
Electronic Thesis and Dissertation Repository

6-3-2024 11:00 AM

Data-Driven Vibration-Based Condition Monitoring: Fundamentals, Applications, and Challenges

Sulaiman A. S. Aburakhia, *The University of Western Ontario*

Supervisor: Shami, Abdallah, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Electrical and Computer Engineering

© Sulaiman A. S. Aburakhia 2024

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Signal Processing Commons](#)

Recommended Citation

Aburakhia, Sulaiman A. S., "Data-Driven Vibration-Based Condition Monitoring: Fundamentals, Applications, and Challenges" (2024). *Electronic Thesis and Dissertation Repository*. 10109. <https://ir.lib.uwo.ca/etd/10109>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Vibration-Based Condition Monitoring (VBCM) is commonly utilized in Prognostics and Health Management (PHM) due to its non-destructive nature and inherent advantages over alternative forms of condition monitoring. Furthermore, the rapid evolution of sensor fabrication and the rise of the Internet of Things (IoT) have facilitated large-scale VBCM systems across diverse domains, including industry, transportation, healthcare, agriculture, and wildlife monitoring. The recent advancements in computing technologies have significantly expanded the potential for VBCM by leveraging the synergy between signal processing and Machine Learning (ML). Accordingly, data-driven VBCM has emerged as a paradigm shift, improving the performance and reliability of VBCM systems. To this end, addressing various attributes of data-driven VBCM becomes increasingly important since it represents the core of current and future VBCM systems. The work presented in this thesis addresses the main aspects of VBCM, including signal processing fundamentals, feature extraction, availability of labeled data, computational complexity, and power efficiency. The methods employed in this thesis span the fields of Digital Signal Processing (DSP) and ML techniques (supervised, Deep Learning (DL)), including signal preprocessing, signal denoising, signal frequency-domain analysis, signal time-frequency domain analysis, feature extraction, signal companding (compression-expansion), and 1-dimensional (1D) convolutional reconstruction autoencoders. These methods address extraction of effective condition-related features, limited availability of labeled data, noise removal, complexity considerations in VBCM systems, and power efficiency of power-constrained sensor nodes in remote VBCM. By addressing the aforementioned problems, the end-to-end performance of VBCM systems can be improved in terms of the size of training data, the reliability of the monitoring process, system delay, memory requirements, and power consumption. To ensure the explainability of the extracted features, the developed methods for the extraction of condition-related features are based on signal processing since feature engineering using signal processing creates explainable features that link meaningfully to

signal conditions or classes compared to DL-based features. The thesis also contributes to the VBCM literature by providing a comprehensive tutorial on signal processing fundamentals, an overview of a typical signal-based ML pipeline, and an application-independent review of feature extraction techniques. The work presented in this thesis presents efficient solutions to the main challenges that face the practical deployment of real-world VBCM systems.

Keywords Prognostics and health management, vibration-based condition monitoring, predictive maintenance, digital signal processing, feature extraction, signal denoising, system delay, signal companding, power efficiency, vibration analysis, peak-to-average power ratio, nonlinear power amplification, wireless sensor networks, machine learning

Summary for Lay Audience

The ongoing technological transformation of data-driven Vibration-Based Condition Monitoring (VBCM) has enormous potential for consumers and businesses. This transformation is centered around integrating Digital Signal Processing (DSP) with Machine Learning (ML) models to facilitate reliable and efficient automated VBCM applications. VBCM utilizes vibration signals generated by various systems to monitor their integrity and predict any abnormal behavior within these systems. For instance, VBCM is commonly adopted in industrial environments where maintenance requirements are predicted based on the machine's or equipment's current condition, ensuring a safe working environment, enhancing productivity, and eliminating costly corrective and preventive maintenance actions. Data-driven VBCM is achieved by training ML models on historical vibration measurement data to learn healthy and abnormal operational conditions. However, despite its advantages, the practical deployment of data-driven VBCM systems faces significant challenges, such as the availability of proper historical data, computational complexity, presence of measurement noise, and high power consumption in vibration sensor nodes. Motivated by the need to develop effective solutions to overcome these challenges, the work presented in this thesis addresses each of the above-mentioned challenges in an effort to accelerate the deployment of data-driven VBCM across various fields and enhance their performance. Accordingly, the work presented in this thesis leverages the advancements of DSP and ML to facilitate reliable and practical data-driven VBCM applications. Specifically, the thesis introduces a similarity-based algorithm along with its open-source software implementation that performs VBCM of rotating machinery using very limited labeled historical vibration data. The software offers a practical solution compatible with other open-source libraries, making it ready for integration within various applications. Additionally, the thesis addresses computational burden and monitoring delay in VBCM systems. These two aspects are crucial for the real-world deployment of VBCM and directly impact safety and financial costs. Specifically, a higher computation burden increases the memory require-

ments, and a long delay in condition prediction could not prevent costly catastrophic failures. Furthermore, the thesis introduces a framework to facilitate power-efficient VBCM in power-constrained Wireless Sensor Networks (WSN). The framework proposes an innovative method to realize various signal processing operations in the framework using lightweight ML models that can be efficiently implemented on microcontrollers. Additionally, the thesis introduces a comprehensive tutorial on the fundamentals of signal processing, as well as a review of signal-based ML pipeline and feature extraction techniques. The aim is to highlight the crucial role of signal processing in VBCM and to bridge the gap between the two interdisciplinary fields of signal processing and ML by enhancing existing knowledge.

Co-Authorship Statement

- **Dr. Abdallah Shami** for his supervision during the work outlined in this thesis and expertise in all the defined research areas. Dr. Shami's contributions include the provision of resources, assistance in the conceptualization, refinement, and execution of the research problems and methodology, as well as the review and editing of the various works presented.
- **Tareq Tayeh** for his expertise in Deep learning (DL) models and his role in reviewing and editing the work outlined in Section 4.1.
- **Ryan Myers** for his industrial perspective on Predictive Maintenance (PdM) and his role in reviewing and editing the work outlined in Sections 4.1 and Chapter 5.
- **Dr. George K. Karagiannidis** for his role in reviewing and editing the work outlined in Chapter 3.
- **Dr. Ismail Hamieh** for reviewing and editing the work outlined in Chapter 6.

Epigraph

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

(وَقُلْ رَبِّ زِدْنِي عِلْمًا) سورة طه : الآية ١١٤ .

Dedication

To my mother and father

To my wife and to our children

To my brother and his family

In loving memory of my grandparents

Acknowledgements

In the Name of ALLAH, the Almighty God, the Most Beneficent, the Most Merciful. First and foremost, I praise ALLAH, Lord of the Worlds, for providing me with the strength, determination, and energy to complete this thesis. Without the grace of Almighty ALLAH, I could not reach this stage.

I would like to express my sincere thanks, appreciation, and gratitude to my supervisor, **Dr. Abdallah Shami**, for his unlimited support, mentorship, and guidance during this journey. I cannot give Dr. Shami the thanks he deserves for his contributions and his full support during the unforeseen circumstances and the obstacles I encountered during these four years. Thank you for everything, for offering me this opportunity, guidance, support, and understanding. I will forever be grateful for you.

I would like to thank the administrative staff at the Faculty of Engineering and Department of Electrical and Computer Engineering (ECE), especially ECE Graduate Coordinator Courtney Harper, for providing me with all the required help throughout this journey.

I would like to express my sincere thanks to Dr. George K. Karagiannidis for offering his expertise in reviewing and editing some parts of the presented work.

I would like to express my gratitude to Dr. Ismail Hamieh for reviewing and editing some parts of the presented work.

I would like to thank Ryan Myers from the National Research Council Canada for this collaboration and our thoughtful discussions.

I would like to express my sincere thanks to my dear friend and research partner, Tareq Tayeh, for his assistance and insightful recommendations.

I would like to express my deepest appreciation to my colleague Dr. Dimitrios Manias, for his help and support.

I would like to thank my former colleagues at the Optimized Computing and Communications (OC²), Dr. MohammadNoor Injadat, Dr. Abdallah Moubayed, Dr. Anas Saci, and Dr. Li Yang. I would like to extend my sincere thanks to my current colleagues, Ibrahim Shaer and Ali Chouman, for their cooperation and support.

I would like to thank my friends, Fuad Helaby, Ahmed Aburakhia, and Mohammed Jwaeid, for their continuous support during this journey.

To my brother, Dr. Walid Aburakhia, and his family, thank you for your encouragement and for always being there to support me.

To my wife, Somaya Alaraj, and our children, Salama, Janna, Ahmad, Mayar, and Mohammed, your boundless love and unwavering support have been the cornerstone of my journey. Your encouragement has been a constant source of strength, and I am deeply grateful for the patience, understanding, and support you have provided.

Finally, to my dear parents, **Ahmad Aburakhia** and **Hakema Aburakhia**, mere words fall short of expressing my immense gratitude for your unconditional love, endless support, and the sacrifices you have made for me. I hold an everlasting appreciation and love for you both.

Table of Contents

Abstract	ii
Summary for Lay Audience	iv
Co-Authorship Statement	vi
Epigraph	vii
Dedication	viii
Acknowledgements	ix
Table of Contents	xi
List of Tables	xviii
List of Figures	xx
List of Abbreviations	xxv
1 Introduction	1
2 Background and Motivation	8
2.1 Background	8
2.2 Motivation	11
2.2.1 Challenge 1	11
2.2.2 Challenge 2	12
2.2.3 Challenge 3	13
2.2.4 Challenges 4 and 5	16
2.3 Experimental Vibration Datasets	17

3 On the Intersection of Signal Processing and Machine Learning: A Use

Case-Driven Analysis Approach	19
3.1 Introduction	19
3.2 Introduction to Signals	24
3.2.1 Signal Characteristics	24
3.2.2 Signal Energy and Power	32
3.2.3 Signal Correlation	35
3.2.4 Signal Convolution	38
3.2.5 Signal Entropy	39
3.3 Signal Transformation and Analysis	42
3.3.1 Fourier Transform	43
3.3.2 Short-Time Fourier Transform	48
3.3.3 Power Spectral Density	55
3.3.4 Wavelet Transform	57
The Continuous Wavelet Transform (CWT)	60
Discrete Wavelet Transform (DWT)	61
Stationary Wavelet Transform (SWT)	66
Wavelet Packet Transform (WPT)	67
3.3.5 Hilbert Transform	70
3.3.6 Hilbert–Huang Transform	74
Empirical Mode Decomposition (EMD)	74
Hilbert Spectral Analysis	78
3.3.7 Wigner–Ville Distribution	81
3.3.8 Conclusive Comparison	85
3.4 A Typical Signal-Based ML Pipeline	86
3.4.1 Preprocessing	88
Signal Smoothing	88

	Signal Denoising	89
	Signal Segmentation	90
3.4.2	Processing	91
3.4.3	Application	92
3.5	Feature Extraction Methods	94
3.5.1	Time Domain-Based Methods	95
3.5.2	Transform-Based Methods	97
	Frequency-Based Methods	97
	Time-Frequency-Based Methods	99
3.5.3	Signal Decomposition-Based Methods	102
3.5.4	Envelope-Based Methods	105
3.6	Conclusion	106
4	Similarity-Based Framework for VBCM using Limited Labeled Data	108
4.1	Similarity-Based Predictive Maintenance Framework for Rotating Machinery	109
4.1.1	Introduction	110
4.1.2	Related Work	112
4.1.3	Methodology	113
4.1.4	Predictive Maintenance Framework for Condition Monitoring of Rotating Machinery	114
	Signal Denoising	115
	Feature Extraction	116
	Similarity Measure	116
4.1.5	Performance Evaluation	118
4.1.6	Results and Discussion	120
4.1.7	Conclusion	123
4.2	SB-PdM: a Tool for Predictive Maintenance of Rolling Bearings Based on Limited Labeled Data	124

4.2.1	Introduction	124
4.2.2	SB-PdM Key Aspects and code functionalities	126
	Key Aspects of SB-PdM Software	127
	SB-PdM code functionalities	128
4.2.3	Performance Evaluation	132
4.2.4	Software impacts	133
4.2.5	Conclusion	134

5 A Hybrid Method for Condition Monitoring and Fault Diagnosis of Rolling

	Bearings With Low System Delay	135
5.1	Introduction	135
5.2	Theoretical Background	139
	5.2.1 System Delay	139
	5.2.2 Wavelet Packet Transform	141
5.3	Literature Review And Related Work	142
5.4	Proposed method for WPT-FFT Features Extraction	145
	5.4.1 Selection of Base Wavelet and Decomposition Level	147
	5.4.2 Complexity Analysis	148
5.5	Performance Evaluation	149
	5.5.1 Experimental Setup	150
	CWRU Bearing Dataset	150
	PU Bearing Dataset	150
	University of Ottawa Bearing Dataset	151
5.6	Results and Discussion	152
	5.6.1 Classifier Selection	152
	5.6.2 Effectiveness of Proposed Features	158
	5.6.3 Experimental Results on CWRU Dataset	159
	5.6.4 Experimental Results on PU Dataset	161

5.6.5	Experimental Results on uOttawa Dataset	162
5.7	Conclusion	162
6	Joint Instantaneous Amplitude-Frequency Analysis of Vibration Signals for Vibration-Based Condition Monitoring	164
6.1	Introduction	164
6.2	Related Work	169
6.3	Joint Instantaneous Time-Frequency Analysis of Vibration Signals	171
6.3.1	The Hilbert Transform	171
6.3.2	Proposed Joint Instantaneous Time-Frequency Analysis	172
	Instantaneous Amplitude-Frequency Mapping	174
	Instantaneous Amplitude-Frequency Correlation	176
	Instantaneous Energy-Frequency Distribution	178
	Complexity Analysis	181
6.4	Performance Evaluation	183
6.4.1	Experimental Dataset	183
6.4.2	Experimental Setup	183
6.5	Results and Discussion	185
6.6	Conclusion	188
7	Signal-Companding AutoEncoder with Compression-Based Activation for an Efficient Vibration-Based Remote Condition Monitoring	190
7.1	Introduction	190
7.2	Background and Motivation	194
7.3	PAPR of Vibration Signals	197
7.3.1	Statistical Distribution of Vibration Signals	197
7.3.2	Statistical Analysis of the PAPR	199
7.4	Review of PAPR reduction techniques	202

7.5	Signal Companding for Reduction of PAPR	205
7.5.1	Conventional Signal Companding	205
7.5.2	Proposed Signal Companding	207
	Signal Compression at the Source	207
	Signal Denoising and Expansion at the Destination	210
7.6	Modeling Nonlinear Power Amplification	212
7.7	Performance Evaluation	213
7.7.1	Experimental Setup	213
	Training the Source Autoencoder	214
	Training the Destination Autoencoder	214
7.7.2	Performance Metrics	214
	PAPR Reduction	215
	Average Power of the Compressed Signal	215
	Nonlinear Distortion	215
	Power Efficiency	218
	Spectral Spreading	219
	Signal Denoising	220
7.8	Results and Discussion	220
7.8.1	Performance Results	222
	PAPR Reduction	222
	Average Power of the Compressed Signal	225
	Nonlinear Distortion	226
	Power Efficiency	228
	Spectral Spreading	228
	Signal Denoising	229
7.8.2	Reliability of Condition Monitoring	229

7.8.3	The Lightweight Nature of the Proposed Activation-Based Compression Autoencoder	232
7.9	Conclusion	233
8	Conclusion and Future Work	235
	Bibliography	239
	Curriculum Vitae	293

List of Tables

3.1	Comparison between STFT and PSD	57
3.2	Comparison between common signal analysis and transformation tools . . .	87
3.3	Inputs, functions, and outputs of each stage of typical signal-based ML framework	93
3.4	A comparison between deep learning-based and signal processing-based feature extraction	95
4.1	Bearing test data	119
4.2	Accuracy results of time, frequency, and time-frequency features	121
5.1	Experimental Datasets	153
5.2	Performance comparison of SVM, XGBoost, and RF classifiers on CWRU dataset: Input segment duration = 0.05 seconds (600 samples), ($k = 3, m = 1$)	154
5.3	Performance comparison between proposed features and other features on CWRU dataset: Input segment duration = 0.05 seconds (600 samples) . . .	154
5.4	Performance results of the proposed method on CWRU dataset	155
5.5	Mean Accuracy (%) on CWRU dataset	155
5.6	Performance comparison between the proposed method and other methods on CWRU dataset	156
5.7	Performance results of the proposed method on the PU dataset	156
5.8	Performance comparison between the proposed method and other methods on PU dataset	157
5.9	Performance results of the proposed method on the uOttawa dataset	157
5.10	Performance comparison between the proposed method and other methods on uOttawa dataset	158
6.1	Proposed envelope representations and extracted features	181

6.2	Experimental dataset	183
6.3	Performance comparison among the proposed method, the STFT-based method, and the HHT-based method	187
6.4	Performance comparison between the proposed method and other DL-based methods on the PU dataset	188
7.1	Comparison between the proposed method and conventional training	209
7.2	Extracted features from expanded vibration signals	230
7.3	Classification results of the operational conditions of the PU dataset	230
7.4	Results of the different training scenarios	230

List of Figures

2.1	Data-driven VBCM.	10
2.2	Typical signal processing-based ML pipeline.	12
2.3	Taxonomy of feature extraction techniques.	13
2.4	Similarity-based framework for VBCM using limited labeled data.	14
2.5	Proposed wavelet spectral energy for VBCM with low computational complexity.	15
2.6	Proposed envelop-based analysis for VBCM with low computational complexity.	15
2.7	remote VBCM-WSN.	16
2.8	Proposed framework of autoencoder-based signal companding and denoising.	17
3.1	Visual representation of chapter contents.	23
3.2	Waveform of a signal, $x(t)$, as it evolves over time t	25
3.3	Sinusoidal waveform $s(t)$	27
3.4	Magnitude of the sinusoidal waveform $s(t)$	28
3.5	$x[n]$, a discrete-time version of signal $x(t)$	30
3.6	Signal $x(t)$ and its autocorrelation.	36
3.7	A comparison between Shannon entropy values of a random vibration $v(t)$, and a square wave.	41
3.8	Composite sinusoidal signal $s_3(t)$	46
3.9	Frequency spectrum of $s_3(t)$	47
3.10	Composite sinusoidal signal $s_{m3}(t)$	49
3.11	Frequency spectrum of $s_{m3}(t)$	50
3.12	STFT spectrogram of the composite sinusoidal signal $s_3(t)$	53
3.13	STFT spectrogram of the composite sinusoidal signal $s_3(t)$	54
3.14	Periodogram and Welch's PSD estimations of vibration signal $v(t)$	56

3.15	Some of the commonly used wavelet functions.	59
3.16	CWT “scalogram” of the composite signal $s_3(t)$	61
3.17	7-level DWT decomposition tree.	64
3.18	7-level DWT decomposition of random vibration signal $v(t)$: Elementary modes and their corresponding frequency spectra.	65
3.19	3-level WPT decomposition tree.	67
3.20	3-level WPT decomposition of random vibration signal $v(t)$: Elementary modes and their corresponding frequency spectra.	68
3.21	Composite signal $c(t)$	72
3.22	Hilbert Transform of the signal $c(t)$	73
3.23	EMD of the random vibration signal $v(t)$: IMFs, residual, and their respective frequency spectra.	77
3.24	EMD of the composite sinusoidal signal $s_3(t)$: IMFs and residual alongside with their respective frequency spectra.	80
3.25	2D heatmap visualization of the HTT of $s_3(t)$	80
3.26	WVD of the composite sinusoidal signal $c(t)$	84
3.27	PWVD of the composite sinusoidal signal $c(t)$	84
3.28	Overview of a typical signal-based ML framework.	87
3.29	Taxonomy of feature extraction techniques.	93
4.1	similarity-based classification framework	113
4.2	(a) STFT and (b) FFT of normal (healthy) operational condition and a faulty operational condition (ball fault of 0.07 inches in diameter)	116
4.3	scatter plot shows (a) all classes and (b) normal and inner race faulty classes	122
4.4	A high-level overview of the SB-PdM software	126
5.1	System delay of vibration-based condition monitoring.	139

5.2	(a) Proposed feature extraction using 3-level WPT. (b) From top to bottom, vibration signal and its WPT-based elementary waveforms “on the left” and corresponding signal spectrum “on the right”.	142
5.3	Feature sample from each class of CWRU dataset with $T_{vin} = 0.05$ seconds, $k = 3$ and $m = 1$	154
5.4	Average system delay τ_d as a function of (a) duration of input segment T_{vin} and (b) size of features vector S	155
6.1	Instantaneous amplitude (IA) and Instantaneous frequency (IF) of a rolling bearing vibration signal, $v_h(t)$	173
6.2	Instantaneous amplitude-frequency mapping (IAFM), Envelope PSD, and envelope spectrum of the healthy vibration pattern, $v_h(t)$	174
6.3	IAFC of the healthy vibration pattern, $v_h(t)$	176
6.4	Joint time-energy-frequency representation of the healthy vibration pattern, $v_h(t)$	178
6.5	IEFD along with the joint time-energy-frequency visualization of the healthy vibration pattern, $v_h(t)$	180
6.6	IAFMs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.	185
6.7	IAFCs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.	186
6.8	IEFDs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.	186
7.1	Overview of a typical remote VBCM system.	195
7.2	HPA range curve.	196

7.3	Histograms of sample vibration signals: (a) Gaussian random vibration, (b) acceleration of a flying aircraft, (c) acceleration measurements of a flying UAV, (d) vibration from an SHM setup, (e) vibration generated by a wind turbine gearbox, and (f) vibration generated by rolling bearings.	198
7.4	Theoretical CCDF and its closed-form approximation for different values of N	200
7.5	Analytical and simulated CCDFs of (a) Gaussian random vibration ($N = 5000$), (b) acceleration of a flying aircraft ($N = 5000$), (c) acceleration measurements of a flying UAV ($N = 50$), (d) vibration from an SHM setup ($N = 5000$), (e) vibration generated by a wind turbine gearbox ($N = 5000$), and (f) vibration generated by rolling bearings ($N = 5000$).	200
7.6	μ -law: (a) compression profile with different values of μ , and (b) peak and average powers as a function of μ	206
7.7	Proposed framework for autoencoder-based companding of vibration signals.	206
7.8	AM-AM and AM-PM conversions of the Rapp SSPA model.	211
7.9	Structures of the reconstruction autoencoders used in the proposed framework.	211
7.10	Signal's amplitude constellation of the PU dataset.	215
7.11	(a) Original and compressed vibration signals and (b) CCDFs of original and compressed signals.	220
7.12	Average power ratios of original and compressed signals with respect to normalized average power of original signals.	221
7.13	Constellations of uncompressed and compressed signals along with HPA response curve.	221
7.14	Original and expanded vibration signals after passing through HPA.	222

7.15	Amplitude constellations after passing through the HPA: (a) uncompanded signals, b) μ -law companded, (c) proposed-companded signals, and (d) original signals of the PU dataset along with uncompressed and expanded signals.	223
7.16	EVM values of uncompressed and expanded signals.	224
7.17	Applied IBO, achieved EVM, and resultant HPA efficiency η in proposed-companding, μ -law companding, and no companding.	224
7.18	Normalized mean PSD plots of original signals, proposed-compressed signals, and μ -law compressed signals after passing through the HPA.	225
7.19	(a) a Denoised-expanded signal and its reference original signal. (b) Average SNR_d of denoised-expanded signals.	225

List of Abbreviations

AM	Amplitude Modulation
AMD	Adaptive Mode Decomposition
ANC	Adaptive Noise Cancellation
AV	Autonomous Vehicle
AWGN	Additive White Gaussian Noise
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CF	Crest Factor
CL	Compression Loss
CM	Corrective Maintenance
CNC	Computer Numerical Control
CNN	Convolutional Neural Network
CoV	Coefficient of Variation
CP	Correlation Peak
CWRU	Case Western Reserve University
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform

DL	Deep Learning
dp4	Daubechies4
DPSD	Discrete Power Spectral Density
DSP	Digital Signal Processing
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EEMD	Ensemble Empirical Mode Decomposition
EMD	Empirical Mode Decomposition
EVM	Error Vector Magnitude
EWT	Empirical Wavelet Transform
FFT	Fast Fourier Transform
GAF	Gramian Angular Field
GAN	Generative Adversarial Network
HHT	Hilbert–Huang Transform
HHT	Hilbert–Huang transform
HOS	High-Order Statistics
HPA	High Power Amplifier

HSA	Hilbert Spectral Analysis
HT	Hilbert Transform
IA	Instantaneous Amplitude
IACF	Instantaneous Autocorrelation Function
IAFC	Instantaneous Amplitude-Frequency Correlation
IAFM	Instantaneous Amplitude-Frequency Mapping
IDFT	Inverse Discrete Fourier Transform
IEFD	Instantaneous Energy-Frequency Distribution
IF	Instantaneous Frequency
IFFT	Inverse Fast Fourier Transform
IIoT	Industrial Internet of Things
IMF	Intrinsic Mode Function
IoT	Internet of Things
IP	Instantaneous Phase
IPB	Input Power Backoff
IR	Inner Raceway
LSTM	Long Short-Term Memory
LTI	Linear Time-Invariant

MAE	Mean Absolute Error
MAF	Moving Average Filter
MER	Mean-to-Entropy Ratio
MFCC	Mel-Frequency Cepstral Coefficient
ML	Machine Learning
MSE	Mean Squared Error
MTF	Markov Transition Field
OFDM	Orthogonal Frequency-Division Multiplexing
OR	Outer Raceway
PAPR	Peak-to-Average Power Ratio
PCI	Brain-Computer Interface
PDF	Probability Density Function
PdM	Predictive Maintenance
PHM	Prognostics and Health Management
PL	Peak Lag
PM	Preventive Maintenance
PSD	Power Spectral Density
PTS	Partial Transmission Sequence

PWVD	Pseudo-Wigner–Ville Distribution
ReLU	Rectified Linear Unit
RF	Random Forest
RMS	Root-Mean-Square
RNN	Recurrent Neural Network
SB-PdM	Similarity-Based Predictive Maintenance
SB-PdM	Similarity-Based Predictive Maintenance
SC	Spectral Centroid
SGF	Savitzky–Golay Filter
SHM	Structural Health Monitoring
SK	Spectral Kurtosis
SLM	Selective Mapping
SM	Smart Manufacturing
SNR	Signal-to-Noise Ratio
SS	Spectral Spread
SSIM	Structural Similarity Index Measure
SSIM	Structural Similarity Index Measure
SSM	Structural Similarity Measure

SSPA	Solid-State Power Amplifier
STFT	Short-Time Fourier Transform
SVM	Support-Vector Machine
SWT	Stationary Wavelet Transform
TinyML	Tiny Machine Learning
UAV	Unmanned Aerial Vehicle
VBCM	Vibration-Based Condition Monitoring
VMD	Variational Mode Decomposition
WPD	Wavelet Packet Decomposition
WPT	Wavelet Packet Transform
WSN	Wireless Sensor Network
WT	Wavelet Transform
WVD	Wigner–Ville Distribution
XGBoost	Gradient Boosting

Chapter 1

1 Introduction

Vibration-Based Condition Monitoring (VBCM) is a widely utilized method in Prognostics and Health Management (PHM) due to its non-invasive nature and apparent advantages compared to other types of condition monitoring. The advancements in sensor technology and the growth of the Internet of Things (IoT) have facilitated large-scale VBCM systems in various fields such as industry, transportation, healthcare, agriculture, and wildlife monitoring. Furthermore, the rapid advancements in computing technologies have significantly expanded the potential for VBCM by leveraging the synergy between Digital Signal Processing (DSP) and Machine Learning (ML). Accordingly, data-driven VBCM has emerged as a paradigm shift, improving the performance and reliability of VBCM systems. To this end, addressing various attributes of data-driven VBCM and the major challenges involved in the practical deployment of data-driven VBCM systems has become increasingly important.

This thesis presents four research problems that address fundamental aspects, practical deployment, and performance improvements of data-driven VBCM systems. The research tackles these problems from a signal-processing perspective, leveraging the synergy between DSP and ML fields to facilitate reliable and efficient solutions. Hence, the thesis, through the first research problem, starts with an attempt to bridge the existing knowledge gap between these two interdisciplinary fields. Despite many attempts in the existing literature to address this gap, the majority are limited to specific applications and mainly focus on feature extraction, often assuming extensive prior knowledge in signal processing. This assumption creates a significant obstacle for a wide range of professionals in the field, hindering the large-scale deployments of data-driven VBCM systems. To overcome these challenges, an integrated approach is adopted that begins by providing a detailed tutorial on the fundamentals of signal processing, thereby giving professionals in the field the neces-

sary background knowledge. Following this, the presented work explores the key stages of a typical signal processing-based ML pipeline and provides an in-depth review of feature extraction techniques. Differing from existing literature, this work offers an application-independent review and introduces a novel classification taxonomy for feature extraction techniques. Furthermore, the presented work encourages a collaborative research culture by making available a public repository containing relevant Python and MATLAB codes for various signal-processing techniques. This effort aims to support collaborative research endeavors and ensure the reproducibility of the results presented.

The second research problem addresses the limited availability of labeled training samples in real-world scenarios. In supervised data-driven VBCM, labeled samples that are large in size and diverse—in terms of class representation—should be available to serve the training purpose. In real-world situations, the available labeled samples are usually limited in their size due to expensive and time-consuming labeling processes. Moreover, faulty or abnormal samples are not abundant since abnormal events, such as faults or anomalies, are infrequent, making it challenging to gather substantial abnormal signal data. As a result, real-world abnormal samples are usually insufficient to represent the corresponding classes during the training phase. This problem is tackled by addressing the classification task as a similarity measure to a reference sample rather than a supervised classification task. Similarity-based approaches require a limited amount of labeled data, meeting real-world industrial applications' requirements. Accordingly, the thesis introduces a similarity-based framework for Predictive Maintenance (PdM) of rotating machinery. A reference vibration signal is generated and labeled according to the machine's operational condition for each operational state of the machine. Consequently, statistical time analysis, Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT), and Wavelet Packet Transform (WPT) are used to extract features from the captured vibration signals. For each feature type, three similarity metrics, namely Structural Similarity Measure (SSM), cosine similarity, and Euclidean distance, are used to measure the similarity between test signals and ref-

erence signals in the feature space. Hence, nine settings in terms of feature type-similarity measure combinations are evaluated. Experimental results confirm the effectiveness of similarity-based approaches in achieving very high accuracy with moderate computational requirements compared to ML-based methods.

The third research problem addresses the computational complexity of data-driven VBCM in terms of memory requirements and system monitoring delay. Generally, accuracy, monitoring delay, and memory requirements represent the main performance requirements in VBCM. Achieving high accuracy with low delay and relaxed memory requirements improves system reliability, prevents catastrophic failures, and reduces computational costs. Furthermore, delay is crucial to remote condition monitoring and time-sensitive applications. While most existing methods focus on accuracy, slight attention has been paid to addressing the delay introduced in the condition monitoring process and memory requirements. Aiming to bridge this gap, the thesis proposes two methods to facilitate low-complexity VBCM. The first method proposes a hybrid method for VBCM of rolling bearings that outperforms existing methods in accuracy, system delay, and memory requirements. Specifically, the work addresses the overall delay in VBCM systems and introduces the concept of system delay to assess it. Consequently, Wavelet Packet Transform (WPT) and Fourier analysis are used to decompose short-duration input segments of the vibration signal into elementary waveforms and obtain their spectral contents, respectively. Accordingly, energy concentration in the spectral components—caused by defect-induced transient vibrations—is utilized to extract a small number of features with high discriminative capabilities. The experimental results show that the proposed method can achieve high accuracy with low system delay and moderate memory requirements. The second proposed method is based on envelope analysis of vibration signals. Specifically, the instantaneous amplitude “envelope” and instantaneous frequency of the vibration signal are jointly utilized to facilitate three novel envelope representations: Instantaneous Amplitude-Frequency Mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and

Instantaneous Energy-Frequency Distribution (IEFD). Maintaining temporal information, these representations effectively capture energy-frequency variations that are unique to the condition of the bearing, thereby enabling the extraction of discriminative features with high sensitivity to variations in operational conditions. Accordingly, a set of six new highly discriminative features are engineered from these representations, capturing and characterizing their shapes. The experimental results demonstrate the effectiveness of the proposed method in detecting and diagnosing various fault types. Moreover, the proposed method has moderate computational complexity, meeting the requirements of real-time applications.

The fourth research problem pertains to the power consumption in sensor nodes and the presence of noise in Wireless Sensor Networks (WSN). These sensor nodes are typically constrained by limited power resources, necessitating control over the Peak-to-Average Power Ratio (PAPR) of the acquired vibration signal. Effective PAPR control is essential to prevent nonlinear distortion and reduce power consumption within the node. Further, preserving the vibration signal's waveform and mitigating the impact of noise are crucial for reliable condition monitoring. The work presented in this thesis addresses this problem through two aspects: First, it provides an in-depth analysis of the PAPR of vibration signals in VBCM systems and evaluates, in the presence of nonlinear power amplification, the impact of PAPR and nonlinear distortion on power efficiency and system performance. Secondly, a lightweight autoencoder-based signal companding scheme aimed at enhancing power efficiency and improving performance is proposed to control the PAPR and mitigate the impact of nonlinear distortion. In the sensor node, the proposed method employs a lightweight reconstruction autoencoder with a compression-based activation function to compress the acquired vibration signal without increasing its average power. In the processing center, a denoising-expansion autoencoder expands the compressed signal while minimizing noise enhancement. Thus, the framework proposes an innovative method to realize various signal processing operations in the framework using lightweight ML models

that can be efficiently implemented on microcontrollers for practical deployment. Experimental results demonstrate the effectiveness of the proposed companding scheme in preventing nonlinear distortion, enhancing power amplification efficiency in the sensor node, improving performance, and restoring PAPR characteristics in the processing center while avoiding noise expansion.

The feature extraction methods developed in the thesis are based on signal processing since feature engineering using signal processing creates interpretable features that link meaningfully to signal conditions or classes, unlike the black-box features of Deep Learning (DL) models. Additionally, in the case of DL features, the performance is highly dependent on the model's architecture, parameters, and training. In contrast, signal processing-based features offer more consistent and interpretable performance.

The remainder of this thesis is structured as follows: Chapter 2 provides background on the topic of VBCM and the research motivation. The tutorial on the fundamentals of signal processing and the review of feature extraction methods are introduced in Chapter 3. Chapter 4 presents the similarity-based framework for PdM of rotating machinery, addressing the limited availability of labeled training samples in real-world scenarios. Chapters 5 and 6 address memory considerations and system monitoring delay and present the proposed low-complexity VBCM methods. Chapter 7 carries out the analysis of the PAPR of vibration signals and introduces the lightweight autoencoder-based signal companding scheme for the effect of remote VBCM. The thesis conclusions and future research directions are presented in Chapter 8.

The contributions of this thesis are summarized as follows:

- Comprehensive tutorial on the fundamentals of signal processing for a diverse readership. Followed by an end-to-end overview of a typical signal processing pipeline and an exhaustive review of feature extraction techniques through a novel taxonomy.
- Introducing a similarity-based framework for condition monitoring of rotating ma-

chinery. The main aspects of the framework are feature extraction and similarity-based classification.

- Analyzing the overall delay in vibration-based condition monitoring and introducing the concept of system delay to assess it and, accordingly, proposing a hybrid wavelet-based method with reduced system delay for VBCM of rotating machinery. The proposed method is highly sensitive to fault-related transients with relatively short durations of input vibration segments.
- Introducing a low-complexity new method for VBCM of rolling bearings based on vibration's envelope analysis, facilitating three novel envelope representations. Maintaining temporal information, these representations effectively capture energy-frequency variations that are unique to the condition of the bearing, thereby enabling the extraction of discriminative features with high sensitivity to variations in operational conditions. Accordingly, a set of six new highly discriminative features are extracted from these representations. The proposed method facilitates a low-complexity VBCM since it utilizes input vibration segments of very short durations and produces six features only. Thus, relaxing memory requirements and reducing monitoring delays which in turn, helps reduce memory costs and prevent costly catastrophic failures.
- To the best of the author's knowledge, this thesis contains the first contribution to the VBCM literature that addresses the PAPR of generated vibration waveform, examines its impact in the presence of nonlinear power amplifications, and proposes controlling the PAPR to enhance power efficiency, mitigate nonlinear distortion, and improve the reliability of condition monitoring in remote VBCM applications. This is achieved by introducing a lightweight autoencoder-based signal companding framework. The presented work uses an innovative method to realize various signal processing operations in the framework using lightweight ML models that can be

efficiently implemented on microcontrollers for practical deployment.

Chapter 2

2 Background and Motivation

This chapter provides a background on data-driven Vibration-Based Condition Monitoring (VBCM) and discusses the motivation behind the research problems addressed in this thesis. Section 1 provides a brief background on data-driven VBCM and identifies the challenges addressed in this thesis. Section 2 briefly introduces the corresponding proposed solutions. Section 3 explains the motivation behind the selection of the experimental vibration datasets used throughout the thesis to implement and evaluate the proposed solutions,

2.1 Background

VBCM can be defined as *a sophisticated signal-based methodology for assessing a system condition based on its inherent vibration patterns*. By monitoring changes in vibration signatures, which reflect a change in the system's current state, VBCM provides a non-invasive, real-time approach to continuously monitor the system's condition. From a historical point of view, the origin of VBCM is connected to condition monitoring of machinery when the relationship between vibration signals and machine condition was first recognized in 1939 [1]. This opened the door for developing effective Predictive Maintenance (PdM) strategies for identifying potential malfunctions and predicting maintenance actions before catastrophic failures occur. Thus, from a historical perspective, the word "system" in the aforementioned definition originally referred to "machinery." Besides PdM of machinery, VBCM has been widely used in Structural Health Monitoring (SHM) as a damage identification strategy in aerospace, civil, and mechanical engineering infrastructure [2]. VBCM-SHM techniques involve the analysis of vibration patterns collected from different sensors placed across a structure to locate and determine the severity of any

structural damage. This, in turn, helps evaluate the structure's safety and make informed decisions about maintenance strategies [3]. The main advantages of VBCM over alternative forms of condition monitoring include [4][1]:

- Vibration sensors are non-intrusive and can be contactless, facilitating non-destructive condition monitoring.
- Real-time acquisition of vibration signals can be conducted in situ, allowing for on-line local condition monitoring.
- Trending vibration analysis can be utilized to identify relevant conditions and conduct comparative analysis across diverse conditions or objects.
- Vibration sensors are cost-effective and widely available, offering various specifications to suit various requirements.
- Vibration waveform responds instantly to changes in the monitored condition and, therefore, is suitable for continuous and intermittent monitoring applications.
- Signal processing techniques can be applied to vibration signals to mitigate corrupting noise and extract weak condition indications from other masking signals.

These advantages, along with the advancements in sensor fabrication and the rise of the Internet of Things (IoT), have extended the use of VBCM to include various domains spanning a diverse range of applications such as wildlife, agriculture, transportation, and healthcare [5–12].

The process of VBCM involves the application of various signal processing methods to analyze vibration signatures and extract relevant features that reflect condition changes in the system's condition. Such signatures often manifest through sudden changes in amplitude, frequency, and phase characteristics of the generated vibration pattern. The advancements in Machine Learning (ML) and computing technologies have significantly expanded the potential of VBCM where data-driven VBCM has emerged as a paradigm shift. The core

of data-driven VBCM is centered around leveraging DSP and ML capabilities, improving the performance and reliability of VBCM systems. Fig. 2.1 shows a process flowchart of

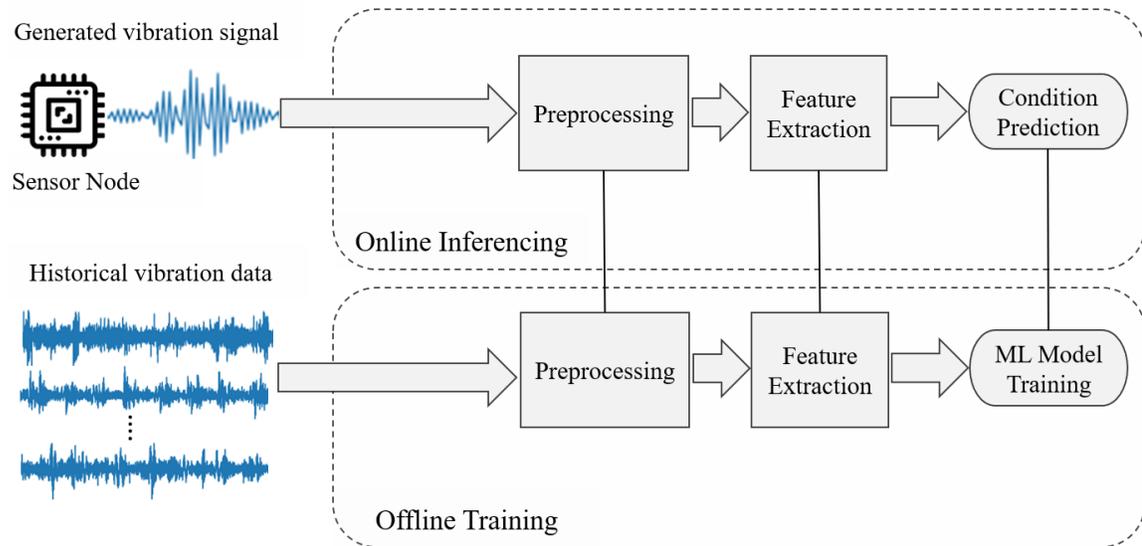


Figure 2.1: Data-driven VBCM.

a typical data-driven VBCM system; the process involves two tasks: offline training of a proper ML model on historical measurement vibration data and online inferencing of the system's current condition. The main aspects of the process can be summarized under three major stages:

- **Signal Preprocessing:** preprocessing involves signal smoothing, signal denoising, and signal segmentation.
- **Feature Extraction:** This stage involves processing the resultant segments to extract appropriate features that serve as inputs to the ML model.
- **Training of ML Model:** In this stage, an appropriate ML model is trained on the extracted features to learn dependencies between various system conditions in the feature space for a specific condition monitoring task, such as signal classification, clustering, or anomaly detection.
- **Condition Monitoring:** In this stage, real-time vibration patterns are acquired and

preprocessed. Accordingly, features are extracted and fed into the trained ML model to predict the system's current condition.

2.2 Motivation

While bringing significant advantages, the practical implementation of data-driven VBCM encounters main challenges, limiting larger-scale deployment and impacting the reliability of the monitoring process. These challenges include:

- Challenge 1: Extensive prior knowledge in signal processing is required.
- Challenge 2: Limited availability of labeled training data since vibration samples of abnormal conditions are often not abundant.
- Challenge 3: Computational complexity in terms of high memory requirements and long monitoring time delay.
- Challenge 4: The Presence of noise in the collected vibration signals affects the reliability of the monitoring process.
- Challenge 5: High power consumption in vibration sensor nodes reduces sensor lifetime.

Accordingly, the work presented in this thesis proposes the following solutions to address these challenges. These solutions are briefly introduced, while further elaboration and implementation details are presented in subsequent chapters.

2.2.1 Challenge 1

Implementing an efficient data-driven VBCM for a specific application requires proper knowledge of signal processing to analyze generated vibration signals and consequently implement effective signal preprocessing and feature extraction methods. The lack of such

knowledge creates a significant obstacle for a wide range of professionals in the interdisciplinary field of data-driven VBCM. Addressing this challenge, Chapter 3 attempts to bridge this knowledge gap through a two-fold approach. First, it introduces a comprehensive tutorial on the fundamentals of signal processing, thereby giving professionals in the field the necessary background knowledge. Second, it conducts a comprehensive review addressing the main stages of a typical signal processing-based ML pipeline and exploring various feature extraction techniques, as illustrated in Fig. 2.2 and Fig. 3.29, which shows the typical signal processing pipeline and taxonomy of feature extraction techniques, respectively. Furthermore, the presented work encourages a collaborative research culture by making

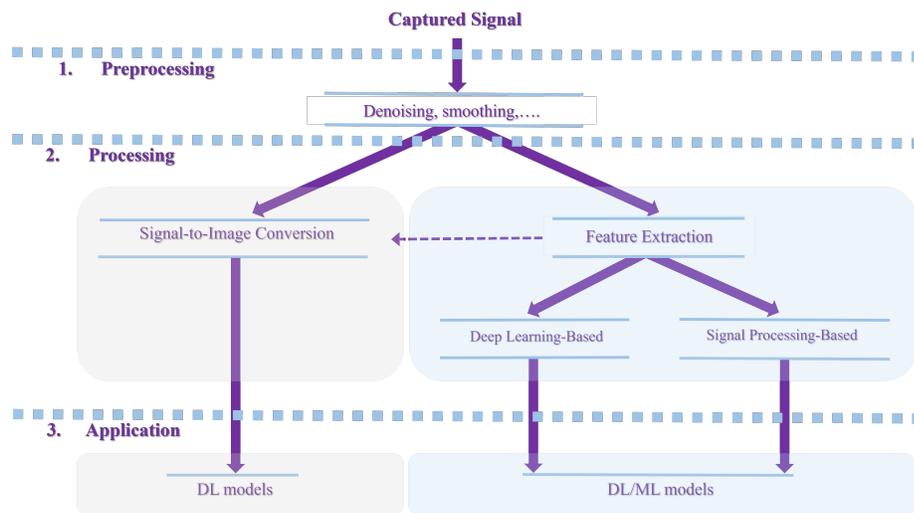


Figure 2.2: Typical signal processing-based ML pipeline.

available a public repository containing relevant Python and MATLAB codes for various signal-processing techniques.

2.2.2 Challenge 2

Utilizing supervised machine learning models in data-driven VBCM requires a substantial and diverse set of labeled samples to train these models effectively. However, in practical scenarios, the quantity of available labeled samples is often limited due to the costly and time-consuming labeling process. Additionally, samples of faults or abnormalities are par-

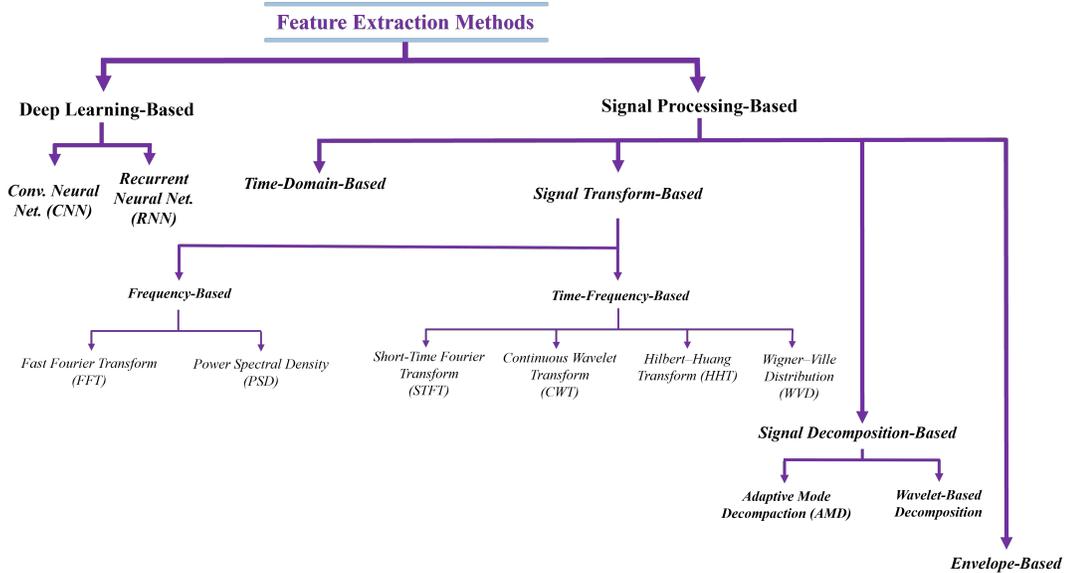


Figure 2.3: Taxonomy of feature extraction techniques.

ticularly not abundant, as such abnormal events occur infrequently. This scarcity makes it challenging to collect a sufficient amount of abnormal signal data. Consequently, real-world samples of abnormal conditions are typically inadequate for accurately representing the respective condition classes during the model training phase. Chapter 4 introduces a similarity-based framework as an alternative solution to perform condition monitoring using limited labeled data by measuring the similarity between an acquired vibration sample and labeled reference samples. In the proposed framework, features are extracted from available labeled reference samples. Accordingly, condition monitoring is achieved by assessing the similarity between the acquired sample and the reference samples in the feature space as illustrated in Fig. 2.4

2.2.3 Challenge 3

Memory requirements and monitoring delay play pivotal roles in the deployment of VBCM, directly influencing monitoring reliability and associated financial expenses. Specifically, an increase in memory demand elevates financial costs, while prolonged delays in condition prediction may fail to prevent costly catastrophic failures. Addressing this challenge,

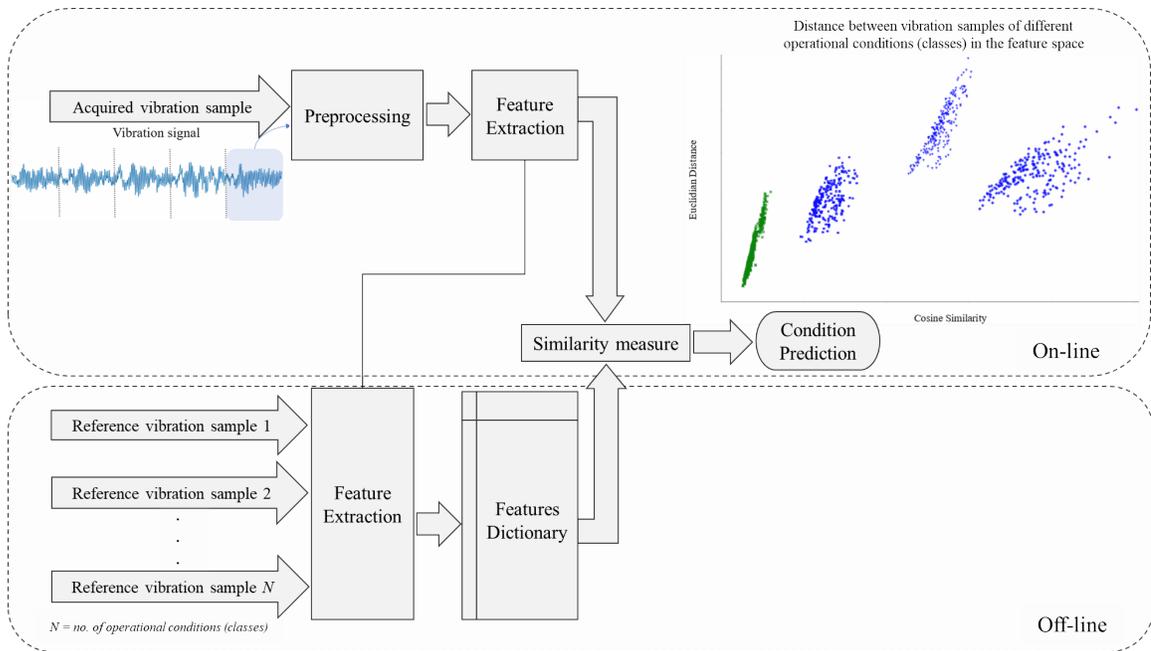


Figure 2.4: Similarity-based framework for VBCM using limited labeled data.

chapters 5 and 6 introduce two methods to facilitate low-complexity VBCM. Chapter 5 proposes a hybrid method that combines wavelet decomposition and Fourier Transform (FT), as shown in Fig. 2.5, to conduct wavelet spectral-energy analysis and extract a few features with high sensitivity to condition changes. The proposed method allows the use of short durations of the generated vibration signal and produces a feature vector of controllable size, thereby relaxing memory requirements and reducing monitoring delay. Chapter 6 introduces a low-complexity method for VBCM of rolling bearings based on envelope analysis of the generated vibration signal. Specifically, the proposed method employs the Hilbert transform (HT) to obtain and then jointly analyze the instantaneous amplitude “envelope” and instantaneous frequency of vibration signals to facilitate three novel envelope representations: Instantaneous Amplitude-Frequency Mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and Instantaneous Energy-Frequency Distribution (IEFD) as illustrated in Fig. 2.6. Accordingly, a set of six new fault-sensitive features are engineered from these representations. Further, the proposed method uses very short durations of the generated vibration signal for condition monitoring, thereby relaxing memory re-

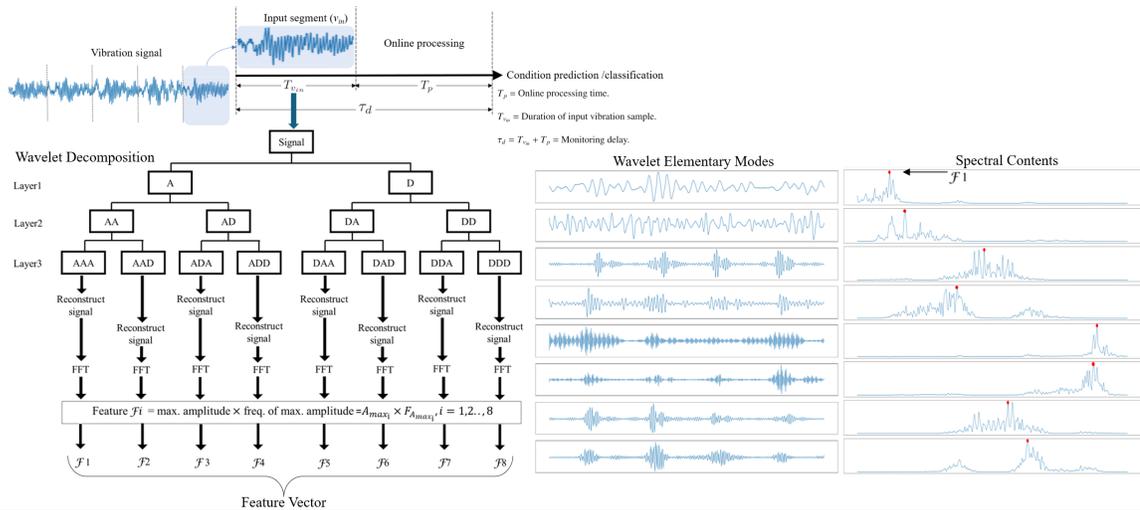


Figure 2.5: Proposed wavelet spectral energy for VBCM with low computational complexity.

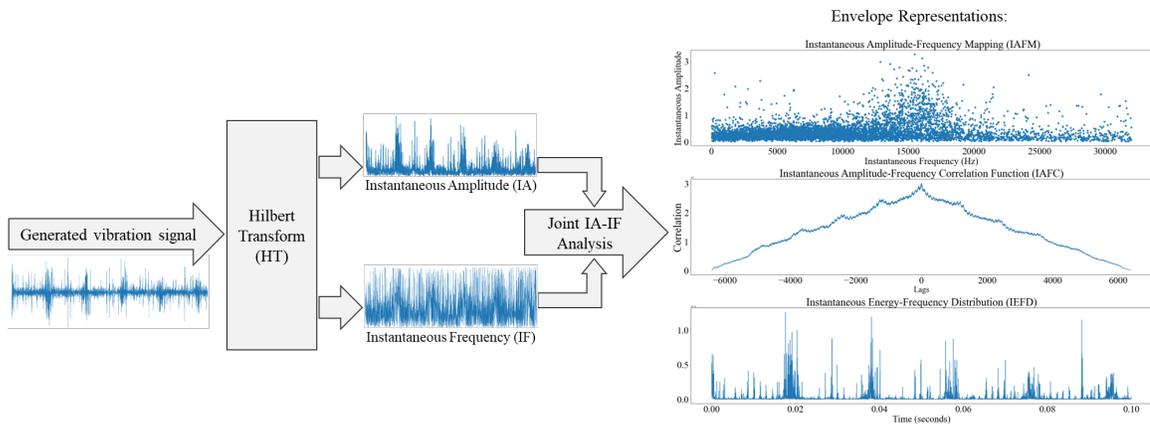


Figure 2.6: Proposed envelop-based analysis for VBCM with low computational complexity.

quirements and reducing monitoring delay.

2.2.4 Challenges 4 and 5

The rapid evolution of sensor fabrication, coupled with advancements in the IoT, has enabled the deployment of large-scale Wireless Sensor Networks (WSNs), facilitating remote VBCM systems comprising distributed sensor nodes as depicted in Fig. 2.7. In

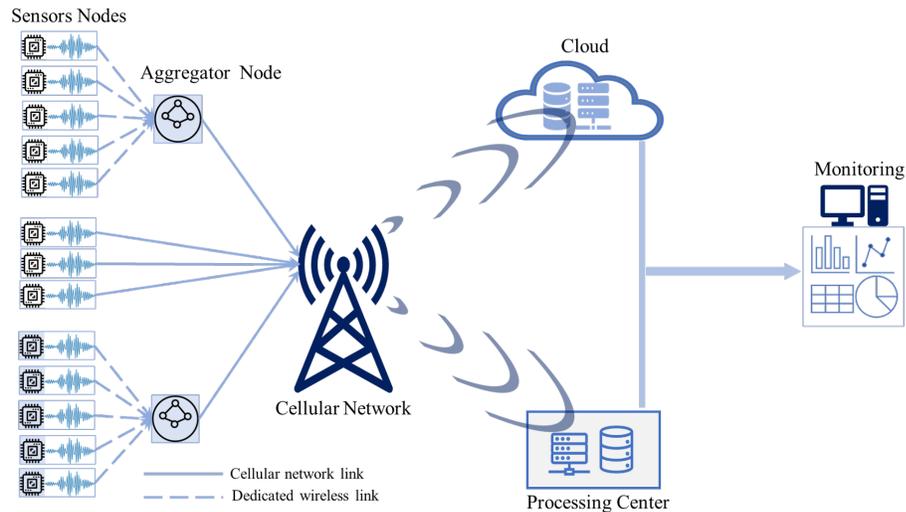


Figure 2.7: remote VBCM-WSN.

such systems, sensor nodes are deployed across various locations, either embedded in objects, placed beneath surfaces, or attached to mobile or airborne objects. These nodes are typically power-constrained, which places the need to maintain low power consumption, thereby extending the lifetime of the sensor node. Additionally, given the noisy environments in which these sensors operate, it is necessary to employ an effective signal denoising mechanism to mitigate the impact of noise accumulated in the acquired signals. The work presented in Chapter 6 tackles these challenges through an innovative approach, as illustrated in Fig. 2.8. Specifically, a lightweight reconstruction autoencoder with a compression-based activation function is used in the sensor node to compress the acquired vibration signal without increasing its average power. The compression operation reduces

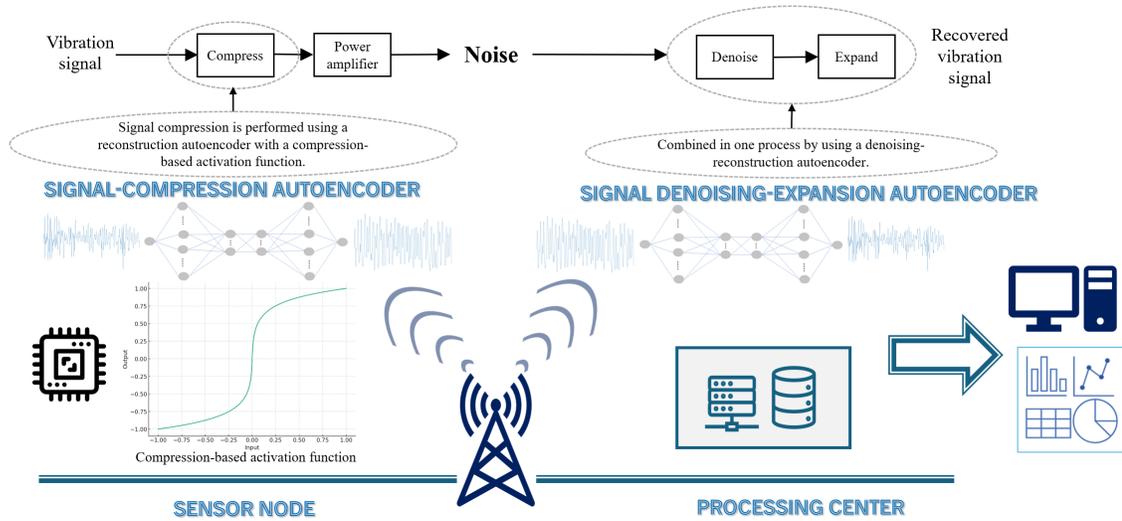


Figure 2.8: Proposed framework of autoencoder-based signal companding and denoising.

the Peak-to-Average Power Ratio (PAPR) to mitigate the impacts of nonlinear power amplification, thereby enhancing power efficiency in the sensor node. In the processing center, a denoising-expansion autoencoder that simultaneously expands and denoises the received, compressed signal is adapted.

2.3 Experimental Vibration Datasets

Throughout this thesis, various public vibration datasets are utilized to implement and evaluate the proposed solutions. These datasets include vibration data of rolling bearings acquired through normal operational conditions and real-life accelerated and simulated faulty conditions. The choice of rolling bearing vibrations over other types of vibrations is motivated by the following reasons:

- Unlike other types of vibration, such as structural vibrations, which are often characterized by simpler dynamics, the vibrations of rolling bearings embody complex vibration patterns due to the intricate interactions between the various components within the bearing assembly, such as balls, raceways, and cages.

- Defects in rolling bearings, such as the inner raceway, outer raceway, or ball defects, produce distinct vibration signatures. These signatures are modulated by varying rotating speed and load conditions, adding layers of complexity to the generated vibration signal.
- The operational behavior of rolling bearings includes nonlinearities due to varying contact forces and load conditions. These nonlinearities influence the vibration patterns, making them more complex compared to structural dynamics.
- Rolling bearing vibrations consist of a broad spectrum of frequency components, including characteristic frequencies related to the bearing geometry and operational speed, as well as harmonic and sideband frequencies caused by various operational conditions.

These complex dynamics in rolling bearings provide rich vibration patterns, making them more suitable for evaluating the proposed data-driven VBCM solutions compared to other types of vibrations.

Chapter 3

3 On the Intersection of Signal Processing and Machine Learning: A Use Case-Driven Analysis Approach

Addressing the first research problem, this chapter¹ presents an integrated article that comprehensively covers the merger between signal processing and Machine Learning (ML). The work presented in this chapter is entitled “On the Intersection of Signal Processing and Machine Learning: A Use Case-Driven Analysis Approach” and is currently in the review stage at the *Proceeding of the IEEE*.

3.1 Introduction

Recent advancements in sensing, measurement, and computing technologies have significantly expanded the potential for signal-based applications, leveraging the synergy between signal processing and ML to improve both performance and reliability. This fusion represents a critical point in the evolution of signal-based systems, highlighting the need to bridge the existing knowledge gap between these two interdisciplinary fields. Despite many attempts in the existing literature to bridge this gap, most are limited to specific applications and focus mainly on feature extraction, often assuming extensive prior knowledge in signal processing. This assumption creates a significant obstacle for a wide range of readers. To address these challenges, this chapter takes an integrated article approach. It begins with a detailed tutorial on the fundamentals of signal processing, providing the reader with the necessary background knowledge. Following this, it explores the key stages of a standard signal processing-based ML pipeline, offering an in-depth review of feature extraction techniques, their inherent challenges, and solutions. Differing from existing literature, this

¹A version of this chapter has been submitted for publication in *Proceeding of the IEEE*.

work offers an application-independent review and introduces a novel classification taxonomy for feature extraction techniques. Furthermore, it aims at linking theoretical concepts with practical applications, and demonstrates this through two specific use cases: a spectral-based method for condition monitoring of rolling bearings and a wavelet energy analysis for epilepsy detection using Electroencephalogram (EEG) signals. In addition to theoretical contributions, this work promotes a collaborative research culture by providing a public repository of relevant Python and MATLAB signal processing codes. This effort is intended to support collaborative research efforts and ensure the reproducibility of the results presented.

The rapid advancements in sensing and measurement represent a paradigm shift in how data is collected, processed, and interpreted. This opens the door for a wide range of signal-based applications, marking a transformative phase across various fields. Moreover, the development of computing technologies and the rise of the Internet of Things (IoT) have paved the way to leverage ML within signal-based applications, offering new insights and achieving unprecedented levels of accuracy and efficiency. This merge between signal processing and ML is expected to play a major role in the next generations of sensor-enabled systems across various areas [13]. The integration of signal processing pipelines into ML models forms the fundamental core of these systems. Further, it represents a critical intersection in their advancement, motivating the research community to address the role of signal processing in ML. However, the diverse landscape of signal types and application requirements shapes the scope of the existing body of work to be application-centric, limiting their scopes to specific applications. For instance, The role of feature extraction in ML has been extensively reviewed within the context of vibration-based Predictive Maintenance (PdM). The work in [14] presents a review that focuses on transforming traditional methods to ML techniques in applying vibration-based damage detection in civil structures. The article highlights traditional methods and presents a comprehensive review of the latest applications of ML algorithms used for this purpose. In [15], a systematic re-

view is conducted on adopting ML for failure prediction in industrial maintenance. The review covers the used datasets, preprocessing, and the training and evaluation of prediction models. In [16], a comprehensive review is presented on signal processing techniques for vibration-based feature extraction in Structural Health Monitoring (SHM). The work in [17] addresses the application of vibration-based condition monitoring techniques for the PdM of rotating Machinery. It provides a comprehensive review of vibration data acquisition and analysis, as well as the methods used for fault interpretation and diagnosis, including data acquisition, data transmission, signal processing, and fault detection. The work in [18] introduces a tutorial on the same topic that describes relevant signal processing methods in this field. Furthermore, the tutorial provides Python and MATLAB code examples to demonstrate these methods alongside explanatory videos.

In the biomedical field, the topic of signal processing in ML is an active area of research. For instance, the work in [19] provides an end-to-end review of biomedical signal processing for health monitoring applications. It introduces a flow for developing biomedical signal processing systems. Further, it covers recent applications, types of low-cost, non-invasive biomedical sensors, signal processing techniques, and future perspectives for building reliable systems. The application of MI-based Brain-Computer Interfaces (BCIs) in controlling external devices through EEG signal processing is addressed in [20]. The article reviews recent ML models, identifies major challenges, and suggests potential solutions by focusing on feature extraction and classification methods. The work in [21] reviews various feature extraction techniques for Electrocardiogram (ECG) Signal analysis. In [22], the main steps in detecting and classifying EEG epileptic seizure activities are addressed, along with a review of related feature extraction techniques. The studies in [23] and [24] deal with the application of emotion recognition using EEG signals, where different existing feature extraction methods are analyzed and compared in terms of classification performance. A comprehensive review of methods and techniques that covers the entire process of EEG signal processing is presented in [25]. The study analyzed numer-

ous articles related to EEG signal processing, identified limitations, and analyzed future development trends. Besides biomedical and BdM fields, the role of signal processing in ML has been addressed in various fields such as audio analysis and recognition [26–32], seismic signal analysis [32–34], and telecommunications[35–42].

Despite this diverse spectrum of articles addressing the role of signal processing within ML, there remains a noticeable gap in the literature, characterized by the following four main limitations in the existing studies:

- **Application-Specific Focus:** The discussions are often confined to specific applications, which narrows the coverage and depth with which the subject is addressed.
- **Limited Audience Reach:** Many articles assume a substantial background in signal processing, which restricts accessibility to a broader audience.
- **Task-Centric Approach:** The exploration of signal processing in ML tends to focus on particular tasks, such as signal preprocessing or feature extraction. This focus neglects other critical tasks like signal segmentation, smoothing, and denoising, which are vital components of a comprehensive signal processing-ML pipeline.

This chapter attempts to bridge this gap by implementing an integrated-article approach that addresses the aforementioned shortcomings through the following contributions:

- **Comprehensive Tutorial for Diverse Readership:** The chapter starts with a comprehensive tutorial on signal processing fundamentals that caters to readers across various domains, offering the interested reader an opportunity to develop a proper background before delving into the review.
- **Application-Independent Approach:** The chapter adopts a broad, application-independent review, providing a comprehensive overview of signal processing in ML that is not shaped to a specific use case.

- **End-to-End Overview of signal processing workflow:** The chapter thoroughly discusses the key tasks in a typical signal processing pipeline, grouping them under three main categories: preprocessing, processing, and application.
- **Exhaustive Review with Novel Taxonomy:** The chapter conducts a detailed review of feature extraction techniques, presented and categorized through a new taxonomy that presents new insights and enriches the reader’s understanding.
- **Public Repository of Signal Processing:** The chapter presents a public repository of the Python codes used throughout the chapter, along with additional codes pertinent to signal processing, thereby encouraging a collaborative research environment and ensuring work reproducibility.

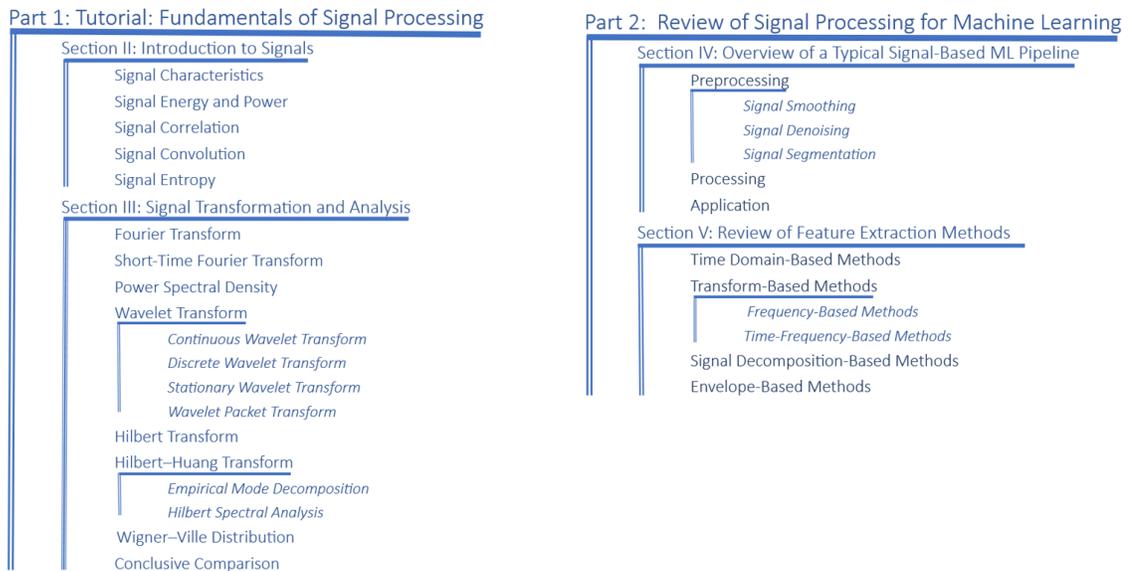


Figure 3.1: Visual representation of chapter contents.

With these contributions, the chapter aims to become a foundational reference on the topic of signal processing in ML, serving a wide range of readers and offering new perspectives on the intersection of signal processing and ML. The chapter allocates a separate section for each of these contributions. The sections are logically grouped under four parts. For the sake of reader convenience, a visual representation that shows the various sections grouped

under each part is depicted in Fig. 3.1. Although these parts are organized sequentially to progressively build knowledge, they are written in a self-contained manner, allowing for selective reading based on interest or need.

The chapter is structured as follows: The tutorial is carried out in Sections 3.2 and 3.3. The typical signal-processing pipeline for signal-based ML applications and the review of feature extraction techniques are presented in sections 3.4 and 4.4, respectively.

3.2 Introduction to Signals

This tutorial provides an in-depth- introduction to signal processing, highlighting concepts, mathematical formulation, applications, advantages, and limitations of common signal processing tools. Additionally, the tutorial offers insights into implementation considerations and highlights programming libraries that offer functionalities for implementing these tools. Throughout the tutorial, illustrative examples are generated using Python and MATLAB codes. These codes are publicly available on the Github site of the Optimized Computing and Communications (OC²) Laboratory.¹

3.2.1 Signal Characteristics

A signal is a function of time; it represents the value of a physical entity or phenomenon as it evolves over time, such as voltage, current, acceleration, etc. Depending on the field, there are more definite definitions of the term “signal”. For instance, within a signal processing context, a signal can be defined as a function that conveys information about the behavior of a system or attributes of some phenomenon [43]. In manufacturing, the term “signal” refers to a physical quantity that carries a certain type of information and serves as a means for communication [44]. Examples of such signals include a vibration signal generated by an accelerometer attached to the rolling-bearing element of a rotating machinery and

¹<https://github.com/Western-OC2-Lab/Signal-Processing-for-Machine-Learning>

a torque signal generated by a torque Sensor in a Computer Numerical Control (CNC) milling machine. The changes in these signals are directly related to the operation of the machine and can therefore be used to communicate the operating status of the machine to the machine operator [44].

Signals are typically represented by their time waveforms; Fig. 3.2. displays the waveform of a given $x(t)$ as it evolves over time t . Amplitude, frequency, and phase are the main

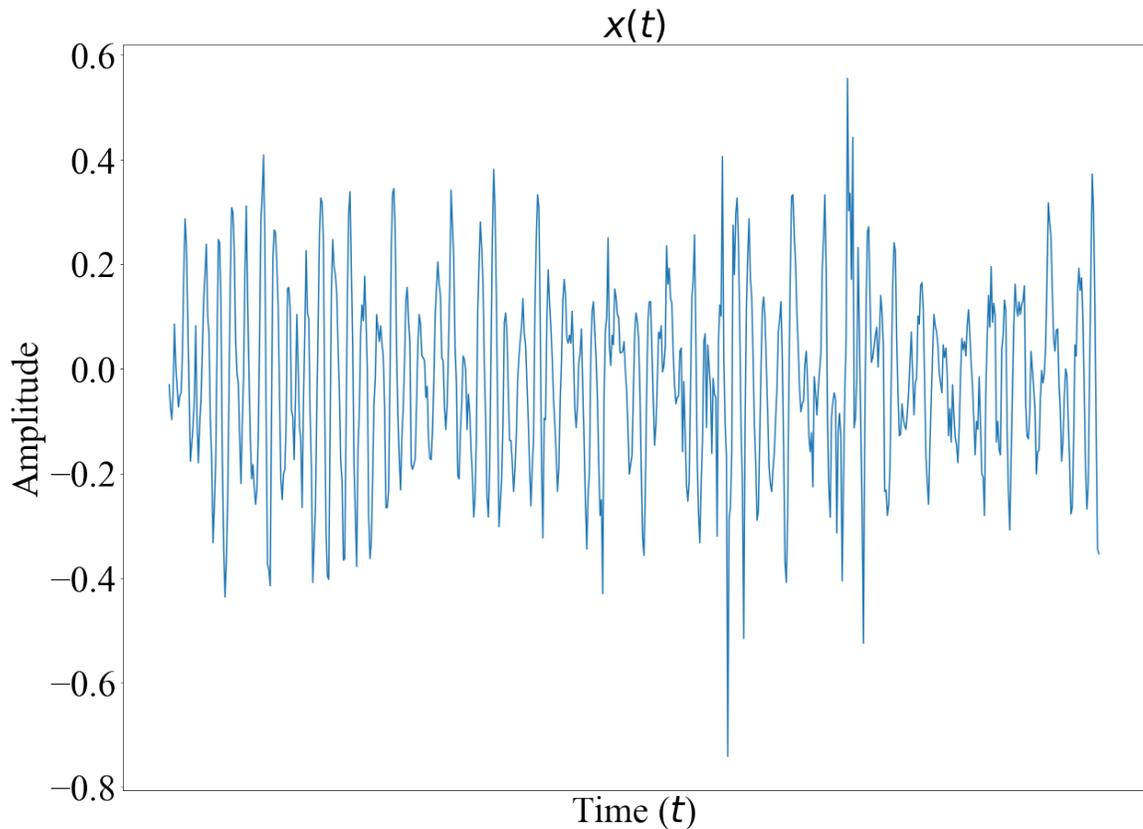


Figure 3.2: Waveform of a signal, $x(t)$, as it evolves over time t .

characteristics of a signal. Amplitude measures the amount and direction of change in the signal with respect to a reference value, as shown in Fig. 3.2. Frequency is the number of cycles or oscillations that occur in a given unit of time. It represents the rate at which the signal oscillates. Frequency is measured in *Hertz (Hz)* where 1 Hz is equal to 1 cycle per second. The phase of a signal refers to the position of a point (time instant) on the signal's waveform cycle. The phase of a wave refers to its position within its cycle and is measured

in either degrees ($0 - 360$) or radians ($0 - 2\pi$). It helps describe the relative position and timing of two signals at a given moment. The concept of frequency of a signal is closely related to the rate at which its phase changes. Specifically, the frequency at a particular moment in time indicates the speed at which the phase is changing at that moment. A higher frequency indicates a faster rate of phase change. Therefore, frequency and phase are mathematically linked through the concept of the derivative. By knowing the phase information of a signal, the frequency can be calculated as the first derivative of the phase with respect to time. The period of a waveform refers to the time it takes to complete one cycle. A periodic signal is a signal that repeats its pattern or the sequence of values exactly after a fixed duration time, known as the period T . This can be expressed mathematically as:

$$x(t) = x(t + T) \quad (3.1)$$

Since the period is the time duration a signal takes to complete one cycle, it is related to the signal's frequency, which represents the number of cycles per second, through the following relation:

$$T = \frac{1}{f} \text{ (seconds)} \quad (3.2)$$

While purely periodic signals do not exist in practice [45], they represent an essential theoretical concept in signals theory [46]. A sinusoidal wave, $s(t)$, represents the basic form of a periodic signal; it is expressed mathematically as:

$$s(t) = A \sin(2\pi ft + \phi) \quad (3.3)$$

Where A is the signal's peak amplitude, f is frequency, and ϕ is the phase. With respect to zero, signal amplitude can be either positive or negative, as shown in Fig.3.3. On the other hand, the magnitude of the signal $|s(t)|$, shown in Fig. 3.4, is the absolute value of the amplitude, it shows the change in the signal regardless of its direction since it is always

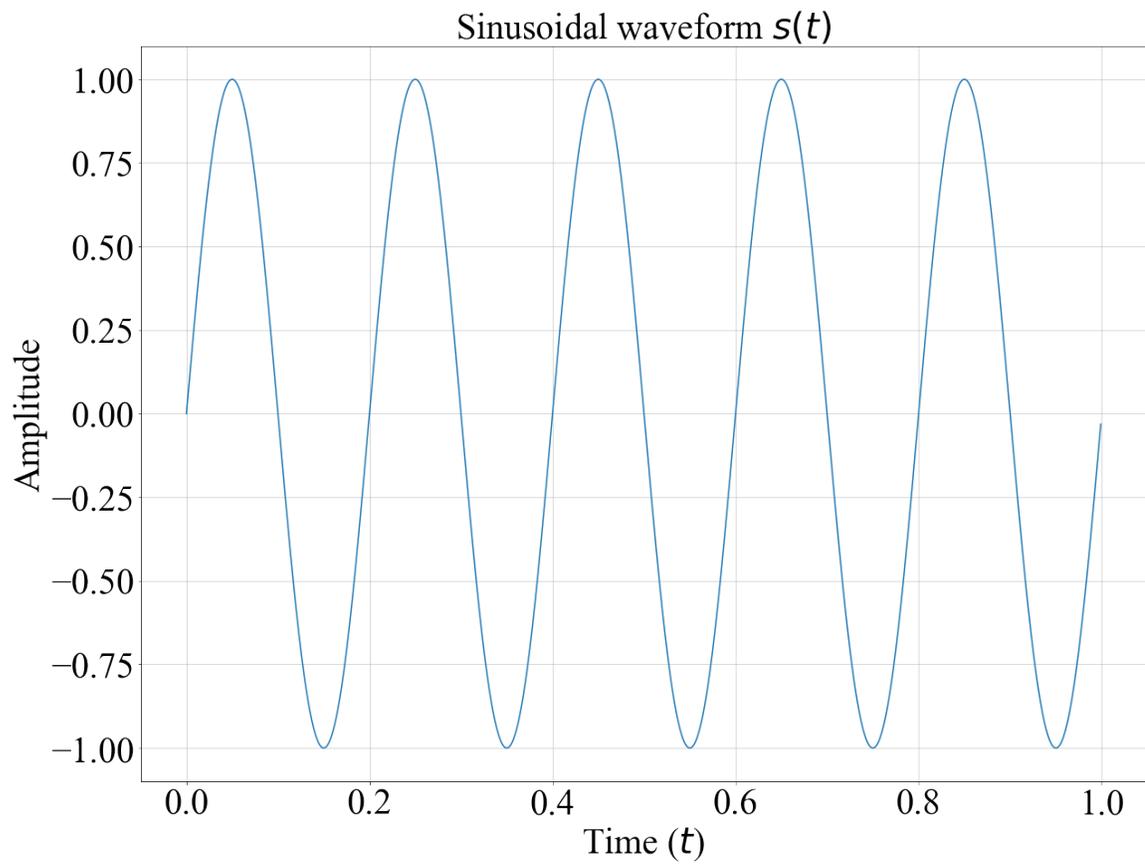


Figure 3.3: Sinusoidal waveform $s(t)$.

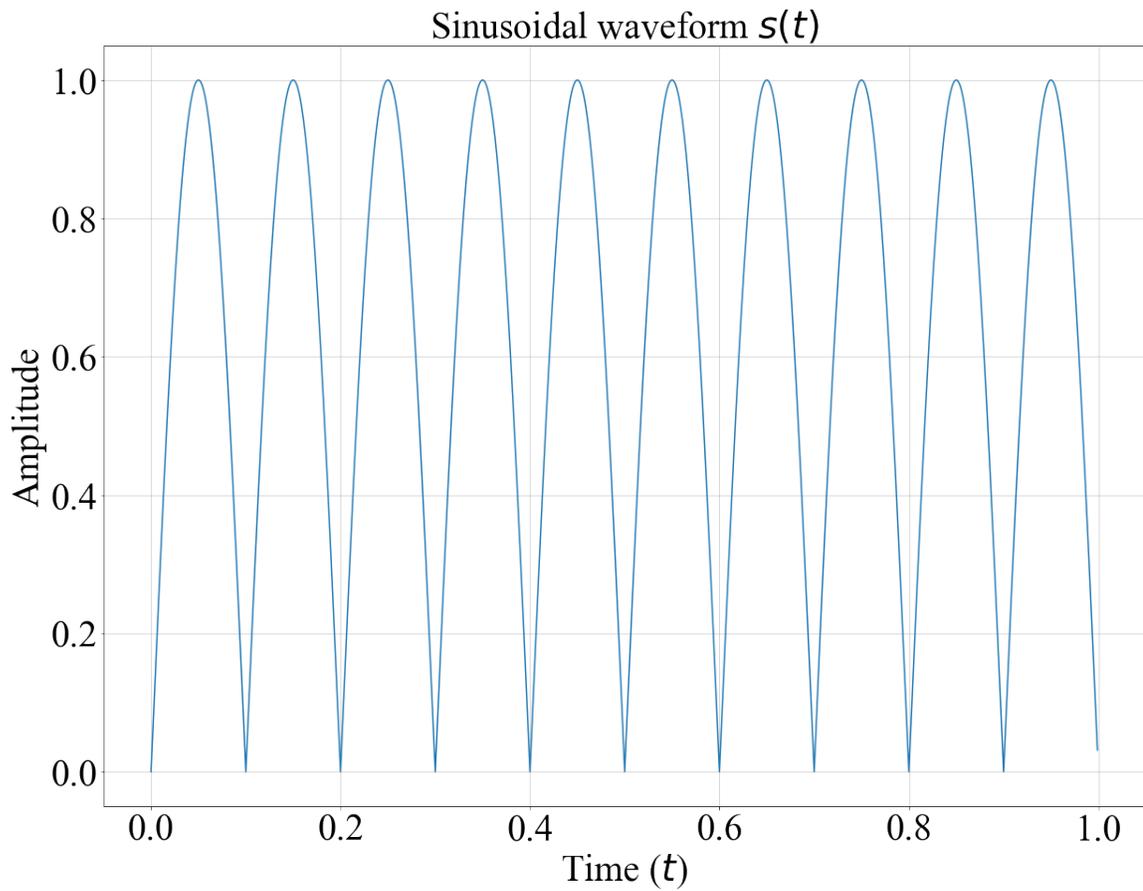


Figure 3.4: Magnitude of the sinusoidal waveform $s(t)$.

a positive quantity. Amplitude values indicate the strength of the signal, reflecting the intensity of the measured physical phenomenon.

As explained earlier, a signal represents a physical entity or phenomenon that changes over time. A signal that remains unchanged over time when measured again under the same conditions is referred to as a deterministic signal. Such a signal has no uncertainty about its value at any given point in time and can be modeled using mathematical expressions. The sinusoidal function is an example of a deterministic function. However, signals that are completely free of unknown and uncontrollable factors and have true deterministic properties are extremely rare [47]. In practical scenarios, signals are non-deterministic as they are random in nature and exhibit various degrees of uncertainty. Since non-deterministic signals are stochastic, they are described in terms of their statistical properties. From this perspective, random signals can be classified into two categories: stationary and nonstationary. Stationary signals generally maintain constant statistical properties over time, while nonstationary signals exhibit changing statistical properties over time.

From the preceding discussion, it is apparent that signal variation describes the changes in the measured physical entity or phenomenon as it evolves over time. This variation is subject to the influence of uncontrollable factors during the measurement period, which manifest themselves as noise or outliers in the signal. Other influencing factors include subsequent processing and operations imposed on the signal, commonly referred to as “systems” in the context of signal processing. Specifically, a system in signal processing is any process or operation that produces an output signal in response to an input signal. In practical situations, real-world systems exhibit mostly nonlinear behaviors² and operate under transient, nonstationary conditions [48]. Therefore, real-world signals in practical situations exhibit nonlinear, time-varying, and nonstationary characteristics. Understanding the

²A linear system in signal theory is defined by two key properties: homogeneity and additivity. Homogeneity means if an input signal $x(t)$ results in an output $y(t)$, then any scaled input $ax(t)$ leads to a scaled output $ay(t)$. Additivity implies that the response to a sum of input signals $x_1(t) + x_2(t)$ is the sum of their responses, $y_1(t) + y_2(t)$.

unique characteristics of a signal holds significant importance, as it directly influences the selection of the proper signal processing tool, ensuring precise and effective analysis of the signal and, consequently, extracting discriminative features.

Another fundamental aspect of signal processing is signal acquisition and the time representation of the signal. From a time perspective, signals are classified as continuous or discrete-time signals. The continuous-time signal $x(t)$ is the signal that has a continuous value over the observation time range, as depicted in Fig. 3.2. A discrete-time version of the signal $x(t)$ is displayed in Fig. 3.5, denoted by $x[n]$. In the discrete-time representation, the independent variable n is discrete in type and takes integer values only. That's it; in contrast to a continuous-time signal, the discrete signal is defined only over discrete time intervals. The discrete-time representation $x[n]$ is obtained by sampling the continuous-

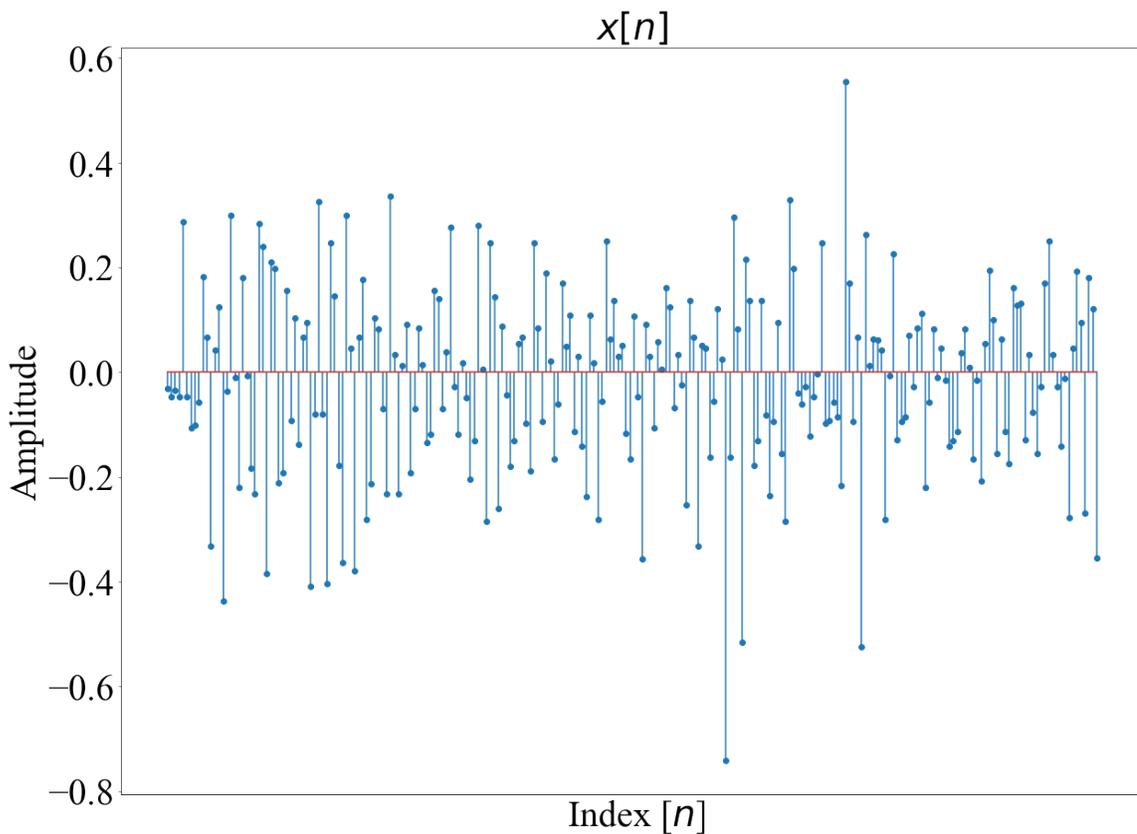


Figure 3.5: $x[n]$, a discrete-time version of signal $x(t)$.

time signal $x(t)$ at time instants separated by T_s , *i.e.*

$$x[n] = x(nT_s), n = 0, 1, 2, \dots \quad (3.4)$$

where T_s represents the sampling interval or sampling period in seconds. Accordingly, the sampling frequency or sampling rate f_s is the number of samples obtained in 1 second; it is given by:

$$f_s = \frac{1}{T_s} \quad (3.5)$$

The unit of f_s is samples per second or hertz. In practice, a measured signal is a sampled version of the actual physical signal since it is acquired by recording a specific number of measurements (samples) every second. The sampling rate plays a crucial role in signal measurement and processing. A higher sampling rate means more measured samples per second and, hence, a higher signal resolution. However, processing time and memory requirements increase as the sampling rate increases. A fundamental concept in signal sampling is the Nyquist theorem [49], which states that a signal can be correctly reconstructed from its sampled discrete-time sequence provided that the sampling rate is equal to or greater than twice the highest frequency (effective bandwidth) of the signal. This implies that the highest frequency, represented as f_{max} , that can be represented accurately in an acquired signal is one-half of its sampling rate f_s ; *i.e.*

$$f_s = 2f_{max} \quad (3.6)$$

This formula is called the Nyquist rate; it provides a lower bound on the sampling rate that is required to reconstruct a signal from its sampled version accurately. Accordingly, a sampled signal is categorized as either undersampled, critically sampled, or oversampled:

- Undersampled signal: is one where the sampling rate is below the Nyquist sampling

rate. Undersampling leads to a distortion in the signal known as aliasing, causing signal components at higher frequencies than the sampling frequency to appear at a lower, aliased frequency [50].

- **Critically sampled signal:** is one where the sampling rate equals the Nyquist sampling rate. In this case, the signal is sampled at the lowest rate, which still allows for a complete reconstruction of the continuous-time signal, assuming the signal bandwidth is limited to half of the sampling frequency. While perfect reconstruction is theoretically still possible, it is less common in practical applications.
- **Oversampled signal:** is one where the sampling rate exceeds the Nyquist rate. Oversampling is widely used in signal measurements because it improves the resolution of the acquired signal. Additionally, with oversampling, the quantization noise (an error that represents the difference between the value of the measured sample and the closest mapped digital value) is spread over a broader frequency range compared to the effective bandwidth of the signal. Accordingly, filtering out high frequencies outside the effective bandwidth effectively reduces noise in the signal, improving the signal-to-noise ratio (SNR).

3.2.2 Signal Energy and Power

As noted above, signal amplitude reflects the intensity of the measured physical phenomenon, making it significant in identifying critical features of the signal, such as energy and power. Energy quantifies the total “work” done by the signal; mathematically, the energy of a signal is the area under the squared magnitude of the signal over a given time interval. for a continuous-time signal $x(t)$, its energy E is defined as:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (3.7)$$

For a discrete-time signal $x[n]$, the energy is given by:

$$E = \sum_{n=-\infty}^{\infty} |x[n]|^2 \quad (3.8)$$

This definition assumes the signal is energy-limited, meaning that its energy is finite over the entire time domain. Signal power measures the average rate at which the signal transmits energy. It is calculated as the average power over a period of time. For a continuous-time signal, its average power P is defined as:

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)|^2 dt \quad (3.9)$$

Similarly, for discrete signals:

$$P = \lim_{N \rightarrow \infty} \frac{1}{2N + 1} \sum_{n=-N}^N |x[n]|^2 \quad (3.10)$$

If the signals' energy E converges to a finite but non-zero value, the signal is classified as an energy signal. Examples include non-zero signals over a limited period but zero elsewhere, such as pulse signals. The average power of an energy signal is zero because the calculation of power involves averaging the finite energy over an infinite period of time. Power signals, on the other hand, are signals that have finite, non-zero average power. A common example is a periodic sinusoidal signal that continues indefinitely; such signals have infinite energy. For signals of finite duration, energy and power can be evaluated over their active time intervals. For periodic signals, due to the repetitive nature of the signal, the computation is simplified by considering one single period T as follows:

$$P = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 dt \quad (3.11)$$

For a discrete-time periodic signal $x[n]$ with period N , the power is:

$$P = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2 \quad (3.12)$$

Instantaneous power refers to the power of a signal at a specific moment in time. Unlike average power, which is calculated over a period of time, instantaneous power gives a moment-to-moment view of the power level of the signal. For a continuous-time signal $x(t)$, the instantaneous power $P(t)$ at time t is defined as the square of the signal's magnitude at that particular time. Mathematically, it is expressed as:

$$P(t) = |x(t)|^2 \quad (3.13)$$

For a discrete signal $x[n]$, where n represents discrete time indices, the instantaneous power $P[n]$ at point n is defined as the square of the magnitude of the signal at that point. Mathematically, this is expressed as:

$$P[n] = |x[n]|^2 \quad (3.14)$$

For a given discrete time signal $x[n]$, represented by N samples, the total energy E is the sum of its instantaneous power values over its discrete time interval:

$$E = \sum_{n=0}^{N-1} |x[n]|^2 \quad (3.15)$$

Accordingly, the average power P is obtained by dividing the energy E by the number of samples N :

$$P = \frac{\sum_{n=0}^{N-1} |x[n]|^2}{N} \quad (3.16)$$

along with power and energy, statistical properties of the signal, such as mean and variance, represent important aspects of signal characteristics and play a major role in signal

processing. For the discrete time signal $x[n]$, its mean μ is equal to:

$$\mu = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad (3.17)$$

The mean represents the central tendency or the average value of the signal over its time period. The signal's variance σ^2 is given by:

$$\sigma^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - \mu)^2 \quad (3.18)$$

The variance measures the spread of the signal's amplitude around its mean value, quantifying the deviation of the signal's samples from its mean. Note that for a zero-mean signal, The variance equals the signal's average power. For such signals, the root-mean-square (RMS) is a useful measure that represents the effective value of the signal because it provides a meaningful measure of the average magnitude of the signal, even though its mean is zero. The RMS is expressed as:

$$x_{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} [x[n]]^2} \quad (3.19)$$

3.2.3 Signal Correlation

Correlation can be viewed as a similarity measure; it quantifies the similarity or the degree of dependence between two signals, providing insights into their common characteristics and patterns. Auto-correlation refers to the correlation of a signal with a delayed version of itself. It represents the degree of similarity between a signal and a time-shifted “lagged” version of itself over a given time interval. Auto-correlation helps to detect repeating patterns in the signal and assess its stability or predictability over time. For a continuous-time

signal $x(t)$, the auto-correlation function is expressed as:

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)x(t - \tau)dt \quad (3.20)$$

where τ represents time shift. For a discrete-time signal $x[n]$, autocorrelation $R_{xx}[k]$, at a lag k is given by defined as:

$$R_{xx}[k] = \sum_{n=-\infty}^{+\infty} x[n]x[n - k], \quad k = 0, \pm 1, \pm 2, \dots \quad (3.21)$$

Note that for zero lag, there is no time shift; hence, the autocorrelation R_{xx} will be at its maximum value,— which equals the signal’s total energy— representing the maximum similarity as depicted in Fig. 3.6 which shows the autocorrelation of the signal $x(t)$.

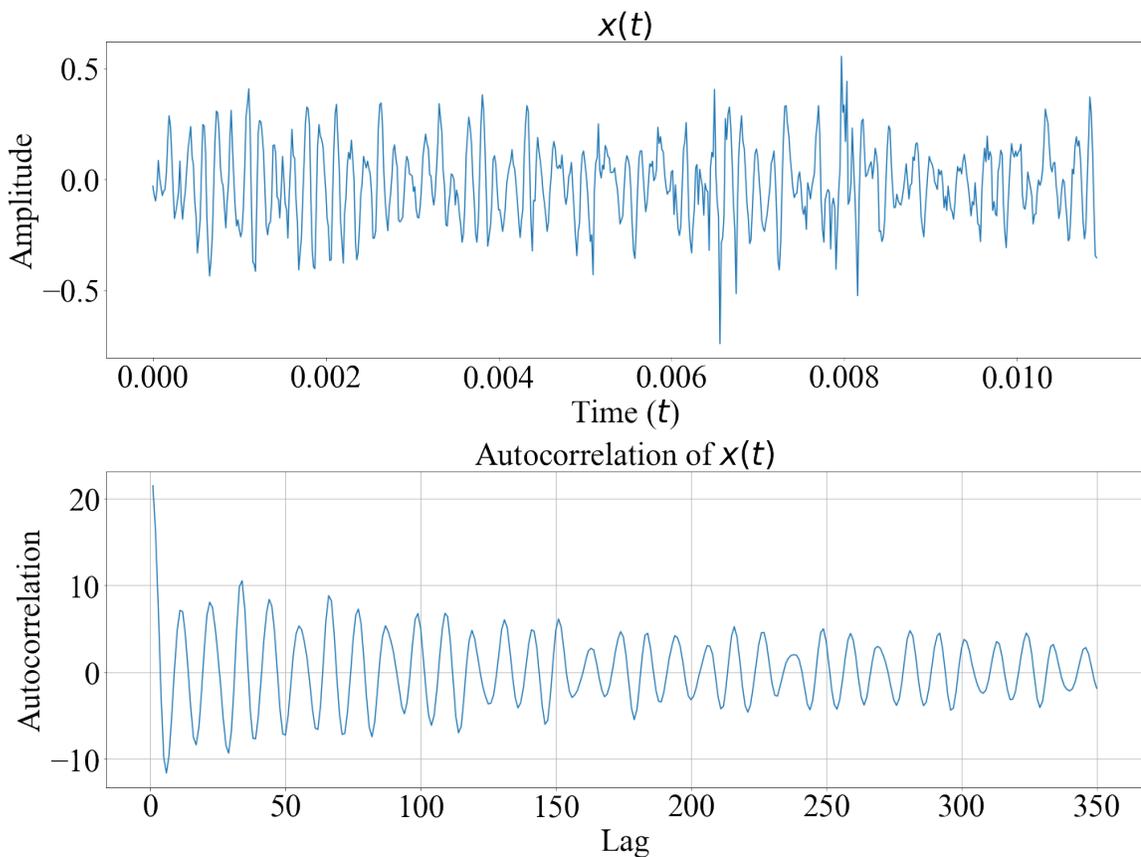


Figure 3.6: Signal $x(t)$ and its autocorrelation.

Cross-correlation extends the concept of auto-correlation to two different signals, measuring the degree of similarity or dependence between them. It quantifies the degree of similarity between one signal and a time-shifted “lagged” version of another signal over a given time interval. For two continuous-time signals $x(t)$ and $y(t)$, their cross-correlation is expressed as:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t - \tau)dt \quad (3.22)$$

For discrete-time signals $x[n]$ and $y[n]$, the cross-correlation R_{xy} at a given time lag k is defined as:

$$R_{xy}[k] = \sum_{n=-\infty}^{+\infty} x[n]y[n - k], \quad k = 0, \pm 1, \pm 2, \dots \quad (3.23)$$

The value of $R_{xy}[k]$ indicates how much of one signal is present in the other signal at a given delay k . Hence, it describes the mutual relationship between the two signals as a function of the displacement of one relative to the other. Cross-correlation is often used to search for desired features or patterns within a signal.

The correlation coefficient between two signals is a statistical measure that quantifies the degree to which the two signals are linearly related. It is a normalized form of the cross-correlation function and is widely used in signal processing and statistics to assess the strength of the linear relationship between two signals [51]. For discrete-time signals $x[n]$ and $y[n]$, each of length N , the correlation coefficient ρ_{xy} is defined as:

$$\rho_{xy} = \frac{\sum_{n=0}^{N-1} (x[n] - \mu_x)(y[n] - \mu_y)}{\sqrt{\sum_{n=0}^{N-1} (x[n] - \mu_x)^2 \sum_{n=0}^{N-1} (y[n] - \mu_y)^2}} \quad (3.24)$$

where μ_x and μ_y are the means of the signals $x[n]$ and $y[n]$, respectively.

The correlation coefficient lies between -1 and $+1$. Values close to $+1$ indicate a strong positive linear relationship, i.e. as one signal increases, so does the other. Values close to -1 indicate a strong negative linear relationship, where one signal increases as the other decreases. Values close to 0 suggest a weak or no linear relationship between the sig-

nals.

3.2.4 Signal Convolution

Convolution is an integral operation on two signals that produces a third signal; it expresses the amount of overlap of one signal as it is shifted over another signal. The convolution of two continuous-time signals $x(t)$ and $h(t)$ is given by:

$$x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau \quad (3.25)$$

where $*$ denotes the convolution operator, which is defined as the integral of the product of $x(t)$ with a time-reversed and shifted version of $h(t)$. The time reversal of $h(\tau)$, represented by $h(-\tau)$, involves reflecting the signal about the vertical axis that represents time origin ($\tau = 0$). Accordingly, the integral is evaluated for all shift values t , producing the convolution signal $x(t) * h(t)$. For discrete-time signals $x[n]$ and $h[n]$ the convolution is expressed as:

$$x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k]h[n - k] \quad (3.26)$$

The convolution process can be visualized as one signal sliding over the other signal, computing the area of overlap as a function of the degree of overlap. In contrast to correlation, which measures the similarity between two signals as a function of the displacement of one relative to the other, the convolution represents the amount of overlap between two signals as one is inverted and shifted over the other signal. Therefore, while correlation measures the similarity between two signals, convolution can be viewed as a measure of the effect of one signal on the other.

Convolution is particularly useful for describing how a Linear Time-Invariant (LTI) system or filter responds over time to an input signal since the output of an LTI system or a filter for any arbitrary input can be determined by convolving the input signal with the system's

impulse response. LTI Systems are systems that are characterized by linearity and time-invariance. Linearity, as explained earlier, implies that the response (output) to a weighted sum of inputs is the weighted sum of the responses to each input. Time-invariance means the system's response to a given input does not change over time. In signal theory, the impulse response of an LTI system or a filter, denoted as $h(t)$ for continuous-time representation or $h[n]$ in discrete-time case, represents the system's output when subjected to an impulse function. In continuous-time representation, the impulse function is denoted as $\delta(t)$ and is defined as a function that is zero at all times except at $t = 0$, where it is infinitely high in such a way that its integral over time is 1. In discrete time, the impulse function is denoted as $\delta[n]$ and is defined as a sequence that is zero at all samples except at $n = 0$, where its value is 1. The impulse response fully characterizes the system's or filter's behavior. Hence, given the impulse response of an LTI system or a filter for any arbitrary input, its output can be determined by convolving the input signal with the impulse response.

3.2.5 Signal Entropy

Entropy is an important concept that originated in information theory as a measure of “information” in a random variable. In signal processing, entropy quantifies the degree of unpredictability or randomness in a signal, thereby reflecting its structure and offering a quantitative measure of various signal characteristics. The fundamental concept of entropy-based applications in signal processing is to treat the signal as a random stochastic process. In this context, the signal samples represent a collection of possible outcomes of the process; each outcome is associated with a probability of occurrence. Hence, for a discrete signal where its samples are represented as a set of possible outcomes x_1, x_2, \dots, x_n with associated probabilities $p(x_1), p(x_2), \dots, p(x_n)$, the Shannon entropy is calculated as follows:

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i) \quad (3.27)$$

This formula calculates the average “information,” “surprise,” or “uncertainty” inherent in the signal’s possible outcomes. In information theory, high entropy suggests that a signal contains a wide range of information. In signal processing, high entropy values are associated with high degrees of unpredictability, uncertainty, irregularity, and randomness, indicating a complex signal with more varying structures compared to predictable or repetitive patterns within the signal. On the other hand, low entropy values often imply a high degree of uniformity and predictability within a signal. This is illustrated in Fig. 3.7, which compares the entropy values of two signals: the random vibration signal $v(t)$ and a square wave. Each signal contains 12800 samples. The signals are segmented into 128 segments so that each segment contains 100 samples. Accordingly, the Shannon entropy (H) is calculated for each segment to quantify uncertainty within each segment and to assess signal irregularity in terms of the variation in entropy values between signal segments. The comparison shows that segments of the random vibration have higher entropy values than those of the square wave, indicating a higher degree of uncertainty. Moreover, the high variation of entropy values among different segments of the random vibration reflects the complex and irregular patterns in the signal. These concepts find useful applications in signal processing, including but not limited to:

- **Measurement of complexity:** Entropy serves as a metric for signal complexity, with higher entropy values indicating more complex signals. Such signals are typically characterized by their low compressibility.
- **Uncertainty Quantification:** Entropy is a direct measure of the uncertainty in a signal. A signal with high entropy suggests less predictability and greater uncertainty about its current state or future values.
- **Regularity Analysis:** Lower entropy values indicate a higher degree of regularity in the signal, which could indicate a fault or abnormal behavior in applications such as condition monitoring and fault detection.

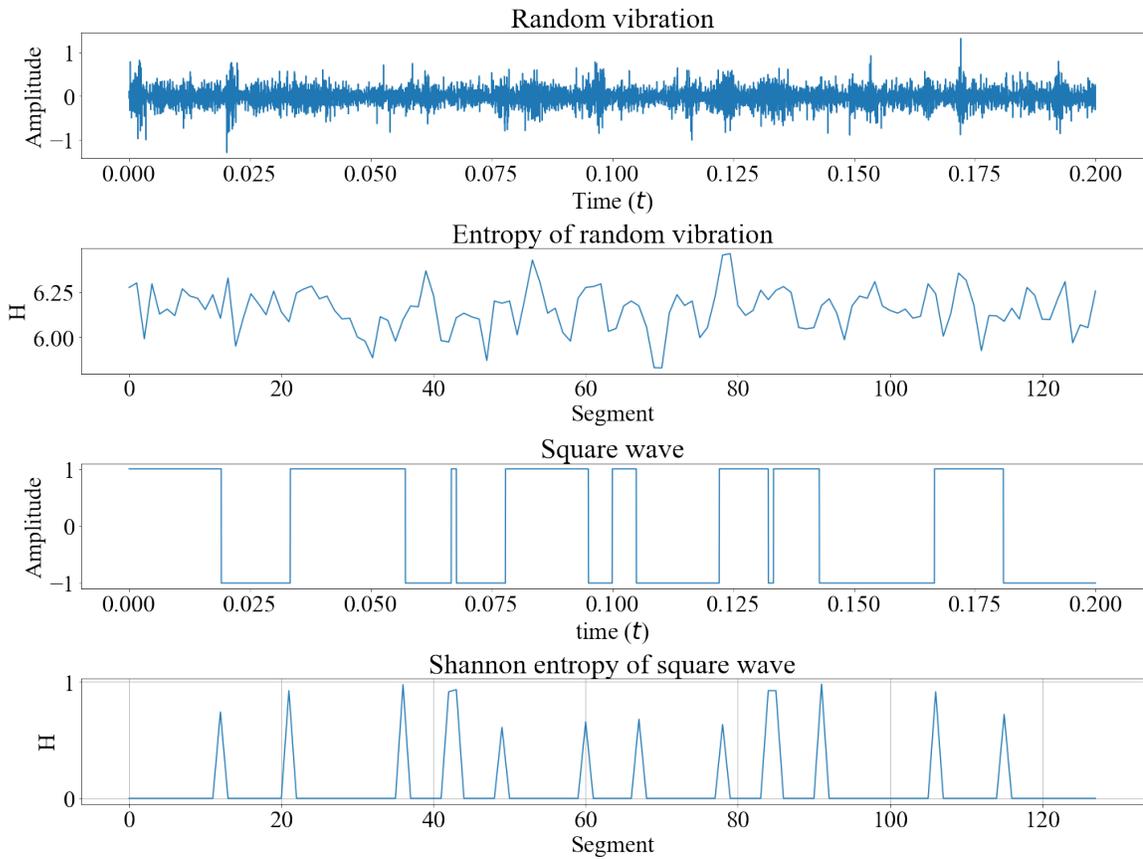


Figure 3.7: A comparison between Shannon entropy values of a random vibration $v(t)$, and a square wave.

- **Randomness Evaluation:** The degree of randomness in a signal can be assessed using entropy. High entropy signals have a high degree of randomness, whereas low entropy signals involve minimal randomness.

The Shannon entropy formula in (3.27), introduced by Claude Shannon in 1948 [52], is the fundamental form of entropy that is widely used in signal processing. Spectral Entropy is another type of entropy that applies Shannon entropy to the power spectrum of a signal, providing insight into its spectral characteristics. Other types of entropy include approximate entropy and sample entropy, commonly used in time-series analysis to measure the regularity and unpredictability of fluctuations in data series. More information on the different types of entropy can be found in [53].

3.3 Signal Transformation and Analysis

Signal transformation is a fundamental concept in signal theory; it plays an important role in many signal processing applications, such as signal analysis, filtering, denoising, compression, and feature extraction. Transforming signals from their time-domain representations into other-domain representations allows the signal to be analyzed in a new paradigm, revealing new aspects of the signal that are not apparent in the time domain. The process of signal transformation involves mapping a signal from its original domain (usually time) to a new domain. This is achieved by computing the inner product of the signal with a set of basic functions “kernels” that serve as the building blocks for representing the signal in the new domain. These functions are chosen based on the properties desired in the transformed domain, for example, sinusoidals in Fourier transforms and wavelets in wavelet transforms. The inner product measures how much of the base function is present in the signal. A large value of the inner product indicates a high degree of similarity or a strong presence of the characteristics represented by the base function in the signal. This section reviews common frequency-domain and time-frequency domain transformations widely used in signal

processing.

3.3.1 Fourier Transform

The Fourier Transform (FT) is one of the most widely used transforms in signal processing. It provides a frequency-domain representation of the signal, revealing its spectral contents and enabling frequency-based analysis. The basic principle underlying the FT is that any waveform can be expressed as a combination of sinusoidal components with different amplitudes, frequencies, and phases. Accordingly, FT attempts to decompose signals into their constituent sinusoidal components. FT is mathematically expressed as:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (3.28)$$

where,

$F(\omega)$ represents the FT the function $f(t)$,

$f(t)$ is the original time-domain signal,

and $e^{-j\omega t}$ is the complex exponential function used in the transform.

The Discrete Fourier Transform (DFT) is the discrete version of the continuous FT, which is used for analyzing discrete signals; it is defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-\frac{j2\pi}{N}kn} \quad (3.29)$$

where,

$X[k]$ represents the DFT of the discrete time-domain signal, $x[n]$,

N is the total number of samples in $x[n]$,

$k = 0, 1, \dots, N - 1$, is the index in the frequency domain,

and $n = 0, 1, \dots, N - 1$, is the index in the time domain.

Both the FT and the DFT are reversible operations where the frequency domain representation of a signal can be transformed back into its original time-domain form by applying

the inverse transform operation. This property is fundamental in signal processing and has various applications, such as signal compression, filtering, and noise reduction.

The Fast Fourier Transform (FFT) and its inverse, IFFT, are computational algorithms that efficiently implement the DFT and its inverse, IDFT. The key advantage of FFT and IFFT is their high computational efficiency. NumPy and SciPy Python libraries offer reliable functions for FFT and IFFT computation. When computing the FFT of a signal, two critical parameters need to be considered: the sampling rate, f_s , and the FFT length, n . The values of these parameters determine the frequency range and frequency resolution of the resulting spectrum. It is essential to choose an appropriate FFT length, usually a power of two (*e.g.*, 256, 512, 1024), as the FFT algorithm performs efficiently when the number of points is a power of two. The FFT length also determines the number of frequency bins in the spectrum, where each bin corresponds to a specific range of frequencies within the total frequency spectrum analyzed by the FFT. The number of unique frequency bins in the FFT output is:

$$\frac{n}{2} + 1 \quad (3.30)$$

The bin width defines the frequency resolution, Δf , of the FFT representing the smallest distinguishable frequency in the spectrum. It is defined as:

$$\Delta f = \frac{f_s}{N} \quad (3.31)$$

Frequency resolution is critical in frequency analysis because it affects the accuracy with which closely spaced frequency components can be resolved in the resulting spectrum. If the FFT length is set equal to the number of samples N in the input signal, the FFT operates directly on the signal as it is. If the FFT length is set greater than N , the signal is typically zero-padded to the desired length. Zero padding does not alter the actual frequency content of the signal but increases the number of frequency bins in the FFT result, leading to a finer frequency resolution in the spectrum. On the other hand, setting the FFT length

to less than N truncates the signal, resulting in loss of information and reduced frequency resolution. The maximum frequency that can be accurately represented in the spectrum is the Nyquist frequency of the signal ($\frac{f_s}{2}$), which equals half of the sampling rate. Frequencies higher than the Nyquist frequency will not be correctly resolved in the spectrum. The output of the FFT function is an array that contains complex Fourier coefficients representing the amplitude and phase information of each frequency bin. The function `fftfreq` in Python can be used to obtain the frequency bins; it takes two inputs, the FFT length n and the sampling period, $T_s = \frac{1}{f_s}$, and returns an array containing frequency bins. Accordingly, the frequency spectrum is constructed by mapping the coefficients to the x -axis and the corresponding frequency bins to the y -axis. Since Fourier coefficients are complex numbers, each bin is expressed either by the magnitude or power of the corresponding Fourier coefficients, resulting in amplitude spectrum or power spectrum, respectively. This gives a two-sided frequency spectrum symmetric around the y axis, showing positive and “negative” frequencies. The negative frequency spectrum is inherent to the Fourier analysis of signals; it is a mirror of the positive “real” spectrum around the y -axis. Thus, it represents redundant frequency information. Reconstructing a positive spectrum involves discarding the second half of coefficients and frequency bins arrays. Thus, to accurately represent the signal’s energy in the frequency domain, it’s necessary to multiply the magnitudes of Fourier coefficients by 2 to restore the energy distribution in the positive frequency since the signal’s energy was equally divided between the positive and negative frequencies. It is worth noting that the number of Fourier coefficients is directly proportional to the length of the input signal, as stated in (3.30). Consequently, the magnitude of the FFT coefficients increases with the length of the signal. Therefore, when longer signal segments are used, the frequency spectrum displays larger values as more components contribute to the sum. Therefore, to reflect the true amplitudes of the signal components, the magnitudes of the Fourier coefficients are usually normalised by the signal length.

To demonstrate the concept of FT in resolving various frequency components, a sinusoidal

signal, denoted as $s_3(t)$ composed of three components, is created as shown in Fig. 3.8. The

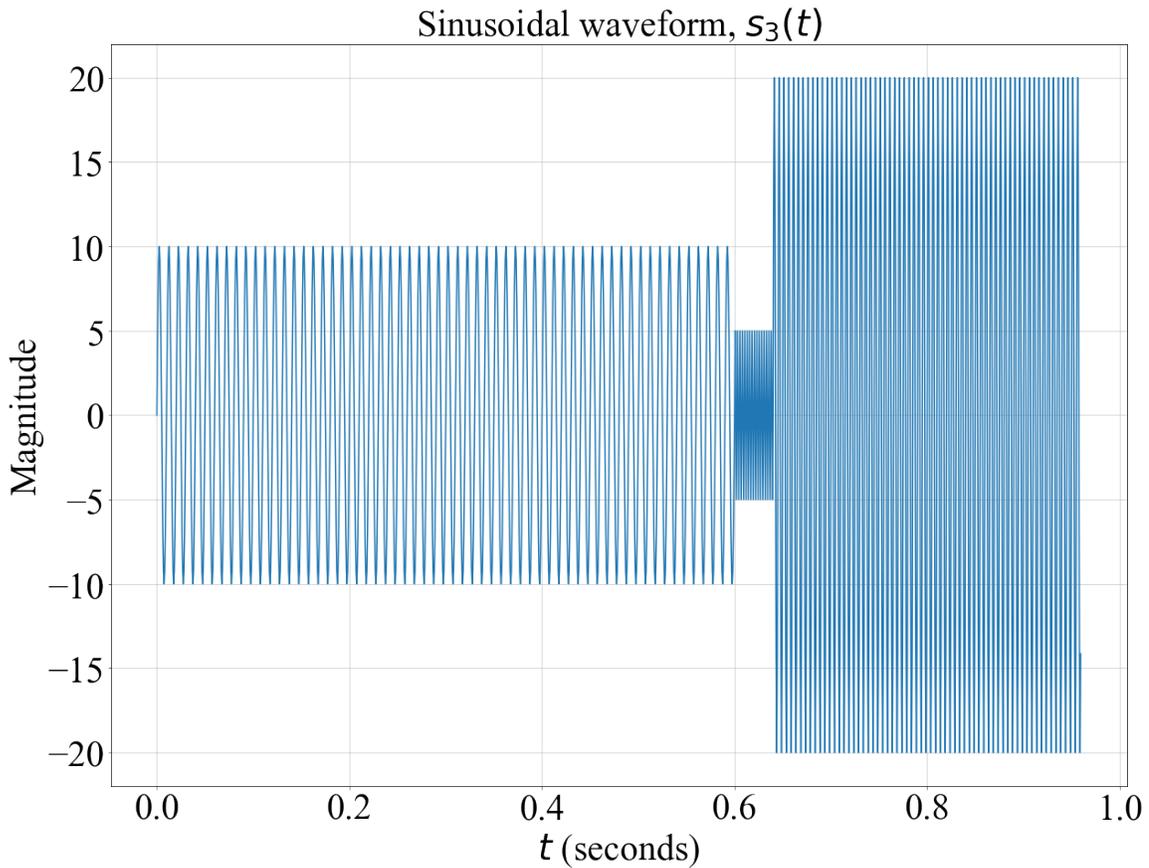


Figure 3.8: Composite sinusoidal signal $s_3(t)$.

used sampling rate is 1600 Hz, and the length of the signal is 1536 data points; accordingly, this corresponds to 0.96 seconds of the signal's time duration. The signal consists of three sinusoidal components with different durations, amplitudes and frequencies. The first component consists of 960 data points and oscillates at 100 Hz with an amplitude of 10. The second component contains 64 data points with a frequency of 400 Hz and an amplitude of 5. The third component consists of 512 data points oscillating at a dominant frequency of 200 Hz and an amplitude of 20. Consequently, the time duration which corresponds to each of the components are 0.6 seconds, 0.04 seconds, and 0.32 seconds, respectively. Compared to the first and third components, the second component represents weak, rapid oscillations of very short duration, simulating a transient component of particular impor-

tance in the composite sinusoidal. The frequency spectrum of the signal is displayed in Fig. 3.9, which clearly shows that the three components in the composite signal at 100 Hz, 200 Hz, and 400 Hz are accurately resolved and represented in the spectrum. The peaks correspond to the amplitudes of the components and therefore convey how much energy is being dissipated within each component. In practical terms, the presence of a high peak in the spectrum indicates the presence of a dominant (energetic) oscillatory component in the time domain signal. The frequency value associated with this peak reflects the oscillation rate of the component. While the FT can identify the three frequency components

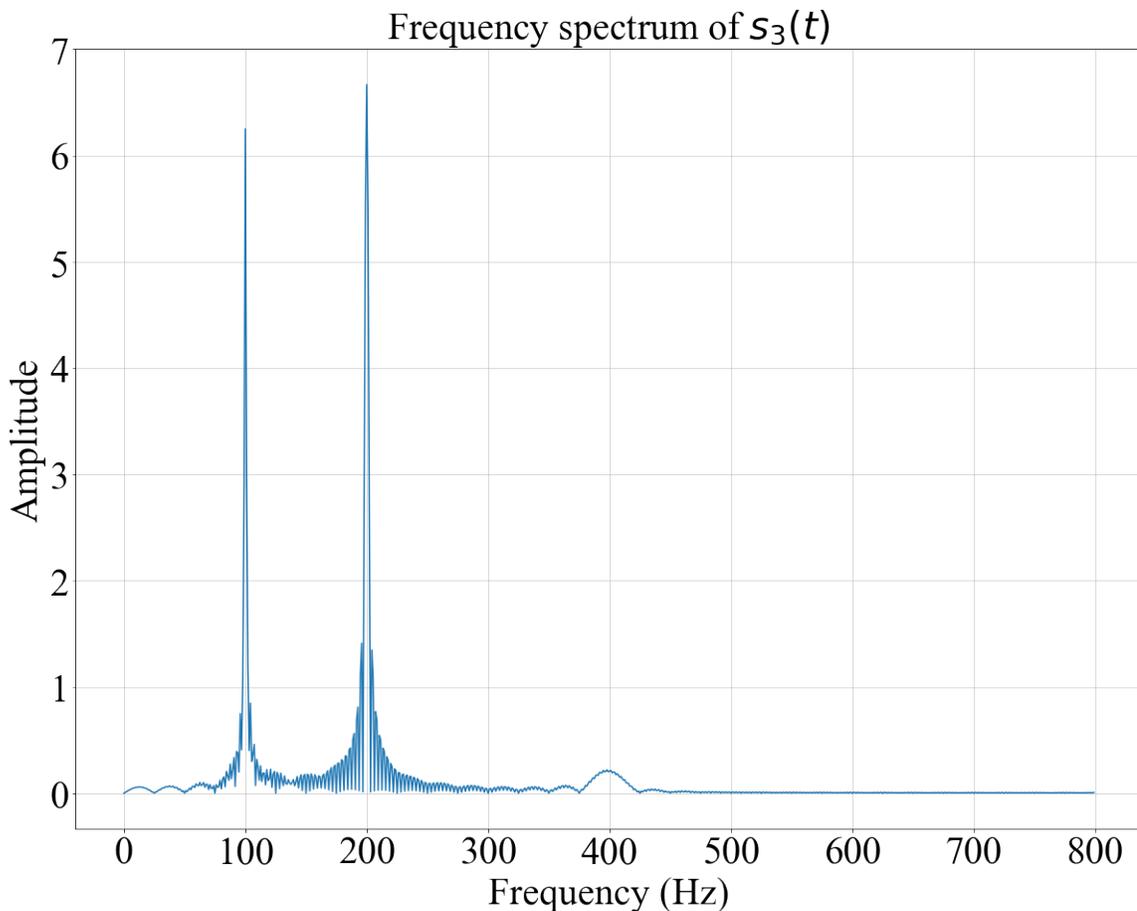


Figure 3.9: Frequency spectrum of $s_3(t)$.

within the signal, its analysis spans the signal's entire duration, which is evident from the infinite integral in the transform equation (3.28). Thus, although it accurately distinguishes between different frequency components, it cannot provide temporal information about the

timing of the occurrence of these frequencies within the signal. This lack of time resolution in the FT limits its applications to pure-frequency analysis only. Additionally, the FT is not ideal for analyzing signals with highly time-localized components, like short bursts with high energy concentration. Such components produce a broad range of frequencies in the frequency spectrum of the signal due to the inherent uncertainty principle associated with Fourier analysis [54], which states that a signal cannot simultaneously have an arbitrarily small duration in time and an arbitrarily narrow bandwidth in frequency. This means that if a signal is highly localized in time, it must spread widely in the frequency domain and vice versa. This concept is demonstrated in Fig. 3.10 and Fig. 3.11, which show a composite sinusoidal signal, denoted as $s_{m3}(t)$, and its frequency spectrum, respectively. The signal $s_{m3}(t)$ is a modified version of the sinusoidal signal $s_3(t)$ where the duration of the second component is reduced to 0.05 (The duration of the first component is increased accordingly), and its amplitude is increased to 30 to simulate localized oscillations of high energy concentration. As depicted in Fig. 3.11, the frequency of the localized component in $s_{m3}(t)$ is broadly spread across the frequency spectrum, which poses a challenge in distinguishing subtle frequency components in practical applications.

3.3.2 Short-Time Fourier Transform

The Short-Time Fourier Transform (STFT) bridges the gap between time-domain and frequency-domain analysis, making it particularly useful for analyzing nonstationary signals whose spectral characteristics evolve over time. Unlike FT, which spans the signal's entire duration for frequency analysis, STFT analyzes the signal over finite, overlapping time windows, thereby preserving temporal information. The computation of the STFT involves dividing the time-domain signal into segments of equal time length, which is achieved by multiplying the signal with a sliding window function. The FFT is then computed for each segment, providing frequency bins and corresponding Fourier coefficients for each

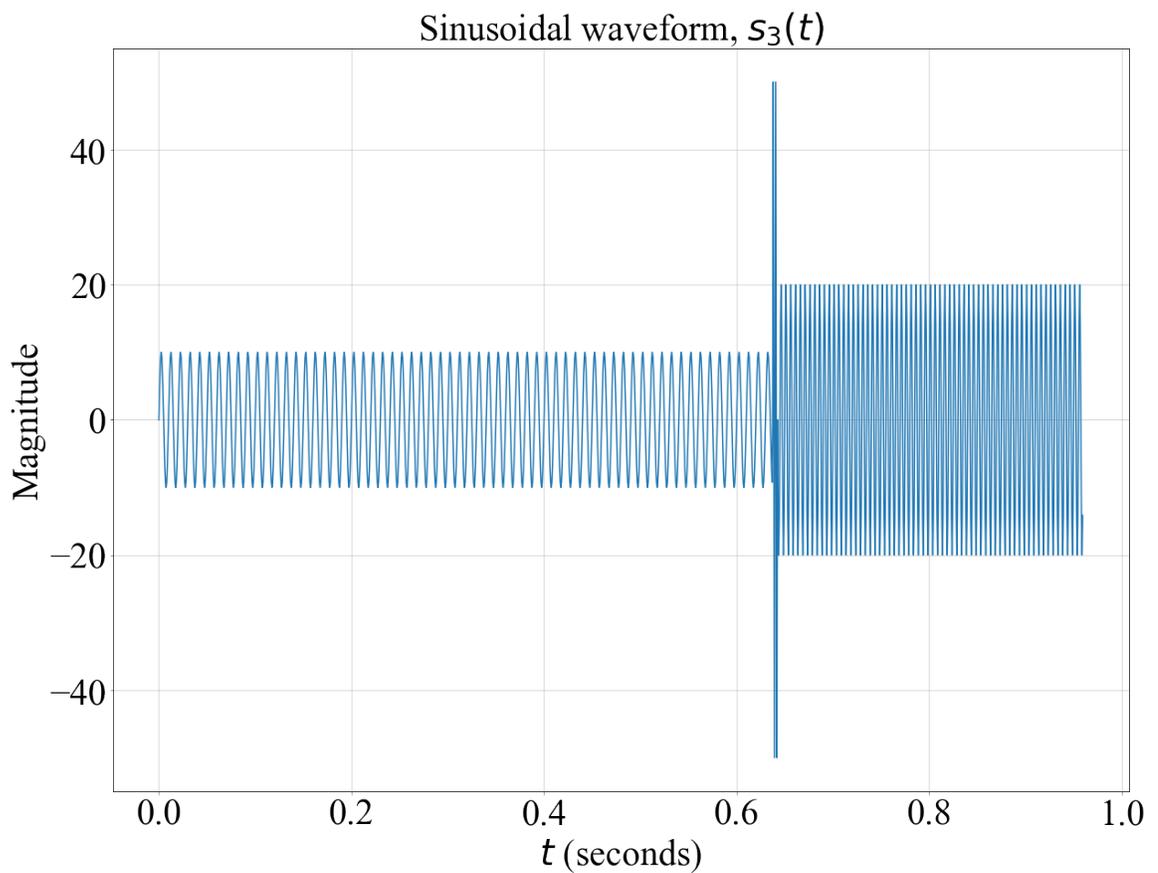


Figure 3.10: Composite sinusoidal signal $s_{m3}(t)$.

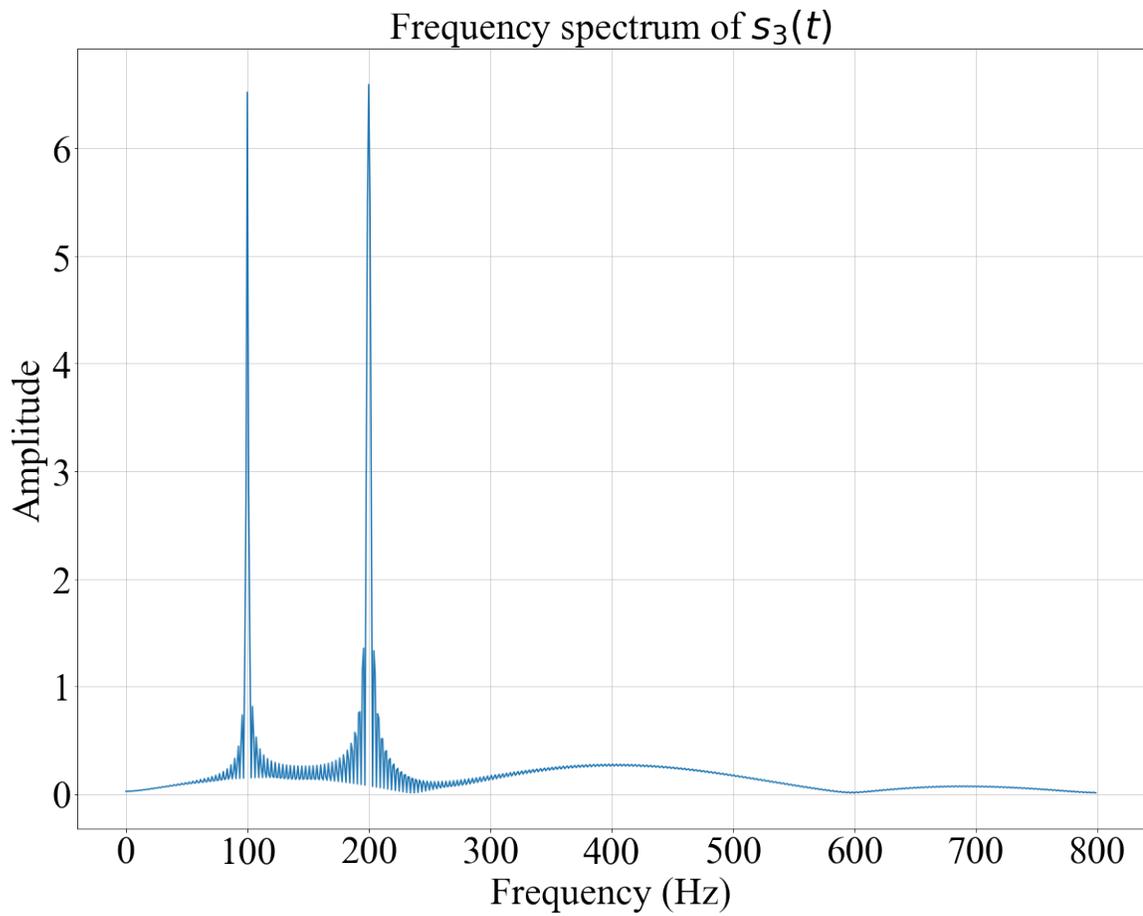


Figure 3.11: Frequency spectrum of $s_{m3}(t)$.

segment. For a continuous signal, $x(t)$, the STFT is mathematically defined as:

$$STFT\{x(t)\}(\tau, \omega) = \int x(t)w(t - \tau)e^{-j\omega t} dt \quad (3.32)$$

where $w(t - \tau)$ is the window function centered at time τ , and ω is the angular frequency.

For discrete signal $x[n]$, the discrete STFT is defined as:

$$DSTFT[n, k] = \sum_{m=0}^{N-1} x[n + m]w[m]e^{-j2\pi\frac{km}{N}} \quad (3.33)$$

where,

N is the size of the window, defining the number of signal samples included in each signal's segment for analysis,

n is the time index around which the window is centered,

k is the frequency bin index,

$w[n + m]$ represents the signal samples within the window,

$w[m]$ is the window function of size N ,

and $e^{-j2\pi\frac{km}{N}}$ is the DFT kernel.

The SciPy Python library provides an efficient implementation of SFTF and its inverse transformation. The output of the STFT function contains three sets:

- The temporal set contains the duration of each segment in seconds. Its size equals the number of segments.
- The frequency set that contains groups of frequency bins corresponding to each time segment. The size of the set is equal to $\frac{n}{2} + 1$.
- Coefficients set: A matrix containing Fourier coefficients in each frequency bin for every segment. Thus, its size is $d_1 \times d_2$, where d_1 and d_2 are sizes of frequency and temporal sets, respectively. STFT features are usually obtained by examining the coefficients matrix. However, since Fourier coefficients are complex numbers, ex-

tracting features in the STFT commonly involves analyzing the magnitudes of these coefficients.

Heatmaps serve as an effective tool for visualizing the STFT output, providing a clear representation of the variations in the frequency content of the signal over time. These heatmaps, commonly referred to as spectrograms, are generated by mapping the temporal set to the x -axis and frequency set to the y -axis. The “intensity” of each time-frequency pair in the $x - y$ plane is represented by the magnitude of the corresponding Fourier coefficients. Fig. 3.12 shows the STFT of the composite signal $s_3(t)$ introduced earlier. The signal is segmented into ten segments using a Hamming window with an overlap of 50% of the segment length. The Fourier coefficients are computed for each segment using an FFT length of 2048. Compared to the signal’s frequency spectrum depicted in 3.9, the STFT spectrogram shows the spectral contents and reveals their temporal characteristics simultaneously, providing time-frequency analysis of various frequency components in the signal. Specifically, The duration of each frequency component in the signal is clearly shown over the timeline (X -axis). Furthermore, the temporal edges, represented by the start and end times of each component, are well-defined, thereby reflecting high temporal resolution that allows the identification of the transient component at 400 Hz. In real-world scenarios, the efficiency of the STFT in time-frequency analysis is significantly influenced by the choice of its parameters, particularly the segment length and the percentage of overlap between adjacent segments. Segment overlap provides continuity and reduces information loss between adjacent segments. A high degree of overlap introduces more redundancy into the segments and improves time resolution, which is particularly useful when a large segment length is used. The segment length determines the resolution of the time-frequency analysis. A smaller segment length enables better time resolution, which helps identify rapid transient patterns in the signal, as demonstrated in Figure 3.12. On the other hand, a larger segment length enhances frequency resolution by allowing more frequency cycles to fit within the segment. However, this comes at the cost of reduced temporal resolution, as

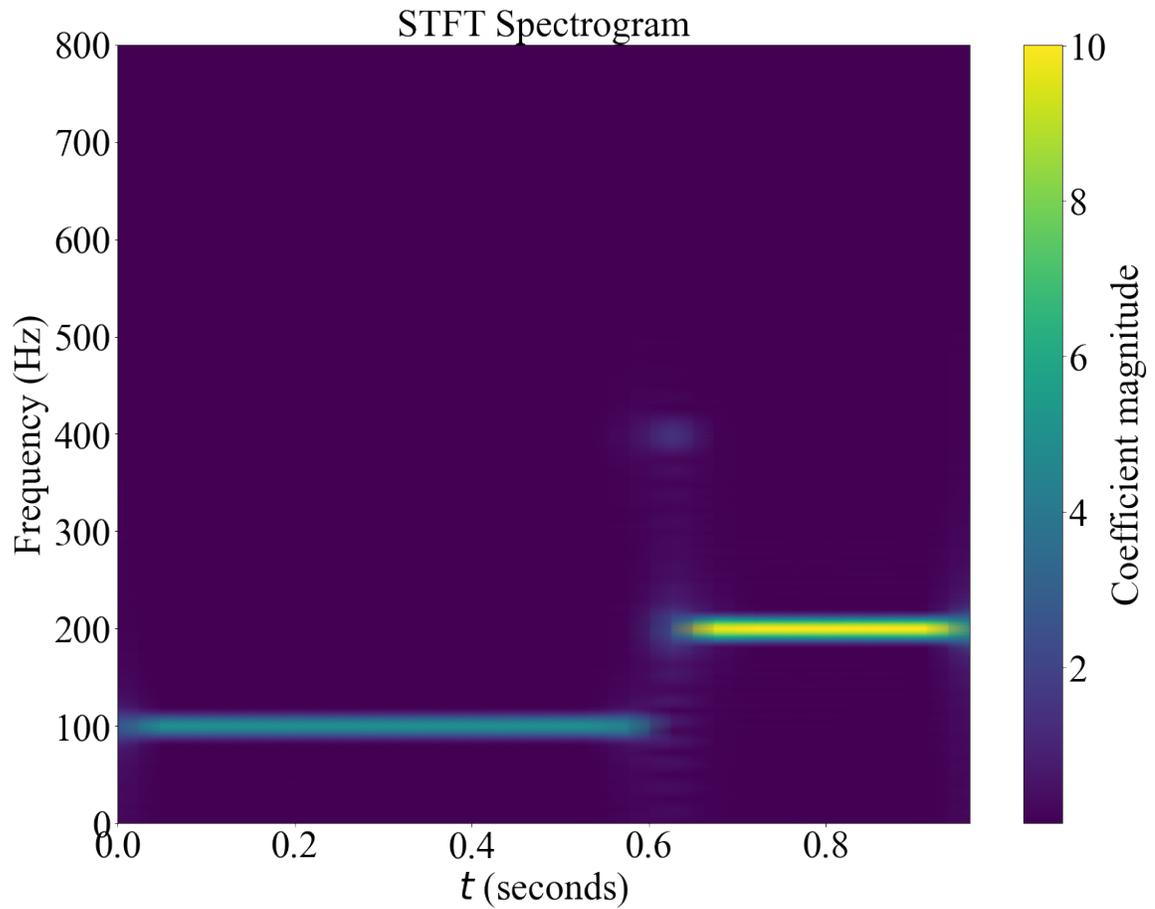


Figure 3.12: STFT spectrogram of the composite sinusoidal signal $s_3(t)$.

transient changes in the signal over short time durations may be overlooked. Therefore, there is a trade-off between time and frequency resolutions in the STFT analysis. This concept is demonstrated in Fig. 3.13, which shows the STFT spectrogram of $s_3(t)$ re-computed by dividing the signal into three segments only instead of ten segments. The frequency resolution is represented by the “thickness” of the timeline. A thinner timeline—reflecting a finer frequency resolution—allows precise identification of the corresponding frequency bin on the y-axis compared to a thick timeline that spans several frequency bins. A comparison of the two spectrograms shows that the use of a longer segment improves the frequency resolution but reduces the temporal resolution, resulting in overlapping temporal edges that obscure the identification of the transient component. The fixed segment length is a major

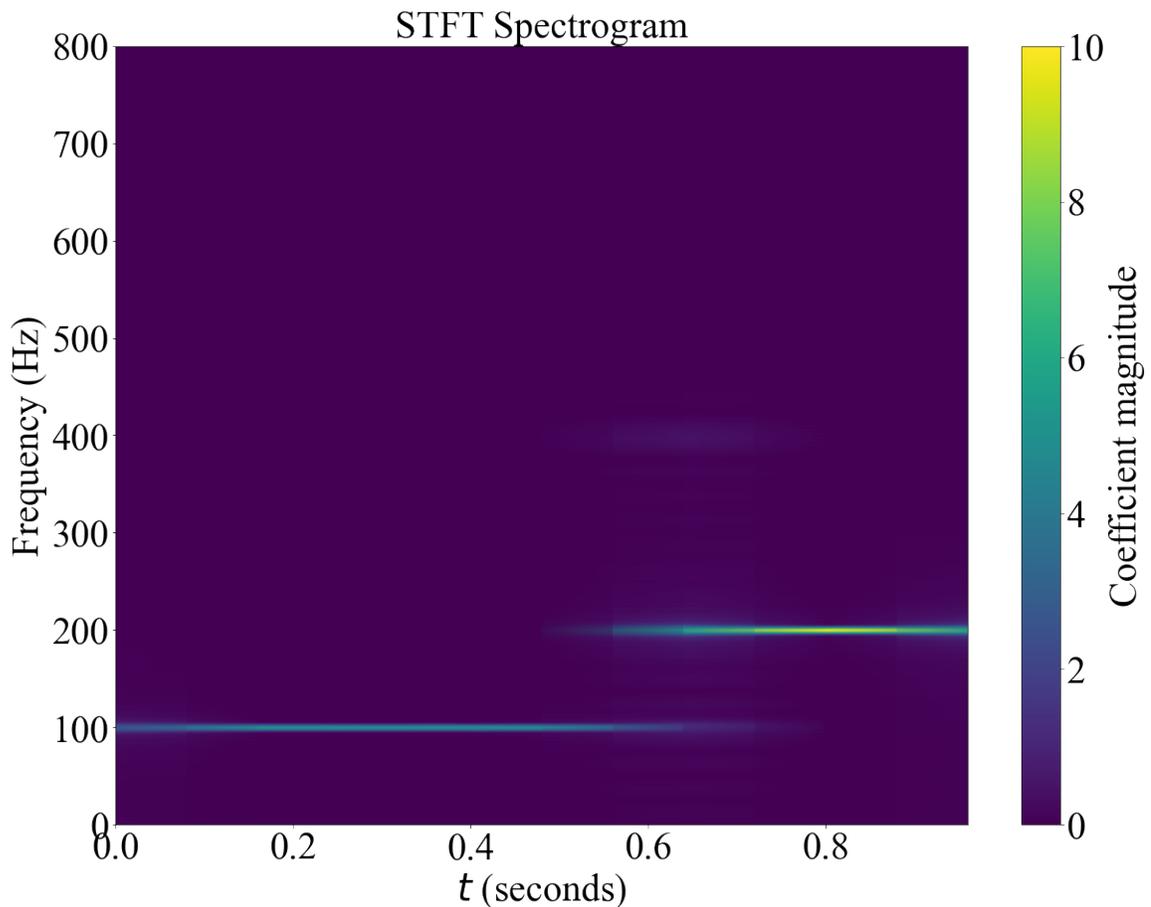


Figure 3.13: STFT spectrogram of the composite sinusoidal signal $s_3(t)$.

limitation in STFT that results in a uniform resolution analysis of the signal and, hence,

leads to an inherent compromise between time and frequency resolutions when applying STFT for time-frequency analysis.

3.3.3 Power Sepctral Density

Power Spectral Density (PSD) describes how the power of a signal is distributed across different frequencies, providing a measure of the power per unit frequency. This is achieved through the normalization of the power spectrum by frequency resolution (frequency bin width), converting the power spectrum into a power spectral density that shows the power distribution per unit frequency. Mathematically, the PSD $S_{xx}(f)$ of a continuous-time signal $x(t)$ is defined as the FT of the autocorrelation function $R_{xx}(\tau)$; it is expressed as:

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f\tau} d\tau \quad (3.34)$$

where,

$$R_{xx}(\tau) = E[x(t)x(t + \tau)] \quad (3.35)$$

and $E[\cdot]$ denotes the expected value. For discrete signals, the discrete PSD (DPSD) is utilized, involving the DFT of the autocorrelation sequence. The DPSD $S_{xx}[k]$ of a discrete-time signal $x[n]$ is given by:

$$S_{xx}[k] = \sum_{n=-\infty}^{\infty} r_{xx}[n] e^{-j\frac{2\pi}{N}kn} \quad (3.36)$$

where $r_{xx}[n]$ is the autocorrelation sequence of $x[n]$.

The SciPy Python library provides two methods for PSD estimation: the periodogram method (`scipy.signal.periodogram`) and the Welch method `scipy.signal.welch` [55]. The periodogram estimation involves obtaining the spectral power of the signal by applying FFT and squaring the magnitude of the resultant FFT coefficients. The PSD estimate is obtained by normalizing the spectral power by the frequency resolution of the FFT. One of the limi-

tations of the periodogram method is its high variance, especially when the number of samples in the signal is limited. The Welch method overcomes this limitation by segmenting the signal into overlapping segments using a sliding window, computing the periodogram for each segment, and averaging the periodograms to obtain the PSD estimate. This effectively reduces noise and provides a smoother estimate than the periodogram method. Fig. 3.14 compares the two methods in estimating PSD of the random vibration signal $v(t)$ introduced earlier. In the Welch method, the signal is divided into four segments with 50% of overlap. It is evident that averaging the estimate over several segments greatly reduces the variance of the estimate, resulting in a smoother PSD. The smoothness of the PSD is of significant importance in power spectral analysis and feature extraction since high estimation variance can hinder major characteristics in the PSD, such as peaks. A smooth PSD allows precise identification of such characteristics, improving the sensitivity of the extracted features.

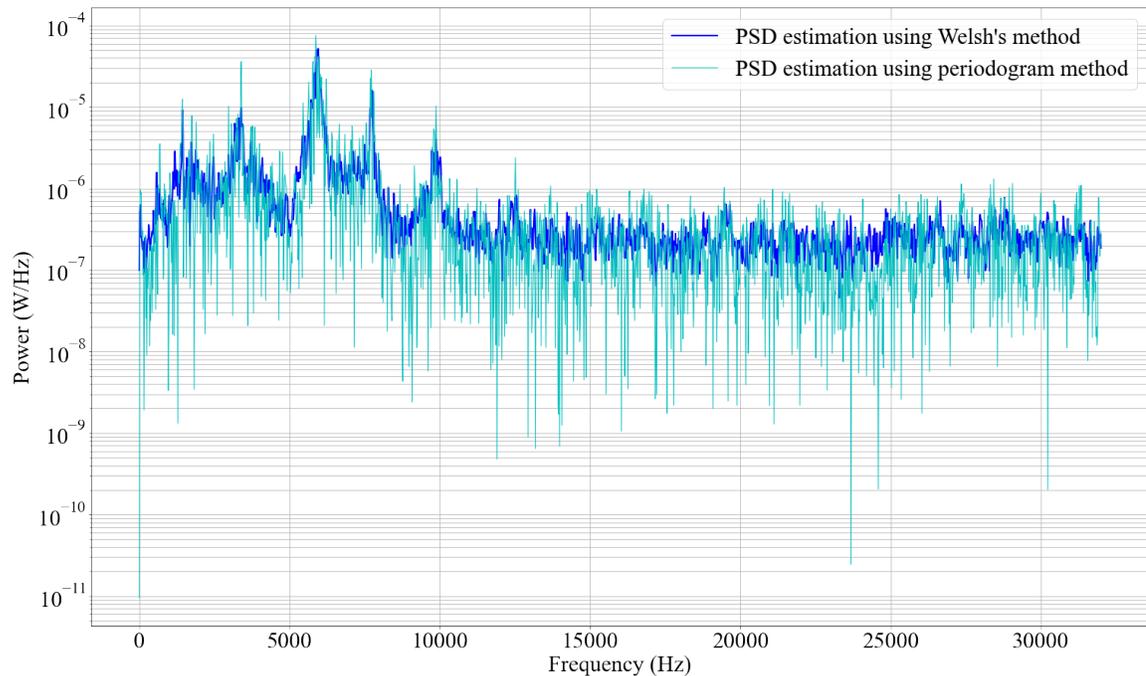


Figure 3.14: Periodogram and Welch's PSD estimations of vibration signal $v(t)$.

Similar to the STFT, the choice of parameters in the Welch method, such as window type,

segment length and overlap, has a significant effect on the PSD estimate. While STFT and the PSD possess similar aspects, they serve different purposes and are based on distinct theoretical foundations. Understanding their differences is crucial, especially in ML and feature extraction applications. Table 3.1 compares STFT and PSD and summarises their main characteristics.

Table 3.1: Comparison between STFT and PSD

Properties	STFT	PSD
Purpose and Concept	STFT is used for time-frequency analysis of signals, especially nonstationary signals where frequency components vary over time	PSD is used to describe the power distribution of a signal across frequency. It is particularly useful in understanding the energy content of different frequency components of a signal
Representation	The result of STFT is a two-dimensional representation of the signal, showing both time and frequency dimensions. This is crucial in analyzing signals whose spectral properties evolve over time	PSD provides a one-dimensional frequency domain representation, indicating how power is distributed across frequencies. Unlike STFT, it does not offer time-domain information
Characteristics	STFT involves a trade-off between time and frequency resolution controlled by the window size. A larger window offers better frequency resolution but poorer time resolution, and vice versa	PSD assumes signal stationary, making it less suitable for analyzing signals with time-varying characteristics
Application	STFT is used to analyze signals with time-varying spectral properties	PSD is more suited for stationary signals or for assessing overall frequency content

3.3.4 Wavelet Transform

The accuracy of time-frequency analysis depends heavily on the localization of the base function in time and frequency domains. In Fourier analysis, the basis functions are sinusoids (complex exponentials) that are perfectly localized in the frequency domain since each corresponds to a single frequency component with no frequency spread. In the time domain, these sinusoidal functions extend infinitely, lacking time localization due to their infinite duration. In the context of wavelets, signal analysis is based on a family of wavelets that are localized in both time and frequency domains. These wavelets are generated from a single mother wavelet through scaling and translation, facilitating signal analysis at different frequencies and time intervals. A wavelet is a mathematical function that satisfies certain conditions, such as finite energy and a zero mean. The term “wavelet” itself means a small wave, which captures the essence of these functions and how they behave. The

main characteristics of a wavelet function include the following:

- **Zero Mean:** This ensures the wavelet oscillates around the zero level; this is crucial for detecting changes in the signal since it allows it to capture both positive and negative variations in the signal effectively.
- **Finite Energy:** This implies that the total energy of a wavelet is finite, ensuring integrability and existence of the inner product with the signal. The energy of a given wavelet function is calculated by integrating the squared amplitude of the wavelet over its entire duration. Mathematically, the total energy of a wavelet function, $\psi(t)$, is expressed as:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (3.37)$$

A wavelet with finite energy is inherently localized in time and has compact support, meaning it is non-zero only over a limited range, and its oscillations are confined to a small region. This time localization allows wavelets to effectively represent signals that have nonstationary or transient components. A wavelet's integrability implies that it won't produce infinite results when integrated over its domain. This is crucial for signal analysis since it is essentially carried out by computing the inner product of the signal with scaled and transformed versions of the wavelet.

Wavelet analysis is conducted through two operations, scaling and translation:

- **Scaling “Dilation”:** This involves stretching or compressing the wavelet. Scaling changes the frequency content of the wavelet, allowing the signal to be analysed in different frequency bands. Specifically, the oscillation frequency of a wavelet is directly affected by the scale. At lower scales, the wavelet compresses and oscillates more rapidly, which allows it to detect high-frequency components or rapid changes in the signal. On the other hand, at higher scales, the wavelet is stretched, resulting in fewer oscillations for the same amount of time, making it more sensitive to low-

frequency components or slow-changing features in the signal.

- Translation “Shifting”: This refers to moving the wavelet along the time axis, enabling the wavelet to analyze different signal parts at a given scale.

These aspects allow high time-frequency localization and provide a multi-resolution analysis through varying scales and translations. The choice of the mother wavelet and scale values is derived by signal characteristics and application-specific requirements. Attributes such as orthogonality, compact support, vanishing moments, and the ability to capture signal discontinuities and singularities play a crucial role in this choice. The wavelet family includes various mother wavelets; each with unique properties that cater to various signal types and application requirements. Fig. 3.15 displays the time-domain waveforms of some of the commonly used wavelet functions.

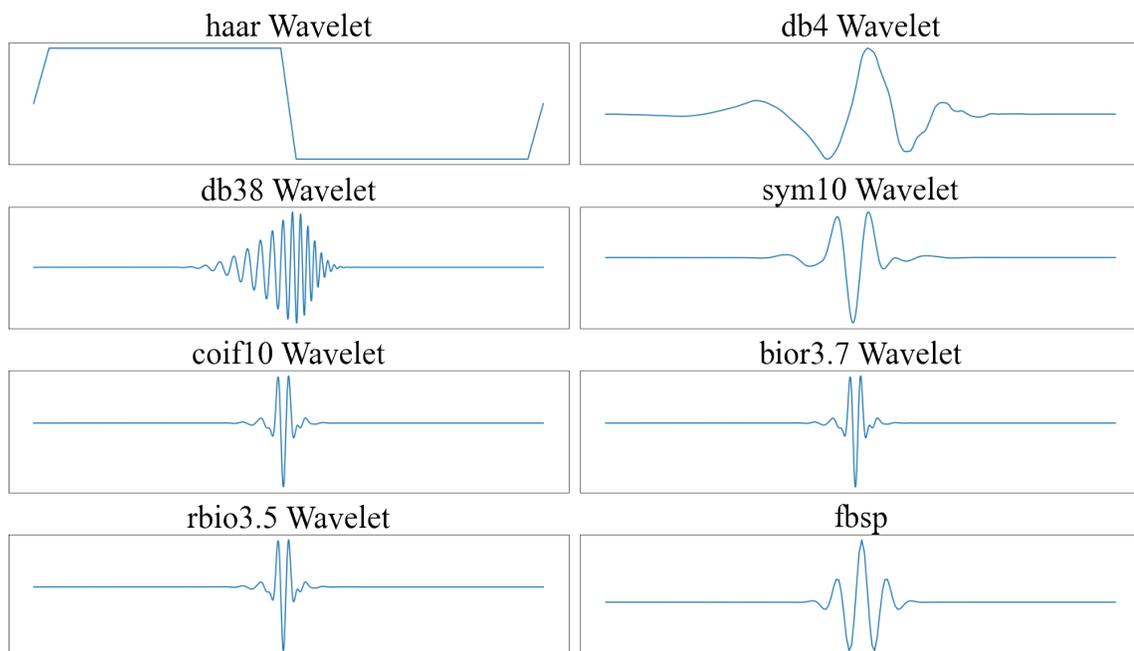


Figure 3.15: Some of the commonly used wavelet functions.

The Continuous Wavelet Transform (CWT)

The CWT of a signal $x(t)$ can be expressed mathematically as follows:

$$CWT_x(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (3.38)$$

where,

a is the scale parameter, b is the translation parameter,

$CWT_x(a, b)$ represents the wavelet coefficient at scale a and translation b ,

$\psi(t)$ is the mother wavelet, and $\psi^*(t)$ is its complex conjugate.

The output of CWT is a matrix of wavelet coefficients, where each row contains the coefficients of the corresponding scale. These coefficients represent the correlation between the signal and a scaled and translated wavelet, indicating how closely the signal matches the wavelet at those specific scales and translations. This emphasizes the importance of selecting a proper wavelet function for signal analysis. The CWT output is commonly visualized through 2D heatmaps, known as scalograms, where the x -axis and y -axis represent time and scale, respectively. The intensity of the heatmap is represented by the magnitude of the wavelet coefficients, which reflect the signal wavelet energy at different scales and translations. The scale parameter inversely relates to the signal's frequency content. High scales correspond to low frequencies, capturing the slower-varying components of the signal. On the other hand, low scales, corresponding to high frequencies, capture rapidly changing components within the signals. Hence, scalograms show how the spectral content varies over time, enabling the identification of transient or nonstationary components within the signal. Fig.3.16 displays the scalogram of the composite signal $s_3(t)$ using a scale range from 1 to 135. The $fbsp$, displayed in the bottom right corner of Fig. 3.15, is employed in the analysis; $fbsp$ is a spline wavelet constructed using a spline function. The selection of this wavelet is based on its sinusoidal-like waveform, resembling the waveform of the composite signal $s_3(t)$. The scalogram demonstrates a multi-resolution signal analysis

at various scale values. In particular, scales greater than 120 effectively capture the slowly varying components (components 1 and 3) of the signal. On the other hand, the rapid change in the signal, represented by component 2, can only be detected at small-scale values (scales between 30 and 40).

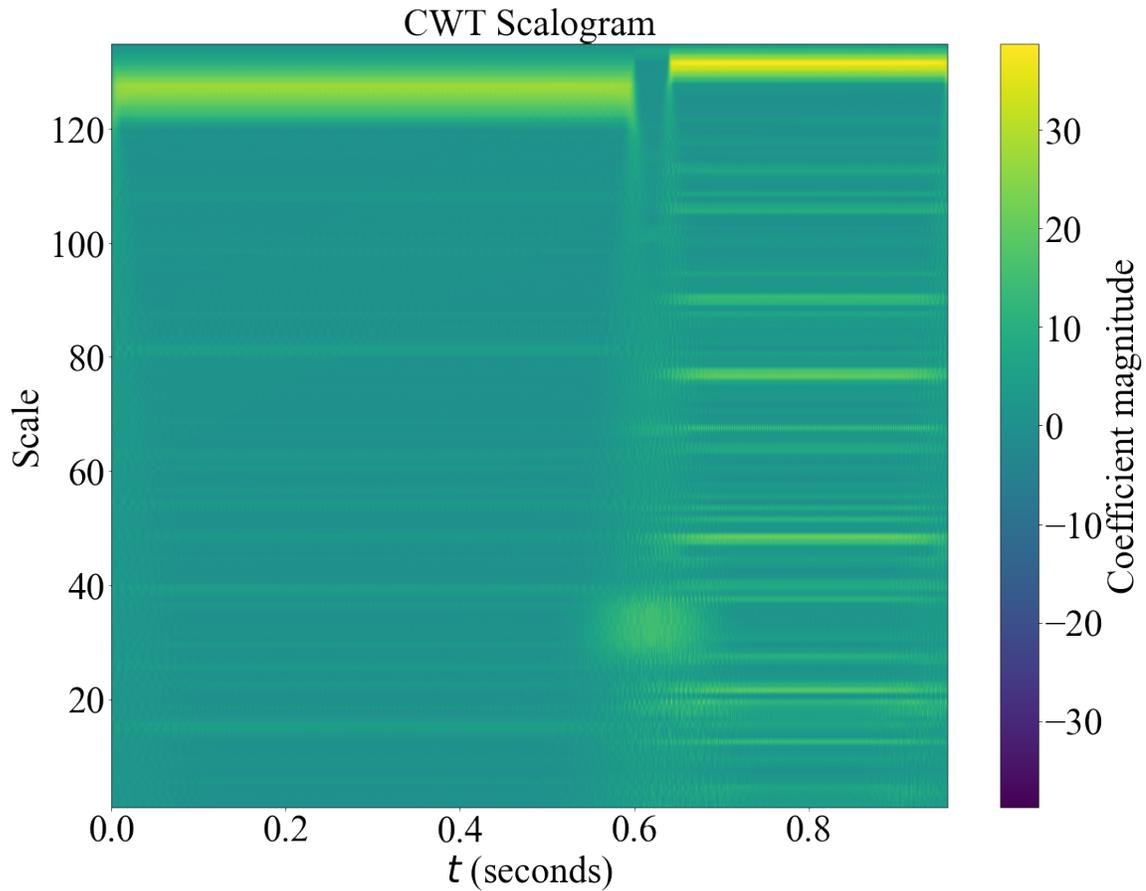


Figure 3.16: CWT “scalogram” of the composite signal $s_3(t)$.

Discrete Wavelet Transform (DWT)

DWT is a wavelet-based signal decomposition tool that uses discrete values for scale and translation, reducing redundancy and computational complexity. In DWT, wavelet multi-resolution analysis is achieved by decomposing the signal into high and low-frequency components, commonly known as elementary modes, using digital filter banks and down-sampling operations. The Key parameters include the type of mother wavelet and the de-

gree of decomposition. Mathematically, the DWT of a discrete signal $x[n]$ can be expressed as:

$$DWT_x[j, k] = \sum_{n=-\infty}^{\infty} x[n]\psi_{jk}[n] \quad (3.39)$$

where,

$DWT_x[j, k]$ represents the wavelet coefficient at the j -th level of decomposition and the k -th position,

$\psi_{jk}[n]$ are the discrete wavelets, defined as scaled and shifted versions of the mother wavelet.

In the decomposition process, the signal undergoes convolution with consecutive low-pass and high-pass filters followed by down-sampling. The main aspects and steps of the decomposition process include:

Mother Wavelet: The first step involves selecting an appropriate wavelet base function, often called the “mother wavelet.”

Scaling and Translation: The decomposition process starts by scaling and translating the mother wavelet to generate a family of wavelet functions that serve as bases to decompose the signal. In the DWT, the mother wavelet is used in a discretized form where scaling and translation are done in discrete steps, typically in powers of two for scaling and integer multiples of that scale for translations. For each level of decomposition, the wavelet is scaled by a factor of 2^j , where j indicates the level of decomposition. This scaling effectively compresses the wavelet to capture signal features at different frequencies. Consequently, the wavelet is translated along the signal in steps that are multiples of the current scale factor 2^j . This stepwise translation allows the wavelet to cover the entire signal, ensuring that all signal parts are analyzed.

High-Pass Filter: This filter represents the mother wavelet since its coefficients are essentially a discrete representation of the mother wavelet. As a high-pass filter, it extracts high-frequency *detail* components from the signal, capturing rapid changes within the signal.

Low-Pass Filter: This filter represents the scaling function, as its coefficients are derived

from the scaled wavelets. As a low-pass frequency filter, it extracts low-frequency *approximate* components that vary slowly over time.

Down-Sampling: After separating the signal into high and low-frequency components, each component is down-sampled by a factor of two. Down-sampling involves keeping every second sample of the filtered signals. This process reduces the sample rate by half, effectively decreasing the data size and ensuring the transform is computationally efficient.

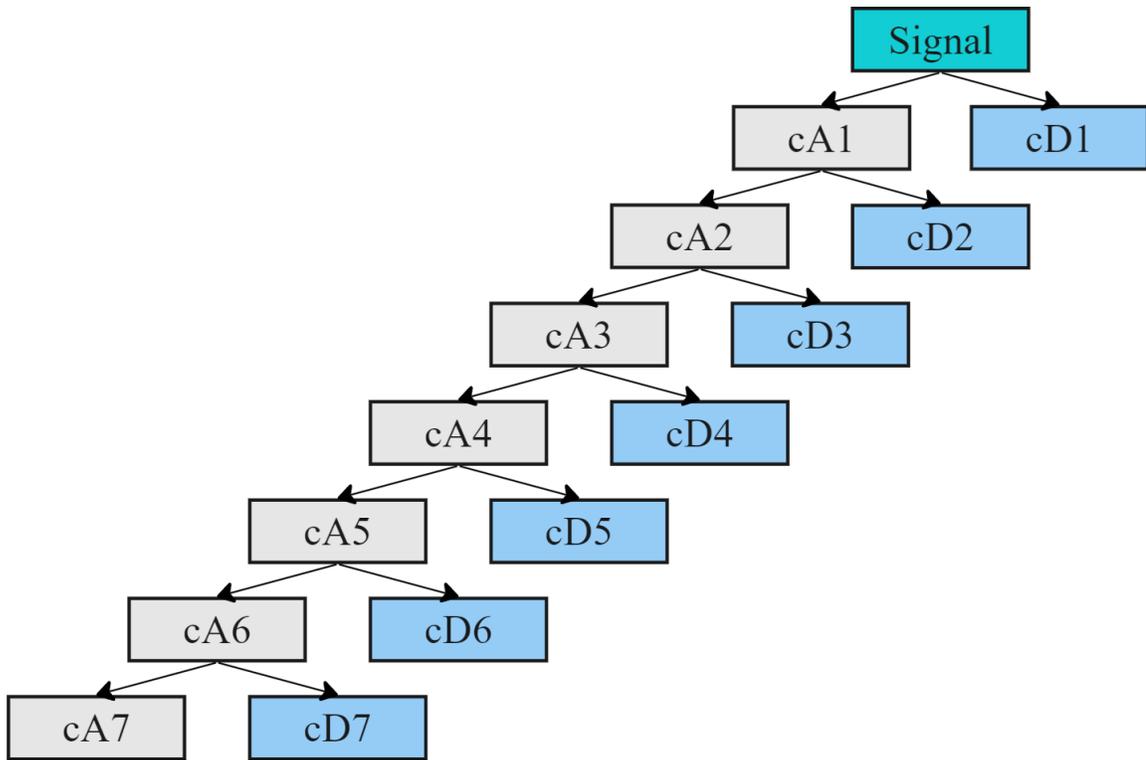
Approximate and Detail Coefficients: Approximation and detail coefficients are outputs of down-sampling operations performed on approximate and detail components, respectively. The approximate coefficients represent a smoothed and down-scaled version of the original signal, while the detail coefficients describe high-frequency content in the signal.

Iterative Analysis: The process of high and low pass filtering followed by down-sampling is performed iteratively on the resultant *approximation coefficients only* until reaching the specified number of levels (j) or until reaching maximum decomposition constrained by signal length. This process creates a j -level tree structure of wavelet coefficients, shown in Fig.3.17.

Signal reconstruction: The original signal can be perfectly reconstructed by applying the inverse DWT, which processes the resultant coefficients through reconstruction filters in an inverse manner to the decomposition process.

Elementary modes: These modes are obtained through the individual reconstruction of the resultant coefficients “coefficients $cD1$ to cDj and cAj as shown in Fig. 3.17”, allowing the processing of the composite signal at the level of its constituent modes. Since these modes form the underlying structure that constitutes the original signal, perfect reconstruction of the original signal is possible through direct summation of these modes.

Fig. 3.18 displays elementary modes of the random vibration signal $v(t)$ along with their corresponding spectra. These modes are obtained through 7-level DWT decomposition of the signal using Daubechies *db4* wavelet (shown in the top right corner of Fig. 3.15). The Daubechies family of wavelets, especially those with higher orders (e.g., *db4*, *db5*), are



cA_n : Approximation coefficients of level n

cD_n : Detail coefficients of level n

Figure 3.17: 7-level DWT decomposition tree.

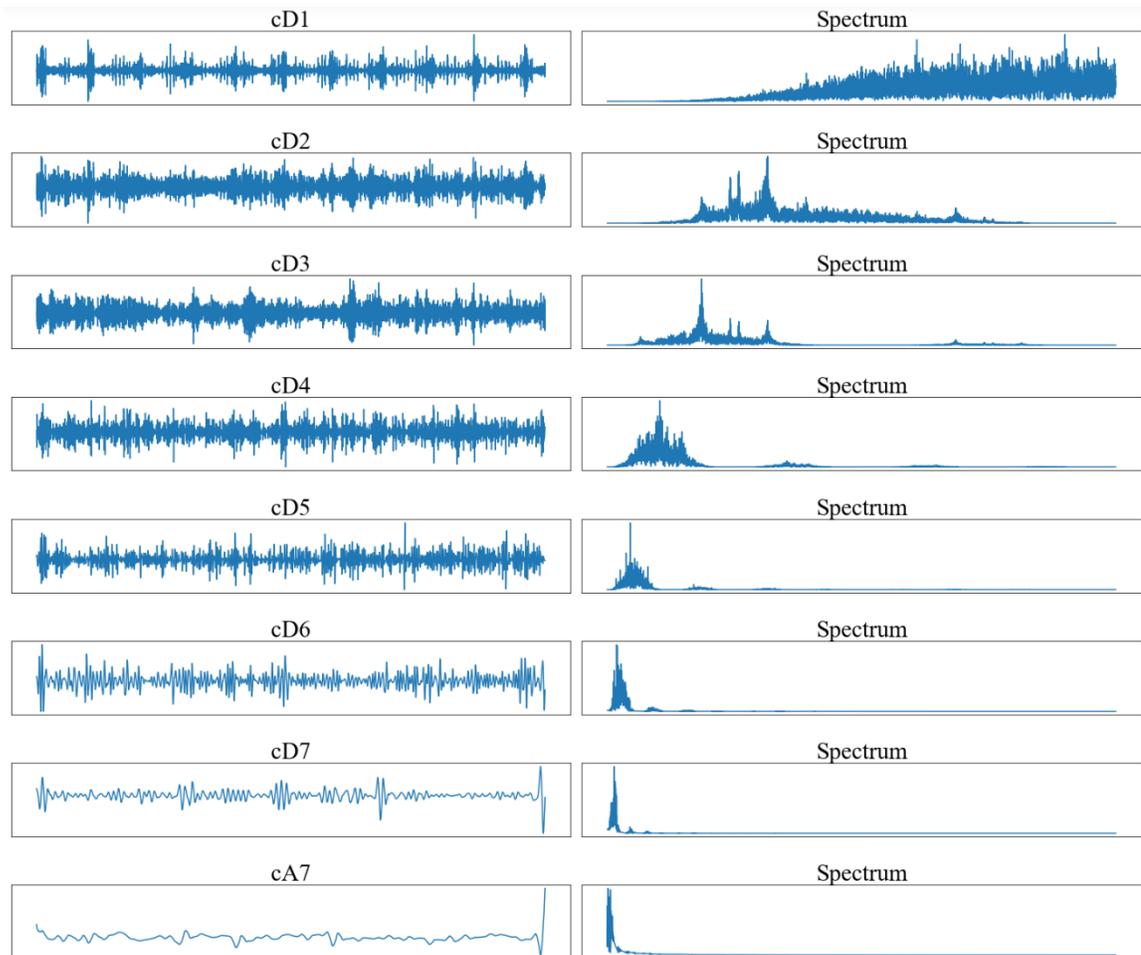


Figure 3.18: 7-level DWT decomposition of random vibration signal $v(t)$: Elementary modes and their corresponding frequency spectra.

compactly supported and well-localized in time, making them highly effective for analyzing signals that contain transient patterns or nonstationary components. This is particularly useful in vibration analysis [56], where detecting sudden changes or faults is essential for condition monitoring and diagnostics. The DWT decomposition of $v(t)$ iteratively extracts slowly varying modes of the signal, corresponding to lower frequencies in the signal at each level. This progressively reduces the resolution of high-frequency details with each level of decomposition, making DWT well-suited for signals where the information of interest is concentrated in lower frequencies.

Stationary Wavelet Transform (SWT)

A major limitation of the DWT is the lack of time invariance, which stems primarily from the down-sampling process. The down-sampling performed at each level makes the output of DWT sensitive to shifts in the input signal. Even small shifts can cause significant changes in the resulting wavelet coefficients, which is problematic in applications where the precise timing of signal features is essential. Furthermore, down-sampling reduces the signal length by half at each level of decomposition, which may lead to the loss of specific signal details, especially at higher decomposition levels. SWT is an extension of the DWT that addresses these limitations by omitting the down-sampling step, thus maintaining the original size of the signal throughout the decomposition levels. Unlike DWT, which provides a critically sampled (non-redundant) representation, SWT provides a more detailed but redundant representation of the signal. SWT is implemented by applying high-pass and low-pass filters without downsampling the filtered signals. Instead, the filters are up-sampled at each level to match the size of the original signal. The up-sampling of filters is done by inserting zeros between the filter coefficients, effectively increasing the size of the filters as the decomposition level increases. This leads to an overcomplete signal representation and increases computational burden compared to DWT.

Wavelet Packet Transform (WPT)

The WPT extends the capabilities of classical wavelet-based decomposition. Unlike the DWT, which limits its decomposition to approximation coefficients, the WPT performs a full decomposition at each level by processing detail and approximation coefficients, as depicted in Fig. 3.19, thereby capturing both low-frequency and high-frequency components. While this process is more computationally intensive, it enhances frequency resolution in the analysis, enabling the detection of subtle signal features that DWT may overlook. Fig.

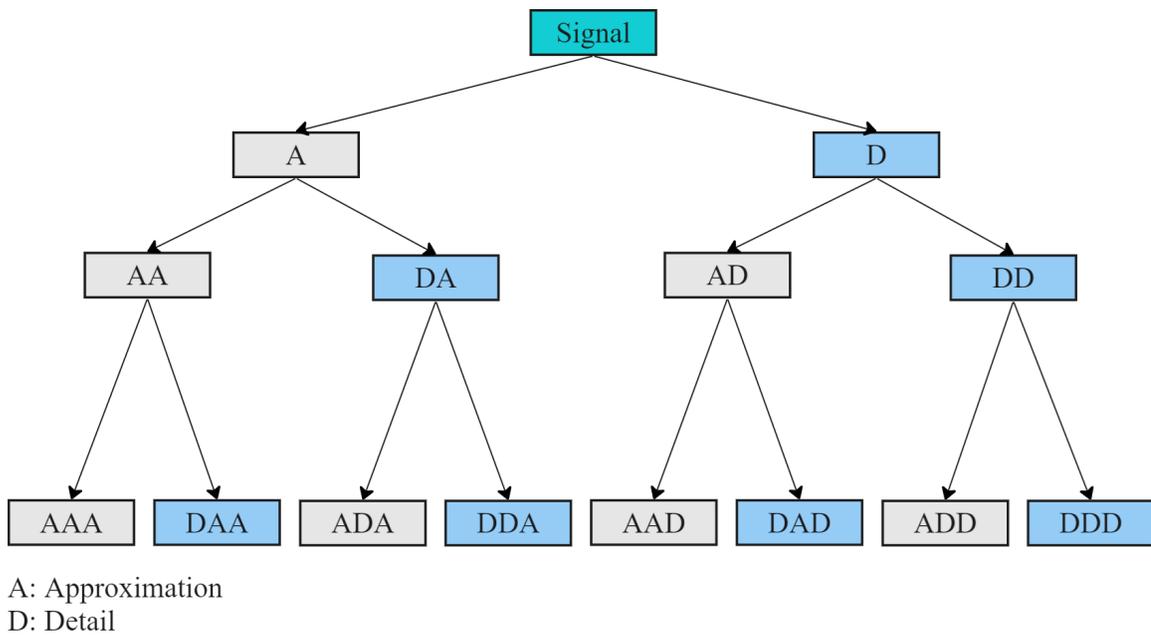


Figure 3.19: 3-level WPT decomposition tree.

3.20 displays elementary modes of the vibration signal $v(t)$, obtained through 3-level WPT decomposition using the *db4* wavelet, along with their corresponding spectral contents. In contrast to DWT, which offers a hierarchical frequency decomposition with good frequency resolution at lower frequencies, WPT provides a more uniform time-frequency localization across the spectrum due to its symmetric decomposition of detail and approximation coefficients at each level. Hence, it maintains a balance between time and frequency resolution throughout the frequency spectrum, allowing for a more precise localization across a wider range of frequencies. Additionally, the spectrum in DWT is skewed towards lower fre-

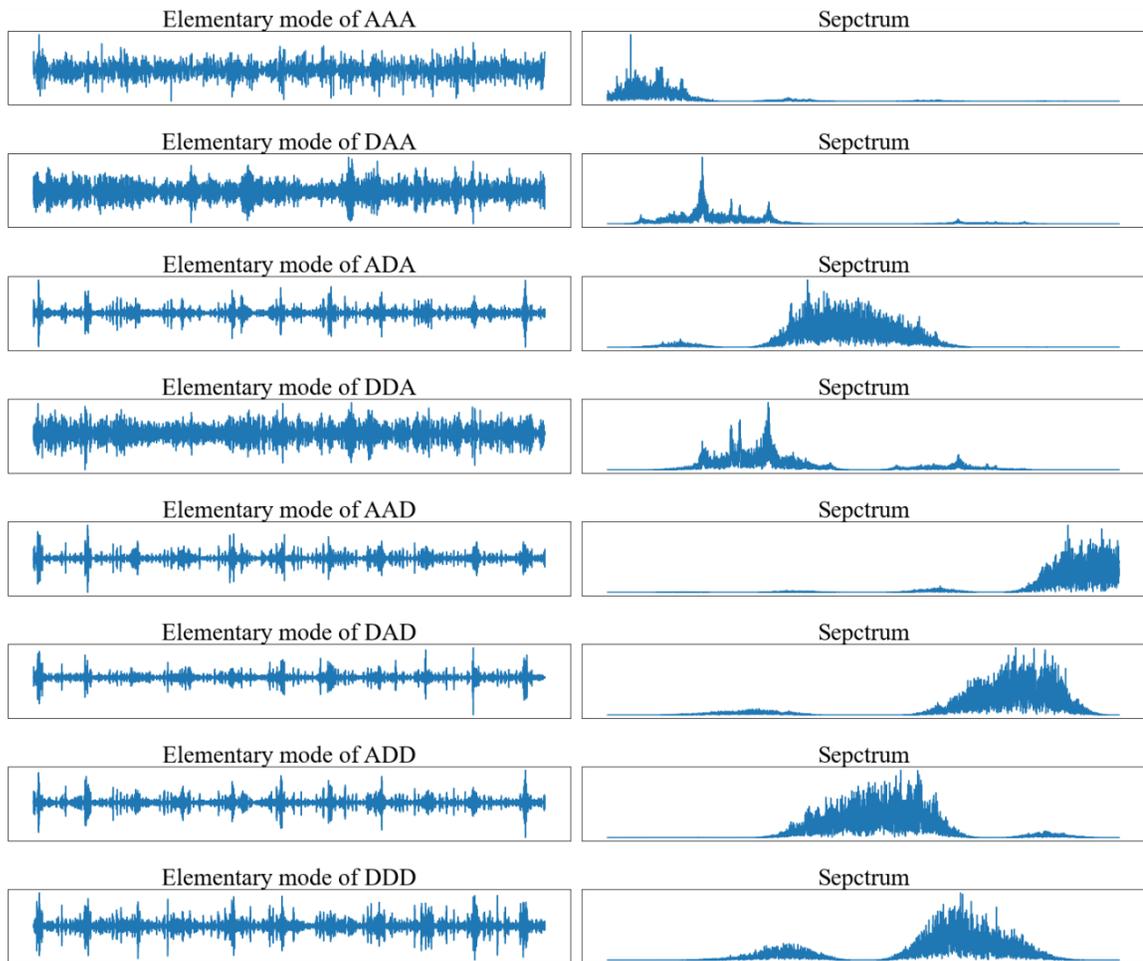


Figure 3.20: 3-level WPT decomposition of random vibration signal $v(t)$: Elementary modes and their corresponding frequency spectra.

quencies with each successive level of decomposition. While this can be advantageous for analyzing signals characterized by low-frequency components, it inadequately represents higher-frequency details. WPT, on the other hand, provides a balanced spectrum with a more equitable representation of both high and low-frequency components.

Wavelet decomposition offers an effective tool for analyzing complex real-world signals that involve transient and nonstationary components, opening the door for a wide range of applications such as:

- **Multi-resolution analysis:** Wavelet decomposition methods attempt to decompose the signal into elementary modes characterized by their high time-frequency localization, offering the capacity for conducting a multi-resolution analysis of the original signal.
- **Feature Extraction:** The obtained modes encompass meaningful information as they exhibit distinct amplitude and frequency characteristics. Accordingly, distinctive features directly related to the inherent structure of the signal can be extracted from these modes or their corresponding coefficients.
- **Noise Reduction:** Noise and signal components often have different characteristics in the wavelet domain. By thresholding certain wavelet coefficients before signal reconstruction [57], noise can be effectively reduced or eliminated.
- **Signal Enhancement:** Wavelet reconstruction can be utilized to enhance signals, where certain characteristics can be amplified or smoothed out as required, improving the quality of the reconstructed signal.
- **Signal Compression:** By keeping only the most significant wavelet coefficients for reconstruction, signals can be compactly represented with minimal loss of quality.

As previously stated, selecting the appropriate mother wavelet and decomposition level are crucial steps in wavelet analysis. Different wavelets have different shapes and properties, such as smoothness, symmetry, and the number of vanishing moments, making the

selection mainly dependent on application requirements and the shape similarity between the wavelet and the signal. The decomposition level determines the resolution at which the signal is analyzed. Higher decomposition levels allow for the analysis of lower-frequency components but also increase computational complexity and could dilute the significance of higher-frequency components. The maximum level of decomposition in DWT and WPT is typically limited by the length of the signal, N . Generally, each level of decomposition reduces the number of data points by half. Thus, the number of maximum possible decomposition levels equals $\log_2(N)$. In SWT, the maximum decomposition level is determined by the signal's length and the filters' length at each decomposition level, which increases with each level due to the up-sampling operation. Further decomposition becomes impractical when the filters become too long relative to the signal.

The Python PyWavelets package provides comprehensive modules for wavelet analysis. It supports custom wavelets and provides various wavelet functions such as CWT, DWT, SWT, WPT and over 100 built-in wavelet filters. Other libraries supporting wavelet analysis include PyCWT, a Python module for continuous wavelet spectral analysis, and `scipy.signal`, which provides CWT and other wavelet functions, particularly for signal filtering.

3.3.5 Hilbert Transform

The Hilbert Transform (HT) is a fundamental operator in signal theory; it is particularly useful in obtaining the analytic signal representation of real-valued signals. An analytic signal is a complex-valued representation of the signal that provides a comprehensive way to describe both its amplitude and phase characteristics. The HT of a signal $x(t)$ is defined as:

$$H\{x(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (3.40)$$

The transform essentially modifies the phase of each frequency component of the signal by $\pm 90^\circ$. The $X_a(t)$ of $x(t)$ is formed by augmenting the signal with its HT $H\{x(t)\}$ as the

imaginary part. Mathematically, it is expressed as:

$$X_a(t) = x(t) + jH\{x(t)\} \quad (3.41)$$

Represented in its polar form, the analytical is expressed as:

$$x_a(t) = A(t)e^{j\theta(t)} \quad (3.42)$$

where, $A(t)$ is the instantaneous amplitude (also known as the amplitude envelope), it is given by:

$$A(t) = |x_a(t)| = \sqrt{x(t)^2 + H\{x(t)\}^2}, \quad (3.43)$$

and $\theta(t)$ is the instantaneous phase; given by:

$$\theta(t) = \arctan\left(\frac{H\{x(t)\}}{x(t)}\right) \quad (3.44)$$

Accordingly, the instantaneous frequency $f(t)$ can be obtained by taking the derivative of the instantaneous phase $\theta(t)$:

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \quad (3.45)$$

The instantaneous amplitude of a signal shows how the signal strength varies with time. The instantaneous phase provides the phase angle of the signal as a function of time. The instantaneous frequency reveals how the frequency content of a signal evolves over time, providing a dynamic view of the signal's spectral properties. By obtaining instantaneous amplitude, phase, and frequency information, the HT serves as an effective tool to identify distinct characteristics, such as common patterns and sudden changes in phase and frequency where relevant features can be extracted accordingly. For instance, HT can be used to obtain instantaneous amplitudes of amplitude-modulated signals such as fault vibration

signals. In these signals, vibrations of damaged bearings are manifested as modulations in the amplitude of the generated vibration signal. Hence, obtaining the signal envelope provides an efficient approach for extracting fault signature frequencies from the envelope's spectrum. Fig. 3.21 displays a composite signal, $c(t)$, that has a time duration of 1 second. The signal is composed of various components so that it exhibits four sudden changes as follows:

- Rapid oscillations of 50 Hz from 0.3 time instant to 0.5 time instant.
- Rapid oscillations of 100 Hz from 0.7 time instant to 0.9 time instant.
- Abrupt phase change of 180 degrees at 0.4 time instant.
- Abrupt phase change of 90 degrees at 0.6 time instant.

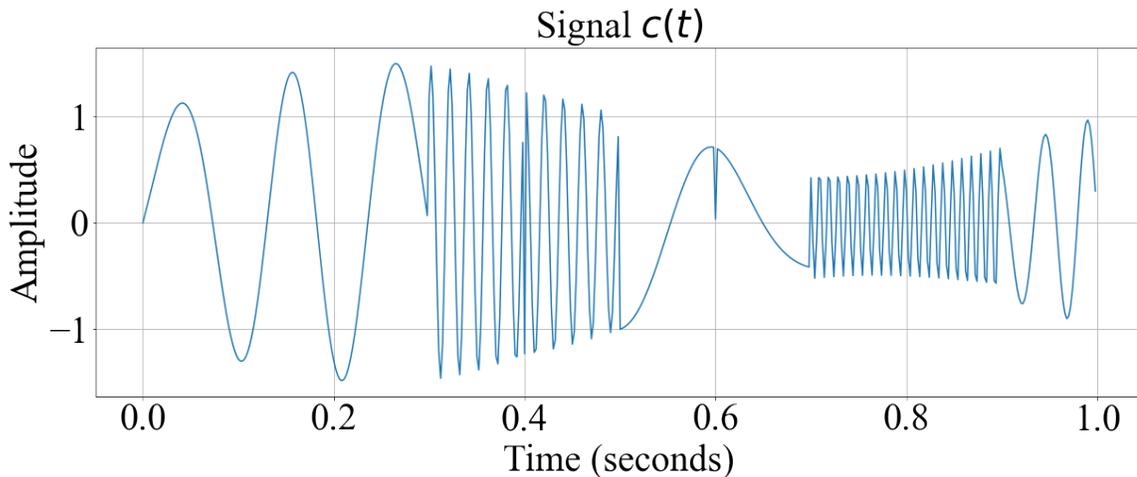


Figure 3.21: Composite signal $c(t)$.

The function `scipy.signal.hilbert` in the `scipy.signal` library is used to compute the HT of $c(t)$. The function returns the analytical signal $c_a(t)$ as an output. The instantaneous amplitude, phase, and frequency are calculated accordingly and displayed in Fig. 3.22. As shown in the plots, the obtained instantaneous information provides critical insights into the signal's behavior, highlighting common patterns and sudden changes. Specifically, the envelope shows the main pattern of the signal, reflecting its energy content over time. The

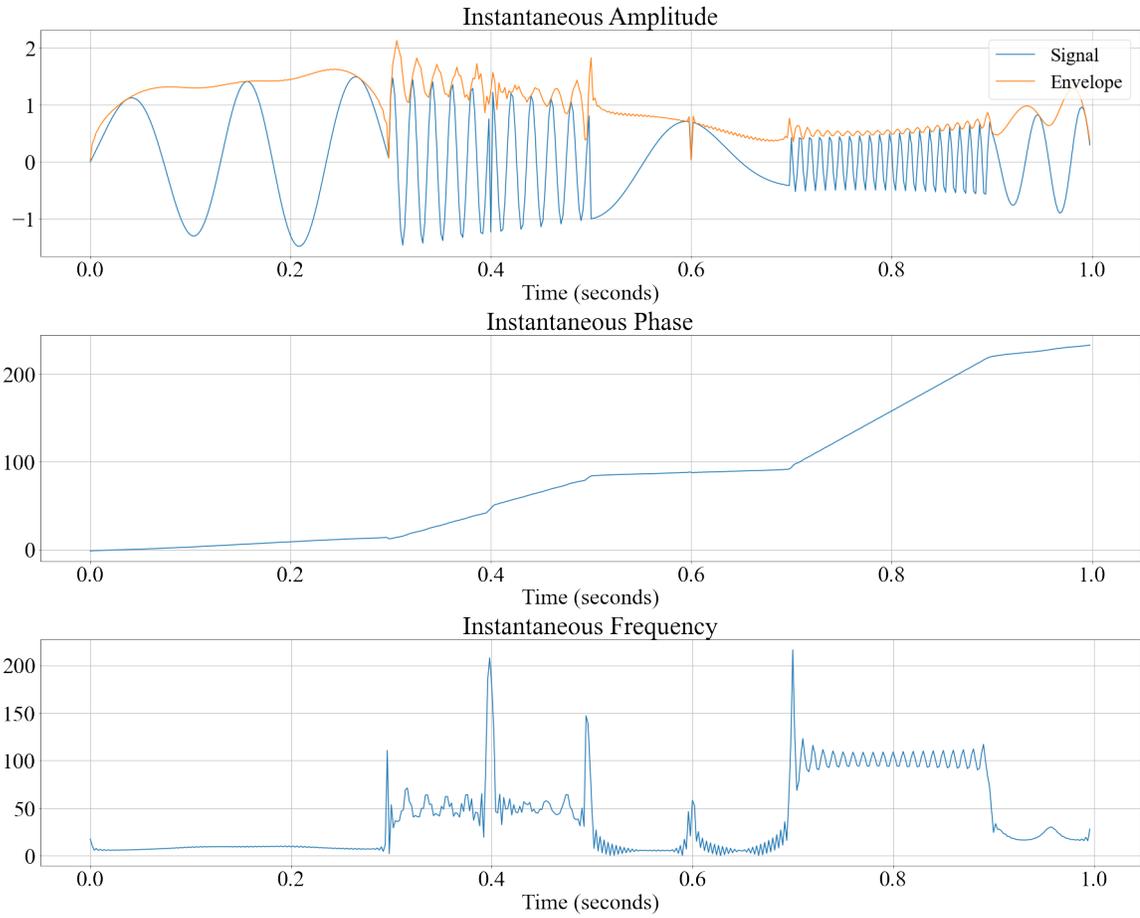


Figure 3.22: Hilbert Transform of the signal $c(t)$.

abrupt phase changes at the 0.4 and 0.6 time instants are manifested as small step changes at these time instants in the instantaneous phase. Moreover, the phase is directly impacted by temporal changes in the signal frequency, as shown in the plot. The instantaneous frequency provides vital information about frequency variations and sudden changes in the signal. Particularly, the instantaneous frequency plot shows the rapid oscillations of 50 Hz and 100 Hz between 0.3 – 0.5 and 0.7 – 0.9 time instants, respectively. Additionally, the sudden changes in frequency and phase are manifested as a spike at the corresponding time instant in the plot. In real-world applications, such spikes represent significant features in the signal that would be of particular interest.

3.3.6 Hilbert–Huang Transform

The Hilbert–Huang Transform (HHT) [58] is a well-known adaptive method for analyzing nonlinear and nonstationary signals. Unlike traditional transforms, HHT does not impose a fixed basis function in signal analysis. Instead, it utilizes an adaptive approach to analyze the signal, making it highly responsive to variations in signal. The HHT is a two-stage process that involves Empirical Mode Decomposition (EMD) of the signal followed by the Hilbert Spectral Analysis (HSA).

Empirical Mode Decomposition (EMD)

EMD is an adaptive decomposition method that decomposes the signal into a set of simpler functions known as Intrinsic Mode Functions (IMFs). EMD forms the basis of Adaptive Mode Decomposition (AMD) methods introduced over the last three decades. In contrast to wavelet-based decomposition, AMD uses adaptive approaches rather than priori basis functions to decompose the signal. Hence, the obtained IMFs are not influenced by a priori basis function [59]. This makes AMD particularly effective when dealing with nonlinear and nonstationary signals due to its ability to adapt to varying signal characteristics. How-

ever, this adaptability comes at the cost of higher computational complexity. The EMD algorithm, introduced in 1998 as part of the HHT, represents the core of all AMD methods. The process of EMD, commonly known as sifting, involves the following steps:

1. Identification of Extrema: Start by identifying all the local maxima and minima of the original signal.
2. Envelope Creation: Construct the upper and lower envelopes of the signal by interpolation between the local maxima and minima, respectively. These envelopes essentially outline the signal's oscillatory amplitude.
3. Mean Envelope Calculation: Compute the mean of the upper and lower envelopes.
4. Extraction of Detail: Subtract the mean envelope from the original signal. This step isolates a component of the signal that potentially qualifies as an IMF.
5. IMF Check: Verify if the extracted component meets the criteria for being an IMF:
 - The number of extrema “local maxima and minima” and the number of zero crossings in the component must either be equal or differ at most by one. This condition ensures that the IMF captures a well-defined oscillatory mode without bias toward upward or downward trends.
 - The mean value of the envelope defined by the local maxima and the envelope defined by the local minima must be zero at any point in the component. This condition guarantees that the IMFs have well-balanced oscillations around zero, reflecting true oscillatory modes rather than trends or biases in the signal.

If the component doesn't qualify for an IMF, return to step 2 and use this component as the new signal.

6. Completion of one IMF: Once an IMF is identified, subtract it from the original signal. This leaves a residue signal.

7. Repetition: Repeat steps 1-6 on the residue signal. This process is iterated, yielding a new IMF at each iteration, and the residual becomes the input for the next iteration.
8. Stopping Criterion: The process terminates when the residual signal becomes a monotonic function from which no more IMFs can be extracted or it becomes sufficiently small in amplitude based on a predefined threshold.

The output of the EMD process of the random vibration signal $v(t)$ is displayed in Fig. 3.23, including IMFs, residual, and their respective spectra. As shown, the EMD adaptively decomposed the signal into ten IMFs of distinct frequency spectra where the frequency contents of IMFs decrease with each successive IMF. While EMD is well-known to be effective in analyzing complex signals, the decomposition process is sensitive to noise since it works directly on the signal's extrema and minima. Additionally, the process requires a complete envelope to decompose the signal into IMFs accurately, which can be difficult to define at the boundaries due to the absence of neighboring data points. This can lead to distortions of IMFs near the ends of the signal, commonly known as the end effect [60], impacting the accuracy and reliability of EMD for signal analysis, particularly when the signal's behavior at the boundaries is critical. Another challenge in EMD is mode mixing [61]; ideally, each IMF should represent a unique frequency component. However, in signals with close spectral proximity or exhibit intermittency characteristics such as abrupt amplitude and frequency changes, the standard EMD algorithm may not be able to separate these features accurately. As a result, a single IMF may contain information from multiple modes, which can lead to a loss of significant physical interpretation of the IMFs. Further improvements to the standard EMD algorithm have been introduced to address these limitations. For example, the Ensemble Empirical Mode Decomposition (EEMD) algorithm [62] was developed by incorporating an ensemble approach and adding finite amplitude white noise to alleviate problems of mode mixing and end effects. Accordingly, the natural oscillatory modes are obtained by averaging the corresponding IMFs obtained

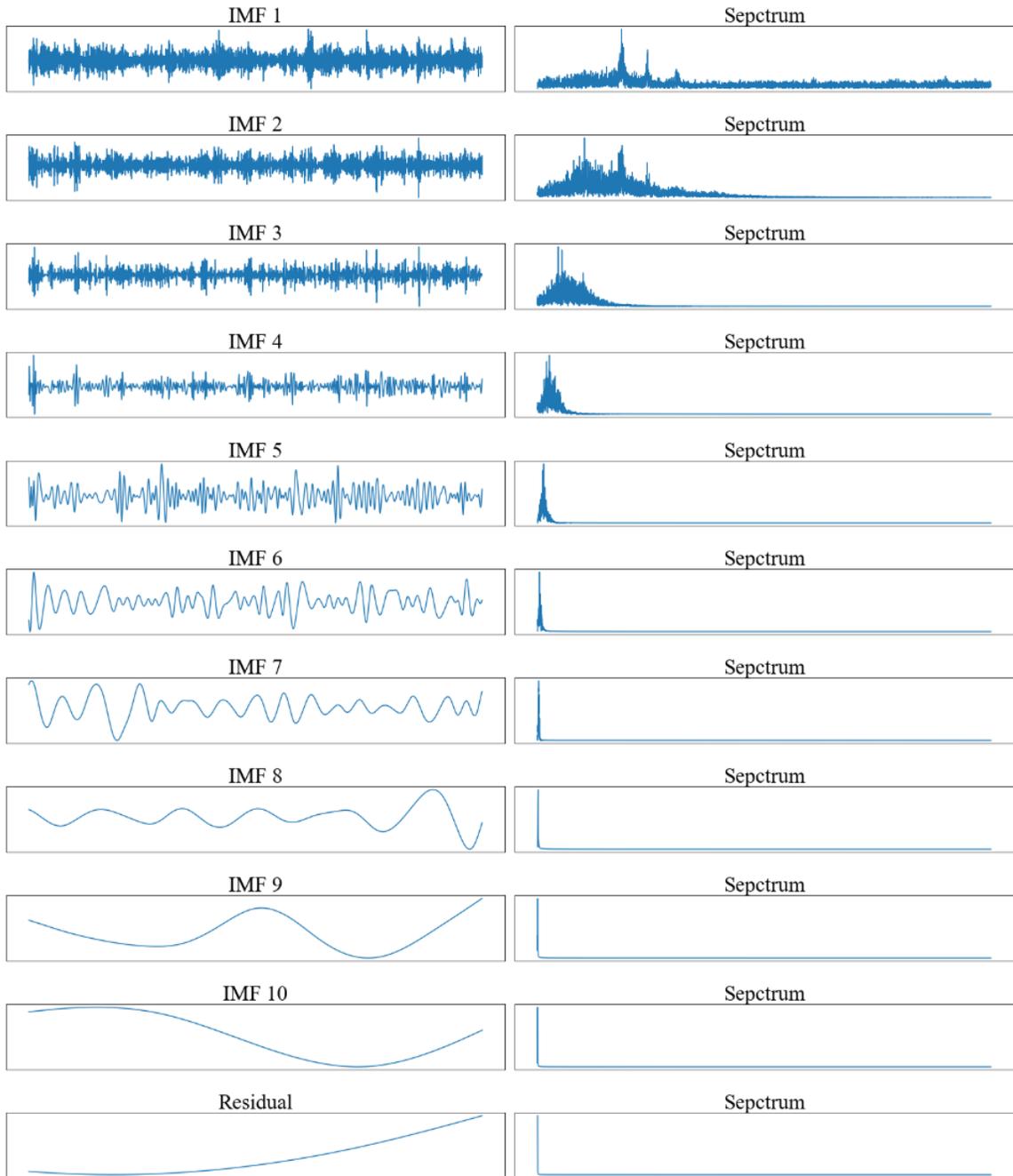


Figure 3.23: EMD of the random vibration signal $v(t)$: IMFs, residual, and their respective frequency spectra.

from an ensemble of the signal and the added noise. Other common AMD algorithms include Variational Mode Decomposition (VMD)[63], local mean decomposition [64], and Empirical Wavelet Transform (EWT) [65]. More details on AMD techniques and their applications can be found in [66][67][68][69].

Hilbert Spectral Analysis

The second step of HHT involves applying the HT to each IMF to obtain its analytic signal and, accordingly, its insinuations amplitude, phase, and frequency. An IMF, due to its defining conditions, ensures that its HT is well-behaved and meaningful for signal analysis [70]. Specifically, due to the equal number of extrema and zero crossings, an IMF exhibits a consistent oscillatory pattern, which is essential for a coherent HT. Further, the symmetry of the IMF's envelope about the zero line guarantees that the IMF doesn't have a bias towards positive or negative values, facilitating a more accurate HT.

Mathematically, the HSA can be expressed as follows: Given a signal $x(t)$, decomposed into n IMFs $c_i(t)$, where $i = 1, 2, \dots, n$, the analytic signal $Z_i(t)$ for each IMF is formed as follows:

$$Z_i(t) = c_i(t) + jH\{c_i(t)\} \quad (3.46)$$

where $H\{c_i(t)\}$ is the HT of $c_i(t)$. The analytical signal in the polar form is expressed as:

$$Z_i(t) = A_i(t)e^{j\theta_i(t)} \quad (3.47)$$

Accordingly, the instantaneous amplitude $A_i(t)$ is:

$$A_i(t) = |Z_i(t)| = \sqrt{c_i(t)^2 + H\{c_i(t)\}^2}, \quad (3.48)$$

and the instantaneous phase $\theta_i(t)$ is:

$$\theta_i(t) = \arctan\left(\frac{H\{c_i(t)\}}{c_i(t)}\right) \quad (3.49)$$

Accordingly, the instantaneous frequency $f_i(t)$ is obtained by taking the derivative of $\theta_i(t)$:

$$f_i(t) = \frac{1}{2\pi} \frac{d\theta_i(t)}{dt} \quad (3.50)$$

The instantaneous frequencies and amplitudes obtained from the HT can be used to construct a time-frequency distribution of the signal. This is accomplished by plotting the instantaneous amplitude or energy against the instantaneous frequencies of the IMFs for each point in time. The result is a distribution that shows how the signal's frequencies evolve over time, highlighting the signal's nonlinear and nonstationary characteristics. The time-frequency-energy distribution is commonly visualized as a 2D heatmap, where the x -axis represents time, the y -axis represents frequency, and the color intensity represents the amplitude or energy level of the signal at each time-frequency point. Considering the composite sinusoidal signal $s_3(t)$, Fig. 3.24 displays the resulting IMFs and residuals from the EMD process, as well as their spectral contents. Consequently, Fig. 3.25 shows the 2D heatmap visualization of the output from the HHT. These results are obtained using `textitemd`, `vmd`, and `hht` built-in MATLAB functions, which, in contrast to Python, provide a mature and user-friendly environment for EMD, VMD, and HHT computations and visualization. It is evident that both EMD and HSA accurately capture all signal components, including the transit oscillations of 400 Hz at 0.6 seconds. They also effectively display the temporal energy distribution of the signal at different frequencies. This highlights the advantage of HHT in adapting to nonlinear and nonstationary signals.

However, the performance of the HHT is tied to the choice of the AMD method, which is fundamental to the accurate decomposition of the signal into its IMFs. Furthermore, the employed stopping criterion for the sifting process [71] and the used interpolation method for envelope estimation [72] have direct impacts on the effectiveness of the HHT. A proper stopping criterion ensures that the decomposition process neither overfits nor underfits the signal, thereby preserving the signal's essential characteristics without introducing arti-

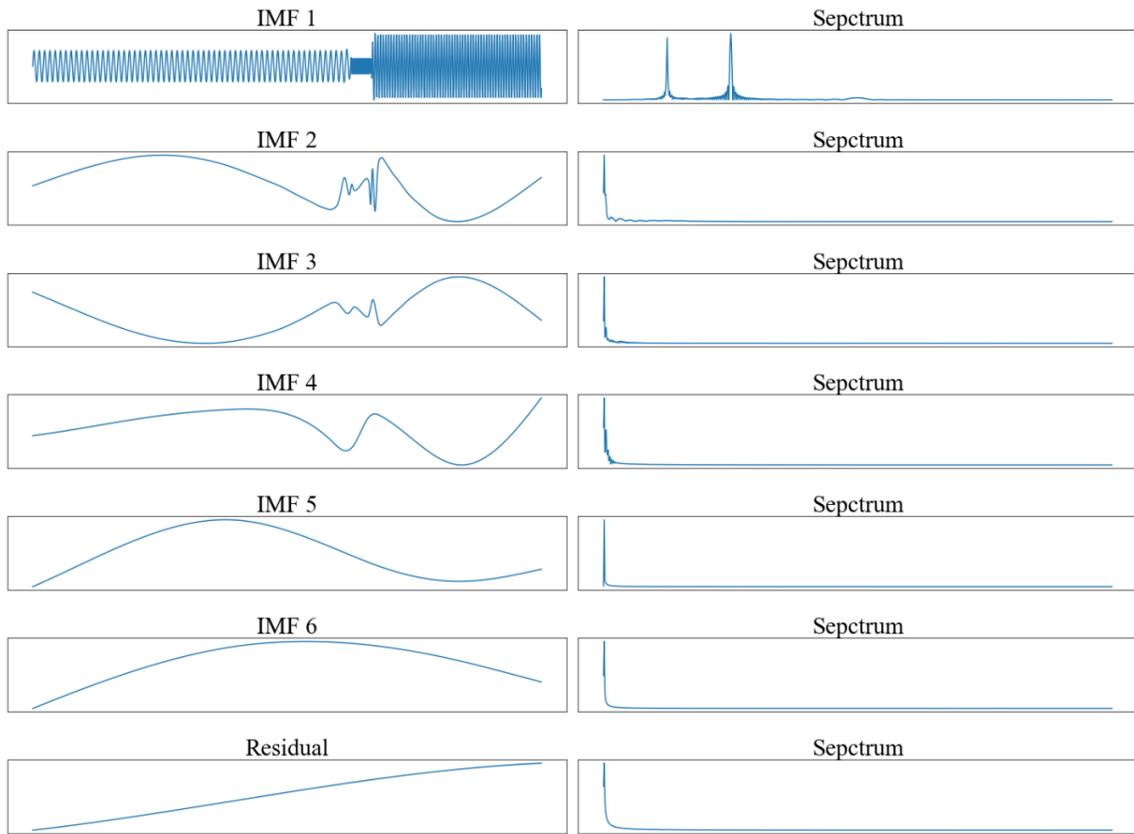


Figure 3.24: EMD of the composite sinusoidal signal $s_3(t)$: IMFs and residual alongside with their respective frequency spectra.

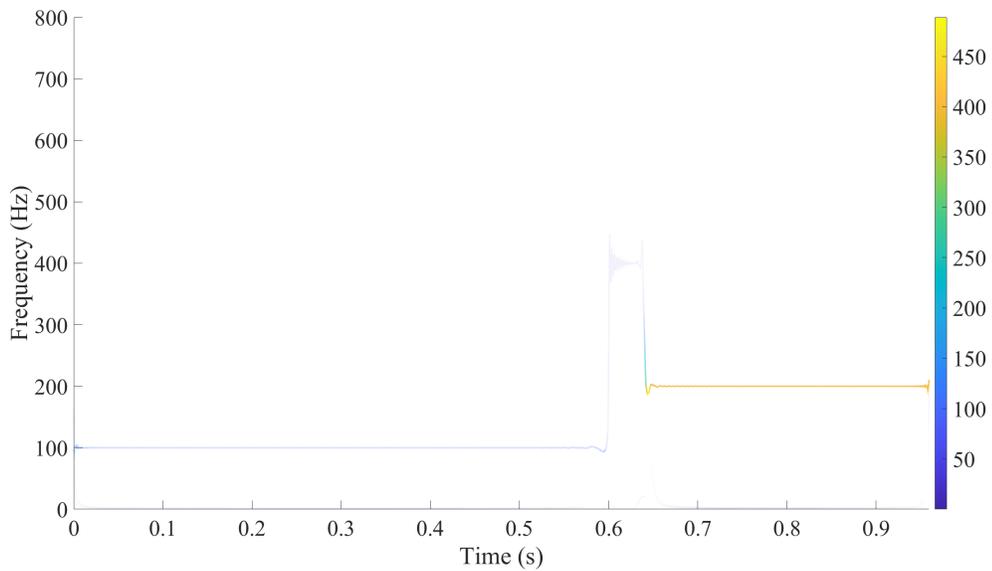


Figure 3.25: 2D heatmap visualization of the HTT of $s_3(t)$.

facts. The interpolation method for envelope estimation influences the HHT’s performance since it determines how well the upper and lower envelopes capture the true oscillatory modes of the signal.

3.3.7 Wigner–Ville Distribution

The Wigner–Ville distribution (WVD), as a time-frequency analysis tool, is characterized by its ability to provide highly resolved energy-time-frequency representations of signals.

The WVD of a continuous-time signal $x(t)$ is expressed as:

$$WVD(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau} d\tau \quad (3.51)$$

where, $x^*(t)$ is the complex conjugate of $x(t)$, τ and f are variables representing time-shift and frequency, respectively.

According to this expression, the kernel function of WVD is given by:

$$x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right), \quad (3.52)$$

which is essentially the Instantaneous Autocorrelation Function (IACF) of $x(t)$, commonly known as Wigner auto-correlation function [73]. The use of the complex conjugate of the signal in the autocorrelation accounts for the magnitude and phase parts of the signal. Note that the traditional autocorrelation function of (3.20) integrates the autocorrelation over time t , providing a global measure of the correlation. In contrast, the IACF provides a local measure of the correlation, which is better suited to reflect local and time-varying features of the signal. Thus, facilitating a more accurate representation of the signal’s characteristics [74]. The Fourier kernel $e^{-j2\pi f\tau}$ in the WVD expression transforms the IACF into the frequency domain with respect to the time-shift variable τ , thereby providing a joint time-frequency representation of the signal. For a discrete signal $x[n]$ with N samples, the

discrete WVD is given by:

$$WVD[n, k] = \sum_{m=-N}^N x[n + \frac{m}{2}]x^*[n - \frac{m}{2}]e^{-j2\pi\frac{km}{N}} \quad (3.53)$$

Here, $x^*[n]$ is the complex conjugate of $x[n]$, and k represents the discrete frequency index.

The term,

$$x[n + \frac{m}{2}]x^*[n - \frac{m}{2}], \quad (3.54)$$

captures the auto-correlation of the signal at different lags.

The output of WVD, $WVD(t, f)$, is essentially a 2D function of time and frequency that is commonly visualized as a 2D heatmap where the value of $WVD(t, f)$ at any point (t, f) reflects signal's energy at that particular time and frequency. The high energy concentration of WVD results from its inherent quadratic (nonlinear) structure since the energy itself is a quadratic representation of the signal [75]. In contrast to FT, STFT, and CWT, which are linear transforms³, the WVD is nonlinear since it involves a product of the signal with a time-shifted version of itself (correlation). This energetic and correlative nature of the WVD [75] makes it a unique tool for energy-time-frequency analysis compared to other transforms, however, the quadratic nature of the WVD presents a major challenge when analyzing multi-component signals due to the high level of cross-term interference [76] that appears in the time-frequency representations. Cross-terms are caused by undesired cross-correlation between various signal components, leading to a false indication of signal components between the desired auto-correlation terms in the WVD representation of the signal. To overcome this problem, window-based approaches [77, 78] and kernel-based approaches [79, 80], such as the Pseudo-Wigner-Ville Distribution (PWVD), are commonly

³In signal theory, a linear transform satisfies the superposition or linearity condition, which states that if the input signal $x(t)$ is a linear combination of some signal components, then, the transform of $x(t)$ is also a linear combination of the transforms of each signal component.

utilized to suppress cross-terms. PWVD is a windowed version of the WVD that applies a smoothing kernel to WVD in either time-domain or frequency domain to reduce the cross-terms at the cost of a certain loss in resolution. The PWVD of a signal $x(t)$ can be expressed as:

$$PWVD(t, f) = \int_{-\infty}^{\infty} h(\tau)x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau}d\tau \quad (3.55)$$

where $h(\tau)$ is the smoothing kernel in the time domain. In the frequency domain, the smoothing operation becomes the convolution operation between the smoothing kernel and the WVD. The smoothing operation essentially averages the signal's energy over the time and/or frequency domains, reducing the presence of cross-terms arising from the quadratic nature of the WVD. However, this comes at the cost of blurring the signal's time-frequency representation. The limitation of WVD resulting from undesired cross-correlation and the advantage of PWVD are demonstrated by comparing the WVD and PWVD of the composite signal, $c(t)$, which, as illustrated earlier, comprises multiple time-varying components. The MATLAB function `wvd` is used to compute WVD and PWVD and visualize the outputs as 2D heatmaps. This built-in function provides straightforward computations and visualizations of various WVD types. The WVD and PWVD results are displayed in Fig.3.26 and Fig.3.27, respectively. The comparison reveals that the WVD representation is impaired by false indications of non-existent signal components due to the cross-correlation between the actual signal components. In contrast, the PWVD, through its smoothing operation, eliminated the interference of cross-terms and accurately resolved all the time-varying frequency components. However, it can be seen that the improvement achieved with the PWVD led to a reduction in frequency resolution, manifested as wider lines in the PWVD's heatmap compared to the WVD's heatmap, thereby affecting the precise estimation of the corresponding frequency values.

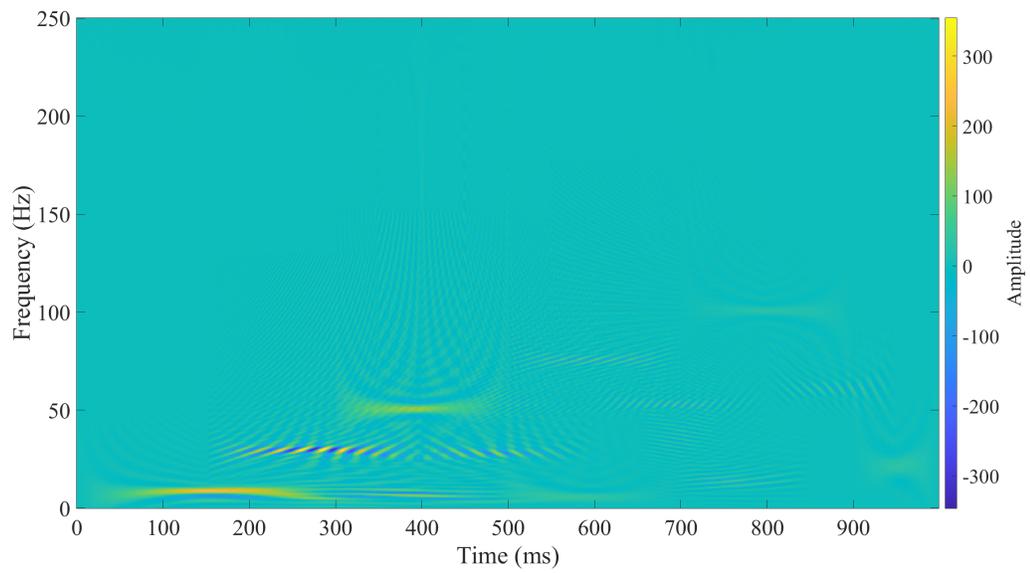


Figure 3.26: WVD of the composite sinusoidal signal $c(t)$.

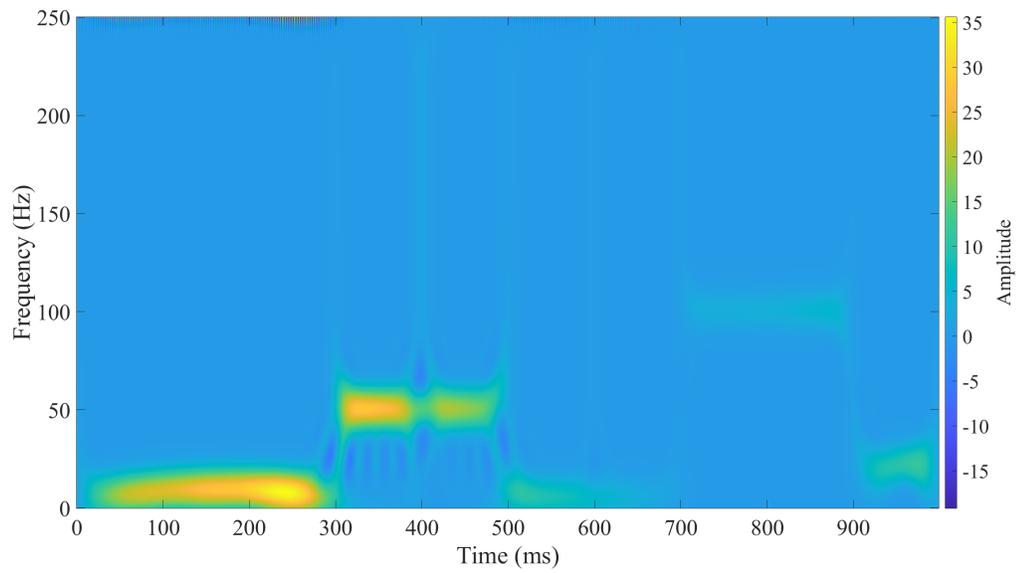


Figure 3.27: PWVD of the composite sinusoidal signal $c(t)$.

3.3.8 Conclusive Comparison

Aimed at providing interested readers a thorough understanding of signal processing fundamentals, this tutorial offered a comprehensive introduction to signal characteristics with an in-depth discussion of prevalent signal transformation and analysis tools. The discussion focused on elucidating the main concepts, mathematical foundations, essential characteristics, advantages, and limitations. Additionally, the tutorial addressed implementation considerations by underscoring the use of programming libraries and built-in functions to facilitate efficient implementations of these tools. Furthermore, the codes employed for generating the illustrative plots throughout this section have been made publicly accessible as previously mentioned.

Throughout the section, it became evident that selecting the proper tool depends entirely on signal nature, application-specific requirements, and available computational resources. For stationary signals, where the primary interest lies in frequency analysis, The FT efficiently obtains the frequency spectrum, thereby enabling the identification of predominant frequency components in the signal. The PSD is particularly advantageous in scenarios involving comparative spectral analysis of multiple signals varying in length and/or bandwidth. It provides a normalized spectral density measure that facilitates effective comparison. The HT is particularly effective in applications where the signal envelope is of primary importance. Moreover, instantaneous phase and frequency information can be efficiently obtained from the analytical signal of the HT.

For signals with time-varying spectral characteristics, the STFT provides a computationally efficient tool for the time-frequency representation of the signal. However, its fixed window size results in a uniform resolution across all frequencies, reducing its effectiveness for applications requiring variable-resolution analysis. In such contexts, Wavelet analysis presents a viable solution due to its ability to perform multi-resolution analysis through scaled and shifted versions of a wavelet base function.

For the analysis of nonstationary and nonlinear signals characterized by rapidly changing components, the HHT provides an adaptive and more robust approach than CWT but with more computational requirements. The WVD offers a powerful method for providing a joint energy-time-frequency representation of the signal, although it is more computationally intensive than the other tools. It is particularly useful in high-resolution analysis scenarios or where energy concentration is a focal requirement.

Signal decomposition techniques allow the analysis of the elementary components of a signal. This makes it possible to extract unique manageable-sized features closely related to the signal's inherent structure. Regarding wavelet decomposition, its performance is highly dependent on the base wavelet function and the level of decomposition. In contrast, AMD methods use an adaptive mechanism rather than an a priori basis function to decompose the signal, thereby generating elementary modes that are not influenced by an a priori basis function. Moreover, AMD methods dynamically adapt to varying signal characteristics, making them effective for analyzing nonlinear and nonstationary signals. However, AMD methods are more computationally intensive than wavelet decomposition. Additionally, AMD methods are sensitive to noise and their efficiency highly depends on the employed stopping criterion and the used interpolation method for envelope estimation. Table 3.2 summarizes the comparison, highlighting the main aspects of each tool regarding signal nature, application requirements, and computational complexity.

3.4 A Typical Signal-Based ML Pipeline

In signal-based Machine Learning (ML) applications, the pipeline of signal processing is typically divided into three core stages: preprocessing, processing, and application, as illustrated in Fig. 3.28. This section addresses each stage, highlighting its purpose and main aspects.

Table 3.2: Comparison between common signal analysis and transformation tools

Tool	Signal Nature	Application	Computational Complexity
Fourier Transform (FT)	Stationary	Frequency analysis	Low
Power Spectral Density (PSD)	Stationary	Spectral analysis	low
Hilbert Transform (HT)	Stationary	Envelope and instantaneous phase/frequency analysis	low
Short-Time-Frequency Transform (STFT)	Nonstationary	Fixed-resolution time-frequency analysis	moderate
Wavelet Transform (WT)	nonstationary	Multi-resolution time-frequency analysis	moderate
Hilbert-Huang Transform (HHT)	Nonstationary and nonlinear	Adaptive time-frequency analysis	High
Wigner-Ville Distribution (WVD)	Nonstationary	High-resolution energy-time-frequency analysis	High
Wavelet Decomposition (WD)	Nonstationary	Signal decomposition	moderate
Adaptive Mode Decomposition (AMD)	nonstationary and nonlinear	Signal decomposition	High

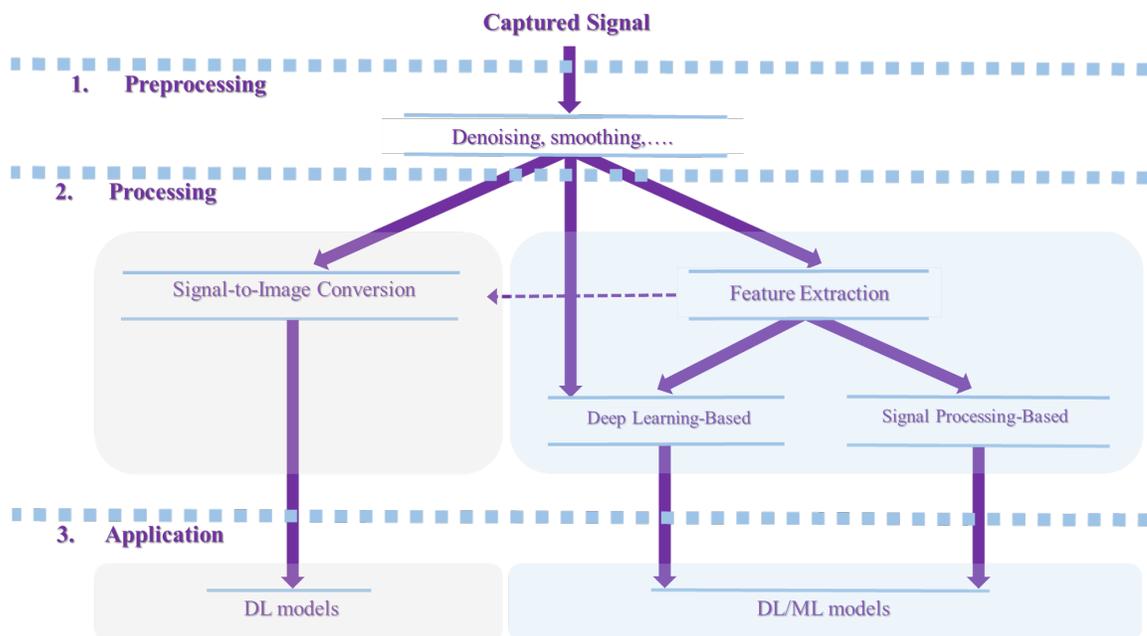


Figure 3.28: Overview of a typical signal-based ML framework.

3.4.1 Preprocessing

The first stage encompasses the preprocessing of acquired signals. Here, it is essential to distinguish between data preprocessing techniques[81], commonly used in ML models, and signal preprocessing, which involves the tasks of signal smoothing, signal denoising, and signal segmentation. Signal preprocessing is performed on the raw signals, whereas data preprocessing techniques, such as normalization and scaling, are applied to the extracted features.

Signal Smoothing

The acquired signal may exhibit rapid amplitude changes between successive samples. Such random changes in amplitude can negatively impact application-level performance. Furthermore, during measurements, the signal is susceptible to outlier samples caused by external factors such as instrument malfunction [82]. Signal smoothing attempts to remove such impairments in the signal by adjusting the amplitudes of individual samples with respect to the amplitudes of adjacent samples. Smoothing acts as an approximation filter that works on N -successive samples at a time and outputs N smoothed samples. Common smoothing methods involve the well-known Moving Average Filter (MAF) filter and the Savitzky–Golay Filter (SGF) [83]. MAF offers a simple method to smooth the signal by averaging neighboring data points while preserving the signal’s shape. It operates by averaging a set of signal data points (window) and sliding this window across the signal to smooth out short-term fluctuations. The SGF uses a different approach to smoothing the signal by performing a local polynomial regression on the signal data points present in the moving window to provide more nuanced smoothing that preserves the original shape and characteristics of the signal. The design of an SGF involves properly selecting the window size and the order of the polynomial [84–87]. Signal smoothing helps refine the signal’s waveform shape and captures its main patterns, which is particularly useful in

many signals-based applications. However, as a lossy operation, smoothing could cause a notable distortion in applications where peak-related features are of main interest. More information on signal smoothing can be found in [88–91].

Signal Denoising

Denoising refers to the process of removing noise from the acquired signal or, in other words, reconstructing the desired signal from its acquired noisy representation. This can be expressed mathematically as follows:

$$w(t) = s(t) + n(t) \quad (3.56)$$

where $w(t)$ denotes the noisy signal, $s(t)$ is the desired signal, and $n(t)$ is the noise component. Noise can be defined as any unwanted signal present in the measurement process other than the desired signal [92]. Signal denoising and signal smoothing are often treated synonymously in the literature [93], as both aim to refine the signal's waveform while preserving meaningful patterns within the signal. However, they have different impacts on the noise and frequency contents of the signal. Smoothing is typically applied in the time domain on a sample-by-sample basis while denoising attempts to remove the whole noise component, $n(t)$.

Furthermore, in smoothing, the rapid fluctuations can be viewed as highly oscillating components of small amplitudes compared to the whole signal over a relatively long duration. From a frequency point of view, such oscillations represent high-frequency components of low amplitudes in the signal's spectrum. Smoothing, therefore, acts as a low-pass filter that removes these high-frequency components from the signal. On the other hand, signal denoising typically involves the use of various signal processing methods to develop more advanced noise filtering techniques.

Signal Segmentation

Depending on the length of the acquired signal, expressed in terms of the number of samples, it might be necessary to divide the signal into smaller segments for the following practical considerations:

- Online deployment: Online applications require working on a pre-defined length of the input signal.
- Real-time and delay-sensitive applications: Given the sampling frequency, f_s in samples per second, the time duration, t_s , of the segment, expressed in *seconds* equals:

$$t_s = \frac{N_o}{f_s} \quad (3.57)$$

Hence, in real-time and delay-sensitive applications, the number of samples, N_o , in the input segment is of paramount importance since time delay increases as a function of N_o [56].

- Computation burden: The computational complexity of signal-based ML systems increases as a function of N_o [56].
- Effectiveness of extracted features: The number of samples, N_o , in the input segment, significantly shapes the discriminative characteristics of the extracted features. Shorter segments may not provide an adequate representation of the signal's characteristics, reducing the reliability of extracted features. Conversely, very long segments could incorporate a high redundancy level, weakening the discriminative capability of the extracted features.

Constant segmentation is commonly used in various applications where the segment length, N_o , is kept constant. Consequently, signals are divided into smaller segments of equal lengths. Determining the value of N_o is application-specific and typically relies on achiev-

ing good trade-offs among the aforementioned practical considerations. Constant segmentation is commonly implemented using a sliding window of size N_o . Another aspect of constant segmentation is the percentage of overlap between successive segments. Introducing overlap between segments increases the number of segments extracted from the signal. The more overlap, the more segments are extracted. However, increasing the overlap increases correlation or dependence among adjunct segments. In general, the extent to which overlap can be used and its percentage depends on the nature of the application and the signal processing techniques used [94–96].

In some applications, the acquired signal exhibits random intervals of activity and inactivity, as observed in various sound and biomedical signals. In such scenarios, constant segmentation becomes unreliable, as it does not count for these inherent intervals. In such situations, adaptive segmentation [97–103] is commonly utilized to dynamically adjust the segment length, N_o . Adaptive segmentation exploits the underlying characteristics of the signal and adapts N_o accordingly. Unlike fixed length segments, adaptive length segments are more discriminative because the signal’s characteristics and inherent patterns change from segment to segment, allowing more discriminative features to be extracted. However, adaptive segmentation involves more computational burden compared to constant segmentation.

3.4.2 Processing

The processing stage involves processing the resultant segments to extract appropriate features that serve as inputs to ML models. The methods and algorithms available in the literature can be broadly categorized into two main approaches: feature extraction and signal-to-image conversion. In feature extraction, signal processing or Deep Learning (DL) can be used to extract relevant features from the input segment. The next section comprehensively covers the topic of feature extraction. Signal-to-image conversion approaches [104–113] transform 1-dimensional (1D) signal segments into 2-dimensional (2D) spatial representa-

tions, allowing to leverage DL models in signal-based applications. Gramian Angular Field (GAF) and Markov Transition Field (MTF) [104] are two well-known techniques that are commonly used to transform time series into images. Advanced techniques involve the use of signal processing to generate 2D time-frequency representations of the signal segments. The essence of these techniques is to represent time-frequency-energy characteristics of the signal and features as width-height-color intensity, mimicking image features. However, to serve training and testing purposes of DL models, a large number of transformed images should be available. Nevertheless, signal-to-image conversion is especially feasible in integrated systems such as Unmanned Aerial Vehicles (UAVs) and Autonomous Vehicles (AVs) [114]. In these systems, computer vision applications, such as object detection and navigation, and signal-based applications are deployed within the same environment. Therefore, it is deemed computationally efficient to convert signals into images and use transfer learning approaches that utilize the trained deep-learning infrastructure for inference tasks of signal-based applications. Moreover, signal-to-image conversion and transfer learning offer a practical solution when the size and type of available signal data are insufficient to fulfill training and testing requirements.

3.4.3 Application

The application is the final stage of a typical signal-based ML pipeline. As a task-specific stage, it uses appropriate ML models [115–117] to address tasks, including classification, clustering, and anomaly detection. Feature preprocessing, hyperparameter tuning, and relevant metrics [118, 119] are often used to improve and evaluate model performance. Table 3.3 serves as a high-level overview of a typical signal-based ML framework, highlighting each stage's purpose, inputs, functions, and outputs.

Table 3.3: Inputs, functions, and outputs of each stage of typical signal-based ML framework

Stage	Inputs	Functions	Outputs
<i>Preprocessing</i>	Acquired signal+impairments (noise / outliers)	Smoothing / Denoising / Segmentation	Segmented signal
<i>Processing</i>	Segmented signal	Feature extraction / Signal-to-image conversion	Features/ Images
<i>Application</i>	Features/ Images	Classification / Clustering / Anomaly detection	Results

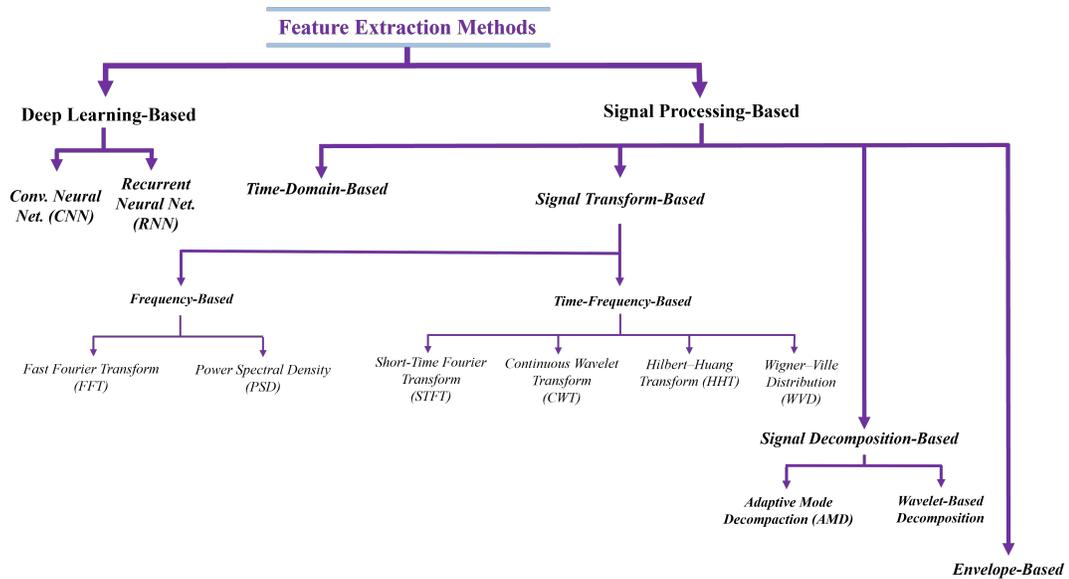


Figure 3.29: Taxonomy of feature extraction techniques.

3.5 Feature Extraction Methods

This section provides a comprehensive review of existing feature extraction methods. The review utilizes a new hierarchical taxonomy, depicted in Fig. 3.29, to categorize and structure various methods that exist in the literature. The taxonomy starts by grouping the existing methods under two main categories: DL-based and signal processing-based. This categorization distinguishes between two mechanisms that exist for feature extraction: feature *learning* and feature *engineering*. The first mechanism relies on DL models— such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and auto-encoders— to *learn* high-level feature representations from signal data [120–129]. The second mechanism, on the other hand, employs signal processing to extract distinctive features from the signal, deliberately *engineered* to emphasize specific attributes and reveal certain characteristics. This deliberate engineering allows meaningful connections to be made between the resulting features and various conditions or classes that exist within the signals. In contrast, DL features are not interpretable since DL models are black-box models, making establishing such meaningful connections impossible. In addition, features extracted by DL models are model-dependent, whereas features based on signal processing are solely signal-dependent. This implies that variations in the model’s architecture, parameters, training, or tuning lead to different feature representations and, consequently, inconsistent performance results. Thus, it can be deduced that in DL-based approaches, performance is model-dependent, while in signal processing-based approaches, performance is mainly feature-dependent. Moreover, DL features would exhibit higher dimensionality and redundancy than signal processing-based features. This requires the use of appropriate dimensionality reduction and feature selection techniques [123, 130, 131], thereby increasing the online processing time and computational burden. Table 3.4 summarizes the aforementioned comparison between signal processing-based and DL-based approaches. A considerable amount of work in the literature uses hybrid approaches where signal processing-

Table 3.4: A comparison between deep learning-based and signal processing-based feature extraction

<i>Aspect</i>	Deep Learning Approaches	Signal Processing Approaches
<i>Feature Extraction Mechanism</i>	Feature learning	Feature engineering
<i>Interpretability of Extracted Features</i>	Not interpretable	Interpretable
<i>Size of Required Dataset</i>	Large	Small-to-medium
<i>Data Dimensionality and Redundancy in Extracted Features</i>	High, depends on DL model	Low-to-medium
<i>Computational Cost</i>	High	Moderate

based features are fed into DL models to further extract highly distinctive representations from these features [41, 132–144]. Compared to DL-based and signal processing-based features, hybrid features have higher discriminative power, which generally leads to better performance. However, employing hybrid approaches involves increased online processing time and higher computational complexity, making practical deployment subject to trade-offs between system complexity and performance requirements. In general, hybrid approaches prove feasible in applications that involve complex signals, such as radio modulation recognition [41, 140–144]. In these applications, signals are randomly modulated, corrupted with high noise levels, and subject to various types of interference.

Focusing on signal processing-based feature extraction, the taxonomy of Fig. 3.29, from a computational point-of-view, divides the existing methods into four main categories: time domain-based methods, transform-based methods, decomposition-based methods, and envelope-based methods. Time domain-based methods extract features directly from time domain waveforms of the signal without further processing. On the other hand, transform-based, decomposition-based, and envelope-based methods use additional steps to process the signal and, consequently, additional computations to extract features. The remainder of this section reviews common methods that exist within each category.

3.5.1 Time Domain-Based Methods

In time-domain-based methods, features are calculated from the signal’s amplitude, representing specific aspects of the signal’s dynamics over its time period. This allows for

the quantification of amplitude variations from one period to another. Conventional methods [145–160] involve extricating shape characteristics— such as maximum, peak-to-peak value, and crest factor— and statistical properties of the input signal. Statistical features describe characteristics of the probability distribution of the input signal; common statistical features involve mean, Root-Mean-Square (RMS) value, variance, and High-Order Statistics (HOS), such as skewness and kurtosis. Compared to first and second-order statistics (mean and variance), HOS features have more distinguishable characteristics as they can identify more complex aspects of the input signal, such as the shape of its distribution. Furthermore, HOS are less sensitive to noise. Besides shape-based and statistical methods, time-domain entropy methods [53, 161–163] are commonly utilized for feature extraction. Entropy methods rely on the observation that signals belonging to different classes or conditions typically exhibit different levels of irregularity or dynamic changes depending on the underlying state. Hence, different entropy measures can be utilized to quantify the irregularity of the input signal. In contrast to statistical features, entropy-based features are more sensitive to changes and complexities that arise within the signal. Thus, they can capture variations and irregularities in the signal that may be difficult to detect using traditional statistical measures. Additionally, entropy-based features are resilient to noise and outliers. In general, time-domain feature extraction approaches work directly on input signals, making them conceptually straightforward and relatively easy to implement. Further, they are computationally efficient since no further processing steps are required, making them advantageous in applications where real-time processing is crucial, especially with limited computational resources. However, time-domain analysis is highly susceptible to noise since noise in the signal would alter its amplitude and mask the dynamic characteristics of the signal. Furthermore, to achieve reliable performance, input segments of relatively long duration are usually required to capture the changes and complexities that evolve within the signal over time.

3.5.2 Transform-Based Methods

Transform-based approaches use advanced techniques to convert signals from the time domain to the frequency domain or time-frequency domain. This allows for the identification of frequency and frequency-temporal information, facilitating robust signal analysis and feature extraction.

Frequency-Based Methods

Frequency-based approaches represent signals in terms of their frequency content, revealing details that are not apparent in the time-domain waveform. Further, in contrast to time-domain analysis, frequency-based analysis allows the isolation and removal of noise or unwanted components by applying appropriate frequency filtering mechanisms to improve the signal-to-noise ratio and improve the reliability of the extracted features. Frequency-based approaches involve the application of the Fourier Transform (FT) to the input signal to compute its frequency content, which can be broadly categorized into two main categories: spectrum-based [24, 146, 164–177] and spectral density-based [174, 178–182]. Spectrum-based approaches use the frequency spectrum of the signal for feature extraction. Specifically, some approaches directly use resultant Fourier coefficients of the input segment as input features [164–168] since these coefficients describe the distribution of the signal's energy over the range of frequencies contained in the signal. Here, dimensionality reduction and feature selection techniques are often used to reduce the size of coefficients and select the most discriminative coefficients.

Other approaches involve the extraction of various frequency-domain features from the frequency spectrum, such as statistical properties, energy, entropy, and correlation coefficients [24, 146, 169–177]. Spectral density-based approaches, on the other hand, utilize the Power Spectral Density (PSD) of the input signal to extract the features [174, 178–182] since it provides a normalized measure of the power per unit frequency compared to the frequency

spectrum. Entropy is frequently utilized in PSD analysis to calculate Spectral Entropy (SE) [179, 180]. SE is a very useful spectral feature that measures the irregularity or randomness of power distribution across signal frequencies and, hence, can be used to quantify the spectral complexity of the input signal. Other methods utilize Mel-Frequency Cepstral Coefficients (MFCCs) as input features [183–192] because they effectively capture the shape of the power spectrum. MFCCs are originally developed for audio signals [193] and are gaining attraction in other various applications as well. A thorough explanation of MFCCs, including computational steps and usage scenarios, can be found in [193]. Higher-order spectra [194–197] are commonly used in frequency-domain analysis to conduct a more comprehensive spectral analysis compared to second-order spectral analysis. Common higher-order spectra include bispectrum and trispectrum [195, 198–207], which correspond to the FT of the signal’s third-order cumulant and the fourth-order cumulant, respectively.

A key advantage of higher-order spectra over power spectrum is their ability to retain phase information [208]. Specifically, the power spectrum, representing the second-order spectrum or the FT of the autocorrelation function, captures the magnitude of frequencies present in a signal but discards phase information. This limitation makes it challenging to identify phase-related properties of the signal, such as frequency coupling or interactions between different frequency components. In contrast, higher-order spectra provide insight into the phase relationships among frequency components by examining moments or cumulants of the signal beyond the second order, enabling the detection of nonlinearities, phase coupling, and other characteristics that are invisible in the power spectrum. However, higher order spectra are more complex to calculate and interpret than traditional power spectral analysis because they require more data for reliable estimation and involve more sophisticated mathematical and computational methods. Furthermore, due to their multi-dimensional nature, visualisation and interpretation of higher order spectra can be challenging. Because of this complexity, specific cross sections or slices of the bispectrum

or trispectrum are often analysed to extract meaningful information. This approach simplifies the analysis by reducing the dimensionality of the data, making it easier to visualise and interpret.

While frequency analysis accurately distinguishes between different frequency components in the signal, the analysis spans the signal's entire duration, lacking the ability to provide temporal information about the timing of occurrence of these frequencies within the signal. Additionally, the FT is not ideal for analyzing signals that contain highly time-localized components, such as short bursts with high energy concentrations, because the components produce a wide range of frequencies in the frequency spectrum due to the inherent uncertainty principle associated with Fourier analysis [54].

Time-Frequency-Based Methods

In contrast to frequency-domain analysis, which lacks time resolution, time-frequency transforms produce representations that map the signal's energy across both time and frequency, allowing for a localized analysis in the time-frequency domain. This approach allows the identification of transient oscillatory components within the signal, providing a more complete understanding of its dynamics. Time-frequency transforms that are commonly used in the literature for feature extraction include Short-Time Fourier Transform (STFT), [209–221] Hilbert-Huang transform (HHT) [222–232], Wigner-Ville Distribution (WVD) [233–241], and Continuous Wavelet Transform (CWT) [213, 239, 242–252]. STFT is efficient in analyzing nonstationary signals whose spectral properties vary over time. However, it uses a fixed segment length in analyzing the signal, imposing a trade-off between time-domain and frequency-domain resolutions. In CWT, on the other hand, time-frequency analysis is based on a family of wavelets generated from a single mother wavelet through scaling and translation, offering a multi-resolution analysis of the signal at different frequencies (scales) and time intervals (translations). Moreover, in contrast to STFT, which uses complex exponentials as basis functions— that extend infinitely in the time

domain—, wavelet base functions are localized in time and frequency, making them more suitable for signals with transient or highly localized components. The HHT uses an adaptive approach for time-frequency analysis. Unlike STF and CWT, HHT does not impose a fixed basis function on the signal analysis but instead uses an adaptive approach to analyze the signal. This adaptivity makes it highly responsive to variations in the signal, facilitating very effective analysis of nonlinear and nonstationary signals. However, as an adaptive analysis method, its performance depends heavily on two factors: the stopping criterion for the sifting process [71] and the choice of the interpolation method for envelope estimation [72].

A proper stopping criterion ensures that the decomposition process neither overfits nor underfits the signal, thereby preserving the signal's essential characteristics without introducing artifacts. On the other hand, the interpolation method for envelope estimation determines how well the upper and lower envelopes capture the actual oscillatory modes of the signal, heavily influencing the HHT's performance. The WVD offers a time-frequency representation with high energy concentration, facilitating detailed analysis of a signal's energy distribution over both time and frequency domains. The high concentration of energy arises from the quadratic nonlinear structure of WVD compared to STFt and CWT. This energetic and correlative nature of the WVD [75] makes it very useful in analyzing single-component signals. However, the quadratic nature of the WVD represents a major challenge when analyzing multi-component signals due to high cross-term interference [76] caused by undesired cross-correlation between various signal components, leading to a false indication of non-existent signal components in the resultant time-frequency representations of the signal. Window-based WVD [77, 78] and kernel-based approaches [79, 80] are commonly used to reduce the cross-terms at the cost of some loss in resolution.

From a feature extraction perspective, time-frequency-based methods can be grouped under three main approaches in which the generated energy-time-frequency mappings are treated

differently:

1. Transformation coefficients as input features: In these approaches, the generated transform energy coefficients, representing signals' energy distribution across time-frequency instants, are used directly as input features for the ML model [209, 210, 222, 227, 230, 231, 242, 251]. Feature selection techniques are commonly used here to select the most distinctive representations, which helps reduce redundancy and relaxes computational requirements for training and inference.
2. Energy-time-frequency 2D heatmaps as imagery data: The generated energy-time-frequency representations are used to visualize 2D heatmaps that map the signal's energy across time-frequency instants. Accordingly, computer vision and DL techniques are leveraged where these heatmaps serve as input images [211–219, 223, 228, 233, 234, 238, 239, 241, 243, 244, 249, 250, 252].
3. Transformation coefficients as signal representations: In these approaches, the generated mappings are treated as transformed representations of the signal. Subsequently, these representations are subjected to further signal-processing techniques to extract features of interest, such as entropy, energy, and statistical properties [220, 221, 224–226, 229, 232, 235–237, 240, 245–248]. Compared to the other two approaches, these methods generally result in a smaller number of features, making them more suitable for applications with low computational requirements.

Spectral Kurtosis (SK) [253, 254] is a popular time-frequency analysis tool in signal processing that is commonly used to locate harmonics, transients, and repetitive impulses in the frequency domain. SK is based on computing the normalized fourth-order moment (kurtosis) of the STFT of the signal at each frequency bin, providing a statistical measure of the signal's characteristics at each time-frequency instant. Essentially, SK quantifies the “peakedness” or impulsiveness of the distribution of the power at each frequency bin. One of its primary applications is the detection of transient components in signals. Transients,

due to their short and sharp nature, create distinct, non-Gaussian patterns in the signal's frequency content. When these features are present, the kurtosis values in the affected frequency bins will be significantly higher than those of normal behavior. Accordingly, high SK values can indicate the presence of impulses or bursts, as they cause a deviation from the Gaussian distribution typically observed in stable operating conditions. This is particularly useful in condition monitoring and fault diagnosis of rotating machinery and mechanical systems [255–257], where transient vibrations often indicate faults. Additionally, SK is an effective tool for distinguishing between harmonic components— which often manifest as consistent, periodic components— and noise in the signal. These harmonics will often show higher kurtosis values, indicating a non-Gaussian distribution, possibly due to the periodic and repetitive nature of the harmonics. In contrast, Gaussian noise tends to have a kurtosis value close to 0, indicating a Gaussian distribution.

3.5.3 Signal Decomposition-Based Methods

Signals generated by physical systems typically consist of multiple components or modes of time-varying nature in amplitude, phase, and frequency. These components contain meaningful information about the underlying structure of the signal. Decomposition techniques attempt to decompose the signal into elementary modes, enabling signal analysis at the level of its constituent components. This, in turn, allows for the extraction of highly distinctive features, that are closely related to the inherent structure of the signal. Common decomposition approaches include wavelet-based decomposition and Adaptive Mode Decomposition (AMD) methods. In wavelet-base decomposition methods [56, 251, 258–278], Discrete Wavelet Transform (DWT), Stationary Wavelet Transform (SWT), and Wavelet Packet Transform (WPT) are commonly used to decompose the signal into elementary modes of high and low-frequency components using digital filter banks. The performance of wavelet decomposition is highly dependent on the base wavelet function and the level of decomposition. AMD methods [260, 265, 279–294] utilize an empirical decomposition

approach where the signal is adaptively decomposed into a set of simpler functions known as Intrinsic Mode Functions (IMFs). Hence, in contrast to wavelet-based methods, AMD methods use an adaptive mechanism rather than a priori basis function to decompose the signal, leading to more “natural” modes that are not influenced by a priori basis function [59]. Moreover, the ability of AMD methods to adapt to varying signal characteristics makes them effective when dealing with nonlinear and nonstationary signals. However, this adaptability comes at the cost of higher computational complexity. Furthermore, AMD methods are generally sensitive to noise because they rely on the extrema and minima of the signal for signal decomposition. Additionally, they suffer from inherent limitations, such as end effect [60] and mode mixing [61]. Various methods have been introduced to address these limitations, such as Ensemble Empirical Mode Decomposition (EEMD) [62] and Variational Mode Decomposition (VMD)[63], local mean decomposition [64], and Empirical Wavelet Transform (EWT) [65].

In the existing literature, signal decomposition is utilized through two approaches for feature extraction; each approach serves a distinct objective:

1. Extraction of highly discriminative features: In these methods, features are extracted from elementary modes resulting from the decomposition process. Accordingly, feature extraction methods can be generally grouped under four main categories: entropy-based [273, 274, 278, 284, 285, 294], energy-based [262, 268, 275, 276], spectral-based [56, 263, 277, 283, 286, 291, 293, 294], and statistical-based [258, 260–262, 264, 265, 271, 272, 286, 294]. Entropy as a measure of uncertainty, is used to quantify the irregularity within the decomposed modes. Energy-based approaches utilize the changes in energy content of the decomposed modes to extract the features. Spectral-based approaches use spectral characteristics of the elementary modes to extract the features. Statistical-based approaches rely on time-domain properties of the modes such as skewness, kurtosis, Root-Mean-Square (RMS), and

crest factor for feature extraction.

2. Obtaining a low-redundancy (highly informative) version of the composite signal [251, 259, 266, 267, 269, 270, 279–282, 287–290, 292]: This is accomplished by decomposing the signal into elementary modes and applying further processing to examine these modes based on predefined ranking criteria. The highest-ranked mode is thus selected and various feature extraction techniques can be applied to the selected mode accordingly. Here, the selected mode represents a low-redundancy and high-informativeness version of the original signal.

An important aspect of the first approach is that it provides a convenient way to construct a feature vector of controllable size and highly discriminative nature [56], which is particularly helpful in applications with limited computational resources. This is facilitated by decomposing the signal into a finite number of modes and extracting a few features from each mode. For instance, decomposing the input segment of a signal into 8 modes and taking the energy or entropy of each mode as a feature leads to a feature vector with a size of 1×8 distinctive features. It is also less computationally intensive than the second approach, making it more suitable for real-time processing and applications with limited processing capabilities. However, by focusing on the highest-ranked mode, the second method ensures that the extracted features are highly informative and less redundant. Moreover, the capability to rank and select modes based on predefined criteria allows for dynamic screening that can be automatically adapted according to the nature of the input signal, which becomes particularly useful in applications that involve multiple signals from various sources, such as advanced diagnostic systems and integrated condition monitoring applications. Additionally, selecting the most informative mode can result in more accurate and relevant feature extraction, potentially improving the performance.

3.5.4 Envelope-Based Methods

Envelope-based feature extraction methods [295–306] primarily focus on analyzing the envelope of the signal to reveal essential characteristics of its amplitude, such as statistical properties, peak frequency, energy, and entropy. In these methods, The HT is commonly used to obtain the signal’s envelope, which is then subjected to various feature extraction techniques. Envelope-based feature extraction is particularly useful in applications where Amplitude Modulation (AM) is a key characteristic of the signal; that’s it, the amplitude of the signal is varied or altered in response to specific factors that are of interest, for instance, impacts generated by defects in rolling bearings [307]; in such situations, envelope-based analysis can effectively extract meaningful features. Generally speaking, as an amplitude-focused approach, envelope analysis offers limited frequency information about the signal. Furthermore, its effectiveness can vary significantly depending on the nature of the signal, as not all signal types are amenable to envelope-based analysis. Additionally, it is highly susceptible to noise that contaminates the signal’s amplitude such as impulsive noise, making it difficult to extract effective features from the envelope [306].

The selection of an appropriate feature extraction method for a given signal depends on several factors, including the nature of the signal, the specific requirements of the application, and the available computational resources. For stationary signals, where the primary interest lies in frequency analysis, Fourier analysis provides a highly efficient approach to identifying dominant frequency components within the signal. In scenarios involving comparative spectral analysis of multiple signals that vary in length and/or bandwidth, PSD is particularly advantageous since it provides a normalized spectral density measure, facilitating effective comparison. On the other hand, signals exhibiting time-varying spectral characteristics necessitate the use of proper time-frequency analysis for effective feature extraction. The STFT offers a computationally efficient tool for time-frequency analysis of the signal. However, its fixed resolution across all frequencies reduces its effectiveness

for applications requiring variable-resolution analysis. In such contexts, wavelet analysis presents a practical solution due to its ability to perform multi-resolution analysis through scaled and shifted versions of a wavelet base function. For analyzing nonstationary and nonlinear signals characterized by rapidly changing components, the HHT provides an adaptive and more robust approach than Wavelet analysis but with more computational requirements. The WVD provides a practical approach for achieving a joint energy-time-frequency representation of the signal, although it is computationally intensive compared to the other transforms. It is particularly useful in high-resolution analysis scenarios or where energy concentration is a focal requirement. Signal decomposition techniques enable signal analysis at the level of its constituent components. This, in turn, allows for the extraction of highly distinctive features of a controllable size that are closely associated with the inherent structure of the signal. The envelope-based analysis is especially effective in applications where the envelope characteristics are of primary interest.

3.6 Conclusion

Addressing the existing gaps related to the role of signal processing in ML, this chapter has undertaken a comprehensive, integrated-article approach to present several contributions that aim to enrich the existing literature. First, the chapter made a solid foundation on the topic through a comprehensive tutorial on signal processing fundamentals. Written for a diverse readership, the tutorial allows interested readers to grasp essential concepts and develop a proper background in signal processing.

Furthermore, the chapter provided a comprehensive overview of a typical signal-processing pipeline, introducing a structured workflow for signal-based ML applications by categorizing tasks into preprocessing, processing, and application phases. Additionally, the chapter introduced an exhaustive review of feature extraction methods through a new taxonomy that clearly distinguishes between two main concepts in feature extraction: feature learning

and feature engineering, thereby offering new insights into the topic of feature extraction. Focused on signal processing-based feature extraction, the chapter reviewed various available techniques in terms of their main aspects, advantages, and limitations. Moreover, this work contributes to the research community by introducing a public repository of relevant Python and MATLAB codes for various signal-processing techniques.

Chapter 4

4 Similarity-Based Framework for VBCM using Limited Labeled Data

In this chapter¹, the second research problem of the thesis is presented, which is related to the limited availability of labeled training data needed for supervised learning-based Vibration-Based Condition Monitoring (VBCM). Classical VBCM methods depend on supervised learning, where a classifier is trained on a labeled dataset to determine the current state of the system. In practical scenarios, the amount and type of labeled data available for training classification models are restricted, which can make it challenging to use.

The first work in this chapter is entitled “Similarity-Based Predictive Maintenance Framework for Rotating Machinery” (Section 4.1). It is published as a technical paper in the *5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*. The presented work at the conference was recognized with the Best Paper Award, highlighting its significance and potential. In this work, the research problem is tackled by addressing the classification task as a similarity measure to a reference sample rather than a supervised classification task. Accordingly, a similarity-based framework for VBCM is proposed and applied to the fault detection and diagnosis of rotating machinery. This work tackles the research problem by addressing the classification task as a similarity measure to a reference sample rather than a supervised classification task. The work introduces a similarity-based framework for VBCM and applies it to the fault detection and diagnosis of rotating machinery.

The second work in this chapter is entitled “SB-PdM: a tool for predictive maintenance of rolling bearings based on limited labeled data” (Section 4.2) and was published as an invited paper in the *Software Impacts* journal. This work presents an open-source implementation

¹A version of this chapter has been published in [308] and [278].

of the similarity-based framework. The SB-PdM software is written in Python and released under the MIT license.

4.1 Similarity-Based Predictive Maintenance Framework for Rotating Machinery

Data-driven techniques are commonly adopted for VBCM in smart manufacturing. Classical approaches use supervised learning, where a classifier is trained on labeled data to predict or classify different operational states of the machine. However, in most industrial applications, labeled data is limited in terms of its size and type. Hence, it cannot serve the training purpose. In this section, this problem is tackled by addressing the classification task as a similarity measure to a reference sample rather than a supervised classification task. Similarity-based approaches require a limited amount of labeled data and, hence, meet the requirements of real-world industrial applications. Accordingly, the section introduces a similarity-based framework for VBCM and applies it to detect and diagnose faults in rotating machinery. For each operational state of the machine, a reference vibration signal is generated and labeled according to the machine's operational condition. Consequentially, statistical time analysis, Fast Fourier Transform (FFT), and Short-Time Fourier Transform (STFT) are used to extract features from the captured vibration signals. For each feature type, three similarity metrics, namely Structural Similarity Measure (SSM), cosine similarity, and Euclidean distance, are used to measure the similarity between test signals and reference signals in the feature space. Hence, nine settings in terms of feature type-similarity measure combinations are evaluated. Experimental results confirm the effectiveness of similarity-based approaches in achieving very high accuracy with moderate computational requirements compared to Machine Learning (ML)-based methods. Further, the results indicate that using FFT features with cosine similarity would lead to better performance compared to the other settings.

4.1.1 Introduction

Predictive Maintenance (PdM) approaches have been widely adopted in recent years for maintenance management in rotating machinery. PdM relies on continuously monitoring the equipment's condition, and actions for maintenance are predicted based on the equipment's actual condition. PdM involves two main tasks: First, extracting useful features from equipment-related data—such as vibration signals generated by rolling bearings of the rotating machinery— that can describe the process integrity well with high sensitivity to any changes within the process. The second task involves utilizing the extracted features to classify or predict normal and abnormal operational conditions with high accuracy. Common feature extraction methods include time domain analysis, frequency domain analysis, and time-frequency domain analysis [309][146]. In time-domain analysis, common statistical properties of the signal, such as kurtosis, skewness, crest factor, and peak, are used as features. Although time domain analysis is considered a simple approach to extract features, it has low sensitivity to process variations. Frequency domain analysis, such as FFT, allows the extraction of spectral-related features that are sensitive to variations in operational conditions. However, frequency domain analysis has no resolution in the time domain. Moreover, its application is limited to stationary signals. On the other hand, time-frequency domain analysis has better temporal and frequency localization compared with the Fourier analysis. Common time-frequency domain analysis methods include STFT, wavelet transform, and Hilbert–Huang Transform (HHT). Regarding the classification task, classical methods utilize supervised learning techniques to train a classifier on the extracted features. These methods usually require large-sized labeled data to fulfill training requirements. However, in most real-world situations, the available labeled data is limited in its size. Moreover, it is difficult to have labeled data that can model all possible classes or operational conditions sufficiently. For example, it is possible to obtain sufficient samples that can model standard or normal conditions. On the other hand, sam-

ples of abnormal conditions are usually not abundant and insufficient to model all possible abnormal operational conditions. Similarity-based approaches [309][310]–[311] offer an alternative solution to perform classification tasks with limited labeled data. In contrast to supervised learning, similarity-based techniques achieve classification tasks by measuring the similarity between a given test sample and a labeled reference sample, which can be achieved using very limited labeled data. This chapter introduces a similarity-based PdM framework for rotating machinery. The monitoring of process integrity is achieved by continuously analyzing the vibration signal generated by the rolling-element bearing. For each operational condition, a reference signal is generated and labeled according to the current operational condition. Then, signal processing-based methods are used to extract features from captured vibration signals. Accordingly, the similarity between test signals and reference signals is measured in the feature space to predict different operational conditions of test signals. The main contributions of the section are summarized as follows:

- Introducing a similarity-based framework for condition monitoring of rotating machinery. The main aspects of the framework are feature extraction and similarity-based classification.
- Three types of features, namely, time, frequency, and time-frequency features, are extracted from vibration signals.
- For each feature type, three similarity metrics are used for similarity-based classification. The three metrics are SSM, cosine similarity, and Euclidean distance.

The section is outlined as follows: The next section provides a review of related work. The methodology and framework are introduced in Section 4.1.3 and Section 4.1.4, respectively. Section 4.1.5 presents the dataset and the experimental setup for performance evaluation, while Section 4.1.6 discusses the results. The section is finally concluded in Section 4.1.7.

4.1.2 Related Work

Data-driven techniques are commonly adapted to perform PdM of rotating machinery using the vibration signals of rolling bearings [56]. However, the majority of the proposed techniques in the literature rely on ML-based classification, and less attention is paid to similarity-based approaches. In [146], a comparative study to evaluate the effectiveness of statistical time domain features with several classifiers is attempted. Results show that the extracted features are effective in identifying bearing faults. Further, it was found that the accuracy of classification increases as the length of the captured signal increases. However, increasing the length of captured signals would increase computational requirements. Moreover, it could delay triggering the faults. In [309], [310], and [312], frequency domain features and time-frequency domain features, along with similarity-based classification approaches, are utilized for rolling bearing condition monitoring. In [309], a labeled reference signal from each operation condition is generated in the first place. Consequentially, FFT is used to extract the features from the signals, and different operational conditions are determined by applying a proposed statistical similarity measure between test samples and reference samples. In [310], a similar approach is used. However, in contrast to [309], SSM [313] is adapted to measure the similarity. Further, Wavelet Packet Decomposition (WPD) is used to improve system robustness against noise and at the same time, increase its sensitivity to local differences in vibration signals. Consequentially, STFT is used to extract features from reconstructed signals. In [312], the spectrogram images of the test vibration signals are compared with spectrogram images of normal baseline vibration signals using SSM. Consequentially, normal and faulty vibration signals are classified by setting a threshold on the resulting SSM scores. The main aspect in [309], [310], and [312] is the use of signal processing for feature extraction along with similarity-based classification, which eliminates the need for machine learning-based trained classifiers. This chapter introduces a framework for similarity-based condition monitoring of rotating machinery and

uses statistical time properties, FFT, and STFT to extract the features from variation signals. Further, in contrast to [309], [310], and [312] where statistical and structural similarities are used, This work uses cosine similarity and Euclidean distance—which are less complex compared to SSM—to measure the similarity between reference and test samples.

4.1.3 Methodology

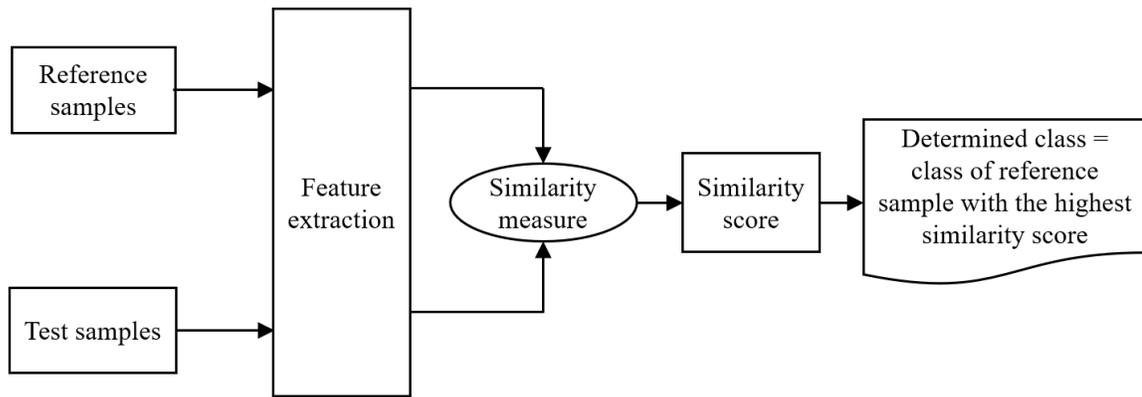


Figure 4.1: similarity-based classification framework

Fig.4.1 shows a flowchart of the similarity-based classification framework. In the first stage, discriminative features are extracted from labeled reference samples and test samples. Consequentially, the similarity between reference samples and test samples is measured in the feature space. Finally, the classification of different operational conditions is achieved by evaluating the resulting similarity scores. In contrast to machine learning-based approaches, the framework utilizes similarity scores to determine operational conditions instead of a trained classifier. The main aspects of the framework are feature extraction and similarity measures. Extracted features should be selected so that they satisfy two main conditions:

- Describe the inherent characteristics of all operational conditions “classes” in the data.
- Have a high discrimination degree between the different operational conditions in the

data.

Once features are extracted, the similarity measure is applied to quantify the similarity between reference samples and test samples in the feature space. To perform the similarity measure, a reference sample from each operational condition “class” should be available. Since the similarity measure provides a quantitative value, it can be used to assess the probability that the reference sample and test sample belong to the same operational condition. The higher the similarity, the higher the probability they belong to the same condition. This can be expressed mathematically as follows: For a given test sample $X_n, n = 1, \dots, N$ and a reference sample $Y_i, i = 1, \dots, m$, the similarity score s_{ni} can be defined as:

$$s_{n,i} = S(F(X_n), F(Y_i)) \quad (4.1)$$

where N is number of test samples, m is number of operational conditions or classes, i is the class of the reference sample, S is the applied similarity measure, and F denotes the feature extraction function. The operational condition C_{X_n} of X_n can be determined according to the below equation;

$$C_{X_n} = i \text{ if } s_{n,i} = M(S) \quad (4.2)$$

where,

$$M(S) = M\{s_{n,1}, s_{n,2}, \dots, s_{n,m}\}, \quad (4.3)$$

$M\{\cdot\}$ denotes the “maximum” or “minimum” operation depending on the similarity measure type.

4.1.4 Predictive Maintenance Framework for Condition Monitoring of Rotating Machinery

In this section, the similarity-based framework is presented in detail. The framework performs PdM by continuously analyzing the vibration signal of the rolling-element bear-

ing. To simulate a noisy environment, vibration signals are corrupted with Additive White Gaussian Noise (AWGN) at different Signal-to-Noise Ratio (SNR) levels. To increase the system's robustness against the noise, the signals are denoised using WPD. For each operational condition, a labeled reference signal is generated. Its label represents the class to which the signal belongs in terms of operational condition. To simulate the effects of a noisy environment, reference, and test signals are corrupted by AWGN at the desired SNR. In the next step, signal denoising is applied. Consequently, features are extracted from denoised signals. In the final stage, similarity in the feature space is measured between each test signal and all labeled reference signals. Accordingly, for a given test signal, its determined class will be the class of the reference signal with the highest similarity score.

Signal Denoising

Generally, noise presence in the signal is characterized by high-frequency components. Thus, signal denoising can be accomplished by decomposing the signal using the WPD technique and filtering out detail coefficients associated with higher frequency sub-bands. This can be achieved by thresholding the detail coefficients so that coefficients below the threshold are set to zero. The denoised signal is then reconstructed using the approximation and the thresholded detail coefficients. In this chapter, Daubechies 4 (db4) wavelet is used to decompose the noisy signals, and the soft thresholding function [314] is applied to the details coefficients. The threshold value is determined according to the below formulas [315]:

$$threshold = \sigma \sqrt{2 \log(N)/N}, \quad (4.4)$$

$$\sigma = \frac{median(|w_k|)}{0.6745}, \quad (4.5)$$

where N is signal length, and $|w_k|$ are wavelet coefficients.

Feature Extraction

Time domain analysis, frequency domain analysis, and time-frequency domain analysis are used to extract features from vibration signals. The aim here is to evaluate and compare the performance of time, frequency, and time-frequency features in a noisy environment. The statistical properties of vibration signals will be used as time-domain features. FFT will be used to obtain spectral components of the signals, and the positive part of the spectrum will be used as frequency domain features. For time-frequency domain analysis, STFT will be used. Fig. 4.2 shows FFT and STFT contents of the bearing's vibration signal of a normal operational condition along with a faulty operational condition. As shown, frequency components of the vibration signal are very sensitive to changes in operational conditions. Thus, in contrast to time-domain features, STFT and FFT provide very useful features with a high discrimination degree for the classification of different operational conditions.

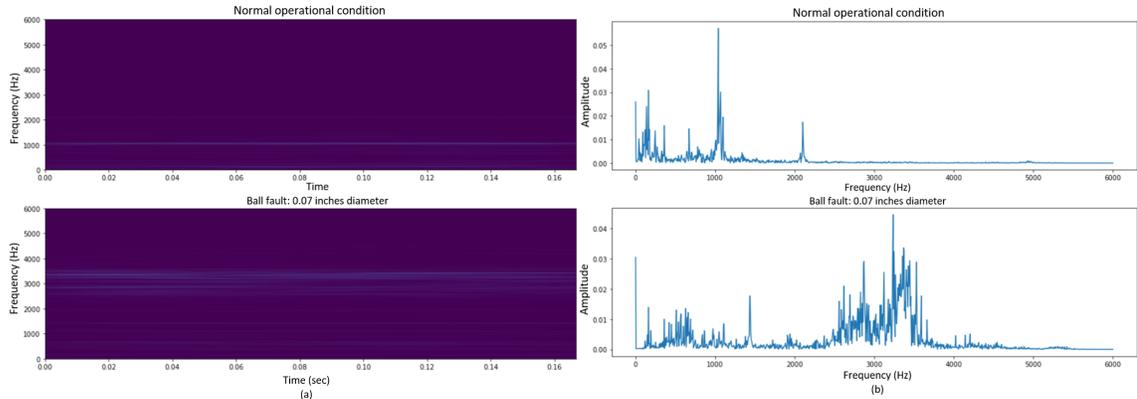


Figure 4.2: (a) STFT and (b) FFT of normal (healthy) operational condition and a faulty operational condition (ball fault of 0.07 inches in diameter)

Similarity Measure

To evaluate the similarity between test signals and reference signals in the feature space, three different similarity measures are used and compared in terms of classification accuracy. The three measures are cosine similarity, Euclidean distance, and SSM [313]. The

cosine similarity C_s between two vectors X and Y of length N is;

$$C_s(X, Y) = \frac{\sum_{n=1}^N X_n \cdot Y_n}{\sqrt{\sum_{n=1}^N X_n^2} \cdot \sqrt{\sum_{n=1}^N Y_n^2}} \quad (4.6)$$

For a given test signal X_n and m reference signals $Y_i, i = 1, ..m$, the class of X_n is determined as follows:

$$\text{Class of } X_n = i \text{ if } C_s(F(X_n), F(Y_i)) = \max\{D\} \quad (4.7)$$

where,

$$D = \{C_s(F(X_n), F(Y_1)), \\ C_s(F(X_n), F(Y_2)), \dots, C_s(F(X_n), F(Y_m))\}. \quad (4.8)$$

The Euclidean distance between two vectors X and Y of length N is:

$$d(X, Y) = \sqrt{\sum_{n=1}^N (X_n - Y_n)^2}, n = 1, \dots, N \quad (4.9)$$

For a given test signal X_n and m reference signals $Y_i, i = 1, ..m$, the class of X_n is determined as follows:

$$\text{Class of } X_n = i \text{ if } d(F(X_n), F(Y_i)) = \min\{D\} \quad (4.10)$$

where,

$$D = \{d(F(X_n), F(Y_1)), \\ d(F(X_n), F(Y_2)), \dots, d(F(X_n), F(Y_m))\}. \quad (4.11)$$

The SSM algorithm [313] is an image quality assessment metric; it provides a perceptual metric to quantify the degradation in image quality caused by image processing such as compression, transmission, etc. The algorithm requires two input images, a processed

image, and its reference image. The structural similarity between the input images is then evaluated by comparing the luminance, contrast, and structure of the two input images. The output is a value between 0 and 1 that quantifies the quality of the processed image with respect to the reference image. The higher the value, the higher the quality. Accordingly, SSM is defined as [313]:

$$SSM(X, Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4.12)$$

where, $\sigma_{xy}, \mu_x, \mu_y, \sigma_x, \sigma_y$ are covariance, means and standard deviations of vectors X and Y ; calculated over a window of size w , in this chapter, the window size is set to 7. $C_{1,2}$ are arbitrary constants to avoid unstable output when either $(\mu_x^2 + \mu_y^2)$ or $(\sigma_x^2 + \sigma_y^2)$ is very close to zero. For a given test signal X_n and m reference signals $Y_i, i = 1, ..m$, the class of X_n is determined as follows:

$$\text{Class of } X_n = i \text{ if } SSM(F(X_n), F(Y_i)) = \max\{D\} \quad (4.13)$$

where,

$$D = \{SSM(F(X_n), F(Y_1)), \\ SSM(F(X_n), F(Y_2)), \dots, SSM(F(X_n), F(Y_m))\}. \quad (4.14)$$

In this section, SSM is used to measure the similarity between reference samples and test samples in the feature space where X and Y represent the extracted features of test samples and reference samples, respectively.

4.1.5 Performance Evaluation

The Case Western Reserve University (CWRU) bearing test data [316] is used to evaluate the performance of the framework. The framework is implemented in Python program-

Table 4.1: Bearing test data

Class	Drive End bearing data					
	Operational condition	Fault diameter	Motor speed (rpm)			
1	Normal		1730	1750	1772	1797
2	Inner Race fault	0.07"	1730	1750	1772	1797
3		0.014"	1730	1750	1772	1797
4		0.021"	1730	1750	1772	1797
5	Ball fault	0.07"	1730	1750	1772	1797
6		0.014"	1730	1750	1772	1797
7		0.021"	1730	1750	1772	1797
8	Outer Race fault	0.07"	1730	1750	1772	1797
9		0.014"	1730	1750	1772	1797
10		0.021"	1730	1750	1772	1797

ming language using Numpy [317], Pandas [318], Scikit-learn [319], SciPy [320], and scikit-image [321], libraries. The CWRU bearing test data consists of vibration signals of normal and faulty bearings. Faults ranging from 0.007 inches in diameter to 0.040 inches in diameter were introduced separately at the inner raceway, the ball, and the outer raceway. Faulted bearings were reinstalled into the test motor, and vibration data was recorded for motor loads of 0 to 3 horsepower (motor speeds of 1797 to 1720 RPM), and digital data was collected at 12,000 samples per second. Table 5.1 shows the operational condition, fault diameter, and motor speed of these vibration signals. According to the operational condition and fault diameter, abnormal operational conditions are classified into 9 classes, as shown in the table. Thus, the dataset consists of 10 operational conditions or classes, one normal operational class, and nine faulty operational classes.

Before processing the signals, each vibration signal is divided into samples of 2000 data points each. With a sampling rate of 12000 KHz, this gives a sampled vibration signal of 0.166 seconds interval [310], which would be short enough to serve the purpose of condition monitoring and reduce computational requirements. The resultant sampled dataset consists of 3019 vibration signals. For each class, one reference signal is selected for each motor speed. This yields a total of 40 reference signals, with 4 reference signals for each class and a total of 2979 test signals.

The next step involves corrupting reference and test signals with AWGN according to the desired SNR. Consequently, signals are denoised using WPD. After the denoising operation, reference and test signals are fed into the feature extraction stage. Time domain features are extracted as proposed in [146]. Frequency domain features and time-frequency domain features are extracted according to settings in [309] and [310], respectively. After the feature extraction stage, reference signals of the same class are combined into one reference signal by taking their mean. This will reduce computational requirements during the next stages. Also, it is found that averaging same-class reference signals at this stage improves the accuracy compared to averaging them before the feature extraction stage. This improvement in accuracy can be justified as follows: Averaging same-class reference signals into one signal after the feature extraction stage will maintain feature-related characteristics of each averaged signal and will reduce the probability of error in the similarity-based classification.

4.1.6 Results and Discussion

The performance of the framework is evaluated in terms of classification accuracy, which quantifies the correctly classified test samples as a percentage of total test samples. Table 5.4 shows accuracy performance under different SNR conditions. As shown, the classification accuracy of the three similarity measures with time-domain, FFT, and STFT features is evaluated at SNR levels of 2 dB, 4 dB, 6 dB, 8 dB, 10 dB, and 20 dB. The aim is to evaluate the effectiveness of the extracted features and the employed similarity measures under extreme, moderate, and low noisy conditions.

The first observation in the results is that STFT and FFT features have similar performance; both achieved very high accuracy (> 98%) under all noisy conditions. It is noticeable that even though FFT has no time resolution compared to STFT, it achieved slightly better performance. From Fig.4.2.a, it is clear that the STFT spectrogram does not reveal any obvious transitions in the frequency within the captured signal's length of 0.166 seconds.

Table 4.2: Accuracy results of time, frequency, and time-frequency features

Extracted Features		Accuracy (%)									Results of [308]	Results of [309]
		Time domain features			Frequency domain features			Time-frequency domain features				
Similarity Measure		SSM	Cosine Similarity	Euclidean distance	SSM	Cosine Similarity	Euclidean distance	SSM	Cosine Similarity	Euclidean distance		
SNR	2 dB	80.52%	73.38%	77.54%	97.89%	98.57%	98.01%	98.01%	97.73%	97.98%	NA	NA
	4 dB	81.70%	76.86%	80.00%	98.06%	99.09%	98.88%	98.38%	98.04%	98.79%	NA	96.35%
	6 dB	84.12%	76.98%	82.08%	98.91%	99.22%	99.07%	98.66%	98.35%	98.85%	NA	96.75%
	8 dB	86.14%	78.75%	84.62%	98.94%	99.13%	99.07%	98.82%	98.54%	99.01%	NA	97.15%
	10 dB	86.77%	81.11%	85.87%	98.97%	99.50%	99.13%	98.85%	98.57%	99.07%	99.30%	97.75%
	20 dB	88.69%	81.64%	87.42%	99.22%	99.57%	99.35%	99.35%	98.73%	99.32%	NA	NA
Average Accuracy		84.66%	78.12%	82.92%	98.67%	99.18%	98.92%	98.68%	98.33%	98.84%	NA	97.00%

Thus, the time resolution aspect of STFT didn't add more discriminative characteristics to the STFT features compared to the FFT features. Instead, it may cause slight redundancy in the extracted features.

The second observation is that time-domain features have lower performance compared to STFT and FFT. In contrast to frequency contents, the statistical time properties extracted from the vibration signal's waveform have low sensitivity to the variations in operational conditions, especially with small lengths of captured signal [146]. Fig.4.3 demonstrates the effectiveness of FFT and STFT features in similarity-based classification for vibration-based condition monitoring. The figure shows two-dimensional scatter plots of test samples with time-domain, FFT, and STFT features. For each feature type, the samples are represented on the $X - Y$ plane by their scores of the two best-performing similarity measures. Specifically, for FFT and STFT features, cosine similarity along with Euclidean distance are used to plot the test samples on the $X - Y$ plane. For time-domain features, SSM score and Euclidean distance are used. The sample colors represent the operational classes, showing how well the samples cluster in the feature space. As shown, unlike time-domain features, FFT and STFT features have large difference margins in terms of similarity scores between the operational classes, which is reflected by the large distance between inter-class samples compared to intra-class samples. This clearly shows that FFT and STFT features can effectively differentiate between classes, which, in turn, demonstrates the high discrim-

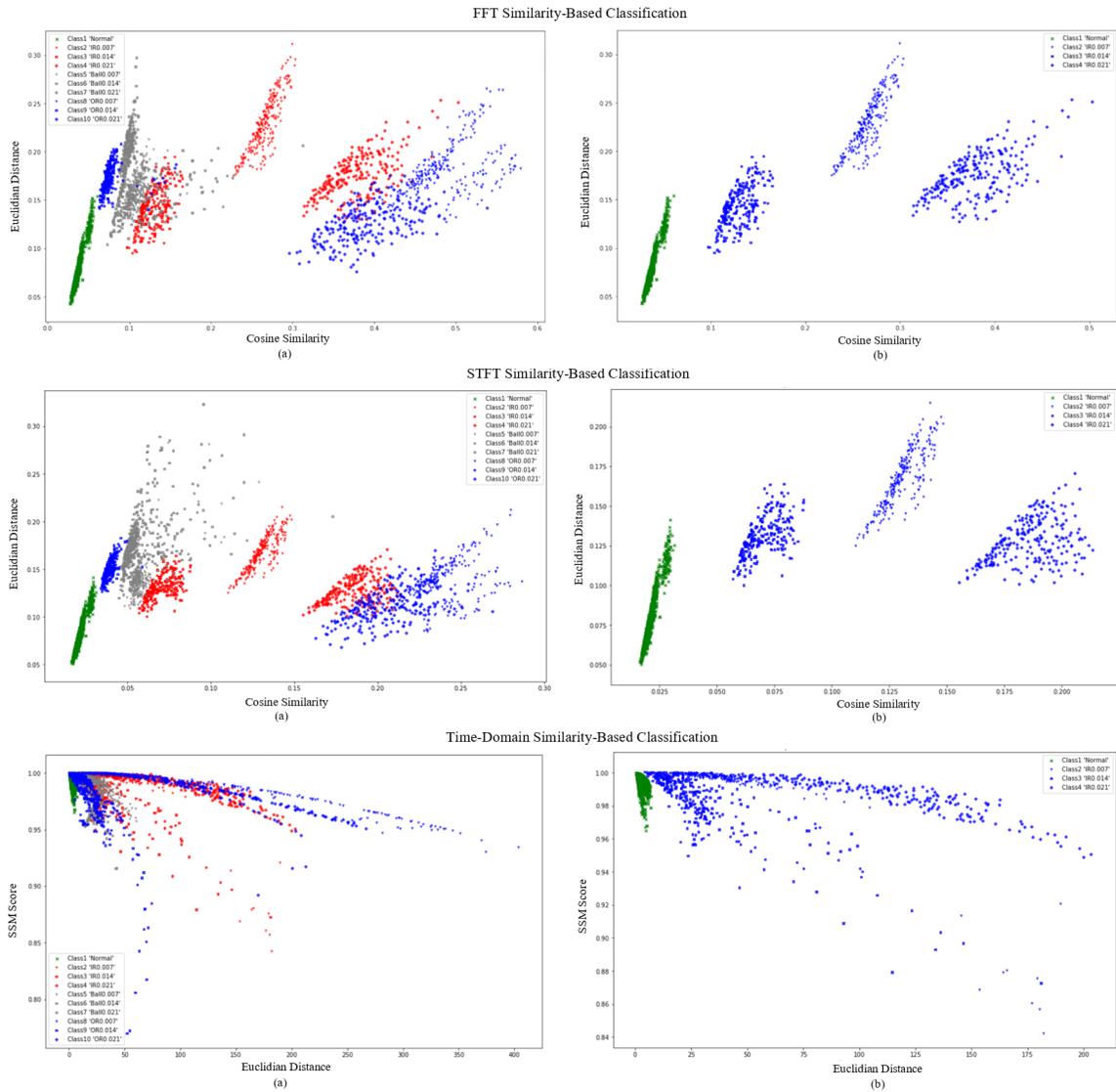


Figure 4.3: scatter plot shows (a) all classes and (b) normal and inner race faulty classes

inative capabilities of FFT and STFT features between the different operational classes compared to time-domain features.

Regarding the effectiveness of the applied similarity measures, all three measures are effective in classifying different operational conditions with FFT and STFT features. While the three similarity measures have very similar performance, the cosine similarity with FFT features achieved the best performance compared to SSM and Euclidean distance with an average accuracy of 99.57%. Table 5.4 also compares the obtained results with recent related work [309][310]. As shown in the table, the obtained FFT results are in conformance with the results of [309] where FFT is used to extract the features. However, the obtained STFT results are better than the results of [310] even though STFT with the same settings is used in this section. In contrast to the signal denoising approach that is used in this section, the signals in [310] are decomposed using WPD and reconstructed using only characteristic fault frequencies and the most impulsive frequency bands. As a result, some of the useful frequency components could be filtered out, which would explain the performance gap between the two approaches.

4.1.7 Conclusion

In this section, a data-driven PdM framework using similarity-based classification is introduced for condition monitoring of rotating machinery. The framework addresses real-world situations of limited availability of labeled data by applying a similarity-based classification where one labeled sample for each operational condition is enough to carry out the classification task. The performance was investigated under different noisy conditions, and WPD was employed to denoise the signals and increase system robustness against the noise. Nine settings of feature type-similarity measure combinations were evaluated under different SNR levels. Experimental results corroborate the effectiveness of similarity-based approaches in vibration-based condition monitoring. Further, results demonstrate the capability of similarity-based approaches in achieving very high accuracy with moderate

computational requirements compared to machine learning-based methods. Moreover, the results indicate that using FFT features with cosine similarity would lead to better performance compared to the other settings.

4.2 SB-PdM: a Tool for Predictive Maintenance of Rolling Bearings Based on Limited Labeled Data

Most of the existing code repositories of condition monitoring software are based on traditional ML-supervised algorithms that require labeled datasets to function. The SB-PdM software has been developed and released publicly to address this limitation with the aim that researchers and professionals in the industry can use and replicate the code in their research and projects and advance the field of VBCM.

Classical PdM methods of rolling bearings rely on supervised learning, which involves training a classifier on a labeled dataset to classify the current state of the bearings. In practical situations, labeled data may be limited in terms of its size and type, which can make it difficult to use for training predictive maintenance models. The proposed Similarity-Based PdM (SB-PdM) software addresses the challenge of limited labeled data by approaching the classification task as a similarity measure to a reference sample rather than using supervised classification. Experimental results have confirmed the effectiveness of the proposed SB-PdM software in achieving very high accuracy while requiring only moderate computational resources compared to traditional ML-based methods.

4.2.1 Introduction

PdM approaches have been widely adopted for the maintenance management of rolling bearings. The main concept of PdM is to continuously monitor the condition of the bearings and predict maintenance actions based on their actual condition. PdM typically in-

volves two main tasks. The first task is to extract useful features from bearing-related data, such as vibration signals, using various signal processing techniques. The second task is to utilize the extracted features to classify or predict normal and abnormal operational conditions, such as identifying potential faults or predicting remaining useful life. Classical methods for PdM typically use supervised learning techniques [322] to train a classifier on the extracted features. However, these methods can be limited by the requirement for large amounts of labeled data to achieve optimal performance. In many real-world industrial situations, labeled data may be limited in size, making it difficult to train accurate predictive models using supervised learning alone. Moreover, it can be challenging to obtain labeled data that can effectively model all possible classes or operational conditions, particularly for abnormal conditions, which may be rare or unpredictable. For example, while it may be possible to obtain sufficient samples that can model standard or normal conditions, samples of abnormal conditions may be insufficient or unavailable to model all possible abnormal operational conditions accurately. Similarity-based approaches [309] [310][312][323] offer an alternative solution to perform classification tasks with limited labeled data. In contrast to supervised learning, similarity-based techniques perform the classification task by 1) measuring the similarity between a given test sample and class-labeled reference samples and 2) utilizing the resultant similarity scores to determine the class to which the test sample belongs. With respect to the availability of open-source software, most existing code repositories for PdM applications are built on ML-supervised classification and assume wide accessibility to labeled datasets. Therefore, the practical situations of data unavailability, such as in the early stages of PdM solution deployment and specialized applications, are often overlooked in this regard. The proposed SB-PdM software aims to address these practical situations and fill the existing gap in the availability of open-source software for PdM applications.

The remainder of this section is organized as follows: Section 4.2.1 highlights the concepts of similarity-based classification and presents details of the SB-PdM software. Section

4.2.2 evaluates the software’s performance using a real-world dataset. Section 4.2.3 highlights the Software’s impacts, and the conclusion is presented in Section 4.2.4.

4.2.2 SB-PdM Key Aspects and code functionalities

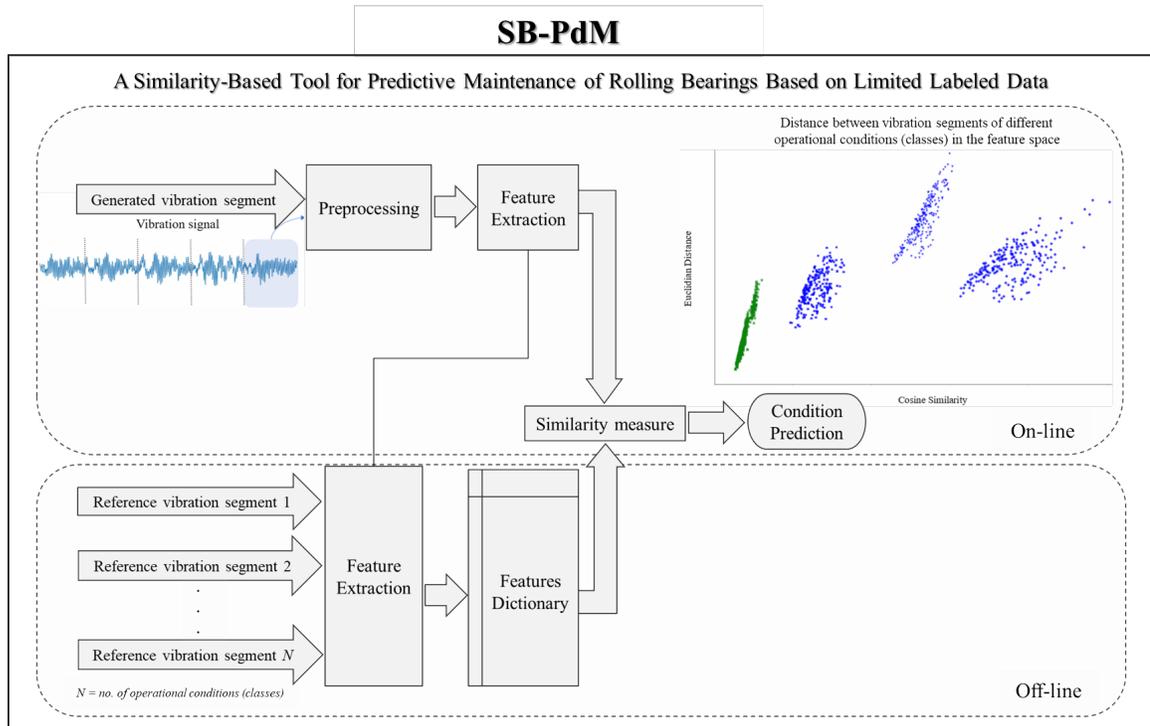


Figure 4.4: A high-level overview of the SB-PdM software

A high-level overview of SB-PdM is illustrated in Fig. 4.4. To perform similarity-based classification, a labeled reference vibration segment should be available for each operational condition “class” as shown in the figure. The monitoring of process integrity is then achieved by continuously measuring the similarity in the feature space between generated vibration segments and the labeled reference vibration segments. Accordingly, the SB-PdM software achieves the classification of different operational conditions by evaluating the resulting similarity scores. The higher the similarity between a vibration segment and a specific labeled reference segment, the higher the probability that they belong to the same condition “class”. It is worth noting that in similarity-based classification, the reference sample represents its corresponding class, and as such, it should sufficiently describe all

variations or patterns within the class [323]. For instance, in a rotating machinery application where the motor speed can vary, it is important to have a reference sample for each speed profile to better model the different operational conditions of the machine.

Key Aspects of SB-PdM Software

The SB-PdM software involves two main stages: **feature extraction** and **similarity measure**. The main highlights of each stage are provided below.

Feature Extraction: The main aspect of the extracted features is that they should be directly related to the operational process with high integrity and sensitivity to any changes within the process. Common feature extraction methods include time domain analysis [146], frequency domain analysis [309], and time-frequency domain analysis [310][56]. In time-domain analysis, statistical properties of the signal, such as kurtosis, skewness, crest factor, and peak value are used as features. Although time-domain analysis is a simple approach for extracting features, it may have low sensitivity to process variations, particularly with short durations of vibration segments. Furthermore, time-domain analysis is generally not effective for use in high-noise environments. Frequency domain analysis, such as FFT, allows for the extraction of spectral-related features that are sensitive to variations in operational conditions. However, frequency domain analysis can capture the frequency content of the signals, but it cannot provide information about the temporal variations in the signals. On the other hand, time-frequency domain analysis, such as wavelet analysis, can provide better temporal and frequency localization compared to Fourier analysis. Common time-frequency domain analysis methods include the STFT, Wavelet Packet Transform (WPT), and Hilbert-Huang transform (HHT). In the SB-PdM software, features are engineered so that they fulfill two main requirements:

- Describe the inherent characteristics of all operational conditions “classes” in the data.

- Have a high discrimination degree between the different operational conditions in the data.

These requirements are fulfilled by utilizing frequency domain analysis and time-frequency domain analysis to extract the features. Specifically, FFT, STFT, and WPT are used to extract features from vibration segments.

Similarity Measure: Regarding the similarity measure, it should provide a quantitative value that represents the similarity score between the two inputs. This similarity score can be used to assess the probability that the two inputs belong to the same class. The higher the similarity score, the higher the probability that they belong to the same class. In SB-PdM, the similarity measure is applied to the extracted features to quantify the similarity between the labeled reference vibration segments and the generated vibration segment in the feature space. Hence, it works as a similarity-based classifier. In contrast to supervised classification, where a labeled dataset is required for training, similarity-based classification requires one labeled reference vibration segment for each operational condition “class”. The SB-PdM uses the SSM [313], cosine similarity, and Euclidean distance to perform similarity measures between the extracted features of labeled reference segments and generated vibration segments.

SB-PdM code functionalities

This section describes the functionalities of the code repository on GitHub [324]. The software is developed in Python programming language and utilizes Numpy [317], Pandas [318], Scikit-learn [319], SciPy [320], scikit-image [321], and PyWavelets [325] packages. The GitHub repository includes the following files.

Reference_Samples.csv and Test_Samples.rar: These are *.csv* files that contain reference-labeled vibration segments and test segments. The vibration segments are extracted from

the Case Western Reserve University (CWRU) bearing data [316]. The data consists of vibration signals of normal and faulty bearings. Faults ranging from 0.007 inches in diameter to 0.040 inches in diameter were introduced separately at the inner raceway, the ball, and the outer raceway to simulate different types of faults. The faulted bearings were reinstalled into the test motor, and vibration data was recorded for motor loads ranging from 0 to 3 horsepower, corresponding to motor speeds of 1730, 1750, 1772, and 1797 RPM. Digital data was collected at a rate of 12,000 samples per second to capture the vibration signals. Based on the operational condition and fault diameter, abnormal operational conditions are classified into nine classes. Thus, vibration data consists of 10 operational conditions or classes, including one normal operational class and nine faulty operational classes that represent different types of faults. Each vibration signal is divided into segments of 2,000 data points each. With a sampling rate of 12000 KHz, this results in a vibration segment of 0.166 seconds interval, which is short enough to serve the purpose of condition monitoring and reduce the computational requirements. The resultant segmented vibration data consists of 3,019 vibration segments. For each class, one reference segment is selected from each of the four motor speed profiles, resulting in a total of 40 reference segments, with four reference segments for each class. In total, there are 2979 test segments, which can be used to evaluate the performance of the SB-PdM software.

SB_PdM_Tool.ipynb: This is Jupiter notebook that contains the SB-PdM code. It includes five functions:

1. *apply_fft*: This Function applies FFT to the input vibration segment “*x*” to obtain its spectral contents. The positive part of the spectrum “*freq_values*” is used as frequency domain features. The parameters “*fs*” and “*num_samples*” provide the sampling frequency of vibrating segments and the number of data points in the input segment, respectively.
2. *apply_stft*: Function to perform STFT to the input vibration segment. STFT pro-

vides time-domain analysis of the input vibration segment, which allows extraction of time-frequency features since STFT determines the frequency contents of the segment as it changes over time. This is achieved by dividing the input time-domain vibration segment “ x ” into local sub-segments of equal time lengths “ seg_length ”, which is typically performed by multiplying the signal with a sliding window function “ $window$ ”. The percentage of overlap between sub-segments is controlled by the parameter “ $num_overlaps$ ”. The spectrum of each segment is then obtained by computing its Fourier transform with a length equal to “ fft_length ”. This results in a matrix consisting of the power spectral density (PSD) “ s ” of each sub-segment. This matrix represents the vibration segment’s spectrogram, which is a two-dimensional representation of the segment spectrum in the time-frequency domain. Hence, the spectrogram shows the resolution of the segment in the time-frequency domain. The STFT features are obtained by taking the absolute value “ $abs(s)$ ” of the resulting PSD matrix of vibration segments. In the software, the STFT parameters are set as follows.

- $seg_length = num_samples/2$.
 - $window =$ Hamming window.
 - $num_overlaps = 95\%$ of $num_samples$.
 - $fft_length = 1028$.
3. *wpt_entropy*: This function extracts WPT features from input vibration segment “ x ”. WPT decomposes the input vibration segment using scaled versions of a base wavelet function into elementary waveforms with high localization in time and frequency. The first step in the function involves decomposing the input vibration segment using num_levels -level WPT decomposition. The parameter *wavelet_function* specifies the wavelet base function to be used for the decomposition. In SB-PdM software, the *Daubechies4 (dp4)* wavelet is used. More details on the selection of the proper

wavelet can be found in [56]. Accordingly, in the second step, 2^{num_levels} elementary waveforms of lower and higher frequency sub-bands are reconstructed from individual wavelet coefficients. In the third step, Shannon entropy is applied to each resultant waveform to obtain its entropy. Hence, a WPT-entropy feature vector “*wpt_entropy_features*” is constructed for each input vibration segment. In the software, *num_levels* is set to 5.

4. *similarity_measures*: Function to perform similarity measures. It applies similarity measures between features of reference segments “*baselines_features*” and features of test segments “*test_samples_features*”. The applied similarity measures are SSM, cosine similarity, and Euclidean distance. For each similarity measure, a record that contains similarity scores between a given test segment and all labeled reference segments is created (*ssim_record*: SSM scores, *dist_record*: cosine similarity scores, and *dist2_record*: Euclidean distance scores). Hence, the length of the score record equals to number of reference segments. The first value in the record represents the similarity score between the given test segment and first reference segment, and the second value represents the similarity between the second reference segment and so on. Accordingly, the function finds the highest similarity score in the record and utilizes the labels vector of the reference segments “*baseline_labels*” to obtain the corresponding label/class which will be the determined class of the test segment. The function returns the following items:

- *y_ssim*: Predicted classes using cosine similarity scores of test vibration segments.
- *y_cos*: Predicted classes using SSM scores of test vibrations segments.
- *y_Ecl*: Predicted classes using Euclidean distance scores of test vibration segments.

5. *similarity_based_classification*: This function recalls *similarity_measures* function

has the same inputs, and provides the same outputs. The parameter “*baselines_no_avg*” is a Boolean type and specifies whether the features of reference segments belonging to the same class should be averaged or not. The default is False (*i.e.* averaging the features). It found that averaging reference segments of the same class after the feature extraction stage and prior to applying similarity-based classification improves the accuracy compared to averaging them before the feature extraction stage. Further, averaging reference segments of the same class will reduce computational complexity during the next stages. The archived results confirm the effectiveness of the software in detecting and diagnosing faults in the rolling bearings.

4.2.3 Performance Evaluation

The performance of the SB-PdM software is evaluated in terms of classification accuracy, which quantifies the correctly classified test samples as a percentage of total test samples. By using FFT features, the achieved accuracy results are 99.4%, 99.6%, and 99.4% with SSM, cosine similarity, and Euclidean distance, respectively. With WPT entropy features, the achieved results are 98.52%, 93.52%, and 95.50% with SSM, cosine similarity, and Euclidean distance, respectively. Lastly, using STFT features, the achieved results are 99.33%, 98.76%, and 99.36% with SSM, cosine similarity, and Euclidean distance, respectively. These results confirm the effectiveness of the SB-PdM software in detecting and diagnosing different faults of rolling bearings. Regarding memory requirements, while WPT entropy achieved a slightly lower performance compared to FFT and STFT, it is more memory-efficient since it results in the lowest size of features compared to the two other methods. Specifically, with 5-levels of decomposition, the WPT entropy extracts 1×32 features from the input vibration segment. On the other hand, FFT and STFT extract 1×1000 features and 515×41 features, respectively.

4.2.4 Software impacts

Building a reliable and comprehensive dataset for the condition monitoring and fault diagnosis of rolling bearings is a time-consuming process that requires costly dedicated resources. Moreover, selecting and tuning an appropriate ML model is a challenging task [326]. Further to add, samples of vibration signals under abnormal conditions are often not as abundant as samples of vibration signals under normal conditions. The SB-PdM software offers a practical solution in such situations because it requires one labeled reference vibration sample for each operational condition “class”, instead of a complete labeled dataset. This flexibility in data availability meets the requirements of many real-world industrial applications, especially in the early phases of solution deployment where only a limited amount of labeled data is available. Moreover, most of the existing code repositories of condition monitoring software are based on traditional ML-supervised classification and require labeled datasets to be available. From this perspective, the SB-PdM is considered an innovative and practical software that fills this gap in the availability of open software that can work with limited labeled datasets. In addition to these advantages, reproducibility and transparency are two other benefits of the SB-PdM software. With the SB-PdM software’s source code publicly available, researchers and professionals in the industry can use and replicate the code in their research and projects and advance the field of VBCM. The SB-PdM software is written entirely in Python, which is an advantage as Python is an easy-to-understand language. Additionally, the software’s compatibility with other libraries and projects, due to Python’s flexibility and widespread use in related applications, makes it easy to integrate into other projects.

Finally, the similarity-based classification method used in the SB-PdM software can be applied as a general approach to solving similar classification problems in other fields, such as anomaly detection, image classification, disease diagnostics, and pattern recognition. As a result, a broader community of researchers and data analysts can benefit from it.

4.2.5 Conclusion

The SB-PdM software, a data-driven PdM tool based on similarity-based classification, is introduced for condition monitoring of rolling bearings. This software tackles the limited availability of labeled data by utilizing a similarity-based classification, requiring only one labeled sample for each operational condition. The software has demonstrated high accuracy with moderate computational requirements compared to machine learning-based methods, as shown in experimental results. These results highlight the effectiveness of the software for vibration-based condition monitoring of rolling bearings.

Possible future research directions for the SB-PdM software include developing more effective signal denoising methods, exploring more efficient techniques for feature extraction, and designing customized similarity measures. Rolling bearings often operate in noisy environments with multiple sources of noise, resulting in corrupted vibration signals. Therefore, understanding noise characteristics and how they affect vibration signals is crucial in developing effective signal-denoising techniques.

Regarding feature extraction, the SB-PdM transforms vibration segments into feature representations with high-discriminative capabilities compared to time-domain representations and measures the similarity between reference and input vibration segments in the feature space. The software currently uses FFT, STFT, and WPT for feature extraction. However, future research directions could explore other methods for extracting more discriminative features with less computational requirements.

A third research direction could investigate the possibility and feasibility, from both accuracy and computational complexity perspectives, of developing a customized similarity measure to be applied directly to the time-domain representations of the vibration segments. The goal here is to simplify the software and reduce memory requirements and processing time by eliminating the feature extraction stage.

Chapter 5

5 A Hybrid Method for Condition Monitoring and Fault Diagnosis of Rolling Bearings With Low System Delay

Addressing the third research problem, this chapter¹ introduces the first contribution related to the computational complexity of Vibration-Based Condition Monitoring (VBCM). Computational complexity, in terms of memory requirements and monitoring delay, plays a pivotal role in the deployment of Vibration-Based Condition Monitoring (VBCM). The presented work introduces a low-complexity hybrid method that combines wavelet decomposition and Fourier transforms to conduct wavelet spectral-energy analysis and extract a few features with high sensitivity to condition changes. The work is published as a regular paper in the *IEEE Transactions on Instrumentation and Measurement* under the title “A Hybrid Method for Condition Monitoring and Fault Diagnosis of Rolling Bearings With Low System Delay”.

5.1 Introduction

VBCM techniques are commonly used to detect and diagnose failures of rolling bearings. Accuracy and delay in detecting and diagnosing different types of failures are the main performance measures in condition monitoring. Achieving high accuracy with low delay improves system reliability and prevents catastrophic equipment failure. Further, delay is crucial to remote condition monitoring and time-sensitive industrial applications. While most of the proposed methods focus on accuracy, slight attention has been paid to addressing the delay introduced in the condition monitoring process. This chapter attempts to bridge this gap and propose a hybrid method for vibration-based condition monitoring and fault diag-

¹A version of this chapter has been published in [56].

nosis of rolling bearings that outperforms previous methods in terms of accuracy and delay. In the first place, the concept of system delay is introduced to assess the overall delay in vibration-based condition monitoring systems. Then, a hybrid method for low-complexity condition monitoring is proposed. It uses Wavelet Packet Transform (WPT) and Fourier analysis to decompose short-duration input segments of the vibration signal into elementary waveforms and obtain their spectral contents. Accordingly, energy concentration in the spectral components—caused by defect-induced transient vibrations—is utilized to extract a small number of features with high discriminative capabilities. Consequently, Bayesian optimization-based Random Forest (RF) algorithm is used to classify healthy and faulty operating conditions under varying motor speeds. The experimental results show that the proposed method can achieve high accuracy with low system delay.

The ongoing automation of industrial manufacturing “commonly referred to as Industry 4.0, smart manufacturing, or Industrial Internet of Things (IIoT)” has many potential benefits for industry and consumers. Managing maintenance is essential in smart manufacturing; effective and efficient maintenance sustains equipment availability and reliability, which, in turn, ensures safety, productivity, quality, and on-time delivery. Maintenance management strategies are generally categorized into three main categories: Corrective maintenance (CM), Preventive Maintenance (PvM), and Predictive Maintenance (PdM). In CM, corrective actions take place after equipment failure. Despite its simplicity, CM is very costly as it involves shutting down the production process and replacing parts. PvM is a time-based scheduled strategy where maintenance is regularly performed. It is an effective strategy as it prevents equipment failure. However, PvM involves unnecessary routine preventive actions, which increases maintenance costs. PdM is based on the continuous monitoring of the equipment’s condition, and actions for maintenance are predicted based on the equipment’s actual condition. This is achieved by adopting predictive approaches to monitor the equipment’s functional-process integrity continuously.

Data-driven-based approaches have been widely adopted in recent years for PdM applications. Generally, data-driven-based PdM can be summarized under two main steps: The first step involves feature extraction, where discriminative features are extracted from equipment-related data. The second step involves utilizing the extracted features to classify or predict normal and abnormal operational conditions accurately. Hence, it is essential to extract features that can represent the integrity of the operational process with high sensitivity to any changes within the process.

In terms of performance evaluation, accuracy, and complexity are commonly used to evaluate the performance of a PdM system. Higher accuracy improves system reliability and less complexity—in terms of memory and processing time—relaxes the computational requirements of the system. Besides accuracy and complexity, delay in predicting the current condition of the running equipment is another important aspect of system performance. Delay is crucial to time-sensitive industrial applications and remote condition monitoring in IIoT as it directly influences the end-to-end latency. Further, it is a critical factor for the early detection of failures. As a real-life example, in one case [327], the failure of a large gearbox caused a three-week shutdown, and extensive repair costs are typically €50,000 to €100,000. After implementing a condition-monitoring system, early detection of the gearbox failure resulted in a repair cost of €5000, saving the customer at least €27,000. Moreover, the company avoided lost production, amounting to around €6000/hour. Such real-life cases demonstrate the important role of early failure detection in condition monitoring systems as it helps to detect the failure in its early stages and prevents catastrophic equipment failure. Hence, it ensures workplace safety and productivity, and reduces maintenance costs. While most of the proposed work in the literature focuses on accuracy and complexity in assessing the performance of condition monitoring systems, little attention has been paid to addressing the overall delay in the condition monitoring process.

This chapter attempts to bridge this gap and introduces a base for overall delay analy-

sis in vibration-based condition monitoring systems. Specifically, the chapter defines and analyzes the overall delay in vibration-based condition monitoring and proposes a hybrid wavelet-based method for vibration-based condition monitoring and fault diagnosis of rolling bearings with low delay. The following are the main contributions of the chapter:

- Analyzes the overall delay in vibration-based condition monitoring and introduces the concept of system delay to assess it.
- Proposes a hybrid wavelet-based method with reduced system delay for vibration-based condition monitoring and fault diagnosis of rolling bearings under varying motor speeds. The proposed method has a high sensitivity to fault-related transients with relatively short durations of input vibration segments.
- The proposed method combines WPT and Fast Fourier Transform (FFT) to decompose the input vibration segment into elementary waveforms with high time-frequency localization and obtain spectral components of these waveforms to achieve high sensitivity with short durations of the vibration signal.
- Accordingly, the proposed method introduces a new technique to extract fault-sensitive features from spectral components of the elementary waveforms. Specifically, the proposed method utilizes high concentration in the spectral energy caused by defect-induced transient vibrations to select the most dominant frequency components (*i.e.*, frequencies with the highest power levels).
- The proposed method allows controlling the size of extracted features through the selection of the number of decomposition levels and the number of selected dominant frequencies, which helps to reduce redundancy in the extracted features. Moreover, this flexibility allows to design and adapt the proposed method according to various operational situations to meet specific application requirements in terms of accuracy and complexity.

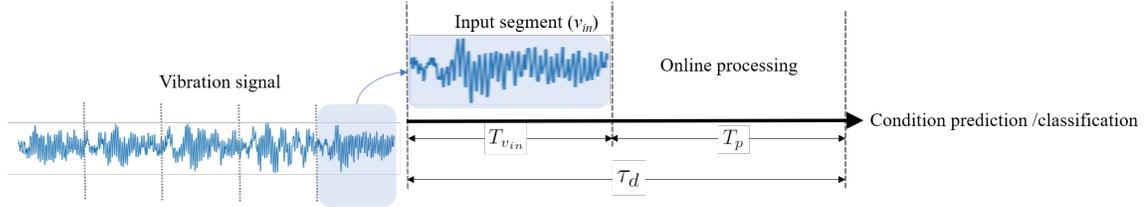


Figure 5.1: System delay of vibration-based condition monitoring.

The rest of the chapter is outlined as follows: The next section provides the theoretical foundation. Section 5.3 presents a literature review on the application of wavelets in condition monitoring and summarizes related work. Section 5.4 introduces the proposed method for vibration-based condition monitoring and fault diagnosis of rolling bearings. Section 5.5 presents the datasets and the experimental setup for performance evaluation, while Section 5.6 discusses the results. The chapter is finally concluded in Section 5.7.

5.2 Theoretical Background

5.2.1 System Delay

The system delay τ_d of a vibration-based condition monitoring system can be defined as the time the system takes to acquire input vibration segment and classify or predict the operational condition of the current state S_c of the equipment. In vibration-based monitoring, the current state S_c is represented by the input segment v_{in} of the generated vibration signal as illustrated in Fig.5.1. Accordingly, the system delay is the sum of the time duration of input segment $T_{v_{in}}$ and the online processing time T_p . The system delay τ_d can be formulated mathematically as follows;

$$\tau_d = T_{v_{in}} + T_p \quad (5.1)$$

The time duration $T_{v_{in}}$ of the input segment v_{in} depends on number of data points in the segment. It can be expressed as:

$$T_{v_{in}} = \frac{N_o}{f_s} \quad (\text{seconds}) \quad (5.2)$$

where N_o is number of data points in v_{in} and f_s is the sampling frequency in samples per second. Online processing time T_p is algorithm-dependent; it involves two tasks, feature extraction (including pre-processing) and condition prediction/classification. Hence, online processing time T_p can be generally viewed as a function of the number of data points N_o in the input segment v_{in} , the size of extracted features S , and available computing resources R_{comp} , *i.e.*,

$$T_p = f(N_o, S, R_{comp}) \quad (5.3)$$

Based on the above formulation, designing a condition-monitoring system with low system delay τ_d involves three main requirements:

1. Extracting features of high sensitivity to fault-related transients to improve system accuracy.
2. Extracting features of small size S .
3. Utilizing input vibration segments of relatively short time duration $T_{v_{in}}$ or equivalently, of small number of data points N_o .

Accordingly, considering fixed computing resources R_{comp} , a reliable design of vibration-based condition monitoring systems with low system delay would address the following two conditions:

$$\text{minimize}(N_o, S) \quad (5.4)$$

and

$$\text{maximize}(\text{Accuracy}). \quad (5.5)$$

For a given vibration-based condition monitoring system, its parameters can be tuned empirically to achieve a good trade-off between Eq. (5.4) and Eq. (5.5). Furthermore, the optimal parameters may be obtained by formulating a multi-objective optimization problem and using Eq. (5.4) and Eq. (5.5) as objective functions.

5.2.2 Wavelet Packet Transform

Wavelet Packet Transform (WPT) offers a reliable approach to engineering a few features from vibration signals with a high discriminative degree. Specifically, WPT is very useful in decomposing the input signal using scaled versions of a base wavelet function into elementary waveforms with high localization in time and frequency. Thus, WPT has a high time-frequency resolution, allowing it to capture of high-frequency transient components in the signal. Fig. 5.2.(a) shows the flowchart of a 3-level wavelet-packet decomposition process. The input signal is decomposed into lower and higher frequency sub-bands at each level. The output of the decomposition process in each level is either the approximation wavelet coefficients “A” associated with lower frequency bands or detail wavelet coefficients “D” associated with higher frequency bands. Accordingly, the elementary waveforms of the signal can be reconstructed using the individual wavelet coefficients. Hence, a highly discriminative feature can be extracted from each reconstructed waveform or its corresponding wavelet coefficients. Numerically, WPT is implemented through the iterative decomposition of the signal by a series of low-pass filters $h(k)$ and high-pass filters $g(k)$ as follows [276]:

$$w_{i+1}^{2s}(t) = \sum_k h(k)w_i^s(2t - k), \quad (5.6)$$

$$w_{i+1}^{2s+1}(t) = \sum_k g(k)w_i^{s+1}(2t - k), \quad (5.7)$$

where $k = 1, 2, \dots$ is decomposition level. The raw signal $x(t)$ is $w_0^1(k)$ and $w_i^s(k)$ is the wavelet decomposition coefficients of node s at level i . $w_{i+1}^{2s}(k)$ and $w_{i+1}^{2s+1}(k)$ are the wavelet coefficients of nodes $2s$ and $2s + 1$ at level $i + 1$, respectively, which correspond to the

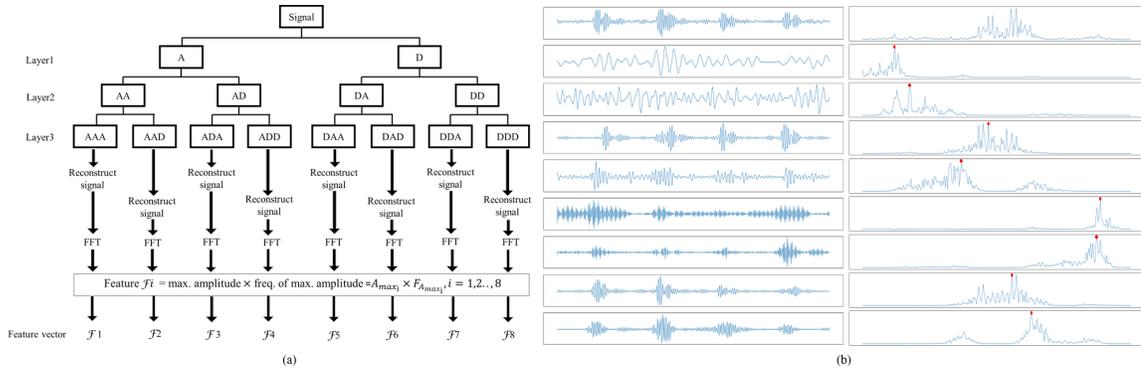


Figure 5.2: (a) Proposed feature extraction using 3-level WPT. (b) From top to bottom, vibration signal and its WPT-based elementary waveforms “on the left” and corresponding signal spectrum “on the right”.

approximation coefficients and detail coefficients. $h(k)$ and $g(k)$ are low pass and high pass filters which are related to the base wavelet function.

5.3 Literature Review And Related Work

Wavelet-based decomposition has been widely adopted in many applications as it offers a flexible tool for analyzing signals with a high time-frequency resolution. During the last two decades, wavelet-based approaches gained high popularity among researchers in the field of machinery diagnostics. Most of the proposed wavelet-based approaches in the literature are combined with other techniques to enhance the discriminative power of the extracted features. Accordingly, wavelet-based approaches can be generally grouped under four main categories: entropy-based [163], energy-based [276][275], spectral-based [328]-[329], and statistical based[330]. Spectral and statistical approaches are commonly applied to the elementary waveforms reconstructed from wavelet decomposition coefficients; entropy and energy approaches are either applied to the wavelet decomposition coefficients or to the reconstructed elementary waveforms. In [331], it was shown that using the reconstructed waveforms leads to a better fault diagnosis than wavelet packet coefficients, especially with a low Signal-to-Noise Ratio (SNR). Entropy is a measure of uncertainty;

it is used to measure the irregularity of time-series data. It is known that vibration signals of healthy bearings have high irregularity compared to faulty bearings [332]. Hence, combining wavelet decomposition and entropy measures helps quantify the degree of irregularity of a vibration signal with a high time-frequency localization. Energy-based approaches utilize the changes in energy contents of wavelet sub-bands to extract fault-related features. However, energy content has low sensitivity to incipient faults since the change in energy content is not significant at the early stages of the fault [333]. This low sensitivity to the incipient faults can be improved by integrating entropy measures with wavelet energy-based approaches [334]. This approach—commonly known as wavelet energy entropy—is based on applying entropy measures to wavelet energy contents. Spectral-based approaches use spectral characteristics of the reconstructed elementary waveforms to extract the features; common techniques to extract the fault-related spectral characteristics from wavelet-decomposed vibration signals involve Fourier transform [328], envelope spectrum [335][336], spectral kurtosis [337], and spectral entropy [329]. Statistical-based approaches rely on time-domain properties of reconstructed waveforms such as skewness, kurtosis, Root-Mean-Square (RMS), and crest factor for feature creation. Time-domain properties have low robustness to noise compared to the other approaches. Moreover, they generally have low sensitivity to defect-induced transient, which requires relatively long input segments of the vibration signal.

Recently, deep learning has been widely used in fault diagnosis and condition monitoring applications. The application of wavelet-based approaches within deep learning-based fault diagnosis falls mainly under two areas: 1) Transforming 1-D vibration signals into 2-D time-frequency input data [338][339]. 2) Utilizing the wavelet concept to design network elements such as activation function [340] and convolution kernel [122]. The main idea of deep learning-wavelet-based approaches is based on utilizing wavelets' powerful time-frequency localization characteristics to extract highly discriminative features and improve the learning process.

As mentioned earlier, most of the proposed vibration-based condition monitoring approaches focus on the accuracy of fault detection and system complexity, while less attention has been paid to addressing the delay of the system in terms of input signal length and size of features vector. In [279], Empirical Mode Decomposition (EMD) was used to decompose the vibration signals. Accordingly, an energy-based analysis is performed to select the most energized components. Further, a statistical analysis was conducted to remove redundant data. In [341], a combined compressive sampling feature ranking method is proposed to learn fewer features from vibration data optimally. In [259], kurtosis value and Fisher discrimination criterion were used to screen wavelet coefficients and extract fault-transient features to enhance the discriminative capability. In [163], an entropy-wavelet-based approach was proposed to extract the features from vibration signals, and the Laplacian score method was used to rank and select the features. In these approaches, attention is paid to reducing feature size only without considering the duration of the input signal or the processing time. In other words, in these approaches, the system is designed to reduce the dimensionality of the extracted features rather than extracting a small number of features. Although dimensionality reduction helps reduce memory requirements and improve the training process, it increases online processing time, increasing system delay. In [132], a deep learning-based method for condition monitoring is proposed, and an input length sensitivity analysis was performed. It was shown that 100% accuracy could be achieved with shorter input lengths compared to other approaches. However, the feature engineering process utilizes vibration signals from two accelerometer sensors and involves FFT, Continuous Wavelet Transform (CWT), and statistical proprieties of the raw signals to construct the features. Thus, resulting in a large-size feature vector. This chapter addresses the system delay from both aspects: the input signal's duration and the extracted features' size. Accordingly, a hybrid WPT-FFT method with low system delay is proposed for vibration-based condition monitoring and fault diagnosis of rolling bearings. The next section introduces the proposed method.

5.4 Proposed method for WPT-FFT Features Extraction

As mentioned earlier, the design of a condition-monitoring system with low delay has three main requirements:

1. Extracting features of high sensitivity to fault-related transients.
2. Extracting features of a small and controllable size.
3. Utilizing input vibration segments of short time durations.

In rolling bearings, the contact between the rolling elements and the defective spot results in repetitive impulse periods. These repetitive impulses appear as a dominant frequency in the spectrum of the vibration signal [342]. Fourier-based analysis such as FFT provides an effective way to transform the vibration signal to the frequency domain and reveal its spectral contents. However, these contents have no time resolution, and hence, Fourier analysis provides a poor representation of signals that are highly localized in time. On the other hand, WPT provides a very efficient time-frequency analysis of vibration signals with a good capture of fault-related transients [343]. Based on this concept, a hybrid WPT-FFT method is proposed to fulfill the aforementioned requirements and extract a small number of highly discriminative features from short-duration vibration signals. The first step involves decomposing the input vibration segment using k -level WPT. Accordingly, 2^k elementary waveforms of lower and higher frequency sub-bands are reconstructed from individual wavelet coefficients. In the second step, FFT is applied to the resultant waveforms to obtain the spectral contents of each waveform. Hence, a feature vector of size $S = 1 \times (m \times 2^k)$ can be reconstructed by utilizing the first m dominant frequency components in the spectrum of each waveform. Fig.5.2 illustrates the proposed method with $k = 3$ and $m = 1$. Specifically, Fig. 5.2.(a) shows the flowchart of obtaining approximation and detail coefficients of a given vibration signal using the 3-level wavelet-packet decomposition process. Accordingly, Fig. 5.2.(b) shows the vibration signal (the topmost signal) and

the reconstructed elementary waveforms, along with their spectral contents obtained via FFT. The first dominant frequency in each spectrum is shown as a red dot. After selecting the first dominant frequency, the feature of each waveform is calculated according to the below formula:

$$\mathcal{F}_i = A_{max_i} \times F_{A_{max_i}}, \quad (5.8)$$

where, $i = 1, 2, \dots, 2^k$, A_{max_i} , and $F_{A_{max_i}}$ are the spectral maximum amplitude of the i th waveform and its corresponding frequency, respectively. For $m > 1$, the first m dominant spectral amplitudes of each waveform and their corresponding frequencies are selected, and the feature of each waveform is calculated according to the below formula:

$$\mathcal{F}_i = A1_{max_i} \times F_{A1_{max_i}}, \dots, Am_{max_i} \times F_{Am_{max_i}}. \quad (5.9)$$

Algorithm 2 shows the pseudo-code of the proposed method. By decomposing the vi-

Algorithm 1 Proposed WPT-FFT Features Extraction

Input: x : input segment of vibration signal of length N_o data points

Parameters: k = WPT decomposition level, m = number of most dominant frequency components to be selected.

Output: $\mathcal{F}[s]$, $s = 1, \dots, S$: features vector of size $S = 1 \times (m \times 2^k)$

Start:

Set k and m

$s = 1$

Obtain 2^k elementary waveforms $w[i]$, $i = 1, \dots, 2^k$ of x using k -level WPT.

for $i = 1, \dots, 2^k$ **do**

 Compute FFT of $w[i]$ and obtain:

A_{FFT} : vector of FFT spectrum amplitudes

$F_{A_{FFT}}$: vector of FFT spectrum frequencies

for $ii = 1, \dots, m$ **do**

$A_{max} = \max[A_{FFT}]$: max. spectrum's amplitude

$F_{A_{max}} = F_{A_{FFT}}[\text{index}(A_{max})]$: Freq. of max. spectrum's amplitude

 Compute feature $f = A_{max} \times F_{A_{max}}$

$\mathcal{F}[s] = f$

$s = s + 1$

 Remove A_{max} from A_{FFT}

 Remove $F_{A_{max}}$ from $F_{A_{FFT}}$

End

bration signal into elementary waveforms and obtaining the dominant frequencies of each waveform, features with high sensitivity to fault-related transients can be calculated from these frequencies according to Eq. (5.8) or Eq. (5.9). This has the effect of applying frequency-selective filters to the spectrum of each waveform so that only dominant frequencies are selected and then scaled by their amplitude values to increase their discriminative capabilities. This way, the remaining ineffectual spectral contents are filtered out, the redundancy is reduced, and the sensitivity to faults is improved. Moreover, since the size S of the features vector equals $1 \times (m \times 2^k)$, the proposed method allows controlling the size of the extracted features by choosing the number of WPT decomposition levels k and the number of selected dominant frequency components m . This, in turn, provides high flexibility and scalability in the proposed method when addressing various operational situations. Moreover, it helps reduce redundancy and meet specific application requirements in terms of accuracy and complexity through proper tuning of (k, m) parameters.

5.4.1 Selection of Base Wavelet and Decomposition Level

For vibration-based condition monitoring, the proper base wavelet is the wavelet that is highly correlated with defect-induced transient vibrations. This leads to high energy concentrated at the corresponding wavelet coefficients only. Due to this high energy concentration, the entropy of the energy distribution of the wavelet coefficients will be minimized. The same concept can be used to determine the appropriate decomposition level k since the proper decomposition level will have the highest energy concentration and minimum entropy. Accordingly, the proper base wavelet and decomposition level can be selected by evaluating the energy and entropy values of the wavelet coefficients for different base wavelets and decomposition levels. The energy of wavelet coefficients at decomposition level i can be expressed as:

$$E_i = \sum_{k=1}^L |w_i(k)|^2, \quad (5.10)$$

where $w_i(k)$ are the wavelet coefficients at level i and L is the total number of the coefficients in that level. The entropy of the energy distribution of the wavelet coefficients can be obtained by the Shannon entropy formula:

$$Entropy_i = \sum_{k=1}^L p_i \log_2 p_i, \quad (5.11)$$

$$p_i = \frac{|w_i(k)|^2}{E_i} \quad (5.12)$$

where p_i is the energy probability distribution of the wavelet coefficients. The energy-to-entropy ratio proposed in [334] relates energy and entropy values of the wavelet coefficients and provides a useful quantitative measure to select the appropriate base wavelet and decomposition level. The energy-to-entropy ratio is expressed as [334]:

$$R(s) = \frac{E_i}{Entropy_i} \quad (5.13)$$

Thus, the higher $R(s)$, the more appropriate base wavelet and/or decomposition level to select.

5.4.2 Complexity Analysis

Considering fixed computational resources, the system delay τ_d of vibration-based condition monitoring depends entirely on three factors:

- duration of the input segment $T_{v_{in}}$,
- algorithm-based computations,
- and size of features vector S .

For the proposed algorithm, these factors include: the number of data points N_o in the input segment v_{in} , decomposition level k , WPT transform, reconstruction of 2^k elementary wave-

forms, 2^k FFT transforms, and number of selected dominant frequencies m . Accordingly, the complexity of the proposed algorithm can be analyzed as follows:

- Complexity of WPT is $O(N_o \log N_o)$.
- Let c represent the complexity of reconstructing one elementary waveform, then the complexity of reconstructing 2^k elementary waveforms is $O(2^k c)$.
- Complexity of FFT computations is $O(2^k N_o \log N_o)$.
- Complexity of feature vector computations is $O(m 2^k)$.

Thus, the system delay and complexity grow as a function of $(N_o, 2^k, m)$.

5.5 Performance Evaluation

Performance of the proposed method is evaluated on the Case Western Reserve University (CWRU) bearing dataset [316], the Paderborn University (PU) bearing dataset [344], and the University of Ottawa (uOttawa) bearing dataset [345]. These datasets are selected to simulate various practical situations regarding defect types, rotational speed conditions, and data sampling rates. Specifically, the CWRU dataset is very useful in benchmarking as it is widely used in the literature. However, it doesn't include combined defects and lacks real damages since faults were artificially generated in the bearings. Vibration signals of CWRU have a sampling rate of 12 KHz. In contrast to CWRU, the PU dataset has real bearing damages with combined defects. Vibration signals were sampled at 64 KHz. uOttawa dataset was released in 2019; experiments were conducted using a machinery fault simulator. The main aspect of this dataset is that it has different healthy and faulty conditions with combined defects and under time-varying rotational speed conditions. Thus, in contrast to CWRU and PU datasets, this dataset has time-varying rotational speeds within the same measurement. Further, vibration signals have a high sampling rate of 200 KHz.

5.5.1 Experimental Setup

CWRU Bearing Dataset

In the CWRU dataset, experiments are conducted using a 2 hp (horse power) Reliance Electric motor and vibration data was collected using accelerometers. Faults ranging from 0.007 to 0.040 inches in diameter were introduced separately at the Inner Raceway (IR), ball, and Outer Raceway (OR). Faulted bearings were reinstalled into the test motor, and vibration data was recorded at motor loads of 0 to 3 horsepower (motor speeds of 1,720 to 1,797 rpm). Digital data was collected at 12,000 samples per second. The dataset used in this chapter consists of vibration signals generated by the accelerometers placed at the drive end of the motor housing. Table 5.1 shows operational condition, fault diameter, and motor speed (rpm) of these vibration signals. According to operational conditions and fault diameter, faulty operational conditions are classified into nine classes. Hence, the dataset consists of ten classes, one healthy operational class, and nine faulty operational classes. Vibration signals are divided into input segments of 300, 600, 1,200, and 2,400 data points to evaluate the proposed method at different time durations of input segments. With a sampling rate of 12,000 samples per second, this gives input segments with T_{vin} values of 0.025 seconds, 0.5 seconds, 0.1 seconds, and 0.2 seconds, respectively. For each time duration, values of $k = 2, 3, 5$ and $m = 1, 2, 3$ are used to decompose the input segment and construct feature vectors. The aim here is to examine the impact of systems parameters (k, m) on the performance and to assess the proposed method at different sizes of features vector.

PU Bearing Dataset

In the PU dataset, experiments are conducted using a 425 W Permanent Magnet Synchronous Motor (PMSM). The dataset used in this chapter is based on measurements conducted at $n = 1,500$ rpm with a load torque of $M = 0.7$ Nm and a radial force on the bearing

of $F = 1,000$ N. Vibration signals were recorded with a sampling rate of 64,000 samples per second by measuring the acceleration of the bearing housing at the adapter at the top end of the rolling bearing module. Regarding bearing defects, the PU dataset includes artificially generated and accelerated-lifetime defects. In this chapter, only accelerated-lifetime defects are used. Accordingly, the dataset has four classes: one healthy class and three faulty classes according to fault type, as shown in Table 5.1. To evaluate the proposed method at different time durations of input segments, each vibration signal is divided into input segments of 1,600 and 12,800 data points. With a sampling rate of 64,000 samples per second, this gives input segments with T_{v_m} values of 0.025 seconds and 0.2 seconds, respectively. For each time duration, values of $k = 3, 5$ and $m = 3, 5, 7$ are used to decompose the input segment and construct feature vectors.

University of Ottawa Bearing Dataset

In the uOttawa dataset, experiments are performed on a SpectraQuest machinery fault simulator (MFS-PK5M). The accelerometer was placed on the housing of the experimental bearing to collect the vibration data with a sampling rate of 200,000 samples per second. The measurements have two experimental settings: bearing health condition and varying speed condition. The health conditions of the bearing include healthy, faulty with an IR defect, faulty with an OR defect, faulty with a ball defect, and faulty with combined defects on the IR, the OR, and the ball. The operating rotational speed conditions are increasing speed, decreasing speed, increasing then decreasing speed, and decreasing then increasing speed. Hence, there are twenty different settings in the measurements. Accordingly, the dataset used in this chapter is arranged under one healthy class and four faulty classes so that each class includes all the five varying speed conditions as shown in Table 5.1. To evaluate the proposed method at different time durations of input segments, each vibration signal is divided into input segments of 5,000 and 40,000 data points. With a sampling rate of 200,000 samples per second, this gives input segments with T_{v_m} values of 0.025 seconds

and 0.2 seconds, respectively. For each time duration, values of $k = 5, 7$ and $m = 5, 7, 9$ are used to decompose the input segment and construct feature vectors.

The *Daubechies 4 (dp4)* base wavelet is selected to decompose input vibration segments and reconstruct the elementary waveforms. After the feature extraction stage, the resultant datasets of features are divided into 80% training samples and 20% test samples. Accordingly, classifiers are trained on these datasets, and Bayesian Optimization with Gaussian Processes [118] is used to tune the hyperparameters. Finally, test samples are used to evaluate the performance of the proposed method. In order to select a suitable classifier, the performances of three classifiers are compared on the CWRU dataset, and the classifier of best performance is selected to evaluate the performance on the three datasets. The three classifiers are Support-Vector Machines (SVM), eXtreme Gradient Boosting (XGBoost), and Random Forest (RF). Python programming language, PyWavelets, and SciPy libraries are used to build the models¹. Results and related discussion are presented in the next section.

5.6 Results and Discussion

5.6.1 Classifier Selection

As mentioned earlier, the performances of SVM, XGBoost, and RF classifiers are compared on the CWRU dataset, and the classifier of best performance is selected to evaluate the performance of the proposed method on the three datasets. SVM creates a hyperplane that splits the input feature space. For inputs belonging to N classes, their feature space is an N -dimensional space representing these inputs and their associated classes. Hence, SVM attempts to create a hyperplane that achieves the best separation between input features according to their classes. XGBoost and RF are decision trees-based classifiers. They

¹Code is available at: <https://github.com/Western-OC2-Lab/Vibration-Based-Fault-Diagnosis-with-Low-Delay>

Table 5.1: Experimental Datasets

	Class	Health condition	Fault diameter	Motor speed (rpm)			
	CWRU dataset	1	Healthy	NA	1730	1750	1772
2		IR faults	0.07"	1730	1750	1772	1797
3			0.014"	1730	1750	1772	1797
4			0.021"	1730	1750	1772	1797
5		Ball faults	0.07"	1730	1750	1772	1797
6			0.014"	1730	1750	1772	1797
7			0.021"	1730	1750	1772	1797
8		OR faults	0.07"	1730	1750	1772	1797
9			0.014"	1730	1750	1772	1797
10			0.021"	1730	1750	1772	1797
PU dataset	Class	Health condition	Fault type	Motor speed (rpm)			
	1	Healthy	NA	1500			
	2	Combined IR and OR faults	Multiple damages				
	3	IR faults	Single, repetitive, and multiple damages				
	4	OR faults	Single and repetitive damages				
uOttawa dataset	Class	Health condition		Speed conditions			
	1	Healthy		Increasing speed, decreasing speed, increasing then decreasing speed, and decreasing then increasing speed			
	2	IR faults					
	3	OR faults					
	4	Ball faults					
	5	Combined IR, OR, and ball faults					

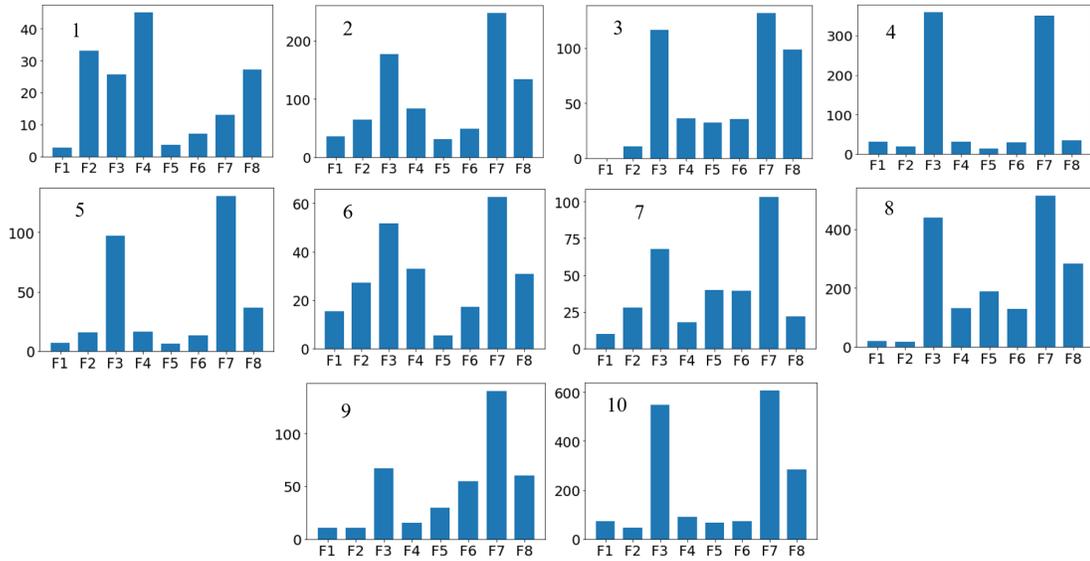


Figure 5.3: Feature sample from each class of CWRU dataset with $T_{vin} = 0.05$ seconds, $k = 3$ and $m = 1$.

Table 5.2: Performance comparison of SVM, XGBoost, and RF classifiers on CWRU dataset: Input segment duration = 0.05 seconds (600 samples), ($k = 3, m = 1$)

Performance Metric	Support Vector Machines (SVM)	eXtreme Gradient Boosting (XGBoost)	Random Forest (RF)
ROC-AUC Score	1.000	1.000	1.000
Accuracy (%)	98.013	98.344	99.172

Table 5.3: Performance comparison between proposed features and other features on CWRU dataset: Input segment duration = 0.05 seconds (600 samples)

Performance Metric	Mean	Crest Factor	Kurtosis	Shannon Entropy	Proposed ($k = 3, m = 1$)
ROC-AUC Score	0.999	0.937	0.963	1.000	1.000
Accuracy (%)	96.689	64.073	74.834	98.510	99.172

both utilize ensemble learning techniques and attempt to predict the target class by combining the estimates from individual decision trees. However, XGBoost builds a single decision tree at a time; each new tree predicts the residuals or errors of the previously trained decision tree. In contrast to XGBoost, RF fits several decision tree classifiers on various subsets of training datasets using different subsets of features. Accordingly, the decisions of individual classifiers are aggregated to obtain a final decision and predict the target class.

Table. 5.2 compares performances of the three classifiers on the CWRU dataset in terms

Table 5.4: Performance results of the proposed method on CWRU dataset

Input Segment		Performance Metric	$k = 2$			$k = 3$			$k = 5$			
N_o	$T_{v_{in}}$ (sec)		$m = 1$	$m = 2$	$m = 3$	$m = 1$	$m = 2$	$m = 3$	$m = 1$	$m = 2$	$m = 3$	
		Size of feature vector S	1×4	1×8	1×12	1×8	1×16	1×24	1×32	1×64	1×96	
300	0.025	AUC Score	0.991	0.996	0.996	0.999	1.000	1.000	1.000	1.000	1.000	
		Accuracy (%)	89.23%	93.10%	93.05%	97.51%	98.18%	98.32%	99.63%	99.11%	99.14%	
		Online Processing Time T_p (sec)*	0.0130	0.0140	0.0140	0.0140	0.0150	0.0150	0.0150	0.0170	0.0260	0.0280
		Peak Memory (MB)*	0.1826	0.1851	0.1875	0.1889	0.1942	0.1949	0.2563	0.2598	0.2672	
		System delay $\tau_d = T_{v_{in}} + T_p$ (sec)	0.0380	0.0390	0.0390	0.0390	0.0400	0.0400	0.0420	0.0510	0.0530	
600	0.05	AUC Score	0.997	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
		Accuracy (%)	93.54%	96.00%	96.20%	99.17%	99.61%	99.56%	99.66%	99.80%	99.85%	
		Online Processing Time T_p (sec)*	0.0140	0.0140	0.0140	0.0150	0.0150	0.0160	0.0270	0.0289	0.0309	
		Peak Memory (MB)*	0.1897	0.1914	0.1923	0.1956	0.2019	0.2039	0.2573	0.2574	0.2687	
		System delay $\tau_d = T_{v_{in}} + T_p$ (sec)	0.0640	0.0640	0.0640	0.0650	0.0650	0.0660	0.0770	0.0789	0.0809	
1,200	0.1	AUC Score	0.999	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
		Accuracy (%)	96.24%	99.11%	98.42%	100.00%	100.00%	99.80%	99.80%	100.00%	99.31%	
		Online Processing Time T_p (sec)*	0.0150	0.0150	0.0150	0.0150	0.0160	0.0160	0.02791	0.0299	0.0319	
		Peak Memory (MB)*	0.2148	0.2158	0.2156	0.2349	0.2353	0.2361	0.2985	0.3063	0.3098	
		System delay $\tau_d = T_{v_{in}} + T_p$ (sec)	0.1150	0.1150	0.1150	0.1150	0.1160	0.1160	0.1279	0.1299	0.1319	
2,400	0.2	AUC Score	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
		Accuracy (%)	95.23%	99.01%	98.41%	100.00%	100.00%	100.00%	99.80%	99.83%	99.80%	
		Online Processing Time T_p (sec)*	0.0150	0.0150	0.0150	0.0155	0.0160	0.0170	0.0340	0.0349	0.0350	
		Peak Memory (MB)*	0.2796	0.2800	0.2803	0.3061	0.3066	0.3087	0.4062	0.4065	0.4089	
		System delay $\tau_d = T_{v_{in}} + T_p$ (sec)	0.2150	0.2150	0.2150	0.2155	0.2160	0.2170	0.2340	0.2349	0.2350	

*Based on a machine with i7-8550u CPU 1.8GHz and 8 GB RAM

Table 5.5: Mean Accuracy (%) on CWRU dataset

Parameters Settings	$k = 2$			$k = 3$			$k = 5$		
	$m = 1$	$m = 2$	$m = 3$	$m = 1$	$m = 2$	$m = 3$	$m = 1$	$m = 2$	$m = 3$
Size of feature vector S	1×4	1×8	1×12	1×8	1×16	1×24	1×32	1×64	1×96
Mean AUC Score	0.996	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Mean Accuracy (%)	93.56%	96.80%	96.52%	99.17%	99.45%	99.42%	99.72%	99.69%	99.53%

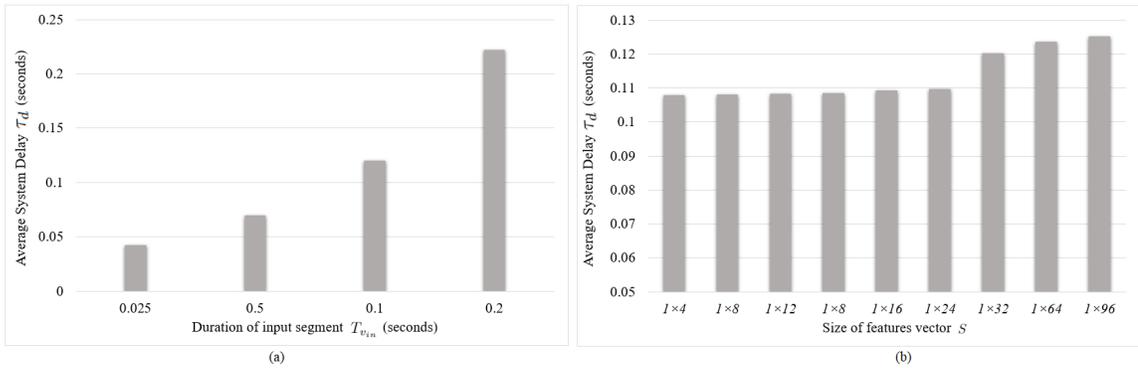


Figure 5.4: Average system delay τ_d as a function of (a) duration of input segment $T_{v_{in}}$ and (b) size of features vector S .

Table 5.6: Performance comparison between the proposed method and other methods on CWRU dataset

Reference	Approach	No. of data points in input vibration segment	Size of features vector	Achieved accuracy (%)
Proposed	WPD-FFT with $k=5, m=1$	300	1×32	99.63%
Proposed	WPD-FFT with $k=3, m=1$	600	1×8	99.17%
Proposed	WPD-FFT with $k=3, m=1$	1200	1×8	100.00%
Proposed	WPD-FFT with $k=3, m=1$	2400	1×8	100.00%
[132]	Convolutional Long Short-Term Memory (CLSTM)	500	$1 \times (((D(K+3)) \times L) + 2D)$, D = no. of accelerometer sensors, L = length of input segment, K = first K components of the raw signal acquired by CWT	100.00%
[122]	Convolutional Neural Network (CNN) with wavelet convolution and deep transfer learning	2048	1×2048	99.73%
[346]	Recurrent Neural Network (RNN) + CNN	4096	1×4096	100.00%
[133]	CNN with Scaled Exponential Linear Unit (SELU) and hierarchical regularization	512	64×64	100.00%
[341]	Compressive sampling and feature ranking with ANN classifier	2000	1×50	100.00%

Table 5.7: Performance results of the proposed method on the PU dataset

Input Segment		Metric	$k = 3$			$k = 5$		
N_o	T_{vm} (sec)		$m = 3$	$m = 5$	$m = 7$	$m = 3$	$m = 5$	$m = 7$
		Size of feature vector S	1×24	1×40	1×56	1×96	1×160	1×224
1,600	0.025	ROC-AUC Score	0.979	0.979	0.977	0.994	0.993	0.992
		Accuracy (%)	86.55%	86.53%	85.99%	91.91%	91.32%	90.29%
12,800	0.2	ROC-AUC Score	0.998	0.998	0.999	1.000	1.000	1.000
		Accuracy (%)	96.88%	96.44%	97.19%	99.56%	100.00%	99.69%

Table 5.8: Performance comparison between the proposed method and other methods on PU dataset

Reference	Approach	Achieved accuracy (%)
Proposed	WPD-FFT with $k = 5$, $m = 5$, length of input segments = 12800 data points	100.00%
[347]	Attention-based Multi-Dimensional Concatenated CNN (AMDC-CNN)	99.80%
[348]	Deep Neural Network (DNN) with Net2Net transformation and domain adaptation	96.24%
[349]	Input Feature Mappings (IFMs)-based Deep Residual Network (ResNet)	99.70%

Table 5.9: Performance results of the proposed method on the uOttawa dataset

Input Segment		Metric	$k = 5$			$k = 7$		
N_o	T_{v_m} (sec)		$m = 5$	$m = 7$	$m = 9$	$m = 5$	$m = 7$	$m = 9$
		Size of feature vector S	1×160	1×224	1×288	1×640	1×896	1×1152
5,000	0.025	ROC-AUC Score	0.996	0.995	0.994	0.998	0.997	0.997
		Accuracy (%)	94.31%	93.94%	93.54%	96.40%	95.81%	95.50%
40,000	0.2	ROC-AUC Score	0.997	0.997	0.997	1.000	0.999	0.999
		Accuracy (%)	96.83%	95.67%	94.67%	98.83%	98.17%	97.83%

of ROC-AUC score and accuracy. Vibration input segments have 600 data points which corresponds to 0.05 seconds. While the three classifiers achieved high and comparable performance levels, tree-based classifiers (XGBoost and RF) have slightly better performance than SVM. However, the RF classifier has a slightly better performance compared to XGBoost. The main advantage of RF over XGBoost is that RF fits several decision trees and trains each tree independently using a random subset of the data. This increases randomness and reduces bias in the training phase, allowing for better generalization. Hence, RF is more robust to overfitting compared to single decision tree-based algorithms such as XGBoost. Accordingly, RF will be used from now onwards to evaluate the performance of the proposed method on the three datasets.

Table 5.10: Performance comparison between the proposed method and other methods on uOttawa dataset

Reference	Approach	Achieved accuracy (%)
Proposed	WPD-FFT with $k = 7$, $m = 5$, length of input segments = 40000 data points	98.83%
[350]	Generalized Broadband Mode Decomposition (GBMD) with Distance Evaluation Technology (DET) for feature screening	96.67%
[351]	Convolutional denoising auto-encoder with Convolutional Long Short-Term Memory (CLSTM)	97.68%
[352]	Long short-term memory (LSTM)	77.00%

5.6.2 Effectiveness of Proposed Features

As mentioned earlier, the proposed method utilizes high energy concentration caused by defect-induced transient vibrations to extract the features according to Eq. (5.8) or Eq. (5.9). For each elementary waveform i , the features are extracted by firstly selecting the first m dominant frequencies from its corresponding vectors of FFT spectrum amplitudes A_{FFT_i} and FFT spectrum frequencies $F_{A_{FFT_i}}$ where $i = 1, \dots, 2^k$. Then, each frequency is multiplied by its amplitude value to increase the discriminative capabilities of the extracted features. Fig.5.3 shows a sample from the features vector for each class of the CWRU dataset with $T_{vin} = 0.05$ seconds (corresponds to 600 data points), $k = 3$, and $m = 1$. As shown, the features have distinct patterns, which reflect the high discriminative degree of the extracted features. Performance comparison on the CWRU dataset is conducted between the proposed features with $k = 3$ and $m = 1$ and other features extracted from vectors of FFT spectrum amplitudes A_{FFT_i} to demonstrate the effectiveness of the proposed features. These features include mean, crest factor, kurtosis, and Shannon entropy. Performance comparison is shown in Table 5.3. Results demonstrate the effectiveness of the proposed features as they achieved the best performance among other features.

5.6.3 Experimental Results on CWRU Dataset

Table 5.4 shows performance results of the proposed method on the CWRU dataset where accuracy, ROC-AUC score, system delay, and complexity—in terms of online processing time and memory requirements—are used to evaluate the performance. In terms of accuracy and ROC-AUC score, the proposed method achieved very high performance (accuracy > 99%, ROC-AUC score = 1.00) with all durations of the input vibration segment. Moreover, it reached 100% accuracy using an input duration of 0.1 seconds (1,200 data points) with 8 features only. These results reflect the high sensitivity of the proposed method to fault-related transients. In terms of the duration of the input segment, it significantly impacts the accuracy. For fixed parameters setting (k, m) , the higher the duration, the higher the accuracy and the higher the delay. However, proper setting of the system parameters (k, m) improves the system sensitivity and achieves a good trade-off between the input duration and the system accuracy. For instance, as shown in the table, with an input duration of 0.025 seconds (300 data points), setting $(k = 5, m = 1)$ brought more than 11% improvement in the accuracy compared to $(k = 2, m = 1)$. On the other hand, with 0.2 seconds (2,400 data points) of input duration, increasing the k value from 3 to 5 has slightly affected the system accuracy. Similarly, with 0.025 seconds (300 data points) of input duration, increasing the m value from 1 to 3 with $k = 3$ slightly affected the accuracy. This slight degradation in the accuracy can be explained by increased redundancy in the extracted features. Extracting more features from vibration signals could either increase redundancy or improve sensitivity to faults in the extracted features. This generally depends on the length of the input vibration segment and fault severity. The results demonstrate that the proposed method can address this issue with high flexibility through k and m parameters.

Regarding the influence of system parameters on system accuracy, Table 5.5 shows mean accuracy and mean ROC-AUC score as a function of system parameters (k, m) . As results generally indicate, the decomposing level parameter k has a considerable impact on system

accuracy compared to the number of selected dominant frequencies m . More specifically, increasing the value of k with fixed m , brought more improvement compared to increasing the value of m with fixed k .

Regarding system delay τ_d , the proposed method achieved excellent performance (accuracy $> 99\%$, ROC-AUC score = 1.00) with a minimum system delay τ_d of 0.042 seconds. As mentioned earlier, with fixed computational resources, the system delay τ_d is a function of input segment duration $T_{v_{in}}$ and size of features vector S . Fig. 5.4 shows empirical sensitivity analysis of average system delay as a function of (a) duration of input segment $T_{v_{in}}$ and (b) size of features vector S . Here, the average system delay at each value of $T_{v_{in}}/S$ is calculated by averaging individual system delays at corresponding values of $S/(T_{v_{in}})$. The results reveal that τ_d is heavily influenced by $T_{v_{in}}$ compared to S . For instance, increasing the input duration from $T_{v_{in}} = 0.025$ seconds to $T_{v_{in}} = 0.2$ seconds (700% increase in input segment duration), increased the average system delay by 424.5%. On the other hand, increasing the size of features vector from $S = 1 \times 4$ to $S = 1 \times 96$ (2300% increase in features size), increased the average system delay by 15.9% only.

According to the analysis in the Theoretical Background Section, for a given condition monitoring system and considering fixed computational resources, the system delay is influenced by the time duration of the input vibration signal (or equivalently, the number of datapoint samples in the input vibration signal) and the size of the extracted features. Hence, as mentioned earlier, a reliable design of a vibration-based condition monitoring system with low system delay would minimize the number of data points in the input vibration signal, minimize the size of extracted features, and maximize accuracy. Thus, to demonstrate the efficiency of the proposed method in this regard, a comparison between the proposed method and other methods is conducted in terms of the number of data points in the input vibration segment, size of the features vector, and achieved accuracy. Table 5.6 shows the comparison results. The length of the input segment and the size of the

feature vector are used to assess the system delay since the chapter establishes the direct relationship between the system delay and the length of the input vibration segment and the resultant feature size, as formulated in the Theoretical Background Section. As shown in the table, the proposed method reached 100% accuracy using an input segment of 1,200 data points with eight features only. On the other hand, other methods require either longer segments of the input signal [346] or use more features [341][132][346][133] compared to our proposed method to reach 100% accuracy. Moreover, the proposed method does not involve extra computations for feature ranking or dimensionality reduction compared to other methods.

5.6.4 Experimental Results on PU Dataset

The performance of the proposed method is further validated on the PU dataset, which includes real accelerated-lifetime damages. Performance results are shown in Table 5.7. As shown, the proposed method achieved an excellent accuracy of 99.56% by extracting 96 ($k = 5, k = 3$) features from input vibration segments of 0.1 seconds duration (12,800 data points). Furthermore, it reached 100% accuracy and a 1.00 ROC-AUC score by extracting 160 features ($k = 5, m = 5$) only. On the other hand, extracting more features (224 features using $k = 5$ and $m = 7$) led to a slight degradation in the accuracy. These results demonstrate the flexibility of the proposed method in tuning the size of extracted features which, in turn, lessens feature redundancy, improves the accuracy, and reduces the complexity. Table 5.8 provides a comparison between the proposed method and some recent works [347]-[349] on the PU dataset of healthy bearings and accelerated-lifetime damages. Results show the superiority of the proposed method in terms of achieved accuracy. Regarding the used approach, the proposed method relies completely on signal processing (WPT and FFT) to extract the features, while other methods utilize Deep neural networks for feature extraction. Hence, the proposed method involves less training and tuning complexity than the other methods.

5.6.5 Experimental Results on uOttawa Dataset

To further validate the effectiveness of the proposed method under varying motor speeds, its performance is evaluated on the uOttawa dataset. Table 5.9 shows performance results where the proposed method achieved excellent performance of 98.83% accuracy and 0.999 ROC-AUC score with input vibration durations of 0.1 seconds (40,000 data points) and 640 features ($k = 7, m = 5$). Table 5.10 shows a comparison between the proposed method and some of the recent works [350]-[352] on the uOttawa dataset. The proposed method achieved the best accuracy among other methods, as shown in the table. These results, along with the previous results obtained on CWRU and PU datasets, demonstrate the effectiveness of the proposed method and affirm its capability of achieving performance requirements under actual operational conditions. Furthermore, the results confirm the flexibility of the proposed method in adapting the size of extracted features according to specific application requirements.

5.7 Conclusion

The work presented in this chapter defined and analyzed the end-to-end delay of vibration-based condition monitoring systems and introduced the concept of system delay to assess it. With fixed computational resources, the system delay depends entirely on the duration of the input vibration segment, computation steps of the used algorithm/s, and the size of the features vector. Accordingly, a low-system delay method is introduced for vibration-based condition monitoring and fault diagnosis of rolling bearings. The proposed method uses a hybrid WPT-FFT approach where it decomposes small durations of the input vibration signal using k -level WPT into 2^k elementary waveforms with high time-frequency localization. Then it obtains the spectral components of these waveforms using FFT. Accordingly, the proposed method utilizes the first m dominant frequency components in the spectrum of each waveform to construct a small number of features with high sensitivity to fault-related

transients. Hence, the proposed method has a high sensitivity to fault-related transients with relatively short durations of input vibration segments. Moreover, it allows controlling the size of extracted features through (k, m) settings, which helps to reduce redundancy and provides high flexibility to customize the proposed method according to specific application requirements. The proposed method has been evaluated on three datasets using different durations of input vibration segments and with different parameter settings. The experimental results show that the proposed method can achieve excellent accuracy in fault diagnosis with low system delay even with combined defects and under varying motor speeds.

Chapter 6

6 Joint Instantaneous Amplitude-Frequency Analysis of Vibration Signals for Vibration-Based Condition Monitoring

Addressing the third research problem, this chapter¹ introduces the second contribution related to the computational complexity of Vibration-Based Condition Monitoring (VBCM). The chapter introduces a low-complexity method for VBCM of rolling bearings based on envelope analysis of the generated vibration signal. The proposed method employs the Hilbert Transform (HT) to obtain and then jointly analyze the instantaneous amplitude “envelope” and instantaneous frequency of vibration signals to facilitate three novel envelope representations. Accordingly, new highly discriminative features are extracted from these representations. The content of this chapter is submitted to the *IEEE Transactions on Instrumentation and Measurement* as a regular paper entitled “Joint Instantaneous Amplitude-Frequency Analysis of Vibration Signals for Vibration-Based Condition Monitoring of Rolling Bearings”. The paper is currently in the review stage.

6.1 Introduction

Vibrations of damaged bearings are manifested as modulations in the amplitude of the generated vibration signal, making envelope analysis an effective approach for discriminating between healthy and abnormal vibration patterns. Motivated by this, the work presented in this chapter introduces a low-complexity method for VBCM of rolling bearings based on envelope analysis. In the proposed method, the instantaneous amplitude “envelope” and

¹A version of this chapter has been submitted for publication in *IEEE Transactions on Instrumentation and Measurement*.

instantaneous frequency of the vibration signal are jointly utilized to facilitate three novel envelope representations: Instantaneous Amplitude-Frequency Mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and Instantaneous Energy-Frequency Distribution (IEFD). Maintaining temporal information, these representations effectively capture energy-frequency variations that are unique to the condition of the bearing, thereby enabling the extraction of discriminative features with high sensitivity to variations in operational conditions. Accordingly, a set of six new highly discriminative features are engineered from these representations, capturing and characterizing their shapes. The experimental results show outstanding performance in detecting and diagnosing various fault types, demonstrating the effectiveness of the proposed method in capturing unique variations in energy and frequency between healthy and faulty bearings. Moreover, the proposed method has moderate computational complexity, meeting the requirements of real-time applications. Further, the Python code of the proposed method is made public to support collaborative research efforts and ensure the reproducibility of the presented work.

Vibration-based condition monitoring (VBCM) can be defined as a signal-based methodology for assessing a system condition based on its inherent vibration patterns. By monitoring changes in vibration signatures, which reflect a change in the system's current state, VBCM provides a non-invasive, real-time approach to continuously monitor the system's condition. The main advantages of VBCM over alternative forms of condition monitoring include [4] [1]:

- Vibration sensors are non-intrusive and can be contactless, facilitating non-destructive condition monitoring.
- Real-time acquisition of vibration signals can be conducted in situ, allowing for on-line local condition monitoring.
- Trending vibration analysis can be utilized to identify relevant conditions and conduct comparative analysis across diverse conditions or objects.

- Vibration sensors are cost-effective and widely available, offering various specifications to suit various requirements.
- Vibration waveform responds instantly to changes in the monitored condition and, therefore, is suitable for continuous and intermittent monitoring applications.
- Signal processing techniques can be applied to vibration signals to mitigate corrupting noise and extract weak condition indications from other masking signals.

The process of VBCM involves two main aspects: feature extraction and condition monitoring. Feature extraction involves the application of various methods to analyze vibration signatures and extract relevant features that reflect condition changes in the system's condition. Such signatures often manifest through sudden changes in amplitude, frequency, and phase characteristics of the generated vibration pattern. Existing feature extraction methods include time-domain, frequency-domain, and time-frequency-domain methods. In time-domain-based methods [145, 146, 148, 153, 159, 162, 163], features are calculated from the signal's amplitude, representing specific aspects of the signal's dynamics over its time period. Common time-domain feature types include shape features and statistical features. Shape features include maximum, minimum, peak-to-peak value, and crest factor. Statistical features describe characteristics of the probability distribution of the signal, such as mean, standard deviation, variance, skewness, and kurtosis. Time-domain feature extraction approaches are generally simple to implement. Further, they are computationally efficient since signal transformation is not required, making them advantageous in applications where real-time processing is crucial, especially with limited computational resources. However, time-domain analysis is highly susceptible to noise since noise presence would mask the dynamic characteristics of the signal. Furthermore, to achieve reliable performance, input vibration segments of relatively large durations are usually required to precisely capture evolving changes and complexities within the segment. Frequency-based methods [168, 171, 174, 178–181, 309] represent the signal in terms of its spectral con-

tents, revealing details that are not apparent in the time-domain waveform. Accordingly, discriminative spectral features can be extracted from the signals' spectrum. In contrast to time-domain analysis, frequency-based analysis allows the identification and removal of noise or unwanted components by applying appropriate frequency filtering mechanisms. In the context of VBCM of rolling bearings, the generated vibration pattern spans a broad spectrum of frequency components, including characteristic frequencies related to the bearing geometry and operational speed, as well as harmonic and sideband frequencies caused by various operational conditions. This makes spectral analysis particularly effective in discriminating between healthy and abnormal vibration patterns. However, in frequency-domain methods, the analysis spans the signal's entire duration, lacking the ability to provide temporal information about the timing of these patterns within the signal. In contrast, time-frequency methods [56, 218, 225, 228, 231, 244, 248, 250, 310] transform the signal into energy-time-frequency representations where the signal's energy is mapped across both time and frequency. This, in turn, allows the identification of time-varying spectral characteristics within the signal. Common time-frequency transforms include short-time Fourier transform (STFT), Hilbert-Huang transform (HHT), and wavelet transform (WT). Despite their effectiveness, especially when handling nonstationary and nonlinear signals, time-frequency methods involve higher computational complexity—in terms of online processing time and memory usage—compared to pure spectral analysis. These factors directly influence the reliability of the monitoring process and associated financial expenses. Specifically, an increase in memory demand increases financial costs, while lengthy delays in condition prediction may fail to prevent costly catastrophic failures. In the context of VBCM of rolling bearings, vibrations of damaged bearings are manifested as modulations in the amplitude of the generated vibration signal [307]. This makes envelope-based analysis [295, 297, 298, 306, 307] an effective approach to facilitate efficient condition monitoring. In envelope-based analysis, the Hilbert transform (HT) is commonly employed to obtain Instantaneous Amplitude (IA) “envelope”, Instantaneous Phase (IP), and Instanta-

neous Frequency (IF). Accordingly, various feature extraction techniques can be used to extract relevant features from obtained instantaneous information.

To this end, this chapter introduces a low-complexity method for VBCM of rolling bearings based on joint analysis of IA and IF information of the generated vibration signal. Specifically, the proposed method employs the HT to obtain and then jointly analyze IA and IF information, facilitating new joint instantaneous amplitude-frequency representations of the vibration signal. Accordingly, a set of six new fault-sensitive features are engineered from these representations. Further, the proposed method uses very short durations of the generated vibration signal for condition monitoring, thereby relaxing memory requirements and reducing monitoring delay. The main contributions of the chapter include:

- Introducing a low-complexity new method for VBCM of rolling bearings based on envelope analysis.
- In the proposed method, instantaneous amplitude “envelope” and instantaneous frequency of the vibration signals are jointly analyzed to facilitate three novel envelope representations: Instantaneous Amplitude-Frequency Mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and Instantaneous Energy-Frequency Distribution (IEFD).
- Maintaining temporal information, the introduced representations effectively capture energy-frequency variations that are unique to the condition of the bearing, thereby enabling the extraction of discriminative features with high sensitivity to variations in operational conditions.
- Accordingly, new highly discriminative features are extracted from these representations.
- The extracted features are engineered to characterize shapes of the proposed instantaneous representations, thereby capturing instantaneous energy-frequency dynamics

of the signal's envelope.

- The proposed method facilitates a low-complexity VBCM since it utilizes input vibration segments of very short durations (0.1 seconds) and produces only six features. Thus, relaxing memory requirements and reducing monitoring delays, which, in turn, helps reduce memory costs and prevent costly catastrophic failures.

The remainder of the chapter is structured as follows: Related work is presented in the next section. Section 6.3 introduces the proposed method. Section 6.4 addresses performance evaluation in terms of the dataset used and experimental setup. Section 6.5 introduces and discusses the obtained results. The chapter is finally concluded in Section 6.6.

6.2 Related Work

The HT is commonly employed in VBCM applications to obtain instantaneous information (IA, IP, and IF) of generated vibration patterns and, consequently, extract distinctive fault-related features [225, 228, 231, 295, 297, 298, 306, 353–372]. The existing HT-based approaches for feature extraction can be generally grouped into three broad categories: Envelope-based spectral analysis [297, 354, 358, 362, 366–369], signal decomposition and envelope reconstruction [298, 353, 355, 357, 360, 361, 363, 364], and time-frequency analysis [225, 228, 231, 306, 356, 359, 362, 365, 370–372]. Envelope-based spectral analysis involves analyzing the spectrum of the obtained envelope (IA) to extract fault-related features. While pure frequency analysis of the signal envelope is effective under steady-state conditions, it falls short in capturing the complexity of bearing vibrations under time-varying speed and load conditions, which is essential for real-time applications.

Signal decomposition and envelope reconstruction approaches utilize signal decomposition techniques such as wavelet decomposition and Adaptive Mode Decomposition (AMD) to decompose the time-domain envelope into elementary modes. A screening process is then

conducted to identify fault-informative modes. Finally, the signal is reconstructed using the identified modes only, and the envelope spectrum is obtained accordingly for fault analysis. The main advantage of decomposition-based approaches is that they attempt to reconstruct a low-redundancy and highly fault-sensitive envelope through decomposition, screening, and reconstruction processes. Additionally, selecting fault-related modes and discarding the remaining modes reduces noise presence in the reconstructed signal, thereby improving the signal-to-noise ratio (SNR) of the obtained envelope compared to the original signal. However, such approaches involve high computational burdens due to decomposition, screening, and reconstruction processes. Further, proper mode screening criteria should be applied to avoid losing useful information.

Time-frequency analysis is particularly useful in analyzing nonstationary vibration patterns whose spectral properties change over time. In the context of HT-based VBCM, HHT and STFT are commonly adapted to perform time-frequency analysis of the signal envelope. Unlike envelope spectral analysis, which utilizes the Fourier Transform (FT) to analyze the spectral contents of the envelope over its entire duration, STFT-based approaches analyze the spectral contents of the envelope over finite short-duration time windows, thereby preserving temporal information. A major limitation in STFT is its fixed segment length, which results in a uniform resolution analysis of the envelope, leading to an inherent compromise between time and frequency resolutions when setting the duration of the time window. HHT employs an alternative approach for time-frequency analysis through two steps: signal decomposition and Hilbert Spectral Analysis (HSA). In contrast to STFT, which relies on fixed-time durations and sinusoidal kernel functions for envelope analysis, HHT utilizes an adaptive approach to analyze the vibration signal, making it highly responsive to variations in vibration patterns. Specifically, HHT uses AMD techniques to decompose the signal into a set of simpler modes known as Intrinsic Mode Functions (IMFs), representing different frequency components of the signal. After the decomposition process, HSA is conducted, where the HT is applied to each mode to obtain its IA and IF infor-

mation, providing a detailed energy-time-frequency representation of the original signal commonly known as the Hilbert spectrum. The adaptive mechanism of HHT and the use of HSA makes it very effective in analyzing complex vibration patterns. The performance of the HHT, however, is highly dependent on the reliability of the decomposition process and its parameters, such as the stopping criterion for the decomposition process and the interpolation method for envelope estimation. Moreover, HHT involves high computational complexity due to its adaptive decomposition mechanisms, which makes it unsuitable for real-time VBCM applications.

This chapter introduces a low-complexity method for VBCM of rolling bearings based on envelope analysis. The proposed method jointly utilizes the IA and IF of the vibration signal to facilitate three novel envelope representations that maintain temporal information to effectively capture energy-frequency variations in the signal envelope. The following section introduces the theoretical background and details of the proposed method.

6.3 Joint Instantaneous Time-Frequency Analysis of Vibration Signals

6.3.1 The Hilbert Transform

The Hilbert transform (HT) is a fundamental operator in signal processing since it provides an efficient way to obtain analytic signal representations of real-valued signals. An analytic signal is a complex-valued representation of a signal that describes its amplitude and phase characteristics. Given signal $x(t)$, its HT is defined as:

$$H\{x(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (6.1)$$

The transform essentially modifies the phase of each frequency component of the signal by $\pm 90^\circ$. Accordingly, the analytical signal $x_a(t)$ of $x(t)$ is formed by augmenting the signal with its HT $H\{x(t)\}$ as the imaginary part:

$$x_a(t) = x(t) + jH\{x(t)\} \quad (6.2)$$

Accordingly, the IA or envelope, $A(t)$ of the signal is given by:

$$A(t) = |x_a(t)| = \sqrt{x(t)^2 + H\{x(t)\}^2}, \quad (6.3)$$

and the IP, $\theta(t)$ is expressed as:

$$\theta(t) = \arctan\left(\frac{H\{x(t)\}}{x(t)}\right) \quad (6.4)$$

Hence, the IF, $F(t)$ can be calculated by taking the derivative of the instantaneous phase $\theta(t)$:

$$F(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \quad (6.5)$$

6.3.2 Proposed Joint Instantaneous Time-Frequency Analysis

Fig. 6.1 shows IA and IF information, obtained through the HT, of a real-world vibration signal of a rolling bearing, denoted as $v_h(t)$. The signal is extracted from the Paderborn University (PU) bearing dataset [344]. $v_h(t)$ represents a healthy vibration pattern and contains 6,400 datapoints that were acquired at a sampling rate of 64,000 Hz; thus, the signal spans a 0.1 seconds of time duration. As shown, the IA and IF information show how the signal's amplitude and frequency evolve over time, respectively. Hence, they separately describe the signal's time-energy and time-frequency characteristics, which are of paramount importance since faults often manifest as changes in the vibration signal's energy and/or frequency content. Accordingly, in this chapter, IA and IF of the vibration signal are jointly

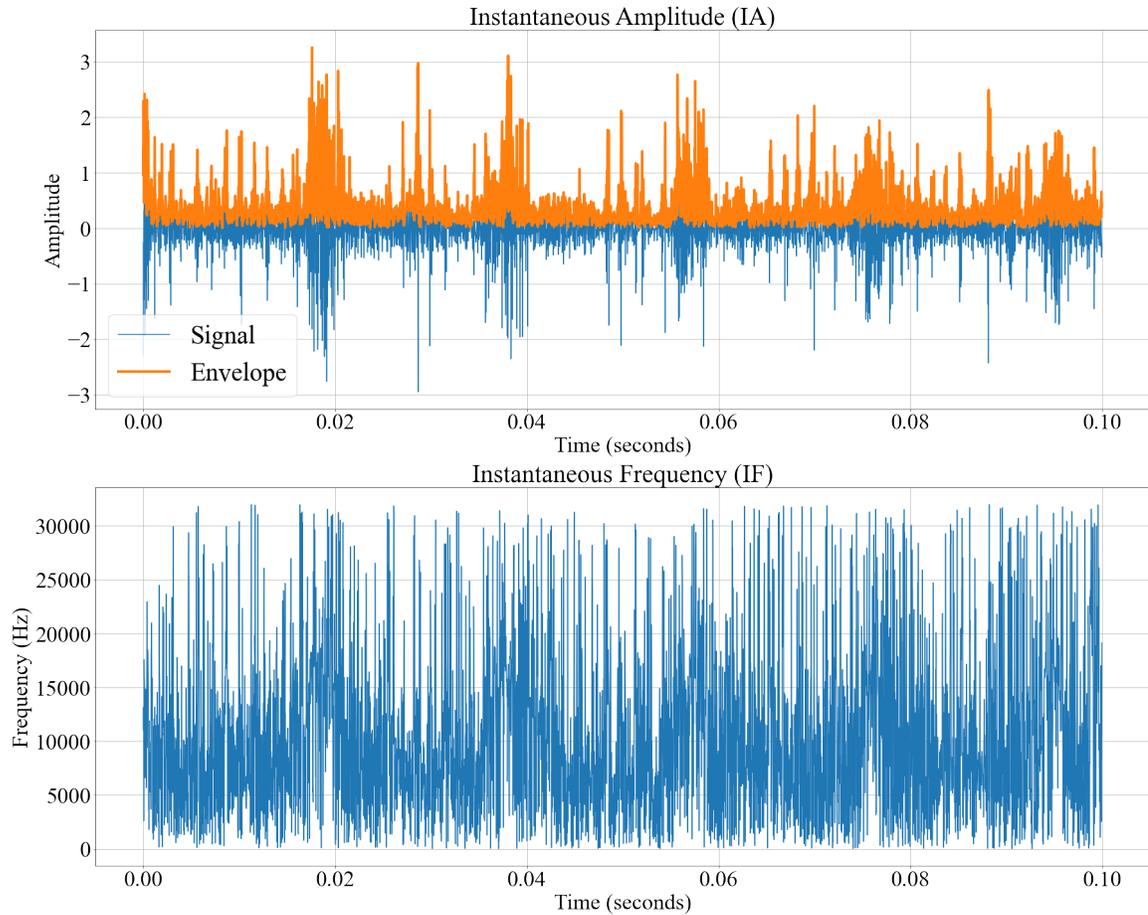


Figure 6.1: Instantaneous amplitude (IA) and Instantaneous frequency (IF) of a rolling bearing vibration signal, $v_h(t)$.

utilized to conduct an efficient instantaneous amplitude-frequency analysis of the signal envelope. The joint instantaneous time-frequency analysis is facilitated through three novel envelope representations: Instantaneous amplitude-frequency mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and Instantaneous Energy-Frequency Distribution (IEFD). Accordingly, highly discriminative features are extracted from these representations, characterizing the shapes of the proposed representations and capturing instantaneous energy-frequency dynamics of the signal's envelope.

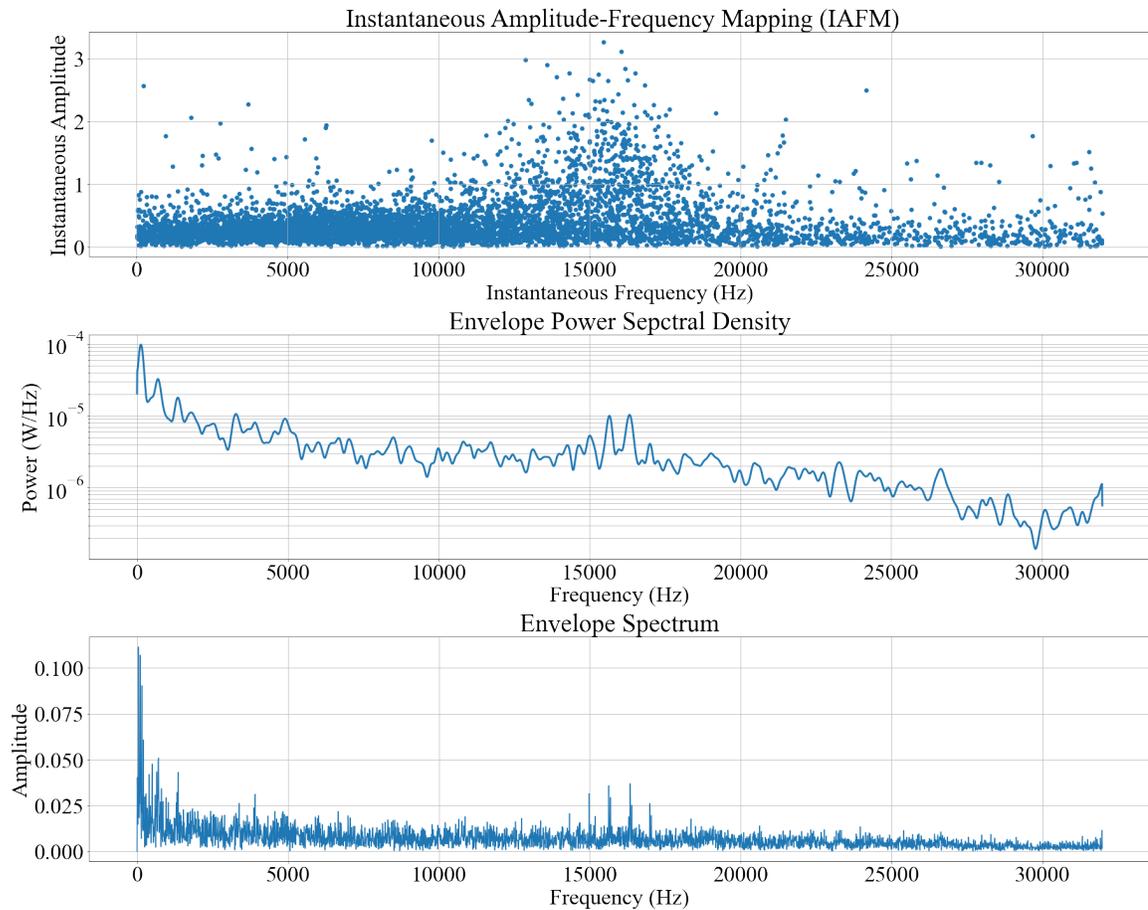


Figure 6.2: Instantaneous amplitude-frequency mapping (IAFM), Envelope PSD, and envelope spectrum of the healthy vibration pattern, $v_h(t)$.

Instantaneous Amplitude-Frequency Mapping

Given IA and IF information, the Instantaneous Amplitude-Frequency Mapping (IAFM) is reconstructed by mapping IA and IF information together where the x -axis represents the IF, and the y -axis represents the IA as depicted in Fig. 6.2. The figure also displays the PSD and frequency spectrum of the envelope. A comparison among the three spectral representations reveals that the proposed IAFM effectively captures the spectral shape of the envelope. In PSD and frequency spectrum, each $x - y$ pair represents a unique frequency-energy value across the spectral representation, which is helpful in precisely identifying dominant frequency components. In contrast, as an instantaneous mapping, the IAFM shows how signal amplitude (energy) values are allocated through various frequencies over

the signal's time duration, thereby providing an energy-frequency density representation of the envelope with high energy resolution. This makes IAFM particularly useful in characterizing energy-frequency behaviors of healthy and faulty vibration patterns through the spread and concentration of energy (amplitude) values across various frequencies. Accordingly, the following features are extracted from IAFM that capture and characterize its shape:

- **Spectral Centroid (SC):** The SC represents the center of gravity of the IAFM and quantifies the average frequency at which the energy is concentrated. It describes the spectral position of dominant oscillations in the generated vibration, thereby identifying various vibration patterns. Given that IA and IF information are represented in the discrete forms $A[n]$ and $F[n]$, respectively, SC is calculated through the following formula:

$$SC = \frac{\sum_{n=0}^{N-1} F[n] \cdot A[n]}{\sum_{n=0}^{N-1} A[n]} \text{ Hz} \quad (6.6)$$

where N is the number of total datapoints in $A[n]$ and $F[n]$.

- **Spectral Spread (SS):** It describes the deviation of instantaneous amplitude-frequency points in the IAFM with respect to the SC. Hence, it provides a measure to assess the variance of the IAFM. SS is obtained by calculating the standard deviation with respect to SC:

$$SS = \sqrt{\frac{\sum_{n=0}^{N-1} (F[n] - SC)^2 \cdot A[n]}{\sum_{n=0}^{N-1} A[n]}} \text{ Hz} \quad (6.7)$$

- **Coefficient of Variation (CoV):** This metric provides insights into the dispersion of the envelope's energy across its frequency components. CoV is expressed as the ratio of the SC to the CC:

$$CoV = \frac{SS}{SC} \times 100 (\%) \quad (6.8)$$

CoV quantifies the dispersion of the IAFM with respect to SC, indicating how spread out the energy is in relation to the mean frequency. A higher CoV implies a wider

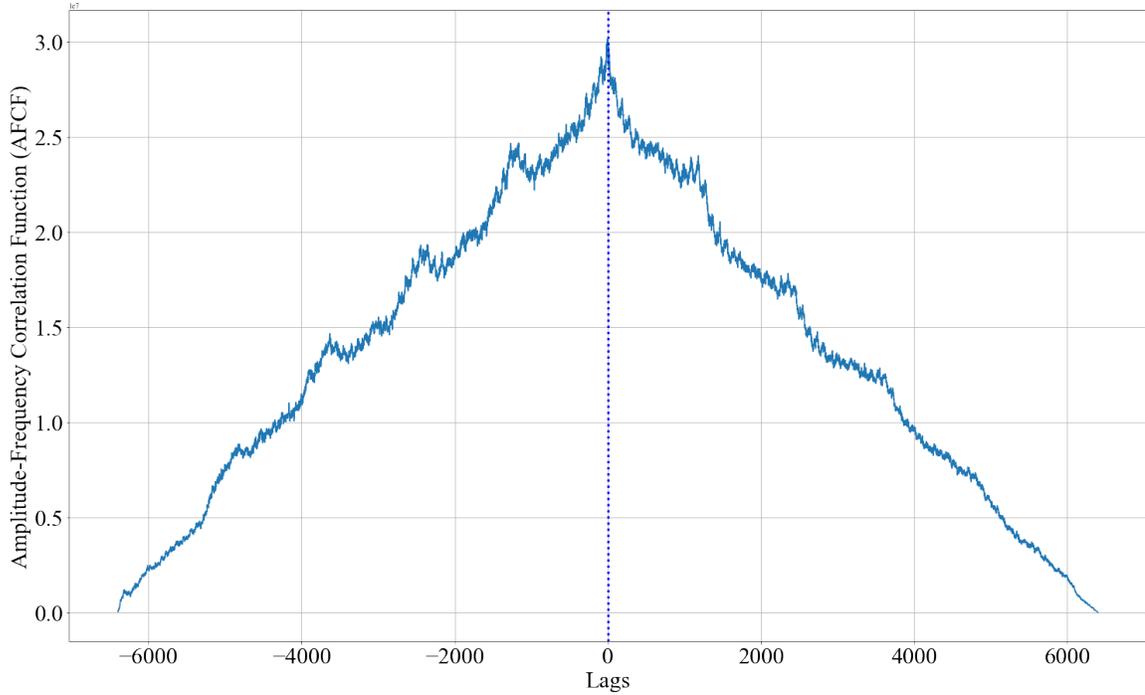


Figure 6.3: IAFC of the healthy vibration pattern, $v_h(t)$

distribution of the envelope's energy across its instantaneous spectrum, suggesting a more dispersed frequency content. Conversely, a lower CoV indicates a more concentrated distribution of frequencies around the SC, implying that the envelope's energy is more narrowly focused around a central frequency.

Instantaneous Amplitude-Frequency Correlation

The second proposed representation is the Instantaneous Amplitude-Frequency Correlation (IAFC), which represents the cross-correlation of IA and IF:

$$\text{IAFC} = R_{AF}[k] = \sum_{n=0}^{N-1} A[n] \cdot F[n - k], \quad k = 0, \pm 1, \dots, \pm N - 1 \quad (6.9)$$

As instantaneous quantities, IA and IF maintain the same temporal dependencies among their values. Accordingly, the IAFC would capture the mutual dynamics between IA and IF that are unique to the system's health state since changes in operational conditions often

manifest as variations in amplitude and frequency. Hence, the IAFC can be interpreted as an envelope representation that describes the relationship between the envelope's energy and frequency variations over time. Since IA and IF are both positive quantities, the IAFC would possess common characteristics as well as condition-specific features. Specifically, given the IAFC of the healthy vibration pattern, $v_h(t)$ depicted in Fig.6.3, the following observations are made:

1. The IAFC is a positive function since IA and IF are positive quantities.
2. The IAFC has an increasing trend until it reaches its peak, after which it shows a decreasing trend.
3. The peak, representing the maximum cross-correlation, typically occurs at full overlap between IA and IF (zero lag). Otherwise, it reflects very high excitations in either amplitude, frequency, or both, which would indicate abnormal behavior.

Accordingly, the following features are extracted from the IAFC to quantify the aforementioned characteristics, thereby capturing various operation conditions:

- Correlation Peak (CP): It represents the maximum value of IAFC, quantifying the maximum cross-correlation. CP is expressed mathematically as follows:

$$CP = \max \{R_{AF}[k]\}, k = 0, \pm 1, \dots, \pm N - 1 \quad (6.10)$$

- Peak Lag (PL): PL is the lag at which the $R_{AF}[k]$ has its maximum value. In other words, it is the lag value corresponding to the CP.

The values of CP and PL characterize the maximum cross-correlation in the IAFC, which would be a discriminative indicator between healthy and faulty conditions, as explained above.

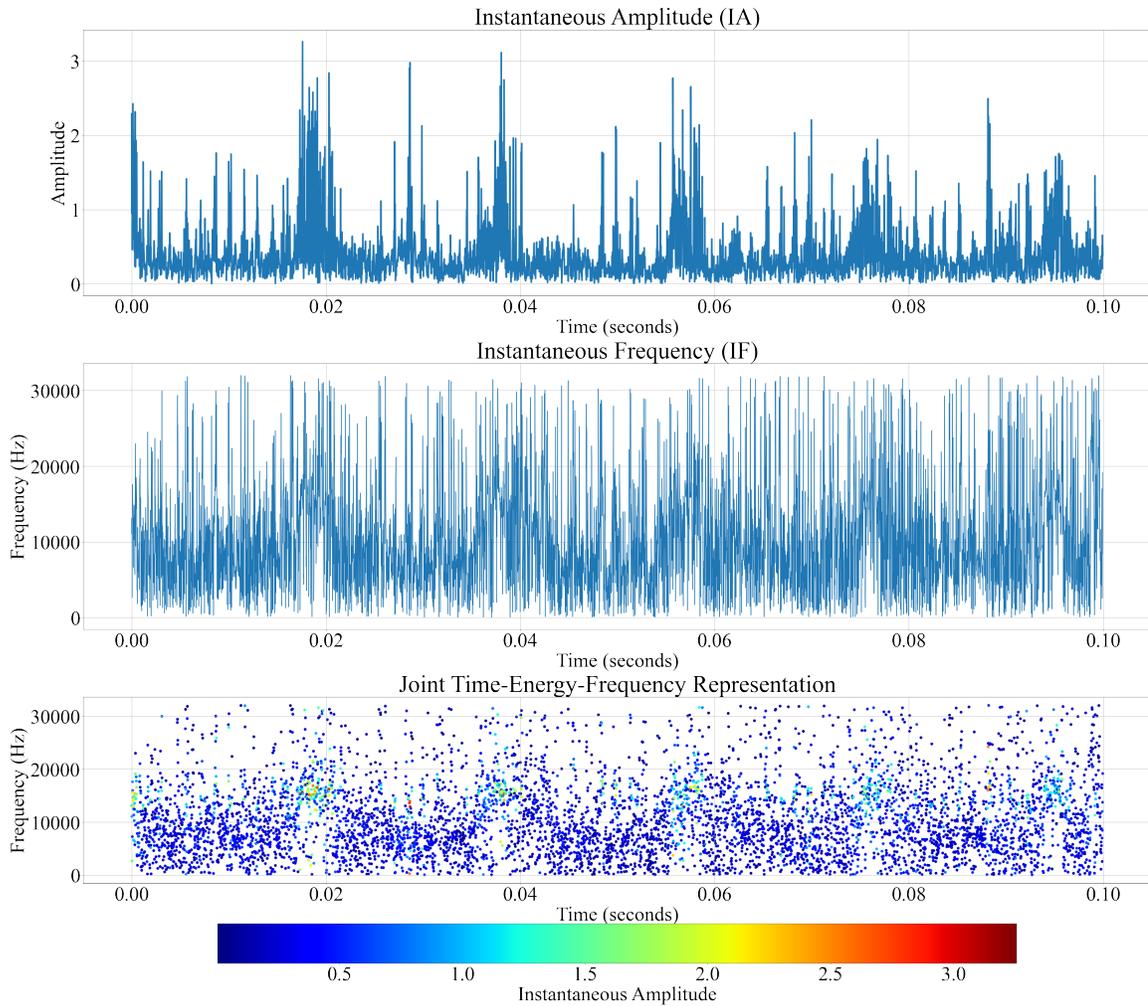


Figure 6.4: Joint time-energy-frequency representation of the healthy vibration pattern, $v_h(t)$

Instantaneous Energy-Frequency Distribution

As previously mentioned, IA and IF information separately show how the envelope’s energy and spectral content evolve over time. Specifically, while IA shows how the energy is distributed over the time period, the IF shows how the frequency changes over that time period. Accordingly, a joint time-energy-frequency representation of the signal can be visualized by mapping the intensity of each time-frequency pair in the IF to its corresponding value of the IA, as demonstrated in Fig. 6.4 for the healthy vibration pattern, $v_h(t)$. This visual heatmap is an effective tool for visualizing the instantaneous components since it

clearly represents the variations in the envelope's energy and frequency over time. To facilitate this visualization efficiently, the third proposed representation is introduced. Namely, Instantaneous Energy-Frequency Distribution (IEFD), which is the product of normalized instantaneous energy (IE_{norm}) and normalized IF (IF_{norm}). The IE_{norm} is obtained by squaring and normalizing the instantaneous amplitude $A[n]$:

$$IE_{norm}[n] = \frac{A[n]^2}{\sum_{n=0}^{N-1} A[n]^2} \quad (6.11)$$

The squaring and normalization processes emphasize larger values of IA more than smaller ones and transform the instantaneous amplitude values into values that represent their proportion of the envelope's total energy over the given time period. Thus, IE_{norm} is more sensitive to points in time where the energy is significantly higher, as these points will contribute disproportionately to the sum. The IF_{norm} is obtained by normalizing the instantaneous frequency $F[n]$:

$$IF_{norm}[n] = \frac{F[n]}{\sum_{n=0}^{N-1} F[n]} \quad (6.12)$$

Normalizing IF makes the representation more sensitive to changes in the frequency content of the envelope. Specifically, the normalization process provides a measure of the relative contribution of each instant's frequency to the overall envelope's spectrum, thereby amplifying the impact of large frequency values in fault analysis. Accordingly, the IEFD is expressed as follows:

$$\begin{aligned} IEFD[n] &= IE_{norm}[n] \times IF_{norm}[n] \\ &= \frac{A[n]^2}{\sum_{n=0}^{N-1} A[n]^2} \times \frac{F[n]}{\sum_{n=0}^{N-1} F[n]} \end{aligned} \quad (6.13)$$

By combining the normalized instantaneous energy and frequency information, the IEFD provides a comprehensive representation that captures crucial dynamics of the generated vibration. Hence, IEFD allows for a localized analysis in the time-energy-frequency domain,

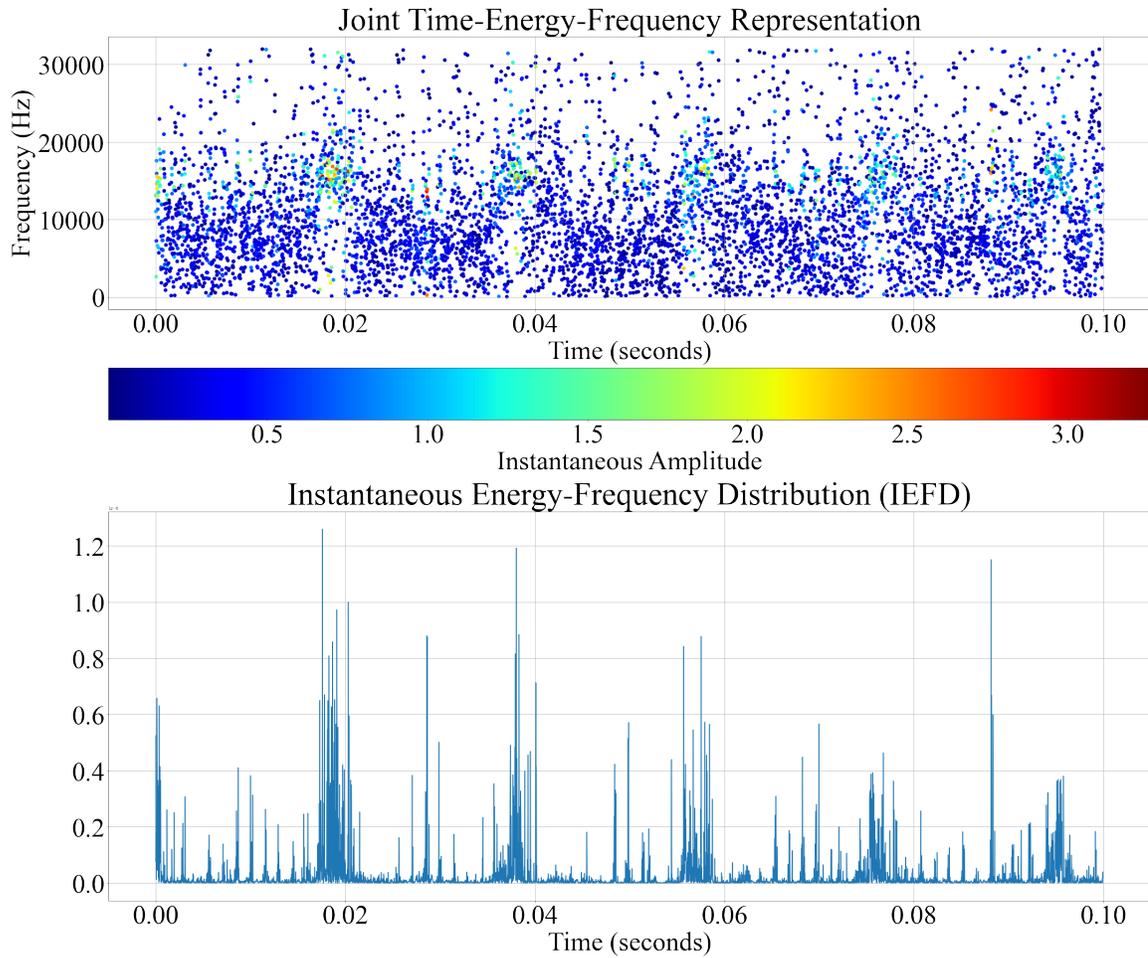


Figure 6.5: IEFD along with the joint time-energy-frequency visualization of the healthy vibration pattern, $v_h(t)$.

enhancing the detection of changes or shifts in operational behavior indicative of faults. Fig. 6.5 shows and compares the IEFD with the joint time-energy-frequency heatmap of the healthy vibration pattern, $v_h(t)$. The heatmap visualization is obtained by mapping the intensity of each time-frequency pair in the IF to its corresponding value of the IA as previously explained. The comparison shows that the IEFD effectively captures the time-varying energy-frequency characteristics of the signal's envelope since its peaks correspond to the most intense time-frequency pairs in the heatmap. These characteristics are highly influenced by changes in the operational conditions since the presence of a fault is often characterized by a high level of regularity in the generated vibration due to the emergence

Table 6.1: Proposed envelope representations and extracted features

Proposed Representation	Extracted Features
Instantaneous amplitude-frequency mapping (IAFM)	Spectral Centroid (SC)
	Spectral Spread (SS)
	Coefficient of Variation (CoV)
Instantaneous Amplitude-frequency correlation (IAFC)	Correlation Peak (CP)
	Peak Lag (PL)
Instantaneous energy-frequency distribution (IEFD)	Mean-to-entropy ratio (MER)

of highly energetic frequency components associated with the fault. Conversely, normal operation conditions would generate more complex and highly irregular vibration patterns. The Mean-to-Entropy Ratio (MER) of the IEFD is used to quantify these characteristics and extract a feature that is sensitive to such changes in operational conditions. The MER represents the ratio between the mean and Shannon entropy of IEFD; it is expressed mathematically as follows:

$$MER = \frac{\frac{\sum_{n=0}^{N-1} IEFD[n]}{N}}{-\sum_{i=0}^{M-1} x_i \log_2 P(x_i)} \quad (6.14)$$

where $i = 0, 1, \dots, M - 1$, M is the number of unique values in the IEFD, and $P(x_i)$ is the probability of each unique value x_i . The mean value of the IEFD quantifies its average concentration in terms of joint energy-frequency content across the observed time period, while entropy quantifies irregularity within the IEFD.

To this end, instantaneous amplitude and frequency information of vibration signals are jointly utilized to facilitate three novel envelope representations. Accordingly, six highly discriminative features are extracted from these representations as summarised in Table 6.1.

Complexity Analysis

Algorithm 2 shows the pseudo-code and computation steps of the proposed method. Accordingly, the complexity of the proposed method can be analyzed as follows:

- Complexity of the HT is $O(N \log N)$ since it is commonly implemented using the fast

Algorithm 2 Proposed Joint Instantaneous Time-Frequency Analysis

Input: vibration signal $x[n]$ of length N data points

Parameters: f_s = sampling frequency (samples/second).

Output: $\mathcal{F}[s]$, $s = 0, \dots, S - 1$: features vector of size $S = 6$

Start:

Compute Hilbert Transform (HT) of $x[n]$: $H\{x[n]\} = HT(x[n], f_s)$.

Compute instantaneous amplitude $A[n]$ and instantaneous frequency $F[n]$, $n = 0, 1, \dots, N$.

% 1- Instantaneous Amplitude-Frequency Mapping (IAFM):

Calculate SS, $\mathcal{F}[0] \leftarrow SS$

Calculate SC, $\mathcal{F}[1] \leftarrow SC$

Calculate CoV, $\mathcal{F}[2] \leftarrow CoV$

% 2- Instantaneous Amplitude-Frequency Correlation (IAFC):

Calculate CP, $\mathcal{F}[3] \leftarrow CP$

Calculate PL, $\mathcal{F}[4] \leftarrow PL$

% 3- Instantaneous Energy-Frequency Distribution (IEFD):

Calculate MER, $\mathcal{F}[5] \leftarrow MER$

End

Fourier transform (FFT) [373, 374], which is the case in both MATLAB and Python SciPy implementations of the HT.

- Complexity of the cross-correlation function IAFC is $O(N \log N)$ considering FFT-based cross-correlation.
- Complexity of the other computations are in the order of $O(N)$.

Thus, the most computationally intensive steps are those involving the HT and the IAFC, each with a complexity of $O(N \log N)$. Therefore, the overall complexity of the proposed method is dominated by these steps, leading to an overall computational complexity of $O(N \log N)$.

Table 6.2: Experimental dataset

	Class	Health condition	Fault type	Motor speed (rpm)
PU dataset	1	Healthy	NA	1500
	2	Combined IR and OR faults	Multiple damages	
	3	IR faults	Single, repetitive, and multiple damages	
	4	OR faults	Single and repetitive damages	

6.4 Performance Evaluation

6.4.1 Experimental Dataset

The performance of the proposed method is evaluated on the Paderborn University (PU) bearing dataset [344]. In contrast to other datasets, the PU dataset has real bearing damages with combined defects. The experiments are conducted using 425 W Permanent Magnet Synchronous Motor (PMSM). The dataset used in this chapter is based on measurements conducted at $n = 1,500$ rpm with a load torque of $M = 0.7$ Nm and a radial force on the bearing of $F = 1,000$ N. Vibration signals are recorded with a sampling rate of 64,000 Hz by measuring the acceleration of the bearing housing at the adapter at the top end of the rolling bearing module. Regarding bearing defects, the PU dataset includes artificially generated and accelerated-lifetime defects. In this chapter, only accelerated-lifetime defects are used. Accordingly, the dataset has four classes: one healthy class and three faulty classes according to fault type, as shown in Table 6.2.

6.4.2 Experimental Setup

In the preprocessing stage, vibration signals of the dataset are segmented into non-overlapping segments of $N = 6,400$ samples. Accordingly, the resultant segment duration is 0.1 seconds, given that the vibration signals are acquired at 64,000 Hz. The segment duration of 0.1 seconds is precise enough to facilitate real-time monitoring with moderate compu-

tational requirements. The segmentation process results in a dataset of 16,005 vibration segments in total. After extracting the features, the resulting dataset is divided into two parts: 11,203 samples for training (70%) and 4,802 samples for testing (30%). A random forest (RF) classifier is then trained on the training dataset. The selection of RF is driven by its better performance compared to other classifiers, as reported in [56]. The performance of the proposed method is compared with two common time-frequency methods: the short-time-Fourier transform (STFT) of the signal envelope and The Hilbert-Huang transform (HHT) of the signal. In the STFT-based method, the envelope of the vibration signal is segmented into three segments using the Hamming window with 50% overlap. The FFT is then computed for each segment with the number of FFT points (NFFT) set equal to 4,096. The proposed features are calculated accordingly, considering that $A[n]$ represents aggregated STFT power coefficients and $F[n]$ represents frequency bins. In the HHT-based method, the obtained IA and IF information of the resulting intrinsic mode functions (IMFs) are aggregated, and the proposed features are calculated accordingly. The following metrics are used to evaluate the performance:

- Prediction accuracy (%) and ROC-AUC score to evaluate the reliability and effectiveness of the proposed method in condition monitoring.
- Online processing time and memory usage to assess computational complexity of the proposed method.

Python programming language, along with *SciPy* and *emd* libraries, are used to build the models. The code is available publicly on the Github site of the Optimized Computing and Communications (OC²) Laboratory¹. The achieved results and related discussion are presented in the next section.

¹<https://github.com/Western-OC2-Lab/Joint-Instantaneous-Amplitude-Frequency-Analysis-for-Vibration-Based-Condition-Monitoring>

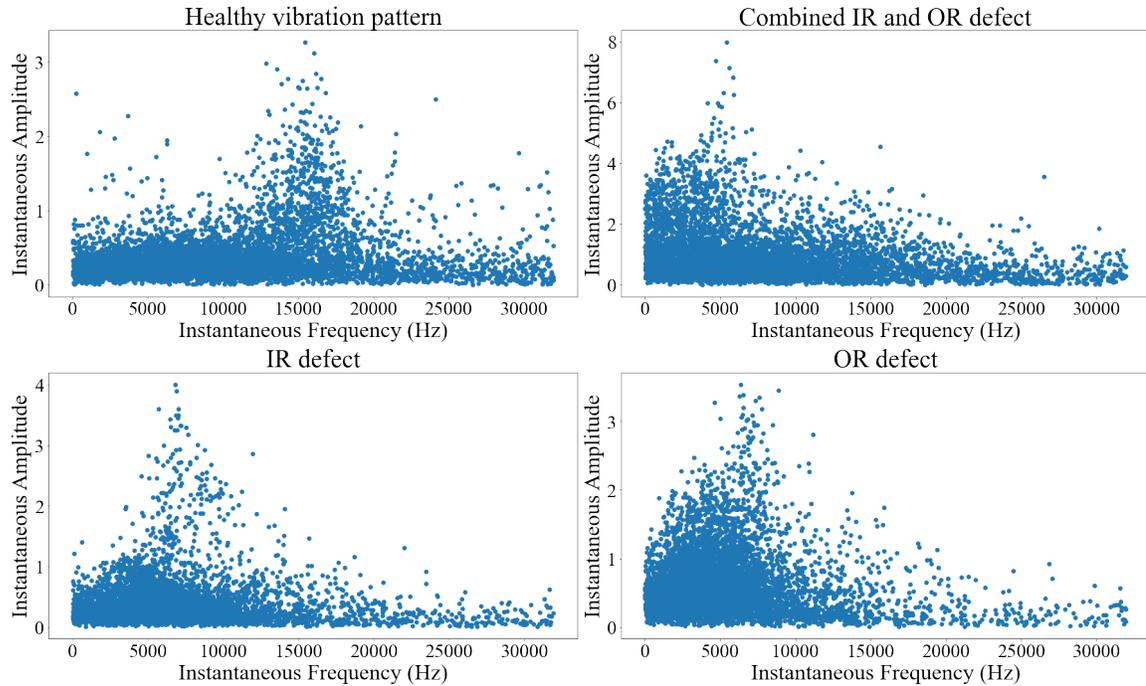


Figure 6.6: IAFMs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.

6.5 Results and Discussion

Fig. 6.6, Fig. 6.7, and Fig. 6.8 compare between IAFMs, IAFCs, and IEFDs, respectively, across healthy and faulty vibration patterns from the PU dataset. It is evident that the proposed representations of vibration signal envelope distinctly capture the unique variations between healthy and faulty bearing conditions by emphasizing the contrast in energy and frequency among various operational conditions. Thus, improving the discriminative characteristics of the extracted features is crucial for effective condition monitoring.

Table 6.3 shows the achieved performance results of the proposed method along with results of HHT-based features and STFT-based features. In terms of accuracy and ROC-AUC score, the proposed method achieved superior performance (accuracy > 99.6%, ROC-AUC score = 1.00), reflecting its reliability and effectiveness in the detection and diagnosis of various fault types. Further, the results demonstrate the effectiveness of the proposed in-

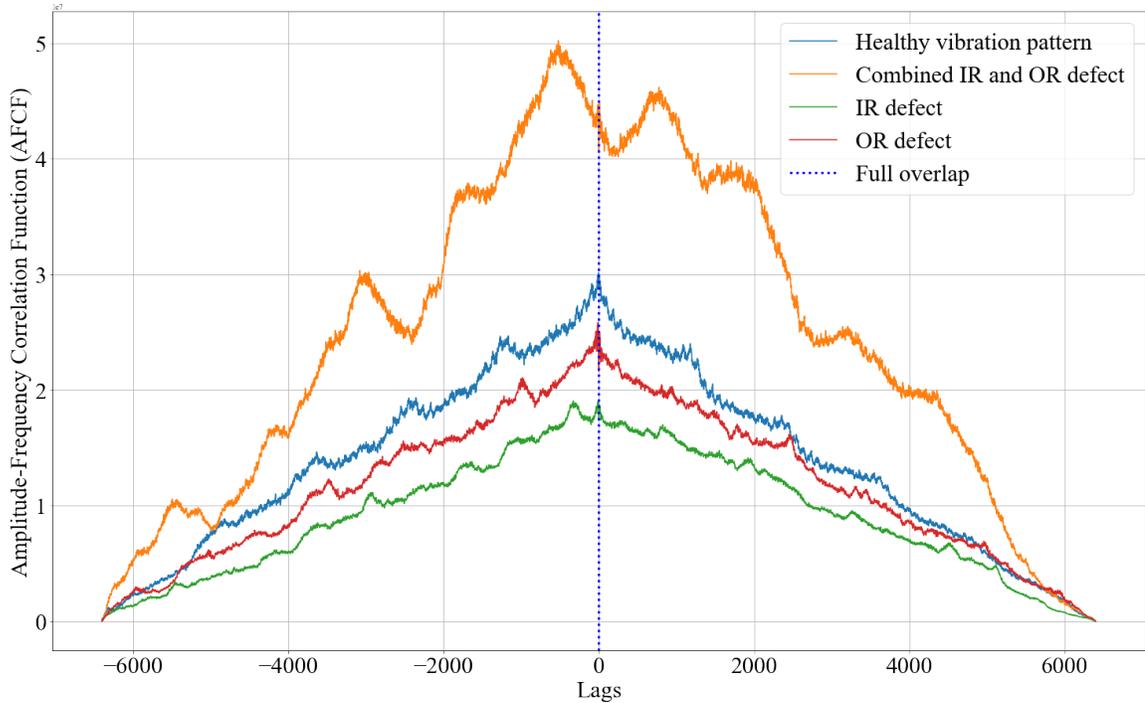


Figure 6.7: IAFCs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.

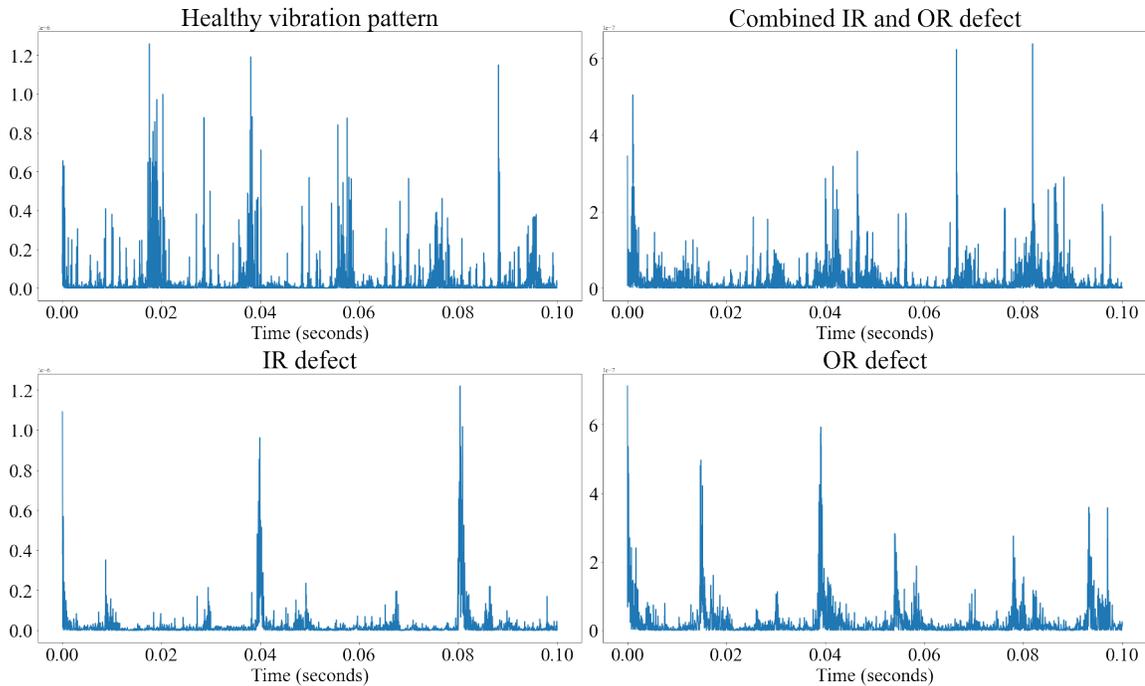


Figure 6.8: IEFDs of healthy, inner race (IR) defect, outer race (OR) defect, and combined IR and OR defect vibration patterns.

Table 6.3: Performance comparison among the proposed method, the STFT-based method, and the HHT-based method

Method	Accuracy (%)	ROC-AUC score	Online processing time (seconds)	Memory usage (MB)
Proposed	99.60%	1.00	0.16	1.1
STFT-based	95.20%	0.997	0.16	0.9
HHT-based	91.46%	0.989	0.24	2.5

stantaneous amplitude-frequency representations in extracting highly fault-sensitive features compared to the other two methods. In contrast to the STFT-based method, which facilitates time-varying spectral representation of the signal envelope, the proposed method effectively utilizes instantaneous frequency, along with the envelope, to jointly facilitate instantaneous amplitude-frequency representations. As the results show, the proposed method outperforms the STFT-based method, indicating that the proposed representations effectively capture more unique variations in energy and frequency between healthy and faulty bearings compared to the STFT-based representation. This could be explained by the limitation of fixed segment length in STFT, which results in an inherent compromise between time and frequency resolutions in the time-energy-frequency representation. Regarding HHT, although it utilizes instantaneous amplitude and frequency information to facilitate time-energy-frequency analysis, it results in lower accuracy compared to the other two methods. This gap in performance can be justified by impairment caused by the EMD process, such as mode mixing, end effects, and over-sifting. Regarding computational complexity, both the proposed method and the STFT-based methods demonstrate comparable and moderate computational requirements compared to the HHT-based method. This is expected since the iterative sifting process used in EMD to extract the IMFs is computationally intensive, especially for long vibration segments. In contrast, the proposed method uses vibration segments of 0.1 seconds of duration and results in six features only, reducing the processing time and relaxing memory requirements.

Table 6.4 presents a comparison, in terms of achieved prediction accuracy (%), with recent

Table 6.4: Performance comparison between the proposed method and other DL-based methods on the PU dataset

Method	Approach	Achieved accuracy (%)
Proposed	Signal processing-based	99.60%
Y. Kim <i>et al.</i> , 2024 [375]	DL-based	99.44%
L. Cui <i>et al.</i> , 2024 [376]	DL-based	97.05%
Z. Zhang <i>et al.</i> , 2024 [377]	DL-based	96.51%
Y. Xue <i>et al.</i> , 2024 [378]	DL-based	99.50%
H. Geng <i>et al.</i> , 2022 [350]	DL-based	96.67%
K. Zhang <i>et al.</i> , 2022 [351]	DL-based	97.68%
A. Sharma <i>et al.</i> , 2022 [348]	DL-based	96.24%
D. Wang <i>et al.</i> , 2021 [347]	DL-based	99.80%
L. Hou <i>et al.</i> , 2020 [349]	DL-based	99.70%

deep learning (DL)-based methods on the PU dataset. The results demonstrate the effectiveness of the proposed method in achieving excellent performance compared to DL-based methods. Moreover, the proposed method is more computationally efficient than DL-based approaches since it utilizes a very short duration of the acquired vibration signal (only 0.1 seconds) and produces six features that are sufficient for fault detection and diagnosis. On the other hand, DL-based approaches usually require extensive training and result in substantial sizes of deep-learned features that necessitate the use of dimensionality reduction feature ranking techniques, which would further increase the computational burden.

6.6 Conclusion

In this chapter, a new method is proposed for vibration-based condition monitoring of rolling bearings that effectively utilizes instantaneous frequency along with the envelope of generated vibration patterns to jointly facilitate three newly introduced envelope representations: Instantaneous Amplitude-Frequency Mapping (IAFM), Instantaneous Amplitude-Frequency Correlation (IAFC), and Instantaneous Energy-Frequency Distribution (IEFD). The introduced representations effectively capture unique variations in energy and frequency between healthy and faulty bearings, thereby enabling the extraction of discrim-

inative features with high sensitivity to variations in operational conditions. Accordingly, a set of six new highly discriminative features are extracted from these representations. The experimental results demonstrated excellent performance in detecting and diagnosing various fault types, marking the effectiveness of the proposed method in capturing unique variations in energy and frequency between healthy and faulty bearings. Moreover, the proposed method has comparable performance to DL-based methods but with more moderate computational requirements attributed to the short duration of the utilized vibration segments, efficient feature extraction, and the small set of resulting features.

Chapter 7

7 Signal-Companding AutoEncoder with Compression-Based Activation for an Efficient Vibration-Based Remote Condition Monitoring

This chapter¹ introduces the fourth research problem within the thesis, focusing on the enhancement of power efficiency and reliability of remote VBCM systems. Specifically, the chapter tackles the problem of power efficiency from a signal's waveform perspective and suggests the control of the Peak-to-Average Power Ratio (PAPR) of the acquired vibration signal in the sensor node to reduce the power consumption. Further, an autoencoder-based method for signal denoising is proposed to improve the reliability of the condition monitoring process. The content of this chapter is submitted to the *IEEE Transactions on Reliability* as a regular paper entitled "Signal-Companding AutoEncoder with Compression-Based Activation for a Power-Efficient Vibration-Based Remote Condition Monitoring". The paper is currently in the review stage.

7.1 Introduction

In remote Vibration-Based Condition Monitoring (VBCM) systems, sensor nodes are constrained by limited power resources, necessitating efficient power management to reduce power consumption within the node. Further, maintaining signal's waveform undistorted is crucial for achieving reliable performance in condition monitoring. This chapter analyzes the PAPR in VBCM systems and evaluates its effects in the presence of nonlinear power amplification on system performance and power efficiency. Aiming to enhance power efficiency and improve performance, the chapter proposes a lightweight autoencoder-based

¹A version of this chapter has been submitted for publication in *IEEE Transactions on Reliability*.

signal companding scheme to control the PAPR and avoid nonlinear distortion. In the sensor node, a lightweight reconstruction autoencoder with a compression-based activation function compresses the acquired vibration signal without increasing its average power. In the processing center, a denoising-expansion autoencoder simultaneously denoises and expands the compressed signal, avoiding noise expansion. Experimental results show this approach effectively prevents distortion, enhances power efficiency, and improves overall system performance.

The rapid evolution of sensor fabrication, coupled with advancements in the Internet of Things (IoT) and computing technologies, has enabled the facilitation of large-scale remote VBCM systems comprising distributed sensor nodes. In these systems, sensor nodes are typically battery-powered and, therefore, have limited power resources [379–383]. Hence, efficient power utilization in key node components, such as signal acquisition, amplification, and transmission, is paramount to maintaining low power consumption. Aiming to reduce power consumption, current research efforts are mainly dedicated to developing power-efficient signal acquisition techniques. Specifically, compressive sensing [384–389] is being adopted to acquire the signal in a compressed form by performing fewer measurements and collecting fewer samples. The ultimate objective is to achieve more power-efficient signal acquisition with less power consumption. Compressive sensing allows the reconstruction of the signal from a few acquired samples. Relying on the sparsity of the acquired signal, reconstruction algorithms utilize sparse optimization to reconstruct the signal from the acquired samples. As a result, the practical use of compressive sensing is limited by the assumption of signal sparsity and costly reconstruction processes [388] that involve time and power-consuming algorithms, making compressive sensing unsuitable for real-time condition monitoring [389].

This chapter tackles the problem of power efficiency from a signal waveform perspective. Specifically, the chapter suggests reducing the power consumption in the sensor nodes by

controlling the PAPR of the acquired vibration signal. The PAPR is the ratio of the peak power to the signal's average power. It directly affects the node's power consumption since it determines the required resolution for analog-digital conversions [390]. Additionally, it determines the required linear range of the power amplification circuit [391], which accounts for the major part of the total power consumption in many systems [392, 393]. To the best of our knowledge, this chapter is the first work that addresses the issue of PAPR in vibration signals and tackles the related problem of nonlinear power amplification in VBCM systems. Specifically, the chapter statistically investigates the PAPR characteristics of vibration signals, evaluates the impact of nonlinear power amplification on the system, and proposes a lightweight framework based on signal companding to reduce the PAPR and ensure linear power amplification of the signals. Companding¹ is a well-known technique in signal processing; it involves signal compression at the source and subsequent expansion at the destination. Signal companding has demonstrated its effectiveness in controlling the PAPR of multi-carrier communication signals, such as Orthogonal Frequency-Division Multiplexing (OFDM). Nevertheless, conventional companding techniques encounter two significant limitations. Firstly, the compression mechanism increases the average power of the compressed signal. Secondly, the expansion operation amplifies the accumulated noise in the compressed signal. To effectively control the PAPR and address these limitations, the proposed framework adopts a two-fold approach. Firstly, it smoothes and compresses the signal using a reconstruction autoencoder with a compression-based activation function. Secondly, it employs a denoising-expanding autoencoder to simultaneously denoise and expand the compressed noisy signal at the destination. This combined approach ensures efficient PAPR control while mitigating the aforementioned limitations. Additionally, the proposed autoencoder structure facilitates an efficient, end-to-end, Deep Learning (DL)-based implementation of the signal processing pipeline in the framework. Specifically, at the source, the smoothing and compression operations are carried out simultaneously

¹The name COMPANDING is a composite of the words COMPRESSing and expANDING.

using a smoothing autoencoder with a compression-based activation function. The simultaneous denoising-expansion of the compressed noisy signal at the destination is achieved through a denoising-reconstruction autoencoder. The main contributions of the chapter include:

- To the best of our knowledge, this chapter is the first contribution to the VBCM literature that addresses the PAPR of generated vibration waveform and examines its impact in the presence of nonlinear power amplification. Accordingly, it proposes controlling the PAPR to enhance power efficiency, mitigate nonlinear distortion, and improve the reliability of condition monitoring.
- Statistically analyzes the PAPR of vibration signals and introduces a closed-form formula that accurately models the statistical distribution of the PAPR.
- Introduces a framework based on signal companding to effectively reduce the PAPR of vibration signals and mitigate the impact of nonlinear power amplification.
- In the sensor node “source”, before the power amplification stage, the framework employs a lightweight reconstruction autoencoder that utilizes a compression-based activation function. The autoencoder function facilitates the simultaneous smoothing and compression of the vibration signal without causing an increase in the average power of the compressed signal.
- In the processing center “destination”, the proposed framework utilizes a denoising-expansion autoencoder to simultaneously denoise and expand the compressed signal while avoiding enhancement “expansion” of the accumulated noise by the expansion operation.
- The proposed architecture, based on reconstruction autoencoders and compression-based activation, allows for simultaneous signal processing, facilitating an efficient end-to-end implementation of the framework.

- Comprehensively evaluates the performance of the proposed framework in the presence of nonlinear power amplification and Additive White Gaussian Noise (AWGN), employing a real-world vibration dataset.
- Adapts new metrics to quantify nonlinear distortion caused by nonlinear power amplification and evaluate power efficiency.

The remainder of the chapter is structured as follows: The next section provides background information and motivation for the problem. Section 7.3 presents the statistical analysis of the PAPR of vibration signals. Section 7.4 reviews signal companding techniques that have been proposed in the OFDM literature. Section 7.5 introduces the proposed autoencoder-based companding framework. Section 7.6 presents the model of nonlinear power amplification used in the experimentation. The experimental setup and performance evaluation metrics are introduced in Section 7.7, while Section 7.8 discusses the obtained results. The chapter is finally concluded in Section 7.9.

7.2 Background and Motivation

Fig. 7.1 shows a high-level architecture of a typical remote VBCM system where sensor nodes are deployed across various locations, either embedded in objects, placed beneath surfaces, attached to mobile or airborne objects and connected to a cloud or a processing center. Typically, sensor nodes employ wireless connections, utilizing existing cellular networks or using dedicated wireless links to aggregator nodes in scenarios where cellular coverage is unavailable or unstable [8, 379, 380, 394–397]. Subsequently, the aggregator node transmits the accumulated signals to the cloud or the processing center via the cellular network. As mentioned earlier, these nodes are typically power-constrained, and hence, efficient power utilization in key components of the node, such as signal acquisition, amplification, and transmission, is critical to maintaining low power consumption in these nodes. The PAPR is crucial in determining how much power the system needs to operate

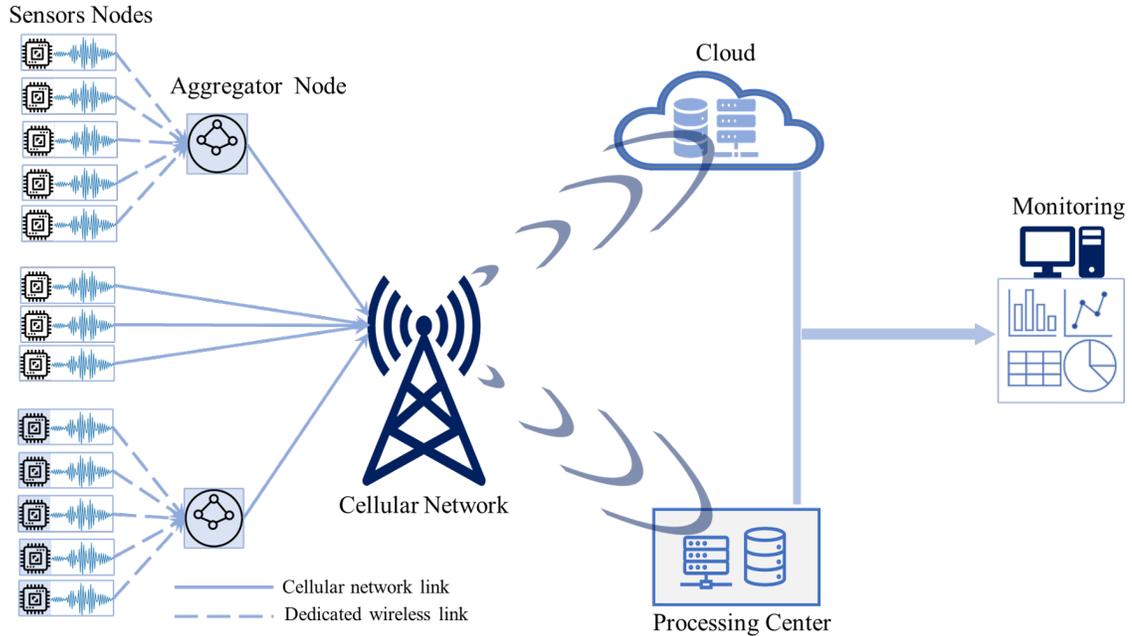


Figure 7.1: Overview of a typical remote VBCM system.

effectively. A smaller PAPR value requires fewer bits and allows the High Power Amplifier (HPA) to operate more efficiently, saving the battery in the system [398]. To achieve maximum power efficiency, the HPA's operating point should be positioned as close as possible to HPA's saturation point [399] as illustrated in Fig 7.2. When peaks of the input signal exceed this designated operating point, the HPA becomes prone to saturation, leading to power wastage, nonlinear amplitude distortion, and spectral spreading induced by abrupt fluctuations in the distorted amplitudes. To prevent these consequences, the HPA circuit must be designed to operate linearly over the PAPR range of the input signal, which tends to be a costly and inefficient solution [400]. Alternatively, a significant Input Power Backoff (IBO) from the HPA's operating point should be applied to restrict the HPA's input power level, ensuring that the entire signal falls within the HPA's linear region. While this approach mitigates nonlinear distortion, it significantly reduces power efficiency, as the HPA operates in a lower-power region. Therefore, it has a high cost in terms of energy efficiency, particularly in battery-power applications [399] such as remotely deployed sensor

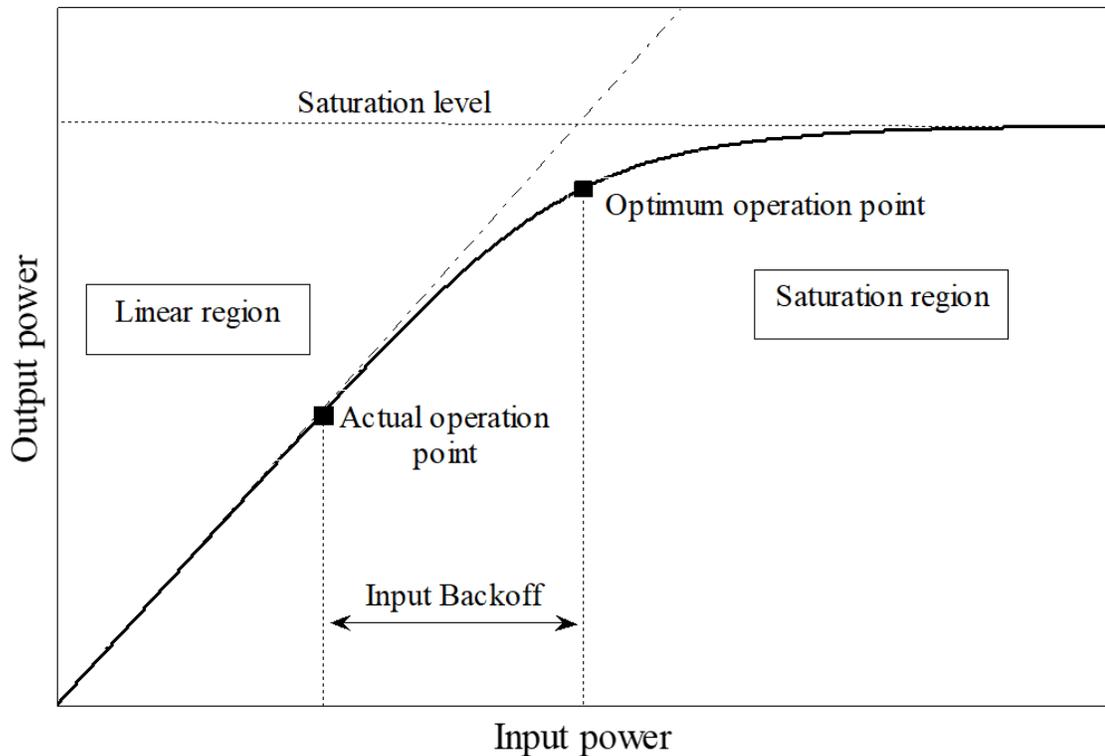


Figure 7.2: HPA range curve.

nodes.

In VBCM systems, the amplitude of the acquired vibration waveform fluctuates according to the monitored condition. Hence, the waveform is anticipated to exhibit a high PAPR due to these fluctuations, which can reach significant magnitudes depending on the monitored condition [8, 394]. To this end, the remainder of the chapter attempts to facilitate power-efficient remote VBCM by addressing the following key aspects:

1. Analyse PAPR characteristics of acquired vibration signals,
2. evaluate the impact of uncontrolled PAPR on the VBCM performance in the presence of nonlinear power amplification,
3. and, accordingly, propose an appropriate remedy solution to control the PAPR.

7.3 PAPR of Vibration Signals

The PAPR quantifies the power ratio between the peak and average amplitudes of a signal.

For a given vibration signal $x(t)$, its PAPR can be expressed as:

$$PAPR_{x(t)} = \frac{\text{Max } |x(t)|^2}{E \{|x(t)|^2\}} \quad (7.1)$$

where $E \{\cdot\}$ denotes the expectation operator. For the finite sampled signal $x[n]$, the PAPR is:

$$PAPR_{x[n]} = \frac{\text{Max}_{n \in [0, N]} |x[n]|^2}{\frac{1}{N} \sum_0^{N-1} |x[n]|^2} \quad (7.2)$$

where N is the number of samples in the vibration signal $x[n]$. The PAPR is usually expressed in dB :

$$PAPR (dB) = 10 \times \log (PAPR) \text{ dB} \quad (7.3)$$

Crest Factor (CF) is another common signal parameter that quantifies a signal's peak amplitude to its Root-Mean-Square (RMS) value. It equals the square root of the PAPR. However, expressed in dB , the CF is equal to the PAPR since:

$$CF (dB) = 10 \times \log (CF^2) \text{ dB} = 10 \times \log (PAPR) \text{ dB} \quad (7.4)$$

A signal with constant power, such as a square wave, has a PAPR of 1 (0 dB). The PAPR of a sinusoidal wave equals 2 dB or 3.01 dB . Determining the PAPR of a random vibration depends on its instantaneous value, which is not predictable. Nevertheless, it is possible to describe the PAPR statistically.

7.3.1 Statistical Distribution of Vibration Signals

The statistical distribution of the vibration samples generated by a VBCM system depends on the characteristics of the monitored condition and is influenced by the surrounding envi-

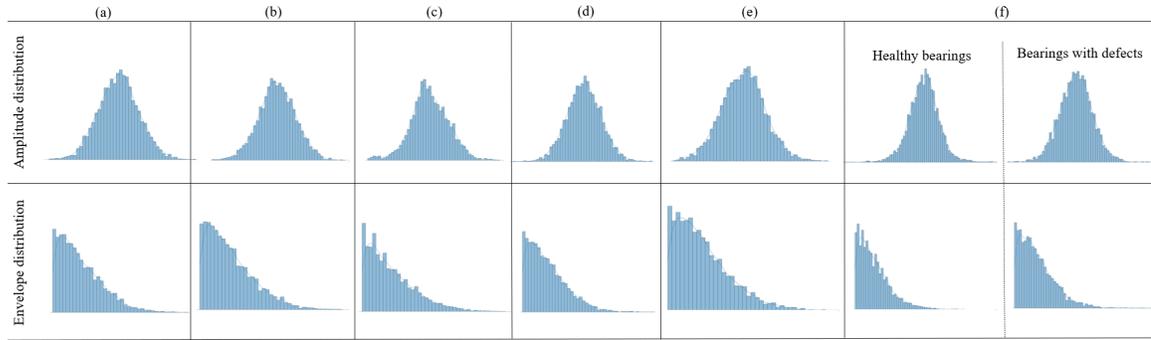


Figure 7.3: Histograms of sample vibration signals: (a) Gaussian random vibration, (b) acceleration of a flying aircraft, (c) acceleration measurements of a flying UAV, (d) vibration from an SHM setup, (e) vibration generated by a wind turbine gearbox, and (f) vibration generated by rolling bearings.

ronment. However, when the number of vibration samples N in the signal is large, the signal will approach the Gaussian distribution with a zero mean and a variance of σ^2 “central-limit theorem.” Therefore, a Gaussian random process can accurately model the vibration signal. Accordingly, the signal’s envelope $|x[n]|$ follows the one-sided Gaussian distribution. Fig. 7.3 shows histograms of vibration signals generated from (a) Gaussian random vibration, (b) exterior of a flying aircraft [401], (c) a flying Unmanned Aerial Vehicle (UAV) [402], (d) a Structural Health Monitoring (SHM) setup [403], (e) a wind turbine gearbox [404], and (f) rolling bearings of a rotating machinery[405]. These vibration sets are chosen to resemble the vibration patterns found in various VBCM applications. They represent healthy or normal vibrations, except vibrations of the rolling bearings (Fig. 7.3.f), which include normal and faulty vibrations. It is worth mentioning that an abnormal operation or a failure influences the instantaneous vibration, and hence, it would alter the amplitude distributions of the generated vibration. Although generalizing the aforesaid assumption of Gaussian nature may not be entirely accurate, the histograms presented in Fig. 7.3 show that this assumption would be valid for a broad range of vibration patterns.

7.3.2 Statistical Analysis of the PAPR

Following the assumption that a vibration signal $x[n]$ follows a Gaussian distribution, its Probability Density Function (PDF) can be expressed as:

$$p(x, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \times \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (7.5)$$

Accordingly, the signal's envelope $|x[n]|$ has a one-sided Gaussian distribution; its PDF is given by:

$$p_e(x, \sigma) = \sqrt{\frac{2}{\pi\sigma^2}} \times \exp\left(-\frac{x^2}{2\sigma^2}\right), x \geq 0 \quad (7.6)$$

The Cumulative Distribution Function (CDF) of the signal's envelope is then obtained by:

$$F_e(x, \sigma) = \int_0^x \frac{\sqrt{2}}{\sigma \sqrt{\pi}} \times \exp\left(-\frac{u^2}{2\sigma^2}\right) du \quad (7.7)$$

using

$$t = \sqrt{\frac{u^2}{2\sigma^2}}, \quad (7.8)$$

the CDF can be written as:

$$\begin{aligned} F_e(x, \sigma) &= \frac{2}{\sqrt{\pi}} \int_0^{\sqrt{x^2/2\sigma^2}} \exp(-t^2) dt \\ &= \operatorname{erf}\left(\sqrt{\frac{x^2}{2\sigma^2}}\right) \end{aligned} \quad (7.9)$$

where $\operatorname{erf}(\cdot)$ is the error function. Accordingly, the probability that the signal's power ratio $P = \frac{x^2}{\sigma^2}$ is above a given PAPR threshold P_o can be obtained using the Complementary Cumulative Distribution Function (CCDF):

$$\begin{aligned} \operatorname{Prob}(P > P_o) &= \text{CCDF} = 1 - (\text{CDF})^N \\ &= 1 - \operatorname{erf}\left(\sqrt{\frac{P_o}{2}}\right)^N \end{aligned} \quad (7.10)$$

where N is the number of samples in the vibration signal $x[n]$. The analytical formula of the CCDF in (7.10) is helpful in studying the PAPR of vibration generated in various

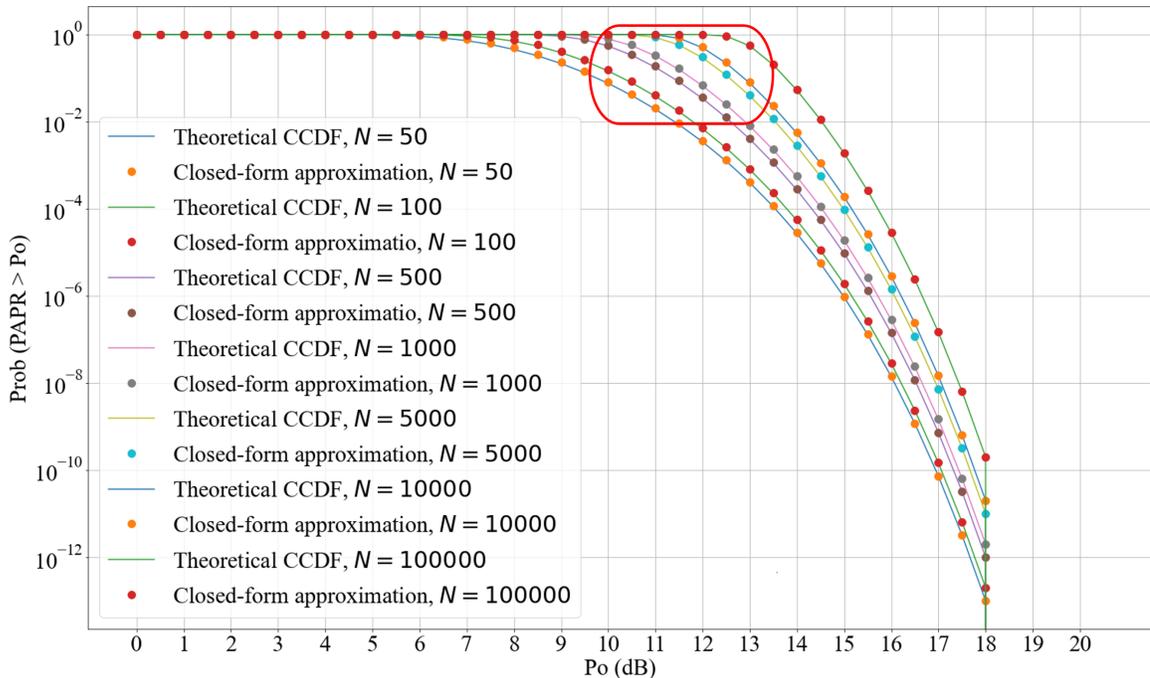


Figure 7.4: Theoretical CCDF and its closed-form approximation for different values of N .

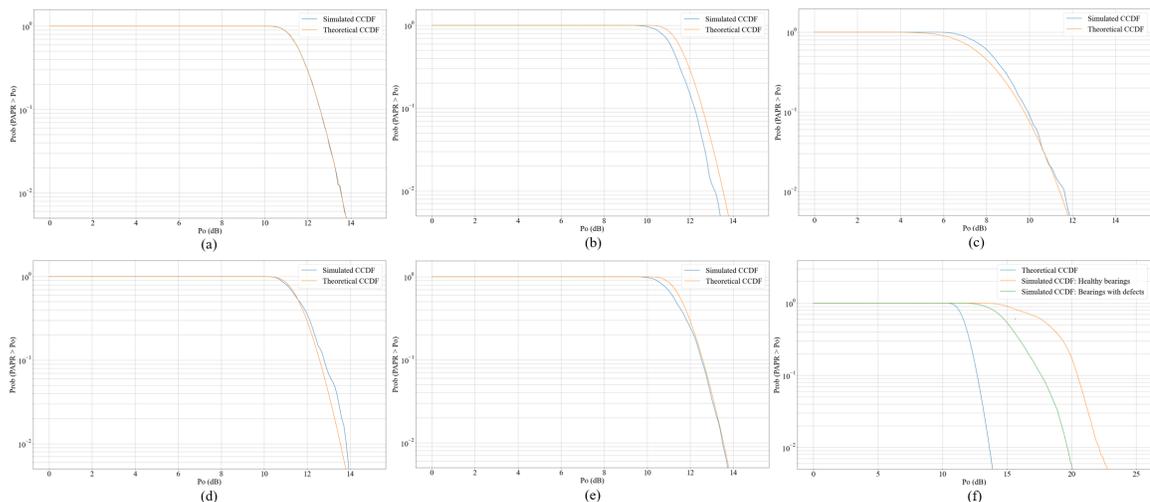


Figure 7.5: Analytical and simulated CCDFs of (a) Gaussian random vibration ($N = 5000$), (b) acceleration of a flying aircraft ($N = 5000$), (c) acceleration measurements of a flying UAV ($N = 50$), (d) vibration from an SHM setup ($N = 5000$), (e) vibration generated by a wind turbine gearbox ($N = 5000$), and (f) vibration generated by rolling bearings ($N = 5000$).

VBCM systems. Furthermore, the CCDF is a useful metric for evaluating the effectiveness of the employed PAPR reduction method. Typically, a simulated CCDF is obtained using PAPR-reduced signals and compared to a simulated CCDF of the original signals to evaluate the reduction achieved in the PAPR. A closed-form approximation of (7.10) can be obtained using the asymptotic series expansion of the complementary error function $erfc(\cdot)$:

Since,

$$erf(\cdot) = 1 - erfc(\cdot), \quad (7.11)$$

(7.10) can be expressed in terms of $erfc(\cdot)$ as follows:

$$Prob(P > P_o) = \text{CCDF} = 1 - \left(1 - erfc\left(\sqrt{\frac{P_o}{2}}\right)\right)^N \quad (7.12)$$

For large values of $\sqrt{P_o/2}$, the complementary error function may be approximated by the asymptotic series expansion:

$$erfc\left(\sqrt{\frac{P_o}{2}}\right) \approx \frac{e^{-P_o/2}}{\sqrt{\frac{P_o\pi}{2}}} \times \left(1 - \frac{1}{P_o} + \frac{1 \cdot 3}{P_o^2} - \dots + (-1)^n \frac{(2n-1)!!}{P_o^n} + \dots\right) \quad (7.13)$$

For $P_o \gg 1$,

$$erfc\left(\sqrt{\frac{P_o}{2}}\right) \approx \frac{e^{-P_o/2}}{\sqrt{\frac{P_o\pi}{2}}} \quad (7.14)$$

Accordingly, a closed-form approximation of the CCDF can be obtained by substituting (7.14) into (7.12):

$$Prob(P > P_o) = \text{CCDF} = 1 - \left(1 - \frac{e^{-P_o/2}}{\sqrt{\frac{P_o\pi}{2}}}\right)^N \quad (7.15)$$

Fig. 7.4 shows plots of the CCDF in (7.10) and its closed-form approximation in (7.15) for

different values of N . The plotted CCDF curves show an exact match between the CCDF and its closed-form approximation. Fig. 7.5 shows simulated CCDFs of the aforementioned vibration sets (refer to Fig. 7.3) along with theoretical CCDFs (7.15). It is evident that the simulated CCDFs align with their corresponding theoretical CCDFs, except for the rotating machinery. The mismatch in the case of rotating machinery could be due to the rotating nature of the bearings and speed fluctuations [394][406]. Additionally, the graphs depicted in Fig. 7.4 and Fig. 7.5 demonstrate that as the number of samples N increases, the likelihood of experiencing a high PAPR increases. Specifically, with the number of samples $N \geq 500$, a PAPR in the range of 10 dB-13 dB is likely to occur. Thus, it can be concluded that vibration signals generally tend to have high PAPR, where peak vibrations that are 10–20 times higher than the average vibrations occur commonly.

7.4 Review of PAPR reduction techniques

To the best of our knowledge, the PAPR of vibration signals and the associated problem of nonlinear power amplification have not been addressed yet in the literature. Nevertheless, the problem of PAPR in multi-carrier communications, particularly in OFDM systems, has been extensively studied since OFDM signals exhibit a high PAPR. The existing techniques can be broadly categorized into three main categories: Symbol structure modification, peak clipping, and signal companding. Structure modification techniques include block coding [407], Selective Mapping (SLM) [408], Partial Transmission Sequence (PTS) [408], and tone reservation [409]. These techniques reduce the PAPR by modifying the structure of the transmitted OFDM symbol. They generally impose restrictions on its parameters and require transmitting side information to reconstruct the symbol at the destination. Therefore, the reduction in PAPR comes at the cost of increased complexity and reduced data rates due to the transmission of side information. Clipping [410] offers a simple approach to reducing the PAPR by hard-limiting the peaks to a pre-defined threshold. Despite its

simplicity, clipping introduces amplitude distortion and spectral spreading. While amplitude distortion is unrecoverable, filtering would reduce spectral spreading. However, the peaks of the filtered-clipped signal could exceed the clipping threshold due to peak power regrowth after filtering. Alternative solutions that help to reduce the clipping distortion involve repeated or iterative clipping [410] and peak windowing [411]. In contrast to clipping, peak windowing applies soft-limiting to the peaks by multiplying the signal with a window-weighting function. As a result, distortion is reduced since the peaks are smoothly and softly limited.

Signal companding is a well-known method in signal processing that involves two steps. First, the signal is transformed into a compressed form at the source. Second, the inverse transform expands the compressed signal at the destination. The compression reduces the signal's dynamic range and allows for efficient signal processing. μ -law and A-law [412] are the most common companding transforms typically applied to speech signals to reduce quantization noise and optimize the required number of bits per sample for analog-to-digital conversion. In μ -law and A-law transforms, signal compression is achieved by applying a logarithmic-based transform to enlarge small amplitudes in the signal. Companding has no restrictions on the symbol's parameters and does not require the transmission of side information. Further, it has better error performance compared to clipping. Using signal companding to reduce the PAPR of OFDM signals was first introduced in [375]. However, the scope of analysis was limited to addressing the effect of μ -law companding on the quantization noise. Reducing the PAPR of the signal by applying μ -law companding will increase the signal's average power. This, in turn, improves the signal-to-quantization noise ratio since the small amplitudes are enlarged. However, considering the nonlinearity of the HPA, reducing the PAPR by increasing the signal's average power will not prevent the nonlinear distortion since the large peaks are not reduced. In fact, it would lead to more distortion in the signal. The main attention of the ongoing research is directed toward addressing this problem by designing the companding function so that the increase

in the signal's average power is avoided [350, 351, 376–378, 413]. The published work in this area can be grouped under two main approaches. The first approach involves using additional transforms and/or optimization algorithms, which obviously increases computational complexity. The second approach involves introducing inflexion points in the signal. This allows for independent scaling of large peaks and small amplitudes, which helps to maintain the signal's average power. However, this approach reduces the data rate since the signal's indexes must be transmitted to apply the inverse operations at the destination. A considerable amount of the recent work focuses on utilizing DL to tackle the problem of PAPR in OFDM [121–123, 414–418]. DL-based approaches are centered around designing and training the DL models to optimally, sub-optimally, or efficiently learn the function of the corresponding conventional PAPR reduction scheme while mitigating the associated drawbacks.

Choosing the appropriate technique among the options mentioned above for reducing the PAPR of vibration signals starts by understanding the distinctions between OFDM and vibration signals. Regarding structure modification approaches, vibration signals are generated by sensors as raw data, unlike OFDM symbols, which are formed based on a pre-determined structure. Therefore, such techniques are not applicable to vibration signals. Clipping introduces unrecoverable distortion in the clipped signal; this can be tolerated in the OFDM signal due to the error correction mechanisms. In VBCM systems, the characteristics of the monitored condition are described by the waveform and the spectrum of the generated vibration signal. This makes clipping distortion critical and intolerable since it introduces distortion in the generated waveform and alters its spectral contents. Compared to structure modification and clipping, signal companding presents a practical solution for reducing the PAPR of vibration signals without impacting the condition monitoring process. However, in order to be adopted for VBCM applications, the companding transform should fulfill the following three requirements:

1. Avoiding the increase in the signal's average power.
2. Denoising the compressed signal to avoid expanding the noise during the expansion of the compressed signal.
3. Avoiding transmission of side information as this will increase the amount of the transmitted data, resulting in more power consumption in the sensor node.

Considering these requirements, the upcoming section introduces the proposed autoencoder-based companding framework.

7.5 Signal Companding for Reduction of PAPR

This section introduces the lightweight companding-based framework proposed for reducing the PAPR of vibration signals. However, it is convenient first to provide a brief overview of conventional signal companding.

7.5.1 Conventional Signal Companding

The most commonly used type of signal companding is the μ -law companding. its compression function $C(x)$ can be expressed as:

$$y = C(x) = A \operatorname{sgn}(x) \frac{\ln(1 + \mu |x/A|)}{\ln(1 + \mu)} \quad (7.16)$$

where, x is the input signal, $\operatorname{sgn}(\cdot)$ is the sign function, A is a normalization constant such that $0 < |x/A| < 1$, and μ is the compression parameter. The expansion (inverse) function is expressed as:

$$\begin{aligned} x' &= C^{-1}(y) \\ &= A \left[\frac{\exp \left\{ \frac{|y|}{A \operatorname{sgn}(y)} \ln(1 + \mu) \right\} - 1}{\mu \operatorname{sgn}(y)} \right] \end{aligned} \quad (7.17)$$

Fig. 7.6.a displays the compression profile of μ -law for different values of compression pa-

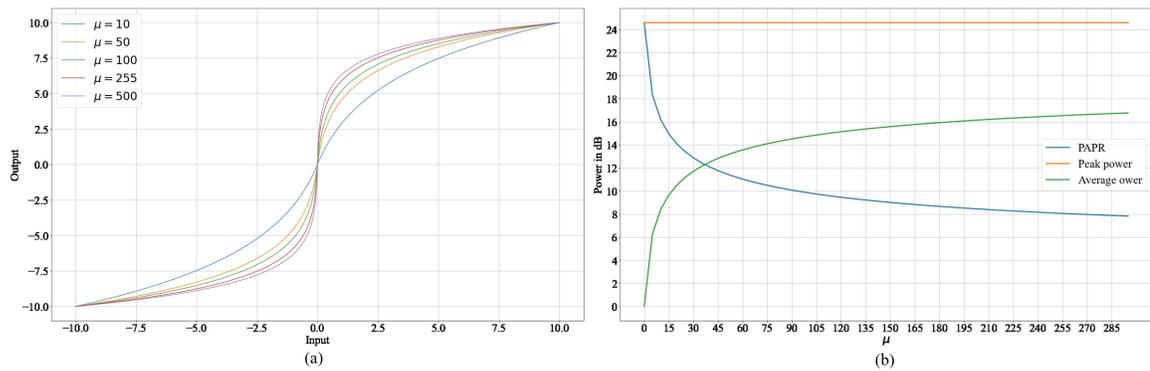


Figure 7.6: μ -law: (a) compression profile with different values of μ , and (b) peak and average powers as a function of μ .

parameter μ . Increasing the μ value leads to more enlargements of small amplitudes, resulting in a higher average power of the signal. Hence, the signal's average power increases as a function of μ as illustrated in Fig. 7.6.b. Since the signal's peaks are maintained unchanged in μ -law companding, the reduction in the PAPR of the signal is achieved solely by increasing its average power. However, to avoid nonlinear distortion in the vibration signal and improve the power efficiency of the HPA, it is required to reduce the PAPR by reducing the signal's peaks instead of increasing its small amplitudes. In other words, it is required to reduce the PAPR of the signal while avoiding any increase in its average power. Another issue with conventional companding is the undesired effect of enhancing the accumulated noise at the destination due to expansion operation [375]-[351]. Thus, it is crucial to reduce the effects of noise enhancement by applying effective denoising to the compressed signal prior to expanding it.

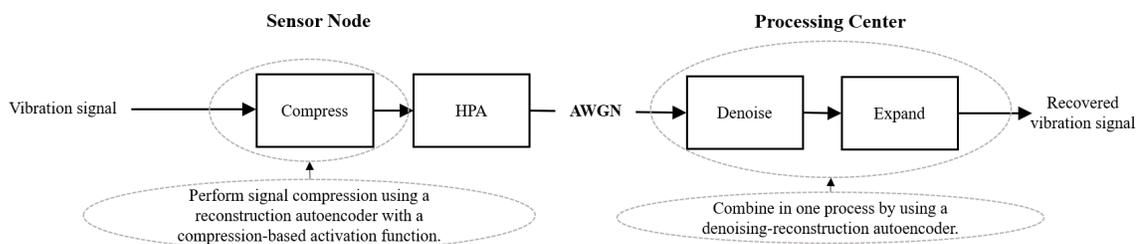


Figure 7.7: Proposed framework for autoencoder-based companding of vibration signals.

7.5.2 Proposed Signal Companding

The main aspects of the proposed framework are illustrated in Fig. 7.7. Specifically, signal compression at the source is achieved using a lightweight reconstruction autoencoder with a compression-based activation function. At the destination, signal denoising and expansion operations are combined in one process by using a denoising-reconstruction autoencoder.

Signal Compression at the Source

In the first place, the raw vibration signals are smoothed to remove measurement noise. Then, a reconstruction autoencoder is trained to learn the smoothing function and reconstruct these smoothed signals “target signals”. By using a compression-based activation function in the autoencoder layers, the autoencoder will learn how to reconstruct the input signals based on the target signal and, at the same time, compress the learned presentations of the input signal. As a result, the output of the trained autoencoder is a smoothed and compressed version of the input signal.

Average power of the compressed signal: During the training process, the autoencoder compresses the signal’s representations in each layer while, at the same time, it learns to minimize the loss between the input signal and the target signal. Here, the average power of the target signal represents an upper bound on the average power of the reconstructed signal. This will avoid any increase in the average power of the output reconstructed (smoothed-compressed) signal. Further, the joint mechanism of reconstruction-compression will maintain the average power of the output signal as close as possible to the average power of the target signal.

Compression loss as a lower bound on the validation loss: To efficiently train the source autoencoder, it is important to consider the following factors:

- The training objective of the source autoencoder is to reconstruct a smoothed and

compressed version of the input signal rather than reconstructing the signal in its original form.

- The compression-based activation function in the autoencoder layers implies that there will be a compression loss or a minimum error floor between the output signal and the target signal caused by the compression mechanism.

Considering these factors, it is not required to minimize the loss until maximum convergence. Instead, it is desirable to train the autoencoder until reaching this error floor, which is determined by compression loss. Theoretically, the compression loss can be defined as the difference between the signal and its “perfectly reconstructed” compressed form. Mathematically, the Compression Loss (CL) can be calculated as the difference between a vibration signal x and its compressed and power-preserved form x_{pc} :

$$\begin{aligned}
 CL &= error(x, x_{pc}), \\
 x_{pc} &= P(AF(x)),
 \end{aligned}
 \tag{7.18}$$

where AF is the compress-based activation function, and P is the power scaling operation. The *error* function can be either the Mean Absolute Error (MAE) or the Mean Squared Error (MSE). To efficiently train the autoencoder, the compression loss can be utilized to set a baseline for the validation loss during the training. It can be empirically determined using the training signals to obtain the average CL according to (7.18). Generally, the effect of compressing a signal and preserving its average power can be approximated by applying a limiter to the signal with a peak-limiting threshold equal to the maximum peak of its compressed form. Accordingly, the clipping noise, which is the power of the clipped portion, can be used as an estimate of the compression loss. Given a target PARR ($PAPR_t$), the maximum peak $Peak_c$ of the compressed and power preserved signal x equals to:

$$Peak_c = \sqrt{PAPR_t \times P_{in}},
 \tag{7.19}$$

where P_{in} is average power of the signal. Following the assumption of the Gaussian nature of x , it can be modeled as a Gaussian random process with a zero mean and a variance $\sigma^2 = P_{in}$. Thus, the probability that, at any given time, the signal x takes the value $Peak_c$ is given by:

$$\begin{aligned} Prob\{x(t) = Peak_c\} \\ = p(x) = \frac{1}{\sqrt{2\pi P_{in}}} \times \exp\left(-\frac{x^2}{2P_{in}}\right) \end{aligned} \tag{7.20}$$

Since maximum peak of the signal is limited to $Peak_c$, the clipping noise (CN) is given by:

$$CN = 2 \int_{Peak_c}^{\infty} (x - Peak_c)^2 p(x) dx \tag{7.21}$$

Using the analysis presented in [391], CN can be approximated as:

$$CN \cong 2 \sqrt{\frac{2}{\pi}} \times \sigma^2 \times (\sqrt{PAPR_t})^{-3} \times \exp\left(-\frac{PAPR_t}{2}\right) \tag{7.22}$$

In contrast to other companding techniques [350, 351, 376–378, 413] that utilize complex signal processing to compress the signal while preserving its average power, the proposed compression method streamlines compression and power-preserving processes, in addition to signal smoothing, into a single online inference task. Furthermore, bounded by the

Table 7.1: Comparison between the proposed method and conventional training

Method			Online Signal Processing		
			Smoothing	Compressing	Power Preservation
Conventional Signal processing-based companding			✓	✓	✓
Proposed Activation-based autoencoder			✗	✗	✗
Method	Activation	Training loss	Signal Processing for Preparation of Target Signals		
			Smoothing	Compressing	Power Preservation
Conventional autoencoder-based companding	Conventional function	Minimize until convergence	✓	✓	✓
Proposed Activation-based autoencoder	Compression-based function	Minimize until reach baseline	✓	✗	✗

compression loss, the proposed method requires less training time compared to traditional minimum-error training, making it more efficient than other autoencoder-based companding methods [122, 414–417] that employ conventional training methods. Table 7.1 provides a high-level comparison between the proposed method and both conventional signal processing-based companding and conventional autoencoder-based companding. Conventional autoencoder-based companding methods use traditional activation functions and involve training the autoencoder to minimize the loss between reconstructed output signals and training target signals that are smoothed, compressed, and power-scaled versions of the input training signals. The comparison demonstrates the lightweight nature of the proposed method, which requires fewer computations and reduced signal processing, making it more power-efficient and less complex than conventional methods.

Signal Denoising and Expansion at the Destination

In contrast to the source autoencoder, the training objective of the destination autoencoder is minimizing the validation loss until maximum convergence. The destination autoencoder is trained in reverse order compared to the source autoencoder. First, the training targets of the destination autoencoder are used as inputs to the **trained** source autoencoder. Accordingly, the output compressed-smoothed signals obtained from the source autoencoder are corrupted with AWGN noise at a desired Signal-to-Noise Ratio (SNR) and used as the training inputs for the destination autoencoder. The autoencoder is then trained using the input and target signals to minimize the validation loss. This way, the autoencoder will learn the expansion mechanism. Additionally, its weights will be tuned to remove the noise. Hence, the autoencoder simultaneously acts as an expanding function and a denoising filter.

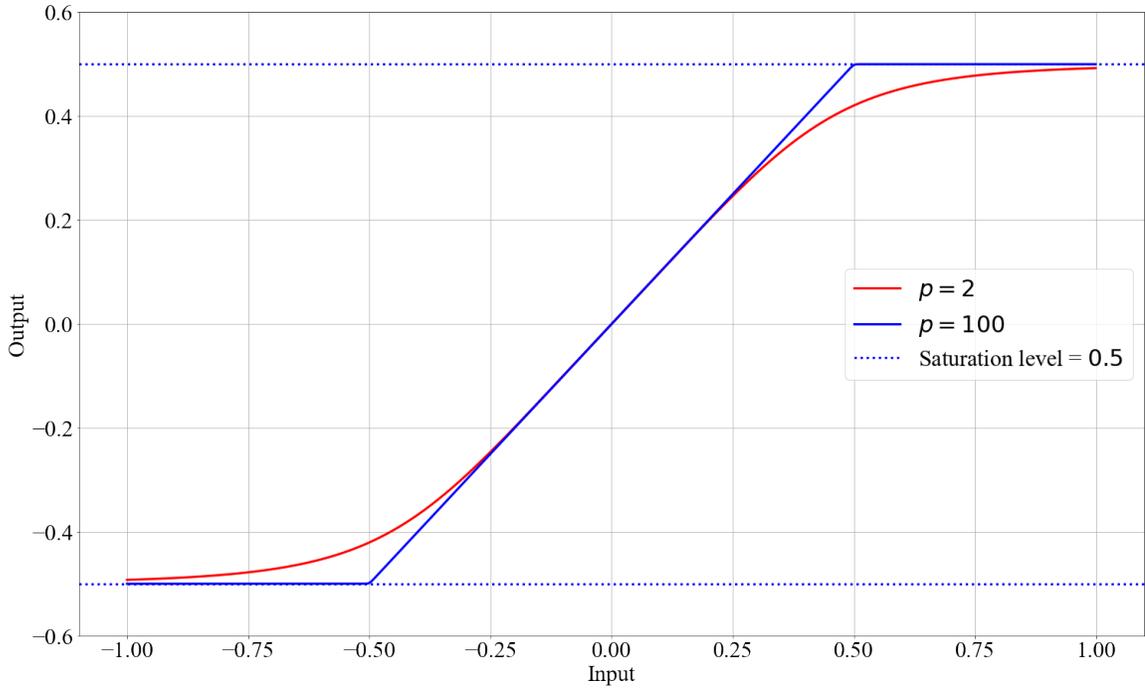


Figure 7.8: AM-AM and AM-PM conversions of the Rapp SSPA model.

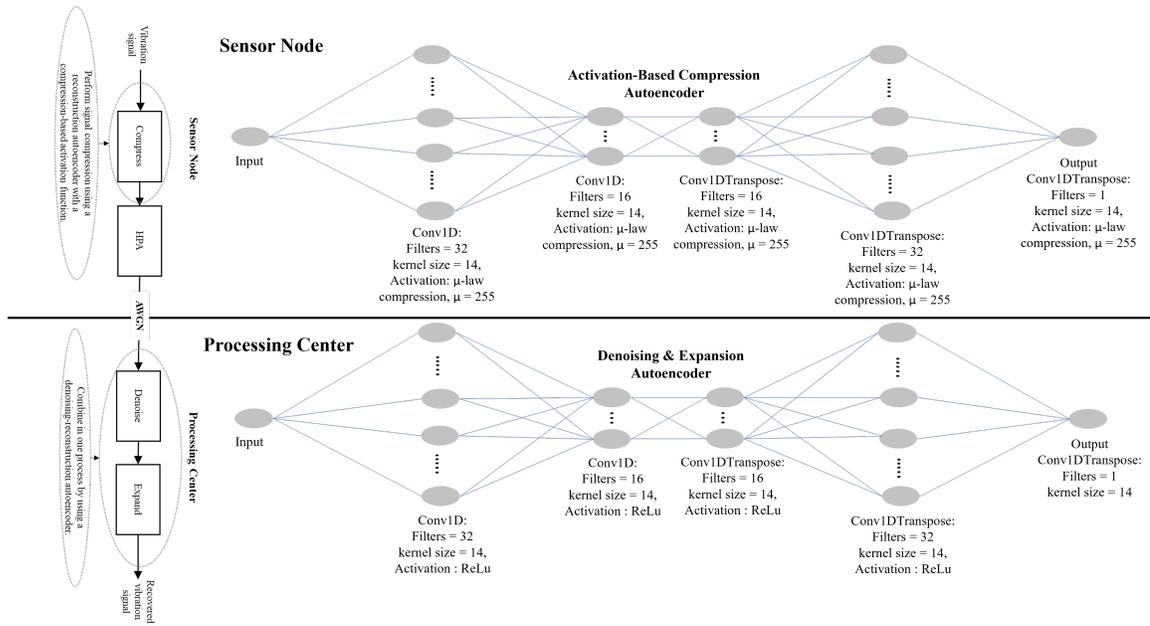


Figure 7.9: Structures of the reconstruction autoencoders used in the proposed framework.

7.6 Modeling Nonlinear Power Amplification

A practical HPA has a limited linear range and exhibits a nonlinear behavior at its saturation point, as deposited in Fig. 7.2. Reliable modeling of the HPA is crucial for the accurate evaluation of nonlinear power amplification effects on the signal. A power amplifier is typically modeled by its amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) conversion functions. The (AM/AM) conversion is used to characterize the amplitude distortion, which is the relationship between the input power (amplitude) and the output power (amplitude). (AM/PM) conversion is used to characterize phase deviation (distortion) caused by amplitude variations. A widely accepted Solid-State Power Amplifier (SSPA) model is the Rapp model [419]. It has a frequency-nonselctive response with a smooth transition from linearity to saturation as input amplitude approaches the saturation level. Its (AM/AM) conversion function is:

$$A_{out} = a \frac{A_{in}}{\left(1 + \left[\left(\frac{aA_{in}}{A_{sat}}\right)^2\right]^p\right)^{1/2p}} \quad (7.23)$$

with, $A_{sat} \geq 0$, $a \geq 0$, and $p \geq 0$

where A_{in} is the input amplitude, A_{sat} is the saturation level, a is the gain, and p is a positive number to control the nonlinearity characteristics of the HPA. The (AM/PM) conversion of the SSPA is small enough and can be neglected [419]. Fig. 7.8 shows the (AM/AM) conversion curve of the model with different values of p . As it is shown, as the value of p increases, the model converges to a hard limiting amplifier. For large values, the model becomes precisely linear until it reaches its output saturation level. A good approximation of existing amplifiers is obtained by choosing p to be in the range of 2 to 3 [420]. In this chapter, the Rapp model with $p = 2$ and $a = 1$ is used to simulate the nonlinear power amplification of vibration signals.

7.7 Performance Evaluation

The vibration signals of the Paderborn University (PU) bearing dataset [405] (Vibration set (f) in Section 7.3.1) are used to demonstrate the effectiveness of the proposed framework in reducing the PAPR, mitigating the effects of HPA nonlinearity, and improving the reliability of the condition monitoring process. This dataset is selected because it includes actual vibration signals from a real system during both healthy and faulty operations. The fault types include Inner Race (IR) defects, Outer Race (OR) defects, and combined defects. To create the training and testing sets, the vibration measurements of the PU dataset are segmented into segments of 6,400 samples. This results in 16,005 vibration signals in total. Accordingly, the dataset is split into 11,202 samples for training (70%), and 4,803 samples for testing (30%), and Adam optimizer (learning rate = 0.001) is used to train the autoencoders. The framework is implemented using Python, Keras library [421], TensorFlow [422], and SciPy library [320].

7.7.1 Experimental Setup

One-dimensional convolutional (Conv1D) layers are used to implement the autoencoders. The structures and the parameters of the autoencoders are shown in Fig. 7.9. The activation function (AF) used in the source autoencoder is based on the μ -law compression and expressed as:

$$AF = \text{sgn}(x) \frac{\ln(1 + 255 \times |x|)}{\ln(1 + 255)} \quad (7.24)$$

The value of the compression parameter μ in the activation function is set to 255 to achieve a target PAPR of 8 dB according to the μ -PAPR curve depicted in Fig.7.6.b.

Training the Source Autoencoder

To train the source activation-based autoencoder, the raw signals of the training set are smoothed in the first place. Then, the autoencoder is trained using (7.24) as the activation function, with the raw signals as the input and the smoothed signals as the target. The training objective is to minimize the MSE validation loss until it reaches the target compression loss of 0.045, which is calculated empirically using the formula of (7.18) as explained in Section 7.5.2.

Training the Destination Autoencoder

In contrast to the source autoencoder, the training objective of the destination autoencoder is minimizing the MSE validation loss until maximum convergence. Two training scenarios that address the absence and the presence of noise are considered:

Noise-free scenario The destination autoencoder is trained in reverse order compared to the source autoencoder. First, the signals of the training set (training targets of the destination autoencoder) are used as inputs to the source autoencoder *that has already been trained*. Subsequently, the compressed-smoothed signals obtained from the source autoencoder are employed as the training input for the destination autoencoder.

Noisy scenario To count for the accumulated noise in practical situations, the obtained (smoothed-compressed) signals from the source autoencoder are randomly and equally corrupted with a zero-mean AWGN of -5 dB and 0 dB SNR levels, and the destination autoencoder is trained accordingly.

7.7.2 Performance Metrics

To show the effectiveness of the proposed framework, its performance is evaluated in the presence of nonlinear power amplification and AWGN against the cases of μ -law compand-

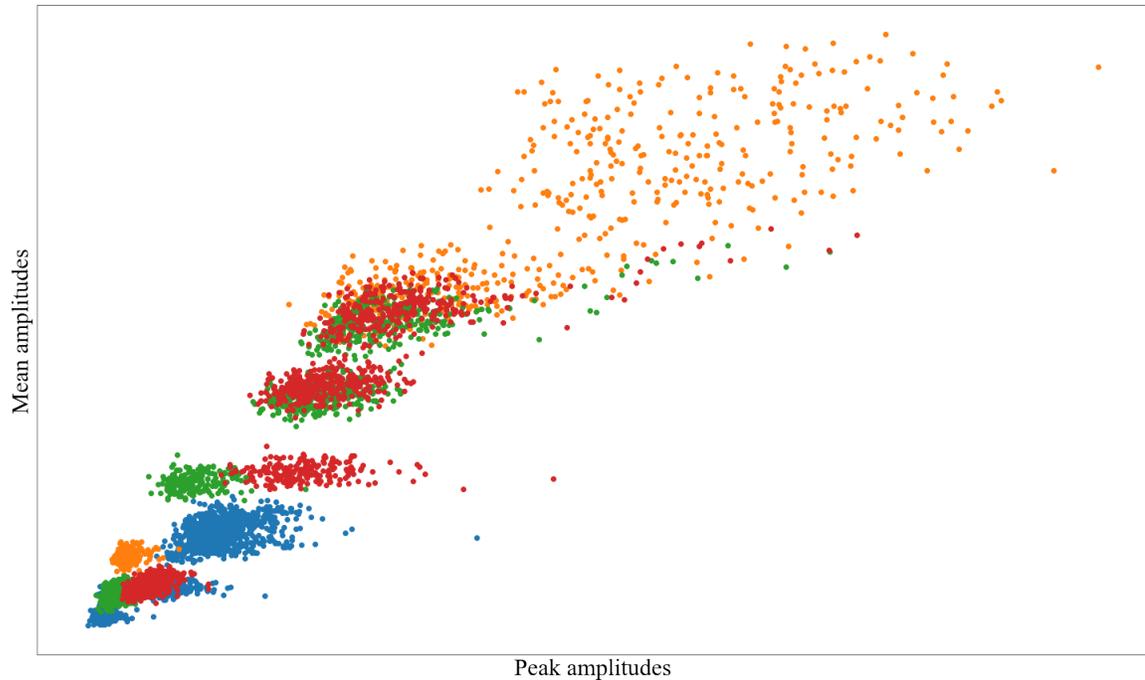


Figure 7.10: Signal’s amplitude constellation of the PU dataset.

ing and no companding. The saturation level (A_{sat}) of the HPA is set to the mean average power of the original “uncompressed” vibration signals. The performance of the proposed framework is evaluated in terms of the following aspects:

PAPR Reduction

The CCDF is used to measure the PAPR reduction capability of the proposed framework.

Average Power of the Compressed Signal

This will assess the increase in the average power resulting from the compression.

Nonlinear Distortion

Quantifying the nonlinear distortion caused by nonlinear power amplification is essential to assess the impact of nonlinear power amplification and evaluate the effectiveness of the

employed signal companding. To achieve this, a proper metric M , needs to be used to measure the distortion D as follows:

$$D = M(x, x') \quad (7.25)$$

where x is the original vibration signal, and x' is the proposed-expanded/ μ -law-expanded/uncompanded signal “processed signal” at the destination. While error metrics such as MSE and MAE can be used directly to measure the error, they are not aligned with the performance objective of restoring the PAPR characteristics in the processing center. This requires the development of a proper metric that is directly mapped to the PAPR characteristics of the signal. Accordingly, a new metric is introduced based on concepts of amplitude constellation and Error Vector Magnitude (EVM). These concepts are widely used in telecommunications systems to represent modulated signals and evaluate system-level performance.

Amplitude Constellation of Vibration Signals: To obtain the amplitude constellation of a given set of vibration signals, each signal x_i is expressed in terms of its peak and mean values of its magnitude amplitude, which represent its PAPR characteristics. This can be expressed mathematically as follows:

$$x_i = (A_{i_p}, A_{i_m}), i = 0, \dots, V - 1, \quad (7.26)$$

where;

$A_{i_p} = \max \{|x_i|\}$ is the signal’s peak amplitude,

$A_{i_m} = \text{mean} \{|x_i|\}$ is the signal’s, mean amplitude,

V is the number of vibration signals in the set.

Accordingly, the constellation can be displayed as a scatter plot on the $x - y$ plane where x and y represent the signal peak amplitudes A_{i_p} and average amplitudes A_{i_m} , respectively.

Fig. 7.10 displays the amplitude constellation of the PU dataset; the points are color-coded according to their health conditions so that the signals of the same health condition share

the same color. The position of a given signal in the constellation indicates both its peak and average amplitude and the distance— in terms of these amplitudes— between the signal and the other signals in the constellation. As mentioned previously, amplitudes of vibration signals are directly related to the monitored condition. Therefore, the constellation can offer crucial information and insight into the health of the monitored system. Further, as the distortion impacts the amplitude of the signal, its position in the constellation will be altered accordingly. This offers the opportunity to visually evaluate the nonlinear distortion by comparing the amplitude constellation, denoted as con_{amp} , of the processed test vibration signals x'_i at the destination to the reference constellation, denoted as con_{ref} , of the original test signals x_i .

Error Vector Magnitude (EVM): While amplitude constellation offers a useful metric to visualize the distortion, The EVM can be utilized to quantify this distortion and evaluate the effectiveness of the signal companding scheme. To obtain the EVM, the error vectors of the processed signals x'_i with respect to their reference test signals x_i are first calculated from the corresponding constellations con_{amp} and con_{ref} . The error vector $error_v$ between two points $x_i = (A_{i_p}, A_{i_m})$ and $x'_i = (A'_{i_p}, A'_{i_m})$ on the constellation is given by:

$$\begin{aligned} error_v &= [err_p, err_m], \\ err_p &= A_{i_p} - A'_{i_p}, \\ err_m &= A_{i_m} - A'_{i_m} \end{aligned} \tag{7.27}$$

Accordingly, the EVM can be calculated as the mean or the RMS value of the magnitudes of these obtained error vectors. It can be expressed as:

$$EVM = \frac{\sqrt{\frac{1}{V} \sum_{i=0}^{V-1} |error_v[i]|^2}}{EVM \text{ Normalization Reference}} \times 100 \tag{7.28}$$

where V is the number of vibration signals in the test set and $|error_v[i]|$ is the magnitude of the i -th error vector. In the above equation, EVM is normalized by *EVM Normalization Reference*,

which equals the maximum magnitude in the reference constellation con_{ref} . Hence, the EVM quantifies the amplitude distortion caused by HPA nonlinearity. In practical situations, the EVM quantifies the combined impact of all signal impairments within a VBCM system (such as distortion and noise effects), enabling measuring the overall system degradation using a single value.

Power Efficiency

The Power efficiency of a given HPA is specific to its type. Nevertheless, the problem of PAPR and its impact on efficiency is common among different HPA types. In this chapter, the efficiency η of class B HPA is utilized to evaluate the power efficiency of the proposed framework. η is defined as [423]:

$$\eta(\%) = \frac{\pi}{4\sqrt{IBO}} \times 100 \quad (7.29)$$

The efficiency is expressed in terms of applied IBO, which is defined as:

$$IBO = \frac{A_{sat}^2}{P_{in}}, \quad (7.30)$$

where A_{sat} is the saturation level of the HPA and P_{in} is the average power of the input signal to the HPA. Expressed in dB , the IBO equals to:

$$IBO(dB) = 10 \times \log_{10} \left(\frac{A_{sat}^2}{P_{in}} \right) \quad (7.31)$$

According to (7.29), a maximum efficiency of 78.5 % is achieved when the input power P_{in} equals the saturation level, A_{sat} ($IBO = 0 dB$). To apply a given IBO to a signal prior to power amplification, the average power P_{in} of the signal is scaled according to:

$$P_{in} = \frac{A_{sat}^2}{IBO} \quad (7.32)$$

As IBO increases, more signal peaks accumulate within the linear range of the HPA, which results in a reduction of nonlinear distortion. This, in turn, improves the EVM and, on the other hand, reduces the HPA efficiency. Accordingly, the performance is evaluated in terms of applied IBO, achieved EVM, and resultant HPA efficiency η .

Spectral Spreading

Spectral spreading or spectral broadening refers to situations when a signal's spectrum becomes wider due to nonlinear processing, such as logarithmic-based compression and nonlinear power amplification. Nonlinearity imposed on the signal's envelope causes an undesirable increase in the power of the side lobes of the Power Spectral Density (PSD). This makes PSD an appropriate measure of spectral regrowth. Accordingly, the mean PSD to is used to evaluate the compressed-amplified signals' spectral spreading. Welch's overlapped segment averaging method [55] is used to estimate the PSD. The method involves segmenting the signal using a moving window and computing each segment's Fast Fourier Transform (FFT). The PSD is then estimated as the average of the computed FFTs over all segments. The following settings are used for Welch's PSD estimation:

- Window: Hamming window of a length equals to $N/2$, where N is the length of the vibration signal. This length is selected to obtain a PSD with a good resolution since reducing the window length would affect the resolution.
- Overlap between segments: 50% overlap. With a window length of $N/2$, an overlap of 50% results in a total of 3 segments, reducing the averaging-error variance compared to using two segments only and, simultaneously, avoiding introducing a high correlation between the segments.
- Number of Discrete Fourier Transform points (NFFT): $NFFT = 8192$. This is calculated using the conventional method where NFFT is set to be equal to 2^P , where P is the smallest power of 2 that is greater than or equal to N , which in this case equals

13.

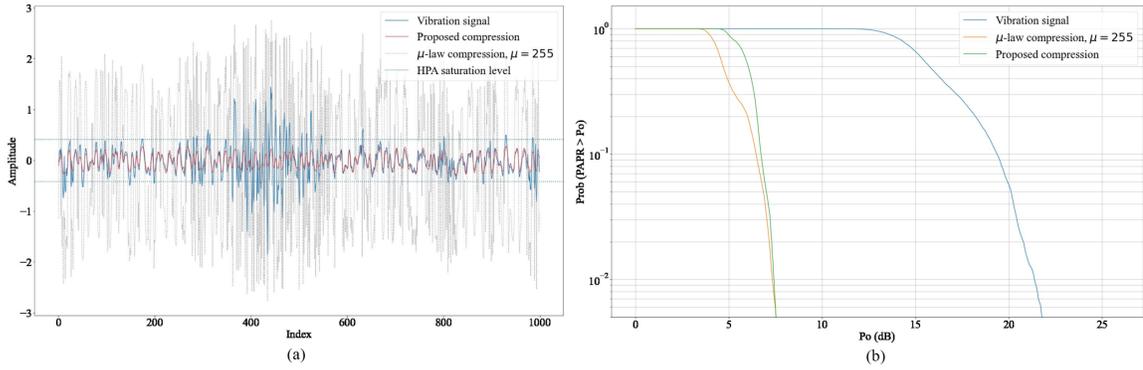


Figure 7.11: (a) Original and compressed vibration signals and (b) CCDFs of original and compressed signals.

Signal Denoising

The SNR of the vibration signals after denoising, denoted as SNR_d is used to assess the effectiveness of the proposed framework in reducing noise. SNR_d is expressed as:

$$SNR_d \text{ (dB)} = 10 \times \log_{10} \left(\frac{\sum_{n=0}^{N-1} |x[n]|^2}{\sum_{n=0}^{N-1} (|x[n] - x'(n)|^2)} \right) \quad (7.33)$$

where,

x : is the original vibration signal,

x' : is the denoised signal,

N : is the length of the vibration signal.

7.8 Results and Discussion

This section introduces and discusses the obtained performance results. It also addresses the reliability of the proposed framework in condition monitoring and demonstrates the lightweight nature of the proposed activation-based compression autoencoder.

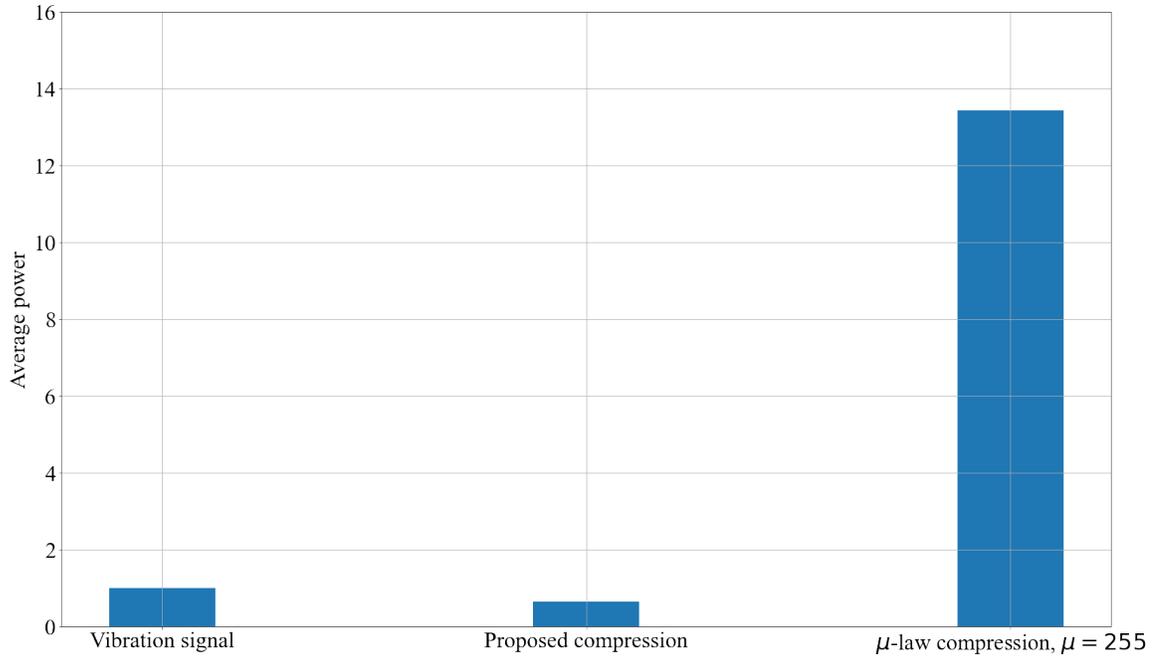


Figure 7.12: Average power ratios of original and compressed signals with respect to normalized average power of original signals.

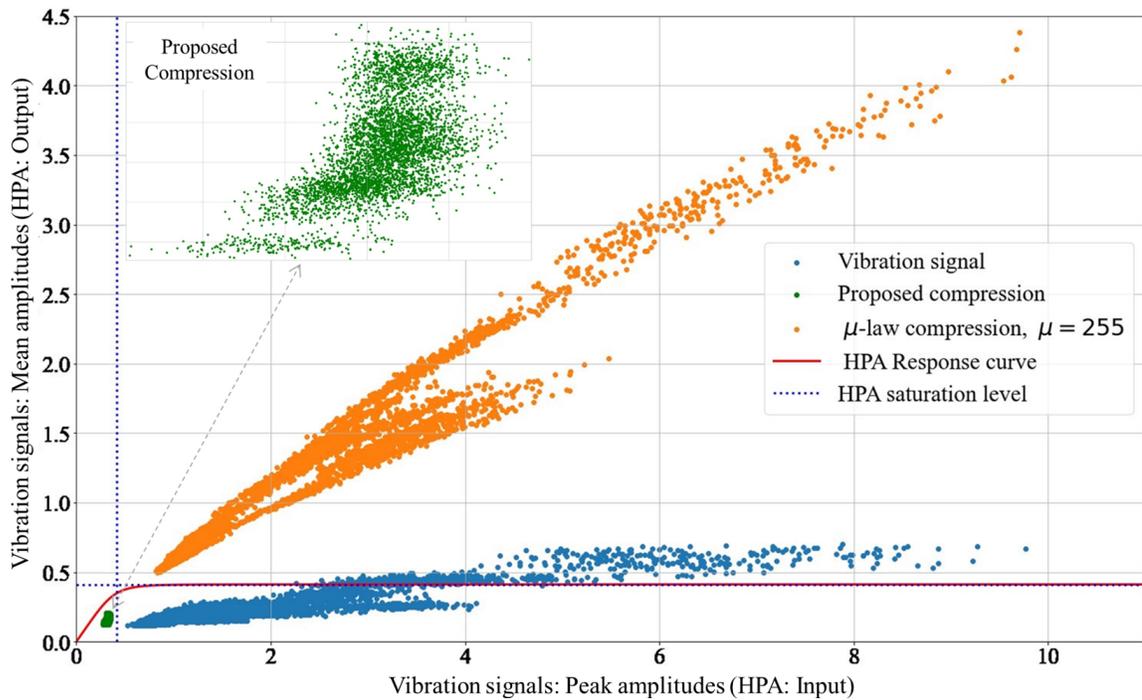


Figure 7.13: Constellations of uncompressed and compressed signals along with HPA response curve.

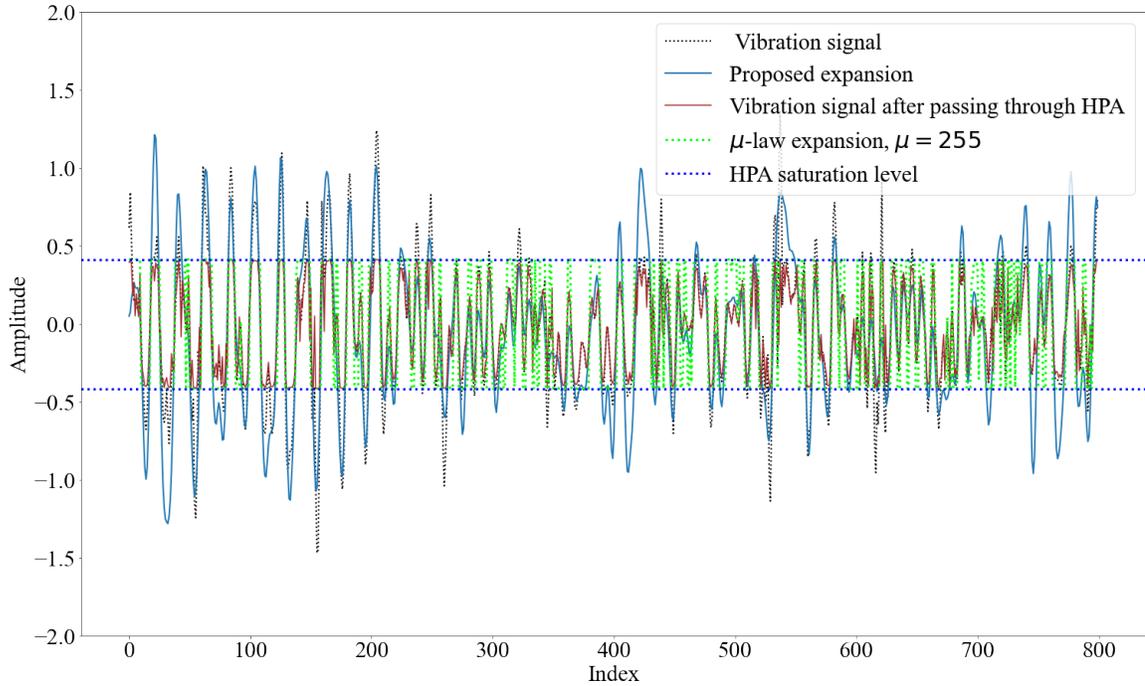


Figure 7.14: Original and expanded vibration signals after passing through HPA.

7.8.1 Performance Results

Before presenting and discussing the obtained results, it is convenient to demonstrate the important role of proper signal companding in mitigating the effects of nonlinear power amplification. In Fig. 7.11.a, it can be seen that the uncompressed vibration signal experiences high-amplitude excitations between the signal’s indices 350 and 550. These amplitude excitations exceed the HPA saturation level, plotted as a dashed horizontal line in the Figure. In the absence of a proper signal companding mechanism, such excitations in the signal’s waveform— directly related to the monitored system and would carry vital information about its current condition— are subject to the nonlinear distortion of the HPA.

PAPR Reduction

As shown in Fig. 7.11.b, μ -law compression and the proposed compression are effective in reducing the PAPR of the test vibration signals. Specifically, the proposed compression

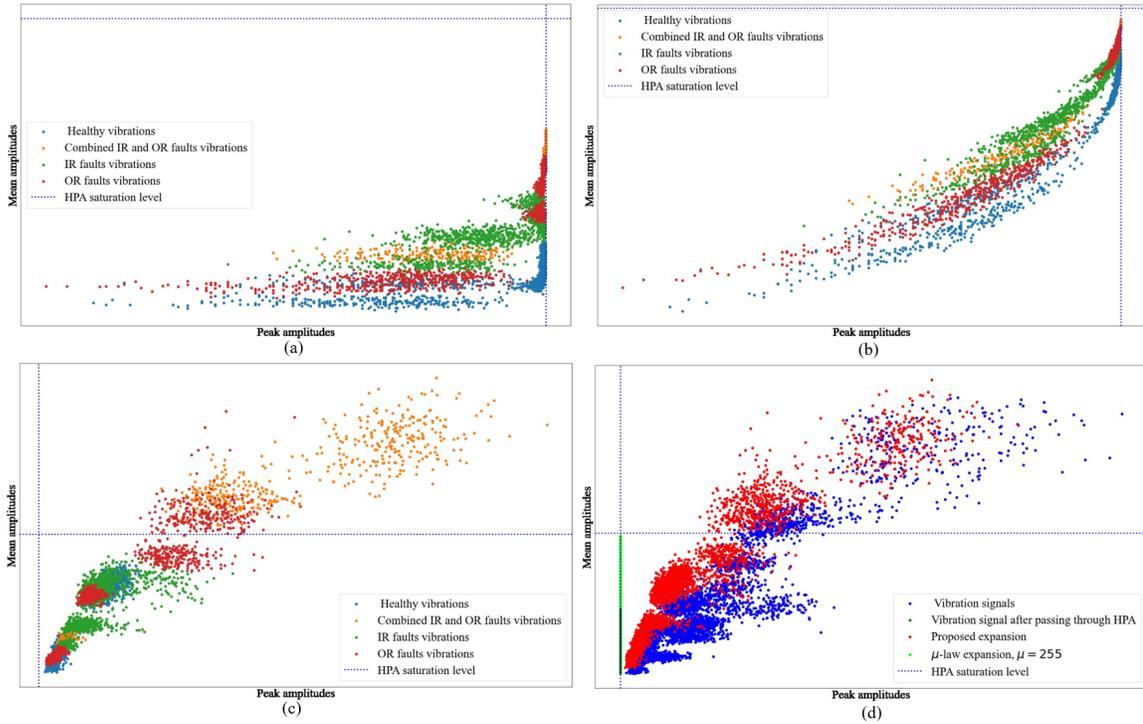


Figure 7.15: Amplitude constellations after passing through the HPA: (a) uncompanied signals, b) μ -law companded, (c) proposed-companded signals, and (d) original signals of the PU dataset along with uncompressed and expanded signals.

and μ -law compression have reduced the probability of exhibiting a PAPR of 8 dB in the test vibration signals from 1.0 to 0.0. However, as previously mentioned, μ -law compression relies on preserving the signal's peak amplitude while increasing its small amplitudes. Consequently, all amplitudes in the μ -law compressed form of the vibration signal surpass the HPA saturation level, as illustrated in Fig. 7.11.a, leading to significant distortion in the compressed signal, as shown later. In contrast, in the proposed compression, the source autoencoder learns how to reconstruct and compress the signal while avoiding the increase in its average power as explained earlier. This is demonstrated in Fig. 7.11.a, which shows that the majority of the amplitudes of the proposed-companded signals are compressed and maintained below the saturation level.

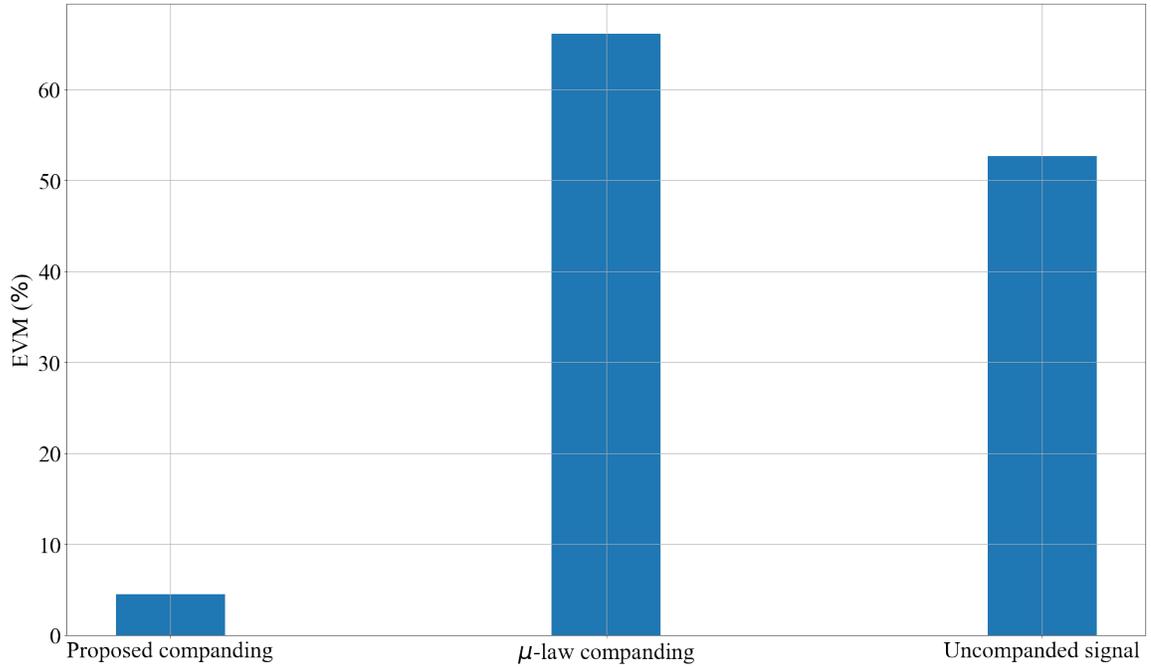


Figure 7.16: EVM values of uncompressed and expanded signals.

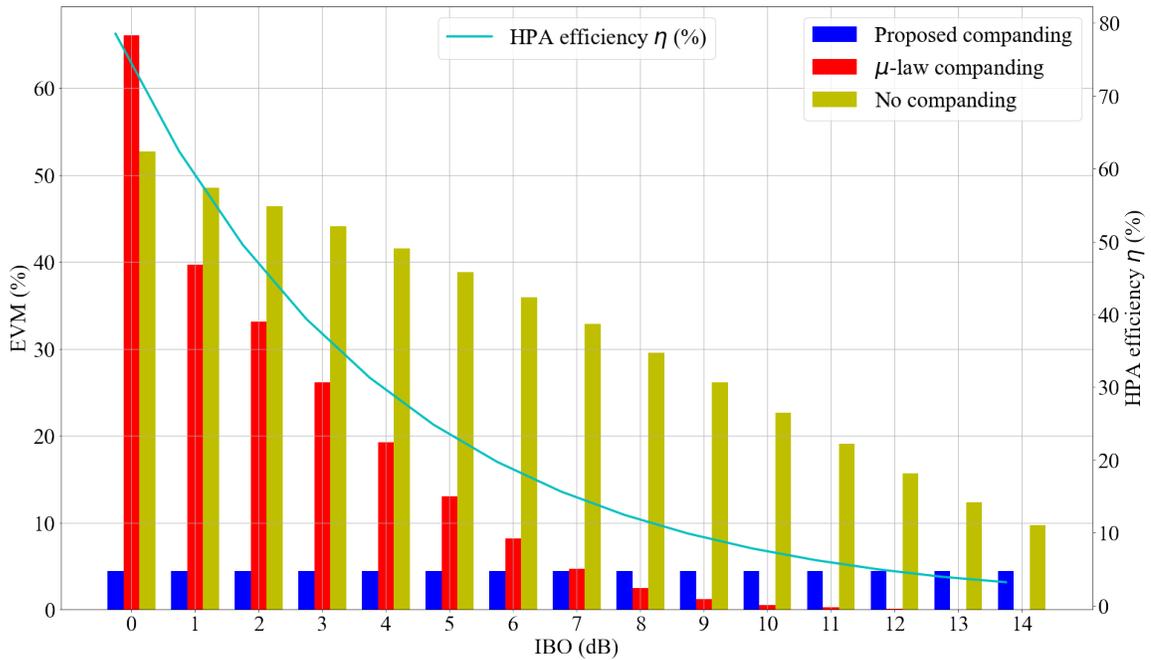


Figure 7.17: Applied IBO, achieved EVM, and resultant HPA efficiency η in proposed-companding, μ -law companding, and no companding.

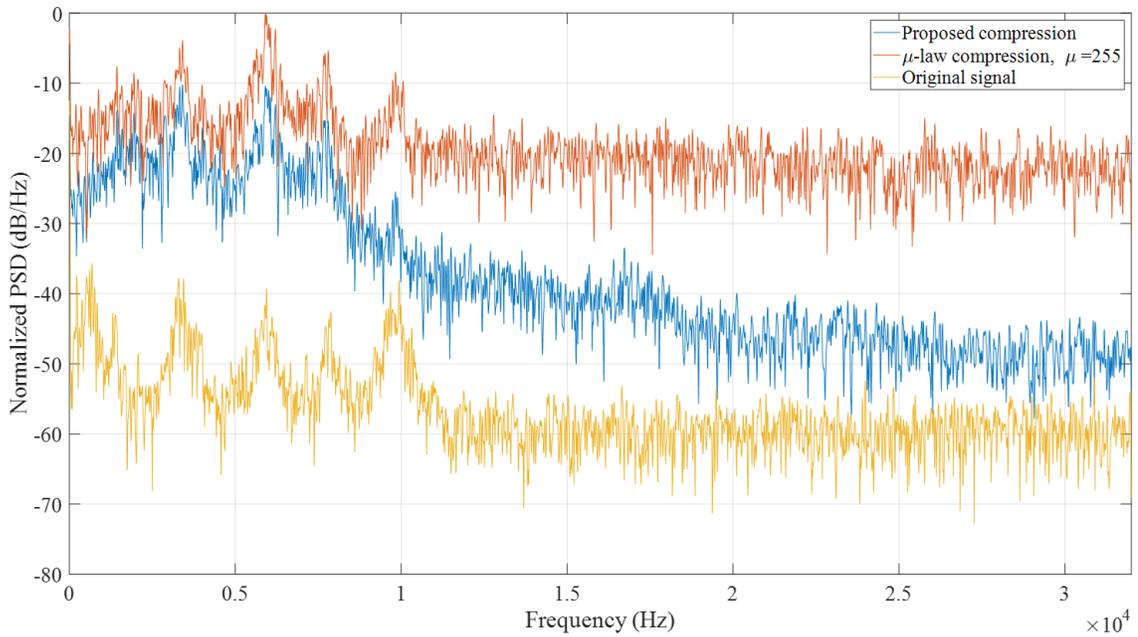


Figure 7.18: Normalized mean PSD plots of original signals, proposed-compressed signals, and μ -law compressed signals after passing through the HPA.

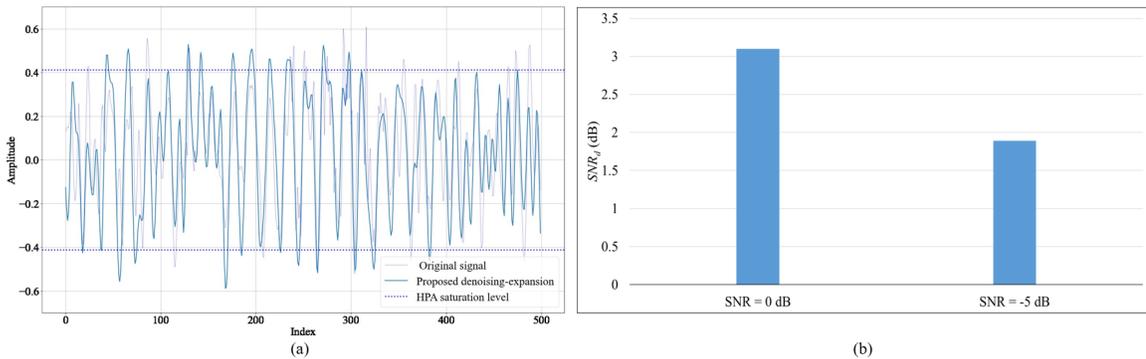


Figure 7.19: (a) a Denoised-expanded signal and its reference original signal. (b) Average SNR_d of denoised-expanded signals.

Average Power of the Compressed Signal

Fig. 7.12 depicts ratios of the mean average power of the compressed test signals with respect to the normalized average power of the original signals. As shown, the μ -law compressed form of the vibration signal, on average, exhibits more than a 13-fold increment in its average power due to the enlargement of its small amplitudes. While this would reduce the quantization noise in analog-to-digital conversion, it would cause severe nonlinear dis-

tortion in the signal when passing through the HPA. Regarding the proposed compression, it slightly reduces the average power of the compressed signal. In terms of HPA nonlinearity, the slight decrease in average power shifts the input power of the compressed signals towards the linear region of the HPA. This is demonstrated in Fig. 7.13, where the amplitude constellations of both the original and compressed test signals are displayed alongside the HPA's response curve. The HPA saturation level is also indicated as a dashed line on both axes. This visual setup provides useful insights into PAPR characteristics of the generated vibration signals, the behavior of the HPA, and companding design requirements. Specifically, considering the peak amplitudes in the plot, it is obvious that all uncompressed signals and their μ -law compressed forms will experience peak distortion after passing through the HPA since their peaks exceed the saturation level of the HPA. As for the mean amplitudes, all of the μ -law compressed forms will experience significant and frequent amplitude distortion after passing through the HPA since their mean amplitudes exceed the saturation level. This is also the case for a considerable part of the uncompressed signals. On the contrary, the proposed framework compressed all the test signals so that their peaks and mean amplitudes fall below the saturation level. As a result, the compressed forms of the vibration signals are not subject to nonlinear distortion of the HPA. However, a slight amplitude distortion is still expected due to the soft limiting nature of the HPA and the imperfect reconstruction of the autoencoders. It is worth mentioning that the visual setup of Fig. 7.13 can be adapted to various systems to gain more insights into the PAPR characteristics, nonlinear behavior, and requirements for PAPR reduction.

Nonlinear Distortion

Fig. 7.14 shows an uncompanded vibration signal before and after the HPA, its μ -law companded form (*the word companded is used here to refer to the signal that is compressed, passed through the HPA at the source, and expanded at the destination*), and its proposed-companded form. As shown, the uncompanded signal and its μ -law companded

form experienced a significant nonlinear distortion as their peak values are restricted to the HPA saturation level. While on the other hand, the amplitudes of its proposed-companded form are free of nonlinear distortion. Since the proposed framework compresses the signal so that its amplitudes fall in the linear region of the HPA, all amplitudes— even the ones that exceed the HPA saturation level— are restored after the expansion. The amplitude consultations of the uncompanded test signals, their μ -law companded forms, and proposed-companded forms are displayed in Fig. 7.15. By co-locating these constellations with the reference constellation as shown in Fig. 7.15.d, a convenient visual comparison to asses the nonlinear distortion can be made. The comparison clearly shows that the uncompanded signals and their μ -law companded forms experience severe nonlinear distortion at the destination while, on the other hand, the proposed companding framework avoids nonlinear distortion and successfully restores the reference constellation—to a large extent— at the destination. A comparison among the obtained EVM values is shown in Fig. 7.16. The comparison demonstrates, in a quantified manner, the effectiveness of the proposed framework in mitigating the effects of nonlinear distortion. Specifically, while the uncompanded vibration signals and their μ -law companded forms suffered from a very high distortion ($> 50\%$ EVM), the proposed-companded forms experienced very low distortion ($< 4.5\%$ EVM). As previously stated, the EVM quantifies the total system degradation experienced by the signals. For the uncompanded vibration signals and their μ -law companded forms, the exhibited distortion is exclusively caused by the nonlinear power amplification. Regarding the proposed framework, the resulting EVM represents a floor value that is not related to nonlinear distortion. Specifically, the factors that contributed to EVM of the proposed framework are:

- Soft limiting nature of the HPA: with $p = 2$, the used SSPA model acts as a soft limiter.
- Autoencoder error: due to imperfect signal reconstruction of the autoencoders during

compression/expansion stages. This error can be reduced by conducting more fine-tuning for hyperparameters of the autoencoders.

- Noise presence and channel effects: While these impairments are not considered in the evaluation setup related to the obtained EVM results, they have a strong influence on the total system degradation in practical situations.

The obtained results from the EVM evaluation show that in the presence of nonlinear devices and the absence of a proper mechanism to reduce the PAPR, VBCM systems could suffer from severe nonlinear distortion. The results also confirm the effectiveness of the proposed framework in mitigating the effects of such distortion.

Power Efficiency

Fig. 7.17 plots applied IBO, corresponding achieved EVM, and the resulting HPA efficiency η . As previously explained, the proposed companding technique allows the compression mechanism to bring the peaks of input signals within the linear range of the HPA (*i.e.* below the defined saturation level, A_{sat}). This avoids nonlinear distortion and eliminates the need for IBO, thereby maximizing the HPA efficiency, as shown in the figure. In contrast, in cases where μ -law companding and no companding are used, enormous IBO values are required to accommodate signal peaks within the linear range of the HPA, which in turn reduces power efficiency significantly. It is worth noting that the shown EVM value of the proposed companding in the figure represents the floor EVM as explained earlier.

Spectral Spreading

Fig. 7.18 shows the mean PSD plots of the test vibration signals, their μ -law-compressed forms, and their proposed-compressed forms. The mean PSD of each of these three sets is calculated by estimating the individual PSD of each signal in the set after passing through the HPA. Accordingly, the mean PSD is obtained by averaging the estimated PSDs. As

shown, the μ -law-compressed forms experienced higher spectral broadening than the proposed-compressed forms. This regrowth in the spectrum is attributed mainly to the nonlinear distortion caused by the HPA. However, it should be mentioned that the logarithmic-based nature of the compression mechanism leads to spectral regrowth in the spectrum of the compressed signal.

Signal Denoising

To evaluate the denoising capability of the proposed framework, the test vibration signals are first compressed using the proposed compression, passed through the HPA, and then corrupted with the AWGN. Consequentially, at the destination, the compressed-amplified-noisy forms of the test signals are passed throughout the trained denoising-expanding autoencoder (trained according to the *Noisy scenario* in Section 7.7.1) to recover the original test signals. The plot in Fig. 7.19.a shows a recovered “denoised and expanded” signal alongside the reference original signal. By comparing both signals in the plot, it can be seen that the proposed framework is effective in removing the corrupted noise, expanding the compressed form, and restoring the original signal. The average SNR_d values of the recovered signals are shown in Fig. 7.19.b. These values quantify the improvement in the recovered signals’ SNR levels and confirm the capability of the proposed framework to remove noise and mitigate the effects of noise expansion. Specifically, considering noisy, compressed signals with 0 dB and -5 dB NR levels, the improvement in the SNR after denoising and expanding these signals are 3.1 dB and 6.9 dB , respectively.

7.8.2 Reliability of Condition Monitoring

To assess the proposed framework’s reliability in condition monitoring, time and frequency features are extracted at the destination from the uncompanded, μ -law companded, and proposed-companded signals of the test PU dataset, considering both the noise-free and noisy scenarios. The dataset comprises four classes of operational conditions: 1- healthy,

Table 7.2: Extracted features from expanded vibration signals

S.N	Feature	Formula
1	Peak power	$Max_{n \in \{0, N\}} x[n] ^2$
2	Mean power	$\frac{1}{N} \sum_{n=0}^{N-1} x[n] ^2$
3	Shannon entropy	$H[x[n]] = - \sum_{i=1}^m p(x_i) \log(p(x_i))$, $p =$ probability, $x_i = m$ different values in $x[n]$
4	Peak spectral power	$Max\{PSD(x[n])\}$
5	Freq. of peak spectral power	
6	Spectral entropy	$H[PSD]$

Table 7.3: Classification results of the operational conditions of the PU dataset

SNR	Accuracy (%)		
	Proposed	μ -law	No Companding
No noise	96.04%	87.6%	93.5%
0 dB	93.6%	55.3%	85.8%
-5 dB	91.1%	47.4%	79.2%

Table 7.4: Results of the different training scenarios

Metrics	Proposed Training		Conventional Autoencoder
	Target val. loss: 0.045 (compression loss)	Target val. loss: 0.015	Achieved val. loss: 0.008
Training epochs	5	25	33
Training time (seconds)	423.58	2171.3	2793.16
EVM (%)	4.58	10.91	23.016
Increment percentage in the average power (%)	No increase	No increase	351.000
Probability of target PAPR (8 dB)	0.0	0.861	0.0

2- IR faults, 3- OR faults, and 4- combined IR and OR faults. The vibration signals in the PU dataset are acquired at a rate of 64,000 samples per second. In the proposed framework, the input vibration segment consists of 6,400 samples. Therefore, with this sampling rate, the duration of the input vibration segment is 0.1 seconds, which is precise enough to enable real-time monitoring. The features extracted from the signals are listed in Table 7.2. The first three features are time-domain features obtained directly from the signal's waveform. The first feature represents the signal's peak power, while the second feature represents its mean power. The third feature is the Shannon entropy of the time-domain waveform, which is used to assess the signal's irregularity under different operating conditions. The last three features are frequency-domain features extracted from the signal's PSD. These features include the peak spectral power, the corresponding frequency, and the spectral entropy. The dataset of resulting features is divided into a 70% training set and a 30% testing set, and a Random Forest (RF) classifier is trained accordingly. Table 7.3 displays the accuracy results in percentage for various scenarios, including noise-free, 0 dB SNR, and -5 dB SNR conditions. In comparison to both μ -law companding and no companding cases, our proposed companding technique achieved superior performance in condition monitoring and fault detection. This superiority arises from its ability to mitigate nonlinear distortion. Obviously, in noisy scenarios, the presence of noise has a negative impact on the accuracy of fault detection. Furthermore, expanding compressed noisy signals without employing effective denoising techniques, as is the case with μ -law companding, heavily degrades the performance due to the undesired effect of noise expansion. In contrast, the proposed method's denoising capabilities effectively mitigate the impact of noise, consequently enhancing the reliability of the condition monitoring process.

7.8.3 The Lightweight Nature of the Proposed Activation-Based Compression Autoencoder

Ensuring computational efficiency during both the training and inference phases is an essential requirement for the proposed activation-based compression autoencoder in the sensor node. The shallow structure of the proposed autoencoder results in a small number of trainable parameters and, hence, less inference time and less computational complexity compared to other deep structures. Particularly, the proposed model involves a total of 19,297 parameters, including 19,105 trainable parameters. Moreover, the compression loss stemming from the compression-based activation serves the purpose of predefining the target validation loss. This, in turn, reduces the training time and enhances performance, as explained earlier. This is demonstrated in Table 7.4, where a comparison between training with target compression loss, training with a target loss of 0.015, and conventional autoencoder-based companding is conducted. In the conventional companding, the Rectified Linear Unit (ReLU) activation function is used. The training procedure involves minimizing the loss between reconstructed output signals and training target signals that are smoothed, compressed, and power-scaled versions of the input training signals. Accordingly, the model is trained to minimize the validation loss until the maximum convergence, considering early stopping criteria with patience of 7. As shown in the table, using the compression loss as a minimum bound for the target validation loss maintains the desired compression characteristics in the compressed signal. This leads to reduced distortion, better PAPR reduction, and prevents any increase in the average power of the signal. Furthermore, it reduces the number of training iterations, resulting in less training time compared to other methods.

7.9 Conclusion

The chapter addressed the issue of PAPR in vibration signals and investigated its effects, in the presence of nonlinear power amplification, on both power efficiency and system performance. Consequently, it introduced a novel compression-based activation autoencoder companding framework designed to control the PAPR. The proposed framework enhances power efficiency and improves system performance significantly. To the authors' knowledge, this chapter represents the first attempt to address this problem in the context of VBCM literature. First, the chapter conducted statistical analysis on the amplitude distribution of vibration signals and presented a closed-form formula to model the CCDF of the PAPR. Accordingly, analytical analysis and empirical investigation of the PAPR were conducted using various vibration datasets. The outcomes confirmed the occurrence of high PAPR values in vibration signals, particularly between 10 *dB* to 13 *dB*, when the number of samples exceeded 500 in the acquired vibration segments. Further, the chapter examined the impact of the high PAPR in the presence of nonlinear power amplification on power efficiency and system performance, where new metrics are adapted to quantify the resulting nonlinear distortion. The findings revealed that in the case of uncontrolled PAPR, HPA nonlinearity induces severe nonlinear distortion, resulting in reduced power efficiency and degraded performance. Accordingly, the chapter proposed a signal-companding-based framework to reduce the PAPR and mitigate the impacts of HPA nonlinearity, which in turn enhances power efficiency and improves performance. In the sensor node, the framework employed a lightweight reconstruction autoencoder with a compression-based activation function that simultaneously smooths and compresses the vibration signal while avoiding any increase in the average power of the compressed signal. Consequently, in the processing center, the proposed framework used an expansion autoencoder that acts as a denoising filter to denoise the compressed signals before the expansion operation, thereby preventing the enhancement of the accumulated noise during signal expansion. Additionally, the

proposed architecture, which incorporates reconstruction autoencoders with compression-based activation, carries out various signal-processing tasks simultaneously. Therefore, it facilitates an efficient end-to-end implementation of the framework. Regarding performance, The experimental results demonstrate that the proposed framework brought substantial enhancement in power efficiency, along with improvements in the performance of condition monitoring. In conclusion, the chapter underscores the importance of implementing an appropriate signal companding mechanism in VBCM systems to mitigate the effects of nonlinear devices, ensuring efficient power consumption and reliable monitoring processes.

Chapter 8

8 Conclusion and Future Work

The work presented in this thesis addressed the main aspects of Vibration-Based Condition Monitoring (VBCM), including signal processing fundamentals, feature extraction, availability of labeled data, computational complexity, the presence of noise in the acquired vibration signals, and power consumption in vibration sensor nodes. The methods employed in this thesis spanned the fields of Digital Signal Processing (DSP) and ML techniques (supervised, Deep Learning (DL)), including signal preprocessing, signal denoising, signal frequency-domain analysis, signal time-frequency domain analysis, signal companding (compression-expansion).

In the first part of the presented work, a data-driven VBCM framework using limited labeled data is developed. The framework addresses real-world situations of limited availability of labeled data by applying a similarity-based classification where one labeled sample for each operational condition is enough to facilitate effective condition monitoring. In the second part, the end-to-end delay of VBCM systems is analyzed. With fixed computational resources, the system delay depends entirely on the duration of the input vibration segment, the computation steps of the used algorithm/s, and the size of the features vector. Accordingly, a low-complexity method is introduced for vibration-based condition monitoring and fault diagnosis of rolling bearings. The proposed method allows the use of short durations of the generated vibration signal and produces a feature vector of controllable size, thereby relaxing memory requirements and reducing monitoring delay. The third part of the thesis addressed the PAPR in vibration signals and investigated its effects, in the presence of nonlinear power amplification, on both power efficiency and system performance. The findings revealed that in the case of uncontrolled PAPR, HPA nonlinearity induces severe nonlinear distortion, resulting in reduced power efficiency and degraded performance. Accordingly, a novel autoencoder-based framework for signal companding and denoising

is designed to control the PAPR and mitigate noise effects, enhancing power efficiency and improving performance.

The experiential evaluation of the introduced solutions underscored their efficiency, evidenced by improved overall performance in monitoring, reduced monitoring delay, relaxed memory requirements, and enhanced power efficiency in sensor nodes. Furthermore, the thesis contributed to the VBCM literature by providing a comprehensive tutorial on signal processing fundamentals and application-independent review of a typical signal-based ML pipeline and feature extraction techniques. As a result, the work presented in this thesis makes available efficient solutions to the main challenges facing the practical deployment of real-world VBCM systems.

The proposed solutions presented throughout the thesis used signal processing-centric approaches for feature extraction and similarity-based VBCM, making them more computationally efficient compared to DL-based approaches. Additionally, feature engineering through signal processing produces explainable features that meaningfully connect to signal conditions or classes, unlike the black-box features generated by DL models. Further, the performance of DL features is influenced by the model's architecture, parameters, and training, resulting in inconsistent performance. Conversely, signal processing-based features provide more consistent and interpretable performance. Moreover, DL Features often result in higher dimensionality and redundancy, necessitating additional dimensionality reduction and feature selection, thus increasing the computational burden. Nevertheless, for similarity-based VBCM, future research directions would explore the feasibility of DL-based approaches. Specifically, the powerful synthesizing capabilities of Generative Adversarial Networks (GANs) could be explored to generate synthetic vibration samples with reliable characteristics to overcome the limitation of labeled training data. The advancement in computational capabilities enables the feasible implementation of these approaches despite their intensive computational requirements. Further, as an offline operation, syn-

thesizing new samples does not impact online inferencing tasks. For feature extraction, future research directions would consider integrating DL models such as Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN) within data-driven VBCM to learn highly discriminative features from acquired vibration signals. Moreover, future research would consider developing efficient signal-to-image transformation approaches, allowing the adaptation of pretrained Convolutional Neural Network (CNN) models as powerful feature extractors. Signal-to-image conversion becomes particularly feasible in integrated systems such as Unmanned Aerial Vehicles (UAVs) and Autonomous Vehicles (AVs). In these systems, computer vision applications, such as object detection and navigation, and VBCM applications are commonly deployed within the same environment. Therefore, it is deemed computationally efficient to convert the acquired vibration signals into images and use transfer learning to utilize pretrained DL infrastructure for inference tasks in VBCM.

The use of DL for signal-denoising applications is an active research field. In this thesis, an effective signal denoising scheme was facilitated through an autoencoder-based structure. From a signal-processing perspective, This approach can be centered around the concept of tunable denoising filters, where the trained autoencoder is viewed as a tuned denoising filter whose coefficients, represented by model weights, are tuned through the training phase to remove noise from the signal. DL-based approaches are approved to be highly effective in noise removal due to their ability to learn complex, noisy patterns and related dependencies within the signal. Moreover, they can be designed to implement various operations on signal simultaneously, facilitating an efficient end-to-end implementation of various signal-processing pipelines, as demonstrated in Chapter 7. However, the effectiveness of DL-based denoising approaches, in general, is limited to noise profiles that are present in the training set. Thus, future research directions would explore the use of Adaptive Noise Cancellation (ANC) filters as they can adapt to varying noise patterns, thereby meeting the requirements of real-time applications that involve varying noise characteristics.

Another promising research topic is the exploration of Tiny Machine Learning (TinyML) to improve the practical implementation of the proposed solutions within resource-constrained environments. TinyML is characterized by its ability to deploy ML models on microcontrollers and other resource-constrained devices, which are typically limited in computational power and memory. Hence, integrating TinyML with these solutions presents a promising opportunity for enhancing their feasibility and efficiency in real-world applications.

Bibliography

- [1] R. B. Randall, *Vibration-based Condition Monitoring*. Wiley, 1 2011.
- [2] C. R. Farrar and K. Worden, *An Introduction to Structural Health Monitoring*. Vienna: Springer Vienna, 2010, pp. 1–17. [Online]. Available: https://doi.org/10.1007/978-3-7091-0399-9_1
- [3] Y. Yang, Y. Zhang, and X. Tan, “Review on vibration-based structural health monitoring techniques and technical codes,” *Symmetry*, vol. 13, no. 11, 2021. [Online]. Available: <https://www.mdpi.com/2073-8994/13/11/1998>
- [4] A. K. N. H. Ahmed, *Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines*. Wiley-IEEE Press, 2020.
- [5] V. Q. C. Tran, D. V. Le, D. R. Yntema, and P. J. M. Havinga, “A review of inspection methods for continuously monitoring pvc drinking water mains,” *IEEE Internet of Things Journal*, vol. 9, pp. 14 336–14 354, 8 2022.
- [6] Q. Meng, P. Lu, and S. Zhu, “A smartphone-enabled iot system for vibration and noise monitoring of rail transit,” *IEEE Internet of Things Journal*, vol. 10, pp. 8907–8917, 5 2023.
- [7] W. Hong, B. Xu, X. Chi, X. Cui, Y. Yan, and T. Li, “Long-term and extensive monitoring for bee colonies based on internet of things,” *IEEE Internet of Things Journal*, vol. 7, pp. 7148–7155, 8 2020.
- [8] M. H. M. Ghazali and W. Rahiman, “Vibration-based fault detection in drone using artificial intelligence,” *IEEE Sensors Journal*, vol. 22, pp. 8439–8448, 5 2022.

- [9] Z. Peng, J. Li, and H. Hao, “Development and experimental verification of an IoT sensing system for drive-by bridge health monitoring,” *Engineering Structures*, vol. 293, p. 116705, 10 2023.
- [10] J. Arshad, T. A. Siddiqui, M. I. Sheikh, M. S. Waseem, M. A. B. Nawaz, E. T. Eldin, and A. U. Rehman, “Deployment of an intelligent and secure cattle health monitoring system,” *Egyptian Informatics Journal*, vol. 24, pp. 265–275, 7 2023.
- [11] A. Singh, N. Nawayseh, Y. K. Dhahi, S. Samuel, and H. Singh, “Transforming farming with intelligence: Smart vibration monitoring and alert system,” *Journal of Engineering Research*, 8 2023.
- [12] T. Shaik, X. Tao, N. Higgins, L. Li, R. Gururajan, X. Zhou, and U. R. Acharya, “Remote patient monitoring using artificial intelligence: Current state, applications, and challenges,” *WIREs Data Mining and Knowledge Discovery*, vol. 13, 3 2023.
- [13] D. K. Iakovidis, M. Ooi, Y. C. Kuang, S. Demidenko, A. Shestakov, V. Sinitsin, M. Henry, A. Sciacchitano, S. Discetti, S. Donati, M. Norgia, A. Menychtas, I. Maglogiannis, S. C. Wriessnegger, L. A. B. Chacon, G. Dimas, D. Filos, A. H. Aletras, J. Töger, F. Dong, S. Ren, A. Uhl, J. Paziewski, J. Geng, F. Fioranelli, R. M. Narayanan, C. Fernandez, C. Stiller, K. Malamousi, S. Kamnis, K. Delibasis, D. Wang, J. Zhang, and R. X. Gao, “Roadmap on signal processing for next generation measurement systems,” *Measurement Science and Technology*, vol. 33, p. 012002, 1 2022.
- [14] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman, “A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications,” *Mechanical Systems and Signal Processing*, vol. 147, p. 107077, 1 2021.
- [15] J. Leukel, J. González, and M. Riekert, “Adoption of machine learning technology

- for failure prediction in industrial maintenance: A systematic review,” *Journal of Manufacturing Systems*, vol. 61, pp. 87–96, 10 2021.
- [16] C. Zhang, A. A. Mousavi, S. F. Masri, G. Gholipour, K. Yan, and X. Li, “Vibration feature extraction using signal processing techniques for structural health monitoring: A review,” *Mechanical Systems and Signal Processing*, vol. 177, p. 109175, 9 2022.
- [17] M. Romanssini, P. C. C. de Aguirre, L. Compassi-Severo, and A. G. Girardi, “A review on vibration monitoring techniques for predictive maintenance of rotating machinery,” *Eng*, vol. 4, pp. 1797–1817, 6 2023.
- [18] O. Matania, L. Bachar, E. Bechhoefer, and J. Bortman, “Signal processing for the condition-based maintenance of rotating machines via vibration analysis: A tutorial,” *Sensors*, vol. 24, p. 454, 1 2024.
- [19] S. Sonawani, K. Patil, and P. Natarajan, “Biomedical signal processing for health monitoring applications: a review,” *International Journal of Applied Systemic Studies*, vol. 10, p. 44, 2023.
- [20] Z. Khademi, F. Ebrahimi, and H. M. Kordy, “A review of critical challenges in mibi: From conventional to deep learning methods,” *Journal of Neuroscience Methods*, vol. 383, p. 109736, 1 2023.
- [21] V. Gupta, M. Mittal, V. Mittal, and N. K. Saxena, “A critical review of feature extraction techniques for eeg signal analysis,” *Journal of The Institution of Engineers (India): Series B*, vol. 102, pp. 1049–1060, 10 2021.
- [22] L. Boubchir, B. Daachi, and V. Pangracious, “A review of feature extraction for eeg epileptic seizure detection and classification,” in *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, 2017, pp. 456–460.

- [23] R. Yuvaraj, P. Thagavel, J. Thomas, J. Fogarty, and F. Ali, “Comprehensive analysis of feature extraction methods for emotion recognition from multichannel eeg recordings,” *Sensors*, vol. 23, p. 915, 1 2023.
- [24] R. Jenke, A. Peer, and M. Buss, “Feature extraction and selection for emotion recognition from eeg,” *IEEE Transactions on Affective Computing*, vol. 5, pp. 327–339, 7 2014.
- [25] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, “Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques,” *Sensors*, vol. 23, p. 6434, 7 2023.
- [26] M. L. Duc, P. Bilik, and R. Martinek, “Harmonics signal feature extraction techniques: A review,” *Mathematics*, vol. 11, p. 1877, 4 2023.
- [27] J. de Lope and M. Graña, “An ongoing review of speech emotion recognition,” *Neurocomputing*, vol. 528, pp. 1–11, 4 2023.
- [28] S. Pangaonkar and A. Panat, *A Review of Various Techniques Related to Feature Extraction and Classification for Speech Signal Analysis*, 2020, pp. 534–549.
- [29] S. Nivetha, “A survey on speech feature extraction and classification techniques,” in *2020 International Conference on Inventive Computation Technologies (ICICT)*, 2 2020, pp. 48–53.
- [30] G. Sharma, K. Umopathy, and S. Krishnan, “Trends in audio signal feature extraction methods,” *Applied Acoustics*, vol. 158, p. 107020, 1 2020.
- [31] A. A. Anthony and C. M. Patil, “Speech emotion recognition systems: A comprehensive review on different methodologies,” *Wireless Personal Communications*, vol. 130, pp. 515–525, 5 2023.

- [32] W. Liu, Y. Liu, S. Li, and Y. Chen, “A review of variational mode decomposition in seismic data analysis,” *Surveys in Geophysics*, vol. 44, pp. 323–355, 4 2023.
- [33] S. M. Mousavi and G. C. Beroza, “Machine learning in earthquake seismology,” *Annual Review of Earth and Planetary Sciences*, vol. 51, pp. 105–129, 5 2023.
- [34] P. Choudhary, N. Goel, and M. Saini, “A survey on seismic sensor based target detection, localization, identification, and activity recognition,” *ACM Computing Surveys*, vol. 55, pp. 1–36, 11 2023.
- [35] A. Rojas and G. J. Dolecek, *A Review of Spectrum Sensing Techniques Based on Machine Learning*, 7 2024, pp. 1–21.
- [36] Q.-V. Pham, N. T. Nguyen, T. Huynh-The, L. B. Le, K. Lee, and W.-J. Hwang, “Intelligent radio signal processing: A survey,” *IEEE Access*, vol. 9, pp. 83 818–83 850, 2021.
- [37] N. Soltanieh, Y. Norouzi, Y. Yang, and N. C. Karmakar, “A review of radio frequency fingerprinting techniques,” *IEEE Journal of Radio Frequency Identification*, vol. 4, pp. 222–233, 9 2020.
- [38] Y. Li and L. Wang, “A survey of radio frequency fingerprint feature extraction and recognition based on deep learning,” in *The International Conference on Forthcoming Networks and Sustainability (FoNeS 2022)*, 2022, pp. 776–782.
- [39] T. Wang, G. Yang, P. Chen, Z. Xu, M. Jiang, and Q. Ye, “A survey of applications of deep learning in radio signal modulation recognition,” *Applied Sciences*, vol. 12, p. 12052, 11 2022.
- [40] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, “An overview of feature-based methods for digital modulation classification,” in *2013 1st International Conference*

- on Communications, Signal Processing, and their Applications (ICCSPA)*, 2013, pp. 1–6.
- [41] S. Peng, S. Sun, and Y.-D. Yao, “A survey of modulation classification using deep learning: Signal representation and data preprocessing,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, pp. 7020–7038, 12 2022.
- [42] S. Xie and J. Ye, “Overview of automatic modulation recognition methods,” in *2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2023, pp. 1–7.
- [43] R. Priemer, *Introductory Signal Processing*. World Scientific, 11 1990.
- [44] R. X. Gao and R. Yan, *Signals and Signal Processing in Manufacturing*. Springer US, 2011, pp. 1–15.
- [45] L. F. Chaparro and A. Akan, *Frequency Analysis: The Fourier Series*. Elsevier, 2019, pp. 241–303.
- [46] S. Braun, *Spectral Analysis, Classical Methods*. Elsevier, 2001, pp. 1208–1223.
- [47] D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, and L. Kaufman, Eds., *Chapter 12 - Signals and Data*. Elsevier, 2003, vol. 2, pp. 157–163. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0922348708702259>
- [48] A. Stepchenko, J. Chizhov, L. Aleksejeva, and J. Tolujew, “Nonlinear, non-stationary and seasonal time series forecasting using different methods coupled with data pre-processing,” *Procedia Computer Science*, vol. 104, pp. 578–585, 2017.
- [49] H. Landau, “Sampling, data transmission, and the nyquist rate,” *Proceedings of the IEEE*, vol. 55, pp. 1701–1706, 1967.
- [50] G. Ellis, *The z-Domain*. Elsevier, 2012, pp. 73–96.

- [51] A. Neupane, N. Raj, R. Deo, and M. Ali, *Development of data-driven models for wind speed forecasting in Australia*. Elsevier, 2021, pp. 143–190.
- [52] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, 7 1948.
- [53] Y. Li, X. Wang, S. Si, and S. Huang, “Entropy based fault classification using the case western reserve university data: A benchmark study,” *IEEE Transactions on Reliability*, vol. 69, pp. 754–767, 6 2020.
- [54] B. Ricaud and B. Torr sani, “A survey of uncertainty principles and some signal processing applications,” *Advances in Computational Mathematics*, vol. 40, pp. 629–650, 6 2014.
- [55] P. Welch, “The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Transactions on Audio and Electroacoustics*, vol. 15, pp. 70–73, 6 1967.
- [56] S. A. Aburakhia, R. Myers, and A. Shami, “A hybrid method for condition monitoring and fault diagnosis of rolling bearings with low system delay,” *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.
- [57] D. L. Donoho and I. M. Johnstone, “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, vol. 81, pp. 425–455, 9 1994.
- [58] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, 3 1998.

- [59] W. Guo and M. J. Zuo, *Adaptive Signal Decomposition Methods for Vibration Signals of Rotating Machinery*. InTech, 5 2017.
- [60] M. Zare and N. M. Nouri, “End-effects mitigation in empirical mode decomposition using a new correlation-based expansion model,” *Mechanical Systems and Signal Processing*, vol. 194, p. 110205, 7 2023.
- [61] Y. Gao, G. Ge, Z. Sheng, and E. Sang, “Analysis and solution to the mode mixing phenomenon in emd,” in *2008 Congress on Image and Signal Processing*, 2008, pp. 223–227.
- [62] Z. WU and N. E. HUANG, “Ensemble empirical mode decomposition: A noise-assisted data analysis method,” *Advances in Adaptive Data Analysis*, vol. 01, pp. 1–41, 1 2009.
- [63] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, pp. 531–544, 2 2014.
- [64] J. S. Smith, “The local mean decomposition and its application to eeg perception data,” *Journal of The Royal Society Interface*, vol. 2, pp. 443–454, 12 2005.
- [65] J. Gilles, “Empirical wavelet transform,” *IEEE Transactions on Signal Processing*, vol. 61, pp. 3999–4010, 8 2013.
- [66] T. Eriksen and N. ur Rehman, “Data-driven nonstationary signal decomposition approaches: a comparative analysis,” *Scientific Reports*, vol. 13, p. 1798, 1 2023.
- [67] M. Civera and C. Surace, “A comparative analysis of signal decomposition techniques for structural health monitoring on an experimental benchmark,” *Sensors*, vol. 21, p. 1825, 3 2021.
- [68] Z. Feng, D. Zhang, and M. J. Zuo, “Adaptive mode decomposition methods and

- their applications in signal analysis for machinery fault diagnosis: A review with examples,” *IEEE Access*, vol. 5, pp. 24 301–24 331, 2017.
- [69] T. Liu, Z. Luo, J. Huang, and S. Yan, “A comparative study of four kinds of adaptive decomposition algorithms and their applications,” *Sensors*, vol. 18, p. 2120, 7 2018.
- [70] G. Lindfield and J. Penny, *Analyzing Data Using Discrete Transforms*. Elsevier, 2019, pp. 383–431.
- [71] A. Tabrizi, L. Garibaldi, A. Fasana, and S. Marchesiello, “Influence of stopping criterion for sifting process of empirical mode decomposition (emd) on roller bearing fault diagnosis,” in *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*, G. Dalpiaz, R. Rubini, G. D’Elia, M. Cocconcelli, F. Chaari, R. Zimroz, W. Bartelmus, and M. Haddar, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 389–398.
- [72] M. Ali, D. M. Khan, I. Saeed, and H. M. Alshanbari, “A new approach to empirical mode decomposition based on akima spline interpolation technique,” *IEEE Access*, vol. 11, pp. 67 370–67 384, 2023.
- [73] L. Cohen, *Time-frequency analysis*. Prentice hall New Jersey, 1995, vol. 778.
- [74] J. E. Torres-Fernández, A. Prieto-Guerrero, and G. Espinosa-Paredes, “Decay ratio estimation using the instantaneous autocorrelation function,” *Nuclear Engineering and Design*, vol. 240, pp. 3559–3570, 10 2010.
- [75] F. Hlawatsch and G. Boudreaux-Bartels, “Linear and quadratic time-frequency signal representations,” *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, 4 1992.
- [76] Y. Wu and X. Li, “Elimination of cross-terms in the wigner–ville distribution of multi-component lfm signals,” *IET Signal Processing*, vol. 11, pp. 657–662, 8 2017.

- [77] P. Flandrin, “Some features of time-frequency representations of multicomponent signals,” in *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 9, 1984, pp. 266–269.
- [78] ———, *Time-frequency time-scale analysis*. Academic press, 1998.
- [79] H.-I. Choi and W. Williams, “Improved time-frequency representation of multicomponent signals using exponential kernels,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 862–871, 6 1989.
- [80] Y. Zhao, L. Atlas, and R. Marks, “The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1084–1091, 7 1990.
- [81] N. Abdennour, T. Ouni, and N. B. Amor, “The importance of signal pre-processing for machine learning: The influence of data scaling in a driver identity classification,” in *2021 IEEE/ACS 18th International Conference on Computer Systems and Applications (AICCSA)*, 11 2021, pp. 1–6.
- [82] N. Mourad, “Robust smoothing of one-dimensional data with missing and/or outlier values,” *IET Signal Processing*, vol. 15, pp. 323–336, 7 2021.
- [83] A. Savitzky and M. J. E. Golay, “Smoothing and differentiation of data by simplified least squares procedures,” *Analytical Chemistry*, vol. 36, pp. 1627–1639, 7 1964.
- [84] S. R. Krishnan and C. S. Seelamantula, “On the selection of optimum savitzky-golay filters,” *IEEE Transactions on Signal Processing*, vol. 61, pp. 380–391, 1 2013.
- [85] Tanu and D. Kakkar, “Accounting for order-frame length tradeoff of savitzky-golay smoothing filters.” *IEEE*, 2 2018, pp. 805–810.
- [86] K. C. B. Rao and B. T. Krishna, “Comparative analysis of integer and non-integer

- order savitzky-golay digital filters,” in *2017 Third Asian Conference on Defence Technology (ACDT)*, 1 2017, pp. 26–31.
- [87] X. Lv, F. Zhou, B. Li, and B. Yan, “Incipient fault feature extraction of rolling bearing based on signal reconstruction,” *Electronics*, vol. 12, p. 3749, 9 2023.
- [88] I. M. A. Wirawan, R. Wardoyo, D. Lelono, and S. Kusrohmaniah, “Comparison of smoothing methods to remove artifacts in emotion recognition based on electroencephalogram signals,” in *2022 Seventh International Conference on Informatics and Computing (ICIC)*, 12 2022, pp. 1–8.
- [89] P. D. Rosero-Montalvo, V. F. López-Batista, and D. H. Peluffo-Ordóñez, “A new data-preprocessing-related taxonomy of sensors for iot applications,” *Information*, vol. 13, p. 241, 5 2022.
- [90] A. Kawala-Sterniuk, M. Podpora, M. Pelc, M. Blaszczynsyn, E. J. Gorzelanczyk, R. Martinek, and S. Ozana, “Comparison of smoothing filters in analysis of eeg data for the medical diagnostics purposes,” *Sensors*, vol. 20, p. 807, 2 2020.
- [91] P. Kowalski and R. Smyk, “Review and comparison of smoothing algorithms for one-dimensional data noise reduction,” in *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, 5 2018, pp. 277–281.
- [92] H. W. Ott, *Noise reduction techniques in electronic systems*, 2nd ed. Wiley-Interscience, 1988, vol. 442.
- [93] V. J. Barclay, R. F. Bonner, and I. P. Hamilton, “Application of wavelet transforms to experimental spectra: smoothing, denoising, and data set compression,” *Analytical Chemistry*, vol. 69, pp. 78–90, 1 1997.
- [94] H. Ashraf, A. Waris, M. Jamil, S. O. Gilani, I. K. Niazi, E. N. Kamavuako, and S. H. N. Gilani, “Determination of optimum segmentation schemes for pattern

- recognition-based myoelectric control: A multi-dataset investigation,” *IEEE Access*, vol. 8, pp. 90 862–90 877, 2020.
- [95] A. Dehghani, O. Sarbishei, T. Glatard, and E. Shihab, “A quantitative comparison of overlapping and non-overlapping sliding windows for human activity recognition using inertial sensors,” *Sensors*, vol. 19, p. 5026, 11 2019.
- [96] B. Fida, I. Bernabucci, D. Bibbo, S. Conforto, and M. Schmid, “Varying behavior of different window sizes on the classification of static and dynamic physical activities from a single accelerometer,” *Medical Engineering Physics*, vol. 37, pp. 705–711, 7 2015.
- [97] V. J. D. Vieira, S. C. Costa, and S. E. N. Correia, “Non-stationarity-based adaptive segmentation applied to voice disorder discrimination,” *IEEE Access*, vol. 11, pp. 54 750–54 759, 2023.
- [98] M. M. Milani, P. E. Abas, and L. C. D. Silva, “A critical review of heart sound signal segmentation algorithms,” *Smart Health*, vol. 24, p. 100283, 6 2022.
- [99] K. Chang, Y. Yoo, and J.-G. Baek, “Anomaly detection using signal segmentation and one-class classification in diffusion process of semiconductor manufacturing,” *Sensors*, vol. 21, p. 3880, 6 2021.
- [100] J. Jirka, M. Prauzek, O. Krejcar, and K. Kuca, “Automatic epilepsy detection using fractal dimensions segmentation and gpdash;svm classification,” *Neuropsychiatric Disease and Treatment*, vol. Volume 14, pp. 2439–2449, 9 2018.
- [101] M. H. M. Noor, Z. Salcic, and K. I.-K. Wang, “Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer,” *Pervasive and Mobile Computing*, vol. 38, pp. 41–59, 7 2017.

- [102] S. Bombiński, K. Błazejak, M. Nejman, and K. Jemielniak, “Sensor signal segmentation for tool condition monitoring,” *Procedia CIRP*, vol. 46, pp. 155–160, 2016.
- [103] A. Prochazka, M. Mudrova, O. Vysata, R. Hava, and C. P. S. Araujo, “Multi-channel eeg signal segmentation and feature extraction,” in *2010 IEEE 14th International Conference on Intelligent Engineering Systems*, 5 2010, pp. 317–320.
- [104] H. O. A. Ahmed and A. K. Nandi, “Vibration image representations for fault diagnosis of rotating machines: A review,” *Machines*, vol. 10, p. 1113, 11 2022.
- [105] C.-L. Yang, Z.-X. Chen, and C.-Y. Yang, “Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images,” *Sensors*, vol. 20, p. 168, 12 2019.
- [106] N. Sobahi, B. Ari, H. Cakar, O. F. Alcin, and A. Sengur, “A new signal to image mapping procedure and convolutional neural networks for efficient schizophrenia detection in eeg recordings,” *IEEE Sensors Journal*, vol. 22, pp. 7913–7919, 4 2022.
- [107] L. Guo, Z. Lu, X. Wen, S. Zhou, and Z. Han, “From signal to image: Capturing fine-grained human poses with commodity wi-fi,” *IEEE Communications Letters*, vol. 24, pp. 802–806, 4 2020.
- [108] P. Singh, P. Pandey, K. Miyapuram, and S. Raman, “Eeg2image: Image reconstruction from eeg brain signals,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6 2023, pp. 1–5.
- [109] K. Gupta and Z. Wei, “Rm-net: Rasterizing markov signals to images for deep learning,” in *2023 International Conference on Cyber-Physical Social Intelligence (ICCSI)*, 10 2023, pp. 654–659.
- [110] C. M. Yilmaz, B. H. Yilmaz, and C. Kose, “A novel signal-to-image conversion approach with ensembles of pretrained cnns for motor imagery eeg signals,” in *2023*

- 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 9 2023, pp. 49–53.
- [111] J. Venton, P. J. Aston, N. A. S. Smith, and P. M. Harris, “Signal to image to classification: Transfer learning for ecg,” in *2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, 7 2020, pp. 1–2.
- [112] B. H. Yilmaz, C. M. Yilmaz, and C. Kose, “Classification of the eeg signals for the cursor movement with the signal-to-image transformation,” in *2019 Medical Technologies Congress (TIPTEKNO)*, 10 2019, pp. 1–4.
- [113] T. Tayeh, S. Aburakhia, R. Myers, and A. Shami, “An attention-based convlstm autoencoder with dynamic thresholding for unsupervised anomaly detection in multivariate time series,” *Machine Learning and Knowledge Extraction*, vol. 4, no. 2, pp. 350–370, 2022. [Online]. Available: <https://www.mdpi.com/2504-4990/4/2/15>
- [114] L. Yang and A. Shami, “A transfer learning and optimized cnn based intrusion detection system for internet of vehicles,” in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 2774–2779.
- [115] A. Moubayed, M. Sharif, M. Luccini, S. Primak, and A. Shami, “Water leak detection survey: Challenges research opportunities using data fusion federated learning,” *IEEE Access*, vol. 9, pp. 40 595–40 611, 2021.
- [116] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, “Machine learning towards intelligent systems: applications, challenges, and opportunities,” *Artificial Intelligence Review*, vol. 54, pp. 3299–3348, 6 2021.
- [117] L. Yang and A. Shami, “Iot data analytics in dynamic environments: From an automated machine learning perspective,” *Engineering Applications of Artificial Intelligence*, vol. 116, p. 105366, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0952197622003803>

- [118] —, “On hyperparameter optimization of machine learning algorithms: Theory and practice,” *Neurocomputing*, vol. 415, pp. 295–316, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231220311693>
- [119] —, “A multi-stage automated online network data stream analytics framework for iiot systems,” *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 2107–2116, 2023.
- [120] V. Narayan, S. Awasthi, N. Fatima, M. Faiz, and S. Srivastava, “Deep learning approaches for human gait recognition: A review,” in *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*, 2 2023, pp. 763–768.
- [121] S. Ahmad, K. Styp-Rekowski, S. Nedelkoski, and O. Kao, “Autoencoder-based condition monitoring and anomaly detection method for rotating machines,” in *2020 IEEE International Conference on Big Data (Big Data)*, 2 2020, pp. 4093–4102.
- [122] M. Liao, C. Liu, C. Wang, and J. Yang, “Research on a rolling bearing fault detection method with wavelet convolution deep transfer learning,” *IEEE Access*, vol. 9, pp. 45 175–45 188, 2021.
- [123] D. Hassan, H. I. Hussein, and M. M. Hassan, “Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis,” *Biomedical Signal Processing and Control*, vol. 79, p. 104019, 2 2023.
- [124] Q. Nawaz, M. Younas, I. Hamid, and S. M. M. Gilani, “Recurrent neural network based human emotion recognition using eeg brain signals,” in *2023 3rd International Conference on Computing and Information Technology (ICCIIT)*, 2 2023, pp. 117–122.
- [125] F. S. Butt, M. F. Wagner, J. Schafer, and D. G. Ullate, “Toward automated feature extraction for deep learning classification of electrocardiogram signals,” *IEEE Access*, vol. 10, pp. 118 601–118 616, 2022.

- [126] S. S. Xu, M.-W. Mak, and C.-C. Cheung, “Towards end-to-end ecg classification with raw signal extraction and deep neural networks,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 1574–1584, 2 2019.
- [127] W. Chen, H. Zhou, L. Cheng, and M. Xia, “Condition monitoring and anomaly detection of wind turbines using temporal convolutional informer and robust dynamic mahalanobis distance,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–14, 2023.
- [128] A. Mohammad-Alikhani, B. Nahid-Mobarakeh, and M.-F. Hsieh, “One-dimensional lstm-regulated deep residual network for data-driven fault detection in electric machines,” *IEEE Transactions on Industrial Electronics*, vol. 71, pp. 3083–3092, 2 2024.
- [129] D. Zhang, Y. Lu, Y. Li, W. Ding, B. Zhang, and J. Xiao, “Frequency learning attention networks based on deep learning for automatic modulation classification in wireless communication,” *Pattern Recognition*, vol. 137, p. 109345, 2 2023.
- [130] B. Zali-Vargahan, A. Charmin, H. Kalbkhani, and S. Barghandan, “Deep time-frequency features and semi-supervised dimension reduction for subject-independent emotion recognition from multi-channel eeg signals,” *Biomedical Signal Processing and Control*, vol. 85, p. 104806, 2 2023.
- [131] W. Jia, M. Sun, J. Lian, and S. Hou, “Feature dimensionality reduction: a review,” *Complex Intelligent Systems*, vol. 8, pp. 2663–2693, 2 2022.
- [132] M. Jalayer, C. Orsenigo, and C. Vercellis, “Fault detection and diagnosis for rotating machinery: A model based on convolutional lstm, fast fourier and continuous wavelet transforms,” *Computers in Industry*, vol. 125, p. 103378, 2 2021.
- [133] Y. Zhang, K. Xing, R. Bai, D. Sun, and Z. Meng, “An enhanced convolutional neural

- network for bearing fault diagnosis based on time–frequency image,” *Measurement*, vol. 157, p. 107667, 2 2020.
- [134] H. V. Dang, H. Tran-Ngoc, T. V. Nguyen, T. Bui-Tien, G. D. Roeck, and H. X. Nguyen, “Data-driven structural health monitoring using feature fusion and hybrid deep learning,” *IEEE Transactions on Automation Science and Engineering*, vol. 18, pp. 2087–2103, 2 2021.
- [135] K. Zaman, M. Sah, C. Direkoglu, and M. Unoki, “A survey of audio classification using deep learning,” *IEEE Access*, vol. 11, pp. 106 620–106 649, 2023.
- [136] Y. Miao, Y. V. Zakharov, H. Sun, J. Li, and J. Wang, “Underwater acoustic signal classification based on sparse time–frequency representation and deep learning,” *IEEE Journal of Oceanic Engineering*, vol. 46, pp. 952–962, 2 2021.
- [137] N. P. G. Seshadri, S. Agrawal, B. K. Singh, B. Geethanjali, V. Mahesh, and R. B. Pachori, “Eeg based classification of children with learning disabilities using shallow and deep neural network,” *Biomedical Signal Processing and Control*, vol. 82, p. 104553, 2 2023.
- [138] H. Altaheri, G. Muhammad, M. Alsulaiman, S. U. Amin, G. A. Altuwaijri, W. Abdul, M. A. Bencherif, and M. Faisal, “Deep learning techniques for classification of electroencephalogram (eeg) motor imagery (mi) signals: a review,” *Neural Computing and Applications*, vol. 35, pp. 14 681–14 722, 2 2023.
- [139] L. Ilias, D. Askounis, and J. Psarras, “Multimodal detection of epilepsy with deep neural networks,” *Expert Systems with Applications*, vol. 213, p. 119010, 2 2023.
- [140] S.-H. Kim, C.-B. Moon, J.-W. Kim, and D.-S. Kim, “A hybrid deep learning model for automatic modulation classification,” *IEEE Wireless Communications Letters*, vol. 11, pp. 313–317, 2 2022.

- [141] S. Zheng, X. Zhou, L. Zhang, P. Qi, K. Qiu, J. Zhu, and X. Yang, “Toward next-generation signal intelligence: A hybrid knowledge and data-driven deep learning framework for radio signal classification,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, pp. 564–579, 2 2023.
- [142] T. Chen, S. Gao, S. Zheng, S. Yu, Q. Xuan, C. Lou, and X. Yang, “Emd and vmd empowered deep learning for radio modulation recognition,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, pp. 43–57, 2 2023.
- [143] W. Li, K. Wang, L. You, and Z. Huang, “A new deep learning framework for hf signal detection in wideband spectrogram,” *IEEE Signal Processing Letters*, vol. 29, pp. 1342–1346, 2022.
- [144] K. Qiu, S. Zheng, L. Zhang, C. Lou, and X. Yang, “Deepsig: A hybrid heterogeneous deep learning framework for radio signal classification,” *IEEE Transactions on Wireless Communications*, p. 1, 2023.
- [145] Y. Lei, Z. He, and Y. Zi, “A new approach to intelligent fault diagnosis of rotating machinery,” *Expert Systems with Applications*, vol. 35, pp. 1593–1600, 2 2008.
- [146] B. R. Nayana and P. Geethanjali, “Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal,” *IEEE Sensors Journal*, vol. 17, pp. 5618–5625, 2 2017.
- [147] P. Geethanjali, Y. K. Mohan, and J. Sen, “Time domain feature extraction and classification of eeg data for brain computer interface,” in *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*, 2 2012, pp. 1136–1139.
- [148] X. Wang, Y. Zheng, Z. Zhao, and J. Wang, “Bearing fault diagnosis based on statistical locally linear embedding,” *Sensors*, vol. 15, pp. 16 225–16 247, 2 2015.

- [149] S. M. S. Moctar, I. Rida, and S. Boudaoud, “Time-domain features for semg signal classification: A brief survey,” *JETSAN 2023*, 2023.
- [150] A. Mert, N. Kılıç, and A. Akan, “Evaluation of bagging ensemble method with time-domain feature extraction for diagnosing of arrhythmia beats,” *Neural Computing and Applications*, vol. 24, pp. 317–326, 2 2014.
- [151] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, “Sensor fault classification based on support vector machine and statistical time-domain features,” *IEEE Access*, vol. 5, pp. 8682–8690, 2017.
- [152] D. Tkach, H. Huang, and T. A. Kuiken, “Study of stability of time-domain features for electromyographic pattern recognition,” *Journal of NeuroEngineering and Rehabilitation*, vol. 7, p. 21, 2 2010.
- [153] B. Sreejith, A. K. Verma, and A. Srividya, “Fault diagnosis of rolling element bearing using time-domain features and neural networks,” in *2008 IEEE Region 10 and the Third international Conference on Industrial and Information Systems*, 2 2008, pp. 1–6.
- [154] Y. Paul, V. Goyal, and R. A. Jaswal, “Comparative analysis between svm amp; knn classifier for emg signal classification on elementary time domain features,” in *2017 4th International Conference on Signal Processing, Computing and Control (ISPCC)*, 2 2017, pp. 169–175.
- [155] R. Barioul, S. Fakhfakh, H. Derbel, and O. Kanoun, “Evaluation of emg signal time domain features for hand gesture distinction,” in *2019 16th International Multi-Conference on Systems, Signals Devices (SSD)*, 2 2019, pp. 489–493.
- [156] F. O.K. and R. R., “Time-domain exponential energy for epileptic eeg signal classification,” *Neuroscience Letters*, vol. 694, pp. 1–8, 2 2019.

- [157] K. Nazarpour, A. R. Sharafat, and S. M. Firoozabadi, “A novel feature extraction scheme for myoelectric signals classification using higher order statistics,” in *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005.*, pp. 293–296.
- [158] K. Nazarpour, A. R. Sharafat, and S. M. P. Firoozabadi, “Application of higher order statistics to surface electromyogram signal classification,” *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 1762–1769, 2 2007.
- [159] J. Chen, B. Xu, and X. Zhang, “A vibration feature extraction method based on time-domain dimensional parameters and mahalanobis distance,” *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, 7 2021.
- [160] A. Saci, A. Al-Dweik, and A. Shami, “Autocorrelation integrated gaussian based anomaly detection using sensory data in industrial manufacturing,” *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9231–9241, 2021.
- [161] E. Ali, R. K. Udhayakumar, M. Angelova, and C. Karmakar, “Performance analysis of entropy methods in detecting epileptic seizure from surface electroencephalograms,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2 2021, pp. 1082–1085.
- [162] W. Jiao, G. Li, Y. Jiang, R. Baim, C. Tang, T. Yan, X. Ding, and Y. Yan, “Multi-scale sample entropy-based energy moment features applied to fault classification,” *IEEE Access*, vol. 9, pp. 8444–8454, 2021.
- [163] C. Liang and C. Chen, “Generalized composite multiscale diversity entropy and its application for fault diagnosis of rolling bearing in automotive production line,” *IEEE Access*, vol. 9, pp. 84 545–84 558, 2021.
- [164] M. Li and W. Chen, “Fft-based deep feature learning method for eeg classification,” *Biomedical Signal Processing and Control*, vol. 66, p. 102492, 2 2021.

- [165] N. F. Güler and S. Koçer, “Classification of emg signals using pca and fft,” *Journal of Medical Systems*, vol. 29, pp. 241–250, 2 2005.
- [166] M. K. Delimayanti, B. Purnama, N. G. Nguyen, M. R. Faisal, K. R. Mahmudah, F. Indriani, M. Kubo, and K. Satou, “Classification of brainwaves for sleep stages by high-dimensional fft features from eeg signals,” *Applied Sciences*, vol. 10, p. 1797, 2 2020.
- [167] E. C. Djamal, W. I. Furi, and F. Nugraha, “Detection of eeg signal post-stroke using fft and convolutional neural network,” in *2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2 2019, pp. 18–23.
- [168] H. Arabaci and A. Mohamed, “A knowledge-based diagnosis algorithm for broken rotor bar fault classification using fft, principal component analysis and support vector machines,” *International Journal of Intelligent Engineering Informatics*, vol. 8, p. 19, 2020.
- [169] C. Altın and O. Er, “Comparison of different time and frequency domain feature extraction methods on elbow gesture’s emg,” *European Journal of Interdisciplinary Studies*, vol. 2, p. 35, 2 2016.
- [170] S. K. Narudin, N. M. Nasir, A. Joret, N. Fuad, B. S. K. K. Ibrahim, and M. S. Huq, “Classification of hand movements of stroke patients using combination of statistical features and artificial neural network,” in *AIP Conference Proceedings*, 2023.
- [171] L. S. Dhamande and M. B. Chaudhari, “Compound gear-bearing fault feature extraction using statistical features based on time-frequency method,” *Measurement*, vol. 125, pp. 63–77, 2 2018.
- [172] M. Murugappan and S. Murugappan, “Human emotion recognition through short time electroencephalogram (eeg) signals using fast fourier transform (fft),” in *2013*

- IEEE 9th International Colloquium on Signal Processing and its Applications*, 2013, pp. 289–294.
- [173] S. A. Shufni and M. Y. Mashor, “Ecg signals classification based on discrete wavelet transform, time domain and frequency domain features,” in *2015 2nd International Conference on Biomedical Engineering (ICoBE)*, 2015, pp. 1–6.
- [174] M. Altaf, T. Akram, M. A. Khan, M. Iqbal, M. M. I. Ch, and C.-H. Hsu, “A new statistical features based approach for bearing fault diagnosis using vibration signals,” *Sensors*, vol. 22, p. 2012, 2022.
- [175] L. Jiang, J. He, H. Pan, D. Wu, T. Jiang, and J. Liu, “Seizure detection algorithm based on improved functional brain network structure feature extraction,” *Biomedical Signal Processing and Control*, vol. 79, p. 104053, 2023.
- [176] P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, “Fourier-based feature extraction for classification of eeg signals using eeg rhythms,” *Circuits, Systems, and Signal Processing*, vol. 35, pp. 3700–3715, 2016.
- [177] M. N. Alam, M. I. Ibrahimy, and S. M. A. Motakabber, “Feature extraction of eeg signal by power spectral density for motor imagery based bci,” in *2021 8th International Conference on Computer and Communication Engineering (ICCCE)*, 2021, pp. 234–237.
- [178] Q. Gao, A. H. Omran, Y. Baghersad, O. Mohammadi, M. A. Alkhafaji, A. K. J. Al-Azzawi, S. H. Al-Khafaji, N. Emami, D. Toghraie, and M. J. Golkar, “Electroencephalogram signal classification based on fourier transform and pattern recognition network for epilepsy diagnosis,” *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106479, 2023.
- [179] A. Zhang, B. Yang, and L. Huang, “Feature extraction of eeg signals using power

- spectral entropy,” in *2008 International Conference on BioMedical Engineering and Informatics*, 2 2008, pp. 435–439.
- [180] S. Babeetha and S. S. Sridhar, *EEG Signal Feature Extraction Using Principal Component Analysis and Power Spectral Entropy for Multiclass Emotion Prediction*, 2023, pp. 435–448.
- [181] D. Konar, S. De, P. Mukherjee, and A. H. Roy, “A novel human stress level detection technique using eeg,” in *2023 International Conference on Network, Multimedia and Information Technology (NMITCON)*, 2 2023, pp. 1–6.
- [182] I. Shaer and A. Shami, “Sound event classification in an industrial environment: Pipe leakage detection use case,” in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, 2022, pp. 1212–1217.
- [183] G. R. D, “Analysis of mfcc features for eeg signal classification,” *International Journal of Advances in Signal and Image Sciences*, vol. 2, p. 14, 12 2016.
- [184] M. Deng, T. Meng, J. Cao, S. Wang, J. Zhang, and H. Fan, “Heart sound classification based on improved mfcc features and convolutional recurrent neural networks,” *Neural Networks*, vol. 130, pp. 22–32, 10 2020.
- [185] Y. Arpitha, G. L. Madhumathi, and N. Balaji, “Spectrogram analysis of eeg signal and classification efficiency using mfcc feature extraction technique,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 757–767, 2 2022.
- [186] M. Boussaa, I. Atouf, M. Atibi, and A. Bennis, “Ecg signals classification using mfcc coefficients and ann classifier,” in *2016 International Conference on Electrical and Information Technologies (ICEIT)*, 5 2016, pp. 480–484.
- [187] S. A. A. Yusuf and R. Hidayat, “Mfcc feature extraction and knn classification in

- ecg signals,” in *2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, 9 2019, pp. 1–5.
- [188] M. Mohan, P. Dhanalakshmi, and R. S. Kumar, “Speech emotion classification using ensemble models with mfcc,” *Procedia Computer Science*, vol. 218, pp. 1857–1868, 2023.
- [189] A. Abbaskhah, H. Sedighi, and H. Marvi, “Infant cry classification by mfcc feature extraction with mlp and cnn structures,” *Biomedical Signal Processing and Control*, vol. 86, p. 105261, 9 2023.
- [190] X. Mu and C.-H. Min, “Mfcc as features for speaker classification using machine learning,” in *2023 IEEE World AI IoT Congress (AllIoT)*, 6 2023, pp. 0566–0570.
- [191] F. Alharbi, S. Luo, H. Zhang, K. Shaukat, G. Yang, C. A. Wheeler, and Z. Chen, “A brief review of acoustic and vibration signal-based fault detection for belt conveyor idlers using machine learning models,” *Sensors*, vol. 23, p. 1902, 2 2023.
- [192] F. Rustam, A. Ishaq, M. S. A. Hashmi, H. U. R. Siddiqui, L. A. D. López, J. C. Galán, and I. Ashraf, “Railway track fault detection using selective mfcc features from acoustic data,” *Sensors*, vol. 23, p. 7018, 8 2023.
- [193] Z. K. Abdul and A. K. Al-Talabani, “Mel frequency cepstral coefficient and its applications: A review,” *IEEE Access*, vol. 10, pp. 122 136–122 158, 2022.
- [194] A. P. Petropulu, “Higher-order spectra in biomedical signal processing,” *IFAC Proceedings Volumes*, vol. 27, pp. 47–52, 3 1994.
- [195] C. Zhu, T. Cao, L. Chen, X. Dai, Q. Ge, and X. Zhao, “High-order domain feature extraction technology for ocean acoustic observation signals: A review,” *IEEE Access*, vol. 11, pp. 17 665–17 683, 2023.

- [196] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, “Application of higher order statistics/spectra in biomedical signals—a review,” *Medical Engineering Physics*, vol. 32, pp. 679–689, 9 2010.
- [197] S. A. Hosseini, “A hybrid approach based on higher order spectra for clinical recognition of seizure and epilepsy using brain activity,” *Basic and Clinical Neuroscience Journal*, vol. 8, pp. 479–492, 11 2017.
- [198] E. B. Assi, L. Gagliano, S. Rihana, D. K. Nguyen, and M. Sawan, “Bispectrum features and multilayer perceptron classifier to enhance seizure prediction,” *Scientific Reports*, vol. 8, p. 15491, 10 2018.
- [199] K. Tan, W. Yan, L. Zhang, Q. Ling, and C. Xu, “Semi-supervised specific emitter identification based on bispectrum feature extraction cgan in multiple communication scenarios,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, pp. 292–310, 2 2023.
- [200] Y. Zhou, R. Zhou, C. Wang, X. Wang, Y. Yu, and X. Zhang, “A modulation recognition method based on bispectrum and ensemble learning,” in *2022 2nd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, 9 2022, pp. 124–128.
- [201] X. Li, M. Yu, Y. Liu, and X. Xu, “Feature extraction of underwater signals based on bispectrum estimation,” in *2011 7th International Conference on Wireless Communications, Networking and Mobile Computing*, 9 2011, pp. 1–4.
- [202] Z. Dong, F. Lv, T. Wan, K. Jiang, X. Fang, and L. Zhang, “Radar signal modulation recognition based on bispectrum features and deep learning,” in *2021 International Conference on Computer Engineering and Application (ICCEA)*, 6 2021, pp. 63–67.
- [203] Y. Liu, Y. Zhou, W. Xin, Y. He, and P. Fan, “Bispectrum analysis for feature extrac-

- tion of pitting fault in wind turbine gearbox,” in *2010 IEEE International Conference on Mechatronics and Automation*, 8 2010, pp. 1298–1301.
- [204] D. G. Bhalke, C. B. R. Rao, and D. S. Bormane, “Musical instrument classification using higher order spectra,” in *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*, 2 2014, pp. 40–45.
- [205] A. C. McCormick and A. K. Nandi, “Bispectral and trispectral features for machine condition diagnosis,” *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 146, p. 229, 1999.
- [206] K. Muthuvel, L. P. Suresh, T. J. Alexander, and S. H. K. Veni, *Classification of ECG Signal Using Hybrid Feature Extraction and Neural Network Classifier*, 2015, pp. 1537–1544.
- [207] M. K. J. A. T. M. P. T. Thangam, and P. S. L., “Hybrid feature extraction technique for electrocardiograph arrhythmia signal classification,” in *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 8 2023, pp. 1518–1524.
- [208] C. K. Kovach and M. A. Howard, “Decomposition of higher-order spectra for blind multiple-input deconvolution, pattern identification and separation,” *Signal Processing*, vol. 165, pp. 357–379, 12 2019.
- [209] M. K. Gourisaria, R. Agrawal, M. Sahni, and P. K. Singh, “Comparative analysis of audio classification with mfcc and stft features using machine learning techniques,” *Discover Internet of Things*, vol. 4, p. 1, 1 2024.
- [210] A.-C. Tsai, J.-J. Luh, and T.-T. Lin, “A novel stft-ranking feature of multi-channel emg for motion pattern recognition,” *Expert Systems with Applications*, vol. 42, pp. 3327–3341, 5 2015.

- [211] T. Huynh-The, Q.-V. Pham, T.-H. Vu, D. B. da Costa, and V.-P. Hoang, “Intelligent spectrum sensing with convnet for 5g and lte signals identification,” in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, 7 2023, pp. 140–144.
- [212] J. Huang, B. Chen, B. Yao, and W. He, “Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network,” *IEEE Access*, vol. 7, pp. 92 871–92 880, 2019.
- [213] Y. Hu, L. Wang, and W. Fu, “Eeg feature extraction of motor imagery based on wt and stft,” in *2018 IEEE International Conference on Information and Automation (ICIA)*, 8 2018, pp. 83–88.
- [214] O. K. Cura, G. C. Yilmaz, H. S. Ture, and A. Akan, “Deep time-frequency feature extraction for alzheimer’s dementia eeg classification,” in *2022 Medical Technologies Congress (TIPTEKNO)*, 10 2022, pp. 1–4.
- [215] A. Buyya, T. Ogeti, G. Suhas, P. Kashapogula, and A. K. Panigrahy, “Arrhythmias classification by using stft-based spectrograms, transfer learning and concatenation of features,” in *2023 4th International Conference for Emerging Technology (IN-CET)*, 5 2023, pp. 1–5.
- [216] H. Liu, L. Li, and J. Ma, “Rolling bearing fault diagnosis based on stft-deep learning and sound signals,” *Shock and Vibration*, vol. 2016, pp. 1–12, 2016.
- [217] P. Ewert, B. Wicher, and T. Pajchrowski, “Application of the stft for detection of the rotor unbalance of a servo-drive system with an elastic interconnection,” *Electronics*, vol. 13, p. 441, 1 2024.
- [218] Y. Hao, C. Zhang, Y. Lu, L. Zhang, Z. Lei, and Z. Li, “A novel autoencoder modeling method for intelligent assessment of bearing health based on short-time fourier transform and ensemble strategy,” *Precision Engineering*, vol. 85, pp. 89–101, 1 2024.

- [219] H. Shayeste and B. M. Asl, “Automatic seizure detection based on gray level co-occurrence matrix of stft imaged-eeeg,” *Biomedical Signal Processing and Control*, vol. 79, p. 104109, 1 2023.
- [220] Z. Suryady, A. W. A. Rahman, and D. Oktarina, “Human emotion classification based on phonocardiography signals (pcg),” in *AIP Conf. Proc.*, 2024, p. 020002.
- [221] N. Sharma, G. G, and R. S. Anand, “Epileptic seizure detection using stft based peak mean feature and support vector machine,” in *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, 8 2021, pp. 1131–1136.
- [222] H. Zehir, T. Hafs, S. Daas, and A. Nait-Ali, “An eeg biometric system based on empirical mode decomposition and hilbert-huang transform for improved feature extraction,” in *2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART)*, 6 2023, pp. 1–4.
- [223] L. Li, S. Wang, Y. Zhao, and F. Wang, “Fault diagnosis for single-phase grounding in distribution network based on hilbert-huang transform and siamese convolution neural network,” in *2023 35th Chinese Control and Decision Conference (CCDC)*, 5 2023, pp. 3571–3576.
- [224] P. Warule, S. P. Mishra, S. Deb, and D. Joshi, “Hilbert-huang transform-based time-frequency analysis of speech signals for the identification of common cold,” in *TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)*, 10 2023, pp. 876–880.
- [225] Y. P. Bórnea, A. L. O. Vitor, M. F. Castoldi, A. Goedtel, and W. A. Souza, “Classification of bearing faults in induction motors with the hilbert-huang transform and feature selection,” in *2023 IEEE 14th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, 8 2023, pp. 567–573.

- [226] D. Walsh, S. Dev, and A. Nag, “Hilbert-huang-transform based features for accent classification of non-native english speakers,” in *2023 34th Irish Signals and Systems Conference (ISSC)*, 6 2023, pp. 1–6.
- [227] B. S. Munir, M. Reza, A. Trisetyarso, and B. S. Abbas, “Feature extraction using hilbert-huang transform for power system oscillation measurements,” in *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 10 2017, pp. 93–96.
- [228] D. Du, J. Zhang, Y. Fang, and J. Tian, “Motor bearing fault diagnosis based on hilbert-huang transform and convolutional neural networks,” in *2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific)*, 10 2022, pp. 1–5.
- [229] D. H. Kisa, M. A. Ozdemir, O. Guren, and A. Akan, “Classification of hand gestures using semg signals and hilbert-huang transform,” in *2022 30th European Signal Processing Conference (EUSIPCO)*, 8 2022, pp. 1293–1297.
- [230] N. M. Bhiri, S. Ameer, I. Jegham, M. A. Mahjoub, and A. B. Khalifa, “Fisher-hht: A feature extraction approach for hand gesture recognition with a leap motion controller,” in *2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 5 2022, pp. 1–6.
- [231] J. Li, Z. Liang, Y. Qi, L. Li, G. Li, N. Zheng, and Z. Wang, “Feature extraction and state identification of transformer dc bias vibration based on hht-svm,” in *2022 7th Asia Conference on Power and Electrical Engineering (ACPEE)*, 4 2022, pp. 746–750.
- [232] J. Chen, H. Li, L. Ma, H. Bo, and X. Gao, “Application of eemd-hht method on eeg analysis for speech evoked emotion recognition,” in *2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 8 2020, pp. 376–381.

- [233] K. Cai, W. Cao, L. Aarniovuori, H. Pang, Y. Lin, and G. Li, “Classification of power quality disturbances using wigner-ville distribution and deep convolutional neural networks,” *IEEE Access*, vol. 7, pp. 119 099–119 109, 2019.
- [234] M. Zulfiqar, M. F. U. Butt, A. Rama, and I. Shafi, “Abnormality detection in cardiac signals using pseudo wigner-ville distribution with pre-trained convolutional neural network,” in *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 12 2019, pp. 1–5.
- [235] G. Wei, T. Jiang, Y. Liu, and W. Liu, “Dynamic gesture recognition using software defined radio based on smooth pseudo wigner-ville distribution,” in *2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 12 2015, pp. 1–5.
- [236] F. Akdeniz, I. Kayikcioglu, I. Kaya, and T. Kayikcioglu, “Using wigner-ville distribution in ecg arrhythmia detection for telemedicine applications,” in *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, 6 2016, pp. 409–412.
- [237] R. Abayneh, M. Mutala, J. Ren, and X. Yang, “Efficient wi-fi device rf fingerprint detection and classification using wigner-ville distribution,” in *2023 8th International Conference on Communication, Image and Signal Processing (CCISP)*, 11 2023, pp. 504–508.
- [238] S. Taran, S. K. Khare, P. V. K. Krishna, and V. S. Kardam, *Focal and non-focal EEG signal classification using the Wigner–Ville distribution and deep feature extraction*. IOP Publishing, 2 2024, pp. 2–1–2–14.
- [239] K. Kumar, P. Kumar, R. K. Patel, M. Sharma, V. Bajaj, and U. R. Acharya, “Time frequency distribution and deep neural network for automated identification of in-

- somnia using single channel eeg-signals,” *IEEE Latin America Transactions*, vol. 22, pp. 186–194, 3 2024.
- [240] X. Zou, H. Mu, R. Wang, K. Fan, Z. Cheng, Y. He, and G. Zhang, “An efficient accuracy improvement method for cable defect location based on instantaneous filtering in time-frequency domain,” *Measurement*, vol. 226, p. 114178, 2 2024.
- [241] S. K. Dash and G. S. Rao, “Arrhythmia detection using wigner-ville distribution based neural network,” *Procedia Computer Science*, vol. 85, pp. 806–811, 2016.
- [242] P. Shegokar and P. Sircar, “Continuous wavelet transform based speech emotion recognition,” in *2016 10th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 12 2016, pp. 1–8.
- [243] C. Peng, W. Cheng, Z. Song, and R. Dong, “A noise-robust modulation signal classification method based on continuous wavelet transform,” in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, 6 2020, pp. 745–750.
- [244] R. N. Toma, F. H. Toma, and J.-M. Kim, “Comparative analysis of continuous wavelet transforms on vibration signal in bearing fault diagnosis of induction motor,” in *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)*, 9 2021, pp. 1–4.
- [245] A. Belkhou, A. Achmamad, and A. Jbari, “Classification and diagnosis of myopathy emg signals using the continuous wavelet transform,” in *2019 Scientific Meeting on Electrical-Electronics Biomedical Engineering and Computer Science (EBBT)*, 4 2019, pp. 1–4.
- [246] R. Singh, S. Gangwar, S. Bartwal, and V. Tomar, “Wavelet-based feature extraction and rnn for ecg signal classification,” in *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*, 6 2023, pp. 1–6.

- [247] J. Fu, X. Xu, and Y. Chen, “Feature selection and optimization method based on cwt,” in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 10 2017, pp. 1–4.
- [248] N. A. Aditiya, Z. Darojah, D. R. Sanggar, and M. R. Dharmawan, “Fault diagnosis system of rotating machines using continuous wavelet transform and artificial neural network,” in *2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, 9 2017, pp. 173–176.
- [249] S. Tang, B. C. Khoo, Y. Zhu, K. M. Lim, and S. Yuan, “A light deep adaptive framework toward fault diagnosis of a hydraulic piston pump,” *Applied Acoustics*, vol. 217, p. 109807, 2 2024.
- [250] J. Dong, H. Jiang, D. Su, Y. Gao, T. Chen, and K. Sheng, “Transfer learning rolling bearing fault diagnosis model based on deep feature decomposition and class-level alignment,” *Measurement Science and Technology*, vol. 35, p. 046006, 4 2024.
- [251] C. Bhagavatula, S. Venugopalan, R. Blue, R. Friedman, M. O. Griofa, M. Savvides, and B. V. K. V. Kumar, “Biometric identification of cardiosynchronous waveforms utilizing person specific continuous and discrete wavelet transform features,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 8 2012, pp. 4311–4314.
- [252] O. Almanza-Conejo, D. L. Almanza-Ojeda, J. L. Contreras-Hernandez, and M. A. Ibarra-Manzano, “Emotion recognition in eeg signals using the continuous wavelet transform and cnns,” *Neural Computing and Applications*, vol. 35, pp. 1409–1422, 1 2023.
- [253] R. Dwyer, “Detection of non-gaussian signals by frequency domain kurtosis estimation,” in *ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 8, 1983, pp. 607–610.

- [254] J. Antoni, “The spectral kurtosis: a useful tool for characterising non-stationary signals,” *Mechanical Systems and Signal Processing*, vol. 20, pp. 282–307, 2 2006.
- [255] Y. Wang, J. Xiang, R. Markert, and M. Liang, “Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications,” *Mechanical Systems and Signal Processing*, vol. 66-67, pp. 679–698, 1 2016.
- [256] S. Liu, S. Hou, K. He, and W. Yang, “L-kurtosis and its application for fault detection of rolling element bearings,” *Measurement*, vol. 116, pp. 523–532, 2 2018.
- [257] J. Wodecki, “Time-varying spectral kurtosis: Generalization of spectral kurtosis for local damage detection in rotating machines under time-varying operating conditions,” *Sensors*, vol. 21, p. 3590, 5 2021.
- [258] I. Ahmad, C. Yao, L. Li, Y. Chen, Z. Liu, I. Ullah, M. Shabaz, X. Wang, K. Huang, G. Li, G. Zhao, O. W. Samuel, and S. Chen, “An efficient feature selection and explainable classification method for eeg-based epileptic seizure detection,” *Journal of Information Security and Applications*, vol. 80, p. 103654, 2 2024.
- [259] Y. Liu, F. Yu, and J. Gao, “Sparse representation-based classification for rolling bearing fault diagnosis,” in *2019 Chinese Automation Congress (CAC)*, 11 2019, pp. 3058–3061.
- [260] J. C. Chan, H. Ma, and T. K. Saha, “Partial discharge pattern recognition using multiscale feature extraction and support vector machine,” in *2013 IEEE Power Energy Society General Meeting*, 2013, pp. 1–5.
- [261] S. Sekkate, M. Khalil, and A. Adib, “Fusing wavelet and short-term features for speaker identification in noisy environment,” in *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 4 2018, pp. 1–8.

- [262] M. Tibdewal, H. Dey, M. Mahadevappa, A. Ray, and M. Malokar, “Wavelet transform based multiple features extraction for detection of epileptic/ non-epileptic multichannel eeg,” in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 1648–1653.
- [263] Y. Xie and T. Zhang, “Feature extraction based on dwt and cnn for rotating machinery fault diagnosis,” in *2017 29th Chinese Control And Decision Conference (CCDC)*, 5 2017, pp. 3861–3866.
- [264] F. J. Floresca, C. K. Tobias, and C. Ostia, “Naïve bayes classification technique for brushless dc motor fault diagnosis with discrete wavelet transform feature extraction,” in *2022 14th International Conference on Computer and Automation Engineering (ICCAE)*, 3 2022, pp. 140–144.
- [265] S. Katiyar and I. Sharma, “Emg signal decomposition methods for physical actions: A comparative analysis,” in *2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS)*, 9 2023, pp. 1–5.
- [266] C. M. Ugwu, C. P. Mukamakuza, and E. Tuyishimire, “Ecg-signals-based heartbeat classification: A comparative study of artificial neural network and support vector machine classifiers,” in *2024 IEEE 22nd World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 1 2024, pp. 000 217–000 222.
- [267] S. Guo, S. Lin, and Z. Huang, “Feature extraction of p300s in eeg signal with discrete wavelet transform and fisher criterion,” in *2015 8th International Conference on Biomedical Engineering and Informatics (BMEI)*, 10 2015, pp. 200–204.
- [268] H. U. Amin, A. S. Malik, R. F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, and W.-T. Chooi, “Feature extraction and classification for eeg signals using wavelet transform and machine learning techniques,” *Australasian Physical Engineering Sciences in Medicine*, vol. 38, pp. 139–149, 3 2015.

- [269] A. Prochazka, J. Kukal, and O. Vysata, “Wavelet transform use for feature extraction and eeg signal segments classification,” in *2008 3rd International Symposium on Communications, Control and Signal Processing*, 3 2008, pp. 719–722.
- [270] Q. Zhao and L. Zhang, “Ecg feature extraction and classification using wavelet transform and support vector machines,” in *2005 International Conference on Neural Networks and Brain*, pp. 1089–1092.
- [271] S. Y. Shah, H. Larijani, R. M. Gibson, and D. Liarokapis, “Epileptic seizure classification based on random neural networks using discrete wavelet transform for electroencephalogram signal decomposition,” *Applied Sciences*, vol. 14, p. 599, 1 2024.
- [272] Q. Gong, K. Peng, Q. Gao, L. Feng, and C. Xiao, “Series arc fault identification method based on wavelet transform and feature values decomposition fusion dnn,” *Electric Power Systems Research*, vol. 221, p. 109391, 8 2023.
- [273] J. Xi and L. Lin, “Multidimensional feature extraction based on vibration signals of rolling bearings,” in *2013 25th Chinese Control and Decision Conference (CCDC)*, 5 2013, pp. 1093–1096.
- [274] L.-Y. Zhao, L. Wang, and R.-Q. Yan, “Rolling bearing fault diagnosis based on wavelet packet decomposition and multi-scale permutation entropy,” *Entropy*, vol. 17, pp. 6447–6461, 9 2015.
- [275] M. J. Gómez, P. Marklund, D. Strombergsson, C. Castejón, and J. C. García-Prada, “Analysis of vibration signals of drivetrain failures in wind turbines for condition monitoring,” *Experimental Techniques*, vol. 45, pp. 1–12, 2 2021.
- [276] X. Ding, L. Lin, D. He, L. Wang, W. Huang, and Y. Shao, “A weight multinet architecture for bearing fault classification under complex speed conditions,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

- [277] I. Attoui, N. Boutasseta, N. Fergani, B. Oudjani, and A. Deliou, “Vibration-based bearing fault diagnosis by an integrated dwt-fft approach and an adaptive neuro-fuzzy inference system,” in *2015 3rd International Conference on Control, Engineering Information Technology (CEIT)*, 5 2015, pp. 1–6.
- [278] S. Aburakhia and A. Shami, “Sb-pdm: A tool for predictive maintenance of rolling bearings based on limited labeled data,” *Software Impacts*, vol. 16, p. 100503, 5 2023.
- [279] Z. Mezni, C. Delpha, D. Diallo, and A. Braham, “A comparative study for ball bearing fault classification using kernel-svm with kullback leibler divergence selected features,” in *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, 10 2019, pp. 6969–6974.
- [280] F. Wang, G. Deng, C. Liu, W. Su, Q. Han, and H. Li, “A deep feature extraction method for bearing fault diagnosis based on empirical mode decomposition and kernel function,” *Advances in Mechanical Engineering*, vol. 10, p. 168781401879825, 9 2018.
- [281] J. Shen, X. Zhang, G. Wang, Z. Ding, and B. Hu, “An improved empirical mode decomposition of electroencephalogram signals for depression detection,” *IEEE Transactions on Affective Computing*, vol. 13, pp. 262–271, 1 2022.
- [282] L. Zhan, F. Ma, J. Zhang, C. Li, Z. Li, and T. Wang, “Fault feature extraction and diagnosis of rolling bearings based on enhanced complementary empirical mode decomposition with adaptive noise and statistical time-domain features,” *Sensors*, vol. 19, p. 4047, 9 2019.
- [283] A. Ahmed, Y. Serrestou, K. Raouf, and J.-F. Diouris, “Empirical mode decomposition-based feature extraction for environmental sound classification,” *Sensors*, vol. 22, p. 7717, 10 2022.

- [284] H. Shang, K. Lo, and F. Li, “Partial discharge feature extraction based on ensemble empirical mode decomposition and sample entropy,” *Entropy*, vol. 19, p. 439, 8 2017.
- [285] S. Nalband, A. Prince, and A. Agrawal, “Entropy-based feature extraction and classification of vibroarthrographic signal using complete ensemble empirical mode decomposition with adaptive noise,” *IET Science, Measurement Technology*, vol. 12, pp. 350–359, 5 2018.
- [286] M. Van and H. Kang, “Bearing-fault diagnosis using non-local means algorithm and empirical mode decomposition-based feature extraction and two-stage feature selection,” *IET Science, Measurement Technology*, vol. 9, pp. 671–680, 9 2015.
- [287] Y. Li, J. hao Luo, Q. yun Dai, J. K. Eshraghian, B. W.-K. Ling, C. yan Zheng, and X. li Wang, “A deep learning approach to cardiovascular disease classification using empirical mode decomposition for ecg feature extraction,” *Biomedical Signal Processing and Control*, vol. 79, p. 104188, 1 2023.
- [288] N. M. El-Kafrawy, D. Hegazy, and M. F. Tolba, *Features Extraction and Classification of EEG Signals Using Empirical Mode Decomposition and Support Vector Machine*, 2014, pp. 189–198.
- [289] R. Jia, F. Ma, H. Wu, X. Luo, and X. Ma, “Coupling fault feature extraction method based on bivariate empirical mode decomposition and full spectrum for rotating machinery,” *Mathematical Problems in Engineering*, vol. 2018, pp. 1–10, 2018.
- [290] H. Yang, T. Ning, B. Zhang, X. Yin, and Z. Gao, “An adaptive denoising fault feature extraction method based on ensemble empirical mode decomposition and the correlation coefficient,” *Advances in Mechanical Engineering*, vol. 9, p. 168781401769644, 4 2017.

- [291] A. Mert, “Ecg feature extraction based on the bandwidth properties of variational mode decomposition,” *Physiological Measurement*, vol. 37, pp. 530–543, 4 2016.
- [292] P. K. Sahu and R. N. Rai, “Fault diagnosis of rolling bearing based on an improved denoising technique using complete ensemble empirical mode decomposition and adaptive thresholding method,” *Journal of Vibration Engineering Technologies*, vol. 11, pp. 513–535, 2 2023.
- [293] Y. Yang, H. Liu, L. Han, and P. Gao, “A feature extraction method using vmd and improved envelope spectrum entropy for rolling bearing fault diagnosis,” *IEEE Sensors Journal*, vol. 23, pp. 3848–3858, 2 2023.
- [294] P. Poludasu, S. J. Bandaru, A. K. Paleru, S. Venkatadas, and P. A. K., “Biometric authentication using ecg signals,” in *2023 4th International Conference for Emerging Technology (INCET)*, 5 2023, pp. 1–6.
- [295] M. Feldman, “Hilbert transform in vibration analysis,” *Mechanical Systems and Signal Processing*, vol. 25, pp. 735–802, 4 2011.
- [296] G. Soto and S. Torres, “Feature extraction based on circular summary statistics in ecg signal classification,” in *2009 International Conference of the Chilean Computer Science Society*, 11 2009, pp. 142–144.
- [297] M. An-bo, Z. Wei, H. Hao-hao, X. Xin-yu, and C. Fu-lei, “Incipient fault feature extraction of main bearing by iterative squared envelope analysis,” in *2017 Prognostics and System Health Management Conference (PHM-Harbin)*, 7 2017, pp. 1–6.
- [298] C. Liu, J. Wang, L. Liang, B. Niu, and D. Liu, “Periodic impact feature extraction for bearings fault based on envelope correlation kurtosis of singular value components,” in *2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 12 2022, pp. 359–363.

- [299] L. Han, J. Yan, S. Fan, M. Xu, Z. Liu, Y. Geng, and C. Guan, “Feature extraction of uhf pd signals based on diode envelope detection and linear discriminant analysis,” in *2019 5th International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST)*, 10 2019, pp. 598–602.
- [300] D. Chowdhury, M. Hasan, and M. Z. R. Khan, “Statistical features extraction from current envelopes for non- intrusive appliance load monitoring,” in *2020 Southeast-Con*, 3 2020, pp. 1–5.
- [301] P. Zuo-wei, A. Hong-wen, L. Shuang-yin, and L. Yi-bing, “Amplitude envelope analysis for feature extraction of power plant blower bearing failure,” in *The 26th Chinese Control and Decision Conference (2014 CCDC)*, 5 2014, pp. 2821–2824.
- [302] S. Marquez-Figueroa, Y. S. Shmaliy, and O. Ibarra-Manzano, “Removing artifacts and optimal features extraction from emg envelope,” in *2020 24th International Conference on Circuits, Systems, Communications and Computers (CSCC)*, 7 2020, pp. 90–95.
- [303] L. Zhou, F. Yang, Y. Zhang, and S. Hou, “Feature extraction and classification of partial discharge signal in gis based on hilbert transform,” in *2021 International Conference on Information Control, Electrical Engineering and Rail Transit (ICEERT)*, 10 2021, pp. 208–213.
- [304] M. R. Mehrjou, N. Mariun, M. Karami, N. Misron, and M. A. M. Radzi, “Broken rotor bar detection in ls-pmsms based on statistical features analysis of start-up current envelope,” in *2015 IEEE 3rd International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, 11 2015, pp. 1–6.
- [305] T. Deghboudj, A. Tlemçani, and A. Derghal, “Identifying broken rotor bars in squirrel cage induction motors: Examination of statistical characteristics derived from

- start-up current envelope,” in *2023 2nd International Conference on Electronics, Energy and Measurement (IC2EM)*, 11 2023, pp. 1–7.
- [306] D.-H. Lee, C. Hong, W.-B. Jeong, and S. Ahn, “Time–frequency envelope analysis for fault detection of rotating machinery signals with impulsive noise,” *Applied Sciences*, vol. 11, p. 5373, 6 2021.
- [307] S. Kim, D. An, and J.-H. Choi, “Diagnostics 101: A tutorial for fault diagnostics of rolling element bearing using envelope analysis in matlab,” *Applied Sciences*, vol. 10, p. 7302, 10 2020.
- [308] S. Aburakhia, T. Tayeh, R. Myers, and A. Shami, “Similarity-based predictive maintenance framework for rotating machinery,” in *2022 5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, 2022, pp. 1–6.
- [309] I. Attoui, “Novel fast and automatic condition monitoring strategy based on small amount of labeled data,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, pp. 1266–1275, 2 2022.
- [310] I. Attoui, N. Boutasseta, and N. Fergani, “Novel machinery monitoring strategy based on time–frequency domain similarity measurement with limited labeled data,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–8, 2021.
- [311] T. Tayeh, S. Aburakhia, R. Myers, and A. Shami, “Distance-based anomaly detection for industrial surfaces using triplet networks,” in *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2020, pp. 0372–0377.
- [312] P. Arun, S. A. Lincon, and N. Prabhakaran, “An automated method for the analysis of bearing vibration based on spectrogram pattern matching,” *Journal of Applied Research and Technology*, vol. 17, 10 2019.

- [313] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, 4 2004.
- [314] L. Jing-yi, L. Hong, Y. Dong, and Z. Yan-sheng, “A new wavelet threshold function and denoising application,” *Mathematical Problems in Engineering*, vol. 2016, pp. 1–8, 2016.
- [315] D. Donoho, “De-noising by soft-thresholding,” *IEEE Transactions on Information Theory*, vol. 41, pp. 613–627, 5 1995.
- [316] C. W. R. U. B. D. C. Website, “Bearing test data.” [Online]. Available: <https://engineering.case.edu/bearingdatacenter>
- [317] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, “Array programming with numpy,” *Nature*, vol. 585, pp. 357–362, 9 2020.
- [318] W. McKinney, “Data structures for statistical computing in python,” 2010, pp. 56–61.
- [319] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Édouard Duchesnay, “Scikit-learn: Machine learning in python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 11 2011.
- [320] *SciPy*. Cambridge University Press, 10 2020, pp. 358–437.
- [321] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner,

- N. Yager, E. Gouillart, and T. Yu, “scikit-image: image processing in python,” *PeerJ*, vol. 2, p. e453, 6 2014.
- [322] L. Yang and A. Shami, “On hyperparameter optimization of machine learning algorithms: Theory and practice,” *Neurocomputing*, vol. 415, pp. 295–316, 11 2020.
- [323] S. Aburakhia, T. Tayeh, R. Myers, and A. Shami, “A transfer learning framework for anomaly detection using model of normality,” in *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2020, pp. 0055–0061.
- [324] A. S and S. A, “Sb-pdm: a tool for predictive maintenance of rolling bearings based on limited labeled data.” [Online]. Available: <https://github.com/Western-OC2-Lab/SB-PdM-a-tool-for-predictive-maintenance-of-rolling-bearings-based-on-limited-labeled-data>
- [325] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O’Leary, “Pywavelets: A python package for wavelet analysis,” *Journal of Open Source Software*, vol. 4, p. 1237, 4 2019.
- [326] L. Yang and A. Shami, “Iot data analytics in dynamic environments: From an automated machine learning perspective,” *Engineering Applications of Artificial Intelligence*, vol. 116, p. 105366, 11 2022.
- [327] S. J. Lacey, “The role of vibration monitoring in predictive maintenance.” [Online]. Available: https://www.schaeffler.com/remotemedien/media/_shared_media/08_media_library/01_publications/schaeffler_2/technicalpaper_1/download_1/the_role_of_vibration_monitoring.pdf
- [328] X. Huang, G. Wen, L. Liang, Z. Zhang, and Y. Tan, “Frequency phase space empirical wavelet transform for rolling bearings fault diagnosis,” *IEEE Access*, vol. 7, pp. 86 306–86 318, 2019.

- [329] L. Li, A. Guo, and H. Chen, “Feature extraction based on ewt with scale space threshold and improved mckd for fault diagnosis,” *IEEE Access*, vol. 9, pp. 45 407–45 417, 2021.
- [330] W. Fan, Q. Zhou, J. Li, and Z. Zhu, “A wavelet-based statistical approach for monitoring and diagnosis of compound faults with application to rolling bearings,” *IEEE Transactions on Automation Science and Engineering*, vol. 15, pp. 1563–1572, 10 2018.
- [331] M. Rostaghi and M. N. Khajavi, “Comparison of feature extraction from wavelet packet based on reconstructed signals versus wavelet packet coefficients for fault diagnosis of rotating machinery,” *Journal of Vibroengineering*, vol. 18, pp. 165–174, 2016. [Online]. Available: <https://www.extrica.com/article/16458>
- [332] Y. Li, X. Wang, Z. Liu, X. Liang, and S. Si, “The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review,” *IEEE Access*, vol. 6, pp. 66 723–66 741, 2018.
- [333] R. Yan and R. X. Gao, “Wavelet domain principal feature analysis for spindle health diagnosis,” *Structural Health Monitoring*, vol. 10, pp. 631–642, 11 2011.
- [334] R. YAN and R. X. GAO, “Base wavelet selection for bearing vibration signal analysis,” *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 07, pp. 411–426, 7 2009.
- [335] L. Song, H. Wang, and P. Chen, “Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, pp. 1887–1899, 8 2018.
- [336] H. Li, T. Liu, X. Wu, and Q. Chen, “A bearing fault diagnosis method based on enhanced singular value decomposition,” *IEEE Transactions on Industrial Informatics*, vol. 17, pp. 3220–3230, 5 2021.

- [337] X. Wang, G. Tang, T. Wang, X. Zhang, B. Peng, L. Dou, and Y. He, “Lkurtogram guided adaptive empirical wavelet transform and purified instantaneous energy operation for fault diagnosis of wind turbine bearing,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–19, 2021.
- [338] P. Liang, C. Deng, J. Wu, G. Li, Z. Yang, and Y. Wang, “Intelligent fault diagnosis via semisupervised generative adversarial nets and wavelet transform,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 4659–4671, 7 2020.
- [339] Z. Shi, J. Chen, Y. Zi, and Z. Zhou, “A novel multitask adversarial network via redundant lifting for multicomponent intelligent fault detection under sharp speed variation,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.
- [340] H. Shao, M. Xia, J. Wan, and C. W. de Silva, “Modified stacked autoencoder using adaptive morlet wavelet for intelligent fault diagnosis of rotating machinery,” *IEEE/ASME Transactions on Mechatronics*, vol. 27, pp. 24–33, 2 2022.
- [341] H. Ahmed and A. K. Nandi, “Compressive sampling and feature ranking framework for bearing fault classification with vibration signals,” *IEEE Access*, vol. 6, pp. 44 731–44 746, 2018.
- [342] W. Caesarendra and T. Tjahjowidodo, “A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing,” *Machines*, vol. 5, p. 21, 9 2017.
- [343] G. Yen and K.-C. Lin, “Wavelet packet feature extraction for vibration monitoring,” *IEEE Transactions on Industrial Electronics*, vol. 47, pp. 650–667, 6 2000.
- [344] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, “Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification,” 2016.

- [345] H. Huang and N. Baddour, “Bearing vibration data under time-varying rotational speed conditions,” 2019.
- [346] A. Shenfield and M. Howarth, “A novel deep learning model for the detection and identification of rolling element-bearing faults,” *Sensors*, vol. 20, p. 5112, 9 2020.
- [347] D. Wang, Y. Li, L. Jia, Y. Song, and Y. Liu, “Novel three-stage feature fusion method of multimodal data for bearing fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.
- [348] A. K. Sharma and N. K. Verma, “Quick learning mechanism with cross-domain adaptation for intelligent fault diagnosis,” *IEEE Transactions on Artificial Intelligence*, vol. 3, pp. 381–390, 6 2022.
- [349] L. Hou, R. Jiang, Y. Tan, and J. Zhang, “Input feature mappings-based deep residual networks for fault diagnosis of rolling element bearing with complicated dataset,” *IEEE Access*, vol. 8, pp. 180 967–180 976, 2020.
- [350] H. Geng, Y. Peng, L. Ye, and Y. Guo, “Fault identification of rolling bearing with variable speed based on generalized broadband mode decomposition and det,” *SSRN Electronic Journal*, 2022.
- [351] K. Zhang, C. Fan, X. Zhang, H. Shi, and S. Li, “A hybrid deep-learning model for fault diagnosis of rolling bearings in strong noise environments,” *Measurement Science and Technology*, vol. 33, p. 065103, 6 2022.
- [352] D. Kumar, S. M. Ujjan, K. Dev, S. A. Khowaja, N. A. Bhatti, and T. Hussain, “Towards soft real-time fault diagnosis for edge devices in industrial iot using deep domain adaptation training strategy,” *Journal of Parallel and Distributed Computing*, vol. 160, pp. 90–99, 2 2022.

- [353] L. Yang and L. Chen, “Fault diagnosis algorithm of printing machine rolling bearing based on hilbert envelope spectrum,” in *2022 International Conference on Computer Network, Electronic and Automation (ICCNEA)*, 2022, pp. 357–361.
- [354] Q. Wu and W. Li, “Fault diagnosis of rolling element bearings based on fdk and hilbert spectrum,” in *2016 IEEE Far East NDT New Technology Application Forum (FENDT)*, 2016, pp. 174–180.
- [355] C. Luo, M. Jia, and Y. Wen, “The diagnosis approach for rolling bearing fault based on kurtosis criterion emd and hilbert envelope spectrum,” in *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2017, pp. 692–696.
- [356] S. Wei, D. Wang, H. Wang, and Z. Peng, “Time-varying envelope filtering for exhibiting space bearing cage fault features,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021.
- [357] Y. Xu, H. Wang, Z. Zhou, and W. Chen, “Diagnosis of weak fault of rolling bearing based on eemd and envelope spectrum analysis,” in *2021 3rd International Symposium on Robotics Intelligent Manufacturing Technology (ISRIMT)*, 2021, pp. 419–422.
- [358] D.-M. Yang, “The envelop-based bispectral analysis for motor bearing defect detection,” in *2017 International Conference on Applied System Innovation (ICASI)*, 2017, pp. 1646–1649.
- [359] S. Osman and W. Wang, “An hilbert-huang spectrum technique for fault detection in rolling element bearings,” in *2018 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, 2018, pp. 549–554.
- [360] H. Xiang, X. Wang, and G. Huang, “Approach to fault feature extractions of rolling

- bearing via eemd and full-vector envelope spectrum,” in *2017 29th Chinese Control And Decision Conference (CCDC)*, 2017, pp. 6492–6497.
- [361] Z. Li, M. Zhang, L. Li, R. Chen, Y. Zhang, and Y. Cui, “Rolling element bearing fault diagnosis based on eemd and index of envelope spectrum sparse ratio,” in *2021 International Conference on Intelligent Transportation, Big Data Smart City (IC-ITBS)*, 2021, pp. 338–341.
- [362] X. Xiaoli, J. Zhanglei, L. Hao, and L. Yuheng, “Fault diagnosis method of rolling bearing based on 1.5-dimensional envelope spectrum,” in *2019 14th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*, 2019, pp. 1163–1168.
- [363] Y. Wu, Z. Hu, W. An, X. Li, X. Wang, P. Du, and X. Yu, “Automatic diagnosis of rolling element bearing under different conditions based on rvmd and envelope order capture,” *IEEE Access*, vol. 7, pp. 91 799–91 808, 2019.
- [364] Y. Ren, W. Li, Z. Zhu, and F. Jiang, “Isvd-based in-band noise reduction approach combined with envelope order analysis for rolling bearing vibration monitoring under varying speed conditions,” *IEEE Access*, vol. 7, pp. 32 072–32 084, 2019.
- [365] E. T. Chelmiah and D. F. Kavanagh, “Hilbert marginal spectrum for failure mode diagnosis of rotating machines,” in *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, 2021, pp. 1–6.
- [366] H. Shi, Y. Qin, Y. Wang, H. Bai, and X. Zhou, “An improved viterbi algorithm for adaptive instantaneous angular speed estimation and its application into the machine fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.
- [367] V. Kannan, H. Li, and D. V. Dao, “Demodulation band optimization in envelope

- analysis for fault diagnosis of rolling element bearings using a real-coded genetic algorithm,” *IEEE Access*, vol. 7, pp. 168 828–168 838, 2019.
- [368] L. Fu, P. Zhang, M. Ding, S. Wang, F. Xu, and Z. Ma, “A novel composite envelope negentropy deconvolution reconstruction method for fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024.
- [369] Z. Chen, Y. Yang, C. He, Y. Liu, X. Liu, and Z. Cao, “Feature extraction based on hierarchical improved envelope spectrum entropy for rolling bearing fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.
- [370] D. Liu, W. Cheng, and W. Wen, “An online bearing fault diagnosis technique via improved demodulation spectrum analysis under variable speed conditions,” *IEEE Systems Journal*, vol. 14, no. 2, pp. 2323–2334, 2020.
- [371] A. Soualhi, K. Medjaher, and N. Zerhouni, “Bearing health monitoring based on hilbert–huang transform, support vector machine, and regression,” *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2015.
- [372] R. Yan and R. X. Gao, “Hilbert–huang transform-based vibration signal analysis for machine health monitoring,” *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2320–2329, 2006.
- [373] L. Marple, “Computing the discrete-time ”analytic” signal via fft,” *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2600–2603, 1999.
- [374] Y.-W. Liu, “Hilbert transform and applications,” *Fourier Transform Applications*, pp. 291–300, 2012.
- [375] Y. Kim and Y.-K. Kim, “Physics-informed time-frequency fusion network with attention for noise-robust bearing fault diagnosis,” *IEEE Access*, vol. 12, pp. 12 517–12 532, 2024.

- [376] L. Cui, X. Tian, Q. Wei, and Y. Liu, “A self-attention based contrastive learning method for bearing fault diagnosis,” *Expert Systems with Applications*, vol. 238, p. 121645, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417423021474>
- [377] Z. Zhang and L. Wu, “Graph neural network-based bearing fault diagnosis using granger causality test,” *Expert Systems with Applications*, vol. 242, p. 122827, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417423033298>
- [378] Y. Xue, C. Wen, Z. Wang, W. Liu, and G. Chen, “A novel framework for motor bearing fault diagnosis based on multi-transformation domain and multi-source data,” *Knowledge-Based Systems*, vol. 283, p. 111205, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0950705123009553>
- [379] S. Gao, X. Zhang, C. Du, and Q. Ji, “A multichannel low-power wide-area network with high-accuracy synchronization ability for machine vibration monitoring,” *IEEE Internet of Things Journal*, vol. 6, pp. 5040–5047, 6 2019.
- [380] C. Zhao, B. Tang, Y. Huang, and H. Fu, “Multilayer joint optimization of packet size and adaptive transmission scheduling of wireless sensor networks for mechanical vibration monitoring,” *IEEE Internet of Things Journal*, vol. 10, pp. 6444–6455, 4 2023.
- [381] M. Z. A. Bhuiyan, J. Wu, G. Wang, Z. Chen, J. Chen, and T. Wang, “Quality-guaranteed event-sensitive data collection and monitoring in vibration sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 572–583, 2017.
- [382] A. Chehri, R. Saadane, N. Hakem, and H. Chaibi, “Enhancing energy efficiency

- of wireless sensor network for mining industry applications,” *Procedia Computer Science*, vol. 176, pp. 261–270, 2020.
- [383] N. W. Bergmann and L.-Q. Hou, “Energy efficient machine condition monitoring using wireless sensor networks,” in *2014 International Conference on Wireless Communication and Sensor Network*, 12 2014, pp. 285–290.
- [384] F. Zonzini, M. Zauli, M. Mangia, N. Testoni, and L. De Marchi, “Model-assisted compressed sensing for vibration-based structural health monitoring,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7338–7347, 2021.
- [385] E. Ragusa, F. Zonzini, L. D. Marchi, and P. Gastaldo, “Vibration monitoring in the compressed domain with energy-efficient sensor networks,” *IEEE Sensors Letters*, vol. 7, pp. 1–4, 8 2023.
- [386] M. Zhang, H. Zhang, C. Zhang, and D. Yuan, “Communication-efficient quantized deep compressed sensing for edge-cloud collaborative industrial iot networks,” *IEEE Transactions on Industrial Informatics*, vol. 19, no. 5, pp. 6613–6623, 2023.
- [387] P. Wei and F. He, “The compressed sensing of wireless sensor networks based on internet of things,” *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25 267–25 273, 2021.
- [388] G.-S. Dong, H.-P. Wan, Y. Luo, and M. D. Todd, “A fast sparsity-free compressive sensing approach for vibration data reconstruction using deep convolutional gan,” *Mechanical Systems and Signal Processing*, vol. 188, p. 109937, 4 2023.
- [389] M. Rani, S. Dhok, and R. Deshmukh, “A machine condition monitoring framework using compressed signal processing,” *Sensors*, vol. 20, p. 319, 1 2020.
- [390] T. Giannopoulos and V. Paliouras, “A novel technique for low-power d/a conversion based on papr reduction,” in *2006 IEEE International Symposium on Circuits and Systems*, 2006, pp. 4 pp.–5002.

- [391] D. Mestdagh, P. Spruyt, and B. Biran, “Analysis of clipping effect in dmt-based adsl systems,” in *Proceedings of ICC/SUPERCOMM’94 - 1994 International Conference on Communications*, pp. 293–300.
- [392] N. Andgart, P. Ödler, A. Johansson, and P. O. Börjesson, “Designing tone reservation par reduction,” *EURASIP Journal on Advances in Signal Processing*, vol. 2006, p. 038237, 12 2006.
- [393] W. Shieh and I. Djordjevic, *OFDM Principles*. Elsevier, 2010, pp. 31–52.
- [394] Z. Hu, J. Zhu, and H. Wang, “Experimental research on uav stability based on vibration signal measurement of uav motors,” in *2021 3rd International Symposium on Robotics and Intelligent Manufacturing Technology (ISRIMT)*, 9 2021, pp. 131–134.
- [395] C.-C. Li, V. K. Ramanna, D. Webber, C. Hunter, T. Hack, and B. Dezfouli, “Sensifi: A wireless sensing system for ultrahigh-rate applications,” *IEEE Internet of Things Journal*, vol. 9, pp. 2025–2043, 2 2022.
- [396] D. Ganga and V. Ramachandran, “Iot-based vibration analytics of electrical machines,” *IEEE Internet of Things Journal*, vol. 5, pp. 4538–4549, 12 2018.
- [397] B.-Y. Ooi, W.-L. Beh, W.-K. Lee, and S. Shirmohammadi, “A parameter-free vibration analysis solution for legacy manufacturing machines’ operation tracking,” *IEEE Internet of Things Journal*, vol. 7, pp. 11 092–11 102, 11 2020.
- [398] H. A. et al., *OFDM-Based MIMO Systems*. Elsevier, 2020, pp. 133–174.
- [399] K. Bulusu, H. Shaiek, D. Roviras, R. Zayani, M. Renfors, L. Anttila, and M. Abdelaziz, *Power Amplifier Effects and Peak-to-Average Power Mitigation*. Elsevier, 2017, pp. 461–489.
- [400] M. Parker, *Orthogonal Frequency Division Multiple Access Wireless Communications*. Elsevier, 2017, pp. 209–230.

- [401] enDAQ, “Aircraft vibration.” [Online]. Available: <https://endaq.com/pages/free-vibration-analysis-files>
- [402] A. Keipour, M. Mousaei, and S. Scherer, “Alfa: A dataset for uav fault and anomaly detection,” *The International Journal of Robotics Research*, vol. 40, pp. 515–520, 2021.
- [403] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, “Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,” *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.
- [404] S. Sheng, “Wind turbine gearbox condition monitoring vibration analysis benchmarking datasets.” [Online]. Available: <https://data.openml.org/submissions/738>
- [405] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, “Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification,” 2016. [Online]. Available: <https://api.semanticscholar.org/CorpusID:4692476>
- [406] 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*, vol. 84, pp. 200–226, 2017.
- [407] T. Li, J. Deng, and J. Ren, “Security and energy efficiency: Breaking the barriers of high peak-to-average power ratio and disguised jamming in nextg iot system design,” *IEEE Internet of Things Journal*, vol. 10, pp. 2658–2666, 2023.
- [408] M. S. Ahmed, S. Boussakta, A. Al-Dweik, B. Sharif, and C. C. Tsimenidis, “Efficient design of selective mapping and partial transmit sequence using t-ofdm,” *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 2636–2648, 2020.

- [409] S. Gokceli, I. Peruga, E. Tirola, K. Pajukoski, T. Riihonen, and M. Valkama, “Novel tone reservation method for dft-s-ofdm,” *IEEE Wireless Communications Letters*, vol. 10, pp. 2130–2134, 10 2021.
- [410] K. Anoh, C. Tanriover, B. Adebisi, and M. Hammoudeh, “A new approach to iterative clipping and filtering papr reduction scheme for ofdm systems,” *IEEE Access*, vol. 6, pp. 17 533–17 544, 2018.
- [411] R. K. Singh and M. Fidele, “An efficient papr reduction scheme for ofdm system using peak windowing and clipping,” in *2015 Third International Conference on Image Information Processing (ICIIP)*, 12 2015, pp. 491–495.
- [412] S. Bhooshan, *Fundamentals of Analogue and Digital Communication Systems*. Springer Nature, 2021.
- [413] A. Mohammed, T. Ismail, A. Nassar, and H. Mostafa, “A novel companding technique to reduce high peak to average power ratio in ofdm systems,” *IEEE Access*, vol. 9, pp. 35 217–35 228, 2021.
- [414] T. Zhang, Z. Tong, W. Zhang, H. Wang, and P. Li, “A novel papr reduction scheme based on joint traditional algorithm and machine learning for co-ofdm systems,” *IEEE Photonics Technology Letters*, vol. 35, no. 8, pp. 418–421, 2023.
- [415] A. Mohamed, A. S. Tag Eldien, M. M. Fouda, and R. S. Saad, “Lstm-autoencoder deep learning technique for papr reduction in visible light communication,” *IEEE Access*, vol. 10, pp. 113 028–113 034, 2022.
- [416] F. Zou, Z. Liu, X. Hu, and G. Wang, “A novel papr reduction scheme for ofdm systems based on neural networks,” *Wireless Communications and Mobile Computing*, 2021.

- [417] X. Tang, Z. Tan, L. Xiao, M. Zhao, and Y. Li, “A new approach to dealing with the papr problem for ofdm systems via deep learning.” *IEEE*, 11 2022, pp. 252–256.
- [418] M. S. Omar and X. Ma, “Mitigating clipping distortion in ofdm using deep residual learning,” in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6 2021, pp. 4925–4929.
- [419] C. Rapp, “Effects of hpa-nonlinearity on 4-dpsk/ofdm-signal for a digital sound broadcasting system,” in *2nd European Conference on Satellite Communication*, 9 1991, pp. 179–184.
- [420] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House, 2000.
- [421] F. Chollet *et al.*, “Keras,” 2015. [Online]. Available: <https://github.com/fchollet/keras>
- [422] M. A. et al., “Tensorflow: A system for large-scale machine learning,” 5 2016.
- [423] B. M. Lee, Y. S. Rim, and W. Noh, “A combination of selected mapping and clipping to increase energy efficiency of ofdm systems,” *PLOS ONE*, vol. 12, no. 10, pp. 1–25, 10 2017. [Online]. Available: <https://doi.org/10.1371/journal.pone.0185965>

Curriculum Vitae

Name: *Sulaiman Ahmed Aburakhia*

Education

- 2020-2024: Doctor of Philosophy, Electrical and Computer Engineering, Western University, London, ON Canada.
- 2023: Certificate in Curriculum, Teaching, and Learning in the STEM Disciplines, Western University, London, ON Canada.
- 2004-2007: Master of Science, Electronics and Communications Engineering, the Arab Academy for Science, Technology and Maritime Transport (AASTMT), Alexandria, Egypt.
- 1998-2003: Bachelor of Science, Electrical Engineering (Honours), October 6 University, Giza, Egypt.

Honours and Awards

- 2024 Graduate Student Award for Excellence in Research, awarded by the Faculty of Engineering, Western University.
- 2021/2022/2023 Ontario Graduate Scholarship (OGS) and Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST).

Graduate Teaching Assistantship

2020-2024: The Department of Electrical and Computer Engineering, Western University.

Courses:

- Computer Networks Applications.

- Communication Theory.
- Software Requirements and Analysis.
- Cloud Computing.

Professional Experience

- 29 November 2015 – 31 May 2019: Regional Manager- Radio Access Network (RAN) Optimization, LCC KSA- Huawei-STC Managed Services Project, Dammam, Saudi Arabia.
- 07 March 2012 – 28 November 2015: Optimization Senior Engineer, LCC KSA- Huawei-STC Managed Services Project, Dammam, Saudi Arabia.
- 05 March 2008 – 10 November 2011: RAN Consultant, Mobile Systems International Consultancy (MSI), Riyadh, Saudi Arabia.

Publications:

- **S. A. Aburakhia**, A. Shami and G. K. Karagiannidis, “On the Intersection of Signal Processing and Machine Learning: A Use Case-Driven Analysis Approach,” *Under Review at Proceedings of the IEEE*
- **S. A. Aburakhia** and A. Shami, “Signal-Companding AutoEncoder with Compression-Based Activation for an Efficient Vibration-Based Remote Condition Monitoring,” *Under Review at the IEEE Transactions on Reliability*.
- **S. A. Aburakhia**, I. Hamieh and A. Shami, “Joint Instantaneous Amplitude-Frequency Analysis of Vibration Signals for Vibration-Based Condition Monitoring of Rolling Bearings,” *Under Review at IEEE Transactions on Instrumentation and Measurement*
- **S. A. Aburakhia** and A. Shami, “SB-PDM: A Tool for Predictive Maintenance of Rolling Bearings Based on Limited Labeled Data,” *Software Impacts*, 16 (2023): 100503.
- **S. A. Aburakhia**, T. Tayeh, R. Myers and A. Shami, “Similarity-Based Predictive Maintenance Framework for Rotating Machinery,” *2022 5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, Cairo, Egypt, 2022, pp. 1-6. **Best Paper Award**.
- T. Tayeh, **S. A. Aburakhia**, R. Myers, and A. Shami, “An Attention-Based ConvLSTM Autoencoder with Dynamic Thresholding for Unsupervised Anomaly Detection in Multivariate Time Series,” *Machine Learning and Knowledge Extraction*, 4, no. 2 350-370, 2022.
- **S. A. Aburakhia**, R. Myers and A. Shami, “A Hybrid Method for Condition Monitoring and Fault Diagnosis of Rolling Bearings With Low System Delay,” in *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-13, 2022
- T. Tayeh, **S. A. Aburakhia**, R. Myers and A. Shami, “Distance-Based Anomaly Detection for Industrial Surfaces Using Triplet Networks,” *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2020, pp. 0372-0377. **Best Paper Award**.
- **S. A. Aburakhia**, T. Tayeh, R. Myers and A. Shami, “A Transfer Learning Framework for Anomaly Detection Using Model of Normality,” *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2020, pp. 0055-0061. **Best Paper Award**.
- **S. A. Aburakhia**, E. F. Badran and D. A. Mohamed, “Distribution of the PAPR for real-valued OFDM signals”, *Proc. Int. Conf Inf Technol.*, pp. 1-8, 2009.
- **S. A. Aburakhia**, E. F. Badran and D. A. E. Mohamed, “Linear Companding Transform for the Reduction of Peak- to-Average Power Ratio of OFDM Signals,” in *IEEE*

Transactions on Broadcasting, vol. 55, no. 1, pp. 155-160, March 2009.

- **S. A. Aburakhia**, E. F. Badran and D. A. E. Mohamed, “A Comparison between Clipping and μ -Law Companding Schemes for the Reduction of Peak-to-Average Power Ratio of OFDM,” *2007 National Radio Science Conference*, Cairo, Egypt, 2007, pp. 1-10.