larger frequency area. The proposed method exploits this phenomenon to identify the bearing resonance frequency region.

D.3.3 Cepstrum pre-whitening

As elaborated in [25], deterministic vibration components originating from shafts and gearboxes are periodic, but not sinusoidal, and thus form multiple harmonics in the frequency domain. In the cepstral domain, these harmonics form a single peak at the quefrency equal to the period of the deterministic signal component. However, bearing vibration CS2 components do not form significant peaks in the cepstrum, as they are not exactly periodic. Therefore, it is possible to eliminate deterministic components by performing a series of liftering operations around the quefrencies of deterministic components [26, 27]. By setting the whole real cepstrum to a zero value, except at zero quefrency, all deterministic components are attenuated. This method is referred as cepstrum pre-whitening

(CPW) and is simplified using the following [25]:

$$x_{cpw} = F^{-1}\left\{\frac{F\{V\}}{|F\{V\}|}\right\}, \tag{D.3}$$

where F and F^{-1} are the forward and inverse Fourier transform, respectively.

Inspecting (D.3) reveals that CPW normalizes the frequency spectrum amplitude, i.e. turns the spectrum flat.

D.3.4 Order tracking and cepstrum pre-whitening

As previously elaborated, resonance vibration is spread in the order domain, while deterministic components are de-spread, and CPW normalizes the amplitude of all frequency components. By combining CPW and order tracking, it is possible to highlight resonance frequency modes in the frequency spectrum. This is visualized in Fig. D.4 and explained hereafter: First, the resonance-to-other ratio (ROR) is defined as the ratio between resonance mode amplitude and the most prominent amplitude of other components in the frequency spectrum.

The raw vibration spectrum is mostly dominated by deterministic low-frequency components from shaft and gearbox vibration, therefore the ROR is lower than 1. By implementing the order tracking algorithm, the ROR is even more reduced because of the spread of resonance frequency and de-spread of deterministic components. Using CPW afterwards causes all frequency components to get the same amplitude, and thus the ROR is equal to 1. The inverse order tracking is used to spread deterministic components and de-spread resonance frequency modes, thus achieving a ROR greater than 1. As the ROR is greater than 1, raising the amplitude to a power increases the ROR even further. To avoid singular dominant peaks in the spectrum, the amplitude is filtered using a forward-backward first order IIR low pass filter, and afterwards raised to a power of 5. The final spectrum is analyzed to identify the most prominent resonance frequency mode, and the original signal is band-pass filtered around this mode to enhance bearing fault diagnosis. While prominent resonance modes are manually chosen in this paper, it is possible to automate the mode splitting using an automatic diagnosis system described in [28].

D.4 Comparison methods

D.4.1 Spectral kurtosis

Spectral kurtosis is a well-known method for identifying a suitable frequency band as a high kurtosis value signifies impulsive signal components such as bearing impact vibration. The fast Kurtogram [21] makes a good estimate of the spectral kurtosis at a fraction of

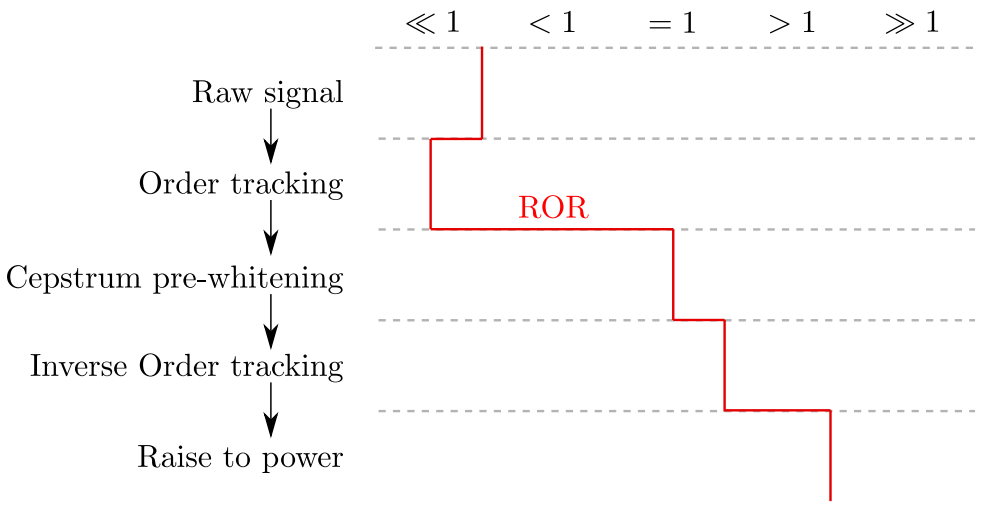

Figure D.4: Visualization of the ratio between resonance mode amplitude versus the most prominent amplitude of other components in the frequency spectrum (ROR) after each step in the proposed method.

the computational cost by calculating the kurtosis at multiple levels of low- and high-pass filtration in an iterative manner. The frequency band with the highest kurtosis should be the most suitable frequency band. The Kurtogram shows kurtosis values at different central frequencies (x-axis) and bandwidth (y-axis), where an increase in level signifies a decrease in bandwidth. Spectral kurtosis may fail if the signal is contaminated by strong non-Gaussian noise or random impacts [22]. The vibration signal is filtered with the band that has the highest kurtosis value, and the envelope order spectrum is calculated.

D.4.2 Cepstrum pre-whitening

The proposed method is also compared to the CPW algorithm in [25]. In that algorithm, the signal is first order tracked, and discrete components are removed by means of CPW. Afterwards, the envelope order spectrum of the resulting signal is computed.

D.4.3 Diagnosis score

When diagnosing a fault with an ES, there are two factors that determine the confidence of the diagnosis: the number of characteristic bearing frequency harmonics, and their prominence compared to the noise floor. With several harmonics it is easier to verify the actual fundamental harmonic frequency, and prominence compared to the noise floor decreases the chance of giving a false diagnosis. To compare the quality of the envelope spectra obtained with the proposed method and the comparison methods, an automatic bearing diagnosis system [28] is applied. This system can identify faults in the bearing using an envelope order spectrum, and a diagnosis score is calculated based on the number

of harmonics and their prominence compared to the noise floor. In a simple form, the score is calculated with

$$S = \sum_{i=1}^{N_h} \frac{Y_{h,i}}{3\mu_{n,i}} i^2 \,, \tag{D.4}$$

where N_h is the number of prominent harmonics identified, $Y_{h,i}$ is the amplitude of the i'th harmonic, and $\mu_{n,i}$ is the mean value of the noise floor around harmonic i. A high number of harmonics is rewarded with a high score due to i^2, and prominence to the noise floor increases the score gained per identified harmonic. Typically, a score lower than 10 can be caused by noise, while a score larger than 100 gives a high confidence of the fault [28]. For a harmonic i to be identified as prominent,

$$Y_{h,i} > 3\mu_{n,i} \text{ and } (Y_{ns,i} > 3\mu_{n,i} \text{ or } Y_{ps,i} > 3\mu_{n,i}) \,, \tag{D.5}$$

where $Y_{ns,i}$ and $Y_{ps,i}$ are the negative and positive side-bands associated with a fault type, respectively. The diagnosis score for each algorithm is directly compared with experimental results to verify the performance of the proposed method.

D.5 Experimental Results

Simulations and experimental datasets from three test rigs are used to validate the effectiveness of the proposed method. The first source [29] features a variable-speed capable test rig installed with bearings that are pre-seeded with three different faults. The second test rig is located at the Case Western Reserve University [30], and the chosen dataset is recorded during VSC with an inner race fault. The final experimental dataset is from an in-house test rig [31] where the bearing has been naturally worn through an accelerated life-time test.

D.5.1 Simulation

A simulation of a bearing fault vibration signal is used to verify the applicability of the proposed method to identify resonance frequency regions. A simulated signal consisting of bearing and shaft vibration, two random impacts, and white Gaussian noise is given by

$$\begin{aligned} x_{sim}(t) = w(t) &+ \sum_{i=1}^{n} A_b(\dot{\theta}) g * \delta(i/O_O - \theta(t)) \\ &+ \sum_{j=1}^{m} A_j(\dot{\theta}) \sin(2\pi j\theta(t) + \phi_j) + I(t), \end{aligned} \tag{D.6}$$

where $w(t)$ is white Gaussian noise, n is the number of outer race fault impulses during the simulation, A_b is the bearing impact vibration amplitude as a function of shaft speed, g is

the impulse response of an underdamped second order system with a resonance frequency of 6000 Hz, δ is the dirac-delta function, O_O is the characteristic fault order for an outer race fault, θ is the shaft position in rounds, $m = 27$ is the number of shaft vibration harmonics, A_j is the amplitude of the j'th harmonic as a function of shaft speed, ϕ_j is a random phase for the j'th harmonic, and $I(t)$ is the vibration from two impacts with high amplitude and a response following an underdamped second order system with a resonance frequency of 3500 Hz.

The bearing and shaft vibration amplitudes are linearly dependent on the instantaneous shaft speed, and the random impacts are added to validate the proposed methods' effectiveness to avoid highlighting frequency regions containing random impacts. Fig. D.5 shows results from using the proposed method on the simulated signal. Figs. D.5 (a) and (b) show a 100 ms long snapshot of the shaft vibration and bearing vibration, respectively, and Fig. D.5 (c) shows the shaft speed during the simulation. The full simulated vibration x_{sim} is shown in Fig. D.5 (d), and its raw frequency spectrum is shown in Fig. D.5 (e). As observed, most of the energy is stored in the low-frequency deterministic shaft components. After order tracking the signal, all shaft deterministic peaks are amplified and sharpened as shown in Fig. D.5 (f). CPW is applied to normalize the spectrum as shown in Fig. D.5 (g). The inverse order tracking causes the resonance frequency region at 6000 Hz to de-spread, and the ROR is therefore increased above unity as shown in Fig. D.5 (h). This allows raising the spectrum to a power to further separate the resonance regions from the rest as shown in Fig. D.5 (i). Here, the amplitude-filtered spectrum is raised to a power of 5 to increase ROR. As observed, there are no high peaks around 3500 Hz, indicating that the proposed method is robust against random impacts. Finally, the most visible area in this spectrum (marked as a red square) is chosen for band-pass filtration of the original vibration signal. A filter bandwidth of 50 orders (500 Hz in this example) is used to retain several harmonics in the final ES. After applying the ES algorithm shown in Fig. D.2, multiple harmonics of the outer race fault order $O_O = 3.54$ are detected using the automatic bearing diagnosis system, as shown in Fig. D.5 (j).

D.5.2 Test rig 1

The first test rig is installed with bearings that are pre-seeded with different faults, and the motor is operated under VSC. The datasets used are provided as supplementary materials in [29], and a simplified schematic of the test setup is shown in Fig. D.6. Both the vibration signal and the shaft angular velocity is available in the repository. The characteristic fault frequencies for the test bearing are: ball-pass inner race order $O_I = 5.41$, ball-pass outer race order $O_O = 3.59$, two times ball (roller) spin order $O_B = 4.75$, and the fundamental cage order $O_C = 0.40$.

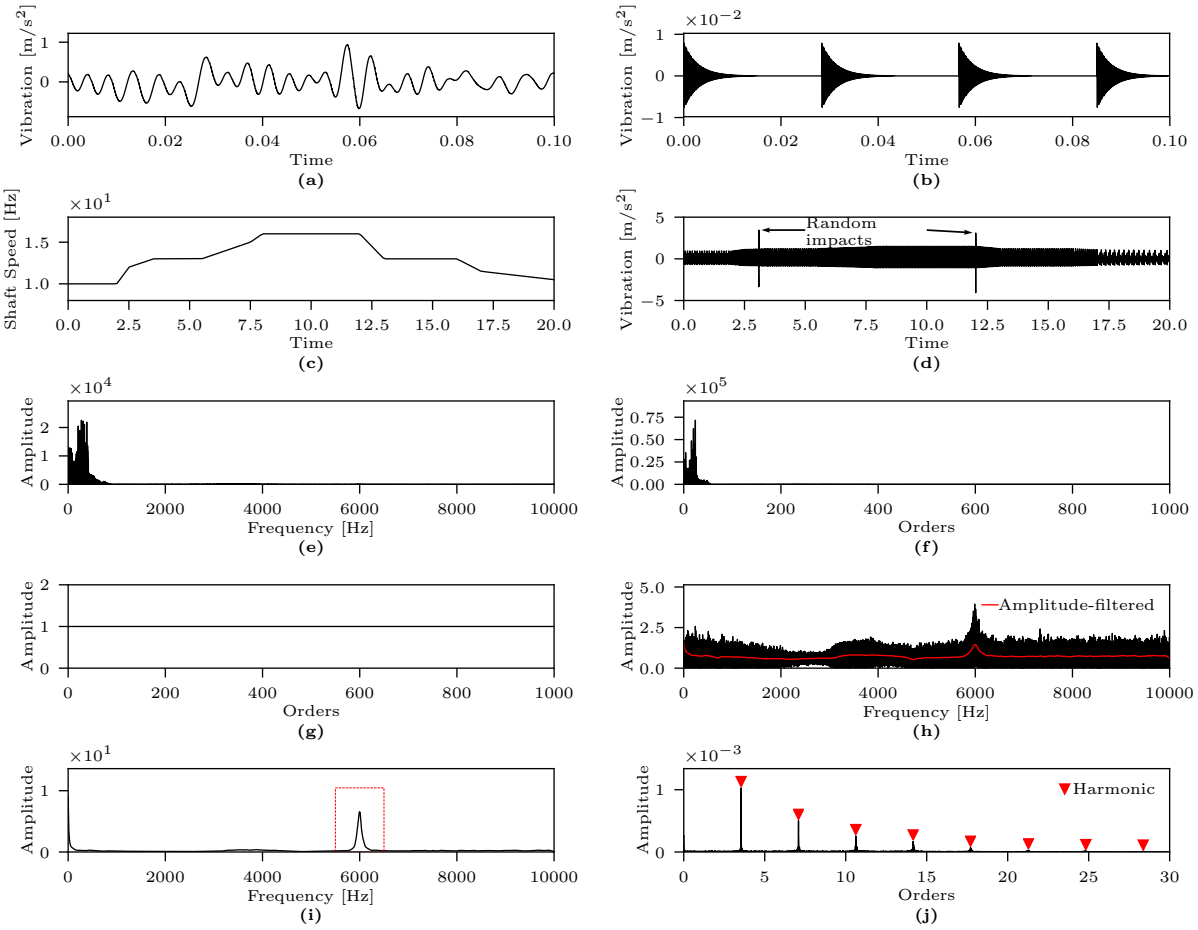

Figure D.5: Resonance band identification using the proposed method on a simulated signal. (a) shaft vibration; (b) bearing vibration; (c) shaft speed; (d) measured vibration signal; (e) frequency spectrum of raw signal; (f) order spectrum; (g) order spectrum after CPW; (h) frequency spectrum after inverse order tracking with a red line showing the amplitude-filtered signal; (i) amplitude-filtered frequency spectrum raised to power of 5, with a suitable band-pass filter region marked in red; (j) envelope order spectrum after band-pass filtration of original signal.

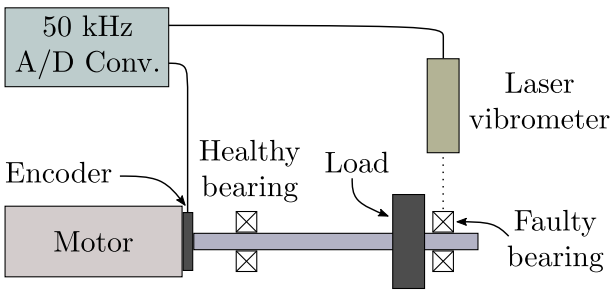

Figure D.6: Simplified schematic of the test setup in

Two times ball order spin is utilized as the roller hits the inner and outer race successively during one spin.

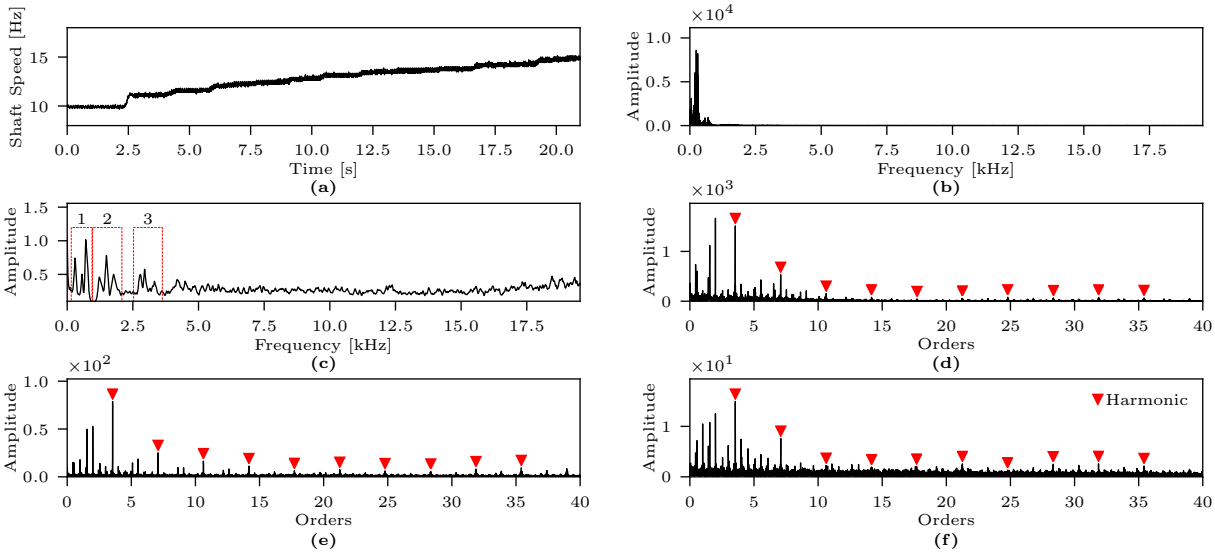

Figure D.7: Results from diagnosing an outer race fault on test rig 1. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where three suitable band-pass filter areas are marked; (d)–(f) envelope order spectra after band-pass filtration at the three marked areas, respectively. Red triangles show identified harmonics related to the fault.

Dataset 1—outer race fault

The first dataset in the repository is recorded while a bearing with an outer race fault is installed on the test rig, and the results are shown in Fig. D.7. Fig. D.7 (a) shows the shaft speed during the measurement and Fig. D.7 (b) shows the frequency spectrum of the raw vibration signal. As shown in Fig. D.7 (b), the low frequency deterministic components dominate the spectrum. After applying the proposed method, the resulting spectrum is shown in Fig. D.7 (c), in which several resonance frequency regions can be chosen due to multiple peaks with high value. Therefore, three marked frequency regions are used for obtaining the ES. Band-pass filtered envelope spectra of the three marked areas are shown in Figs. D.7 (d), (e), and (f), respectively. In these spectra, harmonics of the outer race fault order O_O are shown as red triangles, and multiple prominent harmonics are identified in all three spectra. The diagnosis score for each band is shown in Table D.1, under "DS 1.1". With the high scores on all three bands, the fault diagnosis of an outer race fault is confirmed. These results demonstrate that multiple bearing resonance frequencies may be captured in the vibration signal.

Dataset 2—roller fault

The second dataset from test rig 1 is recorded when the test bearing has a roller fault, and the results are shown in Fig. D.8. The shaft speed varies between 10 and 15 Hz

Table D.1: Diagnosis score per band for each dataset using the proposed method. The best mode for each dataset is written in bold. "DS X.Y" means "Test rig X - dataset Y".

Band \ Dataset	DS 1.1	DS 1.2	DS 1.3	DS 2	DS 3
Band 1	**1615**	**1733**	**448**	41	0
Band 2	1581	287	202	49	0
Band 3	895	328	212	**154**	**166**
Band 4	-	-	-	73	-
Band 5	-	-	-	26	-

as seen in Fig. D.8 (a), and the deterministic low-frequency components dominate the raw vibration spectrum as indicated in Fig. D.8 (b). Applying the proposed method reveals three suitable band-pass filter areas, which are marked in Fig. D.8 (c). With a roller fault, the characteristic vibration patterns are harmonics of two-times roller spin O_B with side-bands given by the fundamental cage O_C. Due to the radial load on bearings, there may be little vibrations when the damaged roller is outside the radial load zone. Therefore, the most visible vibration can be from impacts in the radial load zone, which the damaged roller is passing once per fundamental cage revolution [18]. In case of this dataset, only harmonics of O_C are identified when analyzing the band-pass filtered envelope order spectra in Figs. D.8 (d), (e), and (f). All of the envelope spectra show multiple prominent harmonics at the cage frequency, which strongly verifies the diagnosis. The high diagnosis score given in Table. D.1 under "DS 1.2" also verifies the effectiveness of the proposed method to detect bearing faults.

Dataset 3—inner race fault

The final test is performed with an inner race fault in the bearing, and the results are shown in Fig. D.9. Figs. D.9 (a) and (b) show the shaft speed and the frequency spectrum of the raw vibration signal, respectively. After applying the proposed method, the highlighted resonance modes are shown in Fig. D.9 (c). The three areas marked with red-stapled lines are chosen for band-pass filtered ES analysis due to their prominence. The envelope order spectra for these three areas are shown in Figs. D.9 (d)–(f), respectively. An inner race fault is characterized by harmonics at the inner race frequency O_I with side-bands located 1 order away due to the damaged part rotating in and out of the radial load zone [18]. Therefore, the automatic diagnosis method searches for prominent harmonics and side-bands. With this method, multiple harmonics and side-bands are identified, which can be seen in the envelope spectra. The score for each band is given in Table D.1 under "DS 1.3". Due to low prominence of the higher order harmonics, the

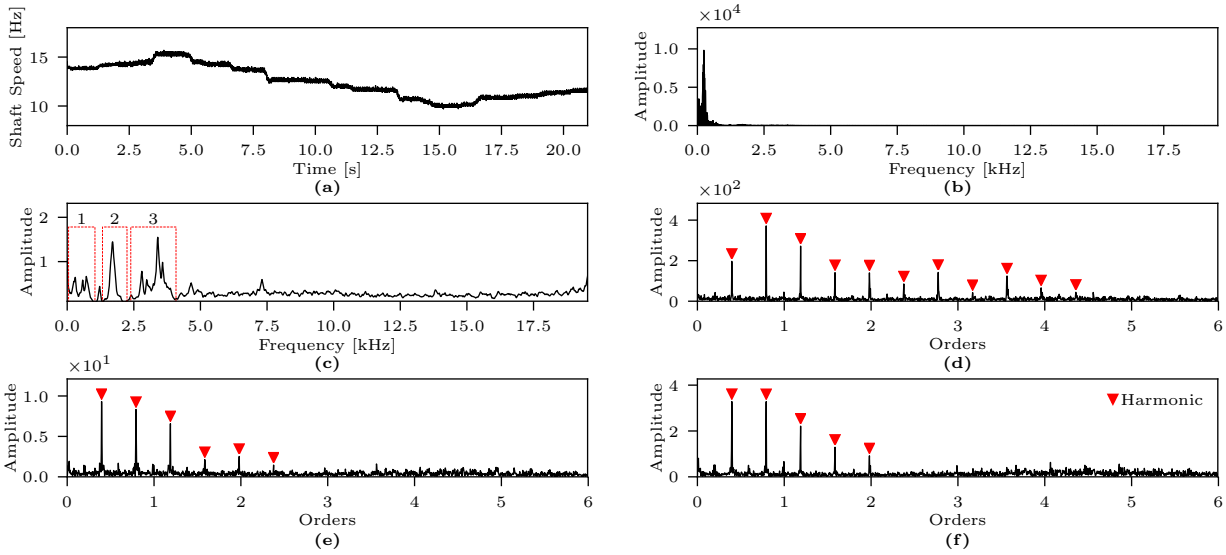

Figure D.8: Results from diagnosing a roller fault on test rig 1. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where three suitable band-pass filter areas are marked; (d)–(f) envelope order spectra after band-pass filtration at the three marked areas, respectively. Red triangles show identified harmonics related to the fault.

maximum score is not as great when compared to the two first datasets. However, with a score greater than 100, the diagnosis confidence is high.

D.5.3 Test rig 2

Data from a second test rig is utilized to further quantify the performance of the proposed method. This test rig is located at the Case Western Reserve University (CWRU), and multiple datasets are available for download [30]. A simplified schematic of the test rig is shown in Fig. D.10. Both fan-end (FE) and drive-end (DE) bearing faults at different severity levels are available in the database, with accelerometer measurements taken at the FE, DE and base (BA).

A study [32] thoroughly analyzed all datasets in the repository and benchmarked multiple fault diagnosis techniques to determine their performance in detection of faults using data from three accelerometers. Dataset 174 proved difficult to diagnose due to speed variation, and consequently the three methods utilized resulted in a mis-diagnosis as reported in Table B3 [32]. A bearing with an inner race fault ($O_I = 5.401$) is installed at the DE when dataset 174 is recorded, and in this research, the DE accelerometer data is used to diagnose the machine.

The time-frequency spectrum of the enveloped vibration signal, shown in Fig. D.11, reveals that the machine is operated under VSC during recording of this dataset. That

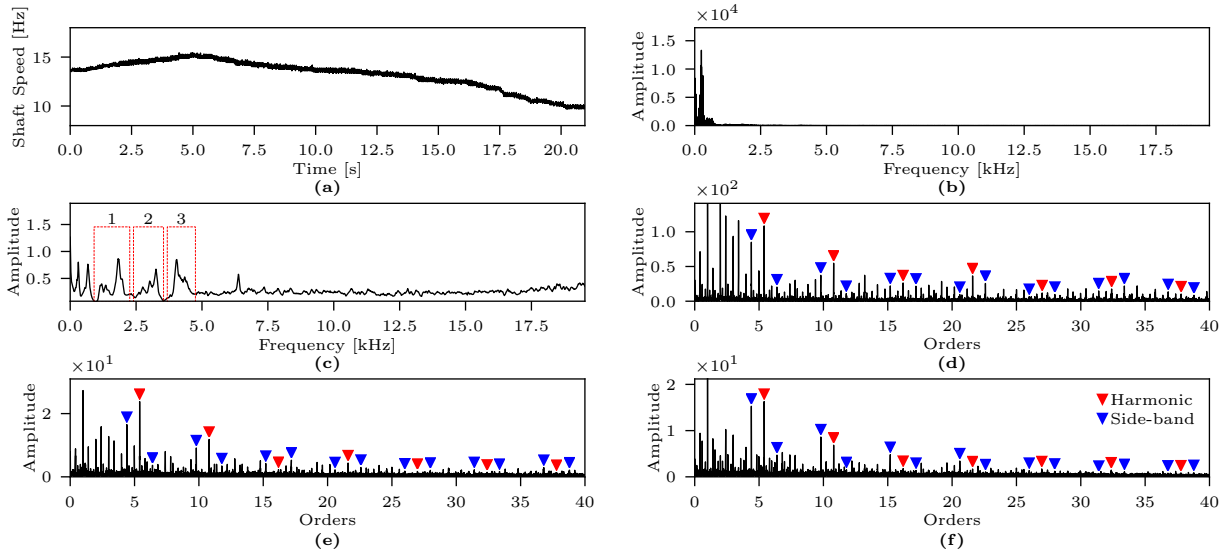

Figure D.9: Results from diagnosing an inner race fault on test rig 1. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where three suitable band-pass filter areas are marked; (d f) envelope order spectra after band-pass filtration at the three marked areas, respectively. Red and blue triangles show identified harmonics and side-bands related to the fault.

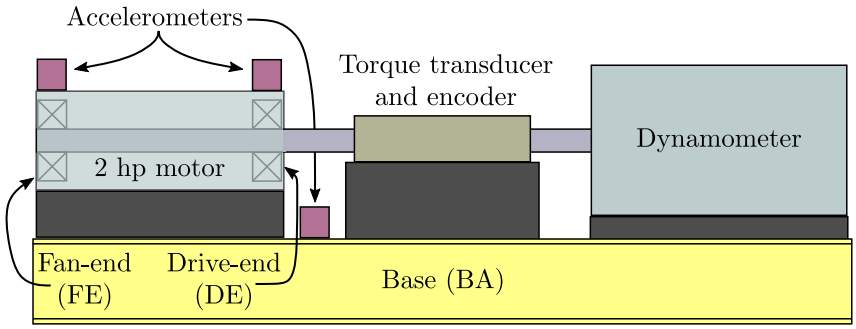

Figure D.10: Simplified schematic of the test rig used at CWRU.

is because visible amplitude ridges show a reduction in frequency over time. The red-stapled line has been identified as the third shaft order harmonic vibration. Integrating this frequency ridge over time and dividing by three gives the relative position of the shaft. Order tracking the raw vibration signal using this position gave high peaks at the assumed shaft order and its second and third harmonics. Additionally, the identified shaft speed is slightly below the supposed fixed-speed of 1797 rpm (29.95 Hz) [30], further indicating that the identified shaft speed is correct.

Results from using the proposed method are shown in Fig. D.12. Fig. D.12 (a) shows the identified shaft speed over time, and the initial speed of about 28.2 Hz is just below the specified shaft speed at 29.95 Hz. Fig. D.12 (b) shows the frequency spectrum of

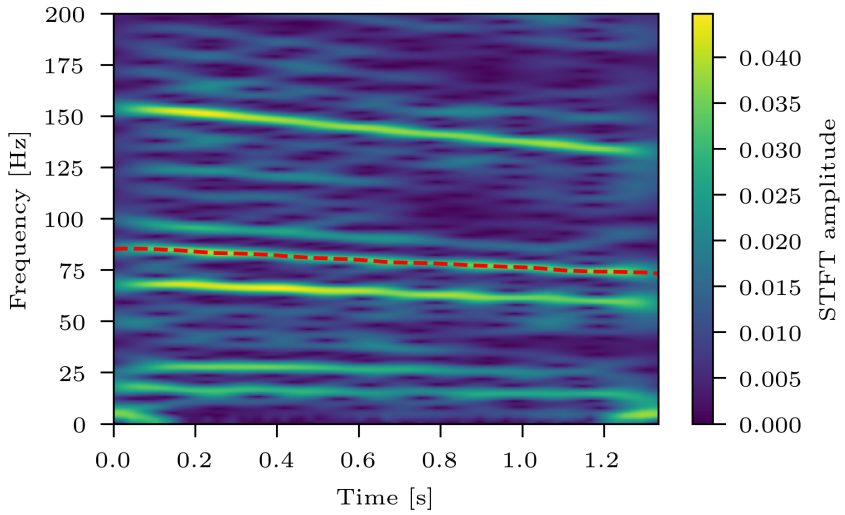

Figure D.11: Time-frequency diagram of the enveloped vibration data. An identified ridge is marked as a red-stapled line. This ridge is identified as the third order of shaft vibration.

the raw vibration signal. After using the proposed method, multiple resonance modes are identified as shown in Fig. D.12 (c). Five bands are chosen for bandpass filtration to properly investigate most of the resonance bands. The band-pass filtered envelope spectra are shown in Fig. D.12 (d)-(h). As there is an inner race on the bearing, side-bands of 1 order should be visible together with the harmonics. However, no such side-bands were identifiable on either of the envelope spectra. This may be the cause of improper order tracking or speed estimation. Therefore, the diagnosis score is based on the harmonics of O_I alone. All the envelope spectra show harmonics of the fault, and one of the bands have a higher score than 100 as indicated in Table D.1 under "DS 2.1". These results show that order tracking is necessary to diagnose the fault in this dataset, and the proposed method aids in identifying suitable band-pass filter specifications.

D.5.4 Test rig 3

Data from an in-house test rig is also used to demonstrate the performance of the proposed method. A 6008 type, 40 mm bore diameter roller element bearing is worn naturally through an accelerated life-time test by applying radial and axial load. Details of the test rig are given in [31], and a schematic is shown in Fig. D.14. The bearing has a dynamic loading capacity of $C = 17.8$ kN and a static load capacity $C = 11$ kN. It was loaded with a radial load of 9 kN and an axial load of 5 kN constantly during the test. The bearing vibration was measured with an accelerometer, the radial shaft movement was measured with an eddy current proximity sensor, and the shaft angular position is acquired using an encoder. The position data from the proximity sensor is used for verification purposes,

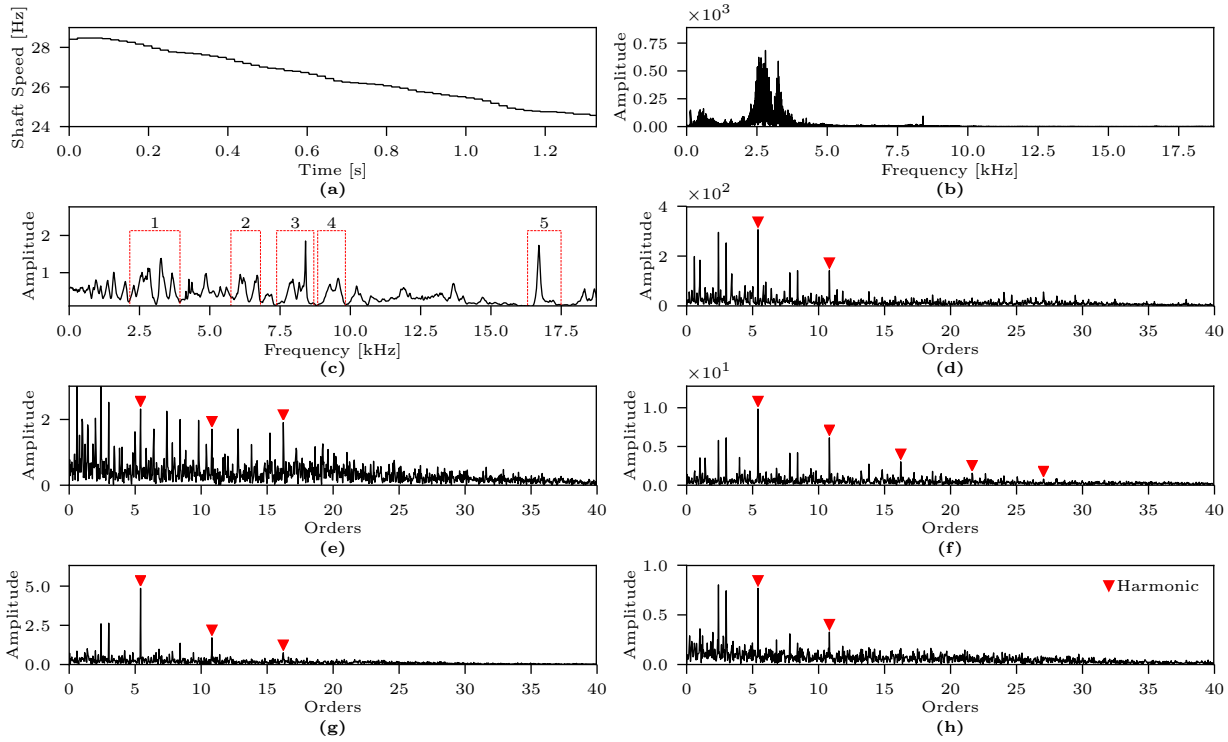

Figure D.12: Results from diagnosing an inner race fault on test rig 2. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where 5 suitable band-pass filter areas are marked; (d)–(h) envelope order spectra after band-pass filtration at each identified mode. Red triangles show identified harmonics elated to the fault.

while the vibration signal is used in this research. During the test, the bearing is run at 500 rpm to accelerate the lifetime. However, every 10 minutes the test rig controller enters a measurement cycle. Within this cycle, vibration and proximity data was measured during various configurations of constant speed and variable speed. After approximately 34.6 million revolutions, an outer-race fault is detected after analyzing data from the eddy current proximity sensor. The characteristic outer race frequency for this bearing is $O_O = 5.12$ orders. In this research, a vibration measurement file captured during variable speed operation right after detecting the outer race fault is used. The shaft speed reference was set to mimic a wave generated using a Pierson–Moskowitz wave spectrum with a mean speed at 50 rpm, significant wave height of 67 rpm, and a significant wave period of 10 s.

Using the proposed method yields the results shown in Fig. D.13. The recorded shaft speed is shown in Fig. D.13 (a), and the raw vibration spectrum is shown in Fig. D.13 (b). The raw spectrum contains peaks mostly at lower frequencies, but there are two single peaks at about 10 kHz and 13 kHz. Using the proposed method yields the spectrum shown in Fig. D.13 (c). In this spectrum, there are three areas of interest for bandpass filtration. The envelope spectra are shown in Figs. D.13 (d)-(f). The spectrum from the

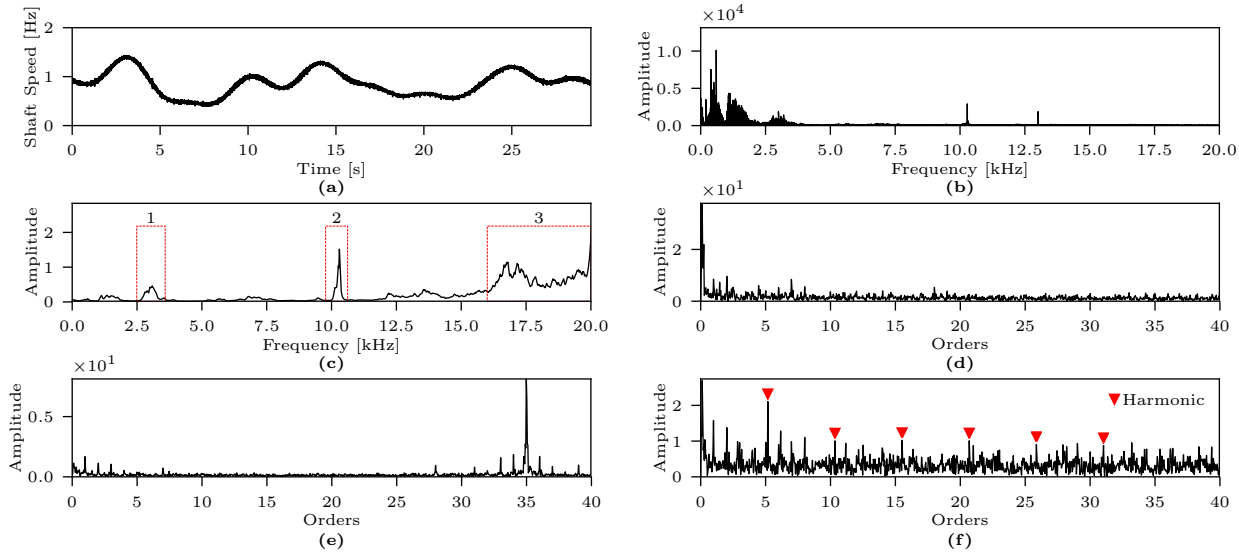

Figure D.13: Results from diagnosing an outer race fault on test rig 3. (a) shaft speed during measurement; (b) frequency spectrum of the raw vibration signal; (c) frequency spectrum after using the proposed method, where 3 suitable band-pass filter areas are marked; (d)–(f) envelope order spectra after band-pass filtration. Red triangles show identified harmonics related to the fault.

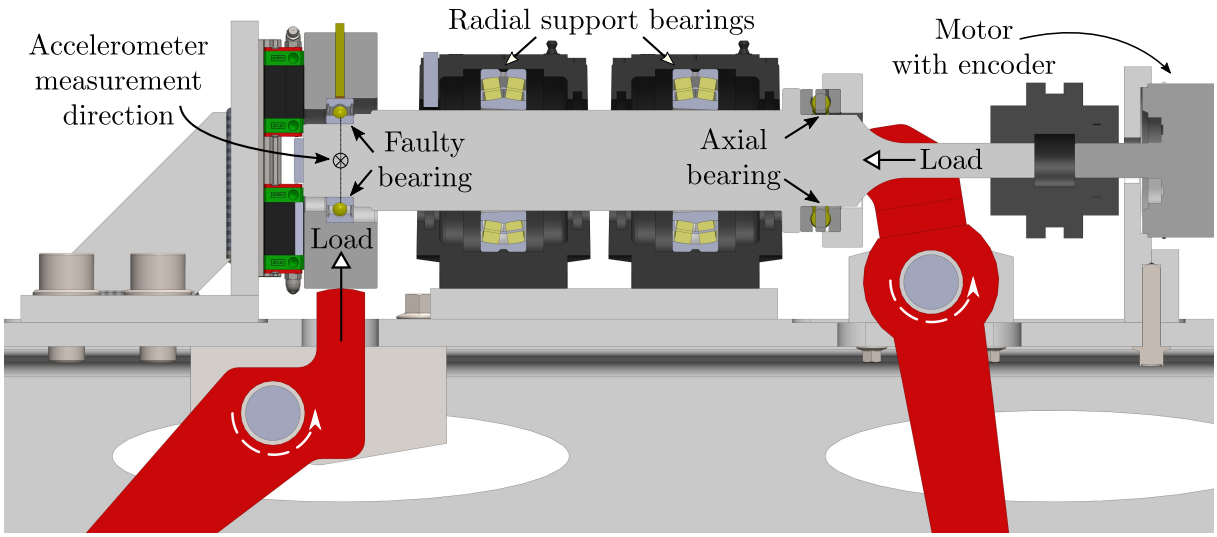

Figure D.14: Simplified schematic of the in-house accelerated life-time test rig

first mode shows no sign of any bearing fault, and this mode may just contain shaft speed invariant noise. There are no signs of bearing fault in the second ES shown in Fig. D.13 (e) either. The only prominent peak is at 35 orders, which most likely sources from the electric motor driving the test rig. The final mode at a high frequency beyond 16 kHz contain some bearing fault related vibrations. The ES in Fig. D.13 (f) shows multiple harmonics of O_O, which verifies the diagnosed fault. Table D.1 shows the diagnosis score, and only the third band has a score greater than 0. However, with a value larger than

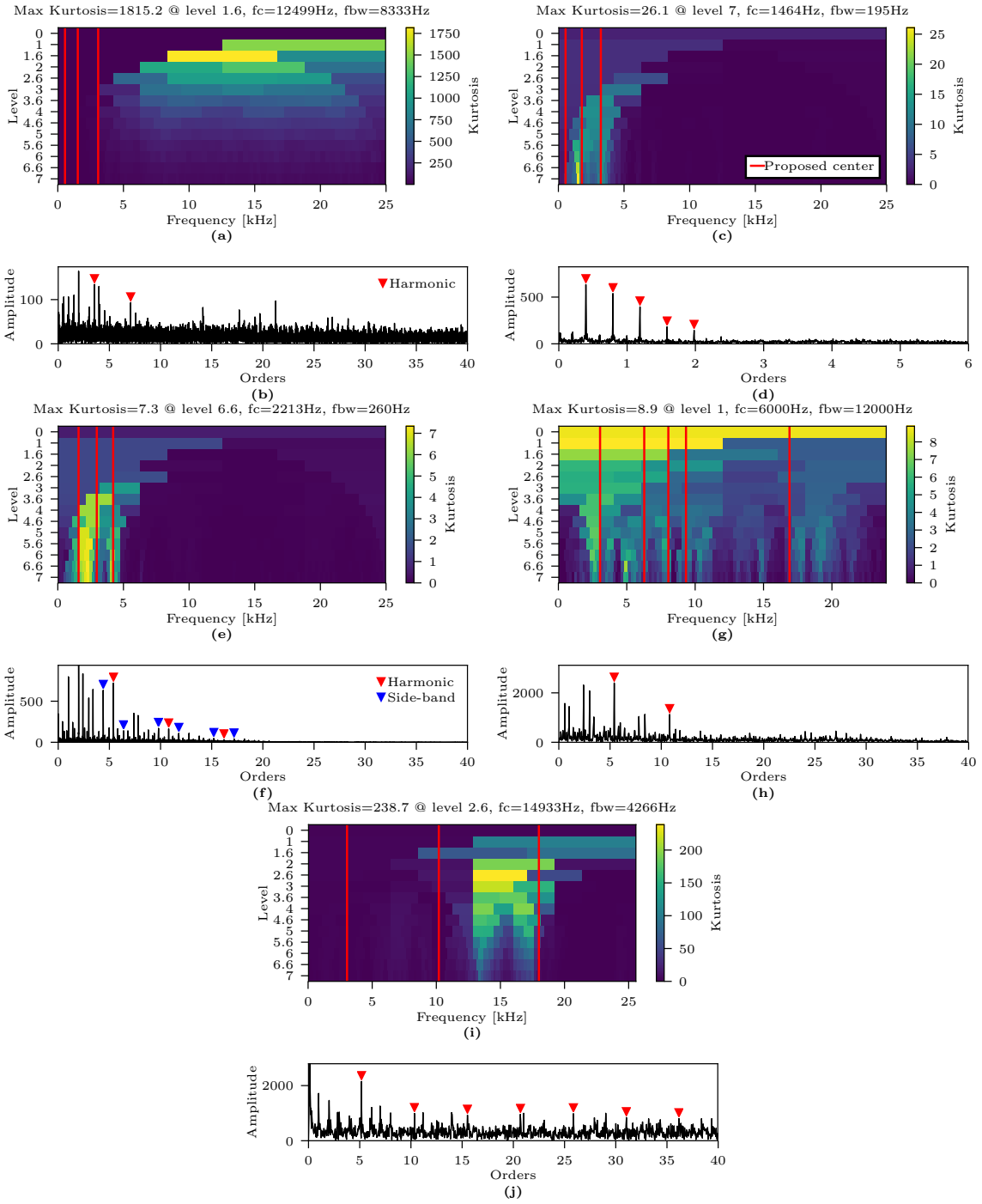

Figure D.15: Results from using the fast Kurtogram detailed in Section D.4.1 on all the experimental datasets. Pairs consisting of a Kurtogram and the resulting ES are given for each dataset. (a)–(b) test rig 1, dataset 1; (c)–(d) test rig 1, dataset 2; (e)–(f) test rig 1, dataset 3; (g)–(h) test rig 2, dataset 1; (i)–(j) test rig 3, dataset 1. In a Kurtogram, red lines indicate central frequencies identified using the proposed method. In envelope spectra, red and blue triangles indicate prominent harmonics and side-bands related to the fault, respectively.

100, there is high confidence with the diagnosis.

D.6 Comparisons

D.6.1 Spectral kurtosis

The fast Kurtogram detailed in Section D.4.1 is tested for all the experimental datasets to check the performance of the proposed method. Results from applying this algorithm are shown in Fig. D.15. Here, a pair consisting of a Kurtogram and an ES is shown for each dataset. Starting with dataset 1 from test rig 1, the Kurtogram is shown in Fig. D.15 (a), while the resulting ES is shown in Fig. D.15 (b). In the Kurtogram, kurtosis values for many combinations of frequency bands are displayed, and the band with the highest kurtosis is deemed the most optimal band for band-pass filtration. Additionally, red lines show central frequencies identified using the proposed method. As seen in Fig. D.15 (a), the highest kurtosis value is found in a high frequency band with a central frequency of 12500 Hz and bandwidth of 8333 Hz. In comparison to the proposed method, this is a much higher frequency band. The corresponding ES is shown in Fig. D.15 (b). Due to the un-optimal filter specifications, there are only two prominent outer race harmonics visible in the spectrum.

For the second dataset with test rig 1, the Kurtogram is shown in Fig. D.15 (c). Here, the highest kurtosis value is identified at the same central frequency as the second mode from the proposed method. The ES in Fig. D.15 (d) hence shows multiple harmonics of the fundamental cage frequency.

Similarly, for dataset 3, the Kurtogram is shown in Fig. D.15 (e). Again, the optimal frequency band with the highest kurtosis value is very close to the ones given by the proposed methods. Therefore, the ES in Fig. D.15 (f) show multiple prominent harmonics and side-bands.

For test rig 2, the Kurtogram is shown in Fig. D.15 (g). Here, a large band centered at 6000 Hz with bandwidth of 12000 Hz is chosen for filtration. In comparison with the red lines, this band encompasses several of the modes given by the proposed method. Therefore, the resulting ES in Fig. D.15 (h) should contain multiple harmonics of the characteristic fault. However, only two prominent harmonics were identified, which is less than the spectra attained with the proposed method.

The final dataset with test rig 3 is also tested, and the Kurtogram is shown in Fig. D.15 (i). An optimal band is identified at a central frequency around 15000 Hz, which is different than the ones given by the proposed method. In this case, the Kurtogram performs better than the proposed method, as the ES in Fig. D.15 (j) indicate multiple

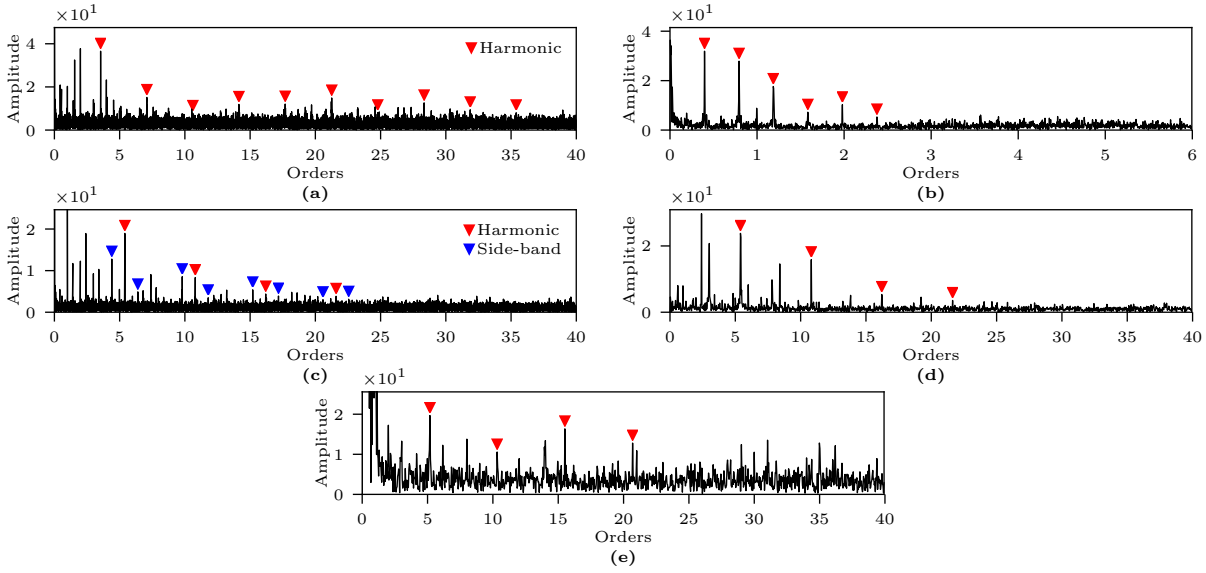

Figure D.16: Results from using the cepstrum pre-whitening method detailed in Section D.4.2 on all the experimental datasets. Envelope spectra with identified harmonics and side-bands are shown for each dataset. (a)–(c) test rig 1, dataset 1 through 3, respectively; (d) test rig 2, dataset 1; (e) test rig 3, dataset 1.

Table D.2: Diagnosis score per dataset for each method. For the proposed method, the best scoring ES is the basis. The best method for each dataset is written in bold.

	Proposed	Kurtogram	CPW
Test rig 1 - outer race	**1615**	26	692
Test rig 1 - roller	**1733**	245	268
Test rig 1 - inner race	**448**	38	49
Test rig 2 - inner race	**154**	41	116
Test rig 3 - outer race	166	**240**	76

harmonics of the outer race fault.

D.6.2 Cepstrum pre-whitening

The cepstrum pre-whitening method elaborated on in Section D.4.2 is also used compare performance with the proposed method. Fig. D.16 shows the resulting envelope spectra from this method. The first ES in Fig. D.16 is from test rig 1, dataset 1 with an outer race fault. With this method, multiple harmonics of the fault are identified, however the harmonics are almost buried in the noise floor. Compared to band-pass filtration, the cepstrum pre-whitening approach is bound to return an increased noise floor, as the white noise covering the entire spectrum is brought to the same amplitude level as other

signals due to spectrum normalization.

The second ES in Fig. D.16 (b) is acquired from the second dataset using test rig 1. Here, multiple harmonics of the cage frequency is detected. Similarly, the ES from dataset 3 in Fig. D.16 (c) also presents multiple prominent harmonics, but not as many when compared to the proposed method.

For test rig 2, the ES in Fig. D.16 (d) shows up to four prominent harmonics, which is about the same as the proposed method. Finally, the ES from test rig 3 is shown in Fig. D.16 (e), and there up to four harmonics are displayed.

D.6.3 Diagnosis score

To quantify the performance of each method, the diagnosis score for all experimental datasets are shown in Table D.2 for the three methods. For the proposed method, the frequency mode resulting in the highest diagnosis score is used. The method with the highest diagnosis score is highlighted with a bold font. As seen, the proposed method scores the highest on the first four experimental dataset, while the Kurtogram method is better for the final dataset. This shows that the proposed method shows an excellent ability to discern resonance frequency vibration from the vibration signal.

D.7 Conclusions

In this paper, a new method for identifying bearing resonance frequency modes in a vibration signal is proposed. The method exploits the resonance vibration distortion caused by order tracking and the spectrum normalization feature of cepstrum pre-whitening (CPW). The algorithm first transforms the signal to the order domain before applying CPW and returning to the time-domain via inverse order tracking. The inverse order tracking causes a de-spread in resonance frequency modes, and since the spectrum is already normalized by means of CPW, resonance mode amplitudes rise above the normalized noise floor. Frequency areas containing prominent peaks in this spectrum are chosen as band-pass filter regions, and the envelope order spectrum of each band-pass filtered signal is analyzed to identify the bearing fault. Simulations and experimental results from three test rigs operating under variable speed conditions validate the proposed method. Comparisons with the fast Kurtogram and CPW with the same datasets show that the proposed method performs better or similarly well on all cases. In the experimental cases, multiple resonance frequency modes were identified using the proposed method, and most of the obtained envelope spectra contained multiple harmonics and side-band related to the fault. The proposed method may also be extended to facilitate automatic resonance mode detection and fault diagnosis.

REFERENCES

[1] W. Zhou, T. G. Habetler, and R. G. Harley. Bearing condition monitoring methods for electric machines: A general review. In *Proc. IEEE Int. SDEMPED*, pages 3–6. IEEE, 2007. doi:10.1109/demped.2007.4393062.

[2] Report of large motor reliability survey of industrial and commercial installations, Part I. *IEEE Transactions on Industry Applications*, 21(4):853–864, 1985. doi:10.1109/tia.1985.349532.

[3] N. Li, Y. Lei, J. Lin, and S. X. Ding. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Transactions on Industrial Electronics*, 62(12):7762–7773, 2015. doi:10.1109/TIE.2015.2455055.

[4] F. Immovilli, A. Bellini, R. Rubini, and C. Tassoni. Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Transactions on Industry Applications*, 46(4):1350–1359, 2010. doi:10.1109/TIA.2010.2049623.

[5] F. Immovilli, M. Cocconcelli, A. Bellini, and R. Rubini. Detection of generalized-roughness bearing fault by spectral-kurtosis energy of vibration or current signals. *IEEE Transactions on Industrial Electronics*, 56(11):4710–4717, 2009. doi:10.1109/TIE.2009.2025288.

[6] B. Yang, R. Liu, and X. Chen. Fault diagnosis for a wind turbine generator bearing via sparse representation and shift-invariant K-SVD. *IEEE Transactions on Industrial Informatics*, 13(3):1321–1331, 2017. doi:10.1109/TII.2017.2662215.

[7] X. Gong and W. Qiao. Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals. *IEEE Transactions on Industrial Electronics*, 60(8):3419–3428, 2013. doi:10.1109/TIE.2013.2238871.

[8] S. Singh and N. Kumar. Detection of Bearing Faults in Mechanical Systems Using Stator Current Monitoring. *IEEE Transactions on Industrial Informatics*, 13(3):1341–1349, 2017. doi:10.1109/TII.2016.2641470.

[9] E. Elbouchikhi, V. Choqueuse, Y. Amirat, M. E. H. Benbouzid, and S. Turri. An Efficient Hilbert–Huang Transform-Based Bearing Faults Detection in Induction Machines. *IEEE Transactions on Energy Conversion*, 32(2):401–413, 2017. doi:10.1109/TEC.2017.2661541.

[10] L. Zhou, F. Duan, D. Mba, and E. Faris. A comparative study of helicopter planetary bearing diagnosis with vibration and acoustic emission data. In *IEEE Int C Prognostics and Health Management (ICPHM)*, pages 246–251. IEEE, 2017. doi:10.1109/ICPHM.2017.7998336.

[11] M. Kang, J. Kim, and J.-M. Kim. An FPGA-based multicore system for real-time bearing fault diagnosis using ultrasampling rate AE signals. *IEEE Transactions on Industrial Electronics*, 62(4):2319–2329, 2015. doi:10.1109/TIE.2014.2361317.

[12] D. Abboud, J. Antoni, S. Sieg-Zieba, and M. Eltabach. Envelope analysis of rotating machine vibrations in variable speed conditions: a comprehensive treatment. *Mechanical Systems and Signal Processing*, 84:200–226, 2017. doi:10.1016/j.ymssp.2016.06.033.

[13] R. B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5):945–962, 2001. doi:10.1006/mssp.2001.1415.

[14] P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1):69–82, 1984. doi:10.1016/0022-460X(84)90595-9.

[15] K. R. Fyfe and E. D. S. Munck. Analysis of computed order tracking. *Mechanical Systems and Signal Processing*, 11(2):187–205, 1997. doi:10.1006/mssp.1996.0056.

[16] J. Shi, C. Shen, X. Jiang, W. Huang, and Z. Zhu. An Auto Instantaneous Frequency Order Extraction Method for Bearing Fault Diagnosis under Time-Varying Speed Operation. In *Int. Conf. on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, pages 623–627, 2017. doi:10.1109/SDPC.2017.122.

[17] J. Wang, Y. Peng, and W. Qiao. Current-aided order tracking of vibration signals for bearing fault diagnosis of direct-drive wind turbines. *IEEE Transactions on Industrial Electronics*, 63(10):6336–6346, 2016. doi:10.1109/TIE.2016.2571258.

[18] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

[19] M. S. Darlow, R. H. Badgley, and G. W. Hogg. Application of High-Frequency Resonance Techniques for Bearing Diagnostics in Helicopter Gearboxes. Technical report, US Army Air Mobility Research and Development Laboratory, 1974.

[20] J. Antoni. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2):282–307, 2006. doi:10.1016/j.ymssp.2004.09.001.

[21] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007. doi:10.1016/j.ymssp.2005.12.002.

[22] T. Barszcz and A. JabŁoński. A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram. *Mechanical Systems and Signal Processing*, 25(1):431–451, 2011. doi:10.1016/j.ymssp.2010.05.018.

[23] R.-B. Sun, Z.-B. Yang, X.-F. Chen, and J.-W. Xiang. Sparse representation based on spectral kurtosis for incipient bearing fault diagnosis. In *Prognostics and System Health Management Conference (PHM-Harbin)*, pages 1–6. IEEE, 2017. doi:10.1109/PHM.2017.8079185.

[24] V. C. M. N. Leite, J. G. B. da Silva, G. F. C. Veloso, L. E. B. da Silva, G. Lambert-Torres, E. L. Bonaldi, and L. E. d. L. de Oliveira. Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Transactions on Industrial Electronics*, 62(3):1855–1865, 2015. doi:10.1109/TIE.2014.2345330.

[25] P. Borghesani, P. Pennacchi, R. B. Randall, N. Sawalhi, and R. Ricci. Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions. *Mechanical Systems and Signal Processing*, 36(2):370–384, 2013. doi:10.1016/j.ymssp.2012.11.001.

[26] N. Sawalhi and R. B. Randall. Signal pre-whitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings. In *COMADEM, Stavanger, Norway*, pages 330–336, May 2011.

[27] R. B. Randall, N. Sawalhi, and M. Coats. A comparison of methods for separation of deterministic and random signals. *International Journal of Condition Monitoring*, 1(1):11–19, 2011. doi:10.1784/204764211798089048.

[28] A. Klausen, K. G. Robbersmyr, and H. R. Karimi. Autonomous Bearing Fault Diagnosis Method based on Envelope Spectrum. *IFAC-PapersOnLine*, 50(1):13378–13383, 2017. doi:10.1016/j.ifacol.2017.08.2262.

[29] C. Mishra, A. K. Samantaray, and G. Chakraborty. Rolling element bearing defect diagnosis under variable speed operation through angle synchronous averaging of wavelet de-noised estimate. *Mechanical Systems and Signal Processing*, 72:206–222, 2016. doi:10.1016/j.ymssp.2015.10.019.

[30] Case Western Reserve University Bearing Data Center Website. 2015. URL: http://csegroups.case.edu/bearingdatacenter/home.

[31] A. Klausen, R. W. Folgerø, K. G. Robbersmyr, and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Modeling, Identification and Control*, 38(3):143–156, 2017. doi:10.4173/mic.2017.3.4.

[32] W. A. Smith and R. B. Randall. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100–131, 2015. doi:10.1016/j.ymssp.2015.04.021.

Paper E

Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

This paper has been published as:

A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings. *IEEE International Conference on Electrical Machines (ICEM)*, Greece, 1912-1918, 2018. doi: 10.1109/ICELMACH.2018.8507056

Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

Abstract **– The prognostics objective is to avoid sudden machinery breakdowns and to estimate the remaining useful life after initial degradation. Typically, physical health indicators are derived from available sensor data, and a mathematical model is tuned to fit them. The time it takes for the model to reach a failure threshold is the estimated remaining useful life. The failure threshold may be determined from historical failure data, but that is not always readily available. ISO standard 10816-3 defines permissible velocity vibration levels for machines that may be used as a failure threshold. However, velocity vibration is not suitable for bearing prognostics due to the effect of integration from acceleration. In this paper, the drawbacks of velocity vibration are explained, and two new failure thresholds using acceleration vibration data are proposed. Results from three run-to-failure tests are provided to show the performance of the proposed failure thresholds.**

E.1 Introduction

Rolling element bearings are used in most rotating machinery. Their purpose is to transfer the shaft load to a stationary housing, and make sure that the shaft rotates smoothly. However, bearing damage is the most common cause of breakdown in rotating machinery, and an unexpected breakdown may result in costly downtime and/or personal harm[1, 2]. It is therefore wise to monitor the machine health using sensor data from i.e. a vibration accelerometer and condition monitoring (CM) techniques. Using signal processing methods, incipient faults in a bearing can be detected before complete breakdown[3]. However, it is not always feasible to perform maintenance as soon as a fault is detected. For remote locations, such as offshore wind farms, weather conditions and pre-planned maintenance trips contribute significantly to feasibility of performing a machine overhaul. It is common that a faulted bearing can be in operation for a certain amount of time after the initial fault. This amount of time is referred to as the remaining useful life (RUL).

For maintenance planning, it is beneficial to estimate the RUL based on the sensor data acquired by the CM system, i.e. make a prediction on how long the machine may operate before an emergency stop is necessary. The general approach is to acquire one or several physical health indicators (PHIs) from available sensor data. A virtual health index (VHI) can be generated from fusion or dimension reduction methods of the chosen PHIs. Most approaches reported in literature use a certain data-driven model[4] to estimate the RUL, where the future trend of a PHI or VHI is estimated using a certain model, and the time/cycles it takes for this model to reach a failure threshold (FT), is the RUL. Some examples of mathematical models are the exponential model, linear model, and the Paris-Erdogan crack-law model. Often, bearing degradation is unstable, and therefore an exponential model may fit best. One common PHI for rotating machinery is the root-mean-square (RMS) of the vibration signal, as it is closely related to the total vibration energy. RMS is also robust against random impacts, as short-lived events contribute little to the mean of squared vibration. References[5, 6, 7] used the RMS for predicting the RUL with different mathematical models, such as the improved exponential model, a stochastic process model, and a Skew-Wiener process. In[8], 14 PHIs are fused into a VHI using the logarithmic Mahalanobis distance (MD), and the RUL is estimated using a Brownian motion (BM) statistical model. While the methods presented in the four papers show excellent performance to predict the RUL, one important issue is not thoroughly discussed: The FT setting. It appears that the FT used to predict the RUL is just set to whatever the PHI/VHI is at the end of the vibration dataset that is utilized, i.e. the FT is reverse-engineered to prove the performance of the model and parameter tuning. In a real CM case where a new machine with no historical failure data is monitored, it is impossible to use the methods described in[5, 6, 7, 8] without setting the FT in advance. One attempt of defining the FT is given in[9], where the FT is said to be set dynamically based on each machine. However, the procedure requires historical failure data sets that must be used for training, which is not always readily available for every machine. Therefore, there is a need for a general procedure to set the FT without relying on historical failure data.

To deal with this unknown FT, this paper proposes methods for defining new FTs for the RMS and the logarithmic MD. These FTs are based on the ISO 10816-3 standard[10] that proposes viable vibration levels for machines operating at speeds faster than 120 rpm. The proposed FTs can be used on any machine given the availability of some vibration data captured during a known healthy state. The rest of the paper is organized as follows. In Section E.2 the ISO 10816-3 standard is explained, and a reason for why the velocity RMS (VRMS) is not optimal to be used directly for RUL estimation is given. Next, Section E.3 provides a method for transforming the VRMS threshold to acceleration

unit. Section E.4 gives an example of this transformation on an experimental dataset. In Section E.5, a method for determining a threshold for the logarithmic MS is presented. Section E.6 provides experimental results for three different datasets, and shows how well the proposed FTs compare to the actual RUL. Finally, Section E.7 concludes the paper.

E.2 ISO standard threshold

ISO standard 10816-3[10] lists acceptable VRMS and displacement RMS (DRMS) levels for rotating machines within the power range between 15 kW and 50 MW. The standard divides vibration levels in four groups: A, B, C and D, where vibration levels in group A belong to new machines, and group D contain vibration levels that cause damage to the machine itself and/or connected components. The RMS value of a signal x is calculated using

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

where T is the period of the signal.

For machinery prognostics purposes, it is of interest to estimate the RUL after the initial damage, and vibration group D in the ISO standard may be used as an FT for the end of serviceable life. However, the vibration level thresholds in the ISO standard are only given for VRMS and DRMS signals for vibration signal components between 10 and 1000 Hz. Common accelerometers typically measure the vibration in acceleration units, and therefore the signal must be integrated to achieve the vibration in a velocity unit. The main issue with using the VRMS FT is the effect of integration itself, as damaged bearings tend to give increased vibration levels at high frequencies, typically higher than 1000 Hz. Integration causes these high-frequency components to be mitigated as it acts like a low-pass filter. Low-frequency vibration components only rise in value once the bearing is critically damaged, and therefore changes in the bearing health may not be detected early if the vibration is monitored in velocity units. An example of this can be seen in Fig. E.1 where the VRMS and ARMS are compared in a real case. The vibration measurements used in this example are collected by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS)[11]. The second run-to-failure test consisting of 982 data files of vibration measurements is utilized. Fig. E.1 (a) shows the VRMS, the initial mean value μ_{VRMS}, and the initial standard deviation (STD) σ_{VRMS}. Initial mean and STD are calculated based on $m_t = 300$ (50 hours) data files in the healthy state as $\mu_{\mathrm{VRMS}} = 0.147$, $\sigma_{\mathrm{VRMS}} = 0.720$, $\mu_{\mathrm{ARMS}} = 0.014$, and $\sigma_{\mathrm{ARMS}} = 0.756$. For the purpose of detecting initial degradation, a statistical alarm is set at a window of 5 times STD, i.e. 5σ. As can be seen, the VRMS value does not surpass this alarm limit before reaching the

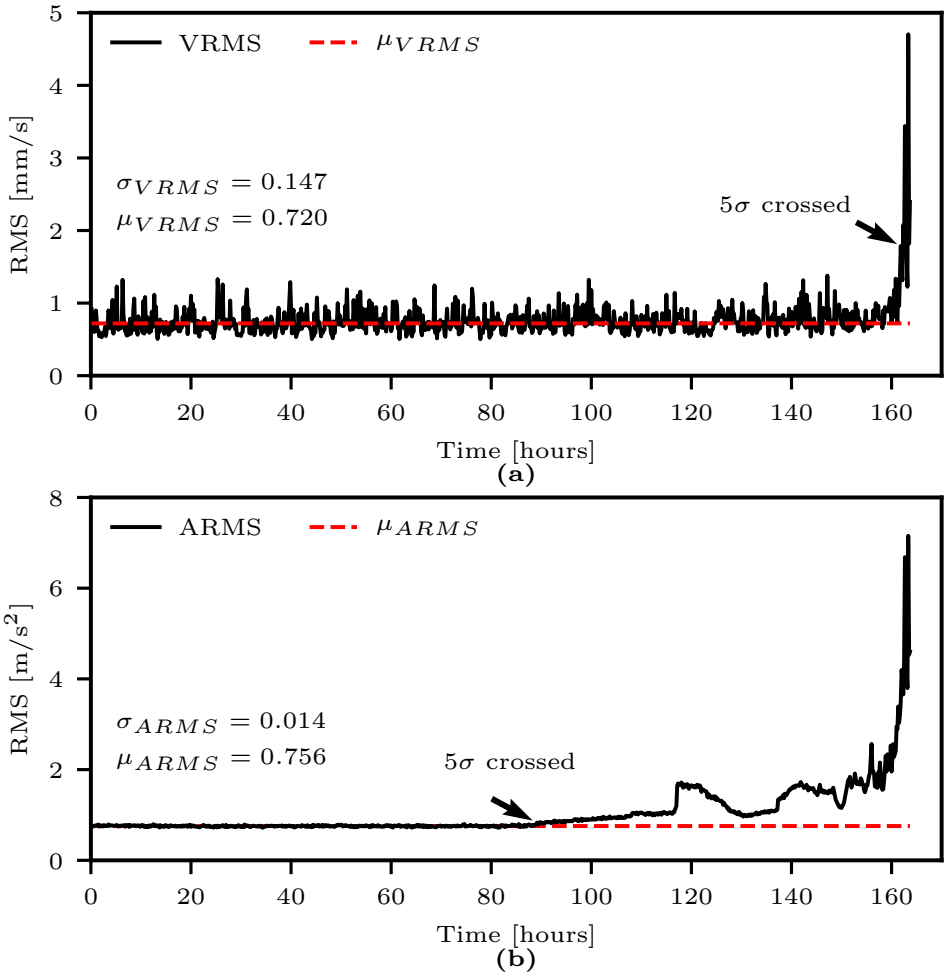

Figure E.1: Difference between RMS values calculated using vibration in velocity and acceleration units. (a) The RMS of the velocity signal. (b) The RMS of the acceleration signal. Both plots show a red-stapled line which is the mean RMS value of the first 300 data files.

last few hours of RUL, and, at this stage, the value increases quickly. From a prognostics point of view, it may be too late to schedule maintenance as there is limited time left before a complete failure. Fig. E.1 (b) shows the ARMS trend. After approximately 90 hours, the ARMS value surpasses the alarm limit, which signifies a trend of degradation in the bearing. The alarm is triggered much sooner when using the ARMS compared to when using the VRMS, and permits a much larger window for scheduling maintenance. Therefore, it is advantageous to transform the VRMS ISO threshold value to acceleration unit to estimate the RUL. The proposed threshold transformation is given in the next subsection.

E.3 Proposed threshold transformation

Using a vibration accelerometer and an analog/digital converter, the signal is acquired in m/s^2 at a sample rate F_s, for a period T, and saved as x_{raw}. The raw vibration signal is first filtered using a 2nd order IIR high-pass filter with a cutoff-frequency at 10 Hz to remove signal drift and comply with the ISO 10816-3 standard[10]. The transfer function of this filter is

$$G_{acc}(s) = \frac{X_{acc}(s)}{X_{raw}(s)} = \frac{s^2}{s^2 + as + b}, \tag{E.2}$$

where X_{raw} and X_{acc} are the frequency transformed raw vibration and filtered vibration, respectively, $s = i\omega$ is the complex variable, ω is the frequency in rad/s, and a and b are the filter constants. The integral of the vibration acceleration is the vibration velocity x_{vel}, and the transfer function is given by

$$G_{vel}(\omega) = \frac{X_{vel}(s)}{X_{raw}(s)} = \frac{1}{s} G_{acc}(s) = \frac{s}{s^2 + as + b}, \tag{E.3}$$

where X_{vel} is the frequency transformed velocity vibration. From ISO 10816-3, the VRMS values are only valid for vibration components up to 1000 Hz, and therefore an extra low-pass filter should be applied at 1000 Hz. However, the authors argue that it is unnecessary to apply a low-pass filter at 1000 Hz because $1/s$ reduces the signal power by approximately 38 dB at 1000 Hz, and therefore an extra low-pass filter should be unnecessary. The RMS is the square root of the mean energy of a signal which can be calculated in the time-domain or in the frequency domain using

$$E\{x\} = \int_0^T |x(t)|^2 dt = \int_{-F_s/2}^{F_s/2} |X(s = i\omega)|^2 d\omega, \tag{E.4}$$

where T is the time signal period, and F_s is the sampling frequency in rad/s. Using the energy definition, the RMS of a signal x is calculated using

$$\mathrm{RMS}\{x\} = \sqrt{\frac{E\{x\}}{T}}. \tag{E.5}$$

To transform the VRMS threshold to an acceleration unit, the RMS ratio (RMS_r) between the ARMS and the VRMS is calculated using

$$\mathrm{RMS}_r = \frac{\mathrm{RMS}\{x_{acc}\}}{\mathrm{RMS}\{x_{vel}\}} = \sqrt{\frac{E\{x_{acc}\}}{E\{x_{vel}\}}}. \tag{E.6}$$

To calculate this ratio analytically, the energies of the two signals are needed. For convenience, the energies are resolved in the frequency domain. One assumption for calculating the energy is that $x_{raw}(t)$ is real-valued only, and therefore the negative frequency components are the complex conjugates of the positive frequency components. The energy

can therefore be calculated by integrating over the positive frequencies and multiply by 2. The energy of the vibration acceleration signal is calculated using

$$E\{x_{acc}\} = 2\int_{\omega_1}^{\omega_2} |X_{acc}(i\omega)|^2 d\omega \tag{E.7}$$

$$= 2\int_{\omega_1}^{\omega_2} \left|\frac{-\omega^2 X_{raw}(i\omega)}{-\omega^2 + ai\omega + b}\right|^2 d\omega \tag{E.8}$$

where the integration bounds ω_1 and ω_2 are given by

$$\omega_1 = 0 \text{ rad/s}, \quad \omega_2 = F_s/2. \tag{E.9}$$

Similarly, the energy of the velocity vibration is calculated using

$$E\{x_{vel}\} = 2\int_{\omega_1}^{\omega_2} |X_{vel}(i\omega)|^2 d\omega \tag{E.10}$$

$$= 2\int_{\omega_1}^{\omega_2} \left|\frac{i\omega X_{raw}(i\omega)}{-\omega^2 + ai\omega + b}\right|^2 d\omega. \tag{E.11}$$

Ideally, the RMS_r would be calculated analytically for all cases using the acceleration and velocity energy given in (E.7) and (E.10), respectively. However, the raw vibration signal, $X_{raw}(s)$, is an unavoidable term in the integration of both energy quantities. X_{raw} is the true vibration signal passing through the dynamic system of the machine itself, the dynamics of the accelerometer, and the characteristics of the analog/digital converter. Therefore, X_{raw} can be significantly different between two machines, even during similar shaft speeds and power outputs. It is not feasible to accurately model X_{raw} for each case, hence RMS_r should be determined experimentally on every machine and after every maintenance overhaul. This is done by calculating the mean VRMS and ARMS values for a machine during a known healthy state, and get the ratio using (E.6). Once the ratio is determined, the acceleration threshold is finally calculated using

$$\text{ARMS}_t = \text{RMS}_r \cdot \text{VRMS}_t, \tag{E.12}$$

where ARMS_t is the ARMS threshold and VRMS_t is the VRMS threshold. In the next section, experimental vibration signal is used to determine the ARMS_t using the methods described in this section.

E.4 Practical test of threshold setting

To test the proposed ARMS_t, the same IMS dataset presented in Section E.2 is used. ISO standard 10816-3 suggests a $\text{VRMS}_t = 4.5$ mm/s for machines larger than 15 kW.

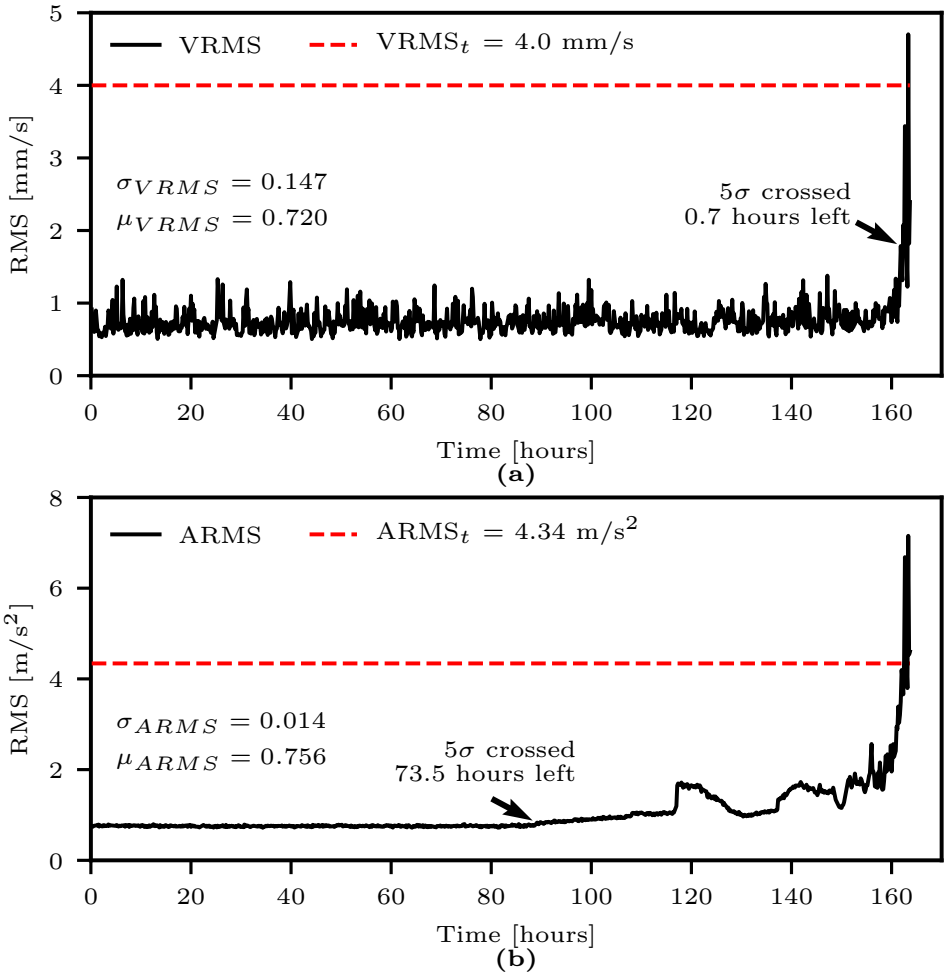

Figure E.2: Comparison of the RMS threshold using velocity and acceleration units for dataset 1. (a) The RMS of the velocity signal. (b) The RMS of the acceleration signal.

Since the machine used in[11] is probably not powerful, the VRMS$_t$ is lowered to 4.0 mm/s. Using the mean and STD values presented in Section E.2, RMS$_r$ is determined as 1085.22 1/s using (E.6). The new threshold is computed by (E.12) as ARMS$_t$ = 1085.22 1/s $\cdot$ 4 mm/s = 4.34 m/s^2. Fig. E.2 shows a comparison of the thresholds where (a) shows the VRMS, and (b) shows the ARMS. As seen on the two graphs, the VRMS and the ARMS reaches their respective thresholds at approximately the same time (163 hours). This result show that the proposed transformation can be used to get a proper threshold for the ARMS with a physical meaning. The alarm-triggered point (5σ) shows how long time is left before the threshold is reached. As can be seen, the maintenance window is much larger using the ARMS when compared to the VRMS: 73.5 hours compared to 0.7 hours.

The presented ARMS threshold may be used directly for RUL estimation using algorithms in[5, 6, 7]. However, some prognostic algorithms combine multiple PHIs to create a VHI for the system. The advantage is that each PHI responds differently to different

degradation trends. Therefore, by fusing multiple PHIs, the initial degradation trend may be easier to detect. In the next section, multiple PHIs are fused using the Mahalanobis distance (MD), and a proposed threshold equation for the MD is outlined.

E.5 Threshold setting for multiple physical health indicators

Some prognostics models use multiple PHIs including the ARMS to determine the RUL. Such a prognostics model using multiple PHIs is proposed in[8], where 14 time-domain PHIs are calculated from the vibration signal. The PHIs are fused using the MD to create a single VHI that represent the data file. According to[8], the MD is linearized by calculating the logarithm of the MD. For brevity, a health index (HI) is defined here as the logarithmic MD. Parameters for a Brownian motion (BM) model are estimated in an on-line fashion using a Kalman filter (KF) to best fit the HI. Since the HI is assumed to be linearly increasing with time once the initial degradation of the bearing has started, this is a valid model to use. The RUL is calculated by using the estimated BM model and predict how long time it would take for the model to reach a certain failure threshold (FT) value. However, in[8], the FT used for estimating the RUL is not discussed. It is impossible to calculate the RUL in a real situation if the FT is not known. Therefore, the rest of this section is dedicated to a method for making a suitable threshold for the HI proposed in[8].

First, the MD is calculated using

$$\mathrm{MD} = \sqrt{(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{S}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \tag{E.13}$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]$ is a vector of the n chosen PHIs, $\boldsymbol{S}$ is the covariance matrix of a set of observations, and $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_n]$ is the mean value of these observations. To determine $\boldsymbol{S}$ and $\boldsymbol{\mu}$, a set of m_t observations during healthy machine condition are used. Finally, the HI is determined using

$$\mathrm{HI} = \log\{\mathrm{MD}\}. \tag{E.14}$$

The FT for the HI is determined using the ARMS and a single-dimension MD (MD_1)

$$\mathrm{MD}_1 = \frac{x-\mu}{\sigma}, \tag{E.15}$$

where σ is the standard deviation of the first m_t observations of x, and μ is the mean value of these observations. Each PHI result in different MD_1 trends, however the assumption for the proposed threshold is that the average MD_1 characteristics for all vibration based

PHIs are similar to the MD_1 of the ARMS. Specifically, that the average maximum value for all MD_1's is the same as the maximum value of the MD_1 of the ARMS. Following this assumption, it should be possible to upscale the $ARMS_t$ to account for multiple PHIs in the MD. To this end, the MD can be estimated by using the ARMS alone, as in

$$\text{MD} \approx \sqrt{\sum_{i=1}^{n} \frac{(\text{ARMS} - \mu_{\text{ARMS}})^2}{\sigma^2_{ARMS}}}, \tag{E.16}$$

$$\approx \sqrt{n} \cdot \frac{\text{ARMS} - \mu_{\text{ARMS}}}{\sigma_{\text{ARMS}}}, \tag{E.17}$$

where $\boldsymbol{S}$ is replaced with σ^2_{ARMS} because, as a covariance matrix, it is diagonal for PHIs that are equal. Given this estimated MD, it is possible to create a threshold for the MD and the HI using the $ARMS_t$ as

$$\text{MD}_t = \sqrt{n} \cdot \frac{\text{ARMS}_t - \mu_{\text{ARMS}}}{\sigma_{\text{ARMS}}} + \mu_{\text{MD}}, \text{ and} \tag{E.18}$$

$$\text{HI}_t = \log\{\text{MD}_t\} + \mu_{\text{HI}}, \tag{E.19}$$

where MD_t and HI_t are the proposed MD and HI thresholds, respectively, and μ_{MD} and μ_{HI} are the mean value of the MD and HI for the first m_t samples, respectively. The mean value is added to the threshold to account for the initial value of the MD and HI. Note that (E.13) and (E.14) are used for computing the MD and HI, respectively, while (E.18) and (E.19) are only used to determine their respective thresholds. In the next section, the proposed HI_t is calculated for multiple test cases.

E.6 Practical test of MD threshold

E.6.1 IMS dataset

The proposed HI_t threshold is here calculated for the IMS bearing dataset introduced in Section E.4. From Section E.4, the $\text{ARMS}_t = 4.34$ m/s^2, $\mu_{\text{ARMS}} = 0.756$ m/s^2, and $\sigma_{\text{ARMS}} = 0.0141$ m/s^2. A total of $n = 14$ PHIs from[8] are utilized to make the HI using (E.14) where the covariance and mean values are calculated from the $m_t = 300$ (50 hours) first data files. The mean value of HI is $\mu_{\text{HI}} = 0.905$ for this dataset. Given these values, the threshold is calculated using (E.19) as

$$\text{HI}_t = \log\{\sqrt{14} \cdot \frac{4.34 - 0.756}{0.0141}\} + 0.905 = 7.7655.$$

Table E.1 shows these calculated values under column "Dataset 1". The HI and the HI_t are both shown in Fig. E.3 for the IMS dataset, and as shown, the HI_t is reached by the HI close to the final data file. This finding is consistent with the ARMS and ARMS_t shown

Table E.1: Threshold value calculations.

Value ↓ \ Dataset →	1	2	3
VRMS_t [m/s]	4.0	3.0	4.0
m_t [files]	300	300	125
RMS_r [1/s]	1085	2803	368
ARMS_t [m/s^2]	4.34	8.41	1.47
σ_{ARMS} [m/s^2]	0.0141	0.0235	0.0049
μ_{ARMS} [m/s^2]	0.756	0.378	0.447
μ_{HI} [-]	0.905	1.053	2.404
HI_t [-]	7.7655	8.208	9.068

in Fig. E.2 (b) in that the computed threshold is reached near the end. The possible advantage of the HI compared to the ARMS in terms of prognostics, is that the general trend of the HI appears to increase linearly over time from first degradation point, while the ARMS exhibits a more exponentially increasing trend near the end. The linear HI trend has an advantage of allowing simpler models for estimating the RUL, such as the BM used in[8]. The choice of RUL estimation model depends on which trend should be used, as the ARMS and HI shows different characteristics.

E.6.2 PHM Challenge dataset

The second dataset used for testing the ARMS_t and HI_t is a part of the Pronostia database[12]. This database was part of the IEEE PHM 2012 Prognostics Challenge that encouraged scientists and engineers to make accurate RUL estimations based on vibration and temperature data. The database consists of 17 accelerated life-time tests conducted during three different operating conditions. The ARMS_t and HI_t are determined using the vibration data in dataset "Bearing1_3", i.e. the third accelerated life-time test during operating condition 1. A total of 2375 vibration data files was collected during this test, and the time interval between two files is 10 seconds. The vibration data in each file was collected for a duration of 0.1 s at a rate of $F_s = 25.6$ kHz.

The $\text{VRMS}_t = 3m/s$ due to the small 250 W motor used on the Pronostia test rig. The first $m_t = 300$ data files (50 minutes) are assumed to be from a healthy machine state. Important values and calculated thresholds are given in Table E.1 under column "Dataset 2". The VRMS and ARMS are shown in Figs. E.4 (a) and (b), respectively, together with their respective thresholds. The VRMS reaches and surpasses the proposed threshold of 3 mm/s at the end of the accelerated life-time test, indicating that the bearing is severely damaged. The ARMS in Fig. E.4 (b) shows very similar characteristics as the VRMS,

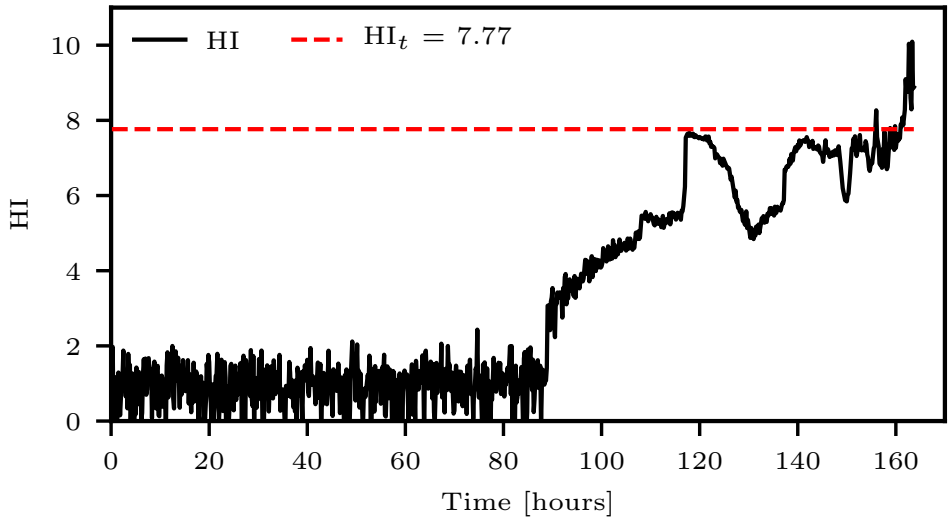

Figure E.3: The HI plotted with the HI$_t$ threshold for dataset 1.

and the alarm at 5σ is triggered almost at the same time for both signals. The main difference is that the ARMS value is less noisy than the VRMS, which should make it easier to predict RUL using the ARMS. In addition, the threshold ARMS$_t$ is reached at the end of the test. The HI is also computed for this dataset using the methods elaborated in Section E.5 and its threshold is given in Table E.1 under column "2". The HI and the HI$_t$ are both shown in Fig. E.5 for this test. As seen, there is a clear advantage of calculating the logarithm of the MD, as the HI is increasing almost linearly after the initial degradation in this test. The HI$_t$ threshold is also very close to the actual final value of the HI at the end of the test, which shows the performance of the proposed threshold calculations. A filtered HI (HI$_f$) is overlaid in Fig. E.5 to make the degradation trend easier to identify. The HI and HI$_t$ can be used in conjunction with the BM algorithm in[8] to estimate the bearing RUL.

E.6.3 Low-speed test rig

The final database used originate from an in-house low-speed bearing test rig. A detailed description of the test rig and the procedures for acquiring vibration sensor data are given in[13]. The dataset is acquired during an accelerated life-time test of a 6008 type bearing, and the vibration signal is measured while the shaft speed is as low as 100 rpm. Each vibration data file was recorded every 18 minutes at a sample rate of 51.2 kHz for a duration of 10 seconds. The test in total lasted 14 full days, i.e. $14 \cdot 24 = 336$ hours, before a torque limit alarm was triggered and forced a stop of the test. The last 507 data files (152 hours) of the accelerated life-time test is used to make the proposed thresholds. In ISO 10816-3, the VRMS threshold is set for components between 10 and 1000 Hz for

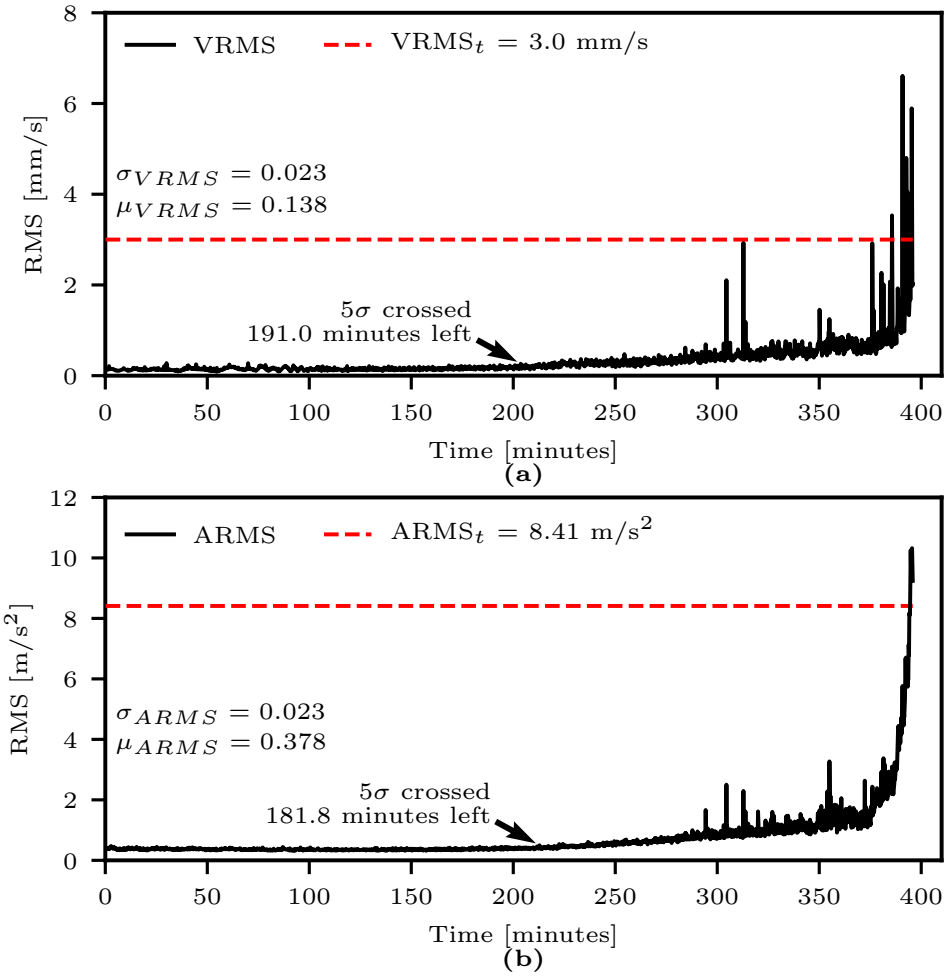

Figure E.4: Comparison of the RMS threshold using velocity and acceleration units for dataset 2. (a) The RMS of the velocity signal. (b) The RMS of the acceleration signal.

machines operating at speeds higher than 600 rpm. For slower machines, between 120 and 600 rpm, the frequency band is set from 2 to 1000 Hz. As the shaft speed in this test is 100 rpm, the 2nd order IIR high-pass filter in (E.2) is configured to a cutoff frequency at $80\% \cdot 100/60 = 1.33$ Hz to preserve the vibration energy at the shaft frequency and higher. The motor power is 1.1 kW, and the VRMS$_t$ is therefore set to 4 mm/s as a conservative value. The $m_t = 125$ (40 hours) first data files are considered to be recorded during a healthy state, and are therefore used for calculating the thresholds. Table E.1 column "Dataset 3" shows calculated values and thresholds for this dataset. The VRMS, ARMS, and their respective thresholds are shown in Fig. E.6. The VRMS exceeds its threshold near the end of the test, and the 5σ alarm is triggered only 2.1 hours before the threshold is reached. In a real application, this will translate to a very short time for RUL estimation and maintenance planning. The ARMS also exceed its threshold at the end, and the initial 5σ alarm is triggered much sooner, at 31.5 hours left before reaching the threshold. It should be noted, however, that the ARMS shows an extremely

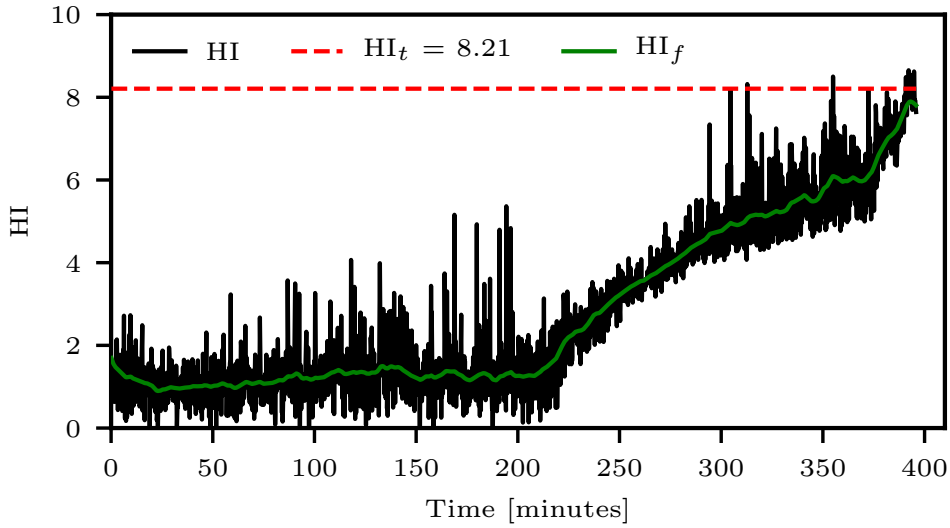

Figure E.5: The HI plotted with the HI_t threshold for dataset 2.

sudden change in degradation ratio near the end, and therefore it might be difficult to fit a degradation model to this signal.

Using the 14 PHIs in[8] the HI for this dataset is calculated using (E.14). The HI is shown in Fig. E.7. As shown, the HI does not indicate a linear trend, but rather an exponentially increasing trend. However, the increase near the end (140 hours) is less sudden when compared to the ARMS in Fig. E.6 (b). Therefore, it may be easier to make a model that follows the trend of the HI compared to the ARMS. In addition, the threshold value is reached near the end, which suggests the validity of the proposed threshold.

E.7 Conclusions

Two failure thresholds (FTs) are proposed in this paper for remaining useful life-time (RUL) estimation of rotating machinery. The FTs are based on the ISO 10816-3 standard for acceptable vibration levels, and historic failure data for each specific machine is not required. The ISO standard provides acceptable levels for velocity RMS (VRMS) vibration which are not suitable for RUL estimation. That is because the initial bearing wear induce increased vibration levels at high frequencies, typically higher than 1000 Hz. Velocity is the integral of acceleration, and this integration causes high-frequency vibration components to be attenuated. The vibration energy at low frequencies only rises significantly once the bearing is critically damaged, and therefore it may be too late to schedule maintenance. To improve the RUL estimation capability of the ISO VRMS threshold, the first presented FT transforms it to acceleration unit. Experimental results show that: 1) the transformed acceleration based RMS (ARMS) threshold is reached by

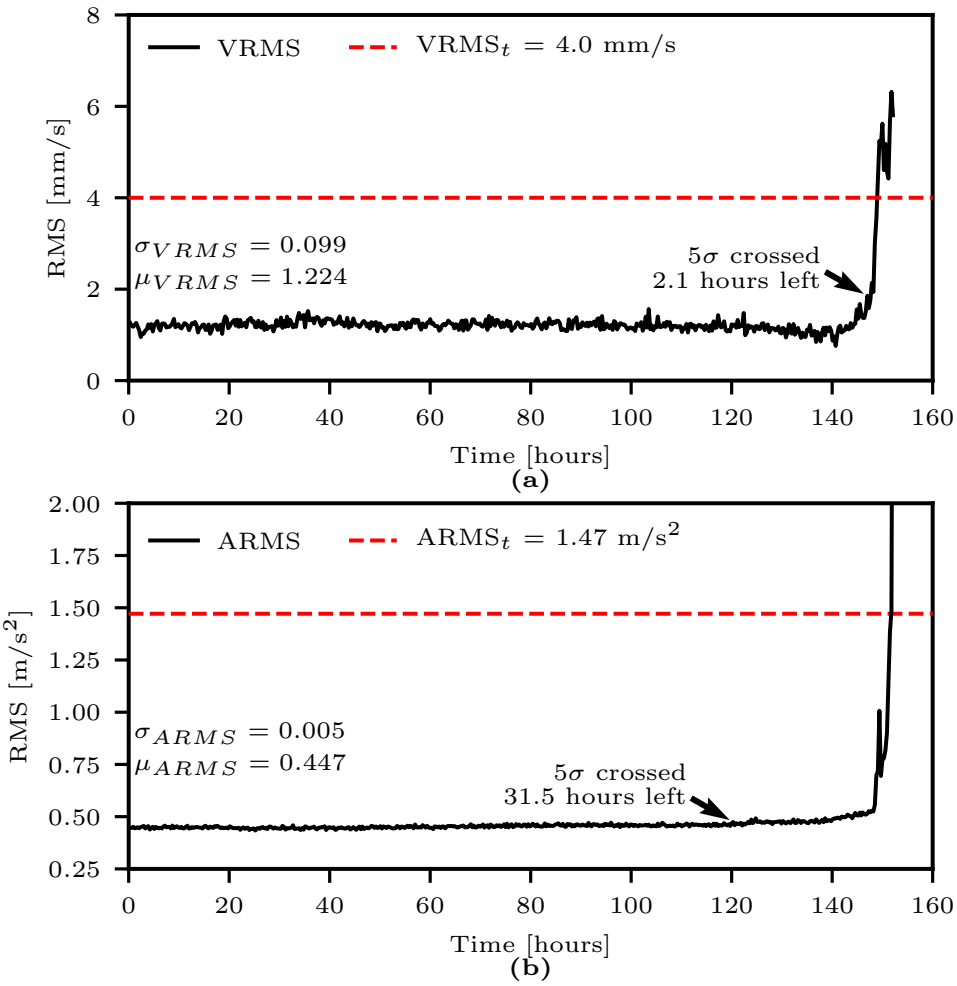

Figure E.6: Comparison of the RMS threshold using velocity and acceleration units for dataset 3. (a) The RMS of the velocity signal. (b) The RMS of the acceleration signal.

the ARMS at approximately the same time as the VRMS reaches its threshold, and 2) that the ARMS show signs of early degradation earlier when compared to the VRMS. The second FT presented is for a health index (HI) based on multiple vibration physical health indicators (PHIs) fused into a single virtual health indicator (VHI) using the Mahalanobis distance (MD). Experimental results show that the HI threshold is suitable, and that the HI presents a different degradation trend than the ARMS. In practice, these FTs can be employed for RUL estimation on any rotating machine in which a tolerable VRMS level is given either by the ISO 10816-3 standard, or from experience by an operator. For future work it may be beneficial to replace some PHIs in the MD, and also to scale up the ARMS threshold for other VHIs than the MD.

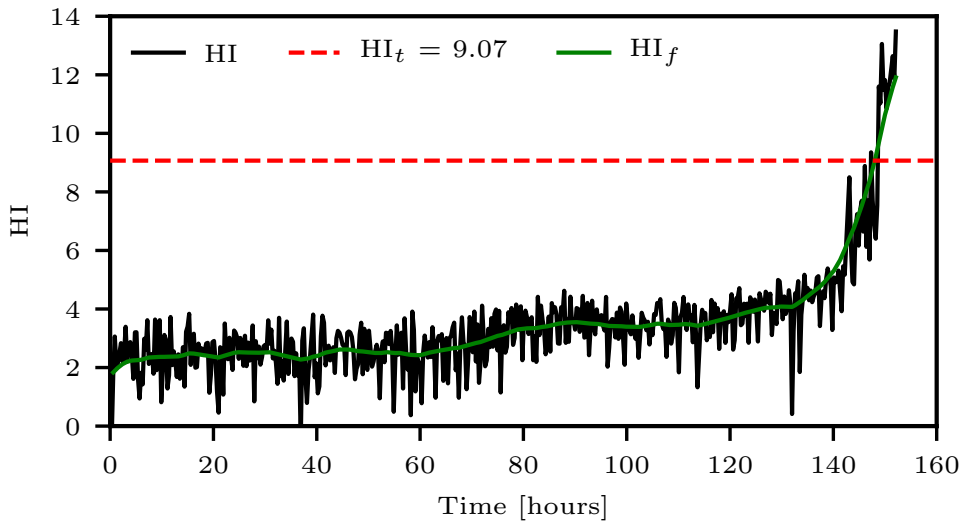

Figure E.7: The HI plotted with the HI$_t$ threshold for dataset 3.

REFERENCES

[1] W. Zhou, T. G. Habetler, and R. G. Harley. Bearing condition monitoring methods for electric machines: A general review. In *Proc. IEEE Int. SDEMPED*, pages 3–6. IEEE, 2007. doi:10.1109/demped.2007.4393062.

[2] Report of large motor reliability survey of industrial and commercial installations, Part I. *IEEE Transactions on Industry Applications*, 21(4):853–864, 1985. doi:10.1109/tia.1985.349532.

[3] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011. doi:10.1016/j.ymssp.2010.07.017.

[4] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 104:799–834, 2018. doi:10.1016/j.ymssp.2017.11.016.

[5] N. Li, Y. Lei, J. Lin, and S. X. Ding. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Transactions on Industrial Electronics*, 62(12):7762–7773, 2015. doi:10.1109/TIE.2015.2455055.

[6] Y. Lei, N. Li, and J. Lin. A new method based on stochastic process models for machine remaining useful life prediction. *IEEE Transactions on Instrumentation and Measurement*, 65(12):2671–2684, 2016. doi:10.1109/TIM.2016.2601004.

[7] Z. Huang, Z. Xu, X. Ke, W. Wang, and Y. Sun. Remaining useful life prediction for an adaptive skew-Wiener process model. *Mechanical Systems and Signal Processing*, 87:294–306, 2017. doi:10.1016/j.ymssp.2016.10.027.

[8] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K.-L. Tsui. A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Transactions on Industrial Informatics*, 12(3):924–932, 2016. doi:10.1109/TII.2016.2535368.

[9] K. Javed, R. Gouriveau, and N. Zerhouni. A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. *IEEE Transactions on Cybernetics*, 45(12):2626–2639, 2015. doi:10.1109/TCYB.2014.2378056.

[10] ISO. Mechanical vibration—Evaluation of machine vibration by measurements on non-rotating parts—Part 3 (Standard No. 10816-3). *ISO*, 1998.

[11] H. Qiu, J. Lee, J. Lin, and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006. doi:10.1016/j.jsv.2005.03.007.

[12] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, and C. Varnier. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12.*, pages 1–8. IEEE Catalog Number: CPF12PHM-CDR, 2012.

[13] A. Klausen, R. W. Folgerø, K. G. Robbersmyr, and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Modeling, Identification and Control*, 38(3):143–156, 2017. doi:10.4173/mic.2017.3.4.

Paper F

Novel RMS Based Health Indicators used for Remaining Useful Lifetime Estimation of Bearings

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

This paper has been submitted as:

A. Klausen, H. V. Khang and K. G. Robbersmyr. Novel RMS Based Health Indicator used for Remaining Useful Lifetime Estimation of Bearings. Under review at *Mechanical Systems and Signal Processing*.

Novel RMS Based Health Indicators used for Remaining Useful Lifetime Estimation of Bearings

Andreas Klausen, Huynh Van Khang and Kjell G. Robbersmyr

*University of Agder
Department of Engineering Sciences
Jon Lilletunsvei 9, 4879 Grimstad, Norway

Abstract **– Estimating the remaining useful life (RUL) of bearings from healthy to faulty is important for predictive maintenance. The bearing fault severity can be estimated based on the energy or root mean square (RMS) of vibration signals. However, the vibration RMS is often not monotonic, which renders a challenge for predicting the RUL. To tackle the problem, a new method is proposed for splitting the vibration signal into multiple frequency bands for RMS calculations. Compared to a digital filter bank, the proposed method is lossless, and the computational burden is much smaller. Some of the resulting RMS trends are more relevant for RUL estimation due to their monotonicity. Suitable RMS trends are identified using the Spearman coefficient, and the RUL is afterwards estimated with the Paris-Erdogan law and particle filters. Historical failure data is not required to set any parameters, which is a clear advantage of the proposed method. Experimental results from two test rigs validate the performance of the proposed method.**

F.1 Introduction

Bearings are common in rotating machines, and bearing defects result in increased vibration, temperature, and friction. Up to 44% failures in the most common motors, namely induction motors, are due to bearing faults [1]. The vibration from severe bearing faults may cause damage to other machine components, such as gears, stators and pump seals, and should therefore be detected. Unscheduled stops can cause long downtime and huge expenses due to maintenance and productivity losses. Therefore, monitoring the bearing health condition is important to avoid emergency shutdowns and plan maintenance.

A condition monitoring (CM) system allows for detecting faults, and therefore enhances machine reliability. On rotating machines, bearing CM systems can be used to detect faults using vibration signals, typically with the envelope spectrum [2]. Continued operation of a machine after detecting the initial bearing fault is beneficial for planning

maintenance, or even necessary if the machine should be shut down in a regulated manner rather than immediately. As such, an estimation of the bearing remaining useful life (RUL) is important to select between regulated or emergency stop.

A bearing health indicator (HI) can be compared to a failure threshold (FT) to estimate the RUL by assuming that a machine should stop if the HI reaches the FT. The future HI trend can be predicted using a mathematical model, and the time until FT is reached is the estimated RUL. The bearing HI can be assessed by examining the level of wear on the bearing rollers and raceways. However, the actual bearing HI is impractical to determine, as it would require an offline inspection after dissembling the bearing [3], resulting in productivity loss. Instead, the HI can be estimated using sensor signals that are related to the amount of bearing wear. Vibration signals [4] can be used for this purpose, as defects in the raceways or rollers increases the vibration energy. The vibration signal is often reduced to features which can afterwards be used as the estimated HI.

HIs can be categorized into physical HIs (PHIs) and virtual HIs (VHIs) [5]. PHIs are generated from primarily physical signals and are directly related to the physics of failure. Examples are the root mean square (RMS) [6], kurtosis [7], and characteristic bearing fault frequency amplitudes [8]. VHIs do not correlate directly with the physics of failure [5] and can be calculated by combining multiple PHIs. The Mahalanobis distance was applied in [9] to combine 14 PHIs into a single VHI, and principal component analysis (PCA) was used in [10] to estimate the principal component of multiple PHIs.

The bearing degradation can be divided into two or multiple health stages (HSs) [3]. During the first HS, there is no apparent degradation, while the second HS often show linear increase, and the third could be unstable growth. The RUL is normally estimated after transitioning from the first to the second HS. This transition can be detected using baseline measurements of the kurtosis value [6] and the RMS signal [11].

After transitioning to the second HS, the future degradation trend can be estimated by a mathematical model which closely resembles the physics of failure. The exponential model [6, 12], Brownian Motion [9], and Paris-Erdogan law [13] have been used to predict the future bearing degradation trend. The Kalman filter [14] and particle filter [15] have been used to update parameters for mathematical models and predict the future degradation trend. Alternatively, data-driven methods allow for tracking the trend without knowing the physics of failure. A Gaussian process model [16] and least squares support vector machine (LSSVM) [10] are examples of data-driven models.

Setting a proper FT is the final prerequisite for predicting RUL using the aforementioned methods. For complex VHIs, historical failure data from a similar setup is often necessary to create a suitable FT [17, 18, 13]. However, for machines with no historical failure data, many of the methods reported in the literature may not work. The

implemented solutions in the referred papers require use of historic failure data to set parameters, rendering a challenge for estimating RUL on new machines. In [19], the FT is set based on the vibration RMS as guided by the ISO standard 10816-3 [20], without involving historic failure data. However, estimating the RUL using RMS is difficult as vibration energy is often not monotonic.

To address the existing challenges, a new approach for extracting the mean degradation trend of the vibration RMS signal is proposed. The vibration signal is split into multiple frequency bands to separate monotonic components from non-monotonic ones. Instead of using a digital filter bank, the new approach utilizes only the discrete Fourier transform of the vibration signal. Compared to digital filters, the new approach yields no information loss, and is less computationally expensive. The RMS is calculated for each frequency band components, and the Spearman coefficient is used to determine monotonic RMS trends suitable for RUL estimation. FTs of the generated RMS trends are determined by extending the FT calculations in [19]. The degradation trend is predicted by updating parameters of the Paris-Erdogan law using a particle filter (PF). Historic failure data is not required, because the FTs are based on general guidelines provided by ISO 10816-3 [20]. The performance of the proposed method is demonstrated on two experimental datasets.

The rest of the paper is organized as follows. Practical differences between velocity- and acceleration-based RMS are discussed in Section F.2. Next, the proposed approach for subdividing the vibration signal into multiple RMS trends is detailed in Section F.3. Afterwards, the algorithms used for RUL estimation are elaborated in Section F.4, and the experimental results are given in Section F.5. Finally, suggestions for further work are given in Section F.6, and conclusions are drawn in Section F.7.

F.2 Velocity vs acceleration RMS

ISO Standard 10816-3 [20] defines levels of velocity-based vibration RMS values for rotating machines with over 15 kW power output. Four levels are given, ranging between A - D: A="standard acceptance for new machines"; B="unlimited operation is possible"; C="short term operation allowed"; D="vibration causes damage". A machine should stop before reaching level D, and therefore the boundary between level C and D can be used as an FT. The RMS of a vibration signal $x(t)$ can be calculated with

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

where T is the measurement period. The vibration signal $x(t)$ is in this study measured with an accelerometer, which gives the vibration in acceleration units. Through integra-

tion, the *velocity* signal $x_v(t)$ is acquired, which is used for checking vibration severity according to ISO 10816-3 [20]. For brevity, $R_a = \text{RMS}(x(t))$ and $R_v = \text{RMS}(x_v(t))$. However, R_v is not suitable for RUL estimation, because early defects generate high frequency resonance vibration, and integration reduces the effect of high frequency components as verified in [19]. To alleviate this limitation, the FT is transformed to the acceleration domain before estimating bearing RUL. This procedure was proposed in [19] and is briefly described in the following.

To experimentally transform the FT from velocity to acceleration units, the mean R_a and R_v values at a known steady state of the machine (baseline) are calculated. Afterwards, the ratio between these is determined with [19]

$$R_r = \frac{\bar{R}_a}{\bar{R}_v}\,, \tag{F.2}$$

where $\bar{R}_a$ and $\bar{R}_v$ are the mean R_a and R_v during baseline measurements, respectively. Using this ratio, the acceleration-based FT $\hat{R}_a$ is calculated using [19]

$$\hat{R}_a = R_r \hat{R}_v\,, \tag{F.3}$$

where $\hat{R}_v$ is the velocity-based FT from the ISO standard.

Vibration data collected by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS) [21] is used to validate this transformation. Accelerometer data from the second run-to-failure test, consisting of 982 vibration measurements, are utilized. Every 10 minutes, 1 second of vibration data was measured at a sample rate of 20480 Hz. At the end of this test, an outer race fault was observed in the dismantled bearing. The motor size is assumed smaller than 15 kW, and therefore the FT is set slightly lower than level D in the ISO standard, at $\hat{R}_v = 4$ mm/s.

The first 30 hours of vibration data (180 datasets with 10 minutes interval) is set as a baseline, and the mean RMS values are $\bar{R}_a = 0.757$ m/s^2 and $\bar{R}_v = 0.755$ mm/s. Using (F.2) and (F.3) results in $\hat{R}_a = 4.01$ m/s^2. Fig. F.1 (a) shows R_v and $\hat{R}_v$, while Fig. F.1 (b) shows R_a and the resulting FT $\hat{R}_a$. As observed in Fig. F.1 (a), R_v starts increasing from the mean value near the end of useful life (after 150 hours), giving a very short time to shut the machine down and plan maintenance. However, R_a starts increasing much sooner ($t \approx 80$ hours), which allows for predicting RUL and plan maintenance in advance. The two trends reach their respective FTs near the end of life, which shows that the transformation can be used for this purpose.

R_a in Fig. F.1 (b) is, however, not suitable for RUL estimation due to oscillations around the mean degradation. Fig. F.2 shows R_a and an artificially created mean trend to highlight the cyclic behavior. For example, the sudden increase of R_a at $t \approx 115$ hours, and the subsequent decrease, render a challenge for determining the actual degradation

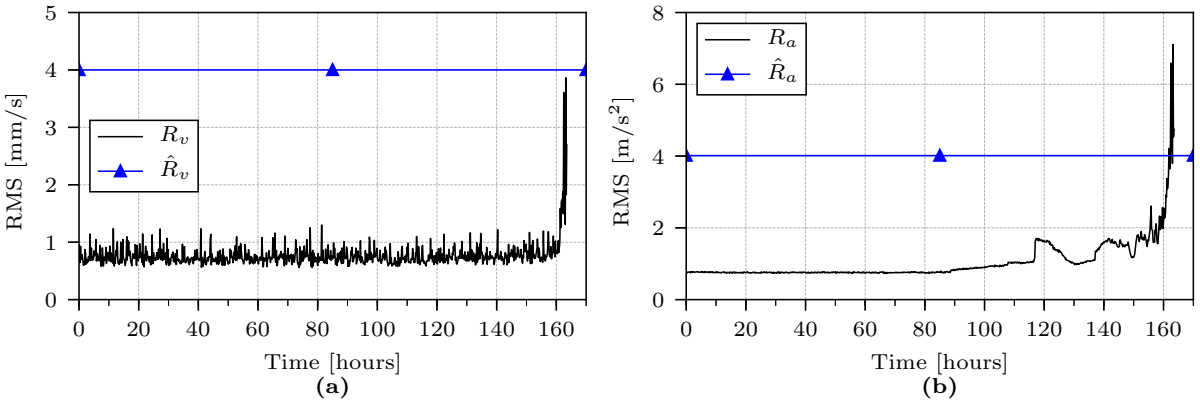

Figure F.1: Comparison between velocity- and acceleration-based RMS for the IMS dataset. (a) velocity-based RMS and its FT; (b) acceleration-based RMS and its FT.

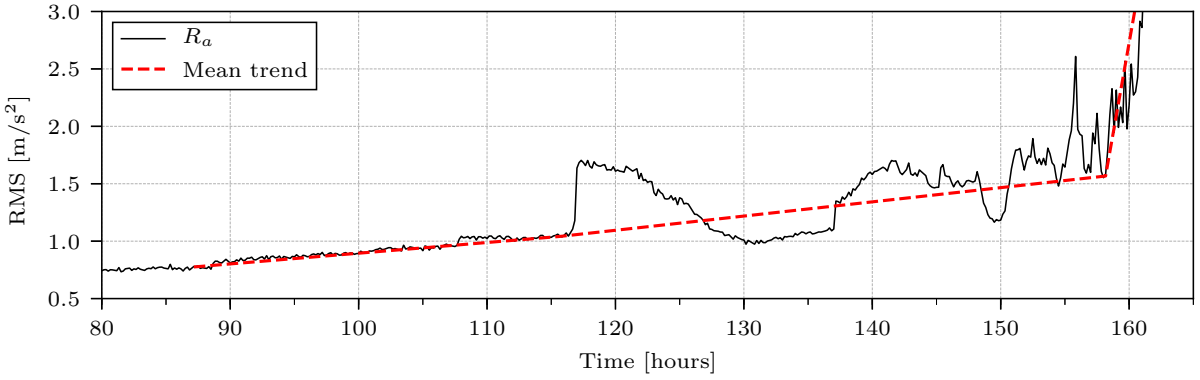

Figure F.2: R_a compared to an artificially created mean trend to highlight oscillations.

trend. A more stable trend should be extracted from the vibration signal to achieve better RUL estimation. In the next subsection, a new approach for extracting more useful RMS data from the vibration signal is proposed.

F.3 Proposed RMS health indicator

F.3.1 Filter bank RMS

A digital filter bank can be used to subdivide a signal into multiple components, where each component contains a frequency sub-band of the signal. This can be achieved by iteratively passing the signal through a low- and high-pass finite impulse response (FIR) filter and decimating each output signal to half the frequency. Such a procedure has some limitations: Digital FIR filters use convolution to filter the signal, and only the overlapping part between the filter kernel and the signal should be preserved to avoid adding artifacts. This means that a high order FIR filter kernel with a sharp frequency

response will remove much of the signal energy. In addition, the energy loss is exponential for each level of filtering. If, however, the FIR filter kernel is small, there can be large frequency overlap between the signal components. Computing a digital filter bank can also be computationally taxing if the input signal is long, and if multiple frequency levels are required. Alternatively, the vibration signal can be split into sub-bands using the discrete wavelet transform (DWT) [22]. A time-domain filter bank is, however, not necessary to acquire the RMS in different frequency bands. The following explains how the spectrum bins from a single discrete Fourier transform (DFT) can be used to directly calculate the vibration RMS in a certain frequency band.

An alternative representation of RMS is given by the energy E of the signal, such as

$$\mathrm{RMS}(x) = \sqrt{\frac{\mathrm{E}(x)}{T}}\,. \tag{F.4}$$

The signal energy can be calculated in both time and frequency domain with

$$\mathrm{E}(x) = \sum_{i=1}^{n} |x_i|^2 \Delta t = \sum_{i=1}^{n} |X_i|^2 \Delta f\,, \tag{F.5}$$

where n is the length of the sampled signal, $x = (x_1, \ldots, x_n)$ is the vibration signal given in discrete time samples, $X = (X_1, \ldots, X_n)$ is the frequency spectrum, Δt is the time interval between samples, and Δf is the frequency step between each spectrum bin.

To explain the procedure, the frequency spectrum is assumed ordered from 0 Hz to Nyquist frequency, and from negative Nyquist up to 0 Hz. Let $X_L = (X_{LP}, Z_{n/2}, X_{LN})$, where X_{LP} contains the spectrum bins of the lower positive frequencies (i.e. 0 Hz to half the Nyquist frequency), $Z_{n/2}$ is $n/2$ zeros, and X_{LN} contains the spectrum bins of the lower negative frequencies. Similarly, let $X_H = (Z_{n/4}, X_{HP}, X_{HN}, Z_{n/4})$, where X_{HP} and X_{HN} are the spectrum bins of higher positive and negative frequencies, respectively. This arrangement represents a single level filter bank that splits X at half Nyquist frequency completely. The total energy of the signal can afterwards be calculated as

$$\begin{aligned} \mathrm{E}(X) &= \mathrm{E}(X_L + X_H) \\ &= \sum_{i=1}^{n} |X_{L,i} + X_{H,i}|^2 \Delta f \end{aligned} \tag{F.6}$$

where $X_{L,i}$ and $X_{H,i}$ are the i'th bin of X_L and X_H, respectively. Let the spectrum bins be given by their complex values, i.e. $X_{L,i} = a_{L,i} + jb_{L,i}$, $X_{H,i} = a_{H,i} + jb_{H,i}$, then (F.6) becomes

$$\begin{aligned} &= \sum_{i=1}^{n} \left(\sqrt{(a_{L,i} + a_{H,i})^2 + (b_{L,i} + b_{H,i})^2} \right)^2 \Delta f \\ &= \sum_{i=1}^{n} \left(a_{L,i}^2 + 2a_{L,i}a_{H,i} + a_{H,i}^2 + b_{L,i}^2 + 2b_{L,i}b_{H,i} + b_{H,i}^2 \right) \Delta f\,. \end{aligned} \tag{F.7}$$

Given that X_L and X_H represent a complete signal separation at half Nyquist frequency, the overlap between X_L and X_H is zero. Therefore, $2a_{L,i}a_{H,i} = 2b_{L,i}b_{H,i} = 0$, and (F.7) is reduced to

$$\begin{aligned}\mathrm{E}(X) &= \sum_{i=1}^{n}(a_{L,i}^2 + a_{H,i}^2 + b_{L,i}^2 + b_{H,i}^2)\Delta f \\ &= \sum_{i=1}^{n}(|a_{L,i} + jb_{L,i}|^2)\Delta f + \sum_{i=1}^{n}(|a_{H,i} + jb_{H,i}|^2)\Delta f \\ &= \mathrm{E}(X_L) + \mathrm{E}(X_H)\,. \end{aligned} \tag{F.8}$$

Eq. (F.8) shows that the energy of the entire signal can be calculated by the energy of separate spectrum bins. It can also be shown as an extension of (F.4) that

$$\begin{aligned}\mathrm{RMS}(X) &= \sqrt{\frac{\mathrm{E}(X_L) + \mathrm{E}(X_H)}{T}} \\ &= \sqrt{\mathrm{RMS}(X_L)^2 + \mathrm{RMS}(X_H)^2}\,. \end{aligned} \tag{F.9}$$

Eq. (F.8) also indicates that the signal can be split into multiple equally-sized frequency bands such that

$$\mathrm{E}(X) = \sum_{i=1}^{N_b}\mathrm{E}(X_{(i-1)n_b+1:in_b}) \tag{F.10}$$

$$\mathrm{RMS}(X) = \sqrt{\sum_{i=1}^{N_b}\mathrm{RMS}(X_{(i-1)n_b+1:in_b})^2}\,, \tag{F.11}$$

where N_b is the number of frequency bands the spectrum is split into, and $n_b = n/N_b$. For brevity, the RMS of a frequency band i is defined as

$$R_i = \mathrm{RMS}(X_{(i-1)n_b+1:in_b})\,, \tag{F.12}$$

and the RMS value for band i at time index k is given by $R_{i,k}$. In summary, the spectrum X is first obtained for the entire vibration signal x, and afterwards the energy of bins belonging to frequency band i is used to calculate R_i using (F.4). The next step is to obtain FTs for all R_i.

F.3.2 Filterbank RMS failure thresholds

To get an FT for each R_i, (F.11) is reconsidered such that

$$\hat{R}_a = \sqrt{\sum_{i=1}^{N_b}\hat{R}_i^2}\,, \tag{F.13}$$

where $\hat{R}_i$ is the FT for R_i. To solve (F.13) for any number of bands N_b, it is necessary to make some assumptions about the signal x. It is assumed that the vibration signal is corrupted with normally distributed noise. Such noise is evenly distributed in the frequency domain, and therefore the energy in all frequency bands should be affected by this noise. It is also assumed that the energy increase from bearing wear is proportional to the baseline noise variance in each frequency band. With these assumptions, the FTs are defined as

$$\hat{R}_i = \mu_i + m\sigma_i \,, \tag{F.14}$$

where μ_i and σ_i are the mean and standard deviation (STD) of R_i during baseline measurements, respectively, and m is a constant scaling factor. To determine m, the conservation of energy given by (F.13) is considered, such that

$$\hat{R}_a = \sqrt{\sum_{i=0}^{N_b} (\mu_i + m\sigma_i)^2} \tag{F.15}$$

$$\hat{R}_a^2 = \sum_{i=0}^{N_b} (\mu_i + m\sigma_i)^2 \tag{F.16}$$

$$= m^2 \sum_{i=0}^{N_b} \sigma_i^2 + m \sum_{i=0}^{N_b} 2\mu_i\sigma_i + \sum_{i=0}^{N_b} (\mu_i)^2 \,. \tag{F.17}$$

Solving (F.17) for m using the quadratic formula yields two possible solutions, m_1 and m_2. Due to squaring of the equation in (F.16), both a positive and negative solution for m are possible. The highest valued solution should be chosen so that the FT is above the mean value, i.e.

$$m = \max(m_1, m_2) \,. \tag{F.18}$$

The rest of this section contains an example of the proposed RMS filter bank. The IMS dataset introduced in Section F.2 is subdivided into $N_b = 32$ frequency bands (320 Hz bandwidth), and R_i is calculated for each band using (F.12). The mean μ_i and STD σ_i of each R_i are determined with the baseline measurements. With these values, the FTs for each R_i are calculated using (F.14) after solving (F.17) for m. Additionally, a degradation alarm is triggered at time index ka_i when $R_{i,ka_i} > \mu_i + 5\sigma_i$.

Fig. F.3 shows four RMS trends at different frequency ranges. Fig. F.3 (a) shows R_{15} from frequency band [4480, 4800] Hz, and the oscillations in the trend are similar to R_a shown in Fig. F.1 (b). On the other hand, R_3 ([640, 920] Hz) in Fig. F.3 (d) is monotonic after the initial alarm is triggered, and is very suitable for RUL estimation. The two other RMS trends in Figs. F.3 (b) and (c) are calculated from frequency bands between the two others, and it is seen that RMS trends from lower frequency bands are more monotonic. The initial degradation alarm is also triggered at different times, and

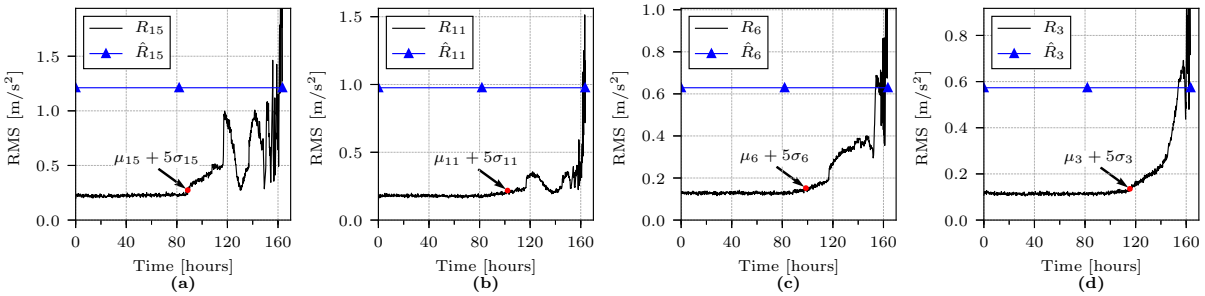

Figure F.3: Collection of RMS trends with FTs for the IMS dataset. (a) R_{15}; (b) R_{11}; (c) R_6; (d) R_3.

RMS trends from high frequency bands appear to trigger it first. In addition, the FTs $\hat{R}_i$ are also shown as blue lines in Fig. F.3. The FTs are reached near the end of useful life and are therefore considered useful for RUL estimation. The next section details how suitable RMS trends can be selected online and used to estimate the RUL.

F.4 Remaining useful life estimation

F.4.1 Overview

Suitable RMS trends for RUL estimation are identified by using the Spearman coefficient [23] in an online manner, which determines the monotonicity of each R_i. A particle filter (PF) is initialized for each R_i trend with a high Spearman coefficient, and the Paris-Erdogan law [24] is used to predict the degradation trend. The model parameters are initialized using a non-linear least squares (NLS) algorithm. For each new sample, the PFs are updated, and the RUL probability density function (PDF) for each PF is estimated. Finally, a weighted PDF combining the RUL estimation of all PFs is used to estimate bearing RUL. A flowchart of the proposed method is shown in Fig. F.4.

F.4.2 Spearman coefficient

The Spearman coefficient [23] is a rank-based correlation between two signals. It can be used to describe how monotonic a signal is, even when the signal is not linearly increasing. The Spearman coefficient between a signal x and time t is calculated as

$$\text{Spearman}(x,t) = \frac{\text{cov}(\text{rank}(x), \text{rank}(t))}{\text{STD}(\text{rank}(x))\text{STD}(\text{rank}(t))}, \tag{F.19}$$

where $\text{cov}(\cdot,\cdot)$ is the covariance of two trends, and $\text{rank}(\cdot)$ is the rank of a signal. Recall that $R_{i,k}$ is a sample at index k of R_i, and ka_i is the index for when $R_{i,ka_i} \geq \mu_i + 5\sigma_i$ for

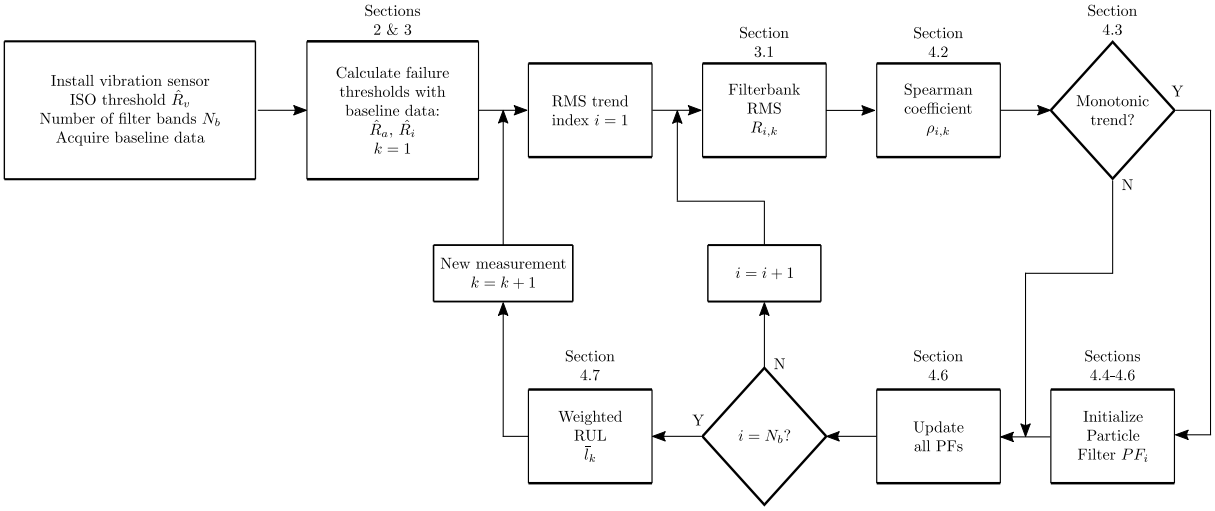

Figure F.4: Flowchart of the proposed method.

the first time. Then, the running Spearman coefficient for R_i is defined as

$$\rho_{i,k} = \begin{cases} \text{Spearman}(R_{i,ka_i:k}, t) & \text{if } k \geq ka_i \\ 0 & \text{else.} \end{cases} \tag{F.20}$$

Eq. (F.20) is used to continually check whether R_i increases monotonously over time.

F.4.3 RMS trend selection

The Spearman coefficient is used to numerically describe how monotonic the RMS trends are. Three criteria must be met by a certain R_i at time index k for it to be used for RUL estimation:

1. $k \geq ka_i + ks$, where ks is the minimum number of samples used to calculate Spearman coefficient.
2. $\rho_{i,k} \geq \hat{\rho}$, where $\hat{\rho}$ is the Spearman threshold.
3. $i \leq \min(I_p)$, where I_p is a set of indices i belonging to R trends previously accepted for RUL estimation.

The reasons of the criteria are as follows. 1) To calculate a stable Spearman coefficient, it is necessary with several samples. Therefore, ks is set as a minimum number of samples. 2) The Spearman coefficient must be higher than the threshold to avoid non-monotonic trends. 3) It is expected that RMS trends of low-frequency bands have less oscillations compared to higher frequency ones. Therefore, the index i of new trends must be smaller than all the other selected ones to avoid estimating RUL on unnecessary many trends. The degradation model chosen to predict each RMS trend is detailed in the next subsection.

F.4.4 Degradation model

The bearing degradation level is assumed to be monotonously increasing and never self-healing. In addition, as an increasing number of defects develop in the bearing, the vibration level increases, which results in an exponential degradation rate. A study [25] shows that bearing crack propagation may be modeled with the Paris-Erdogan law [24] This model describes the crack propagation rate in materials under cyclic load, and is given by

$$\frac{da}{dn_c} = c(\Delta k)^m\,, \quad \Delta k = \Delta\sigma\gamma\sqrt{\pi a}\,, \tag{F.21}$$

where a is the crack size, n_c is the cycle number, c, m and γ are material constants, and $\Delta\sigma$ is the cyclic load amplitude. The material constants can be estimated via experimental testing with a known cyclic load, while measuring the crack length. However, it is impractical to measure the crack length within a bearing during operation. Therefore, a is instead estimated with an HI based on the vibration measurements [26].

Let $\alpha = \gamma c \Delta\sigma\sqrt{\pi}$ and $\beta = m/2$, then the modified Paris-Erdogan law is [13]

$$\frac{da}{dn_c} = \alpha a^\beta\,. \tag{F.22}$$

Eq. F.22 is afterwards re-written in the form of a state-space model as [13]

$$\begin{cases} a_k &= a_{k-1} + \alpha_{k-1} a_{k-1}^\beta \Delta n_c \\ \alpha_k &= \alpha_{k-1} \\ h_k &= a_k + \nu_h\,, \end{cases} \tag{F.23}$$

where $\alpha_{k-1} \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$ is a random variable given by a normal distribution $\mathcal{N}$, β is a constant, $\nu_h \sim \mathcal{N}(0, \sigma_h^2)$ is measurement noise, and Δn_c is the number of cycles since last update.

A PF is applied to update the model parameters based on new samples, but reasonable initial values should be set first. The modified Paris-Erdogan law in (F.23) has five unknown parameters $(a_1, \mu_\alpha, \beta, \sigma_\alpha^2, \sigma_h^2)$ which must be identified. The mean values $\Theta_1 = (a_1, \mu_\alpha, \beta)$ are estimated using a non-linear least squares (NLS) algorithm, and afterwards the variances $\Theta_2 = (\sigma_\alpha^2, \sigma_h^2)$ are set based on other criteria

F.4.5 Initial parameter setting

Three model parameters $\Theta_1 = (a_1, \mu_\alpha, \beta)$ are identified by minimizing an objective function using NLS. The objective function is the mean square error between measurements and model prediction. Let the measurement HI be $h = (R_{i,k-ko:k})^T$, where ko is the number of samples used in the optimization. To obtain the objective function, the ordinary

differential equation in (F.22) is solved for a as

$$a(n_c, \Theta_1) = (C_1(1-\beta) + \alpha n_c(1-\beta))^{1/(1-\beta)}\,, \quad \beta \neq 1\,, \tag{F.24}$$

where C_1 is based on initial conditions. The process noise and measurement noise are assumed zero in the minimization algorithm. C_1 is determined by solving (F.24) with initial values $a = a_1$ and $n_c = 0$. After substitution of C_1, (F.24) becomes

$$a(n_c, \Theta_1) = \left(\alpha n_c\,(1-\beta) + a_1^{1-\beta}\right)^{1/(1-\beta)}\,. \tag{F.25}$$

The unknown parameters Θ_1 are identified using an NLS minimization routine described as

$$\arg\min_{\Theta_1} F(h, \Theta_1)\,, \tag{F.26}$$

$$\text{subject to } \begin{cases} \min(h) \leq a_1 \leq \max(h) \\ 0 < \mu_\alpha < \infty \\ 0.5 < \beta < 1.25\,, \end{cases} \tag{F.27}$$

$$\text{where } F = 0.5 \sum_{i=1}^{ko} (h_i - a(n_c = i-1, \Theta_1))^2\,. \tag{F.28}$$

A trust region reflective algorithm [27] is used in this research to minimize (F.26). The constraints for β are set so that the exponential rate does not get unstable.

The variances in $\Theta_2 = (\sigma_\alpha^2, \sigma_h^2)$ are set based on baseline data and optimized values in Θ_1. The measurement noise variance σ_h^2 is simply identified as the variance of R_i during the baseline measurements, i.e. $\sigma_h^2 = \sigma_i^2$. A particle filter is quite dependent on the choice of process noise, i.e. σ_α^2, and a larger variance gives more headroom for the filter in case the chosen model does not fit very well with measured data. To give enough headroom for the filter, the mean value μ_α is used as the standard deviation, such that $\sigma_\alpha^2 = \mu_\alpha^2$.

F.4.6 Particle filter

After initializing the model parameters, a sequential importance sampling (SIS) PF is applied to further update the parameters in the degradation model. General PF theory is given in [13, 26, 28] for interested readers. A set of initial particles $\boldsymbol{z}_k^j \,\forall\, j \in [1, N_p]$ are initialized at time index k with

$$\boldsymbol{z}_k^j \sim \mathcal{N}\left(\begin{bmatrix} R_{i,k} \\ \mu_\alpha \end{bmatrix}, \begin{bmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix}\right)\,, \tag{F.29}$$

where N_p is the number of particles. The particle weights are initialized as $w_k^j = 1/N_p$. The next model state is predicted using the particles with

$$\boldsymbol{z}_k^j = \begin{bmatrix} a_k^j \\ \mu_\alpha^j \end{bmatrix} = \begin{bmatrix} a_{k-1}^j + (\mu_\alpha^j + \nu_{\alpha,k})^j (a_{k-1}^j)^\beta \Delta t \\ \mu_\alpha^j \end{bmatrix}, \tag{F.30}$$

where $\nu_{\alpha,k}^j \sim \mathcal{N}(0, \sigma_\alpha^2)$ is the process noise which is added due to the uncertainty of bearing load, and Δt is the time between cycles. When a new measurement $R_{i,k}$ is available, the particle weights are updated and normalized with

$$\tilde{w}_k^j = w_{k-1}^j p(R_{i,k} \mid \boldsymbol{z}_k^j) \tag{F.31}$$

$$w_k^j = \tilde{w}_k^j \Big/ \sum_{j=1}^{N_p} \tilde{w}_k^j \ , \tag{F.32}$$

where $\tilde{w}_k^j$ is the un-normalized particle weight and

$$p(R_{i,k} \mid \boldsymbol{z}_k^j) = \frac{1}{\sqrt{2\pi\sigma_h^2}} \exp\left[-\frac{(R_{i,k} - a_k)^2}{2\sigma_h^2} \right]. \tag{F.33}$$

To deal with particle degeneracy, a re-sampling step is utilized if necessary. When

$$N_{eff} = 1.0 \Big/ \sum_{i=1}^{N_p} w_k^j < N_p/2 \,, \tag{F.34}$$

particles are re-sampled according to a systematic re-sampling [29] approach, and particle weights are re-initialized as $w_k^j = 1/N_p$.

The RUL estimated by particle j at time index k is given by

$$l_k^j = \inf\{ l_k^j : a^j(l_k + t_k) \geq \hat{R}_i \mid a_{1:k}^j \} \,, \tag{F.35}$$

where $a^j(l_k + t_k)$ is the state value for particle j at time $t_k + l_k$, $a_{1:k}^j$ is the estimated state value at $1, \ldots, k$ and $\hat{R}_i$ is the FT for R_i. To solve (F.35), the state of each particle j are simulated using the state transition function given by (F.30) up to the time $a^j(l_k + t_k) \geq \hat{R}_i$. With the estimated RUL and weight for each particle, the probability density function (PDF) for RUL l_k is approximated by

$$p(l_k \mid R_{i,1:k}) = \sum_{j=1}^{N_p} w_k^j \delta(l_k - l_k^j) \,, \tag{F.36}$$

where $\delta()$ is the delta-dirac function.

F.4.7 Weighted RUL

The vibration signal is split into N_b frequency bands, where each band is used to calculate RMS R_i. The running Spearman coefficient is used to detect monotonously increasing

RMS trends, and whenever a new trend has a high Spearman coefficient, a new PF is initialized and updated for that trend. In theory, up to N_b PFs can be initialized, and the RUL estimated from all are used in a weighted RUL decision. Let PF_i denote the PF for R_i, and let $w_{i,k}^j$ and $l_{i,k}^j$ denote the weight and RUL, respectively, for particle j at time index k in PF_i. When R_i stops being monotonic, the PF output associated with it should count less in a weighted RUL estimation. Additionally, once the Spearman coefficient of a trend is less than ρ_L, that trend is neglected in the weighted RUL estimation. Then, let the weight of each PF_i be

$$W_{i,k} = \left(\rho_{i,k}^3 - \rho_L^3\right) / \left(1 - \rho_L^3\right) \ , \tag{F.37}$$

where the Spearman coefficient is cubed to prioritize monotonic trends. Afterwards, the weighted RUL PDF at time index k is defined as

$$p(\bar{l}_k \mid R_{1:N_b,1:k}) = \sum_{i=1}^{N_b} W_{i,k} \sum_{j=1}^{N_p} w_{i,k}^j \delta(\bar{l}_k - l_{i,k}^j) \, . \tag{F.38}$$

Instead of using the median of the PDF in (F.38) as the estimated RUL, the weighted mean is instead chosen. The reason is that the median of an even number of PFs will most likely fall under either one of them, even if the medians are far from each other. Instead, the weighted mean of the PDF is chosen as the estimated RUL. This weighted mean is calculated with

$$\text{Weighted mean}(k) = \frac{\sum_{i=1}^{N_b} W_{i,k} \sum_{j=1}^{N_p} w_{i,k}^j l_{i,k}^j}{\sum_{i=1}^{N_b} W_{i,k} \sum_{j=1}^{N_p} w_{i,k}^j} \, . \tag{F.39}$$

The weighted mean and 95% confidence interval (CI) of the weighted RUL PDF are determined after each measurement update to track the estimated RUL over time.

F.5 Experimental results

F.5.1 Test rig 1

To validate the proposed method, the IMS bearing dataset [21] is utilized. $ks = 30$ is the minimum number of samples for Spearman calculation, $ko = 30$ samples are used to initialize PFs, and $\hat{\rho} = 0.9$ is the Spearman threshold. In addition, $N_p = 1000$ particles are used in each PF, and $\rho_L = 0.7$ is the lower Spearman threshold in (4.16). Figs. F.5 and F.6 shows the results of using the proposed method on this dataset. Subplots in the first column show the identified RMS trends with high Spearman coefficients, and the output of the initialized PF. R_i's are shown by black lines, and the median and 95% confidence interval of the initialized PF output are shown as red and red-stapled lines, respectively.

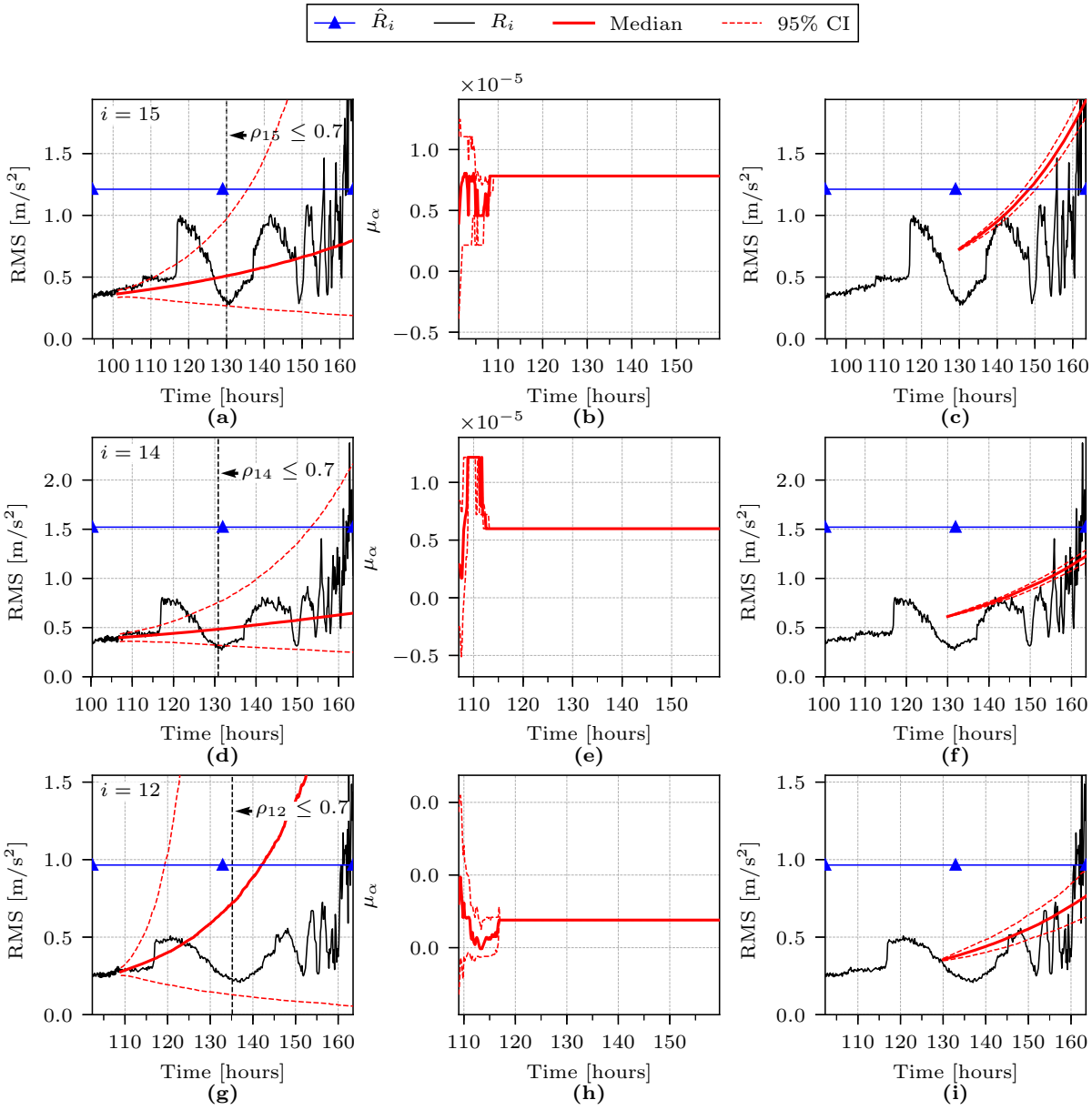

Figure F.5: Identified RMS trends with high Spearman coefficient, and output of the corresponding PFs—Part 1. Rows 1-3 indicate $i = [15, 14, 12]$. (column 1) R_i, FT $\hat{R}_i$, and median and 95% CI of initial PF output; (column 2) μ_α over time for the initiated PF; (column 3) predicted PF trend at $t = 130$ hours.

In addition, blue lines indicate the FT. Each row in Figs. F.5 and F.6 corresponds to a single index i, which is given in the upper left corner of the leftmost subplot. The second column show the median and 95% confidence interval (CI) of the μ_α parameter. The predicted PF output at $t = 130$ hours is given in column 3 to show how the particle filters are converging over time.

At $t = 101$ hours (k = 608), a new PF is initialized for R_{15}, indicating a change in bearing health. This RMS trend, shown in Fig. F.5 (a), is from a high-frequency

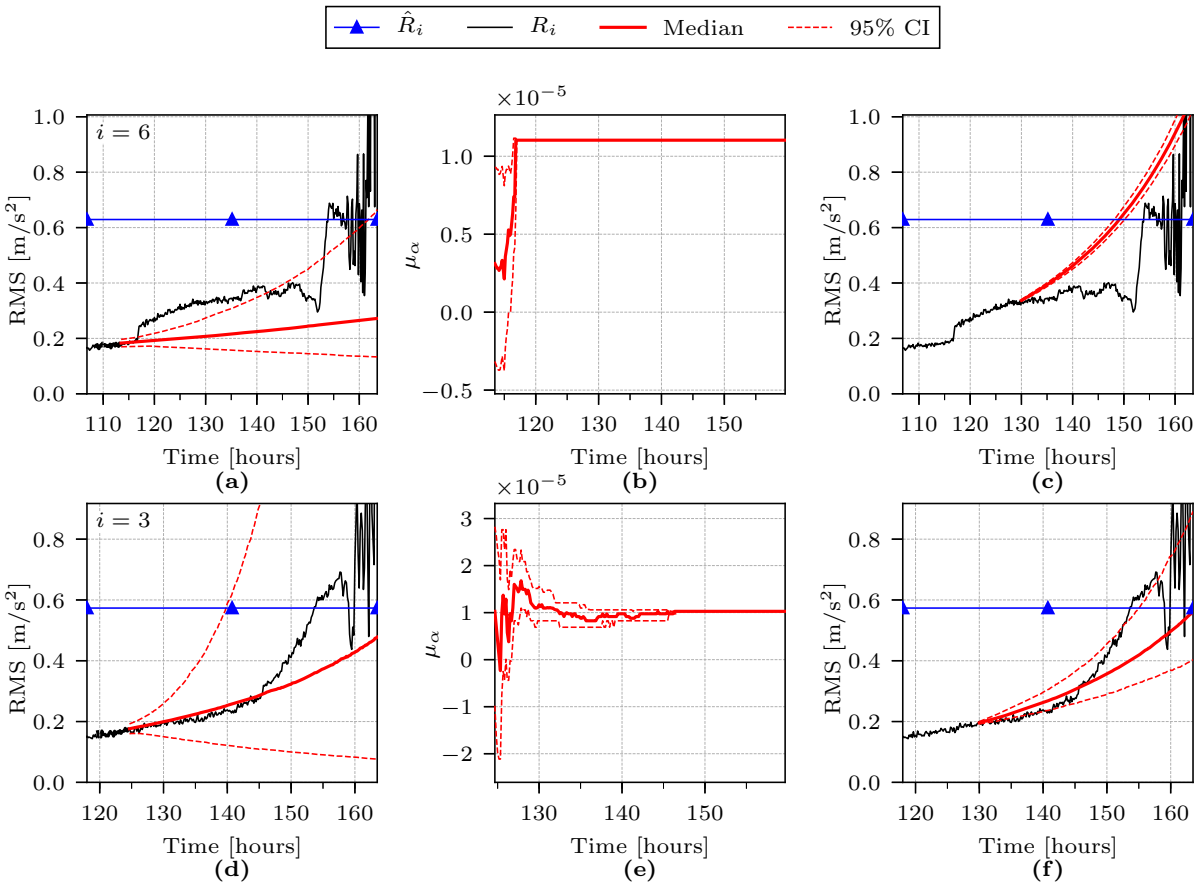

Figure F.6: Identified RMS trends with high Spearman coefficient, and output of the corresponding PFs—Part 2. Rows 1-2 indicate $i = [6, 3]$. (column 1) R_i, FT $\hat{R}_i$, and median and 95% CI of initial PF output; (column 2) μ_α over time for the initiated PF; (column 3) predicted PF trend at $t = 130$ hours.

band within [4480, 4800] Hz. Initially, R_{15} increases monotonously, but at $t = 117$ hours the trend starts to oscillate around a mean value. This behavior is similar to the cyclic behavior observed in the full RMS R_a. μ_α has converged before $t = 110$ hours as seen in Fig. F.5 (b), and therefore the estimated trend will continue until the end of RUL. Fig. F.5 (c) shows the median and 95% CI output of PF_{15} at $t = 130$ hours. A vertical black line in Fig. F.5 (a) shows when the running Spearman value is less than the requirement of 0.7. At this time, the trend is determined not suitable for RUL estimation after all, and its weight W_{15} is 0 as given by (F.37).

R_{14} and R_{12} passes the criteria for RUL estimation at $t \approx 108$ hours, and the initial output of the PFs are shown in Figs. F.5 (d) and (g), respectively. R_{14} and R_{12} are both similar to R_{15}, and gets a low Spearman coefficient eventually as indicated by black-stapled vertical lines due to the non-monotonic behavior of the trend. The parameter μ_α in PF_{14} and PF_{12} also converge as indicated in Figs. F.5 (b) and (e), respectively. Therefore, the converged PF output trends will continue until the Spearman coefficient

values go below 0.7. The converged output of PF_{14} and PF_{12} at $t = 130$ hours are shown in Fig. F.5 (f) and (i), respectively.

At $t = 113$ hours, PF_6 is initialized, and the initial PF output is shown in Fig. F.6 (a). R_6 has less oscillations compared to the three previously identified RMS trends, and therefore the Spearman coefficient never gets below 0.7. The parameter μ_α quickly converges as shown in Fig. F.6 (b). This is because the sudden increase in value of R_6 at $t \approx 117$ hours gives particles with high μ_α value a large weight. The predicted output is shown in Fig. F.6 (c) at $t = 130$ hours. The converged output should reach the FT early at $t \approx 150$, but the PF is continually updated on new samples. Given the uncertainty of bearing load, the PF may follow the future measurement samples.

When $t = 124$ hours, PF_3 is initialized, and R_3 together with the PF output are shown in Fig. F.6 (d). R_3 is the most monotonic trend of the five, and the initial PF output median matches the future samples well. At $t = 130$ hours, μ_α starts to converge as seen in Fig. F.6 (e), and the PF output is shown in Fig. F.6 (f). The model prediction at this point is directed at the end of lifetime, and therefore the RUL is accurately estimated at this point.

The weighed RUL mean and 95% CI for this dataset are shown in Fig. F.7 (a). Here, the true RUL is shown as a black solid line, while the weighted mean and 95% CI are the red and red-stapled lines, respectively. The weighted mean oscillates around the true RUL until the end of life. At $t = 125$ hours, all five PFs have a weight as shown in Fig. F.7 (b). From $t = 130$ hours, the estimated RUL is close to the true RUL. The reason for this is as follows. The median PF_{12} estimate in Fig. F.5 (i) over-estimates the RUL. In addition, the median of PF_6 underestimates the RUL while PF_3 matches the RUL, as indicated in Figs. F.6 (c) and (f). The weighted mean will therefore fall somewhere between these three outcomes. Afterwards, the weights W_{15}, W_{14} and W_{12} decreases towards 0 due to a low Spearman coefficient, and R_6 and R_3 are the only two trends left near the end of life. During this time, the estimated RUL is very close to the true RUL.

To visualize how the weighted RUL PDF changes over time, Fig. F.8 shows a 3D plot where the z-axis is the smoothened PDF. This plot shows why the weighted mean is calculated in comparison to the median. The true RUL is mainly situated between large peaks, while the discretely calculated median would have been situated under either peak with the highest weight. Therefore, it is more natural to estimate the RUL as the center between PFs.

F.5.2 Test rig 2 — in-house setup

Data from the second test rig is used to further validate the performance of the proposed method. A 6008-type roller element bearing is worn naturally in an accelerated life test

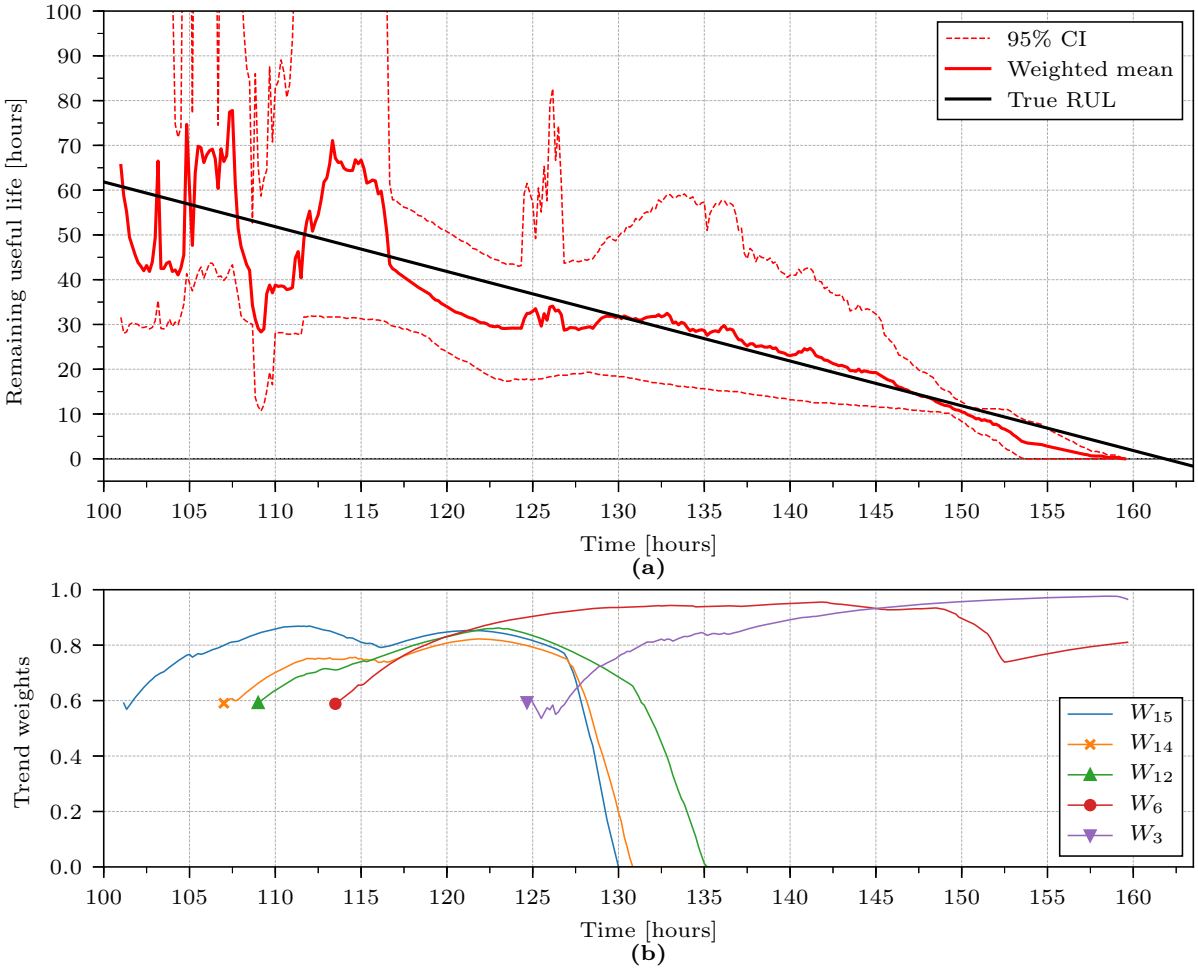

Figure F.7: Weighted RUL of the IMS dataset. (a) weighted mean, 95% CI and true RUL of the IMS dataset; (b) weights for each PF output.

by applying radial and axial loads. The dynamic capacity of the bearing is 17.8 kN, and the static capacity is 11 kN. With constant radial and axial loads of 9 kN and 5 kN, respectively, the bearing lasted approximately 34.6 million revolutions before failing due to an outer race fault. At a shaft speed of 100 rpm, vibration data was sampled every 15 minutes at a rate of 51200 Hz for 6 seconds. The 154 last hours of operation are used to verify the performance of the proposed method. More details of the test rig are given in [30].

The velocity- and acceleration-based RMS are shown in Figs. F.9 (a) and (b), respectively. A 1.1 kW motor is used, and therefore the velocity-based FT is set to $\hat{R}_v = 4.0$ mm/s. The FT $\hat{R}_v$ is not reached entirely at the end of useful life, as shown in Fig. F.9 (a). On the other hand, the transformed acceleration-based FT $\hat{R}_a$ is far from reached in Fig. F.9 (b). The transformation from $\hat{R}_v$ to $\hat{R}_a$ is not analytic and may therefore not be accurate. The test was stopped due to a high rate of change in R_a, and the machine could possibly have been run for more cycles. If the test was run for a few more measurement

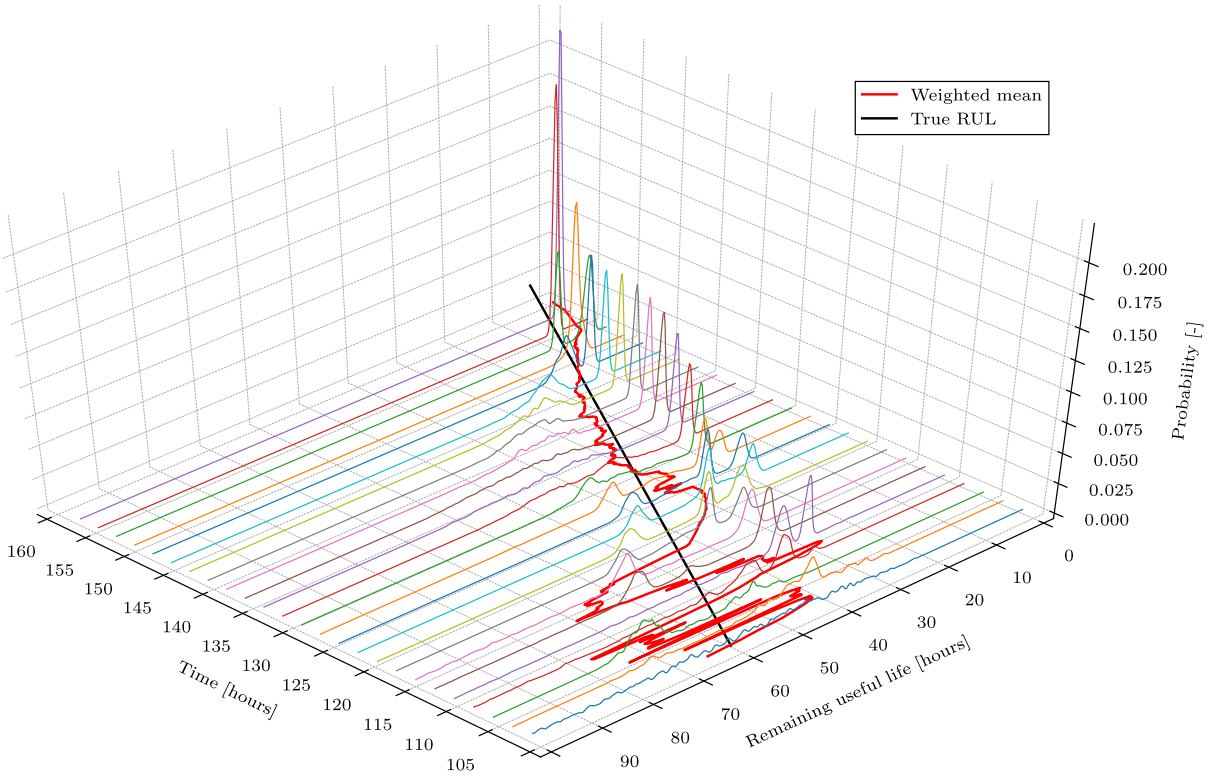

Figure F.8: Weighted RUL PDF of the IMS dataset over time. The weighted mean and true RUL are shown for reference.

cycles, $\hat{R}_a$ might have been reached.

The proposed method is used to split the vibration signal into $N_b = 64$ frequency bands, resulting in a frequency bandwidth of 400 Hz. A few of the RMS trends are shown in Fig. F.10. R_{40} in Fig. F.10 (a) is oscillating, and hence not suitable for RUL estimation. R_{19} and R_{11} in Figs. F.10 (b) and (c) have less oscillations, but do not increase steadily. R_1 in Fig. F.10 (d) on the other hand, increases almost linearly after the alarm is triggered, which makes the trend suitable for RUL estimation.

Using the proposed RUL estimation algorithm, the first and only identified RMS band is R_1 at $t = 120$ hours. The initial PF output and R_1 are shown in Fig. F.11 (a), and μ_α is shown in Fig. F.11 (b). The trend is re-drawn in Fig. F.11 (c) to show the PF output at $t = 130$ hours, which follows the future samples well. Since the predicted trend is similar to new samples, the estimated RUL is accurate. The weighted RUL is shown in Fig. F.12, and it follows the true RUL well. In this case, the estimated FT is reached very close to the actual end of life, and therefore the RUL estimation is accurate. The 3D plot in Fig. F.13 also shows how the weighted RUL PDF changes over time.

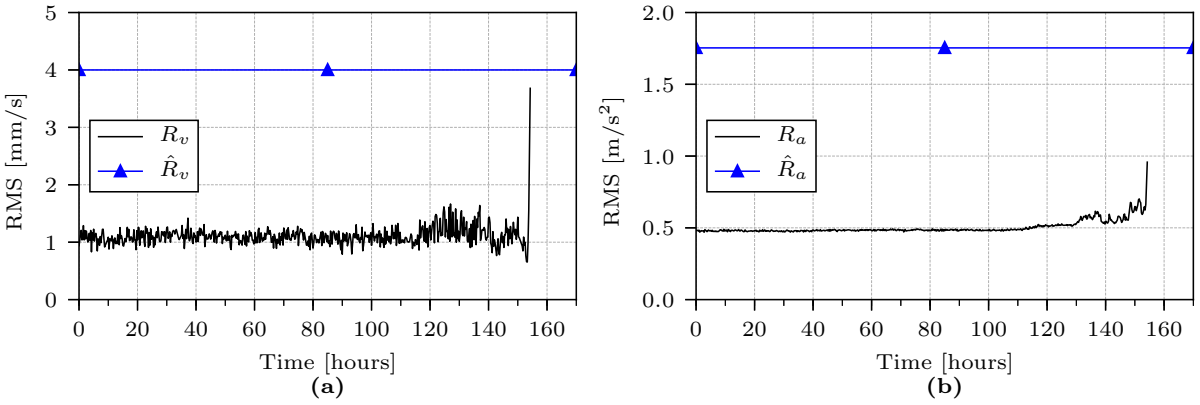

Figure F.9: Comparison of velocity- and acceleration-based RMS of the in-house test rig dataset. (a) velocity-based RMS; (b) acceleration-based RMS.

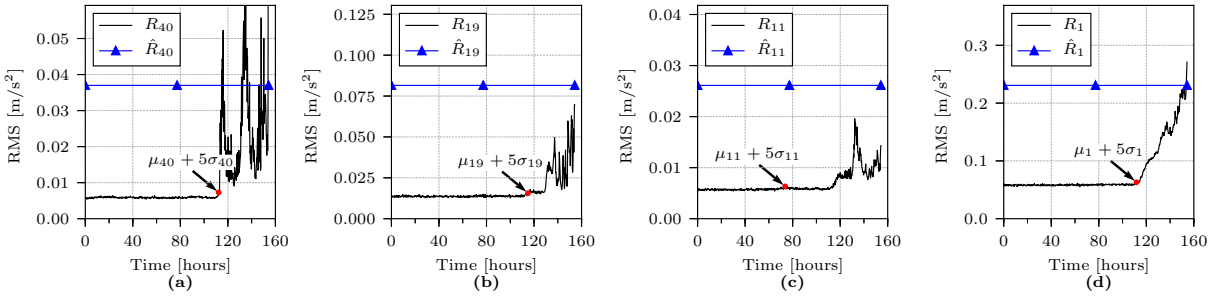

Figure F.10: Collection of RMS trends with FTs for the in-house test rig dataset. (a) R_{40}; (b) R_{19}; (c) R_{11}; (d) R_1.

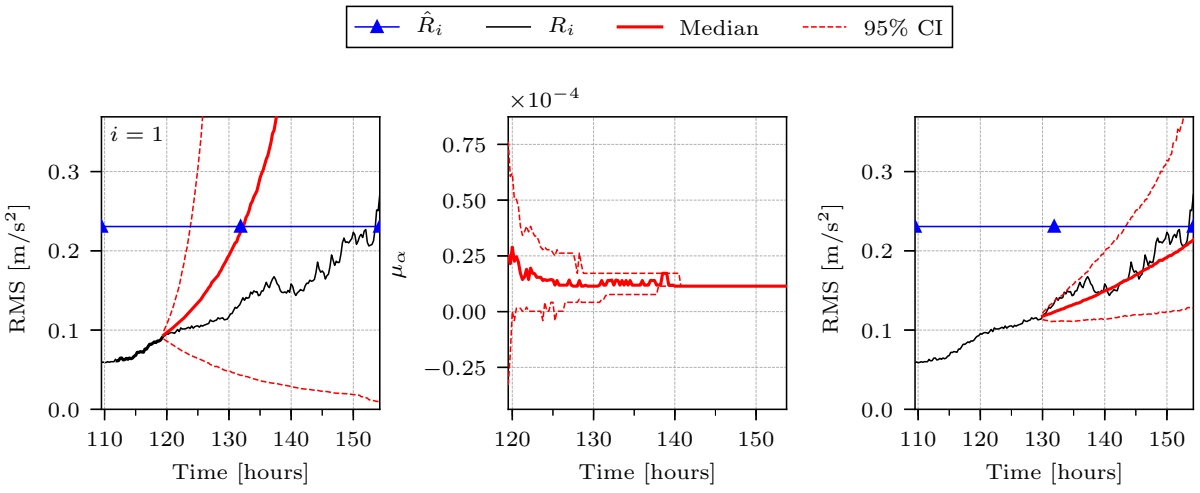

Figure F.11: R_1 and initial median and 95% CI output of the corresponding PF. (a) R_1 with marked part as optimization input, blue line for FT, and red lines for median and 95% CI of PF; (b) α over time for the initiated PF.

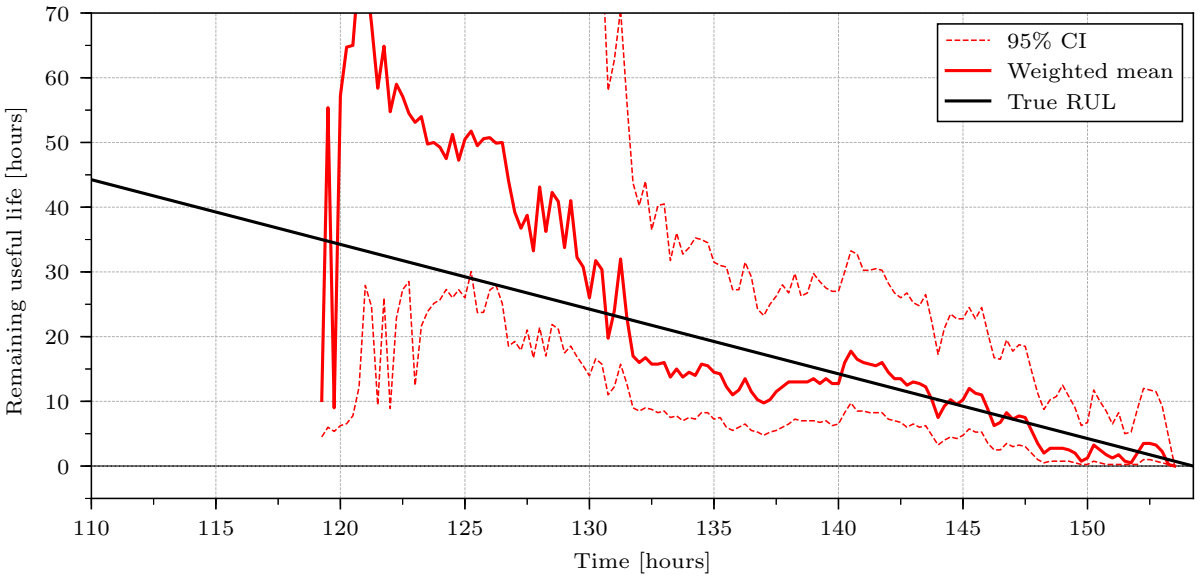

Figure F.12: Weighed RUL of the in-house test rig dataset. The weighted mean, 95% CI and true RUL are shown.

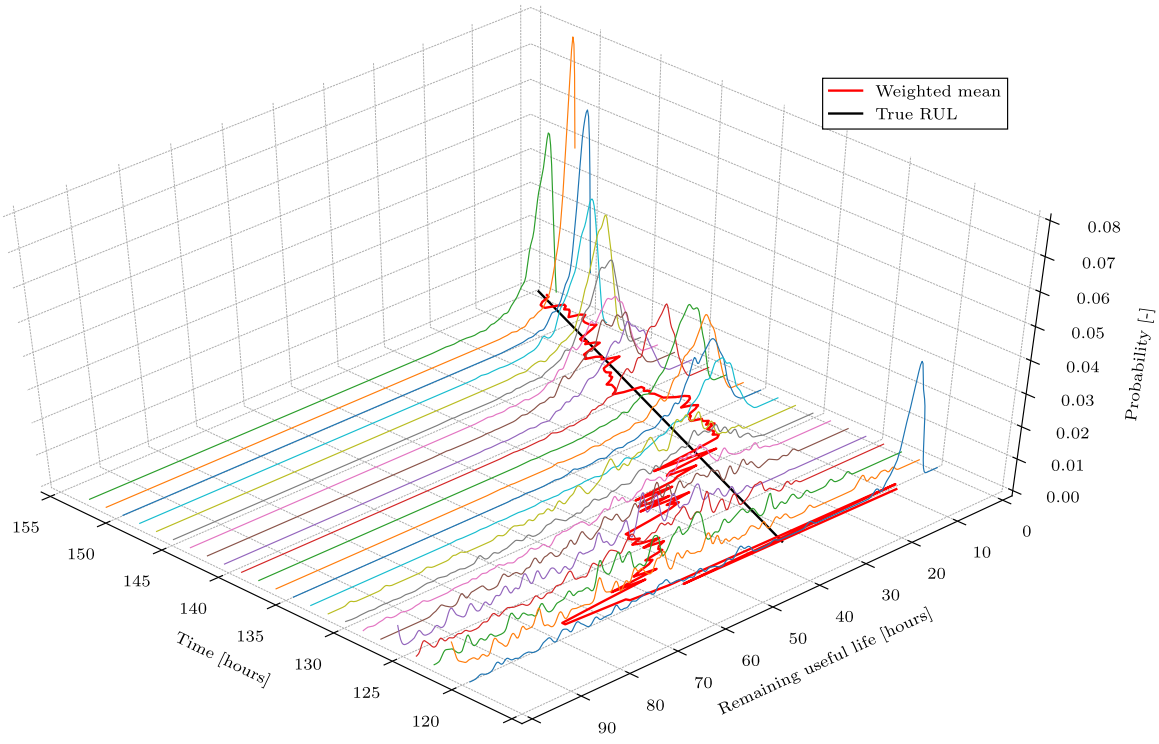

Figure F.13: Weighted RUL PDF of the in-house test rig dataset over time. The weighted mean and true RUL are shown for reference.

F.5.3 Comparisons

In this section, other research using the same IMS dataset are compared with the proposed method. In [31], the RUL was estimated using a feedforward artificial neural network (FFNN). However, training such an FFNN requires historic failure data, and therefore

the results cannot be compared to the proposed method. In [32], the proportional hazard model and logistic regression model were used, but historic failure data of the machine was required as well. Relevance vector machine (RVM) was used in [33], and the results show good estimation of the RUL. But, historical failure data was required to train the RVM. An enhanced phase space warping (PSW) method is proposed in [34] to combine the advantages of physics-based and data-driven techniques. The estimated RUL is close to the actual RUL, but historical failure data from a different test on the same machine was used to determine a few parameters. Soualhi et al. [35] proposed a method using an artificial ant clustering (AAC) technique for classifying faults, a hidden Markov model (HMM) to detect changes in degradation stage, and an adaptive neuro-fuzzy inference system (ANFIS) for RUL estimation. However, historical failure data was necessary to train the AAC. Ahmad et al. [36] presented a hybrid technique for RUL estimation that rectifies RMS fluctuations, and uses least squares minimization to fit a quadratic model to the rectified RMS. The gradient of the quadratic model is used as FT, and it was set based on available historical failure data. The resulting RUL estimation is quite accurate. However, due to usage of historic failure data to set the FT, the results cannot be compared to the proposed method. The authors could not identify publications showcasing RUL estimation on the IMS dataset without use of historic failure data. Therefore, direct RUL estimation comparison is not performed.

F.6 Further work

This section presents opportunities for further work to improve the proposed method. High-frequency RMS bands over 2 kHz may be used for health stage (HS) division. Energy is required to form cracks in materials, and that energy is dissipated during fracture, i.e. strain energy release. It has been observed that the RMS of a high frequency band seem to mimic this physical phenomenon. This information may be used to divide the life of the bearing into multiple health stages (HSs). In the IMS dataset, R_{15} is a high-frequency RMS band that shows signs of energy oscillations, seemingly due to crack formation. Fig. F.14 (a) shows the result of HS division due to strain energy release. The initial degradation alarm is first triggered for R_{15}, which marks the transition to HS 2. The RMS is increasing slowly over time, until R_{15} jumps to a high value ≈ 1 m/s^2. Now the RMS is seemingly high enough for more damage to occur, and during the fracture, R_{15} starts to decrease, which marks the change to a new HS. With the increased amount of damage, bearing impacts also increase vibration energy at lower frequency RMS bands. R_3 in Fig. F.14 (b) is a low-frequency band, and the RMS starts to increase linearly during HS 3. At $t = 130$ hours, R_{15} has reached a local minimum before it starts to increase again and

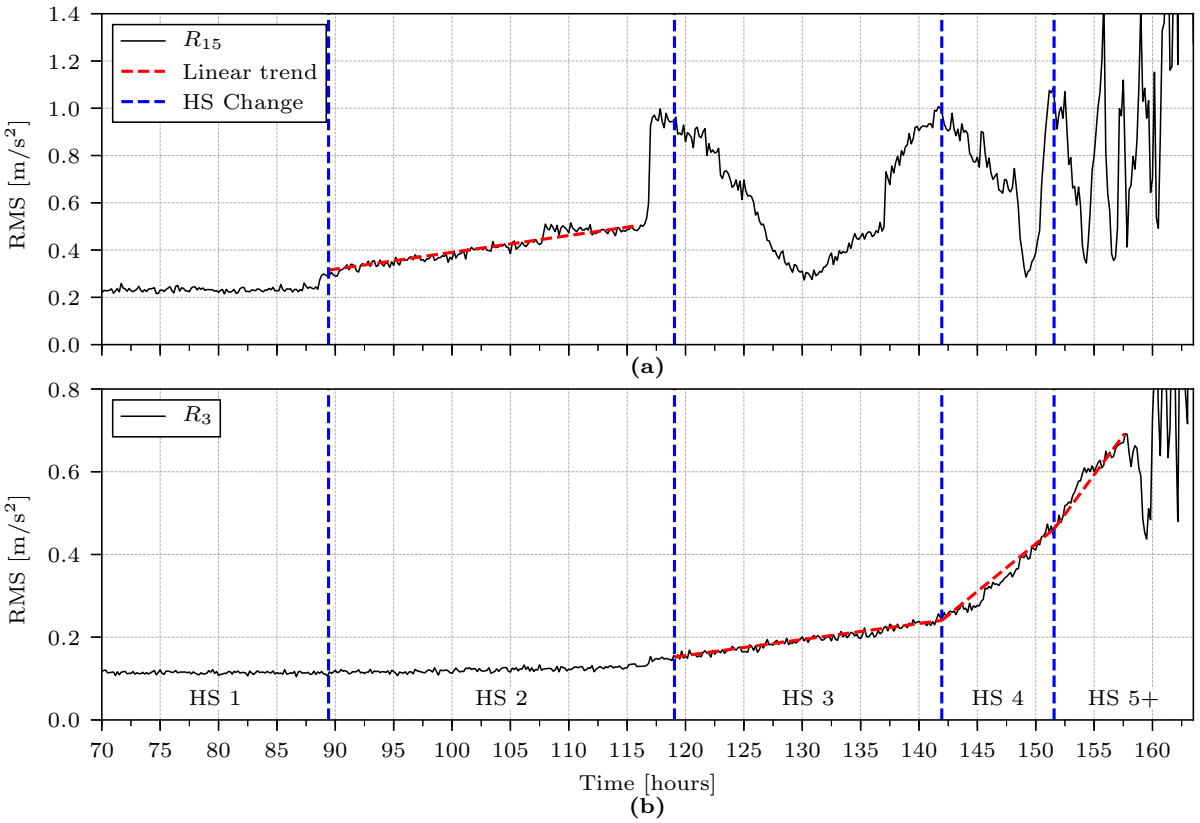

Figure F.14: Health stage division from energy cycles of a high-frequency band on the IMS dataset. (a) High-frequency band R_{15} and identified HS transitions; (b) low-frequency band R_3.

reaches another maximum at $t = 143$ hours. The rate of R_3 increases again afterwards, as there are seemingly more bearing impacts. In HS 4, R_{15} decreases and increases within 10 hours, and the bearing enters yet another HS. From HS 5+, the energy cycles are very short, just a few hours, indicating that critical failure is imminent.

For the second test rig, the RMS in a high-frequency band could indicate HS changes. Fig. F.15 (a) shows R_{40} and the identified peaks at the vertical stapled lines. At first, R_{40} increases until it reaches a top value, after which point it decreases again. This signifies an HS change, and the low-frequency R_1 in Fig. F.15 (b) increases linearly from that point. When then next RMS top point is reached at $t = 135$ hours, the rate of change in R_1 does not increase. Instead, the linear trend stops to increase for 7 hours, until it starts again at $t = 142$ hours. At this point, R_{40} reaches a small peak point. This phenomenon may be investigated in future work, as this information could be used to make a better RUL estimation. For example, the energy oscillation in a high-frequency band can be predicted to assume a higher rate of lower-frequency RMS trends ahead of time.

The RMS filter bank has been used to estimate RUL by identifying monotonic RMS trends, and updating model parameters using a particle filter. However, the presented

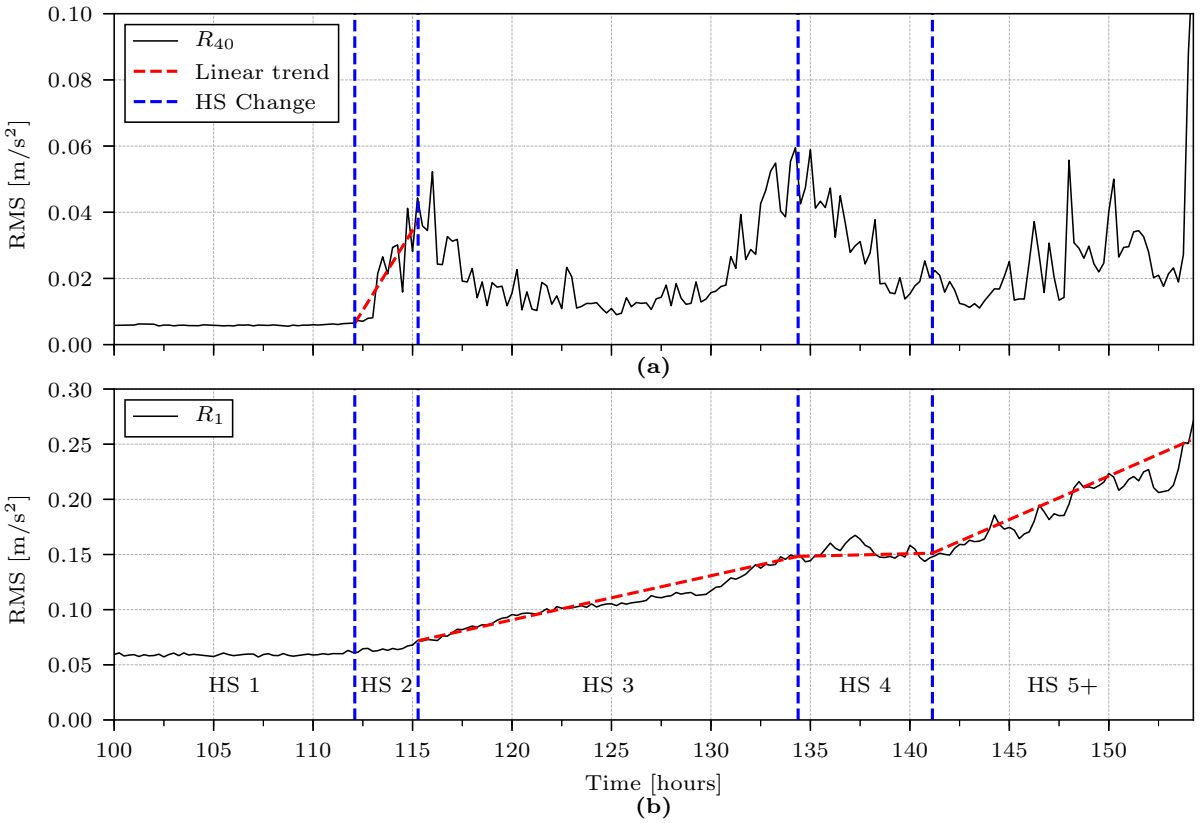

Figure F.15: Health stage division from energy cycles of a high-frequency band on the in-house test rig dataset. (a) High-frequency band R_{40} and identified HS transitions; (b) low-frequency band R_1.

method may be improved in future work. The frequency bands were evenly split with equal frequency bandwidth, and it should be investigated if there are more optimal ways to subdivide the energy bands. Secondly, it should be checked if the presented filter bank could be used for other purposes than calculating the RMS. Finally, there may be better ways to use all the RMS bands for RUL estimation rather than filtering each trend with a PF.

F.7 Conclusions

In this paper, a new method for subdividing the vibration signal into multiple frequency bands for root mean square (RMS) calculations is proposed. The method utilizes a single discrete Fourier transform (DFT) per signal, and individual bins are used to acquire the signal energy within a frequency band. The Spearman coefficient is used to identify monotonic RMS trends that are suitable for remaining useful life (RUL) estimation. It is observed that low-frequency RMS bands are most monotonic, while higher frequency RMS bands show earlier sign of degradation. The failure threshold (FT) for vibration

RMS, developed in earlier research, has been extended for the RMS frequency bands. A particle filter (PF) is applied to estimate parameters for the Paris-Erdogan law, and a single instance is used for every monotonic RMS trend. The RUL is afterwards estimated by weighing the output RUL of all initiated PFs. Experimental results show that the proposed method produces good RUL estimations without the use of historic failure data. The resulting RUL estimation in this research is, however, less accurate compared to referenced work that uses historical failure data for model training. Therefore, more investigation on the RUL estimation algorithm is necessary to improve performance.

REFERENCES

[1] P. Zhang, Y. Du, T. G. Habetler, and B. Lu. A survey of condition monitoring and protection methods for medium-voltage induction motors. *IEEE Transactions on Industry Applications*, 47(1):34–46, 2011. doi:10.1109/tia.2010.2090839.

[2] P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96(1):69–82, 1984. doi:10.1016/0022-460X(84)90595-9.

[3] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 104:799–834, 2018. doi:10.1016/j.ymssp.2017.11.016.

[4] R. K. Singleton, E. G. Strangas, and S. Aviyente. The use of bearing currents and vibrations in lifetime estimation of bearings. *IEEE Transactions on Industrial Informatics*, 13(3):1301–1309, 2017. doi:10.1109/tii.2016.2643693.

[5] C. Hu, B. D. Youn, P. Wang, and J. T. Yoon. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, 103:120–135, 2012. doi:10.1016/j.ress.2012.03.008.

[6] N. Li, Y. Lei, J. Lin, and S. X. Ding. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Transactions on Industrial Electronics*, 62(12):7762–7773, 2015. doi:10.1109/TIE.2015.2455055.

[7] Y. Lei, N. Li, and J. Lin. A new method based on stochastic process models for machine remaining useful life prediction. *IEEE Transactions on Instrumentation and Measurement*, 65(12):2671–2684, 2016. doi:10.1109/TIM.2016.2601004.

[8] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran. Residual life predictions from vibration-based degradation signals: a neural network approach. *IEEE Transactions on industrial electronics*, 51(3):694–700, 2004. doi:10.1109/TIE.2004.824875.

[9] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K.-L. Tsui. A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Transactions on Industrial Informatics*, 12(3):924–932, 2016. doi:10.1109/TII.2016.2535368.

[10] C. Lu, J. Chen, R. Hong, Y. Feng, and Y. Li. Degradation trend estimation of slewing bearing based on LSSVM model. *Mechanical Systems and Signal Processing*, 76:353–366, 2016. doi:10.1016/j.ymssp.2016.02.031.

[11] W. Wang. A model to predict the residual life of rolling element bearings given monitored condition information to date. *IMA Journal of Management Mathematics*, 13(1):3–16, 2002. doi:10.1093/imaman/13.1.3.

[12] D. Wang and K.-L. Tsui. Statistical modeling of bearing degradation signals. *IEEE Transactions on Reliability*, 66(4):1331–1344, 2017. doi:10.1109/tr.2017.2739126.

[13] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, and J. Dybala. A model-based method for remaining useful life prediction of machinery. *IEEE Transactions on Reliability*, 65(3):1314–1326, 2016. doi:10.1109/TR.2016.2570568.

[14] R. K. Singleton, E. G. Strangas, and S. Aviyente. Extended Kalman filtering for remaining-useful-life estimation of bearings. *IEEE Transactions on Industrial Electronics*, 62(3):1781–1790, 2015. doi:10.1109/tie.2014.2336616.

[15] Y. Qian and R. Yan. Remaining useful life prediction of rolling bearings using an enhanced particle filter. *IEEE Transactions on Instrumentation and Measurement*, 64(10):2696–2707, 2015. doi:10.1109/tim.2015.2427891.

[16] P. Boškoski, M. Gašperin, D. Petelin, and Đ. Juričić. Bearing fault prognostics using Rényi entropy based features and Gaussian process models. *Mechanical Systems and Signal Processing*, 52:327–337, 2015. doi:10.1016/j.ymssp.2014.07.011.

[17] S. A. Khan, A. E. Prosvirin, and J.-M. Kim. Towards bearing health prognosis using generative adversarial networks: Modeling bearing degradation. In *International Conference on Advancements in Computational Sciences (ICACS)*, pages 1–6. IEEE, 2018. doi:10.1109/ICACS.2018.8333495.

[18] J. Zhu, N. Chen, and W. Peng. Estimation of Bearing Remaining Useful Life based on Multiscale Convolutional Neural Network. *IEEE Transactions on Industrial Electronics*, 66(4):3208–3216, 2018. doi:10.1109/tie.2018.2844856.

[19] A. Klausen, H. Van Khang, and K. G. Robbersmyr. Novel Threshold Calculations for Remaining Useful Lifetime Estimation of Rolling Element Bearings. In *2018 XIII*

International Conference on Electrical Machines (ICEM), pages 1912–1918. IEEE, 2018. doi:10.1109/ICELMACH.2018.8507056.

[20] ISO. Mechanical vibration—Evaluation of machine vibration by measurements on non-rotating parts—Part 3 (Standard No. 10816-3). *ISO*, 1998.

[21] H. Qiu, J. Lee, J. Lin, and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006. doi:10.1016/j.jsv.2005.03.007.

[22] B. Duong, S. Khan, D. Shon, K. Im, J. Park, D.-S. Lim, B. Jang, and J.-M. Kim. A Reliable Health Indicator for Fault Prognosis of Bearings. *Sensors*, 18(11):3740, 2018. doi:10.3390/s18113740.

[23] C. Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101, 1904. doi:10.1093/ije/dyq191.

[24] P. Paris and F. Erdogan. A critical analysis of crack propagation laws. *Journal of basic engineering*, 85(4):528–533, 1963. doi:10.1115/1.3656900.

[25] P. Rycerz, A. Olver, and A. Kadiric. Propagation of surface initiated rolling contact fatigue cracks in bearing steel. *International Journal of Fatigue*, 97:29–38, 2017. doi:10.1016/j.ijfatigue.2016.12.004.

[26] D. An, J.-H. Choi, and N. H. Kim. Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab. *Reliability Engineering & System Safety*, 115:161–169, 2013. doi:10.1016/j.ress.2013.02.019.

[27] M. A. Branch, T. F. Coleman, and Y. Li. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing*, 21(1):1–23, 1999. doi:10.1137/S1064827595289108.

[28] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002. doi:10.1109/78.978374.

[29] J. D. Hol, T. B. Schon, and F. Gustafsson. On resampling algorithms for particle filters. In *IEEE Nonlinear Statistical Signal Processing Workshop*, pages 79–82. IEEE, 2006. doi:10.1109/nsspw.2006.4378824.

[30] A. Klausen, R. W. Folgerø, K. G. Robbersmyr, and H. R. Karimi. Accelerated Bearing Life-time Test Rig Development for Low Speed Data Acquisition. *Modeling, Identification and Control*, 38(3):143–156, 2017. doi:10.4173/mic.2017.3.4.

[31] A. K. Mahamad, S. Saon, and T. Hiyama. Predicting remaining useful life of rotating machinery based artificial neural network. *Computers & Mathematics with Applications*, 60(4):1078–1087, 2010. doi:10.1016/j.camwa.2010.03.065.

[32] H. Liao, W. Zhao, and H. Guo. Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model. In *Annual Reliability and Maintainability Symposium, RAMS'06*, pages 127–132. IEEE, 2006. doi:10.1109/RAMS.2006.1677362.

[33] A. Widodo and B.-S. Yang. Application of relevance vector machine and survival probability to machine degradation assessment. *Expert Systems with Applications*, 38(3):2592–2599, 2011. doi:10.1016/j.eswa.2010.08.049.

[34] Y. Qian, R. Yan, and R. X. Gao. A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mechanical Systems and Signal Processing*, 83:549–567, 2017. doi:10.1016/j.ymssp.2016.06.031.

[35] A. Soualhi, H. Razik, G. Clerc, and D. D. Doan. Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system. *IEEE Transactions on Industrial Electronics*, 61(6):2864–2874, 2014. doi:10.1109/tie.2013.2274415.

[36] W. Ahmad, S. A. Khan, and J.-M. Kim. A hybrid prognostics technique for rolling element bearings using adaptive predictive models. *IEEE Transactions on Industrial Electronics*, 65(2):1577–1584, 2018. doi:10.1109/TIE.2017.2733487.