UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Engineering, Built Environment and Information Technology

Latent approaches for unsupervised learning-based methodologies in gearbox fault diagnosis applications

by

Ryan Cameron Balshaw

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor (Mechanical Engineering)

in the

Department of Mechanical and Aeronautical Engineering

Faculty of Engineering, Built Environment and Information Technology

University of Pretoria

February 2024

Abstract

Latent approaches for unsupervised learning-based methodologies in gearbox fault diagnosis applications

by

Ryan Cameron Balshaw

Supervisors:	Prof. P. Stephan Heyns, Prof. Daniel N. Wilke, and Dr. Stephan Schmidt
Department:	Mechanical and Aeronautical Engineering
University:	University of Pretoria
Degree:	Philosophiae Doctor (Mechanical Engineering)
Keywords:	Engineering asset management; Gearbox condition monitoring; Fault diagnostics; Temporal preservation; Unsupervised learning; Latent variable models

The monitoring of critical industrial assets is of fundamental importance to the longevity and sustainability of many industrial sectors. Vibration-based fault diagnosis methodologies seek to enhance the monitoring process by providing rich diagnostic information related to the instantaneous asset health state to support maintenance decisions. While conventional signal analysis techniques applied to vibration data have been demonstrated to provide information related to the asset health state, their utility is often impeded by other vibration sources and time-varying operating conditions. Hence, powerful data-driven techniques are used to extract asset health state information to support the decision-making process.

Learning-based methodologies offer the ability to perform effective fault diagnostics by extracting informative features for the fault diagnosis problem. Unsupervised learning latent-based approaches, e.g., latent variable models (LVMs), offer an effective strategy to bypass dataset-dependent feature engineering practices and the requirement for labelled fault conditions inherent to many learning-based methodologies. Temporal preservation is an important approach which advocates for the preservation of time structure to produce an informative latent manifold, and new possibilities and opportunities for condition inference are potentially available from a latent variable perspective. However, this is only the first step to unlocking the potential of LVMs for asset condition monitoring. Improving the application of temporal-preserving LVMs is crucial to advance unsupervised learning-based methodologies for

effective utilisation and adoption in industrial applications. The approaches considered in this thesis address current issues in gearbox condition monitoring, and areas for improvement are identified for temporal-preserving LVMs. Three research directions are established to improve the application of temporal-preserving LVMs for gearbox fault diagnostics. The first direction focuses on leveraging the latent manifold to drive effective condition inference, the second direction explores different LVM training approaches, and the third direction investigates latent health indicators for LVMs developed using different training approaches.

First, a latent indicator framework for temporal-preserving LVMs is proposed, which identifies unique classes to capture several manifold perspectives for latent health indicator derivation. The benefits of an ensemble-based latent condition inference approach are identified, and the latent manifold is demonstrated to be a fruitful source for fault diagnostic information. This enhances LVM-based methodologies for the fault diagnosis task. As the models are fault condition agnostic, it is demonstrated that a diverse set of latent health indicators is necessary to obtain a robust and insightful condition inference approach.

Second, the parameter estimation process for temporal-preserving LVMs is studied to explore methods used to construct the latent manifold. For data-driven models in condition monitoring, two strategies exist to estimate the model parameters: offline and online training. Albeit equally applicable for LVMs, conventional research applications disregard online strategies for LVMs. Online training suits scenarios which lack historical data from a healthy asset. Thus, understanding the implications of using online training for temporal-preserving LVMs is considered. Moreover, two main methods are identified for online training: reconstruction-focused and interpretation-focused methods. Both online and offline models are demonstrated to be effective using basic methods from the proposed latent indicator framework, which reveals their utility in condition monitoring and confirms their ability to perform effective fault diagnostics.

Finally, the proposed latent indicator framework pre-supposes an offline training approach and was only assessed using offline models. Therefore, the latent indicator framework bounds are studied using online models. Implicit assumptions in the proposed latent indicator framework are identified to develop a new category of latent health indicators, and a set of online latent indicators is proposed for online methods. The applicability of the two indicator categories is studied. Diversity is highlighted to be a prominent factor in the indicator categories, and the study advocates for the use of a plethora of diverse indicators for both

offline and online methods.

The research carried out in this thesis indicates that the considered temporal-preserving latent variable approaches are advantageous to the fault diagnosis task. The identified research directions ensure that the latent manifold of offline and online temporal-preserving LVMs can be effectively applied to drive meaningful condition monitoring. By combining offline and online models with a diverse, categorised latent indicator framework comprising of heterogeneous collection latent health indicators, a holistic fault diagnosis strategy for temporal-preserving LVMs can be developed. Temporal-preserving LVMs are established as an effective tool for unsupervised learning-based condition monitoring using gearbox vibration data.

Acknowledgements

To my supervisors, Prof. P. Stephan Heyns, Prof. Daniel Nico Wilke, and Dr. Stephan Schmidt, I would like to express my heartfelt gratitude. They have been exceptional mentors and advisors, consistently providing patient guidance and encouragement. Their unwavering care and attention have profoundly impacted the evolution of this research and my growth as a student and a researcher. This research would not have been possible without their input, support, and counsel. I am forever indebted to each of them and will cherish our interactions and discussions over the years.

To my wife, Kerry, thank you for your endless motivation and constant encouragement, for being a guiding light during the darkest moments, and for undertaking this journey with me. Thank you for your unwavering support and for being by my side through all the ups and downs experienced in my research; you have been a cornerstone throughout this period.

To my family, especially my parents, parents in-law, grandparents, and siblings, thank you for your involvement in my life, the experiences we have shared, the principles and values you have instilled in me, and the sacrifices made to get me to where I am today. I am incredibly fortunate and very grateful to have such a wonderful family.

To my friends, thank you all for your wonderful presence and meaningful contributions to my life. I cherish the moments and many good times we have shared (and will share) over the years.

Finally, I would like to thank the University of Pretoria, the Centre for Asset Integrity Management (C-AIM), and Anglo Gold Ashanti for the financial support that was received during this period.

List of publications based off the candidate's research

Journal articles:

1. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt. Importance of temporal preserving latent analysis for latent variable models in fault diagnostics of rotating machinery, *Mechanical Systems and Signal Processing* 168:108663, 2022.

2. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt. Latent indicators for temporal-preserving latent variable models in vibration-based condition monitoring under non-stationary conditions, *Mechanical Systems and Signal Processing* 199:110446, 2023.

3. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt. A spectral regularisation framework for latent variable models designed for single channel applications, submitted to *SoftwareX*.

Two articles are currently in the internal review process and are going to be submitted to relevant journals once this process is complete. The article are:

1. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt. Temporal-preserving latent variable models: Offline and online training for reconstruction and interpretation of fault data in gearbox diagnostics.

2. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt. Offline and online latent indicators for gearbox fault diagnostics using temporal-preserving latent variable models, Preprint for submission to *Mechanical Systems and Signal Processing*.

Book chapter:

1. R. Balshaw, P.S. Heyns, D.N. Wilke, S. Schmidt, Learning-based methods for vibration-based condition monitoring. *Modelling and Simulation of Complex Systems for Sustainable Energy Efficiency*, volume 20, pages 75-86, Springer International Publishing, 2022.

Table of contents

Abstract . . . . . **i**

Acknowledgements . . . . . **iv**

List of publications . . . . . **v**

List of abbreviations . . . . . **ix**

List of symbols . . . . . **xi**

Chapter 1 Introduction **1**

1.1 Background 1

1.2 Literature survey 3

1.2.1 The nature of gearbox vibration data 3

1.2.2 Learning-based methodologies for gearbox fault diagnostics 8

1.2.3 Signal processing-based methodologies for gearbox fault diagnostics . 18

1.3 Scope of research 25

1.3.1 Directions of research 25

1.4 Layout of thesis 29

Chapter 2 Latent indicators for temporal-preserving latent variable models 31

2.1 Introduction 31

2.2 The implications of temporal preservation 32

2.3 Latent manifold analysis 34

2.3.1 Latent variable models 35

2.3.2 Explicit density estimation 37

2.3.3 Data pre-processing 38

2.3.4 Latent manifold indicators 39

2.3.5 Proposed indicators 41

2.4 Experimental evaluation 47

2.4.1 IMS dataset 47

2.4.2 Gearbox dataset 57

2.5 Conclusion and recommendations 69

Chapter 3 Training approaches for reconstruction-focused and interpretation-focused latent variable models 71
3.1 Introduction 71
3.2 Offline and online latent variable methods 73
3.2.1 Overview 73
3.2.2 Introduction 76
3.2.3 Data pre-processing 79
3.2.4 Latent health indicators 80
3.3 Numerical investigation 82
3.3.1 Sensitivity analysis 83
3.3.2 Latent investigation for offline and online methods 89
3.4 Experimental investigation 91
3.4.1 Experimental bearing dataset 93
3.4.2 Experimental gearbox dataset 99
3.4.3 Comparison between the considered LVMs in an offline and online setting 107
3.5 Conclusion and recommendations 107

Chapter 4 Offline and online latent indicators for latent manifold-based condition inference 111
4.1 Introduction 111
4.2 Latent variable methods in vibration-based condition monitoring 112
4.2.1 Overview 113
4.2.2 Data pre-processing 115
4.2.3 Considered latent variable model methodologies 115
4.2.4 Reconstruction-focused models 117
4.2.5 Interpretation-focused models 117
4.2.6 Latent health indicator categories 119
4.3 Experimental investigation 124
4.3.1 Bearing failure dataset 126
4.3.2 Gear tooth failure dataset 133
4.3.3 Indicator category diversity 142
4.4 Conclusion 145

Chapter 5 Conclusion and recommendations 147

5.1 Thesis conclusion . . . 147
5.2 Recommendations for future research . . . 150

References . . . 154

Appendix A Supplementary material . . . A1
A.1 Model parameters for Chapter 2 . . . A1
A.2 Model parameters for Chapter 3 . . . A1
A.3 Model parameters for Chapter 4 . . . A3
A.4 Signal-to-noise ratio estimation procedure . . . A4
A.5 Relationship between LVM objectives and signal processing . . . A4
A.6 Additional details for the online LVMs . . . A7
A.6.1 Reconstruction-focused model . . . A7
A.6.2 Interpretation-focused model . . . A8

Appendix B Phenomenological gearbox model information . . . B1

Appendix C A spectral regularisation framework for latent variable models designed for single channel applications . . . C1
C.1 Introduction . . . C1
C.1.1 Related work . . . C3
C.2 Model formulation . . . C4
C.2.1 Time-series pre-processing . . . C4
C.2.2 Linear latent variable models . . . C4
C.2.3 Spectral regularisation . . . C5
C.3 Software description . . . C6
C.3.1 Introduction to spectrally-regularised-LVMs . . . C6
C.3.2 Software architecture . . . C7
C.3.3 Software functionalities . . . C8
C.3.4 Software dependencies . . . C8
C.4 Illustrative example . . . C9
C.5 Impact . . . C10
C.6 Conclusion . . . C12

List of abbreviations

Abbreviation	Description	Chapter
AC	Alternative Current	2
AE	Auto-Encoder	1
AIC	Akaike Information Criterion	2
API	Application Program Interface	C
AUC	Area Under the ROC Curve	2
BC	Box-Cox	1, 2
BGC	Bayesian Geometry Compensation	2–4
BPFO	Ball Pass Frequency for the Outer race	2, 3, B
C-AIM	Centre for Asset Integrity Management	2
CBM	Condition Based Maintenance	1
CCR	Cumulative Contribution Rate	3, 4, A
CD	Condition Deviance	3, 4
CI	Condition Indicator	4
CICA	Constrained Independent Component Analysis	1
CM	Condition Monitoring	1–4
DFT	Discrete Fourier Transform	3
EAM	Engineering Asset Management	1
ELBO	Evidence Lower Bound	2
EM	Expectation Maximisation	2
EMD	Empirical Mode Decomposition	1
FIR	Finite Impulse Response	1, A, C
GAN	Generative Adversarial Network	1–4
GI	Gini Index	1, 2
GMM	Gaussian Mixture Model	2
GTM	Generative Topograhphic Map	2
HI	Health Indicator	1–5
HM	Hoyer Measure	1, 2
ICA	Independent Component Analysis	1, 3, 4, A, C
ICS2	Indicator of second-order cyclostationarity	4
IMS	Intelligent Maintenance Systems	2–4, A, C
KDE	Kernel Density Estimation	2, 4

Abbreviation	Description	Chapter
KL	Kullback–Leibler	2
LHI	Latent Health Indicator	1–5
LVM	Latent Variable Model	1–5, A, C
MED	Minimum Entropy Deconvolution	1
NLL	Negative Log-Likelihood	2, 4
OC	Operating Condition	1, 2, 4, 5
PCA	Principal Component Analysis	1, 3, A, C
PCC	Pearson Correlation Coefficient	2, 4
PD	Path Dependent	2–4
PHM	Prognostics and Health Management	1
PI	Path Independent	2–4
PPCA	Probabilistic Principal Component Analysis	2–4, A
Ref.	Reference	1
RMS	Root-Mean-Square	1–4
ROC	Reciever Operating Characteristics	2
RUL	Remaining Useful Life	1
SE	Square Envelope	3
SES	Square Envelope Spectrum	1, 3, 4, A
SI	Smoothness Index	1, 2
SMSE	Synchronous Median of the Square Envelope	3, 4
SNR	Signal-to-Noise Ratio	1, 3, 4, A, B
SRICA	Spectrally Regularised Independent Component Analysis	3, 4, A
SSA	Singular Spectrum Analysis	1, 3
STFT	Short-Time Fourier Transform	1
SUMT	Sequential Unconstrained Minimisation Technique	C
SVM	Support Vector Machine	1
TSA	Time Synchronous Average	2
UP	University of Pretoria	2
VAE	Variational Auto-Encoder	1–5, A

List of symbols

Symbol	Description	Chapter
	Standard symbols	
A^l_{bo}	Amplitude of the l^{th} outer race bearing damage component	B
A^l_{dp}	Amplitude of the l^{th} distributed pinion damage component	B
A^l_{gm}	Amplitude of the l^{th} gear mesh component	B
$A^{SMSE}_{j,healthy}, A^{SMSE}_{j,healthy}$	Root-mean-square of the healthy and damaged portions of the SMSE	3
$\mathbf{a}_j$	D-dimensional vector stored in the columns of $\mathbf{A}$	3, 4, C
$\mathbf{A}$	Generative linear mapping function	3, 4, C
$\mathcal{A}_l$	Indices of the $SES[k_\alpha]$ corresponding to a narrow bandwidth around a specific cyclic frequency (α_c) component	3, 4, A
b	Kernel density estimate model bandwidth	2, 4
B	Batch size	A
$\mathbf{b}(\mathbf{w}_j)$	Fourier representation of the vector $\mathbf{w}_j$	C
$\mathbf{B}^{(GTM)}$	Generative topographic map linear transition matrix	2
$\mathcal{B}_l$	Indices of the $SES[k_\alpha]$ corresponding to a wide bandwidth around a specific cyclic frequency (α_c) component	3, 4, A
$\mathbf{C}$	Data covariance matrix given by $\mathbf{C} = \mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{\mathbf{x}\mathbf{x}^T\}$	3, 4, A
d	Dimensionality of the latent space	2–4, C
D	Dimensionality of the data space	2–4, C
$E[n]$	Envelope signal evaluated at index n	4
$f_{n,j}$	Natural frequency for the j^{th} component of $x_{ng}(t)$	B

f_s	Shaft speed [Hz]	2
$\mathbf{f}_{(\cdot)}$	Discrete Fourier transform amplitude over the full spectral range	3
F_s	Vibration data sampling frequency [Hz]	2
GR	Gear ratio	B
$G(\cdot)$	Non-quadratic negentropy approximation function	C
$\mathcal{G}_{\boldsymbol{\theta}}$	Generative transition function used to parameterise the conditional distribution $p(\mathbf{x}\|\mathbf{z})$, expressed as a function $\mathcal{G}_{\boldsymbol{\theta}}(\mathbf{z})$	1–4, C
$h_j(t)$	Single degree-of-freedom viscously damped impulse response function	B
$H(\cdot)$	Hilbert transform	3, 4
$HI^{(1)}$	Reconstruction error health indicator	2–4
$\mathcal{H}_{\boldsymbol{\phi}}$	Encoding transition function used to parameterise the conditional distribution $p(\mathbf{x}\|\mathbf{z})$, expressed as a function $\mathcal{H}_{\boldsymbol{\phi}}(\mathbf{x})$	1–4, C
$\mathbf{I}$	Identity matrix	2, 3
k_α	Cyclic frequency index [Hz or orders]	3, 4, A
L_{kernel}	Kernel size for the convolutional and de-convolutional network layers	A
L_{pad}	Padding parameter for the convolutional and de-convolutional network layers	A
L_{sft}	Shift parameter	2–4, C
L_{signal}	Length of the stored vibration data	2–4, C
L_{stride}	Stride parameter for the convolution and de-convolution network layers	A
L_w	Window length parameter	2–4, C
L_T	Number of rows in the Hankel matrix representation of a vibration signal	2–4, C
$LHI^{(j)}$	The j^{th} latent health indicator	2–4
$\mathcal{L}_{model}(\cdot)$	Latent variable model objective function	C
$\mathcal{L}_{sr}(\cdot)$	Spectral regularisation term	C

$\mathbf{L}_d, \mathbf{L}_D$	Partial and full diagonal matrix of the eigenvalues of $\mathbf{C}$	3, 4
$M(\cdot)$	Amplitude modulation function	B
$\mathbf{M}_{PPCA}$	Matrix required to compute the mean and covariance of the posterior distribution $p(\mathbf{z}\|\mathbf{x})$ for the PPCA model	3, 4
n	Index variable for the vibration data, where $n \subset \mathbb{Z}^+ = \{1, 2, \cdots, L_{signal}\}$	2–4, C
N	Number of reference dataset signals	2
N_{bo}	Number of outer race bearing damage components	B
N_{dp}	Number of distributed pinion damage components	B
N_{gm}	Number of gear mesh harmonics	B
N_l	Number of integer harmonics	3, 4, A
N_m	Number of observed samples	C
N_o	Number of observed vibration signals	3
N_p	Number of latent prototypical elements	2, 4
N_r	Number of shaft rotations	2
N_s	Number of sample points per shaft rotation	2
N_w	Moving window size used to compute the moving latent root-mean-square and moving latent kurtosis health indicators	2
N_{teeth}	Number of gear teeth for the gear attached to the reference shaft	B
$\mathcal{N}(\nu\|\mu, \sigma^2)$	Univariate Gaussian distribution over the variable ν with a mean of μ and a variance of σ^2	2
$\mathcal{N}(\boldsymbol{\nu}\|\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Multivariate Gaussian distribution over the variable $\boldsymbol{\nu}$ with a mean of $\boldsymbol{\mu}$ and a covariance of $\boldsymbol{\Sigma}$	2
$p(\mathbf{x})$	Model evidence	2–4
$p(\mathbf{x}, \mathbf{z})$	Joint distribution over the observed data variables $\mathbf{x}$ and latent variables $\mathbf{z}$	2–4

$p(\mathbf{x}\|\mathbf{z})$	Generative conditional distribution	2–4
$p(\mathbf{z})$	Latent prior distribution	2–4
$p_{ref}(\mathbf{z})$	Indicates a reference latent distribution	2
$p(\mathbf{z}\|\mathbf{x})$	Latent posterior distribution	2–4
$P_{healthy}^{SMSE}, P_{damaged}^{SMSE}$	Power of the healthy and damaged portions of the SMSE	4
$\mathbf{P}_k$	The k^{th} prototypical element	2, 4
$q(\mathbf{x})$	Approximate model evidence	2
$q_\phi(\mathbf{z}\|\mathbf{x})$	Approximate posterior latent distribution	2
r_{xy}	Sample Pearson correlation coefficient	2, 4
$\mathbf{R}$	Rotation matrix	3, 4
$\mathbb{R}^d$	The Euclidean latent coordinate space of dimension d	2–4, C
$\mathbb{R}^D$	The Euclidean data coordinate space of dimension D	2–4, C
$SE[n]$	Square envelope signal evaluated at index n	3
$SES[k_\alpha]$	Square envelope spectrum evaluated at index k_α	3, 4
T_a	Two-sided alarm threshold	3, 4
T_{bo}^l	Impact time of the l^{th} Dirac function for $x_{bo}(t)$	B
Δt	Data acquisition sampling interval [s]	2
$\mathbf{u}_k$	k^{th} node in the two-dimensional generative topographic map grid	2
$\mathbf{u}_j$	j^{th} latent component of the PCA model	A
$\mathbf{U}_d, \mathbf{U}_D$	Partial and full eigenvector matrix of $\mathbf{C}$	3, 4
$\mathbf{w}_j$	D-dimensional vector stored in the rows of the matrix $\mathbf{W}$	3, 4, C
$\mathbf{W}$	Latent mapping function for the noiseless ICA model	3, 4, C
$x_i[n]$	The i^{th} vibration signal indexed by n	2–4, C
$\overline{x}_{TSA}[n]$	The time synchronous average of signal $x_i[n]$ evaluated at index n	2

$x_{ng}(t)$	Numerical gearbox signal as a function of time t	3, B
$x_{gm}(t)$	Gear mesh signal as a function of time t	3, B
$x_{dp}(t)$	Distributed pinion damage signal as a function of time t	3, B
$x_n(t)$	Noise signal as a function of time t	3, B
$x_{bo}(t)$	Outer race bearing damage signal as a function of time t	3, B
$\mathbf{x}$	Observed data vector	2–4, C
$\mathbf{x}_i$	Vector representation of the i^{th} vibration signal	2
$\mathbf{x}_t$	Data vector representation from $\mathbf{X}$ at sensor time index t	2–4
$\overline{\mathbf{X}}$	Hankel matrix	2–4, C
$\overline{\mathbf{X}}_{ref}$	Reference Hankel matrix	2–4, C
$\mathbf{X}$	Pre-processed Hankel matrix	2–4, C
$\tilde{\mathbf{X}}$	Whitened Hankel matrix	3
$\mathcal{X}$	A set of points in the data vector space	2–4
$\mathbf{y}$	Discrete class variable	2
$\mathbf{z}$	Latent vector	2–4, C
$\mathbf{z}$	Latent vector representation of $\mathbf{x}_t$	2–4, C
$\mathcal{Z}$	A set of points in the latent vector space	2–4
$\mathbb{Z}$	Integer set	2–4
$\mathbb{Z}^+$	Positive integer set	2–4
$\sqrt{-1}$	Imaginary unit	3, 4
$\mathbf{0}$	Zero vector	2–4
$\mathbf{1}$	Constant vector with elements 1	2, 3
	Greek symbols	
α	Penalty enforcement parameter	C
α_c	Characteristic cyclic frequency [Hz|orders]	3, 4
β^{-1}	Local Gaussian covariance for the generative topographic map model	2

γ_j	Scale factor to control the prominence of the j^{th} component of $x_{ng}(t)$	B
$\epsilon_{dp}(t), \epsilon_n(t)$	Zero-mean, unit variance Gaussian white noise components	B
$\boldsymbol{\epsilon}$	Zero-mean Gaussian noise variable	3, 4
ζ_j	Damping ratio for the j^{th} component of $x_{ng}(t)$	B
$\boldsymbol{\theta}$	Parameters of the generative transition function	1–4, C
λ	Eigenvalue of the data covariance matrix $\mathbf{C}$	3, 4, A
λ_{eq}	Equality constraint	C
λ_{BC}	Box-Cox sparsity metric parameter	2
μ	Mean or expected value of a random variable	
$\mu_{m,a}$	Median value used to construct a diagnostic metric confidence interval T_a	3, 4
$\mu_{0,TSA}, \mu_{1,TSA}$	Local mixture model mean components for a synchronous representation of a latent health indicator	2
$\mu_l[t]$	Window average for the l^{th} latent variable signal	2
$\boldsymbol{\mu}$	Mean vector of a multivariate random variable	
$\boldsymbol{\mu}_i$	Feature-wise means of the $\overline{\mathbf{X}}_i$, i.e., the Hankel matrix for the i^{th} vibration signal	3, 4
$\boldsymbol{\mu}_{\mathbf{x}}$	First moment of the observed variable $\mathbf{x}$	3, 4
$\boldsymbol{\mu}_k^{(GMM)}$	Gaussian mixture model component mean vector	2
$\boldsymbol{\mu}_k^{(GTM)}$	Generative topographic map component mean vector	2
$\boldsymbol{\mu}_{ref}$	Constant feature-wise means of the the partitioned $\overline{\mathbf{X}}_{ref}$ matrix	2
$\boldsymbol{\mu}_{\mathbf{z},i}$	Mean vector for the latent representations of the i^{th} vibration data signal	4
$\pi_k^{(GMM)}$	Gaussian mixture model component weights	2
$\rho_{X,Y}$	Pearson correlation coefficient	2

σ	Standard deviation of a Gaussian distributed random variable	
σ_a	Standard deviation value used to construct a diagnostic metric confidence interval T_a	3, 4
σ_l	Variance of the τ_l variable	B
σ_{PPCA}	Standard deviation of the isotropic Gaussian generative conditional distribution for the PPCA model	3, 4
$\sigma_{0,TSA}, \sigma_{1,TSA}$	Local mixture model standard deviation components for a synchronous representation of a latent health indicator	2
$\boldsymbol{\sigma}$	Standard deviation vector	2
$\boldsymbol{\sigma}_{ref}$	Constant feature-wise standard deviations of the the partitioned $\overline{\mathbf{X}}_{ref}$ matrix	2
$\boldsymbol{\Sigma}$	Covariance matrix of a multivariate Gaussian distributed random variable	2, 4
$\boldsymbol{\Sigma}_k^{(GMM)}$	Gaussian mixture model component covariance matrix	2
$\boldsymbol{\Sigma}_k^{(GTM)}$	Generative topographic map component covariance matrix	2
$\boldsymbol{\Sigma}_{\mathbf{z},i}$	Covariance matrix for the latent representations of the i^{th} vibration data signal	4
τ_l	Jitter model variable used to represent bearing slip	B
$\tau_{\mathcal{A}}, \tau_{\mathcal{B}}$	Bandwidth tolerance parameters	3, 4
ϕ	Angular position of an order-tracked signal	3, 4
ϕ_{dp}^l	Phase of the l^{th} distributed pinion damage component	B
ϕ_{gm}^l	Phase of the l^{th} gear mesh component	B
$\boldsymbol{\phi}$	Parameters of the generative transition function	1–4, C
$\boldsymbol{\phi}(\cdot)$	Basis function	2
Φ_k	Cyclic period	3, 4

ψ_{xy}^{j}	Cosine correlation of the j^{th} component in $x_{ng}(t)$	3
ω_n	n^{th} shaft speed sample from the speed distribution $\mathcal{N}(10, 0.1^2)$	B
$\omega(t)$	Phenomenological model shaft speed	B
	Subscripts	
dB	Decibel scale applied to a quantity that measures the signal-to-noise ratio	3, 4, A
eq	Equality	C
h	Indicates a latent variable model developed using offline training	3, 4
i	Index variable indicating record number	2–4, C
j	Index variable	2–4, B, C
k	Parameter index variable	2
l	Index variable	2–4, B, C
ML	Maximum likelihood	3, 4
O	Indicates an online latent health indicator	4
ref	Reference variable	2
t	Index variable indicating sensor time, where $t \in [0, L_T - 1]$	2–4
	Superscripts	
dB	Decibel scale (see above)	3, 4, A
	Operators	
$\det(\cdot)$	Determinant operator	2
$\text{diag}(\cdot)$	Vector to matrix operator	2–4
$DFT\{\cdot\}$	Discrete Fourier transform	3, 4
$\mathbb{E}\{\cdot\}$	Expectation operator	2–4, A, C
$\mathcal{F}(\cdot)$	Fourier operator	3
$\max(\cdot)$	Max operator	3, 4
$\text{median}(\cdot)$	Median operator	3, 4
$\text{tr}(\cdot)$	Trace operator	2
$\rightarrow$	Push-forward	1–4, C

$\lvert \cdot \rvert$	Absolute value	3, 4
$\lVert \cdot \rVert_2$	L_1 norm	3, 4
$\lVert \cdot \rVert_2$	L_2 norm	2–4
$\lVert \cdot \rVert_\infty$	Uniform norm	3, 4
$\otimes$	Convolution operator	B

Chapter 1

Introduction

1.1 Background

Asset Condition Monitoring (CM) has become an ever-crucial facet for maintenance and management decisions applied to critical industrial assets. Not only do industries rely on minimal asset downtime to maximise productivity, but the implications surrounding financial performance, product reliability, and personnel safety and security provide strong motivations for the development of robust CM technologies. The eventual deterioration of an asset's health state is unavoidable. Hence, asset health (condition) monitoring is necessary to address and effectively manage such degradation. Condition Based Maintenance (CBM) represents a maintenance strategy that seeks to address the financial implications of performing asset maintenance by utilising information related to the instantaneous asset health state to guide maintenance decision-making [1, 2]. In any CBM system, there are three stages on which the techniques operate: *(i)* data acquisition, *(ii)* data processing, and *(iii)* maintenance decision-making [1]. Of the three stages, the data processing step requires extensive analyses and academic or industrial practitioner involvement to perform fault diagnosis and fault prognosis and has attracted much interest in recent years.

To supplement a CBM strategy, Prognostics and Health Management (PHM) systems serve to provide an estimation of the underlying asset condition to guide maintenance decision-making based on the recorded and processed asset data [3]. PHM strategies represent an extension of CBM and aim to efficiently detect and diagnose deviations in the instantaneous asset health state, predict when failure will occur, and drive maintenance operations [4]. Fault diagnostics and prognostics are two important aspects of both CBM and PHM strategies, and these aspects rely on the development of asset Health Indicators (HIs) to evaluate the instantaneous asset health state and infer the Remaining Useful Life (RUL) [5, 6]. Fault diagnostics is a precursor to prognostics and emphasises the ability to detect, identify, and trend damage degradation in an asset using CM methodologies [3, 4, 7, 8]. Damage trendability is a necessary preceding step to prognosis once a fault has been detected and correctly identified,

as degradation trends are necessary for subsequent assessment of the relative health of the asset, i.e., to infer the RUL [3]. Hence, the usefulness and robustness of fault diagnosis methodologies directly influence all downstream PHM tasks and are thus critical to effective condition inference. PHM is an essential component of Engineering Asset Management (EAM), whereby PHM maximises the expected operational life of critical assets by facilitating methodologies for effective predictive maintenance to enhance asset longevity, enable effective repair procedures, and improve operational efficiency.

Of the typical mechanical components within industrial assets, there is strong evidence to suggest that gearboxes are of critical importance in mining and power generation industries, e.g., draglines and wind turbines [9]. Industrial gearboxes typically operate in harsh environments, which make them highly susceptible to failure and often result in prolonged downtimes. This places a greater emphasis on the need for robust CM methodologies to enable effective and practical EAM [10]. Additionally, gearboxes comprise of multiple potential failure avenues, e.g., gear tooth cracks, surface pitting, gear spalling, or localised bearing damage [11–13], exhibit complex phenomena because of tooth meshing [14], experience time-varying operating conditions [15–17], and are typically encased in a gear housing for mechanical support and to provide lubrication to interacting internal components. Hence, the gearbox CM problem is a non-trivial endeavour, and it is of critical importance that robust CM methodologies be developed that are responsive to changes in the instantaneous gearbox health state in the presence of variable operating conditions and harsh operating environments.

To extract asset condition information, vibration signatures from rotating machinery offer numerous advantages over alternative data formats, e.g., oil composition analysis [18, 19]. Not only does it allow for non-intrusive monitoring, but rotating machinery vibrations capture periodic dynamic behaviour, which can be directly attributed to components of interest within the asset [10, 18]. Fault signatures often manifest as periodicities in the data, and identifying vibration signatures corresponding to damaged asset components is an all-important task [13]. This has led to interest in frameworks which exploit said periodicity present in vibration signals from rotating machinery, e.g., cyclostationarity analysis [20–24], methodologies which utilise learning-based methods to perform fault diagnosis and prognosis, e.g., supervised learning for fault classification [25–27] or anomaly detection [28, 29], and methodologies which perform advanced signal processing to enhance fault characteristics, or proxies thereof, within the recorded vibration data, e.g., band demodulation [30–33]. The underlying assumption of these

methodologies is that the health state of an asset is a covariate captured within the recorded vibration signals, and it is crucial that these covariates be extracted by CM techniques via HIs to support maintenance decisions [5].

CM methodologies for vibration-based fault diagnostics are important to critical industrial assets. Hence, a literature study into vibration-based CM methodologies directed towards gearbox failures is conducted to establish the focal points within the current literature landscape.

1.2 Literature survey

In this section, a literature study directed towards vibration-based CM methodologies in gearbox fault diagnostics applications is performed. The primary objective is to give context to the gearbox CM problem and the approaches used to perform fault diagnostics. The study is conducted as follows: In Section 1.2.1, attention is given to vibration data obtained from gearboxes and context is given to the gearbox fault problem. Then, the most prominent data-driven approaches explored in the literature are investigated. First, learning-based methodologies are investigated in the context of gearbox CM in Section 1.2.2. Second, signal processing-based methodologies are investigated in the gearbox CM context in Section 1.2.3. These two methodologies are considered data-driven as they are driven to extract or enhance features from gearbox vibration data that are useful for the condition inference task [8].

1.2.1 The nature of gearbox vibration data

Gearbox vibration data acquired via transducers contain information related to the mechanical components within the gearbox casing. The two gearbox components which attract the most interest in the literature are gears and bearings, and it is important to understand both the dynamics of these components and their associated failure mechanisms [34]. Gears produce dominant vibrations due to innate variations in tooth stiffness during the meshing process [11]. Such variations are caused by manufacturing tolerances, which produce deviations from the ideal involute tooth profile, backlash, tooth deflections due to variable load conditions, gear-pinion contact ratios, and wear, e.g., uniform wear during the initial gearbox operating period [35–37]. Gear vibration components tend to dominate the spectra of gearbox vibration signals and produce positive integer components of the gear mesh frequency, which is given by the product of the number of gear teeth and the instantaneous gear shaft speed [11]. Hence, meshing phenomena and the associated gearbox vibration signal content are a function of the

gearbox operating conditions, an important consideration that is discussed further in Section 1.2.1.3.

Bearings also produce vibrations due to geometrical imperfections induced by the bearing manufacturing process, e.g., surface asperities and periodical variations in the effective bearing assembly stiffness due to the variation in roller element position through time [38]. However, the energy associated with the vibrations from bearings are far less prominent than those from gears, which increases the analysis complexity for CM methodologies once the bearing health states changes and fault conditions develop as important diagnostic information is easily masked [37].

1.2.1.1 Gearbox failure mechanisms

Gearbox failure mechanisms are often caused by gears and bearings [10, 14, 34, 39]. As such, both are discussed in turn. While this literature study is focused on gearboxes, interested readers are directed to the comprehensive discussion surrounding failure mechanisms of rotating machinery given by Randall [37]. Gear failure mechanisms manifest within vibration data via changes in the gear mesh stiffness [36] which invoke both amplitude modulation effects [34, 36] and phase modulation effects [37, 40, 41] via changes in the gear shaft speed and changes in the tooth contact region. Zhou *et al.* [42] reviewed different vibration signal modulation effects induced by gearbox faults in the context of rotating machinery. Two prominent gear damage modes are localised and distributed gear damage, whereby the former operates via amplitude modulation around a localised tooth mesh period, while the amplitude modulation for the latter occurs over the whole gear period [36, 43]. Both damage modes produce sidebands within the spectra of the vibration signals around the meshing frequency and its integer harmonics; however, the modulating frequency around the gear mesh frequency components will be different [36].

Bearing failure mechanisms are discussed in this literature study in the context of rolling-element bearings. Such bearings transfer load and facilitate rotational motion via rolling elements placed between two concentric rings [39, 44]. Bearing failure mechanisms typically manifest in damaged regions that are caused by mechanical damage, crack development, corrosion, and wear damage [39, 45]. In vibration data, the damage regions introduce three classifications for bearing faults, namely *(i)* single-point defects, *(ii)* multi-point defects, and *(iii)* distributed defects [8, 46, 47]. Single-point faults, e.g., small pits or spalls, are the most popular in the literature and occur when rolling elements periodically collide with a localised

damage region on the outer or inner race of the bearing or if damage on a rolling element contacts the inner or outer race [13, 37]. The frequency of the impulses generated by the collisions are distinctive to the damage locality [46]. Bearings with damage typically exhibit non-monotonic degradation trends due to changes in the topological and tribological properties of the damaged regions throughout the bearing lifespan [48, 49]. Randall and Antoni [13] provide a succinct tutorial on CM techniques for bearing fault diagnostics.

Outer race faults are expected a priori to manifest within the load zone and are expected to give approximately uniform impulses for each shaft rotation [50]. In contrast, rolling element and inner race bearing fault signatures experience bearing fault amplitude modulation. Bearing fault amplitude modulation occurs via the fault travelling through the radial load zone and via variations in the transmission path if the fault location moves [13]. For the former, the radial load distribution is of interest for rolling element and inner race faults, as this affects the amplitude of the impulses. McFadden and Smith [46] proposed a model which accounts for the radial load distribution for single point defects on the inner race. The energy of the impacts from bearings typically excites a wide range of frequencies, but the energy is not as dominant as energy from other deterministic gearbox components, e.g., gear tooth interactions. Hence, the associated bearing single defect frequencies are often masked by machine noise and gear mesh components which impede their detectability and identification in the spectra of gearbox signals [51].

However, an impulse or a train thereof, whose time duration is extremely short in comparison to the time duration between impulses, excites resonance frequencies within the bearing assembly, thereby causing higher frequency regions to become more prominent in the bearing fault signal. The high-frequency resonance regions operate as carrier frequencies and undergo amplitude modulation with a modulating frequency corresponding to the bearing defect frequency [13, 51]. This modulation is different to the modulation caused by the radial load zone, which modulates the impulse train by the shaft speed for inner race faults or by the cage speed for rolling element faults [50]. This resonance modulation phenomenon is critical to the detection and identification of bearing faults and subsequently leads to envelope analysis techniques which perform demodulation on a filtered gearbox signal [51]. In envelope analysis, the objective is to extract and enhance fault information through the use demodulation methods to extract the modulating frequency, i.e., the bearing defect frequency, from an appropriate resonance band. Envelope analysis is an important technique which is

discussed further in Section 1.2.3.4.

1.2.1.2 Modelling strategies

To understand gearbox faults and the mechanisms surrounding fault generation, physical model-based methods have been investigated in the literature. Such models aim to develop systems that are representative of the dynamic behaviour of physical gearbox assets [52]. There exist two categories for physics-based modelling, namely *(i)* dynamic models and *(ii)* empirical models. Dynamic models refer to systems which simulate the dynamic responses of an asset using physical laws and Newton's laws of motion [52]. Empirical models, i.e., phenomenological models, use empirical methods deduced from observations of the physical asset phenomena and the components therein to describe and generate gearbox vibration data [36].

Dynamic models for gearboxes are grouped into two categories, namely lumped mass models and finite element models, and for both categories the stiffness of the gear mesh is a highly influential time-varying parameter of interest [52]. Lumped mass models consider gearbox components to be rigid bodies modelled as a spring-mass-damper system and Newton's second law is applied to the rigid bodies to obtain a system of differential equations under some initial condition that can then be solved for using numerical integration techniques. Lumped mass models require a model of the time-varying mesh to be known a priori, which manifest either through explicit or implicit modelling, e.g., using the square waveform method [53], the finite element method [54–56], the potential energy method [57], or via iterative mesh stiffness evaluation techniques [58, 59]. Furthermore, incorporating gear and bearing failure mechanisms is an active research endeavour, e.g., Lei *et al.* [60]. Various mesh stiffness integration techniques exist for different gear fault cases [12, 52, 54, 61]. A popular lumped mass model for gearboxes was proposed by Bartelmus [62], and was extended by Wu *et al.* [63] to include tooth crack growth. Examples of gear damage models are proposed by Mohammed *et al.* [64, 65], Chaari *et al.* [66], Lei *et al.* [60], and He *et al.* [67]. Examples of bearing damage models are proposed by Sopanen and Mikkola [68], Sawalhi and Randall [69], Patel *et al.* [70], and Wang *et al.* [71]. Chaari *et al.* [41] investigated the implications of time-varying operating conditions for a gearbox model without gear damage. They demonstrated that variable operating conditions produce spectral phenomena similar to the spectra of gear fault signals, e.g., localised tooth faults [36], and this may directly affect the fault diagnosis task.

Finite element models, alternatively, offer flexibility to different gear profiles and gear

fault cases as the gear mesh stiffness is not required to be known a priori and can be obtained implicitly. However, the results from finite element models are sensitive to contact tolerances, element density, and the assumed element type [52, 72]. Hence, such models are typically used to investigate gear dynamics by iteratively substituting the gear mesh stiffness determined from a finite element result into a lumped mass model or to validate gear mesh stiffness models [60]. Kiral and Karagülle [73] proposed a loading model to incorporate a radial load distribution into bearing finite element models and introduced localised defects by modifying specific radial loads in the loading model to simulate impulsive forces.

Phenomenological models, alternatively, rely on a physical understanding of the gearbox to generate signals that describe the different gearbox components. Such models approximate the underlying gearbox phenomena and are primarily used as they are computationally efficient and can readily be used to investigate the implications of data-driven CM methodologies for different choices of the model parameters. McFadden [74] presented a simple model for gear meshing vibration under constant operating conditions and incorporated amplitude and phase modulation effects to simulate local gear tooth defects. Randall [36] presented different phenomena for gears and the effects of different gear faults under constant operating conditions, McFadden and Smith [46] proposed a model for single point bearing defects and then extended it to multi-point defects in subsequent work [47], Braun and Seth [75] proposed a decomposition for signals with repetitive signatures for rotating machinery, and Borghesani *et al.* [76] investigated different bearing signal models directed towards the localised impact times and the assumptions made therein.

1.2.1.3 Implications of gearbox operating conditions

Time-varying Operating Conditions (OCs) greatly influence the attributes of gearbox vibration data and complicate the diagnosis procedure [17]. This is due to an important feature intrinsic to vibration signals from rotating machinery whereby the statistical properties are not stationary, i.e., the statistical behaviour of the signal is constant in time, but are controlled by some hidden periodic mechanism [20, 77, 78]. This signal classification is called cyclostationarity and is an important component of the literature for vibration signals from rotating machinery discussed further in the fault diagnosis context in Section 1.2.3.4. Gearbox vibration data can be classified as angle-cyclostationary as the components of the data are periodic with respect to the angular shaft position of the gearbox rather than time [20]. If the OCs are constant through time, angle-cyclostationary signals can also be classified as

time-cyclostationary [78, 79].

However, time-varying OCs give rise to three manifestations which violate the notion of cyclostationarity: First, amplitude modulation effects are present due to load variations, and this causes variations in the instantaneous shaft speeds [36, 40, 80]. Second, frequency modulation effects are present due to variations in the instantaneous shaft speed, and this causes spectral smearing as the characteristic gearbox frequencies are functions of the rotating speed [81]. Third, the transmission path between the transducer and the vibration sources introduces amplitude and frequency modulation depending on the gearbox structural characteristics and is speed-dependent [15, 81–83]. Hence, for time-varying operating conditions, gearbox vibration signals are considered cyclo-non-stationary [80].

Amplitude modulation implies that variance-sensitive features extracted from vibration data will be sensitive to time-varying OCs [17, 84]. Stander *et al.* [85] proposed the application of a low-pass filter to the absolute of an analytic signal, i.e., the envelope signal [86], to estimate the modulation function for load modulation normalisation for gearbox vibration signals under time-varying OCs. Schmidt and Heyns [87] proposed using the square root of a moving median filter applied to the square envelope signal as an estimate of the modulation function to attenuate the effects of amplitude modulation induced by time-varying OCs. The implication of frequency modulation is that there is smearing of frequency components in the spectrum. This spectral smearing is accounted for using computed order tracking [88, 89] or by using tacholess order tracking [90, 91] techniques to change the frequency base to shaft orders, i.e., multiples of shaft speed, rather than absolute time frequencies [88, 92].

1.2.2 Learning-based methodologies for gearbox fault diagnostics

Learning-based methodologies represent a category of gearbox condition monitoring that utilise statistical and probabilistic models to conduct fault diagnosis tasks [5]. The increase in computing power, data size and availability, and the rapid advancement and industrial uptake of learning-based techniques have caused a rapid growth in learning-based methodology popularity [93]. In general, a learning-based methodology seeks to bypass any complex decision-making to a model that caters to accomplishing fault diagnosis tasks [94, 95]. Many learning-based works treat fault diagnostics as a pattern recognition problem applied to the condition of rotating machinery, whereby the model is concerned with performing fault identification [7]. Alternatively, the learning-based model operates as a transformation operator on gearbox data and produce HIs that are used for condition inference for the fault diagnosis

task [6, 28]. The latter is more analogous to signal processing-based methodologies, which use advanced expert-driven tools to enhance the covariates related to the instantaneous asset health state.

The terminology learning-based is an umbrella term used to encapsulate statistical learning, machine learning, and deep learning. It is important to the clarity of this study that these terms be defined to avoid any definition ambiguity. Statistical learning is a general framework for models which learn from data to perform predictive or decision-based tasks [96]. Statistical learning encompasses various techniques, e.g., classification, regression, clustering, and dimensionality reduction, and typically makes use of linear models to facilitate automatic decision-making. Machine learning extends statistical learning via the use of nonlinear models to improve the model flexibility and is interested in algorithms which improve model performance to enhance decision-making [97]. For both statistical learning and machine learning, gearbox vibration data is high dimensional, and feature engineering is required to extract hand-crafted time domain, frequency domain, and time-frequency domain signal features to represent the underlying gearbox health state [8, 94, 98–100]. Sharma and Patel [11] and Jardine *et al.* [1] discuss different features that can be extracted from gearbox vibration data.

However, feature selection strategies are non-trivial, capture redundant information, are biased to past knowledge related to features that are sensitive to specific failure mechanisms, and may not generalise between different rotating machinery [6, 7, 94]. Hence, deep learning seeks to circumvent the feature engineering step and decrease the requirement for domain specific knowledge by using complex nonlinear functions with many parameters to learn important features automatically [97]. There are three main frameworks within learning-based methodologies that are popular for rotating machinery diagnostics, namely *(i)* supervised learning, *(ii)* unsupervised learning, and *(iii)* transfer learning. Each of these frameworks are discussed in turn. The connective tissue between each framework resides in the assumptions given to data ability and what can be accomplished using the available data.

1.2.2.1 Supervised learning

Supervised learning refers to a learning-based task whereby a set of data feature vectors and associated labels are used to construct a function which captures the mapping between the input data features and the labels. For gearbox fault diagnostics, the predominant mapping of interest is between vibration data, or features thereof, and the potential fault

conditions of an industrial asset [7, 93, 94]. There are two distinct mapping techniques for supervised learning, namely discriminative techniques, and generative techniques [101–103]. Discriminative techniques, generally, are methods which seek to obtain an accurate mapping between the labels and associated data features or probabilistically model the conditional label distribution given the features [103]. Generative techniques seek to model the joint distribution over the labels and data features and compute the posterior distribution over the labels given some data features using Bayes' rule [102, 103]. Methods in the literature for the fault diagnosis task under these two techniques are discussed in this study.

The prominent[1] discriminative-based supervised learning techniques in the literature are k-nearest neighbours [106, 107], decision trees [108], Support Vector Machines (SVMs) [109], and neural networks [8, 94]. The most prominent generative techniques in the literature are probabilistic graphical models such as Bayesian classifiers, e.g., Naive Bayes [110], and Markov models [101]. Moosavian *et al.* [111] compared the classification performance using k-nearest neighbours and a neural network developed on thirty features extracted from the power spectral density of gearbox signals from an internal combustion engine. Lei [112] used a k-nearest neighbours classification algorithm on weighted time domain and frequency domain features from gearbox vibration data to identify gear crack levels. Wang [113] extended the methodology from Ref. [112] by using the features from the binary wavelet packet transform and Principal Component Analysis (PCA) [114] as a feature dimensionality reduction tool to identify gear crack levels under time-varying OCs. Wan *et al.* [115] investigated the performance implications of different dimensionality reduction strategies for the gear crack level problem. Zhang *et al.* [116] used a C4.5 decision tree on the features from an Auto-Encoder (AE) and adjusted for imbalance in the hand-selected signal features, using a weighted minority oversampling procedure. Praveenkumar *et al.* [117] compared the classification performance of two decision trees built on time domain features and frequency domain features, in particular the area under segments of the Fourier spectra of vibration signals, respectively.

Gryllias and Antoniadis [99] used a hybrid two-stage SVM approach with a one-against-all strategy developed on features from the order-tracked envelope signals from a rolling-element bearing phenomenological model, to detect and classify outer race and inner race bearing fault conditions. Liu and Gryllias [118] used features from the cyclic spectral correlation and

[1]The literature surrounding supervised learning techniques is extensive and cannot, realistically, be introduced, documented, and discussed in full in this study. The work of Liu *et al.* [7], Khan and Yairi [104], Zhao *et al.* [93], Lei *et al.* [94], and Zhu *et al.* [105] all provide excellent reviews on learning-based techniques for condition monitoring.

the cyclic spectral coherence [119] combined with a support vector data description [120] model with artificial outliers to perform bearing fault detection. Muruganatham *et al.* [121] used features from a singular spectrum analysis methodology applied to vibration signals as inputs to a neural network used to classify bearing fault conditions. Unal *et al.* [122] extracted features from the spectra of envelope signals to classify bearing faults using a neural network whose parameters were optimised with a genetic algorithm. Muralidharan and Sugumaran [123] compared the classification performance of a naive Bayes algorithm and a Bayes net algorithm developed on features from the discrete wavelet transform for centrifugal pump faults.

Recent advances in deep learning approaches have been adopted and applied in the condition monitoring literature for supervised fault diagnostics [94, 124]. A noticeable trend in recent research is an increased interest in leveraging deep neural networks, such as AEs, to supplant the conventional feature engineering and dimensionality reduction steps that are often integral to many supervised condition monitoring methodologies [7, 27, 93, 104, 125, 126]. This approach aims to streamline the supervised process by automating the feature extraction role. This interest stems from the need to circumvent potential bias prevalent in feature engineering, which might favour features applicable only to a particular failure condition or industrial asset.

Additionally, there is a focus on employing powerful classification network architectures to enhance the generalisability of a supervised methodology [124]. Lu *et al.* [127] used a stacked denoising AE architecture to improve the fault classification performance by extracting useful hidden feature representations. Mao *et al.* [27] focused on improving the features learnt by the AE by using the labels given to the observed data to improve the hidden feature representation class separability by penalising latent similarity for different fault classes. Wu *et al.* [126] use the Short-Time Fourier Transform (STFT) of gearbox vibration data as input to a deep auto-encoding architecture with convolutional layers to perform classification-based fault diagnosis. AE networks also make use of convolutional layers [128–131], residual network layers [132–134], or are represented by recurrent neural networks [135, 136] or deep Boltzmann machines [137–139] to improve the model performance under different failure conditions and OC states. Ding *et al.* [140] propose using the synchrosqueezed wavelet transform [141] with bicubic interpolation and a transformer architecture [142], which accounts for time-frequency representations to classify different bearing damage fault modes.

1.2.2.2 Unsupervised learning

In the supervised learning scenario, the dependence on labelled vibration data is problematic for four reasons [6, 94]. First, obtaining labelled fault data for all asset states is an expensive endeavour and is often not practical to acquire, which increases the difficulty of acquiring information related to the asset health state. This is particularly prominent for assets with high operation requirements or assets used for critical safety mechanisms [6]. Second, the condition of rotating machinery is predominantly healthy for a significant portion of its operational lifespan. As such, labelled condition information is sparse and places further strain on the faulty data collection process. Third, it is not always possible to obtain labelled data for all fault categories, which presents a discontinuity in many supervised learning techniques which does not account for the possibility of an unseen fault condition class [143]. Finally, the fault labelling process requires human intervention to define the point where a fault is present, which is not a deterministic task as CM research has not been diminished and is still an active research endeavour, i.e., no singular optimal approach or methodology exists. The learnt separability between healthy and the labelled fault conditions is a reflection of the labelling process. Furthermore, the degradation of different gearbox failure modes is continuous through its lifespan and not discrete. Hence, most supervised strategies cannot infer the degradation process without inducing discrete cut-off regions at specific damage levels [8], e.g., Refs. [112, 113, 115]. Unsupervised learning approaches in gearbox fault diagnostics bypass the data label problem by changing the application objective of learning-based models from fault classification to fault detection.

Unsupervised learning techniques represent different fault diagnosis strategies, e.g., discrepancy analysis [144–147], novelty or anomaly detection [28, 118, 148–151], or residual signal analysis using auto-regressive models [144, 152]. There are two primary approaches used to develop unsupervised CM methodologies, namely *(i)* offline training, and *(ii)* online training. The principle upon which these unsupervised strategies reside is that reference, assumed healthy, gearbox vibration data can be used to estimate the parameters of some model, i.e., offline training, thereby allowing the model to act as a representation of the reference gearbox condition [153]. Hence, conventional applications of unsupervised models prefer offline training because it aligns with their primary objective, which is to accurately capture a reference gearbox condition. The model is then used on novel vibration data to evaluate its similarity to the reference data and to trend deviations in the asset condition throughout the operational

lifespan.

This approach is related to density estimation under the assumption that the reference vibration data is concentrated within some region in space, and anomalous data is not concentrated within the same spatial region [154, 155]. This assumption is primarily used to drive unsupervised fault detection techniques. In the discrepancy analysis setting, information related to the periodicity present within vibration data from rotating machinery is exploited to obtain discrepancy measures, i.e., HIs, in terms of time or shaft angle that can be used to infer deviations, e.g., Ref. [145]. Models used for unsupervised fault detection include but are not limited to: Gaussian mixture models [145], one-class SVMs [146] and its deep learning-based extension [156], support vector data description [157, 158] and its deep learning-based extension [159], AEs [160], Variational Auto-Encoders (VAEs) [28], and Generative Adversarial Networks (GANs) [28]. Lu *et al.* [153] investigated and compared different anomaly detection methods in the context of bearing fault detection.

Latent Variable Models (LVMs) are a generative modelling technique concerned with using models to capture the probability distribution of data [161]. The LVM formulation posits that an unobserved, i.e., latent, set of variables effectively describe the intrinsic features of variation in the data. In this setting, the goal is to obtain a suitable probabilistic model that both describes the data distribution of interest and can recover the underlying sources of data variation [97, 161, 162]. The desired outcome of the latent variables is to obtain a more compact and meaningful lower-dimensional representation of the observed data. In the conventional LVM application setting, access to a reference, assumed healthy dataset of vibration signals is defined a priori to estimate the LVM parameters, i.e., offline training is used. In this way, the LVM represents the asset in a healthy state and aim to detect deviations from the captured state.

Additionally, LVMs provide a data space and the latent space, i.e., the latent manifold, for condition inference. LVM-based condition inference techniques predominantly use the observed data space, and this can only provide two HI metrics for condition inference, namely *(i)* the reconstruction error, and *(ii)* the reconstruction likelihood [163]. Ref. [164] demonstrated that temporal preservation is an effective approach for LVMs and advocates for the preservation of temporal aspects, i.e., the time structure within the vibration data, to enhance the effectiveness of LVMs in vibration-based CM applications. Furthermore, temporal preservation is an essential approach to obtain an informative latent manifold, is the first

step towards latent manifold interpretability in a fault diagnostics context, and ensures that LVM-based methodologies can extend beyond the anomaly detection task and into all aspects of the fault diagnosis task. The latter is achieved by exploiting the temporal structure to obtain HI signals. Signal processing-based techniques can then be applied to the HI signals to isolate and identify the fault mode present. Temporal preservation is an important basis for vibration-based LVMs in CM, and discussed further in Section 2.2.

To capture the reference condition, many offline LVMs are developed using reconstruction-focused objectives. Reconstruction-focused objectives identify components in the data which capture the prominent instances of data variance to obtain minimal reconstruction errors. This is optimal for offline training in CM applications as the goal is to best approximate the reference asset distribution devoid of any fault conditions [101, 165, 166]. Models such as PCA, AEs, and VAEs can be identified as reconstruction-focused methods as they seek to maximise the fidelity of the representation of the data distribution by producing a latent manifold which captures the prominent components of variance in the data [101]. GANs, as an adversarially trained method developed by discriminating real from generated data, are implicitly reconstruction-focused as the discrimination process is guided by approximations of the data density ratio, i.e., the ratio between the true and approximate data distributions, to drive model training [167, 168]. Hence, GANs are implicitly reconstruction-focused in the sense that they seek to generate data that is as close to samples from the true data distribution, but their formulation does not attempt to encode information into the latent manifold and this makes the latent manifold uninformative [166, 169].

In the literature, various works have investigated using offline LVMs for the condition monitoring task. Booyse *et al.* [28] demonstrated the practicality of two LVM formulations, namely GANs [167] and VAEs [165], and used the discriminator of a GAN as a proxy to the data space likelihood to infer whether novel vibration data was from the reference dataset. Hu *et al.* [170] used a Wasserstein GAN for fault detection and compared its performance to a conventional GAN. Gugulothu *et al.* [171] proposed an embedded latent HI that uses a Euclidean distance measure to compare a latent representation of an observed data sample to a set of healthy latent representations. Marx and Gryllias [172] used an AE trained on the Square Envelope Spectrum (SES) of healthy vibration data and augmented the available SES to represent expected failure modes. Ref. [172] regularised the AE to enforce latent dissimilarity in the representations of the expected failure modes. Kim *et al.* [173] proposed a reconstruction

along the projection pathway approach for the novelty detection task, which evaluates the Euclidean distance or the Mahalanobis distance between the hidden AE representations of an observed sample and the hidden AE representations of the reconstruction of the observed sample. González-Muñiz *et al.* [29] use the approach from Ref. [173] but propose using just the latent space of an AE for condition inference. Park *et al.* [174], independently from Kim *et al.* [173], proposed using the Manhattan distance between the hidden AE representations of an observed sample and the reconstruction thereof. Additionally, Ref. [174] use the mean-centred cosine similarity for the hidden representations for the novelty detection task. Kwon *et al.* [175] performed novelty detection using a combination of the reconstruction error and a cosine similarity metric applied to the gradient vectors of the AEs. The latent manifold is a representation of the underlying features within gearbox vibration data and can offer unique information that can also be beneficial to the fault diagnosis task.

Alternatively, a portion of the literature focuses on using unsupervised methods as decomposition techniques to recover representations that are useful to the fault diagnosis problem. These methods assume the observed vibration data contains faulty signatures, and the objective is to apply a decomposition approach to the observed signal to obtain a novel signal with improved diagnostic information. Hence, these methods differ in application to conventional unsupervised methods as they are focused on extracting informative components in the data and do not use an unsupervised model to detect deviations from a reference condition, i.e., they use online training.

Additionally, the methods are inherently temporal-preserving, as they assume that the data has temporal structure that enables the decomposition process [176]. These online unsupervised techniques can be broadly categorised into two groups, namely *(i)* reconstruction-focused methods, and *(ii)* interpretation-focused methods. To guide the decomposition process, reconstruction-focused methods use variance as a proxy to informativeness, while interpretation-focused methods use statistical independence and non-Gaussianity as a proxy to informativeness. The former relates to extracting data components, which maximise the variance in the observed data for optimal data reconstruction [101, 177]. The latter uses statistical independence to enforce that the decompositions share minimal information and non-Gaussianity as a measure of informativeness to enforce statistical independence and to obtain proxy measures to extract components from the data which are interpretable [166, 178].

Much academic interest has been shown in the use of Singular Spectrum Analysis (SSA) techniques [177] to decompose a Hankel matrix of vibration data into a sum of variance-weighted rank-one matrices [177]. The SSA decomposition process is performed online and is variance-driven, i.e., reconstruction-focused, and strongly related to PCA [179], SSA has been investigated as a non-parametric alternative to the wavelet transform [180], and the distribution of singular values induced by the SSA decomposition is strongly related to the power spectra of vibration data [181]. Zhao and Jia [176] investigated the implications of re-weighting the rank-one matrices using a periodic modulation intensity factor, which seeks to maximise components which exhibit strong periodicities in the envelope of the decomposed SSA components. Chen *et al.* [182] extended the work of Ref. [176] by first using an auto-regressive model to remove strong deterministic components from the vibration signals. Golafshan and Sanliturk [183] investigated the implications of performing SSA in the time domain and in the frequency domain.

As an alternative to SSA techniques, certain academic works have considered interpretation-focused methods to drive the online unsupervised decomposition process. He *et al.* [184] demonstrated the application of a negentropy-based Independent Component Analysis (ICA) methodology [185] for repetitive transient applications and proposed a methodology to identify the prominent transient components in the decomposed ICA representations. Wang *et al.* [186] used a Constrained Independent Component Analysis (CICA) methodology [187, 188] to extract specific signal components using prior knowledge of the gearbox characteristics and a reference fault signal. Yang *et al.* [189] used the CICA methodology on the order-tracked square envelope signal, where the signal was band-pass filtered using parameters from the kurtogram [190]. Wang *et al.* [191] investigated the use of ensemble empirical mode decomposition combined with an ICA methodology as a decomposition method for bearing fault diagnosis and demonstrated an ability to extract novel signals with improved fault signatures.

1.2.2.3 Transfer learning

An alternative approach to the data label problem discussed in Section 1.2.2.2 is facilitated by transfer learning and domain adaptation tasks. This learning strategy aims to use diagnostic information gained from one application to another novel application, i.e., information re-use, to combat issues such as fault class imbalance, sparse fault labels, cross-domain fault diagnosis, or data covariate shifts [94, 192, 193]. In transfer learning, the domain refers to the observed gearbox signal data or features thereof, and the task refers to the fault condition labels given

the gearbox data. typically, two domains of interest exist in this learning task, the source domain which has a dense set of data and labelled fault conditions, and the target domain where the fault labels are sparse or minimal. For the latter, the lack of labelled data means that the supervised learning model used for intelligent fault diagnosis cannot be effectively developed to perform condition inference [194]. This shortcoming is a central problem of interest for transfer learning methodologies, and various strategies have been employed in the literature improve the performance of a classification model in the sparsely labelled domain [195, 196]. This manifests in two strategies, *(i)* transfer learning for fault feature extraction, and *(ii)* transfer learning for fault classification [195].

In the work of Li *et al.* [196], four application scenarios in transfer learning for fault classification were identified, namely *(i)* generalisation performance improvement, *(ii)* partial domain fault diagnosis, *(iii)* emerging fault detection, and *(iv)* compound fault decoupling. Generalisation performance improvement is focused on improving classification performance in domains where the OCs of the equipment are varying between domains. Partial domain fault diagnosis is concerned with applications where the labels from the target domain exist as a subset of the labels in the source domain. Emerging fault detection is concerned with the alternative scenario from partial domain fault diagnosis, whereby the target domain has more labels of interest than the source domain, i.e., there are fault conditions in the target domain not present in the source domain. Compound fault decoupling is a scenario whereby the target domain labels are non-unique but known, i.e., representative of more than one fault condition, while the source domain has individual labels for each of the possible fault conditions. Regarding transfer learning methodologies, Li *et al.* [196] split transfer learning approaches into three categories: *(i)* instance-based transfer learning, *(ii)* model-based transfer learning, and *(iii)* feature-based transfer learning. The instance-based approach assumes a high likelihood of sparsity in the target domain fault labels, and the objective is to obtain samples in the source domain that can assist with estimating the target domain classification model, where the source domain samples are weighted based off their suitability to the target domain [197–199]. The theoretical weighting for source samples is the likelihood ratio between the marginal domain distributions [196]. The model-based approach operates under the assumption that there are shared features of the observed data in both domains. This approach is then concerned with the ability to impart transferable model knowledge between the two domains by pre-training within the source domain and fine-tuning the classification model to the target domain or using a joint training strategy [200–202]. The feature-based approach uses a mapping function to

produce a domain invariant latent feature space, i.e., perform domain adaptation to produce a feature space where the discrepancy between representations of samples from different domains are minimal [203–206].

1.2.3 Signal processing-based methodologies for gearbox fault diagnostics

Signal processing-based methodologies focus on enhancing the covariates within vibration data related to the instantaneous health state of an asset, using processing strategies based on properties inherent to vibration data from rotating machinery. The ability to uncover rich diagnostic information is of key interest and requires the use of feature extraction techniques to establish a process-transformed vibration signal with fault signatures. These methodologies are predominantly performed in an online setting, i.e., covariate extraction is performed for every vibration data instance. Different methodologies within this category of gearbox condition monitoring are discussed in this section.

1.2.3.1 Conventional signal analysis

Conventionally, gearbox fault detection strategies use time domain indicators as proxy measures, i.e., HIs, to estimate the underlying severity of gearbox faults [8, 11, 19]. Fault detection refers to the process of identifying a deviation from some baseline condition, but the root cause of the deviation is unknown [8]. Time domain indicators typically monitor some statistical property of gearbox vibration data to detect deviations caused by changes in the instantaneous health state. Two popular statistical metrics include the Root-Mean-Square (RMS) and the crest factor of a vibration signal which operate as proxy measures to monitor changes in the signal energy or amplitude induced by gearbox failure mechanisms [19, 207, 208].

However, time-domain indicators of raw vibration signals often struggle to generalise to different failure scenarios and require careful indicator selection depending on the fault of interest, are easily impeded by noisy environments, and certain components, e.g., gear mesh components or driveshaft components, tend to dominate the statistical properties of a signal, which may mask the presence of the underlying fault condition, e.g., a localised bearing fault [5, 104]. Hence, signal processing-based techniques to enhance the underlying fault condition of interest and attenuate non-diagnostic signal information have risen in popularity [8, 13]. These techniques are viewed as pre-processing steps to transform raw vibration data into a representation that is rich with diagnostic information. Subsequently, HIs can be extracted from this transformed vibration data representation for the fault diagnosis task [5, 8]. Wang *et al.* [5] provide a thorough review on HIs from data-driven CM methodologies. Unsupervised

learning-based methodologies, discussed in Section 1.2.2.2, bear the closest similarity to signal-processing based works, whereby both approaches seek to obtain representations with enhanced diagnostic information.

Separation strategies are often used to attenuate the non-diagnostic signal information. Antoni [78] emphasises the importance of separation techniques to recover the first-order and the higher-order cyclostationary signal components, and Randall and Antoni [13] provide a concise overview of different separation strategies for bearing fault detection. Synchronous averaging [209] is a popular tool to estimate the synchronous mean of a vibration signal within a specified period, e.g., one gear rotation period, which can be used to separate the signal into synchronous and non-synchronous components [13, 81]. Angular resampling, i.e., order tracking, is often used to ensure that a set integer number of values exist per specified period [92]. The synchronous average is also a popular indicator for localised gear tooth damage if the specified averaging period corresponds with one rotation of the damaged gear [5].

Nonetheless, as discussed in Section 1.2.1.3, gearbox vibration signals are sensitive to time-varying OCs, and this can degrade the performance of time domain indicators [17]. Zimroz *et al.* [17] analysed the coefficients of a linear regression model used to describe the linear relationship between time-domain features and wind turbine power generation statistics. Ref. [17] found that the regression coefficients are informative diagnostic metrics for bearing damage under time-varying OCs. In the recent literature, sparsity measures have exhibited desirable characteristics for gearbox CM [210]. Antoni [190] popularised the use of kurtosis to quantify repetitive transients caused by localised faults in rotating machinery. Subsequently, sparsity measures such as the spectral kurtosis [190], the spectral negentropy [31, 211], the ratio of the L_2 and L_1 norm [212], the Gini Index (GI) [213], the Hoyer Measure (HM) [210], the Smoothness Index (SI) [214], and the Box-Cox (BC) measure [215] have been used for fault diagnostics. Wang *et al.* [211] discovered an important relationship between well-known sparsity measures exists via unique weighting functions applied to the normalised square envelope signal. Further applications of sparsity measures are discussed in Section 1.2.3.5.

1.2.3.2 Spectral analysis

Spectral analysis, i.e., Fourier analysis enabled by the Fourier transform, is an important tool for rotating machinery CM used to detect the presence gearbox fault frequencies within the spectrum or power spectrum of vibration signals [1]. This technique is useful to detect harmonic components due to periodicities within vibration data from rotating machinery

driven by underlying periodic mechanisms [75, 216, 217]. However, spectral analysis is limited by time-varying OCs and low-frequency deterministic vibration signal components [218]. Time-varying OCs cause spectral smearing if the Fourier transform is applied to a time domain signal, but this is readily accounted for by order tracking the vibration data to transform the signal from the time domain to the angle domain [88]. Applying the Fourier transform to an order-tracked signal produces the order spectrum, which is useful to characterise periodic signal components under time-varying OCs. Low-frequency deterministic components, e.g., gear mesh components, dominate the spectra of gearbox vibration signals due to the concentrated energy associated with repeated tooth impacts and load fluctuations during the gear teeth meshing process. Subsequently, alternative techniques such as envelope analysis [219], cepstrum analysis [86], time-frequency analysis [220] and cyclostationary analysis [20] have become more popular [1, 77]. Additionally, as identified in Antoni [78] for the constant OC scenario, spectral analysis, which is the frequency analysis of the vibration signal and no transformation thereof, is only useful for first-order cyclostationary signal components, e.g., gear mesh interactions, whereas higher-order cyclostationary components, e.g., wear or bearing damage, require more sophisticated analysis techniques such as cyclic spectral analysis [119].

1.2.3.3 Time/frequency analysis

Time-frequency analysis techniques offer the ability to overcome spectral analysis limitations induced by time-varying OCs using a decomposition that is sensitive to both the time domain and the frequency domain. Such decompositions focus on the concentration of signal energy along the frequency axis through time to provide localised spectral information within the global temporal structure of a vibration signal [221]. The STFT is a decomposition which yields localised time-frequency information, from which the spectrogram, i.e., the squared magnitude of the STFT, can be used to analyse the time-frequency relationship within gearbox vibration data [1, 222]. Wang and McFadden [220] applied the spectrogram to the synchronous average, or the residual thereof, to detect localised gear tooth damage. Unfortunately, the spectrogram suffers from a decomposition limitation induced by a localised time window size parameter which gives a time-frequency resolution trade-off [1]. This limitation led to interest in alternative decomposition techniques for non-stationary signals such as the Wigner-Ville distribution [41, 85, 223, 224], the wavelet transform and its variants [141, 143, 147, 216, 225–227], Empirical Mode Decomposition (EMD) [228–233], variational mode decomposition [234–237], and ensemble EMD [238–241].

1.2.3.4 Cyclostationary analysis

Cyclostationary analysis techniques are an important facet of the existing literature, concerned with analysing signals that are generated with some underlying periodic phenomena which induce periodic statistical characteristics. This categorisation is well-suited for vibration data obtained from rotating machinery [20, 78]. In the literature, a distinction between first and second-order cyclostationary signals is made as gearbox vibration signals often exhibit such cyclostationary characteristics, and the interpretation of these characteristics is useful for condition inference purposes [20, 242]. Antoni [78] details that gear and bearing damage are represented by different components within a vibration signal. Gears produce a first-order cyclostationary component periodic with the shaft speed, whereas bearing damage develops as a (pseudo) second-order cyclostationary component approximately periodic at a scalar multiple of the shaft speed [78, 92]. First-order cyclostationary techniques are predominantly concerned with estimating the deterministic, i.e., the periodic mean value, first-order moment of a signal. A popular first-order cyclostationary tool, under the assumption of cyclo-ergodicity, is the synchronous average as discussed in Section 1.2.3.1 [20]. Synchronous averaging requires adaptation to speed variations via order tracking to ensure consistency in the integer number of samples per segment and for phase angle alignment [81, 85, 243, 244]. However, Stander and Heyns [15] and Borghesani *et al.* [92] both identified the implications of the transfer function of the mechanical system as a source of distortion in synchronous averaging applications where order tracking is performed using a tachometer signal. Enhancement techniques, e.g., removing the average tooth meshing vibration [245], are sometimes used to improve the clarity of the synchronous average to detect gear damage [83].

Second-order cyclostationarity is a popular framework to extract information from gearbox vibration signals and has been found to be a powerful methodology to infer gearbox fault information. Second-order cyclostationary signals exhibit periodic second-order moment variations in time and such periodicity is captured in the auto-correlation function [78]. In the frequency domain, this periodic information is captured by the correlation between two frequency components, a cyclic frequency component that is the dual of the auto-correlation time-lag variable and a frequency component that is the dual of the time variable. The former represents information related to modulating components in the signal, while the latter represents the frequencies at which the modulations occur. The spectral correlation, otherwise known as the cyclic power spectrum, is a frequency domain descriptor for second-order

cyclostationary signals captured by a two-dimensional Fourier transform of the auto-correlation function or via the correlation of the signal spectrum with itself [13]. This transformation process yields a bi-spectral map that represents the whole structure of carriers and modulations in a signal [78, 242].

The spectral correlation has attracted much attention to how best to estimate and reduce the computational cost required to evaluate the spectral correlation, e.g., [242, 246–248], and its usefulness for gearbox fault diagnostics, e.g., [119, 249]. The Wigner-Ville spectrum represents the Fourier transform of the auto-correlation function with respect to the time-lag variable and has been actively used for gearbox CM tasks as discussed in Section 1.2.3.3. The spectral coherence represents a normalised version of the spectral correlation and contains a unique connection to the SES [247, 250]. Specifically, the SES represents the marginalisation of the spectral coherence over frequency, and marginalisation over the absolute spectral coherence yields the enhanced envelope spectrum [247].

Envelope analysis is another popular second-order cyclostationary technique. Envelope analysis refers to the calculation of the Fourier transform of the filtered signal envelope (or the square thereof, which is often preferred [251]), where the envelope is obtained through the Hilbert transform or via square rectifying for the purposes of identifying prominent modulating frequencies present in the signal, e.g., bearing fault frequencies [13, 39]. Envelope analysis often relies of signal enhancement strategies to enhance the potential diagnostic content available in the signal, e.g., band-pass filtering [39]. Enhancement strategies are further discussed in Section 1.2.3.5. Time-varying OCs disrupt the cyclostationary assumption and culminated in a angle-time cyclostationarity framework for cyclo-non-stationary signals [80]. Subsequently, recent work has explored the application of the order-frequency spectral correlation and the order-frequency spectral coherence [252] to handle variable OCs and has investigated the implications of time-varying OCs for envelope analysis tasks [253].

1.2.3.5 Signal enhancement strategies

Signal enhancement strategies are typically used to attenuate the prominence of deterministic components to improve the Signal-to-Noise Ratio (SNR) of the damaged component of interest, e.g., impulsive bearing damage [13, 253]. The simplest strategy, for bearing damage scenarios, is to high-pass or band-pass filter the signal in an appropriate frequency range to suppress the lower frequency deterministic components and capture the resonance band associated with the bearing fault information [51, 254]. Deterministic-random separation techniques, which

are a generalisation of the separation techniques discussed in Section 1.2.3.1, seek to separate first-order signal components from the potentially higher-order cyclostationary components to enhance the quality of the diagnostic information in the analysed signals. Both extracted components can then be further analysed to infer the instantaneous health state of an asset. For synchronous averaging, the synchronous component represents the deterministic part, while the non-synchronous component represents the random part [92]. Recent works have proposed methods to improve the synchronous average estimation process, e.g., the generalised synchronous average [255], suppress deterministic components, e.g., cepstral methods [256] or cepstral pre-whitening [257], using auto-regressive models, e.g., Ref. [152], or adaptive filters, e.g., Ref. [251]. A succinct overview and comparison of different deterministic-random separation techniques is discussed in the work of Randall *et al.* [243].

However, even after attenuating the deterministic data components, the fault component of interest may be masked or have a low SNR [13]. This is prominent in bearing fault applications, whereby the signal envelope and its spectrum are generally used to detect faults, but components not related to damage may dominate the spectra. Typically, a bandpass filtering operation is applied to enhance the fault covariates within vibration signals, but the filtering band is generally unknown a priori. Hence, methodologies to enhance fault characteristics have been actively researched and are discussed within two categories: *(i)* band selection approaches, and *(ii)* blind deconvolution approaches. The distinction between the two approaches is that the former typically uses a filter bank and requires end-user inspection to select the appropriate frequency band, while the latter attempts to resolve the appropriate band automatically. Band selection approaches are concerned with identifying specific frequency bands in which certain statistical properties of the vibration signal, e.g., non-Gaussianity, are prominent.

In the literature, band selection approaches can be split into two categories, namely blind and targeted methods. The distinction comes through the requirement for prior system information, e.g., characteristic fault frequencies, which targeted methods exploit [32]. Most approaches use filter banks obtained through the STFT, wavelet packet estimators, or parallel filtering operations with a binary tree or 1/3 binary tree structure to perform multiple band-pass filtering operations, and the most popular method is the fast kurtogram [30, 190, 258]. As a prelude to band selection approaches, the spectral kurtosis was developed as a measure to identify non-Gaussian signal components, which can be estimated using the STFT [30], and

is widely accepted by the academic community for its effectiveness to localise frequency bands with strong impulsive information [13]. The kurtogram was developed as an extension of the spectral kurtosis designed to address uncertainties related to the choice of STFT window length [30, 190] and effectively performs an exhaustive grid search over candidate parameters for the window length to obtain an optimal bandpass filter [13, 190]. Subsequently, the fast kurtogram was developed to reduce the computational complexity of the kurtogram by using a filter bank with a 1/3 binary tree structure [258]. Alternative approaches have been proposed in the literature aimed towards using alternative non-Gaussianity measures as a mechanism to detect impulsiveness and extract repetitive transients. Examples of blind methods include the protogram [259], the sparsogram [260], the infogram [31], the autogram [261], and the IESFOgram [262] to name a few. Examples of targeted methods include the ICS2gram [21, 32], the RCCgram [32, 263], the log-cycligram [32], and the IBFI$_\alpha$gram [33].

Alternatively, blind deconvolution approaches aim to enhance the fault covariates within a signal by designing a Finite Impulse Response (FIR) filter whose coefficients maximise a proxy measure, e.g., non-Gaussianity, to improve the SNR of the faulty component [264]. As with the band selection approaches, the blind and targeted methods can also be identified. Blind methods include Minimum Entropy Deconvolution (MED) [265], optimal MED [266], multipoint optimal MED [267], and maximum correlated kurtosis deconvolution [267]. Endo and Randall [268] used an auto-regressive model to attenuate deterministic gear components and then MED to enhance the bearing fault condition. Sawalhi *et al.* [50] followed a similar approach to Ref. [268] but used the kurtogram to further enhance the fault condition. While not strictly considered a blind deconvolution approach, the spectral kurtosis can operate as a blind detection filter, and has been applied in both gear and bearing fault applications [269, 270].

Sparsity measures have received attention from the academic community in recent years, and different measures have been used for the filtering problem, e.g., the l_0 norm [271], the l_1/l_2 norm [272], the generalised l_p/l_q norm [273], the GI sparsity measure [274], and the BC sparsity measure [275] to name a few. Peeters *et al.* [276] used sparsity measures to enhance the sparsity of the SES. Targeted approaches include the cyclostationary blind deconvolution approach [277], which use an indicator of second-order cyclostationarity [21] for blind deconvolution. Kestel *et al.* [278] proposed the use of sparsity measures applied in narrow bandwidths of the SES between shaft speed harmonics as a targeted approach. One

critical parameter for such approaches is the choice of window length, which is still being actively investigated in the literature [264, 277].

1.3 Scope of research

The scope of research is integral to the cohesion of this thesis and provides a more extensive perspective of the fundamental components on a broader level. In Section 1.3.1, the directions followed in this research are discussed. In Section 1.4, the structure of this thesis is presented considering the research directions followed.

1.3.1 Directions of research

It is evident for the literature study conducted in Section 1.2 that vibration-based gearbox CM is an active and diverse field with various avenues of academic interest. Learning-based and signal processing-based data-driven methodologies are aimed to address the same open issues prevalent in the vibration-based CM field. However, these methodologies approach such issues via fundamentally different avenues. Learning-based methodologies are often driven by supervised objectives to perform fault diagnosis, i.e., health state recognition and fault identification impelled via classification losses [7, 94, 104]. However, the reliance on historical failure information hinders the practicality of such methods. Signal processing methodologies use powerful techniques to extract and enhance the presence of fault signatures in raw vibration data [13, 78]. However, such methodologies require extensive and profound domain understanding of the failure mechanisms present in the data and expert knowledge is required to extract informative fault features, e.g., bearing fault resonance modulation phenomena as a precursor for envelope analysis [51, 219].

Unsupervised learning-based methodologies, e.g., LVMs, can effectively bypass both the requirement for historical failure information. LVMs circumvent the requirement for expert knowledge by learning a probabilistic representation of the vibration data in isolation to produce health indicators. Conceptually, this characteristic renders this category of unsupervised models appealing as they represent an agnostic framework for fault diagnostics with potential to many diverse CM applications [28]. Unfortunately, their potential is not yet fully realised in vibration-based CM as they are often treated as black-box methods limited to fault detection problems using data likelihood measures, e.g., reconstruction errors, as health indicators. Furthermore, conventional applications of LVMs do not preserve time which limits their application to anomaly detection tasks. Temporal preservation is an essential approach

that advocates for the preservation of time structure within vibration data during the model evaluation step to harness the full potential of LVMs [164]. This recently developed approach enables LVMs to expand into the full fault diagnosis problem. Nonetheless, preserving temporal structure is only the first step towards latent interpretation and requires that important aspects of LVMs, e.g., the latent manifold and methods employed for its construction, be revisited, studied, and further advanced for the fault diagnosis task. This is necessary to elevate the efficacy of LVMs for the fault diagnosis task and to develop approaches that advance their applicability as a data-driven learning-based methodology for vibration-based CM.

In this thesis, temporal-preserving LVMs are investigated as an effective learning-based tool to perform gearbox fault diagnostics. The principal mode of focus resides in how to suitably use LVMs with temporal preservation to enhance the knowledge gained in a fault diagnostics context. The investigations performed in this thesis are aimed towards how temporal-preserving LVM-based CM methodologies can be improved. Thus, the central theme running throughout this thesis revolves around the application of temporal-preserving LVMs as a powerful tool for vibration-based CM and the advancement of LVM-based methodologies for the fault diagnosis task. To this end, three research directions are identified to study and advance the application of temporal-preserving LVMs. A preliminary overview of the temporal-preserving LVM application setting is detailed in Figure 1.1 and the fundamental research directions investigated in this thesis are shown. This thesis addresses three primary research directions: the first pertains to the extraction of HIs from the latent space of offline reconstruction-focused temporally-preserving LVMs, the second explores the implications of offline and online LVM training approaches, focusing on online reconstruction-focused and interpretation-focused methods, and the third delves into the extraction of latent HIs specifically within the context of both offline and online LVM methods. These directions will now be examined in more detail.

Direction one: Latent manifold diagnostics

While temporal preservation unlocks the potential of the latent manifold, its utilisation for LVMs requires that the HIs used for LVM-based CM be revisited. In the first research direction, the objective is to study how fault signatures present in the data and latent space of offline temporal-preserving LVMs can be extracted for condition inference purposes. Informative temporal-preserving HIs are crucial to the fault diagnosis task. Conventionally, the data space is a primary source of fault diagnostic information, while the latent manifold is secondary and

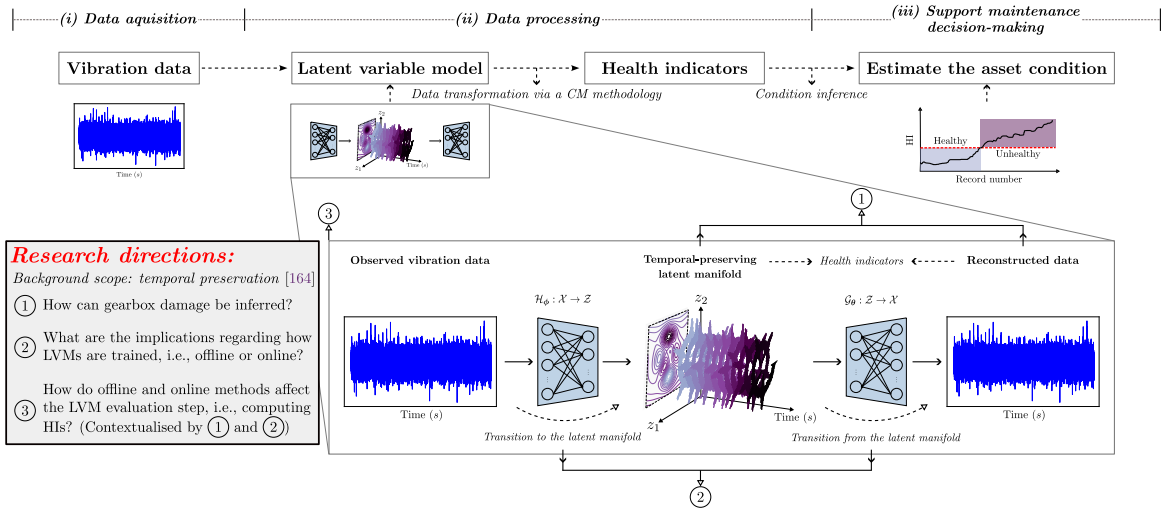


Figure 1.1. An overview is given to the LVM application process used in this thesis in a fault diagnostics-driven CM context. The fundamental questions behind the research performed in this thesis are: (1) how can the data and latent space of temporal-preserving LVMs be used to perform effective fault diagnostics, (2) what are the implications of how LVMs are trained in a fault diagnosis context, and (3) how should HIs be computed when using different LVM training approaches and what are the implications, advantages, and disadvantages that exist in the considered LVM-based CM pipeline. Question (1) is defined in relation to the data and latent space of LVMs, Question (2) is defined in relation to the LVM mapping functions from the data space to the latent space and vice versa, and Question (3) is defined in relation to the whole LVM pipeline, i.e., both the LVM mapping functions and the LVM data and latent space, and represents a synthesis of Questions (1) and (2).

is often under-utilised. The CM implication is that not only is condition inference limited to data space indicators, but the intrinsic features assumed to give rise to the observed data, i.e., the latent variables, are never scrutinised as an informative diagnostic representation.

Hence, a modular ensemble-based Latent Health Indicator (LHI) framework is proposed and presented in this thesis, which exploits the diagnostic potential of the latent manifold to determine how gearbox damage can be inferred when using offline reconstruction-focused temporal-preserving LVMs. The framework defines a set of conceptual classes that can be used to obtain informative LHIs rich with diagnostic information for the LVM-based fault diagnosis task. The framework proposed in this thesis is used as a first step towards LVM interpretability, and the objective is to demonstrate the ability to perform effective gearbox fault diagnostics, i.e., early fault detection and isolation with distinct and comprehensible asset degradation trends, using the latent manifold under constant and time-varying OCs.

Direction two: Training strategies

In traditional applications of LVMs, offline training is used to develop and employ LVMs for the CM task. Training pertains to the parameters of the functions used to map to and from the latent manifold as detailed in Figure 1.1. However, it is possible that access to historical vibration samples that characterise the healthy data distribution are not available a priori. This impedes the application of LVM-based CM methods that require historical data for training. Data-driven methods can be applied in an offline and online training setting, and LVMs are no different. While traditional applications use offline training, online training offers the capability to dynamically adapt the latent manifold, allowing it to continuously refine the latent representation of essential intrinsic information. This additional latent flexibility may lead to better diagnostic information extraction for the CM problem.

Hence, the second research direction aims to explore the implications of employing offline and online training approaches in the context of temporal-preserving LVMs for fault diagnostics. In the dynamic latent manifold scenario induced by online training, two main online LVM methods are considered to evaluate their implications in an online fault diagnostics setting, namely *(i)* reconstruction-focused methods and *(ii)* interpretation-focused methods. LVM methods from the two training approaches are compared in this thesis to understand the implications of offline and online training in a fault diagnostics context. The objective herein is to showcase the applicability of online temporal-preserving LVMs in gearbox fault diagnostics, for both constant and time-varying OCs, while emphasising the complementarity of the online methods used in a fault diagnostics context.

Direction three: Latent indicator categories

In the third research direction, the evaluation step of offline and online temporal-preserving LVMs is critically analysed. The use of online training for temporal-preserving LVMs posits the existence of online latent indicators, i.e., latent indicators that exhibit online characteristics, catering specifically to methods developed using online training. The indicator framework proposed to address the first research direction, primarily defined in the context of offline reconstruction-focused temporal-preserving LVMs, is critically assessed to identify the implicit assumptions made by offline indicators, which potentially reduce their generality to online temporal-preserving LVMs. This represents a combination of the first two research directions, whereby the ability to use the latent manifold for condition inference is studied in-depth for offline and online temporal-preserving LVMs.

Hence, two categories for LHIs are identified using the framework from the first research direction, namely *(i)* offline and *(ii)* online latent indicators. A representative set of online LHIs is proposed in this thesis for online temporal-preserving LVMs, which use implicit assumptions identified in the offline LHIs from the first research direction as constraints. The two LHI categories are then critically compared using both online and offline methods. The purpose of the identified indicator categories is to critically study the implications of latent indicators for offline and online methods. The end goal is to assess if diversity is a pertinent factor to motivate a heterogenous set of latent indicators to maximise on the benefits of using offline and online temporal-preserving LVMs for gearbox fault diagnostics. Diversity enables a holistic fault diagnosis strategy by advocating for a proactive temporal-preserving LVM-based CM approach using multiple latent health indicators for methods from both training domains to support and guide maintenance decision-making.

1.4 Layout of thesis

The research directions considered in this thesis are presented in sequential order. An overview of the research directions and their respective interactions with the chapters in this thesis is presented in Figure 1.2. The abbreviations used in this thesis are presented on page ix. A consolidated list of mathematical symbols is presented on page xi.

In Chapter 2, a latent indicator framework for offline temporal-preserving LVMs is presented that identifies five conceptual classes of LHIs to capture various perspectives of the LVM latent manifold. The indicators obtained from the proposed conceptual classes are investigated using an experimental bearing dataset acquired under constant OCs and using an experimental gearbox dataset acquired under time-varying OCs. The results indicate that the latent manifold is an expressive representation and using an ensemble-based condition inference procedure is beneficial to the fault diagnosis task.

In Chapter 3, the investigation into offline and online training approaches and the implications of two online LVM methods are presented for the vibration-based CM task. The investigation considers a simple set of latent indicators discussed in Chapter 2 to investigate the ability of the considered offline and online methods to perform effective fault diagnostics. The investigation is performed on numerical and experimental bearing failure datasets developed under constant OCs, and on an experimental gearbox dataset developed under time-varying OCs. The results of this investigation indicate that both offline and online LVM methods are

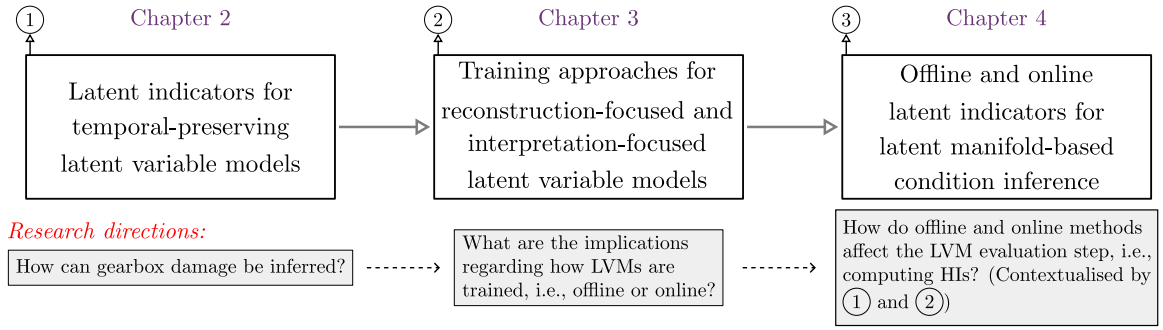


Figure 1.2. An overview of the research directions put forward in Figure 1.1 and their relationship to the layout of this thesis.

useful for the fault diagnosis task. Furthermore, the considered online methods are observed to be complementary in a fault diagnostics setting.

In Chapter 4, latent indicators from the identified offline and online indicator categories are applied to the LVM-based fault diagnosis problem to conduct a critical comparison into the choice of LVM training approach and the methods used therein. The online indicators proposed for online LVMs are presented and compared to the offline indicators presented in Chapter 2. A comparative investigation is conducted in a fault diagnosis context using the experimental datasets considered in Chapters 2 and 3. The results from this investigation highlight the diversity of the considered indicator categories as a function of their application setting and the asset of interest being investigated. Furthermore, access to a heterogeneous set of latent indicators from both indicator categories provide a robust outcome to employ LVMs for fault diagnostics.

Finally, the investigations performed in this thesis are concluded in Chapter 5. An overview of the research directions and the contributions of this thesis due to the research directions are summarised, and recommendations are provided for future work.

Three supplementary appendices are included to provide additional information for this thesis. Information related to the LVM parameters and approaches to compute different signal-to-noise ratio diagnostic metrics are presented in Appendix A. Appendix B presents supplementary information related to the parameters of the phenomenological gearbox model dataset used in the numerical investigation performed in Chapter 3. Appendix C presents a spectral regularisation framework for LVMs used in both Chapter 3 and Chapter 4 for interpretation-focused online LVMs that are sensitive to the issue of source duplication.

Chapter 2

Latent indicators for temporal-preserving latent variable models

2.1 Introduction

Unsupervised learning-based methodologies offer advantages over their supervised counterparts as they bypass the requirement for labelled fault data and are primarily concerned with capturing the distribution of the observed data [6, 94]. Latent variable models in a temporal-preserving framework [164] are a compelling class of unsupervised learning-based methods and provide a latent space and a reconstructed data space to compute health indicators for the condition inference task. However, typical latent variable model applications only rely on the data space as a representative characterisation of the instantaneous asset health state and tend to disregard the latent space in favour of metrics that measure the likelihood of the observed data [28, 163]. This impedes the effectiveness of such models, as inference is limited to and highly dependent on the performance of a sparse set of indicators. The first research direction is concerned with addressing this health indicator limitation by revisiting and advancing the approaches used to perform condition inference by exploiting the temporal-preserving latent manifold.

Therefore, a modular latent indicator framework is proposed in this chapter for models developed using offline training to enable a rich latent variable model-based condition inference procedure. The proposed framework describes different latent manifold perspectives to perform fault diagnostics. This framework identifies five conceptual classes to characterise the latent manifold and the potential avenues in which faulty signatures may manifest therein. Subsequently, temporal-preserving latent health indicators are extracted from the proposed conceptual classes and their effectiveness for the fault diagnosis task is investigated.

The purpose of this investigation is not to propose or identify an optimal latent health indicator. Rather, the objective is to demonstrate that incorporating and using multiple indicators enables an ensemble-based inference procedure, and having direct access to multiple

latent metrics is useful as it allows for interchangeability for any given asset. Thus, offline temporal-preserving latent variable models, as a methodology for the gearbox fault diagnosis task, are enhanced by the proposed latent indicator framework, allowing for improved latent condition inference.

The structure of this chapter is as follows: First, the use of temporal preservation as an approach to enhance the interpretability of the latent manifold is discussed in Section 2.2. Then, the latent indicator framework is presented in Section 2.3. Subsequently, the performance of latent health indicators derived from the proposed conceptual classes is investigated in the fault diagnostics context in Section 2.4. Finally, the chapter is concluded in Section 2.5.

2.2 The implications of temporal preservation

To enhance the condition inference process when using Latent Variable Models (LVMs), the temporal preservation approach proposed by Ref. [164] is used. This approach ensures that the temporal characteristics of the vibration data are preserved in the health indicators obtained from the models. The temporal preservation approach advocates that the sensor time, i.e., the time within a measured signal, must be preserved for informative LVM-based condition inference. Temporal preservation is the mechanism from which LVMs become applicable under constant and time-varying operating conditions. Latent variable model facilitate a direct transformation of the observed vibration data into indicators, and temporal preservation provides temporal structure therein. Without temporal structure, extraction of fault feature information is limited as sensor time is effectively entangled within the latent manifold [164].

The fundamentals of temporal preservation are demonstrated in Figure 2.1 for any methodology that performs a transformation to vibration data. If the output of the input data transformation contains temporal structure, i.e., sensor time is preserved in the output of the transformation, then the methodology, by design, performs temporal preservation. However, if the output is a single scalar or, equivalently, a collection of temporally non-preserving scalars, then the methodology removes the temporal structure present in the input data and is temporally non-preserving.

Temporal preservation is the first step to latent manifold interpretability in condition monitoring applications, and the importance of using temporal preservation for the latent manifold is revealed in Figure 2.2. Without sensor time, as demonstrated through Figure 2.2(a), there is no noticeable distinction between two faults, one localised and one distributed.

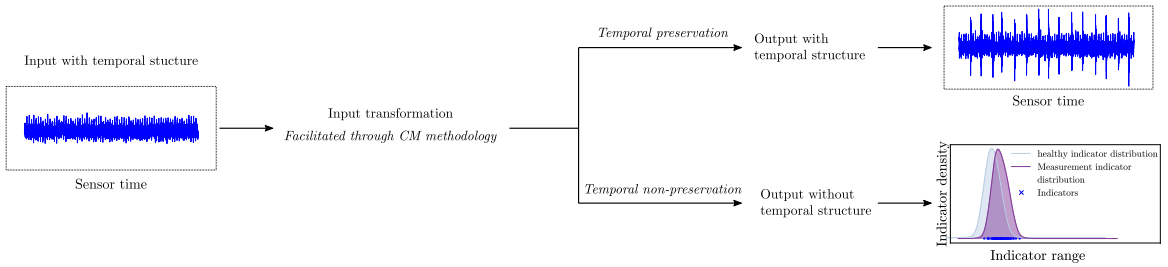
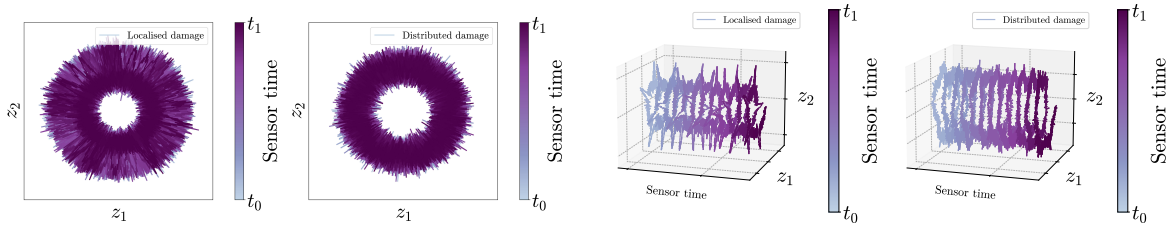


Figure 2.1. The identification of vibration-based condition monitoring methodologies as temporally preserving or temporally non-preserving. The distinction, between temporal preservation and temporal non-preservation, irrespective of the methodology used, is based off the outputs of the methodology.

By preserving sensor time, as presented in Figure 2.2(b), the distinctions between the two fault representations are clear and unique. The temporal characteristics of the manifold representation of the observed vibration data may then be exploited for condition inference.



(a) Manifold without temporal preservation. **(b)** Manifold with temporal preservation.

Figure 2.2. A pictorial demonstration of the latent manifold for two types of vibration data faults with and without sensor time preservation. In (a), the lack of temporal structure corrupts the inference process and there is little distinction between the fault types. In (b), the addition of temporal structure uncovers the clear fault distinctions in the manifold.

To provide context to the investigation conducted in this chapter in relation to the broader field of vibration-based condition monitoring methodologies, Figure 2.3 is presented using the notion of temporal preservation and non-preservation. In Figure 2.3, signal processing methods that are temporally non-preserving refer to methods that collapse sensor time and compute vibration data features at a record time level, e.g., Večeř *et al.* [19]. In contrast, signal processing methods with temporal preservation typically produce a transformed time-series signal with the fault covariates enhanced, e.g., Sawalhi *et al.* [50]. In the learning-based domain, temporally non-preserving methods produce indicator signals that are devoid of any temporal content, e.g., Booyse *et al.* [28], whereas methods that are temporally preserving produce indicator signals with temporal structure, e.g., Ref. [164]. Learning-based approaches

are distinct as signal processing approaches rely on expert knowledge for signal transformation design, whereas learning-based approaches use data-driven models to facilitate the signal transformation. However, this does not imply that the two approaches are mutually exclusive, e.g., Wang *et al.* [279].

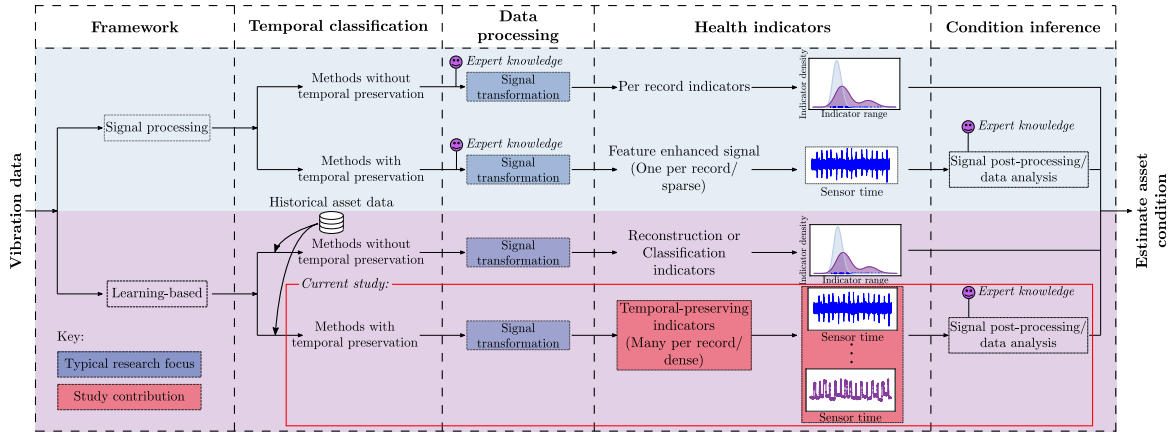


Figure 2.3. The condition monitoring process for signal processing-based and learning-based approaches. The fundamental difference between techniques from the two approaches is driven through the notion of temporal preservation. This investigation uses temporally preserving LVMs and is focused on enriching the condition inference stage through an ensemble of latent health indicators.

2.3 Latent manifold analysis

In this thesis, LVMs are used as a representation transformer that take vibration data and transform it to a set of Health Indicators (HIs) and Latent Health Indicators (LHIs). These HIs and LHIs serve as indicators from which the asset state can be inferred. The LVMs used in this chapter are developed on a reference set of vibration data, and this data is assumed as representative of a healthy asset state. The objective of this chapter is to improve the utilisation of the latent manifold for Condition Monitoring (CM) purposes by developing LHI conceptual classes. This directly adheres to the ambition of the first research direction. These developed LHI classes introduce and capture different manifold perspectives, where each perspective can produce LHIs for condition inference. To detail the methods used to detect damage in the latent manifold, the concepts behind LVMs must first be introduced.

To highlight the usefulness and responsiveness of the latent manifold to anomalous fault data, two reconstruction-focused LVM formulations are considered. The LVM formulations of interest are *(i)* Probabilistic Principal Component Analysis (PPCA) [280], and *(ii)* Variational

Auto-Encoders (VAEs) [165]. The choice of formulation is driven by methodologies that make explicit and tractable model density assumptions, as the focus of this investigation is on the latent manifold and how damage manifests therein, and thus these methods reflect different manifold considerations under a shared model density assumption. Both the PPCA and VAE models are explicit density estimation methods that use Gaussian distributions [101, 165] and are considered to be reconstruction-focused. The latent manifold in PPCA has a linear subspace, while VAEs utilise nonlinear parametric functions, often referred to as neural networks, to obtain a subspace that may be embedded nonlinearly within the observed data.

2.3.1 Latent variable models

LVMs are the foundation of unsupervised learning-based techniques [165]. LVMs are generative modelling techniques and represent a class of statistical models concerned with capturing the distribution over some observed variable set $\mathbf{x}$, where $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^D$ is a D-dimensional observed random variable [101, 161]. As the data distribution is assumed to be unknown, LVMs assume that a set of unobserved random variables $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^d$, where $d \leq D$, has a hidden relationship with the observed variables $\mathbf{x}$, and they describe the intrinsic nature of the data [101, 161]. The latent variable set is referred to as the latent manifold [281]. This assumption gives rise to the generative process that describes how $\mathbf{z}$ forms part of the model [161]. In the CM setting, LVMs learn a probabilistic representation of an asset to detect deviations from a reference asset state [28]. Under the assumption that the reference state captures an asset in a healthy condition and under the assumption of no access to labelled vibration data [94], the detected deviations are attributed to anomalies in the asset due to a change in asset condition.

A LVM defines a joint probability distribution

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x}|\mathbf{z})p(\mathbf{z}), \tag{2.1}$$

where $p(\mathbf{z})$ is the latent prior distribution and $p(\mathbf{x}|\mathbf{z})$ is the generative conditional distribution. Note that the factorisation of the joint distribution in Equation (2.1) defines the generative process of LVMs [161]. The marginal distribution $p(\mathbf{x})$, i.e., the model evidence, over the observed variables is given by

$$p(\mathbf{x}) = \int_{\mathbf{z}} p(\mathbf{x}|\mathbf{z})p(\mathbf{z})d\mathbf{z}, \tag{2.2}$$

which offers a highly flexible model if $\mathbf{z}$ is continuous [282]. Inference is an important aspect of LVMs and the aforementioned generative process is only concerned with sample generation. To infer the latent variables for a data sample $\mathbf{x}$, Bayes' theorem can be used to compute the

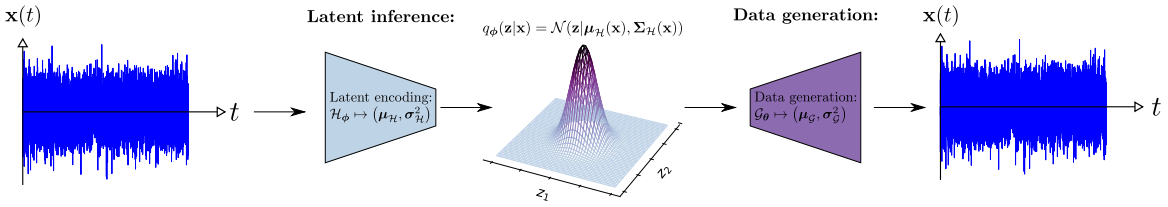


Figure 2.4. A visual LVM illustration of how the posterior and generative distribution interact. This approach is simple, does not take any vibration data pre-processing into account, and primarily serves as visual demonstration of how observed data interacts with different aspects of the LVM.

latent posterior distribution $p(\mathbf{z}|\mathbf{x})$ through

$$p(\mathbf{z}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{x})}. \tag{2.3}$$

LVMs, in practice, utilise linear or nonlinear parametric functions to parametrise the distributions of interest, namely the conditional, posterior, and prior distributions [161]. Typically the prior distribution is assumed to be a simple, fixed distribution and the parametric distributions become $p(\mathbf{x}|\mathbf{z})$ and $p(\mathbf{z}|\mathbf{x})$. The parametric functions of the aforementioned distributions are deterministic and commonly denoted as

$$\mathcal{G}_{\boldsymbol{\theta}} : \mathcal{Z} \mapsto \mathcal{X}, \tag{2.4}$$

$$\mathcal{H}_{\boldsymbol{\phi}} : \mathcal{X} \mapsto \mathcal{Z}, \tag{2.5}$$

where the transition functions $\mathcal{G}$ and $\mathcal{H}$ have parameters $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ respectively. These functions are used to parametrise the generative distribution $p(\mathbf{x}|\mathbf{z})$ and posterior distribution $p(\mathbf{z}|\mathbf{x})$. In Equation (2.4) an example of the transition functions is shown. There are two main types of LVMs, namely *(i)* explicit (likelihood-based) models, and *(ii)* implicit models [283]. Explicit models perform maximum likelihood estimation using the marginal likelihood function to optimise the unknown parameters of the transition functions $\mathcal{G}$ and $\mathcal{H}$. Models in this category require explicit distribution parametrisation. The choice of transition function form is critical for explicit models, as linear functions may allow for exact posterior computation, whereas nonlinear functions, in general, require approximate inference techniques to optimise the model parameters [101, 161, 162]. Typical examples of explicit models include PPCA [280], VAEs [165], normalising flows [284, 285], and energy-based models [286]. Both the PPCA and VAE LVM formulations maximise the marginal log-likelihood lower bound and consist of non-invertible transition functions, while normalising flows maximise the exact log-likelihood and use network architecture constraints to produce bijective transition functions.

Implicit models, alternatively, do not parametrise the distributions of interest explicitly, and represent the data distribution of interest as a model of a latent sampling process [168]. This formulation requires the use of the density estimation by comparison techniques, such as the density ratio $\frac{p(\mathbf{x})}{q(\mathbf{x})}$, to effectively capture $p(\mathbf{x})$, where $q(\mathbf{x})$ is the distribution captured by the implicit LVM [168]. The most prominent implicit model is a Generative Adversarial Network (GAN) [97]. The caveat here is that most implicit models focus on sample generation through $\mathcal{G}_{\boldsymbol{\theta}}$, and neglect the inference stage of LVMs. Bond-Taylor *et al.* [287] provide a succinct review on deep generative modelling techniques.

2.3.2 Explicit density estimation

Explicit LVMs are integral to this chapter to ensure that the latent manifold is accessible. The generative conditional and latent prior distributions are assumed to be Gaussian distributions of the form

$$p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_{\mathcal{G}}(\mathbf{z}), \boldsymbol{\Sigma}_{\mathcal{G}}(\mathbf{z})), \tag{2.6}$$

where the covariance of the generative conditional distribution is given by

$$\boldsymbol{\Sigma}_{\mathcal{G}}(\mathbf{z}) = \operatorname{diag}(\boldsymbol{\sigma}^2_{\mathcal{G}}(\mathbf{z})), \tag{2.7}$$

where $\operatorname{diag}(\cdot)$ is the vector-to-matrix diag operator. The prior is a unit isotropic Gaussian of the form

$$p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I}). \tag{2.8}$$

Finally, the generative transition function is expressed as

$$\left(\boldsymbol{\mu}_{\mathcal{G}}(\mathbf{z}), \boldsymbol{\sigma}^2_{\mathcal{G}}(\mathbf{z})\right) = \mathcal{G}_{\boldsymbol{\theta}}(\mathbf{z}). \tag{2.9}$$

Using the parametrised distributions, maximum likelihood of the marginal distribution $p(\mathbf{x})$ can be used to optimise the parameters of $\mathcal{G}_{\boldsymbol{\theta}}$. However, depending on whether $\mathcal{G}_{\boldsymbol{\theta}}$ is linear or nonlinear, the marginal distrubution may be intractable as it is a realisation of all possible values of $\mathbf{z}$, which may be overtly complicated to compute and possibly mathematically infeasible. PPCA is an LVM with a tractable solution, and thus the latent posterior distribution can be tractably computed [280]. The latent posterior distribution in PPCA is also a Gaussian distribution. However, VAEs commonly use neural networks for increased generative model flexibility. This decision causes Equation (2.2) to be intractable, and thus the posterior distribution cannot be determined [165, 282]. To overcome this issue, a parametric inference model $q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})$ is introduced to approximate $p(\mathbf{z}|\mathbf{x})$ and a lower bound on the marginal distribution, commonly referred to as the Evidence Lower Bound (ELBO) is maximised to

find the optimal model parameters $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ [165]. The ELBO can be found through direct expansion of Equation (2.2) or through variational inference techniques [162, 282].

Regardless of the tractability of the process to obtain the posterior distribution, both the aforementioned LVMs use a latent encoding transition function $\mathcal{H}_{\boldsymbol{\phi}}$ to parametrise a Gaussian posterior distribution

$$q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x}), \boldsymbol{\Sigma}_{\mathcal{H}}(\mathbf{x})), \tag{2.10}$$

where the latent covariance $\boldsymbol{\Sigma}_{\mathcal{H}}$ is given by

$$\boldsymbol{\Sigma}_{\mathcal{H}}(\mathbf{x}) = \text{diag}(\boldsymbol{\sigma}^2_{\mathcal{H}}(\mathbf{x})). \tag{2.11}$$

Finally, the latent transition function is expressed as

$$\left(\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x}), \boldsymbol{\sigma}^2_{\mathcal{H}}(\mathbf{x})\right) = \mathcal{H}_{\boldsymbol{\phi}}(\mathbf{x}). \tag{2.12}$$

PPCA and VAEs both operate on the same principle, whereby a posterior distribution is inferred from the observed data, and a sample can be generated by sampling the posterior $\mathbf{z} \sim q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})$ and then sampling the generative distribution $\mathbf{x} \sim p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z})$. However, the posterior covariance for a PPCA model is not a function of the input data, i.e., $\boldsymbol{\sigma}^2_{\mathcal{H}}$ is constant, while VAE models allow for this flexibility [165, 280].

2.3.3 Data pre-processing

To use LVMs on asset vibration data, the data must be pre-processed to ensure that it is practical for computational usage. In this work, the dimensionality of the i^{th} vibration signal $x_i[n]$ is reduced using a segmentation approach, where $n \subseteq \mathbb{Z}^+ = \{1, 2, \ldots, L_{signal}\}$ is the discrete time variable representative of sensor time, i refers to record time, and L_{signal} is the dimensionality of $\mathbf{x}_i \in \mathbb{R}^{L_{signal}}$ [28, 164]. The signal is transformed into a Hankel matrix $\overline{\mathbf{X}}_i \in \mathbb{R}^{L_T \times L_w}$ [288] by extracting signal segments from the raw vibration signals. This extraction process for the i^{th} vibration signal is given by

$$\overline{\mathbf{X}}_i = \begin{bmatrix} x_i[1] & x_i[2] & \cdots & x_i[L_w] \\ x_i[L_{sft}] & x_i[L_{sft}+1] & \cdots & x_i[L_{sft}+L_w] \\ x_i[2 \cdot L_{sft}] & x_i[2 \cdot L_{sft}+1] & \cdots & x_i[2 \cdot L_{sft}+L_w] \\ \vdots & \vdots & \ddots & \vdots \\ x_i[L_{sft} \cdot (L_T-1)] & x_i[L_{sft} \cdot (L_T-1)+1] & \cdots & x_i[L_{sft} \cdot (L_T-1)+L_w] \end{bmatrix} \tag{2.13}$$

where L_w is the window length, L_{sft} is the shift parameter, and $L_T = \lfloor \frac{L_{signal}-L_w}{L_{sft}} \rfloor + 1$ represents the number of rows in $\overline{\mathbf{X}}_i$. Note that as L_{signal} is typically known a priori via

acquisition procedure used to acquire the vibration data, a primary dependency is induced in L_T on L_w and L_{sft}. These parameters are reported in Appendix A for the experimental investigations performed in this thesis. Note that the rows of Equation (2.13) now represent the feature vectors that are fed into the LVMs. Given a historical, assumed healthy, reference dataset of vibration signals, a reference data matrix $\overline{\mathbf{X}}_{ref} \in \mathbb{R}^{(N \cdot L_T) \times L_w}$ is then developed using N reference vibration signals through

$$\overline{\mathbf{X}}_{ref} = \begin{bmatrix} \overline{\mathbf{X}}_1 \\ \vdots \\ \overline{\mathbf{X}}_N \end{bmatrix}. \tag{2.14}$$

The reference data matrix $\overline{\mathbf{X}}_{ref}$ is then partitioned into a training and validation set that is used for LVM training. In this study, the shift parameter is selected as $L_{sft} = 0.5L_w$ and $\overline{\mathbf{X}}_{ref}$ is split using an $80\% - 20\%$ training-validation split process. A standardisation procedure is applied to all the Hankel matrix representations using the reference Hankel matrix $\overline{\mathbf{X}}_{ref}$ and is given by

$$\mathbf{X}_i = \left(\overline{\mathbf{X}}_i - \mathbf{1}\boldsymbol{\mu}_{ref}^T\right) \text{diag}\left(\boldsymbol{\sigma}_{ref}^{-1}\right), \tag{2.15}$$

which may be given in index notation as

$$X_{i,jl} = \frac{\overline{X}_{i,jl} - \mu_{ref,l}}{\sigma_{ref,l}}, \tag{2.16}$$

where $\boldsymbol{\mu}_{ref} \in \mathbb{R}^{L_w}$ and $\boldsymbol{\sigma}_{ref} \in \mathbb{R}^{L_w}$ are column vectors of the feature-wise means and standard deviations respectively of the partition of $\overline{\mathbf{X}}_{ref}$ used for LVM training, $\mathbf{1} \in \mathbb{R}^{L_w}$ is a constant vector with elements 1, and $\boldsymbol{\sigma}_{ref}^{-1} = [\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_{L_w}}]^T \in \mathbb{R}^{L_w}$. For model evaluation, the temporal preservation approach is used, which effectively re-processes the available data using Equation (2.13) with $L_{sft} = 1$ [164] prior to standardisation step achieved using Equation (2.15).

2.3.4 Latent manifold indicators

In this work, the latent manifold is used for condition inference. To enable LVMs to be applicable to all facets of fault diagnostics, the benefits of the temporal preservation approach proposed in Ref. [164] are incorporated. To detect deviations from the learnt healthy state in the latent space, the following five conceptual LHI classes are proposed, namely *(i)* likelihood-based, *(ii)* discriminative, *(iii)* statistical distance, *(iv)* prototypical approaches, and *(v)* multivariate signal processing approaches. These five classes capture different manifold perspectives, provide LHIs for condition inference, and are all methods that detect deviations

in the latent manifold. The five classes are discussed in turn. Examples of indicators from each class are given in Table 2.1.

1. The likelihood-based class refers to techniques that model the latent space through a reference model $p_{ref}(\mathbf{z})$. For latent anomaly detection, the reference model negative log-likelihood is used to provide an indication of whether new data is from the reference latent manifold.

2. The discriminative class of approaches refers to techniques that adopt a discriminative approach to capturing the reference latent manifold. The distribution of interest is the posterior distribution $p(\mathbf{y}|\mathbf{z})$, where $\mathbf{y}$ is a discrete, class variable. However, the labels given to the observed reference latent data only represent one asset state, which is healthy.

3. Statistical distance-based LHIs are obtained whereby each input $\mathbf{x}$ passed through the transition function $\mathcal{H}_{\boldsymbol{\phi}}$ supplies an output that parametrises a latent posterior distribution $q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x}), \boldsymbol{\Sigma}_{\mathcal{H}}(\mathbf{x}))$. The properties of this posterior distribution are then interrogated to determine whether it exhibits sensitivity to faulty data. There are two avenues of analysis within this approach, namely *(a)* dissimilarity-based comparisons, and *(b)* distribution similarity comparisons. Each of these avenues will now be discussed.

   (a) The dissimilarity-based sub-class inspects how the mean $\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x})$ of the posterior distribution changes. This comparison can occur in two ways, one through changes in the mean with respect to the latent prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I})$, or through changes in the mean within a localised time-frame.The latter approach is analogous to a first-order differences approach, with a step size Δt given by $\Delta t = \frac{1}{F_s}$, where F_s is the signal sampling frequency.

   (b) In the distribution similarity sub-class, metrics that measure the similarity between two distributions are used. As with the mean-based sub-class, the distribution comparison may be between the sample posterior $q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})$ and the latent prior $p(\mathbf{z})$, or within a localised time-frame. For the latter, a comparison between two posterior distributions $q_{\boldsymbol{\phi}}(\mathbf{z}_t|\mathbf{x}_t)$ and $q_{\boldsymbol{\phi}}(\mathbf{z}_{t+1}|\mathbf{x}_{t+1})$ is performed. Note that t indicates the temporally-preserved latent or data representation at any time-point $t \in [0, L_T - 1]$, where $L_T = L_{signal} - L_w$ when temporal preservation is applied.

4. The prototypical class refers to methods that assume that there are a finite number

of prototypical elements in the latent space that characterise the data, which defines a prototype manifold [289]. Under the prototype manifold, the distance between the prototype vectors and $\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x})$ to is measured to obtain a health indicator.

5. In the multivariate signal processing class, each latent dimension of $\mathbf{z}_t = [z_1, z_2, \ldots, z_d]^T$ is treated as an independent component. By collating each latent representation $\mathbf{z}_t$ for a given signal $x_i[n]$ processed into a Hankel matrix using Equation (2.13), a per-dimension latent signal is obtained.

Henceforth, metrics that conduct comparisons to aspects of the manifold that are time independent, such as the latent prior or static representations of the manifold, are referred to as Path Independent (PI) metrics, and metrics that compare within sensor time are referred to as Path Dependent (PD) metrics. For notational simplicity, subscripts $PD|PI$ are used to describe the LHIs that are PD or PI. PD metrics require sensor time, as they evaluate and quantify latent manifold changes within a temporally preserved setting. In Figure 2.5, a visual representation of the LVM HI and LHI classes is shown. Each of the conceptual classes detailed in Figure 2.5 capture a unique perspective of the latent manifold.

In Table 2.1, a list of potential LHI methods identified for the proposed conceptual categories is given. Table 2.1 demonstrates that the number of possible LHIs is numerous, and that the classes are not limited to only one metric. In this work, it is neither feasible nor useful to use all the methods identified in Table 2.1. Thus, an indicator subset is used in this investigation to demonstrate the effectiveness of the latent manifold for vibration-based condition inference in the presence of time-varying Operating Conditions (OCs).

2.3.5 Proposed indicators

In Section 2.3.4, the five conceptual categories for LHI sampling were detailed. However, the conceptual classes provide several LHI methods, and it is infeasible for this study to use and compare every potential method available. Thus, a small representative subset of LHI methods will be used. These metrics are used to show that the latent manifold is indicative of damage, and that the conceptual classes each capture a unique manifold perspective. None of the indicators from the discriminative category are used as these methods are often computationally demanding due to convex optimisation techniques [293]. To use the proposed indicators, the assumption is made that the latent manifold is accessible. As some of the LHI metrics are dependent on models that capture the latent manifold, the latent representations

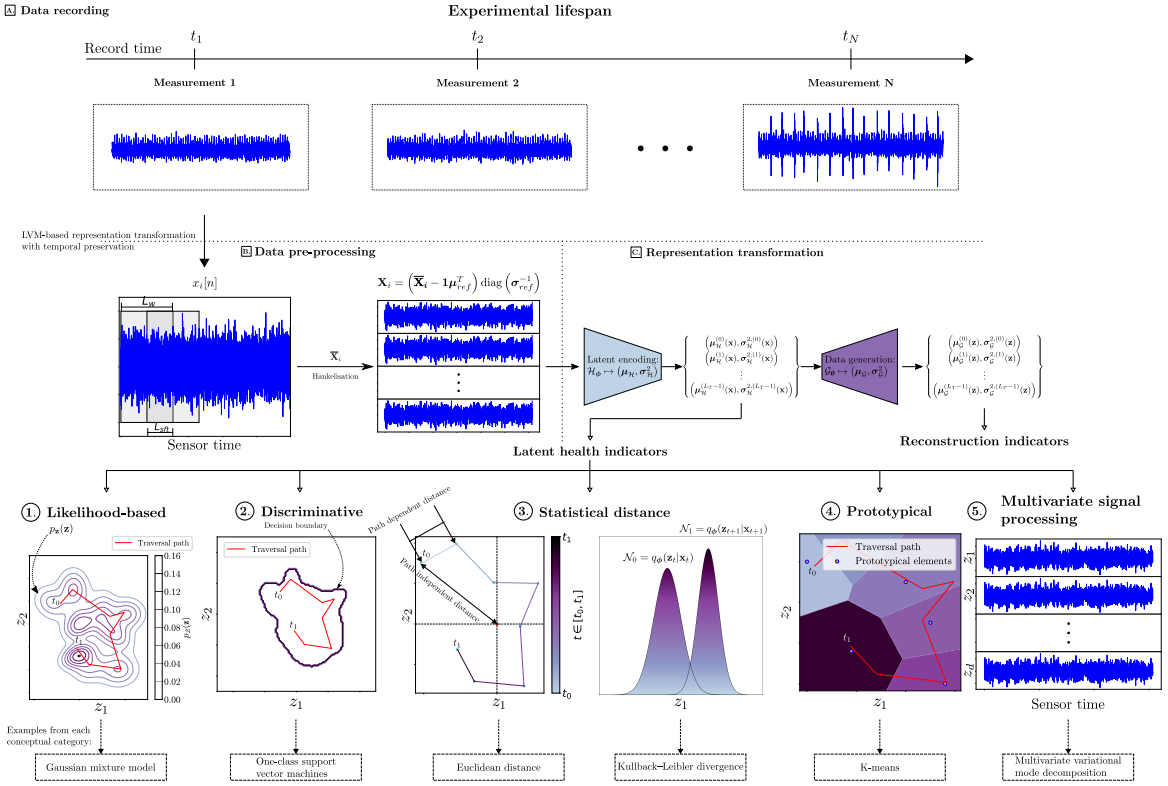


Figure 2.5. A generalised overview of how vibration data interfaces with LVMs and how the data space and latent space is used to obtain condition indicators. The states A., B., and C. indicate the LVM process stages, while the circled numbers 1–5 indicate the five proposed conceptual categories detailed in Section 2.3.4. Note the latent dimensionality d is not specified, and the pictorial perspectives represent an intuitive illustration of the proposed LHI categories.

of the training and validation data are used for reference latent model optimisation. All the proposed indicators exploit the temporal structure induced from the temporal preservation approach. Hence, metrics that are formulated in the PD setting will include reference to t, whereas t is omitted for PI metrics for notational simplicity.

2.3.5.1 Reconstruction indicators

The first metric is the reconstruction HI. This metric is derived through the reconstruction negative log-likelihood of $p(\mathbf{x}|\mathbf{z})$ and is given by

$$HI^{(1)}(\mathbf{x}) = \frac{1}{D}\sum_{l=1}^{D}\frac{(x_l - \overline{x}_l)^2}{\sigma_l^2}, \tag{2.17}$$

where $\overline{\mathbf{x}} = \boldsymbol{\mu}_{\mathcal{G}}$ is the reconstruction mean of $\mathbf{x}$, and $\boldsymbol{\sigma}^2 = \boldsymbol{\sigma}^2_{\mathcal{G}}$ is the reconstruction variance.

Table 2.1. Potential latent manifold indicator methods for each of the proposed conceptual categories. This table demonstrates the types of methods that may be derived from each class, and indicates whether a given method requires some external hyper-parameter selection and/or optimisation intervention. The metrics used in this study are highlighted in bold.

Conceptual class	Method	Requires hyper-parameter/ model parameter optimisation
Likelihood-based	Gaussian model - maximum likelihood estimate for covariance [101]	✓
	Gaussian model - minimum covariance determinant estimate for covariance [290]	✓
	Gaussian mixture model ($LHI^{(1)}$) [101]	✓
	Kernel density estimation ($LHI^{(2)}$) [101]	✓
	Generative topographic mapping ($LHI^{(3)}$) [291, 292]	✓
Discriminative	Support vector data description [293]	✓
	One-class support vector machine [294]	✓
	One-class neural network [295]	✓
	Deep support vector data description [296]	✓
Statistical distance: dissimilarity-based	Manhattan distance (Minowski distance with $p = 1$)	✗
	Euclidean distance (Minowski distance with $p = 2$**)** ($LHI^{(4)}_{PD\|PI}$)	✗
	Chebyshev distance (Minowski distance with $p = \infty$)	✗
	Tanimoto distance (continuous form)	✗
	Kernel distance	✓
	Cosine distance ($LHI^{(5)}_{PD}$)	✗
	Canberra distance	✗
Statistical distance: distribution similarity	**Bhattacharyya distance** ($LHI^{(6)}_{PD\|PI}$)	✗
	Forward KL divergence	✗
	Reverse KL divergence	✗
	Symmetric KL divergence ($LHI^{(7)}_{PD\|PI}$)	✗
	Hellinger distance	✗
	Differential entropy ($LHI^{(8)}_{PD\|PI}$)	✗
	Jensen-Shannon divergence	✗
Prototypical	k-means [101]	✓
	k-medians ($LHI^{(9)}$) [101]	✓
	k-mediods [297]	✓
	Self-organising maps [298]	✓
Multivariate signal: processing and analysis	Univariate empirical/variation mode decomposition [299–301]	✓
	Multivariate empirical/variational mode decomposition [302, 303]	✓
	Minimum entropy deconvolution [50, 265]	✓
	Auto regressive (Box-Jenkins) models [304]	✓
	Moving statistics ($LHI^{(9)}$ and $LHI^{(10)}$)	✓

2.3.5.2 Likelihood-based indicators

For metrics from the likelihood-based category, each metric evaluates the Negative Log-Likelihood (NLL) function

$$LHI_{likelihood}(\mathbf{z}) = -\log p_{ref}(\mathbf{z}), \tag{2.18}$$

where $p_{ref}(\mathbf{z})$ is a reference latent model. This evaluation is PI as the NLL is a proxy metric for an observed sample $\mathbf{z}$ contrasted to the healthy latent manifold. In this work, three reference models are considered, namely, a Gaussian Mixture Model (GMM), a Kernel Density Estimation (KDE) model, and a Generative Topograhphic Map (GTM) model [101]. These models use the latent representations of the data in the training set for model optimisation.

The GMM LHI is given as

$$LHI^{(1)}(\mathbf{z}) = -\log\left(\sum_{k=1}^{K}\pi_k^{(GMM)}\mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_k^{(GMM)},\boldsymbol{\Sigma}_k^{(GMM)})\right), \tag{2.19}$$

where K is the number of Gaussian distribution components used to model the latent manifold. The Expectation Maximisation (EM) algorithm [305] is used to estimate the GMM parameters $\pi_k^{(GMM)}, \boldsymbol{\mu}_k^{(GMM)}$, and $\boldsymbol{\Sigma}_k^{(GMM)}$. The mixture number K is estimated using the silhouette coefficient [306] and the Akaike Information Criterion (AIC) metric [101]. The KDE model LHI is given as

$$LHI^{(2)}(\mathbf{z}) = -\log\left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{(2\pi b^2)^{(d/2)}}\exp\left\{-\frac{\|\mathbf{z}-\mathbf{z}_k\|_2^2}{2b^2}\right\}\right), \tag{2.20}$$

where b is the KDE model bandwidth, $\|\cdot\|_2$ is the Euclidean norm, and K is the number of latent points used as reference centres. The hyper-parameter h is estimated using an exhaustive hyper-parameter grid search and k-fold cross validation on the model log-likelihood. The GTM LHI is given as

$$LHI^{(3)}(\mathbf{z}) = -\log\left(\frac{1}{K}\sum_{k=1}^{K}\mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_k^{(GTM)},\frac{1}{\beta}\mathbf{I})\right), \tag{2.21}$$

where $\boldsymbol{\mu}_k^{(GTM)}$ is a kernel-based regression mapping from a two-dimensional grid of points $\mathbf{u}\in\mathbb{R}^2$ to the latent space and β^{-1} is the local Gaussian variance. The regression mapping $\boldsymbol{\mu}_k$ is defined as

$$\boldsymbol{\mu}_k^{(GTM)} = \mathbf{B}^{(GTM)}\boldsymbol{\phi}(\mathbf{u}_k), \tag{2.22}$$

where $\mathbf{B}^{(GTM)}$ is the GTM mapping matrix, $\boldsymbol{\phi}(\cdot)$ is a basis function, and $\mathbf{u}_k$ is the kth node in the two-dimensional grid. For the GTM model, the number of latent points is set to $K = 625$ to define a 25×25 two-dimensional grid to capture the latent space. The GTM model is a statistical alternative to the self-organising map algorithm [291, 298]. The EM algorithm is used to estimate the model parameters and the basis function of choice is a radially symmetric Gaussian [291, 292].

2.3.5.3 Dissimilarity-based indicators

The dissimilarity metrics of interest here are the Euclidean and cosine distances. Dissimilarity metrics are easily formulated in both the PD and the PI setting. The Euclidean distance is given as

$$LHI^{(4)}_{PD|PI}(\mathbf{z}_{t+1},\mathbf{z}_t) = \|\mathbf{z}_{t+1}-\mathbf{z}_t\|_2^2. \tag{2.23}$$

In the PI formulation, $\mathbf{z}_t$ is set to the zero vector $\mathbf{z}_t = \mathbf{0}$. The cosine distance is given as

$$LHI^{(5)}_{PD}(\mathbf{z}_{t+1}, \mathbf{z}_t) = \cos^{-1}\left(\frac{\mathbf{z}^T_{t+1}\mathbf{z}_t}{\|\mathbf{z}_{t+1}\|_2 \cdot \|\mathbf{z}_t\|_2}\right). \tag{2.24}$$

The cosine distance metric can only be used in the PD setting, as setting $\mathbf{z}_t = \mathbf{0}$ produces an undefined metric.

2.3.5.4 Distribution indicators

To compute statistical distribution differences, an assumption is made that the latent space is Gaussian distributed, which means that closed form solutions are available. This exploitation of the explicit assumption ensures that the statistical distance metrics have closed-form solutions. Metrics from this category are designed to ensure that both the PD and PI settings are accessible. The Bhattacharrya distance, symmetric Kullback–Leibler (KL) divergence and differential entropy are used as representative metrics from this LHI category. The Bhattacharrya distance is given by

$$LHI^{(6)}_{PD|PI}(\mathcal{N}_t \parallel \mathcal{N}_{t+1}) = \tfrac{1}{8}\left(\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t\right)^T \boldsymbol{\Sigma}^{-1}_{t+1}\left(\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t\right) + \tfrac{1}{2}\ln\left(\frac{\det(\boldsymbol{\Sigma})}{\sqrt{\det(\boldsymbol{\Sigma}_t)\det(\boldsymbol{\Sigma}_{t+1})}}\right), \tag{2.25}$$

where $\boldsymbol{\Sigma} = \frac{\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_{t+1}}{2}$, $\mathcal{N}_t = q_{\boldsymbol{\phi}}(\mathbf{z}_t|\mathbf{x}_t)$, $\mathcal{N}_{t+1} = q_{\boldsymbol{\phi}}(\mathbf{z}_{t+1}|\mathbf{x}_{t+1})$, and $\det(\cdot)$ is the determinant function. In the PI setting, $\mathcal{N}_t$ is set to the latent prior $\mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I})$. The symmetric KL divergence metric is the summation of the forward and reverse KL divergence terms and is given by

$$\begin{aligned} LHI^{(7)}_{PD|PI}(\mathcal{N}_t \parallel \mathcal{N}_{t+1}) &= KL(\mathcal{N}_t \parallel \mathcal{N}_{t+1}) + KL(\mathcal{N}_{t+1} \parallel \mathcal{N}_t) \\ &= \frac{1}{2}\left[\mathrm{tr}(\boldsymbol{\Sigma}^{-1}_{t+1}\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}^{-1}_t\boldsymbol{\Sigma}_{t+1}) + (\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t)^T\left(\boldsymbol{\Sigma}^{-1}_t + \boldsymbol{\Sigma}^{-1}_{t+1}\right)(\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t) - 2d\right], \end{aligned} \tag{2.26}$$

where $\mathrm{tr}(\cdot)$ is the trace operator. Finally, the differential entropy of a multivariate Gaussian distribution $\mathcal{N}(\mathbf{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is given as

$$H(\mathcal{N}) = \frac{d}{2}\left(1 + \log(2\pi)\right) + \frac{1}{2}\log\left(\det(\boldsymbol{\Sigma})\right), \tag{2.27}$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{(d\times d)}$. The differential entropy is used to produce a LHI of the form

$$LHI^{(8)}_{PD|PI}(\mathcal{N}_t \parallel \mathcal{N}_{t+1}) = \frac{1}{2}\left[\log\left(\det(\boldsymbol{\Sigma}_{t+1})\right) - \log\left(\det(\boldsymbol{\Sigma}_t)\right)\right]. \tag{2.28}$$

The decision to include $LHI^{(8)}_{PD|PI}(\mathcal{N}_t \parallel \mathcal{N}_{t+1})$ in this investigation was made to ascertain whether the LVMs exhibit sensitivity through $\boldsymbol{\mu}_{\mathcal{H}}$ alone, or through both $\boldsymbol{\mu}_{\mathcal{H}}$ and $\boldsymbol{\sigma}^2_{\mathcal{H}}$. The outputs $\boldsymbol{\mu}_{\mathcal{H}}$ and $\boldsymbol{\sigma}^2_{\mathcal{H}}$ of the latent transition function $\mathcal{H}_{\boldsymbol{\phi}}$ are both dependent on $\mathbf{x}$, and this can be used to determine the sensitivity of $\boldsymbol{\sigma}^2_{\mathcal{H}}$ to damage in vibration data.

2.3.5.5 Prototypical indicators

LHI metrics from this category assume that there are a fixed number of prototypical elements N_p that characterise the latent space $\mathcal{Z}$ [289]. A k-medians algorithm is used to determine the prototypical elements in the latent space [101]. The k-medians LHI is given as

$$LHI^{(9)}(\mathbf{z}) = \min_{k=1,\dots,K} \{\|\mathbf{z} - \mathbf{P}_k\|_2\}, \tag{2.29}$$

where $\mathbf{P}_k$ is the k^{th} prototypical element, $k = 1, \dots, K$. This metric uses the Euclidean distance between $\mathbf{z}$ and the nearest prototypical element for condition inference. The number of prototypical elements K is estimated using the silhouette coefficient [306]. The k-medians metric is PI, and the prototypical elements serve as beacons from which Euclidean manifold dissimilarity within the latent manifold can be measured.

2.3.5.6 Multivariate signal indicators

By treating each latent dimension as an independent data channel, signal processing techniques may be utilised to detect deviations within the latent space. In this work, the moving Root-Mean-Square (RMS) and the moving kurtosis are used for condition inference. The moving RMS, in a discrete-time setting, is given as

$$LHI^{(10)}[t] = \sum_{l=1}^{d} \sqrt{\frac{1}{N_w} \sum_{t}^{t+N_w} z_{l,t}^2}, \tag{2.30}$$

where the summation $\sum_{l=1}^{d}$ is the summation of the windowed RMS terms for each latent dimension, N_w represents the moving window size, and $z_{l,t}$ refers to the l^{th} index of the latent vector $\mathbf{z}_t$. In order to use the moving RMS, the latent representations of vibration data processed using Equation (2.13) is required. The moving kurtosis is given as

$$LHI^{(11)}[t] = \sum_{l=1}^{d} \left[\frac{\frac{1}{N_w} \sum_{t}^{t+N_w} \left(z_{l,t} - \mu_l[t]\right)^4}{\left(\frac{1}{N_w} \sum_{t}^{t+N_w} \left(z_{l,t} - \mu_l[t]\right)^2\right)^2} - 3 \right], \tag{2.31}$$

where $\mu_l[t]$ is the window average

$$\mu_l[t] = \frac{1}{N_w} \sum_{t}^{t+N_w} z_{l,t}. \tag{2.32}$$

In this work, the window size N_w is set to $N_w = 512$ for the moving RMS and the moving kurtosis. The moving RMS and moving kurtosis are both PI metrics.

2.3.5.7 Final reflections

The proposed LHIs are imperative to ensure that all manifolds' perspectives are captured and inspected. Through the five proposed LHI categories, the objective is to showcase that the latent manifold is responsive to damage and that the condition of an asset can be inferred. This ensemble-based approach to latent manifold condition inference can potentially lead to improved detection and interpretation. By investigating the PD and PI characteristics of the latent manifold, manifold deviations through the path travelled or the rate of traversal are accessible. In this way, the proposed conceptual categories allow for different aspects of the latent space to be highlighted.

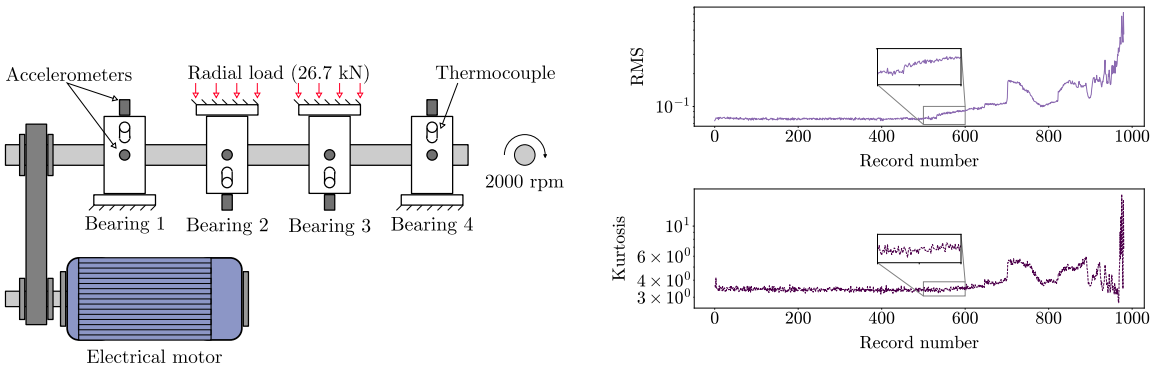
2.4 Experimental evaluation

In this section, the performance of the proposed LHIs will be demonstrated on two experimental datasets. This is imperative to emphasise the responsiveness of the latent manifold to anomalous data. The LVMs considered for this investigation are the PPCA and VAE models [165, 280]. The network architectures, hyper-parameters and all supplementary information required for investigation reproducibility are given in the Appendix A.1. Both gear and bearing fault conditions are considered in the experimental evaluation to encapsulate the prominent gearbox failure modes identified in Section 1.2.1.1. This ensures that the considered temporal-preserving LVMs are applied to representative data for two important gearbox failure cases, where each case exhibits unique fault development mechanisms.

2.4.1 IMS dataset

The dataset was generated by the NSF I/UCR Centre for Intelligent Maintenance Systems (IMS) using an experimental setup as depicted in Figure 2.6(a) [307]. The test rig consists of four rolling element bearings installed on a shaft connected via rub belts to an Alternative Current (AC) motor [216, 307]. Vibration data was recorded using PCB 353B33 high-sensitivity ICP accelerometers attached to the bearing housing of four Rexnord ZA-2115 double row bearings. The dataset contains a total of two different run-to-failure experiments under constant operating and load conditions, with a constant shaft speed of 2000 rpm and a radial load of 26.7 kN exerted on the two central bearings.

The available IMS data was collected from three endurance tests in which bearing faults developed during the test's lifespan. To demonstrate the capabilities of the proposed HIs and LHIs, the vibration data from the second IMS dataset is investigated in this chapter.



(a) IMS experimental setup.

(b) Statistical properties.

Figure 2.6. A diagrammatic representation (a) of the IMS experimental setup and the statistical properties of the recorded vibration signals. In (b), the statistical properties of the raw vibration signals are log-scaled for visualisation purposes.

The recorded data from the second dataset pertaining to bearing one is used as this bearing developed an outer race fault [216, 249]. A total of 984 data recordings were stored over a period of approximately 7 days, and the recordings are each one second long and were digitally sampled with a sampling frequency of 20.48 kHz. The double row bearings had a total of 16 rolling elements with a roller diameter of 0.84 cm, a pitch diameter of 7.15 cm, and a contact angle of 15.17 degrees.

Under the constant operating condition regime, the theoretical Ball Pass Frequency for an Outer race fault (BPFO) is 7.092 shaft orders, or approximately 236 Hz [216, 249]. In Figure 2.6(b), the statistical properties of the recorded data is presented via the RMS and kurtosis of the raw vibration data. The bearing fault degradation is clear in the RMS and kurtosis, and non-monotonic degradation trends are observed through the experimental lifespan. This is attributed to a self-healing phenomena that occurs during the experimental lifespan whereby the outer race fault was smoothed by the rolling elements after defect initialisation for a period of time until further fault propagation occurred [49, 308, 309]. In the literature, it is generally accepted that record 533 is the earliest record for damage identification [249, 310].

2.4.1.1 LVM development

For analysis of the IMS dataset, the first 5% of the available IMS data is assumed as healthy. The healthy data is used for LVM training. The first 49 records are processed using Equation (2.13), split into the training and validation datasets using a 80%–20% split procedure and standardised using Equation (2.15). To perform model evaluation, all of the available data is

processed using the temporal preservation approach, i.e., using $L_{sft} = 1$ for the processing step captured in Equation (2.13) as detailed in Section 2.3.3 [164]. The properties used for training the PPCA and VAE models on the vibration data from the IMS dataset are detailed in Appendix A.1.

2.4.1.2 Analysis methodology

The HI and proposed LHI metrics used in this investigation produce indicator signals. These signals must be post-processed to quantify the responsiveness of the proposed indicators to the presence of a fault. It is emphasised that the analysis methodology is merely a means to facilitate quantitative metric comparison. Hence, this methodology is not proposed to be optimal for CM, only that it allows us to distinguish and quantify the performance of the metrics. For the IMS dataset, the indicator signal RMS and the amplitude of the BPFO component is used for performance quantification purposes. The latter is used as the a priori for the considered IMS dataset is that a bearing fault occurred. As the indicators all have different orders of magnitude, the available indicators are standardised using the records used for LVM training. The standardisation process ensures that different indicators are non-dimensionalised, and removes any interpretation bias to the offset and the scale of the indicators [311, 312].

To allow for a quantitative performance comparison of the different metrics, the Area Under the Reciever Operating Characteristics (ROC) Curve (AUC) is used. As discussed in Ref. [164], the usefulness of the AUC in CM applications for rotating machinery is acutely sensitive to the shaft speed f_s, model window length L_w and the sampling rate F_s of the measured data. The AUC ignores sensor time and is used to quantitatively compare the performance of the considered HIs and LHIs, providing insight into the indicator's sensitivity to the fault present in the data. The reference class distribution is developed using the indicator signals corresponding to the measured vibration data used for LVM training.

To quantify the similarities and differences of the considered HIs and LHIs on an inter-method and intra-class level, the Pearson Correlation Coefficient (PCC), represented by $\rho_{X,Y}$ for two random variables X and Y, will be calculated for the RMS of the indicator signals through record time. The sample PCC, represented by r_{xy}, is a linear correlation measure that is formulated by normalising the covariance between two random variables and is calculated

for N_m sets of paired data $[(x_1, y_1), \ldots, (x_{N_m}, y_{N_m})]$ using

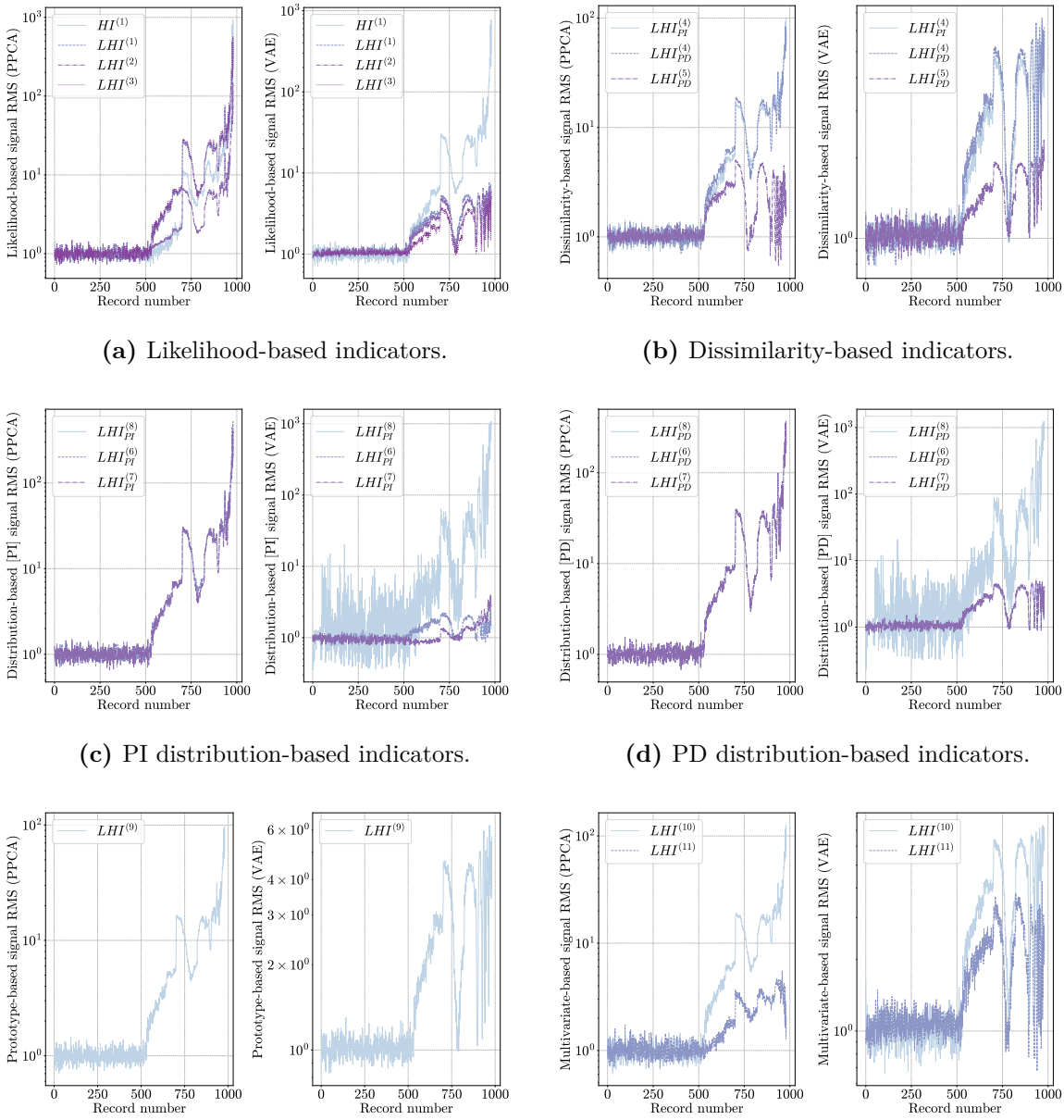
$$r_{xy} = \frac{\sum_{j=1}^{N_m} (x_j - \mathbb{E}\{x\})(y_j - \mathbb{E}\{y\})}{\sqrt{\sum_{j=1}^{N_m} (x_j - \mathbb{E}\{x\})^2} \sqrt{\sum_{j=1}^{N_m} (y_j - \mathbb{E}\{y\})^2}}, \tag{2.33}$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator.

2.4.1.3 Results

In Figure 2.7 the HI and LHI indicator results are shown. Most metrics detect and respond to some change in the vibration data in the region around record number 500. This indicates that the response to anomalous data manifests in both the data space and the latent manifold. The dual nature of the responsiveness to anomalous data in both the data space and the latent manifold is useful for condition inference. As the number of latent manifold metrics surpasses the number of data space metrics, the damage detection process is enhanced by incorporating additional LHIs. Thus, the fusion of both data space and latent manifold indicators is powerful for condition inference. Each of the conceptual categories have metrics that respond to the anomalies present in the data, which emphasises the benefit of considering and using metrics from the conceptual categories proposed in Section 2.3.4. Furthermore, an explicit LVM with linear transition functions is responsive to the anomalous data present in the considered IMS dataset. Thus, both the linear and nonlinear explicit LVMs can capture the latent manifold of the reference IMS vibration data, and both manifolds are indicative of damage.

Due to the number of responsive metrics in Figure 2.7, metrics that appear to be unresponsive to the presence of the anomalous vibration data are discussed. These standardised metrics are $LHI_{PI}^{(8)}$ and $LHI_{PD}^{(8)}$ for the PPCA model seen in Figures 2.7(c) and 2.7(d) respectively. The unresponsiveness of the $LHI_{PI|PD}^{(8)}$ metric is expected for the PPCA model as this metric represents differential entropy of a Gaussian distribution, and this is only a function of $\boldsymbol{\Sigma}_{\mathcal{H}}$. However, the covariance matrix of the PPCA latent posterior distribution is independent of the data [101, 280]. Regardless of whether the PI or PD characteristics of the LHI are inspected, the resulting PPCA model can only respond to damage through changes in the posterior mean $\boldsymbol{\mu}_{\mathcal{H}}(\mathbf{x})$. In the PI and PD formulation, the VAE $LHI^{(8)}$ metric is responsive to the anomalies in the IMS vibration data. This result emphasises the need to consider both the PD and PI perspectives of the latent manifold when deriving LHIs, as the method in which the latent manifold responds to anomalous data is not guaranteed. Hence, metrics that monitor both the characteristics of the path traversed and the path traversal rate in the latent manifold are imperative for condition inference.



(a) Likelihood-based indicators.

(b) Dissimilarity-based indicators.

(c) PI distribution-based indicators.

(d) PD distribution-based indicators.

(e) Prototypical indicators.

(f) Multivariate-based indicators.

Figure 2.7. The RMS of the standardised HI and LHI indicator signals for the PPCA and VAE models. The results for each LHI category is shown for the PPCA and VAE models. In (d), $LHI^{(8)}$ in both the PI and PD formulation for the PPCA model cannot be seen as the presented figures are log-scaled and the standardised metric is zero throughout record number. Note that the model indicator is given in the y-axis label of each sub-figure.

In an inter-method analysis of Figure 2.7, the PPCA model produces stronger latent manifold deviations. The PPCA model metrics consistently have a larger scale change in comparison to the VAE metrics, for metrics that exhibit a response to the anomalous data. This scale change is quantifiable as the indicator signals are standardised. While the reconstruction

indicator, $HI^{(1)}$, for the VAE model appears to be more sensitive to the anomalous data, the latent manifold, while responsive, is less sensitive to the anomalous data. This indicates that a linear latent manifold is sufficient for condition inference purposes, and indicates that the generative transition function, the VAE decoder, is more expressive than the latent posterior transition function.

To quantify the intra-class similarities for the considered metrics, the sample PCC was calculated using the RMS trends seen in Figure 2.7 on a intra-class and inter-method level. The former refers to the correlation for metrics within different classes while the latter refers to the correlation between the metrics from a specific model class with all the metrics from another model. In Figure 2.8, the intra-class correlation is shown. Evidently, from the intra-class correlation, most indicators from both models are highly correlated as the median of both methods is greater than 0.75. However, the tails in Figure 2.8 suggest that there are unique responses in both methods, as the correlation distribution is spread over the r_{xy} domain. This suggests that, for the IMS dataset, most of the indicators presented in this investigation exhibit similar responses when inspecting the indicator signal RMS. In an inspection of Figure 2.7, one would expect that the distribution-based metrics for the PPCA model have strong correlation and hence selecting either $LHI^{(6)}_{PD|PI}$ or $LHI^{(7)}_{PD|PI}$ would have been suitable.

Additionally, the $LHI^{(4)}_{PI}$ and $LHI^{(4)}_{PD}$ metrics exhibit clear similarity, and hence either of these would have been suitable for the dissimilarity-based class. In an inter-method correlation comparison, shown in Figure 2.9, the indicators from the different classes exhibit non-similar responses on an inter-method level. This indicates that the choice of method influences the sensitivity of the indicators to the fault present in the data. This indicates that the metrics used in this investigation are model formulation dependent. The prototype-based indicator from the PPCA model has a condensed r_{xy} range, which suggests that this indicator exhibits a similar response to the indicators obtained from the VAE model. The indicators from the VAE model, on an inter-method level, are unique to the PPCA model, as the r_{xy} distribution is spread over the domain for each of the indicator classes.

To infer the damage present in the IMS dataset, the spectral amplitude of the analytical BPFO component was monitored. In Figure 2.10, the BPFO amplitudes are shown for the HIs and LHIs considered in this work. It is evident that many of the proposed LHI signals contain this spectral component and that it grows over the experimental lifespan. The temporal preservation approach used in the model evaluation stage is critical to ensuring that this

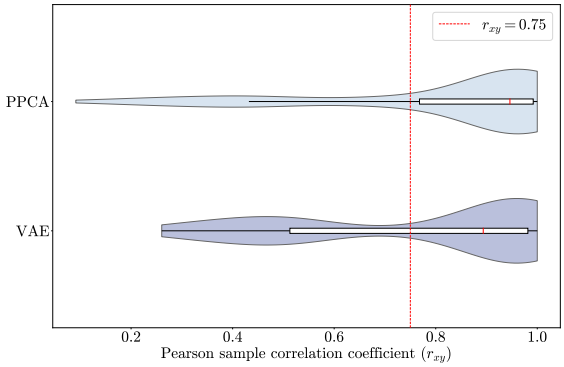
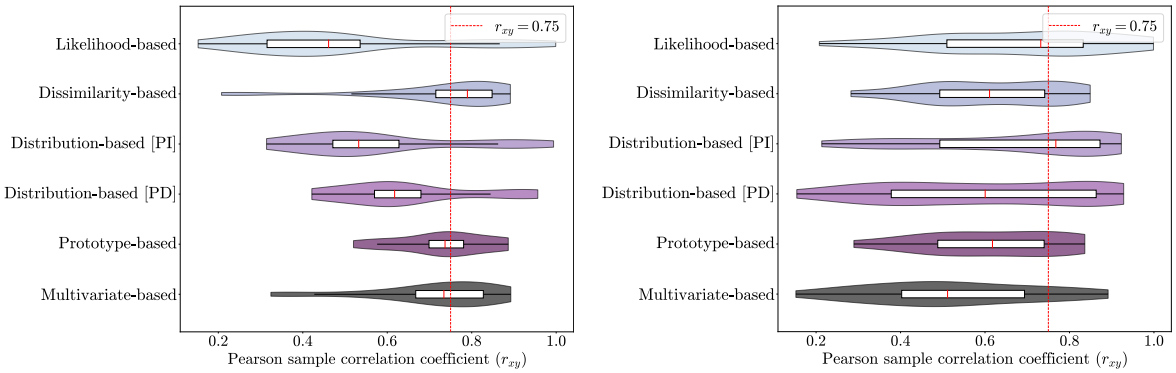


Figure 2.8. Violin plots of the sample PCC on an intra-class level for the IMS dataset. Note that the intra-class comparison is visualised on a per-method level as the r_{xy} terms are computed for all metrics of a given LVM method to compare class-wide similarity.



(a) PPCA method-wide correlation.

(b) VAE method-wide correlation.

Figure 2.9. Violin plots of the sample PCC on an inter-method level for the IMS dataset. Note that the inter-method comparison is visualised on a per-class level as the r_{xy} terms are computed for all metrics of a given indicator class and all metrics from another LVM method to compare method-wide similarity.

spectral component is accessible. Thus, it is possible to not only detect the anomalous data in the IMS dataset, but the cause of the anomalies may be attributed to damage on the outer race of the bearing.

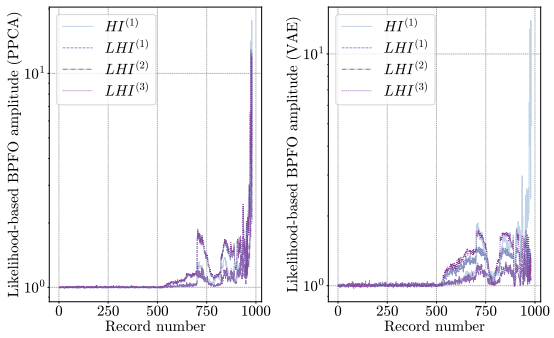
Figure 2.11 depicts the AUC results for the considered HIs and LHIs. The AUC provides a quantitative perspective on the sensitivity of the considered HIs and LHIs for the PPCA and VAE models. In Figure 2.11, sensitivity of the considered HIs and LHIs to the bearing fault development during the experimental lifespan is clear. Furthermore, the variation in manifold sensitivity emphasises the requirement for an ensemble-based latent inference procedure, as different perspectives of the latent manifold produce a variation in the metric response. The

VAE model HI and LHI indicators are more sensitive to the bearing self-healing phenomena present around record 750, which results in an AUC reduction. The strength of the PPCA AUC deviations indicates that facets of the PPCA latent manifold is more sensitive to the anomalous fault data, and the PPCA latent manifold response is significantly more noticeable than the VAE manifold response.

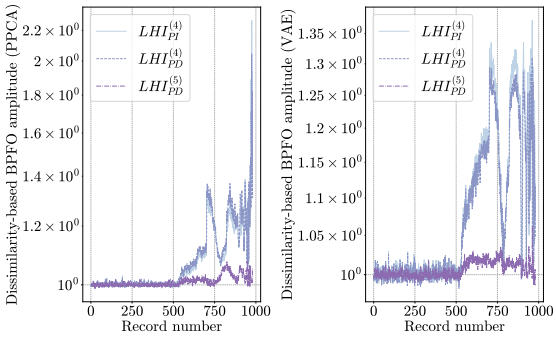
Hence, from the results presented in Figures 2.7, 2.10, and 2.11, it is evident that the proposed HIs and LHIs are useful for condition inference. The results in Figure 2.7 indicate that the metrics are responsive to the fault present in the data, and that the RMS of the LHI is a suitable method for indicator trending. Figure 2.10 serves as an indication of the usefulness of performing temporal preservation during the model analysis stage, as the fault type is isolated and identified. By performing an inter-method and intra-class analysis of the RMS trends of the indicators for the IMS dataset: *(i)* the metrics from the likelihood-based class that monitor the latent manifold offer similar insight, *(ii)* the nonlinear model has a more sensitive response in data space rather than in the latent manifold, *(iii)* the dissimilarity metrics exhibit similar responses across methods and only the $LHI^{(5)}$ metric has a unique response to the other metrics, *(iv)* the $LHI^{(6)}$ and $LHI^{(7)}$ metrics respond similarly, and *(v)* the uniqueness of the PI distribution-based metrics and the multivariate-based metrics is LVM methodology dependent.

2.4.1.4 Comparison

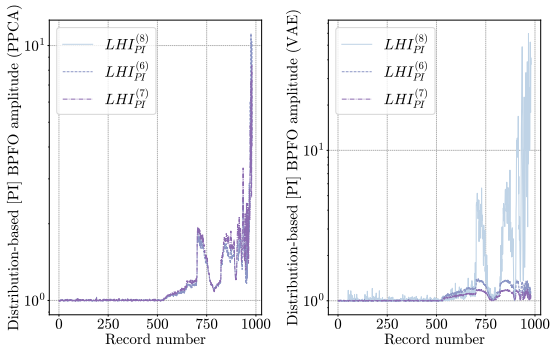
To compare the LHIs used in this investigation to HIs from other works, two methods were used. For the first method, the Mahalanobis distance-based HI proposed in Wang *et al.* [313] is used. For the second method, a set of sparsity measures that provide HIs are used. These sparsity measures utilise the squared envelope spectrum and are designed to exhibit sensitivity to the impulsiveness of repetitive transients. The sparsity measures considered here are the spectral kurtosis [190], the spectral negentropy [31, 211], the Smoothness Index (SI) [214], the Gini Index (GI) [213], the ratio of the L_2 and L_1 norm ($\frac{L_2}{L_1}$) [212], the Hoyer Measure (HM) [210], the Box-Cox (BC) measure [215], and the skewness. Note that a passband $[l, h]$ filtered version of a signal $x_i[n]$ is not considered as the objective is not to obtain a transformed signal that maximises the sparsity measure. Methods focused on the optimal passband are well detailed in the literature, e.g., [32], but are beyond the scope of this work. The BC measure hyper-parameter λ is set to 0.5 as this represents a sparsity measure between the spectral kurtosis and the spectral negentropy.



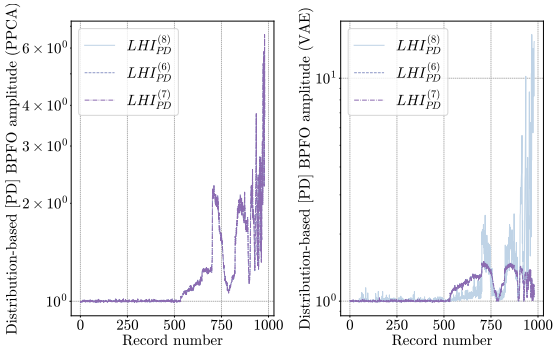
(a) Likelihood-based indicators.



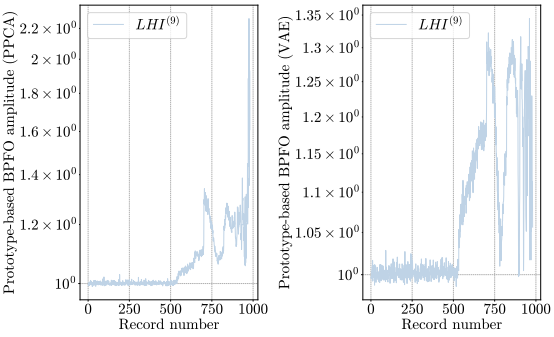
(b) Dissimilarity-based indicators.



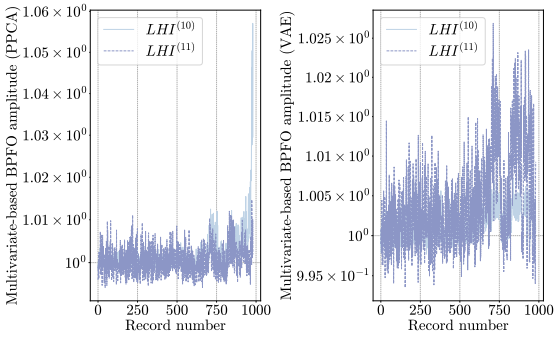
(c) PI distribution-based indicators.



(d) PD distribution-based indicators.



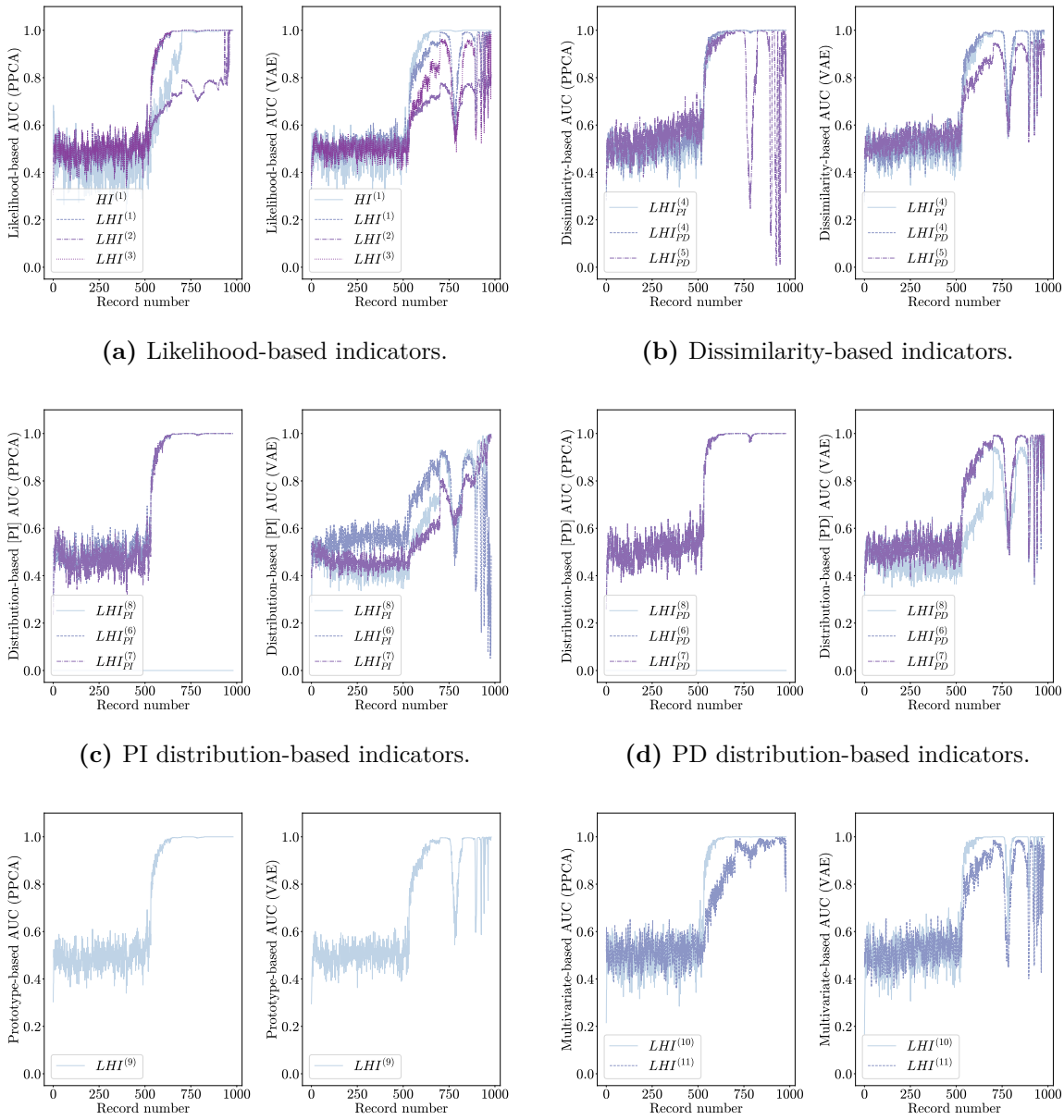
(e) Prototypical indicators.



(f) Multivariate-based indicators.

Figure 2.10. The amplitude of the BPFO spectral component for the standardised HI and LHI indicator signals considered in this work. In (c) and (d), the amplitude of $LHI^{(8)}$ for the PPCA model cannot be seen as the metric is zero and the presented figures are log-scaled. Note that the spectral components for the considered HIs and LHIs were zero-mean adjusted using the values for the records corresponding to the training data and shifted to 1 for comparison purposes.

In Figure 2.12(a) the results of using the HI proposed in Ref. [313] for the IMS dataset is shown, and it is responsive to the bearing fault. However, its performance is inferior to the proposed HIs and LHIs as it is less responsive. Additionally, this HI is designed for bearing



(a) Likelihood-based indicators.

(b) Dissimilarity-based indicators.

(c) PI distribution-based indicators.

(d) PD distribution-based indicators.

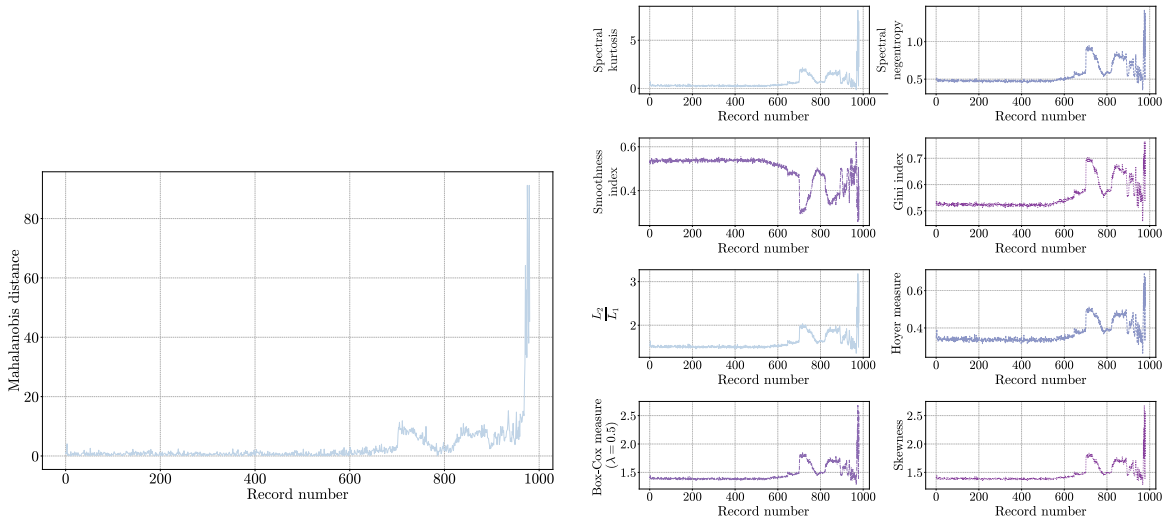
(e) Prototypical indicators.

(f) Multivariate-based indicators.

Figure 2.11. The AUC for the standardised indicator signals considered in this work. In (c) and (d), the AUC is zero as the $LHI^{(8)}$ for the PPCA model is independent of the data and thus exhibits no deviations or changes in value. An AUC of 0.5 indicates that the indicator signal is indistinguishable from the indicators from the reference healthy data.

fault detection, uses hand-crafted time-domain features, and is temporally non-preserving, whereas the proposed LHIs require no prior knowledge of the fault mode and are trained on the observed vibration data alone. Thus, the responses from the proposed LHIs, in comparison with the methodology from Ref. [313], are useful and exhibit a clear response to the bearing

fault present in this dataset.



(a) Mahalanobis distance indicator. **(b)** Sparsity indicators.

Figure 2.12. The Mahalanobis distance metric from Ref. [313] and the considered sparsity metrics applied to the IMS dataset.

The sparsity measures in Figure 2.12(b) exhibit sensitivity to the bearing fault and are sensitive to the change in condition in the region surrounding record 500. Each of the considered sparsity metrics are non-monotonic, attributed to bearing self-healing, and are indicative of a change in asset state. The proposed LHIs, in comparison, exhibit similar sensitivity and are not designed to enhance any property of the observed data. Hence, they are useful to vibration-based CM. Additionally, the fault type can be inferred from the LHIs, which is only possible for the sparsity measures if they are used in a demodulation-band framework [32] or a blind-deconvolution framework [264, 276].

2.4.2 Gearbox dataset

The Centre for Asset Integrity Management (C-AIM) at the University of Pretoria (UP) in South Africa has made available an experimental gearbox fatigue dataset that was obtained under time-varying OCs. The experimental setup was originally designed by Stander and Heyns [40]. The experimental gearbox dataset has representative data for a gearbox in a healthy and unhealthy condition with complex operating conditions and the experimental setup is depicted in Figure 2.13. The test rig consists of an alternator, two step up helical gearboxes, one step-down helical gearbox, and an electrical motor. In this dataset, vibration data was recorded via the use of a 100 mV/g tri-axial accelerometer mounted to the input

shaft bearing housing of the monitored gearbox. Information related to the OCs was recorded using an optical probe and a zebra tape shaft encoder attached to the input gear shaft of the monitored gearbox.

Time-varying OCs were introduced in this dataset to increase the complexity of the diagnostics problem, and the fluctuation in OCs experienced by the setup are depicted in Figure 2.14(a). The statistical properties of the recorded data are presented in Figure 2.14(b). Additionally, impulsive information is present in the recorded data that is not related to the fault mechanism of interest in this dataset and this further complicates the fault diagnosis task. The impulsive component is attributed to contact between a floating bearing and the gearbox casing induced by the axial loads of the helical gearboxes [314].

The gearbox dataset consists of a total of 300 experimental data recordings and can be sequentially separated into a healthy state and a damaged state. In the initial healthy state, a total of 100 recordings were obtained of the gearbox in a healthy condition. Subsequently, the gearbox was disassembled and the gear was artificially damaged by seeding a slot into one of the gear teeth, and re-assembled to complete a second sequence of measurements. A total of 200 records were obtained for the unhealthy state until gear failure occurred. The gear at the beginning and end of the unhealthy state is depicted in Figure 2.15(a) and Figure 2.15(b) respectively. The vibration data and tachometer data was recorded for 20 seconds with a sampling frequency of 25.6 kHz and 51.2 kHz respectively. Following Ref. [164] and Ref. [315], the available data is low-pass filtered in this investigation with a third-order Butterworth filter around 3200 Hz.

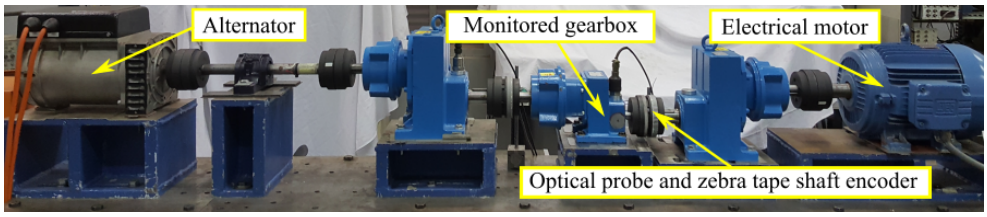
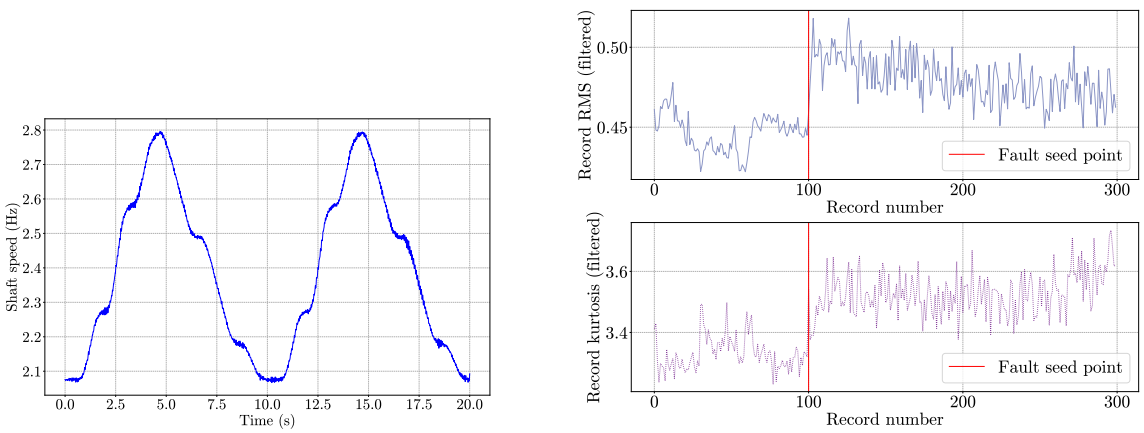


Figure 2.13. The experimental C-AIM gear fatigue test setup.

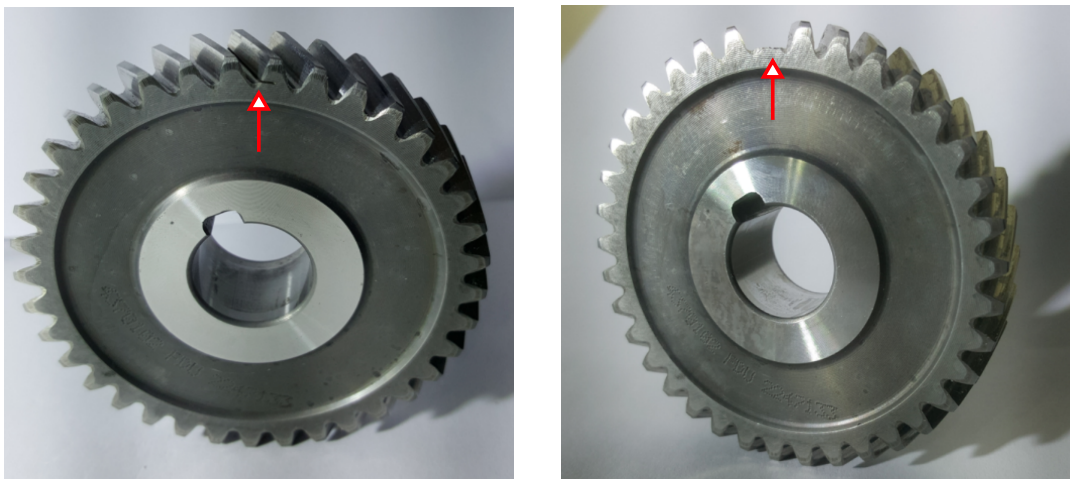
For analysis of the gearbox dataset, 25% of the available healthy gearbox data is used for LVM training. The available healthy records are processed using Equation (2.13), split into the training and validation datasets using a $80\% - 20\%$ split procedure and standardised using Equation (2.15). To perform model evaluation, all the available data is processed using the temporal preservation approach [164]. The use of the temporal preservation approach during



(a) Input shaft operating conditions.

(b) Statistical properties.

Figure 2.14. The input gear operating conditions and the per-record RMS and kurtosis for the vibration data available from the experimental gearbox dataset. It is evident from (a) and (b) that time-varying OCs are present and that the damage is difficult to isolate through the statistical properties of the vibration data alone. However, a change in state is detectable at record 100, attributed to the manual fault seeding procedure used to introduce gear tooth damage.



(a) Seeded gear state.

(b) Final gear state.

Figure 2.15. The helical gear used to introduce localised gear tooth damage for the C-AIM gearbox dataset. (a) and (b) detail the seeded gear tooth failure before and after experiments the vibration data for the unhealthy experimental state was recorded.

model evaluation produces a set HI and LHI signals from which the state of the experimental set-up may be inferred. Appendix A.1 details the properties of the PPCA and VAE models used for the gearbox dataset investigation.

2.4.2.1 Analysis methodology

The HIs and proposed LHI metrics used in this chapter produce indicator signals. As it is known that the fault characteristic present in the gearbox dataset is a gear tooth fault, an analysis procedure is followed to enable quantitative comparison of the metrics considered. The metrics, by design, are not insensitive to the non-stationary OCs present in the gearbox data. Hence, the analysis procedure used here is to quantify and, where applicable, highlight the usefulness of each indicator without sensitivity to the variation in the asset operating conditions. This analysis procedure is merely a means to enable metric comparison, and this procedure is not proposed to be state-of-the-art. To post-process the indicator signals, the Time Synchronous Average (TSA) of the indicator signals is used to ascertain whether the gear tooth damage is detectable. The TSA of signal $x_i[n]$, $n \in \mathbb{Z}$, is given as

$$\overline{x}_{TSA}[n] = \frac{1}{N_r} \sum_{l=1}^{N_r-1} x_i[n + l \cdot N_s],\ 1 \leq n \leq N_s, \tag{2.34}$$

where N_r is the number of shaft rotations, and N_s is the number of points per rotation [209]. The resulting TSA signal is a sequence of N_s points and represents the deterministic component of $x_i[n]$ for one shaft rotation. Each TSA signal is realigned with respect to the tachometer butt joint. The location of the tachometer butt joint is found using the Bayesian Geometry Compensation (BGC) algorithm proposed in Diamond *et al.* [316].

To qualitatively compare the LVMs considered in this work, and to assess the responsiveness of different indicators, a GMM-based algorithm is used to quantify the presence of an isolated fault [101]. In this algorithm, a GMM with two components is used to model the TSA. For the TSA of a healthy indicator signal, the components of the GMM should have little to no observable statistical difference. However, for the TSA of an unhealthy indicator where localised damage is present, the components of the GMM should be statistically different. To quantify the statistical difference between the two local Gaussians of the GMM, the symmetric KL divergence is used. This divergence metric is useful as it offers an analytical solution, does not favour one of the local distributions over the other, and is non-negative. Given parameters of the two local GMM components $(\mu_{0,TSA}, \sigma^2_{0,TSA})$ and $(\mu_{1,TSA}, \sigma^2_{1,TSA})$, the symmetric KL divergence becomes

$$KL_{symmetric} = \tfrac{1}{2}\left(\frac{\sigma^2_{0,TSA}}{\sigma^2_{1,TSA}} + \frac{\sigma^2_{1,TSA}}{\sigma^2_{0,TSA}} + \frac{\left(\mu_{1,TSA}-\mu_{0,TSA}\right)^2}{\sigma^2_{1,TSA}} + \frac{\left(\mu_{0,TSA}-\mu_{1,TSA}\right)^2}{\sigma^2_{0,TSA}} - 2\right). \tag{2.35}$$

The use of the GMM-based algorithm allows for trending of the properties of the TSA over the experimental lifespan. As with the IMS dataset, the TSA of the indicators is standardised

to remove any bias to the scale or offset of the indicators prior to employing the GMM-based algorithm.

To quantify the similarities and differences of the considered HIs and LHIs on an inter-method and intra-class level, the sample PCC, given in Equation (2.33), will be calculated for the symmetric KL divergence trends through record time. To allow for quantitative performance comparison of the different metrics, the AUC is used. However, as the fault in this dataset is localised around one gear tooth, synchronous average indicator values are carefully selected in the region of the tooth in which the fault was seeded as the test set for AUC evaluation. This selection procedure resulted in 200 samples per HI and LHI record for testing and AUC evaluation. The TSA samples per record used in this evaluation is approximately twice the number of samples that would correspond to one tooth window, which accounts for the loading zone of the damaged tooth of interest. This ensures that *(i)* the AUC can be accurately computed for each vibration data record, *(ii)* there is a suitable number of reference healthy samples to represent the healthy tooth window, and *(iii)* that the responsiveness of different LHIs can be quantified for this dataset without succumbing to AUC insensitivity issues. Note that the AUC procedure used is specific to the fault present in the gearbox dataset and cannot be readily applied to every fault class.

Discussion on LVM method considerations

An important distinction must be made between how the variable OC effect is handled in the analysis of the gearbox dataset in this chapter and for temporal-preserving LVM methods in general. The decision to use the TSA was purposefully made to enable a quantitative HI and LHI metric comparison as it is a highly effective gear fault indicator [5] that serves as a means to strengthen the prominence of synchronous components, e.g., signal information produced by amplitude modulation occurring in a gear tooth mesh period, that correspond to the localised gear tooth damage. This is necessary to assess the suitability of the proposed indicators in a fault diagnosis context. However, the TSA is only applied to the HI and LHI signals obtained from a temporal-preserving LVM whose parameters were estimated using historical vibration data and was not applied directly to the observed vibration data. In this way, the TSA serves to enhance the diagnostic information from the outputs of the temporal-preserving LVM, the LHI signals, implying that diagnostic information must be present model facets to begin with. This aligns with the objective of the investigations carried out in this chapter, which is to scrutinise the competency of the latent manifold—a facet of LVMs—in capturing relevant

diagnostic information.

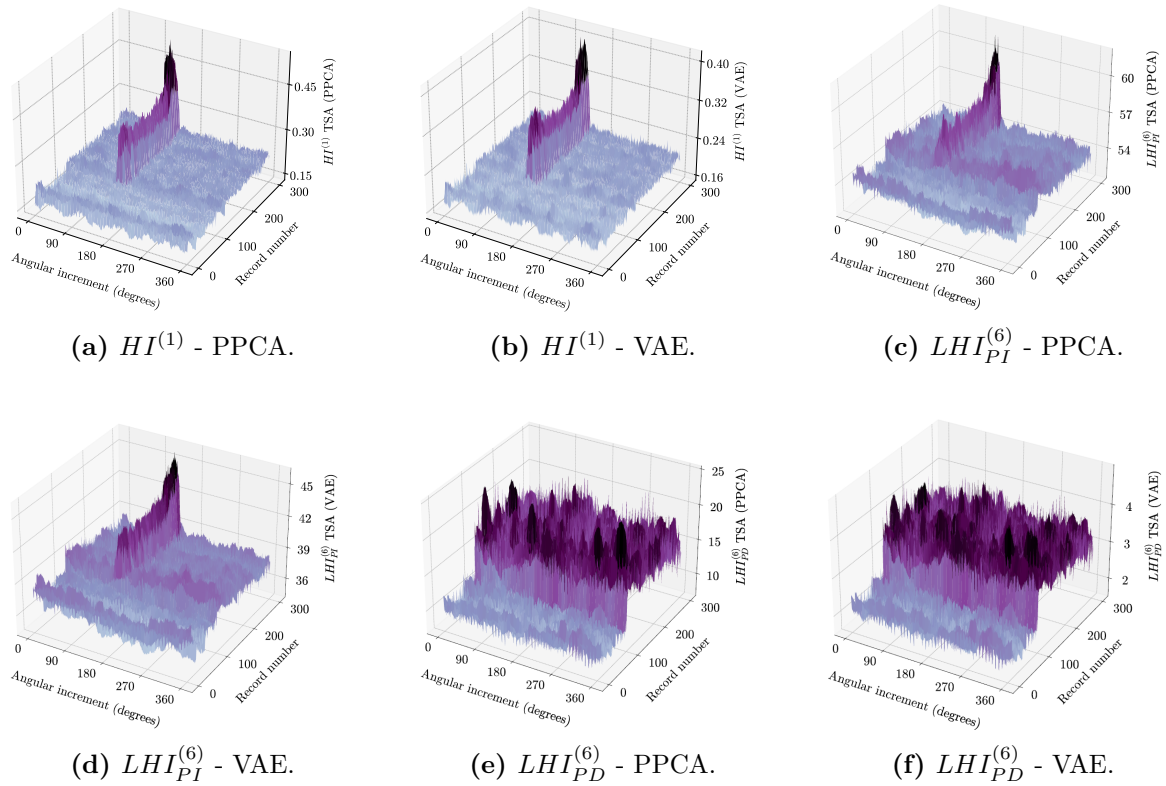
Moreover, the temporal-preserving LVMs are developed on a hankelised representation of the vibration data given via Equation (2.13), a straightforward data processing strategy that does not induce significant alterations to the observed vibration data. Temporal preservation [164] ensures that LVMs are, in general, applicable to any OC scenario induced throughout an asset's lifespan. The critical component to the success of an LVM-based methodology is then how the HIs and LHIs are extracted from the data space and the latent manifold.

Nonetheless, a reliance on synchronous averaging methods posits that temporal-preserving LVMs are sensitive to the amplitude modulation and frequency modulation effects induced by time-varying OCs, as discussed in Section 1.2.1.3. However, this type of sensitivity is a prevalent problem that affects numerous vibration-based CM methodologies developed for time-varying OCs and is a primary reason why normalisation strategies, e.g., [85, 87], have been developed to separate the OC-based amplitude modulation from fault-based amplitude modulation, thereby attenuating the potential downstream effects of variable OCs. Normalisation strategies as a pre-processing step for the LVMs developed on the unfiltered gearbox data were explored as part of the investigations related to the content of this thesis. However, the impulsive data component, discussed in Section 2.4.2, was found to be the primary source of interference for HI and LHI performance in the gearbox dataset investigation. Further comments related on this discussion point are contained in Section 5.2 of this thesis.

2.4.2.2 Results

In Figure 2.16 the TSAs for a selected subset of the HI and LHI metrics considered in this investigation are shown as functions of record number. The selected subset of metrics showcases the TSA trends of metrics from both the data space and the latent manifold for both the PPCA and the VAE model. Evidently, through the inspection of Figure 2.16, the gear tooth fault is detectable by both the PPCA and VAE models. In Figure 2.16(e) and Figure 2.16(f), the TSAs of the $LHI_{PD}^{(6)}$ metric for both LVMs are ineffective for condition inference. This result emphasises the necessity to consider several LHI metrics in the latent manifold, as the method of LVM latent deviation is not guaranteed.

In Figure 2.17 the GMM-based symmetric KL divergence trends using the post-processing methods discussed in Section 2.4.2.1 are shown. These trends are necessary to enable a quantitative metric comparison of the considered HIs and LHIs. Quantitative comparison



(a) $HI^{(1)}$ - PPCA. **(b)** $HI^{(1)}$ - VAE. **(c)** $LHI^{(6)}_{PI}$ - PPCA.

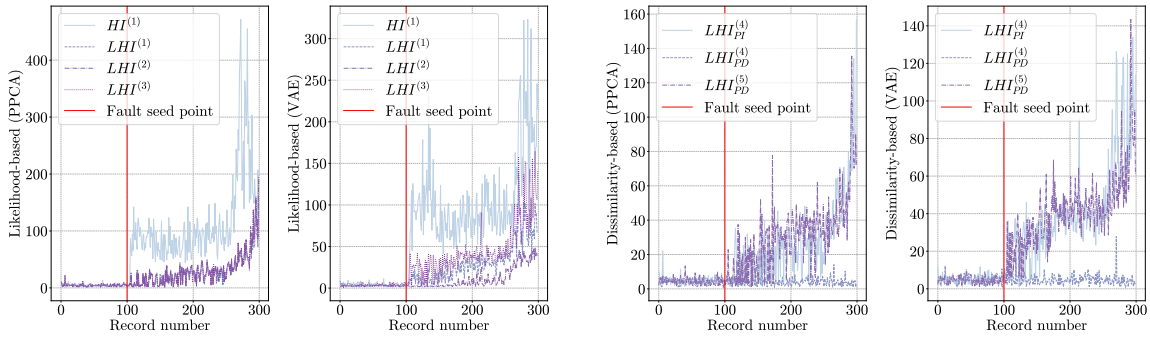
(d) $LHI^{(6)}_{PI}$ - VAE. **(e)** $LHI^{(6)}_{PD}$ - PPCA. **(f)** $LHI^{(6)}_{PD}$ - VAE.

Figure 2.16. The TSAs for a selected subset of the HIs and LHIs considered in this work.

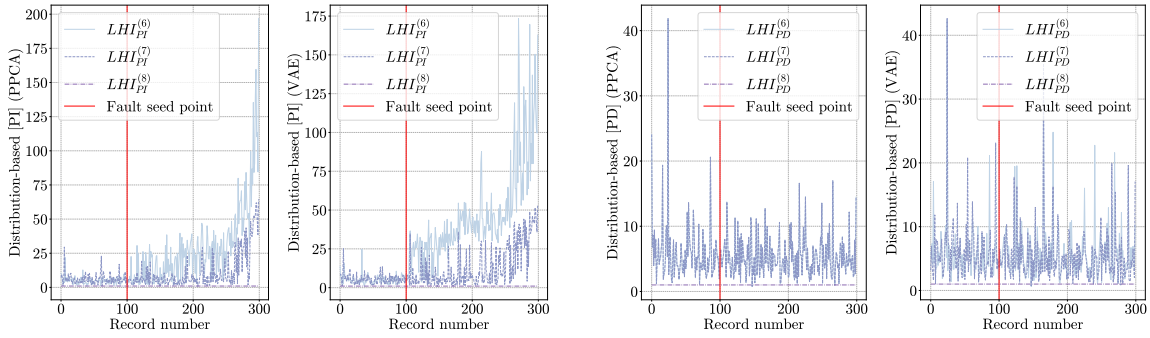
metrics such as monotonicity and trendability, discussed in Ref. [228], are not considered as it is not proposed in this investigation that the post-processing methodology used is a direct requirement for LVM-based condition inference. Figure 2.17 shows that many of the metrics are responsive to the gear tooth fault. In Figures 2.17(b) and 2.17(d), the Euclidean-based PD metrics are seen to be uninformative. However, as seen in Figure 2.17(b), the angular metric $LHI^{(5)}_{PD}$, which is PD, is indicative of the damage present in the vibration data. The $LHI^{(8)}$ metric, which monitors the entropy of the latent posterior distribution, exhibits no response to the gear tooth fault.

This result indicates that both the PPCA and VAE models respond to the damage in the data through the posterior mean $\boldsymbol{\mu}_{\mathcal{H}}$. While this is expected for the PPCA model, the VAE model, by design, had a posterior mean and a posterior covariance that were functions of the input data $\mathbf{x}$. This result emphasises the importance of considering metrics from the conceptual categories proposed in Figure 2.3.4, as they ensure that all facets of the considered LVMs are understood. Figure 2.17 indicates that both the PPCA and VAE models are useful for condition inference on the gearbox dataset. Thus, important considerations must be made

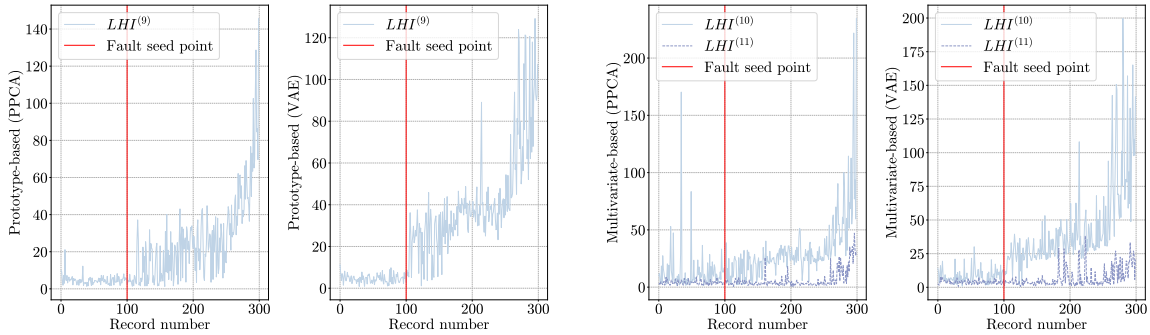
when investigating alternative LVM formulations, and PPCA may serve as a useful baseline for performance comparison and quantification in future work.



(a) Likelihood-based indicators. **(b)** Dissimilarity-based indicators.



(c) PI distribution-based indicators. **(d)** PD distribution-based indicators.



(e) Prototypical indicators. **(f)** Multivariate-based indicators.

Figure 2.17. The TSA GMM-based symmetric KL divergence trends for the HIs and LHIs considered in this work.

In Figure 2.18, the violin plots of the intra-class correlation at a per-method level are shown. Figure 2.18 shows a clear correlation between indicators from different classes, and the metrics with weak correlation most likely belong to indicators evaluated in the PD setting,

namely $LHI_{PD}^{(4)}$, and $LHI_{PD}^{(6-8)}$. This highlights that there is some variety and uniqueness in the responsiveness of the indicators considered in this work.

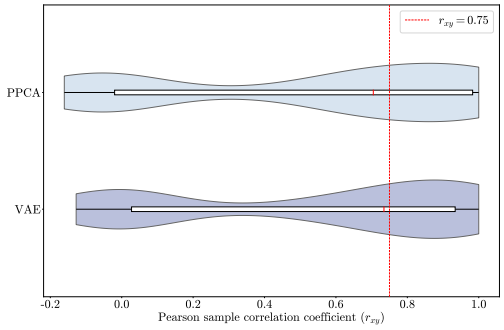
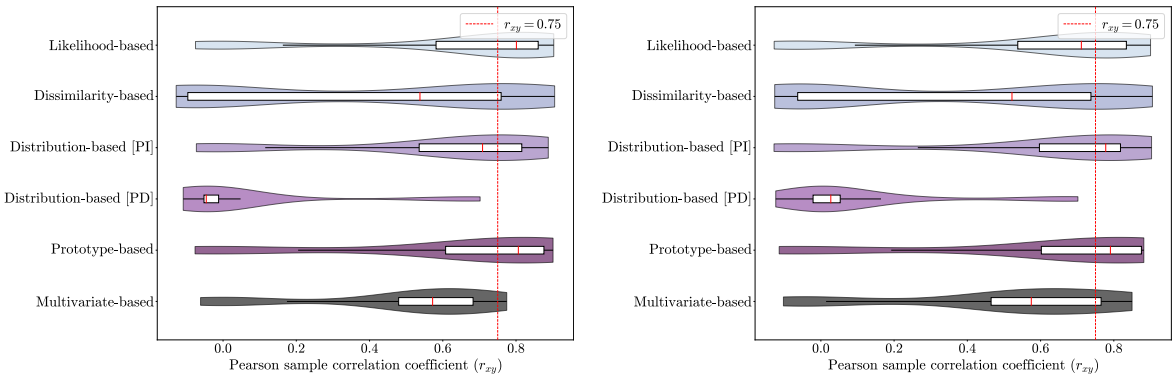


Figure 2.18. Violin plots of the sample PCC on an intra-class level for the gearbox dataset. Note that the intra-class comparison is visualised on a per-method level as the r_{xy} terms are computed for all metrics of a given LVM method to compare class-wide similarity.

Figure 2.17 is used to further analyse the indicators that exhibit strong positive correlation. The $HI^{(1)}$ metric response is unique for the PPCA model, but all other metrics in the likelihood-based class respond similarly. Hence, using a single manifold-based metric from the three considered in this investigation would have been sufficient for condition inference. In the dissimilarity-based class, the $LHI_{PI}^{(4)}$ and $LHI_{PD}^{(5)}$ exhibit similar responses for the VAE model but not for the PPCA model. This indicates that the metrics in this class are unique and offer unique insight for this dataset. For metrics from the distribution-based class in the PI setting, the responsive metrics, $LHI_{PI}^{(6)}$ and $LHI_{PI}^{(7)}$, would exhibit strong positive correlation for both the PPCA and VAE models. This similarity indicates that choosing either of these metrics would have been suitable for condition inference on this dataset. However, Figure 2.17(c) shows that $LHI_{PI}^{(6)}$ is more responsive to the development of the gear tooth fault.

In Figure 2.19 the violin plots are shown to enable an inter-method analysis. From Figure 2.19, is clear that the classes with strong similarity are the dissimilarity-based metrics from the PPCA model, and the distribution-based metrics in the PI setting for the indicators from both models. The distribution-based metrics in the PD setting are outliers in the inter-method analysis, and this is attributed to their non-sensitivity to the presence of the gear tooth fault. This is distinct to the inter-method responses from the IMS dataset and highlights the uniqueness of the indicators proposed in this chapter and the requirement for an ensemble based latent inference procedure.



(a) PPCA method-wide correlation.

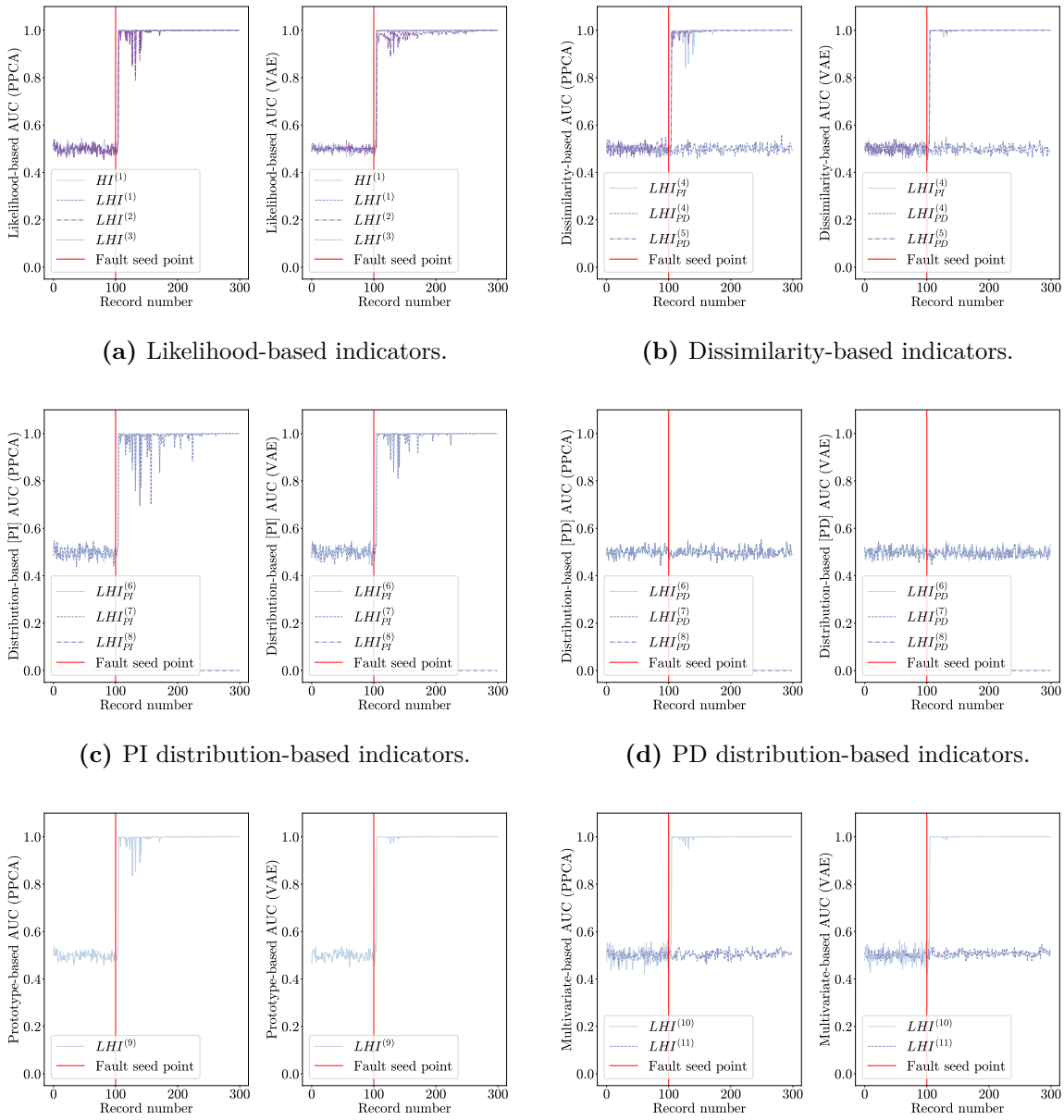
(b) VAE method-wide correlation.

Figure 2.19. Violin plots of the sample PCC on an inter-method level for the gearbox dataset. Note that the inter-method comparison is visualised on a per-class level as the r_{xy} terms are computed for all metrics of a given indicator class and all metrics from another LVM method to compare method-wide similarity.

In Figure 2.20 the AUC responses for the synchronous averaged indicator signals are shown using the approach described in Section 2.4.2.1. Figure 2.20 shows that there are responsive and unresponsive LHIs in the conceptual categories. It is evident that the PD setting is a poor indicator framework for this dataset, but if the PD setting is investigated under a different coordinate system then indicator responsiveness can be obtained through $LHI_{PD}^{(5)}$. This highlights the necessity to consider metrics from different conceptual categories as each represent a different manifold perspective.

Hence, from the results in Figure 2.16 and Figure 2.17, it is possible to detect the gear tooth fault present in the data and establish the point of fault development using the proposed HIs and LHIs. The TSA results ensure that the tooth fault can be detected and isolated, and the analysis methodology detailed in Section 2.4.2.1 ensures that the TSAs can be trended for quantitative performance quantification. The Euclidean-based PD LHIs were found to respond inadequately to the gear tooth fault, which indicates that the manifold response to the tooth fault is to drive data away from the healthy manifold.

By performing an inter-method and intra-class correlation analysis for the gearbox dataset: *(i)* the methods from the likelihood-based class that monitor the latent manifold offer similar insight, *(ii)* there is strong positive correlation for the responses for the metrics considered in this investigation between the PPCA and VAE models, *(iii)* the $LHI^{(6)}$ and $LHI^{(7)}$ metrics exhibit a similar response, and *(iv)* the PD setting is not useful for condition inference for the



(a) Likelihood-based indicators.

(b) Dissimilarity-based indicators.

(c) PI distribution-based indicators.

(d) PD distribution-based indicators.

(e) Prototypical indicators.

(f) Multivariate-based indicators.

Figure 2.20. The AUC for the synchronous average of the indicator signals around the gear tooth of interest. In (c) and (d) the AUC is zero as the $LHI^{(8)}$ metric for both models is insensitive to the tooth fault. An AUC of 0.5 indicates that the TSA of the HI and LHIs is indistinguishable from the indicators of the reference healthy set.

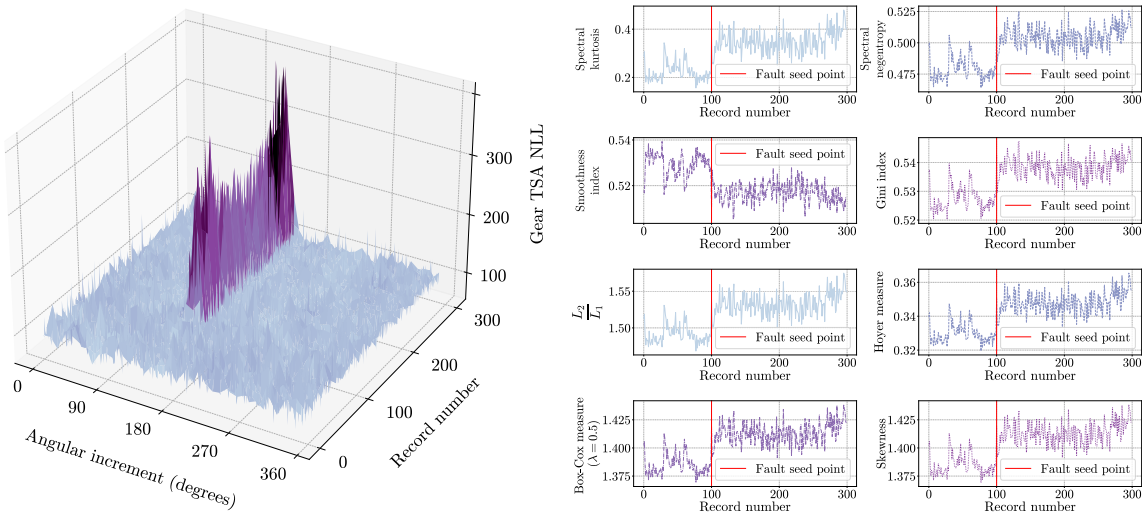
gearbox dataset. This dataset induces complexity through time-varying operating conditions, and the LVMs used in this thesis produce both a data space and a latent manifold that is practical for condition inference purposes. The proposed LHIs capture different perspectives of the latent manifold, and the quantification of the perspectives given in Section 2.3.4 ensures

that the LHIs considered are sufficient for manifold condition inference.

Finally, in a comparison between the indicator responses exhibited for the IMS and gearbox dataset, detailed in Figure 2.7 and Figure 2.17 respectively, the uniqueness of the indicators presented in this chapter is exhibited. Not only do many indicators respond to the damage present in the data, but the responsiveness of the indicators is revealed to be dataset dependent. Furthermore, for the gearbox dataset, indicators are sensitive for vibration data with clear non-stationary OCs. These results advocate for the use of an ensemble-based LVM indicator approach as the latent manifold of LVMs is dependent on the fault conditions, the environment and operating conditions, and the asset of interest. By utilising indicators from each of the proposed conceptual categories, a richer information stream for condition inference becomes accessible and gives rise to an improved vibration-based CM methodology.

2.4.2.3 Comparison

To compare the LHIs used in this investigation to HIs from other works, two HI methods were used. For the first method, the GMM-based NLL HI proposed in Heyns *et al.* [145] was used. Note that the GMM parameters were optimised using the same practices as for the $LHI^{(1)}$ metric. For the second method, the set of sparsity measures from Section 2.4.1.4 were calculated.



(a) TSA of the GMM NLL. **(b)** Sparsity indicators.

Figure 2.21. The GMM NLL metric from Ref. [145] and the sparsity metrics applied to the gearbox dataset.

In Figure 2.21(a) the results using the methodology from Ref. [145] are shown. The proposed LHIs, in comparison with the methodology from Ref. [145], are useful and capable of exhibiting a clear response to the gear fault present in this dataset. While the methodology proposed by Ref. [145] is designed for the detection of gear faults, the LHIs in this investigation exhibit similar responses without any directed pre-processing steps. Additionally, many of the LHIs exhibit a similar response to the GMM NLL, which further advocates for an ensemble-based condition inference approach as this allows for multiple diagnosis avenues.

The sparsity measures in Figure 2.21(b) are noticeably sensitive to non-stationary OCs not due to the gear tooth fault. While a change in condition is present around record 100, this may be attributed to the intervention step required to seed the tooth fault and the presence of the fault is not more pronounced in subsequent records. The lack of a clear degradation trend in the sparsity measures is attributed to the complexity of the gearbox dataset, and this highlights the necessity to consider focused post-processing practices to infer the fault type, as was done for the proposed LHIs. The considered sparsity metrics alone are not indicative enough to motivate asset maintenance.

2.5 Conclusion and recommendations

This chapter proposes a modular latent indicator framework for offline temporal-preserving latent variable models to address the first research direction identified in this thesis. The proposed framework identifies and classifies five unique approaches to characterise the potential avenues of fault manifestation in the latent manifold of offline temporal-preserving LVMs. The considered manifold perspectives are important as the latent manifold is unique to a given asset, and it is not possible to identify an optimal LHI or manifold perspective that generalises to all possible assets. Hence, fifteen LHIs are extracted from the proposed conceptual categories and are assessed in a fault diagnostics context. Two offline LVM formulations are considered that make explicit and tractable LVM distribution parametrisation assumptions to assess the impact of considering a latent manifold contained within a linear or a nonlinear subspace. The fifteen LHIs are applied to the two considered offline reconstruction-focused LVMs and the ability of the LHIs to enable effective fault diagnostics was investigated.

The considered indicators from the proposed latent indicator framework were applied to two experimental datasets: a bearing fault dataset with constant OCs and gear tooth damage with time-varying OCs. The results of this chapter indicate that not only is the latent

manifold sensitive to the presence of faulty signatures in the vibration data, but the LHIs can also be used to accurately infer the source of damage in the presence of constant and time-varying OCs. The performance of the considered LHIs was also compared to results obtained from other methodologies used in the literature. The comparison indicates that the considered LHIs are useful for the fault diagnosis task and capable of exhibiting a clear response to the fault signatures present in the considered experimental datasets.

Thus, the proposed latent indicator framework is effective for gearbox fault diagnostics and demonstrates that a temporal-preserving latent manifold is a useful representation for condition inference in temporal-preserving LVM-based CM methodologies. This addresses the focus of the first research direction concerned with extracting HIs and LHIs from the data space and the latent space to perform effective fault diagnostics with temporal-preserving LVMs. In the next chapter of this thesis, the construction of a temporal-preserving latent manifold and the methods used for this construction are examined by studying the implications of offline and online training approaches. The purpose is to consider gearbox fault diagnosis scenarios where no historical data from a healthy asset is available a priori.

Chapter 3

Training approaches for reconstruction-focused and interpretation-focused latent variable models

---

3.1 Introduction

In recent condition monitoring applications of latent variable models, e.g., Refs. [28, 164], model parameter estimation occurs via offline training. Offline training, used in Chapter 2 of this thesis to investigate the first research direction, assumes the existence of a historical dataset of vibration signals, which serves as a baseline estimate of an asset's condition. Conventionally, a baseline model is constructed from this historical dataset for the asset of interest, and the model is used to determine if novel vibration data is from the same baseline distribution [28]. However, recent works, e.g., Refs. [176, 182, 189, 191], have demonstrated that unsupervised learning-based techniques can extract fault signature information directly from vibration data without relying on historical information as an a priori estimate of the healthy asset condition. In this scenario, online training is used and the latent manifold is treated as a dynamic object through record number to extract the salient features from the vibration data. The anticipation of online training is that the salient latent features capture and enhance the underlying fault signatures within the observed vibration data and, ideally, enable effective gearbox fault diagnosis. Conventional research applications of latent variable models favour offline training, but online training can also be used.

In the second research direction, approaches used to construct the latent manifold are investigated to consider scenarios where historical data is not available from a healthy asset. Such scenarios can invalidate conventional research applications of latent variable models and present a dilemma in the approach followed in Chapter 2 of this thesis. Therefore, the implications of using offline and online training for latent variable models are considered and

investigated in this chapter in the fault diagnosis context. Specifically, an online training approach, which can also benefit from the temporal preservation approach discussed in Section 2.2, is applied to the latent variable model-based fault diagnosis problem, and compared to results from the offline training setting used in Chapter 2. Furthermore, the methods considered during online training are extended to include two main methods: reconstruction-focused and interpretation-focused methods. This facilitates the exploration of different approaches to construct a dynamic latent manifold.

Online training of temporal-preserving LVMs is important for scenarios where relying on reference vibration data is challenging or the data is not be available a priori. Employing online training mitigates the requirement for temporal-preserving LVMs to suitably capture the reference data distribution. Additionally, online training empowers temporal-preserving LVMs with the capability to dynamically characterise the observed vibration data, potentially enabling better extraction and isolation of important latent diagnostic information contained within the data. Moreover, online training presents an approach that shares many characteristics with signal processing methodologies. It allows temporal-preserving LVMs to operate as a non-parametric signal enhancement methodology that extracts salient diagnostic information related to the instantaneous health state of an asset, as dictated by the model objectives. However, it is not suggested that offline temporal-preserving LVMs lack value. Instead, the objective is to enable a comparison and study on the implication of dynamic latent expansion coupled with the chosen LVM approach. These approaches are driven by the choice of online method, i.e., reconstruction-focused or interpretation-focused methods.

Consequently, the purpose of this investigation for the second research direction is two-fold: First, the scenario where the latent manifold can dynamically vary through record time is investigated using online models and the results are compared to the offline training approach to determine the potential of this additional manifold flexibility. Temporal preservation forms the groundwork for this investigation, enabling latent manifold interpretability via latent temporal structure, and basic methods from Chapter 2 of this thesis are adapted to facilitate the training approach comparison. Second, the underlying assumption in offline training is that the historical data represents an asset in a healthy condition [28, 145, 317]. Hence, the scenario where the distribution of the healthy data is not known a priori through a reference dataset is considered and investigated.

The structure of this chapter is as follows: First, the considered latent variable model

formulations and latent health indicators used in this chapter are detailed and discussed in a vibration-based condition monitoring context in Section 3.2. Then, the considered offline and online formulations are evaluated and compared using a numerical gearbox dataset in Section 3.3 and using two experimental datasets in Section 3.4. Finally, the chapter is concluded in Section 3.5.

3.2 Offline and online latent variable methods

The purpose of this chapter is to study the implications of online training for Latent Variable Models (LVMs) in a fault diagnostics context. Models developed for the fault diagnosis task exhibit variations in both the training and evaluation stages based on whether offline or online training is used. Offline training uncouples the model training and evaluation stages as the model parameters are kept static through record time, and novel data is evaluated using health indicators obtained from the trained model. Alternatively, online training couples the model training and evaluation stages, whereby the estimation of the model parameters is dynamic through record time and the evaluation step is performed simultaneously. Online training underlies many signal processing strategies, e.g., blind deconvolution [264, 318], band demodulation [32, 190, 319], and online monitoring [320, 321], but is less common in temporal-preserving latent variable model-based applications and makes it possible to use and apply a new class of algorithms to the fault diagnosis task.

3.2.1 Overview

To facilitate a comparison between the offline and online LVM training approaches as a component of the second research direction, a decision must be made regarding the LVM formulations used. In light of the LVM formulations used in Chapter 2, a Probabilistic Principal Component Analysis (PPCA) model [280] is used in this investigation to represent the offline LVM trained on only reference, assumed healthy, historical vibration data. Henceforth, this model is referred to as the PPCA_h model. Offline training is considered to establish a baseline for result comparison as it represents the conventional LVM application process. In online training, the evolution of intrinsic data features through record time is of interest, and the latent features are a direct function of the LVM formulation.

Two main categories of online temporal-preserving LVM formations are investigated and compared in this study, namely *(i)* reconstruction-focused LVMs and *(ii)* interpretation-focused LVMs. Reconstruction-focused LVMs seek to maximise the variance explained in

some observed vibration data via the latent manifold for high-fidelity sample generation purposes [28, 166]. The latent manifold is driven to capture the dominant components in the observed data for minimal data reconstruction error, e.g., PPCA and Variational Auto-Encoders (VAEs) [101, 165]. Alternatively, interpretation-focused LVMs seek to obtain an informative latent decomposition for some observed vibration data [166, 178]. Such LVMs seek a maximally informative latent manifold driven via statistical independence and non-Gaussianity. Statistical independence implies that the individual latent variables are maximally informative, i.e., provide no useful information to other latent variables, and non-Gaussianity is a key requirement both to enforce statistical independence and to obtain proxy measures for the features of the latent variables to guide the LVM parameter estimation process that makes it easier to interpret the manifold [185, 322, 323]. The LVM-based Condition Monitoring (CM) pipeline and the aspects of investigation in this chapter are highlighted in Figure 3.1. Firstly, the choice of LVM training approach is highlighted and the implication placed on the LVM parameters is denoted. Secondly, the investigation into the choice of online LVM formulation is highlighted as this has implications on the extracted latent information when using the online training approach.

The decision to consider LVMs from both categories when using online training is made as maximal variance may not necessarily represent the most useful information for the fault diagnostics task. Additionally, non-Gaussianity is a popular measure of informativeness for fault signature interpretability in the signal processing literature and has attracted much attention from the academic community to detect damage in rotating machinery [30, 31, 264]. These online methods operate on each sequential data observation and represent different techniques for obtaining an informative data transformation by maximising the non-Gaussianity of the transformed data. Hence, the applicability of different LVM formulations is investigated for the online training approach to establish and compare their suitability in the LVM-based fault diagnostics field. However, this does not necessarily imply that reconstruction-focused LVM formulations lack utility; instead, this investigation may emphasise the complementarity of the characteristics from the considered online LVM formulations.

In this investigation, PPCA is used to represent an online reconstruction-focused LVM, while an Independent Component Analysis (ICA)-based methodology is used to represent an online interpretation-focused LVM. A Spectrally Regularised Independent Component Analysis (SRICA) model is used that includes a spectral regularisation term to combat latent

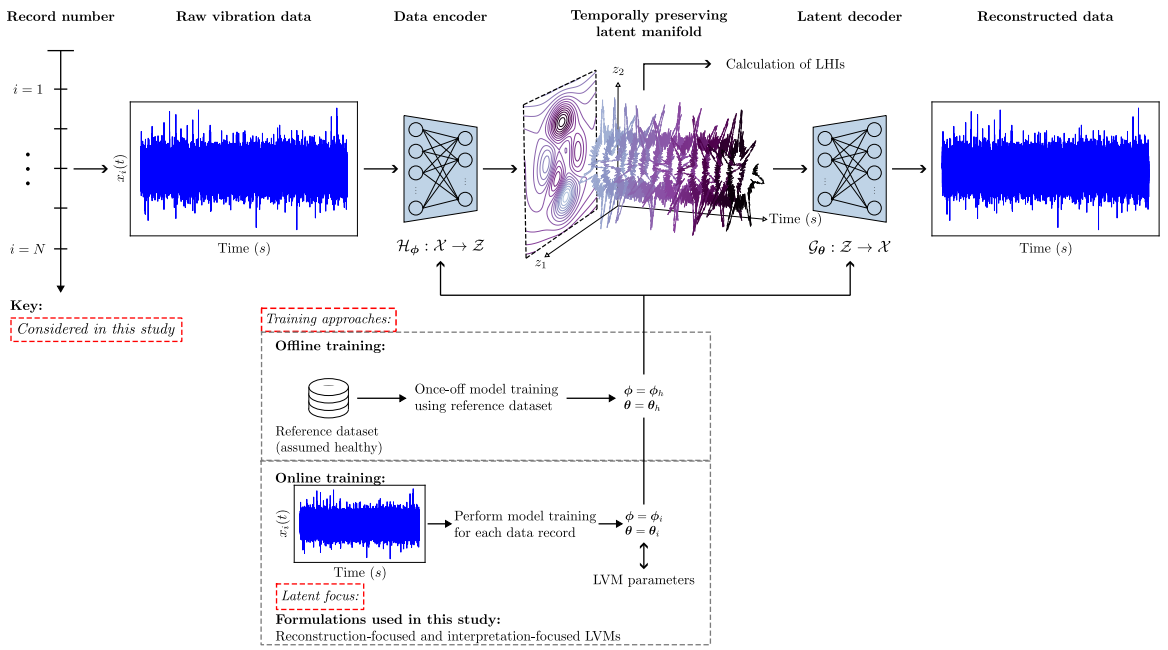


Figure 3.1. The LVM-based CM pipeline in a temporally preserved LVM application setting. Note that no data pre-processing is considered and the latent manifold is set to $\mathbb{R}^2$ with an explicitly disentangled time variable, i.e., sensor time. The choice of LVM training approach directly affects the LVM parameters, whereby for offline training, the parameters are estimated using some reference healthy data, whereas for online training the parameters are iteratively estimated using the i^{th} vibration data record. The subscript of the variable h for the offline training approach and the variable i is reserved specifically for the dependency on record time in the online training approach to simply notation in the remainder of this chapter.

information duplication in negentropy-based ICA models applied to single channel vibration data [178, 322, 324]. The spectral regularisation term used to combat latent information duplication is presented in Appendix C. The principle of this regularisation term is to ensure that the important latent variables are identified without repetition to reduce computational efforts and improve salient feature identification for negentropy-based ICA methodologies. As demonstrated in Chapter 2, Latent Health Indicators (LHIs) are computed directly from the latent manifold, and in the online training approach the LHI quality is directly dependent on the ability of the latent manifold to represent the fault signatures. Hence, the objective is to compare the implication of the chosen LVM formulations in the online training setting against an LVM developed using offline training.

3.2.2 Introduction

LVMs represent a probabilistic framework to describe the data distribution $p(\mathbf{x})$, where $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^D$, using some unobserved latent variables $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^d$, where $d \leq D$. The marginal distribution $p(\mathbf{x})$, often referred to as the model evidence, is given by Equation (2.2) and represents the data distribution as a function of the marginalisation over the latent variables. In common LVM applications, the joint distribution $p(\mathbf{x}, \mathbf{z})$ is assumed to factorise, as demonstrated in Equation (2.1), into a product of the conditional generative distribution $p(\mathbf{x}|\mathbf{z})$ and the prior distribution $p(\mathbf{z})$ over the latent variables. To infer the posterior distribution $p(\mathbf{z}|\mathbf{x})$ for the latent variables given some observed data, Bayes' rule is used, as demonstrated in Equation (2.3).

Different strategies exist to represent the LVM formulation, each of which constitute a unique interpretation of the general LVM objective. Typically, an encoding transition function $\mathcal{H}_{\boldsymbol{\phi}} : \mathcal{X} \to \mathcal{Z}$ and a decoding transition function $\mathcal{G}_{\boldsymbol{\theta}} : \mathcal{Z} \to \mathcal{X}$ are used to parametrise the mapping from $\mathbf{x}$ to $\mathbf{z}$ and vice versa. These functions represent the transfer of information between the data space and the latent manifold, and the parameters of these functions are estimated to allow the LVM to perform data generation and/or capture the intrinsic properties in the data [101]. In this chapter, the encoding and decoding transition functions are developed either using offline or online training approaches, as presented in Figure 3.1, and thus they are dependent on a reference dataset or record time and are denoted by $\mathcal{H}_{\boldsymbol{\phi}_h}$ and $\mathcal{G}_{\boldsymbol{\theta}_h}$ or by $\mathcal{H}_{\boldsymbol{\phi}_i}$ and $\mathcal{G}_{\boldsymbol{\theta}_i}$, respectively.

In explicit LVM formulations, assumptions are given to the form of the distribution of interest for explicit parametrisation, e.g., PPCA [280] that assumes linear Gaussian distributions, or VAEs that assume nonlinear Gaussian distributions [165]. Note that the linearity of the model refers to the parametrisation of the encoding and decoding transition functions used for mapping from $\mathbf{x}$ and $\mathbf{z}$ and vice versa, not the form of the chosen distribution. In implicit LVM formulations, a stochastic procedure is defined to directly generate data and no distribution assumptions are made except for the form of the prior $p(\mathbf{z})$, e.g., Generative Adversarial Networks (GANs) [167, 168]. Typically, LVMs are reconstruction-focused, whereby they are primarily concerned with sample generation, and much research is focused on methods that produce highly expressive samples, e.g., diffusion models [325], or improving the flexibility of the posterior $p(\mathbf{z}|\mathbf{x})$, or the approximation thereof [161, 162], for improved sample generation[326], e.g., Ref. [327]. In the reconstruction-focused setting, the latent variables are driven

to capture the prominent sources of variation in the data for improved generation fidelity. Interpretation-focused methods, alternatively, focus on finding informative decompositions via the latent variables $\mathbf{z}$. ICA is a popular methodology in this category, which seeks to obtain a latent manifold that is statistically independent and non-Gaussian. Certain works, e.g., β-VAEs [328] or InfoGANs [329] have proposed techniques to drive reconstruction-focused methods towards an informative latent decomposition that is statistically independent.

However, in the reconstruction-focused setting, the informative decompositions are primarily driven to be similar to an isotropic Gaussian distribution, while in the ICA-based setting the latent distribution is explicitly assumed to be non-Gaussian. The aversion to Gaussianity for interpretation-focused methods is motivated via the affine property of a Gaussian distribution. Consider the linear transformation of Gaussian random variable $\mathbf{v} = \mathbf{R}\mathbf{u}$, where $\mathbf{u} \sim \mathcal{N}(\mathbf{u}|\mathbf{0}, \boldsymbol{\Sigma})$ with covariance $\boldsymbol{\Sigma} = \mathbf{I}$ and let $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ be a rotation matrix. The transformed Gaussian distribution $\mathbf{v} \sim \mathcal{N}(\mathbf{v}|\mathbf{0}, \mathbf{R}\boldsymbol{\Sigma}\mathbf{R}^T) = \mathcal{N}(\mathbf{v}|\mathbf{0}, \boldsymbol{\Sigma})$ is rotation invariant [101]. Hence, it is impossible to distinguish between latent variables that deviate by a simple latent transformation, and there exist infinitely many combinations of Gaussian latent variables that satisfy the independence property [101]. In this chapter, two linear LVM formulations are considered as representative methods from the reconstruction-focused and interpretation-focused LVM setting. Each method will be briefly discussed in turn. Note that the PPCA$_h$ model used to represent a LVM developed using the offline training approach is considered a reconstruction-focused LVM.

3.2.2.1 Reconstruction-focused latent variable models

The first LVM used in this investigation is PPCA, wherein Principal Component Analysis (PCA) operates as part of the maximum likelihood solution for the LVM formulation [101]. This model is used as a reconstruction-focused method LVM developed using both offline and online training. For the PPCA model, the prior distribution is assumed to be $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I})$, which represents an isotropic Gaussian distribution. The generative process for PPCA is expressed as

$$\mathbf{x} = \mathbf{A}\mathbf{z} + \boldsymbol{\mu}_{\mathbf{x}} + \boldsymbol{\epsilon}, \tag{3.1}$$

where $\boldsymbol{\mu}_{\mathbf{x}}$ represents the first moment of $\mathbf{x}$, $\mathbf{A}$ is the generative transformation matrix, and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2_{PPCA} \cdot \mathbf{I})$ is a zero-mean isotropic Gaussian variable. Note that Equation (3.1) implies that the generative distribution $p(\mathbf{x}|\mathbf{z})$ follows a Gaussian distribution. The posterior

distribution $p(\mathbf{z}|\mathbf{x})$ for the PPCA model is given by

$$p(\mathbf{z}|\mathbf{x}) = \mathcal{N}\left(\mathbf{z}|\mathbf{M}_{PPCA}^{-1}\mathbf{A}^T\left(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}}\right), \sigma_{PPCA}^{-2}\mathbf{M}_{PPCA}\right), \tag{3.2}$$

where $\mathbf{M}_{PPCA} = \mathbf{A}^T\mathbf{A} + \sigma_{PPCA}^2 \cdot \mathbf{I}$. The maximum likelihood solution for $\mathbf{A}$ and σ_{PPCA}^2 for the PPCA model, presented in Ref. [280], gives the latent encoding transformation as

$$\mathbf{z} = \text{diag}\left(\frac{\sqrt{\lambda_1 - \sigma_{PPCA}^2}}{\lambda_1}, \cdots, \frac{\sqrt{\lambda_d - \sigma_{PPCA}^2}}{\lambda_d}\right)\mathbf{U}_d^T\left(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}}\right), \tag{3.3}$$

where $\mathbf{U}_d \in \mathbb{R}^{D\times d} = [\mathbf{u}_1, \cdots, \mathbf{u}_d]$ represents the eigenvector matrix of the covariance matrix $\mathbf{C} = \mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{\mathbf{x}\mathbf{x}^T\}$ arranged in descending order based off the eigenvalues λ_j of $\mathbf{C}$, where $j \in [1, \cdots, d]$. Equation (3.3) represents the mean of $p(\mathbf{z}|\mathbf{x})$ and can be identified as a scaled version of the PCA latent transformation $\mathbf{z} = \mathbf{U}_d^T\left(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}}\right)$ [101]. In the offline training approach, the PPCA$_h$ model parameters are obtained using a reference set of observed vibration data signals. For the online training approach, the PPCA model parameters are obtained for each sequentially ordered vibration signal $x_i(t)$, $i \in [1, N_o]$, where N_o is the potential number of vibration signals.

3.2.2.2 Interpretation-focused latent variable models

The second LVM used in this investigation is an ICA methodology based off the negentropy-based ICA formulation proposed in Hyvärinen [322]. This model is used as an interpretation-focused LVM developed using online training. The noiseless ICA generative model is given by

$$\mathbf{x} = \mathbf{A}\mathbf{z}, \tag{3.4}$$

where $\mathbf{A} \in \mathbb{R}^{D\times d}$. In this ICA formulation, the latent variables are enforced to be independent, i.e.,

$$p(\mathbf{z}) = \prod_{j=1}^{d} p(z_j) \tag{3.5}$$

and the latent variables are assumed to be non-Gaussian [178]. The encoding transformation to obtain the latent variables is represented by

$$\mathbf{z} = \mathbf{W}\mathbf{x}, \tag{3.6}$$

where $\mathbf{W} \in \mathbb{R}^{d\times D} = [\mathbf{w}_1, \cdots, \mathbf{w}_d]^T$ represents a matrix of component vectors. Conventionally, $\mathbf{A}$ is unknown, and ICA methodologies exploit proxy criteria to represent the ICA transform without assuming the probabilistic form of the generative model [178]. A popular ICA methodology is to use an approximation of negentropy [330] as a measure of non-Gaussianity to estimate the ICA model parameters. In the methodology proposed in Appendix C, the

issue of source duplication in single channel time-series applications, discussed in the work of Hyvärinen[178] and Davies and James [324], is addressed via the addition of a spectral regularisation term that enforces the spectral representations of the vectors in $\mathbf{W}$ to be orthogonal. The methodology in Appendix C uses a regularised optimisation algorithm applied to the negentropy approximations from [330] to estimate the SRICA model parameters. To approximate negentropy, the non-quadratic exponential function is used in this investigation as it is a robust estimator of negentropy [178, 322]. In this chapter, the SRICA model is used to represent an interpretation-focused LVM developed using the online training approach and is compared to PPCA within the context of the online training approach and to the PPCA$_h$ model in the context of the offline training approach. In Section A.5, the relationship between the reconstruction-focused and interpretation-focused LVM formulations and signal processing methodologies is discussed. The data pre-processing steps used in this investigation are discussed in the following section.

3.2.3 Data pre-processing

Given a signal $x_i[n]$, where $n \subset \mathbb{Z}^+ = \{1, \cdots, L_{signal}\}$ and L_{signal} is the length of the signal, an initial pre-processing step is required to produce a representation of the raw vibration signal that enables LVM parameter estimation for LVM-based CM. In this investigation, a data Hankelisation step, sometimes referred to as a truncated Toeplitz matrix, is used to obtain a matrix $\overline{\mathbf{X}}_i \in \mathbb{R}^{L_T \times L_w}$ and is given in Equation (2.13). The rows of $\overline{\mathbf{X}}_i$ represent samples of the data vector $\mathbf{x} \in \mathbb{R}^{L_w}$, i.e., $D = L_w$. Note that the pre-processing steps required for LVMs developed using offline training are discussed in Section 2.3.3. For the remainder of this discussion, the pre-processing steps related to the online training approach are presented. After the Hankelisation step, a standard pre-processing strategy that follows is to centre the columns of $\overline{\mathbf{X}}_i$, which ensures that the components of $\mathbf{x}$ are zero-mean. This is given by

$$\mathbf{X}_i = \overline{\mathbf{X}}_i - \mathbf{1}\boldsymbol{\mu}_i^T, \tag{3.7}$$

where $\boldsymbol{\mu}_i \in \mathbb{R}^{L_w}$ is a column vector of the feature-wise means of $\overline{\mathbf{X}}_i$ and $\mathbf{1} \in \mathbb{R}^{L_w}$ is a constant vector with elements 1. Note that the data is not standardised in this chapter as standardisation coupled with online training may remove any inter-record variance that may be useful for condition inference. For the SRICA methodology, a pre-whitening step is used to further pre-process the data matrix $\mathbf{X}_i$. Pre-whitening is a useful processing step to reduce the approximation complexity required to obtain the vectors $\mathbf{w}$ of the independent component matrix $\mathbf{W}$ presented in Equation (3.6) [185]. Pre-whitening is a pre-processing strategy that

removes any second-order correlations in $\mathbf{X}_i$ and is given by a linear transformation via

$$\tilde{\mathbf{X}}_i = \mathbf{X}_i \mathbf{U}_D \mathbf{L}_D^{-1/2} \mathbf{U}_D^T, \tag{3.8}$$

where $\tilde{\mathbf{X}}_i \in \mathbb{R}^{L_T \times L_w}$ is the whitened matrix, $\mathbf{U}_D \in \mathbb{R}^{L_w \times L_w}$ is the full rank eigenvector matrix of the data covariance matrix $\mathbf{C}_i = \mathbb{E}\{\mathbf{x}_i\, \mathbf{x}_i^T\}$, and $\mathbf{L}_D^{-1/2} \in \mathbb{R}^{L_w \times L_w} = \text{diag}\left(\lambda_1^{-1/2}, \cdots, \lambda_{L_w}^{-1/2}\right)$ is a diagonal matrix of the inverse square root of the eigenvalues λ from $\mathbf{C}_i$ [101]. To enable fault diagnostics for the interpretation-focused LVM, the implications of the pre-whitening step must be accounted for as this limits the ability to perform fault trending through record time. The implication of pre-whitening is discussed in Ref. [185]. Note that PPCA has no such issue as the centring process given in Equation (3.7) does not remove any scale variations in the data. In this investigation, a simple re-scaling of the latent components is used for the interpretation-focused LVM via the latent kurtosis i.e.,

$$\mathbf{z} = \text{diag}\left(\sqrt{\text{kurt}\,(z_1)}, \cdots, \sqrt{\text{kurt}\,(z_d)}\right) \mathbf{W}_i \tilde{\mathbf{x}}, \tag{3.9}$$

where $\text{kurt}(\cdot)$ represents the sample kurtosis and $\tilde{\mathbf{x}}$ represents a column vector stored in a row of the $\tilde{\mathbf{X}}_i$ matrix. This re-scaling ensures that the prominence of specific underlying features in the latent manifold can be compared through record time. The use of the square root was made based off its usage in signal processing works, e.g., Ref. [269], and in this investigation, it was found that this scaling gave better results. The choice of scaling used is not proposed to be optimal, and hence may require further investigation. The optimisation of the scaling function is not within the scope of this investigation.

3.2.4 Latent health indicators

To infer the condition of the asset of interest based off a sequential set of vibration signals, a transformation of the latent space and reconstructed data space is required for CM purposes. In this investigation, a simple set of statistical distance-based LHIs from Section 2.3.4 in Chapter 2 are used. The dissimilarity-based sub-class is used in this investigation as it requires no access to a set of healthy latent manifold representations, and the health indicators are simple to calculate. The data space Health Indicator (HI) is the reconstruction error

$$HI^{(1)} = \frac{1}{D} \cdot \|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2, \tag{3.10}$$

which represents the mean squared error between the data sample $\mathbf{x}_t$ and its reconstruction $\widehat{\mathbf{x}}_t$. Note that the subscript $t \in [0, L_T - 1]$ refers to an integer representation of the signal time captured by the signal segments in Equation (2.13). The first LHI considered in this

investigation is the normalised Euclidean distance indicator given by

$$LHI_{PI|PD}^{(1)} = \frac{1}{d} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_2 \tag{3.11}$$

where the Path Independent (PI) and Path Dependent (PD) latent manifold inference settings capture different latent manifold deviation mechanisms, $\mathbf{z}_t = \mathbf{0}$ in the PI setting, and d is the dimensionality of $\mathbf{z}$. For LVMs developed using the offline training approach, the dimensionality of the latent manifold is constant and denoted as $d = d_h$. For LVMs developed using the online training approach, the latent manifold may be a function of the observed vibration data and is denoted as $d = d_i$. The second LHI considered in this investigation is the normalised Manhattan distance

$$LHI_{PI|PD}^{(2)} = \frac{1}{d} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_1, \tag{3.12}$$

which evaluates the L_1 norm. The third LHI considered in this investigation is the normalised Canberra distance

$$LHI_{PD}^{(3)} = \frac{1}{d} \cdot \sum_{j=1}^{d} \frac{|z_{t+1,j} - z_{t,j}|}{|z_{t+1,j}| + |z_{t,j}|} \tag{3.13}$$

which represents a weighted version of Equation (3.12). The fourth LHI considered in this investigation is

$$LHI_{PI|PD}^{(4)} = \frac{1}{d} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_\infty, \tag{3.14}$$

which represents the normalised Chebyshev distance or uniform norm $\|\cdot\|_\infty$. The final LHI considered in this investigation is the cosine distance

$$LHI_{PD}^{(5)} = \frac{1}{d} \cdot \cos^{-1}\left(\frac{\mathbf{z}_{t+1}^T \mathbf{z}_t}{\|\mathbf{z}_{t+1}\|_2 \cdot \|\mathbf{z}_t\|_2}\right), \tag{3.15}$$

which represents a latent dimensionality normalised LHI in a non-Euclidean coordinate system. The $HI^{(1)}$ and $LHI_{PI|PD}^{(1-5)}$ metrics are used to infer instantaneous asset condition and are representative of deviations in the data space and the latent manifold in the presence of some fault condition.

Discussion on the computational efforts of offline versus online methods

The computational aspect of the considered offline and online methods is an important aspect to discuss. Offline methods exhibit greater computational efficiency as they require once-off parameter estimation before model evaluation. Online methods, on the other hand, involve an additional estimation step for each vibration data record as the model estimation and evaluation are coupled. Throughout the investigations performed in each chapter of this thesis, the PPCA model is consistently employed. This decision was based on the model's well-documented

usefulness in the literature and its ability to operate as a sensible performance baseline [101]. The PPCA model has significant computational advantages as it can be efficiently computed using either the singular value decomposition of $\mathbf{X}$ or an eigendecomposition of $\mathbf{X}^T\mathbf{X}$. Hence, in both offline and online training, this reconstruction-focused LVM is highly computationally efficient [182].

In contrast, interpretation-focused methods have increased complexity in the parameter estimation step that requires more computational effort. The choice of approach for interpretation-focused methods introduces computational dependencies that can vary significantly depending on the approach employed. Moreover, deflation-based latent variable optimisation loops, i.e., optimisation loops per latent component, are often necessary to estimate the model parameters. These requirements introduce further computational complexity to interpretation-focused methods. One computational advantage of online methods is the embarrassingly parallel nature of the per-record parameter estimation step, which allows for batched parameter estimation through record time. Nonetheless, it is important this thesis does not primarily focus on maximising the computational efficiency of LVM methods, as this would be better suited to a computation-focused study. While certain computational aspects were refined to ensure that suitably sufficient computational speeds, further efficiency maximisation steps were not adopted and was left for future work.

3.3 Numerical investigation

The phenomenological gearbox model used by Abboud *et al.* [253] is used in this chapter to analyse the performance of the considered offline and online temporal-preserving LVMs in a fault diagnosis setting. The purpose is two-fold. First, the performance and efficacy of the offline and online models under consideration can be compared and validated utilising a representative numerical dataset that guarantees control of the dataset properties. This approach facilitates a comprehension of the considered LVM implications in a detailed manner. Second, the numerical dataset provides explicit access to the underlying vibration data components. This accessibility facilitates a practical investigation into how the considered LVMs construct their respective latent manifolds. Access to the true vibration data components enables a targeted investigation into how the reconstruction-focused and interpretation-focused model objectives decompose and represent the observed vibration data in the latent manifold in offline or online training scenarios.

The phenomenological model generates a numerical gearbox signal $x_{ng}(t)$ using

$$x_{ng}(t) = x_{gm}(t) + x_{dp}(t) + x_n(t) + x_{bo}(t), \tag{3.16}$$

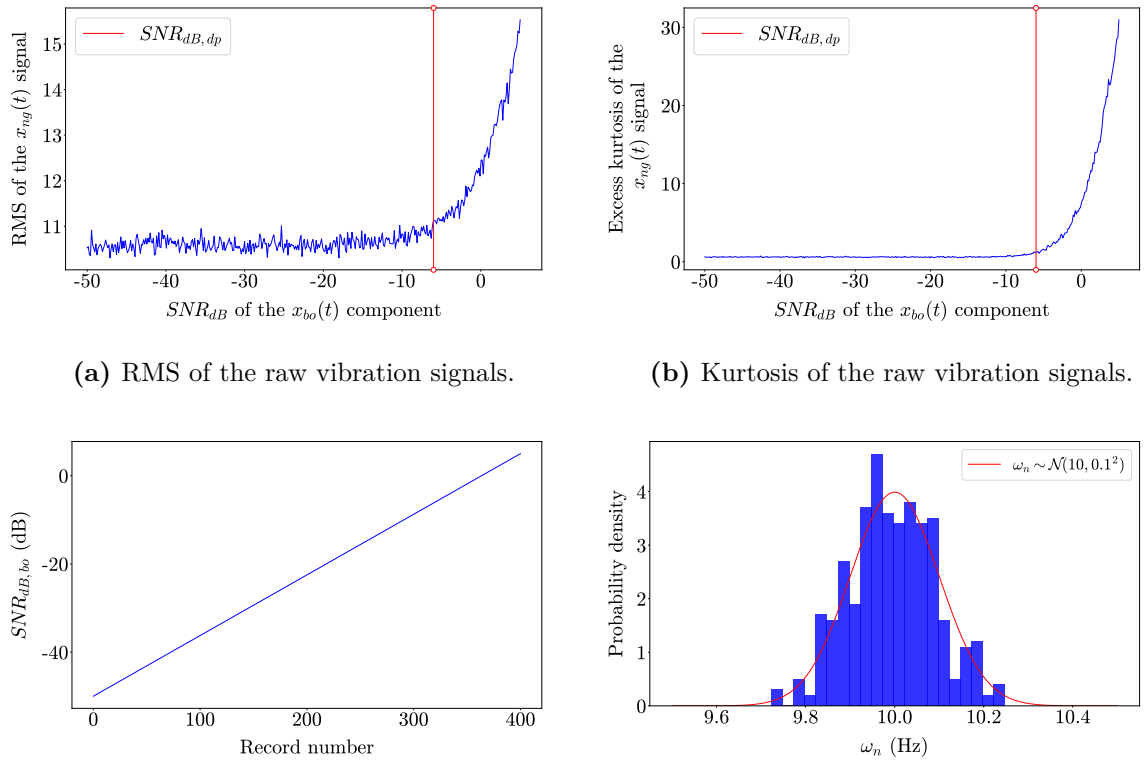
where $x_{gm}(t)$ represents the gear mesh component, $x_{dp}(t)$ represents the distributed pinion damage component, $x_{bo}(t)$ represents the outer race damage component, and $x_n(t)$ represents the noise component. More information related to the model can be found in Appendix B. The Ball Pass Frequency for an Outer race fault (BPFO) is set to 3.57 shaft orders for this investigation. The prominence of the $x_{bo}(t)$ component is gradually increased in the presence of a prominent $x_{gm}(t)$ component and a random second-order cyclostationary component $x_{dp}(t)$, to create a numerical dataset that enables a LVM sensitivity analysis investigation and a latent manifold investigation.

In Figure 3.2 the properties of the numerical dataset are visualised. The Root-Mean-Square (RMS) and excess kurtosis, seen in Figure 3.2(a) and Figure 3.2(b), respectively, indicate that a change in condition can only be detected after the Signal-to-Noise Ratio (SNR) of the $x_{bo}(t)$ component exceeds the $x_{dp}(t)$ component. The SNR of the $x_{bo}(t)$ component was linearly increased relative to the noise component $x_n(t)$ over the record number, as seen in Figure 3.2(c). The range of SNR variation represents a degradation path from a healthy bearing to an unhealthy bearing. To introduce operating condition variability, the shaft speed $\omega_n(t)$ was randomly sampled for the raw vibration signals, as demonstrated in Figure 3.2(d). The line colour in Figure 3.2(d) represents the record number.

3.3.1 Sensitivity analysis

In this section, the performance of the considered offline and online LVMs is compared. The prominence of the $x_{bo}(t)$ is incrementally increased over record number, to represent the experimental lifespan of a gearbox. The statistical properties of the $x_{gm}(t)$ and $x_{dp}(t)$ are kept constant with respect to $x_n(t)$. To develop the considered online LVMs, a window length of $L_w = 512$ and a shift parameter of $L_{sft} = 1$ was used to develop the data matrix $\mathbf{X}_i$ from the raw vibration signals using Equation (2.13). This ensures that the LVMs are temporally preserving and transfer the temporal characteristics of the raw signals to the HIs and LHIs considered for this investigation [164]. For the PPCA and PPCA$_h$ models, the Cumulative Contribution Rate (CCR)

$$CCR = \frac{\sum_{j=1}^{d} \lambda_j}{\sum_{j=1}^{D} \lambda_j}, \tag{3.17}$$



(a) RMS of the raw vibration signals.

(b) Kurtosis of the raw vibration signals.

(c) SNR of the outer race component.

(d) Shaft speed distribution.

Figure 3.2. An outline of the properties of the numerical dataset. (a) demonstrates the Root-Mean-Square (RMS) statistic of the raw vibration signals, (b) demonstrates the excess kurtosis of the raw vibration signals, (c) demonstrates the procedure of increased x_{bo} prominence in the raw vibration signals via the Signal-to-Noise Ratio (SNR), please refer to Equation (B.6) in Appendix B for more information, and (d) demonstrates the speed variation induced in the model.

is used to define the size of the latent space d_h and d_i, respectively. The CCR defined in Equation (3.17) represents the fraction of preserved data variance relative to the total data variance [101]. The CCR is set to 80% for this numerical investigation to ensure that the models capture a large portion of the variance in the vibration data. For the PPCA model, the data matrix $\mathbf{X}_i$ is centred using Equation (3.7) prior to the parameter estimation step. For the SRICA model, both feature centring and pre-whitening, given by Equation (3.7) and Equation (3.8), respectively, were used to pre-process the data matrix $\mathbf{X}_i$ prior to parameter estimation and the dimensionality of the latent manifold is set to $d = 30$. This dimensionality assumes that the observed data can be represented in a linear subspace that is less than 20% of the size of the data space D. For the PPCA$_h$ model, the first 10% of the available data was used as constituents of the reference dataset. The model parameters used for the numerical

dataset investigation performed in this chapter are summarised in Appendix A.2.

To quantify the performance of the considered models, two diagnostic metrics are used. Firstly, the RMS of the $HI^{(1)}$ and $LHI^{(1-5)}$ signals is computed. The RMS is a common statistical feature for CM applications for asset condition trending [19, 331]. Secondly, to infer the prominence of the fault condition present in the numerical dataset, the amplitude of the BPFO component in the Square Envelope Spectrum (SES) of the considered indicators is monitored. This ensures that the change in asset condition detected by the RMS features can be isolated and the cause of condition deviance can be inferred. The metrics considered in this work are dimensionalised, and hence the metrics are non-dimensionalised by standardising using the first 40 records available in the numerical gearbox dataset [311, 312]. This ensures that the responsiveness of the considered indicators can be critically compared without scale bias. A Condition Deviance (CD) point is identified for the considered diagnostic metrics using an alarm condition interval T_a calculated using

$$T_a = \mu_{m,a} \pm 3 \cdot \sigma_a, \tag{3.18}$$

where the median $\mu_{m,a}$ and the standard deviation σ_a of the thresholds are estimated using a percentage of the diagnostic metric values related to the asset in a healthy condition. An abnormal CD point outside of the condition interval is identified when the average of five consecutive metrics exceeds the bounds of T_a.

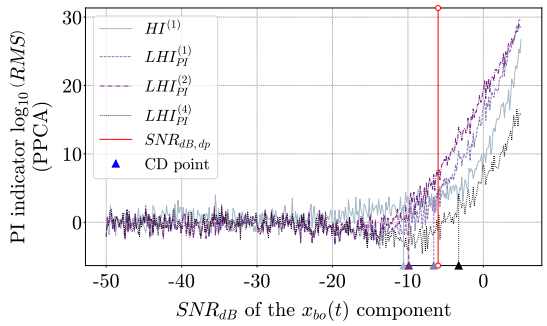
In Figure 3.3 the standardised $\log_{10}(RMS)$ of the temporal indicators is visualised against the SNR of $x_{bo}(t)$ component for the numerical dataset. The RMS metric was log-scaled to improve result interpretability and standardised to enable result comparison. Diagnostic metric standardisation was performed using an estimate of the metric mean and variance from the first 40 available indicator signals. The RMS trend results and the CD identification points indicate that, for the considered online LVMs, unique indicator responses are obtained, which exhibit sensitivity to an increase in the prominence of x_{bo}. The identification of the onset to bearing damage is a function of the choice of indicator and path traversal setting, e.g., $LHI^{(1)}_{PI|PD}$, thereby indicating that the indicator responses are unique. This highlights the need to consider an ensemble of LHIs for LVM-based CM. The $HI^{(1)}$ metric from the PPCA model exhibits the earliest onset to damage in the PI setting, while in the PD setting the $LHI^{(2)}$ exhibits early damage sensitivity. Additionally, the PD indicators from the PPCA model all have earlier identified damage onsets than their PI counterparts.

Figure 3.3 indicates that the SRICA model exhibits indicator sensitivity similar to the PPCA model. However, the indicators respond uniquely when used in the different frameworks, which can be interpreted both from the trends and their respective CD points, thereby highlighting the interdependency between the characteristics of the considered online LVMs and the condition indicators. The $LHI^{(4)}_{PI|PD}$ latent indicators have the earliest onset to damage for the SRICA model, while the $LHI^{(3)}$ metric is the least indicative of damage. As the condition indicators from both online LVM are sensitive to the onset of damage, it indicates that both the reconstruction-focused and interpretation-focused LVM setting developed using the online training approach are suitable for fault detection purposes for the numerical gearbox dataset.

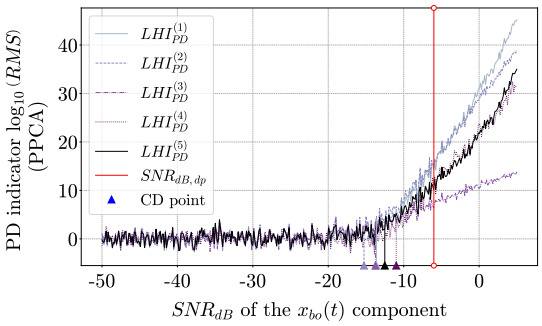
Figure 3.3(e) and Figure 3.3(f) present the results for the PPCA$_h$ model. It is clear that the $HI^{(1)}$ metric is the most sensitive to the onset of bearing damage for the PPCA$_h$ model and provides the earliest CD point. The PI latent indicators for the PPCA$_h$ model exhibit weaker sensitivity than the PD indicators, as the detection of the onset to damage is delayed, as indicated by the identified CD points. Interestingly, the PPCA$_h$ model exhibits obvious fluctuations in the RMS of the HIs and LHIs not due to damage. This sensitivity is attributed to the minor perturbations in the operating conditions, which suggests that the model is sensitive to operating conditions not seen in the reference dataset using for LVM training. Such sensitivity is less obvious in the results from the PPCA and SRICA models.

The results in Figure 3.3 indicate that the proposed scaling procedure for the SRICA model given in Equation (3.9) is suitable for the numerical gearbox dataset. For both of the considered online training-based models, the reconstruction error, $HI^{(1)}$, responds to the presence of the outer race component. This suggests that there is residual data space information not captured by both models that contains bearing fault phenomena. For the PPCA model, the $LHI^{(4)}_{PI|PD}$ metric exhibits the least sensitive response to the bearing fault, while for the SRICA model, the $LHI^{(3))}_{PD}$ metric exhibits weak sensitivity. For the PPCA$_h$ model, the $LHI^{(4)}_{PI}$ and the $LHI^{(3)}_{PD}$ metrics are the least sensitive to the damage.

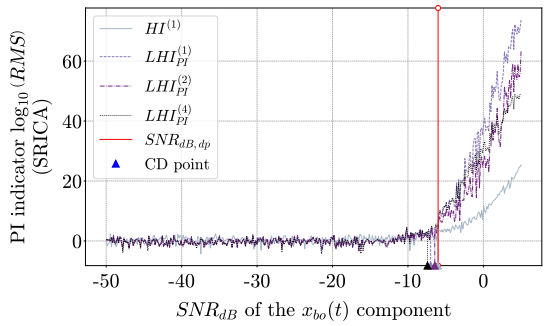
The standardised log-scale amplitude of the outer race component in the SES of the considered latent indicators (i.e., Equations (3.10)–(3.15)) is presented in Figure 3.4. The BPFO component is detectable in the condition indicators from the offline and online LVMs considered in this work. The results seen in Figure 3.4 indicate that the choice of LVM formulation is an important consideration and further validates the scaling approach proposed



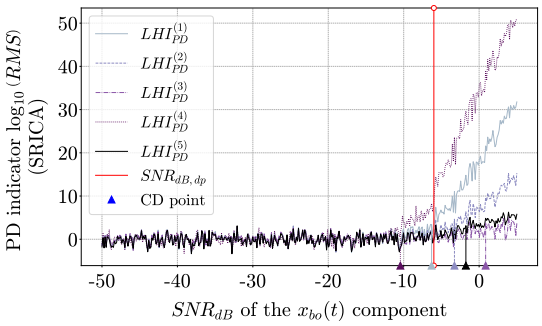
(a) PI indicators - PPCA.



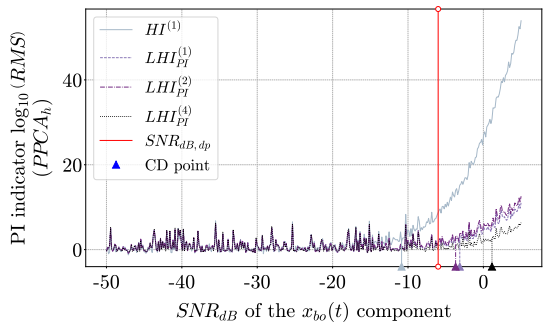
(b) PD indicators - PPCA.



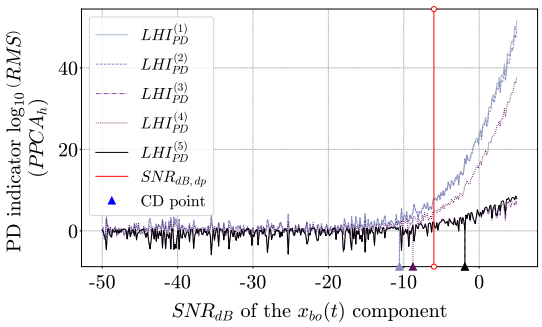
(c) PI indicators - SRICA.



(d) PD indicators - SRICA.



(e) PI indicators - PPCA_h.



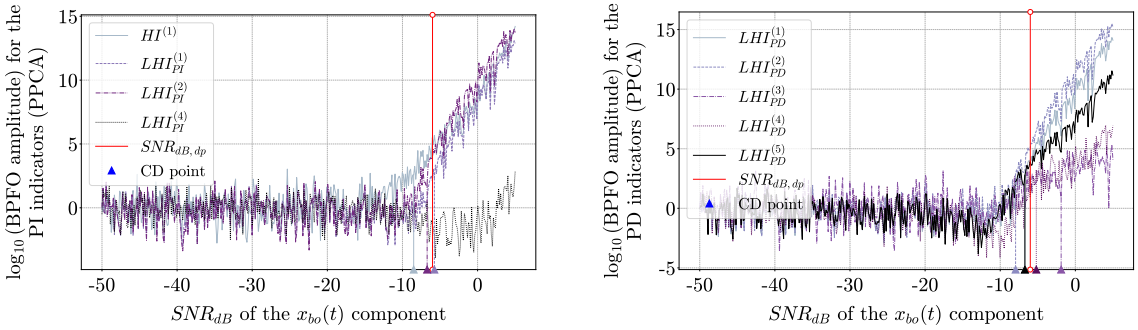
(f) PD indicators - PPCA_h.

Figure 3.3. The standardised $\log_{10}(RMS)$ of the indicators from the considered offline and online LVMs applied to the numerical dataset. The indicators are grouped by LVM formulation and path traversal setting for comparison purposes. The first two rows represent the online models, while the last row represents the offline model. PPCA represents the online reconstruction-focused model, while PPCA_h represents the offline reconstruction-focused model. The indicator RMS is plotted against $SNR_{dB,bo}$ to highlight indicator sensitivity relative to the prominence of outer race component, and the $SNR_{dB,dp}$ level of the distributed pinion damage component is shown to quantify indicator sensitivity to the bearing damage relative to the prominence of this component in the phenomenological data. The Condition Deviance (CD) points are identified using Equation (3.18).

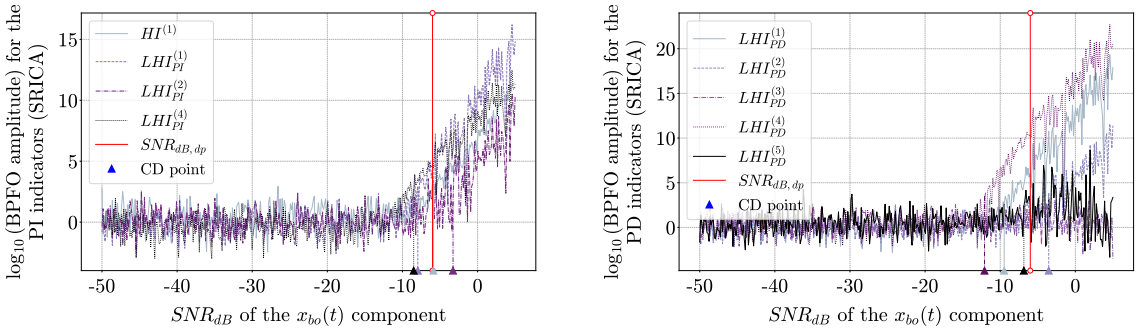
in Equation (3.9) for the numerical gearbox dataset.

Furthermore, the LVMs developed with online training can detect and trend deviations in the asset state, as well as isolate the faulty component of interest for the numerical dataset. This suggests that the bearing damage signatures are captured by the latent manifold. Detection of the BPFO component $LHI_{PI}^{(4)}$ metric is noticeably unobtainable for the PPCA model. The $LHI_{PD}^{(3)}$ and the $LHI_{PD}^{(5)}$ indicators from the SRICA model exhibit minimal sensitivity to the outer race fault, while the remaining condition indicators for the SRICA perform well, giving good responses. This result is also observed in Figure 3.3(c) and Figure 3.3(d). However, for the PPCA$_h$ model, the PD indicators exhibit weaker sensitivity to the detection and trending of the faulty component. This indicates that while the PD indicators are useful to detect damage using the RMS diagnostic metric, as observed in Figure 3.3(f), the fault condition is less identifiable. This result demonstrates two implications: First, the characteristics of the considered offline and online LVMs uniquely influence the health indicators used for condition inference. Second, the latent manifold response to damage cannot be known a priori. Hence, it is necessary to consider multiple LHIs.

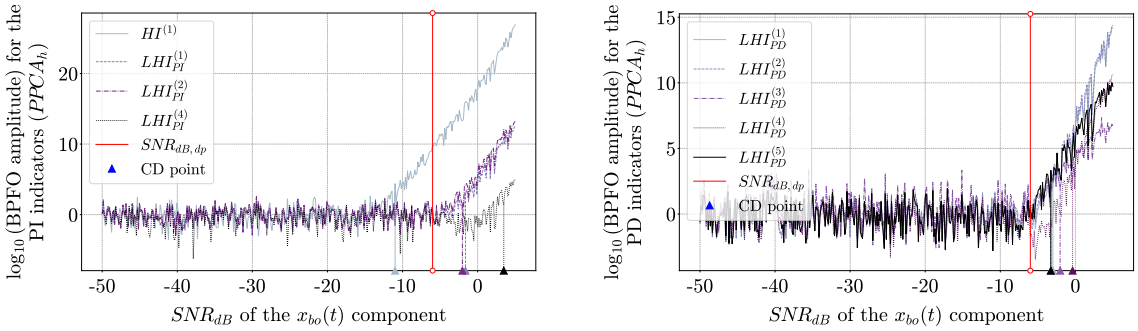
In the comparison between the considered offline and online training approaches used in this investigation, in both the diagnostic metrics used, all of the considered methods respond well to the bearing damage present in the phenomenological model data. However, the following differences were noted: The first observation, seen in Figures 3.3(e) and 3.3(f), is that the PPCA$_h$ model exhibits a noticeable sensitivity to the fluctuation in operating conditions induced in this dataset, which suggests that the offline model is sensitive to data out of its training set, whether it be damage related or operating condition related. However, for the models developed using the online training approach, this sensitivity is less obvious. The second observation is that the SRICA model exhibits an improved ability to capture the fault condition of interest in the LHIs in comparison to the other models, as seen in Figure 3.4(c) and more obviously, in Figure 3.4(d). The exception here is the $HI^{(1)}$ metric from the PPCA$_h$ model, seen in Figure 3.4(e), indicates that the reconstructed data space is a useful tool for condition inference when using the PPCA$_h$ model on this dataset. This highlights the complementarity of the considered methods as the choice of training approach exhibits unique characteristics and implications, and the choice of LVM formulation influences the information gained to perform fault diagnostics.



(a) PI indicators - PPCA. **(b)** PD indicators - PPCA.



(c) PI indicators - SRICA. **(d)** PD indicators - SRICA.



(e) PI indicators - PPCA$_h$. **(f)** PD indicators - PPCA$_h$.

Figure 3.4. The standardised log-scale amplitude of the BPFO component in the SES of the considered indicators for the numerical dataset. The results are log-scaled prior to standardisation for result interpretability purposes.

3.3.2 Latent investigation for offline and online methods

The purpose of this section is to investigate the implications of the considered offline and online LVMs on the learnt latent manifold using temporal preservation [164]. More specifically, the characteristics of the latent variable components (i.e., the components that are present) are

visualised to understand the implications of using the specific LVM models. The transformation process from the raw vibration signals into latent variables for the considered offline and online LVMs is of interest. To investigate this transformation, the correlation of the l^{th} latent component with the components of $x_{ng}(t)$ is considered. The cosine correlation, denoted by ψ_{xy}, is used in this investigation and is determined using the cosine distance between the amplitude of the spectral representations of latent component $\mathbf{f}_{(z_l)} = |\mathcal{F}(\{z_{1,l}, \cdots, z_{L_T-1,l}\})|$, where $1 \leq l \leq d$, and numerical signal component $\mathbf{f}_{(x_j)} = |\mathcal{F}(\{x_1, \cdots, x_{L_T-1}\})|$, where j refers to the components of $x_{ng}(t)$. The cosine correlation is given by

$$\psi_{xy}^{j} = \frac{\mathbf{f}_{(z_l)}^{T}\mathbf{f}_{(x_j)}}{\|\mathbf{f}_{(z_l)}\|_2 \cdot \|\mathbf{f}_{(x_j)}\|_2}, \tag{3.19}$$

where $\mathcal{F}(\cdot)$ represents the Fourier operator and $\mathbf{f}_{(\cdot)} \in \mathbb{R}^{L_T-1}$ represents the discrete Fourier transform's amplitude over the full spectral range. Equation (3.19) represents how similar the latent variables are with the components of $x_{ng}(t)$ and is scaled between 0 and 1. For this investigation, record 305 is used as this represents a point where the SNR of the outer race component, $SNR_{dB,bo} = -8dB$, is slightly less than the SNR of the distributed pinion component $SNR_{dB,dp} = -6dB$.

The spectra of the latent components are shown in Figure 3.5 for the considered offline and online LVMs. Figure 3.5(a) indicates that the PPCA model first extracts the dominant gear mesh components and associated harmonics. This is expected in the context of Section A.5, which highlighted that PPCA seeks latent components that maximise the latent component power. PPCA then extracts a portion of the frequency band corresponding to the distributed pinion component and then a portion of the frequency band corresponding to the outer race-bearing component. This can be attributed to the SNR of these components as the bearing component had a lower variance than the distributed pinion component. The remainder of the PPCA latent manifold then represents the residual frequency bands around the numerical signal components. The variance-based PPCA latent decomposition presented in Figure 3.5(a) draws attention to a notable relationship with Singular Spectrum Analysis (SSA)-based methodologies, which are variance-driven decomposition methods. Particularly, Figure 3.5(a) highlights the smearing of the components over the latent variables, which has implications and potential consequences associated with component selection procedures used in SSA-based methodologies, as discussed in Ref. [176].

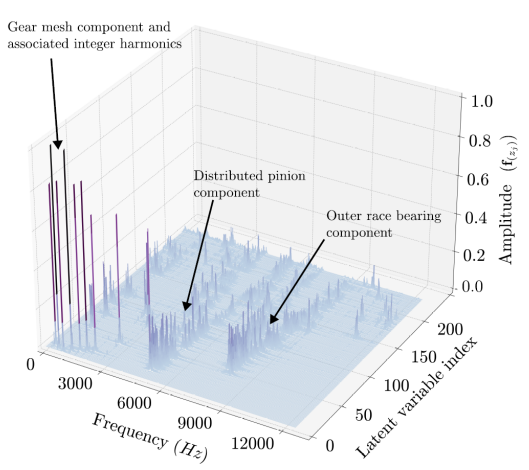
Figure 3.5(a) demonstrates that the PPCA model captures the components that contribute

to the variance in the data. Figure 3.5(b) indicates that the SRICA model, which captures fundamentally different characteristics of the observed data, extracts the outer race component, the gear mesh component and its associated integer harmonics, and then the distributed pinion component. The ordering of the latent components is represented by their respective excess kurtosis values. Each of the first three extracted components represents prominent non-Gaussian latent signals. The remaining latent components are then localised around specific frequency bands in the noise floor of the raw vibration signal, a phenomenon attributed to the spectral regularisation term proposed in Appendix C. Evidently, Figure 3.5(b) demonstrates that the SRICA model focuses on finding informative latent components that are inherently non-Gaussian, and the model seems to group the components better. Figure 3.5(c) presents the latent information of the decomposed signal using the PPCA$_h$ model. The PPCA$_h$ model has less pronunciation of the bearing damage components in its latent manifold, which is expected as the damage was not present in its training set. However, the results seen in Figure 3.3 show sensitivity to the damage, which indicates that the model captures the region of the noise floor in which the resonance band manifests but decomposes it up through its latent manifold as it contains little information beneficial to improving reconstruction.

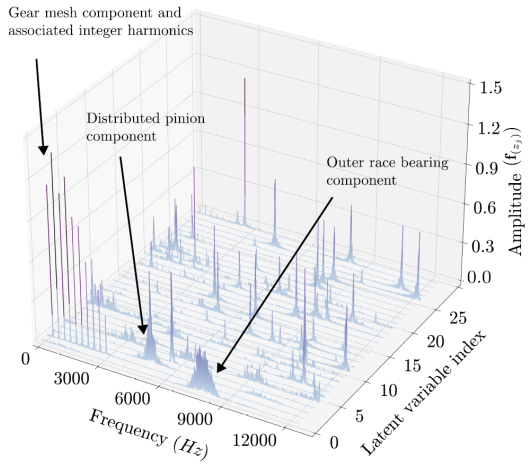
Figure 3.6 presents the cosine correlation, given in Equation (3.19) of the latent components with respect to the numerical signal components from Equation (3.16). Figure 3.6 supplements the observations made for Figure 3.5, whereby the PPCA model extracts components that maximise the explained variance in the observed data, while the SRICA model is focused on informative latent components that are highly non-Gaussian with the notion that this leads to more interpretable latent components. The results indicate that the data components are smeared along the latent variables for the PPCA model, while the data components are more localised for the SRICA model. For the PPCA$_h$ model, it is difficult to isolate the bearing damage component, which indicates that the PPCA$_h$ model decomposes the noise floor in which the resonance band related to the bearing damage manifests in a non-unique manner, and hence isolating its presence using the latent components alone is non-trivial, and the information is smeared.

3.4 Experimental investigation

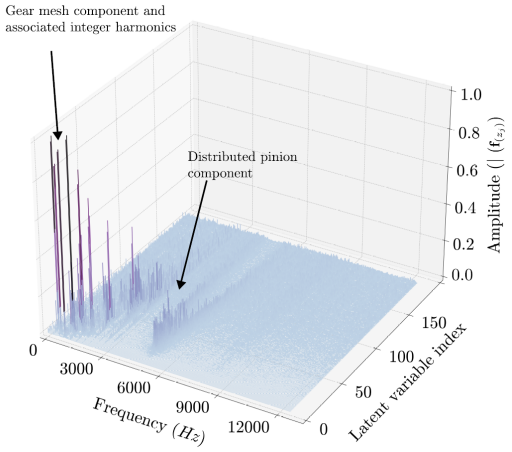
In this section, the considered offline and online LVMs are compared on two experimental bearing and gear tooth fault datasets, which represent different gearbox failure mechanisms



(a) Latent spectra - PPCA.



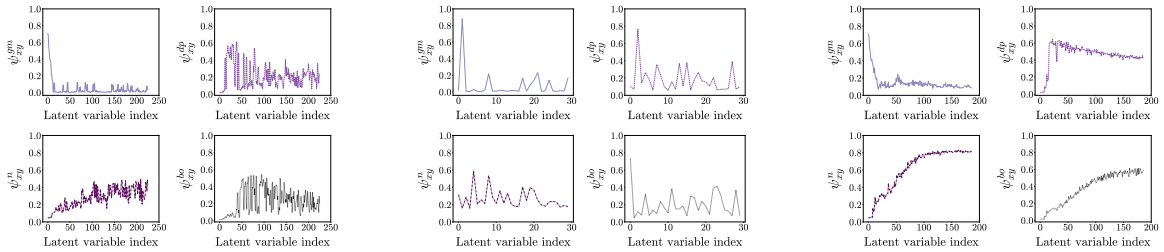
(b) Latent spectra - SRICA.



(c) Latent spectra - PPCA$_h$.

Figure 3.5. The spectral representations of the latent components of the (a) PPCA model, the (b) SRICA model, and the (c) PPCA$_h$ model. In both (a) and (b) the x_{gm}, x_{dp}, and x_{bo} numerical signal components are identified and annotated. The PPCA and PPCA$_h$ models appear to smear the numerical signal components through the latent manifold, whereas the SRICA model has isolated sources that represent the signal components. For reference, the spectral representation of the numerical gearbox signal components is given in Figure B.2.

and operating condition states. The objective is to conduct an LVM formulation comparison in the online training setting using the properties of the HIs and LHIs considered in this investigation. The properties of interest are the ability of the considered HIs and LHIs to detect and trend the damage and the ability to localise the faulty component of interest. The first investigation is performed on a dataset with constant operating conditions and a bearing fault failure that occurred during the experimental lifespan. The second investigation is performed on a dataset with time-varying operating conditions, and a gear tooth failure



(a) PPCA model. **(b)** SRICA model. **(c)** PPCA$_h$ model.

Figure 3.6. The cosine correlation ψ_{xy}^{k} for the (a) PPCA model, the (b) SRICA model, and the (c) PPCA$_h$ model. Note that the superscript k refers to the correlation with the $k = gm$ gear mesh component, the $k = dp$ distributed pinion component, the $k = n$ noise component, and the $k = bo$ outer race bearing damage component.

occurred during the experimental lifespan. The model parameters used for the experimental investigations performed in this chapter are summarised in Appendix A.2.

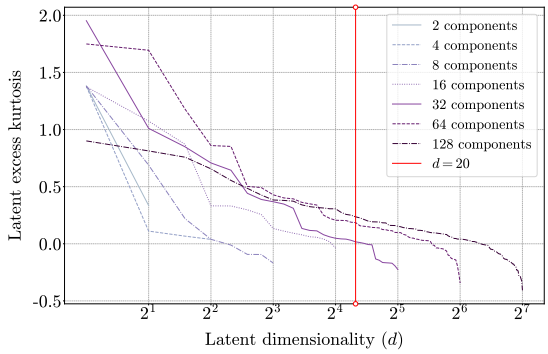
For both datasets, the performance of the LVMs is compared in a fault diagnosis setting, whereby the ability to perform incipient fault detection, isolation, and trending is investigated for LVMs using online training. Additionally, the results are compared to those obtained from the PPCA$_h$ model developed using the offline training approach. This comparison is necessary to investigate the implications of online training in a LVM-based fault diagnostics context. For the considered investigations, the condition interval given in Equation (3.18) is used to identify a CD point.

3.4.1 Experimental bearing dataset

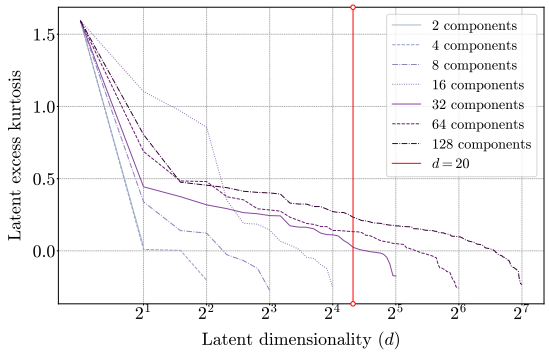
The experimental dataset used in the first part of this investigation is the Intelligent Maintenance Systems (IMS) bearing dataset, discussed in detail in Section 2.4.1. A schematic of the experimental setup is depicted in Figure 2.6(a). The RMS and kurtosis of the raw vibration signals for the considered IMS dataset is visualised in Figure 2.6(b). The fault condition that manifested during the IMS dataset experimental lifespan is an outer race bearing fault with a characteristic fault frequency is 236 Hz.

The LVMs trained on the available 984 raw vibration signals used the Hankelisation-based pre-processing step given in Equation (2.13) with a window length of $L_w = 256$ and a shift of $L_{sft} = 1$. The CCR for the PPCA model was set to $CCR = 80\%$, while the dimensionality of the SRICA model was set to $d = 20$. The first 10% of the available data was used to train the PPCA$_h$ model with a CCR of 80%. To validate the choice of latent dimensionality for

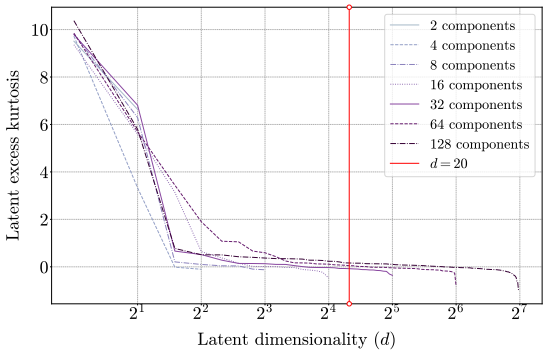
the SRICA model, the excess kurtosis of the latent components for signals at 25%, 50%, and 75% of the experimental lifespan is investigated for a latent dimensionality range of $d = 2$ to $d = 128$ in increments of $d = 2^j$, where $j = 1, \cdots, 7$. Figure 3.7 presents the latent excess kurtosis for the considered latent dimensionality range. It is evident that the latent component kurtosis at 25% and 50% of the experimental lifespan contains little non-Gaussian information of significant interest. However, at 75% of the experimental lifespan, there are prominent components of interest, with a noticeable elbow in the latent kurtosis around $d = 4$. Hence, a latent dimensionality larger than $d = 4$ is sufficient to allow the SRICA model to capture interesting components of the raw vibration data in the online training setting.



(a) 25% of the experimental lifespan.



(b) 50% of the experimental lifespan.



(c) 75% of the experimental lifespan.

Figure 3.7. The latent excess kurtosis for the SRICA models trained at different states of the IMS dataset lifespan. The x-axis of (a), (b), and (c) are $\log_2$ scaled to aid with visualisation.

To conduct a comparison between the considered offline and online LVMs in the offline training setting and the online training setting, the standardised RMS of the $HI^{(1)}$ and $LHI^{(1-5)}$ signals is computed as a diagnostic metric. As was done for the numerical gearbox investigation, the diagnostic metric trends for the indicator signals are non-dimensionalised

using a portion of the first 10% of the available experimental data. To quantify the sensitivity of the considered indicators to the bearing fault present, the SNR of the SES is used as a second diagnostic metric. The Square Envelope (SE) of a signal $x[n]$ is given by

$$SE[n] = |x[n] + \sqrt{-1} \cdot H\left(x[n]\right)|^2, \tag{3.20}$$

where $\sqrt{-1}$ represents the imaginary unit and $H(\cdot)$ represents the Hilbert transform [332]. The SES is given by

$$SES[k_\alpha] = |\text{DFT}\{SE[n]\}|, \tag{3.21}$$

where k_α represents the cyclic frequency and $\text{DFT}\{\cdot\}$ is the Discrete Fourier Transform (DFT). Schmidt *et al.* [33] propose a targeted methodology to determine the SNR based on the SES

$$SNR_{SES} = \sum_{l=1}^{N_l} \frac{\max\{SES[k_\alpha]\}_{k_\alpha \in \mathcal{A}_l}}{\text{median}\{SES[k_\alpha]\}_{k_\alpha \in \mathcal{B}_l}}, \tag{3.22}$$

which represents the prominence of specific spectral components and harmonics thereof with respect to the noise floor in the SES within specific cyclic frequency bands. The noise floor is estimated by the median of the SES in a specific frequency band. The set $\mathcal{A}_l$ represents the indices associated with the l^{th} harmonic of the cyclic frequency α_c of interest within a tolerance range

$$\mathcal{A}_l = \{k_\alpha | k_\alpha \in \mathbb{N}, \text{ and } \alpha_c \cdot (l - \tau_\mathcal{A}) \leq \alpha_k \leq \alpha_c \cdot (l + \tau_\mathcal{A})\}, \tag{3.23}$$

where $\tau_\mathcal{A}$, $0 \leq \tau_\mathcal{A} \leq 1$, represents a cyclic frequency tolerance control parameter [32]. The set $\mathcal{B}_l$ represents the indices of the frequency band with a bandwidth centred around the l^{th} harmonic of α_c

$$\mathcal{B}_l = \{k_\alpha | k_\alpha \in \mathbb{N}, \text{ and } \alpha_c \cdot (l - \tau_\mathcal{B}) \leq \alpha_k \leq \alpha_c \cdot (l + \tau_\mathcal{B})\}, \tag{3.24}$$

where $\tau_\mathcal{B}$ represents the bandwidth parameter $0 \leq \tau_\mathcal{A} < \tau_\mathcal{B} \leq 1$, which captures the components of the SES between the integer harmonics of the cyclic frequency. Figure A.1 in Appendix A provides a visual demonstration of the SNR calculation based on the SES. For this investigation, the first three integer harmonics, i.e., $N_l = 3$, are used as targeted features to compute the SNR_{SES}.

Figure 3.8 presents the standardised RMS of the condition indicators for the considered offline and online LVMs. Indicator standardisation was performed using an estimate of the indicator mean and variance from a portion of the first 10% of the available indicator signals. The indicators from the PPCA model, as presented in Figures 3.8(a) and 3.8(b), exhibit clear sensitivity to the bearing damage present in the considered IMS dataset. Both the PI and

PD settings provide fault-sensitive indicators for the PPCA model. Using the identified CD points, the PPCA model is responsive to the onset of damage around record 533, and the $LHI_{PD}^{(3)}$ metric is the least sensitive to the onset of damage. The reconstruction error, $HI^{(1)}$, exhibits the strongest sensitivity for the PPCA model, which indicates that residual fault information is present in the reconstructed data space. It is also evident that towards the end of the experimental lifespan, the considered latent indicators undergo strong fluctuations from the baseline condition.

Figure 3.8(c) and Figure 3.8(d) present the results from the SRICA model. In comparison to the results from the PPCA model, the indicators from the SRICA model have less sensitivity in the RMS of the condition indicator signals. The PI setting for the SRICA model produces indicators that exhibit better sensitivity to the bearing damage present in the IMS dataset than the PD indicators. In the comparison of the CD points from the PPCA model to the CD points from the SRICA model, it is evident that the PPCA PI indicators have better sensitivity to the onset of bearing damage present in the data when using the RMS as a diagnostic metric. As with the PPCA model, the $HI^{(1)}$ metric from the SRICA model exhibits stronger sensitivity in the RMS of the condition indicator signal. The PD setting for the SRICA model is noticeably insensitive to the onset and progression of damage, whereby the CD points are located around the strong change in condition present around record 700 (as observed in Figure 2.6(b)) is detectable. This suggests that the latent scaling procedure proposed in Equation (3.9) is less optimal for the IMS dataset.

The RMS of the PPCA$_h$ indicator signals, as presented in Figures 3.8(e) and 3.8(f), all exhibit sensitivity to the onset of bearing damage except for the $HI^{(1)}$ metric, the reconstruction error, which shows weakened sensitivity and poorer CD identification. The $LHI_{PD}^{(3)}$ and $LHI_{PD}^{(5)}$ metrics from the PPCA$_h$ model exhibit more significant fluctuations once the fault occurs, indicating that the latent manifold response to damage captured by these metrics is more complex and this causes a drop below the baseline RMS metric range.

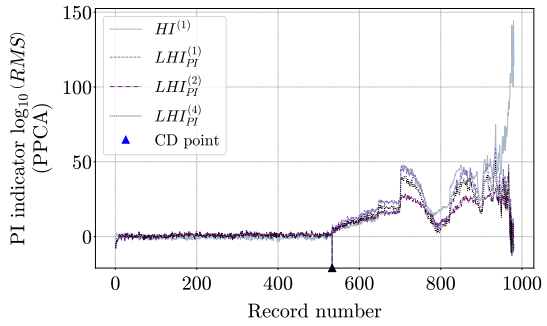
In a comparative analysis of the offline and online reconstruction-focused LVMs under consideration, both the PPCA model and the PPCA$_h$ model exhibit the ability to provide distinct RMS deviations to the onset of the localised bearing damage. However, the sensitive indicators are not consistent between the two methods, e.g., $HI^{(1)}$ and $LHI_{PI}^{(2)}$ in the PI setting and $LHI_{PD}^{(3)}$ and $LHI_{PD}^{(5)}$ in the PD setting. This emphasises the need to use multiple indicators as the best indicators are not known a priori. This is supported by the results

from the SRICA model, which exhibit better sensitivity to the onset of damage from the PI indicators than the PD indicators. This emphasises the need to consider both the PI and PD traversal settings when computing LHIs. The models trained using the online training approach gives good performance in comparison to the PPCA$_h$ model trained on historical vibration data. Hence, it is necessary to further quantify the implications of the LVM training procedure used and LVM formulation by determining the sensitivity to the bearing fault present in the data.

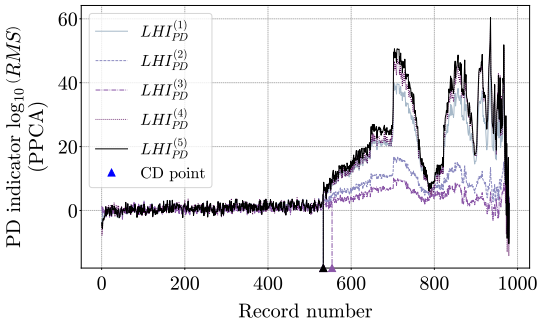
In Figure 3.9 the standardised $\log_{10}(SNR)$ of the considered indicators is presented for the considered offline and online LVMs. The SNR sensitivity investigation is necessary to motivate and quantify the fault isolation characteristics of the considered HI and LHI condition indicators. The condition indicators from the SRICA model, presented in Figures 3.9(c) and 3.9(d), exhibit clear sensitivity to the IMS outer race bearing fault and the indicators from PI setting are more sensitive than the indicators from the PD setting and have less noise. However, the indicators from the PPCA model, presented in Figures 3.9(a) and 3.9(b) exhibit weak sensitivity. This difference is also discernible via the identified CD points, the PPCA model only provide CD detection after record 600.

Figures 3.9(e) and 3.9(f) indicate that the PPCA$_h$ model has improved sensitivity to the fault present in the data over the PPCA model, but the PI indicators from the SRICA model exhibit the best sensitivity overall. For the PPCA$_h$ model, only the $LHI^{(4)}_{PI|PD}$ metrics are able to identify a CD point before record 600. This contrasts the results observed in the RMS of the condition indicators presented in Figure 3.8, thereby highlighting the implication of the characteristics of the considered LVMs and the choice of diagnostic metric. Figure 3.9(c) indicates that the LHIs from the SRICA model in the PI setting all contain information directly related to the fault of interest and the CD points all identify the change in condition around record 533. In the PD setting, seen in Figure 3.9(d), only the $LHI^{(3)}_{PD}$ indicator from the SRICA model is insensitive to the fault condition present in the data and no CD point is obtained for this indicator. However, the SNR diagnostic metric for the LHIs from the SRICA model in the PD setting have more noise, indicating that the PI setting is better suited to CM tasks for the IMS dataset using this LVM formulation.

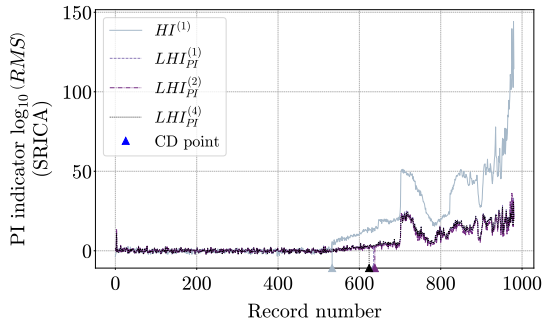
For the PPCA model SNR results, seen in Figures 3.9(e) and 3.9(f), the HI and LHI indicators exhibit clear insensitivity to the change in condition around record 533, and only show a noticeable CD detection around record 700. For the PPCA$_h$ model, CD point



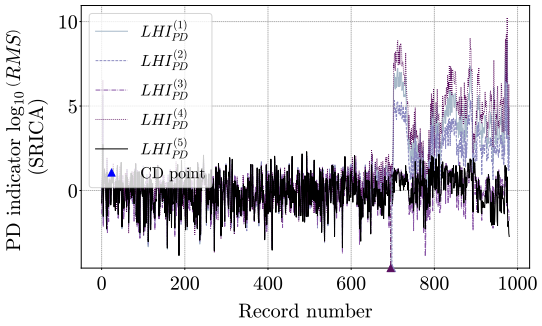
(a) PI indicators - PPCA.



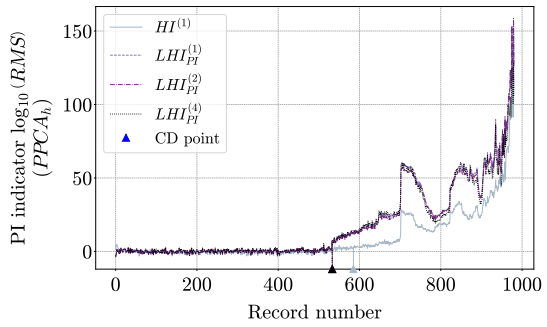
(b) PD indicators - PPCA.



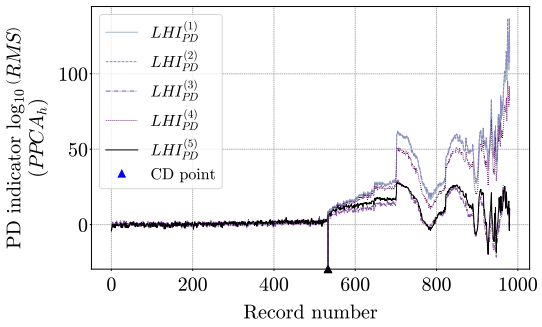
(c) PI indicators - SRICA.



(d) PD indicators - SRICA.



(e) PI indicators - PPCA_h.



(f) PD indicators - PPCA_h.

Figure 3.8. The standardised $\log_{10}(RMS)$ of the indicators for the considered offline and online LVMs applied to the IMS dataset. The indicators are grouped by LVM formulation and path traversal setting for comparison purposes, and log-scaled prior to standardisation for visualisation purposes.

identification occurs earlier than record 700, indicating that although the model has strong indicator deviations in its RMS, the bearing fault frequency is less pronounced. Hence, the SRICA model is observed to be the most sensitive to the bearing damage on the IMS dataset and the characteristics of the LVM better isolate the fault components of interest within the

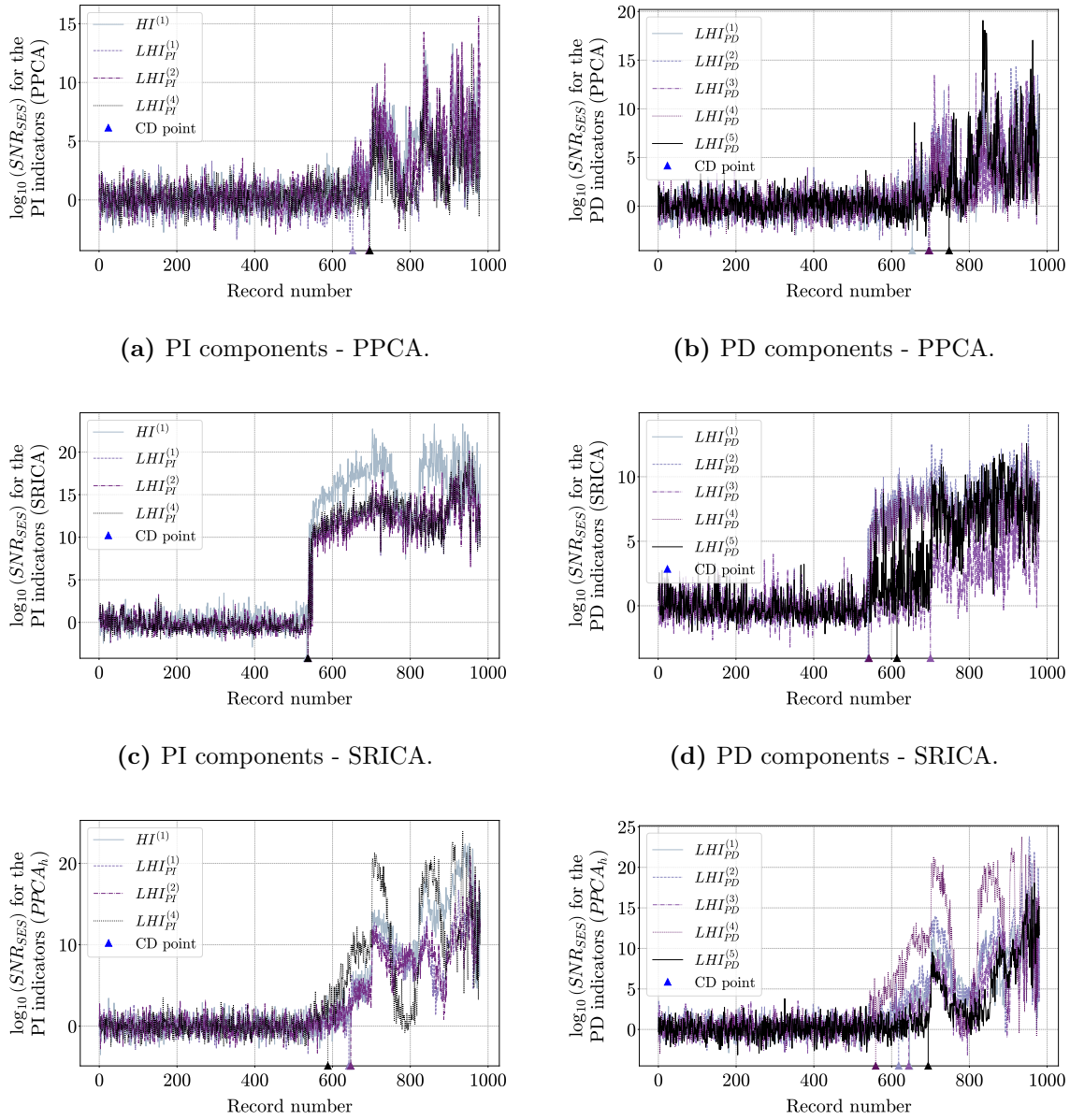
latent manifold. The PPCA and PPCA$_h$ models, in comparison, fail to accurately represent the fault component of interest in the LHIs, indicating that maximal variance is a useful tool to detect deviations due to the presence of a fault, but may not always be useful to infer the underlying fault condition.

The results observed in Figure 3.8 and Figure 3.9 now provide evidence to the complementarity of the reconstruction-focused and interpretation-focused LVMs in the online training setting. The PPCA model was able to clearly respond to the change in the variance of the raw vibration signals over record number but cannot strongly isolate the faulty component of interest in the considered indicators. The SRICA model, in comparison, has weaker sensitivity in the RMS of the considered indicator signals but effectively captures and isolates the bearing fault of interest in the considered PI and PD indicators. This implies that the interpretation-focused LVM is better suited to capture the fault condition present in the IMS dataset. At the same time, the reconstruction-focused LVM is better suited to detecting the change in the variance of the raw vibration signals due to the fault. Additionally, a comparison of the metrics from the PPCA$_h$ model developed using a reference dataset highlights the implications of the training scheme coupled with the model formulation. Together, the presented findings confirm that the online training setting emphasises important considerations surrounding the choice of formulation for LVM-based fault diagnostics, and the comparison to the PPCA$_h$ model highlights that the online training approach is a useful LVM setting in CM applications.

3.4.2 Experimental gearbox dataset

The experimental dataset used in the second part of this investigation is the experimental bearing dataset, discussed in detail in Section 2.4.2. The gear of interest at the point of fault introduction and after the experiments were concluded is shown in Figure 2.15. The time-varying operating condition present during the experimental lifespan is visualised in Figure 2.14(a). Time-varying operating conditions impede the CM task as the instantaneous health state of the gearbox is intertwined with information not related to the health state of the gear.

Additionally, the raw vibration data contains prominent impulsive information for all recorded data, regardless of the gear health state. This is attributed to unseen contact between the casing of the gearbox and a floating bearing when axial shaft loadings were applied on the shaft. The statistical properties of the raw vibration signals are visualised in Figure 3.10.



(a) PI components - PPCA.

(b) PD components - PPCA.

(c) PI components - SRICA.

(d) PD components - SRICA.

(e) PI components - PPCA_h.

(f) PD components - PPCA_h.

Figure 3.9. The standardised $\log_{10}(SNR)$ determined using Equation (3.22) of the indicator signal SES for the considered offline and online LVMs applied to the IMS dataset. The indicators are grouped by LVM formulation and path traversal setting for comparison purposes. The SNR_{SES} metric is determined using Equation (3.22) and the results are log-scaled prior to standardisation for visualisation purposes.

Note that unlike the gearbox dataset used in Chapter 2, no low-pass filtering is applied to the raw vibration data in this chapter. The implication of not filtering the data is observed in the comparison of Figure 2.14(b) with Figure 3.10.

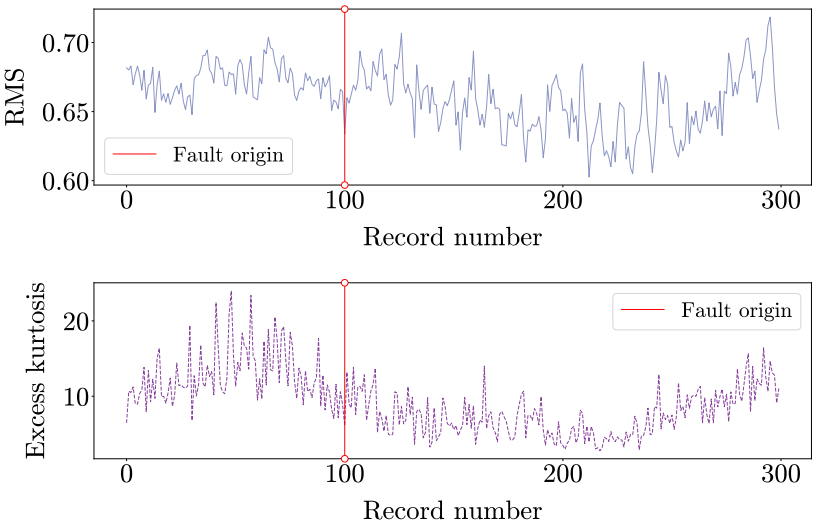
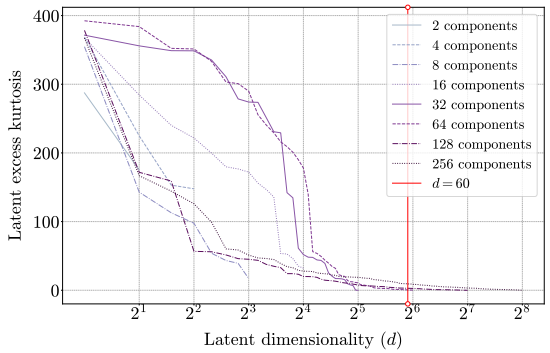


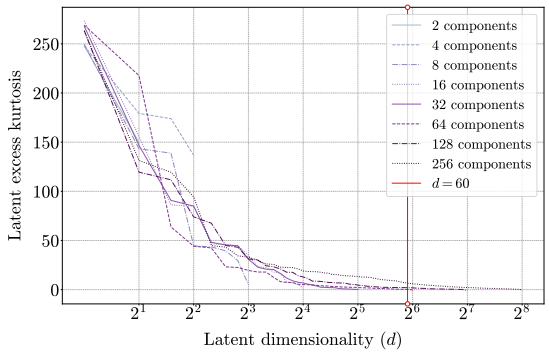
Figure 3.10. The statistical properties of the raw vibration data for the gearbox dataset. It is evident that the operating conditions present and the additional impulsivity, due to the floating bearing, amplify the difficulty of the condition inference task. Furthermore, the excess kurtosis of the raw vibration data is also strongly non-Gaussian throughout the experimental lifespan, and is highly leptokurtic.

The LVMs trained on the 300 raw vibration signals used a window length of $L_w = 512$ and a shift of $L_{sft} = 1$. The CCR for the PPCA model was set to 95%, while the latent dimensionality of the SRICA model was set to $d = 60$. Increasing the CCR was found to significantly improve the performance of the PPCA model for the experimental gearbox dataset. However, increasing the CCR for the IMS dataset investigation did not yield a positive performance improvement. This suggests the CCR is a sensitive parameter that is important to the performance of reconstruction-focused LVMs for the fault diagnosis task. The first 10% of the available data was used to train the PPCA_h model with a CCR of 95%. To validate the choice of latent dimensionality for the SRICA model, the process described in Section 3.4.1 was used, and the latent dimensionality was increased until $d = 256$. Figure 3.11 contains the visualisation of the kurtosis in this regard. The latent source kurtosis approaches an elbow point around $d = 32$, and hence the selected latent dimensionality for the SRICA model is suitable to capture the non-Gaussian components in the raw vibration data.

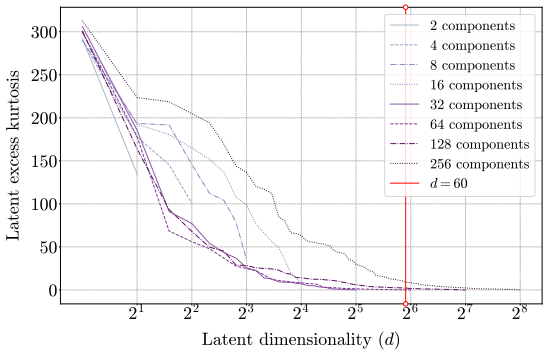
For the gearbox dataset, the Synchronous Median of the Square Envelope (SMSE) [333] is used to conduct the investigation and comparison for the considered offline and online LVMs.



(a) 25% of the experimental lifespan.



(b) 50% of the experimental lifespan.



(c) 75% of the experimental lifespan.

Figure 3.11. The latent excess kurtosis for the SRICA models trained at different states of the experimental gearbox lifespan. The x-axis of (a), (b), and (c) are $\log_2$ scaled to aid with visualisation.

The SMSE of the j^{th} order tracked condition indicator signal from the i^{th} record is given by

$$SMSE_j(\phi, \Phi_k) = \text{median}\left(SE_j(\phi, \Phi_k)\right), \tag{3.25}$$

where median $(\cdot)$ represents the median statistic, $\phi \in [0, \Phi_k)$ refers to the angular position, Φ_k refers to the cyclic period, and $SE_j(\phi, \Phi_k)$ is the order tracked SE of the j^{th} condition indicator signal. As the condition indicator signals are order tracked, deviations related to the gear tooth fault are expected to be synchronous with the gear shaft rotation, i.e., $\Phi_k = 2 \cdot \pi$. The SMSE is an extension of the synchronous average of the SE and is more robust in the presence of impulsive non-synchronous components [333]. Hence, it is well suited for the gearbox dataset investigation. As the gear fault is known a priori, the objective is to ascertain the sensitivity of the condition indicators from the considered offline and online LVMs to this fault. To conduct such an investigation, two diagnostic metrics are used. First, the RMS of

the region within the SMSE corresponding to the gear tooth fault is used. Second, the SNR of the SMSE signals is determined via

$$SNR^{dB}_{SMSE} = 20 \log \left(\frac{A^{SMSE}_{j,damaged}}{A^{SMSE}_{j,healthy}} \right), \tag{3.26}$$

where $A^{SMSE}_{j,damaged}$ and $A^{SMSE}_{j,healthy}$ represent the RMS of the damaged portion and the healthy portion, respectively, of the j^{th} SMSE signal. The SNR^{dB}_{SMSE} metric is used to quantify how prominent the fault component of interest is in relation to the noise floor. Figure A.2 in Appendix A presents a demonstration of the SMSE condition quantification methodology used in this investigation for the experimental gearbox dataset investigation. To ensure that the correct healthy and unhealthy regions are extracted from the SMSE of the indicators signals, each SMSE signal is realigned with respect to the tachometer butt joint. The location of the tachometer butt joint is found using the Bayesian Geometry Compensation (BGC) algorithm proposed in Diamond *et al.* [316].

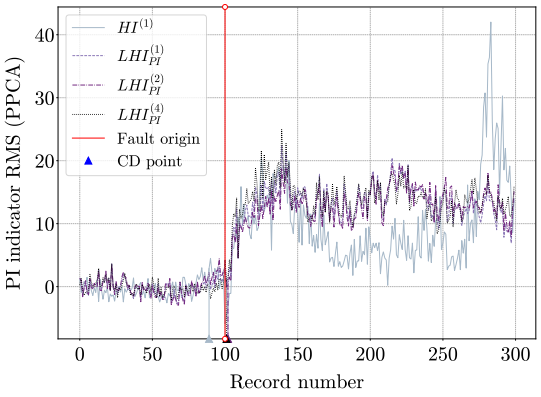
Figure 3.12 presents the standardised RMS of the SMSE metric for the indicators used in this work. The PI indicators from the PPCA model and the SRICA model, presented in Figure 3.12(a) and Figure 3.12(b) respectively, both exhibit a clear variation in condition from the manual fault origin point introduced at record 100. However, the $HI^{(1)}$ metric from the PPCA model erroneously identifies a CD point prior to record 100, which indicates that the PPCA model has some sensitivity to the non-stationarity in the data. The $HI^{(1)}$ metric from the SRICA model exhibits weaker sensitivity to the localised tooth damage, indicating that the latent manifold is more sensitive to the fault condition present in the data. The PD indicators from the PPCA model and the SRICA model, presented in Figure 3.12(b) and Figure 3.12(d) respectively, exhibit an increased sensitivity to the non-stationarity in the data and less sensitivity to the onset of gear tooth damage. For the PD indicators from the PPCA model, this non-stationarity sensitivity is more pronounced as CD points are identified prior to record 100. Additionally, the $LHI^{(3)}_{PD}$ metric undergoes a decrease in value around record 100, indicating that this metric decreases in magnitude around the tooth fault location.

As both LVMs developed using online training are responsive to the onset of gear tooth damage, it can be concluded that the manifestation mechanism for the gear tooth damage is both variance-driven and non-Gaussian in nature. Hence, both of the considered online LVMs are able to effectively capture the variation in asset condition. This highlights the complementarity of the considered online LVMs in a fault diagnosis setting, and enforces

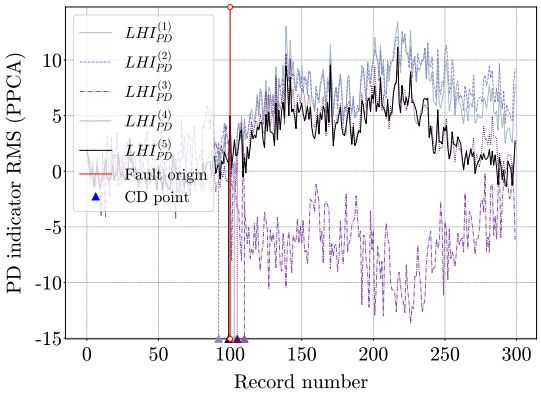
that the choice of LVM formulation is an important consideration for CM using the online training setting. The performance of the PD indicators further highlights the implications of the considered HIs and LHIs used in this work, as damage manifestation locality within the facets of LVMs is not known a priori.

The PPCA$_h$ results presented in Figures 3.12(e) and 3.12(f) also indicate that the PI setting is better suited to the CM task for the gearbox dataset as they capture less non-stationary information and are more sensitive to the onset of damage. For the PPCA$_h$ model, the CD point of the $HI^{(1)}$ metric occurs before record 100, which indicates a clear sensitivity to the non-stationarity in the healthy gearbox data. The RMS of the PI indicators from the PPCA$_h$ model exhibit clearer sensitivity to the onset of the gear tooth fault. The PD indicators from the PPCA$_h$ model are more sensitive to the data non-stationarity, identified via the erroneously identified CD points and the clear variations in RMS metric value in both for both the healthy and unhealthy gearbox data. Hence, the ability to detect and identify an onset to gear tooth damage is effected by the choice of training approach and the LVM formulation used therein and the choice of LHI. While the RMS allows for an investigation into the discernability of the tooth fault, the calculation of the SNR in Equation (3.26) quantifies tooth damage sensitivity relative to the noise floor, which naturally removes any potential sensitivity to non-stationarity present in the data and can better quantify damage detectability performance.

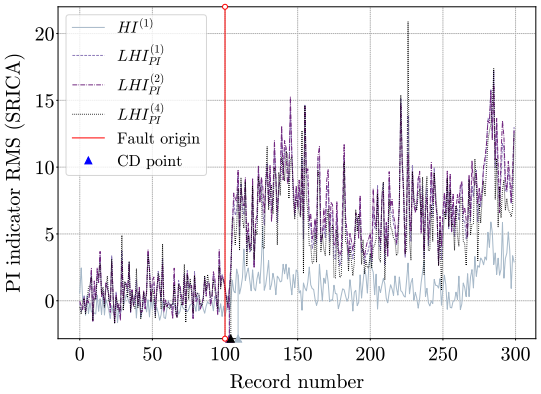
Figure 3.13 presents the SNR of the SMSE for the indicators used in this work. It is clear that the identification of the CD points now occur after record 100, and indicators with worsened sensitivity to the onset of damage can be identified for the considered offline and online LVMs. As the SNR represents the sensitivity of the tooth fault region in the SMSE with respect to the noise floor, it is possible to further compare the indicator responses from the considered offline and online LVMs. In a comparison of the PI indicators, presented in Figures 3.13(a), 3.13(c), and 3.13(e), the PPCA model and the SRICA model exhibit diagnostic metric responses with similar sensitivity to the gear tooth fault. However, the $HI^{(1)}$ metric from the SRICA model, presented in Figure 3.13(a), is less sensitive to the onset of damage than the indicators constructed using the latent manifold. The $LHI_{PD}^{(4)}$ metric response from the PPCA$_h$ model, presented in Figure 3.13(f), performs worse than the other indicators due to the significant fluctuations in the indicator value attributed to sensitivity to the data non-stationarity.



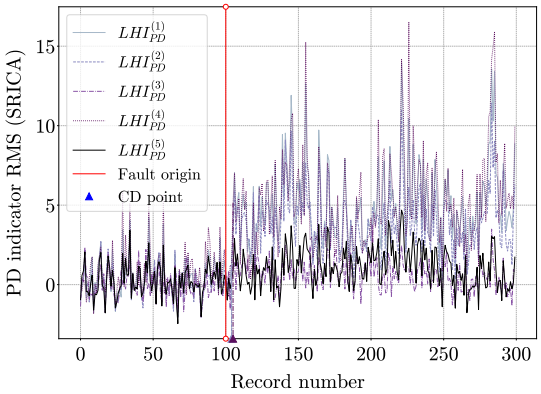
(a) PI indicators - PPCA.



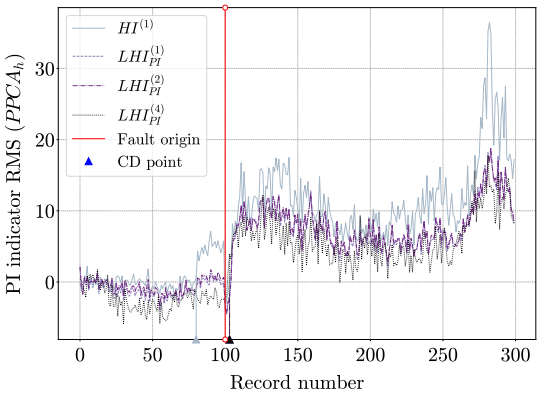
(b) PD indicators - PPCA.



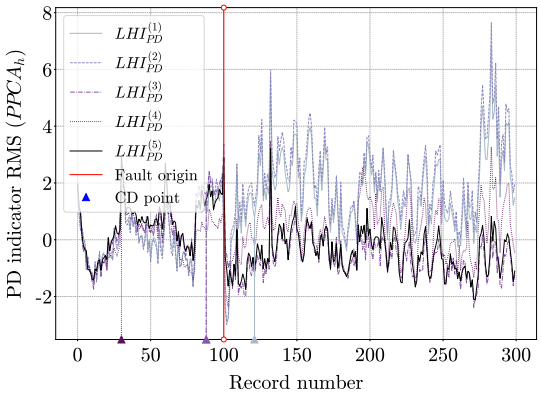
(c) PI indicators - SRICA.



(d) PD indicators - SRICA.



(e) PI indicators - PPCA_h.



(f) PD indicators - PPCA_h.

Figure 3.12. The standardised RMS of the SMSE for the indicators from the considered offline and online LVMs applied to the gearbox dataset. The indicators are grouped by LVM formulation and path traversal setting for comparison purposes.

The PI indicators from the latent manifold of the PPCA and SRICA models, captured in Figures 3.13(a) and 3.13(c) respectively, have improved sensitivity to the onset of damage over the PI indicators from PPCA$_h$ model, seen in Figure 3.13(e), which indicates that they have improved separability between the fault region and the healthy region in the SMSE of the HI and LHI signals. This further emphasises the implication of the LVM formulation used and the choice of LVM training setting. In the case of the gearbox dataset, the choice of offline or online training and the LVM formulation used therein appears to have a less significant impact on the early detection of damage onset compared to the IMS dataset. However, there are still noticeable and identifiable implications. The results seen in Figure 3.13 indicate that the online training setting is a useful setting for LVMs as there are clear responses to the gear tooth fault, and this is substantiated by comparing the results to those obtained from the PPCA$_h$ model.

In a comparison of the PD indicators for the considered offline and online LVMs, presented in Figures 3.13(b), 3.13(d), and 3.13(f), the SNR results indicate that most indicators are sensitive to the change in condition due to the manual fault seeding process, but are in general less sensitive than their PI counterparts. This enforces the implications of the considered LHIs, as each method extracts unique characteristics from the raw vibration data. Furthermore, the SNR results help quantify the RMS results seen in Figures 3.12(b), 3.12(d) and 3.12(f), whereby the prominence of the faulty component can be inferred relative to the healthy region of the SMSE. By normalising the RMS results for the PD indicators, thereby removing the inclusion of any per-record non-stationarity, it is possible to quantify indicator sensitivity from the considered offline and online LVMs to the gear tooth fault.

Thus, the PD indicator improvement captured in the SNR^{dB}_{SMSE} results further highlight that the PD indicators are more sensitive to non-stationarity present in the data and less sensitive to the onset of the gear tooth damage. For the PPCA and the PPCA$_h$ models, the $LHI^{(3)}_{PD}$ and $LHI^{(5)}_{PD}$ metrics (the Canberra distance and the cosine distance respectively) have decreasing SNR trends, which indicates that these metrics respond via a reduction in magnitude around the gear tooth fault. This result was only observable for the $LHI^{(3)}_{PD}$ indicator when using the RMS diagnostic metric for the PPCA model, as demonstrated in Figure 3.12(b). This indicates that the latent response to the gear tooth fault is to project the anomalous data into regions where the latent representations lie closer together. The same observation can be made for the aforementioned metrics from the PPCA$_h$ model. This

result emphasises the need for a two-sided threshold when using the indicators. However, the aforementioned metrics for the SRICA model are insensitive to the onset and progression of the damage. These results further highlight the complementarity of the choice of training framework and LVM framework for CM, whereby the considered offline and online LVMs exhibit uniqueness in both the indicators that are sensitive to damage and in the damage manifestation mechanism within the latent manifold.

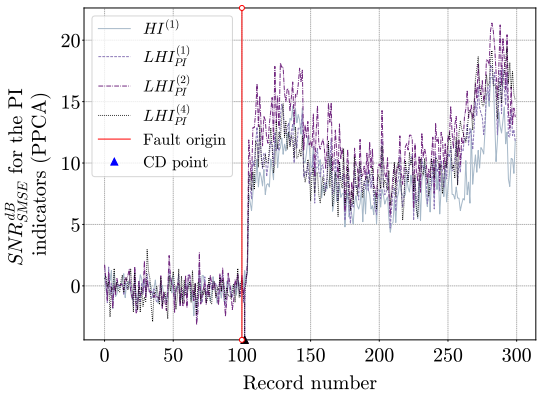
3.4.3 Comparison between the considered LVMs in an offline and online setting

The results from the experimental dataset investigation indicate that the online training approach is able to highlight and emphasise the characteristics of the considered online LVMs. The IMS dataset results indicate that the reconstruction-focused LVM framework has better damage onset detection sensitivity characteristics, while the interpretation-focused LVM framework is better suited to isolating the fault present in the data. The gearbox dataset results highlight that the considered LVM formulations are suitable for both of the considered training approaches. This highlights the complementarity of the considered online LVM frameworks in a fault diagnosis setting. The results also suggest that there is a fault type dependency when using reconstruction-focused and interpretation-focused online LVMs, but the optimal choice of formulation is not known a priori.

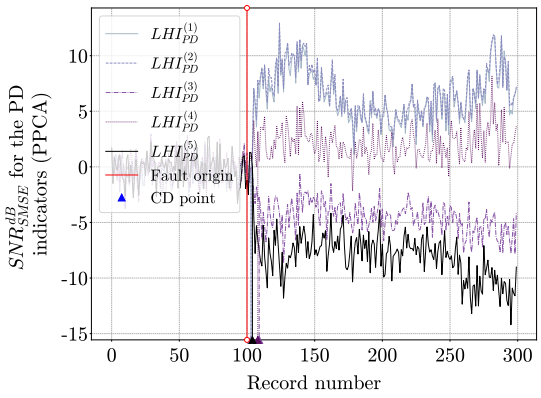
Moreover, the use of multiple latent indicators is necessary to ensure that different perspectives of the damage manifestation mechanisms within the latent manifold can be captured for condition inference. Hence, as the fault detection, trending, and isolation tasks were investigated for the LVMs developed using the online training approach and compared to an LVM developed using offline training, it can be concluded that the online training setting is a useful LVM training approach for reconstruction-focused and interpretation-focused online LVMs in vibration-based fault diagnostics.

3.5 Conclusion and recommendations

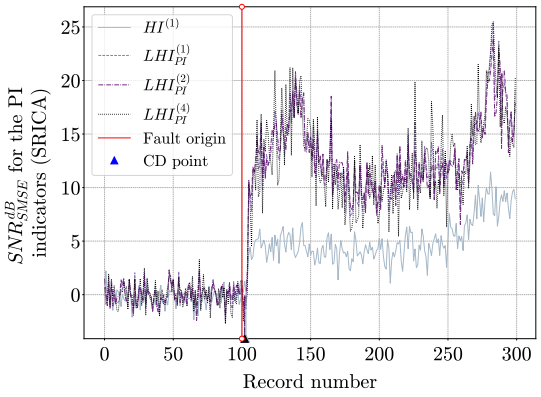
In this chapter, the implication of offline and online temporal-preserving LVM training approaches are investigated in the context of vibration-based gearbox fault diagnostics. The purpose is to address the second research direction, which aims to consider the scenario where there is no available historical asset data available to estimate the LVM parameters in an offline setting. Moreover, the objective is to understand to what extent the training approaches and the methods used therein play a role in the ability to perform fault diagnostics.



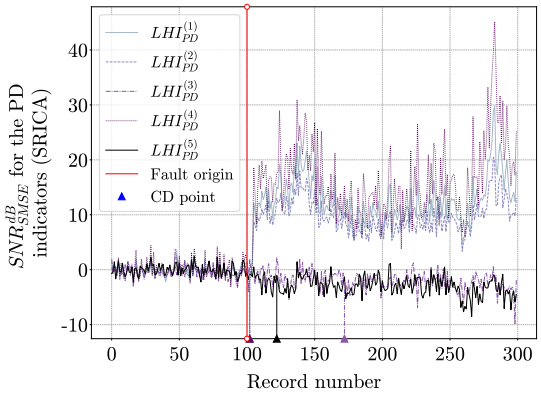
(a) PI components - PPCA.



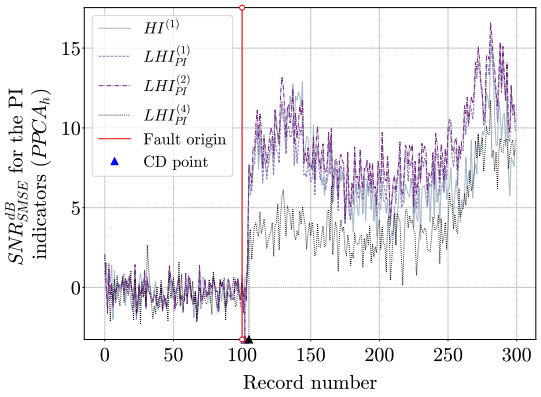
(b) PD components - PPCA.



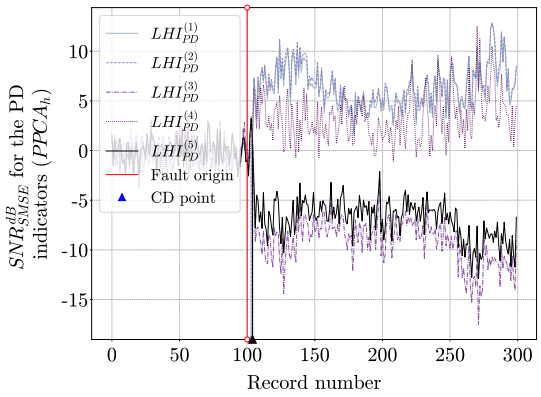
(c) PI components - SRICA.



(d) PD components - SRICA.



(e) PI components - PPCA_h.



(f) PD components - PPCA_h.

Figure 3.13. The SNR of the indicator SMSE for the considered offline and online LVMs applied to the gearbox dataset. The SNR calculation process is demonstrated in Figure A.2 in Appendix A. The indicators are grouped by LVM formulation and path traversal setting for comparison purposes.

Offline and online training approaches are both considered to ensure that the unique LVM characteristics and the implications thereof can be investigated and compared in this thesis for temporal-preserving LVM-based CM applications. A simple set of HIs and LHIs from the latent indicator framework proposed in Chapter 2 were adapted to investigate the performance of the considered offline and online LVMs for different gearbox fault conditions, and to study the implications of online LVM training.

The numerical and experimental investigations performed in this chapter indicate that online training of temporal-preserving LVMs is a useful approach for the vibration-based fault diagnostics task. The study highlights the implications of the choice of LVM formulation in the online training setting and confirms that these models can be used to detect changes in the instantaneous asset health state and identify faults under both constant and time-varying operating conditions. Additionally, using multiple LHIs for condition inference is beneficial, as the optimal LHI for fault diagnosis purposes is not known a priori and this choice directly affects the condition inference process. Reconstruction-focused online LVMs were demonstrated to be a useful tool to detect and trend variations in the instantaneous asset health state, while interpretation-focused online LVMs offered an improved ability to infer the fault condition present in the data.

Furthermore, complementarity in the choice of reconstruction-focused and interpretation-focused online LVM formulation is an important consideration when using online temporal-preserving LVMs, as a suitable online formulation for the condition inference task is not known a priori. The choice of LVM formulation has clear implications on the diagnostic information that can be extracted from the health indicators derived from the latent manifold. The comparison to a reconstruction-focused offline LVM indicated that both offline and online training approaches are viable avenues for condition inference. This indicates that data availability, an issue identified in the second research direction, is an important consideration for LVM development.

Hence, the second research direction has been suitably investigated and the results presented in this chapter indicate that both training modes are useful for gearbox CM. However, a sparse set of LHIs adapted from the first research direction were considered in this investigation. Hence, alternative LHIs must be investigated for online temporal-preserving LVMs to further study the implication of online training for the condition inference task. In the next chapter of this thesis, information from the previous two chapters is combined to

synthesise the third research direction, which considers how LHIs are extracted from the latent manifold of offline and online methods.

Chapter 4

Offline and online latent indicators for latent manifold-based condition inference

4.1 Introduction

Effective and informative health indicators are essential to aid with downstream tasks such as fault diagnostics and prognostics [5, 8]. Different stances are adopted in the literature to obtain health indicators using learning-based methods, and these stances can be connected by the idea that a statistical or probabilistic model is used to facilitate a transformation of the observed data, or some representation thereof, to generate health indicators [5]. Obtaining informative and fault-feature-sensitive health indicators is paramount to incorporating and effectively using learning-based methods in condition monitoring applications. Popular learning-based models are discussed in Section 1.2.2.2 of this thesis.

Learning-based models used to obtain health indicators require a parameter estimation step, and this is conventionally performed using a healthy set of historical vibration data signals, e.g., Heyns *et al.* [145]. This is referred to as offline training and is possible because critical assets are assumed to operate in a healthy condition primarily [94]. In applying temporal-preserving latent variable models, Chapter 2 of this thesis explored and proposed a latent indicator framework consisting of conceptual classes for health indicator derivation using the latent manifold of offline reconstruction-focused models. However, it was demonstrated in Chapter 3 of this thesis that allowing the latent manifold to dynamically adapt to the presence of fault signatures in the observed vibration data by estimating the model parameters iteratively, i.e., using online training, is an effective approach for the fault diagnosis task. This is suitable for scenarios without historical data from a healthy asset. Moreover, reconstruction-focused and interpretation-focused online models were demonstrated to exhibit a complementary capacity for the fault diagnosis task.

Therefore, offline and online methods are considered in this chapter as strategies to estimate latent variable model parameters and the latent health indicators used for condition

inference are thoroughly studied. This chapter encompasses the third research direction, whereby leveraging the latent manifold of offline and online models for fault diagnostics is critically examined and the scope and limitations of the latent indicator framework proposed from the first research direction is investigated. The latent health indicators derived from the framework proposed in Chapter 2 were designed for offline reconstruction-focused models and have several implicit assumptions in their derivation, which reduces their potential applicability to online methods. Thus, two categories of latent indicators are identified and advocated for in this chapter: *(i)* an offline indicator category that applies to the latent indicators from Chapter 2, and *(ii)* an online indicator category that addresses the potential implicit constraints of the offline indicators. Indicators from the latter category are referred to as online indicators, while the indicators from the former category are referred to as offline indicators.

The structure of this chapter is as follows: First, the necessary data pre-processing steps, the considered latent variable models, and the latent indicators available to the considered models from the offline indicator category and the proposed online indicator category are introduced and presented in Section 4.2. Second, an investigation of the considered offline and online indicators is conducted in Section 4.3 for two experimental gearbox failure datasets, which exhibit constant and time-varying operating conditions, respectively. Finally, the chapter is concluded in Section 4.4, and recommendations are made for future work.

4.2 Latent variable methods in vibration-based condition monitoring

In this thesis, temporal-preserving Latent Variable Models (LVMs) transform the vibration data and use Health Indicators (HIs) from the transformed data representation to detect, trend, and isolate fault signatures for the fault diagnosis task. Chapter 2 of this thesis examined how to effectively exploit the latent manifold for the gearbox fault diagnosis task for offline reconstruction-focused methods, and Chapter 3 investigated strategies to construct the latent manifold via offline and online training approaches and considered two distinct classes of online methods. In this chapter, temporal-preserving LVMs are extended to consider categories of Latent Health Indicators (LHIs) for offline and online methods. This aligns with the primary objective of the third research direction, aiming to enhance the adaptability of the latent indicator framework proposed in Chapter 2 by ensuring it is LVM formulation and training strategy agnostic for temporal-preserving LVM-based gearbox fault diagnostics. The

information presented in this section is compartmentalised into four stages. First, an overview of the investigative process is detailed. Second, the necessary data pre-processing steps are discussed. Subsequently, the considered LVM formulations are discussed. Finally, the offline and online indicators used in this investigation are presented and discussed.

4.2.1 Overview

In Chapter 3, online and offline models were compared using a sparse set of latent indicators and the study indicated that the online models performed well, which is especially beneficial if healthy historical data are not available. In Chapter 2, a diverse range of latent indicators were developed for offline reconstruction-focused LVMs and therefore it is essential that models developed using the two training approaches be compared using an equivalent representation of latent indicators. The interpretation of the two training modes in the vibration-based Condition Monitoring (CM) context is unique, whereby the former seeks intrinsic features of the healthy data and relies on these healthy features to perform condition inference, whereas the latter relies on the damage-related variation within the vibration data and a dynamic latent manifold to extract features that can effectively capture different fault conditions.

For the offline indicator framework proposed in Chapter 2, several implicit assumptions exist: First, the assumption made is that a set of latent samples is available obtained from the latent representations of the historical dataset. In the online training approach, the requirement for a reference dataset is negated. Secondly, the latent manifold dimensionality is assumed to be constant throughout the asset lifespan. This assumption is not always valid for models developed under the online training approach as the dimensionality of the latent manifold can be set to be a function of the observed data, e.g., when using the preserved variance ratio for variance-based models [101]. Finally, a distribution-based indicator class is proposed in Chapter 2, and assumes that the encoding transition function parametrises a Gaussian latent distribution. This assumption does not hold for all LVM formulations as the Gaussian assumption was made in light of the offline methods used in Chapter 2, which limits the generalisability to different model formulations.

A visual representation of the purpose of this investigation in the context of temporal-preserving latent variable model-based fault diagnostics is presented in Figure 4.1. The investigation performed in this chapter is focused on how the instantaneous health state of an asset is inferred from models developed using different LVM training approaches, and the objective is to compare each approach critically. This amalgamation of the first two

research questions discussed in Section 1.3.1 presents an opportunity to investigate different configurations of the latent variable model pipeline in a fault diagnostics context. The purpose is to analyse the LVM evaluation step, and to assess online and offline methods critically on an equal playing field.

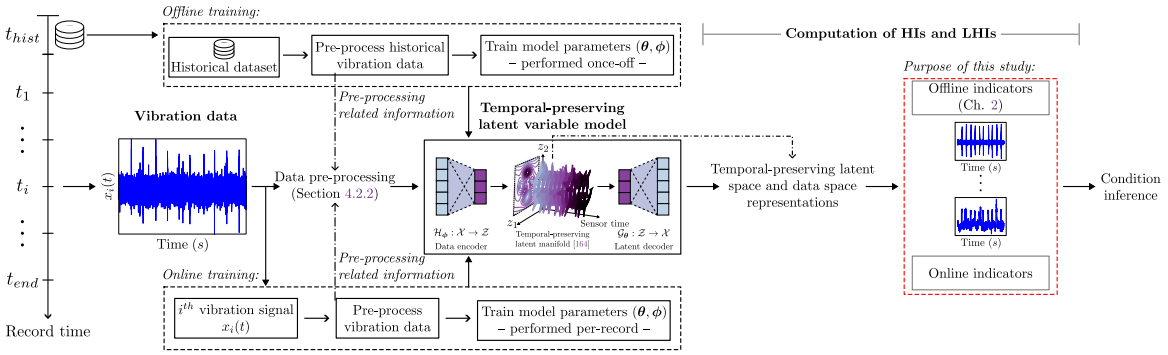


Figure 4.1. The LVM-based condition monitoring process investigated in this chapter. The focus of this study is centred around the computation of Health Indicators (HIs) and Latent Health Indicators (LHIs) using offline and online latent indicators for models developed using offline and online training. Note that the training approaches are separate, i.e., a hybrid approach is not considered. Offline training occurs before record time-step t_1 using historical vibration data, whereas online training is performed iteratively through record time. For the two considered training modes, there is a choice regarding the indicator category used. The offline indicators Chapter 2 are investigated for both modes. The online indicators proposed in this work are designed for LVMs whose parameters are estimated using online training, and hence are applied to online methods.

The training modes considered in this chapter allow for the incorporation of different LVM formulations. Following Chapters 2 and 3, a Probabilistic Principal Component Analysis (PPCA) model [280] represents a LVM developed using an offline training approach and is referred to as the PPCA$_h$ model. In online training, the LHIs computed from the latent manifold are a direct function of how the online training-based LVMs update and encode the information present in the data, and different representative methodologies used to construct the latent manifold can be investigated. Thus, two different LVM formulations are considered due to the complementarity in the choice of online formulation identified in Chapter 3: *(i)* reconstruction-focused LVMs, and *(ii)* interpretation-focused LVMs.

The rationale for the considered online formulations is established in Section 3.2.1 of this thesis. Investigating these two formulations using the online training approach is done to determine if a better suited LVM formulation exists for the fault diagnosis task in the scenario where the latent manifold is dynamically constructed. Additionally, Chapter 3 highlighted

the complementary fault diagnosis capacity of the considered methods for a basic set of indicators from Chapter 2, which suits the third research direction considered in this chapter regarding categories for LHIs. The online LVM formulations considered in this investigation are kept consistent with the investigation performed in Chapter 3, whereby the PPCA model operates as a reconstruction-focused online LVM and the Spectrally Regularised Independent Component Analysis (SRICA) model operates as the interpretation-focused online LVM.

4.2.2 Data pre-processing

As vibration data is representative data format of interest for this study, a segmentation-based Hankelisation strategy is employed to ensure that the observed data can be suitably represented as input features to the LVM [180, 181]. For a signal $x_i[n]$, where $n \subset \mathbb{Z}^+ = \{1, \cdots, N\}$ and N is the length of the signal, the Hankel matrix $\overline{\mathbf{X}}_i \in \mathbb{R}^{L_T \times L_w}$ is obtained via Equation (2.13). The rows of $\mathbf{X}_i$ represent segments of the data vector $\mathbf{x} \in \mathbb{R}^N$. Note that the variable i refers to record time, i.e., the sequential ordering given to the observed vibration data. Subsequently, the data matrix can be further processed depending on the LVM application setting. In the offline training approach, the reference dataset is used to collate the Hankel matrix representations of the reference vibration data and this is given in Equation (2.14). A standardisation procedure is then applied using the reference dataset and this procedure is given in Equation (2.15).

In the online training training approach, the data is not standardised and is merely centred using Equation (3.7), where the vector $\boldsymbol{\mu}_i \in \mathbb{R}^{L_w}$ is a column vector of the feature-wise means of $\overline{\mathbf{X}}_i$. The decision to only centre the data matrix $\overline{\mathbf{X}}_i$ when using the online training approach is sensible as this does not remove any per-record deviations in the data that may be useful for the condition inference process. For the interpretation-focused LVM considered in this investigation, a pre-whitening step is used to pre-process the data matrix $\mathbf{X}_i$ further. The pre-whitening step is given by a linear transformation as demonstrated in Equation (3.8). Pre-whitening is used to reduce the parameter approximation complexity by removing any per-feature correlations, i.e., producing a white covariance matrix [185].

4.2.3 Considered latent variable model methodologies

In this section of this thesis, much of the presented material is strongly connected to the work presented in Chapters 2 and 3. Hence, the discussion is presented for convenience and suitable reference is made to relevant sections of this thesis. In a LVM framework, as discussed in Section 2.3.1, the objective is to describe the distribution $p(\mathbf{x})$ over some observed data

$\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^D$ via the use of a set of unobserved, latent variables $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^d$, where $D > d$. The latent variable set $\mathbf{z}$ is used to capture the intrinsic features of the observed data [101, 282]. This implies that a marginalisation over the latent variables is required to obtain $p(\mathbf{x})$ as given in Equation (2.2), where $p(\mathbf{x}, \mathbf{z})$ represents the joint distribution over $\mathbf{x}$ and $\mathbf{z}$. In most LVM applications, a simple factorisation is applied to the joint distribution $p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x}|\mathbf{z})p(\mathbf{z})$, where $p(\mathbf{z})$ is regarded as the prior latent distribution, to obtain the data generative model $p(\mathbf{x}|\mathbf{z})$ via

$$\mathbf{x} = \mathcal{G}_{\boldsymbol{\theta}}(\mathbf{z}) + \boldsymbol{\epsilon}, \tag{4.1}$$

where $\mathcal{G}_{\boldsymbol{\theta}} : \mathbb{R}^d \to \mathbb{R}^D$ is a deterministic parametric decoding function with parameters $\boldsymbol{\theta}$, and $\boldsymbol{\epsilon}$ represents some data-independent noise. Different LVM formulations exist based on the approach the generative model given in Equation (4.1). Explicit reconstruction-focused LVMs, e.g., PPCA or Variational Auto-Encoders (VAEs) [165, 282], explicitly assume the form of the noise variable to define the conditional distribution $p(\mathbf{x}|\mathbf{z})$ and typically assume an isotropic Gaussian distribution for the prior $p(\mathbf{z})$. Implicit reconstruction-focused models, e.g., Generative Adversarial Networks (GANs) [167], define a stochastic data generation procedure using samples from an assumed form of the prior $p(\mathbf{z})$, and require no assumptions be given to $\boldsymbol{\epsilon}$. Alternatively, additional constraints can be applied to the prior distribution $p(\mathbf{z})$, e.g., statistical independence and non-Gaussianity as stipulated for interpretation-focused models such as the negentropy-based noise-free Independent Component Analysis (ICA) model [185].

In the scenario where latent inference is of interest, Bayes' rule can be used to infer the posterior distribution $p(\mathbf{z}|\mathbf{x})$, as given in Equation (2.3). However, this can often be intractable to evaluate [162, 165, 282]. Hence, latent inference often manifests via a mapping function $\mathcal{H}_{\boldsymbol{\phi}} : \mathbb{R}^D \to \mathbb{R}^d$, i.e., the latent encoder, given as

$$\mathbf{z} = \mathcal{H}_{\boldsymbol{\phi}}(\mathbf{x}), \tag{4.2}$$

which aims to operate as an approximate inverse to the generative model and, in some LVM formulations, parametrises an approximate posterior distribution $q(\mathbf{z}|\mathbf{x})$, e.g., VAEs [282]. To formalise the LVMs used, the reconstruction-focused and interpretation-focused LVM formulations are introduced. These formulations are presented both for convenience and for the purposes of the experimental investigations performed in Section 4.3 of this chapter.

4.2.4 Reconstruction-focused models

The reconstruction-focused LVM considered in this investigation is the PPCA model [280]. While this model has already been introduced in Section 3.2.2.1 of this thesis, certain equations are repeating and extended in this section. This is done for the purposes of accessibility, consistency, convenience, and to enable this section to be self-contained. In the PPCA model, the prior $p(\mathbf{z})$ and the generative distribution $p(\mathbf{x}|\mathbf{z})$ given in Equation (4.1) are assumed to both be Gaussian distributions, and the generative distribution is parametrised with a linear decoding function. The generative distribution for the PPCA model can be expressed via

$$p(\mathbf{x}|\mathbf{z}) = \mathcal{N}\left(\mathbf{x}|\mathbf{A}\mathbf{z} + \boldsymbol{\mu}_{\mathbf{x}}, \sigma^2_{PPCA}\mathbf{I}\right), \tag{4.3}$$

where $\mathbf{A} \in \mathbb{R}^{D\times d}$ defines the linear map from $\mathbf{z}$ to $\mathbf{x}$, $\boldsymbol{\mu}_{\mathbf{x}}$ represents the mean of $\mathbf{x}$, and σ^2_{PPCA} represents a variance of the conditional distribution. The posterior distribution $p(\mathbf{z}|\mathbf{x})$ for the PPCA model is given by

$$p(\mathbf{z}|\mathbf{x}) = \mathcal{N}\left(\mathbf{z}|\mathbf{M}^{-1}_{PPCA}\mathbf{A}^T\left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}}\right), \sigma^{-2}_{PPCA}\mathbf{M}_{PPCA}\right), \tag{4.4}$$

where $\mathbf{M}_{PPCA} = \mathbf{A}^T\mathbf{A} + \sigma^2_{PPCA}\cdot\mathbf{I}$. The maximum likelihood solution for $\mathbf{A}$ and σ^2_{PPCA} is given by

$$\mathbf{A}_{ML} = \mathbf{U}_d\left(\mathbf{L}_d - \sigma^2_{PPCA,ML}\mathbf{I}\right)^{1/2}\mathbf{R}, \tag{4.5}$$

$$\sigma^2_{PPCA,ML} = \frac{1}{D-d}\sum_{i=d+1}^{D}\lambda_i, \tag{4.6}$$

where $\mathbf{U}_d \in \mathbb{R}^{D\times d}$ is a matrix whose columns represent the first d eigenvectors of the covariance matrix $\mathbf{C} = \mathbb{E}\{\mathbf{x}\mathbf{x}^T\}$, $\mathbf{L}_d \in \mathbb{R}^{d\times d} = \text{diag}\left(\lambda_1, \cdots, \lambda_d\right)$ is a diagonal matrix of the first d eigenvalues of $\mathbf{C}$, $\mathbf{R} \in \mathbb{R}^{d\times d}$ is an arbitrary orthogonal matrix assumed to be an identity matrix, i.e., $\mathbf{R} = \mathbf{I}$, and $\sigma^2_{PPCA,ML}$ represents the average variance of the unused latent components [101, 280]. Selecting the dimensionality of the PPCA and PPCA$_h$ model is done using the CCR given in Equation (3.17) [101]. For the PPCA$_h$ model, i.e., the reconstruction-focused model developed using offline training, the model parameters are obtained using the reference set of observed vibration data signals. For the online training approach, the online PPCA model parameters are obtained for every vibration signal that is evaluated.

4.2.5 Interpretation-focused models

The interpretation-focused LVM considered in this investigation is based off an extension of the negentropy-based ICA formulation from the seminal work of Hyvärinen [185]. This model has already been introduced in Section 3.2.2.2 of this thesis, and portions of this information are

presented to provide more context to this model. This is done for the purposes of accessibility, consistency, and convenience to provide further comprehension to the interpretation-focused model. The generative model for the interpretation-focused LVM is assumed to be noiseless and may be given by

$$\mathbf{x} = \mathbf{A}\mathbf{z}, \tag{4.7}$$

where $\mathbf{A} \in \mathbb{R}^{D\times d}$. In the negentropy-based ICA formulation the latent variables are enforced to be independent, i.e., the joint latent distribution factorises

$$p(\mathbf{z}) = \prod_{j=1}^{d} p(z_j), \tag{4.8}$$

and the latent variables $\mathbf{z}$ are explicitly assumed to be non-Gaussian [178]. The encoding transformation to obtain the latent variables is represented by

$$\mathbf{z} = \mathbf{W}\mathbf{x}, \tag{4.9}$$

where $\mathbf{W} \in \mathbb{R}^{d\times D} = [\mathbf{w}_1, \cdots, \mathbf{w}_d]^T$ represents the pseudo-inverse of $\mathbf{A}$. To recover the latent variables, a simple motivation via the central limit theorem can prompt the decision to use non-Gaussianity as a proxy measure to statistical contrast. Under the assumption that the latent variables are independent, Equation (4.7) represents a weighted summation of independent latent variables. This implies, under the central limit theorem, that the distribution over $\mathbf{x}$ is closer to a Gaussian distribution than the latent distribution [185]. Hence, a proxy to non-Gaussianity can be used as a criterion to estimate the matrix $\mathbf{W}$ to obtain an informative latent space. Negentropy is a popular proxy measure and can be represented via

$$J(\mathbf{y}) = H_e(\mathbf{y}_{gauss}) - H_e(\mathbf{y}), \tag{4.10}$$

where $\mathbf{y}_{gauss}$ is a Gaussian random variable with the same variance as $\mathbf{y}$, and $H_e(\cdot)$ represents the differential entropy. Negentropy, in the context of Equation (4.9), can approximated via

$$J(z_j) \approx c \cdot \left(\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{G(\mathbf{w}_j^T\mathbf{x})\} - \mathbb{E}_{\nu\sim p(\nu)}\{G(\nu)\}\right)^2, \tag{4.11}$$

where $J(z_j)$ is the negentropy of the j^{th} latent component, $G(\cdot)$ is a non-quadratic function, $c \in \mathbb{R}^+$ is an arbitrary constant, $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is a standardised Gaussian variable with the same variance as z_j, and $\mathbf{x}$ is assumed to have been pre-processed using Equation (3.7) and Equation (3.8) [330]. The pre-processing steps are important to negentropy-based ICA methods as they simplify the computational effort required to estimate $\mathbf{W}$. Following the approach used in Section 3.2.2.2, the SRICA model is used in this investigation to represent the online interpretation-focused LVM. Additional details for the online LVMs are given in

Appendix A.6. The latent scaling proposed in Equation (3.9) is applied in this investigation.

4.2.6 Latent health indicator categories

In Chapter 2 a latent indicator framework comprising of conceptual classes is proposed for LVMs developed using the offline training approach, and offline latent indicators were derived from the proposed conceptual classes. In this chapter, a set of indicators is proposed for LVMs developed using the online training approach. Chapter 2 proposed five conceptual classes for indicator derivation in a LVM-based indicator framework, and this classification structure if imposed on the online indicators proposed herein for simplicity.

However, due to certain constraints imposed by the online training approach and the LVM formulations considered therein, not all indicator classes are available. Hence, a reformulation in light of these constraints is necessary. Three constraints are applied to obtain the proposed online indicators, namely *(i)* no access to a historical latent representation is provided, *(ii)* the potential of a non-constant latent dimensionality must be accounted for, and *(iii)* no assumptions can be made regarding the explicit posterior distribution $p(\mathbf{z}|\mathbf{x})$, which the latent encoding functions parametrise or represent. The offline indicators and online indicator categories identified in this chapter are now discussed.

4.2.6.1 Category one: Offline indicators

The offline indicators used in this investigation are obtained from Chapter 2 of this thesis, as presented in Table 4.1. The forms of these indicators are not shown in this chapter to reduce redundancy, and interested readers are referred to Section 2.3.5 for more information. An important implication for the models developed under the online training approach is that the indicators in Table 4.1 require a fixed latent dimensionality throughout the evaluation process. For the PPCA model developed using an online training approach, this implies that the CCR given in Equation (3.17) cannot be used to infer the latent dimensionality of the i^{th} vibration signal and a constant latent dimensionality must be used through record time. The implications of restricting the latent dimensionality of the PPCA model developed using the online training approach is discussed in the experimental investigation in Section 4.3.

4.2.6.2 Category two: Online indicators

The latent health indicators derived from the conceptual classes proposed in Chapter 2 using the constraints induced by the online training approach are detailed in this section. As the online training approach induces a dependency on record number, the variable i is

Table 4.1. The offline training framework indicators from Chapter 2 used in this investigation and their application to the LVMs considered in this chapter. As the SRICA model makes no assumptions related to the analytical expression of the latent distribution, the distribution-based indicators cannot be used. Note that for LVMs developed using the online training approach, a constant latent dimensionality is required.

Conceptual class	Indicator name	Reference	LVM formulation application		
			PPCA_h	PPCA	SRICA
Reconstruction-based	Reconstruction error ($HI^{(1)}$)	Equation (2.17)	✓	✓	✓
Likelihood-based	Gaussian mixture model ($LHI^{(1)}$)	Equation (2.19)	✓	✓	✓
	Generative topographic map ($LHI^{(2)}$)	Equation (2.21)			
Statistical distance: dissimilarity-based	Euclidean distance ($LHI^{(3)}_{PI\|PD}$)	Equation (2.23)	✓	✓	✓
	Manhattan distance ($LHI^{(4)}_{PI\|PD}$)	Equation (3.12)			
	Chebyshev distance ($LHI^{(5)}_{PI\|PD}$)	Equation (3.14)			
	Cosine distance ($LHI^{(6)}_{PD}$)	Equation (2.24)			
	Canberra distance ($LHI^{(7)}_{PD}$)	Equation (3.13)			
Statistical distance: distribution similarity	Bhattacharyya distance ($LHI^{(8)}_{PI\|PD}$)	Equation (2.25)	✓	✓	✗
	Symmetric KL divergence ($LHI^{(9)}_{PI\|PD}$)	Equation (2.26)			
Prototypical	K-medians distance ($LHI^{(10)}$)	Equation (2.29)	✓	✓	✓
Multivariate signal processing and analysis	Moving kurtosis ($LHI^{(11)}$)	Equation (2.31)	✓	✓	✓
	Moving RMS ($LHI^{(12)}$)	Equation (2.30)			

used exclusively to indicate a dependency on record time, e.g., the i^{th} vibration signal $x_i[n]$. Furthermore, the latent manifold is assumed to have temporal structure, and the subscript $t \in [0, L_T - 1]$ refers to an integer representation of the signal time induced by the temporal preservation approach [164]. This implies that the Condition Indicator (CI) signals, i.e., the HI and LHI signals, also depend on time, but this time dependency is not shown for brevity. Finally, to avoid confusion between the online and offline indicators, the subscript O is used to identify the online indicators. A feature inherent to the proposed online indicators is that they have components that update through record time and are sensitive to the implications of using the online training approach and the associated indicator constraints.

Likelihood-based

The first LHI proposed is the normalised Mahalanobis distance

$$LHI_O^{(1)} = \frac{1}{d_i} \cdot \sqrt{(\mathbf{z}_t - \boldsymbol{\mu}_{\mathbf{z},i})^T \boldsymbol{\Sigma}_{\mathbf{z},i} (\mathbf{z}_t - \boldsymbol{\mu}_{\mathbf{z},i})}, \tag{4.12}$$

where d_i is the dimensionality of the latent manifold and both $\boldsymbol{\mu}_{\mathbf{z},i}$ and $\boldsymbol{\Sigma}_{\mathbf{z},i}$ represent the latent mean and latent covariance estimated from the LVM latent representation of the i^{th} processed vibration signal.

Dissimilarity-based indicators

The dissimilarity-based metrics investigated in this chapter incorporates a simple normalisation by the latent dimensionality to ensure that they are latent dimension agnostic. To ensure that there is no confusion between the offline and online dissimilarity-based indicators, the same superscript notation as given in Table 4.1 is used. Hence, the denotation of a $LHI_O^{(2)}$ metric is not used in this investigation, for ease of result interpretation. The first metric used is the normalised Euclidean distance

$$LHI_{PI|PD,O}^{(3)} = \frac{1}{d_i} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_2, \tag{4.13}$$

where the Path Independent (PI) and Path Dependent (PD) indicator settings refer to the latent manifold comparison in a sensor time independent or sensor time dependent setting. In the PI setting, the latent variable $\mathbf{z}_t$ is set to a zero vector. The second metric is the normalised Manhattan distance

$$LHI_{PI|PD,O}^{(4)} = \frac{1}{d_i} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_1, \tag{4.14}$$

which is also normalised by the latent dimensionality. The third metric is the normalised Chebyshev distance

$$LHI_{PI|PD,O}^{(5)} = \frac{1}{d_i} \cdot \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_\infty, \tag{4.15}$$

which computes the maximum distance among the available latent dimensions d_i. The fourth dissimilarity-based metric is the normalised cosine distance

$$LHI_{PD}^{(6)} = \frac{1}{d_i} \cdot \cos^{-1}\left(\frac{\mathbf{z}_{t+1}^T \mathbf{z}_t}{\|\mathbf{z}_{t+1}\|_2 \cdot \|\mathbf{z}_t\|_2}\right), \tag{4.16}$$

which computes the distance in the inner product space, i.e., a non-Euclidean space. Equation (4.16) can only be used in the PD setting. The final dissimilarity-based metric is the normalised Canberra distance

$$LHI_{PD}^{(7)} = \frac{1}{d_i} \cdot \sum_{j=1}^{d_i} \frac{|z_{t+1,j} - z_{t,j}|}{|z_{t+1,j}| + |z_{t,j}|} \tag{4.17}$$

which represents a weighted version of the Manhattan distance given in Equation (4.14). Equation (4.17) can only be used in the PD setting.

Prototypical indicators

Indicators obtained from a prototypical manifold assume the existence of a fixed number of static prototypical elements to characterise the latent manifold. A k-means algorithm [101] is used to estimate the locations of the prototypical elements $\mathbf{P}_k^{(i)}$ and the silhouette coefficient [306] is used to determine the suitable number of elements N_p, $k \in [1, N_p]$. The

online indicator representation of the k-means distance is given by

$$LHI_O^{(8)} = \frac{1}{d_i \cdot N_p} \cdot \sum_{k=1}^{N_p} \|\mathbf{z}_t - \mathbf{P}_k^{(i)}\|_2, \tag{4.18}$$

which represents the normalised average distance from the available prototypical elements.

Multivariate signal indicators

In the multivariate signal indicator class, each latent dimension is regarded as an independent data channel and metrics are computed on the individual channels using signal processing-based techniques. Due to the inclusion of temporal preservation, each latent dimension can be treated as a signal with index t. The first indicator in this class is based on the latent signal envelope for the j^{th} latent dimension ($j \in [1, d_i]$), given by

$$E_j[n] = |z_j[n] + \sqrt{-1} \cdot H\left(z_j[n]\right)|, \tag{4.19}$$

where $n \in [0, L_T - 1]$, $\sqrt{-1}$ represents the imaginary unit, and $H(\cdot)$ represents the Hilbert transform [332]. The average envelope indicator is proposed in this chapter and is given by

$$LHI_O^{(9)} = \frac{1}{d_i} \cdot \sum_{j=1}^{d_i} E_j[n]. \tag{4.20}$$

The second indicator used in this class is based off Equation (4.20), but uses targeted features to identify latent signals with contain rich diagnostic information. Hence, each latent dimension is feature-weighted to incorporate information related to the operating conditions and the faulty component of interest, i.e., the characteristic fault frequency is known a priori. The purpose is to ascertain whether exploiting information related to the asset of interest can enhance latent signals useful to the fault diagnostics task. Different targeted indicators are discussed and compared in Smith et al. [32] and Chen et al. [334].

To compute the targeted features, the Square Envelope Spectrum (SES) of the latent signals is required and can be computed via

$$SES_j[k_\alpha] = \text{DFT}\{E_j^2[n]\}, \tag{4.21}$$

where DFT$\{\cdot\}$ represents the discrete Fourier transform. The first targeted measure is the Indicator of second-order cyclostationarity (ICS2) indicator [21], given by

$$ICS2_j = \sum_{l=1}^{N_l} \frac{\max_{k_\alpha \in \mathcal{A}_l} \left(|SES_j[k_\alpha]|^2\right)}{|SES_l[0]|^2}. \tag{4.22}$$

The second targeted measure is the Signal-to-Noise Ratio (SNR) indicator [33], given by

$$SNR_j = \sum_{l=1}^{N_l} \frac{\max_{k_\alpha \in \mathcal{A}_l} \left(|SES_j[k_\alpha]|^2\right)}{\text{median}_{k_\alpha \in \mathcal{B}_l} \left(|SES_j[k_\alpha]|^2\right)}, \tag{4.23}$$

where N_l represents the number of considered harmonics and median $(\cdot)$ represents the median operator. In Equation (4.22) and Equation (4.23), $\mathcal{A}_l$ represents a set of cyclic frequencies in a narrow band around the h^{th} harmonic of a specific characteristic cyclic frequency α_c, and $\mathcal{B}_l$ represents a wider band between the harmonics of interest to capture the noise floor around a specific characteristic cyclic frequency [32, 33, 335]. In this investigation, the first three harmonics of the characteristic frequency are considered, i.e., $N_l = 3$, for all investigations. These sets $\mathcal{A}_l$ and $\mathcal{B}_l$ are given by Equation (3.23) and Equation (3.24), respectively.

The online latent indicator that weights the latent envelope signals by the ICS2 indicator is given by

$$LHI_O^{(10)} = \frac{1}{d_i} \cdot \sum_{j=1}^{d_i} ICS2_j \cdot E_j[n], \tag{4.24}$$

and the online latent indicator that weights the latent envelope signals by the SNR indicator is given by

$$LHI_O^{(11)} = \frac{1}{d_i} \cdot \sum_{l=1}^{d_i} SNR_j \cdot E_j[n]. \tag{4.25}$$

The final two indicators used in this investigation are concerned with the distribution over the j^{th} latent envelope $E_j[n]$, and use a non-parametric estimate of this distribution to compute an online LHI. The distribution of interest is over the envelope and the log-envelope of the j^{th} latent signal. The log-envelope was selected as it has proven to offer benefits in domains where non-diagnostic signal components are present, e.g., cyclostationary components not related to damage or impulsive components [22, 32, 336].

To compute a health indicator for the latent envelope and log-envelope signals, the Negative Log-Likelihood (NLL) from a Kernel Density Estimation (KDE) of the distribution over $E_j[n]$ and $\log E_j[n]$ is used. The KDE model for a univariate variable $\nu \sim p(\nu)$ is given by

$$p_{KDE}(\nu) = \frac{1}{K} \cdot \sum_{k=1}^{K} \frac{1}{\sqrt{2\pi b^2}} \exp\left(-\frac{(\nu - \nu_k)^2}{2b^2}\right), \tag{4.26}$$

where K represents the samples from $[\nu_1, \cdots, \nu_K]$ used to estimate $p(\nu)$, b represents the bandwidth parameter used to control the influence of the samples [101]. In this investigation, the hyper-parameter b is estimated with the Silverman plug-in estimator [337]. Hence, the

KDE indicator for the envelope of the latent signals is defined as

$$LHI_O^{(12)} = -\frac{1}{d_i} \cdot \sum_{j=1}^{d_i} \ln \left[p_{KDE} \left(E_j[n] \right) \right], \tag{4.27}$$

and the indicator for the log-envelope of the latent signals is defined as

$$LHI_O^{(13)} = -\frac{1}{d_i} \cdot \sum_{j=1}^{d_i} \ln \left[p_{KDE} \left(\log_{10} E_j[n] \right) \right]. \tag{4.28}$$

Therefore, several online latent indicators are proposed in this chapter for LVMs developed using the online training approach. In Table 4.2, a summary of the proposed online latent indicators is given.

Table 4.2. The online indicators proposed in this thesis derived from the conceptual classes in Chapter 2 using the constraints induced by the online training approach. A constant latent dimensionality is not a prerequisite for the online indicators.

Conceptual class	Indicator name	Reference
Likelihood-based	Mahalanobis distance ($LHI_O^{(1)}$)	Equation (4.12)
Statistical distance: Dissimilarity-based (normalised)	Euclidean distance ($LHI_{PI\|PD,O}^{(3)}$)	Equation (4.13)
	Manhattan distance ($LHI_{PI\|PD,O}^{(4)}$)	Equation (4.14)
	Chebyshev distance ($LHI_{PI\|PD,O}^{(5)}$)	Equation (4.15)
	Cosine distance ($LHI_{PD,O}^{(6)}$)	Equation (4.16)
	Canberra distance ($LHI_{PD,O}^{(7)}$)	Equation (4.17)
Prototypical	Average k-means ($LHI_O^{(8)}$)	Equation (4.18)
Multivariate signal processing and analysis	Average envelope ($LHI_O^{(9)}$)	Equation (4.20)
	ICS2-weighted average envelope ($LHI_O^{(10)}$)	Equation (4.24)
	SNR-weighted average envelope ($LHI_O^{(11)}$)	Equation (4.25)
	Average KDE NLL - envelope ($LHI_O^{(12)}$)	Equation (4.27)
	Average KDE NLL - log-envelope ($LHI_O^{(13)}$)	Equation (4.28)

4.3 Experimental investigation

In this section, two experimental investigations are performed. In principle, the objective is to thoroughly investigate the implications of using the offline indicators for the LVMs developed using the offline and online training approaches, and compare the results to the online indicators obtained from the LVMs developed using online training. This is done to enable a comparative assessment of the implication of the LVM training framework and the choice of indicator category used to derive LHIs. In both investigations, the performance of the considered indicators for the fault detection, localisation, and trending tasks is examined. The

first dataset considered in this investigation contains bearing damage and constant Operating Conditions (OCs), and the second dataset contains gear tooth damage and time-varying OCs. The datasets are introduced in Section 2.4.1 and Section 2.4.2 of this thesis, respectively. Therefore, different failure modes are investigated and the implications of constant and time-varying OCs are considered to ensure diversity in the investigations performed.

The result analysis process performed in this investigation is presented in Figure 4.2. First, the analysis methodology used for the considered experimental dataset is defined. Subsequently, two diagnostic metrics are used to investigate the performance of the considered indicators for the offline and online LVMs. Third, an investigation into the statistical similarity between the indicator categories and the methods used is conducted to measure the similarity between the considered indicators. Finally, an indicator category diversity discussion is facilitated to emphasise the importance of diversity within the temporal-preserving LVM-based condition inference process.

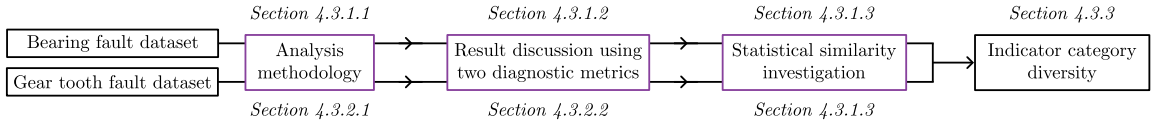


Figure 4.2. The result analysis process followed for the experimental investigation performed in this section.

The LVM formulations considered in this investigation are the PPCA$_h$ model developed using the offline training approach, and the PPCA and the SRICA models developed using the online training approach. The considered offline indicators are computed for the three models mentioned above, and the online indicators are applied to the PPCA model and the SRICA model. The LVMs used for online indicator evaluation are denoted as PPCA$_O$ and SRICA$_O$, to ensure that clarity is given to the investigation process. The purpose herein is to ascertain the implications of using the offline indicators for models developed under the two considered training approaches, and then investigate and compare the performance of the online indicators for the models developed using the online training approach. All parameters related to the LVMs required for investigation reproducibility are given in Section A.3 of Appendix A.

4.3.1 Bearing failure dataset

The LHIs from the LVMs considered in this investigation are compared on a popular benchmark dataset, which contains faulty bearing data. The experimental IMS bearing dataset, discussed in Section 2.4.1, is used in the first part of this investigation. A schematic of the experimental setup is depicted in Figure 2.6(a). In Figure 2.6(b), the statistical properties of the recorded data is presented via the Root-Mean-Square (RMS) and kurtosis of the raw vibration data. To represent the reference dataset for the PPCA$_h$ model, 5% of the available data measurements (49 measurements) is assumed as healthy.

4.3.1.1 Analysis methodology

To investigate the implications of the offline and online indicator sets, scalar diagnostic metrics are required to quantify the performance of the considered offline and online CI signals. The first diagnostic metric used in this investigation is the CI signal RMS metric, a popular CI metric for diagnosis and prognosis [19, 309]. The second diagnostic metric used in this investigation is the SNR of the SES of the CI signals. This metric is chosen as it represents the prominence of the damaged component in the CI signals and quantifies how well the considered indicators attribute their variation in condition, measured through the RMS, to the faulty component of interest. It is advantageous to have CI signals with prominent cyclic spectral information related to the fault condition of interest as this benefits the fault diagnosis task. Equation (4.22) calculates this metric, but it is computed using the CI signals and not the j^{th} latent component. In Section A.4 the procedure to calculate the SNR is demonstrated.

To evaluate the performance of the CI signals based off the two analysis metrics, an alarm threshold approach is used to identify a point of condition deviation. The threshold is given by Equation (3.18). The subset of measurements used is based off the measurements that are identified as part of the offline training set for the PPCA$_h$ model. A Condition Deviance (CD) point is identified when the average of five consecutive diagnostic metric measurements surpasses the alarm threshold T_a.

The final performance quantification metric used is the Pearson Correlation Coefficient (PCC) applied to the RMS and SNR trends through record time. This metric is used to quantify the response diversity of the offline and online indicators on an inter-method level and an inter-class level. The inter-method level refers to the similarity between the indicators from one LVM formulation and training approach with another, and the inter-class level refers

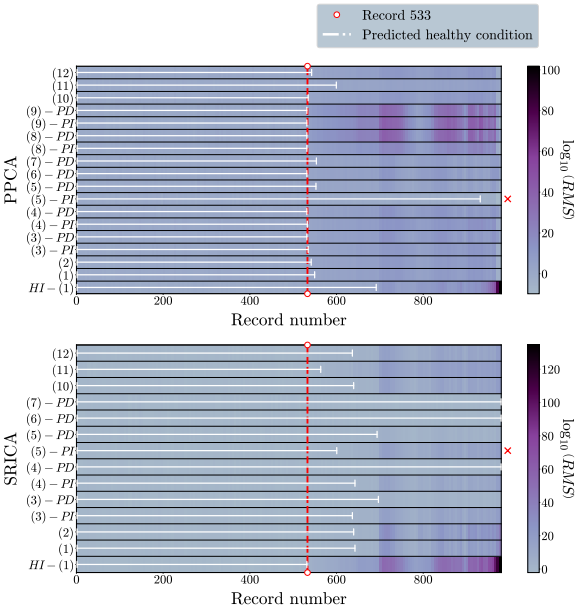
to the similarity between the offline and online indicators. The sample PCC, denoted as r_{xy}, is given in Equation (2.33).

4.3.1.2 Result discussion

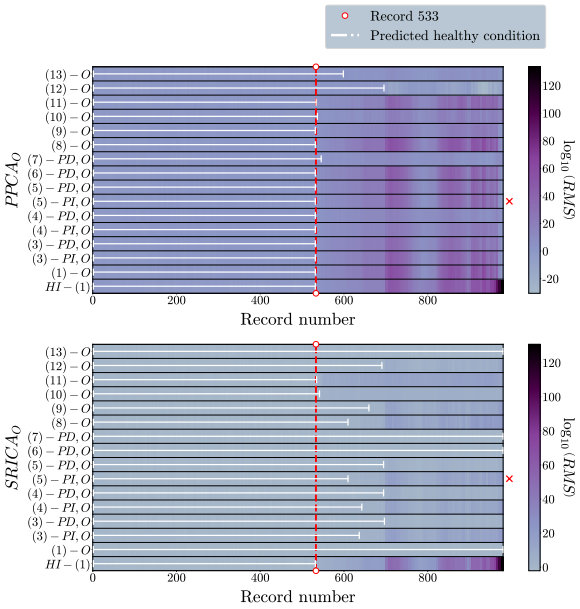
As the number of indicators considered in this investigation is large, it was decided to adapt the visualisation strategy to ensure that indicator legibility was viable without over-complicating and over-inflating the result visualisation process. Figure 4.3 presents the standardised $\log_{10}(RMS)$ diagnostic metric through record number for the considered condition indicators. The RMS diagnostic metric was log-scaled for result interpretability and standardised for comparison. The standardisation occurred using an estimate of the mean and variance from the first 5% of the available diagnostic metric values. This standardisation was done to ensure that the interpretation of the diagnostic metrics is not biased to the baseline scale of one metric over another.

The $\log_{10}(RMS)$ of the offline indicators from the PPCA and SRICA models developed using online training approach are presented in Figure 4.3(a). A constant latent dimensionality is required to evaluate the offline indicators, and the implications of this requirement are evident for the PPCA model. While many of the offline indicators from the PPCA model can detect a deviation point close to record 533, only the distribution-based indicators (i.e., $LHI^{(8)}_{PI|PD}$ and $LHI^{(9)}_{PI|PD}$) exhibit strong sensitivity from the baseline condition. This highlights the diversity in the choice of offline LHI for this model. The distribution-based indicators incorporate the latent posterior covariance ($\sigma^{-2}\mathbf{M}$ in Equation (4.4)) into their formulation. This suggests that there is important diagnostic information captured by the latent posterior covariance of the PPCA model.

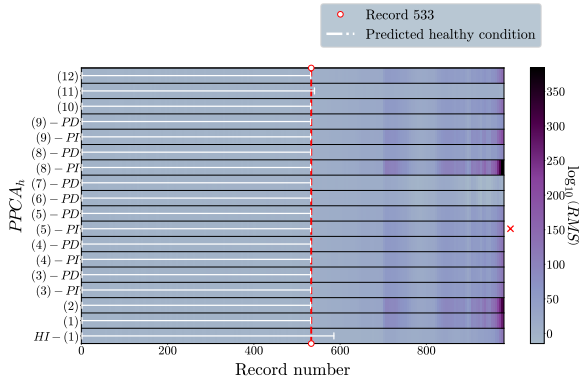
To understand this further, Figure 4.4 presents the latent dimensionality of the PPCA_O model and the constant dimensionality used for the PPCA model. It is clear that the PPCA_O model experiences significant latent dimensionality deviations through record number, and enforcing the static latent dimensionality requirement for the offline indicators into the PPCA model will decrease the magnitude of the preserved variance after record 533. This indicates that the latent posterior variance will undergo deviations in magnitude, which contains diagnostic information. Additionally, simply trending the fluctuation in latent dimensionality or the area under the CCR through record time produces a useful indicator for bearing damage present in this dataset.



(a) Offline indicators - PPCA and SRICA models.



(b) Online indicators - PPCA_O and SRICA_O models.



(c) Offline indicators - PPCA_h model.

Figure 4.3. The standardised $\log_{10}(RMS)$ diagnostic metric applied to the offline and online indicators from the considered LVMs for the bearing dataset investigation. In the y-axis of (a)-(e), the condition indicator name is represented by hyphenation of the subscript and superscript of the considered LHI to improve legibility. The reconstruction error metric ($HI^{(1)}$) is the exception, and is indicated as: $HI-(1)$. The LHIs used for the indicator diversity discussion performed in Section 4.3.3 are identified with a red cross.

The implication of the requirement for a constant latent dimensionality is further emphasised by Figure 4.5. The online indicators from the PPCA_O model are more sensitive to the bearing damage, as demonstrated by a comparison between Figure 4.5(a) and Figure 4.5(b), and incorporating the latent covariance is the only mechanism to obtain similar offline indicator sensitivity to the offline indicators for the PPCA model.

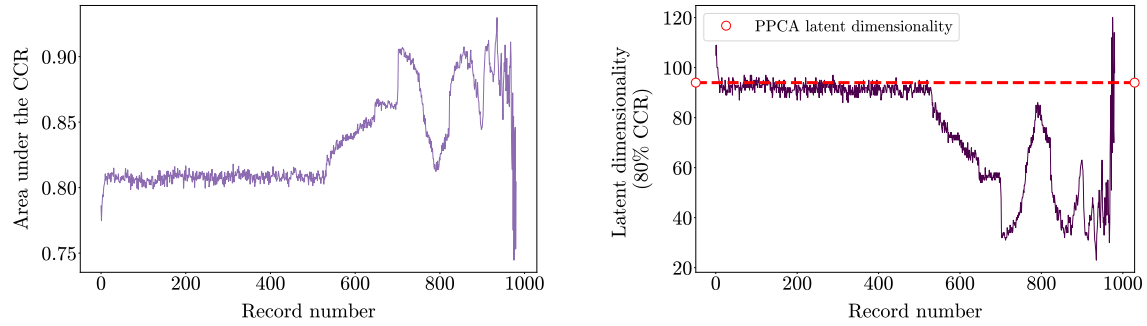
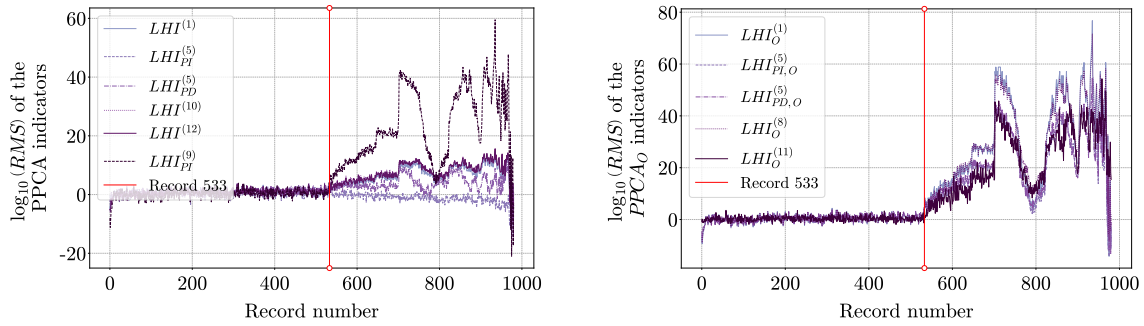


Figure 4.4. The (a) area under the CCR and the (b) latent dimensionality for a CCR of 80% for the PPCA$_O$ model developed using the online training approach. Notice the sensitivity in (a) and (b) to the damage in the bearing dataset. The constant latent dimensionality of the PPCA model, used for offline indicator evaluation, is constant through record number and is set to the average dimensionality of the first 5% of the available data. The area under the CCR, presented in (a), was obtained using a numerical integration technique and normalising by the latent dimensionality to limit the area to a range of $[0, 1]$.

For the SRICA model results obtained using the $\log_{10}(RMS)$ of the offline indicators, seen in Figure 4.3(a), weaker sensitivity and delayed CD identification are observed, with the exception of the $HI^{(1)}$ and $LHI^{(11)}$ indicators. This highlights the diversity in the indicator choice and suggests the SRICA model is sensitive to the choice of offline LHI used for fault detection and trending. In comparison, the online indicator results for the PPCA$_O$ and SRICA$_O$ models are presented in Figure 4.3(b). For the PPCA$_O$ model, improved CD sensitivity is observed for the considered online indicators in comparison to the offline indicators seen in Figure 4.3(a), as supplemented by the results observed in Figure 4.5 for the two indicator classes applied to the reconstruction-focused model. The reconstruction error ($HI^{(1)}$) gives improved performance over both the PPCA and the PPCA$_h$ model, indicating that allowing for latent dimensionality fluctuations through record number is necessary to obtain an informative data space indicator for the reconstruction-focused method. The $LHI_O^{(12)}$ indicator (i.e., the average KDE NLL of the latent envelope) undergoes a decrease in value from record 700, suggesting that the latent envelopes have less discernible impulsive information. This response also highlights the need for a two-sided alarm threshold as is defined in Equation (3.18).

For the online indicators from the SRICA$_O$ model, delayed CD identification points are observed in most indicators except for the ICS2 and SNR targeted latent indicators



(a) Offline indicators.

(b) Online indicators.

Figure 4.5. A subset of the PPCA and PPCA$_O$ standardised $\log_{10}(RMS)$ indicator responses used to demonstrate the implications of enforcing a constant latent dimensionality for the reconstruction-focused model developed using the online training approach. The indicators were selected from each conceptual category proposed in Chapter 2. Only the distribution-based offline indicator ($LHI_{PI}^{(9)}$) for the PPCA model exhibits a sensitivity similar to the online indicators from the PPCA$_O$ model.

(LHI_O^{11}) and $LHI_O^{(12)}$), which detect a change in condition around record 533. This suggests that the latent manifold of the SRICA$_O$ model contains information directly related to the fault condition of interest. However, the latent re-scaling approach used in Equation (3.9) to overcome the pre-whitening step cannot accurately enhance the latent component with prominent diagnostic information for early CD detection via the RMS diagnostic metric when applied to the non-targeted online indicator signals. The reconstruction error metric, $HI^{(1)}$, also is sensitive to the offset of damage, indicating that it is important to consider health indicators derived from both the latent manifold and the data space for condition inference.

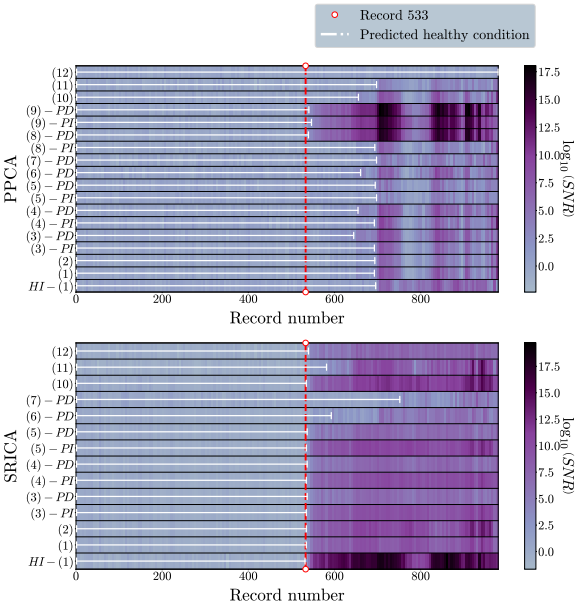
In Figure 4.3(c) the RMS of the offline indicators for the PPCA$_h$ model is presented. The offline indicators and the offline training approach produce clear and consistent CD points, except for the $HI^{(1)}$ metric, i.e., the reconstruction error. Interestingly, many of the PPCA$_h$ indicators, when grouped by conceptual class, appear to exhibit similarities is sensitivity on an intra-class level but not on an inter-class level. This suggests indicator diversity between the conceptual classes but intrinsic similarities between indicators from the same class. Similar results are observed in Section 2.4.1.3 in Chapter 2. The RMS diagnostic metric results indicate, in general, that the reconstruction-focused LVM is better suited to the fault detection task using both offline and online training. However, there is clear indicator diversity as a function of the indicator category, LVM formulation, and training approach used. This

diversity is perceptible in the variability in the identified CD points and in the sensitivity of the considered offline and online indicators.

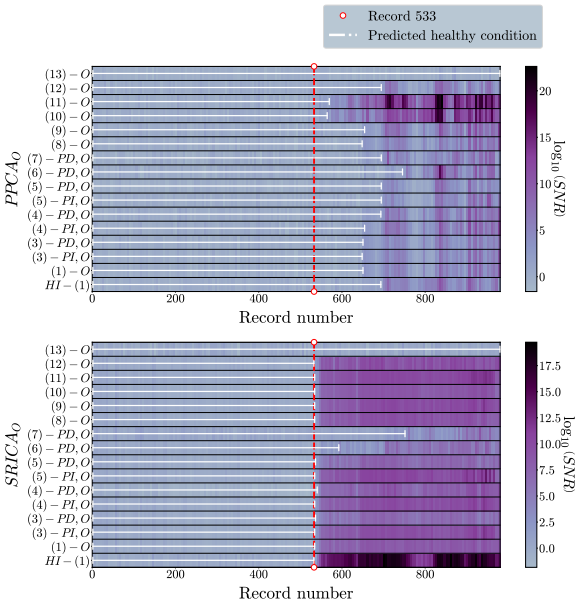
To ascertain and quantify the ability of the considered LVMs from the offline and online training approaches to isolate the faulty component of interest and express this information through the considered offline and online indicators, the SNR of the SES is used as a diagnostic metric. As for the RMS diagnostic metric used for the bearing dataset investigation, the results are log-scaled and standardised to enable consistent indicator comparisons.

The SNR diagnostic metric results for the PPCA, PPCA$_O$, and PPCA$_h$ model indicators, presented in Figures 4.6(a), 4.6(b), and 4.6(c), respectively, clearly indicate that the reconstruction-focused model used in this work is less suitable for the task of inferring the fault condition present in the data. For the reconstruction-focused models developed using the online training approach, i.e., the PPCA and PPCA$_O$ models, only the distribution-based offline indicators ($LHI^{(8)}_{PI|PD}$ and $LHI^{(9)}_{PI|PD}$) and the targeted online indicators ($LHI^{(10)}_{O}$ and $LHI^{(11)}_{O}$) obtain CD identification points before record 600. For the reconstruction-focused model developed using the offline training approach, i.e., the PPCA$_h$ model, only the distribution-based offline indicators (excluding the $LHI^{(8)}_{PI}$ metric) and the $LHI^{(5)}_{PI|PD}$ metrics provide CD identification before record 600.

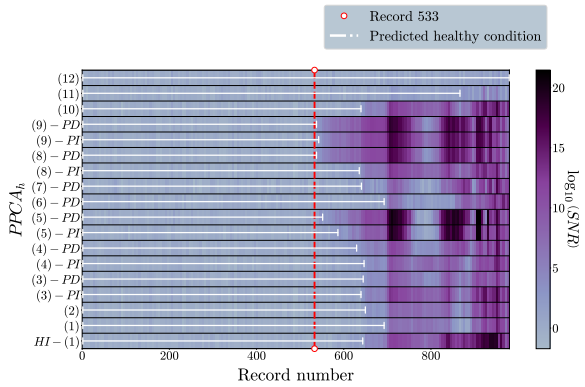
In comparison, the SNR diagnostic metric results for the interpretation-focused LVM developed using the online training approach, presented in Figures 4.6(b) and 4.6(a) respectively, indicate that the fault type can be inferred and the prominence of the faulty component in the SES is dependent on the indicator type and class. The results suggest that the use of online training and alternative LVM formulations induce important considerations when using LVMs for condition inference in this training setting. Notably, the $LHI^{(7)}_{PD,O}$ and $LHI^{(13)}_{O}$ metrics (the Canberra distance and the NLL of the log-envelope respectively) demonstrate poor sensitivity in both diagnostic metrics for the SRICA$_O$ model, as presented in Figures 4.3(b) and 4.6(b) respectively. This indicates that these are non-informative CIs for the considered model for this dataset. There is clear evidence of diversity in the considered offline indicators and online indicators that is dependent on the training approach used and the LVM formulation used. Therefore, investigating and quantifying this diversity is important to this study.



(a) Offline indicators - PPCA and SRICA models.



(b) Online indicators - PPCA_O and SRICA_O models.



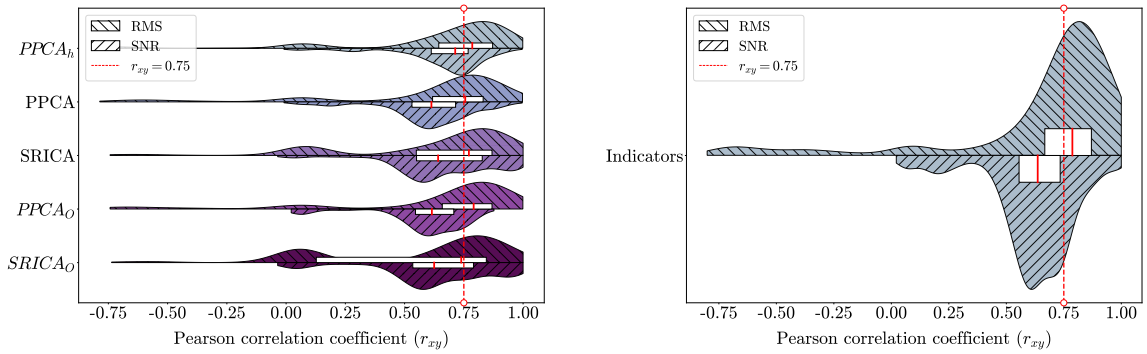
(c) Offline indicators - PPCA_h model.

Figure 4.6. The log-scaled and standardised SNR of the SES diagnostic metric applied to the offline and online indicators from the considered LVMs for the bearing dataset investigation.

4.3.1.3 Statistical similarity investigation

Figure 4.7 presents the PCC on an inter-method and inter-category level applied to both the RMS diagnostic metric and the SNR of the SES diagnostic metric. The purpose of this is to quantify whether there is indicator diversity as a function of the training approach and LVM formulation used, and indicator diversity as a function of the indicator category. Figure 4.7(a) indicates indicator diversity between the considered training approaches and LVM formulations. Hence, there is clear diversity in both the RMS and SNR results as a function of the considered LVM approach to the CM problem.

The RMS metric correlation is, on average, higher than the SNR correlation, emphasising the variability in the ability to localise the fault condition present in the data using the considered LVMs. The negative correlation in the RMS metric is attributed to the $LHI_O^{(12)}$ metric response from the PPCA$_O$ model seen in Figure 4.3(b). Figure 4.7(b) indicates that there is diversity in the choice of indicator category used for condition inference. Diversity in the considered latent indicator categories demonstrates that the choice of LVM training approach and LVM formulation has unique implications for the CM task for this dataset. Furthermore, the considered indicators capture unique perspectives, and the proposed online indicators are useful to supplement scenarios where the implicit assumptions are encoded in the offline indicators from Chapter 2 that reduces their effectiveness. Hence, it is important to consider the indicator category used for the LVM-based condition inference task.



(a) Inter-method similarity.

(b) Inter-category similarity.

Figure 4.7. The PCC of the considered diagnostic metrics for the bearing dataset investigation. (a) indicates the inter-method diversity as a function of the LVM formulation and training approach used, while (b) indicates the inter-category diversity for the two indicator categories regardless of the choice of formulation or training approach.

4.3.2 Gear tooth failure dataset

For the second investigation performed in this chapter, the LHIs from the considered LVMs are compared on a gear fatigue experimental dataset, which is discussed in detail in Section 2.4.2. For localised gear tooth damage, the amplitude modulation induced by the fault is synchronous with the gear shaft rotation and is confined around the order of a tooth-mesh period [36]. As the shaft encoder is attached to the gear shaft, a characteristic cyclic frequency of 1.0 shaft orders is expected. To represent the reference dataset for the PPCA$_h$ model, 10% (10 measurements) of the available healthy gearbox data measurements is assumed to be healthy.

4.3.2.1 Analysis methodology

Synchronous representations of the considered indicators are used in this chapter for the fault diagnostics task as they are an effective approach to enable fault diagnostics in gear damage applications [5, 209, 338]. These representations provide useful information related to specific asset components, and simple post-processing steps can be applied to these representations to enable fault diagnostics [333, 339]. Order tracking is an implicit requirement to obtain synchronous representations [209, 338], and the computed offline and online CI signals are order tracked using angular re-sampling achieved with the provided tachometer information to represent the signals as functions of multiples of the reference shaft speed.

The Synchronous Median of the Square Envelope (SMSE) [333] is the synchronous representation used in this investigation to compare the performance of the offline and online indicators. The SMSE of the i^{th} order tracked CI signal is given by

$$SMSE(\phi, \Phi_k) = \text{median}\left(E^2_{CI}(\phi, \Phi_k)\right), \tag{4.29}$$

where $E^2_{CI}(\phi, \Phi_k)$ is the order tracked square envelope of the j^{th} CI signal, $\phi \in [0, \Phi_k)$ refers to the angular position, and Φ_k refers to the cyclic period. For the considered dataset, gear tooth deviations are expected to be synchronous with the gear shaft rotation, i.e., $\Phi_k = 2 \cdot \pi$. The SMSE is robust estimator in the presence of impulsive non-synchronous components and uses the square envelope to attenuate the implications of phase shift in the synchronous representation due to time-varying OCs [15, 333]. Hence, it is well suited for the gearbox dataset investigation.

As the gear tooth fault is known a priori, the objective is to investigate the sensitivity of the offline and online CIs to this fault type. Two post-processing diagnostic measures are used as performance metrics for this dataset. The first diagnostic measure is the RMS of the SMSE region corresponding to the gear tooth fault. The second diagnostic measure is the relative SNR of the SMSE signals and is determined via

$$SNR_{SMSE} = \frac{|P^{SMSE}_{damaged} - P^{SMSE}_{healthy}|}{P^{SMSE}_{healthy}}, \tag{4.30}$$

where $P^{SMSE}_{damaged}$ and $P^{SMSE}_{healthy}$ represent the power of the damaged portion and the healthy portion respectively. Figure A.2 in Section A.4 demonstrates the calculation procedure for Equation (4.30).

The absolute deviation is computed in the numerator of Equation (4.30) to ensure that

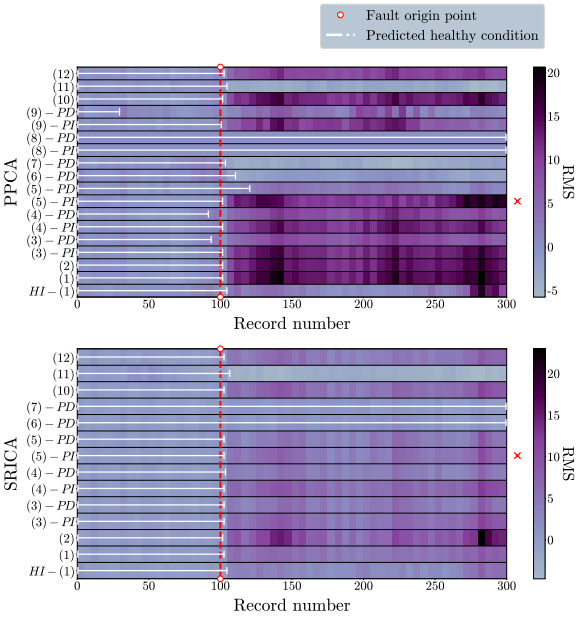
the measure is sensitive to relative deviation of the SMSE from the healthy region. This ensures that the metric is not sensitive to whether the CIs increase or decrease in magnitude around the gear tooth fault region. The RMS is used to detect indicator deviations, and the SNR_{SMSE} is used to compare the prominence of the tooth fault relative to the healthy condition. The correct healthy and unhealthy regions are extracted from the SMSE of the indicators signals using the location of the tachometer butt joint obtained from the Bayesian Geometry Compensation (BGC) algorithm proposed in Diamond *et al.* [316]. The alarm threshold approach given in Equation (3.18) is applied to the diagnostic metrics used in this investigation to identify a CD point. Finally, the Pearson sample correlation coefficient, given in Equation (2.33), of the RMS and SNR_{SMSE} trends is evaluated to assess the diversity in the indicator trends.

4.3.2.2 Result discussion

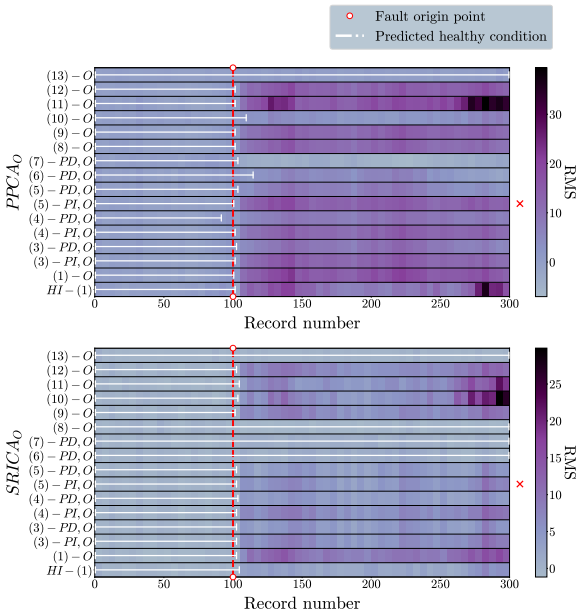
Figure 4.8 presents the RMS of the considered indicators for the gear tooth failure dataset. As was done for the bearing failure investigation, the results of the considered diagnostic metrics are standardised using the first 10% of the healthy indicator signals for consistent result interpretation. Many of the considered indicators are sensitive to the change in gear condition at record 100. This indicates that the considered LVMs are sensitive to the change in condition induced by seeding the tooth fault; however, further investigation is required to ascertain whether the change in the indicators is due to the degradation of the tooth condition or due to the fault seeding process.

The PPCA and PPCA$_O$ results are presented in Figures 4.8(a) and 4.8(b) respectively. The $LHI^{(8)}_{PI|PD}$ indicator from the PPCA model is unresponsive but the $LHI^{(9)}_{PI|PD}$ (an alternative distribution-based offline indicator) exhibits sensitivity to damage, further suggesting that how the latent covariance term σ^{-2} in Equation (4.4) is incorporated into the distribution-based offline CIs is an important consideration. However, the $LHI^{(9)}_{PD}$ metric erroneously identifies a CD point early in the healthy gearbox stage. The PPCA model and the PPCA$_h$ model results, compared to the bearing failure investigation in Section 4.3.1, suggest that the implication of the static latent dimensionality requirement for the offline indicators is less pronounced for this dataset. Additionally, latent and data space health indicators are observed to be sensitive to the fault condition present, indicating that both spaces are informative for the fault diagnosis task.

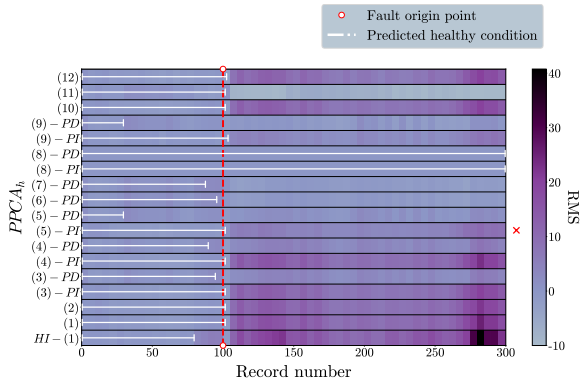
The latent dimensionality fluctuations for the healthy and unhealthy partitions of the



(a) Offline indicators - PPCA and SRICA models.



(b) Online indicators - PPCA_O and SRICA_O models.

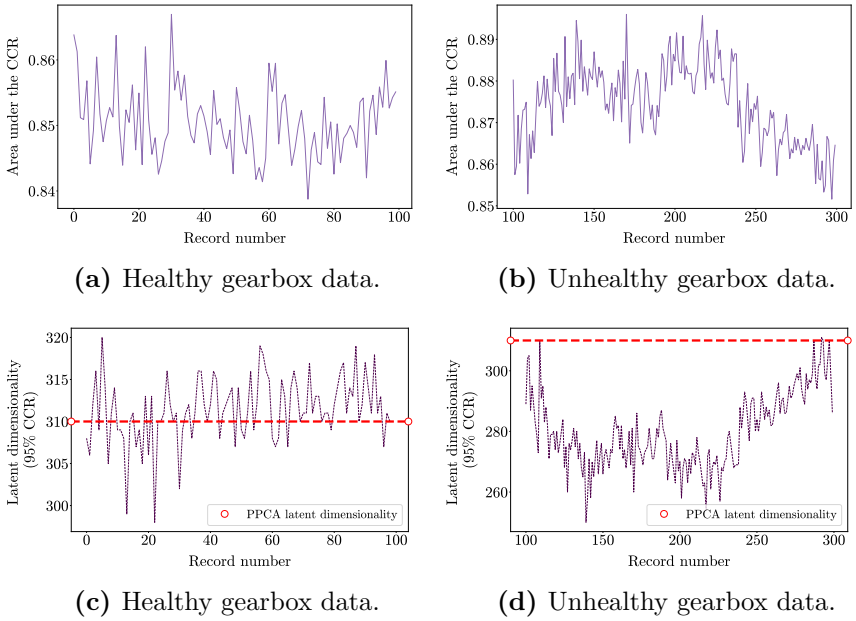


(c) Offline indicators - PPCA_h model.

Figure 4.8. The standardised RMS of the SMSE diagnostic metric applied to the offline and online indicators from the considered LVMs for the gear tooth dataset investigation. The LHIs used for the indicator diversity discussion performed in Section 4.3.3 are identified with a red cross.

gearbox dataset are presented in Figure 4.9. Figures 4.9(a) and 4.9(c) indicate that before introducing the gear tooth fault, the PPCA_O model has latent dimensionality fluctuations attributed to the non-stationarity in the data. Figures 4.9(b) and 4.9(d) indicate that after introducing the gear tooth fault there are noticeable deviations in the latent dimensionality and less latent components are required to represent the variance in the data. These fluctuations are attributed to both the non-stationarity in the data and the degradation of the gear health state over the experimental lifespan. However, the differences in the results presented

in Figures 4.8(a) and 4.8(b) are less noticeable than those observed in the bearing failure investigation.



(a) Healthy gearbox data. **(b)** Unhealthy gearbox data.

(c) Healthy gearbox data. **(d)** Unhealthy gearbox data.

Figure 4.9. The (a), (b) area under the CCR and the (c), (d) latent dimensionality for a CCR of 95% for the PPCA$_O$ model developed using the online training approach. Notice the sensitivity in the latent dimensionality and CCR to the damage in the unhealthy gear tooth dataset. The constant latent dimensionality of the PPCA model, used for offline indicator evaluation, is constant through record number and is set to the average dimensionality of the first 10% of the available healthy data.

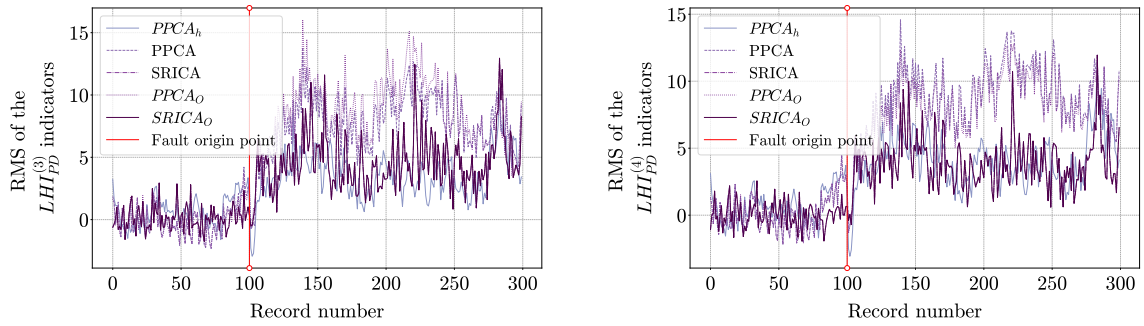
The offline indicator results from the PPCA and PPCA$_h$ models suggest that: *(i)* the selected CCR is suitable to ensure that both models capture the fault information in the latent manifold, and *(ii)* the implications of the choice of LHI obtained from the offline or online indicator categories are less distinct for this dataset. This result emphasises the notion of diversity when using LVMs for condition inference, as the suitability of the indicator category is not known a priori, and both indicator categories have the potential to be useful to the condition inference task. Sensitivity to non-stationarity in the data is detected in Figure 4.8 for all of the reconstruction-focused models, whereby certain indicators identify CD points before record 100. This suggests that the variance-driven LVM formulation has sensitivity to the data non-stationarity, and this sensitivity is detectable in a subset of the offline and online indicators.

Figure 4.8(a) and Figure 4.8(b) indicate that the interpretation-focused LVMs (the SRICA and SRICA$_O$ models) and their associated indicators are sensitive to the gear tooth fault, and

the targeted features ($LHI_O^{(10)}$ and $LHI_O^{(11)}$) are less critical for the RMS-based fault diagnosis task, unlike the bearing dataset investigation. Additionally, the offline and online indicators exhibit no early CD point identification, which indicates that the interpretation-focused LVM is less sensitive to the non-stationarity in the data. It is evident that the implication of using offline or online indicators for the interpretation-focused LVM developed using the online training approach is subtle, and many indicators can detect a change in condition at record 100. The $LHI_{PD,O}^{(6)}$, $LHI_{PD,O}^{(7)}$, $LHI_O^{(8)}$, and LHI_O^{13} indicators are insensitive to the change in condition, indicating that these online indicators are ill-suited to the condition inference task for the SRICA$_O$ model applied to the gearbox dataset. The insensitive metrics are spread throughout the conceptual classes for the online latent indicators, which emphasises the need to consider a diverse set of LHIs from different indicator classes. The data space indicator, $HI^{(1)}$, exhibits less sensitivity than the latent indicators derived from the latent manifold, suggesting that the latent manifold from the interpretation-focused model contains diagnostic information useful for the condition inference task.

The RMS of the offline indicators from the PPCA$_h$ model is presented in Figure 4.8(c), where it is observed that many metrics are useful to the fault diagnostics task. Interestingly, the $HI^{(1)}$ metric (the reconstruction error) produced an early and erroneous CD identification point, suggesting sensitivity to non-fault related information in the data. This sensitivity is also observed for the $LHI_{PD}^{(5)}$ metric and the $LHI_{PD}^{(9)}$ metric. One caveat of the diagnostic metric visualisation approach used in this work is the inability to discern whether the observed CD points are identified due to the fault development operation or the degradation of the tooth fault. It was found that the PD setting for the considered indicators produced results that are more sensitive to the non-stationarity in the data. The RMS diagnostic metric trends of the $LHI_{PD}^{(3)}$ and $LHI_{PD}^{(4)}$ indicators are presented in Figure 4.10 to demonstrate this behaviour. This indicates that the PD setting is more sensitive to fluctuations not due to damage than the gear tooth damage itself. This highlights the diversity in the offline and online indicators due to their derivation approach, whereby the PD indicators are more sensitive to non-fault related information for this dataset. Hence, the performance of the considered indicators is further investigated using the SNR_{SMSE} diagnostic metric to ascertain whether the gear tooth damage can be detected using a metric that potentially removes the sensitivity to the gearbox data non-stationarity.

The SNR_{SMSE} of the considered offline and online indicators for the gear tooth failure

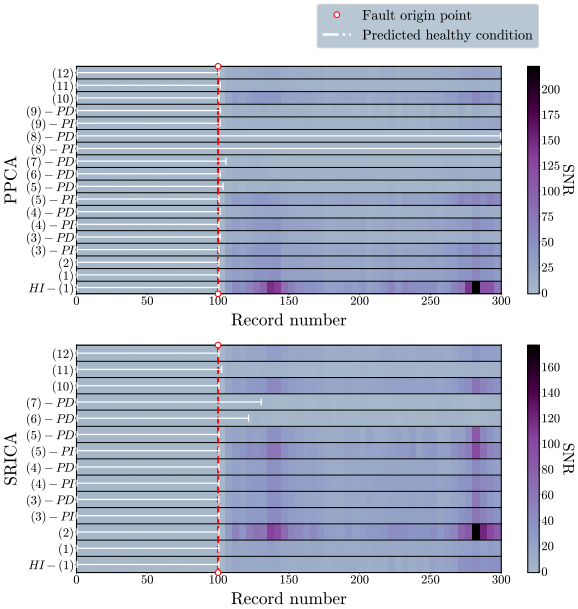


(a) $LHI_{PD}^{(3)}$ indicators. **(b)** $LHI_{PD}^{(4)}$ indicators.

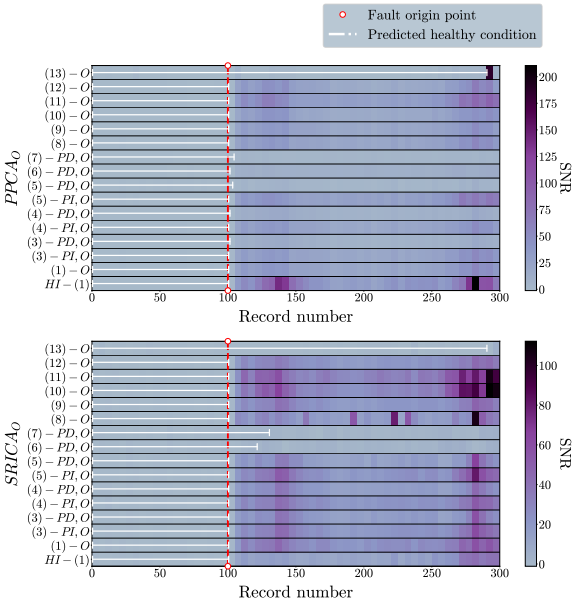
Figure 4.10. The RMS diagnostic metric trends for the (a) $LHI_{PD}^{(3)}$ and (b) $LHI_{PD}^{(4)}$ metrics considered in this chapter. Notice the clear diagnostic metric sensitivity to the non-stationarity in the data for the considered LVMs.

dataset is presented in Figure 4.11. To ensure consistent metric sensitivity comparisons are possible, the metric trends are standardised using an estimate of the mean and standard deviation from the first ten diagnostic measurements. The SNR_{SMSE} diagnostic metric uses an estimate of the power in the noise floor as a normalisation factor, which removes any potential sensitivity to the non-stationarity in the data. The implication of the normalising factor is quantified in Figure 4.12, which presents the SNR_{SMSE} diagnostic metric for the $LHI_{PD}^{(3)}$ and $LHI_{PD}^{(4)}$ indicators. The sensitivity to the data non-stationarity is effectively removed, as observed in the comparison of Figure 4.10 to Figure 4.10, and sensitivity to changes in the instantaneous gear health state is detectable when using the SNR_{SMSE} metric on the PD indicators. This highlights that metrics that are PI in nature, which are the majority of the considered offline and online indicators, are potentially better suited to the condition inference task for this dataset.

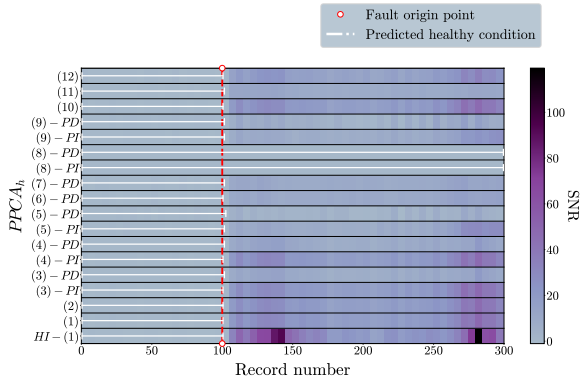
It is evident from Figure 4.11 that both the offline and online categories produce indicators that are useful for the condition inference task. Additionally, there are obvious indicators that respond weakly to damage, which indicates diversity in the selected indicators from the offline and online categories. Figure 4.11(a) indicates that the $LHI_{PI|PD}^{(8)}$ offline indicators are ill-suited to the diagnostics problem for the considered dataset. Figure 4.11(a) also indicates a subtleness in the implication of the LVM formulation for the offline indicators applied to LVMs developed using the online training approach, as both formulations are sensitive to the damage in the dataset. The indicators from the PPCA and SRICA models respond to the



(a) Offline indicators - PPCA and SRICA models.



(b) Online indicators - PPCA_O and SRICA_O models.



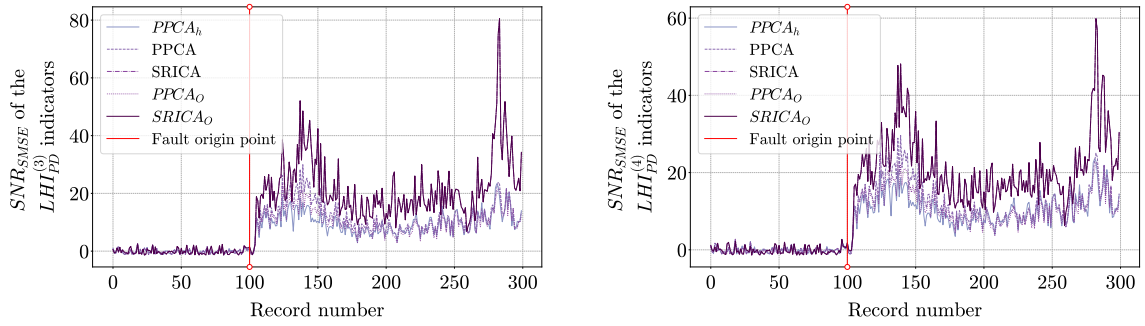
(c) Offline indicators - PPCA_h model.

Figure 4.11. The relative SNR of the SMSE diagnostic metric applied to the offline and online indicators from the considered LVMs for the gear tooth dataset investigation.

damage in the data and identify a point of CD after the fault origin point. Still, the choice of LVM formulation affects the sensitivity of the indicators.

For the *PPCA* model, the offline PD indicators are less sensitive than their PI counterparts; however, this was not observed to be true for the offline PD indicators from the SRICA model. Figure 4.11(b) presents the online indicators for the PPCA_O and SRICA_O models. The $LHI_O^{(13)}$ metric (the NLL of the log envelope) is a poor online indicator for the considered dataset for the LVM formulations developed using online training. For the PPCA_O model, the PD online indicators exhibit weaker sensitivity to the gear tooth fault. However, this is

not, in general, the same for the PD online indicators from the SRICA$_O$ model. Although the $LHI^{(6)}_{PD,O}$ and $LHI^{(7)}_{PD,O}$ give delayed CD identification and are less sensitive, this does not hold for the other considered online PD indicators, e.g., $LHI^{(5)}_{PI,O}$ versus $LHI^{(5)}_{PD,O}$. Figure 4.11(b) indicates that the $LHI^{(8)}_{O}$ metric (the average k-means distance) for the SRICA$_O$ model has spurious regions of anomalous behaviour. Upon further investigation, it was found that these spurious regions corresponded to records where the SMSE of the indicator had an uncharacteristically low magnitude, and the SMSE signals had highly variable baseline offsets throughout record time, thereby indicating sensitivity to non-stationarity in the gearbox data. This is supplemented by the results presented in Figure 4.8(b), whereby no CD point identification was obtained for the $LHI^{(8)}_{O}$ metric.



(a) $LHI^{(3)}_{PD}$. **(b)** $LHI^{(4)}_{PD}$.

Figure 4.12. The SNR_{SMSE} diagnostic metric trends for the (a) $LHI^{(3)}_{PD}$ and (b) $LHI^{(4)}_{PD}$ metrics considered in this chapter. Notice how the sensitivity to non-stationarity in the data, as observed in Figure 4.10, is suppressed. This is a direct result of the normalisation performed when calculating the SNR_{SMSE} metric.

The SNR results for the PPCA$_h$ model are presented in Figure 4.11(c). The $LHI^{(8)}_{PI|PD}$ metric from the PPCA$_h$ model is not useful to the condition inference task for this dataset, and this is a shared result for the offline indicators applied to the PPCA model. Furthermore, in a direct comparison of the offline indicators from the PPCA and PPCA$_h$ model, i.e., an offline indicator comparison between training approaches for the reconstruction-focused LVM, it was found that the general trends and sensitivities were the same between the models, but the PPCA model had improved diagnostic metric sensitivity. The reconstruction error indicator, $HI^{(1)}$, is observed to have the best sensitivity to damage for the reconstruction-focused models, indicating that there is strong diagnostic information contained in the residual data space for the reconstruction-focused methods for this dataset.

Hence, these results directly indicate that the choice of training framework is less important for the offline indicators applied to the reconstruction-focused model for this dataset. However, this result only sometimes holds, as seen in the bearing failure dataset investigation. This is a strong indicator that diversity in the choice of indicator category, LVM formulation, and training approach is important to acknowledge and consider. The results in Figure 4.11 indicate that the offline and online indicators are both useful to the condition inference task, and the implications of the LVM training approach and the underlying assumptions contained within the offline indicators are less crucial to the performance of an LVM-based fault diagnostics methodology for this dataset.

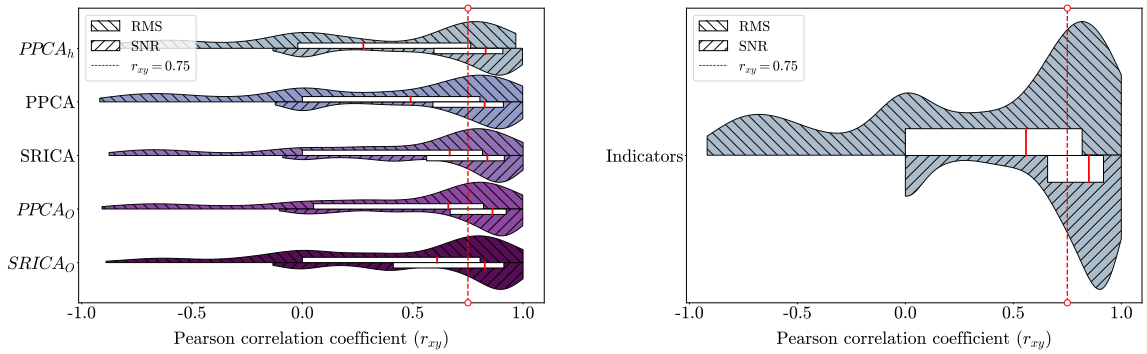
4.3.2.3 Statistical similarity investigation

The PCC metric used to quantify the diversity between the considered indicator categories is presented in Figure 4.13 for the gearbox dataset investigation. The RMS diagnostic metric has multi-modality and weaker correlation, as observed in Figure 4.13(a), while the PCC for the SNR diagnostic metric indicates that most metrics are well correlated. The negative correlation for the RMS metric is associated with indicators that have decreasing RMS trends after the fault origin point. The RMS correlation suggests that there is clear diversity in the indicator responses as a function of the LVM formulation and training approach used.

The SNR correlation observed in Figure 4.13 suggests that after removing the sensitivity to non-stationarity in the data and any variable deviations in magnitude in the SMSE, the offline and online indicators exhibit more similar responses. This highlights that most indicators capture the gear tooth fault but are diverse in their response mechanisms. This equivalence is also captured by the correlation between the offline and online indicators for the two diagnostic metrics, as presented in Figure 4.13(b). The information inferred from the inter-category correlation is much the same as that for Figure 4.13(a), whereby there is less correlation for the RMS trends and increased correlation for the SNR trends. Thus, there is clear diversity within the considered indicator categories, and only by accounting for specific sensitivities in the data and variations in indicator deviation magnitude does this diversity decrease.

4.3.3 Indicator category diversity

The purpose of the experimental investigations conducted in this work was to study the implications of using the two indicator categories for LVMs in a fault diagnosis setting. The online indicators proposed in this work are useful for the fault diagnostics task when using LVMs developed using online training. In the bearing failure dataset investigation, the



(a) Inter-method similarity.

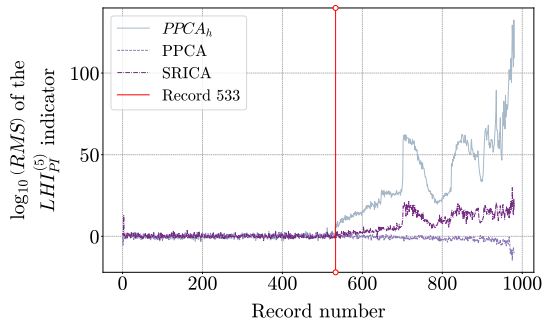
(b) Inter-category similarity.

Figure 4.13. The PCC of the considered diagnostic metrics for the gear tooth dataset investigation. (a) indicates the inter-method diversity as a function of the LVM formulation and training approach used, while (b) indicates the inter-category diversity for the two indicator categories regardless of the choice of formulation or training approach.

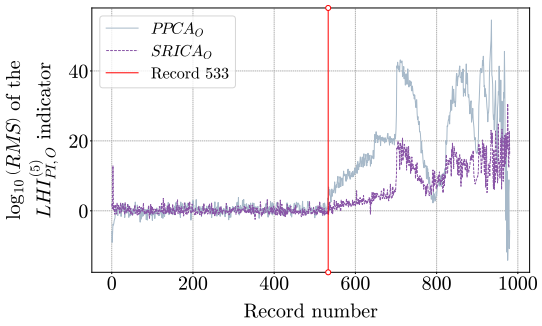
implications of the assumptions inherent to the offline indicator category are made evident for the reconstruction-focused LVM, and using the online indicators for this model produced improved results. For the interpretation-focused LVM, it was observed that online latent indicators that utilise prior characteristic fault frequency information ($LHI_{PD}^{(10)}$ and $LHI_{PD}^{(11)}$) were best suited to enhance the fault trending task using the RMS diagnostic metric. It was also observed that the interpretation-focused LVM developed using online training was better suited to the fault identification task regardless of the indicator category. This suggests that this LVM is better suited to represent the underlying fault condition present in the data within the latent manifold when using the online training approach.

In the gear tooth failure investigation, the implications of the two indicator categories were more subtle, as many of the indicators from both categories were seen to produce favourable results regardless of the LVM formulation or training approach used. Despite this, it is still a positive outcome, as this paves the way for multiple diagnosis avenues to support maintenance decision-making. The implication of the constant latent manifold assumption given to the PPCA model developed using the online training approach was less influential for the performance of the offline indicators applied to the reconstruction-focused model developed using online training.

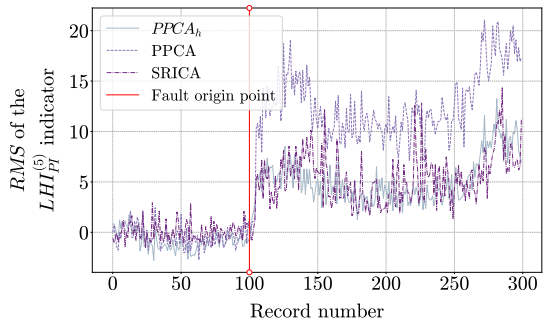
Two categories for latent indicators are identified in this work, and representative offline and online latent indicator sets from these categories are considered. There is clear diversity



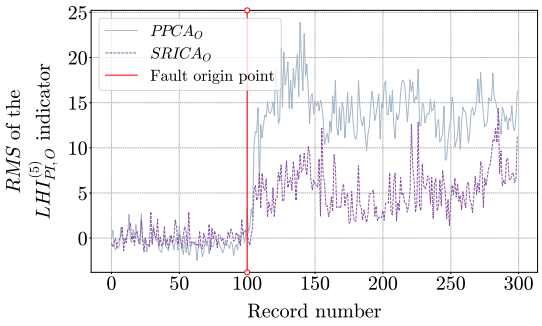
(a) Offline indicator - bearing dataset.



(b) Online indicator - bearing dataset



(c) Offline indicator - gearbox dataset.



(d) Online indicator - gearbox dataset

Figure 4.14. The RMS diagnostic metric responses for the $LHI_{PI}^{(5)}$ metric from the offline and online indicator categories for the considered experimental investigations. This indicator was chosen as it is shared between the offline and online indicator categories and demonstrates the diversity in the results.

in the considered indicators and in the considered indicator categories, and this is succinctly demonstrated in Figure 4.14 for the $LHI_{PI}^{(5)}$ metric, i.e., the Chebyshev distance. For the two categories of indicators, diversity is evident in the indicator responses, and this affects the indicator effectiveness for the fault diagnosis task. This diversity in the indicators from these categories manifests as a function of the fault condition and OCs present in the dataset, e.g., Figure 4.14(a) compared to Figure 4.14(c), whether offline or online LVM training is used, e.g., Figure 4.14(a) compared to Figure 4.14(b), and for the choice of LVM formulation used therein, e.g., Figure 4.14(b). No optimal latent indicator category or latent indicator metric can be identified as the indicator performance is not agnostic to these dependencies.

Hence, diversity supports the use of a plethora of latent indicators for condition inference. Importantly, the diversity is not uncontrollable; instead, it is guided by sensible choices when deriving LHIs from the identified indicator categories, e.g., targeted latent scaling indicators

(Equation (4.25)), and by ignoring non-sensible latent indicators, e.g., the normalised Canberra distance (Equation (4.17)) in a PI setting. Rather, justifiable diversity is obtained as the latent indicators can capture informative perspectives of the latent manifold. An ensemble-based latent indicator approach using numerous offline or online latent indicators allows for unique indicator insight into the underlying asset condition. This enables complementary decision-making to guide maintenance decisions. Furthermore, the likelihood of missing the underlying failure condition decreases as multiple latent fault manifestation mechanisms are accounted for. The choice of indicator category must be addressed as both categories provide a diverse set of latent condition indicators, and this choice must be considered when employing offline or online LVMs as a CM methodology for vibration-based fault diagnostics.

4.4 Conclusion

The purpose of this chapter is to investigate two different indicator categories for offline and online LVMs. This facilitates the purpose of the third research direction identified in Section 1.3.1 of this thesis. Online LHIs are proposed for online LVMs to address implicit offline training-based assumptions in the LHI framework proposed in Chapter 2. The proposed online indicators were compared to the indicators from Chapter 2 to understand the implications of the indicator category used. The comprehensive indicator comparison results indicate that a diverse LHI framework benefits LVM-based condition inference. Furthermore, the implications of using offline indicators for online reconstruction-focused LVMs are apparent, and the investigations performed in this chapter suggest that the choice of indicator category alongside the choice of indicators derived from each category are an important dataset-dependent component of the temporal-preserving LVM pipeline.

Using a diverse range of LHIs allows for a more complete temporal-preserving LVM-based application to the gearbox fault diagnosis problem. This highlights the underlying goal of the third research direction, which is to assess and quantify the importance of diversity in offline and online temporal-preserving LVM-based methodologies for gearbox fault diagnostics. Multiple diagnosis avenues provide a better understanding and reveal different aspects of the underlying asset condition, provide beneficial redundancy to ensure that valuable insights can be obtained, and improve the support given to the decision-making process by offering a comprehensive and flexible description of the underlying asset health state. This emphasises that the research directions considered in this thesis advance the application of temporal-preserving LVMs for

gearbox fault diagnostics.

By conducting a comprehensive comparison with two experimental datasets into the choice of indicator category used for condition inference, the results indicate that a diverse health indicator framework benefits the LVM-based condition inference task. Furthermore, neither indicator category was observed to be better suited for condition inference when using LVMs developed using the online training approach. Rather, both categories had advantages and disadvantages dependent on the dataset investigated. Moreover, the offline training approach is still a useful condition inference avenue to develop and utilise LVMs, which does not discredit the offline indicators or the offline training approach.

Therefore, the LVM training approach and the intrinsic assumptions of the LVM formulation must be carefully assessed when deriving indicators from the available indicator categories, as justifiable diversity is a pertinent factor. The latent manifold is a fertile source for diagnostic information and using it for LVM-based gearbox fault diagnostics is a thoroughly beneficial endeavour. Moreover, a holistic fault diagnosis strategy can be developed using the proposed latent indicator categories for the latent indicator framework, whereby the diagnostic potential of a plethora of LHIs for offline and online models merge different perspectives to supplement and enable proactive and comprehensive temporal-preserving LVM-based gearbox CM.

Chapter 5

Conclusion and recommendations

5.1 Thesis conclusion

This thesis concerns data-driven learning-based techniques for the vibration-based fault diagnosis task. Unsupervised data-driven models are an exciting application of learning-based methods and their adoption in the condition monitoring field is a promising endeavour. In this thesis, temporal-preserving Latent Variable Models (LVMs) are used as an unsupervised learning-based methodology to detect, trend, and isolate gearbox faults, i.e., perform gearbox fault diagnostics. By leveraging the benefits of temporal preservation, this thesis presents approaches to enhance temporal-preserving LVMs for gearbox fault diagnosis applications.

Enhancing the diagnostic information obtained from temporal-preserving LVMs and advancing unsupervised learning-based methodologies are of direct interest in this research, and the research directions considered in this thesis tackle this problem head-on. Concretely, three research directions are identified and thoroughly investigated to enable improved and effective unsupervised learning-based condition inference. The first direction identifies that the latent manifold is an often unused source of diagnostic information, the second direction questions recognises the potential lack of healthy historical data as a basis to investigate online training, and the third direction considers how the latent manifold of offline and online models is used for gearbox fault diagnostics.

To address the first research direction, the latent manifold of offline reconstruction-focused models, i.e., methods focused on preserving the variance in the vibration data, is exploited to extract fault-feature-sensitive latent health indicators. Offline models represent the conventional application mode of unsupervised learning-based methods. However, offline training does not facilitate condition monitoring scenarios where no healthy historical asset data is available and treats the latent manifold as a non-dynamic object through record time. Therefore, the second direction considers the implications of offline and online training and this thesis demonstrates that online training is useful for temporal-preserving LVM-based fault diagnostics. Furthermore, two main online methods are considered, namely reconstruction-

focused and interpretation-focused methods, i.e., methods that seek an interpretable latent manifold, and the complementary nature of these online methods is identified. The third research direction facilitates the identification of categories of latent health indicators for offline and online methods, and the comprehensive offline and online indicator comparison performed in this thesis demonstrates the benefit of using a diverse set of latent health indicators to support maintenance decision-making. The investigations conducted for the considered research directions indicates and reinforces that temporal-preserving LVMs are a powerful methodology for gearbox fault diagnostics and are a competent unsupervised data-driven approach for gearbox fault diagnosis.

A modular latent indicator framework is proposed in Chapter 2 of this thesis to enhance the condition inference task. The proposed framework identified five conceptual classes from which offline latent health indicators are derived, and the latent indicators from these classes were investigated using offline reconstruction-focused models as condition indicators for the fault diagnosis task. Furthermore, the classes demonstrate that the latent manifold is an effective and multi-faceted representation useful for gearbox fault diagnostics. Effectively traversing the latent manifold for vibration data with fault signatures enhances the diagnostic information yield from temporal-preserving LVMs.

Latent indicators from the proposed framework were applied to experimental data acquired under constant operating conditions and time-varying operating conditions, which showed localised bearing and gear tooth failure conditions, respectively. The investigation results highlight that the performance of the latent indicators is dataset and model-dependent, and this dependency underlines the importance of an ensemble-based indicator approach for fault diagnosis applications. Thus, offline LVMs used for the gearbox fault diagnosis are enhanced by the proposed latent indicator framework, allowing for improved latent condition inference.

In Chapter 2, the temporal-preserving LVMs were developed using offline training. However, online models can also be used to consider scenarios lacking healthy historical asset data. Therefore, two distinct LVM training approaches are considered in Chapter 3 of this thesis as different strategies to obtain an informative latent manifold. Under the notion that historical vibration data may not always be available a priori to estimate the model parameters, the implications of an online training approach are investigated. Online training allows for a dynamic and feature-rich latent manifold to be obtained, and for two main online methods, i.e., reconstruction-driven methods and interpretation-driven methods, to be investigated and

compared in a fault diagnosis context to understand their complementarity.

The investigation performed in Chapter 3 compared the performance of offline and online temporal-preserving LVMs. A numerical gearbox dataset and two experimental datasets acquired under different operating conditions were used. The investigation results highlight the complementarity of the considered model formulations when using the online training approach and indicate that the online training approach is a useful strategy for condition inference. Additionally, the results suggest that *(i)* there is a dependency between the fault type present and the applicability of the model formulation, and *(ii)* multiple latent indicators are necessary to ensure and enable effective condition inference. Therefore, online temporal-preserving LVMs are a viable approach for fault diagnostics in scenarios where no prior historical data is available, healthy or otherwise.

Chapter 3 confirmed that online training is useful for the fault diagnosis task, but only a sparse set of indicators were adapted from the latent indicator framework proposed in Chapter 2 and this disparity motivates further investigation. Therefore, the evaluation step of offline and online temporal-preserving LVMs is analysed in Chapter 4 of this thesis. Implicit assumptions encoded into the latent indicators for offline models derived from the latent indicator framework proposed in Chapter 2 are identified, two categories for latent indicators are related, and a set of online latent indicators for online reconstruction-focused and interpretation-focused models is proposed to address the identified implicit assumptions. In the context of the approaches proposed in Chapter 2 and 3, it is logical that the commonalities and differences between the proposed latent indicator framework (i.e., offline or online) and the choice of training approach be explored. This allows for the work conducted in this thesis to come full circle.

The investigation performed in Chapter 4 of this thesis studies the applicability of offline indicators from Chapter 2 to models developed using offline and online training. It compares the performance to a set of proposed online indicators. An experimental dataset obtained under constant operating conditions featuring a bearing fault, and an experimental dataset obtained under time-varying operating conditions featuring a gear tooth fault with additional non-fault-related impulsivity is used. This investigation quantifies the ability of offline and online methods to extract fault signatures from the observed data. It relies on indicators from the two indicator categories to enable performance quantification.

The investigation from Chapter 4 indicates that the applicability of indicators from the identified categories is fault and operating environment-dependent and that the proposed

online indicators are effective for the fault diagnosis task. There is clear diversity in the considered offline and online latent health indicators, which indicates that using an informative and heterogeneous set of latent indicators serves to enhance the LVM-based condition inference task. Moreover, the ability to rely on a multi-faceted characterisation of the latent manifold to obtain a diverse set of latent health indicators for either offline or online methods readily aids maintenance decision-making, and this versatility greatly enhances the capability to ensure asset reliability. Hence, the choice of latent indicator category, training approach, and model formulation used are important considerations for LVM-based fault diagnostics.

The approaches considered in this thesis heighten and enhance the performance of temporal-preserving LVMs to the benefit of the fault diagnosis task and enable a holistic LVM-based fault diagnosis strategy via categories of latent health indicators for a heterogeneous and multi-faceted latent condition inference approach catering to models developed using either offline and online training. Therefore, the research directions raised and investigated in this thesis advance the unsupervised learning-based condition monitoring field.

5.2 Recommendations for future research

It is prudent to offer suggestions for future research opportunities, acknowledge any unanswered questions, and comment on any limitations or shortcomings. The suggestions are as follows:

- Post-processing strategies for ensemble-based latent indicator approaches may be useful to condense and summarise the information gained from different latent health indicators for ease of interpretability and decision-making. Hence, exploring methodologies to accommodate multiple latent indicators and extract fault diagnostic information in an ensemble-based manner needs to be investigated.
- An investigation into a hybrid approach using different LVM formulations can offer a useful methodology for gearbox condition monitoring. For example, a variance-driven model can capture the dominant sources of variance via healthy historical data, e.g., capturing the dominant gear mesh phenomena. A residual signal can then be constructed using the variance-driven temporal-preserving model and analysed online using an interpretation-focused model. Conceptually, this is a LVM-based variant of deterministic-random separation and may be useful for fault diagnostics.
- This thesis primarily concerns the unsupervised learning task, i.e., requiring no access

to labelled fault data. However, incorporating label information into LVMs via semi-supervised learning could be an intriguing research path. The labels need not be fault-condition related but could also be related to the asset operating conditions. The question; however, is to what extent does incorporating this information benefit the fault diagnosis task and how do aspects of the LVM pipeline, e.g., model training objectives and linear or nonlinear transition functions, or asset operating conditions, e.g., harsh operating environments, influence information adoption and incorporation.

- The numerical and experimental data considered in this thesis represented localised bearing and gear tooth damage. However, this is a small percentage of the potential gearbox failure modes experienced in industrial applications. The applicability of temporal-preserving LVMs must be investigated for alternative failure modes to understand the bounds of their capabilities to spur future research into automatic fault identification. Upon initiating this thesis, the scope of comparison was restricted as it was deemed inappropriate to perform extensive comparisons to existing state-of-the-art methodologies without adequate understanding of temporal-preserving LVMs. Instead, approaches to enhance temporal-preserving LVMs were identified and developed, thereby bolstering their applicability as a fault diagnosis methodology. Additional comparisons to existing methodologies using alternative datasets are recommended for future research.

- While the latent manifold of LVMs represents salient data features driven by the model training objectives, not all features are inherently useful to the fault diagnosis task. There are two potential avenues of exploration for accounting for non-fault-related complexity in the data: *(i)* using the latent manifold directly or *(ii)* using additional post-processing. This thesis has partially explored both avenues, e.g., targeted scaling of latent components via Equation (4.25) or synchronous representations to extract and enhance localised gear damage information. However, further investigation is required to understand whether there is an optimal strategy to extract salient latent features useful to the fault diagnosis task.

- Relating to the point above, accounting for variable Operating Conditions (OCs) is an important endeavour. Further research can be broadly categorised into three directions for temporal-preserving LVMs: *(i)* pre-processing the vibration data to account for variable OC-based frequency modulation or amplitude modulation, e.g., order tracking or variable OC amplitude modulation normalisation, *(ii)* normalisation of the Health

Indicator (HI) and Latent Health Indicator (LHI) signals using OC information via a indicator OC post-processing strategy, and *(iii)* identification and enhancement of latent components that carry prominent diagnostic information. Each of these research directions are suggested as they can offer further insight into the effective utilisation of temporal-preserving LVMs.

- Different temporal-preserving LVM formulations have been explored in this thesis. However, the implications of using linear or non-linear reconstruction-focused and interpretation-focused methods were not explored exhaustively, and a critical comparison was not conducted. Hence, it is suggested that linear and nonlinear methods need to be investigated to ascertain whether additional model flexibility yields a more expressive latent manifold, and whether this is beneficial to the fault diagnosis task.

- Further comparisons to state-of-the-art signal processing methodologies are useful to benchmark the performance of LVM-based methodologies in a time-preserving setting to identify similarities, differences, and points of consideration. Thus, it is suggested that this baseline comparison be conducted to establish the relative positioning of the considered methodologies in a fault diagnosis context.

- Temporal-preserving LVMs are an effective methodology for the vibration-based fault diagnosis task. Nonetheless, certain limitations merit discussion. Firstly, while LVMs do learn various informative latent components to construct the latent manifold, it remains difficult to determine which components aid the fault diagnosis task optimally. Second, identifying an appropriate number of latent components poses a limitation and challenge. This issue can impact the performance of the methods. Third, the optimal LVM training objective for effectual asset condition inference continues to be an active area of research. Reconstruction-focused and interpretation-focused methods were considered but they only represent two options in a diverse range of potential methods. Finally, overfitting was observed when training Variational Auto-Encoders (VAEs) on phenomenological gearbox data, with the models favouring dominant deterministic gear information in the healthy training data. The impact therein was a disregard of noise floor spectral content. Preliminary results highlighted this as a potential limitation for offline nonlinear LVMs as the latent manifold for these models seemed ineffective for fault conditions that manifested in spectral regions corresponding to the noise floor of the healthy data.

- Throughout the course of this thesis's development, certain inconsistencies have emerged

and are identified. Notably, the numerical dataset investigation was inconsistently applied. While its initial application in Chapter 3 was intended as a representative approach to study the implication of contrasting LVM objectives on the learnt latent manifold, the inconsistent use throughout the rest of the thesis was a missed opportunity. Considering the potential of numerical datasets described in Section 1.2.1.2, a more consistent application of these datasets should have occurred.

References

[1] A. K. S. Jardine, D. Lin and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20:1483–1510, 2006.

[2] A. Prajapati, J. Bechtel and S. Ganesan. Condition based maintenance: a survey. *Journal of Quality in Maintenance Engineering*, 18:384–400, 2012.

[3] L. Biggio and I. Kastanis. Prognostics and health management of industrial assets: Current progress and road ahead. *Frontiers in Artificial Intelligence*, 3:1–24, 2020.

[4] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao and D. Siegel. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42:314–334, 2014.

[5] D. Wang, K.-L. Tsui and Q. Miao. Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access*, 6:665–676, 2018.

[6] O. Fink, Q. Wang, M. Svensén, P. Dersin, W.-J. Lee and M. Ducoffe. Potential, challenges and future directions for deep learning in prognostics and health management applications. *Engineering Applications of Artificial Intelligence*, 92:103678, 2020. arXiv: 2005.02144.

[7] R. Liu, B. Yang, E. Zio and X. Chen. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108:33–47, 2018.

[8] M. Cerrada, R. V. Sánchez, C. Li, F. Pacheco, D. Cabrera, J. Valente de Oliveira and R. E. Vásquez. A review on data-driven fault severity assessment in rolling bearings. *Mechanical Systems and Signal Processing*, 99:169–196, 2018.

[9] D. McMillan and G. W. Ault. Condition monitoring benefit for onshore wind turbines: Sensitivity to operational parameters. *IET Renewable Power Generation*, 2:60–72, 2008.

[10] J. P. Salameh, S. Cauet, E. Etien, A. Sakout and L. Rambault. Gearbox condition monitoring in wind turbines: A review. *Mechanical Systems and Signal Processing*, 111:251–264, 2018.

[11] V. Sharma and A. Parey. A review of gear fault diagnosis using various condition indicators. *Procedia Engineering*, 144:253–263, 2016.

[12] L. van Eyk, P. S. Heyns and S. Schmidt. A short review of gear fault modelling in a hybrid modelling context. *Modelling and Simulation of Complex Systems for Sustainable Energy Efficiency*, pages 242–258. Springer International Publishing, 2022.

[13] R. B. Randall and J. Antoni. Rolling element bearing diagnostics - A tutorial. *Mechanical Systems and Signal Processing*, 25:485–520, 2011.

[14] A. F. del Rincon, F. Viadero, M. Iglesias, A. De-Juan, P. Garcia and R. Sancibrian. Effect of cracks and pitting defects on gear meshing. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 226:2805–2815, 2012.

[15] C. J. Stander and P. S. Heyns. Transmission path phase compensation for gear monitoring under fluctuating load conditions. *Mechanical Systems and Signal Processing*, 20:1511–1522, 2006.

[16] W. Bartelmus and R. Zimroz. Vibration condition monitoring of planetary gearbox under varying external load. *Mechanical Systems and Signal Processing*, 23:246–257, 2009.

[17] R. Zimroz, W. Bartelmus, T. Barszcz and J. Urbanek. Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings. *Mechanical Systems and Signal Processing*, 46:16–27, 2014.

[18] R. B. Randall. Introduction and background, *Vibration-based condition monitoring: Industrial, aerospace and automotive applications*, chapter 1, pages 1–23. Wiley, 2011.

[19] P. Večeř, M. Kreidl and R. Šmíd. Condition indicators for gearbox condition monitoring systems. *Acta Polytechnica*, 45:35–43, 2005.

[20] J. Antoni, F. Bonnardot, A. Raad and M. El Badaoui. Cyclostationary modelling of rotating machine vibration signals. *Mechanical Systems and Signal Processing*, 18:1285–1314, 2004.

[21] A. Raad, J. Antoni and M. Sidahmed. Indicators of cyclostationarity: Theory and application to gear fault monitoring. *Mechanical Systems and Signal Processing*, 22:574–587, 2008.

[22] P. Borghesani and J. Antoni. CS2 analysis in presence of non-Gaussian background noise – Effect on traditional estimators and resilience of log-envelope indicators. *Mechanical Systems and Signal Processing*, 90:378–398, 2017.

[23] A. Mauricio, J. Qi, W. A. Smith, M. Sarazin, R. B. Randall, K. Janssens and K. C. Gryllias. Bearing diagnostics under strong electromagnetic interference based on integrated spectral coherence. *Mechanical Systems and Signal Processing*, 140:106673, 2020.

[24] K. Feng, W. A. Smith, P. Borghesani, R. B. Randall and Z. Peng. Use of cyclostationary properties of vibration signals to identify gear wear mechanisms and track wear evolution. *Mechanical Systems and Signal Processing*, 150:107258, 2021.

[25] R. Zhang, Z. Peng, L. Wu, B. Yao and Y. Guan. Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence. *Sensors*, 17:549, 2017.

[26] F. Jia, Y. Lei, N. Lu and S. Xing. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mechanical Systems and Signal Processing*, 110:349–367, 2018.

[27] W. Mao, W. Feng, Y. Liu, D. Zhang and X. Liang. A new deep auto-encoder method with fusing discriminant information for bearing fault diagnosis. *Mechanical Systems and Signal Processing*, 150:107233, 2021.

[28] W. Booyse, D. N. Wilke and P. S. Heyns. Deep digital twins for detection, diagnostics and prognostics. *Mechanical Systems and Signal Processing*, 140:106612, 2020.

[29] A. González-Muñiz, I. Díaz, A. A. Cuadrado and D. García-Pérez. Health indicator for machine condition monitoring built in the latent space of a deep autoencoder. *Reliability Engineering & System Safety*, 224:108482, 2022.

[30] J. Antoni. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20:282–307, 2006.

[31] J. Antoni. The infogram: Entropic evidence of the signature of repetitive transients. *Mechanical Systems and Signal Processing*, 74:73–94, 2016.

[32] W. A. Smith, P. Borghesani, Q. Ni, K. Wang and Z. Peng. Optimal demodulation-band selection for envelope-based diagnostics: A comparative study of traditional and novel tools. *Mechanical Systems and Signal Processing*, 134:106303, 2019.

[33] S. Schmidt, A. Mauricio, P. S. Heyns and K. C. Gryllias. A methodology for identifying information rich frequency bands for diagnostics of mechanical components-of-interest under time-varying operating conditions. *Mechanical Systems and Signal Processing*, 142:106739, 2020.

[34] X.-T. Jiao. The characteristics analysis of gear box fault vibration signal. *Proceedings of the 3rd Annual International Conference on Mechanics and Mechanical Engineering*, volume 105, pages 116–121. Atlantis Press, 2017.

[35] Y.-K. Gu, W.-F. Li, J. Zhang and G.-Q. Qiu. Effects of wear, backlash, and bearing clearance on dynamic characteristics of a spur gear system. *IEEE Access*, 7:117639–117651, 2019.

[36] R. B. Randall. A new method of modeling gear faults. *Journal of Mechanical Design*, 104:259–267, 1982.

[37] R. B. Randall. Vibration signals from rotating and reciprocating machines, *Vibration-based condition monitoring: Industrial, aerospace and automotive applications*, chapter 2, pages 25–62. Wiley, 2011.

[38] C. S. Sunnersjö. Varying compliance vibrations of rolling bearings. *Journal of Sound and Vibration*, 58:363–373, 1978.

[39] S. A. McInerny and Y. Dai. Basic vibration signal processing for bearing fault detection. *IEEE Transactions on Education*, 46:149–156, 2003.

[40] C. J. Stander and P. S. Heyns. Instantaneous angular speed monitoring of gearboxes under non-cyclic stationary load conditions. *Mechanical Systems and Signal Processing*, 19:817–835, 2005.

[41] F. Chaari, W. Bartelmus, R. Zimroz, T. Fakhfakh and M. Haddar. Gearbox vibration signal amplitude and frequency modulation. *Shock and Vibration*, 19:635–652, 2012.

[42] P. Zhou, S. Chen, Q. He, D. Wang and Z. Peng. Rotating machinery fault-induced vibration signal modulation effects: A review with mechanisms, extraction methods and applications for diagnosis. *Mechanical Systems and Signal Processing*, 200:110489, 2023.

[43] P. D. McFadden. Low frequency vibration generated by gear tooth impacts. *NDT International*, 18:279–282, 1985.

[44] H. Saruhan, S. Sandemir, A. Çiçek and I. Uygur. Vibration analysis of rolling element bearings defects. *Journal of Applied Research and Technology*, 12:384–395, 2014.

[45] I. El-Thalji and E. Jantunen. A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mechanical Systems and Signal Processing*, 60:252–272, 2015.

[46] P. D. McFadden and J. D. Smith. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*, 96:69–82, 1984.

[47] P. D. McFadden and J. D. Smith. The vibration produced by multiple point defects in a rolling element bearing. *Journal of Sound and Vibration*, 98:263–273, 1985.

[48] I. El-Thalji and E. Jantunen. A descriptive model of wear evolution in rolling bearings. *Engineering Failure Analysis*, 45:204–224, 2014.

[49] I. El-Thalji and E. Jantunen. Fault analysis of the wear fault development in rolling bearings. *Engineering Failure Analysis*, 57:470–482, 2015.

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

[51] P. D. McFadden and J. D. Smith. Vibration monitoring of rolling element bearings by the high-frequency resonance technique — a review. *Tribology International*, 17:3–10, 1984.

[52] X. Liang, M. J. Zuo and Z. Feng. Dynamic modeling of gearbox faults: A review. *Mechanical Systems and Signal Processing*, 98:852–876, 2018.

[53] A. Kahraman. Load sharing characteristics of planetary transmissions. *Mechanism and Machine Theory*, 29:1151–1165, 1994.

[54] F. Chaari, T. Fakhfakh and M. Haddar. Dynamic analysis of a planetary gear failure caused by tooth pitting and cracking. *Journal of Failure Analysis and Prevention*, 6:73–78, 2006.

[55] X. Gu, P. Velex, P. Sainsot and J. Bruyère. Analytical investigations on the mesh stiffness function of solid spur and helical gears. *Journal of Mechanical Design*, 137:063301, 2015.

[56] Y. Pandya and A. Parey. Failure path based modified gear mesh stiffness for spur gear pair with tooth root crack. *Engineering Failure Analysis*, 27:286–296, 2013.

[57] D. C. H. Yang and J. Y. Lin. Hertzian damping, tooth friction and bending elasticity in gear impact dynamics. *Journal of Mechanisms, Transmissions, and Automation in Design*, 109:189–196, 1987.

[58] R. G. Parker, S. M. Vijayakar and T. Imajo. Non-linear dynamic response of a spur gear pair: Modelling and experimental comparisons. *Journal of Sound and Vibration*, 237:435–455, 2000.

[59] T. Eritenel and R. G. Parker. An investigation of tooth mesh nonlinearity and partial contact loss in gear pairs using a lumped-parameter model. *Mechanism and Machine Theory*, 56:28–51, 2012.

[60] Y. Lei, Z. Liu, D. Wang, X. Yang, H. Liu and J. Lin. A probability distribution model of tooth pits for evaluating time-varying mesh stiffness of pitting gears. *Mechanical Systems and Signal Processing*, 106:355–366, 2018.

[61] H. Ma, J. Zeng, R. Feng, X. Pang, Q. Wang and B. Wen. Review on dynamics of cracked gear systems. *Engineering Failure Analysis*, 55:224–245, 2015.

[62] W. Bartelmus. Mathematical modelling and computer simulations as an aid to gearbox diagnostics. *Mechanical Systems and Signal Processing*, 15:855–871, 2001.

[63] S. Wu, M. J. Zuo and A. Parey. Simulation of spur gear dynamics and estimation of fault growth. *Journal of Sound and Vibration*, 317:608–624, 2008.

[64] O. D. Mohammed, M. Rantatalo and J.-O. Aidanpää. Improving mesh stiffness calculation of cracked gears for the purpose of vibration-based fault analysis. *Engineering Failure Analysis*, 34:235–251, 2013.

[65] O. D. Mohammed, M. Rantatalo and J.-O. Aidanpää. Dynamic modelling of a one-stage spur gear system and vibration-based tooth crack detection analysis. *Mechanical Systems and Signal Processing*, 54-55:293–305, 2015.

[66] F. Chaari, T. Fakhfakh and M. Haddar. Analytical modelling of spur gear tooth crack and influence on gearmesh stiffness. *European Journal of Mechanics - A/Solids*, 28:461–468, 2009.

[67] S. He, R. Gunda and R. Singh. Effect of sliding friction on the dynamics of spur gear pair with realistic time-varying stiffness. *Journal of Sound and Vibration*, 301:927–949, 2007.

[68] J. Sopanen and A. Mikkola. Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 1: Theory. *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, 217:201–211, 2003.

[69] N. Sawalhi and R. B. Randall. Simulating gear and bearing interactions in the presence of faults. *Mechanical Systems and Signal Processing*, 22:1924–1951, 2008.

[70] V. N. Patel, N. Tandon and R. K. Pandey. A Dynamic Model for Vibration Studies of Deep Groove Ball Bearings Considering Single and Multiple Defects in Races. *Journal of Tribology*, 132:041101, 2010.

[71] Z. Wang, G. Li, X. Zhou, H. Zhang, Z. Lin and S. Jia. Dynamic analysis of deep groove ball bearing with localized defects and misalignment. *Journal of Sound and Vibration*, 568:118071, 2023.

[72] J. Meagher, X. Wu, D. Kong and C. H. Lee. A comparison of gear mesh stiffness modeling strategies. *Structural Dynamics*, volume 3, pages 255–263, 2011.

[73] Z. Kiral and H. Karagülle. Simulation and analysis of vibration signals generated by rolling element bearing with defects. *Tribology International*, 36:667–678, 2003.

[74] P. D. McFadden. Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *Journal of Vibration and Acoustics*, 108:165–170, 1986.

[75] S. Braun and B. Seth. Analysis of repetitive mechanism signatures. *Journal of Sound and Vibration*, 70:513–526, 1980.

[76] P. Borghesani, W. A. Smith, R. B. Randall, J. Antoni, M. El Badaoui and Z. Peng. Bearing signal models and their effect on bearing diagnostics. *Mechanical Systems and Signal Processing*, 174:109077, 2022.

[77] C. Capdessus, M. Sidahmed and J. L. Lacoume. Cyclostationary processes: Application in gear faults early diagnosis. *Mechanical Systems and Signal Processing*, 14:371–385, 2000.

[78] J. Antoni. Cyclostationarity by examples. *Mechanical Systems and Signal Processing*, 23:987–1036, 2009.

[79] D. Abboud, J. Antoni, M. Eltabach and S. Sieg-Zieba. Angle\time cyclostationarity for the analysis of rolling element bearing vibrations. *Measurement*, 75:29–39, 2015.

[80] D. Abboud, S. Baudin, J. Antoni, D. Rémond, M. Eltabach and O. Sauvage. The spectral analysis of cyclo-non-stationary signals. *Mechanical Systems and Signal Processing*, 75:280–300, 2016.

[81] J. Lin and M. Zhao. A review and strategy for the diagnosis of speed-varying machinery. *2014 International Conference on Prognostics and Health Management*, pages 1–9. IEEE, 2014.

[82] Z. Feng and M. J. Zuo. Vibration signal models for fault diagnosis of planetary gearboxes. *Journal of Sound and Vibration*, 331:4919–4939, 2012.

[83] G. Dalpiaz, A. Rivola and R. Rubini. Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears. *Mechanical Systems and Signal Processing*, 14:387–412, 2000.

[84] J. Urbanek, T. Barszcz, M. Straczkiewicz and A. Jablonski. Normalization of vibration signals generated under highly varying speed and load with application to signal separation. *Mechanical Systems and Signal Processing*, 82:13–31, 2017.

[85] C. J. Stander, P. S. Heyns and W. Schoombie. Using vibration monitoring for local fault detection on gears operating under fluctuating load conditions. *Mechanical Systems and Signal Processing*, 16:1005–1024, 2002.

[86] R. B. Randall. Basic signal processing techniques, *Vibration-based condition monitoring: Industrial, aerospace and automotive applications*, chapter 3, pages 63–141. Wiley, 2011.

[87] S. Schmidt and P. S. Heyns. Normalisation of the amplitude modulation caused by time-varying operating conditions for condition monitoring. *Measurement*, 149:106964, 2020.

[88] K. R. Fyfe and E. D. S. Munck. Analysis of computed order tracking. *Mechanical Systems and Signal Processing*, 11:187–205, 1997.

[89] A. Brandt, T. Lagö, K. Ahlin and J. Tuma. Main principles and limitations of current order tracking methods. *Sound and Vibration*, 39:19–22, 2005.

[90] S. Schmidt, P. S. Heyns and J. P. de Villiers. A tacholess order tracking methodology based on a probabilistic approach to incorporate angular acceleration information into the maxima tracking process. *Mechanical Systems and Signal Processing*, 100:630–646, 2018.

[91] C. Peeters, Q. Leclère, J. Antoni, P. Lindahl, J. Donnal, S. Leeb and J. Helsen. Review and comparison of tacholess instantaneous speed estimation methods on experimental vibration data. *Mechanical Systems and Signal Processing*, 129:407–436, 2019.

[92] P. Borghesani, P. Pennacchi, R. B. Randall and R. Ricci. Order tracking for discrete-random separation in variable speed conditions. *Mechanical Systems and Signal Processing*, 30:1–22, 2012.

[93] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang and R. X. Gao. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115:213–237, 2019.

[94] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li and A. K. Nandi. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138:106587, 2020.

[95] O. Surucu, S. A. Gadsden and J. Yawney. Condition monitoring using machine learning: A review of theory, applications, and recent advances. *Expert Systems with Applications*, 221:119738, 2023.

[96] T. Hastie, R. Tibshirani and J. Friedman. The elements of statistical learning. Springer Series in Statistics. Springer, 2009, pages 508–508.

[97] I. Goodfellow, Y. Bengio and A. Courville. Deep learning. MIT Press, 2016.

[98] Q. Miao and V. Makis. Condition monitoring and classification of rotating machinery using wavelets and hidden Markov models. *Mechanical Systems and Signal Processing*, 21:840–855, 2007.

[99] K. C. Gryllias and I. A. Antoniadis. A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments. *Engineering Applications of Artificial Intelligence*, 25:326–344, 2012.

[100] Ł. Jedliński and J. Jonak. Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing*, 30:636–641, 2015.

[101] C. M. Bishop. Pattern recognition and machine learning. Springer, 2006.

[102] I. Ulusoy and C. M. Bishop. Comparison of generative and discriminative techniques for object detection and classification, pages 173–195. 2006.

[103] A. Ng and M. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes. *Advances in Neural Information Processing Systems*, volume 14, pages 169–187. MIT Press, 2001.

[104] S. Khan and T. Yairi. A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 107:241–265, 2018.

[105] Z. Zhu, Y. Lei, G. Qi, Y. Chai, N. Mazur, Y. An and X. Huang. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery. *Measurement*, 206:112346, 2023.

[106] E. Fix and J. L. Hodges. Discriminatory analysis. Nonparametric discrimination: Consistency properties. *International Statistical Review / Revue Internationale de Statistique*, 57:238, 1989.

[107] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13:21–27, 1967.

[108] S. B. Kotsiantis. Decision trees: A recent overview. *Artificial Intelligence Review*, 39:261–283, 2013.

[109] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.

[110] D. J. Hand and K. Yu. Idiot's Bayes: Not so stupid after all? *International Statistical Review*, 69:385–398, 2001.

[111] A. Moosavian, H. Ahmadi, A. Tabatabaeefar and M. Khazaee. Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing. *Shock and Vibration*, 20:263–272, 2013.

[112] Y. Lei and M. J. Zuo. Gear crack level identification based on weighted K nearest neighbor classification algorithm. *Mechanical Systems and Signal Processing*, 23:1535–1547, 2009.

[113] D. Wang. K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: Revisited. *Mechanical Systems and Signal Processing*, 70-71:201–208, 2016.

[114] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.

[115] X. Wan, D. Wang, P. W. Tse, G. Xu and Q. Zhang. A critical study of different dimensionality reduction methods for gear crack degradation assessment under different operating conditions. *Measurement*, 78:138–150, 2016.

[116] Y. Zhang, X. Li, L. Gao, L. Wang and L. Wen. Imbalanced data fault diagnosis of rotating machinery using synthetic oversampling and feature learning. *Journal of Manufacturing Systems*, 48:34–50, 2018.

[117] T. Praveenkumar, B. Sabhrish, M. Saimurugan and K. I. Ramachandran. Pattern recognition based on-line vibration monitoring system for fault diagnosis of automobile gearbox. *Measurement*, 114:233–242, 2018.

[118] C. Liu and K. C. Gryllias. A semi-supervised support vector data description-based fault detection method for rolling element bearings based on cyclic spectral analysis. *Mechanical Systems and Signal Processing*, 140:106682, 2020.

[119] J. Antoni. Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions. *Journal of Sound and Vibration*, 304:497–529, 2007.

[120] D. M. J. Tax and R. P. W. Duin. Support vector data description. *Machine Learning*, 54:45–66, 2004.

[121] B. Muruganatham, M. A. Sanjith, B. Krishnakumar and S. A. V. Satya Murty. Roller element bearing fault diagnosis using singular spectrum analysis. *Mechanical Systems and Signal Processing*, 35:150–166, 2013.

[122] M. Unal, M. Onat, M. Demetgul and H. Kucuk. Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network. *Measurement*, 58:187–196, 2014.

[123] V. Muralidharan and V. Sugumaran. A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. *Applied Soft Computing Journal*, 12:2023–2029, 2012.

[124] F. Jia, Y. Lei, J. Lin, X. Zhou and N. Lu. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, 72-73:303–315, 2016.

[125] Y. Lei, F. Jia, J. Lin, S. Xing and S. X. Ding. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics*, 63:3137–3147, 2016.

[126] X. Wu, Y. Zhang, C. Cheng and Z. Peng. A hybrid classification autoencoder for semi-supervised fault diagnosis in rotating machinery. *Mechanical Systems and Signal Processing*, 149:107327, 2021.

[127] C. Lu, Z. Y. Wang, W. L. Qin and J. Ma. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Processing*, 130:377–388, 2017.

[128] X. Guo, L. Chen and C. Shen. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement*, 93:490–502, 2016.

[129] X. Wu, Z. Peng, J. Ren, C. Cheng, W. Zhang and D. Wang. Rub-impact fault diagnosis of rotating machinery based on 1D convolutional neural networks. *IEEE Sensors Journal*, 20:8349–8363, 2020.

[130] V. Sinitsin, O. Ibryaeva, V. Sakovskaya and V. Eremeeva. Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model. *Mechanical Systems and Signal Processing*, 180:109454, 2022.

[131] H. Ren, J. Wang, Z. Zhu, J. Shi and W. Huang. Domain fuzzy generalization networks for semi-supervised intelligent fault diagnosis under unseen working conditions. *Mechanical Systems and Signal Processing*, 200:110579, 2023.

[132] M. Zhao, M. Kang, B. Tang and M. Pecht. Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes. *IEEE Transactions on Industrial Electronics*, 65:4290–4300, 2018.

[133] W. Zhang, X. Li and Q. Ding. Deep residual learning-based fault diagnosis method for rotating machinery. *ISA Transactions*, 95:295–305, 2019.

[134] S. Ma, F. Chu and Q. Han. Deep residual learning with demodulated time-frequency features for fault diagnosis of planetary gearbox under nonstationary running conditions. *Mechanical Systems and Signal Processing*, 127:190–201, 2019.

[135] R. Zhao, R. Yan, J. Wang and K. Mao. Learning to monitor machine health with convolutional bi-directional LSTM networks. *Sensors*, 17:273, 2017.

[136] W. Yu, I. Y. Kim and C. Mechefske. Analysis of different RNN autoencoder variants for time series classification and machine prognostics. *Mechanical Systems and Signal Processing*, 149:107322, 2021.

[137] H. Shao, H. Jiang, H. Zhang and T. Liang. Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network. *IEEE Transactions on Industrial Electronics*, 65:2727–2736, 2018.

[138] H. Jiang, H. Shao, X. Chen and J. Huang. A feature fusion deep belief network method for intelligent fault diagnosis of rotating machinery. *Journal of Intelligent & Fuzzy Systems*, 34:3513–3521, 2018.

[139] Y. Pan, D. Cheng, T. Wei and Y. Jia. Rolling bearing performance degradation assessment based on deep belief network and improved support vector data description. *Mechanical Systems and Signal Processing*, 181:109458, 2022.

[140] Y. Ding, M. Jia, Q. Miao and Y. Cao. A novel time–frequency transformer based on self–attention mechanism and its application in fault diagnosis of rolling bearings. *Mechanical Systems and Signal Processing*, 168:108616, 2022.

[141] I. Daubechies, J. Lu and H. T. Wu. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied and Computational Harmonic Analysis*, 30:243–261, 2011.

[142] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017-Decem:5999–6009, 2017. arXiv: 1706.03762.

[143] S. Schmidt and P. S. Heyns. An open set recognition methodology utilising discrepancy analysis for gear diagnostics under varying operating conditions. *Mechanical Systems and Signal Processing*, 119:1–22, 2019.

[144] T. Heyns, S. J. Godsill, J. P. de Villiers and P. S. Heyns. Statistical gear health analysis which is robust to fluctuating loads and operating speeds. *Mechanical Systems and Signal Processing*, 27:651–666, 2012.

[145] T. Heyns, P. S. Heyns and J. P. de Villiers. Combining synchronous averaging with a Gaussian mixture model novelty detection scheme for vibration-based condition monitoring of a gearbox. *Mechanical Systems and Signal Processing*, 32:200–215, 2012.

[146] D. Fernández-Francos, D. Martínez-Rego, O. Fontenla-Romero and A. Alonso-Betanzos. Automatic bearing fault diagnosis based on one-class ν-SVM. *Computers & Industrial Engineering*, 64:357–365, 2013.

[147] S. Schmidt, P. S. Heyns and K. C. Gryllias. A discrepancy analysis methodology for rolling element bearing diagnostics under variable speed conditions. *Mechanical Systems and Signal Processing*, 116:40–61, 2019.

[148] M. Markou and S. Singh. Novelty detection: A review - part 1: Statistical approaches. *Signal Processing*, 83:2481–2497, 2003.

[149] M. Markou and S. Singh. Novelty detection: A review - part 2: Neural network based approaches. *Signal Processing*, 83:2499–2521, 2003.

[150] K. Vos, Z. Peng, C. Jenkins, M. R. Shahriar, P. Borghesani and W. Wang. Vibration-based anomaly detection using LSTM/SVM approaches. *Mechanical Systems and Signal Processing*, 169:108752, 2022.

[151] H. Zhou, Z. Lei, E. Zio, G. Wen, Z. Liu, Y. Su and X. Chen. Conditional feature disentanglement learning for anomaly detection in machines operating under time-varying conditions. *Mechanical Systems and Signal Processing*, 191:110139, 2023.

[152] W. Wang and A. K. Wong. Autoregressive model-based gear fault diagnosis. *Journal of Vibration and Acoustics*, 124:172–179, 2002.

[153] W. Lu, Y. Li, Y. Cheng, D. Meng, B. Liang and P. Zhou. Early fault detection approach with deep architectures. *IEEE Transactions on Instrumentation and Measurement*, 67:1679–1689, 2018.

[154] M. A. F. Pimentel, D. A. Clifton, L. Clifton and L. Tarassenko. A review of novelty detection. *Signal Processing*, 99:215–249, 2014.

[155] L. Ruff, R. A. Vandermeulen, B. J. Franks, K.-R. Müller and M. Kloft. Rethinking assumptions in deep anomaly detection, 2020. arXiv: 2006.00339.

[156] W. Mao, H. Shi, G. Wang and X. Liang. Unsupervised deep multitask anomaly detection with robust alarm strategy for online evaluation of bearing early fault occurrence. *IEEE Transactions on Instrumentation and Measurement*, 71:1–13, 2022.

[157] Y. Pan, J. Chen and L. Guo. Robust bearing performance degradation assessment method based on improved wavelet packet–support vector data description. *Mechanical Systems and Signal Processing*, 23:669–681, 2009.

[158] J. Huang and X. Yan. Related and independent variable fault detection based on KPCA and SVDD. *Journal of Process Control*, 39:88–99, 2016.

[159] L. Kou, J. Chen, Y. Qin and W. Mao. The robust multi-scale deep-SVDD model for anomaly online detection of rolling bearings. *Sensors*, 22:5681, 2022.

[160] L. Guo, Y. Yu, A. Duan, H. Gao and J. Zhang. An unsupervised feature learning based health indicator construction method for performance assessment of machines. *Mechanical Systems and Signal Processing*, 167:108573, 2022.

[161] D. M. Blei. Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Application*, 1:203–232, 2014.

[162] D. M. Blei, A. Kucukelbir and J. D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112:859–877, 2017. arXiv: arXiv:1601.00670v9.

[163] J. An and S. Cho. Variational autoencoder based anomaly detection using reconstruction probability. *Special lecture on IE*, volume 2, pages 1–18, 2015.

[164] R. Balshaw, P. S. Heyns, D. N. Wilke and S. Schmidt. Importance of temporal preserving latent analysis for latent variable models in fault diagnostics of rotating machinery. *Mechanical Systems and Signal Processing*, 168:108663, 2022.

[165] D. P. Kingma and M. Welling. Auto-encoding variational Bayes, 2013. arXiv: 1312.6114.

[166] D. N. Wilke, P. S. Heyns and S. Schmidt. The role of untangled latent spaces in unsupervised learning applied to condition-based maintenance, *Modelling and Simulation of Complex Systems for Sustainable Energy Efficiency*, pages 38–49. Springer International Publishing, 2022.

[167] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2014.

[168] S. Mohamed and B. Lakshminarayanan. Learning in implicit generative models, 2016. arXiv: 1610.03483.

[169] M. Rosca, B. Lakshminarayanan, D. Warde-Farley and S. Mohamed. Variational approaches for auto-encoding generative adversarial networks, 2017. arXiv: 1706.04987.

[170] W. Hu, T. Wang and F. Chu. A Wasserstein generative digital twin model in health monitoring of rotating machines. *Computers in Industry*, 145:103807, 2023.

[171] N. Gugulothu, V. TV, P. Malhotra, L. Vig, P. Agarwal and G. Shroff. Predicting remaining useful life using time series embeddings based on recurrent neural networks. *CEUR Workshop Proceedings*, 2657:1–9, 2017. arXiv: 1709.01073.

[172] D. Marx and K. C. Gryllias. Domain knowledge informed unsupervised fault detection for rolling element bearings. *PHM Society European Conference*, 7:338–350, 2022.

[173] K. H. Kim, S.-H. Shim, Y. Lim, J. Jeon, J.-W. Choi, B. Kim and A. S. Yoon. RaPP: Novelty detection with reconstruction along projection pathway. *8th International Conference on Learning Representations*, 2020.

[174] J. Park, Y. G. Jung and A. Beng Jin Teoh. Discriminative multi-level reconstruction under compact latent space for one-class novelty detection. *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 7095–7102. IEEE, 2021. arXiv: 2003.01665.

[175] G. Kwon, M. Prabhushankar, D. Temel and G. AlRegib. Backpropagated gradient representations for anomaly detection. *European Conference on Computer Vision*, pages 206–226. Springer, 2020.

[176] M. Zhao and X. Jia. A novel strategy for signal denoising using reweighted SVD and its applications to weak fault feature enhancement of rotating machinery. *Mechanical Systems and Signal Processing*, 94:129–147, 2017.

[177] D. S. Broomhead and G. P. King. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20:217–236, 1986.

[178] A. Hyvärinen, J. Karhunen and E. Oja. Independent component analysis. Adaptive and Learning Systems for Signal Processing, Communications, and Control. John Wiley & Sons, Inc., 2001.

[179] M. E. Wall, A. Rechtsteiner and L. M. Rocha. Singular value decomposition and principal component analysis, *A Practical Approach to Microarray Data Analysis*, chapter 5, pages 91–109. Springer, 2002.

[180] X. Zhao and B. Ye. Similarity of signal processing effect between Hankel matrix-based SVD and wavelet transform and its mechanism analysis. *Mechanical Systems and Signal Processing*, 23:1062–1075, 2009.

[181] E. Bozzo, R. Carniel and D. Fasino. Relationship between singular spectrum analysis and Fourier analysis: Theory and application to the monitoring of volcanic activity. *Computers and Mathematics with Applications*, 60:812–820, 2010.

[182] Y. Chen, X. Liang and M. J. Zuo. An improved singular value decomposition-based method for gear tooth crack detection and severity assessment. *Journal of Sound and Vibration*, 468:115068, 2020.

[183] R. Golafshan and K. Yuce Sanliturk. SVD and Hankel matrix based de-noising approach for ball bearing fault detection and its assessment using artificial faults. *Mechanical Systems and Signal Processing*, 70-71:36–50, 2016.

[184] Q. He, Z. Feng and F. Kong. Detection of signal transients using independent component analysis and its application in gearbox condition monitoring. *Mechanical Systems and Signal Processing*, 21:2056–2071, 2007.

[185] A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13:411–430, 2000.

[186] Z. Wang, J. Chen, G. Dong and Y. Zhou. Constrained independent component analysis and its application to machine fault diagnosis. *Mechanical Systems and Signal Processing*, 25:2501–2512, 2011.

[187] W. Lu and J. C. Rajapakse. Constrained independent component analysis. *Advances in Neural Information Processing Systems*. MIT Press, 2000.

[188] W. Lu and J. C. Rajapakse. Approach and applications of constrained ICA. *IEEE Transactions on Neural Networks*, 16:203–212, 2005.

[189] T. Yang, Y. Guo, X. Wu, J. Na and R. F. Fung. Fault feature extraction based on combination of envelope order tracking and cICA for rolling element bearings. *Mechanical Systems and Signal Processing*, 113:131–144, 2018.

[190] J. Antoni and R. B. Randall. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mechanical Systems and Signal Processing*, 20:308–331, 2006.

[191] J. Wang, R. X. Gao and R. Yan. Integration of EEMD and ICA for wind turbine gearbox diagnosis. *Wind Energy*, 17:757–773, 2014.

[192] B. Rezaeianjouybari and Y. Shang. Deep learning for prognostics and health management: State of the art, challenges, and opportunities. *Measurement*, 163:107929, 2020.

[193] J. Chen, R. Huang, Z. Chen, W. Mao and W. Li. Transfer learning algorithms for bearing remaining useful life prediction: A comprehensive review from an industrial application perspective. *Mechanical Systems and Signal Processing*, 193:110239, 2023.

[194] R. Yan, F. Shen, C. Sun and X. Chen. Knowledge transfer for rotary machine fault diagnosis. *IEEE Sensors Journal*, 20:8374–8393, 2020.

[195] C. Li, S. Zhang, Y. Qin and E. Estupinan. A systematic review of deep transfer learning for machinery fault diagnosis. *Neurocomputing*, 407:121–135, 2020.

[196] W. Li, R. Huang, J. Li, Y. Liao, Z. Chen, G. He, R. Yan and K. C. Gryllias. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. *Mechanical Systems and Signal Processing*, 167:108487, 2022.

[197] Y. Song, Y. Li, L. Jia and M. Qiu. Retraining strategy-based domain adaption network for intelligent fault diagnosis. *IEEE Transactions on Industrial Informatics*, 16:6163–6171, 2020.

[198] W. Li, Z. Chen and G. He. A novel weighted adversarial transfer network for partial domain fault diagnosis of machinery. *IEEE Transactions on Industrial Informatics*, 17:1753–1762, 2021.

[199] H. Wang, X. Bai, J. Tan and J. Yang. Deep prototypical networks based domain adaptation for fault diagnosis. *Journal of Intelligent Manufacturing*, 33:973–983, 2022. arXiv: 1912.03688.

[200] R. Zhang, H. Tao, L. Wu and Y. Guan. Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. *IEEE Access*, 5:14347–14357, 2017.

[201] P. Cao, S. Zhang and J. Tang. Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning. *IEEE Access*, 6:26241–26253, 2018. arXiv: 1710.08904.

[202] S. Shao, S. McAleer, R. Yan and P. Baldi. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Transactions on Industrial Informatics*, 15:2446–2455, 2019.

[203] P. Liang, C. Deng, J. Wu and Z. Yang. Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network. *Measurement*, 159:107768, 2020.

[204] Q. Li, B. Tang, L. Deng, Y. Wu and Y. Wang. Deep balanced domain adaptation neural networks for fault diagnosis of planetary gearboxes with limited labeled data. *Measurement*, 156:107570, 2020.

[205] B. Yang, C.-G. Lee, Y. Lei, N. Li and N. Lu. Deep partial transfer learning network: A method to selectively transfer diagnostic knowledge across related machines. *Mechanical Systems and Signal Processing*, 156:107618, 2021.

[206] H. Zhong, S. Yu, H. Trinh, Y. Lv, R. Yuan and Y. Wang. Fine-tuning transfer learning based on DCGAN integrated with self-attention and spectral normalization for bearing fault diagnosis. *Measurement*, 210:112421, 2023.

[207] N. Tandon. A comparison of some vibration parameters for the condition monitoring of rolling element bearings. *Measurement*, 12:285–289, 1994.

[208] A. S. Sait and Y. I. Sharaf-Eldeen. A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. *Conference Proceedings of the Society for Experimental Mechanics Series*, volume 5, pages 307–324, 2011.

[209] S. Braun. The synchronous (time domain) average revisited. *Mechanical Systems and Signal Processing*, 25:1087–1102, 2011.

[210] N. Hurley and S. Rickard. Comparing measures of sparsity. *IEEE Transactions on Information Theory*, 55:4723–4741, 2009. arXiv: 0811.4706.

[211] D. Wang, Z. Peng and L. Xi. The sum of weighted normalized square envelope: A unified framework for kurtosis, negative entropy, Gini index and smoothness index for machine health monitoring. *Mechanical Systems and Signal Processing*, 140:106725, 2020.

[212] D. Wang. Spectral L2/L1 norm: A new perspective for spectral kurtosis for characterizing non-stationary signals. *Mechanical Systems and Signal Processing*, 104:290–293, 2018.

[213] Y. Miao, M. Zhao and J. Lin. Improvement of kurtosis-guided-grams via Gini index for bearing fault feature identification. *Measurement Science and Technology*, 28:125001, 2017.

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

[215] D. Wang, J. Zhong, C. Li and Z. Peng. Box-Cox sparse measures: A new family of sparse measures constructed from kurtosis and negative entropy. *Mechanical Systems and Signal Processing*, 160:107930, 2021.

[216] H. Qiu, J. Lee, J. Lin and G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, 289:1066–1090, 2006.

[217] J. I. Taylor. Identification of Bearing Defects by Spectral Analysis. *Journal of Mechanical Design*, 102:199–204, 1980.

[218] N. Baydar and A. D. Ball. Detection of gear deterioration under varying load conditions by using the instantaneous power spectrum. *Mechanical Systems and Signal Processing*, 14:907–921, 2000.

[219] R. B. Randall. Detection and diagnosis of incipient bearing failure in helicopter gearboxes. *Engineering Failure Analysis*, 11:177–190, 2004.

[220] W. J. Wang and P. D. McFadden. Early detection of gear failure by vibration analysis–I. Calculation of the time-frequency distribution. *Mechanical Systems and Signal Processing*, 7:193–203, 1993.

[221] E. Sejdić, I. Djurović and J. Jiang. Time–frequency feature representation using energy concentration: An overview of recent advances. *Digital Signal Processing*, 19:153–183, 2009.

[222] J. Allen. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 25:235–238, 1977.

[223] W. J. Staszewski, K. Worden and G. R. Tomlinson. Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition. *Mechanical Systems and Signal Processing*, 11:673–692, 1997.

[224] N. Baydar and A. Ball. A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distribution. *Mechanical Systems and Signal Processing*, 15:1091–1107, 2001.

[225] N. G. Nikolaou and I. A. Antoniadis. Rolling element bearing fault diagnosis using wavelet packets. *NDT & E International*, 35:197–205, 2002.

[226] N. Baydar and A. Ball. Detection of gear failures via vibration and acoustic signals using wavelet transform. *Mechanical Systems and Signal Processing*, 17:787–804, 2003.

[227] Z. K. Peng and F. L. Chu. Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography. *Mechanical Systems and Signal Processing*, 18:199–221, 2004.

[228] Y. Lei, J. Lin, Z. He and M. J. Zuo. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 35:108–126, 2013.

[229] R. Yan and R. X. Gao. Hilbert–Huang transform-based vibration signal analysis for machine health monitoring. *IEEE Transactions on Instrumentation and Measurement*, 55:2320–2329, 2006.

[230] H. Li, H. Zheng and L. Tang. Wigner-Ville Distribution Based on EMD for Faults Diagnosis of Bearing, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Volume 4223 LNAI, pages 803–812. 2006.

[231] Z. K. Peng, P. W. Tse and F. L. Chu. A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. *Mechanical Systems and Signal Processing*, 19:974–988, 2005.

[232] K. S. Wang and P. S. Heyns. Application of computed order tracking, Vold-Kalman filtering and EMD in rotating machine vibration. *Mechanical Systems and Signal Processing*, 25:416–430, 2011.

[233] J. Dybała and R. Zimroz. Rolling bearing diagnosing method based on empirical mode decomposition of machine vibration signal. *Applied Acoustics*, 77:195–203, 2014.

[234] S. Mohanty, K. K. Gupta and K. S. Raju. Comparative study between VMD and EMD in bearing fault diagnosis. *2014 9th International Conference on Industrial and Information Systems (ICIIS)*, pages 1–6. IEEE, 2014.

[235] X. Zhang, Q. Miao, H. Zhang and L. Wang. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery. *Mechanical Systems and Signal Processing*, 108:58–72, 2018.

[236] C. Wang, H. Li, G. Huang and J. Ou. Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum. *IEEE Access*, 7:31501–31516, 2019.

[237] H. Li, T. Liu, X. Wu and Q. Chen. An optimized VMD method and its applications in bearing fault diagnosis. *Measurement*, 166:108185, 2020.

[238] S. Ai, H. Li and Y. Zhang. Condition monitoring for bearing using envelope spectrum of EEMD. *2009 International Conference on Measuring Technology and Mechatronics Automation*, volume 1, pages 190–193. IEEE, 2009.

[239] M. Žvokelj, S. Zupan and I. Prebil. Non-linear multivariate and multiscale monitoring and signal denoising strategy using Kernel Principal Component Analysis combined with Ensemble Empirical Mode Decomposition method. *Mechanical Systems and Signal Processing*, 25:2631–2653, 2011.

[240] Y. Lei, M. J. Zuo and M. Hoseini. The use of ensemble empirical mode decomposition to improve bispectral analysis for fault detection in rotating machinery. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 224:1759–1769, 2010.

[241] Jinshan Lin and Qian Chen. Application of the EEMD method to multiple faults diagnosis of gearbox. *2010 2nd International Conference on Advanced Computer Control*, volume 2, pages 395–399. IEEE, 2010.

[242] J. Antoni. Cyclic spectral analysis in practice. *Mechanical Systems and Signal Processing*, 21:597–630, 2007.

[243] R. B. Randall, N. Sawalhi and M. D. Coats. A comparison of methods for separation of deterministic and random signals. *International Journal of Condition Monitoring*, 1:11–19, 2011.

[244] P. D. McFadden. Interpolation techniques for time domain averaging of gear vibration. *Mechanical Systems and Signal Processing*, 3:87–97, 1989.

[245] P. D. McFadden. Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration. *Mechanical Systems and Signal Processing*, 1:173–183, 1987.

[246] P. Borghesani. The envelope-based cyclic periodogram. *Mechanical Systems and Signal Processing*, 58:245–270, 2015.

[247] J. Antoni, G. Xin and N. Hamzaoui. Fast computation of the spectral correlation. *Mechanical Systems and Signal Processing*, 92:248–277, 2017.

[248] P. Borghesani and J. Antoni. A faster algorithm for the calculation of the fast spectral correlation. *Mechanical Systems and Signal Processing*, 111:113–118, 2018.

[249] W. Gousseau, J. Antoni, F. Girardin and J. Griffaton. Analysis of the rolling element bearing data set of the center for intelligent maintenance systems of the University of Cincinnati. *13th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies*, 2016.

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

[251] D. Ho and R. B. Randall. Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*, 14:763–788, 2000.

[252] D. Abboud and J. Antoni. Order-frequency analysis of machine signals. *Mechanical Systems and Signal Processing*, 87:229–258, 2017.

[253] D. Abboud, J. Antoni, S. Sieg-Zieba and M. Eltabach. Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment. *Mechanical Systems and Signal Processing*, 84:200–226, 2017.

[254] M. S. Darlow, R. H. Badgley and G. W. Hogg. Application of high-frequency resonance techniques for bearing diagnostics in helicopter gearboxes, 1974.

[255] D. Abboud, J. Antoni, S. Sieg-Zieba and M. Eltabach. Deterministic-random separation in nonstationary regime. *Journal of Sound and Vibration*, 362:305–326, 2016.

[256] R. B. Randall. A history of cepstrum analysis and its application to mechanical problems. *Mechanical Systems and Signal Processing*, 97:3–19, 2017.

[257] P. Borghesani, P. Pennacchi, R. B. Randall, N. Sawalhi and R. Ricci. Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions. *Mechanical Systems and Signal Processing*, 36:370–384, 2013.

[258] J. Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21:108–124, 2007.

[259] T. Barszcz and A. JabŁoński. A novel method for the optimal band selection for vibration signal demodulation and comparison with the kurtogram. *Mechanical Systems and Signal Processing*, 25:431–451, 2011.

[260] P. W. Tse and D. Wang. The design of a new sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as "two automatic vibration-based fault diagnostic methods using the novel sparsity measurement - Parts 1 and 2". *Mechanical Systems and Signal Processing*, 40:499–519, 2013.

[261] A. Moshrefzadeh and A. Fasana. The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mechanical Systems and Signal Processing*, 105:294–318, 2018.

[262] A. Mauricio, W. A. Smith, R. B. Randall, J. Antoni and K. C. Gryllias. Improved envelope spectrum via feature optimisation-gram (IESFOgram): A novel tool for rolling element bearing diagnostics under non-stationary operating conditions. *Mechanical Systems and Signal Processing*, 144:106891, 2020.

[263] P. Borghesani, P. Pennacchi and S. Chatterton. The relationship between kurtosis- and envelope-based indexes for the diagnostic of rolling element bearings. *Mechanical Systems and Signal Processing*, 43:25–43, 2014.

[264] Y. Miao, B. Zhang, J. Lin, M. Zhao, H. Liu, Z. Liu and H. Li. A review on the application of blind deconvolution in machinery fault diagnosis. *Mechanical Systems and Signal Processing*, 163:108202, 2022.

[265] R. A. Wiggins. Minimum entropy deconvolution. *Geoexploration*, 16:21–35, 1978.

[266] C. A. Cabrelli. Minimum entropy deconvolution and simplicity: A noniterative algorithm. *GEOPHYSICS*, 50:394–413, 1985.

[267] G. L. McDonald and Q. Zhao. Multipoint optimal minimum entropy deconvolution and convolution fix: Application to vibration fault detection. *Mechanical Systems and Signal Processing*, 82:461–477, 2017.

[268] H. Endo and R. B. Randall. Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. *Mechanical Systems and Signal Processing*, 21:906–919, 2007.

[269] N. Sawalhi. The application of spectral kurtosis to bearing diagnostics. *Acoustics - Conference*, pages 393–398, 2004.

[270] F. Combet and L. Gelman. Optimal filtering of gear signals for early damage detection based on the spectral kurtosis. *Mechanical Systems and Signal Processing*, 23:652–668, 2009.

[271] L. Cui, J. Wang and S. Lee. Matching pursuit of an adaptive impulse dictionary for bearing fault diagnosis. *Journal of Sound and Vibration*, 333:2840–2862, 2014.

[272] X. Jia, M. Zhao, Y. Di, C. Jin and J. Lee. Investigation on the kurtosis filter and the derivation of convolutional sparse filter for impulsive signature enhancement. *Journal of Sound and Vibration*, 386:433–448, 2017.

[273] X. Jia, M. Zhao, Y. Di, P. Li and J. Lee. Sparse filtering with the generalized lp/lq norm and its applications to the condition monitoring of rotating machinery. *Mechanical Systems and Signal Processing*, 102:198–213, 2018.

[274] Y. Miao, J. Wang, B. Zhang and H. Li. Practical framework of Gini index in the application of machinery fault feature extraction. *Mechanical Systems and Signal Processing*, 165:108333, 2022.

[275] C. López, D. Wang, Á. Naranjo and K. J. Moore. Box-cox-sparse-measures-based blind filtering: Understanding the difference between the maximum kurtosis deconvolution and the minimum entropy deconvolution. *Mechanical Systems and Signal Processing*, 165:108376, 2022.

[276] C. Peeters, J. Antoni and J. Helsen. Blind filters based on envelope spectrum sparsity indicators for bearing and gear vibration-based condition monitoring. *Mechanical Systems and Signal Processing*, 138:106556, 2020.

[277] M. Buzzoni, J. Antoni and G. D'Elia. Blind deconvolution based on cyclostationarity maximization and its application to fault identification. *Journal of Sound and Vibration*, 432:569–601, 2018.

[278] K. Kestel, C. Peeters, J. Antoni, Q. Leclère, F. Girardin and J. Helsen. Informed sparsity-based blind filtering in the presence of second-order cyclostationary noise. *Mechanical Systems and Signal Processing*, 199:110438, 2023.

[279] D. Wang, Y. Chen, C. Shen, J. Zhong, Z. Peng and C. Li. Fully interpretable neural network for locating resonance frequency bands for machine condition monitoring. *Mechanical Systems and Signal Processing*, 168:108673, 2022.

[280] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61:611–622, 1999.

[281] C. Fefferman, S. Mitter and H. Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29:983–1049, 2016.

[282] D. P. Kingma and M. Welling. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12:307–392, 2019. arXiv: 1906.02691.

[283] P. J. Diggle and R. J. Gratton. Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46:193–212, 1984.

[284] I. Kobyzev, S. J. D. Prince and M. A. Brubaker. Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43:3964–3979, 2021.

[285] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed and B. Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22:1–64, 2021. arXiv: 1912.02762.

[286] Y. Lecun, S. Chopra, R. Hadsell, M. Ranzato and F. J. Huang. A tutorial on energy-based learning. *Predicting structured data*, 1, 2006.

[287] S. Bond-Taylor, A. Leach, Y. Long and C. G. Willcocks. Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44:7327–7347, 2022.

[288] D. Eisenbud. Linear sections of determinantal varieties. *American Journal of Mathematics*, 110:541, 1988.

[289] L. Ruff, J. R. Kauffmann, R. A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T. G. Dietterich and K.-R. Muller. A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109:756–795, 2021.

[290] M. Hubert, M. Debruyne and P. J. Rousseeuw. Minimum covariance determinant and extensions. *WIREs Computational Statistics*, 10, 2018. arXiv: 1709.07045.

[291] C. M. Bishop, M. Svensén and C. K. I. Williams. GTM: The generative topographic mapping. *Neural Computation*, 10:215–234, 1998.

[292] C. M. Bishop, M. Svensén and C. K. I. Williams. Developments of the generative topographic mapping. *Neurocomputing*, 21:203–224, 1998.

[293] D. M. J. Tax and R. P. W. Duin. Support vector domain description. *Pattern Recognition Letters*, 20:1191–1199, 1999.

[294] B. Schölkopf, R. C. Williamson, A. Smola, J. Shawe-Taylor and J. Platt. Support vector method for novelty detection. *Advances in neural information processing systems*, volume 12, pages 582–588. MIT Press, 1999.

[295] R. Chalapathy, A. K. Menon and S. Chawla. Anomaly detection using one-class neural networks, 2018. arXiv: 1802.06360.

[296] L. Ruff, R. Vandermeulen, N. Grönitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller and M. Kloft. Deep one-class classification. *Proceedings of the 35th International Conference on Machine Learning*, volume 10, pages 4393–4402. PMLR, 2018.

[297] L. Kaufman and P. J. Rousseeuw. Finding groups in data: An introduction to cluster analysis. English. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 1990.

[298] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.

[299] N. E. Huang, S. R. Long and Z. Shen. The mechanism for frequency downshift in nonlinear wave evolution. *Advances in Applied Mechanics*, 32:59–117C, 1996.

[300] 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*, 454:903–995, 1998.

[301] K. Dragomiretskiy and D. Zosso. Variational mode decomposition. *IEEE Transactions on Signal Processing*, 62:531–544, 2014.

[302] N. ur Rehman and D. P. Mandic. Empirical mode decomposition for trivariate signals. *IEEE Transactions on Signal Processing*, 58:1059–1068, 2010.

[303] N. U. Rehman and H. Aftab. Multivariate variational mode decomposition. *IEEE Transactions on Signal Processing*, 67:6039–6052, 2019.

[304] G. E. P. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung. Time series analysis: Forecasting and control. English. Wiley series in probability and statistics. John Wiley & Sons, Inc., 2015.

[305] A. P. Dempster, N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39:1–22, 1977.

[306] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[307] H. Qiu, J. Lee, J. Lin, G. Yu and R. T. Services (2007). IMS, University of Cincinnati. Bearing Data Set, NASA Ames Prognostics Data Repository, 2007.

[308] P. J. L. Fernandes and C. McDuling. Surface contact fatigue failures in gears. *Engineering Failure Analysis*, 4:99–107, 1997.

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

[310] D. Abboud, M. Elbadaoui, W. A. Smith and R. B. Randall. Advanced bearing diagnostics: a comparative study of two powerful approaches. *Mechanical Systems and Signal Processing*, 114:604–627, 2019.

[311] D. W. Hogg. Data analysis recipes: Probability calculus for inference, 2012. arXiv: 1205.4446.

[312] T. Y. Lee, J. V. Zidek and N. Heckman. Dimensional analysis in statistical modelling, 2020. arXiv: 2002.11259.

[313] Y. Wang, Y. Peng, Y. Zi, X. Jin and K.-L. Tsui. A two-stage data-driven-based prognostic approach for bearing degradation problem. *IEEE Transactions on Industrial Informatics*, 12:924–932, 2016.

[314] S. Schmidt, R. Zimroz and P. S. Heyns. Enhancing gearbox vibration signals under time-varying operating conditions by combining a whitening procedure and a synchronous processing method. *Mechanical Systems and Signal Processing*, 156:107668, 2021.

[315] Y. Chen, S. Schmidt, P. S. Heyns and M. J. Zuo. A time series model-based method for gear tooth crack detection and severity assessment under random speed variation. *Mechanical Systems and Signal Processing*, 156:107605, 2021.

[316] D. H. Diamond, P. S. Heyns and A. J. Oberholster. Online shaft encoder geometry compensation for arbitrary shaft speed profiles using Bayesian regression. *Mechanical Systems and Signal Processing*, 81:402–418, 2016.

[317] S. Schmidt, P. S. Heyns and J. P. de Villiers. A novelty detection diagnostic methodology for gearboxes operating under fluctuating operating conditions using probabilistic techniques. *Mechanical Systems and Signal Processing*, 100:152–166, 2018.

[318] J.-Y. Lee and A. K. Nandi. Extraction of impacting signals using blind deconvolution. *Journal of Sound and Vibration*, 232:945–962, 2000.

[319] S. Schmidt, P. S. Heyns and K. C. Gryllias. An informative frequency band identification framework for gearbox fault diagnosis under time-varying operating conditions. *Mechanical Systems and Signal Processing*, 158:107771, 2021.

[320] B. Hou, D. Wang, J.-Z. Kong, J. Liu, Z. Peng and K.-L. Tsui. Understanding importance of positive and negative signs of optimized weights used in the sum of weighted normalized Fourier spectrum/envelope spectrum for machine condition monitoring. *Mechanical Systems and Signal Processing*, 174:109094, 2022.

[321] T. Yan, D. Wang, M. Zheng, C. Shen, T. Xia and Z. Peng. Interpretable sparse learned weights and their entropy based quantification for online machine health monitoring. *Mechanical Systems and Signal Processing*, 199:110493, 2023.

[322] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10:626–634, 1999.

[323] P. Comon. Independent component analysis, A new concept? *Signal Processing*, 36:287–314, 1994.

[324] M. E. Davies and C. J. James. Source separation using single channel ICA. *Signal Processing*, 87:1819–1832, 2007.

[325] J. Ho, A. Jain and P. Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 2020-Decem:1–25, 2020. arXiv: 2006.11239.

[326] C. Cremer, X. Li and D. Duvenaud. Inference suboptimality in variational autoencoders:1–14, 2018. arXiv: 1801.03558.

[327] K. Preechakul, N. Chatthee, S. Wizadwongsa and S. Suwajanakorn. Diffusion autoencoders: Toward a meaningful and decodable representation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2022. arXiv: 2111.15640.

[328] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed and A. Lerchner. β-VAE: Learning basic visual concepts with a constrained variational framework. *International conference on learning representations*, 2016.

[329] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever and P. Abbeel. InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets, 2016.

[330] A. Hyvärinen. New approximations of differential entropy for independent component analysis and projection pursuit. *Advances in Neural Information Processing Systems*, 10:273–279, 1998.

[331] J. Zhu, T. Nostrand, C. Spiegel and B. Morton. Survey of condition indicators for condition monitoring systems. *Annual Conference of the PHM Society*, 6:635–647, 2014.

[332] M. Feldman. Hilbert transform in vibration analysis. *Mechanical Systems and Signal Processing*, 25:735–802, 2011.

[333] S. Schmidt, R. Zimroz, F. Chaari, P. S. Heyns and M. Haddar. A simple condition monitoring method for gearboxes operating in impulsive environments. *Sensors*, 20:2115, 2020.

[334] B. Chen, Y. Cheng, W. Zhang, F. Gu and G. Mei. Optimal frequency band selection using blind and targeted features for spectral coherence-based bearing diagnostics: A comparative study. *ISA Transactions*, 127:395–414, 2022.

[335] Z. Chen, A. Mauricio, W. Li and K. C. Gryllias. A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. *Mechanical Systems and Signal Processing*, 140:106683, 2020.

[336] P. Borghesani and M. R. Shahriar. Cyclostationary analysis with logarithmic variance stabilisation. *Mechanical Systems and Signal Processing*, 70-71:51–72, 2016.

[337] B. W. Silverman. Density estimation for statistics and data analysis, volume 60. Routledge, 2018, pages 129–136.

[338] E. Bechhoefer and M. Kingsley. A review of time synchronous average algorithms. *Annual Conference of the Prognostics and Health Management Society, PHM 2009*:1–10, 2009.

[339] W. Wang. Early detection of gear tooth cracking using the resonance demodulation technique. *Mechanical Systems and Signal Processing*, 15:887–903, 2001.

[340] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai and S. Chintala. PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, volume 33, 2019. arXiv: 1912.01703.

[341] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2014. arXiv: 1412.6980.

[342] I. Loshchilov and F. Hutter. Decoupled weight decay regularization, 2017. arXiv: 1711.05101.

[343] S. Schmidt, P. S. Heyns and K. C. Gryllias. A discrepancy analysis methodology for rolling element bearing diagnostics under variable speed conditions. *Mechanical Systems and Signal Processing*, 116:40–61, 2019.

[344] K. Pearson. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2:559–572, 1901.

[345] C. Jutten and J. Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

[346] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.

[347] D. T. Pham and P. Garat. Blind separation of mixture of independent sources through a quasi-maximum likelihood approach. *IEEE Transactions on Signal Processing*, 45:1712–1725, 1997.

[348] S. Haykin. Adaptive filter theory. Pearson, 2014.

[349] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, Š. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman and A. Scopatz. SymPy: Symbolic computing in Python. *PeerJ Computer Science*, 3:e103, 2017.

[350] 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*, 585:357–362, 2020. arXiv: 2006.10256.

[351] L. Rosasco, E. D. Vito and A. Verri. Spectral methods for regularization in learning theory. *Journal of Machine Learning Research*, 2006.

[352] A. N. Tikhonov and V. I. A. Arsenin. Solutions of ill-posed problems. Halsted Press book. Winston, 1977.

[353] G. H. Golub, C. F. V. Loan and C. F. Van Loan. Matrix computations. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2013.

[354] L. L. Gerfo, L. Rosasco, F. Odone, E. D. Vito and A. Verri. Spectral algorithms for supervised learning. *Neural Computation*, 20:1873–1897, 2008.

[355] T. Miyato, T. Kataoka, M. Koyama and Y. Yoshida. Spectral normalization for generative adversarial networks. *International Conference on Learning Representations*, 2018.

[356] W. Xie, J. Lei, B. Liu, Y. Li and X. Jia. Spectral constraint adversarial autoencoders approach to feature representation in hyperspectral anomaly detection. *Neural Networks*, 119:222–234, 2019.

[357] J. A. Snyman and D. N. Wilke. Practical Mathematical Optimization. English, volume 133 of *Springer Optimization and Its Applications*. Springer International Publishing, 2018, pages 1–364.

[358] R. L. Burden, J. D. Faires and A. M. Burden. Numerical analysis. Cengage Learning, 2016.

[359] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. Vanderplas, A. Joly, B. Holt and G. Varoquaux. API design for machine learning software: experiences from the scikit-learn project, 2013. arXiv: 1309.0238.

[360] Python package index - PyPI.

[361] J. D. Hunter. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9:90–95, 2007.

[362] 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 É. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[363] C. da Costa-Luis, S. K. Larroque, K. Altendorf, H. Mary, Richardsheridan, M. Korobov, N. Raphael, I. Ivanov, M. Bargull, N. Rodrigues, G. Chen, A. Lee, C. Newey, CrazyPython, JC, M. Zugnoni, M. D. Pagel, Mjstevens777, M. Dektyarev, A. Rothberg, A. Plavin, D. Panteleit, F. Dill, FichteFoll, G. Sturm, HeoHeo, H. van Kemenade, J. McCracken, MapleCCC and M. Nordlund. tqdm: A fast, extensible progress bar for Python and CLI, 2023.

Appendix A

Supplementary material

A.1 Model parameters for Chapter 2

It is pertinent that all applicable properties of the Latent Variable Model (LVM)s used in Chapter 2 are detailed. Pytorch [340] was used to optimise the Variational Auto-Encoder (VAE) model parameters. The system used to train the models had an Intel i7-8750H processor and a Geforce RTX 2070 graphics card. The properties of the Probabilistic Principal Component Analysis (PPCA) and VAE models for the datasets considered are given in Table A.1 and Table A.2 respectively. The Cumulative Contribution Rate (CCR) was used to obtain the number of latent variables d for the PPCA model [101]. In Table A.3 the network architectures of the VAE models used in Chapter 2 are detailed. The latent encoding and data generation functions were chosen to be mirror images of each other. This was done by selecting the properties of the one-dimensional convolutional and de-convolutional layers such that an equivalent representation expansion or reduction is obtained from one layer to the next. The controllable properties of the convolutional and de-convolutional layers are the stride length L_{stride}, the kernel size L_{kernel}, and the padding length L_{pad}. For the VAE models, the validation model loss was monitored to ensure that no overfitting occurred. The model parameters for the training epoch with the lowest validation loss were retained and used for condition inference. The Adam algorithm [341, 342] was used to optimise the VAE model parameters.

Table A.1. The PPCA model properties used for model training.

Dataset type	L_w	L_{sft}	CCR
IMS	512	256	80%
Gearbox	512	256	80%

A.2 Model parameters for Chapter 3

For the investigations performed in Chapter 3 of this thesis, the parameter reported in Table A.4 were used. Table A.4 reports both the parameters related to the data pre-processing and

Table A.2. The VAE model properties used for model training.

Dataset type	L_w	L_{sft}	Latent dimensionality	Number of training epochs	Learning rate	Batch size (B)
IMS	512	256	$d = 50$	2000	$1e^{-4}$	512
Gearbox	512	256	$d = 50$	2000	$1e^{-4}$	512

Table A.3. The network architectures for both the VAE latent encoding networks and data generation networks. A depth order of 0–6 define the latent encoding network, while the generative network is the reverse of this order. Note that the posterior mean and covariance are both functions of the fourth network depth level. The parameter B is the batch size. The activation functions for the hidden layers of the encoding network are TanH, while for the generative network the hidden layer activation functions are ReLU. Note that ↓ denotes a convolutional layer, while ↑ denotes a de-convolutional layer. The dash symbol (–) is used to indicate that the parameter is not applicable.

Network depth	Layer operator	Output layer dimensionality	L_{stride}	L_{kernel}	L_{pad}	Activation function
0	Input layer	$\mathbb{R}^{B\times L_w}$	–	–	–	Linear
1	↓/↑	$\mathbb{R}^{B\times 32\times \frac{L_w}{4}}$	4	32	14	TanH (↓) / ReLU (↑)
2	↓/↑	$\mathbb{R}^{B\times 64\times \frac{L_w}{16}}$	4	32	14	TanH (↓) / ReLU (↑)
3	↓/↑	$\mathbb{R}^{B\times 128\times \frac{L_w}{64}}$	4	32	14	TanH (↓) /ReLU (↑)
4	Fully-connected	$\mathbb{R}^{B\times \frac{L_w}{2}}$	–	–	–	TanH (↓) /ReLU (↑)
5 (posterior mean)	Fully-connected	$\mathbb{R}^{B\times d}$	–	–	–	Linear
6 (posterior covariance)	Fully-connected	$\mathbb{R}^{B\times d}$	–	–	–	Softplus

evaluation steps and the model properties for the considered LVMs.

Table A.4. The parameters of the LVMs used in the investigations performed in Chapter 3 of this thesis. The dash symbol (–) is used to indicate that the parameter is not applicable to the referenced LVM formulation.

Dataset investigated	LVM formulation	L_w	L_{sft} (Training)	L_{sft} (Evaluation) [164]	CCR	Latent dimensionality (d)
Phenomenological gearbox	PPCA	512	1	1	80%	–
	SRICA	512	1	1	–	30
	$PPCA_h$	512	256	1	80%	–
IMS	PPCA	256	1	1	80%	–
	SRICA	256	1	1	–	20
	$PPCA_h$	256	128	1	80%	–
Gearbox	PPCA	512	1	1	95%	–
	SRICA	512	1	1	–	60
	$PPCA_h$	512	256	1	95%	–

A.3 Model parameters for Chapter 4

All LVM parameters related to the investigations performed in Chapter 4 are detailed in this section. The properties pertaining to the latent dimensionality of the considered LVMs are presented in Table A.5. The parameters related to the pre-processing and evaluation of the available vibration data are presented in Table A.6 for the two considered LVM training approaches.

Table A.5. The properties used for the LVMs considered in Chapter 4 of this thesis related to the latent dimensionality (d) as a function of the indicator category (offline or online).

Experimental dataset	LVM formulation and indicator category				
	PPCA_h (offline indicators)	PPCA (offline indicators)	PPCA_O (online indicators)	SRICA (offline indicators)	SRICA_O (online indicators)
IMS	80% CCR	$d = 94$	80% CCR	$d = 20$	$d = 20$
Gearbox	95% CCR	$d = 310$	95% CCR	$d = 60$	$d = 60$

Table A.6. The parameters for the LVMs used in the investigations performed in Chapter 4 related to the data pre-processing steps discussed in Section 3.2.3 for the LVMs developed using the offline and online training approaches.

LVM training approach	Experimental dataset	L_w	L_{sft} (Training)	L_{sft} (Evaluation) [164]
Offline	IMS	256	128	1
	Gearbox	512	256	1
Online	IMS	256	1	1
	Gearbox	512	1	1

A.4 Signal-to-noise ratio estimation procedure

The process to calculate the diagnostic metrics based off the Signal-to-Noise Ratio (SNR) is demonstrated in Figure A.1 for the Square Envelope Spectrum (SES) of an indicator signal. The SNR of the SES, detailed in Equation (3.22), requires partitioning the SES around specific integer harmonics of a characteristic cyclic frequency. Consequently, the maximum value for the characteristic frequency is obtained and the noise floor is estimated using the median of the partitioned SES segment.

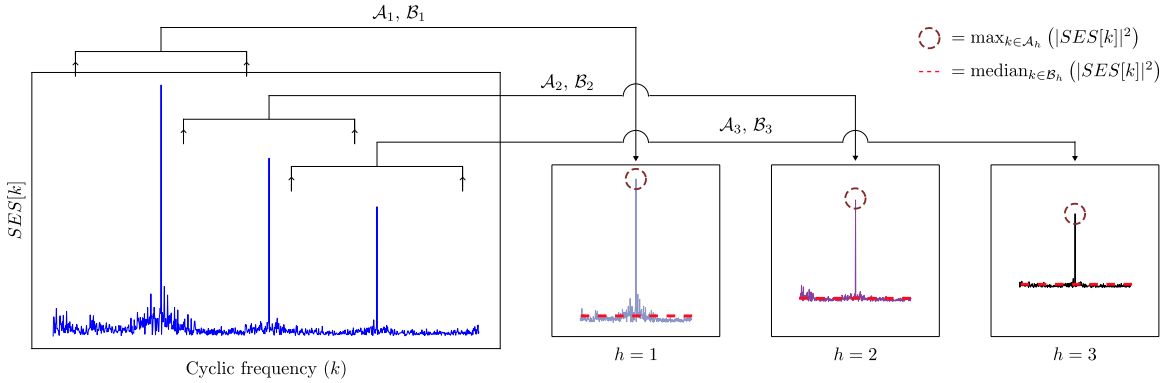


Figure A.1. An example of the procedure required to calculate the SNR of the SES of an indicator signal.

Figure A.2 demonstrates the process for the synchronous representation of an indicator signal. The location of the expected healthy and damaged portions of the synchronous representation is known a priori, and the representative sections are extracted and their power is determined to estimate the SNR given in Equation (3.26) and the relative SNR given in Equation (4.30).

A.5 Relationship between LVM objectives and signal processing

In this section of the supplementary material, a simple connection between the considered Probabilistic Principal Component Analysis (PPCA) and negentropy-based Independent Component Analysis (ICA) methods used in this thesis and existing signal processing methodologies is highlighted. The purpose is to detail what characteristics of the observed data the considered methods extract and represent via the latent manifold. This is important to the online training approach as the Latent Variable Model (LVM)s extract information

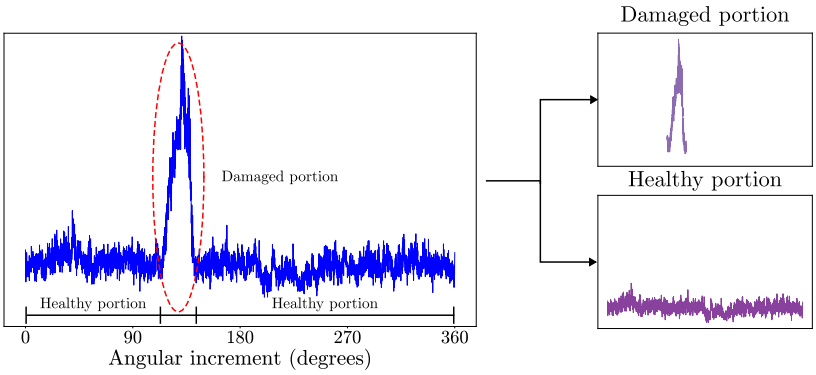


Figure A.2. An example of the procedure required to calculate the SNR from the synchronous representation of an indicator signal.

dynamically through record time, thereby responding to changes in the recorded data based off their specific characteristics of interest.

The Principal Component Analysis (PCA) objective function can be given as

$$\begin{aligned} \max_{\mathbf{u}_j} \quad & \mathcal{L}_{PCA}(\mathbf{u}_j) = \mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{\left(\mathbf{u}_j^T\mathbf{x}\right)^2\} \\ \text{s.t.} \quad & \mathbf{u}_j^T\mathbf{u}_l = \delta_{jl} \end{aligned} \tag{A.1}$$

where δ_{jl} represents the Kronecker delta function, and $\mathbf{x}$ is assumed to be pre-processed using Equation (3.7). Equation (A.1) represents the maximisation of the variance of the j^{th} projection of the data $\mathbf{x}$, i.e., the j^{th} latent variable [114]. Equation (A.1) can be reformulated into a Rayleigh quotient function $R(\mathbf{C}, \mathbf{x})$ for the symmetric covariance matrix $\mathbf{C}$ and can be solved as an eigenvalue problem or via the singular value decomposition of $\overline{\mathbf{X}}$ [101].

The negentropy-based Independent Component Analysis (ICA) objective function can be represented as

$$\begin{aligned} \max_{\mathbf{w}_j} \quad & \mathcal{L}_{ICA}(\mathbf{w}_j) = J(z_j) \approx c \cdot \left(\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{G(\mathbf{w}_j^T\mathbf{x})\} - \mathbb{E}_{\nu\sim p(\nu)}\{G(\nu)\}\right)^2 \\ \text{s.t.} \quad & \mathbf{w}_j^T\mathbf{w}_l = \delta_{jl} \end{aligned} \tag{A.2}$$

where $J(z_j)$ is the negentropy of the j^{th} latent component, $G(\cdot)$ is a non-quadratic function, $c \in \mathbb{R}^+$ is an arbitrary constant, $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is a standardised Gaussian variable with variance $\sigma_\nu^2 = \mathbb{E}_{z\sim p(z_j)}\{(z - \mathbb{E}\{z\})^2\} = \sigma_{z_j}^2$, and $\mathbf{x}$ is pre-processed using Equation (3.7) and Equation (3.8). This objective function is presented in Hyvärinen [330], and a fixed-point algorithm called FastICA was subsequently proposed to estimate $\mathbf{W}$ [322].

The implication of the data Hankelisation step in Equation (2.13) for the considered online LVMs, for the case where $L_{sft} = 1$, is that the Spectrally Regularised Independent Component Analysis (SRICA) and PPCA component vectors operate as a set of Finite Impulse Response (FIR) filters [178, 324]. This allows for strong correlations to be drawn between the LVMs used in this thesis and known signal processing-based methodologies. First, consider the PCA objective function given in Equation (A.1). It is trivial to show, for vibration-based applications, that the optimal solution for the PCA objective given in Equation (A.1) can be reformulated into

$$\begin{aligned}\mathcal{L}_{PCA}(\mathbf{u}_j) &= \mathbf{u}_j^T \mathbf{C} \mathbf{u}_j \\ &= \mathbb{E}\{z_j^2\} \\ &= \lambda_j, \end{aligned} \tag{A.3}$$

where the covariance matrix $\mathbf{C} = \frac{1}{N}\overline{\mathbf{X}}^T\overline{\mathbf{X}}$ also represents the auto-correlation matrix, $(\lambda_j, \mathbf{u}_j)$ is the j^{th} eigenpair that diagonalises $\mathbf{C}$ and $\mathbb{E}\{z_j^2\}$ represents the power of the j^{th} latent signal. In this setting, the PCA objective seeks to maximise the power of the filtered signal, and $\mathbf{u}_j$ is a FIR filter driven to capture components with maximal power.

Secondly, consider the ICA objective function given in Equation (A.2). It is trivial to show that for the choice $G(u) = u^4$ Equation (A.2) becomes

$$\begin{aligned}\mathcal{L}_{ICA}(\mathbf{w}_j) &= \left(\mathbb{E}\{z_j^4\} - \mathbb{E}\{\nu^4\}\right)^2 \\ &= \left(\mathbb{E}\left\{\left(\frac{z_j - \mu_{z_j}}{\sigma_{z_j}}\right)^4\right\} - 3 \cdot \sigma_\nu^4\right)^2 \\ &= \left(\text{kurt}\left(z_j\right) - 3\right)^2, \end{aligned} \tag{A.4}$$

where $\mu_{z_j} = 0$ and $\sigma_{z_j} = 1$ as the data is assumed to be centred and whitened, and $c = 1$. Equation (A.4) represents the square excess kurtosis of latent component z_j. Hence, the ICA objective seeks to maximise the non-Gaussianity of the filtered signal by maximising its kurtosis, or a generalisation thereof, and $\mathbf{w}_j$ operates as a FIR filter that captures the dominant non-Gaussian components of the signal. Hence, it is shown that the considered online LVM methodologies represent fundamentally different latent manifolds, and the learnt latent components capture the sources of maximal power or maximal non-Gaussianity. However, unlike blind deconvolution [264] or band demodulation [32] methodologies, which seek a single transformation, i.e., $d = 1$, that maximises some measure, e.g., the $L2/L1$ norm [212], the considered online linear LVMs seek d latent variables, i.e., d transformations that capture the

intrinsic features in the observed data in alignment with their objective functions.

A.6 Additional details for the online LVMs

In this section, details are given to the PPCA and SRICA models used for online training. The objective is to clarify how aspects of the models are computed and what approaches are employed to estimate the model parameters. For this discussion, it is assumed that a vibration signal $x_i[n]$, with $n \subseteq \mathbb{Z}^+ = \{1, 2, \ldots, L_{signal}\}$, is available and has been processed into a matrix $\mathbf{X}_i \in \mathbb{R}^{L_T \times L_w}$ via Equation (2.13) and has been centred using Equation (3.7). To assist with the legibility of this section, certain equations from the main text of the manuscript are duplicated to provide further context. For the remainder of this discussion, the variable D refers to the dimensionality of the observed data vector $\mathbf{x}$, i.e., $\mathbf{x} \in \mathbb{R}^D$, d refers to the dimensionality of the latent vector $\mathbf{z}$, i.e., $\mathbf{z} \in \mathbb{R}^d$, and the window length parameter L_w defines the dimensionality of $\mathbf{x}$, i.e., $D = L_w$.

A.6.1 Reconstruction-focused model

For the PPCA model, the generative distribution $p(\mathbf{x}|\mathbf{z})$ is given by

$$p(\mathbf{x}|\mathbf{z}) = \mathcal{N}\left(\mathbf{x}|\mathbf{A}\mathbf{z} + \boldsymbol{\mu}_{\mathbf{x}}, \sigma^2_{PPCA}\mathbf{I}\right), \tag{A.5}$$

and the posterior distribution $p(\mathbf{z}|\mathbf{x})$ is given by

$$p(\mathbf{z}|\mathbf{x}) = \mathcal{N}\left(\mathbf{z}|\mathbf{M}^{-1}_{PPCA}\mathbf{A}^T\left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}}\right), \sigma^{-2}_{PPCA}\mathbf{M}_{PPCA}\right), \tag{A.6}$$

where $\mathbf{M}_{PPCA} \in \mathbb{R}^{d\times d} = \mathbf{A}^T\mathbf{A} + \sigma^2_{PPCA} \cdot \mathbf{I}$, σ^2_{PPCA} is a scalar, and $\mathbf{A} \in \mathbb{R}^{D\times d}$ is an unknown matrix. It is clear that the PPCA model can be fully expressed with the $\boldsymbol{\mu}_{\mathbf{x}}$, $\mathbf{A}$, and σ^2_{PPCA} terms. In the work of Tipping and Bishop [280], maximum likelihood estimators for these terms are analytically derived and are given as

$$\boldsymbol{\mu}_{\mathbf{x},ML} = \mathbb{E}\{\mathbf{x}\}, \tag{A.7}$$

$$\mathbf{A}_{ML} = \mathbf{U}_d\left(\mathbf{L}_d - \sigma^2_{PPCA,ML}\mathbf{I}\right)^{1/2}\mathbf{R}, \tag{A.8}$$

$$\sigma^2_{PPCA,ML} = \frac{1}{D-d}\sum_{i=d+1}^{D}\lambda_i, \tag{A.9}$$

where $\mathbf{U}_d \in \mathbb{R}^{D\times d}$ is a matrix whose columns represent the first d eigenvectors of the covariance matrix $\mathbf{C} = \mathbb{E}\{\mathbf{x}\mathbf{x}^T\}$, $\mathbf{L}_d \in \mathbb{R}^{d\times d} = \text{diag}\left(\lambda_1, \cdots, \lambda_d\right)$ is a diagonal matrix of the first d eigenvalues of $\mathbf{C}$, $\mathbf{R} \in \mathbb{R}^{d\times d}$ is an arbitrary orthogonal matrix assumed to be an identity matrix, i.e., $\mathbf{R} = \mathbf{I}$, and $\sigma^2_{PPCA,ML}$ represents the average variance of the unused latent components [101, 280]. As $\mathbf{X}_i$ is centred, the mean vector $\boldsymbol{\mu}_{\mathbf{x},ML}$ is the zero vector. Hence, the

online PPCA model can be fully described by knowledge of the eigenvectors and eigenvalues of $\mathbf{X}_i$. These two parameters are estimated using the singular value decomposition of $\mathbf{X}_i$. Equation (3.3) can be readily obtained by expanding the linear algebra presented in this section.

A.6.2 Interpretation-focused model

For the SRICA model, the objective is to obtain a encoding transformation for the latent variables wherein the latent components are maximally independent designated by Equation 4.8. The encoding transformation for the SRICA model is given by

$$\mathbf{z} = \mathbf{W}\mathbf{x}. \tag{A.10}$$

Non-gaussianity is a popular proxy to independence and can be used to estimate the parameters of $\mathbf{W} \in \mathbb{R}^{d\times D} = [\mathbf{w}_1, \cdots, \mathbf{w}_d]^T$ [330]. Negentropy is one such measure of non-Gaussianity and a popular approach for interpretation-focused methods is to maximise an approximation of negentropy given via

$$\mathcal{L}_{negen}(z_j) = \left(\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{G(\mathbf{w}_j^T\mathbf{x})\} - \mathbb{E}_{\nu\sim p(\nu)}\{G(\nu)\}\right)^2, \tag{A.11}$$

where $G(\cdot)$ is a non-quadratic function, $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is a standardised Gaussian variable with the same variance as z_j, and the data vector is assumed to have been pre-processed using Equation (3.7) and Equation (3.8) [330]. The exponential function

$$G(u) = -\exp\left(-\frac{u^2}{2}\right) \tag{A.12}$$

is used in this thesis as it is a robust estimator of negentropy [178, 322]. To reduce the issue of source duplication identified in Davies and James [324] and Hyvärinen [178], a spectral regularisation term is used. This term is given by

$$\mathcal{L}_{sr}(\mathbf{w}_j) = \alpha \sum_{l=1}^{j-1} \mathbf{b}(\mathbf{w}_j)^T\mathbf{b}(\mathbf{w}_l), \quad j > 1 \tag{A.13}$$

where $\mathbf{b}(\mathbf{w})$ is the Fourier representation of the vector $\mathbf{w}$ and α is a penalty enforcement parameter. More information for this term is provided in Appendix C. Hence, the full SRICA objective for one latent component is given via

$$\begin{aligned} \min_{\mathbf{w}_j} \quad & -\mathcal{L}_{negen}(\mathbf{w}_j) + \mathcal{L}_{sr}(\mathbf{w}_j) \\ \text{s.t.} \quad & \mathbf{w}_j^T\mathbf{w}_j = 1, \end{aligned} \tag{A.14}$$

which can be reformulated into a Lagrangian expression

$$\mathcal{L}(\mathbf{w}_j, \lambda_{eq}) = \mathcal{L}_{model}(\mathbf{w}_j) + \mathcal{L}_{sr}(\mathbf{w}_j) + \lambda_{eq}\left(\mathbf{w}_j^T\mathbf{w}_j - 1\right), \tag{A.15}$$

where λ_{eq} represents a Lagrange multiplier. In this work, Newton's method is applied to the Lagrangian expression to obtain a solution to the SRICA objective. This results in an iterative update process

$$\mathbf{w}_j^{k+1} = \mathbf{w}_j^k + \eta_j^k \cdot \nabla\mathbf{w}_j, \tag{A.16}$$

$$\lambda_{eq}^{k+1} = \lambda_{eq}^k + \gamma_j^k \cdot \nabla\lambda_{eq,j}, \tag{A.17}$$

where η_j^k and γ_j^k are learning rate parameters that are obtained via a univariate line search to satisfy the Armijo condition [166]. The update terms $\Delta\mathbf{w}_j$ and $\Delta\lambda_{eq,j}$ are computed using Newton's method by solving

$$\mathbf{H}(\boldsymbol{\phi}_j^{(k)})\Delta\boldsymbol{\phi}_j^{(k)} = -\nabla\mathcal{L}(\boldsymbol{\phi}_j^{(k)}), \tag{A.18}$$

where $\boldsymbol{\phi}_j^{(k)} \in \mathbb{R}^{d+1} = \begin{bmatrix} \Delta\mathbf{w}_j \\ \Delta\lambda_{eq,j} \end{bmatrix}$ represents the parameter update vector and $\nabla\mathcal{L}(\boldsymbol{\phi}_j^{(k)}) \in \mathbb{R}^{d+1}$ and $\mathbf{H}(\boldsymbol{\phi}_j^{(k)}) \in \mathbb{R}^{(d+1)\times(d+1)}$ represent the gradient vector and Hessian of the Lagrangian expression, respectively. A termination condition is used on the iterative change in $\mathbf{w}_j$ and is given by

$$|\mathbf{w}_j^{(k+1),T}\mathbf{w}_j^{(k)} - 1| \leq \epsilon_{tol}, \tag{A.19}$$

where ϵ_{tol} is a tolerance parameter. As the parameter estimation process is performed sequentially for each vector $\mathbf{w}_j$, $j \in [1, d]$, a orthogonalisation step is included during the parameter estimation process to ensure that the latent components are decorrelated [185]. A Gram-Schmidt orthogonalisation step is used and this is given via

$$\mathbf{w}_{j,orth} = \mathbf{w}_j - \sum_{l=1}^{j-1} \frac{\mathbf{w}_j^T\mathbf{w}_l}{\|\mathbf{w}_l\|_2} \cdot \mathbf{w}_l. \tag{A.20}$$

Appendix B

Phenomenological gearbox model information

The numerical gearbox signal $x_{ng}(t)$ represented via Equation (3.16) is composed of a gear mesh component $x_{gm}(t)$, a distributed pinion damage component $x_{dp}(t)$, a noise component $x_n(t)$, and an outer race bearing damage component $x_{bo}(t)$. This numerical model is derived from the model used by Abboud *et al.* [253].

The gear mesh component of the numerical gearbox signal simulates the interaction between healthy gear teeth in a parallel gearbox

$$x_{gm}(t) = \gamma_{gm} \cdot M(\omega(t)) \cdot h_{gm}(t) \otimes \left(\sum_{l=1}^{N_{gm}} A_{gm}^{(l)} \cdot \sin\left(l \cdot N_{teeth} \cdot \int_0^t \omega(\tau) d\tau + \phi_{gm}^{(l)} \right) \right), \quad \text{(B.1)}$$

where γ_{gm} refers to a scale factor used to control the prominence of $x_{gm}(t)$ with respect to $x_n(t)$, $M(\omega(t)) = \omega(t)$ represents a linear amplitude modulation function that is dependent on the gear shaft speed $\omega(t)$, $h_{gm}(t)$ represents a single-degree-of-freedom viscously damped impulse response function, $\otimes$ represents the convolution operation, N_{teeth} is the number of teeth on the gear connected to the reference shaft, and $A_{gm}^{(l)}$ and $\phi_{gm}^{(l)}$ represent the magnitude and phase of the l^{th} harmonic in a tuple of N_{gm} integer harmonics. The properties of the numerical gearbox signal are similar to those used in Schmidt *et al.* [343]. The number of gear teeth is set to $N_{teeth} = 30$, the number of gear harmonic components is set to $N_{gm} = 10$ to ensure adequate component excitation, the magnitude $A_{gm} = (1, 1.5, 2, 1, 0.5, 0.3, 0.2, 0.1, 0.1, 0.05)$ is a N_{gm}-tuple with dominant second and third component components based off phenomena observed in Ref. [43], and the phase $\phi_{gm} = (0, \cdots, 0)$ is a zero N_{gm}-tuple.

The distributed pinion damage represents a smooth second-order cyclostationary component, which is modulated white Gaussian noise running at integer harmonics of the pinion shaft speed, and is incorporated to complicate the condition inference task

$$x_{dp}(t) = \gamma_{dp} \cdot M(\omega(t)) \cdot h_{dp}(t) \otimes \left(\epsilon_{dp}(t) \cdot \sum_{l=1}^{N_{dp}} A_{dp}^{(l)} \cdot \sin\left(l \cdot GR \cdot \int_0^t \omega(\tau) d\tau + \phi_{dp}^{(l)} \right) \right), \quad \text{(B.2)}$$

where GR represents the gear ratio and $\epsilon_{dp}(t) \sim \mathcal{N}(0, 1)$ represents a zero mean, unit variance

Gaussian variable. The $x_{dp}(t)$ component has $N_{dp} = 3$ integer harmonics, the magnitude $A_{dp} = (1, 2, 3)$ is a N_{dp}-tuple, and the phase $\phi_{dp} = (0, 0, 0)$ is a N_{dp}-tuple. A gear ratio of $GR = 3$ is used for this investigation.

The noise component is amplitude-modulated white noise

$$x_n(t) = \gamma_n \cdot M(\omega(t)) \cdot \epsilon_n(t), \tag{B.3}$$

which consists of a zero mean, unit variance Gaussian function $\epsilon_n(t)$ scaled by the noise scale parameter γ_n and the amplitude modulated by $M(\omega(t))$.

The outer race-bearing fault component is represented by a train of Dirac impulses

$$x_{bo}(t) = \gamma_{bo} \cdot M(\omega(t)) \cdot h_{bo}(t) \otimes \left(\sum_{l=1}^{N_{bo}} A_{bo}^{(l)} \cdot \delta \left(t - T_{bo}^{(l)} \right) \right) \tag{B.4}$$

where $A_{bo} \sim \mathcal{N}\left(\mu_{bo}, \sigma_{bo}^2\right)$ and the impact times $T_{bo}^{(l)}$ are estimated using the expected angle of the impacts, and bearing slip is represented by a jitter model where $\tau_l \sim \mathcal{N}(0, \sigma_\tau^2)$ represents the per-impulse jitter that is localised around each expected impact angle [76]. The Ball Pass Frequency for an Outer race fault (BPFO) is set to 3.57 gear shaft orders. The variation in the impulse amplitude is controlled by sampling $A_{bo} \sim \mathcal{N}\left(1, 0.1^2\right)$, and the jitter variance is set to $\sigma_\tau^2 = 0.1^2$.

The impulse response function $h_j(t)$ is represented by a

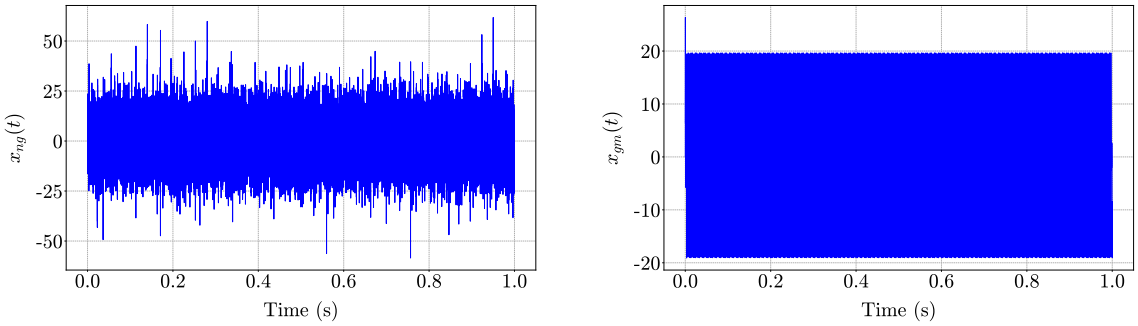
$$h_j(t) = e^{-2 \cdot \pi \cdot \zeta_j \cdot f_{n,j} \cdot t} \cdot \sin\left(2 \cdot \pi \cdot \sqrt{1 - \zeta_j^2} \cdot f_{n,j} \cdot t\right), \tag{B.5}$$

where ζ_j and $f_{n,j}$ represent the damping ratio and the natural frequency for the j^{th} component of $x_{ng}(t)$, respectively. In this investigation, the damping ratio $\zeta_{gm} = \zeta_{dp} = \zeta_{bo} = 0.05$ is constant for all gearbox components, while the natural frequency $f_{n,gm} = 2.5\,\text{kHz}$, $f_{n,dp} = 5\,\text{kHz}$, $f_{n,bo} = 8\,\text{kHz}$, is unique to each gearbox component. The scale factor γ_j applied to the components of the phenomenological model is used to control the prominence of different components with respect to the noise component $x_n(t)$. This is achieved via the Signal-to-Noise Ratio (SNR) in the logarithmic decibel scale

$$SNR_{dB,j} = 10 \log_{10} \left(\frac{\frac{1}{T} \int_0^T |x_j(t)|^2 dt}{\frac{1}{T} \int_0^T |x_n(t)|^2 dt} \right), \tag{B.6}$$

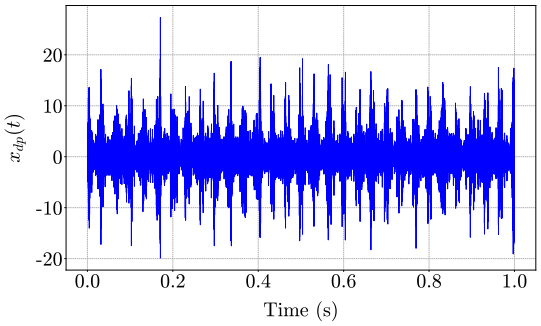
which represents the component power ratio of the j^{th} numerical gearbox component $x_j(t)$ and the noise component $x_n(t)$ respectively.

The $x_{gm}(t)$ component scaled such that $SNR_{dB,gm} = 2dB$, the $x_{dp}(t)$ is scaled such that

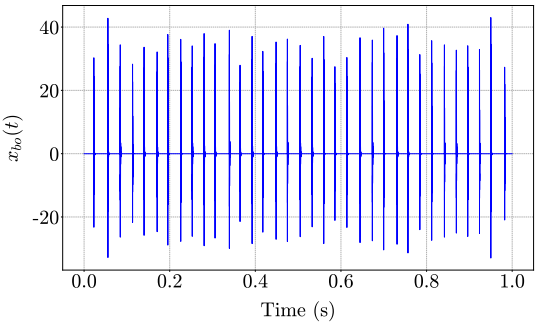


(a) Numerical gearbox signal.

(b) Gear mesh component.



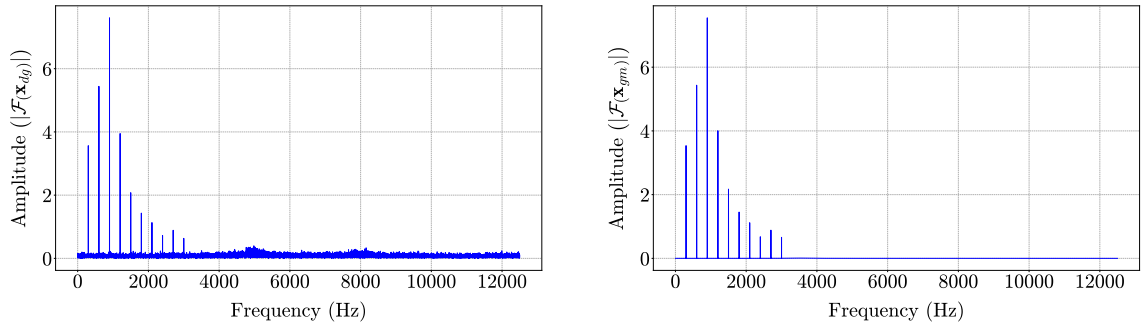
(c) Distributed pinion damage.



(d) Outer race-bearing damage.

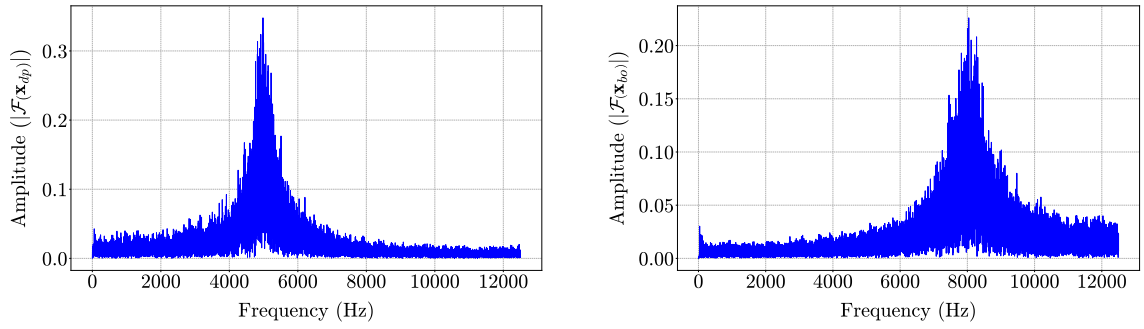
Figure B.1. The numerical gearbox signal $x_{ng}(t)$ and associated components for a bearing component where $SNR_{dB,bo} = -8dB$. Note that $x_n(t)$ is omitted as it is a white noise signal.

$SNR_{dB,dp} = -6dB$, and the $x_{bo}(t)$ is linearly increased in the range $SNR_{dB,bo} = [-50, 5]dB$ over a total of 400 numerical dataset signals. To create variation between the 400 numerical signals, the shaft speed is set to a constant value of $\omega(t) = 2 \cdot \pi \cdot \omega_n$, where $\omega_n \sim \mathcal{N}(10, 0.1^2)$, $n \in [1, 400]$. The numerical gearbox signal and its components are demonstrated in Figure B.1, and the spectral representations of the associated gearbox components are shown in Figure B.2.



(a) Numerical gearbox signal ($x_{ng}(t)$).

(b) Gear mesh component ($x_{gm}(t)$).



(c) Distributed pinion damage ($x_{dp}(t)$).

(d) Outer race-bearing damage ($x_{bo}(t)$).

Figure B.2. The spectral representation of the numerical gearbox signal $x_{ng}(t)$ and its associated components. Record 305 was used which has a SNR for the bearing component of $SNR_{dB,bo} = -8dB$. Note that the spectral representation of $x_n(t)$ is omitted as it is a white noise signal.

Appendix C

A spectral regularisation framework for latent variable models designed for single channel applications

---

Latent Variable Models (LVMs) are commonly used to capture the underlying dependencies, patterns, and hidden structure in observed data. Source duplication is a by-product of the data Hankelisation pre-processing step common to single channel LVM applications, which hinders practical LVM utilisation. In this supplementary document, a Python package titled `spectrally-regularised-LVMs` is presented. The proposed package addresses the source duplication issue via the addition of a novel spectral regularisation term. This package provides a framework for spectral regularisation in single channel LVM applications, thereby making it easier to investigate and utilise LVMs with spectral regularisation. This is achieved via the use of symbolic or explicit representations of potential LVM objective functions which are incorporated into a framework that uses spectral regularisation during the LVM parameter estimation process. The objective of this package is to provide a consistent linear LVM optimisation framework which incorporates spectral regularisation and caters to single channel time-series applications.

C.1 Introduction

Latent Variable Models (LVMs) represent a statistical methodology which aims to extract hidden sources of information, referred to as the latent variables $\mathbf{z} \sim p(\mathbf{z})$, $\mathbf{z} \in \mathbb{R}^d$, from an observed set of random variables $\mathbf{x} \sim p(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^D$. The overarching goal of the LVM framework, as demonstrated pictorially via Figure C.1, is to *(i)* construct a generative model to generate observed samples, and *(ii)* infer the latent variables, commonly referred to as the latent sources, from the observed data [161, 162]. In the linear LVM formulation, a linear encoding and decoding step is used to obtain samples from the posterior distribution $p(\mathbf{z}|\mathbf{x})$ and the generative distribution $p(\mathbf{x}|\mathbf{z})$.

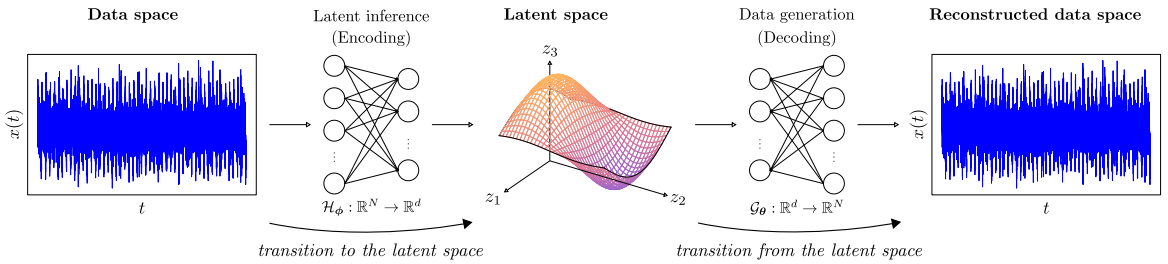


Figure C.1. A pictorial representation of the LVM framework applied to single channel time-series data. Note that the dimensionality of the latent space is presented in $\mathbb{R}^3$ and no data pre-processing steps are considered between the data space and the latent space. Data pre-processing steps are discussed in Section C.2.1.

In common LVM frameworks, e.g., Principal Component Analysis (PCA) [101, 114, 280, 344], a Gaussian assumption is made for the prior $p(\mathbf{z})$, the generative distribution $p(\mathbf{x}|\mathbf{z})$ and posterior distribution $p(\mathbf{z}|\mathbf{x})$, and the latent sources are driven to capture components that best describe the dominant sources in the observed data via minimum mean-square error compression or variance maximisation [101]. In this way, these methodologies are sample generation focused and the latent sources capture the dominant variance in the observed data. The PCA framework deals with second-order statistics and results in uncorrelated latent sources, i.e., the covariance for two sources z_j and z_k, represented by $\text{cov}\,[z_j, z_k]$, for $j \neq k$ is zero and non-zero otherwise. The Independent Component Analysis (ICA) formulation manifests as a LVM with latent sources that are assumed as non-Gaussian and mutually independent [323, 345–347]. In ICA frameworks, the assumptions of statistical independence and non-Gaussianity are more rigorous and enforce that the latent sources should be maximally informative in a non-Gaussian setting [178, 185]. The statistical independence assumption implies that the latent distribution factorises, i.e., $p(\mathbf{z}) = \prod_{j=1}^{d} p_j(z_j)$, where $\mathbf{z} \in \mathbb{R}^d$. In the assumed-Gaussian setting, it is possible to show that the PCA linear transformation can produce a statistically independent latent distribution [101].

In applications when a single channel time-series signal is of interest and available, certain processing steps, e.g., data Hankelisation, must be taken to fully enable the PCA framework [164]. In this setting, as noted in Hyvärinen [178], the linear LVM framework is strongly related to the blind-deconvolution problem [348], and the linear model formulation results in an underdetermined source situation, i.e., the sources extracted by the model may contain duplicate information [178]. To fully recover the latent sources, Davies and James [324] demonstrated that the original sources in the observed data must have disjoint spectral support.

Additionally, if no modifications are made to the standard LVM objective functions, clustering algorithms must be used to combine extracted sources that contain duplicate information [324]. Source duplication is an existing limitation that hinders the practicality of certain LVM objective functions in single channel time-series applications as the full rank solution, i.e., $d = D$, is required to avoid missing any interesting sources. Additionally, the model estimation step may be computationally extensive if $d << D$ as much compute is wasted on duplicate information. Hence, the objective is to bypass this limitation to improve the practicality of linear LVMs. In this work, a framework is proposed which offers ease of objective function implementation and a spectral regularisation term is proposed and implemented to enforce the disjoint spectral support requirement.

The proposed framework caters to spectral regularisation for linear LVM formulations and the `spectrally-regularised-LVMs` package makes spectral regularisation readily accessible for first order and second order optimisers. The package provides simplicity, ease of use, and efficiency to the single channel LVM framework, whereby the burden of repetitive pre-processing and optimisation framework implementation tasks are no longer placed on the user. The proposed framework operates in a deflation-based optimisation setting, i.e., each latent source is solved for one-by-one. Additionally, the methodology is objective function agnostic, which allows users to implement their own cost functions and use the proposed framework. Users can implement their objective functions directly using the symbolic Python package SymPy [349] to automatically obtain the gradient and Hessian of the model objective function, or using standard linear algebra operations enabled though the NumPy library [350]. Additionally, a finite-difference scheme is provided should users wish to obtain approximations of the gradient and Hessian of the model objective function. This provides generality to the `spectrally-regularised-LVMs` package and creates flexibility for future development and package utilisation.

C.1.1 Related work

In the literature, spectral regularisation typically refers to a regularisation technique for learning-based problems which utilises the spectra of an operator or matrix to control the properties of some model [351], e.g., Tikhonov regularisation [352]. For matrices, the spectra refers to its eigenvalue set [353]. Gerfon *et al.* [354] provide a succinct overview of spectral regularisation techniques for supervised learning tasks. Miyato *et al.* [355] use the spectral properties of deep generative models to improve model performance. Specifically, Ref. [355] use

the spectral norm of weight matrices to control the Lipschitz continuity of the discriminator network function for generative adversarial networks to improve model trainability. Xie *et al.* [356] use a spectral loss based off cosine similarity to encourage reconstruction consistency for anomaly detection using adversarial auto-encoders. In this work, spectral regularisation refers to the regularisation of the spectra of the LVM component vectors, where the spectra is given by the Fourier transform, to ensure that it does not capture the same signal information as other component vectors. In this way, spectral orthogonality within the LVM components is encouraged and the signal information captured by the LVM is driven to consist of unique components.

C.2 Model formulation

In this section the pre-processing steps for single channel time-series are detailed, the preliminary formulation of linear LVMs is presented, and the proposed spectral regularisation term is defined.

C.2.1 Time-series pre-processing

For single channel time-series data, a signal $x_i[n]$ is obtained, where $n \subset \mathbb{Z}^+ = \{1, \cdots, L_{signal}\}$ and L_{signal} is the length of the signal, that must be pre-processed to fit into the standard LVM framework. This can be achieved through data Hankelisation to obtain a matrix $\overline{\mathbf{X}}_i \in \mathbb{R}^{L_T \times L_w}$ as given by Equation (2.13). Note that the shift parameter used in Equation (2.13) is typically set to $L_{sft} = 1$ to ensure that the LVM component vectors act as a set of Finite Impulse Response (FIR) filters [324].

C.2.2 Linear latent variable models

The basic linear LVM generative model may be expressed for some observed data $\mathbf{x} \in \mathbb{R}^D$ as

$$\mathbf{x} = \sum_{j=1}^{d} \mathbf{a}_j z_j, \tag{C.1}$$

where $d \leq D$, z_j represents the j^{th} latent source and $\mathbf{a}_j \in \mathbb{R}^D$ represents the i^{th} decoding transition vector. The matrix-vector notation of the generative model is

$$\mathbf{x} = \mathbf{A}\mathbf{z}, \tag{C.2}$$

where $\mathbf{A} \in \mathbb{R}^{D \times d} = [\mathbf{a}_1, \cdots, \mathbf{a}_d]$. In the latent encoding step, a matrix $\mathbf{W} \in \mathbb{R}^{d \times D} = [\mathbf{w}_1, \cdots, \mathbf{w}_d]^T$ can be used to infer the latent vector $\mathbf{z}$ by

$$\mathbf{z} = \mathbf{W}\mathbf{x}, \tag{C.3}$$

where the rows in $\mathbf{W}$ represent the encoding transition vectors. In the LVM model, samples $\mathbf{x} \sim p(\mathbf{x})$ are observed from some unknown distribution $p(\mathbf{x})$, and the goal is to estimate $\mathbf{A}$ or $\mathbf{W}$. Common methodologies prefer to utilise characteristics of the latent sources to estimate the model parameters, e.g., maximum latent variance [114] or maximum non-Gaussianity through measures such as kurtosis or negentropy [185]. The objective function of interest is henceforward generally denoted by $\mathcal{L}_{model}$.

C.2.3 Spectral regularisation

In this work, the use of a spectral regularisation term is proposed for linear LVMs objective functions to combat source duplication and use a general optimisation framework to estimate the model parameters. The assumed form of the general LVM objective function may be written as

$$\begin{aligned} \min_{\mathbf{w}_j} \quad & \mathcal{L}_{model}(\mathbf{w}_j) + \mathcal{L}_{sr}(\mathbf{w}_j) \\ \text{s.t.} \quad & \mathbf{w}_j^T \mathbf{w}_j = 1, \end{aligned} \tag{C.4}$$

where $\mathcal{L}_{model}(\mathbf{w}_j)$ represents the objective function to be minimised, $\mathcal{L}_{sr}(\mathbf{w}_j)$ represents the spectral regulariser that acts as an explicit regularisation term on $\mathcal{L}_{model}(\mathbf{w}_j)$, and the optimisation constraint $\mathbf{w}_j^T \mathbf{w}_j = 1$ is used to ensure that the parameter estimation process focuses on the direction of $\mathbf{w}$ and not its magnitude. The spectral regularisation term encourages the $\mathbf{w}_j$ solution to be spectrally unique with respect to the previous solution vectors $\mathbf{w}_l$, where $l < j$. The general objective function can be reformulated into a Lagrangian expression

$$\mathcal{L}(\mathbf{w}_j, \lambda_{eq}) = \mathcal{L}_{model}(\mathbf{w}_j) + \mathcal{L}_{sr}(\mathbf{w}_j) + \lambda_{eq} \left(\mathbf{w}_j^T \mathbf{w}_j - 1 \right), \tag{C.5}$$

where λ_{eq} represents the Lagrange multiplier. The regularisation term proposed in this work enforces spectral orthogonality and uses the squared modulus of the Fourier domain representation of a vector $\mathbf{w}_j$, denoted as $\mathbf{b}(\mathbf{w}_j)$, to minimise the dot product between $\mathbf{b}(\mathbf{w}_j)$ and the squared Fourier representation for the previously solved $\mathbf{w}_l$ vectors, $l = 1, \cdots, j-1$. This is given as

$$\mathcal{L}_{sr}(\mathbf{w}_j) = \alpha \sum_{l=1}^{j-1} \mathbf{b}(\mathbf{w}_j)^T \mathbf{b}(\mathbf{w}_l), \quad j > 1 \tag{C.6}$$

where α represents a penalty enforcement parameter that controls the importance of the regularisation term. Note that the spectral regularisation term is only applied after the estimation of the first projection vector $\mathbf{w}_1$. To remove ambiguity behind the choice of α, the Sequential Unconstrained Minimisation Technique (SUMT) is used within the model

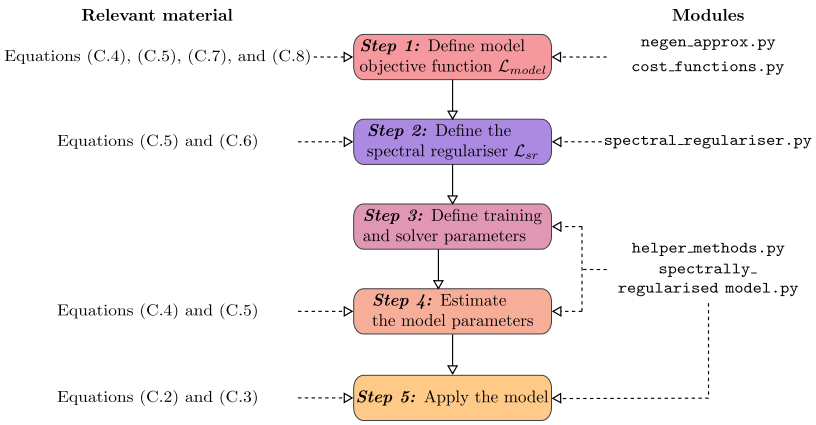


Figure C.2. A summary of the important mathematical aspects that provide context to the software implementation and architecture, the associated package sub-modules, and any relevant material that is provided in this document.

optimisation step to iteratively increase the prominence of the regularisation term during each optimisation iteration [357]. The important terms related to the implemented optimisation algorithm to incorporate the regularisation term in Equation C.6 into the overall LVM objective are detailed in the package documentation. The full derivation is not included here for brevity. The use of the spectral regularisation term in Equation (C.6) allows the basic LVM model to be effectively applied to single channel time-series data. There is no reliance on carefully selected pre-processing strategies, e.g., the spectrogram, and source duplication is discouraged during the parameter estimation process.

C.3 Software description

In this section, a basic description of the software used in the `spectrally-regularised-LVMs` Python package is provided. Figure C.2 provides context to how the relevant mathematical components of the package can be sequentially organised.

C.3.1 Introduction to spectrally-regularised-LVMs

The `spectrally-regularised-LVMs` package provides a framework for linear LVMs with the proposed spectral regularisation term using the Python Application Program Interface (API). In this framework, the objective function for the LVM can either be implemented symbolically or explicitly as a Python function that uses NumPy objects. Alternatively, the objective function for two common LVMs, PCA and ICA, can be called through the sub-modules present in the package. The choice of a symbolic objective function implementation ensures that users need not spend effort on deriving the gradient and Hessian of their objective function,

as this can be done symbolically. For the explicit implementation, users can either provide the necessary first and second order information, or use a finite difference approximation scheme to estimate the gradient and Hessian. As detailed in Figure C.2, there are five steps that are required to formalise the proposed LVMs framework in single channel applications. These five steps are combined within the package to create a framework from which users can utilise LVMs with spectral regularisation under a general model objective function $\mathcal{L}_{model}$. This makes it easier to implement LVMs with spectral regularisation and to apply the LVMs on single channel time-series applications. Given some time-series signal $x[n], n \in \mathbb{Z}^+$, the signal pre-processing, spectral regularisation, and model estimation steps are handled by the framework, thereby automating the involved aspects of the LVMs task.

C.3.2 Software architecture

The `spectrally-regularised-LVMs` module is composed of five sub-modules which were designed to facilitate the steps detailed in Figure C.2. To facilitate step one from Figure C.2, the `cost_functions` sub-module was implemented. This provides four cost function classes to the user of which two are dedicated to two simple LVM formulations, one which allows users to explicitly encode their cost function and its associated derivatives, and one which allows users to define their cost function symbolically using the SymPy package [349] and internally determines the associated gradient and Hessian of the objective function. Additionally, a finite difference approximation scheme can be used to approximate the gradient and Hessian. The `negen_approx` sub-module supports step one by providing estimator classes which provide access to a set of approximation functions for negentropy [330]. The `spectral_regulariser` sub-module supports step two from Figure C.2 by implementing a spectral regulariser class that can be used to determine the regularisation term given in Equation (C.6) alongside its gradient and Hessian. The `helper_methods` sub-module enables step three from Figure C.2, by providing a set of classes which assist with data pre-processing, data batch sampling, quasi-Newton update strategies [357], Gram-Schmidt orthogonalisation [358], and time-series signal Hankelisation [164]. The `spectrally_regularised_model` sub-module captures steps three and four from Figure C.2 and provides the linear LVM model which allows users to define their training parameters and estimate the model parameters using a `.fit`$(\cdot)$ method call to an instance of the `LinearModel` class [359]. Finally, to apply the LVM as detailed in step five from Figure C.2, methods from the `LinearModel` class may be used to facilitate the transition to and from the latent space, as detailed in Figure C.1. A `.transform`$(\cdot)$ method call computes

the transformation from the data space to the latent space, and an `.inverse_transform(·)` method call performs the transformation from the latent space back to the data space. These method calls are available to instances of the `LinearModel` class. This provides a complete framework that automatically performs the necessary steps to estimate the parameters of linear LVMs with the proposed spectral regularisation term and apply the LVM encoding and decoding steps.

C.3.3 Software functionalities

The primary functionality of the `spectrally-regularised-LVMs` package is to provide a complete framework that estimates the linear LVM model parameters that are regularised via the spectral orthogonality term. This is achieved through a cost function initialisation stage and a parameter estimation stage. Each stage is discussed in turn and code listings are given. The cost function initialisation is generalised by offering a symbolic or explicit representation to cater to a wide variety of objective functions. For this example, the PCA objective function is used

$$\begin{aligned}\mathcal{L}_{model}(\mathbf{w}_j) &= -\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{\left(\mathbf{w}_j^T\mathbf{x}\right)^2\} \\ &= -\frac{1}{N_m}\sum_{l=1}^{N_m} z_l^2,\end{aligned} \tag{C.7}$$

where it is assumed that $\mathbf{x}$ is zero-mean and $z_l = \mathbf{w}_j^T\mathbf{x}_l$. In a symbolic representation, the indices of z_l are represented for N_m samples from $p(\mathbf{x})$. the PCA objective function can be implemented following the example in Figure C.3(a). In an explicit representation, the PCA objective function can be implemented using functions that use NumPy objects as detailed in Figure C.3(b).

The parameter estimation stage provides alternative functionalities and advanced options to the user regarding the underlying pre-processing and model estimation steps performed. The main component is to create an instance of the `LinearModel` class and using the `.fit(·)` method to estimate the model parameters. Further information related to the arguments for the parameter estimation stage is given in the documentation. A code listing example of this stage is given in Figure C.4.

C.3.4 Software dependencies

The `spectrally-regularised-LVMs` package is free software distributed under the MIT license, and it is open to contributions on GitHub. The package is hosted on PyPi [360] to enable ease of installation, broader accessibility to the scientific community, and to avoid

```python
import sympy as sp
import spectrally_regularised_LVMs as srLVMs

# Symbolic cost function implementation
cost_inst = srLVMs.SymbolicCost()

z, j, n = cost_inst.get_symbolic_parameters()

loss = -1/n * sp.Sum((z[j])**2, (j))

cost_inst.set_cost(loss)
```

(a) Symbolic objective function.

```python
import numpy as np
import spectrally_regularised_LVMs as srLVMs

# Explicit cost function implementation
cost_inst = srLVMs.ExplicitCost()

obj = lambda X, w, z: -1 * np.mean((X @ w)**2, axis = 0)
grad = lambda X, w, z: -2 * np.mean((X @ w) * X, axis = 0,
                                    keepdims=True).T
hess = lambda X, w, z: -2 / X.shape[0] * (X.T @ X)

cost_inst.set_cost(obj)
cost_inst.set_gradient(grad)
cost_inst.set_hessian(hess)
```

(b) Explicit objective function.

Figure C.3. The code listings detailing how the PCA objective function given in Equation (C.7) can be implemented symbolically or explicitly using functions that use NumPy objects [350].

```python
import numpy as np
import spectrally_regularised_LVMs as srLVMs

# Define a toy signal
x_signal = np.random.randn(10000)

# Define the cost function instance
cost_inst = ...

# Define the model
model_inst = srLVMs.LinearModel(n_sources = 5,
                                cost_instance = cost_inst,
                                Lw = 256,
                                Lsft = 1)

# Estimate the model parameters
model_inst.fit(x_signal)

# Obtain the latent representation Z
Z = model_inst.transform(x_signal)

# Obtain the recovered representation of X
X_recon = model_inst.inverse_transform(Z)
```

Figure C.4. The code listing detailing how to perform parameter estimation in the spectrally regularised framework.

dependency issues when using the package in a user's local Python environment. The package is dependent on the following Python packages: NumPy [350], Matplotlib [361], scikit-learn [362], tqdm [363] and SymPy [349].

C.4 Illustrative example

To demonstrate the capabilities and detail the main functionalities of the `spectrally-regularised-LVMs` package, an illustrative example is used. In this ex-

ample, steps one through five from Figure C.2 are applied to the negentropy-based LVM objective function [178]. This objective function is given as

$$\mathcal{L}_{model}(\mathbf{w}_j) = -\left(\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}\{G(\mathbf{w}_j^T\mathbf{x})\} - \mathbb{E}_{\nu\sim p(\nu)}\{G(\nu)\}\right)^2, \tag{C.8}$$

where $G(\cdot)$ is a non-quadratic function and ν represents a Gaussian variable that is zero mean and unit variance [330]. This objective function is chosen as non-Gaussianity is a well-known measure for highlighting interesting information in data from rotating machinery [30].

A time-series signal from the Intelligent Maintenance Systems (IMS) bearing dataset is used in this example [307]. The intention is to provide evidence of the applicability of the spectrally regularised framework on real-world data. Record 700 from the second IMS dataset is used from the first bearing channel is used as it is expected to contain a vibration signature with a faulty component [249]. The 'Examples' folder located in the Github repository contains a notebook which details this step-by-step use of the proposed package for this example and contains all code for result reproducibility. In Figure C.5 the spectra of the record 700 is shown. The sources from the estimated linear LVM were analysed and their respective spectra with and without spectral regularisation are presented in Figure C.6. Evidently, without spectral regularisation there is clear source duplication and this issue is addressed with the addition of the proposed spectral regularisation term.

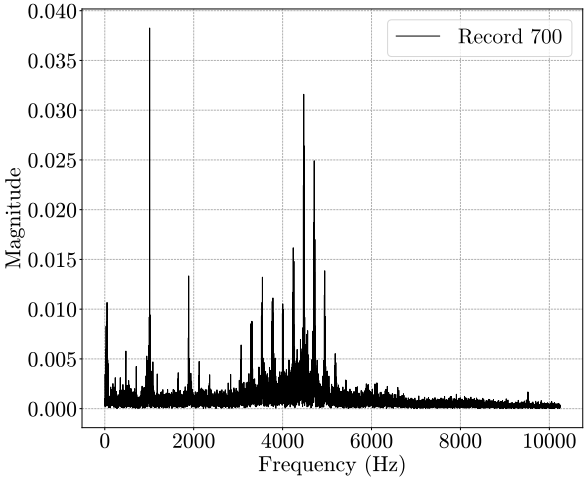


Figure C.5. The signal spectrum of record 700 from the second IMS dataset.

C.5 Impact

The `spectrally-regularised-LVMs` package provides a framework for linear LVMs which utilises spectral regularisation for single channel time-series data applications. This was done

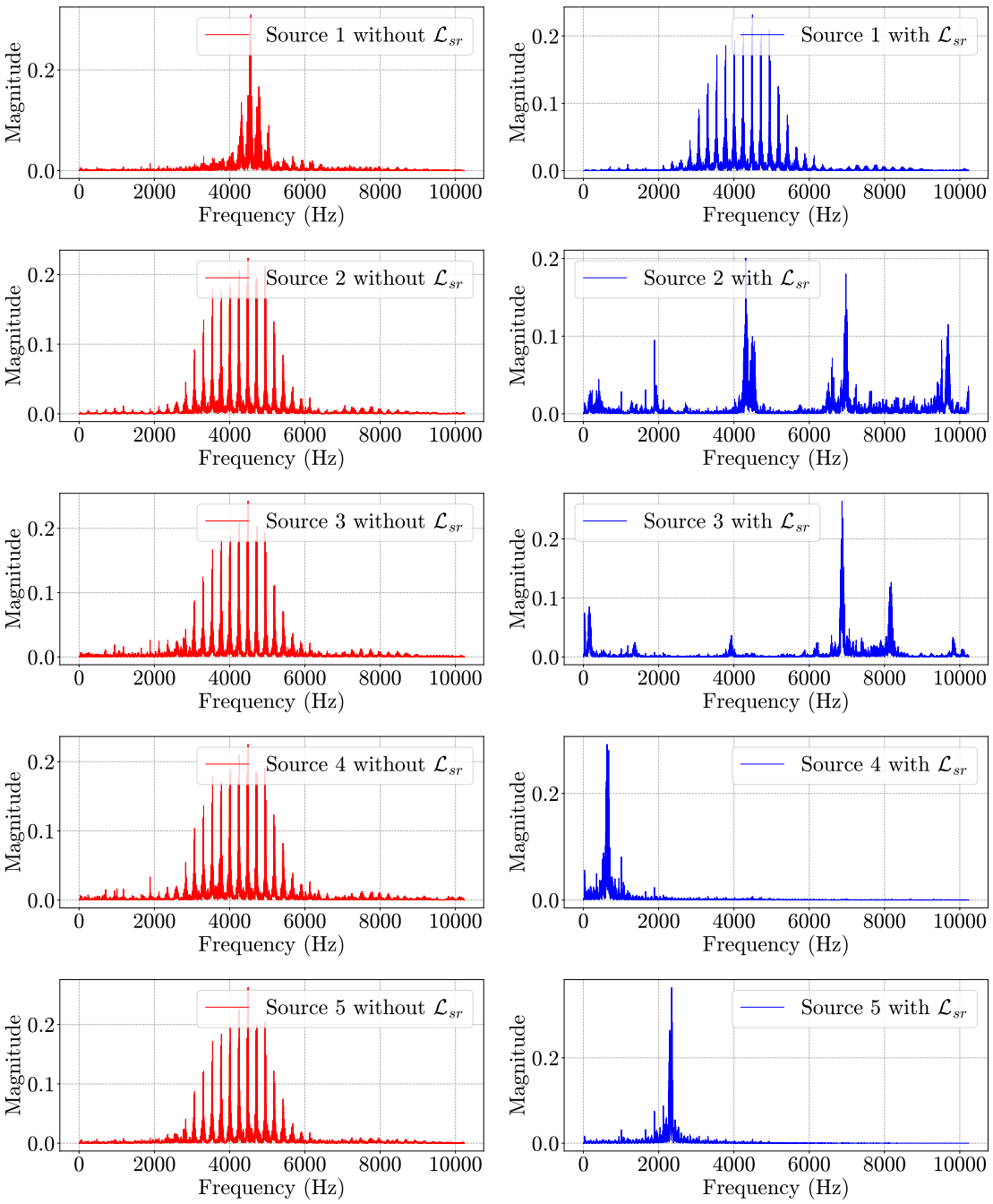


Figure C.6. The spectra of the latent source signals of record 700 from the second IMS dataset without and with the addition of the proposed spectral regularisation term.

to streamline the application process by readily adapting the parameter estimation framework to depend on a generic objective function class instance, where the user can use different model objective functions depending on their required application. Furthermore, explicit control over the model parameter estimation process is provided to ensure that different pre-processing

and parameter estimation approaches can be used. This allows for efficient iteration for a given application through different optimisation strategies, e.g., steepest gradient descent, stochastic gradient descent, or constrained Newton/quasi-Newton methods [357].

The Python package aims to provide researchers with a framework that not only provides a general parameter estimation methodology for linear LVMs, but also removes latent redundancy via the proposed spectral regularisation term. The proposed regularisation term is objective function agnostic in its formulation, and providing access to a Python package for LVM parameter estimation under this term is expected to be beneficial to the broader community. This will allow others to utilise the proposed methodology for various linear LVM objective formulations without having to re-implement all aspects of linear LVMs from scratch.

Finally, while many popular linear LVM applications are available in Python, e.g., PCA and ICA in scikit-learn [362], these implementations cater to a general class of LVM applications, and are not focused on the application of the methodologies to single channel time-series data. The `spectrally-regularised-LVMs` package provides direct access to the linear LVM framework and caters to the referenced methodology application, which is expected to offer fruitful avenues for future research.

C.6 Conclusion

A modular linear LVM framework is proposed and implemented in Python. This framework offers customizable objective functions, facilitates the estimation of the model parameters, and uses a novel spectral regularisation term to combat source duplication. By making the `spectrally-regularised-LVM` framework accessible to the broader community it makes the application and development of new methods easier and I firmly believe it will be beneficial for future research into LVMs for single channel time-series data. Finally, it is envisaged that future research pursuits into single channel time-series data applications will be stimulated and benefit from the `spectrally-regularised-LVMs` package.