On robust statistical outlier analysis for damage identification

Ioannis Matthaiou

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Engineering

Department of Mechanical Engineering

Submission Date

August 2022

Abstract

This thesis aims to contribute towards the development of reliable and accurate damage detection monitoring frameworks, applicable for a range of structural health and condition monitoring problems. Central to this purpose, is to be able to *detect damage patterns* embedded in a system's *vibration signal* responses sufficiently early. This will enable a condition-based maintenance and inspection to be carried out so as to prevent potentially catastrophic events, as related to each application domain.

Firstly, to obviate reliance on data labels, an inclusive outlier analysis study is conducted by means of *robust multivariate statistical* analysis and a range of other (more common) outlier detection techniques, in both multivariate and time-series settings. Given the parametric nature of robust multivariate statistical techniques, it has also been possible to characterise *outliers* according to their influence on a method's estimates.

Secondly, novelty detection is explored, in which a set of samples representing the nominal state of the system, is assumed to be available. This set includes observations from a system with its dynamics being significantly influenced by environmental and operational variability.

Finally, this thesis explored the potential of utilising certain robust techniques as a *pre-processing step* on damage sensitive features (contaminated with outliers) for novelty detection tasks.

Given the large volume of observations, both experimental and computational, different damage sensitive features were extracted, some of which were specific to the range of problems / types of damage being investigated. The performance, in terms of both *sensitivity* in damage detection and *immunity* to environmental and operational variability, was assessed for each damage sensitive feature, in combination to the outlier and novelty detection technique used.

This thesis has introduced to the condition and structural health monitoring fields a range of methods from robust statistics with attractive properties, such as the effective unmasking of outliers.

Acknowledgements

There are many people that made this work possible:

Firstly, I would like to thank my supervisor Professor Nikos Dervilis for his excellent advise and continuous support throughout my PhD.
Many thanks must also go to my PhD co-supervisor Professor Keith Worden for his very helpful comments and insights.
I must also thank Dr Ifigeneia Antoniadou for being my supervisor in the first years of my PhD, which dealt with condition monitoring for engines.
Thanks also to Dr Bhupendra Khandelwal and his research team, with whom I have collaborated with in the first year of my PhD in order to perform experiments and collect data from gas turbine engines.
This PhD would not have been possible if it wasn't for the love of my wife Angelina and son Emmanuel.
Many thanks to my friends and the various kinds of interesting people that I met over the years, and for making this journey more fun.
Finally, I gratefully acknowledge the generous scholarship and financial support received by the Department of Mechanical Engineering at the University of Sheffield.

Contents

1 An introduction to SHM and CM **1**

1.1 Maintenance strategies . . . 2

1.1.1 Run-to-break . . . 2

1.1.2 Preventative maintenance . . . 2

1.1.3 Condition-based maintenance . . . 3

1.2 The "undamaged" and "damaged" system states . . . 3

1.3 Definitions and some motivations for SHM & CM . . . 4

1.3.1 Vibration-based SHM & CM . . . 5

1.3.1.1 Advantages . . . 5

1.3.1.2 International standards and guidelines . . . 6

1.3.1.3 The case for a more "intelligent" SHM & CM strategy . . . 7

1.3.2 The blind source separation problem in SHM & CM . . . 7

1.3.3 Bridges and civil engineering infrastructure . . . 9

1.3.4 Wind turbines, gearboxes and bearings . . . 10

1.3.5 Gas turbine engines with varying operating conditions . . . 12

1.4 SHM & CM via statistical pattern recognition . . . 14

1.4.1 Environmental and operational variability . . . 16

1.5 Damage identification requirements . . . 17

1.6 Aim of the thesis . . . 18

1.7 Organisation of the thesis . . . 19

1.8 Summary . . . 20

2 Vibration signals, their transformations and estimation of common damage-sensitive features **22**

2.1 Stochastic processes and signals . . . 22

2.2 Signal processing for stationary processes . . . 26

2.2.1 Statistical measures . . . 26

2.2.2 Fourier Transform . . . 27

2.2.3 Time series analysis 28
2.3 Signal processing for non-stationary processes 29
2.3.1 Classical time-frequency analysis methods 30
2.3.1.1 Short-time Fourier Transform and Spectral Kurtosis 30
2.3.1.2 Wavelet Transform 31
2.3.2 Adaptive time-frequency analysis methods 33
2.3.2.1 The Hilbert Transform 33
2.3.2.2 Non-linear systems and the requirement for mono-component signals 34
2.3.2.3 EMD-based & iterative decomposition methods 35
2.3.2.4 Iterative filtering and VMD-based decomposition methods 39
2.4 Common DSFs from vibration responses 41
2.4.1 Modal parameters in SHM 41
2.4.2 Modal parameters identification via adaptive signal decomposition . 42
2.4.3 CM in REBs 45
2.4.4 DSFs for REBs using envelope analysis 48
2.4.5 Vibration DSFs in GT engines 51
2.5 Summary 51

3 Datasets description and exploration 53
3.1 SHM on multi-storey structures and bridges 53
3.1.1 The 4 DOF numerical simulation model 53
3.1.1.1 NSim-4DOFLin 54
3.1.1.2 NSim-4DOFNonLin 56
3.1.2 LA-4DOF 58
3.1.3 Z24 61
3.2 CM on rotating machinery & gas turbines 63
3.2.1 WT-REB 63
3.2.2 Gas turbine engines 65
3.2.2.1 TS-SAF 65
3.2.3 TF-SAF/-LBO 68
3.3 Summary 71

4 Damage-sensitive features for damage identification **74**
4.1 The importance of DSFs in DI 74
4.1.1 DSFs used in the literature 75
4.2 Methods for deriving DSFs 75
4.2.1 Features normalisation 75
4.2.2 Visualisation, compression & selection of DSFs 78
4.2.2.1 PCA & its kernel-based equivalent 79
4.2.2.2 Dimensionality reduction with outliers - robust PCA 80
4.2.3 The minimum entropy deconvolution for impact-type of damage . . 82
4.3 The framework of deriving vibration-based DSFs 83
4.4 Modal-based DSFs for LA-4DOF & NSim-4DOFLin/-4DOFNonLin datasets 86
4.5 EA-based DSFs for WT-REB dataset 89
4.5.1 DSFs from the EA procedure on a simulated example 89
4.5.2 DSFs from the EA procedure on the WT-REB dataset 90
4.6 Impact-based DSFs for LA-4DOF, NSim-4DOFLin, NSim-4DOFNonLin & WT-REB datasets 91
4.7 Global-based DSFs for TS-SAF, TF-LBO/-SAF datasets 94
4.8 Summary 99

5 Inclusive outlier analysis **101**
5.1 An introduction to outlier analysis 101
5.2 Outliers influence and robust covariance 104
5.3 Overview of robust statistics and its applications in DI 105
5.4 Methods 107
5.4.1 Extreme value analysis via robust statistics 109
5.4.1.1 Minimum Covariance Determinant 109
5.4.1.2 Deterministic MCD 111
5.4.1.3 Kernel Minimum Regularised Covariance Determinant . . . 111
5.4.1.4 Orthogonalised Gnanadesikan-Kettenring estimator 113
5.4.1.5 Deterministic MM-estimator 114
5.4.2 Exploratory analysis and outlier diagnosis 115
5.4.2.1 rPCA and its outlier map 115
5.4.2.2 Robust clustering via data trimming 116
5.4.3 Robust time series analysis 118
5.5 Evaluation of OD method performance & benchmarking 120
5.6 Results & discussion 121

5.6.1 Important notes regarding the results 122
5.6.2 TF-LBO dataset 123
5.6.3 NSim-4DOFLin & -4DOFNonLin datasets 131
5.6.4 LA-4DOF dataset 138
5.6.5 WT-REB dataset 142
5.6.6 Z24 dataset 145
5.6.7 TS-SAF & TF-SAF datasets 147
5.7 Summary 150

6 Novelty detection and robust pre-processing 152
6.1 Overview of novelty detection and its applications for DI 152
6.2 Robust pre-processing 154
6.3 Novelty detection algorithms 155
6.3.1 Gaussian Mixture Model 155
6.3.2 Isolation Forest 156
6.3.3 One-class Support Vector Machines 156
6.4 Results & discussion 157
6.4.1 Z24 dataset 157
6.4.2 LA-4DOF dataset 159
6.4.3 NSim-4DOFNonLin dataset 161
6.5 Summary 162

7 Conclusions 166
7.1 Vibration-based features for SHM and CM 166
7.2 Robust statistics for damage identification 168
7.3 Diagnosis of outliers in SHM and CM 169
7.4 Enhancing novelty detection by outlier exclusion 169
7.5 Future work 169

A Numerical model for SHM applications 172
A.1 A four degree-of-freedom model 172
A.2 The Duffing model 173
A.3 The ith modal vibration response from a multi-DOF linear system 174

B Methods & Algorithms **176**
B.1 One-class support vector machines 176
B.2 Isolation forests 178
B.3 Time-varying filtering based EMD 179
B.4 M-estimators 180
B.5 Self-adaptive noise cancellation 181
B.6 Minimum Regularised Covariance Determinant 181
B.7 Kernels and the kernel "trick" 182
B.8 C-steps for robust estimators 183

C Robust estimators properties & the framework of novelty detection **184**
C.1 Properties of robust estimators and robustness criteria 184
C.2 An overview of novelty detection 185

D Additional results **193**

E A simulated scenario for CM applications & demonstration of the envelope analysis procedure **196**
E.1 Outer race damage using a simulated example 196
E.2 Envelope analysis on the simulated example 198
E.3 Outer race damage on the WT-REB at damage level 5 a simulated example 200

Bibliography **202**

List of Figures

1.1 Photos of two sophisticated engineering examples demanding the highest possible reliability. (a): General Electric GE9x: World's most powerful commercial aircraft jet engine delivering $134,300$ *lbs* of thrust. Photo from [6]. (b): Akashi Kaikyo: World's largest suspension bridge, designed to resist typhoons of 180 *mph* and earthquakes up to 8.5 on the Richter scale. Photo from [15]. 2

1.2 Damage identification hierarchy and description of each level. 5

1.3 A group of hidden excitation forces are combined together to generate a set of observed mixture vibration responses. 8

1.4 Downtime per failure for each sub-system of five major onshore and offshore wind turbine sub-systems. One third of total downtime is attributed to *gearbox* failures. Figure from [53]. 11

1.5 Rolling element ball bearing diagram, showing the four sub-components. Figure from [10]. 11

2.1 Classification of vibration signals: as stationary and non-stationary. 23

2.2 Simplified diagram of the modulated acceleration signal of a faulty bearing (showing the carrier and impulse signals separately). 47

2.3 Amplitude-modulations in a theoretical signal with carrier frequency of $500Hz$ and modulating frequency of $10Hz$. 47

2.4 The four main steps of the envelope analysis procedure. 48

3.1 Air temperature measurements (see details of Z24 dataset). 54

3.2 A numerical simulation of the 4 DOF model. Damage is progressing exponentially over 8 days (from $t = 1008hr$). 55

3.3 Non-linear spring force profile used for numerical simulations. Damage is progressing exponentially over 8 days. 57

3.4 A numerical simulation of the 4 DOF model with non-linear spring forces in undamaged and damaged states having the same simulation parameters as in previous simulation. 58
3.5 Power spectral densities of acceleration (left) and force input (right) in $10s$ intervals for baseline ($s1$) system state. 61
3.6 Power spectral densities (left) and kernel density estimate (right) for five different system states. 61
3.7 Natural frequencies as a function of time and ambient air temperature. Top plot: red colour denotes the time period in which temperatures dropped below $0°C$. Bottom subplots: show the bilinear relationship that exists between temperature variation and natural frequency. 63
3.8 Diagram depicting the main elements of the experimental laboratory gearbox test-rig. 64
3.9 Concatenated acceleration time-history plot on the vertical acceleration records on Bearing no. 1: Accelerations on different levels of damage, up to damage level 5 ($50-60s$), and the undamaged ($0-10s$) health state (top plot); Undamaged state on four generator loading conditions, starting from no load ($0-10s$) up to the highest load ($30-40s$) (bottom plot). 65
3.10 Spectrum of power density of vertical acceleration records on Bearing no. 1, showing only up to 3 kHz, for damage identification purposes. Left plot: Undamaged and two damage levels, including the highest one, i.e. Damaged L5, with no load (C0); Right plot: Damage level 5 and two loading conditions, C0 and C3 (0.3 kW). 66
3.11 Turbo-shaft gas turbine engine schematic depicting main salient features of the experimental test-rig. 67
3.12 Power spectral density estimates from vibration measurements (two accelerometers) on the turbo-shaft gas turbine engine on two different fuels and AFR. 68
3.13 Power spectral density estimates from vibration measurements (two accelerometers) on the turbo-fan gas turbine engine combustor on 35 different fuels. . 69
3.14 Four statistical moments calculated on the vibration measurements as recorded on five accelerometers. 70
3.15 Kernel density estimates calculated on vibration measurements from Accelerometer 1 on two AFR. 70

4.1 Three feature normalisations applied on the first two NFs of the Z24 dataset. 78
4.2 Block diagram of the inverse filtering process of the MED technique. 82

4.3 Schematic outlining the main components of the framework used to derived vibration-based DSFs. 85
4.4 NFs estimates from MPI, for the LA-4DOF dataset. Regions G2 to G3: undamaged state + EOV changes, regions G3 to G4: damaged state (and its progression levels) and G4 to the end is the system in its damaged only state. Segment time interval is equal to 8 seconds. 87
4.5 NFs estimates from MPI, for the NSim-4DOFLin dataset. Vertical line shows the first sample from the damaged state. Segment time interval is 7.5 hours. 88
4.6 LA-4DOF dataset: 191 amplitudes from the PSD and corresponding robust z-scores. NSim-4DOFLin/-4DOFNonLin datasets: 230 amplitudes from the PSD and their robust z-scores. LA-4DOF is equal to 2 seconds and NSim-4DOFLin/-4DOFNonLin is equal to 2.5 hours. 89
4.7 Envelope spectra, shown between 0 and 2.5 kHz on the four signals considered. The enhancement of the impulses at BPFO is evident when the signal is both pre-processed with an AR model and MED, i.e. bottom-right signal. The BPFO cursors are marked as ':' in the plots. 90
4.8 Envelope spectrum of damage level 2 for C0 and C3. 91
4.9 Kurtosis μ_4 and MED filtering on kurtosis $g(\mu_4)$ on four datasets. Vertical red line signifies the point of damage initiation, while for the WT-REB dataset each red line denotes a damage level progression. Segment lengths for: WT-REB is equal to 2.5 seconds. 92
4.10 Impact-based (kurtosis) DSFs on each IMF decomposed signal of three different adaptive decomposition methods on LA-4DOF dataset. 94
4.11 Impact-based (kurtosis) DSFs on each IMF decomposed signal of two different adaptive decomposition methods on WT-REB dataset. Vertical red lines denote the damage level progression. 95
4.12 PSD estimates (zoomed in up to 10 kHz) for TF-LBO dataset. 97
4.13 Multidimensional feature set MD_2 on TF-LBO (segment time interval is 10 seconds), TF-SAF (segment time interval is 2 seconds) and TS-SAF (segment time interval is 2 seconds) datasets. 98

5.1 The first two NFs of the Z24 dataset, showing the effects of outliers on the sample covariance matrix and potential ways to improve DI using the MCD estimator. 104
5.2 kMRCD diagram of the algorithm's main steps. 113

5.3 TF-LBO: squared Mahalanobis distances (on log-scale) from five robust estimators using $\alpha = 0.55$. The first sample from the damaged state is shown by the vertical line, while the horizontal lines are the two thresholds. For comparison the Mahalanobis distances (on log-scale) using the sample mean and covariance are shown (d^2_{msd}). 125
5.4 Same plots as Figure 5.3 but with $\alpha = 0.8$. 125
5.5 Same plots as Figure 5.3 but with $\alpha = 0.67$. 126
5.6 TF-LBO: a two-dimensional problem (selected two features from the MD_2 DSF matrix) to illustrate the workings of OGK, FMCD and DetMM robust estimators using ellipses on the χ^2 CDF and sensitivity to α value. 127
5.7 TF-LBO: same as in Figure 5.6 but kMRCD workings are illustrated using the contours of its computed SMDs. 129
5.8 TF-LBO: outlier map of rPCA (on log-scale) and the MD_2 DSFs as a function of sample no., rescaled in range $[0,1]$ to aid in visualisation. Thresholds on the SD and ODist, γ_{sd} and γ_{odist}, respectively, were calculated as already explained. 130
5.9 TF-LBO: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.67$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Time series samples are the eight different features from the MD_2 DSFs matrix. 130
5.10 TF-LBO: outlier scores calculated using ocSVM with $\alpha = 0.67$, $\nu = 0.5$ and $\gamma_{rbf} = 1.2$. The first sample from the damaged state is shown by the vertical line, while the horizontal lines are the two thresholds. 131
5.11 NSim-4DOFNonLin: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.75$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Each time series is one of the 10 principal components of rPCA, obtained from PSD amplitudes. 132
5.12 NSim-4DOFNonLin: score plots of the first two principal components of PCA (middle plot) and kPCA (left plot). On the right are the 10 principal components of PCA, computed from the PSD amplitudes. 133
5.13 NSim-4DOFNonLin: two rPCA component scores (Features 3 and 4 in Figure 5.11) showing (on the top) the SMD contours of the kMRCD method and (on the bottom right) the χ^2 ellipse computed from FMCD method. The two bottom left plots show the corresponding SMDs, where the vertical line is the first sample from the damage state. 134

5.14 NSim-4DOFNonLin: top two plots of rPCA components (immune to EOV influence) overlapped with χ^2 ellipses computed using robust and non-robust estimates with varying confidence levels (α values). Bottom plots are the SMDs computed on the same features using DetMM and cMahal estimators. 135
5.15 NSim-4DOFNonLin: the left plot is the rPCA outlier map on a log scale (10 rPCA components). On the top right, the score plot of rPCA's computed on its first two components. On the bottom right, the DSFs, i.e. the PSD amplitudes, unit-normalised. 136
5.16 NSim-4DOFNonLin: ROC curves (on $\log-\log$ scales) for the computed SMDs from five robust estimators and cMahal using two different DSFs: the first 10 principal components from rPCA (right) and PCA (left). 137
5.17 NSim-4DOFNonLin: SMDs computed using the robust and non-robust estimators. Three vertical lines are different thresholds set on the α confidence interval: MC threshold (γ_α^{mc}), on log-normal ($\gamma_{kMRCD,\alpha}$) and on the χ^2 distribution with p DOF ($\chi^2_{p,\alpha}$). 137
5.18 LA-4DOF: DD plots (on log-scales) of the SMDs computed from the five robust estimators on two α values. The detected outliers $\hat{\mathbf{x}}_o$ and inliers $\hat{\mathbf{x}}_i$, are shown as the points above and below the threshold lines, respectively. The diagonal line shift (with respect to the 45 degrees angle) shows deviation from the Gaussian distribution. 139
5.19 LA-4DOF: ROC curves (on $\log-\log$ scales) for the computed SMDs from five robust estimators on two different α values and cMahal. DetMM is showing considerable performance improvement when $\alpha < 0.5$. Impact-based DSFs (kurtosis) are used. 139
5.20 LA-4DOF: SMDs computed from five robust estimators and sample covariance and mean values. Four different thresholds. Three on α confidence interval: MC threshold (γ_α^{mc}), on log-normal ($\gamma_{kMRCD,\alpha}$) and on the χ^2 distribution with p DOF ($\chi^2_{p,\alpha}$). One on the χ^2 distribution with 99% confidence interval: $\chi^2_{p,99\%}$. 140
5.21 LA-4DOF: on the top right: the rPCA outlier map (on a log scale), on the bottom right: the impact-based DSFs (kurtosis) and on the left: the rPCA score plot - first two principal components. 141
5.22 LA-4DOF: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.51$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Each time series is one of the 10 principal components of rPCA, obtained from PSD amplitudes. 141

5.23 LA-4DOF: using as DSFs the first two rPCA components. On the top plot, the samples are z-score normalised, while at the bottom the same samples are robust z-score normalised. Ellipses on the χ^2 distribution from DetMM and the sample covariance and mean are overlapped. 142
5.24 WT-REB: squared Mahalanobis distances (on log-scale) from five robust estimators using $\alpha = 0.51$, apart from DetMM whose $\alpha = 0.2$. The first sample from each level of damage is shown by the vertical line, while the horizontal lines are the thresholds. For comparison the Mahalanobis distances (on log-scale) using the sample mean and covariance are shown (d^2_{msd}). 143
5.25 WT-REB: on the left: the rPCA score plot - first two principal components. On the top right: the rPCA outlier map (on a log scale), on the bottom right: the impact-based DSFs (kurtosis). 144
5.26 WT-REB: on top left: the χ^2 ellipse computed from FMCD method and the SMD contours computed using kMRCD, both with $\alpha = 0.51$. Bottom plots are the SMDs computed on the same two features using kMRCD and FMCD estimators, with the threshold set on the χ^2 distribution with confidence interval on α. On the right, the cluster assignments ($K = 2$) computed from tclust algorithm on the full 3D feature space of impact-based DSFs. The 2D plots are the first two impact-based DSFs. 145
5.27 Z24: on the left: the rPCA score plot - first two principal components. On the top right: the rPCA outlier map (on a log scale), on the bottom right: the four NFs. 146
5.28 Z24: iForest outlier scores on Z24 dataset. Vertical line is the first sample of damage. Horizontal line is the threshold. 146
5.29 Z24: On the right are the clustering results using tclust algorithm ($K = 2$), while on the left are the labelled samples shown on the second and third NFs. 147
5.30 TS-SAF: On the right are the clustering results using tclust algorithm ($K = 1$), while on the left are the labelled samples shown on the first two features from the MD_2 DSFs matrix. 148
5.31 TS-SAF: On the left are the SMDs computed using kMRCD, FMCD and cMahal methods on the first two rPCA components. Vertical lines are the thresholds. On the right are the corresponding χ^2-based ellipses and SMDs contours (for kMRCD) on the same rPCA components. 148

5.32 TF-SAF: on the left are the first two rPCA components of the MD_2 DSFs matrix. On the top right are the SMDs contours computed by the KMRCD method with $\alpha = 0.75$. And on the bottom right are the clustering analysis results computed from the tclust algorithm with $K = 1$ and $\alpha = 0.75$. 149
5.33 A_{ROC} metric as calculated on four different datasets for: five robust estimators of scatter and location, the classical Mahalanobis distance, ocSVM and iForest. 150

6.1 Z24: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively. Vertical line is the first sample of the damage state. 157
6.2 Z24: A_{ROC} metric for the case where robust pre-processing is applied with NLTS and tclust with $\alpha = 0.85$, and compared to the case where no robust pre-processing is applied $\alpha = 1$. 158
6.3 Z24: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively (first and third NFs). Vertical line is the first sample of the damage state. Outliers removed from the training set $\hat{\mathbf{x}}_{o,trn}$ using tclust are also shown. 159
6.4 Z24: The corresponding 2D results shown in Figure 6.3 analytically. 159
6.5 LA-4DOF: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively (84 PSD amplitudes)). Vertical line is the first sample of the damage state. Outliers removed from the training set $\hat{\mathbf{x}}_{o,trn}$ using tclust are also shown. 161
6.6 LA-4DOF: on the top are the outlier scores computed by ocSVM and iForest algorithms and on the bottom the corresponding histograms on $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$ samples, showing also the thresholds applied. Red dots $\hat{\mathbf{x}}_o$ are the estimated outliers for each method. 162
6.7 NSim-4DOFNonLin: A_{ROC} metric for the case where robust pre-processing is applied with tclust on $\alpha = 0.75$ and $\alpha = 0.9$, and compared to the case where no robust pre-processing is applied $\alpha = 1$. 163
6.8 NSim-4DOFNonLin: Illustration of the outliers removed using two α values: 0.75 and 0.9. 164
6.9 NSim-4DOFNonLin: on the top two plots are the outlier scores computed using ocSVM and iForest, while in the middle are the corresponding histograms. On the bottom are the 230 PSD amplitudes used as DSFs. 165

A.1 The four DOF numerical model used for generating the simulated data under different damage scenarios and EOV influence. 173

C.1 Machine learning general framework diagram. 187

D.1 LA-4DOF: Damping ratio estimates from system identification (Chapter 4). 193
D.2 NSim-4DOFLin: Damping ratio estimates from system identification (Chapter 4). 193
D.3 NSim-4DOFNonLin: The first 10 PCA components of the PSD amplitudes (Chapter 5). 194
D.4 iForest outlier scores on NSim-4DOFNonLin and LA-4DOF datasets (Chapter 5). 194
D.5 NSim-4DOFNonLin: Clustering analysis using tclust (Chapter 5). 194
D.6 NSim-4DOFNonLin: SMDs computed using robust estimators on the first 10 rPCA components of PSD amplitudes as DSFs (Chapter 5). 195
D.7 Z24: outlier scores and corresponding histograms - no robust pre-processing has been applied (Chapter 6). 195

E.1 Simulated damage signal - single impulse amplitude modulated. 196
E.2 Overall acceleration from three simulation scenarios shown in time- and frequency-domains (zoomed in from 1.9 to 2 seconds and 0 to 5 kHz): Undamaged rolling element bearing with low SNR (left plots), Onset of damage with low SNR (middle plots) and Late damage with high SNR (right plots). 197
E.3 Power spectra of pre-processed acceleration signal: AR residual (left), MED applied on the raw signal (middle) and MED applied on the AR residual (right). It particularly shows the reduction of discrete low frequency components not associated with BPFO when MED is applied and the enhancement of impulses at BPFO when the AR residual is considered instead of the raw signal. 199
E.4 Kurtogram up to level $k = 10$ calculated on four signals: the raw acceleration signal (top-left), the AR residual (top-right), the signal pre-processed by MED only (bottom-left) and the AR residual signal pre-processed by MED (bottom-right). The maximum SK calculated on the four signals shows us same frequency bandwidths between $\ddot{x}$ and $\varepsilon_{\ddot{x}}$, but, different ones between the other two signals. 200
E.5 Envelope spectrum of damage level 5. 201

List of Abbreviations

A_{ROC} Area under the receiver operating characteristic curve

ALIF Adaptive local iterative filtering

APU Auxiliary power unit

AR Autoregressive

BIC Bayesian Information Criterion

BPFO Bearing passing frequency of the outer race

BSS Blind source separation

CBM Condition-based maintenance

CM Condition monitoring

cMahal Classical Mahalanobis squared distance

DetMCD Deterministic minimum covariance determinant

DetMM Deterministic MM

DFT Discrete-time Fourier transform

DI Damage identification

DOF Degree of freedom

DR Damping ratio

DSF Damage sensitive feature

EA Envelope analysis

EM Expectation Maximisation

EMD Empirical mode decomposition

EMD-SSSC Empirical mode decomposition with soft sifting stopping criterion

EOV Environmental and operational variability

FFT Fast Fourier transform

FMCD Fast minimum covariance determinant

FN False negative

FP False positive

GMM Gaussian mixture model

GP Gaussian process

GT Gas turbine

HT Hilbert transform

ICA Independent component analysis

ICEEMDAN Improved complete ensemble empirical mode decomposition with adaptive noise

iForest Isolation Forest

iid Independent and identically distributed (random variable)

IMF Intrinsic mode function

IRF Impulse response function

KDE Kernel density estimate

kMRCD Kernel Minimum Regularised Covariance Determinant

kPCA Kernel principal component analysis

LA-4DOF Los Alamos four degree of freedom

LBO Lean blow out

LSTM Long short term memory

LTI Linear time invariant

MA Moving average

MAC Modal assurance criterion

MAD Median absolute deviation

MC Monte Carlo

MCD Minimum covariance determinant

MED Minimum entropy deconvolution

MLE Maximum likelihood estimate

MPI Modal parameter identification

MS Mode shapes

ND Novelty detection

NDE Non-destructive evaluation

NExT Natural excitation technigue

NF Natural frequency

NLTS Nonlinear least trimmed squares

NSim-4DOFLin Four degree of freedom linear numerical simulation model

NSim-4DOFNonLin Four degree of freedom nonlinear numerical simulation model

ocSVM One-class Support Vector Machine

OD Outlier detection

OGK Orthogonalised Gnanadesikan-Kettenring

PCA Principal component analysis

PDF Probability density function

PP Projection pursuit

REB Rolling element bearing

RMS Root mean square

ROC Receiver Operating Characteristic

rPCA Robust PCA

SAF Sustainable alternative fuels

SDO Stahel-Donoho outlyingness

SHM Structural health monitoring

SK Spectral kurtosis

SMD Squared Mahalanobis distance

SPR Statistical pattern recognition

SSI Stochastic subspace identification

SV Support vector

SVM Support vector machine

SVMD Successive variational mode decomposition

TF-LBO Turbo-fan lean blow out

TF-SAF Turbo-fan sustainable alternative fuel

TS-SAF Turbo-shaft sustainable alternative fuel

TVF-EMD Time varying filter empirical mode decomposition

VMD Variational mode decomposition

WT-REB Wind turbine rolling element bearing

YJ Yeo-Johnson (transform)

Nomenclature

α Inliers proportion

$\alpha(t)$ Unless otherwise specified, time varying ratio between the damaged and undamaged stiffness values

α^* Highest possible breakdown point for robust estimators ($= 0.5$)

β Air temperature profile switching variable

$\boldsymbol{\mu}$ Unless otherwise specified, is a multivariate measure of center

$\boldsymbol{\mu}_K$ A robust estimation of location/centre of estimator K, e.g. K=MCD

$\boldsymbol{\theta}_{NLTS}$ The fitted NLTS parameters vector to the time series

$\chi^2_{99\%}$ Multivariate ellipse defined at the 99th percentile of the χ^2 distribution

$\chi^2_{p,\alpha}$ The χ^2 threshold calculated based on the α level of significance and p degrees of freedom

$\gamma(t)$ Rate of damage progression for the numerical simulation model

γ^{mc}_{α} The Monte Carlo threshold based on the α level of significance

$\gamma_{kMRCD,\alpha}$ The threshold calculated for the kMRCD method based on the α level of significance

γ_{mc} Monte Carlo threshold for outlier detection

γ_{rbf} Radial basis function kernel scale

$\hat{\mathbf{x}}_{o,trn}$ Declared outliers in the training set

$\hat{H}$ Tree height of the isolation forest

$\hat{y}$ Fitted or estimated NLTS time series

λ_k The kth eigenvalue

$\mathbf{a}_k$ The kth eigenvector of kPCA

$\mathbf{f}_s$ Spring force vector

$\mathbf{p}_n$ Impact-based damage sensitive features

$\mathbf{q}_n$ Global-based damage sensitive features

$\mathbf{s}_n$, $\boldsymbol{\omega}_n$, $\boldsymbol{\zeta}_n$ Modal-based damage sensitive features

$\mathbf{u}_k$ The kth eigenvector

$\mathcal{A}$ Outliers underlying distribution

$\mathcal{E}_{\Delta}$ Differential entropy

$\mathcal{G}$ Inliers underlying distribution

$\mathcal{H}$ Hilbert transform

μ Unless otherwise specified, is a univariate measure of center

μ_{mcd} Univariate robust MCD mean

ω_d Damped natural frequency

ω_n Natural frequency

$\phi(\cdot)$ Feature map

$\phi(t)$ Unless otherwise specified, time-varying phase

π_k The $k^{th} mixtureweight$

ψ Wavelet basis function

ρ A robust estimator loss function

Σ Unless otherwise specified, is a multivariate measure of scale

σ Unless otherwise specified, is a univariate measure of scale

Σ_K A robust estimation of scatter of estimator K, e.g. K=MCD

σ_{mcd} Univariate robust MCD scale

τ Time lag variable

$\tilde{\mathcal{K}}$ Gram matrix

$\tilde{Y}$ Linear PCA projections

ε A Gaussian white noise term

$\mathbf{x}$ Unless otherwise specified, a p-dimensional vector of system responses or acceleration measurements of length N. Also specified as X. Additionally, $\mathbf{x}_n$ refers to the nth observation

$\mathbf{x}_*$ Candidate outlier observation vector of $\mathbb{R}^p$

$\mathbf{x}_o(\gamma_{uv})$ Estimated outliers using the univariate threshold

$\mathbf{x}_{trn}$ Training set matrix

$\mathbf{x}_{tst}$ Testing set matrix

ζ Damping ratio

$\{M, C, K\}$ Mass, damping and spring matrices

$\{m, c, k\}$ mass, damping and spring coefficients

A_c Carrier signal amplitude

A_m Modulated signal amplitude

c_k Unless otherwise specified, the extracted k^{th} IMF

c_{ER} Eigenvalue ratio restriction of the tclust algorithm

C_{xx} Autocovariance function

d_K^2 A robust estimation Mahalanobis distance of estimator K, e.g. K=MCD

d_{msd}^2 Squared Mahalanobis distance

f_s Sampling rate

f_c Carrier signal frequency

f_m Modulated signal frequency

f_{shaft} Rotating shaft signal frequency

H An $M \times M$ unknown system mixing matrix

h Number of samples that are inliers. Also referred to as the h-subset

j Imaginary variable ($= \sqrt{-1}$)

K Unless otherwise specified, is the total number of clusters

$k(\mathbf{x}_n, \mathbf{x}_m)$ Kernel function between two vectors $\mathbf{x}_n$ and $\mathbf{x}_m$

L Time segment length of the original signal

O_k Kurtosis objective function of the MED

$ODist_i$ The ith orthogonal distance of the $\tilde{\mathbf{y}}_i$ projected observation of the $\mathbf{x}_i$ multidimensional data point

p Unless otherwise specified, number of features / variables

P_x Power spectral density

p_{max} Maximum AR order

$p_{optimal}$ Optimal AR order

s Unless otherwise specified, M exogenous and endogenous sources of system excitation

S_x Discrete-time Fourier transform

SD Score distance of rPCA

t Time index

$T(t)$ Air temperature time history

U Loadings matrix from PCA

Y_{rPCA} Matrix of all rPCA projections

Chapter 1

An introduction to SHM and CM

Some of the most well-known examples of advanced engineering systems that are currently operational include the world's largest suspension bridge, the Akashi Kaikyo in Japan, and the most powerful turbofan jet engine, to-date, the GE9X (Figure 1.1). What these two engineering achievements have in common is their immense physical scales, pushing the boundaries of what is physically possible, without compromising on performance.

One way towards innovation is to make structures lighter and more cost-efficient. For instance, by simply reducing the safety factor during the design phase. However, this will lead to an increase in the uncertainty around a system's operational safety, e.g. the maximum permissible load on a given structure. In those cases, the safety factor is therefore used to account for our lack of knowledge or understanding of many things, including the environment the system is expected to be operating in.

Reliability is an integral part of innovation, and sophisticated safety-critical systems, like aircraft engines, need to employ a means of *fully automated identification* of *early* deviations from their intended functions. These identification measures need to be applicable throughout the whole system's service life, which includes many factors that may influence the successful implementation for such an automated procedure.

For these and many other reasons, the scientific research fields of Structural Health and Condition Monitoring, *SHM* & *CM*, respectively, were developed many decades ago. However, many of their challenges still remain largely unsolved, as of today. A major challenge is the effect of environmental and operational variability (EOV) on the damage identification (DI) performance of any SHM & CM implementation.

This chapter includes a brief description of the fields of SHM and CM, including how EOV may influence the ability to perform DI accurately and reliably. Firstly though, an overview of the three broad classes of maintenance strategies is provided in the next section.

(a)

(b)

Figure 1.1. Photos of two sophisticated engineering examples demanding the highest possible reliability. (a): General Electric GE9x: World's most powerful commercial aircraft jet engine delivering $134,300$ *lbs* of thrust. Photo from [6]. (b): Akashi Kaikyo: World's largest suspension bridge, designed to resist typhoons of 180 *mph* and earthquakes up to 8.5 on the Richter scale. Photo from [15].

1.1 Maintenance strategies

There are three main maintenance strategies in which to evaluate and ensure the safety of a system*. Starting from the simplest and least ideal, i.e. *run-to-break*, to the most advanced and suitable for our purposes, i.e. *condition-based maintenance*.

1.1.1 Run-to-break

As the name suggests, this approach has been used in the past, in systems which were run until a component broke down. The failed component was then replaced or repaired and then the system was back online again. At best, this will lead to the system being shut down for a while. This may be suitable for small machines, in which monitoring may seem a waste of resources, e.g. a printer device. However, a run-to-break maintenance strategy for a jet engine is clearly undesirable.

1.1.2 Preventative maintenance

Maintenance is performed immediately before the expected time of failure is reached, for a given component. Therefore, this provides a safer maintenance solution for safety-critical systems, for instance, in a jet engine, as long as the expected time of failure can be *accurately* predicted. Because any maintenance is done regardless of condition, the procedure inevitably introduces waste as many components could still have been used for longer. This is also costly as it requires frequent shut-downs for the maintenance to be carried out, e.g.

*Unless otherwise specified, the word *system* or *systems* in this thesis, will be used as general reference to any type of structural or mechanical system.

in a power plant this is not ideal. In components, such as rolling element bearings, where many factors influence their condition, the uncertainty in time-to-failure is high enough for this maintenance strategy to be successful.

1.1.3 Condition-based maintenance

Any maintenance is carried out according to the condition of a system and its constituent subsystems. It normally has dedicated or embedded sensing, capable of evaluating its health on a continuous basis. The success of this strategy, and hence the suitability for early DI, is heavily dependent on the correct implementation, i.e. starting from the hardware (sensors, etc.) to the software level, of the condition-based maintenance (CBM) strategy. If implemented correctly, the potential economic benefits are substantial, as quantified in a landmark case study of Swedish paper mill in [19].

Because of the ability to evaluate the health of a system and its working conditions, a better maintenance plan can be placed, reducing downtime to a minimum, increase safety and operational efficiency. Many of CBM methodologies have also the potential to be performed online and at high sample rates, and this is a major advantage is systems where, for instance, the time-to-failure interval is relatively small.

1.2 The "undamaged" and "damaged" system states

In this thesis, the definition of "damage" and "failure" is the one used in C. Farrar's and K. Worden's book [67]: "*Damage will be defined as intentional or unintentional changes to the material and/or geometric properties of these systems, including changes to the boundary conditions and system connectivity, which adversely affect the current or future performance of these systems.*" While, "*failure occurs when the damage progresses to a point where the system can no longer perform its intended function*". Failure is, therefore, to be prevented, as it leads to undesirable outcomes, e.g. sudden building collapse, machinery breakdown, etc. This can be achieved by a successful implementation of a CBM strategy.

All man-made systems contain imperfections, and therefore, as time goes by it is only inevitable that they will develop major defects that impact their performance negatively, e.g. surface flaking in bearing races. However, prior to reaching such performance degradation, a system's overall dynamic state may be considered as the one that was originally intended to be used for, i.e. its "undamaged" system state.

Also, when discussing about *engine performance monitoring* or engine CM, it is important to note that the above definitions of "damaged" and "undamaged" system states do not apply. This is because there will not be such actual damage to the engine. Instead, "damage" and "undamaged" system states will be used to refer to a system in its unintended and

intended states, respectively, *with respect to* its characteristics *under typical engine operation*. For instance, the air-to-fuel ratio that a given engine was designed to operate with. Similarly, DI in the engine performance monitoring context, will be referring to the identification of such *atypical engine operating conditions*, e.g. *relatively* large vibration amplitudes.

1.3 Definitions and some motivations for SHM & CM

When talking about SHM & CM, it is necessary to state their main differences and similarities. Although, both make use of the CBM strategy and share most of the methodologies for analysis and sensing technologies, SHM refers to the "*process of implementing a damage identification strategy for aerospace, civil or mechanical infrastructure*" [67]. On the other hand, CM is concerned with the detection of damage in *rotating or reciprocating machinery* [33], having moving components that may be interacting with each other. Both SHM & CM are *non-destructive in-situ* procedures that share similar philosophies for DI, including sensing, data transmission and analysis methods. CM is a mature technology and has played a key role in ensuring the safety of safety-critical equipment, e.g. power plants, for a few decades now. Both SHM & CM have four levels of hierarchy to perform DI [157], as shown in Figure 1.2.
There are a number of potential benefits for implementing SHM or CM in a system. Some of the most important are:

- Identification of damage prior to failure: This is probably the most significant advantage, since it will prevent potential loss of life, expensive repairs and other costs (e.g. related to downtime). In addition, further maintenance costs can be reduced due to eliminating routine inspections;

- Gaining insight into system dynamics: The interaction between full-scale in-situ systems and their environments, e.g. wind speed fluctuations, offer invaluable knowledge of system dynamics, which cannot be gained in a laboratory setting;

- Objective assessment and quantification of system states: It is important to objectively assess a system's state, in a quantifiable way, with reference to its previous state(s). This is in contrast to the subjective nature of Non-Destructive Evaluation (NDE).

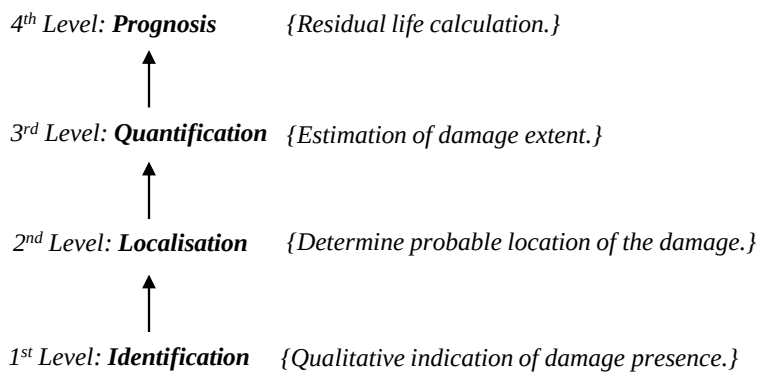

Figure 1.2. Damage identification hierarchy and description of each level.

1.3.1 Vibration-based SHM & CM

1.3.1.1 Advantages

One way to perform SHM & CM is through vibration measurements. Other ways include, acoustic emissions and strain-based monitoring (using fibre optic sensors) [67]. Vibration-based monitoring has been used for DI since the 1980s, and it is currently, one of the most commonly used diagnostic[†] measures to carry out SHM & CM [67], [33]. Throughout the thesis, *vibration* will refer to the absolute amplitude of acceleration responses. These responses are measured using an acceleration transducer or accelerometer [33], which is attached at a specific location on the system, e.g. the housing of a rolling element bearing. There are several benefits in using vibration measurements in SHM & CM, including:

- Identification of overall system state: Damage is known to influence the system's dynamic properties, i.e. its stiffness and energy dissipation [67]. Any measured vibration responses (due to internal or external force excitations), will directly reflect changes in the dynamic properties of the system;

- Making use of advanced signal processing: Acceleration transducers sample data in the rate of several or even tens of kHz. At these levels of sampling rates there is an abundance of data points for utilising advanced signal processing analysis, e.g. adaptive signal decomposition. Also, unlike other kind of signals, e.g. lubricant analysis, vibration measurements with high sampling rates, enable online monitoring to be implemented;

[†]This term refers to the identification of not only the existence but also the type of damage, i.e. localisation (see Figure 1.2).

- Identification of the type of damage (localisation): Any kind of machinery under rotating motion, will have well-known operating conditions. Vibration measurements, will provide a direct measure of the rotational frequency of each component and its harmonics, which can be used to determine the location of damage;

- Capturing characteristic reference patterns: Any kind of machinery or structure in its undamaged state will vibrate with a characteristic pattern. Hence, this provides a reference state in which all future decisions, regarding DI, are to be based upon.

1.3.1.2 International standards and guidelines

For vibration-based CM there are a number of international standards, e.g. ISO-7919 ("*Evaluation of machine vibration by measurements on rotating shafts*") and ISO-10816 series ("*Evaluation of machine vibration by measurements on non-rotating parts*") [103], with the latter being probably the most widely adopted in industry. These standards provide certain vibration acceptance levels (also referred to as measurement quantiles) on different frequency bands as applied to the type of machinery, measurement locations, etc. In general, these standards were developed based on extensive experience and collaboration from specialists all over the world, and therefore, are expected to offer the best possible diagnostic information for early DI.

On the other hand, there is a limited number of international standards in SHM, especially for highly-localised DI, e.g. cracks, where the number of sensors is relatively small as compared to the surface area being monitored [40]. One of those standards, as described in [40], provides guidelines for sensor selection and placement, as well as, ways for DI on the whole-structural level by means of dynamic system properties. Such properties include mode shapes (MS), natural frequencies (NFs) and the modal assurance criterion (MAC). In the European Union, the Structural Assessment, Monitoring and Control network, developed several guidelines for SHM spanning several application domains including bridges, buildings and railway infrastructure [55].

These standards, however, should only be used as general guidelines for the design of a more intelligent diagnostic system, able to infer damage more reliably and accurately. This is particularly true in situations where EOV influences the dynamic properties significantly and unpredictably, for instance, as in the case of offshore wind turbines. As discussed in [33], using widely-applicable and general features from vibration signals, such as, the root-mean-square (RMS) of the amplitude (specified in ISO-10816-6), is unlikely to provide early signs of damage.

1.3.1.3 The case for a more "intelligent" SHM & CM strategy

This leads to a discussion of some of the most important considerations for implementing a vibration based SHM & CM. Those are as follows:

- Damage is a *local phenomenon*, and as demonstrated in many studies using data from experiments, e.g. in [138], higher-frequency vibration modes (e.g. about $30-40\ Hz$) are typically more sensitive to damage presence. For example, as discussed in [40], a crack covering 1% of a beam's cross-sectional area, results in a resonance frequency change (for the *first mode* of vibration) of just 0.1%.
  Therefore, and unless damage has progressed significantly, the lower-frequency vibration modal properties (e.g. NFs) will not be impacted significantly. This implies that early DI will not be feasible using the lower-frequency modal properties, given that the transducer is far from the damage location. This is to be expected, for instance, from a large structure like a bridge, which may span more than a few kilometres in length [14], while damage can be on the scale of a few centimetres (initially);

- Any significant EOV influence on the system, e.g. freezing temperatures on an aircraft wing during flight conditions, will alter its dynamic characteristics or modal properties (damping and stiffness) [169]. Thus, its vibration responses are expected to be different under EOV influence. Also, EOV influence can be similar in magnitude or many times greater than that of damage, and this has been reported numerous times in the literature [169], [20].

Therefore, the implemented vibration-based SHM & CM must be both insensitive to EOV influence, while at the same time identify any type of damage, even highly-localised ones, sufficiently early for action to be taken. These two objectives demand a more sophisticated approach to SHM & CM.

1.3.2 The blind source separation problem in SHM & CM

Both in SHM & CM, the DI problem may be conducted using only system responses. That is, the only observations are system responses, measured through vibration signals. In general, system responses are expected to be excited from *three different sources*:

1. Exogenous forces acting on the system, e.g. aerodynamic forces or crosswinds on a bridge deck;

2. Endogenous forces due to damage, e.g. impact forces as opening and closing cracks;

3. Other endogenous forces generated by a number of elements interacting with each other in the system (applicable mostly in CM applications), e.g. elastic forces due to contact stiffnesses of gear pairs.

These three different excitation forces will each yield a distinct system vibration response. When all three forces are present at the same time, the system response will contain a *mixture* of all them. Therefore, the main challenge in DI is to determine whether or not any force generated, as seen on the system's responses, is *due to damage presence or due to changes from any other sources of excitation.*

As mentioned previously, in CM for rotating machinery, apart from the exogenous forces acting on the system and the endogenous ones due to damage, there will also be a number of forces generated by the interactions from the different elements or components, e.g. gears. Due to the fact that it is more desirable to examine the health state of specific components [25], this poses an additional challenge for detecting damage in CM applications.

For a linear time invariant (LTI) system, e.g. a gearbox or a structure, its impulse response function (IRF) $h_{i,j}$, will completely characterise the transfer path between its ith response x_i and the jth excitation force or source s_j, with $i = j = 1, ..., M$ independent components. Thus, in discrete-form, the ith vibration response can be written as the sum of M *convoluted mixtures*, as discussed in [22]:

$$x_i = \sum_j h_{i,j} * s_j + \varepsilon \tag{1.1}$$

More compactly Equation 1.1 may be written as,

$$\mathbf{x} = H * \mathbf{s} + \boldsymbol{\varepsilon} \tag{1.2}$$

where, $H \in \mathbb{R}^{M \times M}$ is an unknown mixing matrix and $\mathbf{s} = [s_1, ..., s_M]^T$ is the sources vector, while $\boldsymbol{\varepsilon} \in \mathbb{R}^M$ is a vector of observation white noise terms. In Figure 1.3, a simplified diagram of a multi-input multi-output (MIMO) system is shown. In order to identify damage

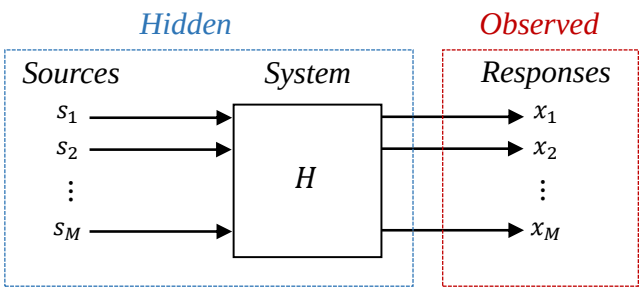

Figure 1.3. A group of hidden excitation forces are combined together to generate a set of observed mixture vibration responses.

and distinguish it from the rest of the sources, which implies to find the independent sources $\mathbf{s}$, H is required. This can be done with a deconvolution procedure. However, since H is unknown, the problem can be cast formally within the framework of *blind source separation* (BSS) [33]. BSS is a methodology for recovering the different *independent sources* that excite the system, only from its responses.

In the frequency domain the problem in Equation 1.2 turns into multiplication. This can then be solved by assuming M independent sources, e.g. using independent component analysis (ICA), on each of the frequency bins within the Nyquist frequency range [146]. However, there is a particular issue: the indeterminacy of BSS [198], which must be considered whenever such an approach is followed.

Therefore, given $\mathbf{x} \in \mathbb{R}^N$, i.e. a vector containing time-series acceleration data from an experiment, the DI problem can be solved by obtaining certain characteristics from it, i.e. a particular damage sensitive feature (DSF) set. This will enable any DI algorithm to determine whether or not the system is in its *damaged* or *undamaged* health states, without actually recovering the sources of excitation or their contributions [22]. This is true whether DI is to be performed in CM or SHM applications. This is also the essence of *feature extraction* in SHM and CM, as it will be covered extensively later on.

It must be noted that damage may not necessarily introduce another source of excitation in the system, but, may only modify the IRF between any type of excitation force and the corresponding responses. For instance, in case of a crack in a beam, it will only result in a localised stiffness reduction. In these cases, one will need to determine whether any changes in $\mathbf{x}$ are the result of changes in $\mathbf{s}$ and/or the transfer path H.

1.3.3 Bridges and civil engineering infrastructure

Bridges and other civil engineering infrastructure, e.g. multi-storey buildings, experience structural deterioration as time goes by, e.g. progressive corrosion of steel reinforcing bars in concrete structures. These may lead to sudden collapse of a structure, since the load carrying capacity will be significantly reduced. One of the most recent and unfortunate examples of structural deterioration and subsequent sudden failure in bridges is the "Morandi" bridge in Genoa, Italy [127]. After 51 years of service life, this concrete bridge partially collapsed in 2018, causing loss of life. The collapse is attributed to poor maintenance practises, lack of redundancy, but also, the inability of certain locations to be manually inspected and evaluate its structural condition. Due to pitting corrosion of steel tendons, the bridge collapsed [133]. This is a perfect example in which all three collapse causes (i.e. lack of redundancy, poor maintenance and inability to inspect) would have been prevented if SHM was in place. From the year 2000, there have been at least 115 bridge collapses. This is a

sign of the infrastructure ageing and a major indication that implementing SHM procedures is *immediately* needed to prevent further disaster.

1.3.4 Wind turbines, gearboxes and bearings

In the last decade or so, the wind power generation industry has seen an exponential growth. The installed capacity of offshore wind energy, in particular, has grown from approximately 4 GW in 2011 to more than 18 GW in 2017 [53]. Offshore wind turbines, as the name suggests, are installed in the sea, thus, operating in harsh environments. Their maintenance, is therefore, many times higher than that of their onshore counterparts. The operation and maintenance cost of offshore wind contributes to about one third of the total energy production cost [161]. Although it has a great potential for contributing to a clean power generation mix, mainly due to the high wind speed resource at sea, this type of cost is significantly high for it to scale-up easily in capacity.

Both onshore and offshore wind turbines have mechanical moving parts, one of which is the *gearbox*. Located between the blades by an *input* shaft and the electric generator by an *output* shaft, its purpose is to convert the rotational speed from low to high (and the shaft torque from high to low), i.e. from $5-15$ to $1000-1800$ rotations per minute (RPM) [2]. A comprehensive survey published in 2019 [53], collected data from 18 different wind turbine reliability studies, to investigate the failure rates[‡] and downtimes[§] of each wind turbine sub-assembly, including shafts, bearings, gearbox, generator, control system and tower. The gearbox has been found to have a relatively low failure rate (constituted of just 5% in onshore and 8% in offshore wind of the total failure rates) as compared to the most critical sub-assembly, which is the electrical system in onshore (27.5%) and the pitch system in offshore (25%). Nevertheless, the reported *downtime* of an offshore or an onshore wind turbine due to gearbox failure is one of the highest ones, as compared to the rest of the main sub-assemblies (see Figure 1.4).

One of the most frequently failing components in a gearbox is the *rolling element bearing* (REB). A study conducted by the National Renewable Energy Laboratory on 750 wind turbine gearbox damage records between 2009 and 2015 [12], has found that more than 75% of the total gearbox failures occur due to axial cracks on REBs that form within the high- and intermediate-speed stages. For that reason, the usual practise in wind turbine gearbox monitoring, is to mount acceleration transducers, as near as possible to those stages [69].

REBs can be seen in almost every type of rotating machinery in existence. Their failure

[‡]Calculated as the percentage of the total number of failure occurrences per turbine per unit time.

[§]Calculated as the percentage of the total time interval that the wind turbine does not produce enough power due to a certain sub-system failure.

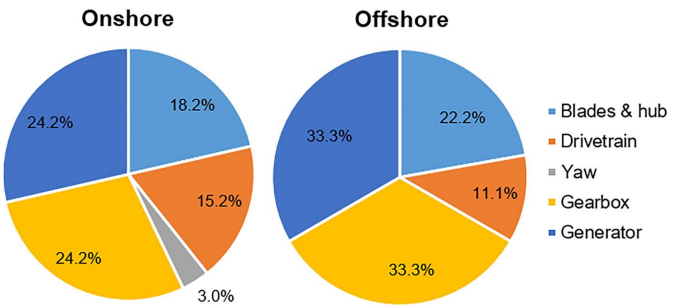

Figure 1.4. Downtime per failure for each sub-system of five major onshore and offshore wind turbine sub-systems. One third of total downtime is attributed to *gearbox* failures. Figure from [53].

can compromise the overall reliability of a plant, e.g. an industrial gas turbine engine, by substantially increasing friction forces between rotating parts. REBs usually operate in harsh environments (high speeds, contaminating environments, etc.), and as in the wind turbine gearboxes case, they are a frequent cause of failure [33].

REBs consist of two circular races, i.e. the outer and inner races, rolling elements, and the cage, which separates each element from coming in contact with each other. The rolling elements are usually made of hardened steel, and are manufactured such that they have an extremely high surface finish. Due to inevitable manufacturing imperfections though, e.g. during the grinding and polishing processes, there are radial variations of a few microns in length [107]. When the bearing elements are in motion, a vibration wave is generated in the bearing. Rolling elements, outer and inner races all have their characteristic vibration frequencies (see Chapter 2). For that reason condition monitoring of REBs is considered a relatively "easier" task than, for instance, SHM on a structure. The rolling-ball bearing is a common type of REB and is shown in Figure 1.5.

Note that, vibration from REBs will start to become dominant, with a few orders of

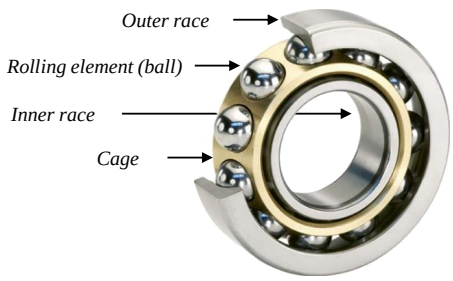

Figure 1.5. Rolling element ball bearing diagram, showing the four sub-components. Figure from [10].

magnitude above the noise source , e.g. instrumentation noise, when these imperfections grow and/or when a defect starts to form, e.g. in one of the outer races of the bearing.
In REBs, a typical damage scenario, is material fatigue. Fatigue causes micro-level spalls[¶] on the inner and outer races as well as on the rolling ball [177]. Fatigue spalls will appear suddenly and unexpectedly due to numerous factors, including overloading the bearings from improper clearances and from other components along the drivetrain failing, as well. For instance, the exertion of high localised loads from an unbalanced rotating shaft, will increase the friction force between the races and rolling balls. As such, with insufficient lubrication film, the shear stresses between the two elements in contact will result in micro-spalling. Improper maintenance or hard-material contamination of the bearing will accelerate the spalls formation. Although such damage is not catastrophic, by itself, when spalls grow, material tends to break off from the surface, e.g. from the outer race. Each time a rolling ball enters and exits a spall, will cause vibrations. When spalls grow, vibrations may increase within the drivetrain, which will compromise its performance and reliable operation.
The type of damage that this thesis is concerned with is fatigue spalling on the outer race of the bearing. This damage mechanism can wear the bearing very quickly [33], and therefore, needs to be detected at the early stages of damage progression, when the spall is relatively small and generates distinct periodic impulses. Extending spalls have shown not to generate impulses when a rolling ball is passing through them, due to their surfaces being smoothed out with time [149]. Hence, for the identification of this specific type of damage, an early identification is required. It should be noted here, that this thesis is concerned with DI and not the diagnostic capability in a drivetrain. However, when monitoring the condition of mechanical systems such as gearboxes, it is common practise to investigate each part in isolation to the rest of machine elements, and regardless of the acceleration measurement location. Hence, diagnostics are inherent in the analysis. The reason behind it, is due to each source s_M having a discrete or characteristic frequency associated with it. It is also one of the reasons that frequency-domain representations of vibrations in mechanical systems is a powerful tool to explore, at least, during the initial stages of the analysis.

1.3.5 Gas turbine engines with varying operating conditions

Gas turbine (GT) engines are critical components in both the aerospace and power generation industries. Their extremely high safety and reliability requirements, as well as high costs, due to possible overhaul and downtime, has prompted manufacturers to incorporate advanced CM strategies into their designs. In aircraft engines, the so-called engine health monitoring (EHM) systems comprise of embedded data acquisition systems and on-board

[¶]Spalls in REBs is the material removal that can be as small as a few hundred microns

monitoring for key engine parameters. Hourly sampled data, including vibrations, core temperatures and pressures, on certain flight regimes are being transmitted to a ground station for analysis [101]. These low-sampling measurements can be analysed further to infer system state.

The GT technology has been under tremendous research and development over the last decade, or so, with the aim to increase its efficiency and sustainability. A 2019 report by the International Air Transport Association (IATA) [57], stresses the importance of using sustainable non-hydrocarbon derived alternative fuels for powering aircraft, as part of the industry's future environmental goals, as well as, to reduce operating costs for airliners. Recently large consortia[‖], were assembled to develop sustainable alternative fuels (SAFs) for the aviation industry. A large body of literature exists in the use of SAF for GT engines. Studies such as [158] and [36], have shown that certain thermochemical properties, e.g. aromatic content, of different SAFs have significant impact on GT engine performance and its operating characteristics. Each SAF having very different flammability characteristics from each other, may cause a range of issues within combustion including auto-ignition.

Research is also being conducted on operating GT engines on high air-to-fuel mass flow ratios (AFRs), i.e. lean conditions, mainly to reduce fuel consumption and lower exhaust gas temperature, which in turn reduces Nitrogen Oxide (NOx) emissions [56]. Apart from performance requirements, a combustion device must operate safely at all times. Combustion instabilities, such as thermo-acoustic instabilities and lean blow-out (LBO) can occur. It has been shown in several key studies, e.g. in [136] and [82] that combustion instability favours *lean* combustible mixtures.

Thermo-acoustic instabilities is a phenomenon that occurs when dynamic combustion pressure and unsteady heat release rate are in-phase, adding energy to each other, i.e. their dynamically coupled. This causes self-excited oscillations. In the case of LBO, large amplitude oscillations in heat release rate prior to flame extinction occur. Both phenomena are *undesirable* as they will lead to engine structural deterioration inside the hot section very quickly, and furthermore to failure [113]. Detecting such events early is key in the quest of maintaining reliable combustor operation while achieving these performance improvements with SAF and lean conditions. More in-depth analysis and theoretical understanding of combustion instabilities can be found in [112].

[‖]For instance, JETSCREEN and HIGFLY. More information on both projects can be found on the relevant website https://cordis.europa.eu/.

1.4 SHM & CM via statistical pattern recognition

One way to perform SHM or CM is through the use of "model-based" approaches, by employing a set of physical laws in order to develop an *approximate* representation of a system. After such a model is built, it is then possible to *infer* whether there is damage in the system, by means of calculating the *residual* between expected dynamic characteristics and the equivalent from the measured data.

Usually model-based approaches employ Finite Element Method (FEM) models for structural engineering problems or Computational Fluid Dynamic (CFD) models (based on the Finite Volume Method) for fluid-flow problems. For CFD, the well-known Navier-Stokes equations (time-dependent conservations of *mass*, *momentum* and *energy*) are typically used to analyse the time-varying flow inside a combustion chamber (CC). Such high fidelity models can capture adequately the variation in velocity, pressure, temperature and density for *laminar* flows. In most engineering applications, however, the flow is most often than not *turbulent*. This requires a much finer mesh to improve (i.e. to the extend possible that is dependent on the Reynolds number (Re) [135]) the model's accuracy in predicting fluid flow. This comes at the expense of prohibitively high-computation resources. For instance, there are 10^{13} degrees-of-freedom (DOF) for a typical Reynolds number of 10^6, since DOF $\approx Re^{9/4}$ [172]. The solution to high Re flows, using CFD, is usually obtained by implementing a suitable turbulence model for a certain flow regime, e.g. $k-\epsilon$. In this case, the computational effort is reduced considerably as flow quantities such as velocity and pressure are now time-averaged [129].

It is clear that there are certain and rather strong set of assumptions, that need to be applied in model-based approaches. Similarly, for the FEM models many assumptions must also be made. For instance, there is the inherent assumption that material properties are isotropic throughout [200], which is not at all true for composite materials.

Apart from the above, phenomena such as thermo-acoustic instabilities in CC, consisting of non-linear interactions between heat transfer modes, chemical reactions and gas dynamics among other factors, cannot be captured sufficiently by a set of physical laws alone [52]. Also, due to inter-system variability, e.g. no two engines will produce identical vibration levels, a generic theoretical model developed for detecting unwanted engine scenarios will be of no practical use.

Another way to perform either SHM or CM is by taking advantage of the exponential growth in data availability through the use of *intelligent data-driven* methodologies. The field of *pattern recognition* implemented through advanced *machine learning* algorithms,

e.g. Gaussian Processes (GP) [150] and Support Vector Machines (SVM) [30], is particularly attractive for the purposes of developing automated DI frameworks. This approach is the so-called statistical pattern recognition (SPR) approach to SHM & CM and offers many advantages as compared to the model-based approach. This includes the ability to incorporate complex non-linear dynamic relationships in a more flexible model such as the SVM, by making relatively mild assumptions about the data generating mechanism, i.e. the monitored system.

As a simple example, a statistical hypothesis test can be used to determine whether the calculated mean values between the previous $\mu_X(n-1)$ and current $\mu_X(n)$ engine measurements arise from the same distribution under a given significance level, formulated as:

$$\begin{aligned} H_0 &: \mu_X(n-1) = \mu_X(n) \\ H_1 &: \mu_X(n-1) \neq \mu_X(n) \end{aligned} \tag{1.3}$$

where, H_0 is the null hypothesis and H_1 is the alternative hypothesis. In the latter case, immediate action may need to be undertaken in order to prevent potential failure as indicated by the data itself. Hence, a major limitation of this approach, is the requirement for obtaining high-quality data from the system. However, with the SPR approach, any uncertainty related to any part of the SHM process, e.g. the implemented data acquisition system and/or EOV, can be incorporated into the learning or statistical model by means of *Bayesian* approaches.

The SPR for SHM & CM includes the following four steps, as described in detail in [67]:

1. *Operational evaluation*: It includes an initial investigation of the benefits of SHM or CM in a system, its environmental and operational conditions and expected damage types, e.g. fatigue loading;

2. *Sensing, data-acquisition and cleansing*: Includes, specification of sensing equipment and location and data pre-processing, i.e. cleansing and removal of unwanted or erroneous data. In terms of sensing specification, a typical approach is to instrument the system using a sparse sensor array, using ambient excitation sources, e.g. traffic loading on a bridge;

3. *Feature extraction*: This includes, data transformation, compression or selection and calculation of certain characteristics that relate, for instance, to the expected damage type. Suppressing the influence of EOV on the data by means of certain data transformations, e.g. finding a sub-space in which EOV does not seem to alter the data (at least significantly), may also be examined at this stage.

   At this stage, the *curse of dimensionality* [30] must be taken into consideration. The

curse of dimensionality is directly related to the loss of *generalisation* performance for the statistical model. That is, given unseen data points, the model will make unstable and/or inaccurate predictions. This will happen, for instance, when the model has a large number of non-zero coefficients, in relation to the (simpler) underlying problem that needs to be solved (i.e. *overfitting*). Also, using highly correlated features tends to cause similar issues as demonstrated in [181].

To avoid that, *feature selection*, e.g. by means of dimensionality reduction techniques is required for the successful implementation of the next step in the SPR process. That is, both *low-dimensional* and high-information density** DSFs are sought;

4. *Statistical model development*: In the final stage of the SPR process, a decision must be made whether the system is damaged or not, based on the DSF from the previous stage. This will be either a feature *classification* or *discrimination* approach. Firstly, a statistical model is *learned* using the available DSFs (*training data*). Then, given unseen data points, converted as DSFs (*testing data*), decisions about the system state can be made.

   Note that, the above approach assumes that training data are from the system in its undamaged state (also called semi-supervised learning), and it will be demonstrated in Chapter 6. However, if we suspect that the training data include *both* damaged and undamaged states, then no statistical model development needs to be carried out. Instead, the discrimination procedure is being done inclusively. That is, the task is to *discover* observations in the existing dataset that are *"different"*†† from the rest, without any other information. This is the unsupervised learning procedure, which will be demonstrated in Chapter 5.

This thesis will be dealing with the last two steps of the SPR problem in SHM & CM.

1.4.1 Environmental and operational variability

A key element for the wide adoption of SHM, may lie on the high-*robustness* or *immunity* of the DSFs to EOV influence. This is necessary, in order to prevent any "masking" effects, e.g. increase in FP rates due to the simultaneous sensitivity to damage and EOV, which is usually the case. For instance, EOV by means of temperature changes in the ambient air, can modify the boundary conditions of a structure like a bridge (by thermal expansion), in addition to its stiffness properties, causing changes in its dynamic characteristics. For instance, consider the Alamosa Canyon bridge in the US, the authors in [62] have reported

**More information using a smaller number of features for the DSF matrix.

††The term "different" is used rather loosely in this context, because most of the outlier detection methods will make very distinct assumptions about these "differences".

a first mode frequency variation of $\approx 5\%$ on a typical *daily* cycle.
As it has been observed in many other studies, the EOV influence on system dynamics "behaves" in a similar way to the effects of damage [104], as far as vibration signals are concerned. EOV influence will cause an, otherwise, stationary process to behave in a non-stationary manner. Such changes can be readily observed in the time-scale of hours. Whereas, operational changes in machinery, may cause transient non-stationarity events in its dynamical behaviour, but, the time-scale is only a few seconds. So that when measurements do exist from EOV, its effects can be *removed* by estimating its trend (given a suitable set of observations), while operational changes can be *accommodated* through switching-state model approaches, e.g. threshold-switching dynamic regression models. Hence, as discussed in [187], detection of damage can be easily tackled by placing a $3-\sigma$ control limit on the EOV-free signal. On the other hand, when EOV is *unmeasured*, EOV influence is assumed to exist in a subspace, far from that of the damage influence. In Chapter 4, ways to solve the EOV influence are examined from the DSF point-of-view, while in Chapter 6 a new way is being proposed. In this thesis, the methodologies considered do not use any information regarding EOV or its influence on system dynamics.

1.5 Damage identification requirements

Prior to developing a fully-automated DI procedure, it is necessary to explicitly define its requirements. These are briefly described as follows:

- *Early DI*: An essential component in DI is the ability determine *sufficiently early* whether or not the system being monitored follows the intended or designed behaviour at any given point in time. Therefore, each sampled data point at any given time, or even a sample of data points within a short time-interval (for the case that our temporal resolution is sufficiently high) needs to be analysed and a decision must be made for that purpose.

- *Handling of non-stationarity*: This is probably the most important requirement in this thesis, since it is also considered the key-enabling technology for the wide and practical adoption of SHM in industry [67], as well as, the continuous success of CM in challenging environments. As the underlying statistical properties of data change over time due to operating in continuously varying and uncertain environments (EOV), the monitoring system must be either *immune* or *adapt* accordingly.

- *Unsupervised learning*: Identifying damage in a dataset may be conducted *inclusively*, where no labels are provided for all the examples, i.e. whether damaged or not. In

contrary, in semi-supervised learning, a separate training and testing set is provided, i.e. for statistical model development, as mentioned above. That is, a set of examples from the undamaged system is provided for training.
A major drawback of semi-supervised learning may be its infeasibility to obtain a *"sufficient" representation* of the system in its undamaged state, when such examples are inadequate.
On the other extreme, binary classification, i.e. supervised learning, will require labels for both damaged and undamaged examples. For the former, examples from the damaged system are rare and may be impossible to obtain in an SHM & CM context. Also, there is an infinite number of ways a system may fail.
Therefore, the system state may be inferred directly, i.e. inclusively, using unsupervised learning methods.

- *Sensitivity to damage*: When specifying the requirements of any DI, it is necessary to talk about performance in terms of the number of false positive (FP) and false negative (FN) rates. A high sensitivity to a certain DSF may trigger many false alarms when there is no presence of damage (high FP rate). However, signalling damage early may be needed to prevent catastrophic failures. Therefore a compromise needs to be made between the two and/or a purposely designed set of DSFs computed from the raw signals.

1.6 Aim of the thesis

The aim of this thesis is to formulate reliable and accurate DI frameworks and demonstrate their applicability on different datasets / problem cases. This thesis proposes the use of high-breakdown robust estimators of multivariate location (or mean) and scatter (or covariance), robust clustering and non-linear time series, for the identification and analysis of different types of outliers. To enhance DI, which means to reduce FP and, perhaps more crucially (for a safety-critical system), FN rates, DSFs were extracted.
Outliers in an existing dataset, may be present due to events related to both damage and EOV influence on the system dynamics. For that purpose, to *distinguish* between these two types of outliers, and thus, detect damage reliably, it is necessary to adopt certain methodologies.
It is particularly important to distinguish between the two types of outliers, in order to minimise misclassification rates. For instance, EOV influence tends to mask the effects of damage on system dynamics, by shifting or enlarging the distribution of data points

representing the system in its undamaged state. Any attempt to infer damage using distributional properties, e.g. covariance structure, from a single dataset representing a system under significant EOV influence, is not expected to be fruitful.
As mentioned in Section 1.5, one of the requirements for the fully-automated DI procedure is to identify damage in its earliest possible stage, while maintaining its high-robustness to EOV influence.
In terms of detecting damage, this thesis proposes two different methodologies:

1. The first methodology uses a set of DSFs, designed specifically to highlight the type of damage that is being expected, while at the same time being relatively insensitive to EOV influence. Their performance, in terms of DI, will be examined in detail and compared with common DSFs, e.g. natural frequencies;

2. The second methodology aims to investigate the potential of using outlier-free training sets for novelty detection tasks. These outlier-free training sets are obtained by using two multivariate and multimodal robust statistical methods. In that way, the influence of EOV, as seen on more common DSFs, e.g. natural frequencies and spectral amplitudes, is reduced with minimal any feature engineering.

Given the above methodologies, it is expected that the proposed monitoring strategy will yield satisfactory results for *different sets of problems* in both SHM & CM applications. In general, these two methodologies are the most novel parts of this thesis.
Moreover, this thesis aims to identify the distinct characteristics of the two types of outliers, i.e. those that arise from damage events and significant EOV influence, by means of subspace outlier analysis. The aim is to help to draw conclusions that apply to all sets of problems analysed in this work, and pave the way to find solutions that "filter-out" or distinguish EOV influence from damage. This task has also not been explored in the literature, in the field of SHM & CM.
The performance of various robust-based estimators is compared with more commonly used outlier and novelty detection approaches, which includes one-class SVM and mixture models.

1.7 Organisation of the thesis

This thesis is organised as follows:

- Chapter 2: literature review on the available signal processing methods for analysing time-series vibration signals and for extracting important features that can be used further for the identification of damage.

- Chapter 3: exploratory analysis of eight different datasets. In this chapter, the main characteristics of each dataset are discussed, including details of the data acquisition system used. Also, some basic vibration signal analysis, both in time- and frequency-domains, is conducted to better understand these datasets.

- Chapter 4: analysis and comparison for a range of DSF vectors proposed, in terms of their sensitivity to damage presence and EOV influence on the system.

- Chapter 5: inclusive outlier detection is performed, using the proposed DSFs from the previous chapter. A range of methods for estimating "robust" (or outlier-free) versions of multivariate location and scatter estimates are examined. As well as, robust clustering and time series analysis is presented and discussed. Moreover, part of this chapter is dedicated in the diagnosis of estimated outliers that exist in the data. Finally, a comparison using well-established outlier detection techniques is conducted.

- Chapter 6: proposes one way to obtain an outlier-free data representation for the undamaged state of the system to be used for training. Two robust methods, which avoid restricting the distribution of the training set, from a potentially multi-modal to a single elliptical one, are utilised for the purpose of obtaining an EOV-/outlier-free training data. This training set, is then used for semi-supervised learning tasks using three well-known techniques. The performance of this methodology, in terms of DI accuracy and sensitivity, is compared to the case where EOV influence presence in apparent on the DSFs.

- Chapter 7: main conclusions of the thesis are discussed, along with possible future work.

1.8 Summary

In this introductory chapter the following were discussed:

- The concept of SHM and CM, and several examples motivating the use of data-driven monitoring strategies, i.e. wind turbine gearbox rolling element bearings, alternative fuels on gas turbine engines for aircraft propulsion, bridges and other large civil engineering infrastructure;

- In terms of the characteristics for a monitoring system, four different requirements are specified and will be pursued throughout the thesis;

- Alongside a brief overview of the SPR framework for data-driven SHM and CM, the problem of EOV has also been mentioned and is recognised in the literature as a key issue for the successful implementation of monitoring strategies;
- Both SHM and CM, using vibration measurements, can be defined in terms of the blind source separation problem;
- The four main advantages of vibration-based DI, as well as, the associated challenges that such measurements impose on the reliability of the monitoring strategy;
- An overview of the objectives and methodologies of the thesis.

Chapter 2

Vibration signals, their transformations and estimation of common damage-sensitive features

2.1 Stochastic processes and signals

In vibration-based SHM & CM, determining the appropriate set of tools to analyse the different vibration signals, plays an important role in developing a robust monitoring strategy. This stage is fundamental to the design of DSFs that need to have the appropriate set of properties: sensitivity to damage and suppression of EOV influence. Therefore, it is first necessary to determine the characteristics of the signal that is being analysed, as this will yield the suitability of the methods that can be used.
With that in mind, an appropriate classification of the different vibration signals expected in SHM & CM contexts will be provided, similar to the one used in [166]. This will lead to a review of the various techniques available in *time*, *frequency* and *time-frequency* domains.

Vibration signals are obtained as a sequence of real valued signals, observed at *successive* and *equally spaced* time intervals. This may also be referred to as a *time series*. A single time series is considered as one particular *realisation* of a *stochastic process*[*] $X(t)$, where $-\infty < t < \infty$. Any finite or infinite set of realisations generated by the same stochastic process is called an ensemble $\{X(t)\}$.
To describe such a *continuous* stochastic process, the first and second (central) statistical moments are usually used [166]. The former is the mean function $\mu_x(t) = \mathbb{E}[X(t)] = \int_{-\infty}^{\infty} xp(x,t)dx$, while the latter is the variance function $\sigma_x^2(t) = \mathbb{E}[(X(t) - \mu_x(t))^2]$, i.e. $\sigma_x^2(t) = \int_{-\infty}^{\infty}(x - \mu_x(t))^2 p(x,t)dx$, where x is any value that the stochastic process $X(t)$ may take at any given time t. Note that, $p(x,t)$ is the time-dependent evolution of the

[*] A random variable that evolves over time.

probability density function (PDF) of the stochastic process taking any x value.

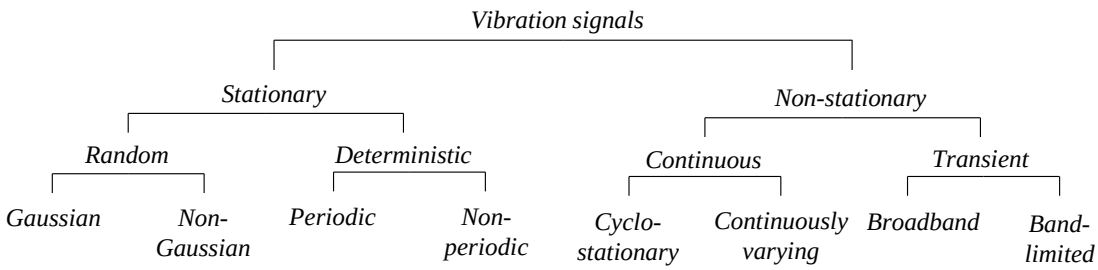

Figure 2.1. Classification of vibration signals: as stationary and non-stationary.

The hierarchical classification tree in Figure 2.1, separates signals into the *type* of stochastic process, i.e. those that are *stationary* and *non-stationary*. A description of the most important signal types is provided below:

- *Stationary*: Typically, a stationary stochastic process is one whose statistical properties are said to be in "steady state". That is, its statistical properties remain the same along $-\infty < t < \infty$. A weakly stationary, or simply, stationary stochastic process satisfies the following two assumptions, $\forall\ t$:

  1. The PDF is time-invariant, $p(x,t) = p(x)$, so that the first two statistical moments are constants: $\mu_x(t) = \mu_x$ and $\sigma_x^2(t) = \sigma_x^2$;
  2. Given $X(t_1) = x_1$ and $X(t_2) = x_2$ the *joint* PDF $p(x_1,t_1;x_2,t_2) = p(x_1,t_1+\tau;x_2,t_2+\tau)$ for any time lag value τ. That is, the joint PDF for any given time shift τ remains the same.

  For a stationary process, an important measure that describes the degree in which different pairs of samples can be linearly related to each other for any time shift τ, can be computed by the *auto-covariance function*. That is, by generalising the expression of $\sigma_x^2(t)$, the linear relation of the time series at t_1 with itself at t_2 is given as [167],

  $$C_{xx}(\tau) = \mathbb{E}[(X(t_1)-\mu_x)(X(t_2)-\mu_x)] = \mathbb{E}[(X(t)-\mu_x)(X(t+\tau)-\mu_x)] \tag{2.1}$$

  Note that, almost all physical processes exposed to environmental and operational changes, as well as component degradation, etc., will exhibit some degree of non-stationarity. However, given the fact that there is the need to compute *time-averaged* quantities, the assumption of stationarity is a necessary condition to do so. This is because, its statistical properties remain constant across all t. The stationarity assumption may also be justified by the fact that it typically yields relatively good estimates, in practise [166];

- *Non-stationary*: Given the above stationarity assumptions, if a time-series does not satisfy both of them, then the generating process can be automatically classified as non-stationary. It is relatively difficult to assess if a process is non-stationary, but, there are a number of statistical tests that are helpful in making such a decision, e.g. see [29].

  Consider for instance, the model $x(t) = \alpha x(t-1) + \varepsilon(t)$ as $t \geq 0$ with $\varepsilon(t)$ being a stationary process. This implies that, $x(t)$ will be stationary if $|\alpha| < 1$. Otherwise, it is a non-stationary process. Such a signal may be generated by a pure random walk or an autoregressive (AR) model of order 1, i.e. when $\alpha = 1$. This is referred to as a *non-mean-reverting* process, which drifts away from the expected value of $x(t)$, and as time goes to infinity its variance does, too [179];

- *Random*: Its value at any time sample t is a realisation of a stationary stochastic process, as described previously. Such a type of signal, has *no discrete frequency components* (or it has a flat frequency spectrum).

- *Deterministic*: On the other hand, deterministic signals are entirely composed of discrete frequency sinusoids [33]. Therefore, their value can be calculated analytically at any given point in time;

- *Continuous*: Its *spectral components vary* continuously with time, and thus, these signals are usually treated using time-frequency domain methods;

- *Transient*: These signals exist on *finite* time-intervals having finite energy, e.g. an impulse vibration response of a structure due to an impact. Such signals are analysed as a whole (i.e. not in parts as in continuous non-stationary signals).

The above classification helps in identifying the range of methods applicable for analysing each type of signal. At the lowest branch of the tree, however, there are two types of signals that need to be discussed: the *cyclo-stationary* and *Gaussian* signal. The former is important to the analysis of CM signals, while the latter is fundamental to a range of techniques used for pattern recognition and machine learning.

The cyclo-stationary signal, is one whose n^{th} order statistic is periodic. For instance, a white noise process that is modulated by a periodic amplitude is considered an nth-order cyclo-stationary signal if its mean (where $n = 1$) or auto-covariance (where $n = 2$) functions are periodic [195]. Although such a signal is not necessarily periodic, it is possible that there is a periodic component buried in it. Most signals generated by *rotating machinery*, such as gearboxes, are of the cyclo-stationary type. Methods such as *spectral correlation* are useful for analysing them, as will be shown later on.

Probably the most important stationary stochastic process is the *Gaussian Process* (GP). A relatively simple definition for a GP is that the joint PDF between all time samples $t = t_1, t_2, ...$ and so on, for the stochastic process $X(t)$, must be a multivariate (or p-dimensional) Gaussian [150], [99]. A GP is one of the most convenient models that can be used to describe a physical process. This is partly because it can be fully parametrized by only two functions, i.e. its mean and covariance. And partly because, many physical systems, machinery vibrations, can be modelled by a Gaussian. This due to the *central limit theorem* (CLT), which states that when adding k independent and non-identically distributed random variables together, in the limit when $k \to \infty$, the result of the summation will tend to be Gaussian distributed.

Given a multivariate mean $\boldsymbol{\mu} \in \mathbb{R}^p$ and covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$, a multivariate Gaussian can be used to simulate realisations for the random stochastic process $X \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ according to:

$$p(\mathbf{x}\,|\boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{p/2}} |\Sigma|^{-1/2} exp\left\{-1/2(\mathbf{x} - \mu_x)^T \Sigma^{-1} (\mathbf{x} - \mu_x)\right\} \tag{2.2}$$

where, $\mathbf{x}$ is now a p-dimensional vector, e.g. having p DSFs.

The white noise signal, is an example of a random signal that is generated by a Gaussian distributed stochastic process, having a mean equal to 0. The white noise signal ε, has a flat spectrum, i.e. it is constant over all frequencies, while its autocorrelation function is equal to zero. This random signal is an important one as it is widely used in many techniques to account for, e.g. prediction errors in an AR model, assuming its sample values are both identical and independently distributed (iid).

In practise, vibration data from experiments, will be composed of many different types of signals. For instance, the frequency component of a constant-speed rotating machinery is *deterministic*, but, it may also be modulated by the transmission path frequency component, which is *transient*. *Random* white noise from instrumentation and other sources are also to be expected. This is explained previously in the BSS analogy in Chapter 1. Therefore, one will need to relate the expected type of signal to the specific subsystem or component that is being monitored, e.g. a shaft's rotation to its periodic signal characteristic.

A major challenge in dealing with non-stationary signals is that they cannot be easily predicted. As mentioned previously, their statistical properties may not revert to a finite value, e.g. an explosive AR time series [75]. The usual approach to this problem is to first transform a non-stationary time series into a stationary one by means of a mathematical

operation, e.g. time-differencing. This leads to the so-called ARIMA (Auto-Regressive Integrated Moving Average) models [180]. More details on this later. What is important to note here, is the fact that when dealing with non-stationary signals (having time-varying statistical properties), any inferences made at the *current* time interval will not necessarily hold true for a *future* time interval. For that reason, when strong non-stationarity is suspected, e.g. due to significant EOV influence, then a more appropriate set of methods must be employed. In the following two sections, a number of methods are reviewed for both stationary and non-stationary signals.

2.2 Signal processing for stationary processes

A number of signal processing methods used extensively in SHM & CM, including Fourier-based transformations, assume that the stochastic process generating the time series is stationary. Some of them are discussed in this section.

2.2.1 Statistical measures

Many useful statistical measures can be obtained from a stationary random vibration signal, in order to aid the DI process. As mentioned previously, a Gaussian type of signal can be fully characterised by its first two statistical moments. To characterise the *asymmetry* of a distribution *with respect to the Gaussian*, the third statistical moment is used, i.e. skewness. Similarly, its degree of flatness or *impulsivity*, is given by the fourth statistical moment, i.e. kurtosis.

Consider a univariate time series vector $\mathbf{x} \in \mathbb{R}^N$ sampled at N discrete time intervals, ranging from $n = 1, ..., N$. The expression for skewness and kurtosis is given in Equation 2.3, where $l = 3$ and $l = 4$, respectively:

$$\mu_l = \frac{1/N \sum_{n=1}^{N} \left(x(n) - 1/N \sum_{n=1}^{N} x(n) \right)^l}{\sigma_x^l} \tag{2.3}$$

When the centre of gravity of the probability distribution leans towards the right of its mean, it is negative skewness, while when leaning towards the left of its mean, it is positive skewness. Note that for the Gaussian PDF (or simply Gaussian), $\mu_4 = 3$. So that, any *significant* deviation from that value indicates that the distribution of the sampled signal is no longer Gaussian. Given the wide range of physical processes that can be modelled or assumed to be sampled from a Gaussian [166], it is expected that any *irregularities* in the vibration responses due to damage presence will cause deviations from this assumption. Therefore, kurtosis and skewness can be used for DI.

Other measures, including spectral peak, spread and entropy will also be examined as potential DSFs for *system-agnostic* SHM & CM applications. That is, for systems where there is no information available about their dynamics and/or their possible damage mechanisms. These additional measures will be presented and discussed in Chapter 4. These can also be called as *global DSFs*, because they provide a single statistic for the system as a *whole*.
Kurtosis-based measures, such as spectral kurtosis, are considered very effective for DI, whenever damage manifests itself in the form of *impacts*, e.g. in CM for REBs [79].

2.2.2 Fourier Transform

Some of the most frequently used methods, both in CM research and practice, are based on the *Fourier Transform* (FT) [33]. FT computes the dot product, i.e. the *similarity*, between a given family of periodic basis functions and the sampled vibration signal vector $\mathbf{x}$. In that way, the FT represents the signal, in terms of its magnitude and phase at each frequency.
In rotating machinery, this *decomposition* is important because the main purpose is to identify damage on a sub-assembly/-system level, e.g. gears. Each sub-system, rotating with a frequency associated with the running speed of the system (i.e. a multiple of it), will have at least one frequency component with non-zero energy. Thus, it is possible to *isolate* that sub-system from the rest of the sub-systems, so that, its state can be examined directly and more optimally [33]. For the discrete-time vector $\mathbf{x} \in \mathbb{R}^N$, the FT is *discrete-time* FT (DFT) of the sequence,

$$S_x(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk}, \qquad k = 0, ..., N-1 \tag{2.4}$$

where, the term $\frac{k}{N}$ is the normalised discrete frequency. The corresponding discretised analysis frequency for $\mathbf{x}$ is related to the sampling frequency f_s (in Hz) according to $f = \frac{k}{N} f_s \; Hz$ or $f = k\Delta_f \; Hz$, where Δ_f is the frequency bin width (in Hz).
The *maximum* number of analysis sinusoids is dictated by f_s due to the *Shannon/Nyquist sampling criterion*, which states that at least two samples are needed to represent a single sinusoidal cycle [166]. Aliasing (higher frequency components are shown as lower ones) will result if frequencies above $f_s/2$ are used to represent a signal. For instance, a signal component with an oscillation of $f_s/2 + f_1$ appears in the frequency spectrum at the aliased frequency of $f_s - f_1$, due to the DFT mirroring effect around $f_s/2$.
Note that the DFT has high computational complexity of $O(N^2)$. In practise, an efficient way to compute it is through the *Fast Fourier Transform* (FFT) algorithm. In Matlab, the *fft.m* function can be used. It works by decomposing the previous DFT of the sequence of

length N into two DFT calculations, each of length $N/2$. It is beyond the scope of the thesis to discuss the specific details of the FFT. However, an important consideration is that the FFT works most optimally when the length of the sequence is a power of 2 (reducing the number of calculations to $O(Nlog_2N)$).

As already discussed in Section 2.1, vibration data obtained from experiments will contain many different types of signals, including random ones such as white noise. In fact, "random" vibration analysis is a term that is used for analysing this type of data.

The main difficulty in dealing with such data using DFT comes from the fact that its bin width Δ_f is a function of its sample rate f_s and length N. This implies that certain frequency components in the signal may not be successfully appear in the frequency spectrum, for instance, when Δ_f is large enough and the noise floor is high.

The *Power Spectral Density* (PSD), on the other hand, is *normalised* to the bin width. This prevents any changes to the signal magnitude in the spectrum. The PSD can also be used to compare different vibration signals, having different $\Delta_f = \frac{f_s}{N}$ characteristics. For those reasons, the PSD is fundamental to the analysis of random vibration signals.

More specifically, the PSD shows the distribution of power across the Nyquist frequency range. It is calculated as the DFT of the autocorrelation function [166], or simply,

$$P_x(k) = \frac{1}{\Delta_f}|S_x(k)|^2, \qquad k = 0, ..., N-1 \tag{2.5}$$

where, the term $|S_x(k)|^2$ is the squared magnitude of the DFT. Given a signal from an accelerometer, the units of P_x are therefore, $[m/s^2]^2/Hz$.

By taking the squared magnitude, any phase information will be lost. Unless otherwise specified, in this thesis, the PSD is calculated using the Matlab function *welch.m*, which implements the Welch's PSD method [166]. In particular, Welch's PSD takes the average of k PSD estimates over the positive Nyquist range by dividing the time signal into k segments. The averaging reduces frequency components from noise sources.

2.2.3 Time series analysis

Time series analysis allows the derivation of a class of models that describe the data generating process in a compressed and effective manner. These models can be used for predicting new instances, or simply, their *coefficients* (derived from data) can be used as DSFs. In general, they describe the temporal relationship between pairs of observations in the vibration data vector $\mathbf{x}$. Given the fact that these models are parametric, they are interpretable, too. Classical time series models are of the *autoregressive moving average* (ARMA) type, given as,

$$x(n) = -\sum_{i=1}^{p} a_i x(n-i) + \sum_{i=1}^{q} b_i \varepsilon(n-i) + \varepsilon_n, \quad \varepsilon_n \sim \mathcal{N}(0, \sigma_\varepsilon^2) \tag{2.6}$$

where, p, q are the orders of the autoregressive (AR) and moving average (MA) parts, respectively. The coefficients a_i, b_i, are, respectively, the derived parameters for the AR and MA parts, respectively. Thus, when $q = 0$ Equation 2.6 reduces to an AR(p) model that uses its p lagged values of $\mathbf{x}$ as predictors. And similarly, when $p = 0$ Equation 2.6 reduces to an MA(q) model that uses its q lagged error terms (i.e. uncorrelated white noise).
Typically, and unless the error terms are correlated, only an AR process is used for modelling a time series. In that case, a model of order p can be determined via a suitable information criterion function like the *Bayesian Information Criterion* (BIC). Additionally, its *partial auto-correlation function* [132] can be calculated to verify whether or not any lagged terms greater than p have negligible values. BIC is used to penalise the sum of squared error terms for adding more parameters in the AR model. This is consistent with the model-building principle of parsimony [132].
Model parameters a_i are calculated by formulating the problem as a system of Yule-Walker equations, first, and then solve them recursively, typically using the Levinson-Durbin algorithm [167].
In the multivariate case, $\{\mathbf{x}(n) \in \mathbb{R}^p; n = 1, ..., N\}$, *cross-correlations* between each time series need to be analysed using *vector ARMA* (VARMA) types of models. Such models may have better prediction accuracies than ARMA models, since any linear inter-dependencies between each time series is evaluated.

2.3 Signal processing for non-stationary processes

The above methodologies, imply that the observations in the time series vector $\mathbf{x}$ come from a stationary process. In FT, for instance, the sinusoidal analysis functions used in the transformations are constant over the whole time-interval that the dot product is computed. Hence, FT methods are only meaningful for signals that are composed of purely sinusoidal components that remain time-invariant.
In many practical cases as in vibration analysis, a time series signal will be time-varying, e.g. when EOV influence is present or when N is not large enough. A non-stationary time series has a more complicated structure, and although it may consist of a certain number of discrete frequencies, the calculated PSD may contain power at other frequencies, too. Therefore, *temporal localisation* of the frequency decomposition is required for such signals to be represented accurately in the frequency-domain. The main premise for *time-frequency* or *time-scale* analysis methods is to expand the feature space from its original time-domain 1-D representation towards a 2-D one.
Since no additional information can be *created* for a signal [190], a time-frequency domain

representation, will lead to a DSF matrix that is highly-redundant. This violates the requirement for developing DSFs that have high-information density, which is an important element for the successful implementation of the SPR paradigm (see Section 1.4). However, in this expanded feature space, it is expected (although there are no guarantees) that there will be a much lower-dimensional subspace in which the majority of DSFs can be easily obtained. This is also true for the EOV influence, which may be isolated into a certain manifold in this new feature space [190].

Additionally, the methods presented in this section are *not restricted to linear* stochastic processes, for instance, as suggested by the AR model. As it will be seen in Chapter 3, most datasets examined in this thesis are generated from systems with significant non-linear characteristics. One example is the four-DOF numerical simulation based on the Duffing stochastic process model, which generates an *amplitude-dependent* frequency of oscillation.

2.3.1 Classical time-frequency analysis methods

2.3.1.1 Short-time Fourier Transform and Spectral Kurtosis

The *Short-Time FT* (STFT), extends the classical FT, to simultaneously analyse the signal in both time- and frequency-domains. This is done by sliding an *analysis window* w of length l_w and computing the discrete DFT (using FFT) of the signal S_x, within each windowed segment. To simplify notation, assume l_w is a multiple of N and that there is 0 overlapping samples between each subsequent segment. The result will produce $\left\lfloor \frac{N}{l_w} \right\rfloor$ DFTs,

$$S_x(m,k) = \sum_{n=0}^{N-1} x(n)w(n-mM)e^{-j\frac{2\pi}{N}nk}, \quad k = 0,...,N-1 \tag{2.7}$$

where, $S_x(m,k)$ is the complex-valued two-dimensional representation of $x(n)$, where k is the kth Fourier coefficient centred around the mth time sample. Note that the *spectrogram* is a time-frequency (2D) visualisation plot, which is obtained by computing the STFT magnitude, i.e. $|S_x(m,k)|$. In Matlab the built-in *stft.m* function is available.

In many cases involving transients in the signal, the STFT can be combined with the *spectral kurtosis* (SK) to help identify them. Typically, transients contain a certain degree of periodicity that needs to be identified from a mixture of different types of signal in a random vibration time series. The difficulty is that these periodicities may be hidden in the raw time series.

Note that signals of this type are 1st or 2nd order cyclo-stationary and are thus non-stationary. As discussed in [147], the periodic and non-periodic signal parts may be decoupled and, preferably, analysed separately. SK helps to determine the most impulsive *frequency band*. In brief (more details can be found in [149]), SK based on the STFT,

computes the fourth statistical moment across different frequency bands in the spectrum. Given the complex-valued function $S_x(m,k)$ from the STFT, SK as a function of frequency:

$$SK_x(k) = \frac{\langle |S_x(m,k)|^4 \rangle_m}{\langle |S_x(m,k)|^2 \rangle_m^2} - 2 \tag{2.8}$$

where, $\langle \cdot \rangle$ is the averaged operator over the sample index m. Note $K(k)$ is normalised with -2 to make it zero for a Gaussian random signal. High values (in the order of 10 to 30) are expected for impulsive signals.

Critical to the above calculation is the *fixed* analysis window length used for the STFT. As the *uncertainty principle* states, the time and frequency resolutions, Δt and Δf, respectively, are bounded: $\Delta f \Delta t \geq \frac{1}{4\pi}$. This limitation, therefore, will have a major impact on the effectiveness to accurately localise transient events within the time-frequency spectrum. Note that l_w will need to be shorter than the temporal spacing between each generated impulse (or transient), while also cover most of its length sufficiently.

For these reasons, SK_x is usually represented on a 2D plot called the *kurtogram*, i.e. $SK_x(k, l_w)$, to determine a suitable l_w. Thus, on its x-axis is the frequency $k = 0, ..., N/2$ and on its y-axis is the frequency resolution Δf. Since it is not practical to consider every l_w possible, it is implemented as a 1/3 binary tree filter bank using the *fast kurtogram* approach, proposed in [23]. The SK is then calculated on each of the decomposed frequency bands. The decomposition is done using a dyadic filter bank structure. A level r decomposition, results in 2^r frequency bands within the Nyquist range. As suggested by the corresponding Matlab implementation in [23], the maximum level of decomposition for computing the kurtogram could be calculated as $r_{max} = log_2(N) - 6$. As it will be shown later on, the kurtogram will used to identify the frequency band in which the signal is most impulsive.

2.3.1.2 Wavelet Transform

As noted above, the analysis of non-stationary signals requires *variable* time-frequency resolution. In this case, the window length is large on lower frequencies; offering good frequency resolution, i.e. small Δf, on *slow-varying* signal components. Conversely, the window length is small on higher frequencies; offering good temporal resolution, i.e. small Δt, on *fast* signal amplitude changes. This is achieved by means of the *Wavelet Transform* (WT), which utilizes basis functions with *varying window lengths* [117].

The coefficients of WT $w_{m,k}$ are calculated using the dot product between a set of complex-valued basis functions $\psi^*_{m,k}$ and the signal $x(n)$, i.e.

$$w_{m,k} = \langle x, \psi^*_{m,k} \rangle = \sum_n x(n) \psi^*_{m,k}(n), \quad m, k \in \mathbb{Z} \tag{2.9}$$

where, $\psi(n)$ is the prototype function (also called the mother wavelet) from a specific family of wavelets, from which all other kernel functions are derived as *dilated* and *translated* versions of it.

The basis function $\psi^*_{m,k}$ is, therefore, a function of two parameters (integers): the dilation $m \approx 1/frequency$, which defines also the analysis window length, and its translation k, which is responsible for shifting it through the sampled data. This means that in WT, there is no fixed set of basis functions, e.g. sines and cosines as in FT. Instead, the analysis takes a more flexible approach; $\psi(n)$ can now be *chosen* from a set of available function families or designed to better suit the signal $\mathbf{x}$ [48].

However, there are certain requirements for a wavelet basis function [48], including:

- Is an *oscillatory* function, i.e. $\sum_n \psi(n) = 0$, $\forall n$, that is *compactly supported*;
- Of *finite power* also, i.e. $\sum_n |\psi(n)|^2 = 0$, $\forall n$;
- And satisfies the *admissability condition* (requirement for an inverse WT to exist).

For the *discretised wavelet* $\psi_{m,k}$, a *dyadic sampling grid* is normally used [196]. That is, the scale m and translation k take only values of powers of 2. Therefore, at each scale m, k is run over the samples:

$$\psi_{m,k}(n) = 2^{\frac{m}{2}} \psi^*(2^m n - k) \tag{2.10}$$

Note that the *redundancy* in coefficients (using a *continuous WT* (CWT)) is greatly reduced with the above dyadic grid representation of the *discrete WT* (DWT). A special case now arises when the basis function used here is *orthogonal* to its own dilations and translations, guaranteeing linear independence between the coefficients, i.e.

$$\langle \psi_{m,k}, \psi_{m,p} \rangle = \delta(k - p) \tag{2.11}$$

where, the Kronecker delta function δ, is unity only when $k \neq p$, and zero, otherwise. The family of wavelets that is often used for its orthogonality properties is the Daubechies [67]. In practise, the DWT coefficients $w_{m,k}$, are calculated by convolving the time series signal with a subsequent set of half-band (containing half of the frequency content) low- and high-pass filters [117], yielding the low- and high-pass sub-bands, $(0, f_s/2]$ and $(f_s/2, f_s]$ respectively, at different *levels of decomposition*. In the DWT, the low-pass sub-band is decomposed further in the same way after *decimation by 2* (half the samples are eliminated, following the Nyquist criterion), while the high pass sub-band is not analysed any further. The signal after the first level of decomposition will have twice the frequency resolution than the original signal, since it has half the number of samples. This procedure is known as two-channel sub-band coding filter bank [117] and provides an effective way to calculate

the coefficients of WT by means of conjugate quadrature mirror filters.
Since the DWT *does not* analyse the high-pass sub-band further, as it does with the low-pass one, the *wavelet packet transform* (WPT) (see, for instance, SHM application in [173]) has been developed. In general, it follows the same procedure as the DWT, which offers a better frequency resolution on the *higher frequencies* by decomposing the corresponding sub-band further. The above methodology is also termed as *multi-resolution analysis*.
In general, the WT achieves a *trade-off* between frequency and time resolutions. That is, at higher frequencies (smaller m), there are more sample points corresponding to good time resolution. On the other hand, at lower frequencies (larger m), there are less sample points corresponding to good frequency resolution. In most cases, in engineering, such large scale components occur on large durations, e.g. a constant-speed rotating input shaft in an industrial gas turbine, hence their temporal resolution is not an issue. On the other hand, transients such as impacts are only analysed using short scales.

2.3.2 Adaptive time-frequency analysis methods

This section will review a number of methods for adaptive time-frequency analysis. These methods do not use a fixed set of basis functions (e.g. as in WT), rather they obtain a set of analysis functions from the *data* itself.

2.3.2.1 The Hilbert Transform

An alternative method for extracting time-frequency information from a non-stationary signal, is through the use of the *Hilbert Transform* (HT). The HT can be used to obtain the signal's time-varying spectral characteristics at any given point in time, i.e. its *instantaneous frequency* (IF) and *instantaneous amplitude* (IA) [33]. Perhaps most importantly, it does that without the need for a fixed basis function. For a non-stationary signal, where its statistical properties (and frequency characteristics) can vary with time, the estimation of IF and IA will provide a better temporal and frequency representation of the signal, without the need to sacrifice accuracy in either time or frequency domains.
The instantaneous properties of the signal are available only when a real-valued signal $x(t)$ (i.e. given in continuous time t), is converted into a *complex-valued* signal. In that way, the phase and magnitude information are accessible. Note that, a complex-valued signal occupies a three-dimensional space: time, real, and imaginary and can be represented by a time-varying complex exponential function of the form:

$$z(t) = |z(t)|e^{j\phi(t)} = x(t) + j\hat{x}(t) \tag{2.12}$$

where, j is the complex operator, $\phi(t)$ is the time-varying phase angle in $rads/s$ and $|z(t)|$ is the time-varying amplitude.

HT extracts and ads the phase-quadrature ($1/\pi$ cycle shift) component $j\hat{x}(t)$ to the $x(t)$, in order to create the complex-valued signal, also called the "analytic signal", as defined in Equation 2.12. Formally, the $\hat{x}(t)$ is the HT of the signal and is defined as the convolution operation:

$$\hat{x}(t) = x(t) * \frac{1}{\pi t} \tag{2.13}$$

The FT of Equation 2.13, illustrates the HT operation in the frequency-domain:

$$S_x(f) = -jsgn(f)S_x(f) \tag{2.14}$$

where, $-jsgn(f) = -j$ for $f > 0$, $-jsgn(f) = j$ for $f < 0$ and $-jsgn(f) = 0$ for $f = 0$. Hence, as Equation 2.14 illustrates, the HT is implemented as a *finite impulse response* (FIR) filter, which shifts the phase of positive frequency components by $-\pi/2$ and of negative frequency components by $+\pi/2$[†]. Its magnitude, on the other hand, is equal to 1.

The analytic representation of the signal will contain only the positive frequency components of $x(t)$, since the negative frequency components cancel out during the HT filtering operation. In fact, the Matlab built-in function *hilbert.m* calculates the HT by taking the FT of $x(t)$, doubling the positive frequency components of the signal and setting the negative ones equal to 0. The next step is the inverse FT of the signal, in order to get the complex-valued copy of it, as in Equation 2.13. More details of this algorithm in [121].

Given the complex-valued signal $z(t)$, the IA and IF are simply given as:

$$\begin{aligned} IA : |z(t)| &= \sqrt{x(t)^2 + \hat{x}(t)^2} \\ IF : \dot{\phi}(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} &= \frac{d}{dt}\left[tan^{-1}\left(\frac{\hat{x}(t)^2}{x(t)^2}\right)\right] \end{aligned} \tag{2.15}$$

2.3.2.2 Non-linear systems and the requirement for mono-component signals

The estimation of $\dot{\phi}(t)$ is only meaningful for *mono-component* or *narrow-band* signals, having "concentrated" frequency characteristics [31]. Strictly speaking, in *adaptive signal processing*, a signal is expected to consist of a finite number of K amplitude- and frequency-modulated mono-components. In continuous time, this is expressed as,

$$x(t) = \sum_{k=1}^{K} A_k(t)cos\left[\int \dot{\phi}_k(t)dt\right] \tag{2.16}$$

[†]Note that the positive frequency components are defined between between $f = [0, fs/2)$ and negative ones when $f > fs/2$.

where, for each kth mono-component, both $A_k(t)$ and $\omega_k(t)$ need to be slowly-varying (i.e. $\forall\ k$), for the HT analysis to be valid [68]. That is, such a signal is guaranteed to have a positive IF. On the contrary, a fast-varying IF may be negative, which is not meaningful in physical terms.

For highly non-linear systems with varying frequency components at very different time-scales, a negative IF calculation is common [68]. Such non-linear processes, exhibit phenomena like *jumps* and *harmonic distortion*, as demonstrated in Appendix A using the Duffing mass-spring-damper solution to a Fourier series expansion.

It is obvious that when dealing with non-linear and non-stationary processes signal *pre-processing* is needed prior to HT analysis. One of the first attempts used band-pass filtering, to create a narrow-band signal [31]. However, there are issues associated with this approach, including the requirement to parametrize the filter. Thus, a more flexible/adaptive approach, suitable for non-linear and non-stationary systems, is needed.

The next section provides an overview of six different adaptive signal processing / decomposition techniques, which can be used to obtain mono-component signals for subsequent HT analysis.

2.3.2.3 EMD-based & iterative decomposition methods

For the purposes of utilising the powerful analysis offered by HT, in their landmark paper [89] Huang et al. came-up with what is called the *empirical mode decomposition* (EMD). This algorithm is used as an extraction tool for *intrinsic mode functions* (IMFs), which are mono-component signals. Note that there are two main requirements for a valid IF and IA representation of any given signal:

1. The number of extrema (maxima and minima) and zero crossings must be equal or differ at most by one;

2. At any time instant the mean value of the envelope defined by the local maxima and the envelope defined by the local minimum must equal to zero.

The above requirements can be used as the definition of an IMF, i.e. it is any function that satisfies them both. The EMD is simply a way of obtaining the IMFs of a signal, similar to a filter-bank method. However, the basis functions are now derived from the data itself, which makes it an *adaptive method*. The EMD follows a *sifting procedure*, aiming to decompose a signal into a finite number of IMFs, as described in some detail in Table 2.1. The Matlab built-in function *emd.m* is used to implement this method.

Note that, unless otherwise stated, the signal that is being analysed is represented by the one-dimensional time series vector $\mathbf{x} \in \mathbb{R}^N$. Equivalently, using the discrete-time sampled

Table 2.1. The sifting algorithm of the EMD.

1.	Set $k=0$ and $r_0(n)=x(n)$. Calculate the extrema of $x(n)$, i.e. $\forall\ n$ samples.
2.	Interpolate (using cubic spline) between all minima and all maxima, to obtain lower and upper envelopes, $e_{min}(n)$ and $e_{max}(n)$, respectively.
3.	Compute mean envelope, $m(n)=(e_{min}(n)+e_{max}(n))/2$.
4.	Compute the IMF candidate, $c_{k+1}(n)=r_k(n)-m(n)$.
5.	Is c_{k+1} an IMF? If *yes*, then c_{k+1} is the $k+1$ IMF. Compute residual $r_{k+1}(n)=x(n)-\sum_{i=1}^{k}c_i(n)$. Increase sifting number index, i.e. $k=k+1$ and use r_k in step 2. If *no*, then repeat step 2-5 by using c_{k+1}.
6.	Terminate procedure until the final residual $r_K(n)$ is a monotonic function.

one-dimensional representation of $x(n)$ with $n=1,...,N$.

The above sifting algorithm, eliminates riding waves and transforms the waveform of the signal in each kth sifting iteration to be more symmetric with respect to the zero mean. To prevent the information loss due to excessive reduction in amplitude and physical meaning, a stopping criterion ψ is used. One of the most common criteria is the Cauchy, which uses threshold ρ, such that,

$$\psi=\frac{||c_{k+1}(n)-c_k(n)||_2^2}{||c_k(n)||_2^2}<\rho,\quad \rho\in(0.2,0.3) \tag{2.17}$$

It is a well known fact that the EMD may produce IMFs with overlapping frequency content. This issue is known as the "mode mixing" problem of the EMD. This is typical when the signal contains an intermittency, e.g. white noise or abrupt signal changes. Recall that in EMD, the local mean $m(t)$ is computed via interpolation, which in most cases is a cubic spline. However, the splines are unstable when an intermittency is present, and as such mode mixing and/or splitting does occur [144].

The mode mixing has been demonstrated many times in the literature, including in [118], when the authors used EMD to decompose a mono-component signal with an added transient. In general, the mode mixing phenomenon in adaptive signal processing is when an IMF contains oscillations of disparate scales, or when a component of similar scale is present in different IMFs [97].

The ensemble empirical mode decomposition (EEMD) [191] addresses this problem by adding Gaussian white noise to the entire signal. Applying EMD on the noisy signal, will generate K IMFs $\{c_k(n);k=1,...,K\}$. For an ensemble of M times, the same procedure is repeated (creating noisy versions of the signal and applying EMD to it). Averaging through all M members in the ensemble for each of the K IMFs *separately*, gives the averaged i^{th} IMF, i.e. $\bar{c}_i=1/M\sum_{m=1}^{M}c_{i,m}$.

The idea behind EEMD was inspired by the fact that the EMD acts as a dyadic filter bank when applied to a purely white noise signal. This means that the frequency bands are uniformly populated in the whole time-frequency space. When a signal is added to this uniformly distributed white noise background, the different frequency components of the signal are projected onto proper bands of reference, defined by the background noise. If the background is not white noise, then "mode mixing" occurs.
As each member of the ensemble consists of a different realisation from a Gaussian PDF, the white noise component will reduce according to σ_ε^2/M, where σ_ε^2 is the noise variance. This means that as M becomes large, the IMFs of the signal will be the only components that will persist through the averaging process.
Notice that, EEMD needs specification of σ_ε and M. Moreover, to ensure equal number of IMFs between each of the M members of the ensemble, the authors in [191] proposed to use a relatively small number of iterations (for the sifting process). However, enforcing this type of constraint can result in inadequate signal decomposition.

For a more robust signal decomposition, the *improved complete ensemble empirical mode decomposition with adaptive noise* (ICEEMDAN) [49], the *the time-varying filter empirical mode decomposition* (TVF-EMD) [109] and *the empirical mode decomposition with soft sifting stopping criterion* (EMD-SSSC) [115], were developed. The basic premise of ICEEMDAN involves generating M realisations of a white noise term ε_m ($\{m = 1, ..., M\}$) to create M noisy copies of the original signal $x(n)$. Each mth realisation is decomposed up to its 1st IMF using EMD (Table 2.1). Both the added white noise and the "controlled" EMD decomposition, diminishes the probability of mode mixing [178]. The procedure is outlined in Table 2.2. The Matlab implementation is publicly available on one of the authors' personal website [3].
Note that the EMD uses a *hard* stopping criterion, based on a threshold applied on the standard deviation of two consecutive sifting iterations (Equation 2.17). It has been proven in many studies, e.g. in [115], that this approach does not decompose the signal into ideal IMFs, i.e. satisfying the two main requirements set above. This implies that such a stopping criterion is prone to mode mixing. For this reason, the authors in this paper, have developed a *soft* stopping criterion for the EMD. This criterion involves the computation of the *kurtosis* (or excess kurtosis) and RMS values for each iteration. Whenever two consecutive sifting iterations show an increase in both kurtosis and RMS values for the decomposed IMF, the sifting process stops. Otherwise, it stops until the maximum number of iterations is reached. The method was applied on DI for REB damage, and have shown promising results. The Matlab implementation for this algorithm is provided by one of the authors in

Table 2.2. The sifting algorithm of the ICEEMDAN.

1. Obtain M realisations of $x(n)$: $x_m(n) = x(n) + \beta_0 S[\varepsilon_m(n)]_1$, where $S[\cdot]_1$ is the EMD decomposition up to the 1st IMF, $m = 1, ..., M$ and β is a scaling constant for the noise term ε.
2. Use EMD, to decompose each $x_m(n)$ up to the 1st IMF, i.e. $c_{1,m}(n)$.
3. The 1st IMF is the average, i.e. $c_1(n) = 1/M \sum_{m=1}^{M} c_{1,m}(n)$.
4. The first residual is calculated as $r_1(n) = x(n) - c_1(n)$.
5. Use EMD to decompose $r_2(n) = r_1(n) + \beta_1 S[\varepsilon_m(n)]_1$.
6. The kth IMF is thus, $c_k(n) = 1/M \sum_{m=1}^{M} S[r_{k-1}(n) + \beta_{k-1} S[\varepsilon_m(n)]_1]_1$.
7. The kth residual $r_k = r_{k-1} - c_k(n)$.
8. For $k = 1, ..., K$ repeat $(5) - (7)$ until the Kth residual is a monotonic function.

[8].

On the other hand, the TVF-EMD is an EMD-based technique that approaches the signal decomposition problem by constructing a low pass filter, with *time-varying cut-off frequency*. This filter is used to separate the signal into its local low and high frequency parts. Subsequently, the method successively applies the time-varying filter on the local low frequency part of the signal, until a stopping criterion is reached. At each iteration, the calculated local high frequency part will be a locally narrow-band signal, which is amenable to analysis with the HT (i.e. solving the mode mixing of the EMD). Each local high frequency component will satisfy the requirements for being an IMF[‡]. The development of the method was inspired by the fact that the *mean envelope* of the EMD algorithm ($m(n)$), is also a time-varying filter.

The authors of the TVF-EMD algorithm demonstrated in their paper [109], that a *B-spline polynomial* is a special type of low-pass filter, whose cut-off frequency is determined by its *knot spacing*. In addition, its order determines the roll-off specification of the filter (or filter attenuation rate). Note that the most important feature of the TVF-EMD algorithm is the computation of the time-varying filter. Whereas, the sifting process is the same one as in Table 2.1 of the original EMD algorithm. Table 2.3 outlines the steps of the time-varying filter estimation. Some further details are also provided in Appendix B.

Therefore, using the knot spacing of the B-spline filter, the EMD algorithm is carried out as usual, for each iteration. A stopping criterion that defines whether the decomposed signal is "sufficiently" narrow-band or needs to be filtered further, is calculated as the ratio of the

[‡]It should be noted that the method is not restricted to the definition of the IMF, such as requiring symmetric upper and lower envelopes. Instead, it uses a more flexible definition for the local high-frequency narrow-band signal. However, to avoid confusion the term IMF will be used for a signal that is sufficiently narrow-band for a valid HT analysis.

Table 2.3. The main steps of the time-varying filter calculation used for the TVF-EMD algorithm.

1. Calculate IF and IA using HT.
2. Find the *extrema timings*, i.e. time points for all local minima and maxima of IA.
3. *Interpolate* the two sets of time points, separately.
4. Step(3) allows the calculation of a new set of IFs, which further enables the computation of the *bisecting frequency*, which is the frequency that separates the local low and high frequency parts of the signal.
5. The rate of change for the bisecting frequency is limited by a fixed value (using a realignment algorithm) to prevent abrupt signal changes (intermittencies) that can cause mode mixing.
6. The knots of the B-spline approximation filter are calculated as the *extrema timings* of the function $h(t) = cos\left[\int \dot{\phi}_{bis}(t)dt\right]$, where $\dot{\phi}_{bis}(t)$ is the bisecting frequency (from Step(5)).

Loughlin instantaneous bandwidth [116] and the weighted IF average (see Appendix B). Note that the Matlab implementation of the TVF-EMD can be found in [13].

2.3.2.4 Iterative filtering and VMD-based decomposition methods

Similar to the TVF-EMD, the *adaptive local iterative filtering* (ALIF) [44], is a technique that utilises a special type of low-pass filter to replace the mean envelope $m(n)$. In particular, ALIF computes a *moving average filter*, whose *length* is a function of the *input signal*. For this reason it is suitable for non-stationary signals. In comparison, moving average filters with a fixed length assume similar statistics within their averaging window. More specifically in ALIF, a filter function, is convolved with the signal to generate the moving average filter. This filter function poses some ideal properties, including smoothness. The authors used the solution to the Fokker-Planck equation for this filter function, as it satisfies these properties. The moving average filter is then used to calculate IMFs by using a similar sifting procedure, as in the EMD algorithm. The method is implemented in Matlab using the code from the University of L'Aquila's website [1].

The *variational mode decomposition* (VMD) technique [63] and its variants (e.g. the multivariate one [182]), calculate each IMF[§] simultaneously, instead of iteratively via an optimisation procedure. In particular, if K is the number of expected IMFs in the original signal, then each IMF c_k is assumed to be compact around a centre frequency ω_k, whose

[§]In VMD [63] the term "mode" is used instead of an IMF, which has a slightly different definition. For the purposes of this thesis, the term IMF will be used for narrow-band and mono-component signals, calculated from any decomposition method used.

bandwidth is calculated in the frequency spectrum. Using Lagrange multipliers $\lambda(t)$, the optimisation problem can be described as,

$$\begin{aligned}\mathcal{L}(\mathbf{c}_k, \omega_k, \lambda) = \alpha \sum_{k=1}^{K} \left\| \frac{\partial}{\partial t} \left\{ \left[\delta(t) + j\frac{1}{\pi t} \right] * c_k(t) \right\} e^{-j\omega_k t} \right\|_2^2 \\ + \left\| x(t) - \sum_{k=1}^{K} c_k(t) \right\|_2^2 + \langle \lambda(t), x(t) - \sum_{k=1}^{K} c_k(t) \rangle\end{aligned} \tag{2.18}$$

where, α is a penalty factor and $\delta(t)$ the Dirac delta function. The saddle point of the above Lagrangian is the solution to minimising the *difference* between the *sum* of IMFs $\mathbf{c}_k(t) \in \mathbb{R}^K$ and the signal $x(t)$. To *assess* the bandwidth of each candidate IMF, i.e. whether it satisfies the IMF or narrow-band properties (as discussed previously) the following main steps are followed:

1. The analytic signal of each $c_k(t)$ is calculated using the HT. All negative frequencies are then discarded (as they do not have physical meaning);

2. Each candidate IMF spectrum $c_k(t)$ is shifted to its estimated centre frequency ω_k by multiplying it with a complex exponential tuned to ω_k;

3. Then the bandwidth is computed using the L^2-norm of the gradient.

Similar to both the EMD and ALIF, a stopping criterion on the standard deviation of the IMFs is constructed such that the algorithm finishes to update $\lambda(t)$ when this is below a certain threshold. The Matlab built-in function *vmd.m* is used.

One of the issues with the VMD arises from the fact that it convolves the signal's IMFs with a complex exponential. A system like a GT engine cannot be assumed to be stationary on long-time scales, and therefore, VMD may only be applied on smaller time-segments of the signal (where stationarity may be assumed). Also, given the VMD requires the specification of K, this poses some restrictions on the kind of problems that SHM & CM deals with. That is, in most real-world applications, e.g. a GT engine, the number of IMFs cannot be possibly known due to the complexity of the system that is being monitored. Any misspecification of K may result in mode mixing.

The successive variational mode decomposition (SVMD) [137] has been recently proposed to tackle the issues of the VMD. Briefly, this is done by including additional criteria on the optimisation problem such that it minimizes (or eliminates) the overlap in the frequency spectrum of each IMF. Note that VMD operates as a spectral separation technique, so that it also assumes that the IF of each IMF has not overlapped with each other. This may also be a potential cause of mode mixing in the original VMD, which is expected to be solved

by SVMD.

Note that, a time-series generated by a non-stationary process can also be modelled directly using, parametric and non-parametric techniques. One of the first attempts for the former category was the threshold autoregressive (TAR) model, which assumes piecewise linearity [180]. Non-parametric ones include the aforementioned GPs. However, since each time-series may be sampled with a rate in excess of 1 kHz on the time-scale of a few minutes, modelling the data vector directly will not be possible. That is, to prevent issues including modelling noise in the data, a set of DSFs need to be calculated first that provides certain key characteristics for each time-segment of the whole time-series. Hence, the above IMF-decomposition techniques will provide a way to obtain DSFs, as described in the next section.

2.4 Common DSFs from vibration responses

This section outlines the identification procedure used for estimating the modal parameters from structures for SHM purposes. For CM on REBs, the procedure of the envelope analysis (EA) will be also discussed, as a potential set of DSFs for an automated DI. Both of these sets of DSFs are commonly used in SHM & CM applications and will be analysed further in a subsequent chapter of this thesis. In Chapter 4, additional types of DSFs found in the literature are discussed.

2.4.1 Modal parameters in SHM

In vibration-based SHM it is typical to extract modal parameters, i.e. NFs, MS and damping ratios (DRs), from structural accelerations. In large and complex structures, e.g. an offshore wind turbine engine, localised inspection of each of its parts is rather infeasible to make. Therefore, monitoring its modal parameters is one solution towards DI.

Moreover, they are particularly useful measures to be used beyond the 1st level in the damage identification hierarchy, i.e. can be used for both localisation and quantification purposes (see Figure 1.2). In the literature, it is very common to use NFs and MS and their variants, e.g. displacement of modal curvatures, as DSFs (see recent examples in [122], [87] and [134]). MS and curvatures, for instance, have much higher sensitivity to localised events of damage, but, the structure needs to be instrumented with a dense array of sensors for their estimation [153]. On the other hand, both NFs and DRs can provide a direct measure of the structure's health state using a single vibration transducer [35], i.e. for *localised DI*. The NF tends to *decrease* as damage progresses, while DR *increases*. As mentioned in Chapter 1 and also as described more in-depth in [67], in order to enhance damage sensitivity, it is necessary to look for changes in the higher NFs, since damage is a

local phenomenon.

In comparison to NFs and MS, DRs have not been investigated as extensively as DSFs in SHM applications. This is mainly due to the difficulties in estimating them. Some recent studies, e.g. [110], [39] and [128], have studied the DI problem for SHM using estimated DRs from vibration responses. Although the concept of using DRs for SHM problems is not new, many issues still remain. For instance, estimation of DRs in the presence of noise and non-linearity, is both unstable and inaccurate. For instance, in [141] the authors of the Z24 bridge experiment found high uncertainties in DR estimation using the stochastic subspace identification method. This finding was also reported in other studies, e.g. in [111], where the authors studied the variation of DR estimates on a structure subjected to wind-induced structural vibrations and compared it with NF estimation.

However, in structures where a crack is formed in its early stages, a considerable amount of vibration energy will be dissipated (mostly as frictional dissipation mechanism in the form of heat), and thus, the DR will increase proportionally, too. The sensitivity to damage in such subtle changes to the geometrical properties of the structure are expected to have minimal influence to its NFs (as reported, for instance, in [40]). Since DR is also correlated with damage, it will be examined as a candidate DSF.

2.4.2 Modal parameters identification via adaptive signal decomposition

The modal parameters (NFs and DRs) estimation or *identification* (MPI), as it is normally referred to, may be approached either in time, frequency or time-frequency domains. Typically, in MPI, the system is excited by an impulse force (e.g. impact hammer test), which results in the corresponding *impulse vibration response* [143]. For these applications, Frequency Response Functions (FRFs) are used for MPI. In practise, for SHM & CM the system will, in most part, be excited by ambient forcing, e.g. wind-induced vibrations in a high-rise building or traffic loading in a bridge. Therefore, *output-only methods* will be discussed, as more suitable MPI tools for SHM & CM. Table 2.4 groups some of the available output-only MPI techniques according to their domain, and summarizes each group's main merits.

Several damage scenarios that occur in practise, including opening and closing cracks, will force a structure into a non-linear dynamic region [67]. In these cases, techniques like ARMA and peak picking may not be applicable for MPI. One of the first approaches for MPI, applicable to systems that are not necessarily linear or stationary, was achieved using the WT along with the HT, e.g. in [100], for computing the time-varying nature of DRs and NFs. Most notably in [192], the authors introduced a methodology for MPI using the

Table 2.4. Some output-only MPI methods for estimating DRs and NFs.

Domain	*Techniques*	*Main merits*
Time	ARMA models, random decrement technique (RDT), stochastic subspace identification (SSI), proper orthogonal decomposition.	Limited to lightly damped systems with well-defined spacing between the modes. Applied to linear and stationary systems (with time-invariant modal properties) and are prone to errors due to noise in the measurements.
Frequency	Peak-picking, Frequency domain decomposition, Half-power bandwidth.	Same as above.
Time-frequency	Using adaptive signal decomposition methods along with an IA and IF estimation method, e.g. the HT (see Section 2.3.2).	Applicable to non-linear and non-stationary processes, but, more complex to implement, e.g. HT is prone to end-effects and high fluctuations in IF estimation due to differentiation of its instantaneous phase angle.

EMD algorithm, along with the HT and the RDT[¶]. This methodology was applied for SHM applications. Other studies, including the one in [42], have used a very similar approach for DI in bridge. More recently, VMD-based MPI was also employed for structures in [197] and [26], having a priori knowledge of the expected number of vibration modes.

Note that, for a single DOF linear, lightly-damped and undamaged system without EOV influence, its NFs and DRs will largely remain constant over time [34]. On the other hand, a non-linear system (like the Duffing oscillator - see Appendix A) will exhibit amplitude-dependent modal properties, as already discussed. Therefore, it is important to utilize these type of methods for MPI in systems that exhibit any degree of non-linearity and non-stationarity. This is true for systems in both damaged and undamaged states.

In general, the basic premise for MPI using adaptive decomposition techniques, like the VMD and EMD, is to decompose the vibration acceleration signal $x(t)$ into a set of M modal vibration responses and K IMFs, including the final residual $\varepsilon_K(t)$,

$$x(t) = \sum_{m=1}^{M} \tilde{x}_m(t) + \sum_{k=1}^{|K-M|} c_k(t) + \varepsilon_K(t) \tag{2.19}$$

[¶]RDT is a technique that is used to obtain the free decay vibration response from a system that is excited by ambient vibration forces.

As can be shown, for a multi-DOF system the analytical expression for its ith *free decay modal vibration response* (see Appendix A and Equation A.5) is centred around a single NF. This is inline with the definition of a mono-component signal, which is a requirement for HT analysis. To obtain the free decay vibration response of the system, the natural excitation technique (NExT) [38] is used in this thesis, due to the fact that it produced more stable MPI results than the RDT. The NExT Matlab implementation is available on the GitHub repository in [7].

In order to determine the NFs and DRs for a system using any of the above algorithms (EMD-/VMD-based or ALIF) the following procedure is summarized, as shown in Table 2.5.

Table 2.5. General methodology followed for MPI.

1. Decompose the vibration signal using an adaptive technique, e.g. VMD, into its $K+M$ constituent signals (M modal vibration responses and K IMFs). If M is required by the algorithm (and/or known beforehand), as in VMD, then $K=0$.
2. Obtain the $M+K$ free decay vibration responses, by applying NExT on all of the decomposed $K+M$ signals. Let's denote $x_i^f(t)$, with $i=1,...,M+K$ being the index of the decomposed signals.
3. Obtain the analytical signal of $x_i^f(t)$: $z(t)=x_i^f(t)+j\mathcal{H}\left[x_i^f(t)\right]$, where $\mathcal{H}$ is the HT operator.
4. It can be shown that the analytical signal expression for $x_i^f(t)$, from Step(3), is *equivalent* to the expression for the ith mode of vibration (see Equation A.5 in Appendix A), which is in the form of a real-valued exponential, $A(t)e^{-\zeta\omega t}cos(\omega t)$. Equating these two expression, and allowing only for light damping, i.e. $\zeta \ll 1$, the *free decaying amplitude* is given as: $\ln|z(t)|=-\zeta\omega_n t+z_0$, where, ω_n is the NF, ζ the DR and z_0 the non-oscillatory part of the response (i.e. a constant).
   Computing the *slope* of $\ln|z(t)|$ versus t, i.e. m_z, enables the calculation of DR as $\zeta=\frac{1}{\sqrt{1+(\frac{\omega_d}{m_z})^2}}$.
   While, the phase angle is given by the following expression: $\phi(t)\approx\omega_d t$. So that the *slope* of $\phi(t)$ versus t yields the damped NF ω_d.
   For both DR and NF a *robust* linear least-squares fit (see Chapter 5) is used for slope estimation, in order to reduce the influence of outlying data points on its computation.

Note that Step(4) is repeated for all $M+K$ decomposed signals to obtain all DRs and NFs. Apart from VMD, some vibration modes may be shared among any of the $M+K$ decomposed signals, since M is unknown (for instance, in ICEEMDAN). Therefore, post-processing of MPI is required to keep only *unique vibration modes*. Critical to the above implementation for the stability of DRs and NFs estimates, is the *segment length* of free

decay response, which is required by the NExT algorithm. An estimate of the *average segment length* is found by first calculating the *absolute residual* between the actual NF and the estimated one from the MPI method above. Then, this absolute residual is used as an *objective function* to a general optimisation procedure, i.e. particle swarm[‖], in order to find the optimal segment length that minimizes it. Note that the average segment length was computed on several different segments of the N-dimensional time series vector $\mathbf{x}$.

An important note regarding the use of HT. As it has been reported numerous times in the literature before, for instance, in N. Huang's very insightful paper in [90] and in [91], HT has certain inherent limitations. These include its "end-effects" and the Bedrosian and Nuttall theorem. The former may be solved by removing a certain part of the edges of the signal, i.e. after the HT is applied, as it was done in this thesis. The latter limitation imposes a significant constraint, i.e. the frequency spectra of the amplitude-modulated part of the signal $|z(t)|$ and the frequency-modulated part $e^{j\phi(t)}$ must not overlap, otherwise, negative frequencies are expected in IF estimation. Hence, as alternatives the following algorithms can be used instead [70]: the energy separation algorithm using the Teager energy operator, the normalised Hilbert Transform, the generalised zero-crossing, the empirical amplitude- and frequency-modulation decomposition and the direct quadrature.

2.4.3 CM in REBs

In CM for REBs, it is necessary to distinguish between REB vibrations, vibration from other machine elements within the same drivetrain and noise. The latter, may represent random, uncorrelated source of vibration with a white noise profile (or iid). While, other sources of vibration will have their own distinct characteristic frequencies. Distinguishing between these three vibration sources, will allow successful identification of damage in REBs (presented as the BSS problem in Chapter 1).

Concerning the outer race damage in the form of spalls, the characteristic frequencies generated with each impact can be *estimated*, as shown in one of the earliest works in [126]. Theoretically, the ball passing frequency of the outer race (BPFO) is given by the following kinematic relationship:

$$BPFO = \frac{n_b f_{shaft}}{2}\left(1 - \frac{d}{D} cos\alpha\right) \tag{2.20}$$

where, n_b is the number of rolling balls, α the bearing contact angle between the ball and outer race, and d and D are the ball and pitch diameters, respectively. Hence, this characteristic frequency is said to be *non-synchronous* to the fundamental of the drivetrain f_{shaft}. At the BPFO, the acceleration transducer will record a *sharp impulse*, each time a rolling ball hits the spall. This impulse, being of significantly high frequency, is capable of exciting

[‖]Particle Swarm optimisation is part of Matlab's Global Optimisation Toolbox.

a broadband range of resonances (up to $100kHz$ as shown in [130]) across the transfer path, i.e. from the damage location to the acceleration transducer. Sometimes, acceleration transducers will be tuned such that their resonances are within this excitation frequency range, in order to increase their sensitivity to the generated impulses [33]. Therefore, determining the frequency of the impulses, provides diagnostic information and DI capability. As also discussed in Chapter 1, reference values are available to investigate the damage severity for each component.

Although there are characteristic frequencies that will point to the specific damage in a drivetrain, those may be *hidden*, e.g. when computing the PSD of the *raw* vibration signal. Therefore, a certain amount of pre-processing needs be applied on the raw signal, by taking into consideration the following:

- The impulses generated at the early damage stages, are of low energy, in comparison to the discrete frequency components generated by other machine elements, e.g. a rotating shaft and its harmonics, as well as, gear meshes.
- In the kinematic relationship (Equation 2.20), the contact angle α varies by about $1-2\%$ in reality [149]. This is called *roller slippage*, and will change the BPFO, resulting in spectral leakage. Therefore, it will weaken the diagnostic signal further, at least for the highest frequency harmonics [50].
- The vibration response will contain amplitude-modulations by the excited resonances.
- White noise will be generated by other sources, e.g. instrumentation.

Modulations in amplitude due to excited resonances, will cause "distortions" in the frequency spectrum, by forming sidebands. This is beneficial for DI purposes, since the structural resonances generated by each impact, and modulated in amplitude at each BPFO, will be *amplified* by the impulse component of the signal. As illustrated from the simple schematic in Figure 2.2, the sharp impulses generated with a period of $T_{bpfo} = BPFO^{-1}$, correspond to the modulating signal, while structural resonances with a period of $T_{\omega b}$ correspond to the carrier signal, i.e. "carrying" the impulse signal. In general, when the amplitude or frequency of a signal is non-constant and varies as a function of time, it is called a modulated signal. In the first case, it is called amplitude modulation, while in the second case it is frequency modulation.

In the simplest case, where there is a single sinusoidal modulating signal in the form of $A_m sin(2\pi f_m t)$, with frequency f_m and amplitude A_m, and a single carrier signal in the form of $A_c sin(2\pi f_c t)$, with frequency f_c ($f_c \gg f_m$) and amplitude A_c, their combined signal

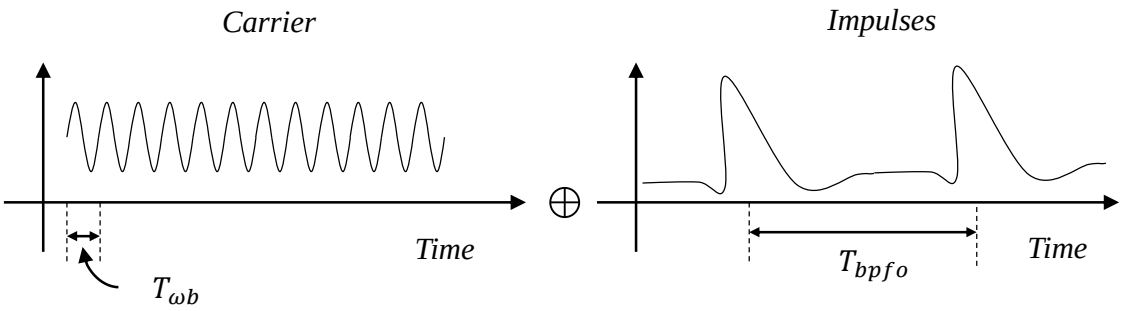

Figure 2.2. Simplified diagram of the modulated acceleration signal of a faulty bearing (showing the carrier and impulse signals separately).

$x(t)$ can be shown (using trigonometric identity of two sinusoids) to comprise the following three frequency components:

$$\begin{aligned} x(t) &= [A_c + A_m sin\,(2\pi f_m t)]\, sin\,(2\pi f_c t) \\ \therefore x(t) &= A_c sin\,(2\pi f_c t) + \frac{A_m}{2} sin\,(2\pi(f_c + f_m)t) + \frac{A_m}{2} sin\,(2\pi(f_c - f_m)t) \end{aligned} \tag{2.21}$$

A simplified example of a theoretical signal with a single carrier frequency $f_c = 500Hz$ and a single modulating frequency $f_m = 10Hz$ is shown in Figure 2.3. In the frequency domain, the two components around f_c (the sidebands) are spaced by $f_m = 10Hz$, apart. In an outer race damage, the *frequency spacing* will be the BPFO, while the red line in the time-domain is the *signal envelope* with period T_{bpfo}.

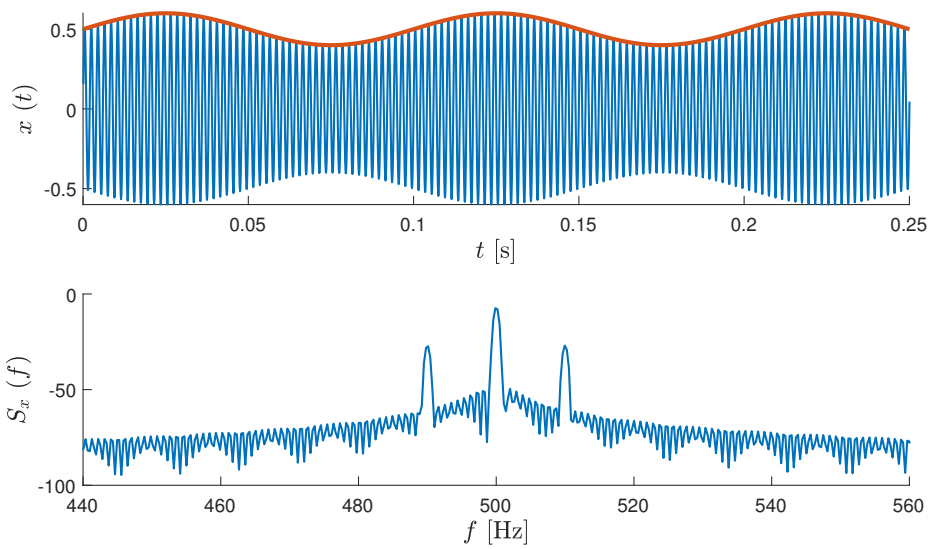

Figure 2.3. Amplitude-modulations in a theoretical signal with carrier frequency of $500Hz$ and modulating frequency of $10Hz$.

2.4.4 DSFs for REBs using envelope analysis

In order to solve the problem of DI for BPFO damage, several methods were proposed in the literature. Among those methods, Cepstrum and EA [102] are well-known methodologies, each one having its own distinct characteristics and application domains. In brief, Cepstrum analysis aims to reveal *periodic components* in the frequency spectrum through calculating the spectrum of the logarithm of the energy spectrum. On the other hand, there are certain advantages associated with EA, such as the flexibility to incorporate different algorithms into its framework and its proven/robust effectiveness as a *demodulation tool*. Additionally, as the REB signal is of cyclo-stationary type, there have been developed and applied many other techniques to tackle this characteristic specifically. One very promising method for this application is the *spectral correlation density function* [24].
Although EA was developed in the 1970s by the authors in [54], it has been the *benchmark technique* for most of the CM research studies for REBs [140]. Its main aim was to overcome the limitations of poor frequency resolution by shifting the analysis of high-frequency carrier resonances towards a lower-frequency range [126].

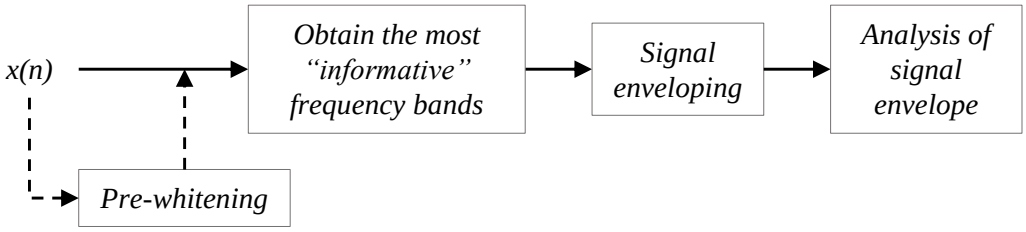

Figure 2.4. The four main steps of the envelope analysis procedure.

The main idea behind EA, as applied to DI on the outer-race, is to isolate the amplitudes at BPFO, i.e. the impulsive nature of the signal, which is the *signal envelope*, from the resonances (i.e. the carrier). As it was shown in Figure 2.3, the signal envelope** (shown as red line) represents the *extrema* of the signal in the time-domain. The first step towards obtaining the envelope of the signal is to bandpass filter it around the frequency of one of the excited structural resonances f_c. Figure 2.4 shows an overview of the processes involved in EA, as follows directly from the discrete-time raw acceleration signal $x(n)$. Each step in EA is described as follows:

1. **Pre-whitening**: This is a pre-processing step, which becomes important when other elements are present in the drivetrain, e.g. gears. The vibration energy from REBs at any damage frequency (not only at BPFO) will be weak, and thus, the generated

**Mathematically, signal envelope is defined as the square root of the squared envelope of the signal.

impulses will be masked by other more dominant discrete frequency sources. This is true especially in the very early stages of damage progression when the vibration signal at BPFO is close to the noise floor. As mentioned previously, there will also be some variation in the contact between the mating elements of the bearing, also called roller slippage. Although small, and usually up to about $1-2\%$ as has been observed in [149], it will cause random fluctuations in the time-period between each generated impulse. Coupled with cage clearances, the signal generated by the damage will be (approximately) 2nd order cyclo-stationary signal or "pseudo-cyclo-stationary". Pre-whitening will be required to separate the *random part of the signal* (REB impulses) from the discrete part of the signal (gear meshing frequencies and harmonics, etc.). There are several techniques to do that, as described in the excellent review paper in [148]. However, some of the most effective ones for an *automated DI*, is pre-whitening by means of obtaining the *residual signal* from the estimation of the *AR model* of the signal. A study by NREL[††] [164] found that pre-whitening enhances the signal's sensitivity of the acceleration step response when a rolling ball enters a spall. It should be noted here that the range of order p of the AR model used to pre-whiten the signal should be chosen carefully. That is, to *avoid inclusion of the impulses* as part of the deterministic signal and to enhance the impulsiveness of the residual signal. For the first requirement, the maximum order p_{max} should be calculated as:

$$p_{max} = \min_{1 \leq i \leq N_{imp}} \left[\frac{1}{BPFO} \right]_i f_s - c \tag{2.22}$$

where N_{imp} is the total number of impulses in the signal and c is a scalar, usually less than 10 (determined from numerical simulations). So that, p_{max} is computed by finding the *minimum time period between each impulse* from all N_{imp} impulses.
Given the above range, of possible AR orders $p \in [1, p_{max}]$, the second requirement is satisfied by choosing the *optimal order* (among all p_{max} orders) such that the *kurtosis value* is at its maximum:

$$p_{optimal} = \max_{1 \leq i < p_{max}} \left[\frac{N^{-1} \sum_{n=1}^{N} (x_n - \bar{x})^4}{(N^{-1} \sum_{n=1}^{N} (x_n - \bar{x})^2)^2} \right]_i \tag{2.23}$$

Therefore, the residual signal from the AR model (see Equation 2.6) may now be computed by fitting an AR model of order $p_{optimal}$ to $x(n)$, first.

2. **Obtaining the most "informative" frequency bands**: Although pre-whitening is a first step in enhancing the damage signal, the key step in EA is to identify the

[††]National Renewable Energy Laboratory

frequency band that BPFO is most dominant in the spectrum. One of the earliest approaches [126] is to bandpass filter the signal around one of the impulse resonant frequencies, which amplifies the BPFO (or any bearing damage signal) and its harmonic series. This results in filtering out both the low frequencies (gear mesh, shaft rotations, etc.) and the higher frequencies (where white noise levels are high), within the whole Nyquist range. In that way, it further enhances the SNR.

Many researchers have tested different combinations for the lower- and upper-pass cut-off frequencies for the FIR filter. Most notable is the recommendation by [33], which mentions to obtain a frequency band that corresponds to the biggest dB difference. This means that a reference signal of the undamaged bearing is required, which will not necessarily be available. On the other hand, the author in [88] recommends "paying attention" to higher frequency bands where the energy is high, and specifies from experience that it is common to filter-out frequencies below 2.5 kHz and above 5 kHz. In addition, an impact hammer test on the bearing case is also common practise in order to identify the structural resonances and then specify the band for the filter.

The application of SK (see Equation 2.8), however, has been seen as a more rigorous way to parametrize the bandpass FIR filter, which also enables an automated DI to be carried out. Using SK, enables the wide applicability of the EA method on systems of varying characteristics, e.g. different REB geometric features and shaft speeds. More specifically, SK has given rise to the *kurtogram*, which helps in identifying the bandpass bandwidth automatically as the one that *maximizes* it.

Moreover, adaptive signal decomposition may be performed to investigate the different frequency bands of the system in isolation and identify the ones with the highest damage indicator, e.g. highest kurtosis. This will be explained in Chapter 4, but, not in the context of EA.

3. **Signal enveloping**: The next step is to obtain the envelope of the bandpass filtered signal. As it was shown, using the HT the analytic representation of the signal can be obtained. This in turn, can be used to get IA, which is the signal's envelope. However, it is important that the identified bandpass frequency, computed in the previous step, is sufficiently narrow-band, as per the requirements for HT analysis.

4. **Analysis of signal envelope**: The final step in EA, is to get the frequency spectrum representation, e.g. using Welch's PSD, of the signal's envelope in order to obtain a "clean" version of the amplitude at $BPFO$ and its harmonics. This step is also called the envelope spectrum.

In Appendix E, the EA procedure, as presented above, is studied using a numerical simulation dataset (representing outer-race damage on an REB) and its corresponding envelope spectrum is also discussed in Chapter 4. Note that, although the amplitude of each impulse may progress as a function of time, e.g. due to bearing deterioration, it will in theory, remain constant within short time intervals. This is because outer-race damage exists within a specific "load zone" of the bearing. The load zone can be defined as the normal or radial load applied on the bearing due to the shaft's mass. As such, the acceleration transducer will record the same magnitude each time a ball strikes the spall. For that reason, there are no shaft rotational-speed amplitude modulations, as in the case of diagnosing inner-race damage. In the latter, due to being rotated with the shaft, there will be sidebands around the ball passing inner race frequency, which are repeated in its harmonics as can be shown in a typical power spectrum, e.g. in [28]. Diagnosis in outer-race damage is generally considered more straightforward than in inner-race, due to the former being located within a fixed (or nearly fixed) load zone. What is most important, however, as far as bearing defects are concerned, is to be able to extract the impulsivity of the raw vibration signal. Using EA is an approach to achieve this.

2.4.5 Vibration DSFs in GT engines

Typically, in vibration-based CM on GT engines and variable-speed rotating machinery, the set of DSFs that is mostly used for automated DI are the *tracked-orders* [101], [45]. This is because, with any of the above mentioned vibration analysis methods, e.g. adaptive signal decomposition or PSD, vibrations will be a function of the speed the system is rotating, i.e. its shaft. Such variation in rotating speed will also change the vibration characteristics of a GT engine, e.g. the fundamental frequency of the main shaft and its harmonic series. Therefore, by *measuring the rotational speed* of the system using tachometer probes, it is possible to *normalise* the vibration characteristics of the GT engine. This enables to *track* each frequency component *regardless* of the rotational speed. In GT engines, the first four (fundamental and its first three harmonic series) are used as DSFs due to the fact that they contain the most energy [176]. In this thesis, the rotational speed of the GT engine is constant or there is no rotating part, e.g. in a combustor, so that vibration characteristics are dependent only on the operating conditions (AFRs and SAFs).

2.5 Summary

In this chapter the types of vibration signals and relevant analysis methods were reviewed for the purpose of identifying a suitable set of analysis tools for different SHM & CM applications. More specifically, this chapter may be summarized using the following points:

- Vibration signals can be categorised according to their generated processes, which can either be stationary or non-stationary, linear or non-linear. For each group, a number of techniques were discussed.

- Structural vibration signals cannot be assumed to be generated from stationary and linear processes, e.g. due to EOV influence. Also, a common non-linearity that occurs in structures is that of amplitude-dependency. For these reasons, adaptive signal decomposition techniques, are preferred as analysis tools for these types of signals.

- A very sensitive measure, which is correlated with impacts, is that of kurtosis and its variations, e.g. SK. It can be used whenever damage manifests itself as non-linearities in the system in the form of impulses, e.g. on an outer-race damage in REBs.

- A methodology has also been presented to *identify modal properties* of a structure, i.e. by first decomposing the signal into its vibration modes (plus IMFs) and then computing the instantaneous properties of the signal through the HT. In that way, a measure of the DRs and NFs can be provided for further analysis to examine their potential as DSFs.

- In DI on the REB outer-race, it is crucial to *amplify* the relatively small vibration amplitude at the BPFO. This is indeed necessary, because in a broadband vibration spectrum that may include both high-amplitude noise and other periodic frequency components, the amplitudes at BPFO may be buried in the raw signal spectrum. EA, with the aid of the SK (to identify band-pass filter cut-off frequencies), the HT (to obtain the signal envelope) and an AR model (to remove periodicities in the signal) seems as a promising technique.

Following the above review, a description of the available datasets will follow to further understand the specific problems this thesis is dealing with.

Chapter 3

Datasets description and exploration

This chapter provides a description of the datasets that are used in this thesis, in order to demonstrate the different DI methods proposed. The datasets include measurements from one or more accelerometers, attached on the system that is being monitored. Namely, the four datasets are:

1. Los Alamos four DOF laboratory experiment (LA-4DOF) (Section 3.1.2)
2. Wind turbine gearbox REB laboratory experiment (WT-REB) (Section 3.2.1)
3. Z24-bridge controlled field experiment (Z24) (Section 3.1.3)
4. Gas turbine engine laboratory experiment (TF-LBO, TS-SAF and TF-SAF) (Section 3.2.2)

Additionally, Section 3.1.1 presents a numerical simulation model of a lumped parameter mass-spring-damper system with *linear* and *non-linear* restoring forces, denoted as NSim-4DOFLin and NSim-4DOFNonLin, respectively. The problem of EOV influence on system dynamics will also be clearly illustrated, following this exploratory analysis.

3.1 SHM on multi-storey structures and bridges

3.1.1 The 4 DOF numerical simulation model

A bookshelf-type structure is one of the most typical experimental arrangements, used in numerous SHM studies, e.g. in [105]. This is mainly because it can represent a wide range of structures of interest, for instance, multi-storey buildings. Given the popularity of the bookshelf-type structure, a similar numerical simulation model was developed. Its basic arrangement is described in Appendix A.

3.1.1.1 NSim-4DOFLin

In expanding the 4-DOF model shown in Appendix A, the influence of an *actual air temperature profile* is added. This in turn, causes it to become a non-stationary generating process. In addition, a *damage model* is added, which significantly alters its dynamic properties:

1. As discussed in Chapter 1, the modal properties of the system are mostly *dependent* on temperature variations and gradients. This implies that the relationship between the nominal spring stiffnesses $\bar{\mathbf{k}}$, and the thermal field $T(t)$, can be modelled as:

$$\mathbf{k}(t) = \bar{\mathbf{k}} - \beta(T)\ \bar{\mathbf{k}}\ T(t) \tag{3.1}$$

   where, β is a switching function between two values, β^- and β^+, i.e. it depends on whether $T(t)$ is below (switching to β^-) or above zero (switching to β^+). This distinction is necessary, given that from past experiments, e.g. on bridges, a notable increase in NFs has been observed when $T \leq 0$ [169]. Therefore, to simulate this scenario a much higher β^- as compared to β^+ is specified.
   Similar approaches for modelling the influence of temperature on the modal properties can be found in other recent SHM studies, for instance, in [165]. The raw ambient temperature measurements from the Z24 dataset are shown in Figure 3.1. Since the sampling time of those measurements is at 1-hour intervals, Gaussian white noise was added to the temperature measurements with a standard deviation of $0.001°C$ in order to add short-term variability (given the numerical simulation step size was in the time-scale of seconds).

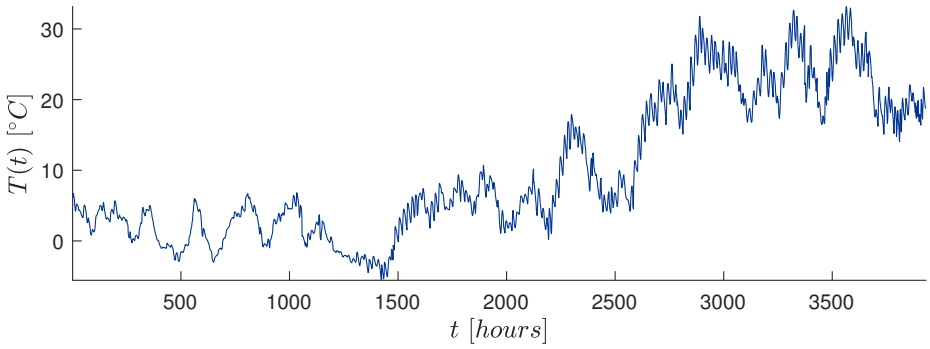

Figure 3.1. Air temperature measurements (see details of Z24 dataset).

2. Damage is modelled with a time-decaying function, of the form $\alpha_k(t) = e^{-\gamma_k(t)}$. This can be thought of as a time-varying ratio between the damaged and undamaged stiffness values. Hence, the rate of damage progression over time can be controlled by changing the value of γ_k. Additionally, *impacts* in the form of short-duration external force excitations, increasing with the same rate as $\alpha_k(t)$, were included.

Although the NSim-4DOFLin model has linear restoring forces, it will certainly be *non-linear* and *non-stationary*, due to the above two additional influences on its dynamics. An illustrative numerical simulation example is shown in Figure 3.2 where the DRs (ζ) show significant fluctuation around their nominal values on periods where $T(t) \leq 0$. When the progressive damage occurs, at $t = 1008\ hr$, the most profound change occurs on the DR corresponding to the second NF. As expected a similar pattern takes place on the NFs, too.

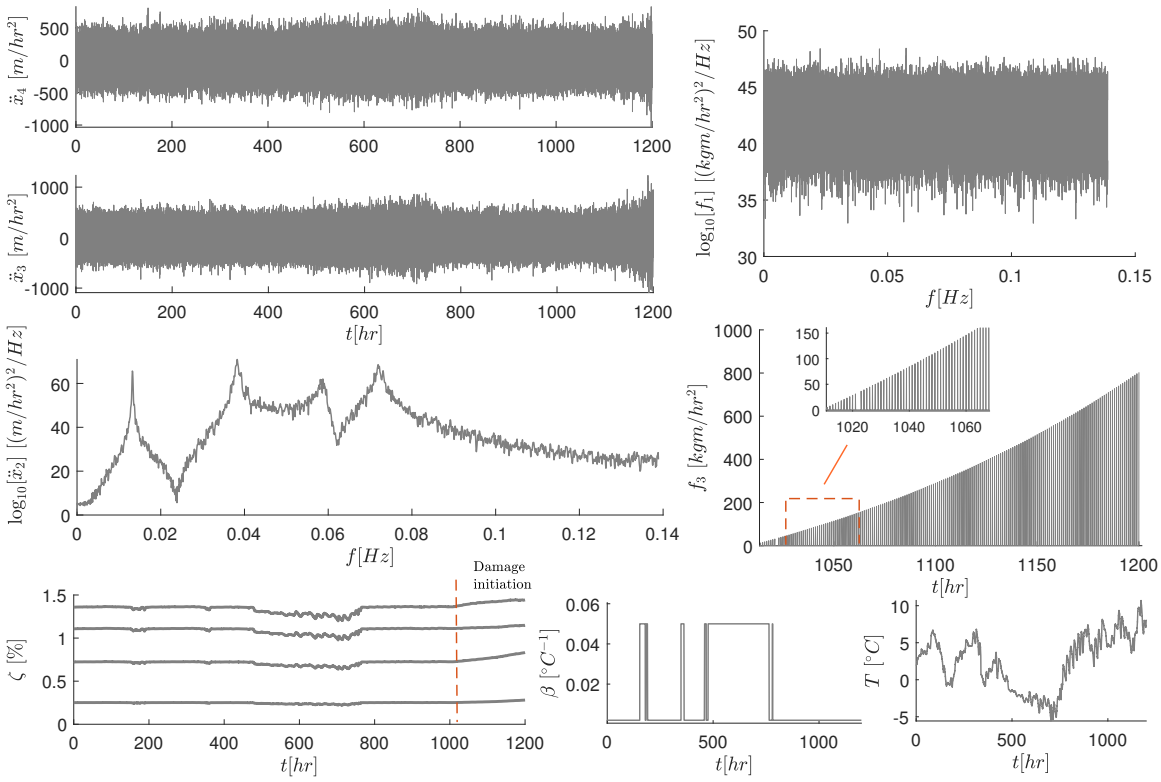

Figure 3.2. A numerical simulation of the 4 DOF model. Damage is progressing exponentially over 8 days (from $t = 1008hr$).

Note that damage is simulated as: an exponential stiffness decrease, alongside an external random impulse excitation force, $f_3(t)$, that grows at the same exponential rate. The damage is simulated between m_2 and m_3, i.e. on $k_3(t)$. Also, the time-histories of acceleration of m_3 and m_4 show a change in magnitude when damage occurs, but, it is *indistinguishable* from the changes that occur due to a *freezing temperature*. The PSD magnitude of the input force shows equally distributed excitation across the whole spectrum, i.e. from 0 to $.14Hz$, as expected from a white noise signal. The main numerical simulation parameters for this example can be found in Table 3.1.

Table 3.1. Numerical simulation example 1: main parameters of the four degree-of-freedom mass-spring-damper model.

Parameter	Value	
$\bar{\mathbf{k}}$	$\{6\times10^5\}$	$[kgm/hr^2]$
c_{1-4}	1×10^1	$[Nhr/m]$
m_{1-4}	0.8	$[kg]$
γ_k	5×10^{-3}	$[1/hr]$
β^+	0.002	$[^\circ C^{-1}]$
β^-	0.05	$[^\circ C^{-1}]$
$\{\sigma_{F_1},\mu_{F_1}\}$	$\{2.5\times10^2,0\}$	$[kgm/hr^2]$

3.1.1.2 NSim-4DOFNonLin

The above presented numerical simulation model provides a good starting point for conducting a basic investigation of DI using, for instance, modal parameters as DSFs, influenced by both EOV and progressive damage. In expanding the applicability and testing the effectiveness of DI methodologies it is necessary to account for *stronger non-linearities* in system dynamics, as well as a more *realistic damage scenario*.

As explained in [67] (pages 245-247), under certain environmental and operating conditions the dynamic response of a structure can transition from linear to non-linear, when damage is present. An illustrative example is a *bilinear spring stiffness* coefficient model that is used as a simplified approximation for breathing crack simulations in beams.

With this in mind, the spring force $\mathbf{f}_s(t)$, acting on each mass, can be modelled to include many non-linearities encountered in practical problems. Firstly, the equations of motion must be re-written (see Appendix A) as:

$$\begin{bmatrix}\ddot{\mathbf{x}}\\ \dot{\mathbf{x}}\end{bmatrix}=\begin{bmatrix}-M^{-1}C & 0\\ I & 0\end{bmatrix}\begin{bmatrix}\dot{\mathbf{x}}\\ \mathbf{x}\end{bmatrix}+\begin{bmatrix}M^{-1}\mathbf{f}_{s,L}(t)-M^{-1}\mathbf{f}_{s,R}(t)\\ 0\end{bmatrix}+\begin{bmatrix}M^{-1}\mathbf{f}\\ 0\end{bmatrix} \tag{3.2}$$

where, the spring forces on the left $\mathbf{f}_{s,L}(t)$ and on the right $\mathbf{f}_{s,R}(t)$ of each mass are calculated based on the vector of relative displacements $\Delta\mathbf{x}=\{-x_1(t),x_1(t)-x_2(t),x_2(t)-x_3(t),x_3(t)-x_4(t)\}$. Whereas, damping is modelled the same as in NSim-4DOFLin. Each spring force, either on the left or on the right of each mass was calculated according to the following piecewise function:

$$\mathbf{f}_s(t)=\begin{cases}\mu\mathbf{k}(t)\Delta\mathbf{x}+(\mathbf{k}(t)-\mu\mathbf{k}(t))\Delta x_*+\nu\mathbf{k}(t)\Delta\mathbf{x}^3, & \Delta\mathbf{x}\geq\Delta x_*\\ \mathbf{k}(t)\Delta\mathbf{x}+\nu\mathbf{k}(t)\Delta\mathbf{x}^3, & 0\leq\Delta\mathbf{x}<\Delta x_*\\ \boldsymbol{\alpha}_k(t)(\mathbf{k}(t)\Delta\mathbf{x}+\nu\mathbf{k}(t)\Delta\mathbf{x}^3), & -\Delta x_*<\Delta\mathbf{x}<0\\ \boldsymbol{\alpha}_k(t)(\mu\mathbf{k}(t)\Delta\mathbf{x}-(\mathbf{k}(t)-\mu\mathbf{k}(t))\Delta x_*+\nu\mathbf{k}(t)\Delta\mathbf{x}^3), & \Delta\mathbf{x}\leq\Delta x_*\end{cases} \tag{3.3}$$

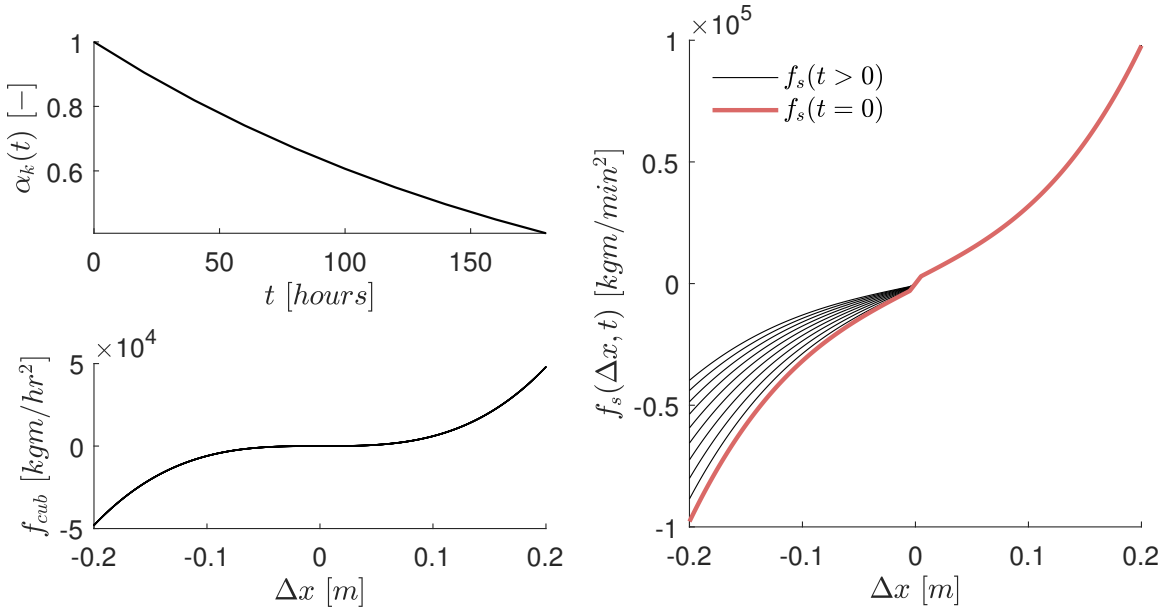

Figure 3.3. Non-linear spring force profile used for numerical simulations. Damage is progressing exponentially over 8 days.

where, $\boldsymbol{\alpha}_k(t)$ is applied only when $\Delta x(t) < 0$, which indicates when the crack opens, and although it is defined as a vector here (as a generalisation), it will usually be applied on a single $\Delta x(t)$, unless otherwise specified. Also, ν is the ratio between $\bar{\mathbf{k}}$ and cubic stiffness term, and similarly, μ is the ratio between $\bar{\mathbf{k}}$ and the non-linear region of the spring force. The displacement threshold, Δx_*, indicates the transition to the non-linear region. As shown in the following spring force profile example in Fig. 3.3, the cubic force term, f_{cub}, is more dominant at higher $|\Delta x|$, where the the non-linear region starts at .5 cm, while $\nu = 10$ and $\mu = .4$. Note the decrease in negative $f_s(t)$ value as time (and damage) progresses according to the time-decaying ratio α_k over a certain number of realisations for a time-period of 8 days. So the negative spring force resistance decreases almost by half towards the end of this simulation.

As a note here, if the spring force has been modelled with just the cubic term, the problem would have been equivalent to a Duffing *stochastic process* (i.e. by replacing the forcing function with white noise in Equation A.2 in Appendix A). In [84] the authors provided a solution to modelling this type of system with a TAR (amplitude-dependent time-series) model. However, the presented model is more involved than the single DOF Duffing stochastic process.

The results from a typical numerical simulation scenario are shown in Figure 3.4. Similar to the previous numerical simulation example, damage is considered as a progressive stiffness degradation between m_2 and m_3, in addition to the external impulse forces acting on m_3.

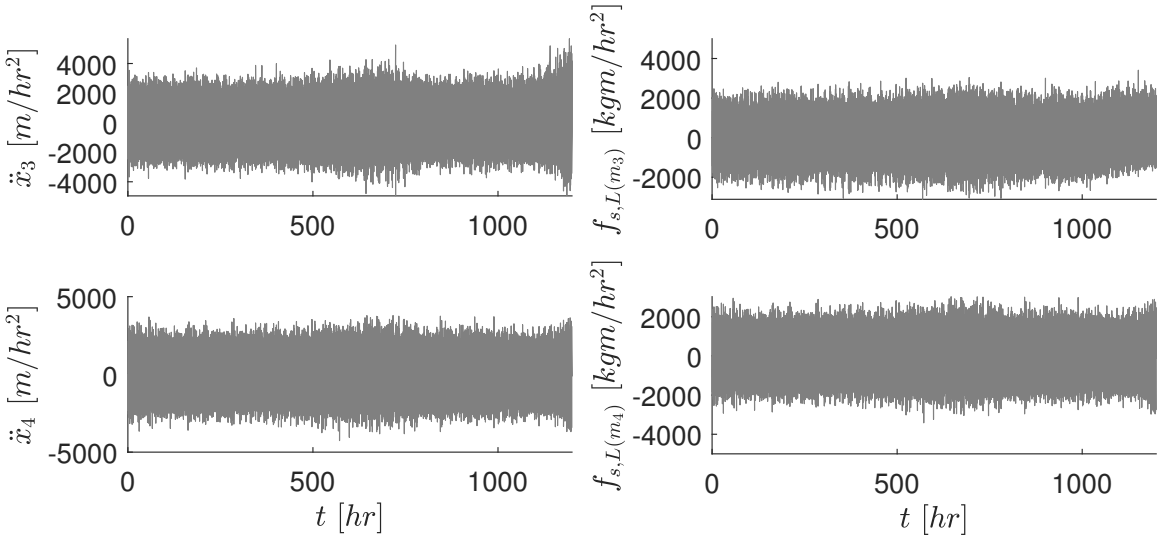

Figure 3.4. A numerical simulation of the 4 DOF model with non-linear spring forces in undamaged and damaged states having the same simulation parameters as in previous simulation.

Unless otherwise specified, the parameters of the simulation are the same as in the previous example. As can be seen, the left spring force applied on mass 4, $f_{s,L(m_3)}$ decreases according to the damage ratio, α_k, as shown previously in Figure 3.3. In turn, this influences the acceleration amplitude of m_3, although this may not be easily observable on the plot here. On the other hand, there is no significant visual change in m_4 acceleration, instead it follows the undamaged state very closely. Note the reduction in spring force (and acceleration) due to damage is only on the negative region. The influence of temperature on acceleration time-histories is also evident from the plots.

The equations of motion in both numerical simulation examples presented above were solved with a fixed time-step $h = 1 \times 10^{-4}$ hr, as a compromise between stability of numerical integration solutions and computational speed. The fixed step solution is required, for instance, when an AR time-series model is used, which assumes equally-spaced time-intervals between the discrete observations. The 5th order explicit Runge-Kutta ordinary differential equation solver [37] was employed to obtain the solution at each time-step. This was needed due to the numerical model's *higher stability* requirements. Note that the above two numerical simulations examples will be used as two case studies in investigating DI methodologies.

3.1.2 LA-4DOF

The numerical simulation model presented above, has a very similar arrangement to the three-storey benchmark laboratory experiment contacted by the Los Alamos National Laboratory. A detailed description of this experiment can be found in the technical report in

[72]. A notable difference of this experiment, is the way damage and EOV are being *simulated*. Here the authors considered stiffness drop on one or more of the support columns (as EOV influence), while damage is considered as impacts (only) of *varying amplitudes*. On the other hand, in NSim-4DOFNonLin and NSim-4DOFLin models, damage is simulated both as a stiffness decrease *and* force impulses.

The experimental setup, uses an electromagnetic shaker device to exert band-limited white noise forces, ranging from $20Hz$ to $150Hz$ in frequency, with RMS value of approximately $20N$, on the base floor of the structure. The sampling frequency was set to $322.6Hz$ for both accelerometers and force transducer. The accelerometers are attached opposite to the shaker and in the middle of each horizontal floor. As the structure is supported on linear rails, only x-axis motion is considered, as in the NSim-4DOFLin and NSim-4DOFNonLin models. Each time-history from these sensors was recorded for 25.6 seconds, which results in a frequency resolution of $0.0391Hz$.

In order to simulate the effects of temperature on the system dynamics, the stiffness coefficient of one or more of the support columns was decreased by 50%. The effects of traffic are also simulated by adding a specific mass of $1.2kg$ on different floors. On the other hand, damage is considered as an opening and closing crack and simulated by a suspended column (on the 3rd floor) that is impacting a bumper. The gap between the bumper and the suspended column varies in length to simulate damage progression.

As shown in the Table 3.2, the seventeen different system structural states recorded can be separated into four main groups. In the first group, the system is in its undamaged state without EOV influence. In the second group, the system is undamaged but influenced by temperature and added mass, while in the third group the system is damaged at various progressions (without EOV influence). Finally, in the fourth group the system is damaged at various progressions with extra mass added on different floors. Note that in any given state $s1 - s17$ there are 9 examples measured at different times. Thus, for each of the five sensor measurements there are 81920 sample points in each state.

For the purposes of this thesis, measurements recorded from the accelerometer located on the top floor, i.e. furthest away from the base excitation force, will be used. It has been shown in previous studies on LA-4DOF, e.g. in [74] and in [72], that damage-sensitivity decreases significantly when considering measurements from the two bottom floors, due to those sensors being located furthest away from the impact source. Figure 3.5 shows the overlapped PSDs for $s1$ state, as calculated on different realisations of this experiment using a time-interval of 10 seconds for each one. On each realisation, the peaks show three modes of vibration, at $\approx \{30.5, 53.7, 70.8\}Hz$, which were consistent with the numerical results from [72]. Due to the excitation signal being band-pass filtered, by the authors of

Table 3.2. Structural system states and data groupings for Los Alamos bookshelf frame laboratory experiment.

State #	Description	Group
$s1$	Undamaged.	1
$s2$	Undamaged with added mass on the base floor.	2
$s3$	Undamaged with added mass on the 1st floor.	
$s4$	Undamaged with column 1BD stiffness decrease.	
$s5$	Undamaged with columns 1AD + 1BD stiffness decrease.	
$s6$	Undamaged with column 3BD stiffness decrease.	
$s7$	Undamaged with column 3AD + 3BD stiffness decrease.	
$s8$	Undamaged with columns 2AD + 2BD stiffness decrease.	
$s9$	Undamaged with columns 2BD stiffness decrease.	
$s10$	Damaged with progression level 1 (least severe).	3
$s11$	Damaged with progression level 2.	
$s12$	Damaged with progression level 3.	
$s13$	Damaged with progression level 4.	
$s14$	Damaged with progression level 5 (most severe).	
$s15$	Damaged with progression level 1 + added mass on the base floor.	4
$s16$	Damaged with progression level 1 + added mass on the 1st floor.	
$s17$	Damaged with progression level 4 + added mass on the 1st floor.	

this experiment, the lowest vibration mode is not available. As it was shown in Figure 3.2, the first NF provides the lowest sensitivity of damage as compared to the other three NFs, and therefore, can be neglected. Additionally, the frequency spectrum of the excitation force signal shows to be relatively flat over the frequency range of interest, apart from the amplitude drops at frequencies of the three vibration modes (due to the dynamic coupling effects).

The sensitivity of the three modes of vibration due to EOV and damage levels is illustrated in Figure 3.6. In particular, the PSDs show a shift in frequency of the 3rd and 4th modes with EOV when the system is undamaged. On the other hand, the impact-type of damage causes an *energy transfer* from the three vibration modes towards higher frequencies. In other words, an amplitude drop in each of the peaks can be seen, with an increase in amplitude on higher spectral components. This observation becomes more pronounced as the damage progresses, i.e. to the highest level $s14$. Now, by monitoring such changes, e.g. using a kernel density estimate (KDE) as in Figure 3.6, one can observe that as the damage progresses, this type of damage skews the PDF. Therefore, deviations from the Gaussian, e.g. central moments, may be used for DI. This is also a sign of inducing *non-linearity* in

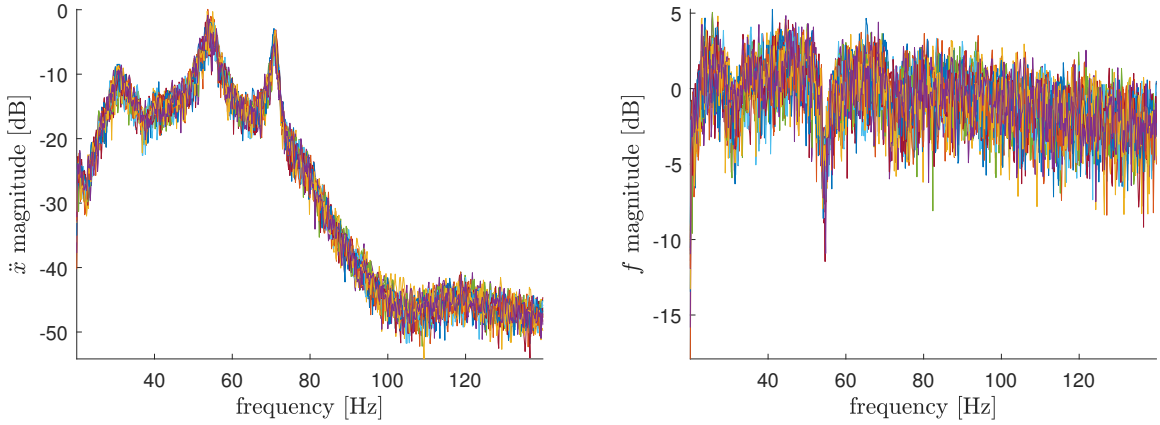

Figure 3.5. Power spectral densities of acceleration (left) and force input (right) in $10s$ intervals for baseline ($s1$) system state.

system dynamics.

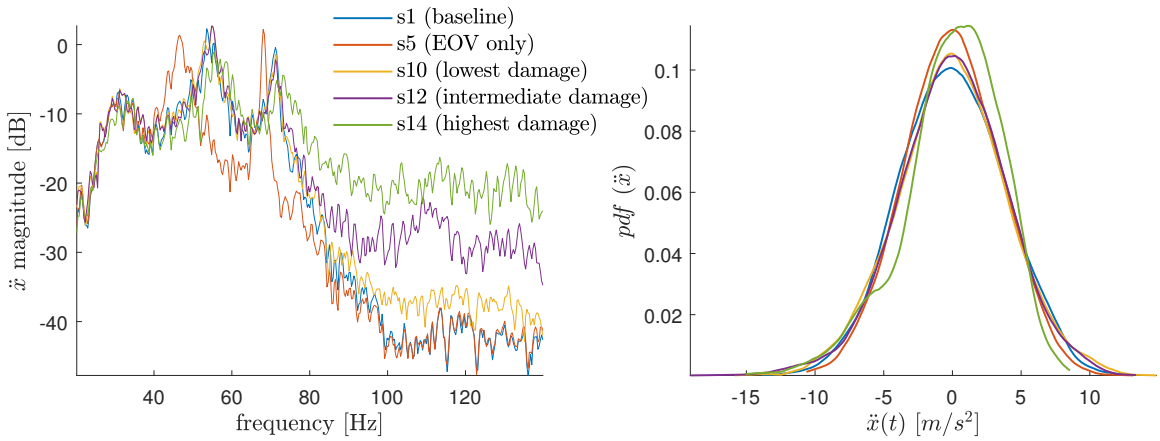

Figure 3.6. Power spectral densities (left) and kernel density estimate (right) for five different system states.

On the other hand, Gaussian assumptions may apply for EOV influence ($s5$), although the density is more peaked than $s1$. Thus, a higher kurtosis value than the baseline. As the impact-type damage introduces amplitude drop, mainly on the 3rd and 4th NFs, a spectral window surrounding those components, i.e. between $40Hz$ and $80Hz$ can also provide sensitive information for damage. At the same time, the frequency shifts due to EOV may also play an important role in the development of a reliable DI framework.

3.1.3 Z24

The Z24-bridge experimental campaign is another well-known benchmark study in SHM. The monitoring of this bridge took place between 11 November 1997 and 11 September 1998. The main aim of this project was to quantify the influence of both EOV and structural

damage on the dynamical characteristics of the bridge. From 10 August 1998 until the end of the monitoring period, i.e. a month, the bridge was progressively damaged. Within this monitoring time span, environmental parameters such as ambient air temperature, wind speed, humidity and soil temperature were measured with a sampling period of 1 hour, using 49 different sensors, e.g. thermocouples, at several locations along the bridge. It was found and confirmed (as previous studies have claimed) that temperature, in particular, had a profound influence on the stiffness characteristics of the bridge. This is mainly because it alters the *Young's modulus of concrete*, which in turn affects the modal properties of the structure. This change can be approximated with a bilinear model [152], similar to the expression used in the NSim-4DOFLin/-4DOFNonLin models in Equation 3.1. A detailed description of the Z24 experiment can be found in [142].

The bridge's dynamics were monitored using 16 accelerometers at different locations along the bridge and orientation axes. Every 1 hour, accelerations were measured with a sampling frequency of $100Hz$. From prior modal experiments on the structure, the first four modes of vibration were expected in the range of $0-12Hz$. An automated modal analysis procedure was performed based on the stochastic subspace identification algorithm, which is an output-only MPI algorithm, as discussed in Chapter 2. That is, the bridge was excited by natural sources only, which in this case it was vehicles passing along the bridge and wind-structure interactions.

In Figure 3.7 the first four NFs of the bridge are shown for the whole monitoring period. On the top part of the plot, the four NFs are shown as a function of time, where the marker colour changes with temperature, i.e. it is dark red when temperature dropped below $0°C$ and dark blue when above. The change in each NF as a function of temperature can be more readily observed on the bottom subplots in this figure. The bilinear relationship is evident from those plots. In particular, when it drops below $0°C$ the rate of change increases substantially, where NFs increase. When temperature is above the freezing point, a decrease of all four NFs is also observed, though this rate is much smaller than in freezing temperatures.

Note that there are only 3932 samples in the record, which does not correspond to the 7320 samples for the full 305 days of the monitoring period. This is because, there were periods in which the monitoring system stopped working and data couldn't be collected [142]. Note that the temperature profile that was shown previously in Figure 3.1, corresponds to the same time instances as in the NFs of Figure 3.7.

Damage is introduced in the bridge just after $t = 2496$ hours, in the form of pier settlements, initially, using a settlement system. On the latest stages of damage, several tendons were

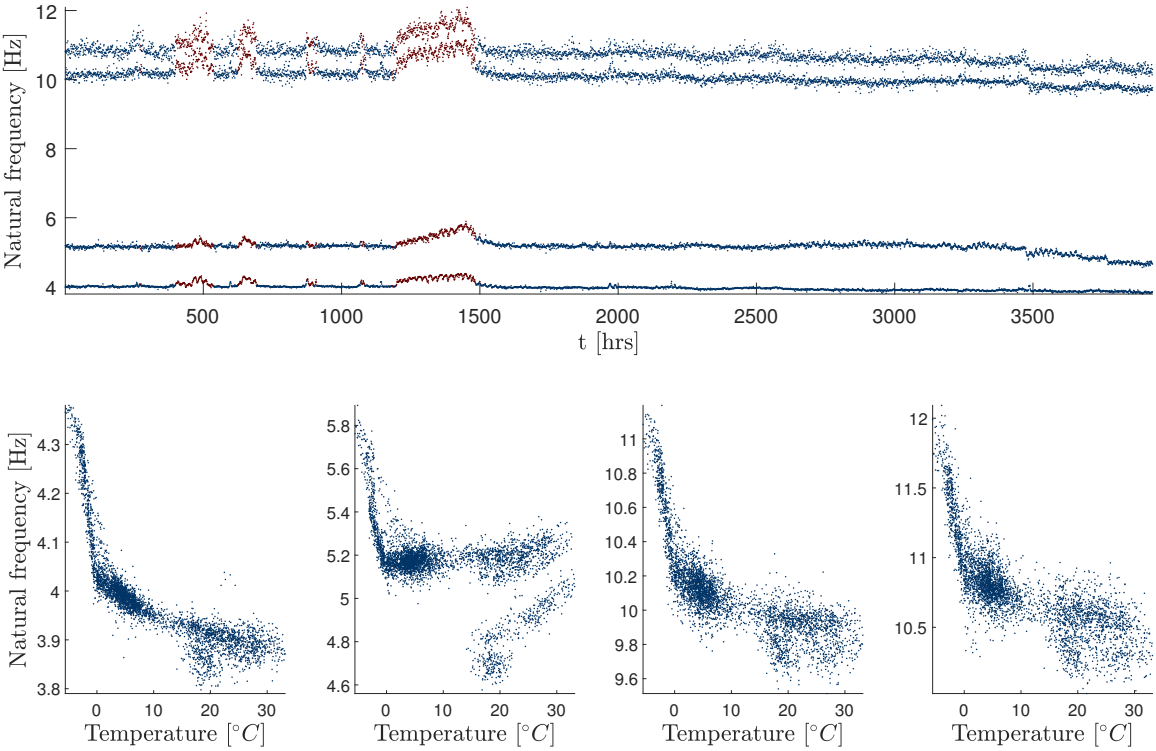

Figure 3.7. Natural frequencies as a function of time and ambient air temperature. Top plot: red colour denotes the time period in which temperatures dropped below $0°C$. Bottom subplots: show the bilinear relationship that exists between temperature variation and natural frequency.

subsequently ruptured. More details can also be found on the KU Leuven's Structural Mechanics website [16].

3.2 CM on rotating machinery & gas turbines

3.2.1 WT-REB

Bearing vibrations in this experimental campaign were collected from a laboratory gearbox wind turbine test-rig. The main shaft rotational speed is set at 45 Hz, powered by an electrical induction alternative-current motor. A three-axis accelerometer was attached to the housing of one of the REBs and data were collected at a sampling rate of 25 kHz. A single speed reduction gearbox with a 2.8 ratio is also located next to the generator. The generator is used to provide electrical power on four settings: $\{0.0, 0.1, 0.2, 0.3\}$ kW. All major components of the experimental setup are shown in the diagram in Figure 3.8. Note that Bearing no. 1 was the instrumented REB, from which acceleration measurements were recorded.

Damage was manufactured on the *outer race* of Bearing no. 1 by drilling five different groove

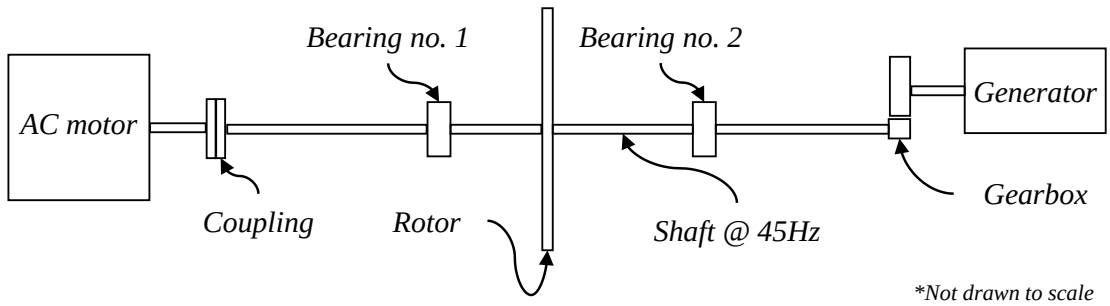

Figure 3.8. Diagram depicting the main elements of the experimental laboratory gearbox test-rig.

sizes of up to $1mm$ in width and depth. So that level 5 of damage progression, corresponds to the largest groove size, while level 1 to the smallest. For each of the *four* generator *loading conditions* and *five damage progression levels*, 10 seconds of acceleration time-histories were collected. Note that, this is a controlled experiment, having well-known operating conditions and damage modes and progression levels. Note that the DI frameworks that this thesis will be demonstrating in subsequent chapters are *applicable* to any damage mode, e.g. inner race, outer race or rolling ball. However, the fact that the focus is on outer race damage makes the problem more straightforward.

In Figure 3.9 and 3.10, the time- and frequency-domain plots, respectively, are shown for a small subset of the vertical accelerations recorded. It can be seen from Figure 3.9 on the top plot that as damage progresses beyond 30 s, i.e. damage levels 3 to 5, the standard deviation/RMS value increases substantially, in comparison to the first two levels of damage, i.e. from 10 s to 30 s. Furthermore, the standard deviation/RMS value drops as the load increases from 0.0 kW $(0-10\ s)$ to 0.3 kW $(30-40\ s)$, as indicated in Figure 3.9 on the bottom plot.

Note that the expected BPFO on this particular speed profile and REB geometry is 233 Hz. In Figure 3.10 the expected BPFO and its harmonic series (up to $12\times BPFO$) are shown as vertical dotted lines, also called cursors. Cursors are typically used to aid the CM process using visual tools like PSD plots. Some indication of outer race damage may be seen from the 6th to the 8th BPFO harmonic, on the highest damage level (Damaged L5). However, on an intermediate damage level (Damage L2), any indication of damage remains hidden. Also, on the right plot in Figure 3.10, the increase in generator load tends to reduce the previously seen peaks on the BPFO harmonics (i.e. for Damage L5). This is significant, because it means that even at the highest level of damage, the indications of damage may be masked on high loading conditions. This demands more sophisticated DI procedures to be implemented that maintain their damage sensitivity, regardless of the load.

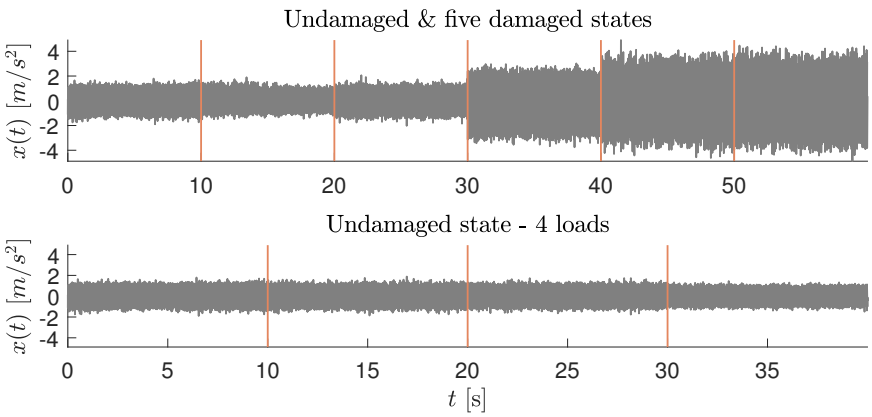

Figure 3.9. Concatenated acceleration time-history plot on the vertical acceleration records on Bearing no. 1: Accelerations on different levels of damage, up to damage level 5 $(50-60s)$, and the undamaged $(0-10s)$ health state (top plot); Undamaged state on four generator loading conditions, starting from no load $(0-10s)$ up to the highest load $(30-40s)$ (bottom plot).

3.2.2 Gas turbine engines

The author of this thesis was actively involved in testing a variety of SAFs* and AFRs on two experimental facilities housing: a small turbo-shaft GT engine, i.e. an aircraft Auxiliary Power Unit (APU) and a single combustion chamber from turbo-fan GT engine. The author also worked on designing these experimental tests, including the LabVIEW implementation. These tests were mainly meant to investigate the impact on engine performance using such atypical operating conditions, i.e. SAFs and relatively high AFRs. From these tests, three datasets were collected:

- *TF-LBO*: Testing the LBO limits of a turbo-fan GT engine on three AFRs by measuring structural accelerations;
- *TF-SAF*: Investigating the GT engine performance of a smaller turbo-fan GT engine using 35 SAFs by collecting structural accelerations, pressure, temperature measurements, among others;
- *TS-SAF*: On the APU GT engine, 2 different AFRs and 2 SAFs were tested, measuring structural accelerations.

3.2.2.1 TS-SAF

This experimental facility houses a Honeywell GTCP85, which is an APU of turbo-shaft gas turbine engine type. The operating principle of this type of engine follows a typical Brayton

*Consisting of various fuel blend mixtures of conventional jet, bio and other sustainable fuels.

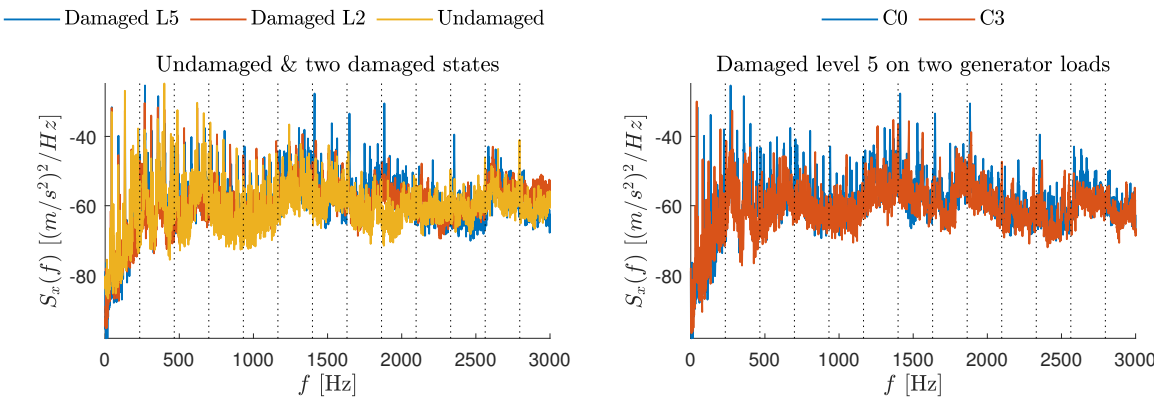

Figure 3.10. Spectrum of power density of vertical acceleration records on Bearing no. 1, showing only up to 3 kHz, for damage identification purposes. Left plot: Undamaged and two damage levels, including the highest one, i.e. Damaged L5, with no load (C0); Right plot: Damage level 5 and two loading conditions, C0 and C3 (0.3 kW).

cycle. As can be shown in the schematic diagram of the engine in Figure 3.11, the engine draws ambient air from the inlet at 1 bar through the centrifugal compressor $C1$, where it raises its pressure by accelerating the fluid and passing it through a divergent section. The fluid pressure is further increased across a second centrifugal compressor $C2$, before being mixed with fuel into the combustion chamber CC and ignited to add energy into the system (in the form of heat) at constant pressure. The high temperature and pressure gasses are then expanded across the turbine, which drives the two compressors, a $32kW$ generator G that provides aircraft electrical power and the engine accessories EA, e.g. fuel pumps, through a speed reduction gearbox. Several gears exist at different ratios to the main shaft speed, which itself is rotating at a constant speed of $\approx 700 \pm 1.6\ Hz$. For instance, there is a speed reduction ratio of $9.8:1$ for the tachometer, $2.6:1$ with the generator drive and $6.8:1$ with the output drive shaft.

The air outlet valve, allows the extraction of high temperature, compressed air at $\approx 232°C$ and $338kPa$ of absolute pressure to be passed to the aircraft cabin and to provide pneumatic power to start the main engines. This allows the engine to be tested on different operating modes as the AFR that goes into the combustion chamber can be modified. When the air outlet valve opens, a decrease in turbine speed will take place if there is no addition of fuel to compensate for the lost work. The energy loss arises from the decrease in work done w_{c2} to the engine's working fluid as it passes through the second compression stage. The amount of lost work is proportional to the mass of air outlet m_{ao}, and can be expressed as $w_{c2} = m_{ao}c_p dT$, with c_p being the heat capacity of the working fluid and dT the temperature differential across the second compression stage. Since the shaft speed must remain

constant, the fuel flow controller FFC achieves this by means of regulating the pressure in the fuel line, thus, injecting different mass fuel flow into the combustion chamber.

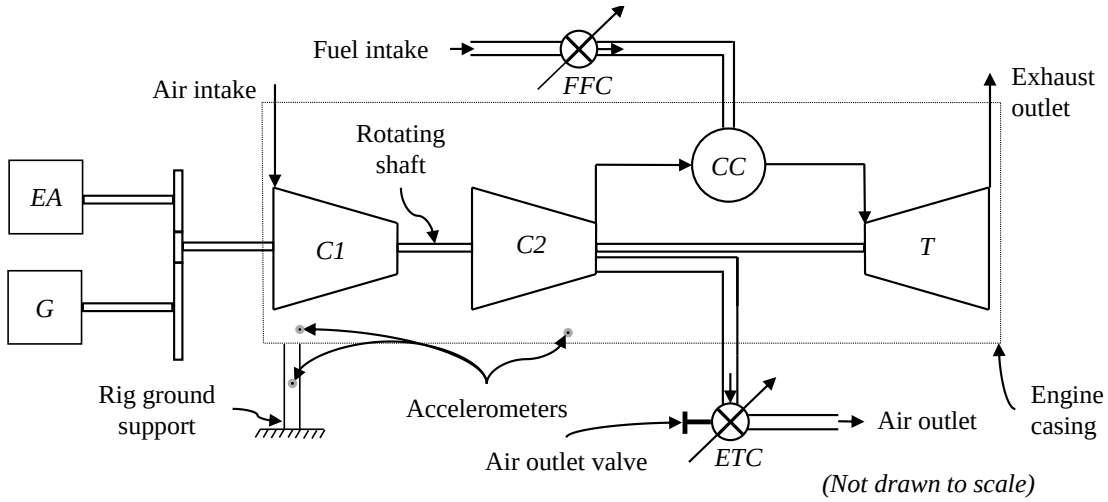

Figure 3.11. Turbo-shaft gas turbine engine schematic depicting main salient features of the experimental test-rig.

Three Dytran 3225F1 single-axis accelerometers (sensitivity of $10mV/g$) were attached on the experimental-rig's ground support structure. Those were located as close as possible to the hot section of the engine, in order to capture the dynamics of combustion processes, from the point-of-view of structural vibrations. Data were sampled on a National Instruments 9234 module at 51.2 kHz, i.e. its maximum sampling rate. The total time duration for these experiments was 20 s for each SAF and AFR combination. Note that measurements were recorded during steady-state engine operation.

In Figure 3.12, the structural acceleration PSDs, recorded using an SAF and a conventional Jet-A fuel on two different AFRs are shown. Two frequency regions of interest ($(8.5, 10.5)$ kHz and $(20.4, 21.1)kHz$) are shown, in which the acceleration amplitude was relatively high. The two AFRs are AFR1 ≈ 136 and AFR2 ≈ 84. As observed on the plots, there are a few processes, in which the fuel composition causes both frequency and amplitude changes. For instance, at around $9kHz$, with SAF and AFR2, significant amplitude drops are observed. Looking at around $10.5kHz$, Accelerometer 2 shows the exact opposite to be happening. At higher frequencies, the two accelerometers show a significant frequency shift with the fuel change on both AFRs. As this frequency shift is typical across the whole spectrum it means that engine main shaft speed as the fuel changes from the conventional Jet-A to the SAF (on the same AFR) decreases, thus, generating less power output ($power\ output \propto rotational\ speed$).

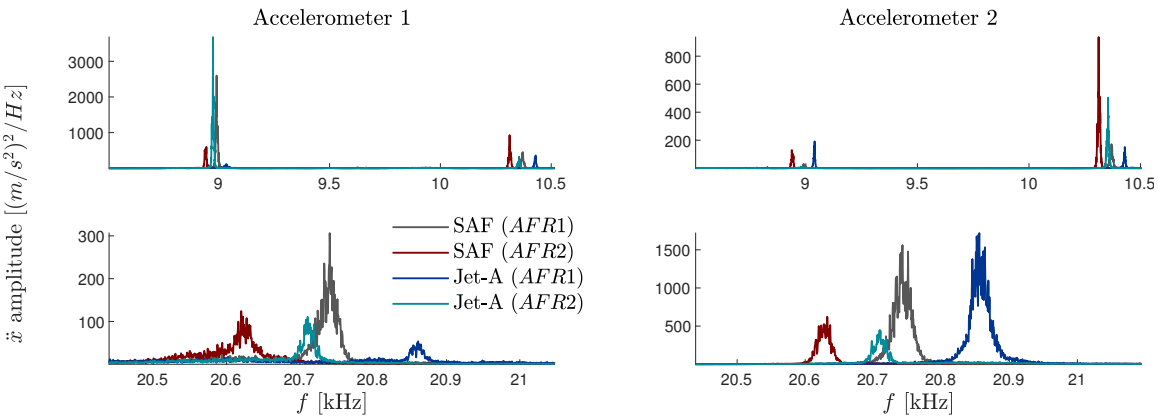

Figure 3.12. Power spectral density estimates from vibration measurements (two accelerometers) on the turbo-shaft gas turbine engine on two different fuels and AFR.

3.2.3 TF-SAF/-LBO

This experimental test-rig houses an annular-type combustion chamber (TF-SAF0), which was originally part of a turbo-fan engine manufactured by Rolls Royce for aircraft propulsion. In addition, the facility houses a smaller-scale in-house designed and built combustor (TF-SAF). With these experimental test-rigs a number of SAFs were investigated. In addition, various AFRs, to test LBO limits were tested (TF-LBO).

1. *TF-SAF0*: 25 alternative fuels were tested, using the same data acquisition as in the TS-SAF experiments. Additionally, four pressure sensors (PCB 106B52) were incorporated and were located at several locations across the GT engine. Unfortunately, instrumentation errors such as unsuitable mounting positions of pressure sensors, and location of accelerometers meant that measurements were not sensitive to the physical processes taking place during combustion. However, to alleviate this problem, the same data acquisition system was transferred to the smaller-scale combustor (TF-SAF);

2. *TF-SAF*: 35 alternative fuels were tested, with the sensors being located sufficiently closer to the combustion chamber this time. This is to measure any subtle differences in GT engine performance between the SAFs. Measurements of vibration were being recorded, each for $\approx 60s$, having a constant $AFR \approx 90$ and exhaust gas temperature of $\approx 405°C$. Five accelerometers were mounted around the combustor's cold section, which means some redundancy was expected. To evaluate this redundancy the pairwise correlation of the PSD estimates between each of the 5 sensors was evaluated across the 35 fuels. It was observed that the calculated correlation coefficient between

two pairs of fuels was greater than 0.9, so that, So that either Accelerometer 2 or 4 should be removed from further analysis. Table 3.3 shows the calculated correlation coefficient values for one of the SAFs.

Table 3.3. Calculated correlation coefficients from PSD estimates between pairs of the five accelerometers for one fuel (similar trends were observed for the rest of the fuels).

Accelerometer #	*1*	*2*	*3*	*4*	*5*
1	1.00	**0.97**	0.09	0.68	0.72
2		1.00	0.11	0.68	0.73
3			1.00	0.16	0.07
4				1.00	**0.94**
5					1.00

In terms of variation in acceleration amplitudes across 35 fuels, including conventional kerosene-based ones, some changes were observed across a relatively narrow frequency band. As seen in PSD plots in Figure 3.13, Acceleration 3 and 5 show amplitude changes on different frequency ranges. Note that the most variation in amplitude between the fuels occurs within the frequency band presented in these two plots;

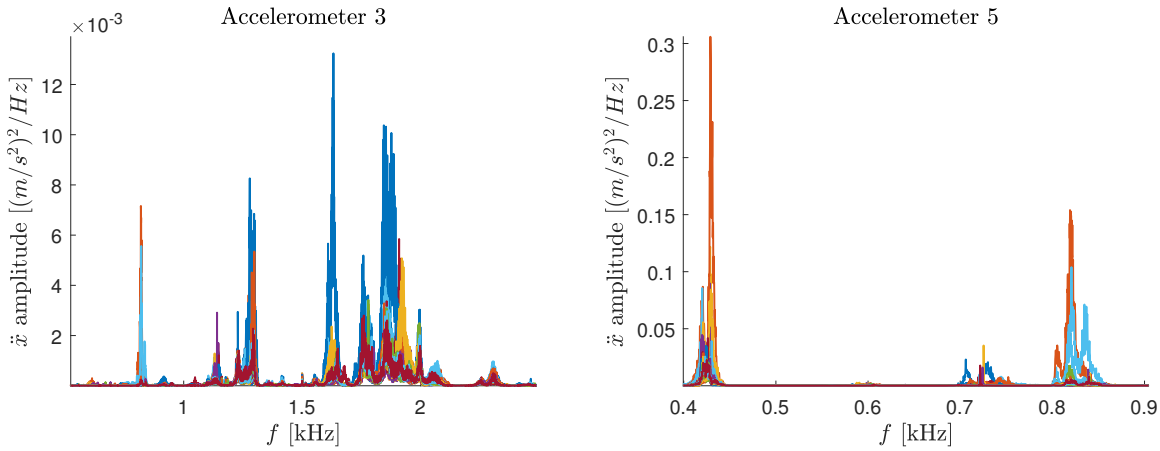

Figure 3.13. Power spectral density estimates from vibration measurements (two accelerometers) on the turbo-fan gas turbine engine combustor on 35 different fuels.

3. *TF-LBO*: Using the same data acquisition system, as previously, three different AFRs were considered using Jet-A1 fuel by altering the air mass flow rate $\approx \{300, 360, 420\}$ kg/s and the corresponding fuel mass flow rate $\approx \{.56, .66, .75\} \times 10^{-3}$ kg/s. As AFR in-

creases, it was observed that the overall vibration amplitude level as measured on each of the five accelerometers rises, too. Without emphasising on particular regions in the frequency-domain, as in previous datasets, the four statistical moments can provide adequate insights into the vibration amplitude differences as AFR increases. As seen in Figure 3.14, both AFRs 1 and 2 have skewness values ≈ 0 and kurtosis to ≈ 3, which suggests Gaussian-like distributions. On the other hand, AFR3 differs in distribution shape. In Figure 3.15 the KDE estimates from data on AFR 1 and 3 are shown. Thus, the distribution shape of AFR 3 is bimodal and more dispersed, as compared to AFR 1, which is Gaussian-like and "peaked".

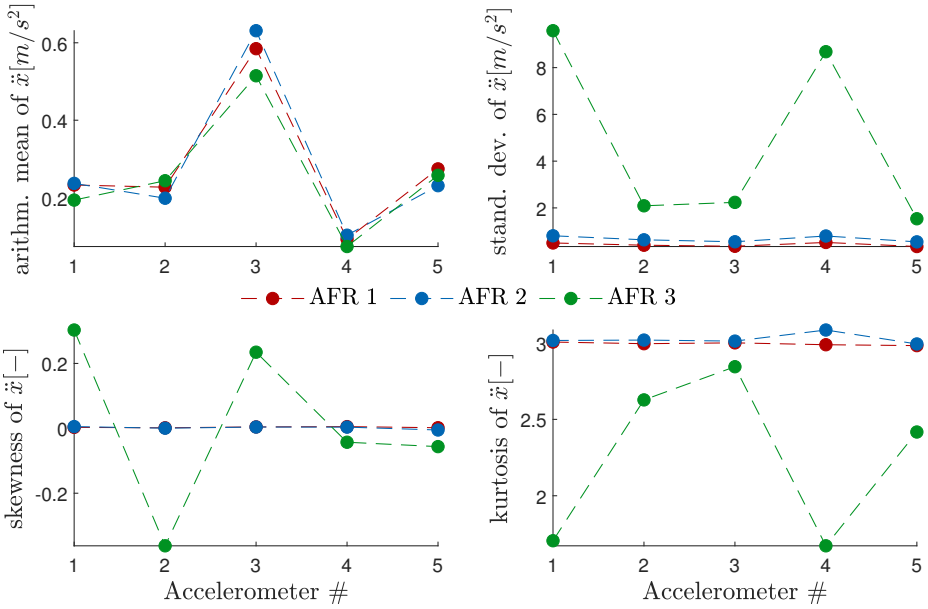

Figure 3.14. Four statistical moments calculated on the vibration measurements as recorded on five accelerometers.

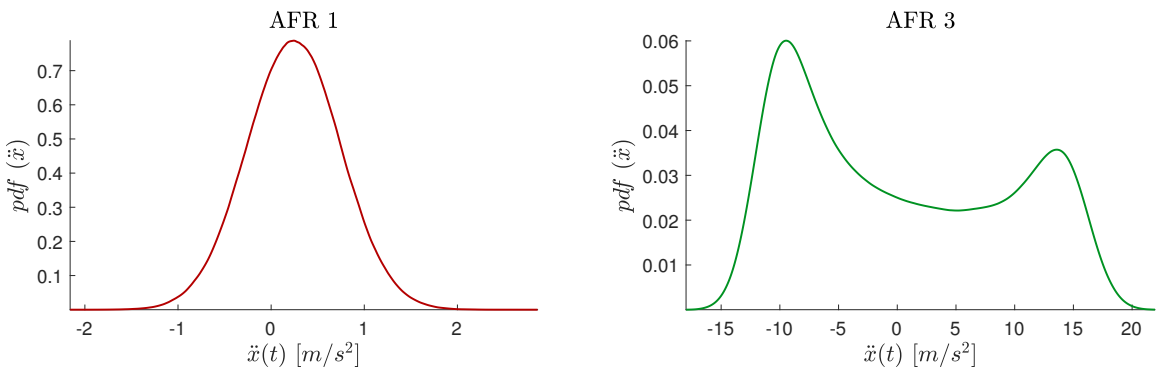

Figure 3.15. Kernel density estimates calculated on vibration measurements from Accelerometer 1 on two AFR.

3.3 Summary

The datasets presented above cover a broad spectrum of challenges in SHM & CM. DI in this thesis ranges from specific *localised* damage, for instance, REB outer race defects to more *generic* system changes, e.g. structural vibration amplitude changes with SAF on a GT engine. The latter presents a more ambiguous problem, so that any DSF may use global characteristics, e.g. central moments. In contrast, for CM on REBs, specific frequency ranges were analysed to reveal changes due to damage (e.g. on the outer race). The unique properties of each dataset and the significance behind using it in this thesis, are summarised below.

For the *NSim-4DOFLin* and *NSim-4DOFNonLin* datasets:

- Basis for testing the proposed DI methods.
- Specifically, allows for tuning parameters of the proposed DI methods, e.g. MED filter coefficients.
- Conduct outlier diagnostics using by using DSFs that are influenced by EOV.
- Perform modal identification of parameters using adaptive signal decomposition methods and use those for testing the proposed DI methods.
- Experimental air temperature profile was implemented - realistic EOV influence.
- Bilinear spring force profile, simulating a crack - NSim-4DOFNonLin.

For the *LA-4DOF* dataset:

- Impact-type of damage of progressive severity via a length-adjustable bumper.
- EOV influence on system dynamics has been accounted for by loosening the connections of the support columns, i.e. stiffness coefficient variation, and added masses to the structure.
- Simpler EOV influence on system dynamics, and well-controlled impact-type of damage allows testing the sensitivity of the proposed DI methods and DSFs.
- Modal parameter identification and outlier diagnostics, as in NSim-4DOFLin and NSim-4DOFNonLin datasets.

For the *WT-REB* dataset:

- As in typical CM problems, the DI problem focuses on certain frequencies related to the defect. In this case, the outer race fault bearing frequency.

- The above allows for both detection and diagnosis of damage.
- Challenges in DI as the indicators become increasingly less sensitive due to the increase in operating load - similar to the "masking" effects caused by EOV influence.
- Allows for benchmarking the proposed DI algorithms against established techniques used in the industry for CM in REBs, e.g. the envelope analysis technique.

For the *Z24* dataset:

- Not impact-based type of damage, but, it progresses in time.
- A set of DSFs, i.e. system's first four NFs, are provided, which are highly influenced by EOV.
- Outlier diagnostics using robust methods offer insights into the effects of EOV influence and damage as measured by changes in NFs.

For the *TF-LBO* dataset:

- The DI problem is posed as one in which unstable operating conditions of the combustor must be identified. Specifically, air-to-fuel ratio leading to lean blow-out and flame extinction.
- There are no operating conditions that will influence the DI problem, e.g. no "masking" effects as in the previous datasets.
- Simpler problem to test DI methods to detect the presence of unstable GT engine operations as measured by vibration sensors.

For the *TF-SAF* and *TS-SAF* datasets:

- In both TF-SAF and TS-SAF datasets, the aim is to identify structural vibrations that differ significantly from the typical engine condition. For instance, the effects of a specific alternative fuel on the vibration characteristics.
- As with TF-LBO dataset, there are no "masking" effects to the DI problem.
- For TS-SAF, the engine type is a turbo-shaft one. This is significant because it allows testing for the proposed DI problem on a different type of GT engine.

Data exploration conducted on these datasets offered an invaluable insight into the design of suitable DSFs for each problem. The main conclusions from this chapter are:

- *CM on REBs*: Rolling slip creates smearing effects in the frequency spectrum, thus, it weakens the amplitude at the defect frequency. There are many parameters in the EA procedure; AR model order and band-pass filtering cut-off frequencies using SK. Thus, the tuning of these parameters needs to be carefully conducted. Kurtosis and its variants is an important measure that must be included in the DSF vector. Changes in operating conditions in the form of increasing the generator loads were also shown to influence the amplitude at the BPFO. Therefore, it needs to be addressed in the DSF development step;

- *CM on GT engines*: The observable changes in the probability distribution, e.g. the bi-modality of the higher AFR in testing the LBO limits and its higher-order statistical moments enable the separation of the nominal and the extreme vibration behaviour. For this specific test scenario, Accelerometer 1 showed to be most sensitive to the AFR ratio changes, as far as the higher-order moments are concerned. For the rest of the datasets, i.e. TS-SAF and TF-SAF datasets, some significant differences within certain frequency regions could be observed from the PSD plots, e.g. between $\sim 20.5kHz$ and $\sim 20.9kHz$ (TS-SAF) and between $\sim 0.5kHz$ and $\sim 2.2kHz$ (TF-SAF). Therefore, spectral characteristics, e.g. spectral means, may also be proved important for automated DI and DSFs development;

- *SHM on structures and bridges*: NFs are undoubtedly influenced by damage, as proven from the Z24 dataset. However, they are also influenced by temperature, as it was observed in the same experiment. This influence can be modelled with a bilinear relationship. Hence, temperature change causes non-linear behaviour, although this occurs on larger time-scales than damage. One way to distinguish between the two is that damage introduces local changes in the modal properties, whereas environmental variability is spread along the structure (global change). Non-linearity is also introduced as opening and closing crack, in the LA-4DOF experiment. Kurtosis in this case is also extremely important, since, as in the WT-REB case, damage is in the form of impacts; it changes the peakedness of the probability distribution. The structure, under the influence of a relatively weak ambient excitation force, may assumed to be linear. This is especially true for short time-scales. Modal properties such as DRs have also been proven to be valuable DSFs, but, their estimation may be a challenge.

These datasets provide an appropriate testing ground for investigating the capability of automated DI methodologies under the influence of EOV on the system dynamics. The next chapter will be concerned with the design of DSFs.

Chapter 4

Damage-sensitive features for damage identification

4.1 The importance of DSFs in DI

Deriving a suitable set of DSFs, which is the *feature extraction* stage in the SPR paradigm (see Chapter 1), is a key step in the development of an automated and effective DI framework. In particular, DSFs derived from data, need to be *immune* to EOV influence, such that to improve the *robustness* of the SHM & CM strategies.

The methods examined in Chapter 2, and in particular the adaptive signal decomposition methods, will be used as part of the DSFs derivation process. Given their *filter bank properties*, the different modes of vibration can be separated. Thus, in comparison to broadband signal analysis, e.g. using amplitudes from the PSD (across the whole Nyquist frequency range) as DSFs, the sensitivity to damage is expected to improve by *selecting the most relevant modes*. Also, each narrow-band mode / IMF can be interrogated separately from the rest, which has the potential to unmask any damage presence. Selecting the most relevant or informative ones, will also eliminate redundancy. Hence, adaptive signal decomposition techniques can be thought of as *dimensionality reduction techniques for time series data*.

Unlike other methodologies that investigate the use of deep learning (DL) architectures for modelling data with temporal dependencies [77], this thesis takes the more traditional approach of "handcrafting" a set of features. As such, this thesis makes the case that by using handcrafted features, one can incorporate *fundamental information* of the physical characteristics of the system that is being monitored. It should, nevertheless, be acknowledged the fact that there are certain techniques in time-series analysis using DL, e.g. Long Short Term Memory (LSTM) and bidirectional LSTM Networks, that have shown impressive accuracy in recent literature. For instance, in [43], the authors have explored the possibility of using the temporal dependencies learned by LSTM (trained on raw input data) to build a more accurate representation in activity recognition tasks. Whereas, the generalisation

capability of a single-layer Feed-forward ANN was transferred to LSTM (as the latter is prone to overfitting).

4.1.1 DSFs used in the literature

In vibration-based CM for REBs, it is very common to calculate the energy value of each kth IMF, i.e. $\sum_{n=1}^{N} c_k(n)^2 \; \forall \; K$ IMFs, in order to construct a K-dimensional DSFs matrix, either using VMD- as in [21], or EMD-based methods as in [175] and [194]. In the latter work, the authors went one step further to compute the differential entropy of the normalised energy of each IMF. On the other hand, in [171], the Hilbert marginal spectrum was constructed using EMD-derived IMFs. This in turn, was used to find the *frequency bands* that are closer to the characteristic frequency of three types of REB damage: outer-race (i.e. the BPFO), inner-race and ball damages. In [25], the authors used the mode mixing problem for DI purposes in a gearbox. In particular, given that mode mixing occurs when an intermittency is present in the signal, e.g. an impulse-type of damage, the *number of IMFs* generated by the EMD increases (as it is an iterative technique). Therefore, the authors discussed the possibility of using the number of IMFs as potential DSFs.

For vibration-based structural DI under EOV influence a variety of DSFs were proposed in the literature. In [59], the authors used the peak amplitudes of the FRFs from a wind turbine blade. AR models were also used for the LA-4DOF dataset in [159] and [108], while well-known benchmark studies in SHM use modal-based DSFs (mostly NFs derived from system identification methods like SSI), e.g. the Z24 [152] and the Wooden Bridge [104] datasets. NFs as DSFs were also used in other studies utilising different sets of experimental data, e.g. in [83]. Another common approach, is to track NFs by selecting a subset of the amplitudes from the power spectrum around the NF-region(s) of interest. In that way, amplitudes within a specific frequency band, can be used to monitor changes in modal characteristics as shifts in magnitude, as was done in [187] (although not with vibration signals). These DSFs, however, are highly sensitive to EOV changes (temperature variation in this case). Therefore, NFs and DRs, which were estimated using the MPI method as outlined in Chapter 2, are examined as alternatives for DI in structures.

4.2 Methods for deriving DSFs

4.2.1 Features normalisation

Given the *scale differences* from the various types of features that can be computed, e.g. statistical measures such as central moments and information theoretical ones such as entropies, *feature normalisation* may proved to be important. That is, when a particular

algorithm, in the statistical model development stage, is not *location and scale invariant* this will cause many issues. In the context of robust outlier analysis (see Appendix C), this property is essential for the effectiveness of this type of models.

Given that the range of a given type of feature may differ significantly from the other(s), choosing not to *normalise the features into similar dynamic range of values*, the statistical model will eventually place more emphasis on the ones that take larger values [168]. Hence, the statistical model will be *biased* towards the features with the largest dynamic ranges [30]. An example of feature normalisation in CM is given in [66]. The authors of this paper have used the normalised PSD amplitudes from a range of different frequency bands of interest, to monitor the condition of REBs. These normalised PSD amplitudes were computed by removing the mean and dividing by the standard deviation of all the amplitudes in the spectrum.

Giving an *equal weighting* to the features in the DSFs matrix is, therefore, necessary. This is particularly true when the statistical model relies on distance measures to infer whether a data point comes from the undamaged or damaged system states. Such distance measures include the squared multivariate distances or *Mahalanobis distances* [18], which are examined in Chapter 5 of this thesis.

As demonstrated in many studies, e.g. in [139], it is common to *test* different types of feature normalization methods, in order to find the one that achieves the best performance. In general, the choice of the specific feature normalization will be dependent on many factors. One of the most important ones being the choice of the statistical model used to fit our data, given that this model is not location or scale invariant. For that reason, this thesis is examining three feature normalisation methods, as shown in Table 4.1.

The first one, being the *z-score*, which subtracts the mean and divides by the standard deviation of the data or feature vector $\mathbf{x} \in \mathbb{R}^N$ (i.e. a single feature type such as the kurtosis values of N data points)*. However, these two statistics, i.e. the mean μ_x and standard deviation σ_x are influenced by data points that may have "significantly" different values (e.g. very large values in the absolute sense) than the rest of the data, also called *outliers*. These outliers will skew the z-score estimates towards their direction (more details in Chapter 5). An example of what can be considered an outlier, is the change in NFs seen when $T \leq 0$, in the Z24 dataset (see Chapter 3 for details). If we want to unmask these outliers, the z-score may be replaced with what is called a *robust z-score*, where it replaces the sample mean and standard deviation of a certain feature vector $\mathbf{x}$ with its median $med(\mathbf{x})$ and median absolute deviation (MAD). This converts the whole analysis procedure into a location and

*Note that N is not necessarily the same as the one used in Chapter 2 to refer to the discrete-time sample index, i.e. $n = 1, ..., N$.

scale invariant one.

Table 4.1. Feature normalisation methods used.

Feature normalisation	*Description*
Z-score	Scale each feature vector $\mathbf{x}$ so that it has properties of the Gaussian standard distribution, i.e. $\mathbf{x}_s = \frac{\mathbf{x}-\mu_x}{\sigma_x}$.
Robust z-score	Same as above but replace the mean μ_x and standard deviation σ_x with robust estimates of scale and location, calculated with median and median absolute deviation (MAD), respectively, i.e. $\mathbf{x}_{rs} = \frac{\mathbf{x}-med(\mathbf{x})}{MAD}$.
Yeo-Johnson	Power transform that takes into account outliers in the data and brings the feature vector $\mathbf{x}$ closer to a symmetric distribution. Unlike other power transformations, YJ allows both positive and negative values in $\mathbf{x}$.

The third feature normalisation method used in this thesis is the Yeo-Johnson (YJ) normalisation, which has recently been proposed in [151]. YJ is a *power transformation* method[†], that seeks to transform the feature vector $\mathbf{x}$ as close to a symmetric distribution as possible, e.g. a Gaussian distribution. At the same time it enables both negative and positive feature values, something which has not been possible with other power transformation methods. It involves computing and modifying the four central statistical moments of the feature vector $\mathbf{x}$. More importantly, YJ accounts for potential outlying data in the feature vector $\mathbf{x}$, by computing a new set of values for its mean and standard deviation using a re-weighted maximum likelihood estimation (MLE). The concept is similar to an *M-estimator*, that is presented in Appendix B.

In Figure 4.1, the three feature normalisation methods are applied on the first two NFs of the Z24 dataset. Alongside that normalised DSFs, the $\chi^2_{99\%}$ ellipses are shown to illustrate any particular changes to data distribution. It is clear from the plots that, applying YJ causes a shift in the data distribution, so as to achieve the so-called "*central normality*" [151]. This means that the most central part of the data or the majority of the data resemble a symmetric distribution, like the Gaussian. In most techniques for outlier analysis, e.g. Mahalanobis distances, the Gaussian distribution requirement for the data is essential. In Chapter 5, most techniques used for outlier analysis incorporate a symmetrical distribution assumption for the majority of the data. Note that YJ is implemented as part of the *cellWise: Analyzing Data with Cellwise Outliers* software package [4] in R, which was developed by the Robust Statistics research group at KU Leuven [9].

[†]There are many other power transformation methods, e.g. Box-Cox transform [98].

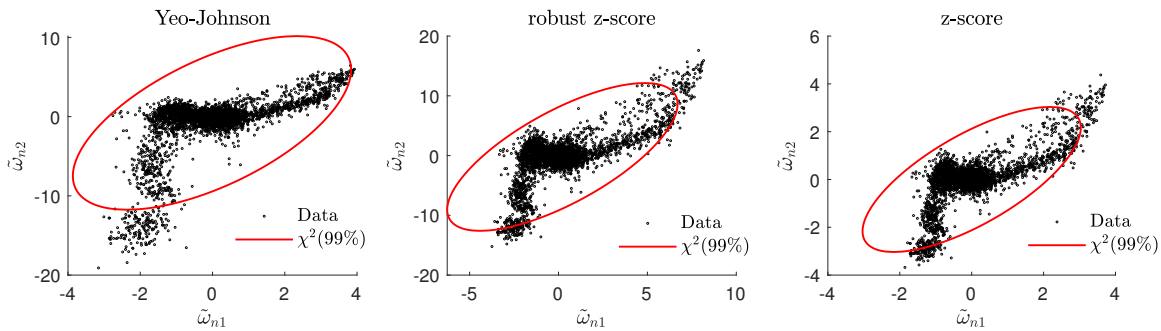

Figure 4.1. Three feature normalisations applied on the first two NFs of the Z24 dataset.

On the other hand, for novelty detection (ND) tasks (see Chapter 6) feature normalisation is used in a slightly different way. This is because in ND there is a known dataset available from the undamaged system state, e.g. the feature vector $\mathbf{x}_{un}$. Therefore, its mean $\mu_{x_{un}}$ and standard deviation $\sigma_{x_{un}}$ may be used to normalise any *future unseen feature vector* $\mathbf{x}_*$ as $\tilde{\mathbf{x}}_* = \frac{\mathbf{x}_* - \mu_{x_{un}}}{\sigma_{x_{un}}}$. This a common procedure in ND tasks, e.g. in [186], where the authors normalised the amplitudes of the PSD of vibration signals in a similar way, prior to using them as feature vectors to train a Self-Organising Map algorithm. This helps in enhancing the separability between inlying and outlying data points, which may be present in $\mathbf{x}_*$, as already demonstrated by the author of this thesis in [123] and [124].

4.2.2 Visualisation, compression & selection of DSFs

Very often, the set of DSFs used in SHM & CM applications is usually a high-dimensional one, i.e. much greater than 2. For instance, the number of DSFs derived from the EA procedure, can be equal or even exceed the number of available examples. Therefore, it is infeasible to *visualise* these DSFs on a plot. Additional issues with high-dimensional DSFs include data overfitting and high computational requirements. The latter will occur, for instance, when inverting a large covariance matrix to calculate thresholds in a Monte Carlo (MC) simulation approach (see Chapter 5 for details on threshold calculation using MC). Also, most methods used in Chapter 5 and 6 for outlier detection (OD) and ND, respectively, require that the number of examples be much greater than the number of dimensions / features [156].

In terms of *selecting* the most relevant features for the statistical model development stage, there are several approaches, as discussed in the review paper in [131]. For instance, in the *forward selection* approach, features are added gradually to test their performance and weights are added to each feature vector, according to a certain *loss function* minimization argument. In unsupervised learning, e.g. in OD, where samples are not labelled, this loss

function may not be possible to be defined. Therefore, it is necessary to resort to dimensionality reduction methods as a means of both obtaining a reduced set number of features used in the DSFs matrix and for visualising them on a plot. The only drawback is that the selection process is now being conducted on the basis of the dimensionality reduction method itself, and not on the specific problem. This will be explained in the next few paragraphs.

4.2.2.1 PCA & its kernel-based equivalent

One of the fundamental assumptions behind dimensionality reduction techniques is that most *variance* in a dataset exists on a *lower-dimensional manifold*, i.e. lower from the dimensionality of the original dataset. The variance, is therefore, directly connected to the "most significant" information that can be extracted from a dataset. A good review of dimensionality reduction techniques, which compares a large number of them on different datasets, is given in [183]. The authors of this paper have generously provided a Matlab Toolbox (link is provided in their paper) that implements all these techniques.
One of the first and most widely-used techniques is principal components analysis (PCA) [30]. This method works by transforming a data matrix X[‡] by projecting them into a new set of orthogonal axes, i.e. $X \in \mathbb{R}^{N \times p} \rightarrow \mathbb{R}^{N \times k}$, where $k \leq p$. The orthogonal set of k axes, with $k = k_1, ..., k_p$, are in decreasing order in terms of the *amount of data variance each axis captures*. Ideally, a method like PCA, does capture most of the data variance using the first few axes (or principal components), so that visualisation can be carried out, too. In PCA, the orientation of the kth principal component is calculated by solving for the eigenvectors $\mathbf{u}_k$ (normalised to have unit length) of the data covariance matrix $\Sigma(X)$, where the variance in each kth principal component is given by the kth eigenvalue λ_k. The projection of X (assume it has been normalised first using, for instance, a z-score method), is given as the linear combination of the $p \times k$ loadings matrix $U = [\mathbf{u}_1, ..., \mathbf{u}_p]^T$, i.e. the contribution of each $\mathbf{u}$ (sorted from the largest to the smallest k eigenvalues $\lambda_k \geq \lambda_{k-1} \geq, ..., \geq \lambda_1$) on the kth principal component, or the so-called k^{th} PCA subspace,

$$\tilde{Y} = XU, \quad \tilde{Y} \in \mathbb{R}^{N \times k} \tag{4.1}$$

PCA assumes data in X are Gaussian distributed, due to the use of the sample covariance $\Sigma(X)$. Additionally, the iid assumption is necessary, so that there exists directions (the principal components) that diagonalize the covariance matrix.
The fact that the projections $\tilde{Y}$ are *linear* combinations of the p features in X and their

[‡]This can be the matrix with p feature vectors, i.e. the DSFs matrix.

individual contributions to each kth eigenvector, may not be valid for some problems. This is because, the pairwise relationship between the features in X may exist in non-linear high-dimensional manifolds, e.g. can be exponentially decaying or higher order polynomials. Therefore, the estimated principal axes will not be able to describe "sufficiently" the variance that exists between the feature vectors, which may lead to wrong conclusions about the data characteristics.

Kernel PCA (kPCA)[§] was developed as a generalisation for PCA, such that, it can be used to extract k linear or *non-linear* principal components, where $k = k_1, ..., k_N$, i.e. up to N components. Generally, kPCA employs a kernel function $k(\mathbf{x}_n, \mathbf{x}_m) = \phi(\mathbf{x}_n)^T \phi(\mathbf{x}_m)$ with a feature map $\phi(\cdot)$ such that it projects the *original feature space* $\mathcal{X}$ to a new *higher-* or *infinite-dimensional feature space*, i.e. $\mathcal{X} \rightarrow \mathcal{F}$. In general, kernel methods like kPCA employ the *kernel "trick"* in order to compute the non-linear "similarities" between $\mathbf{x}_n$ and $\mathbf{x}_m$. Given that $\mathcal{F}$ may be potentially infinite-dimensional, the kernel "trick" makes it possible to *implicitly* map / project the data. Further details on the kernel "trick", the *RBF kernel* and the *Gram matrix* $\tilde{\mathcal{K}}_{\mathbf{x},\mathbf{x}}$ are provided in Appendix B.

It can be shown that by replacing the dot product with a kernel function $k(\mathbf{x}, \mathbf{x})$, the following eigenvalue problem presents itself,

$$\tilde{\mathcal{K}}_{\mathbf{x},\mathbf{x}} \mathbf{a}_k = \lambda_k N \mathbf{a}_k \tag{4.2}$$

where, $\mathbf{a}_k \in \mathbb{R}^N$ is the kth eigenvector (it can be shown that the eigenvector is a linear combination of the mapped features, i.e. $\mathbf{a}_k = \sum_{n=1}^{N} a_{k,n} \phi(\mathbf{x}_n)$). For an arbitrary data point $\mathbf{x}_* \in \mathbb{R}^p$, its projection onto the kth principal component will be,

$$y_k(\mathbf{x}_*) = \sum_{n=1}^{N} a_{k,n} k(\mathbf{x}_*, \mathbf{x}_n) \tag{4.3}$$

The Gram matrix (see Appendix B) is used instead of the kernel matrix for a non-centred mapped dataset $\phi(\mathbf{x})$.

4.2.2.2 Dimensionality reduction with outliers - robust PCA

In the presence of outliers, the principal axes computed using either PCA or kPCA, will be "pulled" towards observations that are sufficiently far from the bulk of the data, since this is where the variance is maximal. Therefore, it will not be possible to view potential patterns

[§] An argument for using kPCA has to do with the fact that the feature space is "expanded" up to N *non-linear* principal components, as compared to a projection of up to p *linear* principal components, as in PCA. The extra "redundancy" in these eigendirections is expected to provide a means to identify data patterns and at the same time noise (that may be eliminated). In a way, this may be seen as similar to the filter bank methods, where a univariate time-series signal is decomposed into its spectral contents and noise (see Chapter 2). In addition, in datasets with just a few features, the dimensionality reduction procedure of the PCA may result in substantial information loss, since most of its structure may exist over all eigendirections.

in the data, since these outliers / "extreme values" will dominate the analysis. Therefore, a *robust version of PCA* is needed, which can be used to perform dimensionality reduction on the majority of the data that are not considered as outliers. Then, it is possible to visualise potential patterns that may exist in the data and obtain DSFs that are free from such extreme values.

One of the earliest and most straightforward approaches to the robust PCA problem, is to replace the sample covariance matrix, with a *robust covariance estimator*, e.g. a minimum covariance determinant (MCD) (see Chapter 5), so that the eigenvalues are computed as in the original PCA formulation. Note that, the MCD estimator is limited to problems where p (the dimensionality of the original dataset) is *much less* than the number of non-outlying data points. An important requirement in dimensionality reduction, however, is to be able to handle problems where p can be equal to the number of examples.

In [95] the authors proposed a robust PCA approach that combines the robustness properties of the MCD (see Chapter 5 regarding MCD and Appendix C regarding robustness properties) with projection pursuit (PP), so that it can handle high-dimensional data, while also being insensitive to outliers. In their formulation, the PP is used as an initial step in order to reduce the dimensionality of the problem, while the MCD is used to compute robust estimates of multivariate mean and covariance that can be used in eigenvalue decomposition. In particular, PP utilizes a robust measure of location and scale, in this case the univariate MCD, and the Stahel-Donoho outlyingness (SDO) for each observation $\mathbf{x}_n$, so that,

$$SDO(\mathbf{x}_n) = \max_{\nu} \frac{|\nu^T \mathbf{x}_n - \mu_{MCD}(\nu^T \mathbf{x}_j)|}{\sigma_{MCD}(\nu^T \mathbf{x}_j)} \tag{4.4}$$

where the robust location and scale estimates, $\mu_{MCD}(\nu^T \mathbf{x}_j)$ and $\sigma_{MCD}(\nu^T \mathbf{x}_j)$, respectively, are computed from all N data points ($j = 1, ..., N$). SDO computes the outlyingness of $\mathbf{x}_n$ by projecting it onto unit length vectors $\nu \in \mathbb{R}^p$ and taking the maximum over all directions. In practise p may be large, so that the algorithm randomly samples up to 250 directions.

It is then possible to calculate the linear projections, using the original PCA method (as above), for a *subset* of h data points, whose SDO is the smallest among all N data points. Then only the $k(< p)$-projections (or dimensions), whose variance is greatest, are retained. An improved robust subspace estimation is then obtained by first computing the residual or *orthogonal distances* (ODist) between each ith projected example $\tilde{\mathbf{y}}_i \in \mathbb{R}^k$ with $i = 1, ..., h$ and data point $\mathbf{x}_i$ as,

$$ODist_i = ||\mathbf{x}_i - \tilde{\mathbf{y}}_i|| \tag{4.5}$$

Subsequently, the $k \times k$ covariance matrix is calculated from all h data points, i.e. $\mathbf{x}_i$ with $i = 1, ..., h$, whose $ODist_i$ is less than a threshold. This new covariance matrix is used to

obtain the improved k-dimensional robust subspace. In the last step, the projected data points onto this new subspace, are used to compute the multivariate mean and covariance using the MCD estimator. Then, the eigenvalue decomposition is performed to obtain the p-dimensional robust loadings matrix $U_{rPCA} = [\mathbf{u}_{1,rpca}, ..., \mathbf{u}_{k,rpca}]$, being a collection of k orthogonal unit normalised eigenvectors. The robust PCA (rPCA) scores are then computed using the linear operation,

$$Y_{rPCA} = XU_{rPCA}, \quad \{Y_{rPCA} = [\mathbf{y}_1, ..., \mathbf{y}_N]^T; \mathbf{y}_n \in \mathbb{R}^k\} \tag{4.6}$$

where, X is the DSFs matrix. The rPCA Matlab implementation is provided as part of the LIBRA toolbox, developed by the Robust Statistics research group at KU Leuven [9].

4.2.3 The minimum entropy deconvolution for impact-type of damage

In the BSS problem, which was presented in Chapter 1, Equations 1.1 and 1.2 described the process of convolution for an LTI MIMO system. Assume that this system has only a single excitation force $s_i(n)$, which generates the corresponding response $x_j(n)$ (in discrete-time n). Note that the response is corrupted by white noise $\varepsilon(n)$. The *aim* of the minimum entropy deconvolution (MED) technique developed by Wiggins [185], is to construct an inverse FIR filter f_c (with $c = 1, ..., C$ being its coefficients) for the system's IRF $h_{i,j}$. Such a filter can then be used to obtain the output signal $y_i(n)$ being as close as possible to the input excitation $s_i(n)$. This is illustrated in the diagram in Figure 4.2.

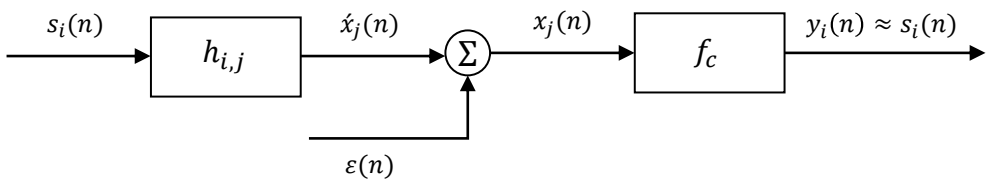

Figure 4.2. Block diagram of the inverse filtering process of the MED technique.

More specifically, the inverse filter f_c of the MED *deconvolves* the effect of the transmission path $h_{i,j}$ from response signal $x_j(n)$, by *assuming* that the source $s_i(n)$ is an *impulse-type of signal*. This will be the case, for instance, when an REB is damaged, generating a signal that has impulsive character (along with other components, e.g. periodicities). In practise, this is implemented by finding the maximum of the *kurtosis objective function* O_k of the output signal $y_i(n)$, by varying the coefficients of the filter f_c:

$$O_k(f_c) = \frac{\sum_{n=1}^{N} y_i(n)^4}{\left[\sum_{n=1}^{N} y_i(n)^2\right]^2} \tag{4.7}$$

where, the number of the coefficients is a user supplied parameter. Note that MED solves this problem iteratively, where in each iteration it changes the filter values until they reach certain convergence criteria, as outlined in [33]. The Matlab implementation developed by the authors in [125] is used. Typically, around 100 coefficients are used, as determined by trial-and-error using a grid search approach for a variety of problems used in this thesis. Of course, this is a crude approach and further enhancement of the kurtosis may be obtained by using it together with a genetic algorithm.

The MED technique, has been applied in CM for *enhancing the impulses* generated by damage in REBs [160] and in gearboxes [64]. Although there have been many similar works in the literature, these two papers, were the first to introduce MED into CM applications. Most of works in the literature demonstrated considerable improvements in identifying early impact-type of damage buried in noise and other more dominant sources of energy. More recently, in [199] an alternative to the MED technique, called multi-point optimal minimum entropy deconvolution, was implemented for CM in REBs and claimed success.

For vibration-based SHM applications, e.g. DI in opening and closing cracks, the author is not aware of any studies that implemented MED into their strategies, for unmasking signals to identify impact-type of damage.

4.3 The framework of deriving vibration-based DSFs

This section outlines the methodology used to derive DSFs from time series acceleration datasets. Apart from the Z24 dataset[¶], all other experimental and numerical data are in the form of accelerations, sampled in different rates: from 322.6 Hz (the LA-4DOF dataset) up to 51.2 kHz, and for a few tens of seconds to a few minutes. This yielded time series vectors in the tens of millions of data points, for each condition, e.g. for AFR3 in the TF-LBO dataset. Therefore, given that this thesis will not use techniques like DL, that are designed to handle large sets of sampled time series data (e.g. LSTM networks), and the fact that such high-sampling rates are not necessary for SHM & CM applications, a *data partitioning* procedure was first used.

In particular, the process of extracting DSFs starts with partitioning each univariate vibration time series into N equal-length segments, with $n = 1, ..., N$ the segment index. Each time series segment is chosen as a short-time non-overlapping window, e.g. 1-second for TF-LBO/-SAF & TS-SAF datasets, offering a good trade-off between frequency- and time-resolutions. For each nth time series segment $x_n(l)$, consisting of L data points ($l = 1, ..., L$), either: 1. a set of IMFs was calculated or 2. amplitudes from the power spectrum / PSD

[¶]The NFs of the Z24 dataset were estimated using the SSI method [141], [142].

were obtained. In the latter case, $x_n(l)$ was band-pass filtered within the frequency band of interest, e.g. for the LA-4DOF dataset $[30, 70]$ Hz to cover the first three NFs. Additionally, if the number of resulting amplitudes from the PSD was high (e.g. more than 100 - 100-dimensional problem) the data points were typically downsampled. Therefore, each nth segment, which is a univariate time series of length L samples, is converted into a *p-dimensional set of DSFs* versus sample number n. Note that as N gets larger (as consequence of choosing smaller time segment intervals), some algorithms that rely on inverting $N \times N$ matrices, as in the case of kernel methods (see Appendix B), will increase in computational time considerably. Therefore, the segment length L will be adjusted accordingly and will be specified explicitly for each dataset, throughout the thesis.
Following the data partitioning stage, *three* different types of DSFs were derived:

1. *Impact-based DSFs*: Changes related to the *peakedness* of the distribution or a system's *entropy*, measured through kurtosis and relative entropy, respectively;

2. *Modal-based DSFs*: Tracks changes in *NFs* and *DRs*, either by estimating these parameters directly, i.e. using the MPI procedure (see Chapter 2) or by selecting a set of amplitudes from the PSD;

3. *Global-based DSFs*: Monitoring is performed using a multiple set of "global" DSFs: statistical, modal and information-theoretic measures. Each global DSF captures different aspects of the system's characteristics, with the purpose of detecting *global changes*.

The process that is followed to calculate these types of DSFs is shown schematically in Figure 4.3. The estimated NFs and DRs, $\boldsymbol{\omega}_n$ and $\boldsymbol{\zeta}_n$, respectively, both in $\mathbb{R}^M$, are modal-based DSFs, along with T amplitudes (band-pass filtered within a specific frequency band) obtained from the power spectrum $\mathbf{s}_n$. For impact-based DSFs, the vector $\mathbf{p}_n \in \mathbb{R}^M$ are the sample Kurtosis (set $l = 4$ in Equation 2.3) or differential entropy (see Table 4.2) values, calculated from each IMF $c_n(t) \in \mathbb{R}^M$ that has either been filtered by MED first or used in its raw version. Whereas, the vector $\mathbf{q}_n \in \mathbb{R}^{D_2}$ includes global-based DSFs. So that, for each dataset considered, a different type of DSFs may be used, as follows:

1. LA-4DOF, NSim-4DOFLin & NSim-4DOFNonLin: 1. Modal-based DSFs $\mathbf{s}_n$, $\boldsymbol{\omega}_n$, $\boldsymbol{\zeta}_n$ to monitor changes in vibration modes and 2. impact-based DSFs $\mathbf{p}_n$, as damage is in the form of impacts;

2. WT-REB: 1. Impact-based DSFs $\mathbf{p}_n$ to monitor impulsive signals due to REB damage and 2. amplitudes from EA;

3. TF-SAF/-LBO & TS-SAF: Global-based DSFs (MD_2) $\mathbf{q}_n$ to monitor general GT engine changes.

Note that using the adaptive decomposition methods, EMD-/VMD-based and ALIF, the number of IMFs $\tilde{M}$ extracted between each nth segment may differ (apart from the VMD). To construct the DSF matrix, only M IMFs were selected, according to the *decision rule*: 1. find the minimum number of M IMFs by considering all N segments (i.e. $\mathbf{c}_n \in \mathbb{R}^M \ \forall \ n = 1, ..., N$), 2. select the ones with the maximum value of kurtosis $\times$ RMS value. The latter value has been used to identify both the significant IMFs and, in the case of impact-type damage, the ones that are most impulsive.

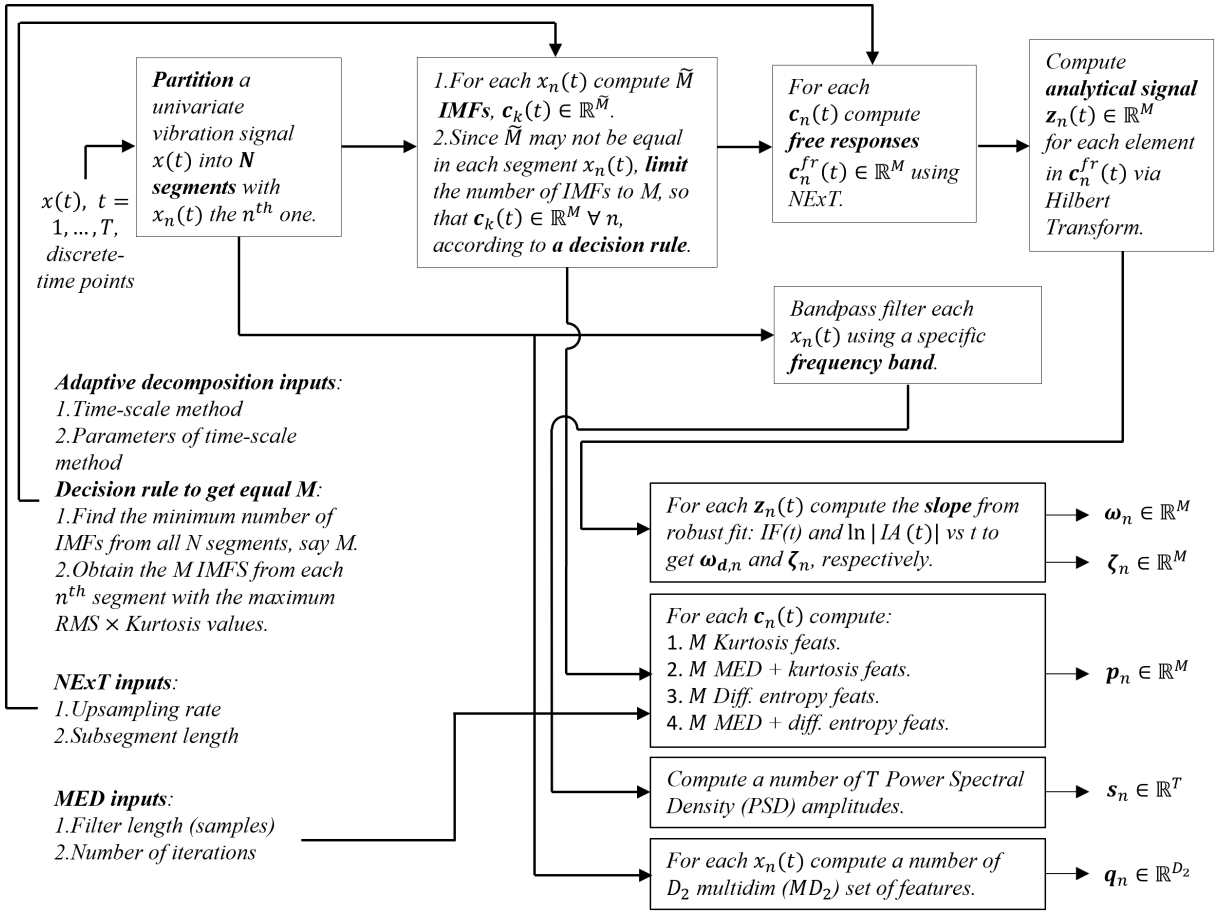

Figure 4.3. Schematic outlining the main components of the framework used to derived vibration-based DSFs.

In terms of the bandpass filter specification the upper and lower cut-off frequencies, used to compute $\mathbf{s}_n$, varies between each dataset. From Chapter 3, the frequency band of interest for the LA-4DOF dataset (see Figure 3.6), where the amplitude changes with respect to the *damage level*, but, not with respect to the EOV change, is from around 100 Hz and onwards.

Due to the relatively low sample rate, $\mathbf{s}_n$ is a 191 element vector, so no downsampling is performed. In comparison, for the NSim-4DOFLin/-4DOFNonLin datasets the acceleration time series is filtered using a low-pass cut-off frequency of 0.12 Hz. As such, tracking changes in these four NFs, results in $\mathbf{s}_n$ being a 230 element vector, which may need to be downsampled for some of the methods used in the next part of the thesis (not all methods can handle high-dimensional features efficiently).

Moreover, it is expected that using adaptive decomposition methods, the non-linear system dynamic characteristics will be better represented than, for instance, using amplitudes from the spectral lines. Note that each nth segment is only a few seconds in length, so that non-stationarity due to EOV in structures may be negligible. On the other hand, non-linearity can be expected to cause "significant" changes in the time-scale of a few seconds, given a high-enough excitation force. For instance, non-linearity in the LA-4DOF dataset takes place due to impact-type of damage, while for the NSim-4DOFNonLin dataset both impact-type of damage and non-linear restoring force will cause it to behave non-linearly. However, for the WT-REB, TF-SAF/-LBO and TS-SAF datasets, both non-stationarity and non-linearities may be "significant" within each segment, even when its time-scale is only a few seconds in length. This is due to the *abrupt nature* of the changes in operating conditions, e.g. generator loadings.

4.4 Modal-based DSFs for LA-4DOF & NSim-4DOFLin/-4DOFNonLin datasets

The NFs and DRs of LA-4DOF and NSim-4DOFLin/-4DOFNonLin datasets were estimated using the MPI procedure in Table 2.5, by following the framework that was shown in Figure 4.3. Firstly, the estimated ω_d (NFs), before and after a locally weighted linear regression smoothing filter is applied[‖], are shown in Figures 4.4 and 4.5.

Although there is some discrepancy between the estimated and actual NFs, this is rather unimportant for DI applications. What is needed from these DSFs, is high-sensitivity with respect to damage and its progression levels. As can be seen, VMD and CEEMDAN on the NSim-4DOFLin, show promising results on the NF estimates, in terms of DI. For the LA-4DOF dataset, there is some sensitivity due to damage (region G4 until the end), for instance, the sudden drop in NFs as seen from the plots of the SVMD and VMD methods. However, EOV influences NFs comparably in both magnitude and patterns observed. In regions where there is significant change in EOV, e.g. when $T \leq 0$ in the NSim-4DOFLin dataset (sample no between 200 and 300), EOV influence is more dominant than damage. In

[‖]The estimated parameters from the MPI were "noisy", so that a smoothing filter was applied.

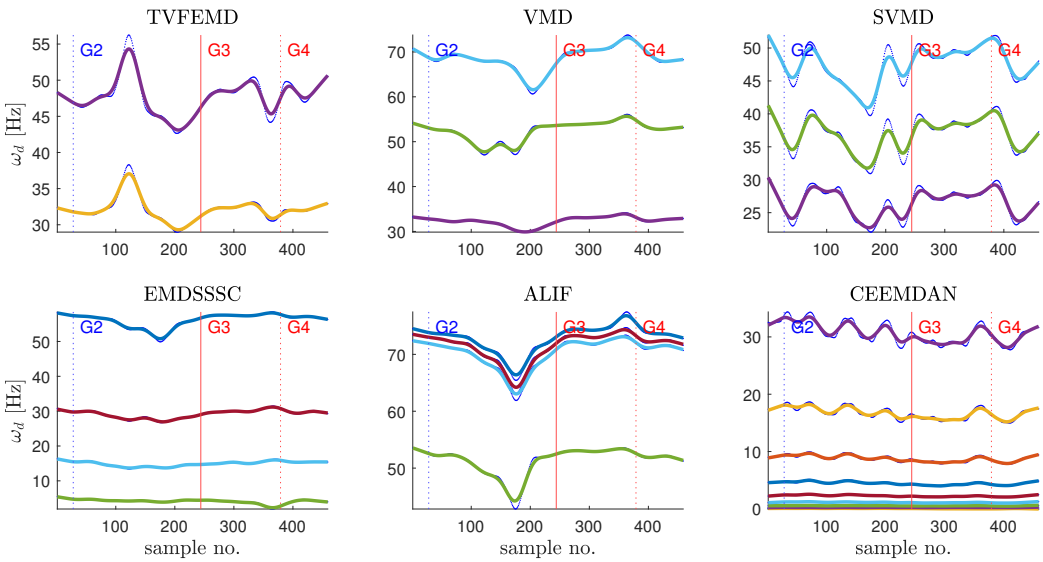

Figure 4.4. NFs estimates from MPI, for the LA-4DOF dataset. Regions G2 to G3: undamaged state + EOV changes, regions G3 to G4: damaged state (and its progression levels) and G4 to the end is the system in its damaged only state. Segment time interval is equal to 8 seconds.

these cases, it may mask data points representing the damage. In general, as NFs estimates showed, there will be potentially no *significant distinguishing feature* that a statistical model may use to discriminate between damaged and undamaged system states. Note also that from further testing these sets of DSFs, i.e. using the DI techniques presented in Chapter 5, it was found that they were unsatisfactory for the purposes of achieving accurate and reliable system monitoring. Moreover, in both ALIF- and CEEMDAN-based NFs estimates, the mode mixing phenomenon is readily observed. For instance, in the ALIF-based estimates, the higher-frequency (around 70 Hz) has been found in three different IMFs. Monitoring the same vibration mode in more than one DSF is undesirable as it introduces redundancy in the features (see Chapter 1).

In contrast, the DR estimates, shown in Appendix D in Figures D.1 and D.2 for the same two datasets, have shown negligible sensitivity to EOV. Unfortunately, apart from the ALIF-based DR estimates, this is true for damage, too. This is due to the fact that significant smoothing had to be applied on the DRs estimates using the six decomposition methods, which rendered them relatively flat. This extensive smoothing was necessary, however, since the "raw" estimates were particularly noisy as a result of *both* IF and IA fast variations.

The above analysis, yielded modal-based DSFs that are either sensitive to EOV influence *and* damage or insensitive to both of those parameters. In particular, DRs have shown to be particularly unsuitable as DSFs for DI in these two datasets. Note that similar outcomes were observed on the NSim-4DOFNonLin dataset, too.

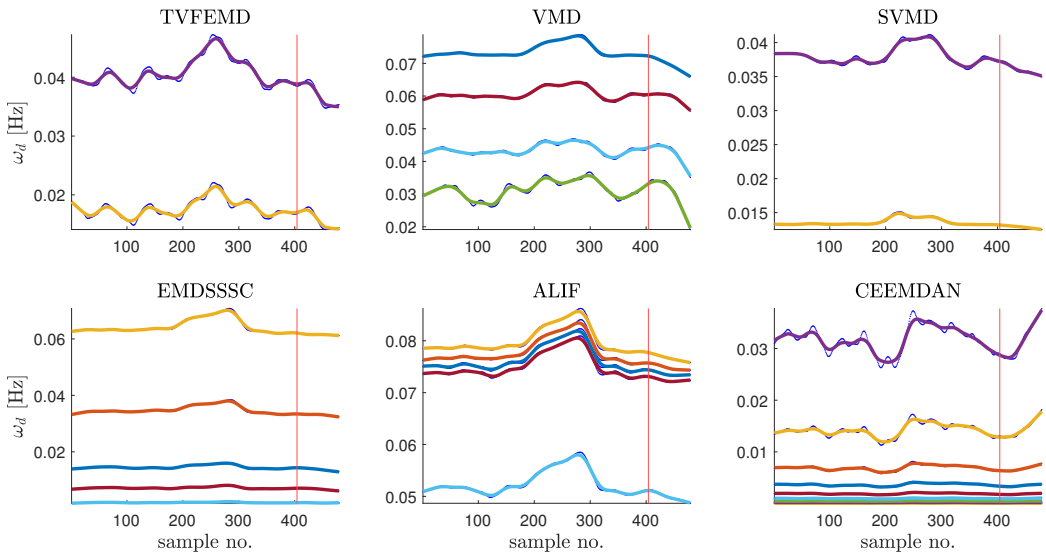

Figure 4.5. NFs estimates from MPI, for the NSim-4DOFLin dataset. Vertical line shows the first sample from the damaged state. Segment time interval is 7.5 hours.

The amplitudes from the PSD and their robust z-scores, $\mathbf{s} \in \mathbb{R}^T$, for these three datasets, are shown in Figure 4.6. For the LA-4DOF dataset, the data were band-pass filtered between 100 Hz and 140 Hz, while for the NSim-4DOFLin/-4DOFNonLin datasets, the data were low-pass filtered below .1 Hz. For the former, the higher frequency region has shown marginal influence from EOV (e.g. between around sample no. 700 and 800), while for the latter it is used to track changes in NFs, which consequently show more EOV influence. However, when the robust z-score is applied to normalise these DSFs, damage is more readily inferred in all three datasets. This is especially true for the NSim-4DOFLin/-4DOFNonLin datasets, where damage has been masked by significant EOV influence, although some influence may still be relatively high, e.g. between around sample no. 400 and 550 when $T \leq 0$.

The amplitudes from the PSD are high-dimensional features that need to be subjected to dimensionality reduction or downsampling prior to further analysis. Note that, although it is not important for the purposes of demonstrating the different DSFs, the number of samples was increased by a factor of more than 3 for these three datasets in Figure 4.6. This is because a smaller length was used for each segment (e.g. from about 8 seconds to 2.5 seconds for LA-4DOF), which resulted in a higher number of partitions N.

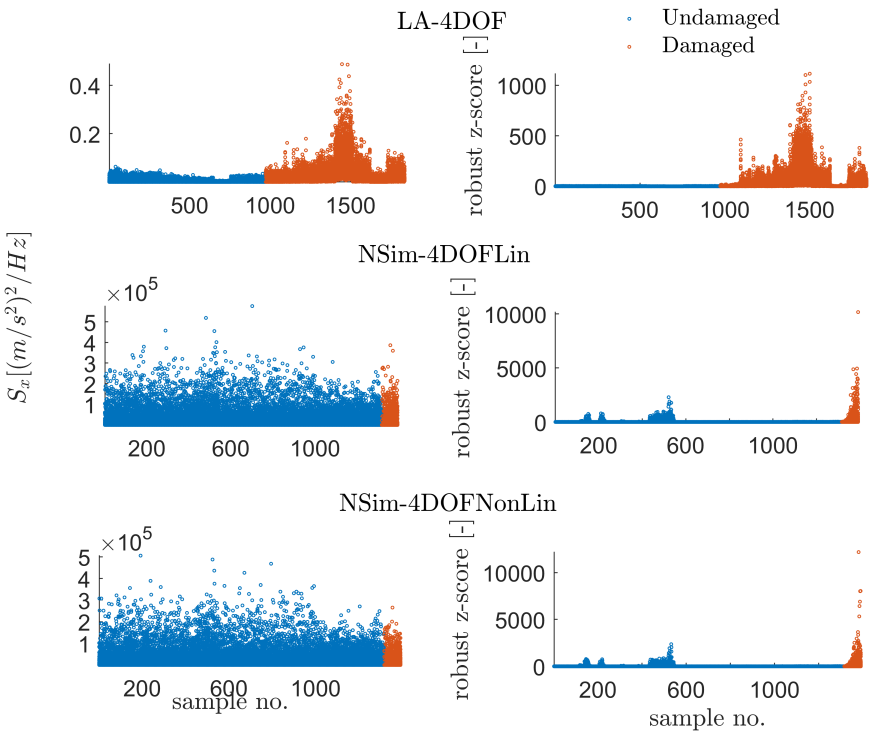

Figure 4.6. LA-4DOF dataset: 191 amplitudes from the PSD and corresponding robust z-scores. NSim-4DOFLin/-4DOFNonLin datasets: 230 amplitudes from the PSD and their robust z-scores. LA-4DOF is equal to 2 seconds and NSim-4DOFLin/-4DOFNonLin is equal to 2.5 hours.

4.5 EA-based DSFs for WT-REB dataset

4.5.1 DSFs from the EA procedure on a simulated example

The EA procedure, as presented in Chapter 2, was demonstrated in Appendix E on a simulated example representing a simplified version of a fixed-rotating machinery. This included a periodic rotating component and its harmonic series and an REB signal with and without outer race damage and white noise.

As seen in Appendix E, the procedure yielded the *spectral bandwidths* that *maximised the SK*. Therefore, in order to further enhance the impulsiveness of the signal, an FIR band-pass filter was specified with these parameters. Then, the HT is used to obtain the analytic signal representation and its envelope is subsequently calculated as its absolute value. The envelope spectrum is then calculated, as the PSD of the envelope signal. As can be seen in Figure 4.7, the deconvolved with MED AR residual signal shows the BPFO and at least 5 harmonics very clearly in the spectrum. In comparison, the envelope of the AR residual contains no amplitude at BPFO. While, the raw signal that was filtered with the MED contains the fundamental shaft frequencies and its harmonics, as well as white noise, in

addition to a weaker BPFO component (masking effect of the DSF). On the other hand, there seems to be only white noise in the envelope spectrum of the original signal, signifying the importance of pre-processing in damage detection of REBs. Note the frequency smearing effect of the 2% in rolling slippage is also evident on the envelope analysis spectrum of the MED filtered AR residual signal. Hence, these sets of DSFs show particular promise in early DI, even in the presence of high-amplitude white noise and weak components at BPFO relative to the rest of the periodicities. What remains to be seen, is its performance under changing operating conditions.

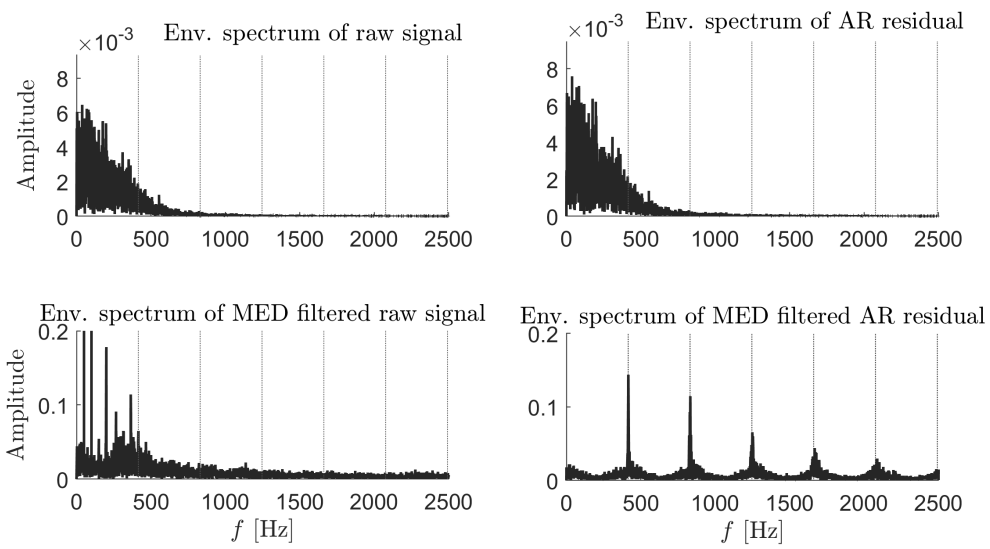

Figure 4.7. Envelope spectra, shown between 0 and 2.5 kHz on the four signals considered. The enhancement of the impulses at BPFO is evident when the signal is both pre-processed with an AR model and MED, i.e. bottom-right signal. The BPFO cursors are marked as ':' in the plots.

4.5.2 DSFs from the EA procedure on the WT-REB dataset

The EA procedure was also applied on the WT-REB dataset in order generate a set of DSFs. The earliest damage level that could be highlighted was damage level 2 and the results are shown in Figure 4.8 for the no load (C0) and highest load operating conditions (C3). In comparison to the "raw" power spectrum in Figure 3.10, no indication of damage, e.g. amplitudes at BPFO cursors, was shown. Additionally, as indicated in the four plots in Figure 4.8, the amplitude of the envelope spectrum of the MED filtered AR residual $g(\varepsilon_{\ddot{x}})$ is the highest, when looking at the BPFO and its harmonics. This is due to the fact that MED further enhanced the impulsive character of the signal output, after it has been pre-whitened with an AR model of order $p = 95$.

In contrast, applying an AR on the raw vibration signal $\ddot{x}$, produced only a relatively

weak amplitude at the BPFO. On the other hand, applying an MED on the raw vibration signal $g(\ddot{x})$ (bottom left plot), produced only half of the peak amplitude as compared to the MED on the AR residual $g(\varepsilon_{\ddot{x}})$. The impulses generated after applying MED on either the raw or AR residual signals, increase by approximately 40%. Hence, it can be concluded that MED is an invaluable technique for identifying damage early in REBs. Note that the small deviation from the BPFO and its harmonic series as the load increases is due to the reduction in rotational speed of the main shaft**. Therefore, changes in generator load impose an additional challenge on DI, as it modifies the spectral characteristics of the vibration signal (in addition to the statistical characteristics that were seen earlier in Figure 3.9).

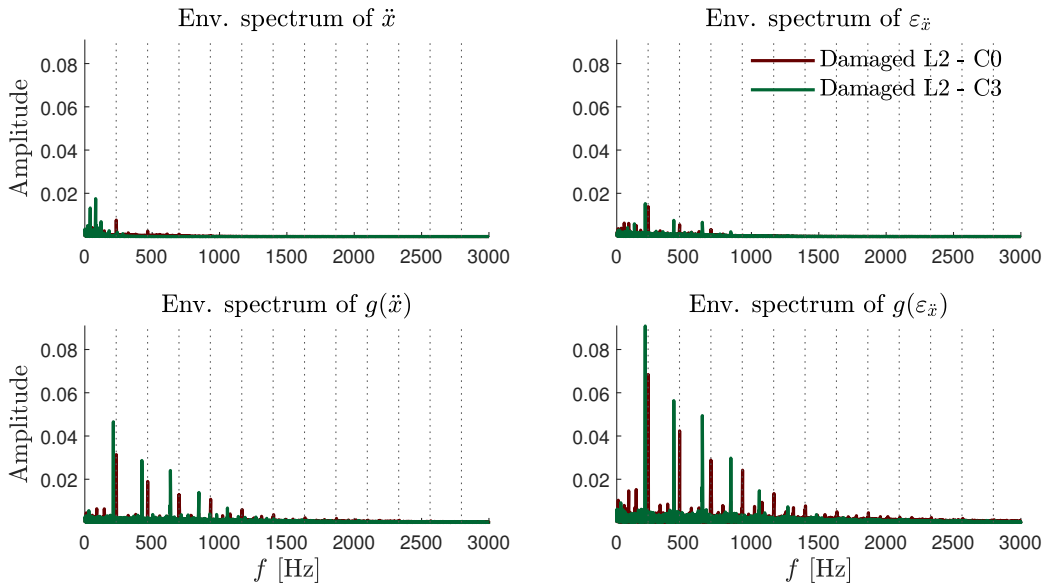

Figure 4.8. Envelope spectrum of damage level 2 for C0 and C3.

Although the EA procedure is sensitive to changes in operating conditions, incorporating the MED technique is particularly useful at the earliest damage levels. However, when the damage has significantly progressed, any pre-processing / filtering on the raw signal will reduce the amplitudes at the BPFO (see Appendix E).

4.6 Impact-based DSFs for LA-4DOF, NSim-4DOFLin, NSim-4DOFNonLin & WT-REB datasets

Impact-based DSFs $\mathbf{p} \in \mathbb{R}^M$, both using kurtosis and differential entropy, were derived by following the framework shown in the diagram in Figure 4.3. The main premise of using

**This can be explained from the fact that at higher generator electrical load demand, the strength of the magnetic field in the armature increases. This in turn, generates an opposing torque on the shaft, causing rotational speed reduction.

differential entropy as DSFs is to quantify for the *uncertainty* that is being observed in the vibration responses. In REBs, this quantity will reach its maximum value when the vibration energy is evenly distributed (as seen from the spectrum), which implies more system uncertainty. This corresponds to a system in its undamaged state. However, when there is damage, the vibration energy will be more concentrated around the resonances, which will make the differential entropy to decrease, which implies more *order*. In structures, the differential entropy will rise when a system changes from linear to non-linear. This is referred to as Axiom VIII in SHM, which which states that *"Damage Increases the Complexity of a Structure"* [67].

These DSFs were, firstly applied on the raw vibration signal for each nth segment. Kurtosis has been found to be particularly sensitive to damage for these four datasets, since it manifests itself in the form of impacts. Figure 4.9 shows the computed kurtosis values μ_4 applied on all N partitions of the raw vibration signal. As the plots on all four datasets demonstrate, the impulsivity of these kurtosis values has been further enhanced by using the MED filter $g(\mu_4)$[††]. What is most important, is the fact that these DSFs show the damage progression level on all four datasets from its earliest point of initiation (vertical red line), while being relatively insensitive to EOV influence. In terms of using differential entropies, the results did not show such desirable properties as kurtosis.

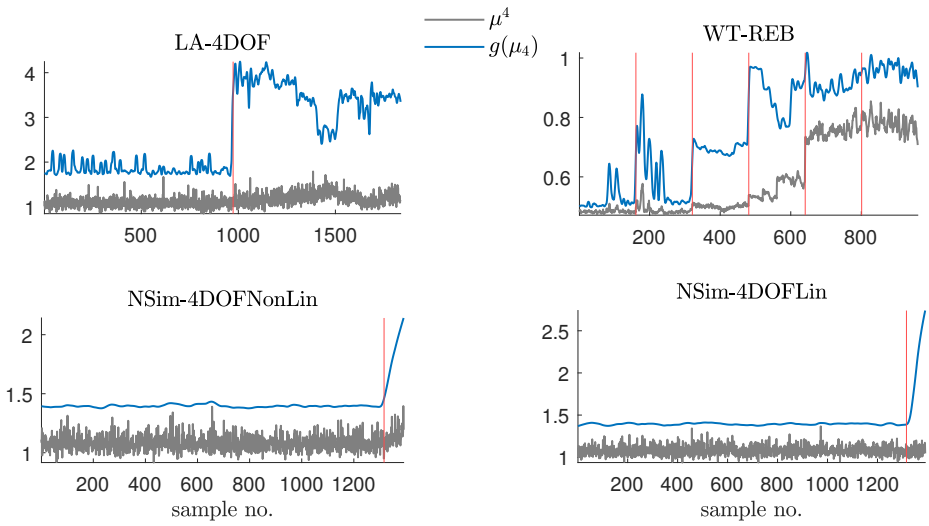

Figure 4.9. Kurtosis μ_4 and MED filtering on kurtosis $g(\mu_4)$ on four datasets. Vertical red line signifies the point of damage initiation, while for the WT-REB dataset each red line denotes a damage level progression. Segment lengths for: WT-REB is equal to 2.5 seconds.

Referring to Figure 4.9 the following should also be noted for each dataset:

[††]Due to the fact that MED filtering results in noisy outputs from the signals, the results shown in Figure 4.9 are after a locally weighted linear regression smoothing filter was applied.

- LA-4DOF: Undamaged system: sample no. $0-108$, undamaged system and EOV influence: sample no. $109-976$, damaged system: sample no. $977-1516$ and damaged system and EOV influence: sample no. $1517-$;

- NSim-4DOFLin & NSim-4DOFNonLin: Undamaged system and EOV influence: sample no. $0-1315$ and damaged system and EOV influence: sample no. $1316-$;

- WT-REB: Undamaged system: sample no. $0-160$, damaged system (level 1): sample no. $161-320$, damaged system (level 2): sample no. $321-480$, damaged system (level 3): sample no. $481-640$, damaged system (level 4): sample no. $641-800$ and damaged system (level 5): sample no. $801-$. For each of the six system states there is a progressive increase in generator load for every 40 samples. For instance, sample no. $0-40$: no generator load (C0), sample no. $41-80$: first level of generator load (C1), sample no. $81-120$: second level of generator load (C2) and sample no. $121-160$: third level of generator load (C3). The same pattern repeats for each of the subsequent five system states.

Similarly, for each decomposed IMF the kurtosis values were computed for each of the N segments. The most promising results, in terms of damage-sensitivity and EOV influence immunity, for the LA-4DOF dataset, are shown in Figure 4.10. Most remarkably, using MED filtering applied on the kurtosis values for each IMF, the differences between damaged and undamaged system states have been made evident from the plots (no statistical model may need to be applied further to discriminate between the two states). This is especially true for the VMD-based methods, while for the TVF-EMD, the two IMFs are shown to be gradually diverging from each other as damage progresses. Also, for TVF-EMD the minimum number of IMFs was two, while for VMD-based it was three.
Therefore, instead of using a univariate DSF, e.g. $g(\mu_4)$ from the raw vibration signal, it is possible to compute kurtosis values for each IMF separately. As it is shown, for instance, on the SVMD plot in Figure 4.10, each IMF has a slightly different pattern - showing different aspects of our data. What is more, using multivariate DSFs will offer diagnostic capability, i.e. to gain insights from each dataset, as it is demonstrated in Chapter 5.
The two VMD-based methods have also shown particularly promising results on the WT-REB, too. In comparison to the EA method, these sets of DSFs are able to identify the earliest signs of damage (damage level 1), as seen in Figure 4.11. At that damage level, the impulses at BPFO are relatively weak, relative to the overall vibration signal. Therefore, a subsequent increase in load, i.e. load C2 and C3, weaken them further, which in turn causes negligible rise in $g(\mu_4)$. Moreover, in the SVMD method, the IMF 2 shows steady increase in its $g(\mu_4)$ value as the damage progresses, with insignificant influence due to load

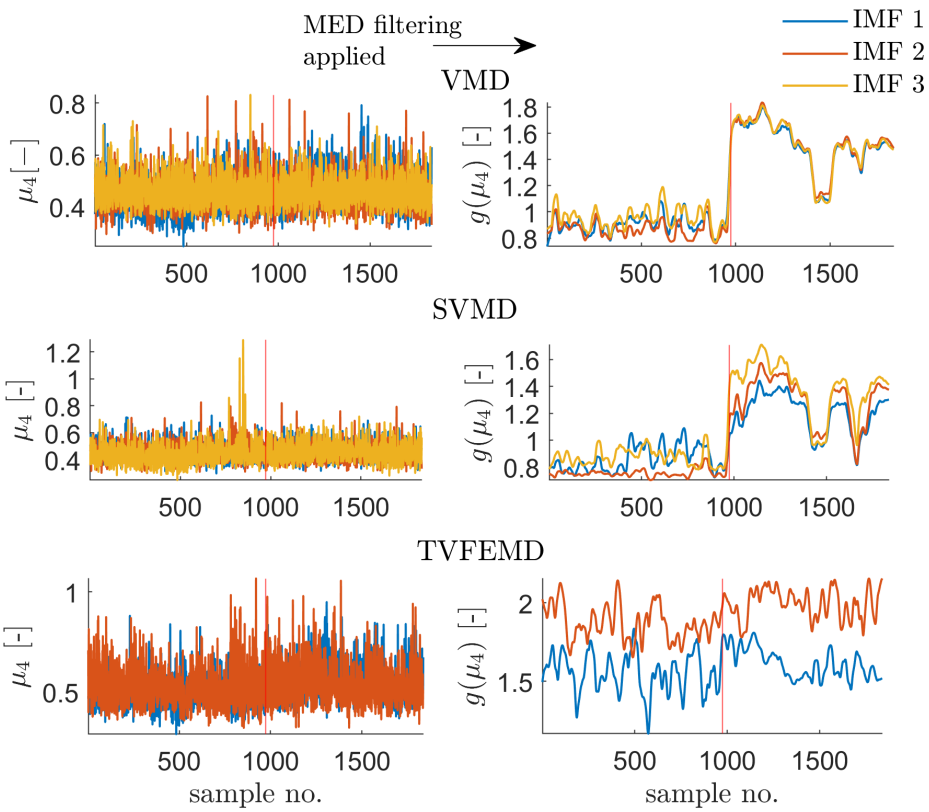

Figure 4.10. Impact-based (kurtosis) DSFs on each IMF decomposed signal of three different adaptive decomposition methods on LA-4DOF dataset.

changes. Hence, using this multivariate set of DSFs, instead of the univariate $g(\mu_4)$ on the raw vibration signal, the monitoring strategy will benefit, since different data characteristics are obtained. In this case, it will be the subspace of the three different IMFs considered using these two VMD-based methods.

4.7 Global-based DSFs for TS-SAF, TF-LBO/-SAF datasets

A total of 26 DSFs were computed for each of the N partitions (with a time-interval of 2 seconds for each nth segment) obtained from the TS-SAF, TF-LBO and TF-SAF datasets. In particular, the dimension of the DSFs matrix for TF-LBO is 117×26, for the TS-SAF is 36×26 and for the TF-SAF is 665×26. The multidimensional set of features MD_2 combine different dataset characteristics, whereby some of them may be irrelevant to the DI problem. This issue can be addressed using a dimensionality reduction approach, for instance, using PCA to obtain uncorrelated DSFs, thus, eliminating redundancy further in the process of DI. The main purpose of MD_2 DSFs is to enable monitoring of system changes, which are largely highly-unpredictable and unknown. This is in contrast to impact-based DSFs, in which a specific type of damage (impacts or impulses) is expected. Following, the procedure

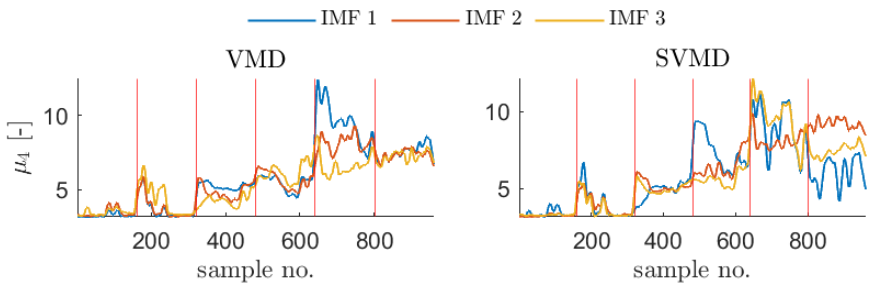

Figure 4.11. Impact-based (kurtosis) DSFs on each IMF decomposed signal of two different adaptive decomposition methods on WT-REB dataset. Vertical red lines denote the damage level progression.

outlined in Figure 4.3, the $MD-2$ DSFs comprise of 26 different features, are shown in Table 4.2.

Table 4.2. The MD_2 set of global-based DSFs.

Feature	*Formula*
Mean band power	$\mu_f = 1/m \sum_m S_x(m)$.
Spectral entropy	$\mathcal{E}_s = -\sum_m \frac{S_x(m)}{\sum_m S_x(m)} \ln(\frac{S_x(m)}{\sum_m S_x(m)})$, with m the spectral bin and $S_x(m)$ the PSD amplitude (Equation 2.5).
Spectral centre	$\tilde{s}_\mu = \frac{\sum_m f_m S_x(m)}{\sum_m S_x(m)}$, $f_m : m^{th}$ bin frequency.
Spectral spread	$\tilde{s}_2 = \sqrt{\frac{\sum_m (f_m - \tilde{s}_\mu)^2 S_x(m)}{\sum_m S_x(m)}}$.
l^{th}-order spectral moment	$\tilde{s}_l = \sqrt{\frac{\sum_m (f_m - \tilde{s}_\mu)^l S_x(m)}{\tilde{s}_2^l \sum_m S_x(m)}}$.
Spectral crest	$\tilde{s}_{cst} = \max_m(S_x(m)) / (\sum_m S_x(m))$.
Spectral peak	$\tilde{s}_{pk} = \max_m(S_x(m))$.
Differential entropy	$\mathcal{E}_\Delta = -\sum_n p_k(x)\Delta_x \ln(p_k(x))\Delta_x - \ln(\Delta_x)$.
l^{th}-order central moment	See Equation 2.3, i.e. $l = 2,3,4$.
ARMA coefficients	See Equation 2.6.
Max. log squared difference	$\max \Delta_{\log,x^2} = \log((x_n^2(t) - x_n(t+1))^2), t = 1,...,T-1$
Mean Teager-Kaiser energy	$mean(\psi) = (x_n^2(t) - x_n(t+1)x_n(t-1)), t = 2,...,T-1$.
Sum of sequential variation	$\log \tilde{j}_{sv} = \sqrt{\sum_t (x_n(t) - x_n(t-1))^2}$, $\forall\ t = 2,...,T$.

To calculate the differential entropy $\mathcal{E}_\Delta$ the method from [30] is used. In particular, different quantisation or discretisation approaches of the data need to be considered, e.g. choice of appropriate bin width, which will result in different values of $\mathcal{E}_\Delta$. In this thesis, the differential entropy is calculated by first discretising the data $x_n(t)$ into equal sized bins, using an automated algorithm (*histcounts.m* function in Matlab). Given the bin size Δ_x, and the empirical probabilities for each nth bin $p_n(x)$, the expression for computing the differential entropy $\mathcal{E}_\Delta$ is given in Table 4.2. For the spectral entropy $\mathcal{E}_s$ the frequency bin width is

determined in the power spectrum calculation.
The range of p and q parameters for the ARMA model has been selected based on an initial visual examination of the partial and autocorrelation functions, which provided an approximation on the number of lags for an AR and MA process (more details in [132]). The order of the ARMA model is selected via the minimum of the Bayesian Information Criterion (BIC), i.e. $\min_{p\in\{0,1,...,12\},q\in\{0,1\}}\left[-2\ln\hat{L}(\boldsymbol{\theta}|x)+\ln(N)(p+q)\right]$. Note that $\hat{L}(\boldsymbol{\theta}|x)$ is the optimised log-likelihood estimate of the ARMA model coefficients vector $\boldsymbol{\theta}=\{p_1,...,p_{12},q_1\}$. Since an ARMA model assumes stationarity in the data generating process, a unit root test (*kpsstest.m* Matlab function) is used for testing for stationarity. If a time-series $x_n(t)$ is found to be non-stationary using the KPSS test [106][‡‡], a first-order differentiation takes place before the number of ARMA coefficients are selected using the BIC value.
The different characteristics that each MD_2 feature computes are as follows:

- Spectral entropy: signal complexity in the power spectrum;
- Mean band power: concentration of energy in the power spectrum;
- Spectral spread: signal variation in the power spectrum;
- l^{th} spectral moment: for $l=3$ this is spectral skewness representing symmetry around the mean band power, and $l=4$ is spectral kurtosis representing peakedness of the power spectrum;
- Spectral crest: a measure of the peakedness of the power spectrum;
- Spectral peak: amplitude of the peak power;
- Differential entropy: as discussed previously;
- ARMA coefficients: measures temporal dependencies between the samples;
- mean Teager-Kaiser energy: computes the mean variation in signal energy using three adjacent sample sequences. Note that Teager-Kaiser takes into account changes in frequency and amplitude of the signal, such that any changes from either of those will lead to an increase in its value;
- Sum of sequential variation: computes the sum of variations between two adjacent signal sequences.

[‡‡]That is, by assessing the null hypothesis with a significance value of 0.05 that the series is stationary over a range of lags

The last two features, i.e. the mean Teager-Kaiser energy and the sum of sequential variation are commonly used features to track temporal changes in electroencephalogram signals, e.g. in [41]. For the purpose of improving the spectrum-based DSFs, e.g. spectral moments, it is first necessary to investigate the regions in which any *prominent changes* (either in amplitude increase or decrease or shift in frequency peaks) occur.

For the TF-SAF dataset, the main variation in PSD amplitudes is observed in frequencies approximately below 2.5 kHz (see Figure 3.13). For the TS-SAF dataset, variations due to changes in operating conditions are mainly concentrated in two frequency regions in the PSD. That is, the first region is between 8.5 kHz and 10.5 kHz and the second region is between 20 kHz and 21 kHz (see Figure 3.12). On the TF-LBO dataset, there are specific peaks in the PSD where the change in AFR is evident. As seen in Figure 4.12, there are two prominent *common peaks*, i.e. at 0.26 kHz and 0.67 kHz, which increase in amplitude with the AFR. A third peak is also noticeable for AFR 2 at 8.1 kHz and for AFR 3 at 8.7 kHz. Note that above 10 kHz there was no observable change in the PSD. Two frequency regions can then be defined: the first being below 1 kHz and the second being between 8 kHz and 9 kHz.

To simplify the analysis, the following frequency regions are considered for computing the MD_2 DSFs:

- TS-SAF: 8 $kHz \leq$ frequency region $\leq$ 21 kHz.
- TF-LBO: 0 $kHz \leq$ frequency region $\leq$ 9 kHz.

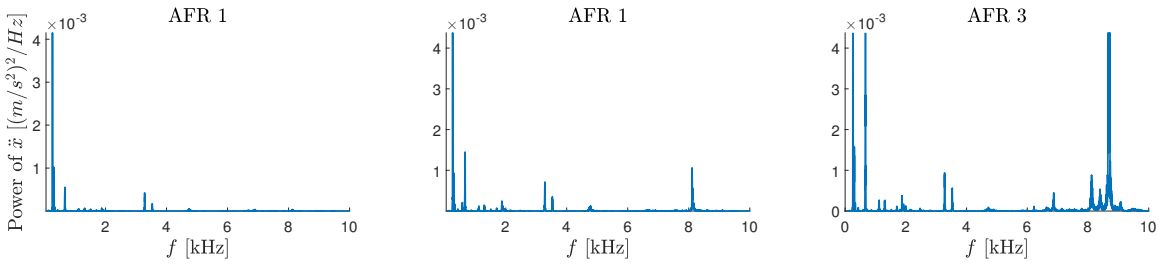

Figure 4.12. PSD estimates (zoomed in up to 10 kHz) for TF-LBO dataset.

In Figure 4.13, the variation of MD_2 DSFs with respect to the sample no., for each of the three datasets is shown. Each DSF was normalised using the z-score and its robust equivalent, which is necessary due to the wide range of scales of MD_2 DSFs. In order to gain an insight into the results that these plots are showing it is first necessary to account for the following:

- TF-SAF: Conventional jet fuel: $0-37$, SAF fuels: $38-$;

- TS-SAF: SAF @ AFR 1: sample no. $0-9$, SAF @ AFR 2: sample no. $10-18$, Jet-A @ AFR 1: sample no. $19-27$ and Jet-A @ AFR 2: sample no. $28-$;

- TF-LBO: AFR 1: sample no. $0-40$, AFR 2 : sample no. $41-80$ and AFR 3: sample no. $81-$.

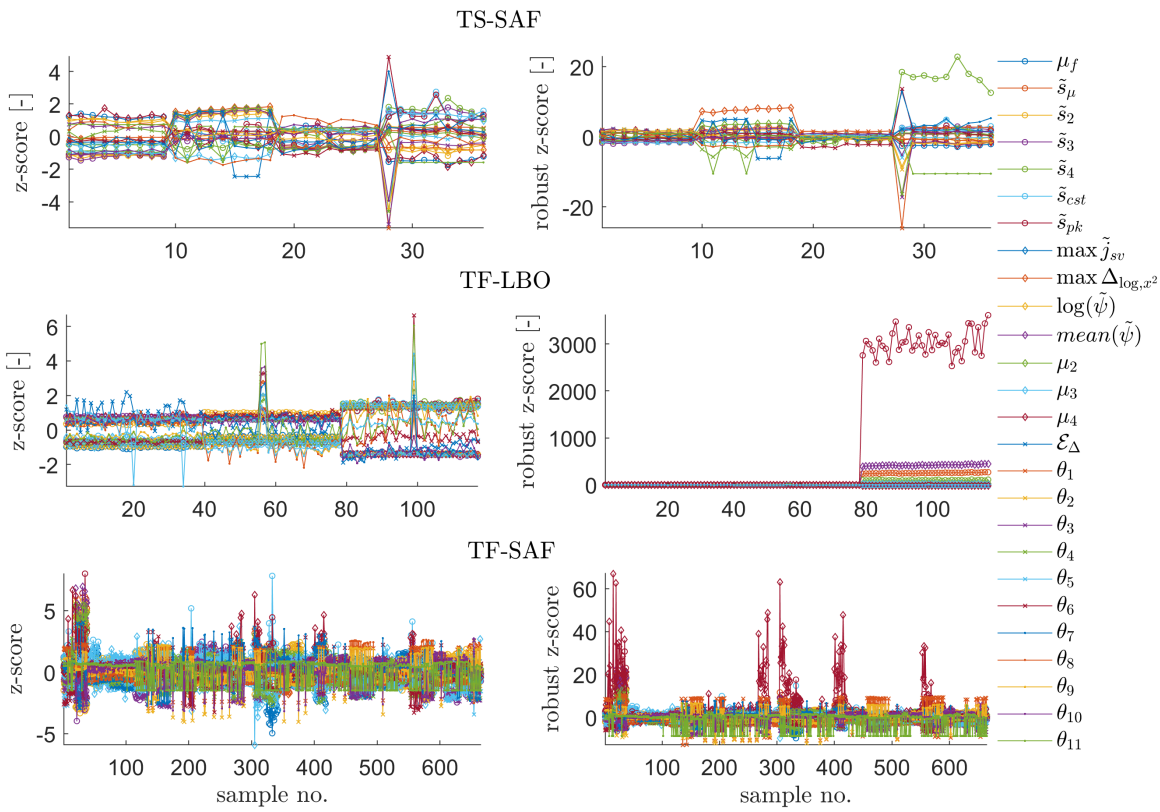

Figure 4.13. Multidimensional feature set MD_2 on TF-LBO (segment time interval is 10 seconds), TF-SAF (segment time interval is 2 seconds) and TS-SAF (segment time interval is 2 seconds) datasets.

Moreover, the following data classification / labelling has been used (it will be used later on in the subsequent chapters to quantify performance in DI):

- TF-SAF: Conventional jet fuel may be used to represent the undamaged[§§] engine state, while the rest are labelled as the damaged ones. This means that 2 out of a total of 35 engine states may represent the undamaged states;

- TS-SAF: Jet-A @ AFR 1 can be considered as the only undamaged state out of a total of four engine conditions;

[§§]Recall that damaged and undamaged for the GT engine datasets refer to abnormal and normal conditions, respectively. As explained in Chapter 1, this is to maintain consistency throughout the thesis when referring to the various system states across the various datasets.

- TF-LBO: AFR 1 and AFR 2 are considered as the undamaged states due to the fact that no LBO is taking place on these two engine states.

Returning to the plots in Figure 4.9, the AFR 3 engine condition on the TF-LBO dataset, shows significant differences from the other two conditions. And this is more clearly illustrated using the robust z-score method, whereby the spectral peak $\tilde{s}_{pk}$, the mean of the Teager-Kaiser energy $mean(\tilde{\psi})$ and the mean band power $\tilde{s}_{\mu}$ are the most discriminatory DSFs, as the AFR changes from AFR 2 to AFR 3. Such findings are consistent with the previous analysis using the four statistical moments in Figure 3.14. The ARMA coefficient θ_6 is also seen as an important one to infer changes between AFR 1 and AFR 2, where the flame is still stable (LBO limit is not reached). Note that the number of ARMA model coefficients, as decided from the BIC criterion, were 11 or 12 for the AR and 0 or 1 for the MA, i.e. for all datasets. For consistency, a zero was added to the matrix if the number of coefficients was not equal among the observations.
For the TS-SAF dataset, it may be shown that AFR changes result in more significant variations in the DSFs than the change from jet to SAF fuel. Looking at the robust z-scores, in particular, the spectral kurtosis $\tilde{s}_4$ shows to be particularly strong feature to discriminate between the other three engine states. The ARMA coefficient θ_{11} is also an important feature that can distinguish between Jet-A AFR 1 and AFR 2 conditions. Similarly, the variation in the the sample kurtosis μ_4 for the TF-SAF dataset, may be used to monitor changes in the fuel mixture.

4.8 Summary

In this chapter three different types of DSFs were examined and analysed on the datasets described in Chapter 3: impact-, modal- and global-based DSFs. Each one has its own particular characteristics that are suitable for different problems. Impact-based DSFs target damage that manifests in the form of *impacts*, while modal-based DSFs can be used to *track changes in NFs, DRs* and certain *frequency bands*. On the other hand, when no particular information is available a set of DSFs, integrating features from both frequency- and time-domains seems to be an attractive choice to identify changes that occur *globally within the system* being monitored. More specifically, the following can be deducted from this chapter:

- For the LA-4DOF, NSim-4DOFLin/-4DOFNonLin and WT-REB datasets, the impact-based DSFs, and especially the kurtosis value, together with the MED filtering of each IMF separately, provides a satisfactory discrimination among damaged and undamaged system states;

- In comparison to the rest of the adaptive decomposition methods, the IMFs decomposed using the SVMD, seems to be the most well suited for the impact-based DSFs in the three aforementioned datasets, when combined with kurtosis features. Generally, the combination between the type of DSF and IMF extraction method is an important one;

- As mentioned, the MED filtering is, in general, a particularly valuable pre-processing step to enhance impulses in a signal, so much so, that in some cases considered, e.g. the LA-4DOF dataset, the discrimination between the two system states increases dramatically. However, this means that it also enhances impulses in the signal from the undamaged system state, and thus, careful tuning is required. For instance, the length of the MED filter must be adjusted accordingly;

- Modal parameters (NFs and DRs) estimates for the LA-4DOF and NSim-4DOFLin/-4DOFNonLin datasets were extensively pre-processed using smoothing filters. Apart from the NFs estimates for the NSim-4DOFLin/-4DOFNonLin dataset (which were also influenced by EOV change), all other MPI estimates did not show promising outcomes for DI. It may be the case that the MPI methodology needs significant improvement to produce more reliable estimates. For instance, estimates that are not as noisy, perhaps using a different method to calculate the IA and IF estimates, as mentioned in Chapter 2;

- Tracking modal characteristics was also done using DSFs as the amplitudes from the PSD, within a specific frequency band of choice. These DSFs are especially effective for early DI, when the band-pass filter is parametrised to keep only the frequencies that are suitable, i.e. where modal characteristics are not altered by EOV changes, but, at the same time are affected by damage;

- For the GT engine datasets, one or more DSFs from the MD_2 feature set, shows sensitivity to changes in the operating conditions. In TS-SAF and TF-SAF datasets, kurtosis-based features ($\tilde{s}_4$ and μ_4) look promising for detecting change in engine conditions. This is true for all three different experiments conducted. It can also be said that, changes in the AFR cause higher vibration-based changes, as compared to the fuel mixture change.

Chapter 5

Inclusive outlier analysis

The framework of unsupervised learning is particularly attractive for DI in SHM & CM applications. This is because there might be no prior information regarding the distributional properties of either the damaged or undamaged system states. Consequently, within this framework, the techniques that can be used to discriminate between the two system states, draw their own *assumptions* regarding these distributional properties. Engineering judgement and insight to the problem that is being analysed, e.g. understanding of the physical processes taking place in a damaged bearing, are required to interpret the outcomes of such analysis.

This chapter discusses the use of *inclusive outlier analysis*, which means that the dataset cannot be assumed to have come from the undamaged state only. Instead, both damaged and undamaged states may be expected to be present in the dataset. Given that there are no labels attached to the observations in the dataset, a discriminatory technique like the Mahalanobis distance can be used to identify any outliers presence. In an ideal situation, any outliers identified, will correspond to the damaged state only, i.e. $DI = OD$.

Therefore, this chapter demonstrates how techniques from robust statistics can provide a principled way of identifying outliers, along with their diagnostic capabilities, for the wide range of datasets considered in this thesis. The performance of these techniques is compared to the more mainstream methods of subspace, linear models and statistical/probabilistic outlier analysis that have been used extensively in SHM & CM problems [71].

5.1 An introduction to outlier analysis

In the statistical and machine learning literature an outlier, according to [85], "*...is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.*" Therefore, using this definition, an outlying data point is one that was generated by a different probability distribution than the one that

generated the *majority* of the data points. Hence, outliers are universally considered as the minority observations that do not fit the bulk of the data, and which may be in *sparse regions* and *far away from the data cloud*.
Novelty detection (ND) or one-class classification (see Chapter 6) is the task of recognising, with a given probability, whether a data point is generated by the *reference or "normal" data distribution*. Therefore, this framework utilizes any available knowledge regarding the "normal" data distribution to classify a new data point. In outlier detection (OD) or unsupervised learning, there is no knowledge about the distributional properties of either of the two classes. Therefore, the task reduces to determining data points that *deviate significantly* from the majority of the data distribution. Regardless of the approach taken or algorithm choice, a *model* of the "normal" data distribution class needs to be determined, from which any deviations of subsequent observations will be inferred from.
The problem of outlier detection is an ill-defined one, since setting up an outlier threshold, i.e. deciding what constitutes a *significant* deviation, becomes only relevant with regards to the problem that is being analysed, as discussed in great detail in [18]. For instance, when does an *inherent* imperfection becomes an issue, in order to flag the system as damaged? In addition, what *reference* can we use to draw our decisions from. And also, given the usually small subsets describing such a reference condition, how can we guarantee its validity as an adequate representation of the "normal" model? Maybe such an assumed model is not necessarily appropriate, given the varied environment a system is interacting with and the relatively limited number of available observations. Fortunately, this thesis is concerned with *controlled* experimental observations. Therefore, the [although incomplete due to the finiteness of the observations] defined "normal" system state, is well-known a priori and inferences can and need to be drawn from it. This is true for all datasets, apart from TF-SAF and TS-SAF, where the conditions of "normal" system state are not defined.

Given multidimensional samples $\mathbf{x} \in \mathbb{R}^{N \times p}$, for each nth sample (or segment - see Chapter 4 for details) with dimensionality p, an *outlier score* will be calculated. This outlier score quantifies the level of outlyingness for each $\mathbf{x}_n$, either as a probability of fit to a particular distribution or a distance measure from it. When a binary decision (outlier/inlier) is required to be made, the outlier score is compared to a pre-defined threshold γ.
In the context of SHM & CM, an *ideal threshold* or an *optimal decision boundary* γ is defined as the one that perfectly separates the data points from the two system states: damaged and undamaged. In the robust statistics literature, the notion of α-outliers is used as a starting point to guide the OD procedure, as discussed in [27]. Essentially, a robust statistical model assumes that the *minority* of the observations with proportion $1-\alpha$, are the

outliers, since those have low probability of being generated by the majority data points (with proportion α)*. Thus, a statistical model is used to describe patterns in the data, from which the *majority* of the samples (the inliers) are assumed to be generated from it. In the Mahalanobis distance d^2_{msd} model, the distribution is assumed to be a p-dimensional Gaussian $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and where each $\mathbf{x}_n$ is assumed to be drawn independently (iid). Using this model, an outlier may be one that,

$$out(\alpha, \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})) = \left\{ \mathbf{x}_n : d^2_{msd} = (\mathbf{x}_n - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \boldsymbol{\mu})^T > \gamma (= \chi^2_{p,1-\alpha}) \right\} \tag{5.1}$$

The Mahalanobis distance, is one of the simplest algorithms to use, but, can perform surprisingly well for detecting outliers at the extremes of the Gaussian. It is a *distribution-dependent* method [18] that relies on *linear* correlations between each pairwise combination of the p features in $\mathbf{x}$, through its covariance matrix Σ. The outlyingness of each sample is determined as the distance to the multivariate mean $\boldsymbol{\mu}$.
Note that γ in Equation 5.1 can be calculated by considering the fact that when p variables are drawn independently from a Gaussian, the sum of their squared values is approximately χ^2-distributed with p DOF. Therefore, the χ^2 cumulative probability distribution tables are used to determine the outliers within a level of significance, dictated by α.
A threshold can also be estimated using asymptotic frameworks. One such framework uses Monte Carlo-type simulation to compute, the α-quantile of all the extreme values calculated. It starts by constructing an $N \times p$ matrix, where each element is an independent draw from a zero-mean unit-standard Gaussian. It then computes the d^2_{msd} of all N elements and the one with the largest d^2_{msd} is stored. This is repeated for a large number of times (in this thesis is normally $\approx 1 \times 10^5$) and the α-quantile (e.g. the 99%) of all these extreme values is taken as the threshold. Such methodology was first introduced in SHM in [189], which demonstrated successfully the use of *extreme-value statistics* in setting up OD thresholds.

In outlier analysis there is a diverse range of methodologies (see [18]), some of which were developed specifically for OD (direct methods), e.g. isolation Forest (iForest) and others were adapted for it (indirect methods), e.g. one-class SVM (ocSVM). The method choice in both OD approaches is critical: on the one hand, direct methods place no assumptions on the inliers' underlying distribution $\mathcal{G}$, but, assume certain properties on the outliers' underlying distribution $\mathcal{A}$. Whereas, indirect methods place assumptions only on $\mathcal{G}$.

*In robust statistics, α is the ratio of *expected* uncontaminated observations to the total observations in the dataset. For instance, a value of .99 means that only the 1% of observations are expected to be outliers.

5.2 Outliers influence and robust covariance

As demonstrated previously, methods like PCA and the Mahalanobis distance, rely on the sample covariance Σ, which in turn, is not robust to outliers. Given a certain proportion of outliers in a dataset, such non-robust methods, may be influenced in two ways [156]:

1. *Masking effects*: certain outliers that are significantly different or extreme from the majority of the samples, may mask other samples that may also be considered as outliers. For instance, samples representing a progressive level of damage may hide samples representing an earlier damage. Samples representing significant EOV influence may mask other samples from damage (as discussed in Chapter 1). For instance, the orthogonal axes computed from PCA will be "pulled" towards the most outlying / extreme samples;

2. *Swamping effects*: identify inlying samples as outliers, due to model estimates, e.g. covariance, being shifted due to extreme values.

For the successful implementation of SHM & CM methods, the problem of OD needs to be the same as DI, where outliers are only the samples representing the damaged system state. However, given the EOV influence on the system dynamics, this is particularly challenging. As an illustration, Figure 5.1 shows the masking effect of outliers on the covariance matrix. This 2D dataset is the first two NFs from the Z24 dataset, whereby the selected samples are chosen as the NF values that showed minimal fluctuation from their mean values with respect to the sample no, i.e. sample no. $1700 - 2100$ (see Figure 3.1). Whereas, the full dataset includes all 3932 samples.

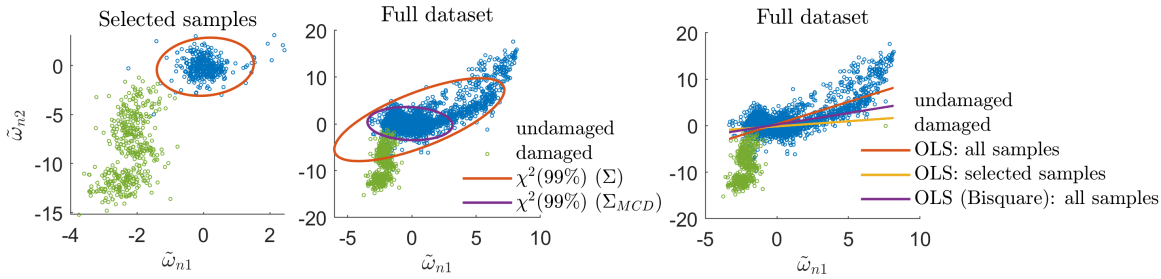

Figure 5.1. The first two NFs of the Z24 dataset, showing the effects of outliers on the sample covariance matrix and potential ways to improve DI using the MCD estimator.

On the left plot, the selected samples are used to compute the 99% error ellipse from the inverse cumulative χ^2 distribution with 2 degrees-of-freedom. Given that these selected samples do not contain data from where the EOV influence was significant, the simple Mahalanobis distance may be used to identify the samples representing damage. For instance,

this error ellipse may be used as a suitable threshold for that purpose.
While, for the middle plot the full dataset is now used to compute the covariance matrix Σ, which is then used to draw the corresponding error ellipse. As can be seen, the EOV influence has expanded the covariance matrix Σ and has shifted the multivariate centre of the ellipse towards those samples. This is due to the fact that the determinant of the covariance matrix Σ becomes larger due to the outliers presence. Using the MCD estimator, however, a new multivariate centre and scatter matrix Σ_{MCD} is estimated using only the $\alpha = 75\%$ of the samples that satisfy certain criteria (further details on the MCD estimator below). As such, drawing the 99% ellipse of the MCD estimator, it will consider samples from *both* EOV influence and damage state as outliers. The error ellipse of the MCD encompasses only those observations that are representative of the bulk of the data.
Moreover, outliers will have an influence on the ordinary least-squares (OLS) fit, too. Outlying observations with large residual values will shift the OLS slope, due to the quadratic error term that needs to be minimized. Hence, one of the earliest attempts has been to apply a weighting function, depending on the squared distances from the line fitted on the majority of the data (e.g. $\alpha = 75\%$ samples). As seen in the right plot in Figure 5.1, the OLS fit with a Bisquare function *approaches* the fit of the OLS in which the selected samples are only included. This is an example of an M-estimator (see Appendix B) with Tukey's Bisquare function. Although it may be feasible to use an M-estimator and the Bisquare function, it does have its limitations, as will be discussed below.
For DI, methods like the Mahalanobis distances, will lead to undesirable outcomes, whenever large outlying observations are present in the dataset. This is because these methods are based on the sample means and covariance matrices. Therefore, this thesis aims to utilize robust methodologies throughout the whole analysis; from pre-processing up to statistical model building for DI.

5.3 Overview of robust statistics and its applications in DI

Robust statistics were initially developed to treat outliers, not as points for further analysis, but, as unwanted noise that influences the accuracy of the statistical model representing the data [92]. In the last decade or so, methodologies such as rPCA and its associated outlier maps [95] provided unprecedented opportunities for *outlier diagnosis*, offering more insights into their behaviour with respect to the assumed statistical model. The essence of robust statistics, is to eliminate or minimize the influence of outliers on the model's estimates, e.g. $\boldsymbol{\mu}, \Sigma$.

Robust statistics is a mathematically well-grounded framework that is essential for modelling datasets with outliers. From an engineering point-of-view the maturity of these methods and its strong theoretical proofs are important, in order to increase confidence for our results and to gain further insight from them [†].

All robust estimators are based on the Tukey-Huber model [119]. This model formalizes the departure from an *elliptical and symmetric* distribution assumption. This model is a mixture of two distributions,

$$F = \alpha \mathcal{G} + (1-\alpha)\mathcal{A} \tag{5.2}$$

where, outliers are drawn from $\mathcal{A}$. Note that $\mathcal{A}$ may be *any* distribution. If $\mathcal{G}$ is assumed to be Gaussian then the model will be able to accommodate other elliptical and symmetric distributions, like the Student-t distribution. Note that the above model places only assumptions on $\mathcal{G}$. The data distribution F may be called *approximately* Gaussian with the assumption failing at its tails [119]. In this case, there exists a *single elliptical region* in the multivariate or univariate space in which the outliers exist at the extremes of this distribution. Robust statistical estimators, typically require the specification of α. This in turn, enables them to compute a *subset of h samples* or the h-subset, which is the set of samples that were found to be *inlying* according to the assumed underlying distribution $\mathcal{G}$. Note that $h = \lfloor \alpha N \rfloor$ and is lower-bounded $h > \lfloor 0.5N \rfloor$, i.e. the majority of the samples. Note that to avoid the determinant of the estimated robust scatter becoming equal to 0, $h > p$ if a selection of $h \approx \lfloor 0.5N \rfloor$ is to be made. A general recommendation is to have $N > 5p$, which ensures good generalisation of the estimates, too [154].

Robust statistics have not been widely used as a potential solution to the DI problem in SHM & CM applications. In the last decade or so, only a limited number of papers were published, and those will be reviewed. In one of the first works [61], the authors used the MCD and the minimum volume enclosing ellipsoid robust scatter and centre estimators for the purpose of identifying damage on several experimental datasets, including the Z24 dataset. The authors showed that these two robust estimators were relatively more sensitive to changes in the nominal system state, i.e. due to both EOV influence and damage, than the Mahalanobis distances is. Similarly, the authors in [80] used the rPCA method to identify damage from vibration data on a wind turbine blade, claiming greater sensitivity to damage, as compared to PCA. Several other robust estimators of multivariate centre and

[†]In contrast, more recent frameworks, such as DL, model highly complex interactions between each of the features and time instances and this poses great challenges for justifying their inferences for instance, as more layers are added in the network and/or why a particular activation function works better than the others. Hence, the dilemma is between enhancing interpretability versus enhancing statistical model flexibility/applicability by increasing complexity

scatter were evaluated in [193] for the purposes of DI. The authors in that paper used cross-PSD DSFs (i.e. between pairs of sensors) from a composite structure having embedded fibre Bragg gratings. Very recently [65], a robust version of multidimensional scaling technique has been used as a solution to DI using a high-dimensional dataset containing outlying observations. Moreover, the diagnostic capabilities of robust estimators, and in particular the least-trimmed squares (LTS) regression model, were used in [60]. The authors showed on 2D plots of standardised LTS residual versus robust MCD distances, that damage is classified as a different outlier than EOV influence does.

5.4 Methods

This section reviews the different robust estimator methods that will be used throughout the thesis for DI. The problem of DI will be approached using *three* different types of analysis:

- *Extreme value analysis via robust statistics*: Calculation of multivariate squared distances using high-breakdown robust estimators of multivariate centre and scatter. Outliers are identified as samples at the extremes;

- *Robust time series analysis*: Residuals estimation using a robust time series model. Outliers are identified as samples that do not adhere to the laws of the time series model, having residuals above the confidence interval set;

- *Exploratory analysis and outlier diagnosis*: Explore differences between the undamaged and damaged system states using robust clustering and subspace analysis. In particular, the variation of samples obtained from the undamaged system state under significant EOV influence will be investigated.

Additionally, the relative benefits of each method considered in this chapter and their applicability to different dataset types can be briefly summarised below:

- *Classical Mahalanobis Squared Distance*: As mentioned previously, it is a distribution-dependent method, which can be used to compute multivariate distances for each data point. It has a zero breakdown value, so that, it cannot be used for outlier detection reliably, e.g. when there is a large proportion of outlier samples and when their squared distances are extreme as compared to the rest of the samples. This is because both the multivariate mean and covariance matrix will be influenced significantly in such cases.

- *Minimum Covariance Determinant*: High-breakdown estimator (up to 50%) that can be used to estimate a robust scatter matrix. For outlier detection, it performs very well

on datasets where the inlying samples come from a unimodal ellipsoidal distribution. Its outlier fraction can be adjusted accordingly.

- *Deterministic Minimum Covariance Determinant*: Tackles the issue of obtaining an initial estimate of the inlying samples by utilising six different scatter estimators. This is important in high-dimensional datasets, where outliers may be "hidden" more easily. Unimodal assumption for the inlying samples is also assumed.

- *Orthogonalised Gnanadesikan-Kettenring*: Must be used in datasets when the expected outlier proportion is comparable or close to 50% of the total number of samples. This is because it has a non-adjustable outlier fraction parameter. Rather, it is always set at its breakdown point. Also, for high dimensional datasets, e.g. when the number of variables is close to 100, the algorithm is not very efficient and alternatives must be considered, e.g. the previous two algorithms. However, when there is a high number of samples for a moderate-dimensional problem, the algorithm is more efficient than the Minimum Covariance Determinant. As with the previous two robust estimators, this algorithm also assumes unimodal elliptical distribution for the inlying samples.

- *Kernel Minimum Regularised Covariance Determinant*: Unlike the three aforementioned robust estimators, this algorithm estimates a "kernelised" scatter matrix. This means that, the algorithm can estimate a robust scatter matrix by considering both non-ellipsoidal and multi-modal shape for the inlying samples. This is true when a non-linear kernel matrix, e.g. the radial basis function, is used for the estimation. Especially useful for datasets with a high number of features and small number of samples as the kernel matrix is $N \times N$.

- *Deterministic MM*: As with S-estimators, this algorithm is based on loss functions to calculate the scatter matrix of the inlying samples, i.e. the h-subset. This allows the algorithm to be applied in datasets where the number of outlier samples far exceed the inlying ones. The formulation is not based on computing different versions of the determinant of the covariance matrix. Therefore, there are no issues of it exploding or imploding. This is also discussed in Appendix C.

- *Robust Principal Component Analysis*: Together with its outlier map, it can be used to compute the type of outliers that exist in a high-dimensional dataset. Therefore, it is especially useful for classifying the outliers in datasets that have a large number of dimensions. Also, it can be used for dimensionality reduction in datasets with a large proportion (up to 50%) of extreme values. However, since it is based on the

Minimum Covariance Determinant algorithm, it is limited to unimodal and ellipsoidal distribution datasets (for the h-subset).

- *Robust clustering via data trimming - tclust*: Can be used to estimate K clusters that are ellipsoidal in shape. These clusters may also be contaminated by outliers, so instead of using methods like Gaussian Mixture Models, tclust can be used to reliably estimate the multivariate means and boundaries of each cluster. Also, its flexibility in defining the ellipsoidal shape makes it a better alternative than other robust cluster methods, e.g. trimmed k-means.

- *Non-linear Least Trimmed Squares*: This is a robust time series modelling algorithm, which enables additional non-linearities, e.g. trends to be modelled. It can be applied on each DSF separately, in order to compute outliers that do not adhere to the time series model estimate. The benefit of using such method is that it does not place any assumptions on the multivariate distribution of the inlying samples in the features space. Its computational performance won't be affected by the number of dimensions as it models each DSF separately from the rest.

- *One-class Support Vector Machine*: This algorithm estimates the inlying distribution as the data points that are further from the origin. Kernels, such as the radial basis function, can be used to model non-linear relationships between the chosen DSFs. As it is based on a different formulation than the multivariate and time series robust methods, it offers an alternative method that can be used for benchmarking and comparison. It also has high computational efficiency in high-dimensional datasets. It is also highly sensitive to its hyper-parameters, e.g. the kernel scale.

- *Isolation Forest*: The Isolation Forest algorithm does not place any assumptions on the inlying data distribution, like all the aforementioned methods do. It only assumes that outliers are far fewer and exist in less dense regions than the inliers. It is highly applicable for datasets, where no particular information for the dataset is known beforehand, e.g. the proportion of outliers and/or the distributional properties of the inliers. Additionally, it is straightforward to implement, unlike the one-class Support Vector Machines that require significant hyper-parameter tuning.

5.4.1 Extreme value analysis via robust statistics

5.4.1.1 Minimum Covariance Determinant

The main idea behind the MCD [94] is to obtain a subset of h samples, in which the determinant of the scatter matrix is minimal. This means that h out of N samples will

be enclosed within a symmetrical unimodal ellipse, with its multivariate centre $\boldsymbol{\mu}_{MCD}$ and scatter Σ_{MCD} being the robust MCD estimates.

The MCD yields maximal breakdown point α^* (see Appendix C), such that $h = [(N + p + 1)/2]$. As it is usually the case with high-breakdown robust estimators, the MCD lacks statistical efficiency, as compared to the classical estimates. For that reason, its two estimates are weighted as shown in Equation 5.3,

$$\begin{bmatrix} \boldsymbol{\mu}_{MCD} \\ \Sigma_{MCD} \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N} w_{MCD}\, \mathbf{x}_n \,/\, \sum_{n=1}^{N} w_{MCD} \\ c1 \frac{1}{N} \sum_{n=1}^{N} w_{MCD} (\mathbf{x}_n - \boldsymbol{\mu}_{MCD})(\mathbf{x}_n - \boldsymbol{\mu}_{MCD})^T \end{bmatrix} \tag{5.3}$$

where, $w_{MCD} = 1$ if $d^2_{MCD}(\boldsymbol{\mu}_{MCD}, \Sigma_{MCD}) \leq \chi^2_{p,99\%}$ and $w_{MCD} = 0$, otherwise, c_1 is a consistency factor. Notice that the weighting function w_{MCD} may serve as a binary decision for OD / DI using the $\chi^2_{p,99\%}$ threshold. In this case, the use of χ^2 assumes that the h samples are drawn from a Gaussian. Note that w_{MCD} can be any other function of choice with similar robustness properties, including a bounded IF. Although it is not a smooth IF function, this is the most widely used and tested in the literature, and is readily available in all current software implementations.

The MCD is in practise implemented using the Fast-MCD algorithm (FMCD) [155], since it would otherwise require evaluating all $\binom{N}{h}$ subsets, making it a computationally intensive procedure for large N. Briefly, FMCD works as follows:

1. Randomly draws a number of subsets (≈ 500) of length $p+1$ (there is a higher probability to obtain an outlier-free subset, as compared to drawing h subsets).

2. Compute estimates of multivariate location and scatter for the subsets.

3. Apply C-steps (see Appendix B) to improve these estimates. Such that, at the current step k a smaller or equal scatter determinant will be found, as compared to the previous step: $|\Sigma_k| \leq |\Sigma_{k-1}|$.

4. The solution converges when $|\Sigma_k| = |\Sigma_{k-1}|$ or $|\Sigma_k| = 0$.

The solution of the FMCD is only locally optimal, but, still retains all the robustness properties of the original MCD, i.e. of bounded IF, affine equivariance and high-breakdown point. The FMCD formed the basis for the development of a number of methods that this thesis will be investigating for DI: rPCA, Deterministic MCD (DetMCD) and Kernel Minimum Regularised Covariance Determinant (kMRCD). Matlab implementations are available on the KU Leuven's website [9], which includes the popular LIBRA[‡] toolbox.

[‡]LIBRA: a MATLAB Library for Robust Analysis.

Table 5.1. The main steps of the DetMCD algorithm.

1.	Compute robust z-scores by subtracting each column of $\mathbf{x}$ by its median and divide by its Q_n to get $Z \in \mathbb{R}^{N \times p}$.
2.	Get $k = 6$ different scatter estimates $S_k(Z)$, including *Spearman's correlation coefficient*, the *Gnanadesikan-Kettenring robust correlation* (see Equation 5.6) and the covariance matrix of the $N/2$ observations with the smallest L_1-norm.
3.	Each $S_k(Z)$ estimate goes through a "refinement" step to ensure potential errors, e.g. small eigenvalues, are avoided.
4.	C-steps are applied on each refined estimate $\check{S}_k(Z)$, to obtain the h-subset with the smallest determinant, and therefore, compute the multivariate estimates of scatter and location, $\check{S}_{h,k}(Z)$ and $\check{\boldsymbol{\mu}}_{h,k}(Z)$, respectively.
5.	Among the $k = 6$ estimates choose the one, say S_{k^*} that has the lowest determinant, so that the raw DetMCD estimates are $S_{DetMCD} = \check{S}_{h,k^*}(Z)$ and $\mathbf{s}_{DetMCD} = \check{\boldsymbol{\mu}}_{h,k^*}(Z)$.
6.	As in the MCD, a weighting step is added on the raw estimates to get Σ_{DetMCD}, $\boldsymbol{\mu}_{DetMCD}$.

5.4.1.2 Deterministic MCD

In FMCD, drawing many random subsets, results in a high probability to obtain at least one that is outlier-free. In cases when $p > 10$ and $\alpha \approx 0.5$, drawing many random subsets (≈ 500) may not be sufficient to obtain an outlier-free subset. Therefore, there may be a better alternative to *guide* this search for the h-subset, by means of using $k = 6$ initial scatter estimates S_k. Each S_k represents different scatter estimates with certain robustness properties, in order to reduce or eliminate the influence of outliers on its estimation. The steps of the DetMCD algorithm [96] are briefly described in Table 5.1. Step 1 ensures the algorithm's scale and location invariance (the method is not affine equivariant). The Matlab function *detmcd.m* that implements this algorithm is included in the LIBRA toolbox [9].

5.4.1.3 Kernel Minimum Regularised Covariance Determinant

The Kernel Minimum Regularised Covariance Determinant (kMRCD) [163] has been recently proposed as a methodology that combines the computational capabilities of the Minimum Regularised Covariance Determinant (MRCD) and the concept of *kernel functions* (see Appendix B for both MRCD and discussion on kernel matrices) to obtain:

- a *regularised* version of the scatter matrix for a dataset with $p > h$, which it can be inverted without its determinant becoming equal to zero (as in the MCD formulation);
- a *flexible* decision boundary to define the inliers' region, which is not restricted to the elliptical distribution assumption that is typically placed on $\mathcal{G}$.

For an h-subset, where $k = 1, ..., h$ are the h selected indices from $n = 1, ..., N$, a matrix of centred feature mappings may be defined as $\Phi_h = \left[\phi(\mathbf{x}_1) - \boldsymbol{\mu}^h(\phi(\mathbf{x}))_{\mathcal{F}}, ..., \phi(\mathbf{x}_h) - \boldsymbol{\mu}^h(\phi(\mathbf{x}))_{\mathcal{F}}\right]^T$. Each $\phi(\mathbf{x}_k)$ is a feature map, so that $\mathcal{X}$ (*data space*) $\rightarrow \mathcal{F}$ (*feature space*), where $\mathcal{F}$ is an m-dimensional space with $m > p$. Also, $\boldsymbol{\mu}^h(\phi(\mathbf{x}))_{\mathcal{F}}$ is the m-dimensional multivariate location vector, given as $\boldsymbol{\mu}^h(\phi(\mathbf{x}))_{\mathcal{F}} = 1/h \sum_{k=1}^{h} \phi(\mathbf{x}_k)$.

As in the MRCD formulation, a regularised scatter is defined as a convex combination of between the original scatter and the target matrices, given a scalar parameter of regularisation ρ. The equivalent regularised scatter using kernels can be provided as,

$$\Sigma_{reg}^h(\phi(\mathbf{x})) = \frac{1-p}{h-1}\Phi_h^T\Phi_h + \rho I_{m\times m} \tag{5.4}$$

where, $I_{m\times m}(= T)$ is the identity matrix in $\mathcal{F}$. Since, $\mathcal{F}$ may be a potentially infinite dimensional space, i.e. $m \rightarrow \infty$, the solution to the problem needs to be redefined as,

$$\Sigma_{kMRCD} = \arg\min_h(|K_{reg}^h|) = \arg\min_h(|(1-\rho)K^h + (h-1)\rho I_{h\times h}|) \tag{5.5}$$

where, the kernel matrix of the h-subset $K^h = \Phi_h\Phi_h^T$. The question now arises, whether the two problems, as described in Equations 5.4 and 5.5, are equivalent. It turns out that they are, since $|K_{reg}^h| = c_k|\Sigma_{reg}^h|$, where c_k is a constant, so that the optimisation as defined in Equation 5.5 can be carried out. Note that, the Mahalanobis distance computed by the kMRCD for the kth data observation, $d_{kMRCD}^2(\phi(\mathbf{x}_k), \boldsymbol{\mu}^h(\phi(\mathbf{x})), K_{reg}^h)_k$ requires invertibility of the $N \times N$ kernel matrix, instead of the $m \times m$ scatter Σ_{reg}^h.

As with the DetMCD, kMRCD begins by computing four initial estimates of multivariate location and scatter that are robust to different types of outliers in the data. So that, it initially obtains four h-subsets, $\tilde{\mathbf{h}} = \tilde{h}_1, ..., \tilde{h}_4$, before proceeding to the "kernelized" equivalent of the refinement step of the DetMCD. Briefly, the main steps of the kMRCD algorithm are as shown, in the diagram in Figure 5.2.

For each of the four initially estimated h-subsets, i.e. $\tilde{\mathbf{h}} = \{\tilde{h}_1, ..., \tilde{h}_4\}$, compute C-steps as follows:

1. Calculate $K_{reg}^{\tilde{h}} = (1-\rho)K^{\tilde{h}} + (\tilde{h}-1)\rho I_{\tilde{h}\times\tilde{h}}$, for each $\tilde{h}$;
2. Compute N Mahalanobis distances $\forall n$, d_{kMRCD}^2;
3. Redefine each $\tilde{h}$ using the h-indices with the smallest d_{kMRCD}^2;
4. Repeat the previous steps until convergence.

Note that the dataset $\mathbf{x}$ is firstly, robustly standardised using the univariate MCD formulation. Then, the value of ρ is determined using the eigenvalues of the kernel matrix K^h,

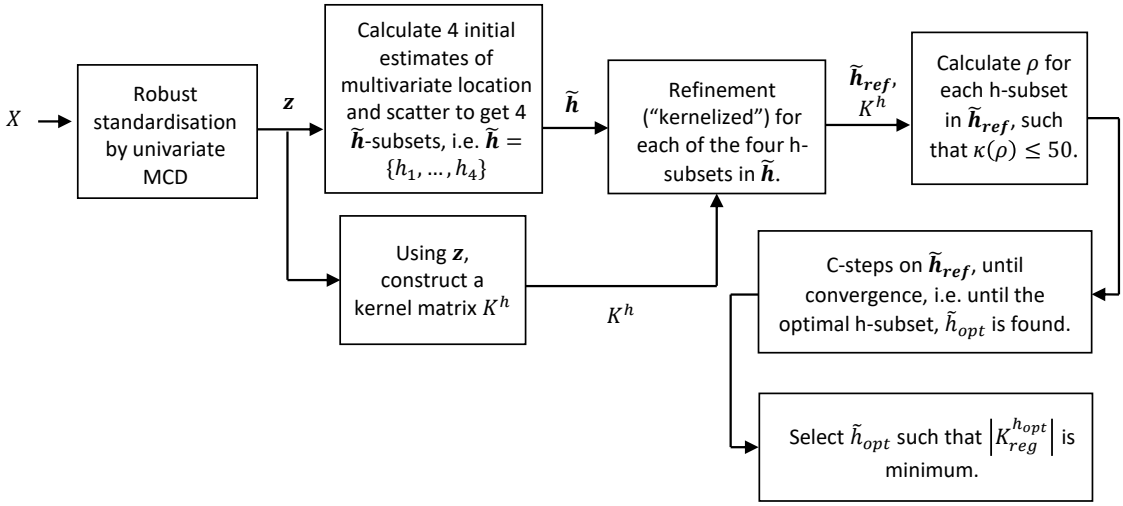

Figure 5.2. kMRCD diagram of the algorithm's main steps.

satisfying $\kappa(\rho) \leq 50$, where the expression is provided analytically in the paper. The method is only location and scale invariant, while it maintains the high-breakdown point, bounded IF of the original MCD. The kMRCD Matlab function *kMRCD.m* and its associated files, are provided on KU Leuven's website [9].

5.4.1.4 Orthogonalised Gnanadesikan-Kettenring estimator

The Orthogonalised Gnanadesikan-Kettenring (OGK) [120] estimator is only scale and location invariant (not affine equivariant as MCD). However, its high-breakdown point ($\alpha^* \to 0.5$) makes the method an interesting alternative to use.
The algorithm is based on a robust correlation matrix between pairs of features in $\mathbf{x}$,

$$U_{ij} = \frac{1}{4}\left[\tilde{\sigma}(\mathbf{x}_i + \mathbf{x}_j)^2 - \tilde{\sigma}(\mathbf{x}_i - \mathbf{x}_j)^2\right], \quad \forall\, i \neq j \tag{5.6}$$

where, $\tilde{\sigma}$ is any univariate robust scale estimate, e.g. Q_n and MAD. In the case of MAD the algorithm reaches its highest breakdown $\alpha^* = 0.5$. The eigenvalues and eigenvectors V of U are then computed, so that for each nth observation the data are projected as,

$$\mathbf{y}_n = V^T A^{-1} \mathbf{x}_n \tag{5.7}$$

where, $A = diag(\tilde{\sigma}(\mathbf{x}_1), ..., (\tilde{\sigma}(\mathbf{x}_p))$. Using these projections, the OGK multivariate location and scatter estimates ($\boldsymbol{\mu}_{OGK}$, Σ_{OGK}) can be expressed as,

$$\begin{aligned} \boldsymbol{\mu}_{OGK}(\mathbf{x}) &= (AV)\boldsymbol{\nu} \\ \Sigma_{OGK}(X) &= (AV)L(AV)^T \end{aligned} \tag{5.8}$$

where, $L = diag(\tilde{\sigma}(\mathbf{y}_1), ..., (\tilde{\sigma}(\mathbf{y}_p))$ are the univariate robust variances of matrix Y, which contains each nth data point $\mathbf{y}_n \in \mathbb{R}^p$ in its rows. The corresponding univariate robust location estimates are $\boldsymbol{\nu} = (\tilde{\mu}(\mathbf{y}_1), ..., \tilde{\mu}(\mathbf{y}_p))$, e.g. medians.
OGK follows a weighting step, as in FMCD, to improve its efficiency, where it utilizes the computed estimates, $d^2_{OGK} = d_{MSD}(\mathbf{x}, \boldsymbol{\mu}_{OGK}, \Sigma_{OGK})$. The Matlab *robustcov.m* built-in function includes the OGK as an option.

5.4.1.5 Deterministic MM-estimator

A Deterministic MM (DetMM) robust estimator [93] is an MM-estimator [119] of multivariate location and scatter $(\boldsymbol{\mu}_{MM}, \Sigma_{MM})$, whereby six different scatter estimates, as in DetMCD, are used to obtain initial h-subsets, instead of random sampling, as in FMCD. The steps 1-3, as outlined in Table 5.1, are similar for the DetMM algorithm.
MM-estimators are more *statistically efficient* (when $\mathcal{G}$ is assumed to be Gaussian distributed) extensions of S-estimators of multivariate location and scatter $(\boldsymbol{\mu}_S, \Sigma_S)$. In particular, S-estimators solve the following optimisation problem,

$$\arg\min |\Sigma_S| \quad s.t. \quad \frac{1}{N}\sum_{n=1}^{N} \rho\left(\sqrt{(\mathbf{x}_n - \boldsymbol{\mu}_S)\Sigma_S^{-1}(\mathbf{x}_n - \boldsymbol{\mu}_S)}\right) = b \tag{5.9}$$

where, the loss function $\rho(d^2_S)$ is selected to be: symmetric, twice continuously differentiable, with $\rho(0) = 0$, strictly increasing in a finite interval $[0, k(>0)]$ and a constant when $[k, \infty)$. A typical choice of ρ that satisfies these conditions and achieves close to .5 breakdown point, is Tukey's Bisquare function, which was demonstrated in the context of regression, previously in Figure 5.1. As in the MCD method, the optimisation problem is non-convex, with several local minima. As such, random sampling is usually performed through iterative solutions (as in FMCD) until convergence of the objective function in Equation 5.9. This is done through the *FastS* algorithm that follows similar steps as in FMCD.
An MM-estimator, inherits the high-breakdown point of the S-estimator, but, utilizes two loss functions, ρ_0, ρ_1. If $\rho_0 = \rho_1$, then the method is an S-estimator. More specifically, the basic procedure to compute $\boldsymbol{\mu}_{MM}, \Sigma_{MM}$ pairs is outlined in the following steps,

1. For a specific loss function $\rho_0(d^2(\mathbf{x}, \boldsymbol{\mu}_S, \Sigma_S))$ (as in Equation 5.9), calculate S-estimate pairs $\boldsymbol{\mu}_S, \Sigma_S$.

2. Compute a scale value $\sigma_{\rho_0} = |\Sigma_S|^{-2p}$.

3. For a different loss function ρ_1 (having better efficiency than ρ_0), solve the following optimisation problem to obtain an estimate for the scatter matrix $\tilde{S}$ and $\boldsymbol{\mu}_{MM}$:
   $1/N \sum_{n=1}^{N} \rho_1 \left(\left[(\mathbf{x}_n - \boldsymbol{\mu}) \tilde{S}^{-1} (\mathbf{x}_n - \boldsymbol{\mu}) \right] / \sigma_{\rho_0} \right)$
   $\forall \boldsymbol{\mu} \in \mathbb{R}^p$ and $\forall \tilde{S} \in \mathbb{R}^p$ $s.t.$ $|\tilde{S}| = 1$.

4. The scatter matrix of MM-estimator is $\Sigma_{MM} = \sigma^2_{\rho_0}\tilde{S}$.

Note that DetMM is a high-breakdown, near-affine equivariant and permutation invariant (since it eliminates the initial random sampling) robust estimator, which is more statistically efficient than the FMCD.
The DetMM Matlab function *detMM.m* is provided on KU Leuven's Robust Statistics website [9]. The implementation is based on the MM-estimator function *MMmult.m*, which is part of the *Flexible Statistics and Data Analysis (FSDA)* Matlab toolbox, developed by the Robust Statistics Academy at the University of Parma [5].

5.4.2 Exploratory analysis and outlier diagnosis

5.4.2.1 rPCA and its outlier map

The rPCA algorithm was presented in Chapter 4. Its main objective is to project *only* the non-outlying data points on a new subspace, which in turn makes the rest of the data (the outliers) more distinguishable. Simple distance measures, e.g. Mahalanobis distances, can then be used to identify these outliers as points that are outside the χ^2 error ellipse.
Let, $\lambda^1_{rPCA}, ..., \lambda^k_{rPCA}$ be the collection of k eigenvalues of the MCD scatter matrix. A useful measure known as *score distance* (SD) for each data point can be computed $SD(\mathbf{x}_n) = \sqrt{\sum_{j=1}^{k} y^2_{n,j}/\lambda_j}$, i.e. measuring the squared *robust* distance of $\mathbf{x}_n$ projection to the centre of all the N projections.
Note that rPCA has the same properties of PCA, i.e. it is a *location* and *orthogonal equivariance* technique, meaning that any kind of rotation and/or shift of transformation applied on the original dataset, it will cause the corresponding rotation of the loadings matrix and/or robust location shift, so that *SD will not be affected.*

An *outlier map* is a 2D plot of ODist (see Equation 4.5 for its definition) versus SD, where it is possible to classify the outlier *type*, as well as, the regular observations. To determine outliers, a typical χ^2 cut-off value is used for the SD direction, while for OD a scaled version of it is used.
Since most of the methods in robust multivariate analysis are parametric and highly interpretable, they offer the capability to classify each outlier, according to the model/method being used. There are three main classes of outliers, where all share one property, i.e. located away from the data majority, but they differ on their influence on the statistic being estimated, e.g. the slope of the ordinary least squares regression.

- *Good leverage points* have similar statistical properties as the data majority. In rPCA they lie close to the subspace on which the data majority is projected on. In regression,

these points will have the same slope as the data majority. In general, these points help in defining a model for the non-outlying samples.

- *Bad leverage points* have different statistical properties as the data majority, so that they may become "hidden" or "disguised" as inliers when non-robust methods are being used. The influence of such outliers on the estimation has been discussed previously using the masking and swamping effects. In general, these points may shift the statistic being estimated towards them, i.e. far away from the data majority. In rPCA, they have both large orthogonal and score distances.

- *Vertical/orthogonal outliers* are outliers that can be hidden when non-robust methods are computed, as observed in [60] in the context of SHM. In rPCA they are located far away from the projected rPCA subspace (large orthogonal distance). At the same time, they may be located within the data majority cloud in the projected rPCA subspace. This implies that they remain hidden, while they may have some influence on the definition of the principal components.

The rPCA outlier map in which the orthogonal and score distances are plotted against each other are typical tools used for diagnosing outliers, as shown in [154]. Additionally, departure from normality may be seen on a distance-distance (DD) [155] plot. That is, a DD-plot shows the Robust Mahalanobis distances against the classical ones d^2_{msd}, since the latter assumes all the samples are Gaussian-distributed due to the MLE-derived estimates.

5.4.2.2 Robust clustering via data trimming

Generally, the aim of clustering analysis is to partition the p-dimensional samples in $\mathbf{x}$ into k disjoint clusters, where in each kth cluster the samples representing it have common characteristics [30]. This may be used in the estimation of statistical parameters, such as k multivariate centres and scatter matrices. The possibility of outlying observations also needs to be accounted for in cluster analysis.
Note that the estimation of a robust multivariate scatter and its location imply the following:

- (C1): A single elliptical and symmetrical distribution is sufficient for the inclusion of the inlying samples;

- (C2): Linear interdependency between each feature.

Clustering analysis aims to address (C1), in part, where there may be *several* elliptical distributions in the p-dimensional space. This is important because it will allow outlying observations that lie *between* these elliptical distributions (also called as "bridge" data

points) to be detected, as those that exist in *sparse regions*. In that way, $\mathcal{G}$ can be modelled as a *multi-modal* one, rather than a single ellipse that encompasses the non-outlying observations.

In the context of SHM / CM this is interesting; EOV influence on the dynamics can certainly be a source of multi-modality, as seen in the Z24 dataset. In this case, multiple elliptical regions in the p-dimensional (or feature) space may be used in order to provide a more accurate description for its undamaged state. Note that this places no assumptions on the outliers, other than, they exist in sparse regions in the feature space of $\mathbf{x}$.

In [78] the authors presented a robust clustering methodology that tackles both the problem of OD, as well as, the estimation of a number of elliptical distributions in the feature space, called clusters from now on. Similar methodologies were presented in [51], using a trimmed k-means approach for robustly defining multiple clusters alongside a proportion of outliers. However, trimmed k-means, inherits properties of the original methodology, one of which is the assumption of spherical cluster shapes of roughly equal sizes. This limits its applicability and OD capabilities. As discussed in the paper [78], apart from the spherical assumption, another major difficulty in dealing with cluster analysis in the presence of outliers, i.e. robust cluster analysis, is the *heterogeneity* of the clusters. That is, the Mahalanobis distances computed by members of one cluster, may be at different scales than the ones of another cluster, imposing difficulties in formulating a global optimisation problem, in the first place.

One of the earliest methods that dealt with the variation in cluster scales, has been to normalise the covariances of each cluster in order to have a determinant of unity. In the paper, however, the authors proposed to place a restriction on the ratio of the largest to the smallest covariance eigenvalues, the so-called *eigenvalue-ratio* (ER) restriction, in order to get more "informative" clusters.

Let the total number of clusters be K, and $H = \{\mathbf{h}_1, ..., \mathbf{h}_K\}$ be a collection of h-subsets so that h_k is the subset of *inlying samples* assigned to the kth cluster. Using the *spurious-outlier* model [76], the authors suggest to optimise the following likelihood function,

$$\left[\prod_{k=1}^{K}\prod_{n\in h_k}\pi_k g(\mathbf{x}_n\,|\boldsymbol{\mu}_k,\Sigma_k)\right]\left[\prod_{n\notin H}\tilde{g}_n(\mathbf{x}_n)\right] \tag{5.10}$$

where, the second term in the bracket includes only the $\lfloor(1-\alpha)N\rfloor$ outliers and may be drawn from *any* probability density function $\tilde{g}_n \in \mathbb{R}^p$. Whereas, the non-outlying observations $\mathbf{x}_n$ are treated as iid draws from an elliptical and symmetric probability density function $g(\cdot)$, with its multivariate location and covariance matrix, $\boldsymbol{\mu}_k, \Sigma_k$, being different for each cluster. To some extend, the Tukey-Huber model presented in Equation 5.2, is of

a similar form, whereby it places some restrictions on the distribution properties for the non-outlying observations and no such restriction on the outliers' distribution. Also, each cluster is multiplied by a coefficient π_k, so that $\sum_k \pi_k = 1$. Note that if $\Sigma_k = \sigma^2 I$, with I being the identity matrix, the solution reduces to the trimmed k-means method. Also, when $k = 1$, i.e. a single cluster expectation, the problem is *equivalent* to the one being solved by the FMCD.

The solution to Equation 5.10, solves an Expectation-Maximization (EM) problem to get MLE estimates for $\boldsymbol{\mu}_k$ and Σ_k [30]. In order to clarify, the method places no restrictions on the outlying part of the model, it is therefore, *not considered in any part of the procedure.* The EM problem is combined with C-steps, in order to impose the necessary restrictions on the ER, and progress iteratively to find the optimal solution, among all $p \times K$ candidate eigenvalues λ. For a user defined constant $c_{ER} \geq 1$, the following restriction is formulated,

$$\frac{\max_{m=1,\ldots,p,k=1,\ldots,K} \lambda_{k,p}}{\min_{m=1,\ldots,p,k=1,\ldots,K} \lambda_{k,p}} \leq c_{ER} \tag{5.11}$$

The scalar constant c_{ER} places the necessary restrictions on each kth cluster covariance. When $c_{ER} = 1$ the method is similar to the trimmed k-means with equal weights on each cluster, and is the most restrictive scenario. On the other hand for a large constant, $c_{ER} > 50$, the problem is rather unrestricted, *leaving more freedom to handle the variation in scattering* between each cluster. The method is implemented in Matlab *tclust.m* function as part of the FSDA toolbox [5]. Among the user defined parameters, perhaps the three most important are: the expected amount of contamination $(1-\alpha)$, c_{ER} value and the number of clusters K.

5.4.3 Robust time series analysis

It is important to note that *temporal relationships* do exist in $\mathbf{x}$, since all features were derived from *time series vibration data* taken on fixed-window time-intervals. Up to now, all presented methods dealt with the problem of detecting *point* outliers [18], in the multivariate context. Although these multivariate outlier analysis methods, e.g. MCD, take into consideration the relationship of the p-features, they completely disregard any relationship for each pth feature individually *with respect to its sample number.*

Recently, in [154], the authors presented a time series analysis methodology, which generalizes to non-linear problems, too. The method combines an *alternate least squares* (ALS) approach, alongside a least trimmed squares (LTS) methodology (implemented via the Fast LTS algorithm), in order to solve a non-linear LTS (NLTS) *univariate* time series modelling problem via robust fitting.

A brief explanation of the method will be attempted here. Note that, this thesis deals with

sample numbers $n = 1, ..., N$, due to the preceding stage of DSF extraction (see Chapter 4). Therefore, given a univariate time series at sample number n, x_n, a NLTS robust estimator for a general model $f(\boldsymbol{\theta}, n)$, where $\boldsymbol{\theta}$ is the vector containing its parameters, can be cast as the following optimisation problem,

$$\boldsymbol{\theta}_{NLTS} = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{h} [x_n - f(\boldsymbol{\theta}, n)]_j^2 \tag{5.12}$$

where, the term $[\cdot]_j^2$ is the jth smallest squared residual. This includes up to h most *"central"* residuals, which is equivalent to the h-subset for the multivariate robust estimators of scatter and centre.

In general, conventional time series, e.g. AR analysis models, may yield erroneous results when *structural changes* occur. These structural changes may be point outliers and/or *collective* outliers, which can be grouped with respect to the sample number, and/or *level shifts*. Therefore, apart from OD, both the *time* of occurrence of a level shift as well as its *amplitude* are important considerations for a generalised (and flexible) time series analysis approach. Similar to the multivariate analysis, a robust fit is obtained using the h samples, which yields the optimal parameters vector $\boldsymbol{\theta}_{opt}$ for the considered model, i.e. $f(\boldsymbol{\theta}_{opt}, n)$, whereby it minimizes the objective function in Equation 5.12. Using this model, it is possible to compute a *T-scaled* score (similar to a z-score method), as,

$$[x_n - f(\boldsymbol{\theta}_{opt}, n)] / \tilde{\sigma}^2(\mathbf{x}_i, f(\boldsymbol{\theta}_{opt}, n), h) \tag{5.13}$$

where, $\tilde{\sigma}(\mathbf{x}_i, f(\boldsymbol{\theta}_{opt}, n), h)$ is the scale of the squared residual between the time series model $f(\boldsymbol{\theta}_{opt}, n)$ and the ith feature as a function of sample number $\mathbf{x}_i(n)$. A confidence interval of 99% is typically used.

The method assumes a model of the form $x_n = f(\boldsymbol{\theta}, n) + \varepsilon_n$, where $\boldsymbol{\theta}$ is a d-dimensional vector, which includes all d parameters of the model $f(\cdot)$. In general, an LTS estimator would have been used for a time series model that is linear in its parameters. However, $f(\cdot)$ is non-linear in its parameters (although some may be zero), and in particular, when expanded, it yields the following expression,

$$x_n = \sum_{a=0}^{A} \alpha_a n^a + \left[\sum_{b=1}^{B} (\beta_{b,1} cos(2\pi bn) + \beta_{b,2} sin(2\pi bn)) \right] \left(1 + \sum_{c=1}^{C} \gamma_c n^c \right) + \delta_1 I(n \geq \delta_2) + \varepsilon_n \tag{5.14}$$

where, ε_n is a stationary stochastic process with finite variance σ^2, which may be modelled as $\varepsilon_n \sim \mathcal{N}(0, \sigma^2)$. The first term is a polynomial trend of order A, whereas the second term is a seasonal component, with a polynomial-type of variation of order C. The level shift amplitude δ_1 switches between 1 and 0 using the indicator function $I(\cdot)$, which depends on the

Table 5.2. Confusion matrix for outlier detection.

	Predicted undamaged	*Predicted* damaged
Actual undamaged	**T**rue **N**egative	**F**alse **P**ositive
Actual damaged	**F**alse **N**egative	**T**rue **P**ositive

level shift location δ_2. The above model is non-linear in its parameters, i.e. $\beta_{b,1}, \beta_{b,2}, \gamma_c, \delta_2$, so that only a non-linear time series model can be used to compute the parameter estimates vector $\boldsymbol{\theta}$. The NLTS estimator fits the parameters $\boldsymbol{\theta}_{NLTS}$ to the model given in Equation 5.14, by (1) using ALS to alternate between two linear LS fits and (2) obtaining the h-subset using C-steps (see above MCD estimator). The procedure is initialised by setting $\gamma_c = 0, \forall\, c = 1, ..., C$, which converts the problem in Equation 5.14 into a linear one (in its parameters $\alpha, \beta_{b,1}, \beta_{b,2}, \delta_1$) and applying LS fit on a small subset of the data with at least $p-1$ observations. The Matlab function *LTSts.m* is part of the FSDA toolbox [5], and implements the NLTS method. Its user defined parameters include A, B, C, δ_2, the fraction of outliers expected $1-\alpha$, which determines the h-subset length.

5.5 Evaluation of OD method performance & benchmarking

In a practical scenario, it is not possible to evaluate the performance of an OD method within the framework of unsupervised learning, simply because there are no labels available. Thankfully, every dataset used in this thesis is labelled, which makes it possible to quantify the performance of the different methods. This is usually done by striking a compromise between false positive and negative rates, depending on the application. For instance, in a life-threatening situation, in which *a single* true outlier matters, it is more preferred to have more false alarms (false positives), than it is to have any misses (false negatives). The confusion matrix, as shown in Table 5.2 is helpful in finding out, to which extend, observations from the actual system states (damaged and undamaged), are being classified correctly. Note that, the damaged state is the positive class.

To better summarize the evaluation of a binary classifier's performance, different measures such as *precision* and *recall* can be computed as in Table 5.3. Those two metrics are commonly used on datasets where positive class/damaged/outliers observations are rare. These metrics are a function of the threshold γ being used. F_1-score uses the *harmonic mean* to strike a balance between Precision and Recall, since they may be inversely proportional, as α takes different values. If we choose to increase Recall, on its own, we are creating a

Table 5.3. Performance measures for outlier detection.

Measure	*Description*	*Formula*
Precision (α)	number of true outliers predicted to the total outliers predicted.	$\frac{TP}{TP+FP}$
Recall (α)	number of true outliers predicted to the total of true outliers.	$\frac{TP}{TP+FN}$
F_1**-score** (α)	Precision and Recall harmonic mean.	$\frac{TP}{TP+.5(FP+FN)}$

conservative monitoring system, as we minimise FN at the expense of higher FP. On the other hand, if even a small amount of FP is undesirable, and FN is unimportant, then Precision can be used on its own. However, since both of these two measures are important in SHM and CM, then F_1-score will be used as performance measure. Note that in heavily imbalanced datasets, TN (see Table 5.2) is not regarded as important, in terms of outlier detector performance. Hence F_1-score should be used alongside other measures, too.

In cases where the dataset is not skewed towards the undamaged examples, the False Positive Rate, i.e. $FPR(\alpha) = FP/(FP+TN)$, can be plotted against Recall to generate the *Receiver Operating Characteristic* (ROC) curve. The ROC(γ) is a monotonically increasing curve, and a random classifier has performance along its diagonal. The area under the curve (AUC) of the ROC A_{ROC} is computed to provide a single measure of the effectiveness of the classifier to *distinguish* between the two classes. In the upper limit, when $A_{ROC} = 1$, an OD method with its threshold γ, discriminates *all* the undamaged examples from the damaged ones.

One final important consideration relates to *benchmarking* of the above robust methods' performance with respect to DI. That is, how their performance, mostly in terms of discriminating between the two system states, measures against more commonly used OD methods. For that purpose, two widely-used OD methods will be employed for OD: 1. ocSVM and 2. iForest. Note that these two methods are not restricted to representing the decision boundary as a unimodal ellipse, which may or may not offer benefits for DI. Therefore, their performance will be explored further. In Appendix B these two methods are discussed in some extend, including ways to setup thresholds, which is different from robust methods.

5.6 Results & discussion

In this section, the results for DI using the framework of inclusive outlier analysis are presented for all datasets considered. Given the wide range of problems and methods considered, an attempt has been made to include only those results that showed promise

with regards to DI and/or provided useful insights on the datasets. From the three types of analysis approaches, which were discussed previously, a range of different conclusions will be drawn, independently.

It is important to note that the various methods used to produce the results in this section have their own merits and thus, they may be suitable for different range of problems. Therefore, it is not the main purpose of the analysis to draw direct comparisons between the methods, e.g. between OGK and kMRCD. What is most important, is to demonstrate how robust statistics may be used as an alternative method for DI and general data exploration in SHM and CM problems. This implies that comparison between robust and non-robust techniques, e.g. rPCA and PCA, will need to be made explicitly.

5.6.1 Important notes regarding the results

Details related to the implementation of each method used in this chapter are as follows:

- ocSVM: In ocSVM, the hyperparameters of the *RBF kernel* (see Appendix B) were selected empirically and/or using the following metric: $\sqrt{median||\mathbf{x}_i - \mathbf{x}_j||}$. This corresponds to the median of the pairwise distances.

- iForest: Given the iForest has no hyperparameters, as ocSVM does, no tuning of the algorithm was performed. However, the number of the ensemble iTrees, was chosen as ≈ 1000, which provided fast implementation for the algorithm, balanced with satisfactory results;

- tclust: The restriction factor c_{ER} (see Equation 5.11) is chosen by trial-and-error for each dataset. In most datasets, $c_{ER} > 100$, which represents an unrestricted problem (see previous explanation on tclust algorithm).

- NLTS: In NLTS, the values of each of its parameters A, B, C, δ_2 (see Equation 5.14) were chosen according to the dataset being investigated. Additionally, the confidence interval for its fitted line, is chosen as a typical one 97.5%, unless otherwise specified.

Details of the thresholds used for each method in this chapter are as follows:

- MC: Most datasets are high-dimensional, and therefore, MC threshold estimation using 1×10^5 simulations required significant computational resources. For that reason, the Sheffield Advanced Research Computer [11], which is one of The University of Sheffield's High Performance Computing systems has been used to derive the MC thresholds of all robust estimators, apart from kMRCD. In particular, the following main specifications were used in a parallel computational environment: 5 G per CPU

core of *real memory*, 8 CPU cores of symmetric multiprocessing, which yielded a total computational time of close to 2 *hours*. Note that for kMRCD, since its scatter matrix is of $N \times N$, derivation of a MC is still impractical using the Sheffield's High Performance Computing: taking more than 8 hours with a similar set of specifications to complete the 1×10^5 simulations;

- χ^2-based: A threshold on the χ^2 CDF with a confidence interval equivalent to the proportion of inliers α is used. This is because a χ^2 distribution assumption may be approximated for the inliers, but, not for the outliers. While, if we suspect that outliers may also be elliptically distributed (which is typically not the case), a typical confidence interval, e.g. 99%, may be used;

- kMRCD: A threshold on kMRCD is computed based on the following empirical approximation: higher-dimensional SMDs, e.g. infinite-dimensional for an RBF kernel, are better approximated with a log-normal distribution, rather than a χ^2 one. Similarly, a confidence interval equivalent to the proportion of inliers α may be used. A sample $\mathbf{x}_*$ will be flagged as an outlier if its SMD computed from the kMRCD estimates $\log SMD_{kMRCD}(\mathbf{x}_*)$ exceeds a certain confidence interval (e.g. α) of the standard normal distribution,

$$\frac{\log[SMD_{kMRCD}(\mathbf{x}_*)] - \mu_{mcd}(\log[SMD_{kMRCD}(\mathbf{x}_*)])}{\sigma_{mcd}(\log[SMD_{kMRCD}(\mathbf{x}_*)])} > z(\alpha) \tag{5.15}$$

  where, μ_{mcd} and σ_{mcd} are the univariate robust MCD mean and scale estimates, respectively;

- iForest: As Equation B.8 shows, the expected path length from an ensemble of ≈ 1000 iTrees is computed. This in turn is used to compute the iForest outlier scores, which classifies outliers as samples that exceed a confidence interval, which in this case is α;

- ocSVM: As Equation B.7 shows, the bias term ρ of a projected sample $\mathbf{x}_*$ provides the threshold value for the expected outlying samples in the dataset.

5.6.2 TF-LBO dataset

The TF-LBO dataset, has some important characteristics that need to be considered. Its three states: AFR 1, 2 and 3 yield different vibration responses for the combustor. This implies that, the identification of LBO (AFR 3 state) from two different reference states is not as straightforward. This is similar to DI in structures, whereby EOV influences the undamaged state, to the point that outliers from that state cause masking / swamping of the damaged state.

The results in Figures 5.3 to 5.5, show the squared Mahalanobis distances (SMDs) computed using five robust estimators. As a comparison, the SMDs from the sample covariance and multivariate mean are computed. The results may be summarised as follows:

1. As the parameter α changes from 0.55 to 0.8 in Figures 5.3 and 5.4, respectively, we notice the effects of masking. Outlying samples from AFR 1 and 2 are clearly shown on the plots of the SMDs computed using FMCD and DetMCD estimators. As the α value increases, more samples are expected to be inlying, thus, forcing samples from AFR 3 into its inlier region;

2. Given a suitable α value ($= 0.67$), since AFR 3 occupies 1/3 of the total number of samples, the SMDs reject AFR 3 samples as being members of the inlying region. At the same time, the SMDs with $\alpha = 0.67$ calculate the robust scatter and mean estimates by excluding less points from AFR 1 and 2 system states, as compared to when $\alpha = 0.55$;

3. LBO identification from the SMDs calculated using the sample covariance and multivariate mean (d^2_{msd}), will be unsuccessful due to masking effects from AFR 1 and 2 outlying samples;

4. Notice from Figures 5.3 to 5.5, the SMDs computed using OGK robust estimator are invariant to α. This is because its estimates $\boldsymbol{\mu}_{OGK}$ and Σ_{OGK} are based on the univariate robust scale estimates of Q_n or MAD (see Equation 5.8). This implies that the α value of OGK robust estimator is always close to 0.5. This explains the high rejection rates seen in the plots, which are comparable to when $\alpha = 0.55$ for other robust estimators, e.g. FMCD, as seen in Figure 5.3;

5. Notice that due to the variation in α the thresholds $\gamma_{p,\alpha}$ and $\gamma_{kMRCD,\alpha}$ change accordingly.

In order to gain an insight into some of the most important characteristics of robust estimators, a number of two-dimensional plots were produced. This was done by selecting two features from the MD_2 DSFs matrix, in which AFR 1 and 2 appeared as two disjointed sets of samples. This means that, from the point-of-view of these two features, one may have justifiably assumed that these two sets of samples could have been generated by two different distributions. At the same time, with these two features, samples from AFR 3 were shown as furthest away from both AFR 1 and 2.

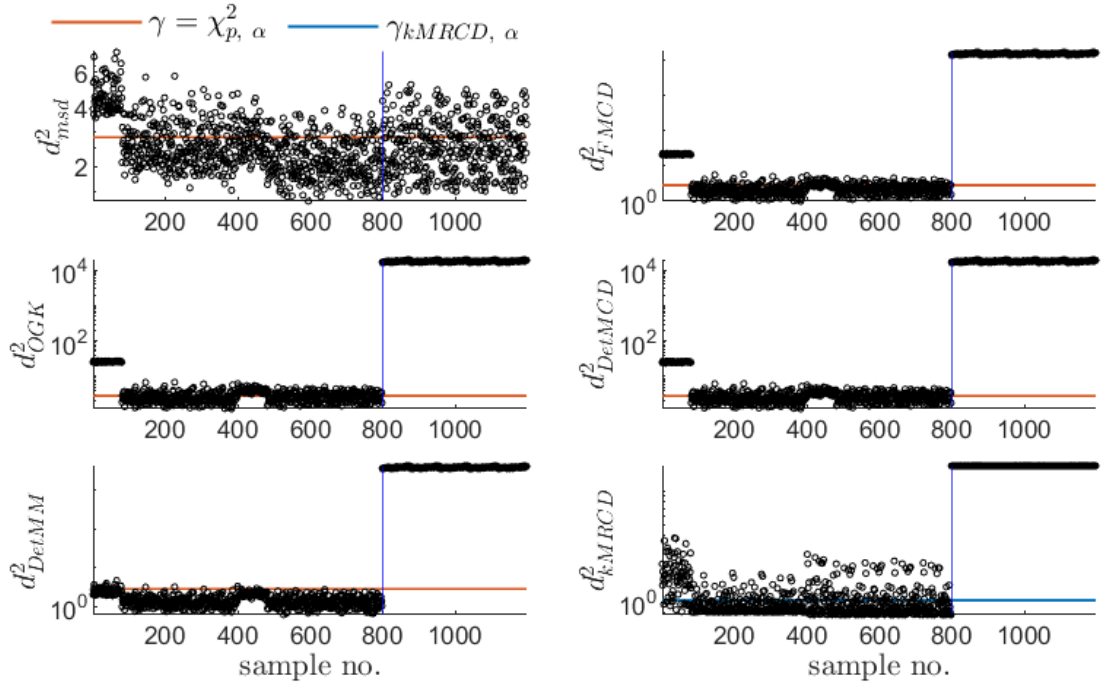

Figure 5.3. TF-LBO: squared Mahalanobis distances (on log-scale) from five robust estimators using $\alpha = 0.55$. The first sample from the damaged state is shown by the vertical line, while the horizontal lines are the two thresholds. For comparison the Mahalanobis distances (on log-scale) using the sample mean and covariance are shown (d^2_{msd}).

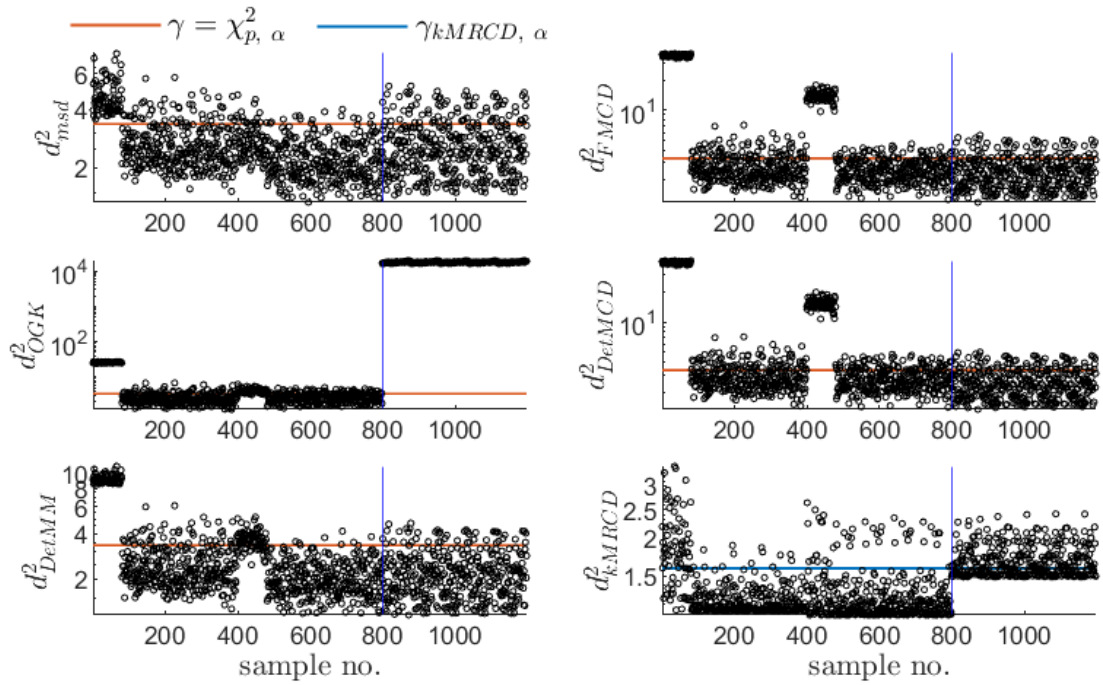

Figure 5.4. Same plots as Figure 5.3 but with $\alpha = 0.8$.

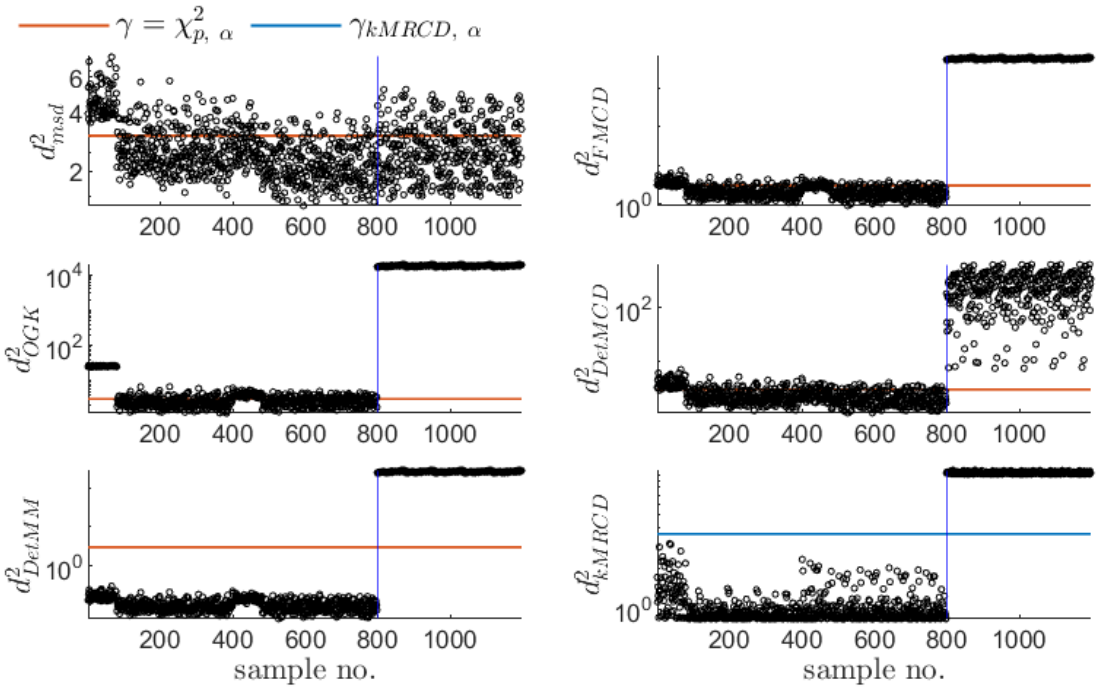

Figure 5.5. Same plots as Figure 5.3 but with $\alpha = 0.67$.

The plots in Figure 5.6 show the two-dimensional features space overlapped by the $\chi^2_{p,\alpha}$ ellipsoids of three robust estimators: OGK, FMCD and DetMM. Whereas, in Figure 5.7 the same feature space is overlapped with the contour plots from the SMDs computed using the kMRCD. As a comparison, the $\chi^2_{p,\alpha}$ ellipse calculated from classical SMDs (cMahal) is shown. Note that these ellipses are equivalent to the decision boundary or the threshold that was shown previously in the SMDs plots. In summary, the following can be deducted from these two figures:

1. On the cMahal plot, the sample covariance matrix is calculated using the samples from all three system states (AFR 1,2 and 3). This caused the corresponding ellipse to expand towards AFR 3 samples that are furthest away. This is an example of a swamping effect, where outlying samples (AFR 3) shifted the ellipse such that a proportion of inlying samples (AFR 1 and 2) are being rejected as belonging to the reference states;

2. On the same plot, the first threshold is based on a typical confidence level, e.g. 99%, on the χ^2-distribution with DOF $p = 2$. Thus, its ellipse encompasses the 99% of the samples from all three system states. Using such a threshold, AFR 1 samples will be identified as outliers, while AFR 3 (the actual outliers) will be inlying. The second threshold, is based on the same χ^2 distribution assumption, but, with a confidence

level of 67%($= \alpha$). Thus, a rejection of 33% of the most outlying samples is made using the same centroid as the 99% ellipse. Using this ellipse, it will incorrectly identify samples from AFR 1 and 2 as outliers (swamping effects);

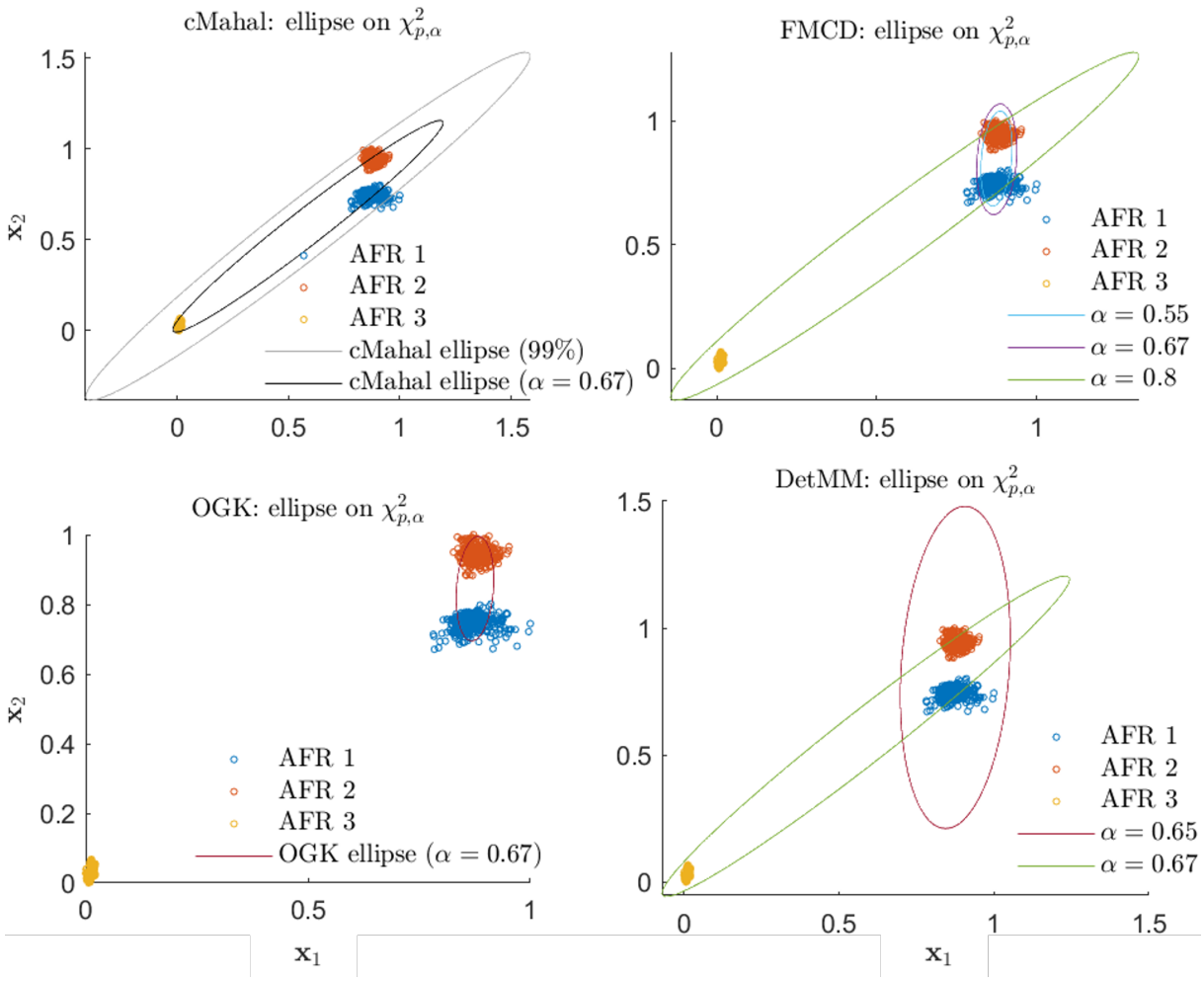

Figure 5.6. TF-LBO: a two-dimensional problem (selected two features from the MD_2 DSF matrix) to illustrate the workings of OGK, FMCD and DetMM robust estimators using ellipses on the χ^2 CDF and sensitivity to α value.

3. The next three plots in Figure 5.6 show the sensitivity of the parameter α on the FMCD and estimators. The ellipses were computed using the same χ^2-distribution assumption with varying values of α, where for these robust estimators, it also implies different values for scatter and location. While, for the OGK estimator, as explained previously, its estimates are not a function of α, but, depend on the univariate robust scale estimate method used. On the other hand, DetMM is particularly sensitive to α: a small change from .67 to .65 results in a very different ellipsoid shape that encompasses only the samples from AFR 1 and 2 and rejects AFR 3. Similarly, for

the FMCD ellipse, as α decreases from 0.8 to 0.67, it correctly encloses only AFR 1 and 2. However, as $\alpha \to 0.5$, which is the breakdown point, it will begin rejecting samples from AFR 1 and 2, as well. That is, similar estimates as with the OGK method. Therefore, correct / suitable specification of α is critical for the performance of robust SMDs: too tight will result in rejection of sample points from the reference state, while too loose will result in masking and/or swamping effects;

4. The previous plots in Figure 5.5 showed that DetMM correctly identified all samples using the correct specification of α. However, in the two-dimensional problem shown in Figure 5.6, this is not the case: an α value of .67 showed that this was not suitable - a further decrease to .65 was required to reach the same result. This may be explained by the fact that in the two-dimensional problem only two features from the DSF matrix were used. Thus, certain features that were in the original DSF matrix also played a major role in helping to identify the correct reference states (AFR 1 and 2), whilst rejecting the true outliers (AFR 3);

5. In Figure 5.7, the contours of the SMDs computed using the kMRCD with three different α values show the critical importance of that parameter. For $\alpha = 0.8$, kMRCD encompasses samples from AFR 3, as well, while for $\alpha = 0.55$ it incorrectly rejects AFR 1 samples from the inlying set. On the other hand, with $\alpha = 0.67$ it includes all data points from AFR 1 and 2, while some samples from AFR 3 have high SMDs. Note that, the ellipse defined for both FMCD ($\alpha = 0.67$) and OGK estimators in Figure 5.6, it may be shown that it will result in a higher number of FP rate (incorrectly rejecting both AFR 1 and 2 samples as inlying), as compared to kMRCD ($\alpha = 0.67$). This is an advantage of kMRCD since it employs an RBF kernel (to project the data), so that it may encompass all samples from AFR 1 and 2 using a (more flexible) non-linear decision boundary.

The classification of outliers is also an important part of this analysis. As described previously, this is done by plotting the orthogonal and score distances against each other, i.e. the outlier map of rPCA. In Figure 5.8, the outlier map of rPCA with $k = 2$ (the number of principal components) and $\alpha = 0.67$ is shown. As seen, the majority of the samples from AFR 3 are classified as bad leverage points, which implies that these samples will influence non-robust estimates such as the sample covariance matrix, as observed previously. On the other hand, rPCA correctly identifies almost all samples from AFR 1 and 2 as regular points. This observation, may further strengthen the previous argument made in point (4), regarding the influence of certain DSFs on the diagnostic performance of outliers. That is, by including all features from the MD_2 DSFs matrix, AFR 1 and 2 are more "easily"

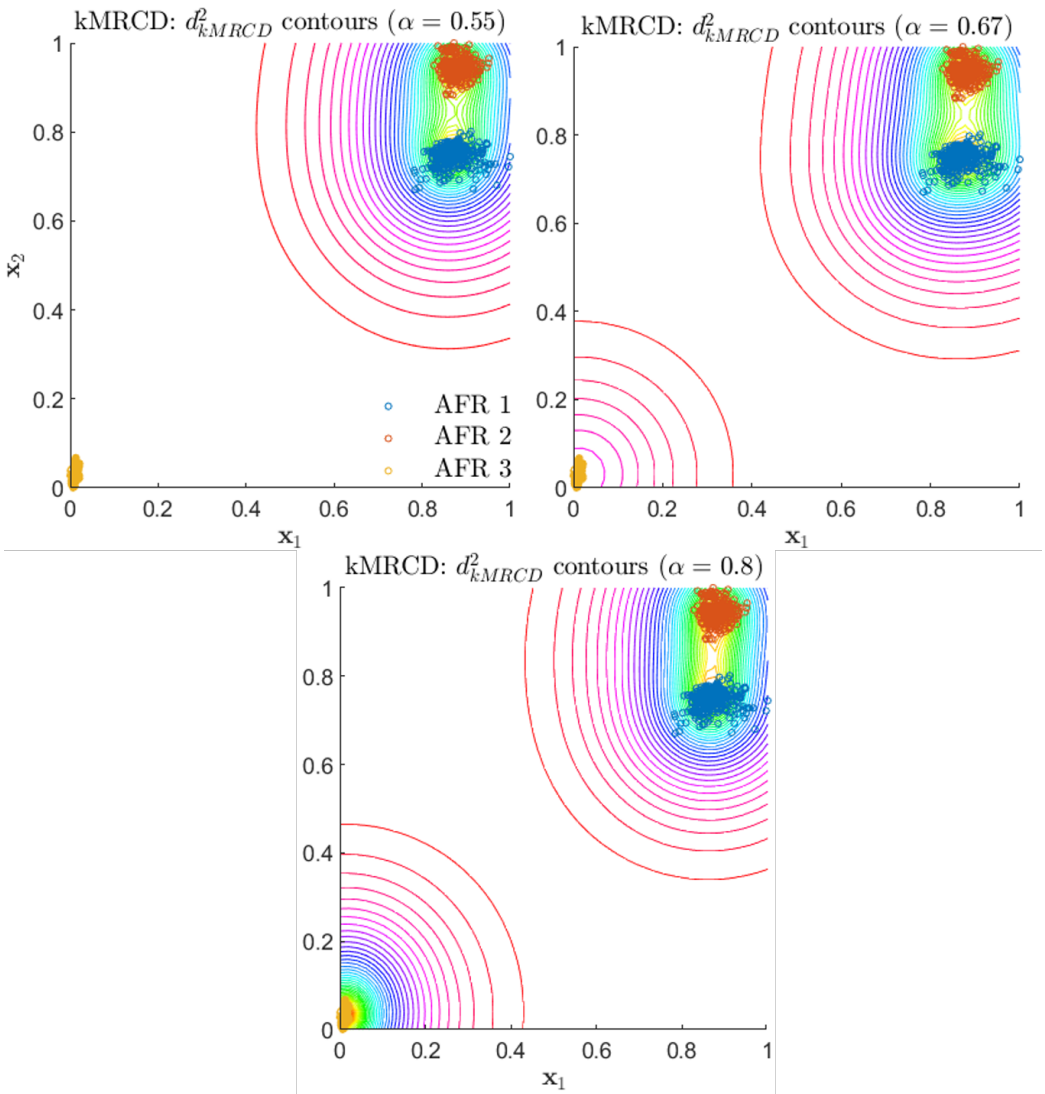

Figure 5.7. TF-LBO: same as in Figure 5.6 but kMRCD workings are illustrated using the contours of its computed SMDs.

distinguishable from AFR 3.

Moreover, an NLTS time series model was built as a quadratic regression model ($A = 2$). Each time series model was fitted on each feature vector from the MD_2 DSFs matrix, individually. The fitted estimates $\hat{y}$, the actual inliers $\mathbf{x}$ and estimated outliers (using the default-univariate threshold) $\mathbf{x}_o(\gamma_{uv})$, are shown on eight different features in Figure 5.9. In each time series model the fitted quadratic curve (with three coefficients) is tracking well the changes from AFR 1 and 2 (e.g. in Feature 1). At the same time, the rather abrupt changes in amplitudes of these features associated with AFR 3 state, are correctly identified as outliers. The A_{ROC} metric of each NLTS model is $\approx .99$. Not that a quadratic time series model in Features 3, 4, 5 and 7 is not required: a simple linear regression model ($A = 1$)

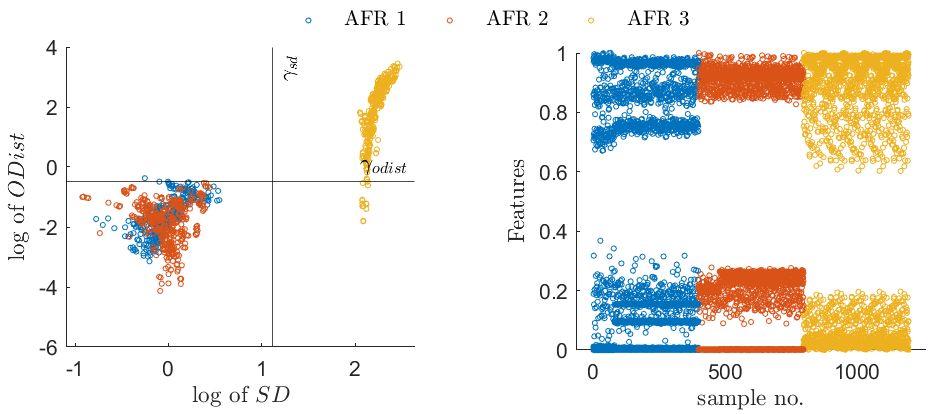

Figure 5.8. TF-LBO: outlier map of rPCA (on log-scale) and the MD_2 DSFs as a function of sample no., rescaled in range $[0,1]$ to aid in visualisation. Thresholds on the SD and ODist, γ_{sd} and γ_{odist}, respectively, were calculated as already explained.

should have been sufficient to identify the outliers.

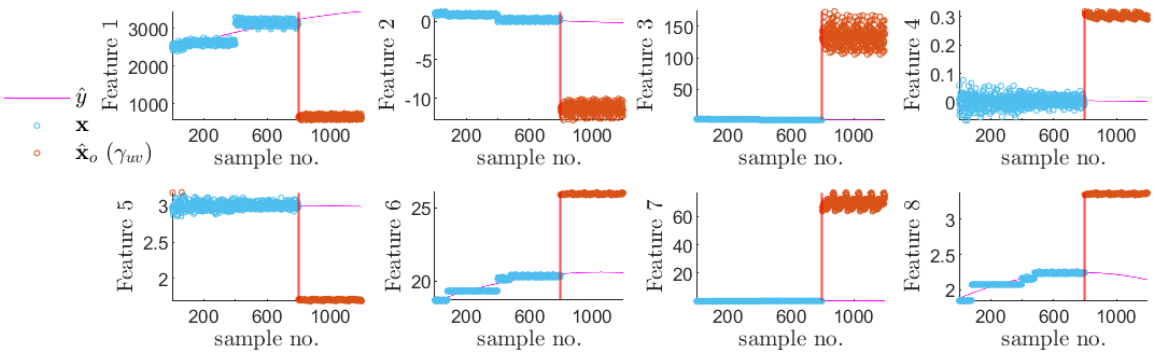

Figure 5.9. TF-LBO: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.67$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Time series samples are the eight different features from the MD_2 DSFs matrix.

The two benchmark methods: ocSVM and iForest, were also applied in this dataset. In iForest the outlier detection accuracy suffered: $A_{ROC} = 0.77$, while for the ocSVM the same metric was 0.98. Note that for the latter method, this value was achieved by first computing the robust z-scores of each MD_2 feature. This increased the outlier scores computed from AFR 3 samples, while at the same time it reduced the outlier scores of the rest of the samples, especially for AFR 1 state. In Figure 5.10 the outlier scores computed using ocSVM with $\alpha = 0.67$, $\nu = 0.5$ and $\gamma_{rbf} = 1.2$ are shown. Note that, when no pre-processing was applied its performance fell considerably, almost as close to a random classifier (≈ 0.5). This highlights the importance of utilising methods, e.g. FMCD and DetMM, that have robust properties, like affine equivariance: see Appendix C for a discussion on robust properties of

estimators.

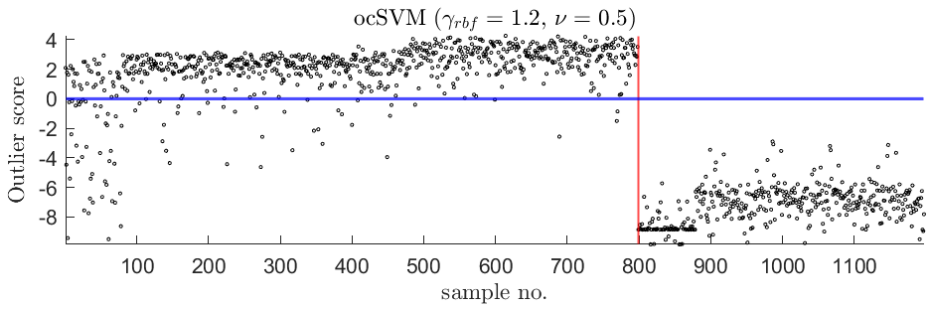

Figure 5.10. TF-LBO: outlier scores calculated using ocSVM with $\alpha = 0.67$, $\nu = 0.5$ and $\gamma_{rbf} = 1.2$. The first sample from the damaged state is shown by the vertical line, while the horizontal lines are the two thresholds.

5.6.3 NSim-4DOFLin & -4DOFNonLin datasets

Using these two simulation-based datasets, the amplitudes of the PSD were calculated on each segment time interval of 8 hours. Subsequently, a downsampling by two was performed to obtain the final matrix with dimensions 479×115. Due to the high-dimensionality of the problem, rPCA was employed and its first 10 principal components were chosen to form the final modal-based DSFs matrix. Note that the implementation of rPCA limits the maximum number of obtainable components to 10. However, in these two datasets the cumulative variance of the first 10 components was $\approx 95\%$. This implies that a projection onto these 10 axes of rPCA, contains almost (95%) all available information from the dataset.

The particular choice of the time interval for each segment, the downsampling used and the low-pass filtering of $\leq 0.1\ Hz$ provided sufficient monitoring of the variation in NFs. That is, both as a function EOV influence (temperature changes) and damage. This is because, understanding EOV influence is equally important as rejecting it, for instance, by using impact-based DSFs (see Chapter 4 for details on impact-based DSFs and how EOV influence may be rejected). Due to the simplicity of the dataset, e.g. no experimental errors, and noise contamination, it is possible to explore EOV influence more optimally, as well as, find way for its rejection.

For brevity, the analysis of the results will be focused on the NSim-4DOFNonLin dataset. This is for two main reasons: 1. it represents a more realistic case due to its non-linear restoring forces, and 2. no particular further insights could be gained by analysing its linear counterpart, i.e. the NSim-4DOFLin dataset[§].

In Figure 5.11, 10 time series models using NLTS with $A = 2$ were fitted on each of the 10 features separately. What is most interesting in these plots, is the fact that on the last 3

[§]SMDs for NSim-4DOFLin may be shown in Figure D.6 (Appendix D)

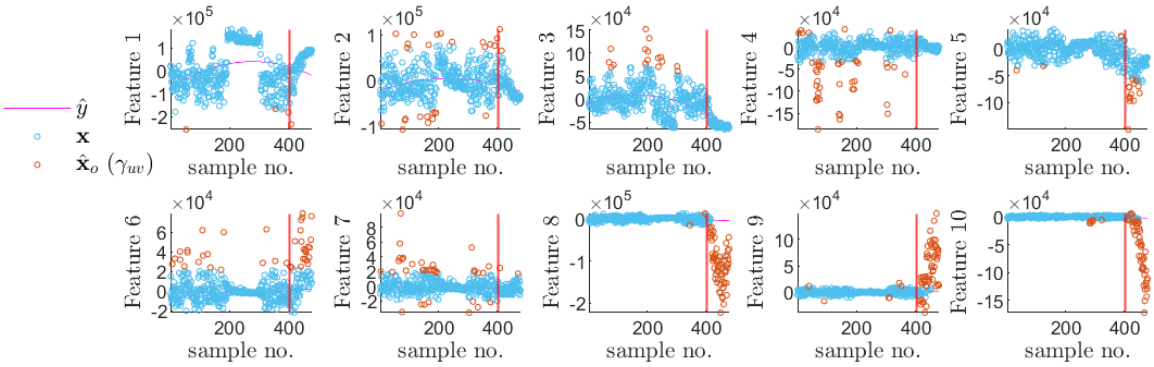

Figure 5.11. NSim-4DOFNonLin: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.75$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Each time series is one of the 10 principal components of rPCA, obtained from PSD amplitudes.

features, Features $8-10$, the EOV influence is almost non-existent. This is an important observation, which highlights some of the important attributes of working on the subspace of the feature space, which is spanned by a few principal components. In this subspace, which exists in a linear manifold of rPCA space, samples representing EOV influence and/or damage may be isolated. This possibility has been explored in the SHM literature in the past [58], [169], for obtaining DSFs as the minor components of PCA, i.e. the components with the least variance. These components showed immunity to EOV influence, while being sensitive to damage.

As the NLTS fitted lines $\hat{y}$ suggest on the last three rPCA axes, there is a negligible amount of outliers prior to damage. As the highest variance is expected in the first principal components, Feature 1 shows to be influenced by both EOV (sample no $\approx 200-300$) and damage. Thus, we may choose to isolate EOV on Feature 1 and damage on Features 8 to 10. Similar observations can be made by looking at Figure 5.12. In particular, the first two principal components of PCA (middle plot) show that most of the samples representing damaged and undamaged states ($T > 0$) are overlapped. At the same time, samples from EOV influence (undamaged with $T \leq 0$) are weighted heavily in these two axes of PCA. In Appendix D, in Figure D.3 the first 10 principal components of PCA show that there is no axes in which either damage or EOV influence may be isolated. This is another example of masking effects, where larger EOV influence variation hides smaller / subtler changes due to damage. Also, in kPCA all three system states are completely overlapped on its first two principal components. This may suggest a few things: 1. more principal axes are required in kPCA to discriminate between the system states, and 2. better tuning of its kernel width is required to reveal such differences. On the right plot, the first 10 principal components from PCA show that the amplitude of these features will be greater for EOV influence than

with damage, which makes DI a challenge.

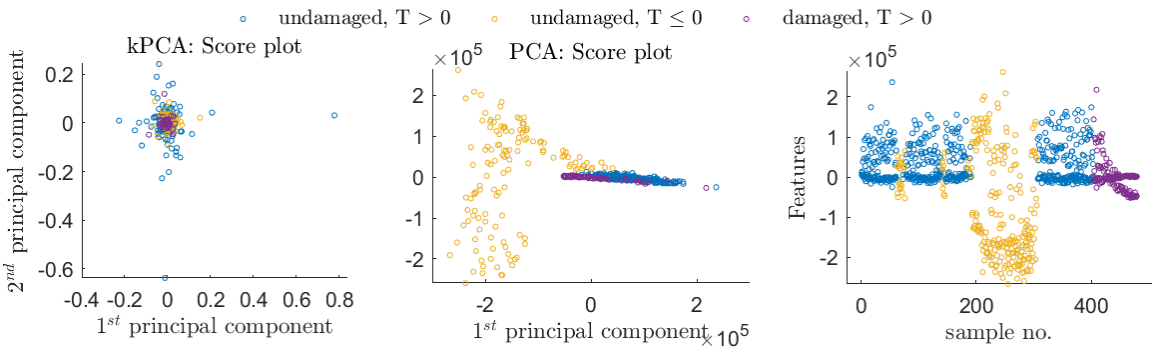

Figure 5.12. NSim-4DOFNonLin: score plots of the first two principal components of PCA (middle plot) and kPCA (left plot). On the right are the 10 principal components of PCA, computed from the PSD amplitudes.

The influence of outliers, and in particular, the masking effects of this dataset can be clearly seen on the two-dimensional plots, as it was done in a previous dataset: TF-LBO. In this dataset, however, we can select the features from the DSF matrix that are least immune to EOV influence, but, show some indications of damage. As it was shown in Figure 5.11, this can be done by selecting the minor components of rPCA, which in this case are the last three axes: Feature $8-10$. On the other hand, in Feature 3 and 4 we observed a large EOV influence, while damage was not discernible from these plots.
In Figure 5.13 the following remarks can be made, using Features 3 and 4:

1. The SMDs computed using kMRCD show as extreme values most of the samples from the undamaged condition with significant EOV influence ($T \leq 0$). While, samples representing the damaged state are considered as the "inlying" ones.;

2. Similar observations can be made for the FMCD, which draws a tight ellipse (using the χ^2 assumption with $\alpha = 0.75$ as the confidence level) around the samples representing the damaged state and the undamaged state at $T > 0$;

3. Due to its non-ellipsoidal assumption, kMRCD is able to identify more samples from the damaged state as outliers, in comparison to FMCD;

4. The SMDs of the FMCD and kMRCD methods, show the masking effects very clearly, whereby it is not possible to admit a threshold and discriminate all (or the majority) of the samples from damaged and undamaged states.

While, in Figure 5.14 the following remarks can be made, using Features 8 and 9:

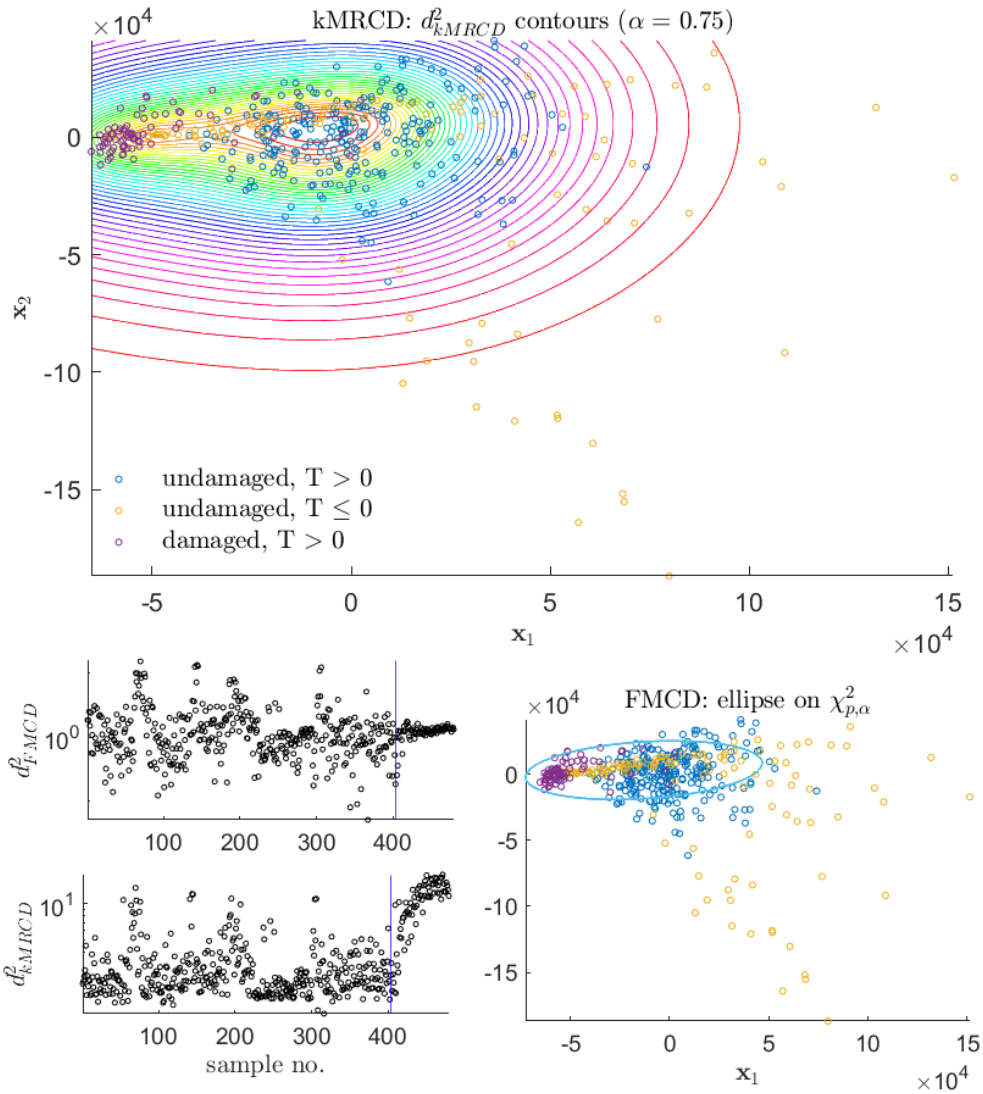

Figure 5.13. NSim-4DOFNonLin: two rPCA component scores (Features 3 and 4 in Figure 5.11) showing (on the top) the SMD contours of the kMRCD method and (on the bottom right) the χ^2 ellipse computed from FMCD method. The two bottom left plots show the corresponding SMDs, where the vertical line is the first sample from the damage state.

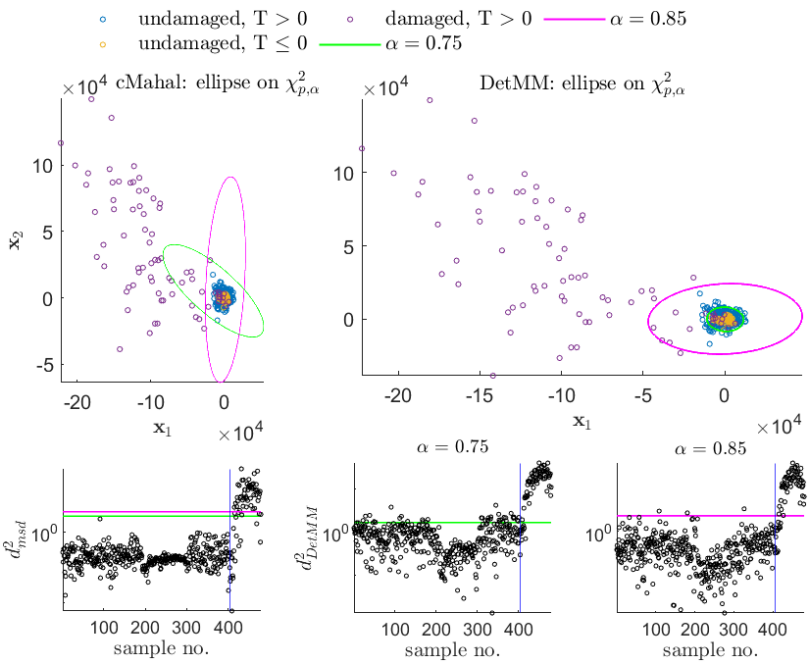

Figure 5.14. NSim-4DOFNonLin: top two plots of rPCA components (immune to EOV influence) overlapped with χ^2 ellipses computed using robust and non-robust estimates with varying confidence levels (α values). Bottom plots are the SMDs computed on the same features using DetMM and cMahal estimators.

1. As can be seen, these two-dimensional features have high immunity to EOV influence and high sensitivity to damage. Using these two features as DSFs, the classical Mahalanobis distance will have no problem in identifying damage at an early point using the χ^2 threshold. This is especially true when the threshold is computed using the correct amount of expected outliers, i.e. $\alpha = 0.85$. While some samples from the damaged state are within its inlier region, the error rate will be sufficiently small;

2. In the same figure, the SMDs and ellipse (although it is skewed differently) of the DetMM estimator indicate similar performance as with the classical Mahalanobis technique.

Note also that a robust z-score, computed on the 10 rPCA components, using iForest have produced similar outcomes. This is because, a robust z-score highlights the true outlying samples by normalising all samples accordingly. To maintain clarity of the results, the outlier scores of iForest are shown in Figure D.4 (Appendix D).

From the outlier map in Figure 5.15, we observe that EOV influence exists orthogonal to the 10-dimensional subspace of rPCA. On the other hand, samples from the damaged state are lying on that 10-dimensional subspace but are further away from regular samples (undamaged, $T > 0$). As such the former are classified as orthogonal outliers, while the

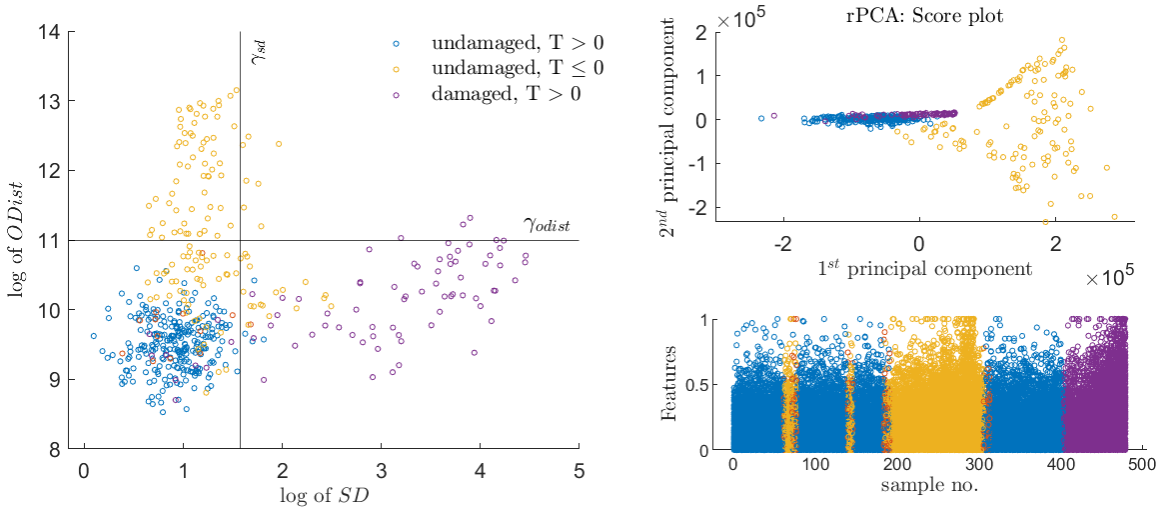

Figure 5.15. NSim-4DOFNonLin: the left plot is the rPCA outlier map on a log scale (10 rPCA components). On the top right, the score plot of rPCA's computed on its first two components. On the bottom right, the DSFs, i.e. the PSD amplitudes, unit-normalised.

latter are good leverage points. By observing the 10 principal components of PCA (see Appendix D in Figure D.3), it is clear that these orthogonal outliers had a major influence in defining the principal axes, since variance is maximised in these directions.

As a comparison, the 10 DSFs computed using the PSD amplitudes of rPCA and PCA are compared in terms of A_{ROC} performance in Figure 5.16. As it is shown, all estimators including cMahal have better A_{ROC} metrics when the 10 rPCA components are used as DSFs. Note that all robust estimators, apart from DetMM are specified with $\alpha = 0.75$. Surprisingly, the DetMM estimator, has shown considerable improvement in performance when the specified outliers fraction $1-\alpha$ exceeds its theoretical maximum breakdown value α^* of 0.5. In general, as $\alpha \to 0.5$, the determinant of the scatter matrix becomes smaller, forming a tighter ellipsoid around the inliers. The fact that the DetMM estimator yielded superior performance (more precisely, in terms of A_{ROC}), has to do with the fact that its contamination factor can be higher than the 50% maximum breakdown point. This is also observed in the WT-REB dataset, where this is explained further.

A more conservative threshold may be specified using the MC simulation that was discussed previously, with a confidence interval set at α value. As shown in Figure 5.17, the MC threshold γ_α^{mc} results in a decrease in FP rates for the robust estimators, as compared to the χ^2-based threshold. On the other hand, for cMahal it will result in a considerable increase in FN rate. For kMRCD, the MC threshold was impractical to calculate due to the high-computational resources required, as explained previously.

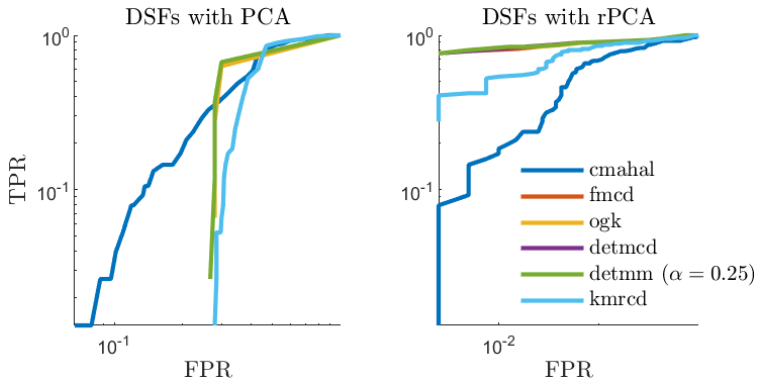

Figure 5.16. NSim-4DOFNonLin: ROC curves (on log − log scales) for the computed SMDs from five robust estimators and cMahal using two different DSFs: the first 10 principal components from rPCA (right) and PCA (left).

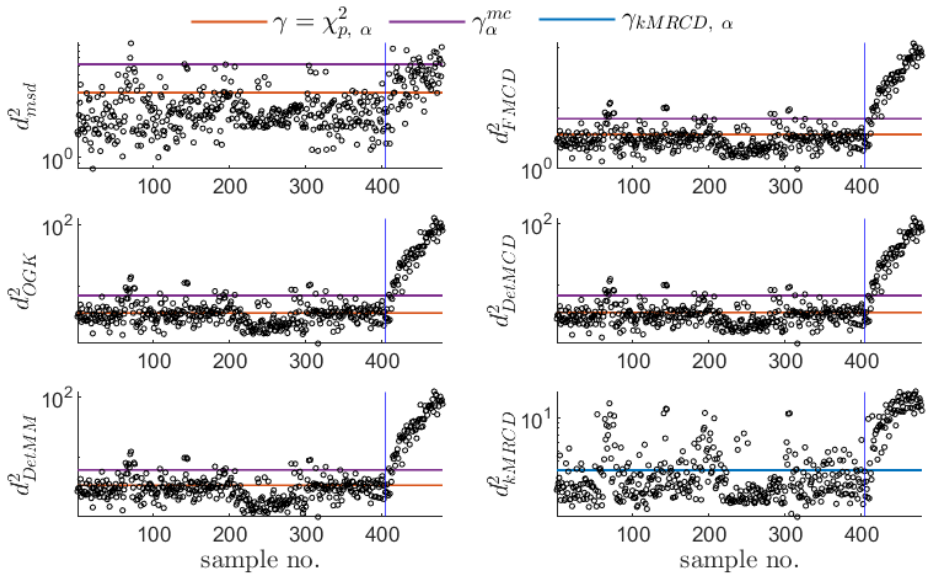

Figure 5.17. NSim-4DOFNonLin: SMDs computed using the robust and non-robust estimators. Three vertical lines are different thresholds set on the α confidence interval: MC threshold (γ^{mc}_{α}), on log-normal ($\gamma_{kMRCD,\alpha}$) and on the χ^2 distribution with p DOF ($\chi^2_{p,\alpha}$).

5.6.4 LA-4DOF dataset

In LA-4DOF dataset the impact-based DSFs are used, i.e. for each nth segment (of time interval equal to 8 seconds) as the kurtosis values of each decomposed IMF from an adaptive decomposition technique, such as VMD. These kurtosis values were then filtered using the MED technique, as it is discussed in Chapter 4. Although the details of the analysis are not presented here (for clarity), it was found that the SMDs calculated using DSFs from SVMD produced the most satisfactory results, in terms of DI (A_{ROC}). Therefore, and unless otherwise specified, the IMFs used to compute the impact-based DSFs are derived using the SVMD method.

In Figure 5.18 the impact of specifying the correct α value is again examined on the DD plots. These plots are the SMDs of the five robust estimators against the SMD from the sample covariance and mean. The most important information that can be gained from these plots is that they highlight deviation from the MLE estimates (sample mean and covariance). Where the distance from the diagonal line may be used as a reference point for that deviation. As Figure 5.18 shows as α decreases from .75 to ≈ 0.5, more points deviate from the Gaussian distribution assumption (furthest away from the diagonal). This is because robust estimators place no assumptions on the estimated outliers $\hat{\mathbf{x}}_o$. On the other hand, the samples that are estimated as inliers $\hat{\mathbf{x}}_i$ (samples that are below the threshold lines γ) remain close to the diagonal line, especially, for DetMM, DetMCD and FMCD estimators when $\alpha = 0.75$. Note that for kMRCD inliers will not generally be expected to be lying along or close to the diagonal since its estimated inliers are not expected to follow an elliptical distribution assumption as the other robust estimators do. And for OGK, there is no change with α, since it does not depend on it.

The comparison in performance, in terms of A_{ROC} of the calculated SMDs from all five robust estimators and the classical Mahalanobis, is shown in Figure 5.19 for two α values. When α is selected arbitrarily, e.g. $\alpha = 0.75$, all robust estimators will have better performance than using the sample covariance and mean (cMahal). On the other hand, when α is chosen appropriately to the expected outlier proportion, $A_{ROC} \to 1$. As discussed in NSim-4DOFNonLin dataset, DetMM performs well when its value decreases below the theoretical breakdown limit of α^* (in this dataset $\alpha = 0.2$).

Figure 5.20 shows the computed SMDs of the five robust estimators with $\alpha = 0.51$ and their non-robust equivalent. As the plots show, an MC threshold is most appropriate for the robust estimators[¶]. For kMRCD its threshold has been calculated differently, as mentioned

[¶]An MC threshold is computed using the dimensions of the training matrix, so that it is parameterised for the given dataset, instead of using a fixed threshold on the χ^2. Also, since the extreme values are computed using the robust estimates, it is also parameterised to fit each robust estimator.

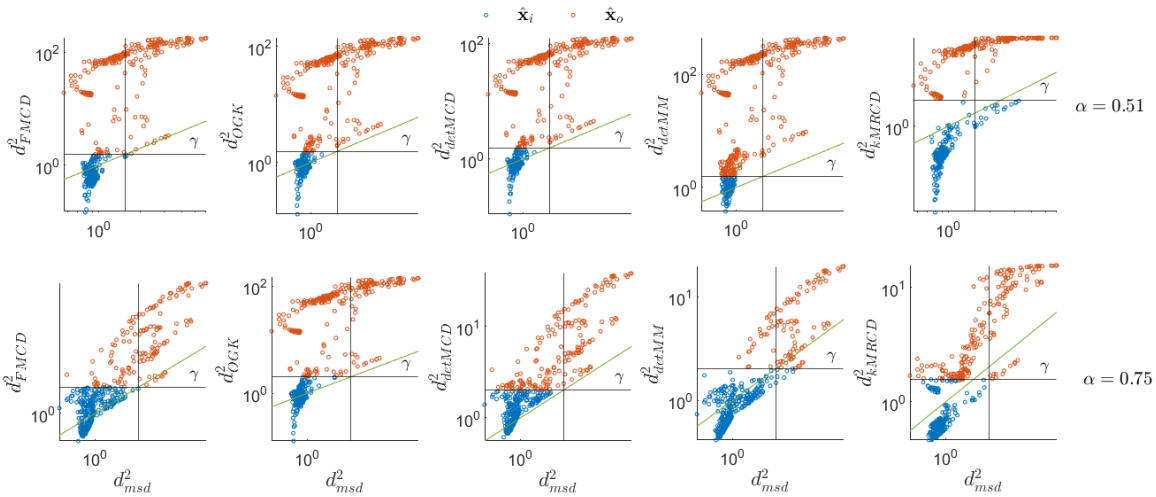

Figure 5.18. LA-4DOF: DD plots (on log-scales) of the SMDs computed from the five robust estimators on two α values. The detected outliers $\hat{\mathbf{x}}_o$ and inliers $\hat{\mathbf{x}}_i$, are shown as the points above and below the threshold lines, respectively. The diagonal line shift (with respect to the 45 degrees angle) shows deviation from the Gaussian distribution.

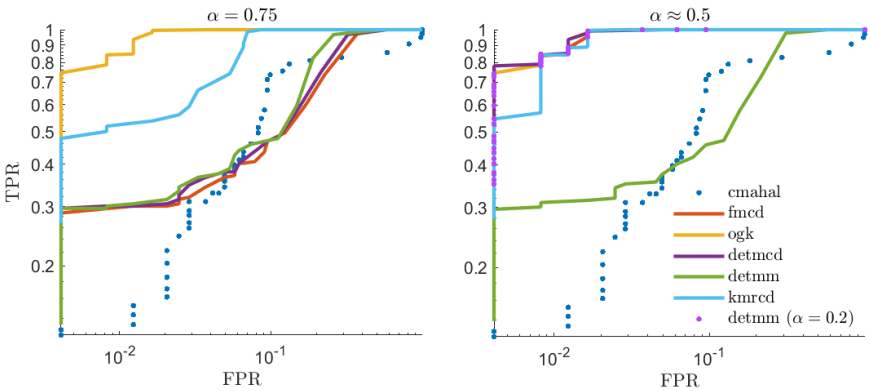

Figure 5.19. LA-4DOF: ROC curves (on $\log - \log$ scales) for the computed SMDs from five robust estimators on two different α values and cMahal. DetMM is showing considerable performance improvement when $\alpha < 0.5$. Impact-based DSFs (kurtosis) are used.

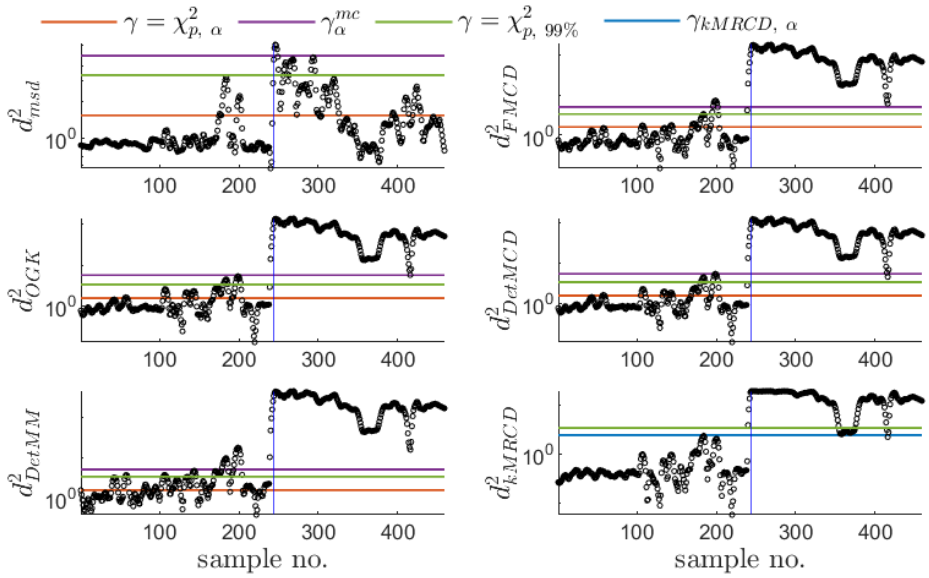

Figure 5.20. LA-4DOF: SMDs computed from five robust estimators and sample covariance and mean values. Four different thresholds. Three on α confidence interval: MC threshold (γ^{mc}_{α}), on log-normal ($\gamma_{kMRCD,\alpha}$) and on the χ^2 distribution with p DOF ($\chi^2_{p,\alpha}$). One on the χ^2 distribution with 99% confidence interval: $\chi^2_{p,99\%}$.

previously. As expected the 99% threshold from the χ^2 distribution is more conservative than the one computed with $\alpha = 51\%$. Note that, like the χ^2-based threshold, the MC threshold also implies that the estimated inlying samples are drawn from a Gaussian distribution. However, since the extreme values are computed using the robust estimators, its value is different for each one, i.e. more suitable for each one, too. As the outlier map in Figure 5.21 shows, samples representing the damaged state may be classified as bad leverage points (using this set of imapct-based DSFs). These points are expected to have significant influence on non-robust estimators, e.g. PCA, skewing the estimates towards their directions. Therefore, utilising a robust estimator, e.g. rPCA has the potential to minimize their influence on these estimates.

As it was shown previously in NSim-4DOFNonLin dataset, PSD amplitudes provide an opportunity to explore further methods towards achieving accurate and reliable DI. For that purpose, the first 10 rPCA components of the PSD amplitudes were also used as DSFs. For visualisation, the PSD amplitudes were shown in the previous chapter in Figure 4.6. However, a segment time interval of 8 seconds (instead of 2.5 seconds - as in Figure 4.6) provided better (more than 3 times) frequency resolution, producing 333 amplitudes. A band-pass filter was applied between 20 and 100 Hz so that EOV influence was also present in the data. Note that its NFs are $\approx \{30.5, 53.7, 70.8\} Hz$ and that damage is most profound above ≈ 80 Hz (see Figure 3.6).

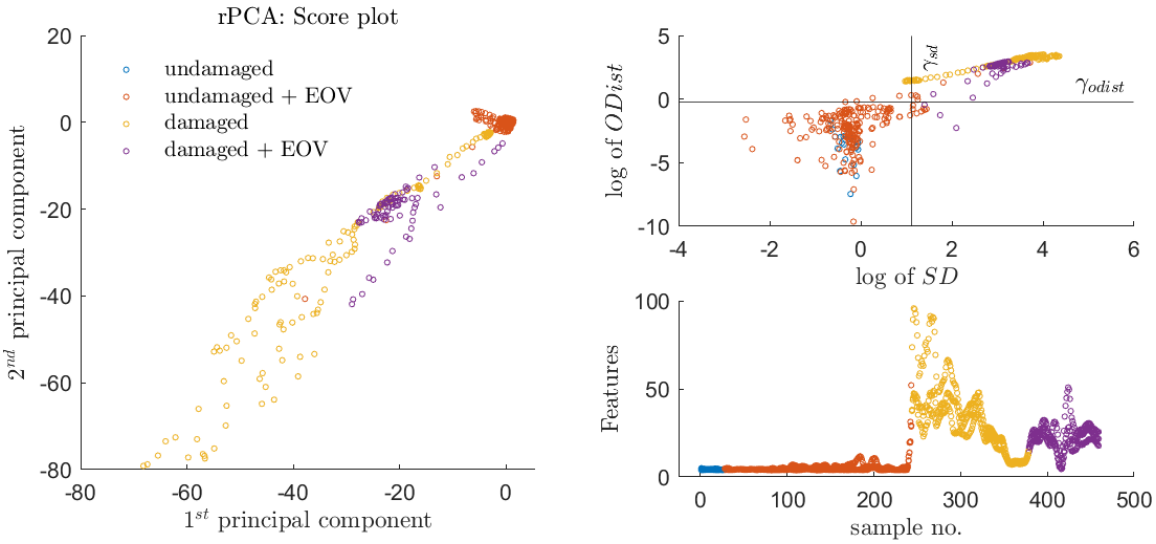

Figure 5.21. LA-4DOF: on the top right: the rPCA outlier map (on a log scale), on the bottom right: the impact-based DSFs (kurtosis) and on the left: the rPCA score plot - first two principal components.

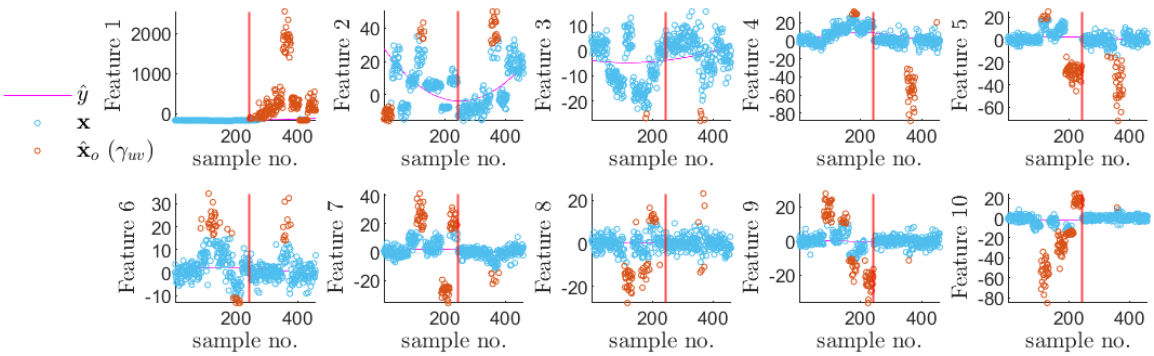

Figure 5.22. LA-4DOF: NLTS fitted line $\hat{y}$ with $A = 2$ and $\alpha = 0.51$. Showing also actual inlier samples $\mathbf{x}$ and estimated outliers $\mathbf{x}_o(\gamma_{uv})$, which were computed using the default-univariate threshold. Each time series is one of the 10 principal components of rPCA, obtained from PSD amplitudes.

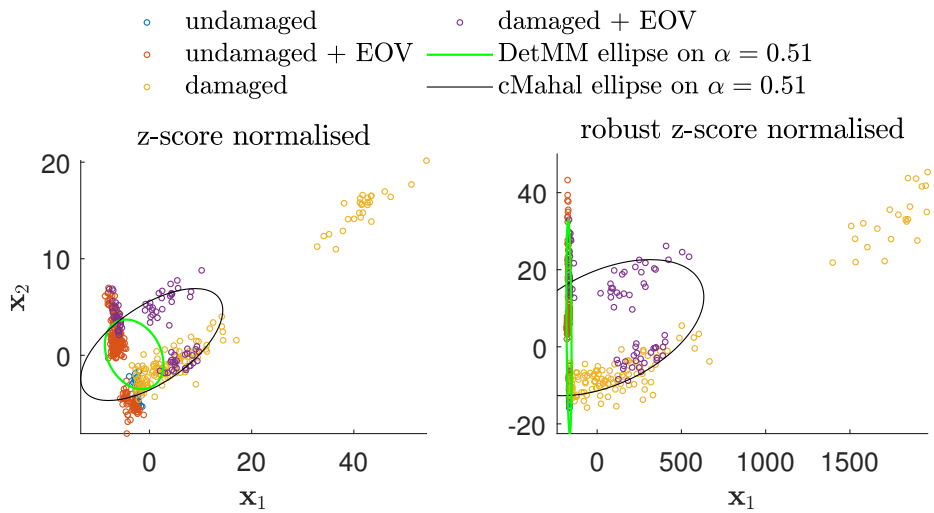

Figure 5.23. LA-4DOF: using as DSFs the first two rPCA components. On the top plot, the samples are z-score normalised, while at the bottom the same samples are robust z-score normalised. Ellipses on the χ^2 distribution from DetMM and the sample covariance and mean are overlapped.

In Figure 5.22 10 NLTS time series models are shown as quadratic univariate regression models ($A = 2$). Each model is fitted on the first 10 rPCA components, while they have all been robust z-score normalised. As seen in Feature 1, the samples representing damage have been isolated from undamaged samples having significant EOV influence. Therefore, utilising a simple model, e.g. $A = 2$, it is possible to reliably identify damage at an early stage.

Note that, using a different scale, in this case a robust z-score has been critical in isolating damage samples in a single feature with rPCA. In particular, subtracting the median and dividing by the MAD scale in each of the 10 columns of the DSFs matrix, has amplified the samples (the rows of the DSFs matrix) where amplitude is high. On the left of Figure 5.23, the first two DSFs (Feature 1 and 2) are shown with a z-score scale. On the right, these two DSFs are normalised using a robust z-score. As it was seen in Figure 5.22, the amplitude of the samples representing damage in Feature 1 are much larger than the rest of the samples. As such, on the right plot they lie far from the rest (the undamaged samples). DetMM forms a tight ellipse around the undamaged samples, while cMahal is "expanded" towards the "damaged" samples.

5.6.5 WT-REB dataset

As in the LA-4DOF dataset, impact-based DSFs (kurtosis) using the IMFs from the SVMD method were used for DI. The SVMD was, once more, chosen on the basis of the A_{ROC} metric by direct comparison with the rest of the adaptive decomposition methods. As

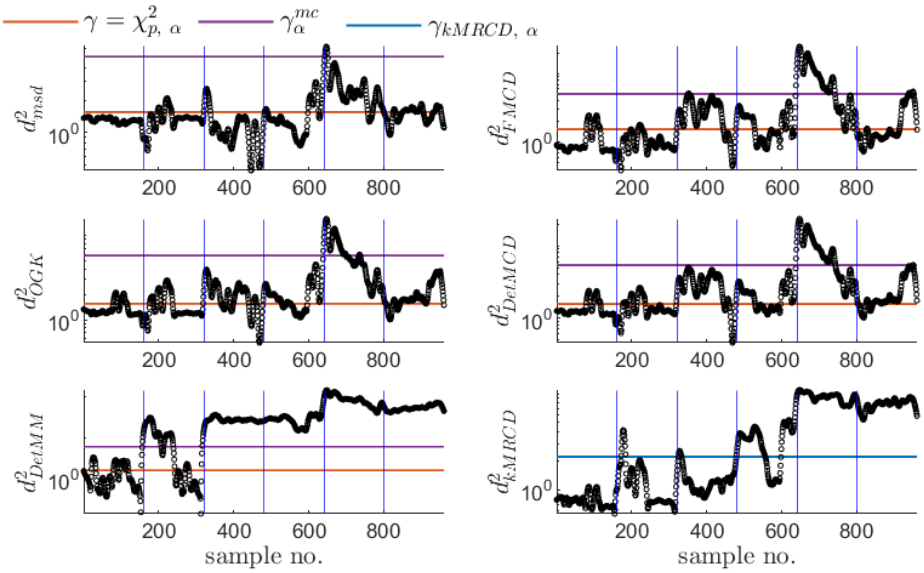

Figure 5.24. WT-REB: squared Mahalanobis distances (on log-scale) from five robust estimators using $\alpha = 0.51$, apart from DetMM whose $\alpha = 0.2$. The first sample from each level of damage is shown by the vertical line, while the horizontal lines are the thresholds. For comparison the Mahalanobis distances (on log-scale) using the sample mean and covariance are shown (d^2_{msd}).

Figure 5.24 shows, DetMM estimator identifies damage as early as level 1 (sample no 160), which could not be achieved using the EA method. In particular, as it was demonstrated previously in Chapter 4 and seen also in Figure 3.10, the loading condition results in both a shift and an amplitude decrease of BPFO and its harmonics. This in turn, reduces the diagnostic capability of EA dramatically. Using these impact-based DSFs, however, the loading condition has negligible effect from damage level 2 and onwards (sample no. 320−), where damage signals are rather significant at BPFO. On the other hand, the weaker damage signal of damage level 1, decreases to the noise floor when the load increases (or equivalently, is being masked by EOV influence).

Note that the particular performance of the DetMM is attributed to the fact that it is possible to reduce $\alpha \to 0$, as discussed previously. In this dataset, where the number of outliers far exceed the number of inliers, i.e. outlier percentage is $\approx 83\%$, a method like the DetMM makes it possible to perform DI more robustly than the rest of the methods. The formulation of DetMM is not based on the estimation of the determinant of a subset of the samples. But rather, DetMM uses loss functions that allow the outlier fraction to fall below 50% without any issues, e.g. the determinant becoming zero (see also AppendixC). Additionally, techniques like kMRCD also performed well in this dataset, owing to its non-ellipsoidal shape assumption for the inliers.

The rPCA outlier map and score plot (of the first two principal components) for this dataset

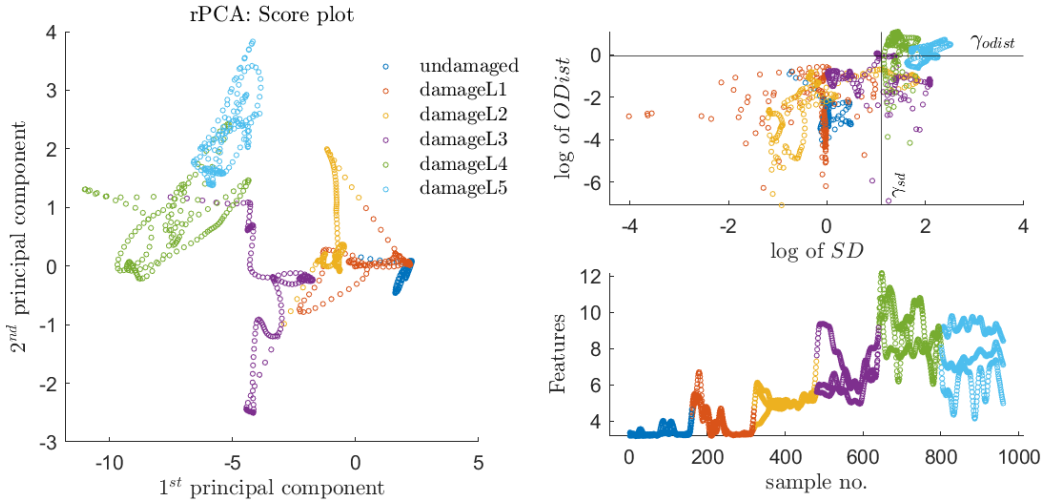

Figure 5.25. WT-REB: on the left: the rPCA score plot - first two principal components. On the top right: the rPCA outlier map (on a log scale), on the bottom right: the impact-based DSFs (kurtosis).

(Figure 5.25), reveals that damage levels $4, 5$ are further away from the other 3 damage levels and the undamaged system state. As the outlier map shows, these final two damage levels may be classified as bad leverage points. At the same time, regular points are damage levels 1 and 2, along with the undamaged state, while part of damage level 3 is classified as good leverage points and the rest as regular points. Therefore, when damage occurs, the first two levels of damage will lie very close to the undamaged state (and may not be identified easily). The samples at the next damage level will then proceed to become good leverage points (with minimal influence on non-robust estimators) and finally end-up as a bad leverage points (with significant influence on non-robust estimators).

In line with the above observations, using only two out of the three impact-based DSFs, it is possible to visualise some of the workings of robust estimators, in Figure 5.26. In particular, given the flexibility of the kMRCD technique, it is able to define a decision boundary such that damage level 3 is detected reliably. On the other hand, the χ^2 ellipse defined by a method that assumes a unimodal elliptical distribution for the inliers, e.g. the FMCD will consider a large proportion of damage level 3 as inliers. On the right plot, clustering analysis using tclust technique (applied on all three DSFs) shows that the undamaged state and the first two damage levels may be considered as coming from the same cluster $\hat{\mathbf{x}}_o$. On the other hand, damage level 3 and 4 form their own cluster, i.e. Cluster 1, while Cluster 2 is comprised of damage level 5 and a small fraction of damage level 4. This is similar to the outlier map, which showed that the undamaged state and the first two damage levels were classified as regular points, while the rest were classified as outliers.

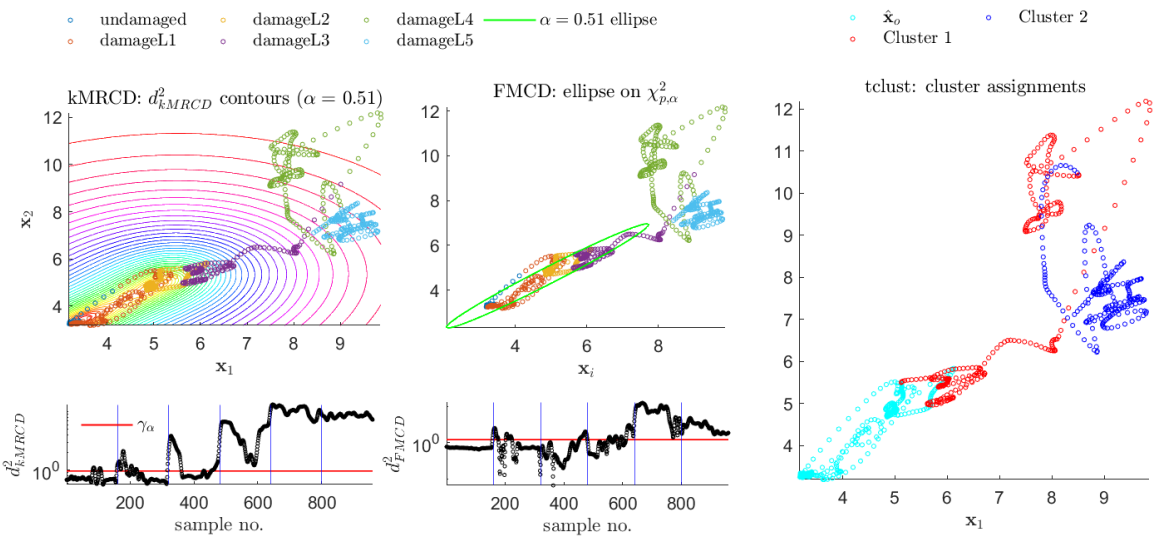

Figure 5.26. WT-REB: on top left: the χ^2 ellipse computed from FMCD method and the SMD contours computed using kMRCD, both with $\alpha = 0.51$. Bottom plots are the SMDs computed on the same two features using kMRCD and FMCD estimators, with the threshold set on the χ^2 distribution with confidence interval on α. On the right, the cluster assignments ($K = 2$) computed from tclust algorithm on the full 3D feature space of impact-based DSFs. The 2D plots are the first two impact-based DSFs.

5.6.6 Z24 dataset

The four NFs of the Z24 dataset are used for DI without any further processing, e.g. filtering operations with a deconvolution technique like MED. Given that this is a dataset from a field experiment it is interesting to explore EOV influence and damage as they influence its first four NFs. The rPCA outlier map in Figure 5.27 shows the fact that samples representing the undamaged state with significant EOV influence, i.e. when $T \leq 0$, will have the greatest influence on non-robust estimates. Therefore, any estimates will be skewed towards them, thus, masking samples representing damage that are orthogonal outliers and their influence will not be as significant. As the rPCA score plot shows on the first two principal components, the undamaged state is sparser than the other two states. Hence, techniques that rely on sparseness properties for the outlying samples will, in most part, detect outliers as points from the undamaged state with $T \leq 0$. As the outlier scores of the iForest algorithm show in Figure 5.28, when $T \leq 0$ their value increases considerably, i.e. greater than samples representing damage.

Figure 5.29 shows (on NF 2 and 3) the results from the clustering analysis that was carried out on all four NFs. The purpose of clustering analysis is to determine how samples from the different system states can be grouped together. Cluster 2 comprises mostly of samples representing damage, while the outlying samples $\hat{\mathbf{x}}_o$ are mostly samples representing the

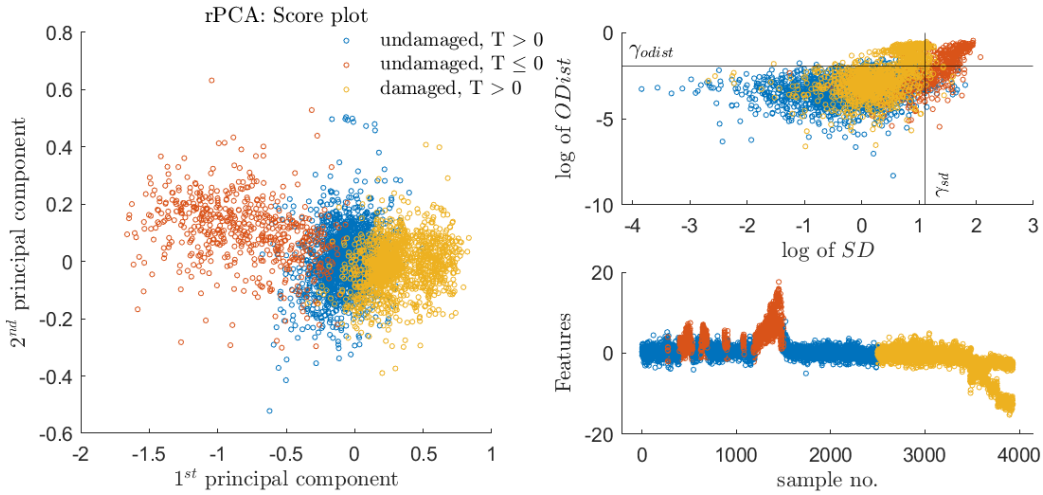

Figure 5.27. Z24: on the left: the rPCA score plot - first two principal components. On the top right: the rPCA outlier map (on a log scale), on the bottom right: the four NFs.

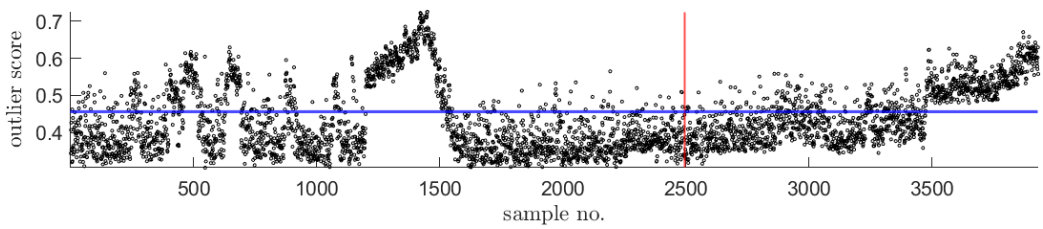

Figure 5.28. Z24: iForest outlier scores on Z24 dataset. Vertical line is the first sample of damage. Horizontal line is the threshold.

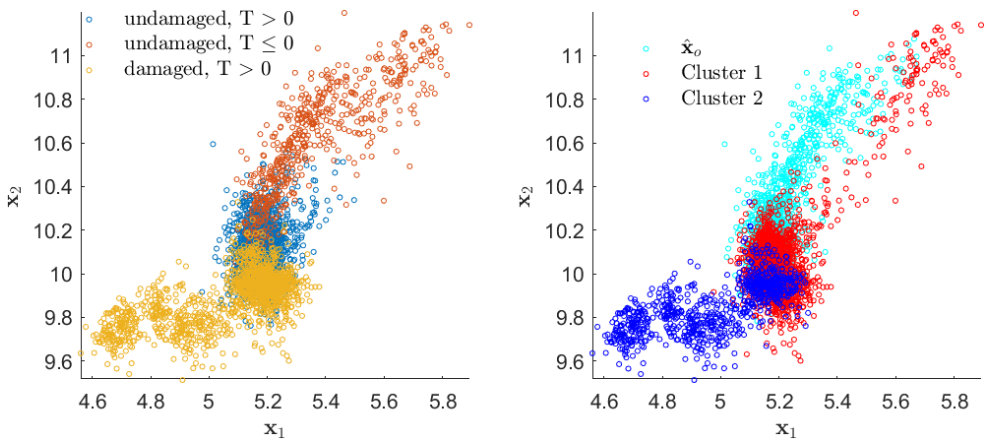

Figure 5.29. Z24: On the right are the clustering results using tclust algorithm ($K = 2$), while on the left are the labelled samples shown on the second and third NFs.

undamaged state, having significant EOV influence. Hence, in line with the previous observation made: most methods consider EOV influence samples as outliers. As it is shown also in clustering analysis, these samples do not seem to be members of a single cluster as they exist in lower-density regions in this four-dimensional space of features. Therefore, they are shared between different clusters: in this case $\hat{\mathbf{x}}_o$ and Cluster 1 has samples from the undamaged state when EOV influence is significant $T \leq 0$.

5.6.7 TS-SAF & TF-SAF datasets

In these two datasets: TS-SAF and TF-SAF a number of alternative fuel mixtures and AFRs were tested on two different engines, as described in Chapter 3. The purpose is to identify operating conditions in which engine performance, in terms of vibration characteristics, has shown significant deviation from the rest (or the majority conditions). In both datasets, the global-based MD_2 DSFs were calculated on each nth segment of the acceleration data gathered from a single sensor. Note that the segment time interval for TF-SAF has remained unchanged from the previous results that were presented in Chapter 4, which is 2 seconds. While, the segment time interval for TS-SAF has decreased to 0.5 seconds, which increased the number of samples N by 4 times. This was deemed necessary if we were to draw some inferences from this dataset, e.g. identify groups.

Figure 5.30 shows the clusters that were identified using the tclust algorithm, with $K = 1$ and $\alpha = 0.75$ (on a plot with the first two DSFs from MD_2 - $\mathbf{x}_1$ and $\mathbf{x}_2$). The outlying cluster $\hat{x}_o$ comprises of samples representing SAF @ AFR 1. The next set of results, in Figure 5.31 show the identification of outliers using the first two rPCA components (for visualisation purposes). On the plot at the top, the sample covariance matrix is expanded towards SAF @ AFR 1 samples. On the plot in the middle, FMCD ($\alpha = 0.75$) identifies as

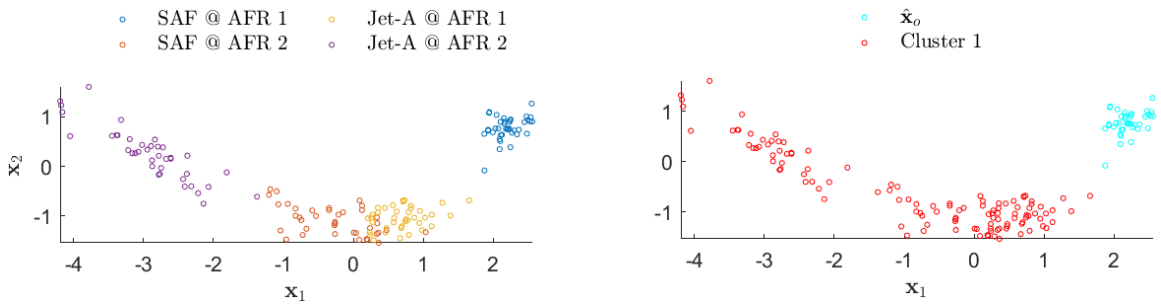

Figure 5.30. TS-SAF: On the right are the clustering results using tclust algorithm ($K = 1$), while on the left are the labelled samples shown on the first two features from the MD_2 DSFs matrix.

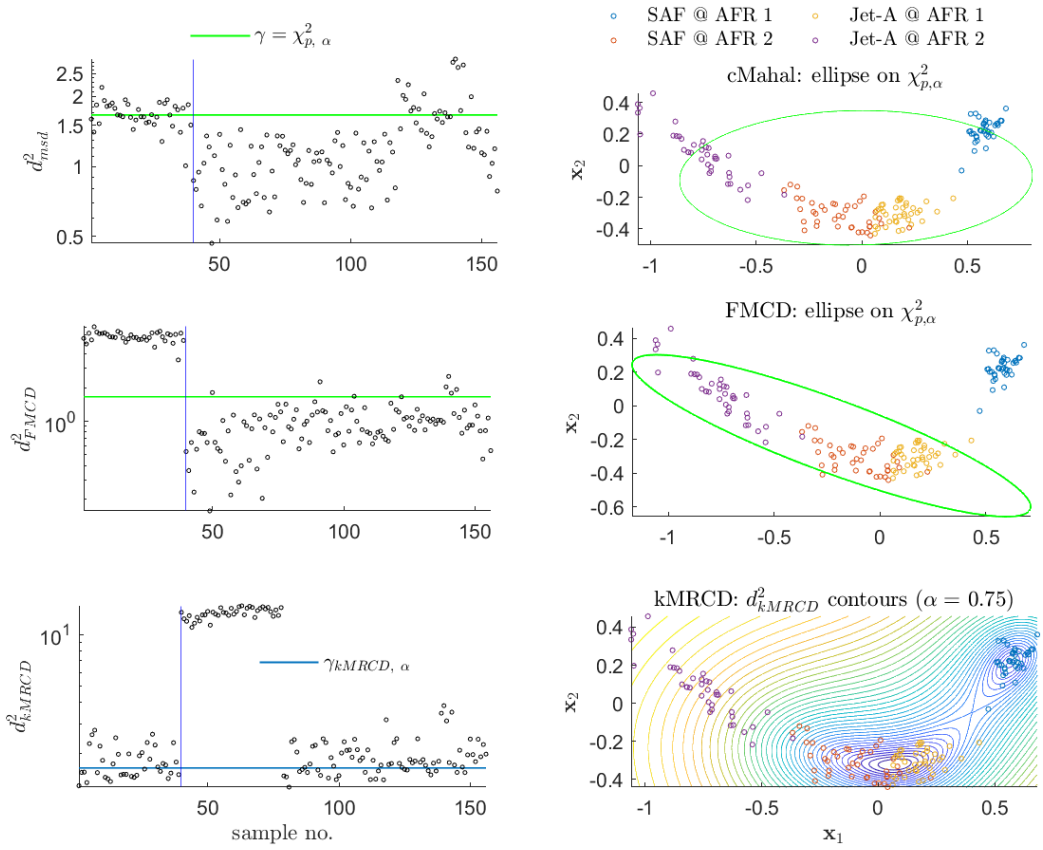

Figure 5.31. TS-SAF: On the left are the SMDs computed using kMRCD, FMCD and cMahal methods on the first two rPCA components. Vertical lines are the thresholds. On the right are the corresponding χ^2-based ellipses and SMDs contours (for kMRCD) on the same rPCA components.

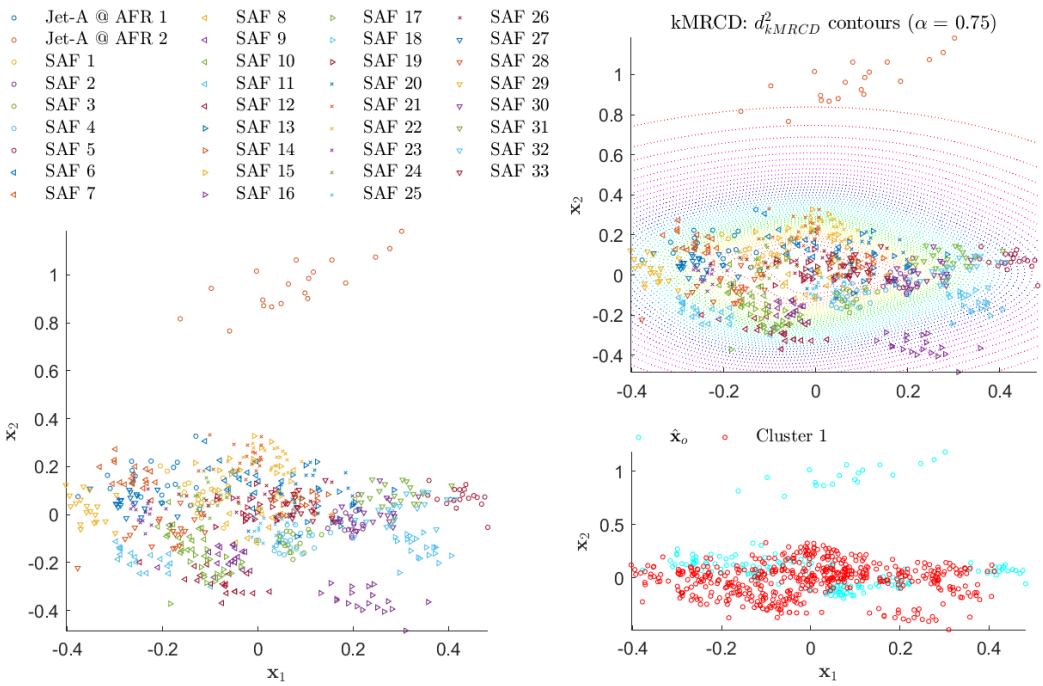

Figure 5.32. TF-SAF: on the left are the first two rPCA components of the MD_2 DSFs matrix. On the top right are the SMDs contours computed by the KMRCD method with $\alpha = 0.75$. And on the bottom right are the clustering analysis results computed from the tclust algorithm with $K = 1$ and $\alpha = 0.75$.

outliers the same condition as the tclust algorithm, i.e. SAF @ AFR 1. On the other hand, kMRCD ($\alpha = 0.75$) encompasses in its 2D rPCA space all conditions apart from Jet-A @ AFR 2. Therefore, these two robust estimators are in disagreement, with regards to the inlying samples. Note that both SAF @ AFR 1 and Jet-A @ AFR 2 samples, may be considered as deviant conditions as they are far away from the most central part of the data. Nevertheless, given that tclust operates in the full dimensional feature space of MD_2, it is more justifiable to classify SAF @ AFR 1 as the outlier / damage engine state.

Clustering analysis with $K = 1$ was also performed on the TF-SAF dataset, and the results are shown in Figure 5.32. Also, in the same figure the contours of the computed SMDs from kMRCD are shown, while $\alpha = 0.75$ is used for both techniques. The most important finding in this dataset is that the samples representing Jet-A @ AFR 2 are very distinct (in the sense of both clustering analysis and kMRCD) from the rest[‖]. Additionally, the engine condition SAF 5, which is an alternative fuel mixture from a sustainable source, was also found to be an outlier in this dataset.

[‖]Note that without further details on the dataset, e.g. fuel mixture compositions and actual AFR values, it has been impossible to draw further inferences from TF-SAF and TS-SAF datasets.

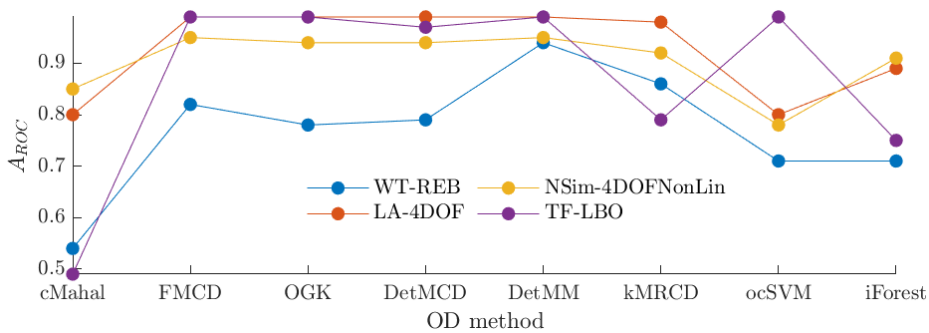

Figure 5.33. A_{ROC} metric as calculated on four different datasets for: five robust estimators of scatter and location, the classical Mahalanobis distance, ocSVM and iForest.

5.7 Summary

The results from inclusive outlier analysis were presented and analysed in this chapter. As each dataset represents a different problem overall, having certain distinct characteristics (e.g. high-dimensionality and specific set of DSFs), the analysis presented for each one varied. It should be noted that the A_{ROC} of robust estimators of scatter and location, was overall higher on most of the datasets. The A_{ROC} metric is shown in Figure 5.33 for the following datasets and associated DSFs: impact-based DSFs for WT-REB and LA-4DOF, global-based DSFs for TF-LBO, while for NSim-4DOFNonLin the first 10 rPCA components are used. Some general findings from this chapter can be made as follows:

- For all the aforementioned case studies and methods analysed in this chapter, among the most critical factors for improving DI performance, has been undoubtedly the choice of DSFs including feature normalisation and dimensionality reduction. With regards to the latter, it has been demonstrated that non-robust estimators like PCA are susceptible to outlier influence, i.e. masking and swamping effects. Such outliers may be samples representing the undamaged state at $T \leq 0$, where significant EOV influence on the system dynamics has been observed. These outliers will mask samples representing damage, since the latter are generally of weaker amplitude;

- With regards to the WT-REB dataset, although the choice of DSFs considerably improved DI performance, there have been two main challenges that needed to be overcome: 1. the under-representation of the undamaged state, and 2. the four different loading conditions that had a profound influence on the signals. These challenges, were mostly overcome, using a high-breakdown estimator like the DetMM with $\alpha \approx 0.2$. As it was shown, identification of damage level 1 was possible, something that has not been achieved using the benchmark EA method. At the same time, from

damage level 2 and onwards the outliers are all correctly identified, with minimal influence from the loading condition. In general, DetMM with $\alpha < 0.5$ may be used in such cases where the outlier fraction is greater than the inlier one;

- For robust estimators of scale and location, the parameter value α has been shown to be particularly influential for their definition of the correct inlier region;
- Using a robust subspace method, like rPCA, it has been possible to isolate damage on a few principal components;
- The computation of SMDs using a kernel matrix, as it is done with the kMRCD method, has been proven particularly helpful in cases where the decision boundary couldn't be established by a single ellipsoidal shape;
- A threshold value on χ^2 distribution with a confidence interval equal to the α value has been shown to be particularly suitable. Also, an MC threshold based on α was shown to be appropriate in most of the datasets;
- The fitted NLTS models considered were mostly simple quadratic regression models, where each one was fitted on a single rPCA component. Damage was then identified on the component that showed most sensitivity to damage (and immunity to EOV influence) - exceeding the 97.5% confidence interval of the fitted line;
- The tclust algorithm has been shown to be suitable for DI in problems where the dimensionality of the dataset exceeds two features. Additionally, when $K > 1$ suitable clusters may also be determined from the samples.

Chapter 6

Novelty detection and robust pre-processing

6.1 Overview of novelty detection and its applications for DI

In the previous chapter (Chapter 5), the identification of outliers (or damage) has been performed inclusively, i.e. as samples that deviate from the majority. This led to the unsupervised learning framework for DI, whereby there has been no pre-defined undamaged class samples.

In this chapter, the problem of DI is conducted as a semi-supervised learning framework, where samples are available (and known) a priori, from a system in its undamaged state. The task for this kind of problems, is essentially the estimation of the data distribution through the extracted features for that class only. Subsequently, this leads to the formation of the decision boundary. Samples that lie outside the decision boundary are outliers: ideally those will be the damaged samples only.

The field of ND [145] is well-studied and has been applied in SHM widely using the SPR paradigm [67], as outlined in Chapter 1. In ND, a statistical model is developed from samples obtained when the system is in its "normal" or "undamaged" state. This is also called the model of normality, and in order to develop such model the undamaged state may be well-defined in the dataset, i.e. samples must be in abundance (see Appendix C). In contrast, samples from the damaged state are generally rare, and in the SHM & CM contexts are typically not available: hence the importance of ND. It should be noted, however, that fitting a specific kind of model assumption to the undamaged samples, e.g. a Gaussian distribution, may be proven problematic / unrealistic [169].

In the literature of SHM & CM there have been numerous examples of utilising a wide range of methods for ND. A basic overview of ND and some details regarding the methods commonly used is provided in Appendix C. For instance [188], the authors used transmissibility ratios from acceleration responses as features to detect damage. For modelling the

undamaged system state, the authors used KDE and auto-associate Neural Networks (for comparison). In [188] and [170], extreme value statistics were employed to model the tails using one of the three theoretical distributions. Instead of assuming the whole undamaged state samples arise from a Gaussian, it has been shown that modelling only the tails provided much more accurate thresholds. This methodology was employed on a system that is initially linear in its undamaged state and transitions to non-linear in this damaged state, at a given operational and environmental load condition. More recently, in [83] the authors presented a neural network-based model, which was trained to estimate the relationship between system NFs from the undamaged state and air temperature variation. The latter, were the inputs to the model. This provided a way to mitigate the influence of EOV on the vibration responses. On the other hand, in [73] assumed no such knowledge/information from EOV. A model of normality was developed using only the NFs from the Z24 dataset when the bridge was in its undamaged state. The model of normality was a mixture of Gaussian distributions, whereby it provided the clusters for each undamaged system state (high EOV influence, low EOV influence, etc.). For DI the SMDs of a test set were computed utilising the developed clusters.
Apart from SHM studies, the ND framework has been used for a wide range of other problems of engineering nature, for instance, in CM for GT engines. In [86], the authors used data from vibrations to train an ocSVM model using the vibration tracked orders. Also, in [46] the ocSVM model was used to detect the impending combustion instability in an industrial combustor using pressure measurements and high-speed imaging as input DSFs. These methodologies were extended in a later paper by the same author in [47] to calibrate the novelty scores of the ocSVM into conditional probabilities.

As it was observed in the preceding chapters of this thesis, EOV influence, e.g. the influence of NFs from air temperature changes in the Z24 dataset, poses challenges for reliable DI. As it has been proven already, a suitable set of DSFs, e.g. impact-based DSFs, can provide a significant increase in reliability for DI. This is due to the fact that the monitoring system becomes less influenced by EOV influence, while it is more sensitive to the damage type. However, the wide applicability and computation of certain DSFs, like NFs, makes it necessary to seek ways to improve DI reliability in different ways. As it was mentioned previously in one of the papers [83]: to model the NFs responses as a function of EOV changes. In some cases, this may be impractical since it will require a range of sensors to be installed to measure several (not only temperature) sources of possible EOV influence. Also, this requires long-term monitoring of a system, to collect significant trends. In this chapter, a different approach is proposed that utilise robust statistical methods as pre-processors to

the ND problem.
In summary, the main task in ND is to infer the model of normality $\mathcal{L}(\boldsymbol{\theta})$ using available samples from the undamaged state, i.e. $[\mathbf{x}_n \in \mathbb{R}^p; n = 1, ..., N_{train}]$, where $\boldsymbol{\theta}$ are its hyper-parameters. Novelty (or outlier) scores $z(\mathbf{x}_*)$ are then computed on test data instances $[\mathbf{x}_i^* \in \mathbb{R}p; i = 1, ..., N_{test}]$. Note that n is the segment number as it was presented in Chapter 4.

6.2 Robust pre-processing

This chapter proposes to apply two of the methods that were used in Chapter 5: tclust and NLTS, for the purpose of robust pre-processing samples from the undamaged system state. In particular, the main objectives of this chapter can be summarised in the following three points:

- *Reduce the effects of EOV* on the derived DSFs that represent the undamaged state, by removing any points that are outlying in the training set. In that way, it is expected that the "clean" training set will represent the undamaged state more effectively than the one contaminated with outliers / EOV influence;

- In a multivariate scenario, both tclust and NLTS, *avoid the use of the single (elliptical) distribution assumption* as other methods do, e.g. FMCD and OGK. This is beneficial since it enables the definition of a multimodal decision boundary for the training set. For tclust, a number of K clusters will be defined and for NLTS the outlier removal is done on each pth dimension separately. Note that the elliptical distribution, albeit a *multi-modal* one, for the robust pre-processed samples in the training set still holds true;

- As the *definition of outliers is different for NLTS and tclust*, the performance of using these two methods to obtain a training set of observations that is "outlier-free" or "free from EOV influence" will also be assessed.

Specifically, both ocSVM (with RBF kernel) and iForest algorithms will be trained from samples that are found to be inliers according to either tclust or NLTS. Each combination of methods will then be compared to the case when all available samples (i.e. including EOV influence) are used for training. Additionally, a Gaussian Mixture Model (GMM) is used as an alternative method, providing also a probabilistic treatment of the ND problem. A GMM has been used widely for ND tasks [145], and will also be used alongside tclust and NLTS. Some basic theory of the GMM is provided below.
The primary way to examine the feasibility of robust pre-processing techniques, in terms

of improving ND, is through the use of A_{ROC} metric. This is due to the fact that this metric is not dependent on a fixed threshold, which means that an overall assessment of the separability (obtained from each method) between undamaged and damaged health states can be carried out. A percentage change of the A_{ROC} curve from the reference case study, i.e. the one with no pre-processing applied, is mainly used to test this feasibility.

6.3 Novelty detection algorithms

6.3.1 Gaussian Mixture Model

A Gaussian Mixture Model (GMM) belongs to the probabilistic approaches for ND (see Appendix C for the classification of methods). Unlike methods such as KDE, a GMM is parametric, in the sense that it assumes that the model of normality, is generated by a mixture of Gaussian distributions. This Gaussian mixture has a set of parameters $\boldsymbol{\theta} = \theta_1, ..., \theta_K$, where $\theta_k = [\boldsymbol{\mu}_k, \Sigma_k]$, i.e. the kth multivariate mean and covariance, that need to be estimated from training samples. Note that K is the total number of Gaussian distributions in the mixture.

More formally, given a set of p-dimensional samples $\mathbf{x}_i$, a GMM is defined as a linear combination of the mixture component weights π_k ($\sum_{k=1}^{K} \pi_k = 1$) and a p-variate Gaussian probability density function,

$$p(\mathbf{x}_i) = \sum_{k=1}^{K} \pi_k p(\mathbf{x}_i \,|\theta_k) \tag{6.1}$$

Assuming i.i.d conditions hold, the likelihood function of all $i = 1, ..., I$ observations is,

$$p(\mathbf{x}) = \prod_{i=1}^{I} \sum_{k=1}^{K} \pi_k p(\mathbf{x}_i \,|\theta_k) \tag{6.2}$$

To solve Equation 6.2 the EM algorithm is used [30]. In short, the EM algorithm is an iterative procedure for computing, relatively efficiently, the MLE for $\boldsymbol{\theta}$, given a known number of K Gaussian densities.

Following the EM procedure, the optimal likelihoods of each kth Gaussian density, given all the N_{train} samples in $\mathbf{x}_n$, are obtained. Hence, given all K likelihoods and K weights π_k, the unconditional probability of observing a data point drawn from the testing set $\mathbf{x}_i^*$ ($i = 1, ..., N_{test}$), is given using Equation 6.2 (where $\boldsymbol{\theta}$ are now the optimised ones). Note that $p(\mathbf{x})$ is a density, which may not have a direct probabilistic interpretation for setting up a threshold. For this reason, a threshold on $\int p(\mathbf{x}) d\mathbf{x}$ (its CDF), is sometimes applied [145].

For robust pre-processing, the number of Gaussian mixtures K in GMM will be selected to be equal to the number of clusters in tclust. Additionally, the "full" covariance structure

parametrisation will be used for each Gaussian mixture, unless otherwise specified. This will allow more flexibility, e.g. determine interdependencies between each feature pair, albeit having more parameters $\boldsymbol{\theta}$ to optimise.

6.3.2 Isolation Forest

The iForest algorithm that was presented in Chapter 5, worked for unsupervised learning tasks (see Appendix B for further details on the iForest method). In ND, however, its implementation will differ, given the fact that there is a separate training and testing phase. Therefore, in iForest the following methodology is proposed to identify outliers:

- Given N_{train} samples from the training set (undamaged state only), the same procedure that was used in Chapter 5 can also be used in ND to compute the expected path length $\mathbb{E}\left[\hat{H}(\tilde{\mathbf{x}})\right]$ from an ensemble of $D \approx 1000$ iTrees;
- Based on this expected path length, outlier scores from the training set can be calculated, as in Equation B.8;
- Similarly, given N_{test} samples from the testing set, a new expected path length is calculated and outlier scores are, once more, obtained using Equation B.8;
- The *decision* as to whether a sample from the testing set is an outlier is based on the *outlier scores computed from the training set*. That is, given undamaged state samples in the training set, a threshold (e.g. 99%) can be set on the *maximum outlier score*. However, this approach usually results in a high rate of FN, so that a threshold can be set on the outlier scores, computed from the training set, which are three scaled median absolute deviations (MAD) from their median value. Any outlier scores computed from the testing set that exceed this threshold are classified as outliers;

6.3.3 One-class Support Vector Machines

As discussed in Chapter 5, ocSVM determines the maximum margin from the origin. The origin in this case is where the outliers lie, whereas the rest of the feature space belongs to the inliers. Using the decision function in Equation B.7 an unseen / test sample $\mathbf{x}_*$ may be classified as inlier or outlier, according to the *decision boundary constructed using the training set*. This decision boundary is defined by the training data points that are on the outer most part in the features space, i.e. they become the training data support or *support vectors*. In the current implementation, when an outlier score for a given sample is negative, this sample will be considered as an outlier, and vice versa. So that, a threshold value of

0 is applied on the outlier scores of the ocSVM to distinguish between inliers and outliers. Additionally, since the fraction of damaged state samples are zero in the training set, the bias term ρ is not shifted, as it was done in inclusive OD in Chapter 5. Note also that, as it was done in the previous chapter, the parameter γ_{rbf} is cross-validated, in terms of A_{ROC} curve, using $15-20\%$ of the training data.

6.4 Results & discussion

In this section, the results of performing ND for DI are presented and discussed. Specifically, three datasets have been used: Z24, LA-4DOF and NSim-4DOFNonLin datasets, using a set of DSFs that both damage and EOV influence were apparent. For the Z24, its first four NFs are used, while for the other two datasets PSD amplitudes were extracted as DSFs. These PSD amplitudes were in frequency ranges that were affected by EOV and damage.

6.4.1 Z24 dataset

The first dataset that will be used to implement the proposed ND strategy is the Z24 dataset, as was presented in the previous chapters. To give equal weighting to each feature in the DSFs matrix (the first four NFs) of the Z24 dataset, each one was normalised according to the robust z-score method (presented in Chapter 4). The normalised DSFs for this problem are shown in Figure 6.1. Also, in the same figure, the training and testing sets are labelled as follows: $\mathbf{x}_{trn}$ is used to compute the model parameters and $\mathbf{x}_{tst}$ is used to test the classification performance, in terms of A_{ROC} metric. The results from the three ND methods using this dataset are: 85% for GMM, 90% for ocSVM and 91% for iForest. In Appendix D, the outlier scores of these three methods are shown in Figure D.7.

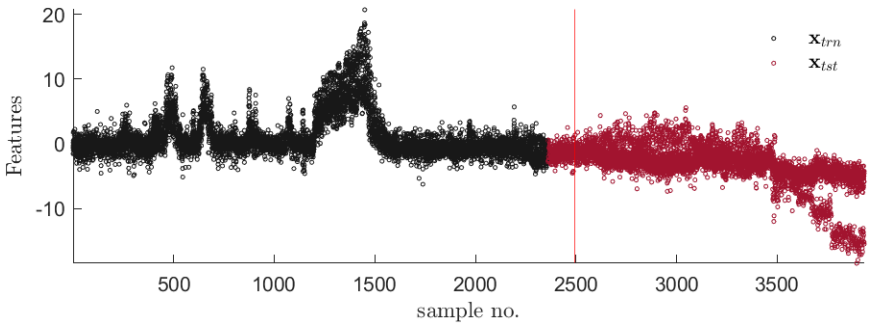

Figure 6.1. Z24: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively. Vertical line is the first sample of the damage state.

In order to visualise the workings of these methods, two features were purposely selected from the DSFs matrix: the first and third NFs. Applying NLTS and tclust with $\alpha = 0.85$, resulted in the A_{ROC} results shown in Figure 6.2. We can observe that all three ND

techniques can generally improve A_{ROC} value for ocSVM as compared to the case where no robust pre-processing (either NLTS or tclust) was applied, i.e. $\alpha = 1$. Whereas, for iForest and GMM the performance remains almost unchanged. Note that, with regards to the NLTS, a different threshold was applied to the current default one. In the current default implementation, the threshold is the confidence interval of the predicted fitted line (see Chapter 5) applied on each of the p features individually (as you will expect from a univariate method). However, it is necessary to maintain equal number of dimensions in the DSFs matrix. Therefore, an outlier in one of the dimensions automatically becomes an outlier in all dimensions, so that all p features share the same outliers (even if in one of those is an inlier). For that reason, however, this method tends to remove a large proportion of samples from the training set, and is sometimes unsuitable for problems with a small number of training samples.

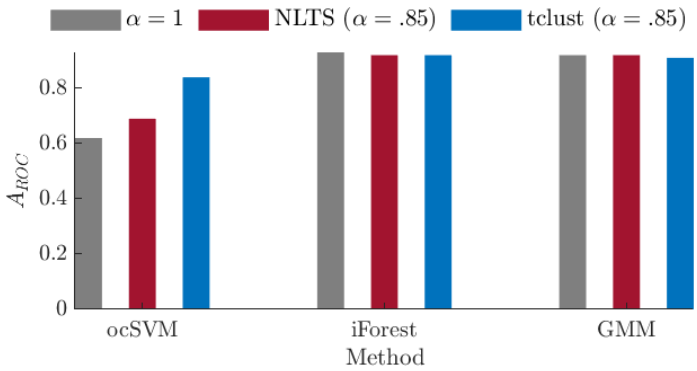

Figure 6.2. Z24: A_{ROC} metric for the case where robust pre-processing is applied with NLTS and tclust with $\alpha = 0.85$, and compared to the case where no robust pre-processing is applied $\alpha = 1$.

The corresponding results of applying tclust with $\alpha = 0.85$ on the first and third NFs of the Z24 dataset are shown in Figure 6.3. As can be observed, most of the training samples that tclust has removed $\hat{\mathbf{x}}_{o,trn}$, are samples with significant EOV influence, i.e. when $T \leq 0$. This result is also shown on a 2D plot in Figure 6.4. However, it was not the removal of samples from EOV influence that increased the value of A_{ROC}, but, rather the removal of samples around the undamaged state at $T > 0$. In particular, the samples from the undamaged ($T > 0$) that are overlapping with the samples from the damaged state. It can be shown that this, effectively, yields a more "concentrated" training set $\mathbf{x}_{trn}$ (belonging to the undamaged state), which is less overlapped with the damaged state. Note that, we assumed knowledge of the number of clusters for the tclust method, which implies that the number of clusters selected for GMM is also 2. Also, the support vectors (SVs) of the ocSVM method are generally selected to be sparse, so that $\nu = 0.01$ for this problem. This

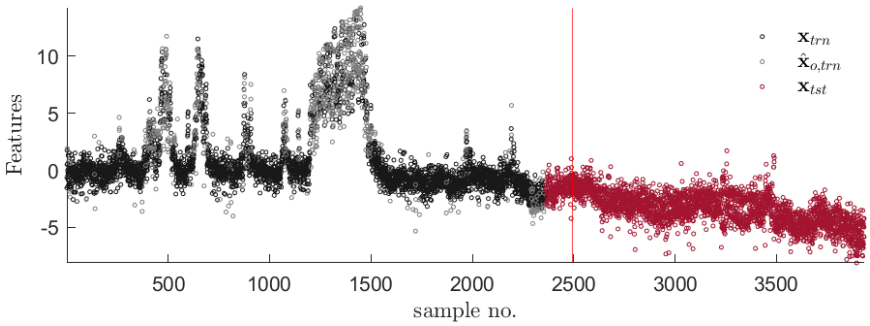

Figure 6.3. Z24: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively (first and third NFs). Vertical line is the first sample of the damage state. Outliers removed from the training set $\hat{\mathbf{x}}_{o,trn}$ using tclust are also shown.

is required for preventing overfitting problems, typically associated with drawing complex decision boundaries in the context of ocSVM.

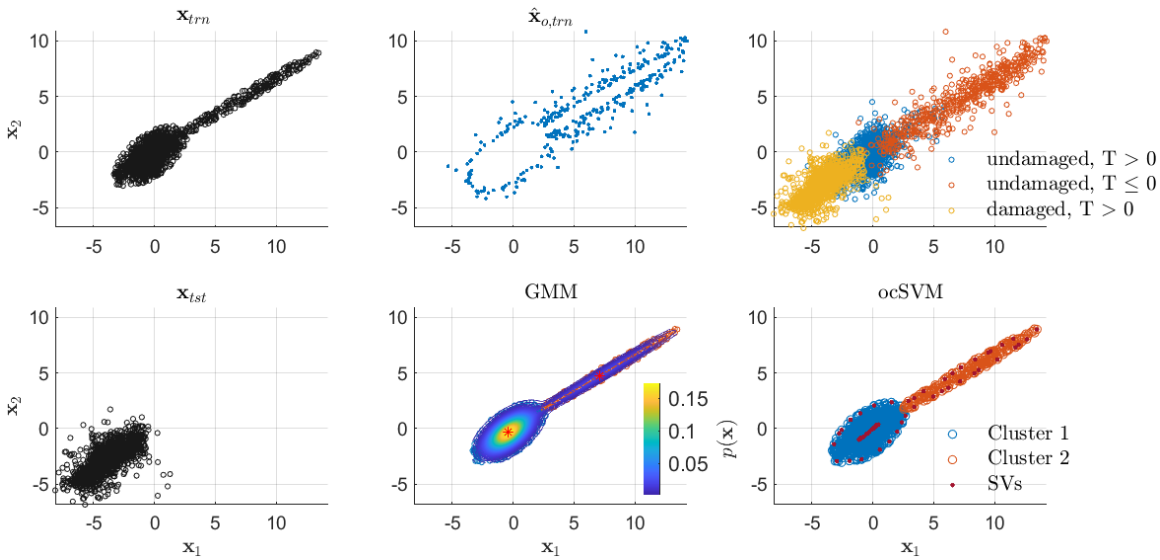

Figure 6.4. Z24: The corresponding 2D results shown in Figure 6.3 analytically.

The main parameters of this study are also shown in Table 6.1.

6.4.2 LA-4DOF dataset

In this dataset, a time interval of 4 seconds is utilised to segment each acceleration time series. In each of the 1224 segments extracted, 84 amplitudes from the PSD between $25-130$ Hz were taken as DSFs. Within this region, as it was shown in Chapter 3, both damage and EOV influence are apparent. This resulted in a 1224×84 DSFs matrix, where 45% of those were used for training purposes, since half of the total number of samples (612)

Table 6.1. Parameters for ND in Z24 dataset.

Parameter	*Value/description*
Inlier fraction, α	0.85
DSFs	robust z-scores of 4 NFs
NLTS parameters	$A = 2$ (the rest are equal to 0)
tclust number of clusters, K	2
ocSVM parameters, ν and γ_{rbf}	$\nu = 0.01$, γ_{rbf}: computed using holdout
$\mathbf{x}_{trn}$ ratio	0.6
GMM number of clusters, K	2

belong to the damaged state*. Note that, methods such as GMM, with many parameters to optimise, especially when $K > 1$, require a large number of samples for training [145]. As it was found out, GMM couldn't be used in this dataset since the number of samples available for training was small, with respect to the DSFs matrix dimensionality. The main parameters used for this study are shown in Table 6.2.

Table 6.2. Parameters for ND in LA-4DOF dataset.

Parameter	*Value/description*
Inlier fraction, α	0.9
DSFs	z-scores of 84 PSD amplitudes: $25-130\ Hz$
NLTS parameters	$A = 2$ (the rest are equal to 0)
tclust number of clusters, K	2
ocSVM parameters, ν and γ_{rbf}	$\nu = 0.05$, γ_{rbf}: computed using holdout
$\mathbf{x}_{trn}$ ratio	0.45

One of the most critical factors in achieving high accuracy in this dataset is the use of an appropriate feature normalisation method. As it has been discussed in Chapter 4, scaling the testing set samples with respect to the training set's standard deviation and mean, is a common methodology in ND tasks. This has increased the performance (A_{ROC} value) of around .25 to .81 for ocSVM. On the other hand, the corresponding performance of iForest remains invariant to the choice of feature normalisation, i.e. it has remained .89.
The results from robust pre-processing using NLTS resulted in a significant rejection of training samples (i.e. the outliers), which has impacted negatively the performance of

*A further decrease in the segment time interval, e.g. 0.5 seconds, would have resulted in a much higher number of samples, but, at the expense of very poor frequency resolution.

both ocSVM and iForest. Therefore, only the tclust algorithm was used for robust pre-processing with $\alpha = 0.9$. This α value was chosen arbitrarily small, to avoid removing a large percentage of the training samples. The DSFs after tclust was applied on the training samples are shown in Figure 6.5. The A_{ROC} values after applying tclust on the training data are: .80 for ocSVM and .90 for iForest. Therefore, some marginal improvement for iForest, while for ocSVM the performance decreased slightly.

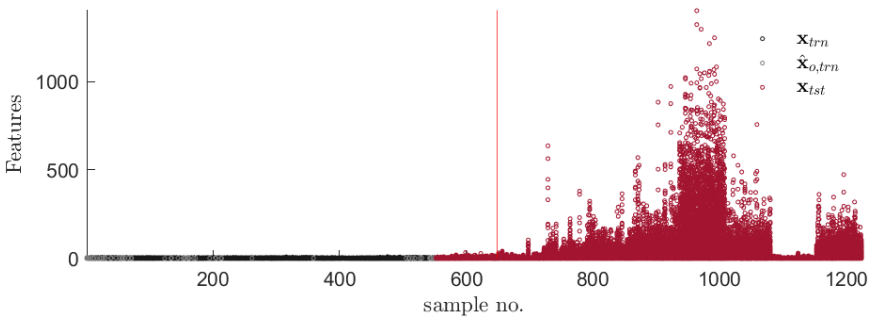

Figure 6.5. LA-4DOF: Training and testing sets, $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$, respectively (84 PSD amplitudes)). Vertical line is the first sample of the damage state. Outliers removed from the training set $\hat{\mathbf{x}}_{o,trn}$ using tclust are also shown.

The outlier scores of these two methods are also shown on the top of Figure 6.6, while on the bottom are the histograms of the training and testing sets, illustrating the placement of the threshold line. In both of these methods the outlier scores are high (in absolute terms) for the majority of the samples in the testing set $\mathbf{x}_{tst}$. These are samples that were predicted as the outliers $\hat{\mathbf{x}}_o$ (red dots in the outlier score plots). For ocSVM, everything below 0 is considered an outlier, while for iForest this threshold is set according based on the deviation from the median of the threshold computed on the training set (as mentioned above).

6.4.3 NSim-4DOFNonLin dataset

In this dataset, the number of segments was increased to 1391, where the time interval for each segment is now 2.5 hours. As in LA-4DOF the amplitudes from the PSD were extracted. In this dataset, however, the amplitudes from the whole Nyquist frequency range were obtained, which produced a DSFs matrix of 1391×230. For visualisation purposes, PCA was applied and the first two principal components were extracted as features. It has been observed that, when tclust is applied on this two-dimensional dataset an improvement in A_{ROC} performance could be achieved. The results by using two different values of α for tclust are shown in Figure 6.7. In comparison to the case where no robust pre-processing was applied ($\alpha = 1$), A_{ROC} was increased for iForest and GMM when $\alpha = 0.9$, but, decreased

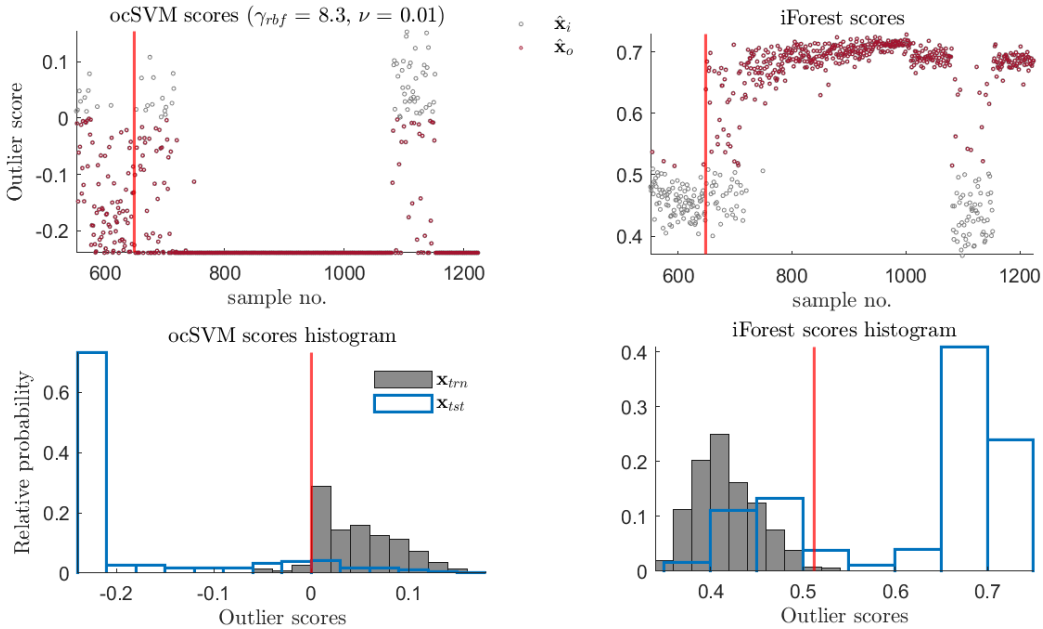

Figure 6.6. LA-4DOF: on the top are the outlier scores computed by ocSVM and iForest algorithms and on the bottom the corresponding histograms on $\mathbf{x}_{trn}$ and $\mathbf{x}_{tst}$ samples, showing also the thresholds applied. Red dots $\hat{\mathbf{x}}_o$ are the estimated outliers for each method.

when $\alpha = 0.75$. However, the opposite happened for ocSVM.

From Figure 6.8, the effect of increasing the value of α on the dataset is clearly shown. At $\alpha = 0.75$ value the samples that represent the undamaged state when $T \leq 0$ are mostly eliminated from the training set $\mathbf{x}_{trn}$. On the other hand, when $\alpha = 0.9$ only the samples from the same state that are furthest away from the cluster are removed. In this case, tclust with $K = 1$ considers a single cluster of data points to be the undamaged state at $T > 0$ and all the rest are outliers.

Note that, both ocSVM and iForest could result in very high A_{ROC} values using the uncompressed set of DSFs, i.e. all 230 PSD amplitudes. Note that there is generally not any indication of damage by looking at the DSFs, as shown in Figure 6.9. However, after damage initiation the outlier scores computed from iForest increase significantly, whereas, they quickly fall below zero for ocSVM. Note that some marginal improvement in performance of less than 1% could also be achieved when tclust is applied as robust pre-processor to the whole (230) PSD features space.

6.5 Summary

To summarize, in this chapter robust pre-processing, was used as a way to define a training set that is free from EOV influence. This "uncontaminated" training set was expected to

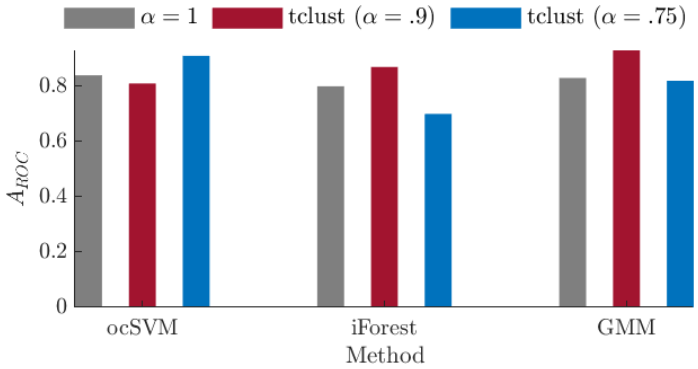

Figure 6.7. NSim-4DOFNonLin: A_{ROC} metric for the case where robust pre-processing is applied with tclust on $\alpha = 0.75$ and $\alpha = 0.9$, and compared to the case where no robust pre-processing is applied $\alpha = 1$.

improve the performance of three ND methods: ocSVM, iForest and GMM. Two robust statistical methods were used: the robust clustering algorithm tclust and the non-linear version of least trimmed squares NLTS.

Using three different datasets, the following conclusions can be drawn from this analysis:

- When training samples, i.e. from the undamaged state, are overlapping with samples from damaged state (in the test set), robust pre-processing by means of either tclust or NLTS can provide a better separation between them. This was observed when significant EOV influence was present;
- The amount of outlier removal from the training set, as dictated by the specification of the α value must not be too excessive so as to cause performance reduction, as observed by decreasing α from 0.9 to 0.75;
- In particular, tclust will remove scattered samples in the feature space that are mostly samples representing significant EOV influence, which is of particular benefit to SHM studies;
- Note that, for each ND method (ocSVM, iForest and GMM) the performance varied with respect to α value chosen, for robust pre-processing;
- Finally, it was observed that tclust requires significant computational resources when the dimensionality of the problem is high, e.g. greater than a 100, which limits its applicability slightly.

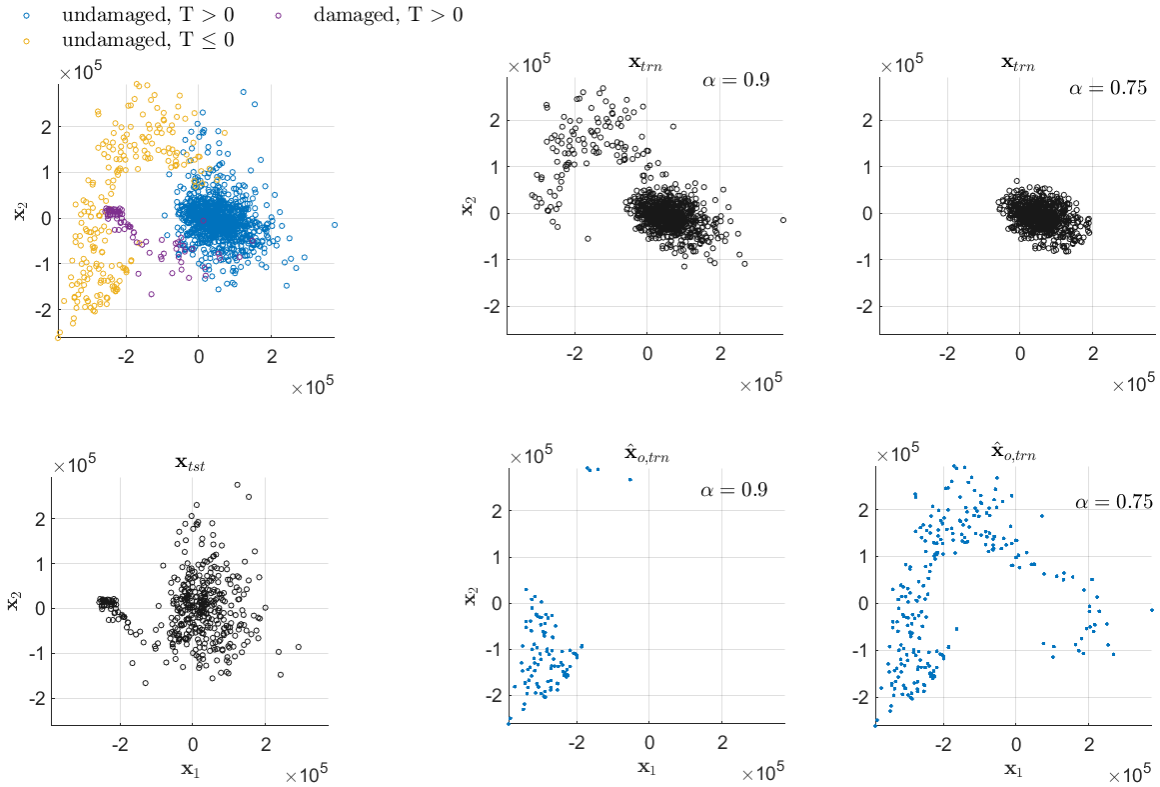

Figure 6.8. NSim-4DOFNonLin: Illustration of the outliers removed using two α values: 0.75 and 0.9.

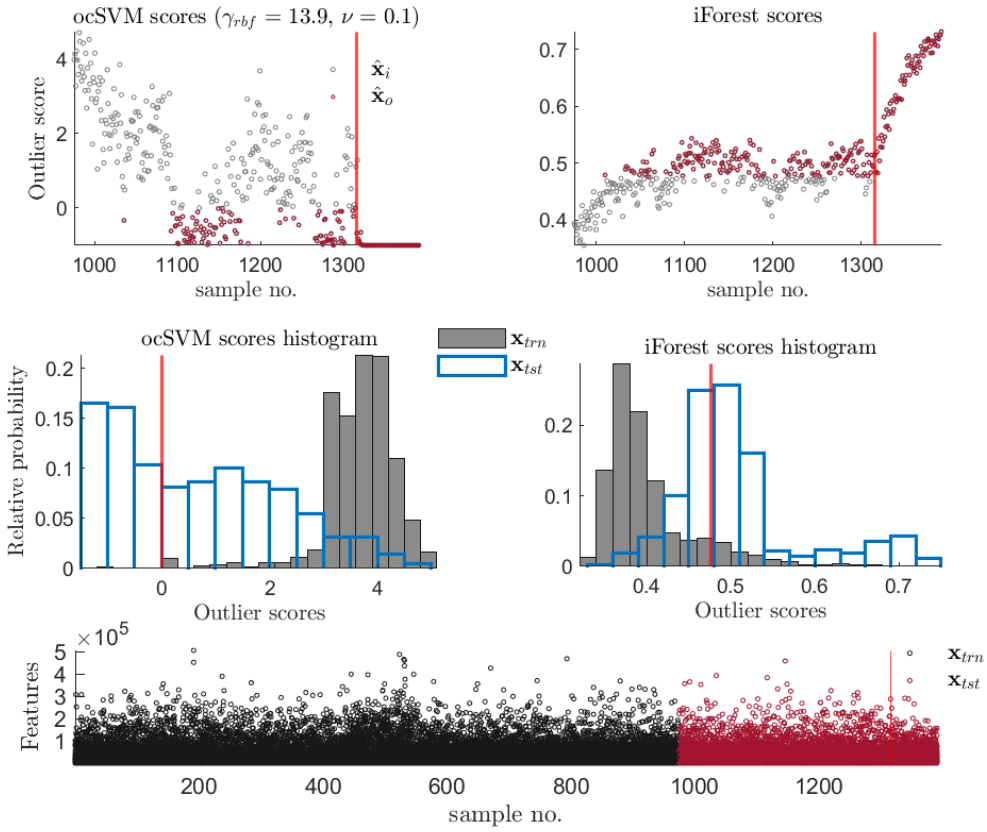

Figure 6.9. NSim-4DOFNonLin: on the top two plots are the outlier scores computed using ocSVM and iForest, while in the middle are the corresponding histograms. On the bottom are the 230 PSD amplitudes used as DSFs.

Chapter 7

Conclusions

This thesis employed a data-driven approach to the problem of *damage identification* on a number of structural health and condition monitoring applications, each one having a distinct set of characteristics. The main focus of this work is on key areas related to the development of monitoring frameworks that maximise *reliability* and *accuracy*.
First and foremost, data representations or *features*, were calculated from large volumes of experimental observations. These features have attractive properties for data-driven structural health and condition monitoring applications, e.g. immunity of environmental influence on system dynamics. A number of features were proposed in Chapter 4, and explored for outlier and novelty detection tasks in Chapters 5 and 6. Secondly, the particular difficulty and scarcity in obtaining labelled sets of data, i.e. the ground truth, for conducting a novelty detection task for damage identification, has been approached by an *inclusive outlier detection framework*. This has been done with the aid of robust statistical techniques, which some of them were used to improve novelty detection tasks. This has been done by defining a training set that is free from outlying samples, which includes environmental influence.

7.1 Vibration-based features for SHM and CM

In the first part of the thesis, in Chapters 2 and 3, certain data attributes were computed from large volumes of vibration time-series measurements. These attributes, called *damage sensitive features* (DSFs), were classified into: global-based, modal-based and impact-based DSFs. The latter, being the most "optimal" one for detecting the *specific damage mode* on three different datasets: rolling element bearing (WT-REB) outer race fault, four-degree of freedom numerical simulation (NSim-4DOF) and four-degree of freedom Los Alamos benchmark study (LA-4DOF). On the other hand, global-based DSFs measure general system attributes as: *general changes* in the vibration signal characteristics. These sets of

DSFs are typically necessary when no system or damage information is available. On the other hand, modal-based DSFs are used for *benchmarking*, as they are the most common attributes used in SHM applications.

Impact-based DSFs were, mainly, calculated as the kurtosis values (relative entropies were also explored) of a small subset of decomposed intrinsic mode functions (IMFs). The IMFs that were selected were those with the highest kurtosis and root-mean-square values. These IMFs were decomposed using six different adaptive decomposition methods, for comparison purposes. The final part, after obtaining the kurtosis values of a small segment (a few seconds of < 10), was to apply a filtering algorithm, i.e. the minimum entropy deconvolution (MED), designed to *enhance signal impulses*. A particular combination stood-out as the one that improved the identification of damage: the kurtosis values of a subset of IMFs decomposed using Successive Variational Mode Decomposition (SVMD). Note that smoothing-out (using averaging filters) the MED filtered DSFs was necessary due to the sharp spikes that resulted from its application. The final result was a set of DSFs that was highly damage sensitive, for instance, could detect an early outer race fault, something that could not be achieved using other well-known techniques, e.g. envelope analysis (presented in Chapter 2). The fact that these sets of DSFs could also be applied to a *variety of problems*, in which damage manifests itself in the form of impacts, has also been demonstrated on both the NSim-4DOFNonLin and LA-4DOF datasets.

Global-based DSFs have been used for analysing vibration signals in gas turbine engine experimental datasets. A particular subset of these DSFs, which are *kurtosis-derived* ones, enabled the early detection of engine conditions changes (air-to-fuel ratios and fuel mixtures).

Modal-based DSFs (PSD amplitudes), on the other hand, were useful in identifying damage early, while minimizing influence from changes in environmental and operational conditions. This is true, when, for instance, in the LA-4DOF dataset, a *high-pass filter* was applied in order to include only those power spectrum amplitudes on frequencies $> 100\ Hz$. This means that natural frequencies, which are sensitive to *both* damage and environmental changes, are excluded from the analysis. For modal-based DSFs, a system identification procedure was also investigated, in order to estimate *damping ratios* and *natural frequencies* using decomposed IMFs from adaptive decomposition techniques. In general, these estimates were both unsatisfactory and relatively insensitive to damage, and so, these particular sets of DSFs were not used for further analysis.

In general, impact-based DSFs have very low dimensionality (< 5), which makes them a much more attractive choice, than, for instance, power spectrum amplitudes (> 100). For

the latter, dimensionality reduction needs to be performed prior to statistical model development, e.g. in Gaussian Mixture Models when the number of samples is comparable to the number of dimensions.

7.2 Robust statistics for damage identification

In the first part of exploring damage identification techniques, inclusive outlier analysis was presented in Chapter 5. Given the particular difficulty in obtaining labelled datasets in systems that operate in harsh environments, e.g. gas turbine engines and offshore wind power plants, conducting inclusive outlier analysis is the only way towards damage identification. Some of the findings from this methodology, as presented in Chapter 5 are discussed below. One of the most important aspects of inclusive outlier analysis, was the derivation of a suitable set of DSFs, which can target the specific type of damage for a given problem (i.e. impacts in the case studies analysed). This has been proven to be critical for reliable and robust damage identification.

Apart from the choice of DSFs, in the WT-REB outer race fault dataset, a high-breakdown robust estimator (DetMM), demonstrated excellent performance in outlier detection. Specifically, as the fraction of damage observations is higher than the undamaged ones in this dataset, DetMM estimates of robust covariance and multivariate mean were challenging to obtain. However, by specifying the fraction of inliers around 0.2 in DetMM, it has been possible to detect the *earliest possible damage level*. While, for the next levels of damage progression, the influence of the operational loads was, mainly, negligible. Such results, could not be achieved using other techniques, e.g. the benchmark envelope analysis method. Using a robust version of PCA, i.e. rPCA, it has been demonstrated that damage may be isolated from EOV influence in a few principal components. This is an important result for SHM purposes: it shows that when DSFs are influenced by EOV, such as modal-based DSFs, a technique like rPCA may be used to improve damage identification. As it was also demonstrated, in PCA it was not possible to isolate EOV influence or damage in a few principal components, as it was done in rPCA. This is because when significant EOV influence is present, PCA will have its principal components in the direction of these samples, effectively masking samples from damage. This effect was also demonstrated in Chapter 5. Among the methodologies used, kernel minimum regularised covariance determinant (kMRCD), which is the latest advancement of the well-known minimum covariance determinant (MCD) method, offered consistently good detectability of outliers. This is true given a radial basis function (RBF) kernel is specified.

It should also be noted that, the fraction of outliers specified for the robust statistical techniques, is of particularly critical importance. Although in most practical cases this will

certainly be unknown (i.e. the fraction of damaged samples expected), it is anticipated that the analyst will need to experiment with this parameter in order to better explore the datasets and draw conclusions from them.

7.3 Diagnosis of outliers in SHM and CM

Apart from its robustness properties, as a dimensionality reduction technique, rPCA enabled the *diagnosis of outliers*, as demonstrated on several datasets in Chapter 5. This was made possible by constructing the outlier map of rPCA, in which the calculated score and orthogonal distances can be explored (on a 2-D plot). In terms of significant EOV influence on the system's dynamics, e.g. when temperature falls below zero degrees Celsius, outlying observations were seen as having large orthogonal distances for NSim-4DOFNonLin dataset. On the other hand, for the Z24 dataset, orthogonal outliers were samples from its damaged state. Generally, it is important to note that the choice of DSFs influenced the outliers classification. For instance, it has been observed in the LA-4DOF and TF-LBO that when damage (or significantly different condition in the engine's case) is very apparent from the dataset, it will be classified as a bad leverage point.

7.4 Enhancing novelty detection by outlier exclusion

In Chapter 6, two techniques that were presented in the previous chapter, are used as *"robust" pre-processors* on DSFs that show influence from EOV. These techniques are the NLTS and tclust, and are used as *outlier-removal tools* for obtaining an "outlier-free" training set using for novelty detection tasks. In particular, it was found that when a moderate amount of outliers are removed and when those outliers are overlapped with the samples related to damage, then identification performance increases, albeit by a small percentage. However, this performance change differs from the technique used for novelty detection and care should be taken to avoid excessive training samples removal, as this will degrade performance.

7.5 Future work

Some of the findings and assumptions made in this thesis, can be used to further guide future research into the development of damage identification frameworks. In particular, future work may focus on the following:

- The interaction between robust pre-processing methods and other novelty detection algorithms, including how their parameters (α, γ_{rbf}, etc.) can be simultaneously

optimised to achieve higher A_{ROC} values. On this point, more techniques should be tested, e.g. kMRCD as robust pre-processing;

- The outlier map of rPCA should be further utilised to understand the classification of condition related to EOV influence and damage. This implies that more datasets and DSFs should be examined in order to draw valid conclusions regarding the three possible outlier classifications: orthogonal outliers, bad or good leverage points. Ultimately, this should be used to further develop techniques to understand and mitigate EOV influence;

- The choice of intrinsic mode functions (IMFs) has been made according to the maximum root-mean-square and kurtosis values. From the analysis, it was found that the choice of IMFs (both the number and which ones to use) plays an important role in enhancing damage sensitivity, and at the same time EOV influence. Further work can be carried out for a more "informed" or "targetted" decision with regards to which IMFs are more important for the damage identification task;

- The application of the deconvolution filtering technique, in order to enhance signal impulsivity, requires the application of a smoothing filter as a final step for spike removal. In order to improve damage sensitivity, so that any change in the signal impulsivity is not smoothed out by the smoothing filtering operation, it may be necessary to employ alternatives for signal impulsivity enhancement. Although in this thesis, many other methods were tested to act as alternatives to MED, e.g. the maximum correlated kurtosis deconvolution, their performance was poor as compared to the MED. A range of other filtering algorithms do exist, albeit there is a need for setting their parameters correctly;

- As with the filtering procedures, many adaptive decomposition methods have particularly high sensitivity to the specification of their parameters. In particular, methods like the adaptive local iterative filtering (ALIF), are not straightforward to tune at all, and thus, most of the times some of the decomposed IMFs may be redundant (e.g. mode mixing issues) or being trivial in their values (i.e. close to zero). On the other hand, methods such as SVMD are more stable and can be tuned much more easily, i.e. their filtering operation is not as sensitive to their parameter(s). Hence, more work may be conducted on the outcome of deriving DSFs from methods like ALIF, in which their parameters have been optimised. This is a difficult endeavour, however, given that these algorithms have just been developed, there is little or no guidance as to how to achieve maximal performance and the fact that there might be

significant characteristic differences between time segments from a monitored system. The latter, implies that optimal parameters will need to be calculated for each time segment, which is both impractical and undesirable for monitoring applications, as it will possibly introduce significant time lag. Note that a time lag in the time-scale of seconds may not be critical, while a time lag in the time-scale of hours may have undesirable consequences;

- The unsatisfactory results from the attempt to estimate damping ratios (DRs) and natural frequencies (NFs) using adaptive decomposition methods, may be caused from a range of different factors. Apart from the parameters' tuning that was mentioned above, the differences in the number of IMFs, have proven to be of critical importance to the estimation of DRs and NFs. This can be shown by looking at the results in Chapter 4, in which only the VMD gave satisfactory estimates, since it allows the user to specify the number of modes. However, the high sensitivity of EOV influence, coupled with relatively low sensitivity, in these parameters makes them undesirable as DSFs. Although this is the case, in the NSim-4DOFNonLin dataset results the estimated DRs look promising enough using VMD, and thus, more research may be carried out to understand these sets of results resulting from this system identification exercise. Note that, an attempt was made to "optimise" the parameters of the Natural Excitation Technique (NExT) using general purpose optimisation algorithms, e.g. using particle swarm optimisation. However, as previously mentioned, given the potential significant characteristic differences between each time segment, this optimisation on one a single segment may yield unsatisfactory results. Therefore, new ways will need to be proposed towards better system identification using this combination of techniques;

- In this thesis, a comprehensive set of results was provided for inclusive outlier analysis, including five different robust multivariate location and scatter estimates, robust time-series analysis (via NLTS), robust clustering analysis (via tclust) and robust subspace analysis (via rPCA). In clustering analysis, one may provide a means to construct classifiers, i.e. by specifying the number of clusters as equal to two. In the ideal case, cluster 1 will be EOV influence, cluster 2 will be regular observations / undamaged system system and cluster 0 will be outliers / damage events. Therefore, having such groupings / labelling multi-class classification could be constructed and further analysis can be done on unseen data points to demonstrate the validity of such an approach for damage detection.

Appendix A

Numerical model for SHM applications

A.1 A four degree-of-freedom model

A four DOF numerical simulation model is used to investigate SHM methodologies. It represents a bookshelf-type structure, which is a very common setup in SHM studies. In a bookshelf-type structure, lateral excitation forces are acting on it, causing it to bend. Since the structure is in shear, the *shear building assumption* applies. The dynamics of such structural arrangement can be conveniently modelled as a lumped mass-spring-damper system. The main assumptions underlying the model are as follows:

1. No rotation of the floors is allowed. *
2. Floors are rigid and at 90 degree angle to the support columns. *
3. No mass contribution from support columns, i.e. total mass is concentrated at the levels of each floor.
4. The base is assumed to be mounted on ball bearings, and thus, it can have different stiffness and damping coefficient values from the rest of the floors.
5. Equivalent stiffness and damping coefficient values between each of the support columns.
6. Purely lateral excitation force is acting on the system that can be, for instance, Gaussian white noise $F_1 \sim \mathcal{N}(\mu_{F_1}, \sigma^2_{F_1})$. Hence, stationary random excitation input is assumed (assuming uniform excitation over the whole frequency spectrum).

The schematic diagram of the four DOF lumped mass-spring-damper numerical simulation model is shown in Figure A.1. Damage can be simulated as a stiffness reduction between any two given masses.

*Requirement for shear building model.

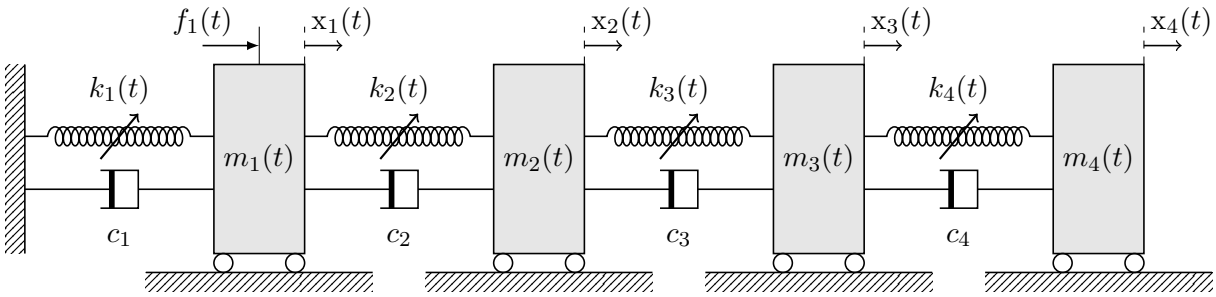

Figure A.1. The four DOF numerical model used for generating the simulated data under different damage scenarios and EOV influence.

In state-space form, the numerical model can be written as:

$$\begin{bmatrix} \ddot{\mathbf{x}} \\ \dot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} -M^{-1}C & -M^{-1}K \\ I & 0 \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{x} \end{bmatrix} + \begin{bmatrix} M^{-1}\mathbf{f} \\ 0 \end{bmatrix} \tag{A.1}$$

where, M, C, K, are the mass, damping and stiffness matrices, respectively, while $\mathbf{f} = \{f_1, 0, 0, 0\}$ is the force excitation vector. The composition of the state-space matrices allows the calculation of the natural frequencies $\boldsymbol{\omega_n}$ (and mode shapes) of the system. And, since damping is non-zero, the damping ratios $\boldsymbol{\zeta}$ and damped natural frequencies $\boldsymbol{\omega_d}$ can also be found through the geometric relationships in the complex plane. Analysis using this numerical simulation model can therefore be performed by using one or more of the system states, $\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}$ in time- and/or in frequency-domain or any of its modal properties.

A.2 The Duffing model

The expression in Equation A.2, can be viewed as a simple mass-spring-damping system excited by a sinusoidal forcing function and whose spring constant, i.e. $k + \beta x^2$, varies as a function of the displacement, x, of the mass,

$$\frac{d^2x}{dt^2} + c\frac{dx}{dt} + (k + \beta x(t)^2)x(t) = \gamma cos(\omega t + \phi) \tag{A.2}$$

where, γ is the amplitude of the forcing function, c is the energy dissipation constant, β and k are the restoring force coefficients to the equilibrium position, ω is the frequency of oscillation and ϕ the phase.

The above expression suggests that the system's frequency of oscillation will not be a constant one within a single period. In particular, it exhibits *amplitude-dependent* frequency of oscillation, since it depends on the input excitation force amplitude. This concept is also termed as intra-wave frequency modulation [91].

Note that, for a perfectly linear process, $\beta = 0$, whereas if $\beta \to 0$ perturbation methods can

be performed to linearise the process and provide a sufficient approximation to the physical characteristics of this system [184]. However, for a highly non-linear generating process, $\beta \neq 0$, and the system will exhibit non-linear characteristics (especially when $\beta \gg 0$), which in this case it will be in the form of *jump phenomena*. That is, as the forcing frequency ω varies, the system's amplitude response will reach its bifurcation points[†], and a sudden jump upwards or downwards in its value will take place.

In terms of signal processing, this non-linearity causes several challenges. Let's assume a steady-state periodic solution to the 2nd order differential equation, which can be expanded using a Fourier-series[‡] of the form:

$$x(t) = \sum_{i=1}^{\infty} \left[\tilde{X}_i cos(i\omega t) + \acute{X}_i sin(i\omega t) \right]; \tag{A.3}$$

Expanding the above up to the *first harmonic*, and considering only the cosine terms, it can be shown that a third-order harmonic, i.e. $cos(3\omega t)$, appears in the solution:

$$\tilde{X}_1 \omega^2 cos(\omega t) + \tilde{X}_1 csin(\omega t) + \tilde{X}_1 kcos(\omega t) + \tilde{X}_1^3 \beta \left[\frac{1}{4} \left\{ \frac{3}{4} cos(\omega t) + \mathbf{cos(3\omega t)} \right\} \right] = \gamma cos(\omega t + \phi) \tag{A.4}$$

That is, the cubic restoring force result in cosine waves at *integer multiples* of the input frequency ω to appear in the solution. This is referred to as *harmonic distortion* [81], and it is a well-known phenomenon in non-linear systems.

The higher-order harmonic terms appearing in the solution, in order to approximate the non-sinusoidal waveforms, do not provide a physical interpretation of the system. This is simply because such frequency components are not generated by the physical process, but, are rather a way to simulate the non-linearities present, which are only valid mathematically. Obtaining a FT of the above (e.g. PSD or FFT), will yield spectrum with the fundamental frequency at ω, and many multiples of it.

A.3 The ith modal vibration response from a multi-DOF linear system

Through modal expansion, a multi-DOF linear system, can be decomposed into M modes of vibration, where the orthogonality properties of the ith mode shape vector $\boldsymbol{\psi}_i$ are being utilised, for the purposes of deriving the decoupled i^{th} free acceleration response of the generalised modal coordinate $q_i(t)$:

$$q_i(t) = \frac{F_0 \psi_{pi} \omega_{n,i}}{m_i \sqrt{1-\zeta_i^2}} e^{-\zeta_i \omega_{m,i} t} cos\left(\omega_{d,i} t + \phi_i + \pi/2\right) \tag{A.5}$$

[†]The point in a bifurcation diagram where system switches from stable to unstable.

[‡]Admitting a solution of the form of Fourier series expansion allows any signal to be approximated using orthogonal basis functions and in an infinite number of ways.

where, F_0 is the excitation force applied on the pth DOF, ψ_{pi} is the pth element of $\boldsymbol{\psi}_i$, $\omega_{m,i}$ is the mth modal NF, $\omega_{d,i}$ is the modal damped frequency, ζ_i is the modal DR, ζ_i is the modal mass and θ_i is the phase lag of the ith mode of vibration. As can be seen from Equation A.5, each modal response is centred around a specific NF $\omega_{m,i}$.

Appendix B

Methods & Algorithms

B.1 One-class support vector machines

SVM are also called *maximum margin classifiers*, because their main purpose is to find a linear decision hyperplane that maximizes its *smallest perpendicular distance* to *any* of the observations in X, i.e. the margin. Support Vector Machines (SVM) were initially developed as a binary class classification method, offering the flexibility of an ANN, while overcoming its pitfalls. Using a kernel function to expand the original data space $\mathcal{X}$ into a higher dimensional feature space $\mathcal{F}$, is related to adding more layers to an ANN. Therefore, an algorithm can be adapted to match the characteristics of the data, as mentioned previously in kPCA and kMRCD methods. In ocSVM, a quadratic *convex* optimisation problem is solved, which guarantees a *global optimal solution* to the problem of placing a linear *decision hyperplane*, represented with the equation. An attempt will be made to explain the binary version of SVM, before, proceeding to the ocSVM objective.

As with linear discriminant models, a $p-1$ linear *decision hyperplane* of the form $y(\mathbf{x}_n) = \mathbf{w^T}\mathbf{x}_n + \beta$ can be contemplated, such that $t_n y(\mathbf{x}_n) \geq 0, \forall n = 1, ..., N$, with $t_n \in \{-1, 1\}$. Since, the perpendicular distance of a point $\mathbf{x}_n$ to the linear decision surface is $y(\mathbf{x}_n)/||w||$, there exists a combination of $\mathbf{w}$ and β in which the margin is maximized,

$$\arg\max_{\mathbf{w},\beta}\{||\mathbf{w}||^{-1}\min_{n}[t_n y(\mathbf{x}_n)]\}, \forall n = 1, ..., N \tag{B.1}$$

The optimisation argument in Equation B.1 can be simplified further by adopting a *normalised constraint*, $t_n y(\mathbf{x}_n) \geq 1, \forall n = 1, ..., N$, so that given two points as $y(\mathbf{x}^-) = -1$ and $y(\mathbf{x}^+) = 1$,

$$\begin{aligned} \arg\max_{\mathbf{w},\beta}\{||\mathbf{w}||^{-1}\mathbf{w}[\phi(\mathbf{x}^+) - \phi(\mathbf{x}^-)]\}, \forall n = 1, ..., N \\ \therefore \arg\max_{\mathbf{w},\beta} 0.5\{||\mathbf{w}||^{-1}\} \end{aligned} \tag{B.2}$$

Note that a feature map $\phi(\cdot)$ can be used as the general case, i.e. where kernel functions are considered in order to find a linear decision hyperplane in $\mathcal{F}$ that separates the two

classes. Converting Equation B.2 into a *quadratic programming optimisation*, an equivalent representation can be given as,

$$\arg\min_{\mathbf{w},\beta} 0.5\{||\mathbf{w}||^2\} \tag{B.3}$$

Using Lagrange multipliers $\mathbf{a}$, the dual representation of the maximum margin binary classifier is given in terms of the kernel function between two data points $\mathbf{x}_n, \mathbf{x}_m$ $k(\mathbf{x}_n, \mathbf{x}_m)$,

$$L(\mathbf{a}) = \sum_{N}^{n=1} a_n - 0.5 \sum_{n=1}^{N} \sum_{m=1}^{N} a_n a_m t_n t_m k(\mathbf{x}_n, \mathbf{x}_m) \tag{B.4}$$

This means that the kernel trick can be applied, which is beneficial in cases where the feature space is much larger than N. For a new data point $\mathbf{x}_*$, the trained model will classify it using the following rule: $y(\mathbf{x}_*) = \sum_{n=1}^{N} a_n t_n k(\mathbf{x}_*, \mathbf{x}_n) + \beta$.

An important property of SVM is that they satisfy the *Karush-Kuhn-Tucker* conditions of the optimisation problem in Equation B.4 [30]. In brief, for each ith data point has either $a_i = 0$ or $t_i y(\mathbf{x}_i) = 1$, which means that points that have $a_i > 0$ will satisfy $t_i y(\mathbf{x}_i) = 1$, so these are the *support vectors* and only these observations are used for classifying new data points, while the rest are excluded, yielding a *sparse representation*. Given a set of indices of the support vectors $s = 1, ..., S$ β can be computed using the constraint $t_s y(\mathbf{x}_s) = 1$.

However, it cannot be expected that data points in X will all be classified correctly, even if a kernel function is obtained that finds a non-linear decision boundary to separate them. Therefore, introducing a penalty term $C(> 0)$ and slack variables $\xi(\geq 0)$ where $\xi/||\mathbf{w}||$ corresponds to the perpendicular distance from the calculated decision boundary (to get the maximum margin), the primal optimisation problem can be re-written as,

$$\begin{gathered} \arg\min_{\mathbf{w},\beta} 0.5\{||\mathbf{w}||^2 + C\sum_{n=1}^{N} \xi_n\} \\ s.t.: t_n(\mathbf{w}^T \phi(\mathbf{x}_n) + \beta) \geq 1 - \xi_n, \forall n =, 1..., N \end{gathered} \tag{B.5}$$

Note that the Lagrange optimisation problem, presented in Equation B.4 is identical to the separable case, while the constraints slightly different.

The ocSVM formulation by Schölkopf et al. [162], trains a *one-class classifier* considering that points in X are all of the same class. In other words, it estimates a decision boundary using the support vectors from X to define the positive class, while the origin is considered as the only member of the negative class. Therefore, ocSVM finds the maximum margin from the origin, which holds true only for an RBF kernel [32]. Let $\beta = 1 - \rho$, gives $\mathbf{w}^T \phi(\mathbf{x}_*) - \rho = 0$, where ρ is the bias term, the primal objective to be solved is,

$$\begin{gathered} \arg\min_{\mathbf{w},\beta} 0.5\{||\mathbf{w}||^2 + \frac{1}{\nu N}\sum_{n=1}^{N} \xi_n - \rho\} \\ s.t.: t_n(\mathbf{w}^T \phi(\mathbf{x}_n) \geq \rho - \xi_n, \forall n =, 1..., N \end{gathered} \tag{B.6}$$

where, ν is both an upper bound on the outlier fraction and a lower bound of support vectors, so that it can be tuned accordingly. The decision of ocSVM for a new data point $\mathbf{x}_*$ is such that,

$$y(\mathbf{x}_*) = \rho - \sum_{n=1}^{N} a_n k(\mathbf{x}_*, \mathbf{x}_n) \tag{B.7}$$

Matlab's built-in implementation *fitcsvm.m* allows to define the fraction of outliers by modifying the bias term to exclude the percentage of outlying data points, as specified by the user.

B.2 Isolation forests

In iForest [114], the aim is to find feature-axes parallel *partitions* in the data space $\mathcal{X}$ to *isolate* each observation *recursively*, from the rest of the data, one at a time. Data are split using: 1. randomly chosen splitting points in the data space and 2. randomly chosen features from X. The method is therefore, not dependent on any *distance measure* nor on *density estimation assumptions* for points in $\mathcal{G}$ to infer whether observations are outliers or not (as in ocSVM). Instead, the following two assumptions are (needed) to be made regarding the observations that are found to be outliers, according to iForest:

- outliers exist in sparse regions in the space of $\mathcal{X}$;
- and, they are far fewer than the non-outlying observations.

As such, an iForest is one of the simplest methods to implement and to tune and the least computationally expensive to run, while most of the times works surprisingly well in practise. The method identifies outliers, as those observations that have the *shortest sequences* of data splits in an isolation tree (iTree). For an outlying observation, an iTree's branch height will be smaller (i.e. smaller splitting operations), than the height of the rest of the non-outliers. The branch height of the iTree, is then considered as an *outlier score* for the iForest formulation. By isolating its observation from the rest of the data, it means that an iTree has at most (can be terminated prematurely) N leaf nodes. However, the outliers are normally found in the first few data splits (or dimensions).
The procedure that each iTree follows is:
For $i = 1$ to D:

1. Select the pth feature from X, randomly;
2. Sub-sample $\mathbf{x}_p$, to get $\tilde{\mathbf{x}}_p \in \mathbb{R}^{m<N}$;
3. Split the data using random partitioning between the range $\min(\tilde{\mathbf{x}}_p)$ and $\max(\tilde{\mathbf{x}}_p)$;

4. Repeat steps 1-3 until all observations are isolated, if an early leaf node termination is not specified.

In iForest, a multiple number of iTrees are created, say $D \approx 1000$, to average out the final outcome. As the number of created iTrees increases, the prediction outcome, or equivalently, the averaged iTree height has less variance. Let $\hat{h}_i(\tilde{\mathbf{x}})$ be the heights of the ith iTree for all $\tilde{\mathbf{x}}$, an outlier score for each data point in X, can be calculated using an *ensemble* of iTrees, i.e. $\mathbb{E}\left[\hat{H}(\tilde{\mathbf{x}})\right]$, with $\hat{H}(\tilde{\mathbf{x}}) = \hat{h}_i(\tilde{\mathbf{x}}), ..., \hat{h}_D(\tilde{\mathbf{x}})$,

$$s(X, \hat{H}(\tilde{\mathbf{x}})) = 2^{-\frac{\mathbb{E}[\hat{H}(\tilde{\mathbf{x}})]}{c(N)}} \tag{B.8}$$

where, $c(N)$ depends on the data length N. For instance, if $s \to 1$ then indicates outliers with high probability and when $s \leq 0.5$ are considered as non-outlying points. iForest is implemented as a built-in Matlab function *iforest.m*. The hyperparameters are: ensemble size D and sub-sampling size m. A threshold can be set on the expected fraction of outliers that exist in X.

B.3 Time-varying filtering based EMD

Recall that a B-spline approximation $g_m^n(t)$, can be constructed for a signal $x(t)$, given a knot spacing m and order n through convolution with a pre-filter p_m^n whose cut-off frequency is determined by m:

$$g_m^n(t) = [p_m^n * x]_{\downarrow m} * b_m^n(t) \tag{B.9}$$

where, $b_m^n(t)$ is a post-filter required for approximation reconstruction, after it is being decimated with the pre-filter p_m^n by a factor of m. The above is a special type of low-pass filtering, with a cut-off frequency determined by knot spacing m. Hence, Step 1 of the procedure is to obtain the knot spacing that can be used to calculate $\dot{\phi}_{bis}(t)$. This is done as shown Table B.1.

Note that abrupt changes in the bisecting frequency may result in mode mixing. Therefore, the authors proposed a *realignment algorithm*. This algorithm limits the rate of change of the cut-off frequency within two adjacent maxima by a fixed value. This methodology applies only for signals that do not possess "strong" frequency modulations, i.e. fast-changing ones.

After realigning the bisecting frequency, the *extrema timings* of the oscillation of the instantaneous phase of the signal, i.e. $cos\left[\int \dot{\phi}_{bis}(t)dt\right]$, where $\dot{\phi}_{bis}(t)$, are calculated. These are the *knots* of the B-spline approximation filter. The next step involves the utilization of the instantaneous parameters found (see Table B.1), to define the stopping criterion. This determines whether the signal is "sufficiently" narrow-band or needs to be filtered further, and

Table B.1. Step 1 in TVF-EMD: calculating the local cut-off frequency.

1.	Compute IF $\phi(t)$ and IA $\|z(t)\|$ using HT.
2.	Find the local extrema of $\|z(t)\|$.
3.	Let $z(t)$ being a signal of two components, i.e. $z(t) = \alpha_1(t)e^{j\phi_1(t)} + \alpha_2(t)e^{j\phi_2(t)}$. It can be shown that, $\|z(t)\|^2 = \alpha_1^2(t) + \alpha_2^2(t) + 2\alpha_1(t)\alpha_2(t)cos[\phi_1(t) - \phi_2(t)]$ and a similar expression for $\dot{\phi}(t) = f(\dot{\phi}_1(t), \dot{\phi}_2(t), \dot{\alpha}_1(t), \dot{\alpha}_2(t), \phi_1(t), \phi_2(t), \alpha_1(t), \alpha_2(t), \|z(t)\|^2)$. The goal is to compute the bisecting frequency between $\phi_1(t)$ and $\phi_2(t)$. Assuming both $a_1(t)$ and $a_2(t)$ are locally narrow-band (vary slowly), the local extrema are dictated (approximately) by the term $cos[\phi_1(t) - \phi_2(t)]$.
4.	Compute $\alpha_1(t)$, $\alpha_2(t)$ and subsequently $\dot{\phi}_1(t)$, $\dot{\phi}_2(t)$ by using the assumption in Step(3) and an interpolation approximation to simplify things.
5.	Then the bisecting frequency is simply the average of $\dot{\phi}_1(t)$ and $\dot{\phi}_2(t)$.

is calculated as the ratio of the Loughlin instantaneous bandwidth [116] and the weighted IF average, i.e.

$$\psi(t) = \frac{\sqrt{\frac{\dot{\alpha}_1(t)^2 + \dot{\alpha}_2(t)^2}{\alpha_1(t)^2 + \alpha_2(t)^2} + \frac{\alpha_1^2(t)\alpha_2^2(t)(\dot{\phi}_1(t) - \dot{\phi}_2(t))^2}{(\alpha_1^2(t) + \alpha_2^2(t))^2}}}{\frac{\alpha_1^2(t)\dot{\phi}_1(t) + \alpha_2^2(t)\dot{\phi}_2(t)}{\alpha_1^2(t) + \alpha_2^2(t)}} \tag{B.10}$$

where, the Loughlin bandwidth is in the numerator and the weighted IF average is at the denominator. It can be observed that as the term $|\dot{\phi}_1(t) - \dot{\phi}_2(t)|$ increases, there will be greater separation between the two components in the signal. In this case, the Loughlin bandwidth is greater. This bandwidth measure can be seen as the standard deviation in the weighted IF average at any given time. In general, the above stopping criterion means that when $\psi(t)$ decreases, the bandwidth of the signal will decrease, too. Given a threshold ξ a signal is considered narrow-band (locally) if $\psi(t) \leq \xi$.

B.4 M-estimators

This type of estimators are generalisations of the maximum likelihood estimate (MLE). Therefore, $\boldsymbol{\theta}_{r,M}$ will be "nearly optimal" for the assumed two-mixture model F, in Equation 5.2. Let X be *iid* and F a symmetric and elliptical distribution, then the likelihood expression may be written as a function h_d for computing N multivariate distances,

$$p(X|\boldsymbol{\theta}) = |\Sigma|^{-N/2} \prod_{n=1}^{N} h_d((\mathbf{x}_n - \boldsymbol{\mu})\,\Sigma^{-1}(\mathbf{x}_n - \boldsymbol{\mu})^T) \tag{B.11}$$

where $|\Sigma|$ is the determinant of Σ. Differentiating the above expression, with respect to $\boldsymbol{\theta}$, yields the MLE estimates,

$$\boldsymbol{\theta}_{r,M} = \begin{bmatrix} \sum_{n=1}^{N} w_1(d_{rmsd}^2)\,\mathbf{x}_n \,/\, \sum_{n=1}^{N} w_1(d_{rmsd}^2) \\ 1/N \sum_{n=1}^{N} w_2(d_{rmsd}^2)(\mathbf{x}_n - \boldsymbol{\mu}_{r,M})(\mathbf{x}_n - \boldsymbol{\mu}_{r,M})^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{r,M} \\ \Sigma_{r,M} \end{bmatrix} \tag{B.12}$$

where, $w_1(d^2_{rmsd})$ and $w_2(d^2_{rmsd})$ are functions of the robust distances (using the new estimates $(\boldsymbol{\mu}_{r,M}, \Sigma_{r,M})$), thus, they poses certain robustness properties including a bounded IF. Typical functions include the Bisquare, as demonstrated on a toy dataset, for a regression problem, previously.

B.5 Self-adaptive noise cancellation

As part of the EA procedure, a pre-whitening step is usually required in order to remove periodicities in the signal. One such approach is to apply a self-adaptive noise cancellation (SANC) procedure [33]. SANC separates two uncorrelated components from a signal, using another (reference) signal that is related to the corresponding part of the original signal by a transfer function. As such, the reference signal can be subtracted from the original signal. This is the basic adaptive noise cancellation procedure. In the case of DI in REBs, SANC first delays the original signal, where the delay is larger than the correlation length of the random component of the signal (bearing vibration). Therefore, the transfer function between the deterministic part of the signal (gears, etc.) and its delayed version is calculated and then subtracted from the original signal, leaving only the bearing vibration components and other noise sources.

B.6 Minimum Regularised Covariance Determinant

A short description of the MRCD will be provided, as it established the basis for the development of the kMRCD, which is described in Chapter 5. In MRCD, the scatter matrix of an h-subset is a convex combination of scatter and target matrices, given a scalar parameter. A "...predetermined and well-conditioned symmetric and positive definite" target matrix T, is combined with a scatter matrix $S^h(\mathbf{z})$ of an h-subset of a robustly standardised (e.g. subtract median and divide by MAD) dataset $\mathbf{z}_n \in R^{N\times p}$ by using a scalar regularisation parameter ρ to get the regularised scatter,

$$\Sigma^h_{reg}(\mathbf{z}) = \rho T + (1-\rho)c_1 S^h(\mathbf{z}), \quad 0 \leq \rho \leq 1 \tag{B.13}$$

where, c_1 is a consistency scalar constant. Note that for a robustly standardised dataset, T can be chosen as the identity matrix $I_{p\times p}$, depending on whether the features in the original dataset X are correlated. The problem of finding an h-subset, out of all H-subsets, can be expressed as the following optimisation procedure,

$$\Sigma_{MRCD}(\mathbf{z}) = \arg\min_{h\in H}\{|\,\Sigma^h_{reg}(\mathbf{z})|\} \tag{B.14}$$

To reach the above objective, C-steps (see Chapter 5 for a description) are being employed.

B.7 Kernels and the kernel "trick"

The kernel "trick" has been used in Support Vector Machines (SVM), but, the concept has far reaching potential. This is because whenever a dot product is required to be computed between two p-dimensional data points, $\mathbf{x}_n^T \mathbf{x}_m$, a kernel can be used to replace it. Given any $\mathbf{x}_n$ and $\mathbf{x}_m \in \mathcal{X}$ a function $k(\cdot,\cdot)$ is called a *kernel* on $\mathcal{X}$ such that $\mathcal{X} \rightarrow \mathcal{F}$ using a feature map ϕ,

$$k(\mathbf{x}_n, \mathbf{x}_m) = \phi(\mathbf{x}_n)^T \phi(\mathbf{x}_m) \tag{B.15}$$

The above computation is being carried out in $\mathcal{X}$ and produces a scalar value. That is, $\mathbf{x}_n$ and $\mathbf{x}_m$ are *not explicitly* transformed in $\mathcal{F}$, in which $\phi(\mathbf{x}) \in \mathbb{R}^{d>p}$. By maintaining the same complexity as in the original dot product computation, it makes subsequent operations feasible. This is especially true when $d \rightarrow \infty$ using a specific type of kernel called radial basis function (RBF),

$$k(\mathbf{x}_n, \mathbf{x}_m) = \exp(-\gamma_{rbf} ||\mathbf{x}_n - x_m||^2) \tag{B.16}$$

where, γ_{rbf} is the kernel bandwidth, which determines the significance in the relationship between $\mathbf{x}_n$ and $\mathbf{x}_m$. At the two extremes, when $k(\mathbf{x}_n, \mathbf{x}_m) \rightarrow 1$ means that the two data points are very similar (i.e. approaching the mean of the Gaussian), while when $k(\mathbf{x}_n, \mathbf{x}_m) \rightarrow 0$ the two data points are almost unrelated (i.e. at the tails of the Gaussian). For the RBF kernel, this exponential-type of similarity is given by the Euclidean distance between the two points, and its importance can be tuned via γ_{rbf}, i.e. it is a *hyperparameter*. By taking Taylor's expansion of the term that contains the dot product $\exp(-2\mathbf{x}_n^T \mathbf{x}_m)$ the RBF kernel can be shown be equivalent to an *infinite sum over polynomial kernels*.

Now, given N data points, in the p-dimensional space $\mathcal{X}$, a kernel function with a feature map ϕ can be used to construct an $N \times N$ kernel matrix $\mathcal{K}_{\mathbf{x},\mathbf{x}}$, which contains all N kernel calculations for any $\phi(\mathbf{x})$,

$$\mathcal{K}_{\mathbf{x},\mathbf{x}} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \geq 0 \tag{B.17}$$

The condition $\mathcal{K}_{\mathbf{x},\mathbf{x}} \geq 0$ is called Mercer's condition and is a requirement for $\mathcal{K}_{\mathbf{x},\mathbf{x}}$ in order to be a valid kernel matrix, along with the fact that it needs to be a symmetric matrix. The RBF kernel is considered to be the most important kernel choices, and will be used throughout the thesis, i.e. whenever a kernel method such as the kMRCD or kPCA is used. Note that it is usually the case that $\sum_i^N \phi(\mathbf{x}_i) \neq 0$, so that the *Gram* matrix needs to be used instead of the kernel matrix, i.e.

$$\tilde{\mathcal{K}}_{\mathbf{x},\mathbf{x}} = \mathcal{K}_{\mathbf{x},\mathbf{x}} - \mathbf{1}_N \mathcal{K}_{\mathbf{x},\mathbf{x}} - \mathcal{K}_{\mathbf{x},\mathbf{x}} \mathbf{1}_N + \mathbf{1}_N \mathcal{K}_{\mathbf{x},\mathbf{x}} \mathbf{1}_N \tag{B.18}$$

where, $\mathbf{1}_N$ is the $N \times N$ matrix with all elements equal to $1/N$.

B.8 C-steps for robust estimators

C-steps are core components of the FMCD algorithm, and are used to obtain the h samples with the smallest robust squared distances, and therefore the smallest $|\Sigma_{MCD}|$. In summary the C-steps algorithm of the MCD proceeds as follows,

1. Randomly draw a subset of length $p+1$
2. Compute multivariate mean and scatter, $\boldsymbol{\mu}_0$, Σ_0 (if $|\Sigma_0| = 0$, add more observations incrementally).
3. Calculate Mahalanobis distances $d^2_{msd}(\mathbf{x}_n, \boldsymbol{\mu}_0, \Sigma_0)$, $\forall n = 1, ..., N$.
4. Sort observation indices by d^2_{msd} and select the h observations with the smallest value.
5. Using the h observations compute a new multivariate mean and scatter, $\boldsymbol{\mu}_1$, Σ_1. It is guaranteed that $|\Sigma_1| \leq |\Sigma_0|$.
6. Repeat steps $2-5$, M times until $|\Sigma_M| = |\Sigma_{M-1}|$

Appendix C

Robust estimators properties & the framework of novelty detection

C.1 Properties of robust estimators and robustness criteria

To evaluate the performance of a robust estimator three different criteria are used: the *influence function* (IF), *maximum bias curve* (mbC) and associated *breakdown point* (α^*), and *statistical efficiency*.

The IF is defined as the measure of influence of an infinitesimal amount of contamination on the robust estimator, as a function of its variation in the feature space. Let $\hat{\boldsymbol{\theta}}(\mathcal{G})$ be an estimated statistic of the probability distribution $\mathcal{G}$, the distortion of $\hat{\boldsymbol{\theta}}$ given slight deviations from $\mathcal{G}$, as $\alpha \to 1$ (see Equation 5.2, can be computed as the change of $\hat{\boldsymbol{\theta}}$ with respect to the variation of point $\mathbf{x}$ in p-dimensional space,

$$IF(\mathbf{x};\hat{\boldsymbol{\theta}},\mathcal{G}) = \lim_{\alpha\to 1} \frac{\hat{\boldsymbol{\theta}}(\alpha\mathcal{G}+(1-\alpha)\delta_{\mathbf{x}}) - \hat{\boldsymbol{\theta}}(\mathcal{G})}{(1-\alpha)} \tag{C.1}$$

where, $\delta_{\mathbf{x}}$ is the point-mass probability for $\mathbf{x}$. A robust estimator needs to have both a *smooth* and *bounded* IF, in order to ensure the estimate of $\hat{\boldsymbol{\theta}}$ is stable and changes only slightly with shifts in $\mathbf{x}$. This is also referred to as *local shift insensitivity*.

Realistically, a proportion of outliers may exist in the data so that $\alpha < 1$. Hence, the calculation of the IF is a special case of point-mass contamination. The mbC, on the other hand, examines the *amount of contamination*, and more specifically, the worst possible one for $\hat{\boldsymbol{\theta}}$ at the assumed distribution $\mathcal{G}$,

$$mbC(\alpha;\hat{\boldsymbol{\theta}},\mathcal{G}) = \sup_{G} ||\hat{\boldsymbol{\theta}}(\alpha\mathcal{G}+(1-\alpha)\delta_{\mathbf{x}}) - \hat{\boldsymbol{\theta}}(\mathcal{G})|| \tag{C.2}$$

where, for a robust estimator mbC is bounded when $\alpha < 1$. For instance, the OLS regression (see Figure 5.1) has an unbounded mbC for any $\alpha \neq 1$. On the other hand, the univariate median has a bounded mbC until $\alpha = 0.5$, where the estimates of $\hat{\boldsymbol{\theta}}(\mathcal{G})$ becomes no longer

valid or it breaks down, i.e. $mbC \to \infty$.

To quantify this limit, the breakdown point α^* is defined as convenient measure for investigating different robust estimators. The value of α^* for the classical Mahalanobis distance is 0, while the highest possible that can be obtained by any robust estimator is 0.5. For instance, in calculating the scatter matrix of $\mathcal{G}$, $\Sigma_G = (\mathbf{x}_n - \boldsymbol{\mu}_G)(\mathbf{x}_n - \boldsymbol{\mu}_G)^T$, with eigenvalues $\lambda_1 \geq ... \geq \lambda_p$, α^* is defined as the minimum fraction of outliers (m/N) that may cause any $\lambda_i \to \infty$ or its determinant to become 0, i.e. explosion or implosion of Σ_G. Based on the previous mbC definition,

$$h(\hat{\boldsymbol{\theta}}, \mathcal{G}) = \min_G \left\{ m/N; mbC(\alpha; \hat{\boldsymbol{\theta}}, \mathcal{G}) = \infty \right\} \tag{C.3}$$

The increase in α^* for an estimator leads to a proportional decrease in statistical efficiency. Generally, any robust estimator will be less efficient than a standard estimator given the assumptions of the latter hold true for $\mathcal{G}$. A typical example is the MLE for a Gaussian-distributed $\mathcal{G}$, which implies the least variance (or uncertainty) than any other estimator. On the other extreme of the spectrum, the univariate median has the highest α^* possible, while when compared to the standard mean, its efficiency is much less, given that $\mathcal{G}$ is elliptically and symmetrically distributed.

Important properties of robust estimators include *affine equivariance*. This is usually an important consideration for multivariate analysis (i.e. where each column of X is related), since the data matrix X may be subjected to linear transformations, e.g. rescaling and rotations. Given a $p \times p$ nonsingular matrix A and a vector b of length p, so that $X = \mathbf{x}_n (n = 1, ..., N)$ is linearly transformed as $A\mathbf{x}_n + b$, a robust estimator is affine equivariant if and only if the following two expressions hold true,

$$\begin{aligned} \boldsymbol{\mu}(A\mathbf{x}_n + b) &= A\boldsymbol{\mu}(\mathbf{x}_n) + b \\ \Sigma(A\mathbf{x}_n + b) &= A\Sigma(\mathbf{x}_n)A^T \end{aligned} \tag{C.4}$$

The above ensures that the robust multivariate distance equivalence of the Mahalanobis, i.e. the robust distance measure, $d^2_{rmsd} = (\mathbf{x}_n - \boldsymbol{\mu}_r)\Sigma_r^{-1}(\mathbf{x}_n - \boldsymbol{\mu}_r)^T$, where $\boldsymbol{\mu}_r$ and Σ_r were calculated using a robust estimator, is *affine invariant*. This implies that the outlyingness of $\mathbf{x}_n$ from $\mathcal{G}$, will not be influenced by linear transformations of the matrix columns/features. In contrary, methods, including PCA and Deep Neural Networks (DNN), are not affine equivariant and thus, it is common practise to run the analysis multiple times, using different linear transformations, to investigate the sensitivity of the results.

C.2 An overview of novelty detection

Many applications ranging from medical diagnostics to system network intrusion, have benefited greatly from novelty detection, in terms of improving their safety and reliability. A

relatively recent review paper [145] classifies all available novelty detection techniques into five categories, as summarised in Table C.1.

Table C.1. The five types of novelty detection.

Class	Brief description
Probabilistic	Density estimates of the undamaged condition; defining a novelty threshold at low density regions.
Distance-based	Analysis of clusters of data by assuming that the undamaged condition is tightly concentrated in space and that the opposite class is far from its nearest neighbours.
Reconstruction-based	Measures the distance between an unseen data point and the output of a model trained using undamaged class.
Domain-based	Calculates the location of the boundary that describe class data in the feature space, without making assumptions about their distribution a priori.
Information content-based	The information content, e.g. entropy, of the data provides a measure of uncertainty. This measure will drop significantly if an unseen data point from class $\mathcal{A}$ is removed from the data set.

In machine learning an unknown function $f(\mathbf{x})$ is to be approximated by learning certain patterns from a training data set $\{(\mathbf{x}_n, y_n); n = 1, ..., N\}$. In the general case, each $\mathbf{x}_n$ is a d-dimensional vector, while a label y_n is the target variable. In SHM and CM, $\mathbf{x}_n$ are plentiful, whereas the system's health state, i.e. y_n is not. This is an inverse problem, in which we are working from the output to find an approximation for the underlying generating mechanism behind it. The schematic diagram in Figure C.1, illustrates the components of a generic machine learning procedure [17].

In the most ideal of cases, a small fraction of training data are provided with their corresponding labels from the system's undamaged state, i.e. one-class classification problems. The hypothesis set, $\mathcal{H}$, contains several candidate formulas, $h \in \mathcal{H}$, that are chosen by the learning algorithm, $\mathcal{C}$, to best represent the training data. From every possible candidate formula, $h(\mathbf{x})$, the learning algorithm chooses the one that closely approximates $f(\mathbf{x})$. This is the final hypothesis function, $g(\mathbf{x})$, that will closely approximate the unknown function, if most of the data patterns are successfully identified. The goal is to get the final hypothesis function, so that accurate predictions for any unseen observation $\mathbf{x}^*$ can be made. The hypothesis set, $\mathcal{H}$, and learning algorithm, $\mathcal{C}$, comprise the learning model $\mathcal{L}$.

The simplest hypothesis set that one can consider is the perceptron. The perceptron can be used for classification tasks when the training data are linearly separable. For reference

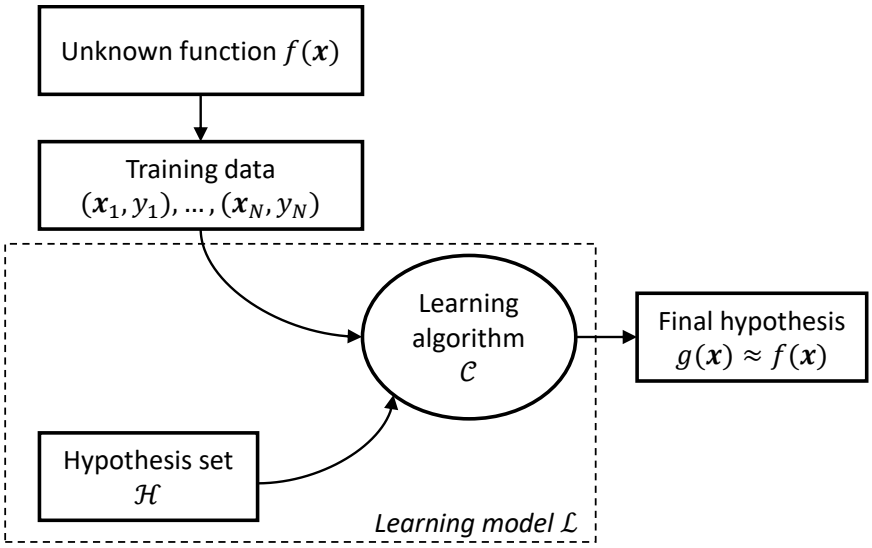

Figure C.1. Machine learning general framework diagram.

purposes, a short description of the perceptron will be presented at this point. Consider a D-dimensional training data matrix $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_N]$, where each row contains an experimental observation, and its corresponding labels vector $\mathbf{y} = [y_1, ..., y_N]$. This data set can be used to train the perceptron to obtain an accurate hypothesis function. A candidate hypothesis of the perceptron for the Nth data point, $\mathbf{x}_N$, is given as:

$$h(\mathbf{x}_N) = sgn(\mathbf{w}^T \mathbf{x}_N - \beta) \tag{C.5}$$

where, β is a bias term and $\mathbf{w}$ is a D-dimensional weight vector. The above equation implies that the hypothesis for $\mathbf{x}_N$ will take the following values:

$$h(\mathbf{x}_N) = \begin{cases} +1, & \text{if} \quad \mathbf{w}^T \mathbf{x}_N - \beta \geq 0 \\ -1, & \text{if} \quad \mathbf{w}^T \mathbf{x}_N - \beta < 0 \end{cases} \tag{C.6}$$

It is possible at this stage that the initial guess of both β and $\mathbf{w}$ will classify wrongly the observations, i.e. $sgn(\mathbf{w}^T \mathbf{x}_N - \beta) \neq y_N$. For that reason, β and $\mathbf{w}$ are updated (using a learning strategy $\mathcal{C}$) until all points are classified correctly (as they are linearly separable). This is the simplest example of classification. More advanced methods, such as support vector machines (SVM), can be used to classify non linearly separable data by finding the maximum in-between separation boundary, as described in Chapter 2. In the case of SVM, $\mathcal{H}$, consists of a set of non-linear candidate formulas that are selected by its accompanied, $\mathcal{C}$, which is a quadratic program (QP).

The machine learning procedure that is followed to estimate g is called the training phase, while the procedure that is used to make predictions using g is called the testing phase.

Regarding the first, it can be seen that prior to obtaining g, the training data are passed through a stage called feature extraction. The purpose of this stage is to transform the data into a format, which can improve the capability of $\mathcal{L}$ to learn the patterns in the training data. For instance, by separating better the observations corresponding to different labels. This new format, is called features and is assembled into an $N \times D$ feature matrix. Its rows correspond to the N observations and its columns to the D features. An example of feature extraction that can be considered is by taking the Fourier Transform of each observation in $\mathbf{X}$. In general, feature extraction comprises three sequential stages: pre-processing, feature transformation and selection. In brief, pre-processing refers to data normalisation or re-scaling, whereas feature selection, as the name suggests, is the process of choosing low-dimensional informative features.

As a final step, the final hypothesis is evaluated for its generalisation capability at the performance evaluation stage. Generalisation, is one of the most important measures for assessing model performance in machine learning. This is because it measures the ability to "learn" the patterns rather than "memorising" the data. Memorisation refers to a well-known problem encountered in statistical modelling, where the noise in the training data is learned along with its underlying function. As an example, consider learning a sinusoidal function with a high Signal-to-Noise Ratio (SNR). We can find a polynomial function that describes this function by adding higher order terms. As the order increases the sum of squared errors between the polynomial and the noisy sinusoidal function diminishes. However, this increase in the number of polynomial coefficients results in over fitting and is a direct consequence of the curse of dimensionality (the dimensionality increases as more coefficients are added to the polynomial), as described by Bishop [30]. It is expected that during the performance evaluation stage, this complex polynomial function will diverge from the actual sinusoidal function, giving a very large prediction error, when an unseen data point $\mathbf{x}^*$ presents itself. That is why feature selection is necessary. The performance evaluation stage is conducted by separating the observations in a training, a validation and a testing data set. For instance, 75% for the first two and the rest are used for the testing phase. Using the validation data set, the final hypothesis obtained from $\mathcal{L}$ (using the training data set), is tested. It is possible that during validation, the error obtained using $g(\mathbf{x})$ to estimate the corresponding labels $\mathbf{y}$, can be found to be unacceptable. In this case, a new $g(\mathbf{x})$ is obtained, either by choosing different features or changing the hypothesis set $\mathcal{H}$ (i.e. the method). The procedure then goes to the testing phase if the validation results look promising. In the testing phase, the performance of $\mathcal{L}$ is checked for a second time, since over fitting is possible to occur when validation is performed many times. Cross-validation is another important performance evaluation approach that does not require a

separate validation data set. This is advantageous in cases when the observations are very limited, as in most of the SHM and CM applications, and an accurate learning model must be obtained.

The ability of the learning model to find $g(\mathbf{x}) \approx f(\mathbf{x})$, is highly dependent on the amount and quality of the training data, as well as, our tolerance for errors. If the data are too noisy or do not represent well the system characteristics, e.g. inadequate sensor readings, then a suboptimal hypothesis function will be selected by $\mathcal{C}$. The following example will aim to provide a simple explanation as to the extend in which a learning model can learn the experimental patterns well. Consider a simple experiment with two possible events: $E1$ and $E2$. The theoretical probability of the first happening is μ (and the second is $\mu - 1$). If μ were known, the future output of the experiment would have been calculated. However, these two probabilities are unknown, and instead we are only given N observations that we can use to calculate the fraction of $E1$ happening in this experiment, which is ν. The law of large numbers tells us that when $N \to \infty$, then $\nu \to \mu$ within a tolerance ϵ that we can specify. As described in [17], Hoeffding's Inequality is of fundamental importance to machine learning:

$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq 2e^{-2\epsilon^2 N} \tag{C.7}$$

In other words, the probability that the absolute value of the difference between the empirical and theoretical probabilities being greater than a given tolerance, is dependent only upon the number of training data observations N and the squared tolerance value ϵ. If the requirements are very stringent then learning might be proved infeasible even if N is very large. The number of observations for training is an important parameter, and later it will be shown that an attempt has been made to increase N artificially.

At this stage, it is important to distinguish the three main approaches of machine learning. This serves as a starting point for discussing the methods that will be presented in subsequent sections. The first is called supervised learning, because each observation is always given a corresponding label or value (by a supervisor). Its task is to match the characteristics of the observations to this label or value by obtaining the final hypothesis to closely approximate the unknown generating function. If the task is to match the observation to some categorical variables, e.g. "good", "very good", etc., then this is a classification problem. If instead, the observations need to be matched to some other continuous values, e.g. to relate the temperature to the power output of an engine, then the problem is called regression.

In most real-life problems, however, only a certain proportion of labels is known beforehand or not known at all even. In these problems, the frameworks of unsupervised or

semi-supervised learning are used. The former, being the most challenging one, utilises the features extracted to learn the topologies of the distribution of the data in the D-dimensional space. Therefore, clusters can be defined and unseen data are assigned in one of these clusters per their similarities, e.g. Euclidean distances. Although this is an ambiguous approach to obtaining the final hypothesis, its practical usefulness is high and there is currently a lot of research in this area. On the other hand, the fundamental problem that concerns semi-supervised learning approaches, is the prediction of labels through finding similarities between unlabelled and labelled observations. Given that two data points are found to be "similar" in the D-dimensional space, the labels will be "propagated" on neighbouring data points until convergence. This strategy relies on the assumption that common characteristics between two labelled and unlabelled observations can be used as a measure of their similarity, and can be grouped together. Thus, having a good metric for measuring the similarity is critical in this machine learning approach.

The final machine learning approach is called Reinforcement learning and is mainly concerned in finding the optimal way for decision making for any dynamic system that interacts with its environment. The machine learning procedure shown in fig. C.1 is still applicable, however, the learning process is performed sequentially, i.e. a function of time. In addition, there are no labels provided, but, there is a feedback input to learning model $\mathcal{L}$ regarding the progress that the system is making in reaching the final hypothesis. In this case, the final hypothesis is the objective that needs to be reached following several sequential steps of trial-and-error. The fundamental property of Reinforcement learning is that current system states at t, directly influence the next step $t+1$. Given that states are observable (or partially observable), the problem can be viewed as a Markov Decision Process, as described in more detail by Sutton and Barto [174]. Therefore, this approach differs significantly from the other three machine learning approaches, since training data observations are not assumed to be independent identically distributed. Therefore, only the first three machine learning approaches are applicable to obtain the final hypothesis in this work.

It is the final hypothesis that will be used, in this work, to monitor a system's state. This is where the strength of a well-trained learning model, $\mathcal{L}$, lies. Although a learning model lacks transparency, it can capture the complex physics of the system very well, while offering the flexibility to analyse the problem on different parameters. At the same time, they are simpler to analyse, as only the inputs and outputs can be "seen". Using such model, it is possible to use it to identify optimal engine parameters that meet certain performance requirements, e.g. increase in thermal efficiency and safety, as discussed previously. Moreover, a sensitivity analysis can be conducted by using the learning model as a tool to investigate key parameters that affect its performance.

In novelty detection, the model is designed using training data $\mathbf{X}$ from the nominal or design-intended system state. For an unseen observation $\mathbf{x}^*$, a prediction is made using the final hypothesis and compared to a novelty threshold q, to give the following:

$$g(\mathbf{x}^*) = \begin{cases} > q, & \text{system state is novel} \\ \leq q, & \text{system state is nominal} \end{cases} \tag{C.8}$$

It is apparent that the value of q is equally important as the prediction $g(\mathbf{x}^*)$ in providing accurate predictions of the system's state. For convenience from now on, the novel system state will be referred to as class $\mathcal{A}$ and the nominal state as class $\mathcal{N}$. The reader might be wondering whether it is possible to use a supervised learning approach, where each observation can be labelled as either class $\mathcal{A}$ or $\mathcal{N}$ to solve this problem. When implementing such a system in practise, information regarding the former class is usually very rare or even not available at all, whereas there can be a lot of observations from class $\mathcal{N}$. More importantly, even if some observations are available for training, class $\mathcal{A}$ observations will tend to be very different from each other as the system is inherently complex. Therefore, the only assumption that can be made is about the distribution of class $\mathcal{N}$ given the training data $\mathbf{X}$, because this class is usually very well-understood, as discussed also in [176]. One of the key concepts in novelty detection is the prognostic capability that it offers in detecting a potential unwanted scenario that might lead to component damage. Thus, the essence of novelty detection is to identify the deviations from the nominal system state by means of monitoring some suitable precursors. This prognostic capability depends on: the sensitivity of the measured signal and the sensitivity of the model. However, a high sensitivity in both will cause too many false alarms (predicting an observation as class $\mathcal{A}$, while being class $\mathcal{N}$), and a trade-off needs to be made between the two. This trade-off can be provided by setting an appropriate value of the novelty threshold q.

In regression, the model is designed by capturing the relationship between the dependent and independent parameters. The relationships that are captured, e.g. with an artificial neural network (ANN), are non-linear in the coefficients, which distinguishes it from the simple polynomial fitting. In the context of gas turbine engine analysis, a learning model for regression makes no assumptions, thus, it can capture non-linear characteristics such as combustion instabilities, and might outperform CFD-based models in obtaining realistic results. Moreover, using Bayesian machine learning the uncertainty in prediction can be calculated. Accounting for uncertainty in dynamic noisy environments, enables better decisions to be made and can provide a better insight into the dynamics of the system. Using Bayes rule, we can do inference about a certain hypothesis H (for instance, it can be the

learning model's parameters), from the data:

$$\mathbb{P}[H|\ data] = \frac{\mathbb{P}[data|\ H]\ \mathbb{P}[H]}{\mathbb{P}[data]} \tag{C.9}$$

The improvement of the learning model can also be achieved using data fusion methods. These methods combine several measurements together to increase the prediction accuracy of the model. It is particularly suited for online novelty detection, because measurements from different sensors are already available.

Appendix D

Additional results

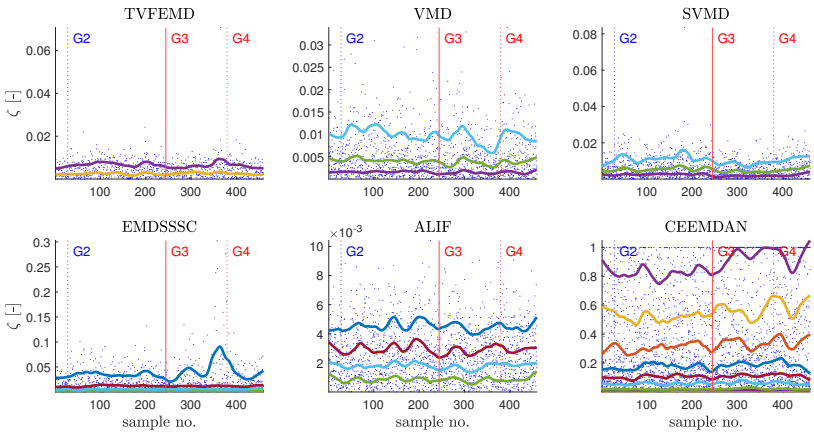

Figure D.1. LA-4DOF: Damping ratio estimates from system identification (Chapter 4).

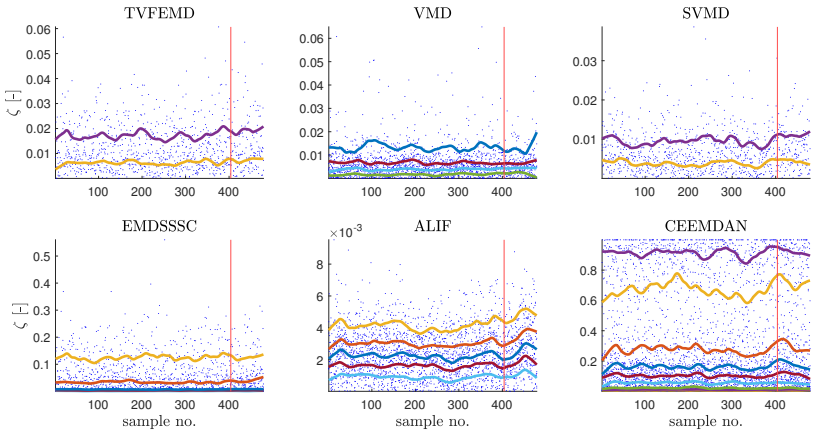

Figure D.2. NSim-4DOFLin: Damping ratio estimates from system identification (Chapter 4).

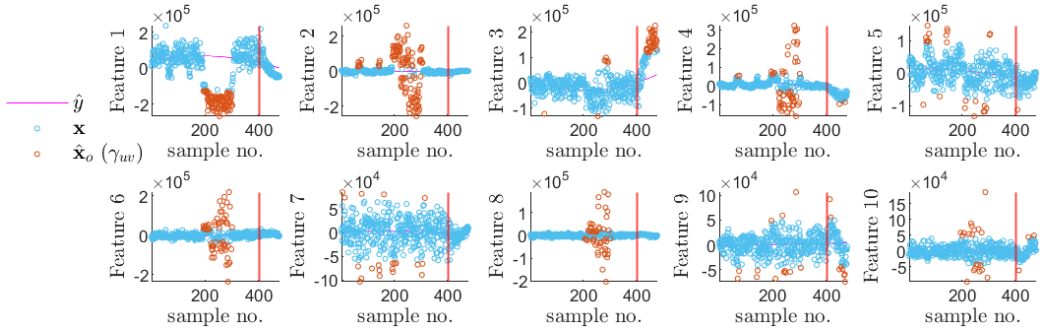

Figure D.3. NSim-4DOFNonLin: The first 10 PCA components of the PSD amplitudes (Chapter 5).

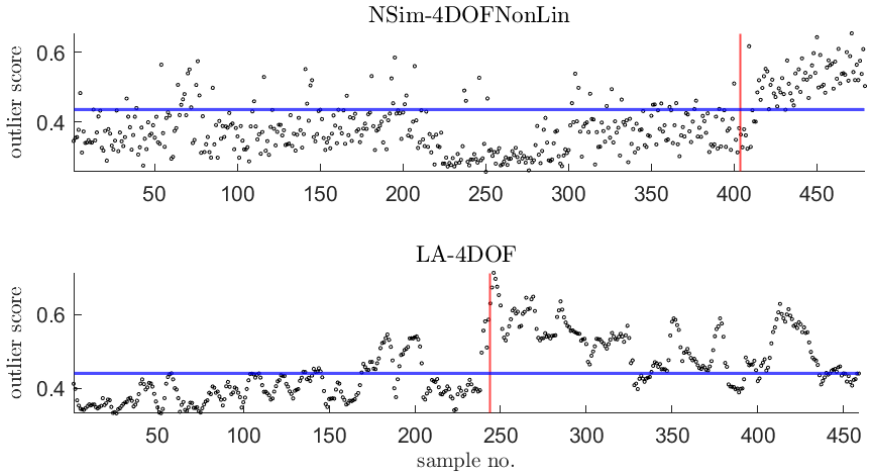

Figure D.4. iForest outlier scores on NSim-4DOFNonLin and LA-4DOF datasets (Chapter 5).

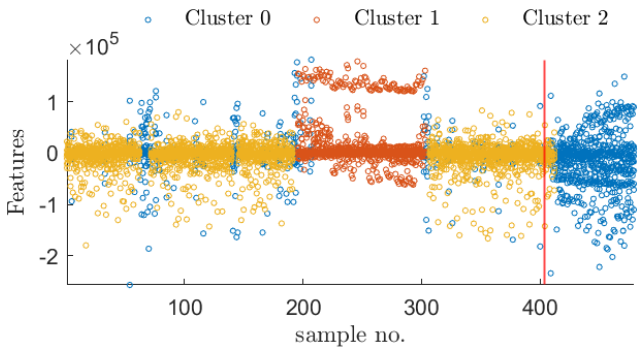

Figure D.5. NSim-4DOFNonLin: Clustering analysis using tclust (Chapter 5).

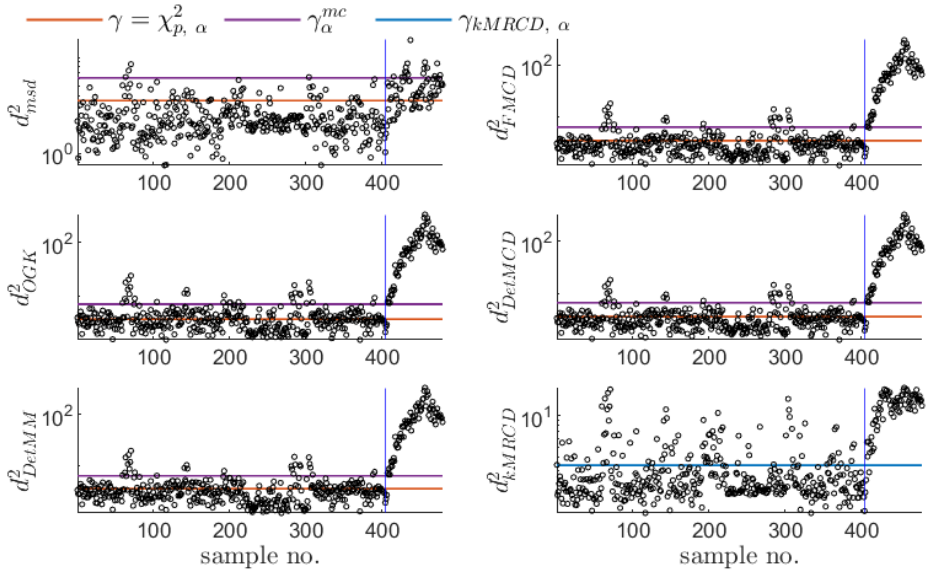

Figure D.6. NSim-4DOFNonLin: SMDs computed using robust estimators on the first 10 rPCA components of PSD amplitudes as DSFs (Chapter 5).

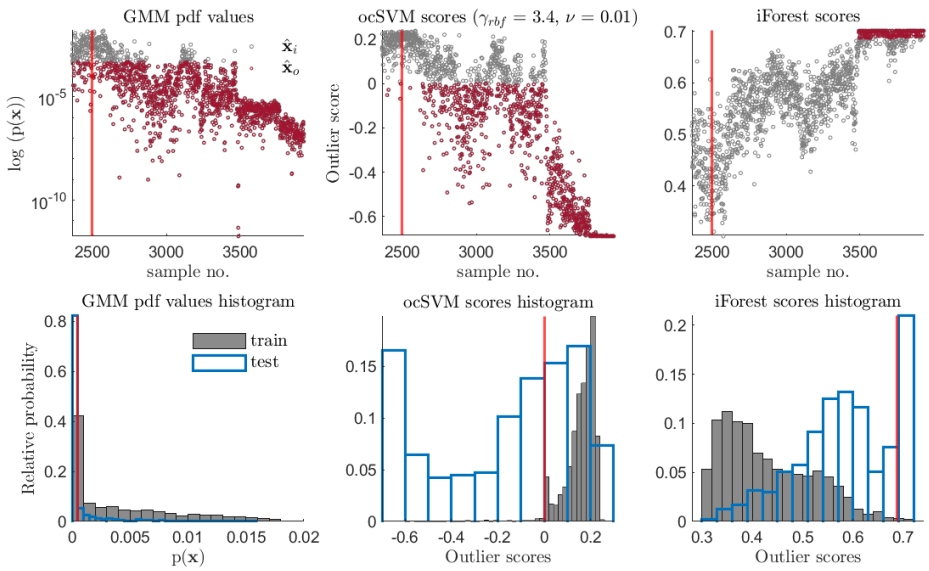

Figure D.7. Z24: outlier scores and corresponding histograms - no robust pre-processing has been applied (Chapter 6).

Appendix E

A simulated scenario for CM applications & demonstration of the envelope analysis procedure

E.1 Outer race damage using a simulated example

A simulated example, representing a fixed-speed rotating machinery consisting a shaft and a bearing with and without an outer race damage, is used to examine the EA procedure. Damage is present only on the outer race of an REB, and includes vibrations from the main shaft and its five harmonics, in the form of constant-amplitude sinusoidal signals. Damage, is considered as an amplitude modulated signal (having half-Gaussian envelope), as shown in Figure E.1, with a specific modulation frequency $f_{n,imp} = 3kHz$.

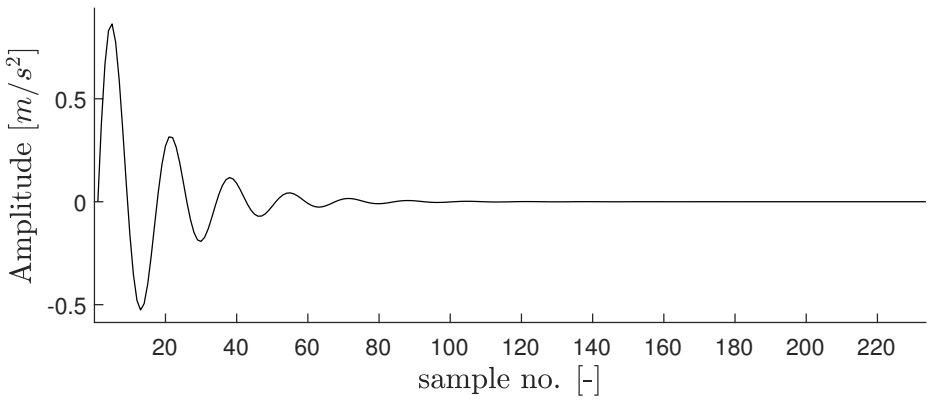

Figure E.1. Simulated damage signal - single impulse amplitude modulated.

The time-domain (shown for only $0.1s$) and frequency-domain (power spectra with a frequency resolution of 3 Hz) plots from three different simulation scenarios are shown in Figure E.2.

The differences of the three simulations are:

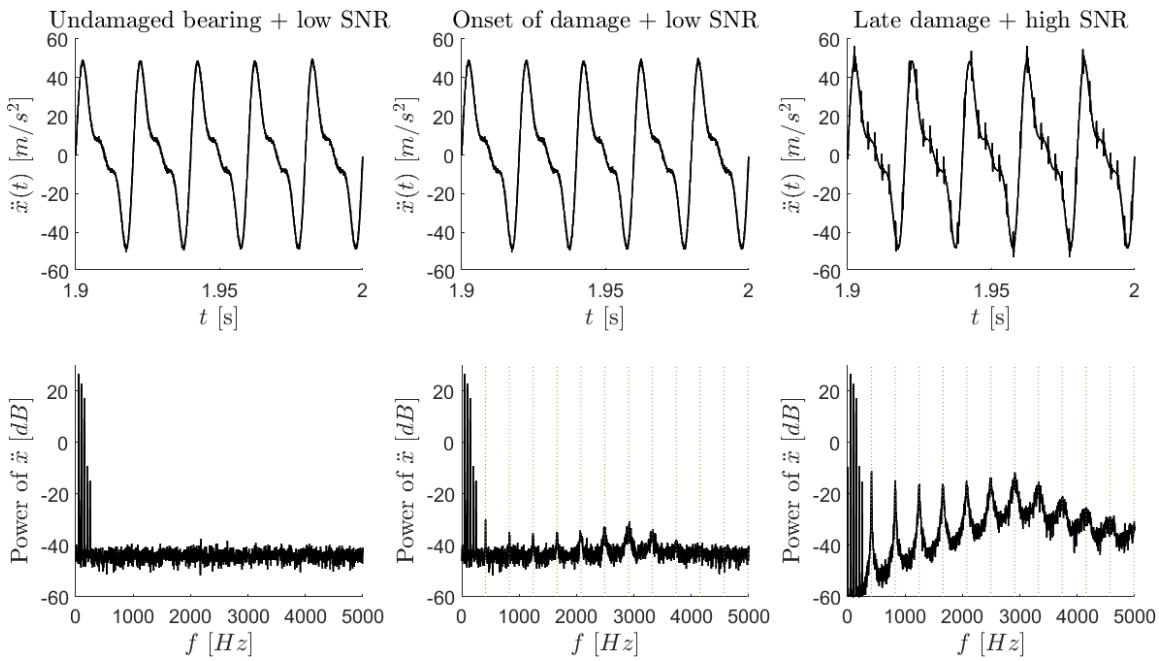

Figure E.2. Overall acceleration from three simulation scenarios shown in time- and frequency-domains (zoomed in from 1.9 to 2 seconds and 0 to 5 kHz): Undamaged rolling element bearing with low SNR (left plots), Onset of damage with low SNR (middle plots) and Late damage with high SNR (right plots).

- Undamaged bearing and low SNR (Sim A): Undamaged system with noise added to the overall signal equal to 0.6 m/s^2.
- Onset of damage and low SNR (Sim B): System at an early stage of damage, having same noise as in Sim A and amplitude peak impulses equal to 1.1 m/s^2.
- Late damage and high SNR (Sim C): System on a late stage of damage, i.e. having amplitude impulses 10 times larger than in Sim B, while the overall noise level has been reduced by 10 times, as well.

The main parameters, common to all three simulation scenarios, are shown in Table E.1. The frequency smearing is evident on the power spectra, due to the bearing *rolling slip* of 2 %. This, of course, implies that the impulse train produced using the prototype impulse signal, presented in Figure E.1, will have *unequal spacing* between each impulse. The standard deviation of the BPFO, calculated with 2 % rolling slip, is 8.3 Hz. Thus, when referring to the value of the BPFO, from now on, only its mean value will be meaningful, i.e. ≈ 415 Hz, as calculated with Equation 2.20.

The acceleration impulses increase in amplitude as time progresses; sign of deterioration of the bearing. In the ideal scenario (Sim C), there is also a high SNR and the impulses

Table E.1. Main parameters for rolling element bearing simulation.

Parameter	Symbol	Value	Unit
Sampling rate	f_s	50	kHz
Number of rolling elements	n_b	18	$-$
Simulation time span	t_{span}	2	s
Pitch diameter	D	250	mm
Rolling ball diameter	d	20	mm
Shaft rotational speed	f_{shaft}	50	Hz
Contact angle	$\alpha_r b$	15	$degrees$
Impulse resonance frequency	$f_{n,imp}$	3	kHz
Rolling slip	β_{rb}	2	%

are of high amplitude, as compared to Sim B. This means that damage can be more easily detected in Sim C, which is also evident by looking at the time-domain plots. In Sim B, it can be observed that the dB scale and the cursors placed at the mean of BPFO in the power spectrum, help in revealing and identifying the characteristic damage frequency. The sidebands (due to amplitude modulation) around the impulse resonance of 3 kHz clearly indicate that damage exists on the outer race as the spacing is exactly 1×BPFO, 2×BPFO, and so on. However, as shown in the spectrum, the fundamental shaft rotation and its five harmonics are also present in the lower frequency regions, and whose spectra are non-synchronous to the BPFO. Hence, there is a need to eliminate those additional frequency components from the spectrum and at the same time amplify the signal of the bearing fault, to prevent potential masking effects, e.g. from additional white noise added to the system, in order to successfully identify damage at its early stages.

E.2 Envelope analysis on the simulated example

As part of the EA procedure (see Chapter 2), a pre-whitening step is required for removing periodic components in a signal, e.g. gear-pairs. One approach that has been used in the literature is the SANC algorithm (see Appendix B for a brief explanation of this algorithm). However, the success of this procedure is influenced heavily on the choice of its parameters, namely filter length/order, and delay of the samples.

Instead, with the aid of an *AR model* and by assuming an underlying stationary process within a short time-scale, e.g. of 2 seconds, the deterministic part of the vibration signal was estimated. Subsequently, subtracting it from the original signal, yields the residual signal $\varepsilon_{\ddot{x}}$. The coefficients of an AR model can be computed using the well-known Yule-Walker equations by means of the Levinson-Durbin recursion algorithms. This is, in practise, done

using MATLAB built-in function *lpc.m* to obtain the coefficients and then filter the signal using the function *fftfilt.m* to obtain the residuals. The number of coefficients, or the order p of the AR model / FIR filter, is chosen such that the kurtosis value of the AR residual signal is maximized, as in Equation 2.23. This yields $p = 110$, since $p_{max} \approx 120$, in order to avoid including the BPFO impulses as part of the AR model, as explained in Chapter 2. The AR residual signal contains both white noise and the impulses at BPFO. To further enhance the impulsiveness of the signal, another FIR filter is designed as in [160], i.e. the MED filter (see Chapter 4 for further details on this technique). The PSD plots, shown in Figure E.3, illustrate the importance of the MED (shown as the function $g(.)$ in the plots). In particular, MED amplifies the vibration impulses at BPFO and harmonics, which in turn reduces the importance of other components, e.g. the shaft rotation and harmonics. The MED is applied to both the pre-processed AR residual signal, i.e. $g(\varepsilon_{\ddot{x}})$ and the "raw" vibration signal, i.e. $g(\ddot{x})$. Clearly, applying MED on the AR residual $g(\varepsilon_{\ddot{x}})$ both the impulsiveness at BPFO can be revealed and the discrete frequencies, not related to REB damage, will be reduced.

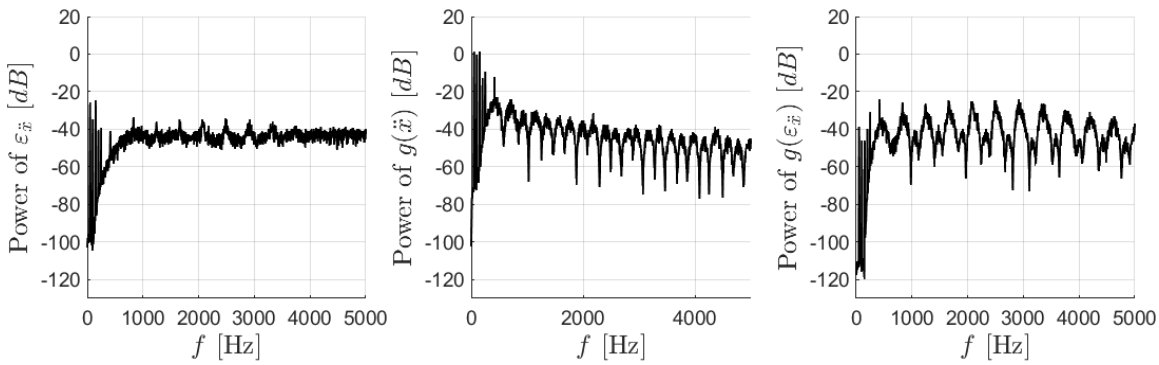

Figure E.3. Power spectra of pre-processed acceleration signal: AR residual (left), MED applied on the raw signal (middle) and MED applied on the AR residual (right). It particularly shows the reduction of discrete low frequency components not associated with BPFO when MED is applied and the enhancement of impulses at BPFO when the AR residual is considered instead of the raw signal.

In the next stage of the analysis, the SK is computed for the pre-processed signal $g(\varepsilon_{\ddot{x}})$ and the corresponding kurtogram is shown in Figure E.4. For comparison purposes, the kurtogram is also shown for the raw signal $\ddot{x}$, its AR residual $\varepsilon_{\ddot{x}}$ and its deconvolved (using MED) version $g(\varepsilon_{\ddot{x}})$. Interestingly, for $g(\varepsilon_{\ddot{x}})$ the calculated SK suggests an optimal filter band of $[0, 12.5]$ kHz using $2^{k=1}$ spectral bands, i.e. level of decomposition $k = 1$. It is possible that the highest SK value is calculated when including the *full harmonic series* of the BPFO and not only the frequency band around the impulse resonance of 3 kHz, as it

was originally expected. On the other hand, the maximum value of SK for both $\ddot{x}$) and $\varepsilon_{\ddot{x}}$ is calculated on a narrow range of $[21.827, 21.875]$ kHz using $2^{k=10}$ spectral bands, while for $\varepsilon_{\ddot{x}}$ the corresponding spectral band is $[0, 8.333]$ kHz at decomposition level $k = 2.6$.

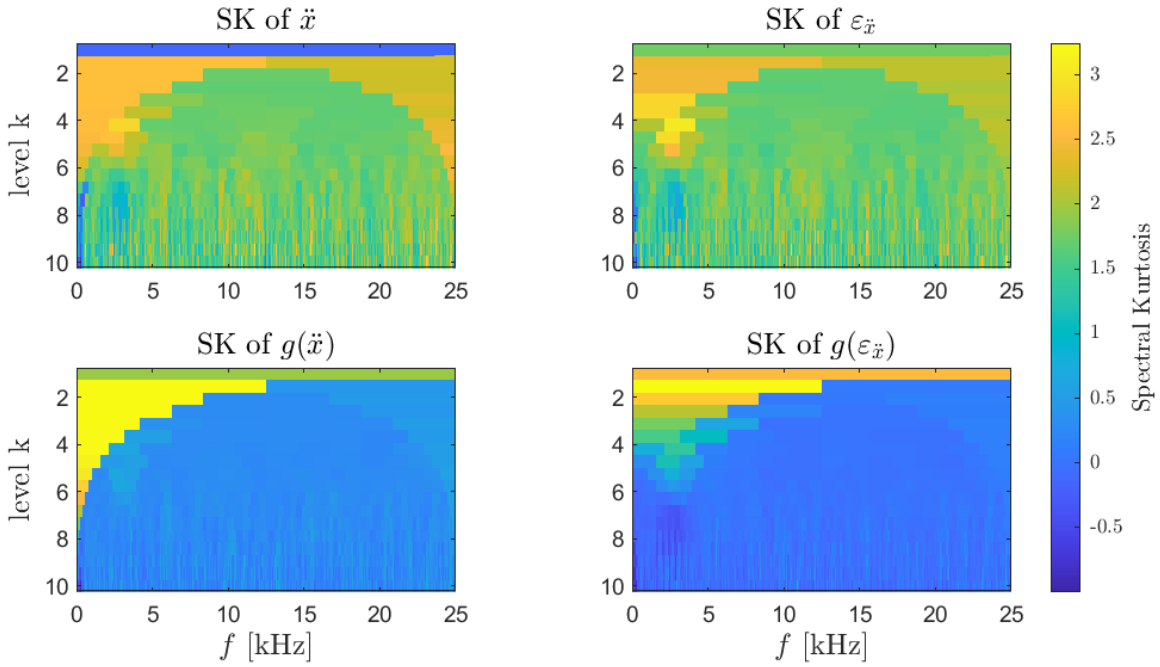

Figure E.4. Kurtogram up to level $k = 10$ calculated on four signals: the raw acceleration signal (top-left), the AR residual (top-right), the signal pre-processed by MED only (bottom-left) and the AR residual signal pre-processed by MED (bottom-right). The maximum SK calculated on the four signals shows us same frequency bandwidths between $\ddot{x}$ and $\varepsilon_{\ddot{x}}$, but, different ones between the other two signals.

E.3 Outer race damage on the WT-REB at damage level 5 a simulated example

Note that at significant damage levels, the EA will reduce the amplitudes of the envelope spectrum at BPFO, as shown in Figure E.5.

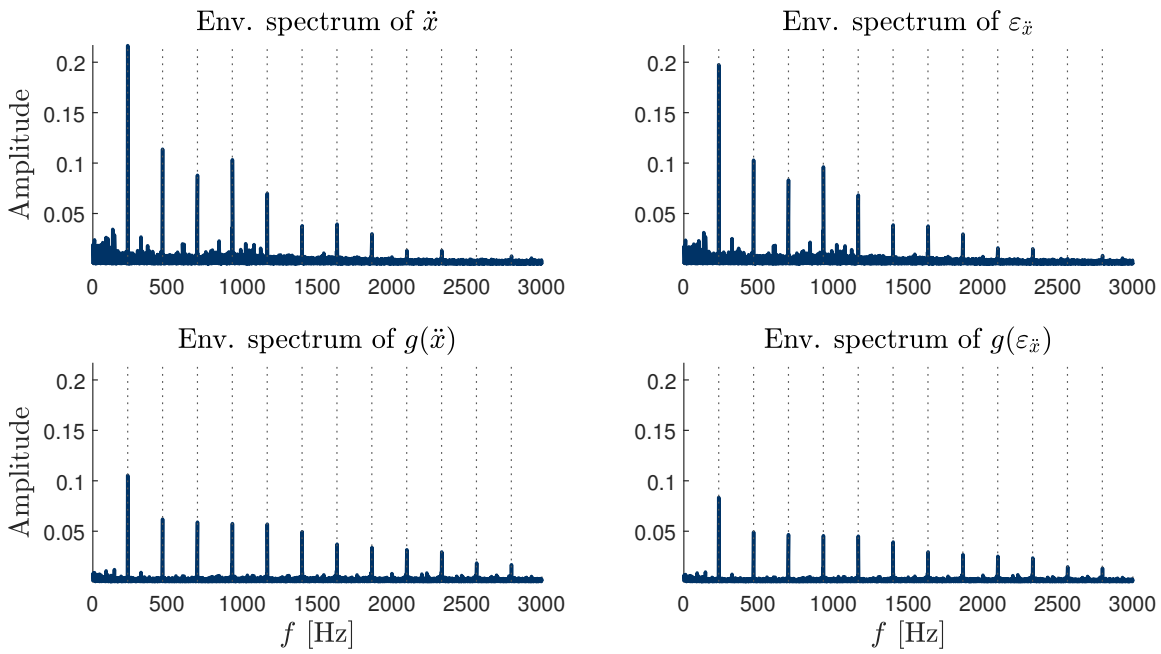

Figure E.5. Envelope spectrum of damage level 5.

Bibliography

[1] Adaptive local iterative filtering. http://people.disim.univaq.it/~antonio.cicone/Software.html. Accessed: July 2022.

[2] Advanced wind turbine drivetrain trends and opportunities. https://www.energy.gov/eere/articles/advanced-wind-turbine-drivetrain-trends-and-opportunities. Accessed: July 2019.

[3] Ceemdan codes. http://bioingenieria.edu.ar/grupos/ldnlys/metorres/. Accessed July 2022.

[4] cellwise: Analyzing data with cellwise outliers. https://CRAN.R-project.org/package=cellWise. Accessed: July 2021.

[5] Flexible statistics and data analysis (fsda). http://rosa.unipr.it/fsda.html. Accessed: July 2022.

[6] Ge9x recognized as most powerful jet engine in the world. https://www.aerospacetestinginternational.com/videos/ge9x-recognised-as-most-powerful-jet-engine-in-the-world.html. Accessed: July 2019.

[7] Natural excitation technique. https://windengineeringuis.github.io/. Accessed: July 2022.

[8] Robust empirical mode decomposition. https://www.mathworks.com/matlabcentral/fileexchange/70032-robust-empirical-mode-decomposition-remd. Accessed: July 2022.

[9] Robust@leuven. https://wis.kuleuven.be/statdatascience/robust. Accessed: July 2022.

[10] Rolling element bearing testing. https://www.abdynamics.com/en/applications/production-testing/rolling-element-bearing-testing. Accessed: July 2019.

[11] Sharc. https://docs.hpc.shef.ac.uk/en/latest/sharc/. Accessed: May 2022.

[12] Statistics show bearing problems cause the majority of wind turbine gearbox failures. https://www.energy.gov/eere/wind/articles/statistics-show-bearing-problems-cause-majority-wind-turbine-gearbox-failures. Accessed: July 2019.

[13] Time-varying filtering based emd. https://uk.mathworks.com/matlabcentral/fileexchange/63300-time-varying-filter-based-empirical-mode-decomposition-tvf-emd. Accessed: July 2022.

[14] Top 10: World's longest bridges. https://www.sciencefocus.com/science/top-10-worlds-longest-bridges/. Accessed: July 2022.

[15] The world's 10 toughest buildings. https://www.popularmechanics.com/technology/infrastructure/g2237/worlds-10-toughest-buildings/?slide=5. Accessed: July 2019.

[16] Z24 bridge benchmark. https://bwk.kuleuven.be/bwm/z24. Accessed: July 2022.

[17] Yaser S Abu-Mostafa, Malik Magdon-Ismail, and Hsuan-Tien Lin. *Learning from data*, volume 4. AMLBook New York, NY, USA:, 2012.

[18] Charu Aggarwal. *Outlier Analysis*. Springer, 2017.

[19] Basim Al-Najjar and Imad Alsyouf. Enhancing a company's profitability and competitiveness using integrated vibration-based maintenance: A case study. *European journal of operational research*, 157(3):643–657, 2004.

[20] S Alampalli. Influence of in-service environment on modal parameters. In *Proceedings-SPIE the international society for optical engineering*, volume 1, pages 111–116. Citeseer, 1998.

[21] Xueli An and Yongjun Tang. Application of variational mode decomposition energy distribution to bearing fault diagnosis in a wind turbine. *Transactions of the Institute of Measurement and Control*, 39(7):1000–1006, 2017.

[22] Jérôme Antoni. Blind separation of vibration components: Principles and demonstrations. *Mechanical systems and signal processing*, 19(6):1166–1180, 2005.

[23] Jerome Antoni. Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21(1):108–124, 2007.

[24] I Antoniadis and G Glossiotis. Cyclostationary analysis of rolling-element bearing vibration signals. *Journal of sound and vibration*, 248(5):829–845, 2001.

[25] Ifigeneia Antoniadou. *Accounting for nonstationarity in the condition monitoring of wind turbine gearboxes*. PhD thesis, University of Sheffield, 2013.

[26] Abdollah Bagheri, Osman E Ozbulut, and Devin K Harris. Structural system identification based on variational mode decomposition. *Journal of Sound and Vibration*, 417:182–197, 2018.

[27] Claudia Becker, Roland Fried, and Sonja Kuhnt. *Robustness and complex data structures: festschrift in honour of Ursula Gather*. Springer Science & Business Media, 2014.

[28] Inigo Bediaga, Xabier Mendizabal, Aitor Arnaiz, and Jokin Munoa. Ball bearing damage detection using traditional signal processing algorithms. *IEEE Instrumentation & Measurement Magazine*, 16(2):20–25, 2013.

[29] Julius S Bendat and Allan G Piersol. Random data analysis and measurement procedures, 2000.

[30] Christopher M. Bishop. Pattern recognition and machine learning (information science and statistics), 2006.

[31] Boualem Boashash. Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals. *Proceedings of the IEEE*, 80(4):520–538, 1992.

[32] Abdenour Bounsiar and Michael G Madden. One-class support vector machines revisited. In *2014 International Conference on Information Science & Applications (ICISA)*, pages 1–4. IEEE, 2014.

[33] S Braun. Rb randall, vibration based condition monitoring, john wiley and sons, chichester, uk (2011), 2011.

[34] S Braun and M Feldman. Time–frequency characteristics of non-linear systems. *Mechanical systems and signal processing*, 11(4):611–620, 1997.

[35] Rune Brincker. Some elements of operational modal analysis. *Shock and Vibration*, 2014, 2014.

[36] Dan Burnes and Alejandro Camou. Impact of fuel composition on gas turbine engine performance. *Journal of Engineering for Gas Turbines and Power*, 141(10):101006, 2019.

[37] JC Butcher. On fifth order runge-kutta methods. *BIT Numerical Mathematics*, 35(2):202–209, 1995.

[38] Juan Martin Caicedo, Shirley J Dyke, and Erik A Johnson. Natural excitation technique and eigensystem realization algorithm for phase i of the iasc-asce benchmark problem: Simulated data. *Journal of Engineering Mechanics*, 130(1):49–60, 2004.

[39] Rhara Cardoso, Alexandre Cury, and Flávio Barbosa. A robust methodology for modal parameters estimation applied to shm. *Mechanical Systems and Signal Processing*, 95:24–41, 2017.

[40] Peter Cawley. Structural health monitoring: Closing the gap between research and industrial deployment. *Structural Health Monitoring*, 17(5):1225–1244, 2018.

[41] Soumya Chatterjee. Detection of focal electroencephalogram signals using higher-order moments in emd-tkeo domain. *Healthcare technology letters*, 6(3):64–69, 2019.

[42] J Chen, You Lin Xu, and RC Zhang. Modal parameter identification of tsing ma suspension bridge under typhoon victor: Emd-ht method. *Journal of Wind Engineering and Industrial Aerodynamics*, 92(10):805–827, 2004.

[43] Zhenghua Chen, Le Zhang, Zhiguang Cao, and Jing Guo. Distilling the knowledge from handcrafted features for human activity recognition. *IEEE Transactions on Industrial Informatics*, 14(10):4334–4342, 2018.

[44] Antonio Cicone, Jingfang Liu, and Haomin Zhou. Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis. *Applied and Computational Harmonic Analysis*, 41(2):384–411, 2016.

[45] David A Clifton, Peter R Bannister, and Lionel Tarassenko. A framework for novelty detection in jet engine vibration data. In *Key engineering materials*, volume 347, pages 305–310. Trans Tech Publ, 2007.

[46] Lei Clifton, David A Clifton, Yang Zhang, Peter Watkinson, Lionel Tarassenko, and Hujun Yin. Probabilistic novelty detection with support vector machines. *IEEE Transactions on Reliability*, 63(2):455–467, 2014.

[47] Lei A Clifton, Hujun Yin, David A Clifton, and Yang Zhang. Combined support vector novelty detection for multi-channel combustion data. In *2007 IEEE International Conference on Networking, Sensing and Control*, pages 495–500. IEEE, 2007.

[48] Albert Cohen and Nira Dyn. Nonstationary subdivision schemes, multiresolution analysis, and wavelet packets. In *Wavelet Analysis and Its Applications*, volume 7, pages 189–200. Elsevier, 1998.

[49] Marcelo A Colominas, Gastón Schlotthauer, and María E Torres. Improved complete ensemble emd: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14:19–29, 2014.

[50] Joelle Courrech and Mark Gaudet. Envelope analysis-the key to rolling-element bearing diagnosis. *Brüel & Kjær Application Notes*, 1998.

[51] Juan Antonio Cuesta-Albertos, Alfonso Gordaliza, and Carlos Matrán. Trimmed k-means: an attempt to robustify quantizers. *The Annals of Statistics*, 25(2):553–576, 1997.

[52] FE Culick and Paul Kuentzmann. Unsteady motions in combustion chambers for propulsion systems. Technical report, NATO Research and Technology Organization Neuilly-Sur-Seine (France), 2006.

[53] Cuong Dao, Behzad Kazemtabrizi, and Christopher Crabtree. Wind turbine reliability data review and impacts on levelised cost of energy. *Wind Energy*, 22(12):1848–1871, 2019.

[54] Mark S Darlow, Robert H Badgley, and GW Hogg. Application of high-frequency resonance techniques for bearing diagnostics in helicopter gearboxes. Technical report, Mechanical Technology Inc Latham NY, 1974.

[55] Werner Daum. Guidelines for structural health monitoring. In *Handbook of technical diagnostics*, pages 539–541. Springer, 2013.

[56] Somnath De, Arijit Bhattacharya, Sirshendu Mondal, Achintya Mukhopadhyay, and Swarnendu Sen. Investigation of flame behavior and dynamics prior to lean blowout in a combustor with varying mixedness of reactants for the early detection of lean blowout. *International Journal of Spray and Combustion Dynamics*, 11:1756827718812519, 2019.

[57] Alexandre de Juniac. International air transport association annual review, 2019.

[58] Arnaud Deraemaeker and Keith Worden. A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mechanical systems and signal processing*, 105:1–15, 2018.

[59] N Dervilis, M Choi, I Antoniadou, KM Farinholt, SG Taylor, RJ Barthorpe, G Park, K Worden, and CR Farrar. Novelty detection applied to vibration data from a cx-100 wind turbine blade under fatigue loading. In *Journal of Physics: Conference Series*, volume 382, page 012047. IOP Publishing, 2012.

[60] N Dervilis, K Worden, and EJ Cross. On robust regression analysis as a means of exploring environmental and operational conditions for shm data. *Journal of Sound and Vibration*, 347:279–296, 2015.

[61] Nikolaos Dervilis, Ifigeneia Antoniadou, Robert J Barthorpe, Elizabeth J Cross, and Keith Worden. Robust methods for outlier detection and regression for shm applications. *International Journal of Sustainable Materials and Structural Systems*, 2(1/2), 2016.

[62] Scott W Doebling, Charles R Farrar, and Randall S Goodman. Effects of measurement statistics on the detection of damage in the alamosa canyon bridge. Technical report, Los Alamos National Lab., NM (United States), 1996.

[63] Konstantin Dragomiretskiy and Dominique Zosso. Variational mode decomposition. *IEEE transactions on signal processing*, 62(3):531–544, 2013.

[64] H Endo and RB 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(2):906–919, 2007.

[65] Alireza Entezami, Hassan Sarmadi, Masoud Salar, Carlo De Michele, and Ali Nadir Arslan. A novel data-driven method for structural health monitoring under ambient vibration and high-dimensional features by robust multidimensional scaling. *Structural Health Monitoring*, page 1475921720973953, 2021.

[66] Ehsan Tarkesh Esfahani, Shaocheng Wang, and V Sundararajan. Multisensor wireless system for eccentricity and bearing fault detection in induction motors. *IEEE/ASME Transactions on Mechatronics*, 19(3):818–826, 2013.

[67] Charles R Farrar and Keith Worden. *Structural health monitoring: a machine learning perspective*. John Wiley & Sons, 2012.

[68] Michael Feldman. Hilbert transform in vibration analysis. *Mechanical systems and signal processing*, 25(3):735–802, 2011.

[69] Yanhui Feng, Yingning Qiu, Christopher J Crabtree, Hui Long, and Peter J Tavner. Monitoring wind turbine gearboxes. *Wind Energy*, 16(5):728–740, 2013.

[70] Zhipeng Feng, Dong Zhang, and Ming J Zuo. Adaptive mode decomposition methods and their applications in signal analysis for machinery fault diagnosis: a review with examples. *IEEE access*, 5:24301–24331, 2017.

[71] Eloi Figueiredo, Gyuhae Park, Charles R Farrar, Keith Worden, and Joaquim Figueiras. Machine learning algorithms for damage detection under operational and environmental variability. *Structural Health Monitoring*, 10(6):559–572, 2011.

[72] Eloi Figueiredo, Gyuhae Park, Joaquim Figueiras, Charles Farrar, and Keith Worden. Structural health monitoring algorithm comparisons using standard data sets. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2009.

[73] Eloi Figueiredo, Lucian Radu, Keith Worden, and Charles R Farrar. A bayesian approach based on a markov-chain monte carlo method for damage detection under unknown sources of variability. *Engineering Structures*, 80:1–10, 2014.

[74] Eloi Figueiredo and Adam Santos. Machine learning algorithms for damage detection. In *Vibration-Based Techniques for Damage Detection and Localization in Engineering Structures*, pages 1–39. World Scientific, 2018.

[75] Wayne A Fuller. *Introduction to statistical time series*. John Wiley & Sons, 2009.

[76] María Teresa Gallegos and Gunter Ritter. A robust method for cluster analysis. *The Annals of Statistics*, 33(1):347–380, 2005.

[77] John Cristian Borges Gamboa. Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*, 2017.

[78] Luis A García-Escudero, Alfonso Gordaliza, Carlos Matrán, and Agustin Mayo-Iscar. A general trimming approach to robust cluster analysis. *The Annals of Statistics*, 36(3):1324–1345, 2008.

[79] Anthimos Georgiadis, Xiaoyun Gong, and Nicolas Meier. Vibration analysis based on the spectrum kurtosis for adjustment and monitoring of ball bearing radial clearance. In *MATEC Web of Conferences*, volume 211, page 06006. EDP Sciences, 2018.

[80] Fahit Gharibnezhad, Luis E Mujica, and José Rodellar. Applying robust variant of principal component analysis as a damage detector in the presence of outliers. *Mechanical Systems and Signal Processing*, 50:467–479, 2015.

[81] Damian A Gonzalez and John C Mccall. Design of filters to reduce harmonic distortion in industrial power systems. *IEEE Transactions on Industry Applications*, (3):504–511, 1987.

[82] Hiroshi Gotoda, Hiroyuki Nikimoto, Takaya Miyano, and Shigeru Tachibana. Dynamic properties of combustion instability in a lean premixed gas-turbine combustor. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(1):013124, 2011.

[83] Jianfeng Gu, Mustafa Gul, and Xiaoguang Wu. Damage detection under varying temperature using artificial neural networks. *Structural Control and Health Monitoring*, 24(11):e1998, 2017.

[84] Valérie Haggan and Tohru Ozaki. Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, 68(1):189–196, 1981.

[85] Douglas M Hawkins. *Identification of outliers*, volume 11. Springer, 1980.

[86] Paul Hayton, Bernhard Schölkopf, Lionel Tarassenko, and Paul Anuzis. Support vector novelty detection applied to jet engine vibration spectra. *Advances in neural information processing systems*, 13, 2000.

[87] Rongrong Hou, Yong Xia, and Xiaoqing Zhou. Structural damage detection based on l1 regularization using natural frequencies and mode shapes. *Structural Control and Health Monitoring*, 25(3):e2107, 2018.

[88] D Howieson. Vibration monitoring: Envelope signal processing. *Diagnostic instruments, SKF Reliability Systems*, 2003.

[89] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry 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(1971):903–995, 1998.

[90] Norden E Huang, Zhaohua Wu, Steven R Long, Kenneth C Arnold, Xianyao Chen, and Karin Blank. On instantaneous frequency. *Advances in adaptive data analysis*, 1(02):177–229, 2009.

[91] Norden Eh Huang. *Hilbert-Huang transform and its applications*, volume 16. World Scientific, 2014.

[92] Peter J Huber. *Robust statistics*, volume 523. John Wiley & Sons, 2004.

[93] M Hubert, PJ Rousseeuw, D Vanpaemel, and T Verdonck. A deterministic algorithm for s-estimators and mm-estimators of multivariate location and scatter, 2013.

[94] Mia Hubert and Michiel Debruyne. Minimum covariance determinant. *Wiley interdisciplinary reviews: Computational statistics*, 2(1):36–43, 2010.

[95] Mia Hubert, Peter J Rousseeuw, and Karlien Vanden Branden. Robpca: a new approach to robust principal component analysis. *Technometrics*, 47(1):64–79, 2005.

[96] Mia Hubert, Peter J Rousseeuw, and Tim Verdonck. A deterministic algorithm for robust location and scatter. *Journal of Computational and Graphical Statistics*, 21(3):618–637, 2012.

[97] Li Jingsong, Liu Quan, and Su Hang. The study of the intermittency test algorithm to eliminate mode mixing. In *2011 4th International Congress on Image and Signal Processing*, volume 5, pages 2384–2387. IEEE, 2011.

[98] John Johnston and John DiNardo. Econometric methods. 1963.

[99] Benjamin Kedem and Konstantinos Fokianos. *Regression models for time series analysis*. John Wiley & Sons, 2005.

[100] T Kijewski and A Kareem. Wavelet transforms for system identification in civil engineering. *Computer-Aided Civil and Infrastructure Engineering*, 18(5):339–355, 2003.

[101] Stephen P King, Andrew R Mills, Visakan Kadirkamanathan, and Dave A Clifton. *Equipment health monitoring in complex systems*. Artech House, 2017.

[102] Hans Konstantin-Hansen and Henrik Herlufsen. Envelope and cepstrum analyses for machinery fault identification. *Sound and Vibration*, 44(5):10, 2010.

[103] Horst Kuemmlee, Thomas Gross, and Josef Kolerus. Machine vibrations and diagnostics the world of iso. In *Industry Applications Society 60th Annual Petroleum and Chemical Industry Conference*, pages 1–13. IEEE, 2013.

[104] Jyrki Kullaa. Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mechanical Systems and Signal Processing*, 25(8):2976–2989, 2011.

[105] Bibin Kurian and Ranjith Liyanapathirana. Machine learning techniques for structural health monitoring. In *Proceedings of the 13th International Conference on Damage Assessment of Structures*, pages 3–24. Springer, 2020.

[106] Denis Kwiatkowski, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178, 1992.

[107] SJ Lacey. Using vibration analysis to detect early failure of bearings. *Insight-Northampton-Including European Issues*, 49(8):444–446, 2007.

[108] K Lakshmi and A Rama Mohan Rao. A robust damage-detection technique with environmental variability combining time-series models with principal components. *Nondestructive Testing and Evaluation*, 29(4):357–376, 2014.

[109] Heng Li, Zhi Li, and Wei Mo. A time varying filter approach for empirical mode decomposition. *Signal Processing*, 138:146–158, 2017.

[110] Xiao Li and QS Li. Monitoring structural performance of a supertall building during 14 tropical cyclones. *Journal of Structural Engineering*, 144(10):04018176, 2018.

[111] Zhijun Li, Maria Q Feng, Longxi Luo, Dongming Feng, and Xiuli Xu. Statistical analysis of modal parameters of a suspension bridge based on bayesian spectral density approach and shm data. *Mechanical Systems and Signal Processing*, 98:352–367, 2018.

[112] Tim C Lieuwen. *Unsteady combustor physics*. Cambridge University Press, 2012.

[113] Timothy C Lieuwen and Vigor Yang. *Combustion instabilities in gas turbine engines: operational experience, fundamental mechanisms, and modeling*. American Institute of Aeronautics and Astronautics, 2005.

[114] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *2008 eighth ieee international conference on data mining*, pages 413–422. IEEE, 2008.

[115] Zhiliang Liu, Dandan Peng, Ming J Zuo, Jianshuo Xia, and Yong Qin. Improved hilbert–huang transform with soft sifting stopping criterion and its application to fault diagnosis of wheelset bearings. *ISA transactions*, 125:426–444, 2022.

[116] Patrick J Loughlin and Keith L Davidson. Instantaneous bandwidth of multicomponent signals. In *Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, volume 3807, pages 610–624. International Society for Optics and Photonics, 1999.

[117] Stephane G Mallat. *A wavelet tour of signal processing.* Elsevier, 1999.

[118] Danilo P Mandic, Naveed ur Rehman, Zhaohua Wu, and Norden E Huang. Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis. *IEEE signal processing magazine*, 30(6):74–86, 2013.

[119] Ricardo A Maronna, R Douglas Martin, Victor J Yohai, and Matías Salibián-Barrera. *Robust statistics: theory and methods (with R).* John Wiley & Sons, 2019.

[120] Ricardo A Maronna and Ruben H Zamar. Robust estimates of location and dispersion for high-dimensional datasets. *Technometrics*, 44(4):307–317, 2002.

[121] Lawrence Marple. Computing the discrete-time" analytic" signal via fft. *IEEE Transactions on signal processing*, 47(9):2600–2603, 1999.

[122] Gabriele Marrongelli, Carmelo Gentile, and Antonella Saisi. Anomaly detection based on automated oma and mode shape changes: application on a historic arch bridge. In *International Conference on Arch Bridges*, pages 447–455. Springer, 2019.

[123] Ioannis Matthaiou. Gas turbine engine condition monitoring and performance prediction. Technical report, The University of Sheffield, Sheffield, UK, 2017.

[124] Ioannis Matthaiou, Bhupendra Khandelwal, and Ifigeneia Antoniadou. Vibration monitoring of gas turbine engines: machine-learning approaches and their challenges. *Frontiers in Built Environment*, 3:54, 2017.

[125] Geoff L McDonald and Qing Zhao. Multipoint optimal minimum entropy deconvolution and convolution fix: Application to vibration fault detection. *Mechanical Systems and Signal Processing*, 82:461–477, 2017.

[126] PD McFadden and JD Smith. Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review. *Tribology international*, 17(1):3–10, 1984.

[127] Pietro Milillo, Giorgia Giardina, Daniele Perissin, Giovanni Milillo, Alessandro Coletta, and Carlo Terranova. Pre-collapse space geodetic observations of critical infrastructure: the morandi bridge, genoa, italy. *Remote Sensing*, 11(12):1403, 2019.

[128] T Mimura and Akira Mita. Automatic estimation of natural frequencies and damping ratios of building structures. *Procedia Engineering*, 188:163–169, 2017.

[129] Bijan Mohammadi and Olivier Pironneau. Analysis of the k-epsilon turbulence model. 1993.

[130] Amiya Ranjan Mohanty. *Machinery Condition Monitoring: Principles and Practices.* CRC Press, 2014.

[131] Luis Carlos Molina, Lluís Belanche, and Àngela Nebot. Feature selection algorithms: A survey and experimental evaluation. In *2002 IEEE International Conference on Data Mining, 2002. Proceedings.*, pages 306–313. IEEE, 2002.

[132] Douglas C Montgomery, Cheryl L Jennings, and Murat Kulahci. *Introduction to time series analysis and forecasting.* John Wiley & Sons, 2015.

[133] Maurizio Morgese, Farhad Ansari, Marco Domaneschi, and Gian Paolo Cimellaro. Post-collapse analysis of morandi's polcevera viaduct in genoa italy. *Journal of Civil Structural Health Monitoring*, 10(1):69–85, 2020.

[134] Francesco Mosti, Giuseppe Quaranta, and Walter Lacarbonara. Numerical and experimental assessment of the modal curvature method for damage detection in plate structures. In *Structural nonlinear dynamics and diagnosis*, pages 59–68. Springer, 2015.

[135] Bruce Roy Munson, Theodore Hisao Okiishi, Wade W Huebsch, and Alric P Rothmayer. *Fluid mechanics.* Wiley Singapore, 2013.

[136] Vineeth Nair and RI Sujith. Multifractality in combustion noise: predicting an impending combustion instability. *Journal of Fluid Mechanics*, 747:635–655, 2014.

[137] Mojtaba Nazari and Sayed Mahmoud Sakhaei. Successive variational mode decomposition. *Signal Processing*, 174:107610, 2020.

[138] Ana C Neves, Ignacio González, Raid Karoumi, and John Leander. The influence of frequency content on the performance of artificial neural network–based damage detection systems tested on numerical and experimental bridge data. *Structural Health Monitoring*, 20(3):1331–1347, 2021.

[139] Jiaqi Pan, Yan Zhuang, and Simon Fong. The impact of data normalization on stock market prediction: using svm and technical indicators. In *International Conference on Soft Computing in Data Science*, pages 72–88. Springer, 2016.

[140] Diego Galar Pascual. *Artificial intelligence tools: decision support systems in condition monitoring and diagnosis.* CRC Press, 2015.

[141] B PEETERS. System identification and damage detection in civil engineering. *PhD Dissertation, Katholieke University*, 2000.

[142] Bart Peeters and Guido De Roeck. One-year monitoring of the z24-bridge: environmental effects versus damage events. *Earthquake engineering & structural dynamics*, 30(2):149–171, 2001.

[143] Bart Peeters and Guido De Roeck. Stochastic system identification for operational modal analysis: a review. *J. Dyn. Sys., Meas., Control*, 123(4):659–667, 2001.

[144] Mirko Piersanti, Massimo Materassi, Antonio Cicone, Luca Spogli, Haomin Zhou, and Rodolfo Gerardo Ezquer. Adaptive local iterative filtering: A promising technique for the analysis of nonstationary signals. *Journal of Geophysical Research: Space Physics*, 123(1):1031–1046, 2018.

[145] Marco AF Pimentel, David A Clifton, Lei Clifton, and Lionel Tarassenko. A review of novelty detection. *Signal Processing*, 99:215–249, 2014.

[146] Theodor D Popescu. Blind separation of vibration signals and source change detection–application to machine monitoring. *Applied Mathematical Modelling*, 34(11):3408–3421, 2010.

[147] Maurice Bertram Priestley. *Spectral analysis and time series: probability and mathematical statistics.* Number 04; QA280, P7. 1981.

[148] RB Randall, N Sawalhi, and M Coats. A comparison of methods for separation of deterministic and random signals. *International Journal of Condition Monitoring*, 1(1):11–19, 2011.

[149] Robert B Randall and Jerome Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical systems and signal processing*, 25(2):485–520, 2011.

[150] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer, 2003.

[151] Jakob Raymaekers and Peter J Rousseeuw. Transforming variables to central normality. *Machine Learning*, pages 1–23, 2021.

[152] Edwin Reynders, Gersom Wursten, and Guido De Roeck. Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. *Structural Health Monitoring*, 13(1):82–93, 2014.

[153] Raimund Rolfes, Stephan Zerbst, G Haake, Jonannes Reetz, and Jerome P Lynch. Integral shm-system for offshore wind turbines using smart wireless sensors. In *Proceedings of the 6th International Workshop on Structural Health Monitoring*, volume 200, pages 11–13. Citeseer, 2007.

[154] Peter Rousseeuw, Domenico Perrotta, Marco Riani, and Mia Hubert. Robust monitoring of time series with application to fraud detection. *Econometrics and statistics*, 9:108–121, 2019.

[155] Peter J Rousseeuw and Katrien Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3):212–223, 1999.

[156] Peter J Rousseeuw and Mia Hubert. Robust statistics for outlier detection. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 1(1):73–79, 2011.

[157] A Rytter. Vibration based inspection of civil engineering structures ph. d. *Aalborg University, Aalborg, Denmark*, 1993.

[158] Ashwin A Salvi, Dennis Assanis, and Zoran Filipi. Impact of physical and chemical properties of alternative fuels on combustion, gaseous emissions, and particulate matter during steady and transient engine operation. *Energy & fuels*, 26(7):4231–4241, 2012.

[159] Adam Santos, Eloi Figueiredo, MFM Silva, CS Sales, and JCWA Costa. Machine learning algorithms for damage detection: Kernel-based approaches. *Journal of Sound and Vibration*, 363:584–599, 2016.

[160] Nader Sawalhi, RB Randall, and Hiroaki 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(6):2616–2633, 2007.

[161] Matti Scheu, Denis Matha, Matthias Hofmann, and Michael Muskulus. Maintenance strategies for large offshore wind farms. *Energy Procedia*, 24:281–288, 2012.

[162] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, 2001.

[163] Joachim Schreurs, Iwein Vranckx, Mia Hubert, Johan AK Suykens, and Peter J Rousseeuw. Outlier detection in non-elliptical data by kernel mrcd. *Statistics and Computing*, 31(5):1–18, 2021.

[164] Shuangwen Sheng. Wind turbine gearbox condition monitoring round robin study-vibration analysis. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

[165] Haichen Shi, Keith Worden, and Elizabeth J Cross. A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. *Mechanical Systems and Signal Processing*, 103:381–397, 2018.

[166] Kihong Shin and Joseph Hammond. *Fundamentals of signal processing for sound and vibration engineers*. John Wiley & Sons, 2008.

[167] Robert H Shumway, David S Stoffer, and David S Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.

[168] Dalwinder Singh and Birmohan Singh. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97:105524, 2020.

[169] Hoon Sohn. Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):539–560, 2007.

[170] Hoon Sohn, David W Allen, Keith Worden, and Charles R Farrar. Structural damage classification using extreme value statistics. *J. Dyn. Sys., Meas., Control*, 127(1):125–132, 2005.

[171] Abdenour Soualhi, Kamal Medjaher, and Noureddine Zerhouni. Bearing health monitoring based on hilbert–huang transform, support vector machine, and regression. *IEEE Transactions on instrumentation and measurement*, 64(1):52–62, 2014.

[172] Sauro Succi. Computational fluid dynamics (ac274). 2015.

[173] Z Sun and CC Chang. Structural damage assessment based on wavelet packet transform. *Journal of structural engineering*, 128(10):1354–1361, 2002.

[174] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.

[175] A Tabrizi, L Garibaldi, A Fasana, and S Marchesiello. Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine. *Meccanica*, 50(3):865–874, 2015.

[176] Lionel Tarassenko, David A Clifton, Peter R Bannister, Steve King, and Dennis King. Novelty detection. *Encyclopedia of structural health monitoring*, 2009.

[177] Timken Team. Bearing damage analysis reference guide. Technical report, The Timken Company, 2020.

[178] María E Torres, Marcelo A Colominas, Gastón Schlotthauer, and Patrick Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. In *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 4144–4147. IEEE, 2011.

[179] Kostas Triantafyllopoulos and Giovanni Montana. Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, 8(1):23–49, 2011.

[180] Ruey S Tsay and Rong Chen. *Nonlinear time series analysis*, volume 891. John Wiley & Sons, 2018.

[181] Eugene Tuv, Alexander Borisov, George Runger, and Kari Torkkola. Feature selection with ensembles, artificial variables, and redundancy elimination. *The Journal of Machine Learning Research*, 10:1341–1366, 2009.

[182] Naveed ur Rehman and Hania Aftab. Multivariate variational mode decomposition. *IEEE Transactions on Signal Processing*, 67(23):6039–6052, 2019.

[183] Laurens Van Der Maaten, Eric Postma, Jaap Van den Herik, et al. Dimensionality reduction: a comparative. *J Mach Learn Res*, 10(66-71):13, 2009.

[184] Wojciech Wawrzynski. Duffing-type oscillator under harmonic excitation with a variable value of excitation amplitude and time-dependent external disturbances. *Scientific Reports*, 11(1):1–15, 2021.

[185] Ralph A Wiggins. Minimum entropy deconvolution. *Geoexploration*, 16(1-2):21–35, 1978.

[186] MLD Wong, LB Jack, and AK Nandi. Modified self-organising map for automated novelty detection applied to vibration signal monitoring. *Mechanical systems and signal processing*, 20(3):593–610, 2006.

[187] Keith Worden, Tara Baldacchino, Jennifer Rowson, and Elizabeth J Cross. Some recent developments in shm based on nonstationary time series analysis. *Proceedings of the IEEE*, 104(8):1589–1603, 2016.

[188] Keith Worden, Graeme Manson, and David Allman. Experimental validation of a structural health monitoring methodology: Part i. novelty detection on a laboratory structure. *Journal of sound and vibration*, 259(2):323–343, 2003.

[189] Keith Worden, Graeme Manson, and Nick RJ Fieller. Damage detection using outlier analysis. *Journal of sound and vibration*, 229(3):647–667, 2000.

[190] Keith Worden, Wieslaw J Staszewski, and James J Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111, 2011.

[191] Zhaohua Wu and Norden E Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01):1–41, 2009.

[192] Jann N Yang, Yu Lei, S Lin, and N Huang. Hilbert-huang based approach for structural damage detection. *Journal of engineering mechanics*, 130(1):85–95, 2004.

[193] Mike Yeager, Bill Gregory, Chris Key, and Michael Todd. On using robust mahalanobis distance estimations for feature discrimination in a damage detection scenario. *Structural Health Monitoring*, 18(1):245–253, 2019.

[194] Yang Yu, Cheng Junsheng, et al. A roller bearing fault diagnosis method based on emd energy entropy and ann. *Journal of sound and vibration*, 294(1-2):269–277, 2006.

[195] FA Zakaria, Sofiane Maiz, Mohamed El Badaoui, and M Khalil. First-and second-order cyclostationary signal separation using morphological component analysis. *Digital Signal Processing*, 58:134–144, 2016.

[196] Dengsheng Zhang. Wavelet transform. In *Fundamentals of image data mining*, pages 35–44. Springer, 2019.

[197] Mingjie Zhang and Fuyou Xu. Variational mode decomposition based modal parameter identification in civil engineering. *Frontiers of Structural and Civil Engineering*, 13(5):1082–1094, 2019.

[198] Wenliang Zhou and David Chelidze. Blind source separation based vibration mode identification. *Mechanical systems and signal processing*, 21(8):3072–3087, 2007.

[199] Xiangyu Zhou, Yibing Li, Li Jiang, and Li Zhou. Fault feature extraction for rolling bearings based on parameter-adaptive variational mode decomposition and multi-point optimal minimum entropy deconvolution. *Measurement*, 173:108469, 2021.

[200] Olgierd Cecil Zienkiewicz, Robert Leroy Taylor, and Jian Z Zhu. *The finite element method: its basis and fundamentals.* Elsevier, 2005.