EARLY ANOMALY DETECTION AND CLASSIFICATION WITH STREAMING SYNCHROPHASOR DATA IN ELECTRIC ENERGY SYSTEMS

A Dissertation

by

YANG CHEN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Le Xie
Committee Members,	P. R. Kumar
	Alan Palazzolo
	Chanan Singh
	Jianzhong Tong
Head of Department,	Miroslav M. Begovic

December 2015

Major Subject: Electrical Engineering

ABSTRACT

The large-scale streaming data collected from the increasing deployed phasor measurement unit (PMU) devices poses significant difficulties for real-time data-driven analytics in power systems. This dissertation presents a dimensionality-reduction-based monitoring framework to make better use of the streaming PMU data for early anomaly detection and classification in power systems.

The first part of this dissertation studies the fundamental dimensionality of large-scale PMU data, and proposes an online application for early anomaly detection using the reduced dimensionality. First, PMU data under both normal and abnormal conditions are analyzed by principal component analysis (PCA), and the results suggest an extremely low underlying dimensionality despite the large number of raw measurements. In comparison with prior work of utilizing multi-channel high-dimensional PMU data for power system anomaly detection, the proposed early anomaly detection algorithm employs the reduced-dimensional data from PCA, and detects the occurrence of an anomaly based on the change of core subspaces of the low-dimensional PMU data. Theoretical justification for the algorithm is provided using linear dynamical system theory. It is demonstrated that the proposed algorithm is capable to detect general power system anomalies at an earlier stage than would be possible by monitoring the raw PMU data.

The second part of this dissertation investigates the classification of a special anomaly in power systems, low-frequency oscillation, which may cause severe impacts on power systems while at the same time is difficult to be accurately classified. We present a robust classification framework with online detection and mode estimation of low-frequency oscillations by using synchrophasor data. Based on per-

sistent homology, a cyclicity response function is proposed to detect an oscillation, through the use of the low-dimensional features (pre-PCA features) extracted from PCA. Whenever the cyclicity response exceeds a numerically robust threshold, an oscillation can be detected. After the detection, PCA is applied again to extract the low-dimensional features (post-PCA features) from the multi-channel transient PMU data. It is shown that the post-PCA features preserve the underlying modal information in a more robust way in comparison to raw synchrophasor measurements. Based on the post-PCA features, fast Fourier transform (FFT) and Prony analysis can be subsequently applied to extract modal information of the oscillation. The proposed classification framework offers system operators a data-driven analytical tool for fast detection of low-frequency oscillation and robust mode estimation against high measurement noise.

DEDICATION

To my family

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest respect and gratitude to my advisor, Professor Le Xie, for his wonderful guidance, persistent support, and endless encouragement over the years. Dr. Xie has truly become a friend and life mentor for me during the past five years at Texas A&M, and working with him is my greatest honor. His deep insight of seeing the big picture and valuable research skills have always inspired me. Because of him, I become a better researcher that I never thought I could be before. Especially, I would like to thank Dr. Xie for his great help, support and understanding with my career decision. Without his guidance and persistent help throughout my Ph.D. process, I would have never completed this dissertation.

Next, I wish to thank my committee members, Dr. P. R. Kumar, Dr. Alan Palazzolo, Dr. Chanan Singh and Dr. Jianzhong Tong, for their interest and discussions in my defense, and for their suggestions and time. Especially, I would like to thank Dr. P. R. Kumar, for all the informative discussions, the wonderful guidance, and the persistent help throughout my Ph.D. process. I would also like to thank Dr. Yuliy M. Baryshnikov from University of Illinois at Urbana-Champaign for the informative discussions.

I am indebted to my internship supervisor, Dr. Jianzhong Tong, at Pennsylvania-New Jersey-Maryland (PJM) Interconnection for providing me the internship opportunity and kindling my interest in the power industry.

I would like to thank the Electric Reliability Council of Texas (ERCOT), PJM Interconnection, and Bonneville Power Administration (BPA) for the synchrophasor

data provided for my Ph.D. research.

I am also grateful to the current and past members of Dr. Xie's research group, my office mates and friends: Meng Wu, Chen Yang, Yun Zhang, Dae-hyun Choi, Anupam A. Thatte, Yingzhong Gu, Omar A. Urquidez, Hao Ming, Xinbo Geng, Yuanyuan Li, Benjamin P. Wiseman, Sadegh Modarresi, Fan Zhang, James Carroll, Meichen Chen, Haiwang Zhong, Xiaowen Lai, and Yang Bai. I thank them for the technical discussions and for making my Ph.D. journey enjoyable.

I would like to thank the staff in Department of Electrical and Computer Engineering, particularly Tammy Carda, Jeanie Marshall, Anni Brunker, Melissa Sheldon, and Nancy Reichart.

Special thanks to my family for their endless love and support. I would like to thank my parents for their unconditional love, constant support and endless encouragement, especially my mom for being here several times to accompany me and take care of me. I also wish to thank my uncle and my grandmother for their strength that influences me a lot.

Above all I dedicate my work to my grandfather who passed away in March 2015.

NOMENCLATURE

AC	Alternating Current
AVR	Automated Voltage Regulator
BPA	Bonneville Power Administration
CAISO	California Independent System Operator
DAEs	Differential and Algebraic Equations
DFR	Digital Fault Recorder
DFT	Discrete Fourier Transform
DOF	Degree of Freedom
ERCOT	Electric Reliability Council of Texas
FDR	Frequency Disturbance Recorder
EMS	Energy Management System
FNET	Frequency Monitoring Network
FFT	Fast Fourier Transform
GPS	Global Positioning System
IEDs	Intelligent Electronic Devices
ISOs	Independent System Operators
PJM	Pennsylvania-New Jersey-Maryland
LTI	Linear Time invariant
MRI	Magnetic Resonance Imaging
NASPI	North American SynchroPhasor Initiative
PCs	Principal Components
PCA	Principal Component Analysis
PDC	Phasor Data Concentrator

PMU	Phasor Measurement Unit
PSS/E	Power System Simulator for Engineering
RTDMS	Real-time Dynamics Monitoring System
SCADA	Supervisory Control And Data Acquisition
SMIB	Single Machine Infinite Bus
UTC	Coordinated Universal Time
WAMPAC	Wide-area Monitoring, Protection and Control
WECC	Western Electricity Coordinating Council

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

NOMENCLATURE vii

TABLE OF CONTENTS ix

LIST OF FIGURES xii

LIST OF TABLES xv

1. INTRODUCTION 1

1.1 Motivation and Overview 1
1.2 Synchrophasor-based Anomaly Detection in Power Systems 4
1.2.1 Prior Work 4
1.2.2 Main Contribution 6
1.3 Synchrophasor-based Monitoring of Low-frequency Oscillation in Power Systems 7
1.3.1 Prior Work 7
1.3.2 Main Contribution 8
1.4 Dissertation Outline 9

2. BACKGROUND 11

2.1 Synchrophasor Technology 11
2.1.1 Synchrophasor Measurement Concept 11
2.1.2 Synchrophasor Device and Network 12
2.1.3 Synchrophasor Applications in Power Systems 14
2.1.4 Synchrophasor Deployment 16
2.2 Dimensionality Reduction 17
2.2.1 Principal Component Analysis 18
2.2.2 Isometric Feature Mapping 19
2.3 Power System Anomaly Detection and Classification 20

3. PMU-BASED EARLY ANOMALY DETECTION VIA DIMENSIONALITY REDUCTION . . . 23

3.1 Introduction . . . 23
3.2 Dimensionality Analysis of Synchrophasor Data . . . 25
3.2.1 Linear Dimensionality Analysis of Synchrophasor Data . . . 25
3.2.2 Nonlinear Dimensionality Analysis of Synchrophasor Data . . . 31
3.2.3 Choice of Dimensionality Reduction Technique . . . 32
3.3 Proposed Early Anomaly Detection Algorithm . . . 33
3.3.1 PCA-based Adaptive Training . . . 33
3.3.2 Robust Online Monitoring . . . 37
3.3.3 Accuracy of Early Anomaly Detection . . . 44
3.3.4 Choice of Measurements . . . 45
3.4 Numerical Examples . . . 46
3.4.1 Linear Dimensionality Reduction . . . 46
3.4.2 Online Anomaly Detection . . . 49

4. PMU-BASED ROBUST MONITORING OF LOW-FREQUENCY OSCILLATIONS . . . 61

4.1 Introduction . . . 61
4.2 Scatter Plot based Visualization of Power System Oscillations . . . 63
4.2.1 PCA based Adaptive Training . . . 63
4.2.2 Scatter Plot based Visualization . . . 65
4.3 Cyclicity Response based Oscillation Detection . . . 66
4.3.1 Delay Embedding Theorem . . . 66
4.3.2 Persistent Homology and Cyclicity Response . . . 67
4.4 PCA-enabled Mode Estimation Approach . . . 71
4.4.1 Post-PCA on Oscillatory Data . . . 71
4.4.2 Mode Estimation with Post-PCA Feature . . . 72
4.5 Robust Monitoring Framework of Low-frequency Oscillation . . . 76
4.5.1 Choice of Measurements . . . 78
4.5.2 Accuracy of Oscillation Detection . . . 78
4.6 Numerical Examples . . . 79
4.6.1 Scatter-plot-based Oscillation Visualization . . . 79
4.6.2 Cyclicity Response based Oscillation Detection . . . 84
4.6.3 PCA-based Mode Estimation . . . 86

5. CONCLUSIONS . . . 96

5.1 PMU-based Early Anomaly Detection . . . 96
5.1.1 Summary . . . 96
5.1.2 Future Work . . . 97

5.2 PMU-based Robust Monitoring of Low-frequency Oscillations 98
5.2.1 Summary 98
5.2.2 Future Work 99

REFERENCES 101

LIST OF FIGURES

FIGURE Page

2.1 Traditional PMU installation [56]. 13

2.2 Traditional phasor measurement system [57]. 14

2.3 Phasor applications [15]. 15

2.4 PMU map in North America as of Oct. 2014 [66]. 16

3.1 Demonstration of PCA-based dimensionality reduction with bus frequency measurements under normal condition. 27

3.2 Demonstration of PCA-based dimensionality reduction with bus frequency measurements during an oscillation. 28

3.3 Raw profile and scatter plot comparison of PCA-based oscillation detection with bus frequency measurements. 29

3.4 Demonstration of PCA-based dimensionality reduction with voltage magnitude measurements under normal condition. 30

3.5 Demonstration of PCA-based dimensionality reduction with voltage magnitude measurements during an oscillation. 30

3.6 Raw profile and scatter plot comparison of PCA-based oscillation detection with voltage magnitude measurements. 31

3.7 Residual variance using Isomap. 32

3.8 Overview of the early anomaly detection algorithm. 34

3.9 Implementation of the early anomaly detection algorithm. 34

3.10 Topology of PSS/E 23-bus system [82]. 47

3.11 Cumulative variance preserved by PCs for PSS/E data. 48

3.12 Cumulative variance of an oscillation. 50

3.13 Cumulative variance of a line fault. 50
3.14 Cumulative variance of the 1^{st} unit tripping. 51
3.15 Cumulative variance of the 2^{nd} unit tripping. 51
3.16 Timeline of three simulated system anomalies in PSS/E. 52
3.17 Line 152-202 tripping & closure in PSS/E. 54
3.18 Unit 3011 tripping in PSS/E. 55
3.19 Bus 211 input change in PSS/E. 56
3.20 Oscillation in Western U.S.. 57
3.21 Line fault in Western U.S.. 58
3.22 Unit trippings in TX. 59
4.1 Flowchart of oscillation visualization algorithm. 64
4.2 Bus frequency profile and scatter plot during a ringdown oscillation. . 67
4.3 Demonstration of hierarchical clustering. 69
4.4 Demonstration for the computation of persistent homology. 69
4.5 Implementation of the robust monitoring algorithm. 77
4.6 Scatter plot visualization of unit 3011 tripping. 80
4.7 Scatter plot visualization of control input change of unit 101. 81
4.8 Scatter plot visualization of unit 211 tripping. 81
4.9 Scatter plot visualization of line 151-201 tripping. 82
4.10 Scatter plot visualization of a unit tripping. 83
4.11 Scatter plot visualization of another unit tripping. 83
4.12 Scatter plot visualization of a ringdown oscillation. 84
4.13 Scatter plot visualization of a braking anomaly. 85
4.14 Classification of oscillations using cyclicity response. 85

4.15 Classification of non-oscillatory anomalies using cyclicity response. . . 86

4.16 Bus frequency measurements during a sustained oscillation. 87

4.17 FFT results for the sustained oscillation on: (a) raw PMU measurements, and (b) post-PCA feature. 90

4.18 Bus frequency measurements during a ringdown oscillation. 90

4.19 FFT results of raw PMU measurements for ringdown oscillation. . . . 91

4.20 FFT of post-PCA feature for ringdown oscillation. 92

4.21 Topology of 2-area system [5]. 93

4.22 Bus frequency deviation for synthetic case. 93

4.23 FFT of raw PMU measurements for synthetic case. 94

4.24 FFT of post-PCA data for synthetic case. 94

LIST OF TABLES

TABLE Page

1.1 Historical incidents related to low-frequency oscillations [3] 2

3.1 Dynamic models in PSS/E system 48

3.2 Results of dimensionality reduction analysis 49

3.3 Comparison of detection time for bus frequency cases 60

4.1 Official NASPI results of the sustained oscillation 87

4.2 Results comparisons of the sustained oscillation 88

4.3 Official NASPI results of the ringdown oscillation 91

1. INTRODUCTION*

1.1 Motivation and Overview

Since the early 1960s, low-frequency oscillations have led to many system separations [1], and even system-wide blackouts such as the 1996 Western Electricity Coordinating Council (WECC) Blackout induced by the undamped inter-area oscillation of a 0.25 Hz mode [2]. Besides the 1996 WECC Blackout, there are many other historical incidents induced by power system low-frequency oscillations. As illustrated in Table 1.1, low-frequency oscillations have led to severe damages on power system reliable operations. This drives both researchers and engineers to explore power system low-frequency oscillation and to develop effective ways for its detection.

In power systems, there are typically two types of low-frequency oscillations: 1) *Inter-area oscillation*, which refers to two coherent groups of generators swinging against each other at 0.1-1.0 Hz. The reason for inter-area oscillations could be either weak tie-line connections between the generating source and the load center, or a forced oscillation at an inter-area mode. 2) *Local oscillation*, which corresponds to one generator swinging against the rest of the system at 1.0-2.0 Hz. Typically, when automatic voltage regulators (AVRs) operate at a high output and feed into weak transmission networks, a local oscillation may occur.

*This section is in part a reprint of the material in the following papers: (1) Reprinted with permission from Le Xie, Yang Chen, and P. R. Kumar, "Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784-2794, Nov. 2014. Copyright 2014, IEEE. (2) Reprinted with permission from Yang Chen, Harish Chintakunta, Le Xie, Yuliy M. Baryshnikov, and P. R. Kumar, "Robust Detection and Mode Estimation of Power System Low-frequency Oscillations using Synchrophasor Data," *IEEE Transactions on Power Systems*, to be submitted. (3) Reprinted with permission from Yang Chen, Le Xie, and P. R. Kumar, "Dimensionality Reduction and Early Event Detection Using Online Synchrophasor Data," *IEEE Power and Energy Society General Meeting 2013*, pp. 1-5, 2013. Copyright 2013, IEEE.

Table 1.1: Historical incidents related to low-frequency oscillations [3]

Time	**Location**	**Further Consequences**	**Oscillating Mode (Hz)**
1980	United Kingdom		0.5
1984, 1989, 1990	Taiwan		0.78-1.05
1996	Western U.S., Canada	System Separations	0.224
1997	Scandinavia		0.4
Mar. 6, 2003	China	Blackouts	0.4
Aug. 14, 2003	U.S.	Blackouts	0.17
Sep. 28, 2003	Italy	Blackouts	0.55

To analyze low-frequency oscillations, traditional approaches start from a detailed dynamic model of the entire interconnection for inter-area oscillations [4, 5, 6], and a single-machine-infinite-bus (SMIB) representation for local oscillations [7]. Then further modal analysis, eigen analysis, and participation factor matrix methods can be applied for oscillation detection and analysis. However, the increasing penetration of spatially dispersed and temporally variable resources into power grids poses difficulties in achieving an accurate dynamic model of the system, and further results in an increasing modeling complexity for oscillation analysis. Even if an accurate dynamic representation can be achieved with new modeling techniques, the corresponding computational complexity will become another key issue that deteriorates the accuracy of model-based oscillation analysis.

On the other hand, synchrophasors have shown great potentials for improving wide-area monitoring, protection and control (WAMPAC) [8, 9]. The 30 Hz (and

higher) sampling rate and the synchronization provided by the global positioning system (GPS) enable monitoring and control of low-frequency oscillations in power grids through synchrophasor measurements [10, 11, 12]. As the synchrophasor technology exhibits great superiority in enhancing situational awareness, more and more PMUs and other intelligent electronic devices (IEDs), such as frequency monitoring network (FNET) [13] and frequency disturbance recorder (FDR) [14], are rapidly being brought online. Correspondingly, the availability of massive data is increasing significantly. Just one phasor data concentrator (PDC) collecting data from 100 PMUs of 20 measurements each at 30 Hz sampling rate generates over 50 GB data one day [15]. With the increasing amount of PMU data, it has become a challenge to determine how to best manage and leverage the increasing amount of data from synchrophasors for real-time operational benefits.

Dimensionality reduction has been recognized recently in power systems for its adaptive machine learning features [16, 17, 18]. PCA, as one of the premier linear dimensionality reduction methods, reduces dimensionality by preserving the most variance of the original data [19, 20, 21]. Its fast computational feature itself provides much attraction in the areas of coherency identification [22], extraction of fault features [23], and fault location [24].

In this dissertation, PCA is applied to the raw synchrophasor measurements to reduce the dimensionality for both early detection of general anomalies, and robust monitoring of low-frequency oscillations. The key part of this dissertation is to extract the underlying dimensionality of massive PMU data in wide-area power systems by use of PCA. The indication of operating condition changes from the change of underlying dimensionality is utilized to develop an early anomaly detection algorithm. The extracted pre-PCA and post-PCA features from PMU data are employed for robust monitoring of low-frequency oscillations.

The contributions of this dissertation are as follows: (1) The underlying low-dimensional characteristic of large-scale PMU data is investigated through PCA, with an online application of early anomaly detection proposed. Theoretical justification using linear dynamical system theory is provided to present the connection between PMU-based data-driven analytics and first-principle model-based analysis in power systems. (2) PCA-based scatter plot is employed to visualize the topological patterns during oscillatory and non-oscillatory anomalies. Built upon that, a robust monitoring framework for power system low-frequency oscillations is presented, where cyclicity response calculated by pre-PCA features serves effectively in oscillation detection, and the post-PCA features provide high accuracy and robustness of mode estimation against measurement noise. (3) A PMU-based data-driven analytical tool is developed for real-time anomaly detection and classification, requiring no prior knowledge of system topology/model.

1.2 Synchrophasor-based Anomaly Detection in Power Systems

1.2.1 Prior Work

In power systems, any incident that violates system normal operating conditions can be defined as an anomaly. Therefore, power system faults, islanding, generator/line outages, and oscillations, etc., can all be categorized as anomalies. Data mining tools have shown great potentials for anomaly detection in power systems, ranging from detection of cyber attacks [25, 26], voltage events [27], to predicting-aided state estimation [28], circuit breaker maintenance [29], and detection of sensitive buses [30]. Several conventional data mining techniques, including decision tree (DT), artificial neural networks (ANNs) and neuro fuzzy models, are first summarized in [31]. Each of these techniques is specifically associated with one or several applications in power systems, such as security assessment, fault detection, stability

assessment, economic dispatch, etc. As indicated in [31], till the early 2000, DT is the mainstream in data mining in power systems. Following [31], an overview on the applications of data mining in power systems in terms of visualization, clustering, outlier detection, and classification, is recently provided in [32], where it has been emphasized that fast, easy and interpretable data mining methods is becoming more and more attractive in power systems.

Among all the data sources in power systems, varying from geographic information system (GIS), markets, dynamics, to smart meters, PMUs, IEDs, PMUs provide time-synchronized measurements with high resolution (30 Hz and higher). There have been a large body of literature discussing about utilizing PMUs data to improve anomaly detection. [33] introduces FFT-, Yuly-Walker-, and matrix-pencil-based methods to find power system events, and identify common characteristics extracted from the events. The Lyapunov exponents of the voltage phasors are utilized to monitor the short-term voltage stability [34]. A PMU-based adaptive technique for transmission line fault detection and location is proposed using the discrete Fourier Transform (DFT) [35, 36]. The phasor angle measurements are employed together with the system topology to detect line outages [37]. [38] applies a Fourier-based ringdown analysis to detect electromechanical oscillatory modes in real-time. By only using PMU measurements, a fully adaptive fault location algorithm for series-compensated lines is proposed in [39], requiring no knowledge on the models of series-compensation devices.

Given the large-scale streaming PMU data into power systems, it becomes more and more significant for the system operators to be able to process and analyze these data in real-time. Dimensionality reduction starts to play an important in such a situation. In [40], PCA is first performed on synchrophasor data for dimensionality reduction. Using the participation weights and principal components (PCs), the

reconstruction errors are utilized to extract correlations of different variables and therefore reduce the dimensionality. The reconstruction accuracy is very high for the global variables such as bus frequency, due to the global characteristic. However, for some local variables, such as voltage magnitude or reactive power, the reconstructions may not exhibit high accuracy. Recently, PCA is applied in [41] for a combined usage of event detection and data archival reduction. Both voltage magnitude and real power events can be effectively detected using the real PMU data from a campus microgrid [41]. However, the events considered all have large deviations from normal operating conditions, and may easily be detected from the raw measurements. Under this condition, the event detection seems to be unnecessary.

1.2.2 Main Contribution

In Section 3, we first explore the underlying dimensionality of large scale PMU data, and then present a PCA-based algorithm, which lends itself to earlier anomaly detection than would be possible by monitoring the raw measurements.

Motivated by the increasing deployment of synchrophasors and the resulting large-scale PMU data, PCA is first applied on the high-dimensional PMU data to select the "pilot PMUs." Based on PCA, the dimensionality reduction analysis provides a significantly lower dimensional "signature" of the states in the overall power system. At the occurrence of a system anomaly, an alert from the early anomaly detection algorithm is issued whenever a large value of the proposed anomaly indicator, induced by the change of the core subspaces of the PMU data, is detected.

The main contributions of the early anomaly detection algorithm in this section are:

- It introduces an online data-driven analytical tool, which requires no knowledge of the system model/topology;

- It implements the dimensionality reduction at the adaptive training stage to extract the key features of the embedded high-dimensional PMU data;
- It performs anomaly detection using a much reduced number of PMUs as "pilots," which is computationally desirable in real-time operations;
- It is theoretically justified using linear dynamical system theory;
- For the online anomaly detection, it does not require lengthy buffering of data, which is required in the alternative approaches based on frequency-domain analysis;
- It is capable of detecting system anomalies at an earlier stage than would be possible by monitoring the raw PMU data.

1.3 Synchrophasor-based Monitoring of Low-frequency Oscillation in Power Systems

1.3.1 Prior Work

Considering the severe impacts of low-frequency oscillations on power system reliable operations, more and more efforts have been devoted into real-time low-frequency oscillation monitoring and analysis. However, due to the integration of new components in power grid, the accuracy of the model-based oscillation monitoring and analysis approaches [4, 42, 43] has been deteriorated, with increasing modeling and computational complexities. This results in a great significance of developing data-driven analytics for low-frequency oscillation monitoring and analysis.

The 30 Hz (and higher) sampling rate and the synchronization provided by GPS enable the monitoring and control of low-frequency oscillations in power grids through synchrophasor measurements [10, 11, 12]. [44] summarizes and compares the

performance of three data-driven methods, subspace identification, spectral independent component analysis, and wavelet transform, in their estimation of damping of electromechanical oscillations. Fourier spectral analysis has been applied to synchrophasor data to estimate the eigenvalues for monitoring the inter-area oscillation [45]. In [46], an adaptive stochastic subspace identification algorithm is proposed to estimate the mode and damping of low-frequency oscillations via a fast computational feature. Three distributed communication and computational architectures with centralized coordination are introduced in [47] for wide-area oscillation monitoring with PMU measurements. Oscillation monitoring system (OMS) with three signal processing engines, Prony's method, matrix pencil method, and Hankel total least squares (HTLS) method, is developed to extract the modal information and mode shape of low-frequency oscillations [48], and a corresponding OMS test engine is developed in Washington State University [49].

Besides the research work, [50] describes the basic principles of the oscillation trigger in the BPA disturbance monitor. Modal analysis for grid operation (MANGO) tool is developed by Pacific Northwest National Laboratory (PNNL) for mode estimation and damping improvement [51].

1.3.2 Main Contribution

In Section 4, a novel framework based on persistent homology and PCA is proposed for low-frequency oscillation monitoring. PCA-based scatter plot is first applied to visualize power system low-frequency oscillations. Then a robust monitoring algorithm is proposed, which lends itself to purely data-driven oscillation detection and mode estimation. Compared with existing literature, the proposed mode estimation approach utilizes both the pre-PCA and post-PCA features, and therefore provides a faster oscillation detection, and avoids the channel selection issues of

traditional Fourier-based analysis methods.

The main contributions of this section are as follows.

- Visualization of power system low-frequency oscillations is presented to differentiate oscillations from other types of anomalies by use of the topological-pattern-based approach.
- Cyclicity response with pre-PCA features provides high accuracy in detecting low-frequency oscillations from other non-oscillatory anomalies, and benefits real-time data analytics in power systems.
- Mode estimation with post-PCA feature preserves robustness of mode estimation against high measurement noise.
- The proposed robust monitoring algorithm is purely online data-driven, requiring no knowledge of system model or topology.

1.4 Dissertation Outline

The rest of the dissertation is organized as follows. Section 2 presents an overview of synchrophasor technology, including synchrophasor concept, device, network, applications in power systems, and deployment. The linear and nonlinear techniques of dimensionality reduction are then introduced. Traditional model-based anomaly detection and classification techniques are briefly reviewed.

In Section 3, dimensionality reduction is considered for power system early anomaly detection using PMU data. We first analyze the dimensionality of the synchrophasor data sampled during both normal operating conditions and contingencies. With the extremely low dimensionality of PMU data obtained from PCA, we propose an online application for early anomaly detection. The change of core subspaces of the PMU

data in the low-dimensional space indicates the occurrence of an anomaly. Theoretical justifications are provided for the proposed algorithm using linear dynamical system theory.

In Section 4, the scatter-plot-based visualization of oscillations is first introduced. Secondly, a persistent-homology-based cyclicity response is proposed, lending itself to oscillation detection through the use of the pre-PCA features. A pre-defined threshold can be determined by using historical eventful PMU data. Cyclicity response exceeding the threshold indicates the occurrence of an oscillation. After an oscillation is detected, PCA is applied again to extract the post-PCA features from multi-channel transient PMU measurements. Theoretical justification is provided to show that the post-PCA features still retain the modal information as is in the raw measurements. FFT and Prony analysis are finally applied to the resulting post-PCA features for mode estimation.

In Section 5, we summarize our main contributions in this dissertation and provide some future research directions.

2. BACKGROUND*

2.1 Synchrophasor Technology

2.1.1 Synchrophasor Measurement Concept

A *phasor* is a complex number representing both the magnitude and phase angle of voltage and current sinusoidal waveforms (50 or 60 Hz) at a specific point in time. *Synchrophasors* are the phasor measurements time stamped against the coordinated universal time (UTC) time by GPS. By precisely time-aligning measurements in different locations, the GPS synchronization enables synchrophasors to provide an interconnection-wide comprehensive view of the entire power grid stress and dynamics.

Compared to the traditional supervisory control and data acquisition (SCADA) system, the synchrophasor technology has three main advantages:

- *High resolution.* With a 30 Hz (and higher) sampling rate, synchrophasors provide a much detailed visibility of power systems than the traditional SCADA system, which samples the system every 2-4 seconds. This "magnetic resonance imaging (MRI) quality [52]" high-resolution technology enables synchrophasors for wide area monitoring, real-time dynamics and stability monitoring, and improvements in state estimation, protection, and controls.

- *Synchronization by GPS.* This feature further enables the wide-area monitoring

*This section is in part a reprint of the material in the following papers: (1) Reprinted with permission from Yang Chen, Le Xie, and P. R. Kumar, "Integrating PMU-data-driven and Physics-based Analytics for Power System Operations," *2014 48th Asilomar Conference on Signals, Systems, and Computers*, 2014. Copyright 2014, IEEE. (2) Reprinted with permission from Le Xie, Yang Chen, and P. R. Kumar, "Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784-2794, Nov. 2014. Copyright 2014, IEEE.

applications of synchrophasors.

- *Phase angle measurements.* Synchrophasors can directly provide the phase angle measurements at the high hub-second resolution. This has greatly improved the accuracy of phase angle measurements from traditional state estimator, and further enables the PMU-based state estimation with higher accuracy.

IEEE standard C37.118.1 [53, 54] specifies synchrophasor measurement definitions and performance requirements.

2.1.2 Synchrophasor Device and Network

2.1.2.1 Synchrophasor Device

A *PMU* is a standalone electronic device that uses the state-of-the-art digital signal processors to measure the 50/60 Hz alternating current (AC) voltage and/or current signals to provide synchrophasor measurements. Typical sampling rate of PMUs is 48 samples per cycle (2880 samples per second). An analog to digital converter is first utilized to digitize the analog AC waveforms for each phase. Then high-speed synchronized samples are created by a phase-lock oscillator along with a GPS reference source with 1 microsecond accuracy [55]. After that, the voltage and current phasors are computed by digital signal processing techniques.

Besides the PMU devices, the synchrophasor measurement functionality can be achieved by any device incorporating the PMU functionality, such as digital fault recorders (DFRs) and digital relays, i.e., the PMU-enabled IEDs. Other unrelated functions of the device must be shown not to affect the performance of the PMU component, and equally importantly the PMU functions must not affect the other functions of the device. The main components of a PMU or PMU-enabled IED include analog input signal interface, data acquisition system, phasor estimation module, and post-processing module for output data [15].

PMUs and PMU-enabled IEDs are typically installed in a substation or at a power plant. Figure 2.1 illustrates a typical installation of a PMU device.

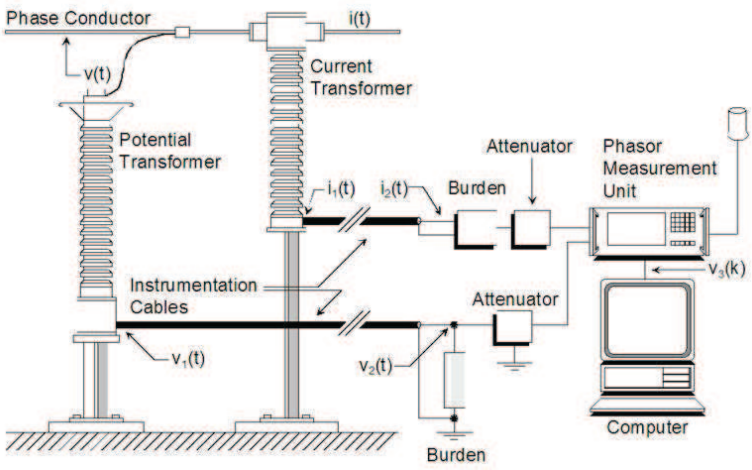


Figure 2.1: Traditional PMU installation [56].

2.1.2.2 Synchrophasor Network

As shown in Figure 2.2, a simple synchrophasor network consists of several PMUs and one PDC.

Typically, many PMUs located at various key substations gather data and send it in real time to a PDC at a location where the data is aggregated and analyzed. If multiple IEDs in a substation provide synchrophasor measurements, a PDC may be locally deployed at the substation. Software, such as real-time dynamics monitoring system (RTDMS), is utilized to compute and display the synchronized measurements, such as frequencies, primary voltages, currents, MWs and MVARs for reliability engineers. If there are many PDCs belonging to different utilities, a common central PDC, i.e., SuperPDC, will be employed to aggregate data across the utilities, in

order to provide an interconnection-wide snapshot of the power grid measurements [55].

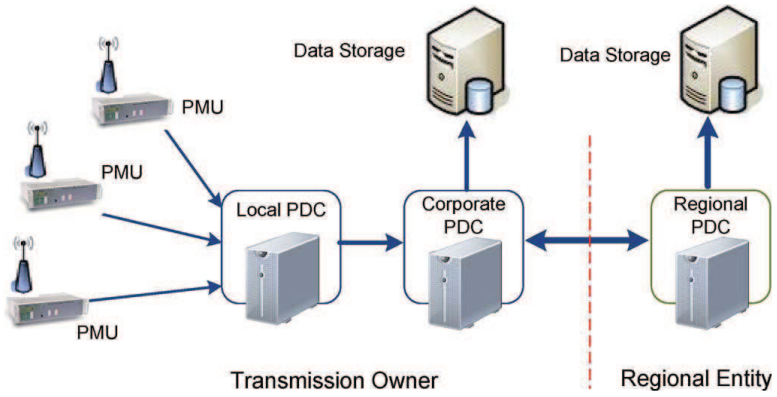


Figure 2.2: Traditional phasor measurement system [57].

2.1.3 Synchrophasor Applications in Power Systems

According to [15], the synchrophasor-based applications can be classified into three categories:

- *Online applications to support real-time operations.* California Independent System Operator (CAISO) started to implement synchrophasor data into control room and use RTDMS for WECC wide-area visualization and monitoring from 2008 [58]. Proposed in [59], the Mode Meter has been tested with field measurement data [51] from WECC, and has been utilized as a supplementary online oscillation monitoring tool. The PSGuard wide-area monitoring and control system [60] from ABB is capable to monitor phase angle stability, voltage stability, and line thermal limits, etc.

- *Offline applications to improve system planning and analysis.* The synchronized wide-area data are essential for disturbance analysis, as evidenced by

the 2003 U.S. Eastern blackout investigation and the 2006 European disturbance investigation [61]. The high-speed observations from PMUs can also be employed by planners to improve static and dynamic system models through model calibration and validation [62].

- *Response-based applications to enable wide-area control.* With the instantaneous and high-resolution measurements, PMUs can be used to activate local or centralized control of corrective measures for angular stability, voltage stability, and low-frequency oscillation.

From the perspective of different users, Figure 2.3 summarizes the potential phasor data applications.

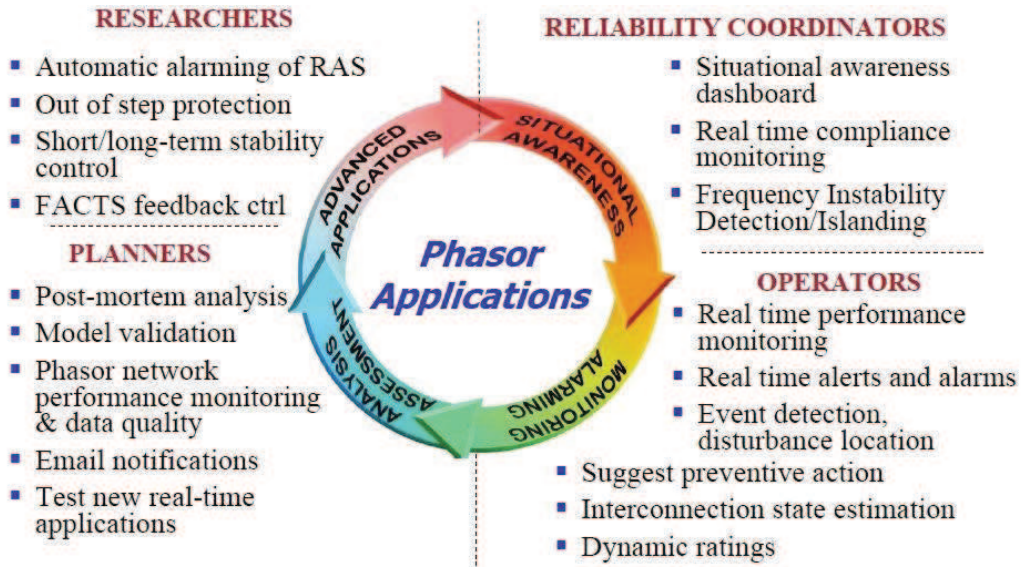


Figure 2.3: Phasor applications [15].

2.1.4 Synchrophasor Deployment

As reported in [63, 64], there are about 1400 (pure) PMU devices installed across North America from 2009 to 2014, let alone the number of PMU-enabled IEDs, and the micro-PMUs in the distribution system. Figure 2.4 illustrates the PMU map with synchrophasor data flows in North America as of October 2014. Besides, as reported in [65], from 2014 to 2023, the utilities are expected to invest nearly $107 billion in the development of synchrophasors and wider-area situational awareness systems.

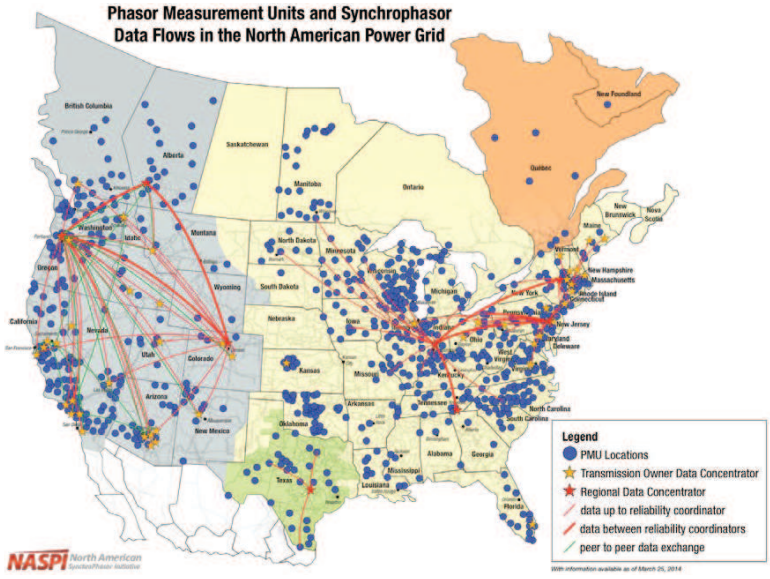


Figure 2.4: PMU map in North America as of Oct. 2014 [66].

Given the increasing deployed synchrophasor devices, it has become a challenge to determine how to best manage and leverage the increasing amount of data from synchrophasors for real-time operational benefits. Just one PDC collecting data from 100 PMUs of 20 measurements each at 30-Hz sampling rate generates over 50 GB of

data one day [15].

2.2 Dimensionality Reduction

With the high-dimensional streaming data from PMUs, it is advantageous for us to discover or impose the true structure on the data for real-time analysis. Therefore, it is of great significance to explore the *intrinsic dimensionality* of the PMU data.

The transformation of the high-dimensional datasets into a meaningful representation of the reduced dimensionality, ideally, the intrinsic dimensionality of the data, is referred as the *dimensionality reduction* [67]. The *intrinsic dimensionality* is defined as the minimum number of parameters required to account for the observed properties of the data [68]. Dimensionality reduction is important in many domains, since it facilitates classification, visualization, and compression of high-dimensional data, by mitigating the curse of dimensionality and other undesired properties of high-dimensional spaces [67].

The problem of dimensionality reduction can be defined as follows. Define a dataset matrix $\mathbf{Y} \in \mathbb{R}^{N \times n}$, which consists of N data vectors $\mathbf{y_i}$, $i = 1, 2, \ldots, N$, and has dimensionality n. Assume the intrinsic dimensionality of dataset $\mathbf{Y}$ is d, where $d < n$, and often $d \ll n$. In geometric terms, the intrinsic dimensionality means that the points in dataset $\mathbf{Y}$ are lying on or near a manifold with dimensionality d that is embedded in the n-dimensional space [21]. A dimensionality reduction technique aims to transform $\mathbf{Y}$ into a new dataset $\mathbf{Y_{new}}$ with dimensionality d, while preserving as much geometry of the data as possible.

There are mainly two categories of the dimensionality reduction techniques: 1) linear techniques, which are famous for the fast computational feature; and 2) nonlinear techniques, which are capable to capture the nonlinear degree of freedom (DOF) of the data. This section only introduces the two dimensionality reduction tech-

niques that are utilized in this dissertation. The readers are referred to [67] for more techniques.

2.2.1 Principal Component Analysis

PCA, as one of the premier linear dimensionality reduction methods, reduces dimensionality by preserving the most variance of the original data [19, 21]. Its fast computational feature alone is greatly attractive in the areas of coherency identification [22], extraction of fault features [23], and fault location [24], aside from its considerable benefits for visualization.

Mathematically, PCA is defined as an orthogonal linear transformation, aiming at finding a linear transformation $\mathbf{U}$ from the original high-dimensional dataset to a new coordinate system such that the largest variance by some projection of the data lies on the first coordinate (first principal component, or PC_1), the second largest variance on the second coordinate (second principal component, or PC_2), and so on. Equivalently, for a high-dimensional dataset $\mathbf{Y} \in \mathbb{R}^{N \times n}$, PCA is designed to find a linear transform $\mathbf{U}$ such that

$$\underset{\mathbf{U}}{\text{argmax}} \ \mathbf{U}^T \cdot \mathbf{\Sigma}_{\mathbf{Y}-\mu} \cdot \mathbf{U}, \tag{2.1}$$

where $\mathbf{\Sigma}_{\mathbf{Y}-\mu}$ represents the covariance matrix of the zero mean dataset $(\mathbf{Y} - \mu)$, and $\mu = E[\mathbf{Y}]$ is the mean of $\mathbf{Y}$.

This linear mapping $\mathbf{U}$ is formed by the d principal eigenvectors of the covariance matrix $\mathbf{\Sigma}_{\mathbf{Y}-\mu}$ as the columns, and each column is called a principal component (PC). The low-dimensional representations $\mathbf{Y}_{\mathbf{new}}$ of the original high-dimensional data can be computed by a linear mapping on the transformation $\mathbf{U}$, i.e.,

$$\mathbf{Y}_{\mathbf{new}} = (\mathbf{Y} - \mu) \cdot \mathbf{U}. \tag{2.2}$$

The main drawback of PCA is that the size of the covariance matrix is proportional to the dimensionality of the dataset. Therefore, for very high-dimensional dataset, the computation of the eigenvectors might be infeasible. Furthermore, simple PCA [69] and probabilistic PCA [70] have been proposed to address this problem.

2.2.2 Isometric Feature Mapping

Isometric Feature Mapping (Isomap) is a global nonlinear technique that attempts to preserve the global properties of the data, while in the meantime constructs the nonlinear transformation between the high-dimensional dataset $\mathbf{Y}$ and the low-dimensionality representation $\mathbf{Y_{new}}$. Based on Multidimensional Scaling (MDS), which only captures the Euclidean distance, Isomap aims to preserve the intrinsic geometry of the data, as captured in the geodesic manifold distances [71]. *Geodesic distance* is defined as the distance between two points measured over the manifold [71].

Given dataset $\mathbf{Y}$, Isomap [71] can be summarized in the following three steps:

(1) *Construct neighborhood graph G.* For each data point y_i, connect it with its k nearest neighbors.

(2) *Compute shortest paths.* With the neighborhood graph G, the shortest path between two data points, which forms a good estimate of the geodesic distance, can be computed using some algorithm. A pairwise geodesic distance matrix $\mathbf{D_G}$ can be obtained when all data points have been processed.

(3) *Construct d-dimensional embedding.* The low-dimensional representation $\mathbf{Y_{new}}$ can be computed by applying MDS on the resulting distance matrix.

Some major drawbacks of Isomap are: 1) it is computationally inefficient; 2) it is topological instable; 3) it may fail when the manifold is nonconvex [67].

2.3 Power System Anomaly Detection and Classification

In the area of data mining, *anomalies* are defined as patterns in data that do not belong to a well-defined notion of normal behavior [72]. The problem of *anomaly detection* refers to find such abnormal patterns that deviate from the expected behaviors in the datasets.

To better study the anomalies in power systems, it is worthwhile to understand the normal behaviors in power systems, namely, the normal operating conditions. According to [73], under normal operating conditions, the power system is characterized by:

- Almost symmetrical three-phase voltages and currents;
- Operational currents remaining below the pre-set levels (including some permissible overload);
- Voltage level within permissible range around the nominal value;
- Frequency of the signals equal or very close to the nominal 50/60 Hz;
- Harmonics content within permissible limits.

Violations of any of the above normal operating conditions may result in the occurrence of one or more anomalies, for which particular procedures and approaches are of great necessity in power system for detection and classification. Several power system anomalies with one or more violations of normal operating conditions are classified as follows:

- Asymmetrical three-phase conditions in power systems are often associated with power system ***faults***, of which protective relays are mainly utilized for

the indication and detection of the faults and location. However, the existence of the mis-operations of the protective relays, and the delayed update of the relay status sometimes fail the accurate detection of such faults in a timely manner.

- ***Generator or line tripping*** usually results in the operational currents exceeding the pre-set levels, and/or the voltage or frequency exceeding certain range around the nominal values. Such tripping anomalies can be indicated by the open/closure status of the equipped circuit breakers. Similarly, the delayed update of the status and mis-operations also exist for the circuit breakers, which make the traditional detection method inaccurate for timely detection.

- The harmonics content exceeding certain limits will result in ***power system oscillations***, where low-frequency oscillations attract the most attention because of the severe damage to the components and the whole system, and the difficulty for detection.

  Power system *low-frequency oscillations* contains the following two main types:

  * *Inter-area oscillation*, which refers to one generator/plant group oscillates against other groups at a frequency of 0.1 to 1.0 Hz. It usually occurs in the weak interconnected power systems through long tie lines. Sometimes, a forced oscillation may occur, with a bad control action driving an oscillation at an inter-area mode. In the early era of power system analysis, inter-area oscillations were analyzed and controlled by modal and eigen-analysis with a detailed representation of the entire interconnection [1, 4, 5] .

  * *Local oscillation*, in which one generator oscillates against the rest of the

whole system at a mode of 1.0 to 2.0 Hz. Such oscillations are usually caused by the AVRs of generators operating at high outputs and feeding into weak transmission networks [74]. The SMIB model is well applied in the local oscillation analysis, with a dynamic model for the oscillating generator, and the rest of the system being modeled as a constant voltage source with constant frequency [1].

Nowadays, with the increasing penetrations of renewable energy, electric vehicles, energy storage, and other new components into the power grid, it is of great difficulty to obtain an accurate dynamic representation of the entire interconnection for one power system to study the inter-area oscillation. Even for local oscillation analysis, only using the SMIB model is not accurate enough. Even though an accurate dynamic representation can be achieved by advanced modeling techniques, the resulting increasing computational complexity will also deteriorate the accuracy of the model-based analytics. Therefore, new approaches are of great necessity to capture the properties of oscillations.

All the drawbacks of traditional approaches and procedures in the ineffective detection and classification of power system anomalies drive us to explore in a parallel aspect of traditional model-based analytics, and to research on the measurement-based approaches for early anomaly detection and classification with the synchrophasor technology.

3. PMU-BASED EARLY ANOMALY DETECTION VIA DIMENSIONALITY REDUCTION*

3.1 Introduction

This section describes the synchrophasor based algorithm for early detection of system-wide anomalies by use of dimensionality reduction techniques.

With the increasingly deployed synchrophasors and the massive data it generated as discussed in Section 2.1.4, it becomes a major challenge for large-scale power system operators to utilize the data in a timely manner. As an example, with only 120 PMUs installed, the Tennessee Valley Authority (TVA) needs to manage 36 GB data per day [40]. It becomes very difficult for the system operators to directly use the raw data for real-time decision making.

The increasing deployment of synchrophasors and the large volume of PMU data raise several open questions: 1) What is the underlying dimensionality of the massive PMU data in wide-area power systems? 2) Does the underlying dimensionality change as the system operating conditions change? 3) Can such a change of dimensionality indicate the occurrence of an anomaly in power system real-time operations? 4) Is there any fundamental connection between the PMU data-driven analytics and the model-based analysis of power systems? These are the new questions that conventional model-based approaches alone cannot address. It is therefore an urgent need to leverage inherent correlations among the PMU data for dimensionality reduction.

*This section is in part a reprint of the material in the following papers: (1) Reprinted with permission from Le Xie, Yang Chen, and P. R. Kumar, "Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784-2794, Nov. 2014. Copyright 2014, IEEE. (2) Reprinted with permission from Yang Chen, Le Xie, and P. R. Kumar, "Dimensionality Reduction and Early Event Detection Using Online Synchrophasor Data," *IEEE Power and Energy Society General Meeting 2013*, pp. 1-5, 2013. Copyright 2013, IEEE.

The dimensionality reduction method has been recognized in power systems for its adaptive machine learning features. In this section, by exploring the underlying dimensionality of PMU data, we propose theoretical justifications for an early anomaly detection algorithm. Based on PCA, the dimensionality reduction analysis provides a significantly lower dimensional "signature" of the states in the overall power system [17]. At the occurrence of a system anomaly, an alert from the early anomaly detection algorithm is issued whenever a large value of the proposed anomaly indicator, induced by the change of the core subspaces of the PMU data, is detected.

The main contributions of this section are:

- It proposes an online data-driven approach without requiring any knowledge of the system model/topology;
- It implements the dimensionality reduction at the adaptive training stage to extract the key features of the embedded high-dimensional PMU data;
- It performs anomaly detection using a much reduced number of PMUs as "pilots," which is computationally desirable in real-time operations;
- It is theoretically justified using linear dynamical system theory;
- For the online anomaly detection, it does not require lengthy buffering of data, which is required in the alternative approaches based on frequency-domain analysis;
- It is capable of detecting system anomalies at an earlier stage than would be possible by monitoring the raw PMU data.

3.2 Dimensionality Analysis of Synchrophasor Data

In this section, we explore the underlying dimensionality of streaming PMU data by use of dimensionality reduction. Both the linear technique, PCA (introduced in Section 2.2.1), and the nonlinear technique, Isomap (introduced in Section 2.2.2), are applied.

The notations utilized in this section are described as follows. Let p denote the number of available PMUs across the whole power network, each providing ℓ measurements. It is anticipated that there could be up to thousands of PMUs in interconnected power systems, with each PMU providing up to 20 different measurements[†] at each sample [15, 75, 76]. At each time sample, a total of $N_T := p \times \ell$ measurements are collected, indicating the difficulty of online data analytics. Define the *measurement matrix* $\mathbf{Y_{n\times N}} := [\mathbf{y^{(1)}}, \ \ldots, \ \mathbf{y^{(N)}}]$ containing the N measurements. Each measurement has n past samples, i.e., $\mathbf{y^{(i)}} := [y_1^{(i)}, \ \ldots, \ y_n^{(i)}]^T$, $i = 1, \ \ldots, \ N$.

3.2.1 Linear Dimensionality Analysis of Synchrophasor Data

This part describes the results from PCA on the synchrophasor data and the demonstration of results by use of real PMU data from Western U.S.

3.2.1.1 Results from PCA

When performing PCA on the measurement matrix $\mathbf{Y}$, a loading plot, a score plot, and cumulative variances will be produced to interpret the results.

On the *loading plot*, the projected data "sitting together" indicates the similarity in the behaviors among the N variables. The axes of the loading plot are formulated by the loading of PC_1, PC_2, etc.

[†]Considering the different numbers of measurements provided by different types of IEDs, this dissertation utilizes the industry grade PMU devices, which provide 20 measurements at each sample [40].

The *score plot* is also called the *scatter plot*, where the projected data "sitting together" indicates the similarity in the behaviors of the n samples in the time history. The axes of the score plot are formed by PCs. Usually in order to visualize the scatter plot, a 2D plot with PC_1 and PC_2 as the axes, or a 3D plot with PC_1, PC_2 and PC_3 can be utilized.

As discussed in Section 2.2.1, the linear mapping from PCA is formed by the principal eigenvectors of the covariance matrix. Consider the eigenvectors form a basis for the data. Then the eigenvalues represent the distribution of the source data's energy among each of the eigenvectors. The *cumulative energy content* or *cumulative variance* for the j^{th} eigenvector is the sum of the energy contents across all of the eigenvalues from 1 through j. In other words, the cumulative variance preserved by the j^{th} PC has an added value of the variances from 1 to j [77].

3.2.1.2 Case Studies

In this part, the real PMU data from Western U.S. are utilized to demonstrate the results from PCA. A total number of 44 PMUs are utilized, with measurements of bus frequency and voltage magnitude at a sampling rate of 60 Hz.

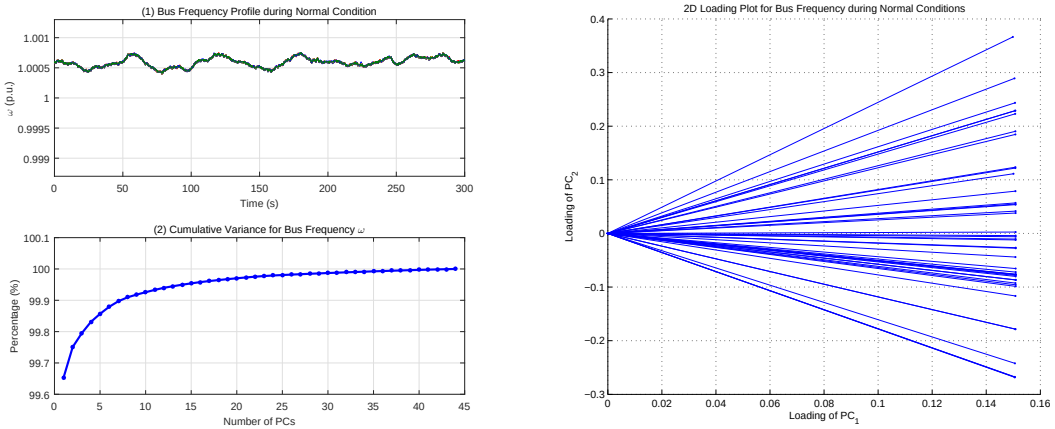
The purpose of the analysis is to

- demonstrate why the dimensionality of large-scale PMU data can be reduced by PCA, and
- illustrate how PCA can indicate the occurrence of a system-wide anomaly with PMU data.

3.2.1.2.1 Analysis with Bus Frequency Measurements

5 minutes of bus frequency data under normal operating conditions are depicted in Figure 3.1(a.1), with the corresponding cumulative variance shown in Figure 3.1(a.2).

As can be observed, only the first 2 or 3 PCs can preserve over 99% cumulative variance, indicating the high redundancy of the bus frequency measurements. Similarly, from the loading plot in Figure 3.1(b), the end points of all the radials cluster as a line, indicating the similar behaviors among the 44 measurements.



(a) Profile and cumulative variance.

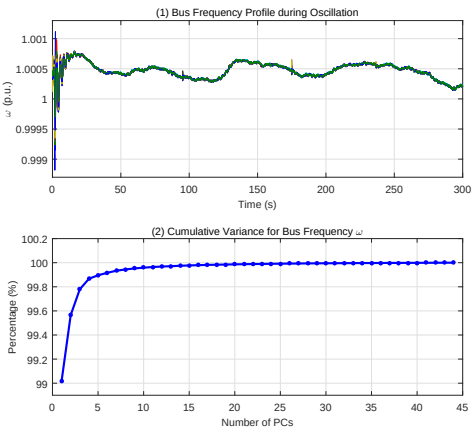
(b) 2D loading plot.

Figure 3.1: Demonstration of PCA-based dimensionality reduction with bus frequency measurements under normal condition.

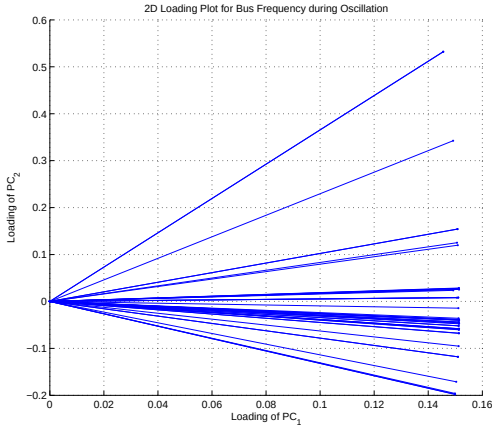
Figure 3.2 illustrates the results using eventful data. 5 minutes of bus frequency data starting from a wind-farm oscillation are utilized. Similarly as the case with normal PMU data, the first 2 or 3 PCs still preserve a high amount of cumulative variance, above 99.5% as shown in Figure 3.2(a.2). Even with an oscillation, the loading plot in Figure 3.2(b) still clusters as a line, indicating the redundancy.

From Figures 3.1 and 3.2, it can be observed that PCA can significantly reduce the dimensionality of the bus frequency data by preserving a large amount of cumulative variance, i.e., preserving most of the information contained in the original high dimensional data. In other words, PCA performs well in reducing the dimensionality

of the bus frequency measurements.



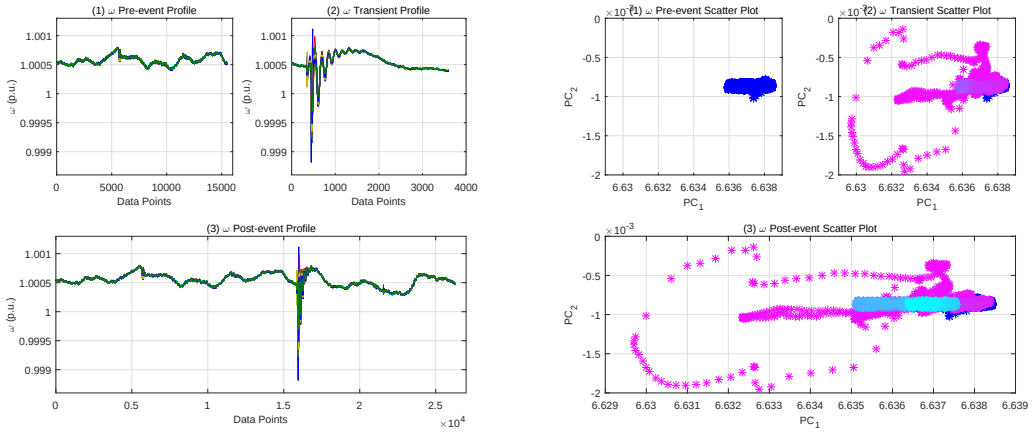
(a) Profile and cumulative variance.



(b) 2D loading plot.

Figure 3.2: Demonstration of PCA-based dimensionality reduction with bus frequency measurements during an oscillation.

Figure 3.3 illustrates the bus frequency profile before, during and after the oscillation and the corresponding scatter plot. The change of color in the scatter plot Figure 3.3(b) indicates the time trajectory. As can be observed, before the anomaly, the bus frequency measurements in Figure 3.3(a.1) are around nominal values of 1.0005 p.u., while the scatter plot in Figure 3.3(b.1) shows a small cluster in the 2D space. With the occurrence of the oscillation in Figure 3.3(a.2), the scatter plot starts to deviate from the original cluster with a changing color to pink. After the anomaly, when the system recovers to a new operating condition in Figure 3.3(a.3), the scatter plot again comes back to a new cluster as shown in Figure 3.3(b.3) in cyan. The change of subspace with the pink color clearly indicates the occurrence of the anomaly, further illustrating the capability of PCA to detect system-wide anomalies.



(a) Profile. (b) Scatter plot.

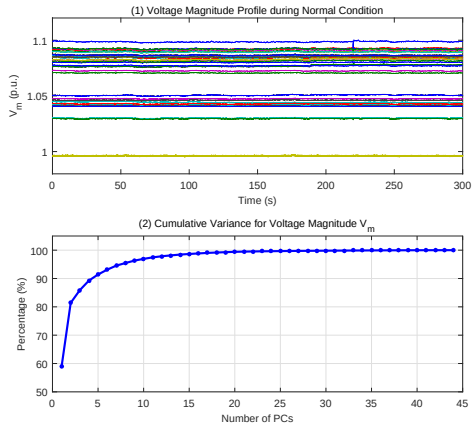
Figure 3.3: Raw profile and scatter plot comparison of PCA-based oscillation detection with bus frequency measurements.

3.2.1.2.2 Analysis with Voltage Magnitude Measurements

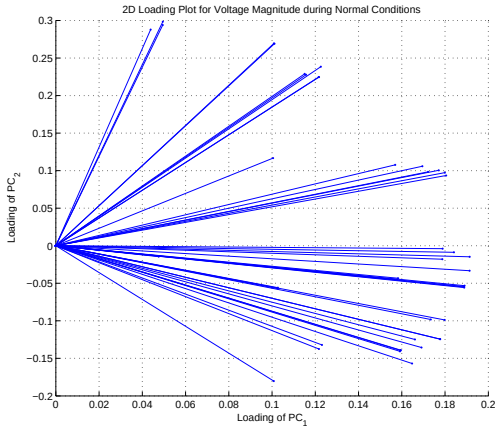
In this part, we illustrate the PCA results by use of voltage magnitude measurements. Similarly as the bus frequency case, 5 minutes data under normal condition and following an oscillation are utilized.

Different from the bus frequency case, the first 2 PCs for voltage magnitude only preserve about 80% cumulative variance in Figure 3.4(a) with normal data, and 87% in Figure 3.5(a) with oscillatory data. From the corresponding loading plots in Figures 3.4(b) and 3.5(b), the separability of the radials, compared to the ones in Figures 3.1(b) and 3.2(b), indicates that the voltage magnitude measurements have much less redundancy compared to the bus frequency. One reason for this is that bus frequency is well-known to be a global variable in power system, while voltage magnitude is a local variable because of the voltage levels.

Figure 3.6 demonstrates the PCA results for the scatter plot of voltage magnitude measurements with the raw profile plot. Similarly, the change of subspace in

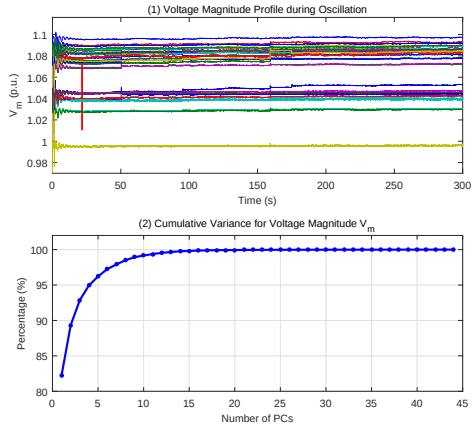


(a) Profile and cumulative variance.

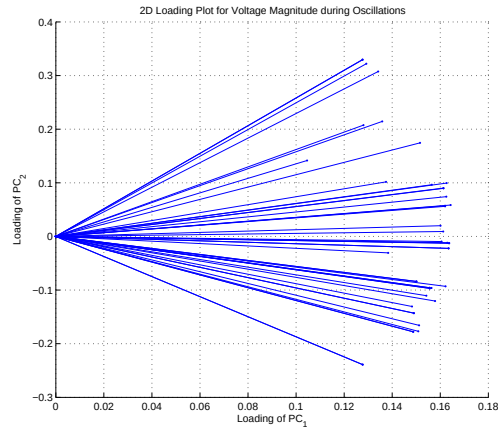


(b) 2D loading plot.

Figure 3.4: Demonstration of PCA-based dimensionality reduction with voltage magnitude measurements under normal condition.

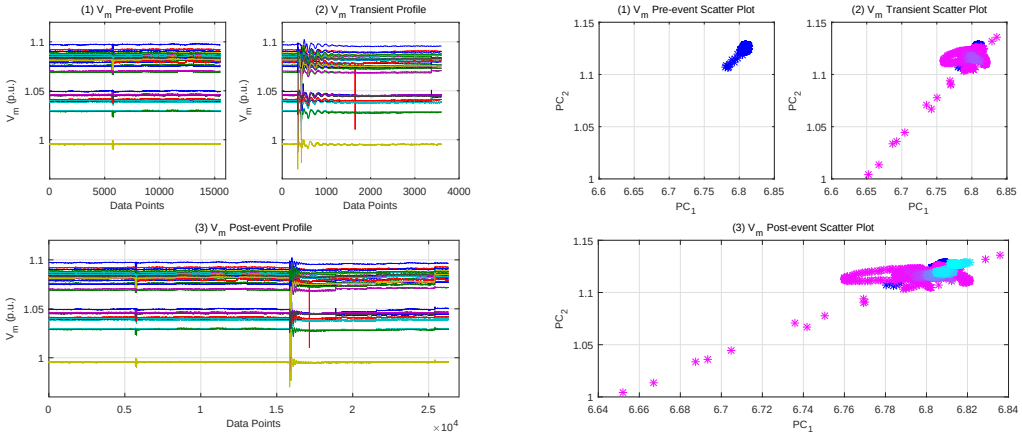


(a) Profile and cumulative variance.



(b) 2D loading plot.

Figure 3.5: Demonstration of PCA-based dimensionality reduction with voltage magnitude measurements during an oscillation.



(a) Profile. (b) Scatter plot.

Figure 3.6: Raw profile and scatter plot comparison of PCA-based oscillation detection with voltage magnitude measurements.

the scatter plot indicates the occurrence of the anomaly. However, because of the localized property of voltage magnitude, this indication is not as clear as the ones from bus frequency measurements. Therefore, in the following case studies, the bus frequency measurements will be utilized for demonstrations.

3.2.2 Nonlinear Dimensionality Analysis of Synchrophasor Data

In this part, Isomap will be utilized to explore the nonlinear dimensionality of the synchrophasor data.

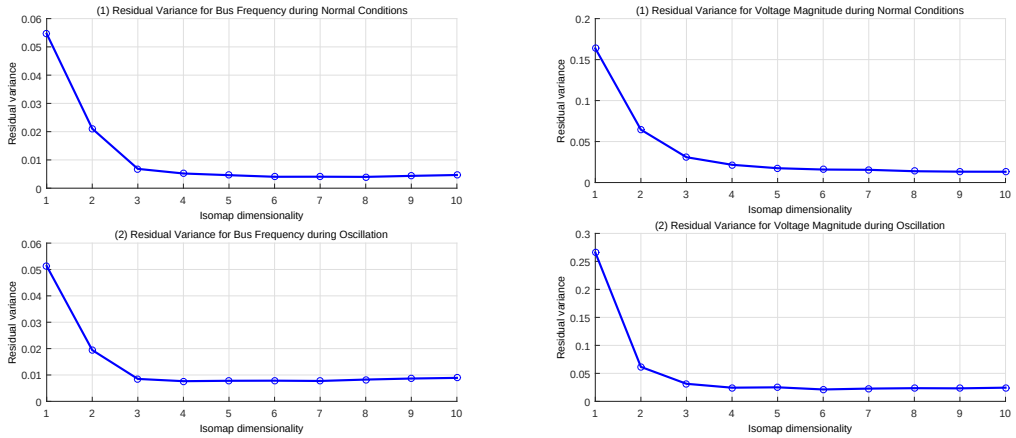
As discussed in [71], Isomap is capable to capture the intrinsic DOF from high-dimensional data by the "knee point" of the residual variance.

The same set of data from Western U.S. are utilized for both bus frequency and voltage magnitude analysis. The nonlinear dimensionality will be analyzed with data from both normal conditions and oscillations.

Figure 3.7(a) illustrates the residual variances for the bus frequency measurements under both normal condition and oscillation. As can be observed, the "knee point"

occurs at Isomap dimensionality of 3. This demonstrates that given the 44 PMU measurement sets, the intrinsic DOF of the data tends to be 3.

Similarly as the bus frequency case, the "knee point" obtained from the residual variance using the voltage magnitude measurements in Figure 3.7(b) also occurs around Isomap dimensionality of 2 or 3.



(a) With bus frequency measurements. (b) With voltage magnitude measurements.

Figure 3.7: Residual variance using Isomap.

3.2.3 Choice of Dimensionality Reduction Technique

Sections 3.2.1 and 3.2.2 have illustrated that the dimensionality of high-dimensional PMU data can be significantly reduced by use of both linear and nonlinear techniques.

As shown by PCA in Section 3.2.1.2, only 2 or 3 PCs can preserve a high amount of cumulative variance from both bus frequency and voltage magnitude measurements. Similarly, the "knee point" in the residual variance in Isomap in Section 3.2.2 also suggests the intrinsic DOF of the high-dimensional PMU data is around 3. In other words, the results from PCA match the ones from Isomap, indicating both linear

and nonlinear techniques are effective in reducing the dimensionality of PMU data.

Considering the purpose of utilizing PMU data to improve power system real-time monitoring and control, the computational efficiency of the selected technique should be high. PCA, as a traditional linear technique, is famous for its fast computational feature. For Isomap, although it's capable to capture the intrinsic DOF, it is very computational time-consuming. Therefore, PCA is utilized to further develop the early anomaly detection algorithm by use of the streaming PMU data.

3.3 Proposed Early Anomaly Detection Algorithm

If the massive PMU data essentially lie in a much reduced dimensional space, independent system operators (ISOs) or vendors can leverage the change in the underlying subspaces of the PMU data to visualize and detect system anomalies at an early stage. In this section, we propose such an algorithm with the following features: (a) only a reduced number of PMUs are needed; (b) it is online implementable; (c) it is theoretically justified using linear dynamical system theory; (d) the implementation of the algorithm can be done without knowledge of any underlying physical model of the system; (e) it can detect a system anomaly at a very early stage. Figure 3.8 provides an overview of the proposed early anomaly detection algorithm to be implemented in power systems.

This proposed early anomaly detection algorithm with two parts is shown in Figure 3.9.

3.3.1 PCA-based Adaptive Training

The PCA-based adaptive training stage is presumed to have been taken in normal operating conditions. Assume the *training period* is T_{trn}.

The PCA-based dimensionality analysis is described as follows:

1) **Form the PMU measurement matrix Y**.

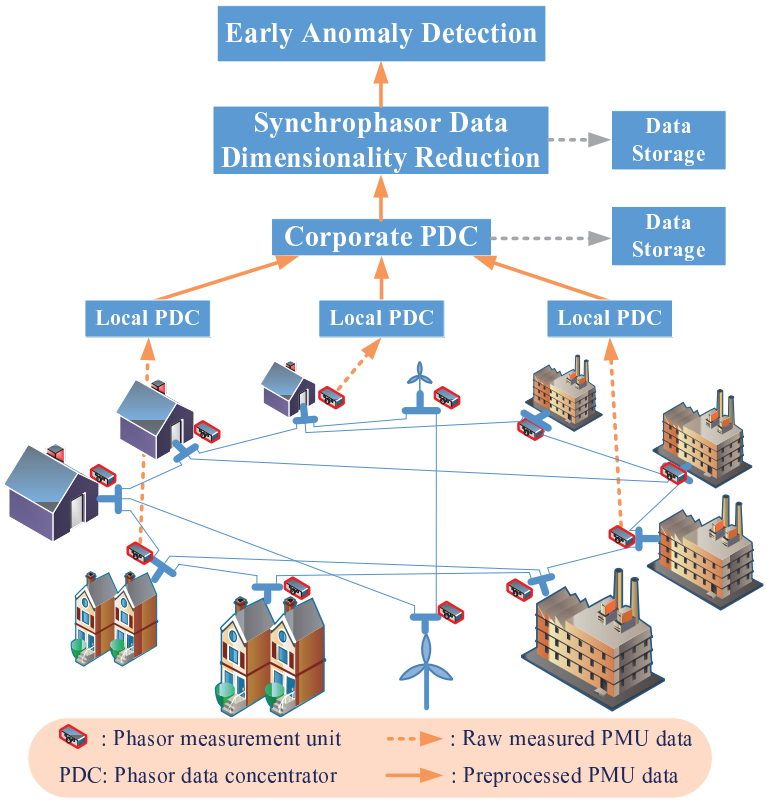


Figure 3.8: Overview of the early anomaly detection algorithm.

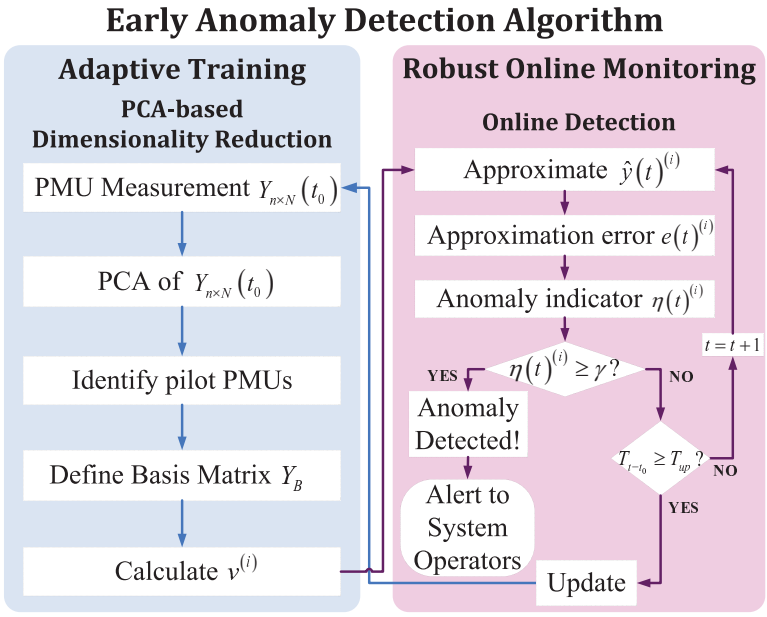


Figure 3.9: Implementation of the early anomaly detection algorithm.

At current time t_0, the PMU data in T_{trn}, under normal operating conditions, are employed to form the measurement matrix $\mathbf{Y_{n\times N}(t_0)}$ for training.

2) **Perform PCA on Y**.

3) **Identify "pilot" PMUs.**

Pilot PMUs are defined as the PMUs whose projections in the low-dimensional space are as orthogonal as possible. In other words, the projections of pilot PMUs form the linear basis in the low-dimensional space. The approaches to determine the pilot PMUs and the number of them are described as follows.

- Out of the N PCs from PCA, select the highest m, which preserve a cumulative variance satisfying $\sum_{i=1}^{m} var_i \geq \tau$. τ is a pre-defined variance threshold, and $m \ll N$. For the purpose of visualization, m is usually selected as 2 or 3.

- Form a new m-dimensional subspace from the top m PCs.

- Project the original N-dimensional data onto the m-dimensional PC-based space. Select m projected variables such that for projections $p^{(i)}$ and $p^{(j)}$,

$$\cos\theta = \frac{p^{(i)} \cdot p^{(j)}}{|p^{(i)}| \cdot |p^{(j)}|} \approx 0, \tag{3.1}$$

  where $i, j = 1, \ldots, m$, i.e., the m projections should be as orthogonal to each other as possible. Then the m PMUs corresponding to the m projections are denoted as "pilot[‡]" PMUs. The remaining $(N - m)$ PMUs are denoted as "non-pilot[§]" PMUs.

[‡]In reality, some practical concerns could also be included in the determination of the pilot PMUs. For example, for some topologically and physically significant buses, their installed PMUs can be enforced to be pilot PMUs.

[§]Excluding the pilot PMUs from the total N available PMUs, the rest are denoted as non-pilot PMUs. In reality, some PMUs can also be enforced to be non-pilot PMUs if they are historically eventful.

4) **Form the basis matrix $\mathbf{Y_B}$.**

Define $\mathbf{Y_B} := [\mathbf{y_b^{(1)}}, \ \ldots, \ \mathbf{y_b^{(m)}}] \in \mathbb{R}^{n \times m}$, containing m pilot PMUs as columns. $\mathbf{Y_B}$ forms a linear basis for the original N measurements.

5) **Calculate linear regression coefficients $\mathbf{v^{(i)}}$.**

Represent the non-pilot PMUs $\mathbf{y^{(i)}}$ in terms of $\mathbf{Y_B}$ by linear regression, where $\mathbf{y^{(i)}} \subseteq \mathbf{Y}$ and $\mathbf{y^{(i)}} \nsubseteq \mathbf{Y_B}$, $i = 1, \ldots, N - m$. Let $\mathbf{v^{(i)}} := [v_1^{(i)}, \ \ldots, \ v_m^{(i)}]^T$ be the vector of the linear regression coefficients for the approximation, i.e.,

$$\mathbf{y^{(i)}} \approx \sum_{j=1}^{m} v_j^{(i)} \cdot \mathbf{y_b^{(j)}} = \mathbf{Y_B} \cdot \mathbf{v^{(i)}}. \tag{3.2}$$

Considering the large number of training PMU data in the dimensionality analysis, it follows that $n \gg m$. Therefore, using $\mathbf{y^{(i)}}$ and $\mathbf{Y_B}$ from the training data, $\mathbf{v^{(i)}}$ can be calculated by solving the over-determined problem (3.2) as

$$\mathbf{v^{(i)}} := (\mathbf{Y_B}^T \mathbf{Y_B})^{-1} \mathbf{Y_B}^T \mathbf{y^{(i)}}, \tag{3.3}$$

in which the squared approximation error $\|\mathbf{y^{(i)}} - \mathbf{Y_B} \cdot \mathbf{v^{(i)}}\|^2$ is minimized [78].

Using (3.3), each non-pilot PMU measurement vector can be represented in terms of the pilot PMUs. The dimensionality of the PMUs across the whole network can therefore be reduced from N to m, where $m \ll N$. In such a case, ISOs or the vendors can utilize the pilot PMUs to approximate some selected non-pilot PMUs and detect the changes of system operating conditions in real-time operations.

6) **Adaptive update.**

The update signal is sent from the second stage of robust online monitoring.

Denote the *update period* as T_{up}. T_{up} is a system-dependent variable, and usually can be chosen as 3-5 minutes. The adaptive update mechanism is designed as follows:

- If there is no anomaly occurring during a period of T_{up}, the training procedure is adaptively updated every T_{up} time units.
- If an anomaly is detected within T_{up}, the training procedure is updated immediately after the system recovers from the anomaly.

3.3.2 Robust Online Monitoring

The robust online monitoring utilizes the basis matrix $\mathbf{Y_B}$ at current time t and the coefficients $\mathbf{v^{(i)}}$'s calculated from the adaptive training stage to approximate the measurements of some selected non-pilot PMUs at the same time. Under normal operating conditions, the predictor coefficients $\mathbf{v^{(i)}}$'s provide accurate approximations of non-pilot $\mathbf{y^{(i)}}$'s because of the usage of normal-operating-condition data in the training procedure. Whenever an anomaly occurs, the spatial dependencies inside the power system will change, resulting in the deterioration of the approximations, leading to large approximation errors. Whenever a significant approximation error is noticed, an anomaly alert is declared for the purpose of corrective control.

Assume that the *real-time approximation* of the i^{th} non-pilot PMU is

$$\hat{y}(t)^{(i)} := \mathbf{Y_B^{meas}}(t) \cdot \mathbf{v^{(i)}}, \tag{3.4}$$

where $\mathbf{Y_B^{meas}}(t)$ is the real-time measurements of $\mathbf{Y_B}$ at time t, and $\mathbf{v^{(i)}}$ is adopted from the adaptive training in Section 3.3.1.

Define the *relative approximation error* of the i^{th} non-pilot PMU as

$$e(t)^{(i)} := \frac{\tilde{y}(t)^{(i)}}{y(t)^{(i),meas}} \times 100\%, \tag{3.5}$$

where $y(t)^{(i),meas}$ represents the real-time measurement of the ith non-pilot PMU at time t, and $\tilde{y}(t)^{(i)} := \hat{y}(t)^{(i)} - y(t)^{(i),meas}$ is the *absolute approximation error*. The

occurrence of anomalies can be monitored by using $e(t)$ of some selected non-pilot PMUs.

Numerically, because of the per unit scale of power system variables, $e(t)^{(i)}$'s may be too small to be accurately identified at the occurrence of anomalies. We therefore propose a real-time ***anomaly indicator*** $\eta(t)^{(i)}$ for the i^{th} non-pilot PMU, for the purpose of early anomaly detection, as

$$\eta(t)^{(i)} := \frac{e(t)^{(i)}}{e^{(i)}_{normal}}, \tag{3.6}$$

where $e^{(i)}_{normal}$ is the mean value of $e(t)^{(i)}$ calculated under normal operating conditions. $\eta(t)$ represents the *normalized relative approximation error*. Whenever $\eta(t)$ becomes larger than a pre-specified threshold γ, an anomaly alert is issued. Given the fact that the PMU samples at a rate of 30 Hz or higher, an alert can be issued within several samples after the occurrence of an anomaly. Such a swift alert is capable of quickly identifying system anomalies in real-time situations.

Proposition 1 (PCA-based Early Anomaly Detection)**.** *Using the proposed anomaly indicator (3.6), a system anomaly can be detected within several samples of PMUs, whenever, for some selected non-pilot PMU i, the anomaly indicator satisfies*

$$\left|\eta(t)^{(i)}\right| \geq \gamma, \tag{3.7}$$

where γ is a system-dependent threshold that can be calculated using historical eventful PMU data.

Proof

A *system anomaly* is generalized in this dissertation as a change of system topology (such as the generator or line tripping), operating conditions or control inputs.

Large-scale power systems can be described by a coupled set of nonlinear DAEs [79]

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u_o}(t), \mathbf{h}(t), \mathbf{q}), \tag{3.8}$$

$$\mathbf{0} = \mathbf{g}(\mathbf{x}(t), \mathbf{u_o}(t), \mathbf{h}(t), \mathbf{q}), \tag{3.9}$$

where $\mathbf{x}(t)$ and $\mathbf{u_o}(t)$ represent the power system dynamic state and input vectors, respectively. $\mathbf{h}(t)$ defines the algebraic variables, i.e., the real and reactive power injections. $\mathbf{q}$ denotes the time invariant system parameters. Differential equation (3.8) consists of all the system dynamics including generators, wind turbines, loads, etc. Algebraic equation (3.9) represents the real and reactive power balance equations.

We linearize the nonlinear DAEs (3.8)-(3.9) around one system equilibrium point (one operating condition), and eliminate the algebraic equations by Kron Reduction [80]. The resulting continuous linear time invariant (LTI) state space model is

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \boldsymbol{\alpha}(t), \tag{3.10}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \boldsymbol{\varepsilon}(t), \tag{3.11}$$

where $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are the state and measurement vectors, respectively, with corresponding system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, which usually satisfies $\mathbf{D} \approx \mathbf{0}$ in power systems. $\mathbf{u}(t)$ is the augmented input vector including the original system inputs $\mathbf{u_o}(t)$ with the net injections $\mathbf{h}(t)$ of real and reactive power [80]. $\boldsymbol{\alpha}(t) \sim N(\mathbf{0},\ \mathbf{Q})$ and $\boldsymbol{\varepsilon}(t) \sim N(\mathbf{0},\ \mathbf{R})$ are assumed to be uncorrelated white noises representing the modeling and measurement errors, respectively.

Assume: 1) A zero-order hold of $\mathbf{u}(t)$; 2) A continuous integration of $\boldsymbol{\varepsilon}(t)$; 3)

$\mathbf{D} \approx \mathbf{0}$. The discretization of (3.10) and (3.11) with sampling time T yields [81]

$$\mathbf{x}[k+1] = e^{\mathbf{A}T}\mathbf{x}[k] + \mathbf{A}^{-1}\left(e^{\mathbf{A}T} - \mathbf{I}\right)\mathbf{B}\mathbf{u}[k] + \boldsymbol{\alpha}[k], \tag{3.12}$$

$$\mathbf{y}[k] = \mathbf{C}\mathbf{x}[k] + \boldsymbol{\varepsilon}[k], \tag{3.13}$$

where

$$\begin{aligned} \boldsymbol{\alpha}[k] &\sim N\left(\mathbf{0},\ \mathbf{Q_d}\right),\ \mathbf{Q_d} = \int_{\tau=0}^{T} e^{\mathbf{A}\tau}\mathbf{Q}e^{\mathbf{A}^T\tau}d\tau, \\ \boldsymbol{\varepsilon}[k] &\sim N\left(\mathbf{0},\ \mathbf{R_d}\right),\ \mathbf{R_d} = \mathbf{R}. \end{aligned} \tag{3.14}$$

Recursively substituting (3.12) into (3.13), the general expression for the measurement column vector at time k can be represented as

$$\begin{aligned} \mathbf{y}[k] &= \mathbf{C}(e^{\mathbf{A}T})^{k-1}\mathbf{x}[1] + \sum_{l=1}^{k-1}\mathbf{C}(e^{\mathbf{A}T})^{l-1}\mathbf{A}^{-1}(e^{\mathbf{A}T} - \mathbf{I})\mathbf{B}\mathbf{u}[k-l] + \boldsymbol{\varepsilon}[k] \\ &= \mathbf{y_x}[k] + \mathbf{y_u}[k] + \mathbf{y}_{\boldsymbol{\varepsilon}}[k], \end{aligned} \tag{3.15}$$

where $\mathbf{x}[1]$ stands for the first system state in the training data, and $\mathbf{u}[\cdot]$ represents the inputs at each time step before time k.

To generalize the proof, we further assume: a) Each measurement represents one PMU; b) A total number of N measurements are analyzed, each having n samples for training, i.e., $\mathbf{Y}(t_0) \in \mathbb{R}^{n\times N}$. Therefore, the k^{th} sample/row of $\mathbf{Y}$ can be represented as $\mathbf{Y}(k) := [y^{(1)}[k],\ \ldots,\ y^{(N)}[k]] \in \mathbb{R}^{1\times N}$. Denote the observation matrix as $\mathbf{C} := [\mathbf{c^{(1)}},\ \ldots,\ \mathbf{c^{(N)}}]^T$, where $\mathbf{c^{(i)}} := [c_1^{(i)},\ \ldots,\ c_M^{(i)}]$. M is the total number of system states, usually huge and unknown in reality. Correspondingly, $y^{(i)}[k] = \mathbf{c^{(i)}}\mathbf{x}[k] + \boldsymbol{\varepsilon^{(i)}}[k]$, $i = 1,\ \ldots,\ N$.

In order to prove the capability of the proposed algorithm for early anomaly

detection, assume: (i) All the PMU data for adaptive training are under normal operating conditions. Equivalently, (i.1) $\mathbf{u}[k] = \mathbf{u_0}$ is a constant input vector for $k = 0, \ldots, N-1$; (i.2) the initial condition $\mathbf{x}[1]$ stays the same; (i.3) the system matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ stay the same. (ii) Only one system anomaly[¶] occurs at time $t > N+1$.

Using (3.15), the general form for the i^{th} measurement/column in $\mathbf{Y}$ can be represented as

$$\begin{aligned}
\mathbf{y^{(i)}} &= \begin{bmatrix} y^{(i)}[1] \\ y^{(i)}[2] \\ \vdots \\ y^{(i)}[n] \end{bmatrix} = \begin{bmatrix} \mathbf{c^{(i)}} \\ \mathbf{c^{(i)}} e^{\mathbf{A}T} \\ \vdots \\ \mathbf{c^{(i)}} \left(e^{\mathbf{A}T}\right)^{n-1} \end{bmatrix} \mathbf{x}[1] + \begin{bmatrix} \boldsymbol{\varepsilon}^{\mathbf{(i)}}[1] \\ \boldsymbol{\varepsilon}^{\mathbf{(i)}}[2] \\ \vdots \\ \boldsymbol{\varepsilon}^{\mathbf{(i)}}[n] \end{bmatrix} \\
&+ \begin{bmatrix} \mathbf{c^{(i)}_{u,\,1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,1}} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,2}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,2}} & \mathbf{c^{(i)}_{u,\,3}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,1}} & \mathbf{c^{(i)}_{u,\,2}} & c^{(i)}_{u,\,3} & \cdots & \mathbf{c^{(i)}_{u,\,n-1}} \end{bmatrix} \begin{bmatrix} \mathbf{u_0} \\ \mathbf{u_0} \\ \mathbf{u_0} \\ \mathbf{u_0} \\ \vdots \\ \mathbf{u_0} \end{bmatrix} \\
&= \mathbf{c^{(i)}_x} \mathbf{x}[1] + \mathbf{y^{(i)}_{\boldsymbol{\varepsilon}}} + \mathbf{c^{(i)}_u} \mathbf{U_0}, \qquad (3.16)
\end{aligned}$$

where $\mathbf{c^{(i)}_{u,\,j}} = \mathbf{c^{(i)}} \left(e^{\mathbf{A}T}\right)^{j-1} \mathbf{A}^{-1} \left(e^{\mathbf{A}T} - \mathbf{I}\right) \mathbf{B}$.

Without loss of generality, assume the m basis vectors in $\mathbf{Y_B}$ are the first m

[¶]We only consider the detection of a single anomaly using the early anomaly detection algorithm. The analysis and detection of multiple anomalies or cascading anomalies is a future avenue of research.

columns in $\mathbf{Y}(t_0)$. Therefore, using (3.3) and (3.4), $\mathbf{y^{(i)}}$ can be represented as

$$\begin{aligned}\mathbf{y^{(i)}} &\approx \sum_{j=1}^{m} v_j^{(i)} \cdot \mathbf{y_b^{(j)}} \\ &\approx \sum_{j=1}^{m} v_j^{(i)} [\mathbf{c_x^{(j)}} \mathbf{x}[1] + \mathbf{y_\varepsilon^{(j)}} + \mathbf{c_u^{(j)}} \mathbf{U_0}] \\ &= \mathbf{c_x^{(i)}} \mathbf{x}[1] + \mathbf{y_\varepsilon^{(i)}} + \mathbf{c_u^{(i)}} \mathbf{U_0},\end{aligned} \tag{3.17}$$

where $i = m+1, \ \ldots, N$. Equivalently,

$$\begin{aligned}&[\mathbf{c_x^{(i)}} - \sum_{j=1}^{m} v_j^{(i)} \mathbf{c_x^{(j)}}] \mathbf{x}[1] + [\mathbf{y_\varepsilon^{(i)}} - \sum_{j=1}^{m} v_j^{(i)} \mathbf{y_\varepsilon^{(j)}}] + [\mathbf{c_u^{(i)}} - \sum_{j=1}^{m} v_j^{(i)} \mathbf{c_u^{(j)}}] \mathbf{U_0} \\ &= \mathbf{\Delta c_x} \mathbf{x}[1] + \mathbf{\Delta y_\varepsilon} + \mathbf{\Delta c_u} \mathbf{U_0} \approx \mathbf{0}.\end{aligned} \tag{3.18}$$

As stated in (3.3), the $\mathbf{v^{(i)}}$'s are calculated by minimizing the squared error. Assume the calculation of $\mathbf{v^{(i)}}$'s is of absolute accuracy. Therefore, (I) the errors in the calculations of the $\mathbf{v^{(i)}}$'s are zero; (II) $e_{normal}^{(i)}$ in (3.6) is almost zero. Consequently, for high dimensional training data, the three terms in (3.18) can be assumed to be zero, respectively, i.e.,

$$\mathbf{\Delta c_x} \mathbf{x}[1] \approx \mathbf{0}, \ \mathbf{\Delta y_\varepsilon} \approx \mathbf{0}, \ \mathbf{\Delta c_u} \mathbf{U_0} \approx \mathbf{0}. \tag{3.19}$$

Now using (3.18) and (3.19), we will prove the capability of the early anomaly detection algorithm to detect the following three types of system anomalies:

Control Input Changes: For the control inputs $\mathbf{U_0}$ in (3.18), there are $n \times M$ linear equalities. These equalities, which are not necessarily linearly independent, form an over-determined condition. Under this over-determined condition, the initial input vector $\mathbf{u_0}$ can be theoretically calculated by minimizing the squared error.

Under normal operating conditions, $\mathbf{\Delta c_u U_0} \approx \mathbf{0}$ holds from (3.19). When one of the control inputs changes, the new input vector $\mathbf{U_{new}}$ will not lie in the null space of $\mathbf{\Delta c_u}$. Consequently, a large nonzero term $\mathbf{\Delta c_u U_{new}}$ will violate the zero approximation of (3.18), and thus impact the approximation error (3.5).

Initial Condition Changes: Consider the term related to $\mathbf{x}[1]$ in (3.18). There are n linear equalities to solve for $\mathbf{x}[1]$, in an over-determined manner, $n \gg M$. Under normal operating conditions, $\mathbf{\Delta c_x x}[1] \approx \mathbf{0}$ holds as assumed in (3.19). A change of the initial condition will make the new condition $\mathbf{x}[1]_{\mathbf{new}}$ lie outside the null space of $\mathbf{\Delta c_x}$. This will result in a large nonzero term $\mathbf{\Delta c_x x}[1]_{\mathbf{new}}$, which violates the zero approximation of (3.18) and results in a large approximation error in (3.5).

System Topology Changes: During normal operating conditions, $\mathbf{x}[1]$ and $\mathbf{U_0}$ can be theoretically calculated by the over-determined $(n + n \times M)$ equalities. In other words, they lie in the null space of $\mathbf{\Delta c_x}$ and $\mathbf{\Delta c_u}$, respectively. A change of topology from $\mathbf{A}$ into $\mathbf{A_{new}}$ will yield changes of $\mathbf{\Delta c_{x,\ new}}$ and $\mathbf{\Delta c_{u,\ new}}$, as shown in (3.16). These changes will further induce changes in the corresponding null spaces, in which $\mathbf{x}[1]$ and $\mathbf{U_0}$ will consequently not lie. As a result, a large nonzero term $(\mathbf{\Delta c_{x,\ new}} \mathbf{x}[1] + \mathbf{\Delta c_{u,\ new}} \mathbf{U_0})$ will violate the zero approximation of (3.18), and the approximation error (3.5) will be large.

For the above three types of system anomalies, the occurrence of any one anomaly will result in a nonzero approximation error (3.5), which serves as the numerator of the anomaly indicator $\eta(t)$ in (3.6). With an almost zero denominator $e^{(i)}_{normal}$ calculated from normal operating conditions, $\eta(t)$ will become huge at the occurrence of any one of the system anomalies.

For some selected non-pilot PMUs, historical data with known system anomalies can be utilized to calculate the system-dependent threshold γ. Whenever $\left|\eta(t)^{(i)}\right| \geq \gamma$ in (3.7), a system anomaly will be issued, and an alert will be declared for the purpose

of further corrective control. □

Remark 1. *The achievement of the early detection depends on the choice of the system-sensitive threshold. With numerous simulations and engineering experience, a threshold of* $\gamma = 10$ *is effective for an early detection.*

3.3.3 Accuracy of Early Anomaly Detection

It is worthwhile to discuss the accuracy and robustness of the early anomaly detection algorithm.

3.3.3.1 Mis-detection (True-Negative)

Because PCA is a linear dimensionality reduction method, the deterioration of the approximation becomes severe in face of system anomalies. Therefore, with the proposed anomaly indicator based on the approximation error, the online anomaly detection yields high robustness. With the real PMU data from Texas, Eastern and Western Interconnections and synthetic PMU data from power system simulator for engineering (PSS/E), there is no mis-detection occurring.

3.3.3.2 False Alarm (False-Positive)

From previous studies with both real and synthetic PMU data, the proposed indicator has created false alarms when it encounter 1) missing data issues and 2) system high-stress conditions. Therefore, it is assumed that the PMU data utilized here are pre-processed without any missing data issues. When the system is in high stressed conditions, the high fluctuations in the PMU data may result in false alarms. However, considering the low probability of the occurrence of high-stressed conditions, the proposed algorithm still has high accuracy and robustness in achieving the detection of system anomalies.

3.3.4 Choice of Measurements

In this part, we will discuss how to choose appropriate measurements for the proposed algorithm.

PCA aims at finding out the low-dimensional embedding from the high-dimensional variables by preserving the most variances. As illustrated in Figure 3.9, PCA is utilized to determine both the pilot PMUs and the linear regression coefficients, which are essential to achieve an accurate online anomaly detection. Therefore, choosing a correct set of variables is significant for the performance of the proposed algorithm.

Commercial-grade PMU is capable of providing 20 measurements at each sample [40], including bus frequency, real power, reactive power, sequence and phase values of voltage and current, etc. In the following, two possible choices of variables will be discussed, of which the latter is chosen in this dissertation.

3.3.4.1 Measurements from Multiple Categories

Assume a *measurement category* corresponds to one kind of power system variables, such as bus frequency, voltage magnitude, etc.. This case discusses that such mixed categories of measurements will deteriorate the performance of the mode estimation.

Consider the physical meanings of power system variables. Each measurement category has a unique variance profile. For example, bus frequency, as a global variable, has a similar behavior throughout the whole power grid, yielding a small variance. For voltage magnitude, as a local variable, it has different profiles on different voltage-levels, resulting in a large variance. This can be demonstrated from the separation of the radials in the loading plots in Figures 3.1(b), 3.2(b), 3.4(b), and 3.5(b). Therefore, combination of multiple categories will change the variance of each, even though these categories are sampled by the same set of PMUs. The performance

of PCA on such combined categories will be deteriorated, and the accuracy of online anomaly detection will be further affected.

3.3.4.2 Measurements from Single Category

Based on the discussion above, in this dissertation, we apply measurements from a single category for the proposed algorithm. Considering the variance differences among the categories of power system variables, we conduct PCA on each category of measurements independently. In other words, we assume that at each round of analysis, $N := p$.

3.4 Numerical Examples

In this section, we illustrate the efficacy of the PMU-based early anomaly detection algorithm, including the dimensionality reduction, the adaptive training, and the early anomaly detection. Both synthetic PMU data generated from Siemens PSS/E [82] and real PMU data from Western and Texas Interconnections are utilized.

3.4.1 Linear Dimensionality Reduction

Synchrophasor data from normal operating conditions are utilized in the adaptive training procedure. Assume that the length of the training data is $T_{trn}^{PSSE} = 250$ s for PSS/E PMU data, $T_{trn}^{west} = 300$ s for the western PMU data, and $T_{trn}^{TX} = 250$ s for the Texas PMU data[‖].

3.4.1.1 Dimensionality Reduction of Synthetic PSS/E Data

A 23-bus 6-generator system in PSS/E is utilized to generate PMU data. Figure 3.10 serves as a demonstration of the system topology, which is not necessary in the early anomaly detection algorithm. Table 3.1[**] lists the dynamic models [82] for

[‖]In reality, more synchrophasor data under normal operating conditions could be utilized in the adaptive training to obtain more accurate and robust training models.

[**]GENROU: Round rotor generator model.
GENSAL: Salient pole generator model.

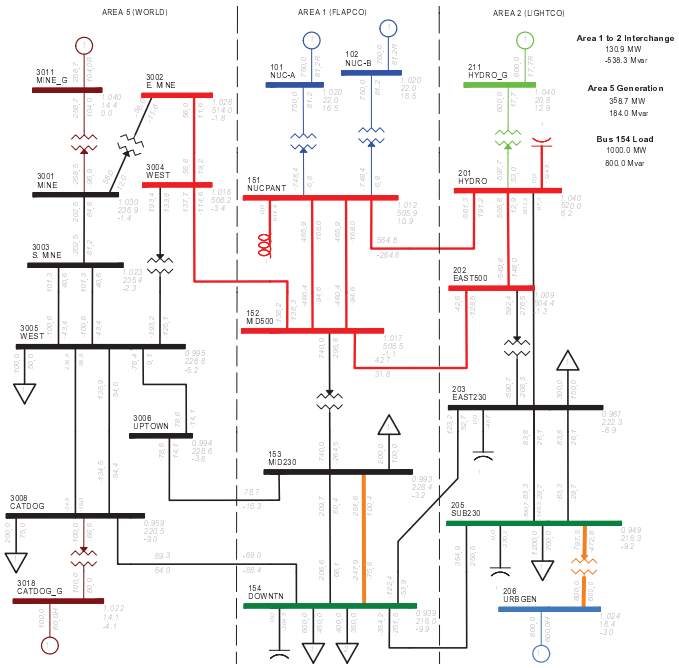


Figure 3.10: Topology of PSS/E 23-bus system [82].

the 6 generators employed in the PSS/E system. Assume each bus has one PMU installed, and the sampling rate is 30 Hz. To mimic the industry-grade PMUs, noise is added to the synthetic training data so that the signal-to-noise ratio (SNR) is 92dB.

The cumulative variance calculated from PCA is shown in Figure 3.11. The results of the dimensionality reduction analysis are shown in Table 3.2 including 1) the original number of PMUs, indicating the high dimension of PMU data, 2) the threshold of the cumulative variance, 3) the number of pilot PMUs, and 4) the basis matrix. As can be observed, by selecting certain threshold for the cumulative

IEEET1: 1968 IEEE type 1 excitation system model.
SCRX: Bus or solid fed SCR bridge excitation system model.
SEXS: Simplified excitation system model.
TGOV1: Steam turbine-governor model.
HYGOV: Hydro turbine-governor model.
N/A: No model for the component.

Table 3.1: Dynamic models in PSS/E system

Bus No.	Generator Model	Exciter Model	Turbine/Governor Model
101	GENROU	IEEET1	TGOV1
102	GENROU	IEEET1	TGOV1
206	GENROU	IEEET1	TGOV1
211	GENSAL	SCRX	HYGOV
3011	GENROU	SEXS	N/A
3018	GENROU	SEXS	N/A

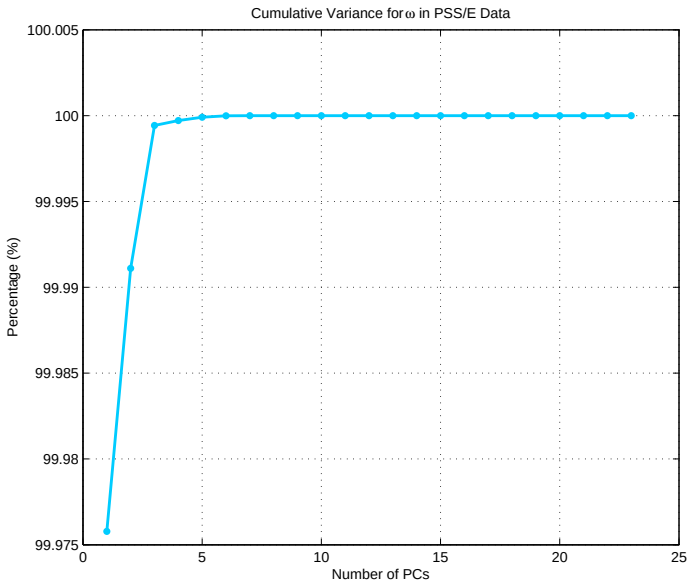


Figure 3.11: Cumulative variance preserved by PCs for PSS/E data.

variances, two pilot PMUs can be identified in the basis matrix for online anomaly detection.

Table 3.2: Results of dimensionality reduction analysis

Anomaly	Original No. of PMUs	Threshold of Cumulative Variance	No. of Pilot PMUs	Basis Matrix
PSS/E	23	99.99%	2	$\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B},\ \mathbf{PSS/E}} = [\omega_{206},\ \omega_{102}]$
Oscillation	44	99.65%	2	$\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B},\ \mathbf{OS}} = [\omega_{10},\ \omega_{18}]$
Line Fault	45	99.85%	2	$\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B},\ \mathbf{LF}} = [\omega_{10},\ \omega_{26}]$
Unit Tripping 1	7	99.9%	2	$\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B},\ \mathbf{UT1}} = [\omega_{5},\ \omega_{3}]$
Unit Tripping 1	7	99.97%	2	$\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B},\ \mathbf{UT2}} = [\omega_{5},\ \omega_{3}]$

3.4.1.2 Dimensionality Reduction of Real PMU Data

Three types of anomalies, oscillation from western data, line faults from western data, and unit tripping from Texas data, are utilized as illustrations. Neither system topology nor model could be obtained due to the confidentiality of the interconnected areas and the modeling complexity of the system components. The cumulative variances of ω for the anomalies are shown in Figures 3.12-3.15. The results of the dimensionality reduction analysis are shown in Table 3.2 with specified basis matrices.

3.4.2 Online Anomaly Detection

In this section, the eventful data of the above anomalies are utilized to demonstrate the early anomaly detection algorithm with respect to its purely data-driven capability, i.e., *without* the knowledge of system topology or model.

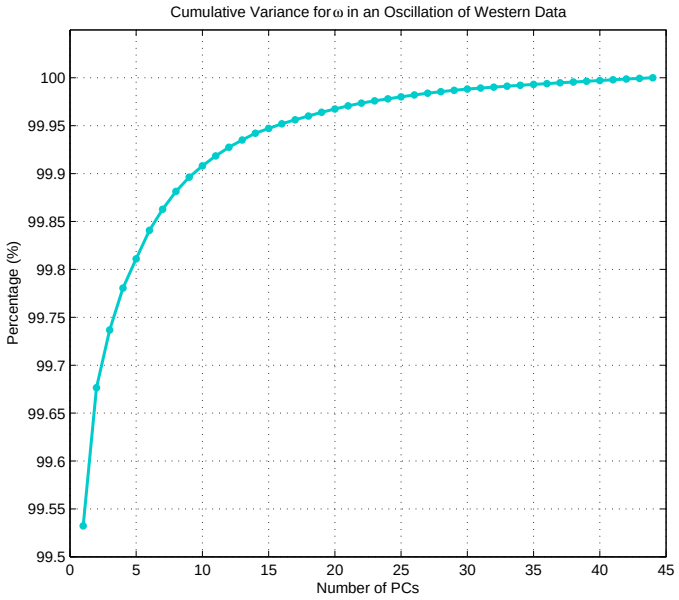


Figure 3.12: Cumulative variance of an oscillation.

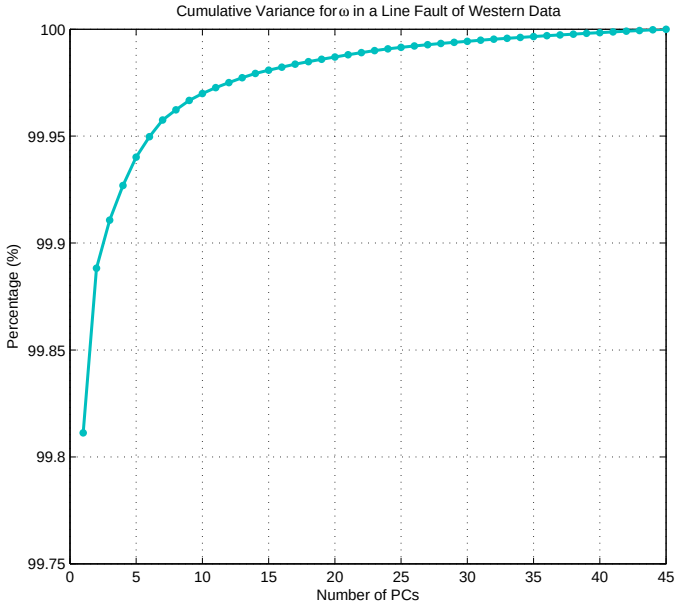


Figure 3.13: Cumulative variance of a line fault.

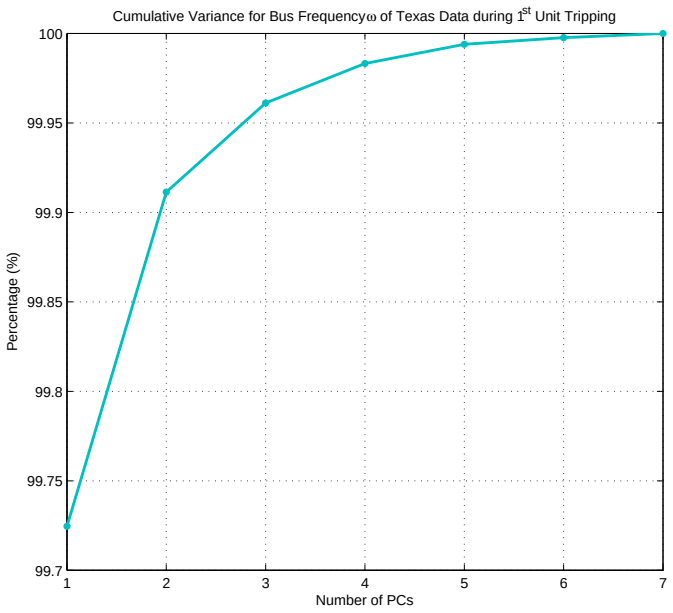


Figure 3.14: Cumulative variance of the 1^{st} unit tripping.

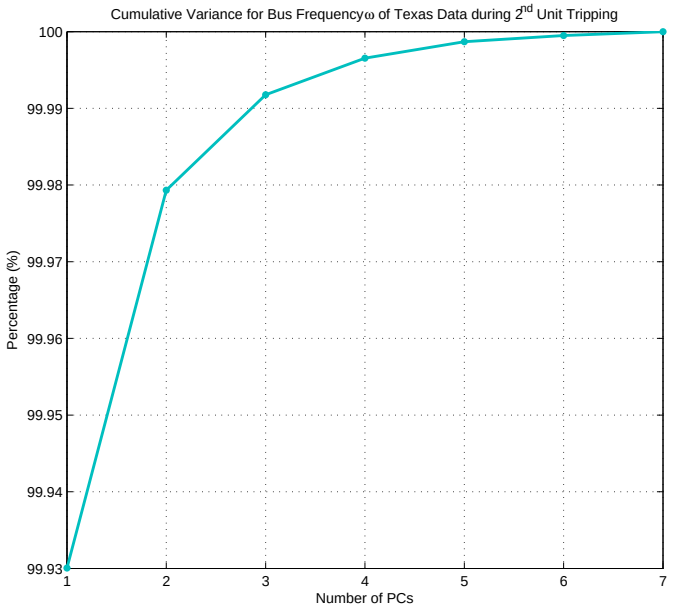


Figure 3.15: Cumulative variance of the 2^{nd} unit tripping.

3.4.2.1 Online Anomaly Detection of Synthetic PSS/E Data

Three types of system anomalies, line tripping, unit tripping, and control input change, are simulated in the PSS/E 23-bus system, with the anomaly details shown in Figure 3.16.

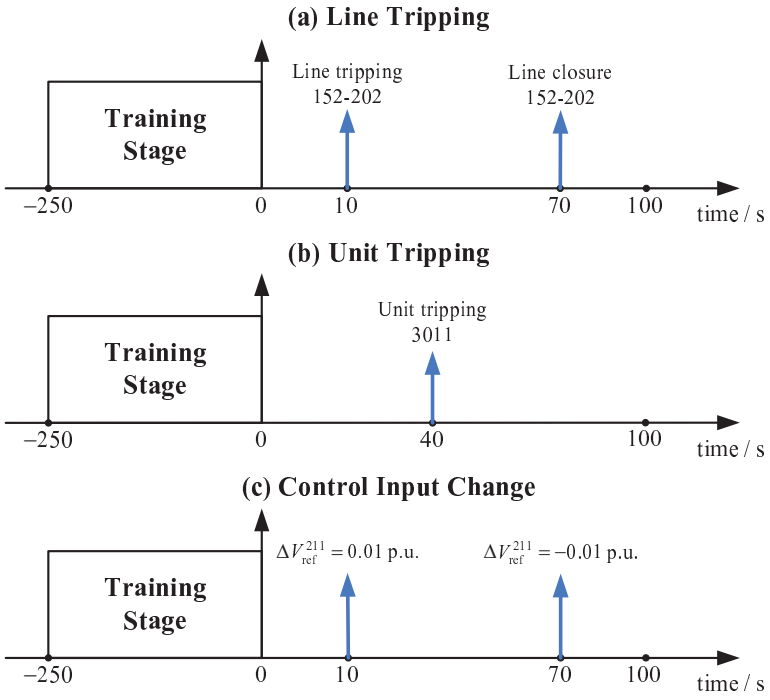


Figure 3.16: Timeline of three simulated system anomalies in PSS/E.

In order to demonstrate the efficacy of the adaptive training, assume: 1) The training procedure conducted in Section 3.4.1.1 works for all the three types of system anomalies before the first adaptive training takes place; 2) It takes 10 s for the system to return to normal operating conditions after an anomaly; 3) The updating period is $T_{up} = 40$ s; 4) If update is needed, the retraining period is the same as the original training period, i.e., $T_{retrn} = T_{trn} = 250$ s.

3.4.2.1.1 Line Tripping

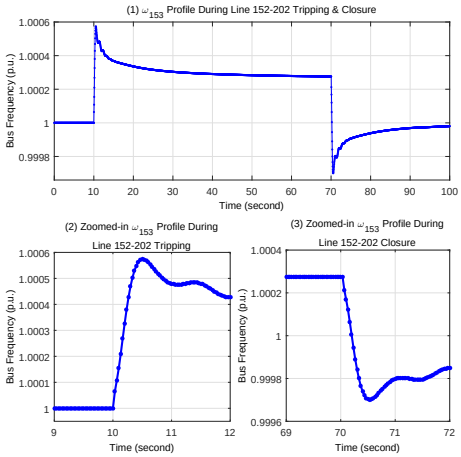
As shown in Figure 3.16(a), assume the transmission line connecting buses 152 and 202 (Line 152-202) is tripped at $t = 10$ s, following by a closure of Line 152-202 at $t = 70$ s. The total length of data is 100 s.

The bus frequency profile for bus 153 during the anomalies is shown in Figure 3.17(a). As can be observed, it takes about 10 s for the system to recover from either anomaly, i.e., the system recovers to normal operating conditions at $t = 20$ s and $t = 80$ s, respectively. With $T_{up} = 40$ s, the training model is updated at time $t = 60$ s with the latest 250 s data. The updated basis matrix is $\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B,PSS/E}} = [\omega_{3011},\ \omega_{101}]$. The Line 152-202 closure is detected using the updated basis matrices.

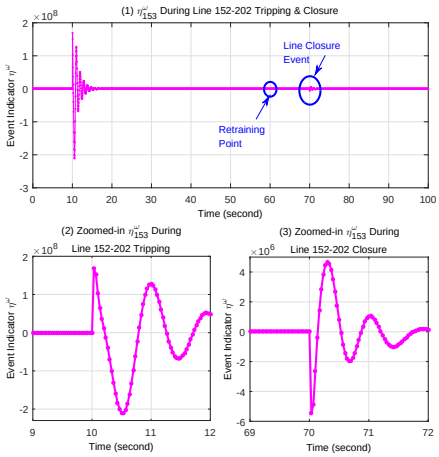
Figure 3.17(b) illustrates the anomaly indicator η^{ω}_{153} of bus 153, which can detect both anomalies. A zoomed-in view of the early detection of the tripping is presented in Figures 3.17(b.2) and 3.17(b.3), showing the capability to detect the anomaly almost instantly, within 40 ms. Line 152-202 tripping at $t = 10$ s results in a huge value of η^{ω}_{153} at the next step $t = 10.033$ s. This indicates a large approximation error of the linear representation. Comparatively, as can be seen from the frequency profile in Figure 3.17(a.2), at time $t = 10.033$ s, the bus frequency deviation is $\Delta\omega_{153} \approx 0.00005$ p.u., which is too small to be identified as an anomaly. When a relatively large deviation $\Delta\omega_{153} \approx 0.0004$ p.u. is detected, it is already 250ms later than the occurrence time of the anomaly. Similar results can be observed for the line closure at time $t = 70$ s. In this sense, the advantage of the proposed algorithm is illustrated.

Another observation is from the comparison of Figures 3.17(b.2) and 3.17(b.3): the maximum deviation of the anomaly indicator in Figure 3.17(b.3) is much smaller than that in Figure 3.17(b.2). The reason comes from the adaptive training, i.e.,

the retraining takes the eventful data into consideration, and therefore improves the accuracy of the training model. However, from Figure 3.17(b.3), the capability to early detect system anomalies is not affected by this improvement. In reality, two system anomalies will not occur as close as those in this case. Therefore, the retraining data will not always contain the eventful data. In addition, by choosing an appropriate length of training data, this kind of improvement can also be avoided, and the training model can be accurate and robust enough to detect the anomaly at an early stage, as shown in Figure 3.17(b.3).



(a) ω_{153} profile.



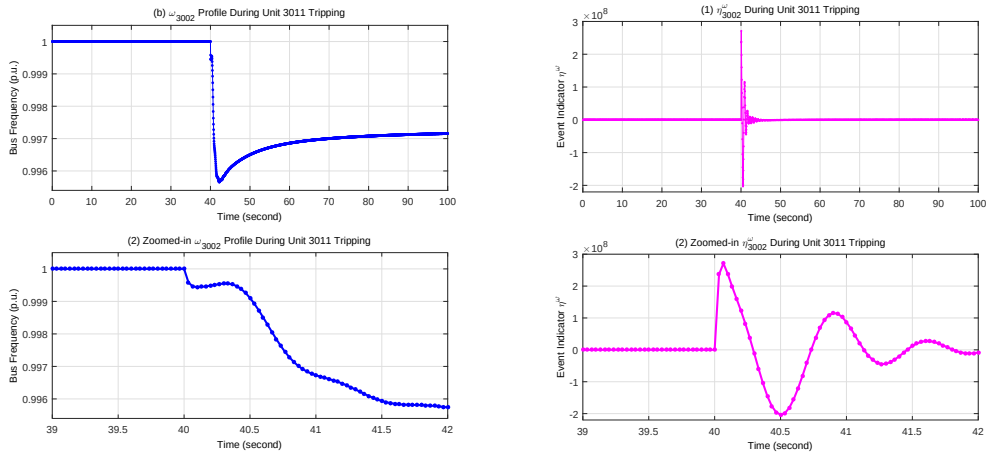
(b) η^{ω}_{153}.

Figure 3.17: Line 152-202 tripping & closure in PSS/E.

3.4.2.1.2 Unit Tripping

As shown in Figure 3.16(b), a unit at bus 3011 is tripped at $t = 40$ s. Figure 3.18(a) shows the bus frequency profile of bus 3002, with the anomaly indicator η^{ω}_{3002} in Figure 3.18(b).

In Figure 3.18(a.2), from the bus frequency deviation $\Delta\omega_{3002} \approx 0.0005$ p.u. at



(a) ω_{3002} profile. (b) η^{ω}_{3002}.

Figure 3.18: Unit 3011 tripping in PSS/E.

$t = 40.033$ s, it is difficult to directly detect the unit tripping. However, from Figure 3.18(b.2), unit 3011 tripping yields a huge value of the anomaly indicator η^{ω}_{3002} provided by bus 3002 at $t = 40.033$ s. This clearly demonstrates that the proposed anomaly indicator is capable of early anomaly detection by magnifying the difference between the quantities for the normal condition and the contingency.

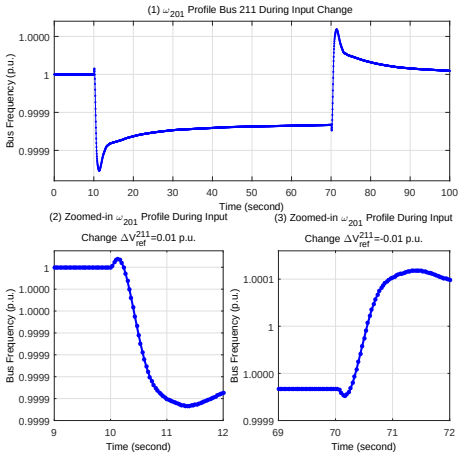
3.4.2.1.3 Control Input Change

As shown in Figure 3.16(c), the voltage regulator set-point of bus 211 is changed by 0.01 p.u. and -0.01 p.u. at $t = 10$ s and $t = 70$ s, respectively. In this case, the training models are updated at time $t = 60$ s with the latest $T_{retrn} = 250$ s data. The updated basis matrix is $\mathbf{Y}^{\boldsymbol{\omega}}_{\mathbf{B,PSS/E}} = [\omega_{211},\ \omega_{3011}]$. The second anomaly with $\Delta V^{211}_{ref} = -0.01$ p.u. is detected using the updated basis matrix.

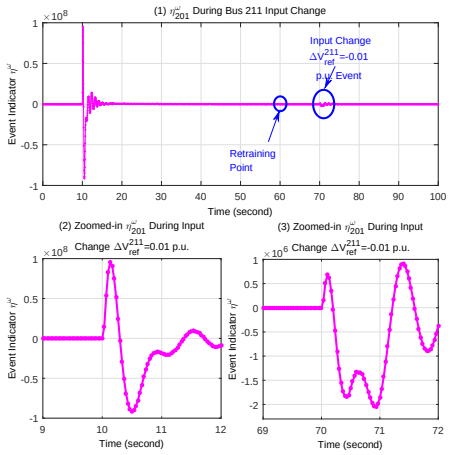
In Figure 3.19(b.1), the anomaly indicator η^{ω}_{201} provided by bus 201 is capable of indicating both anomalies. From Figure 3.19(b.2), the control input change by $\Delta V^{211}_{ref} = 0.01$ p.u. is detected around $t = 10.1$ s with $\eta^{\omega}_{201} \approx 10^8$, while from the bus frequency profile in Figure 3.19(a.2), the bus frequency deviation $\Delta\omega_{201} < 0.0001$

p.u. at $t = 10.1$ s cannot be detected efficiently. Similar results can be observed from Figures 3.19(a.3) and 3.19(b.3).

In this case, because of the adaptive training, the maximum deviation of the anomaly indicator in Figure 3.19(b.3) is smaller than that in Figure 3.19(b.2). However, this does not impact the efficacy and capability of the anomaly indicator for the early anomaly detection, as shown in Figure 3.19(b.3).



(a) ω_{201} profile.



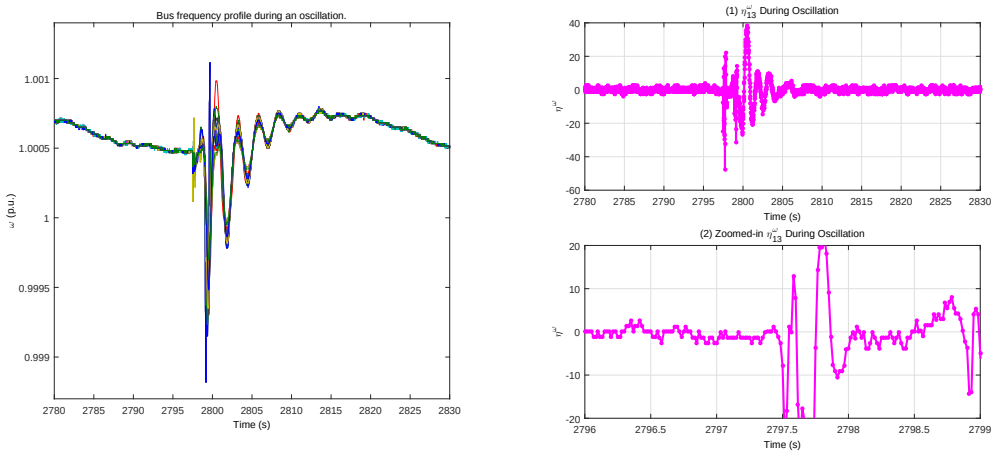
(b) η_{201}^{ω}.

Figure 3.19: Bus 211 input change in PSS/E.

3.4.2.2 Online Detection of Oscillation

The oscillation shown in Figure 3.20(a) is induced by the uncertainties from the wind farm generations in the western interconnection, while Figure 3.20(b) illustrates the anomaly indicator of PMU #13. As can be observed from Figure 3.20(b.2), with a threshold of $\gamma = 10$, the anomaly can be detected at time $t = 2798.52$ s. However, at the same time, from Figure 3.20(a), the frequency deviation from the normal operating condition is less than 0.0001 p.u., which is too small to be considered as

an anomaly. Therefore, the benefit of utilizing the proposed indicator is shown for an effective anomaly detection at an earlier stage than would be possible by monitoring the raw measurements.

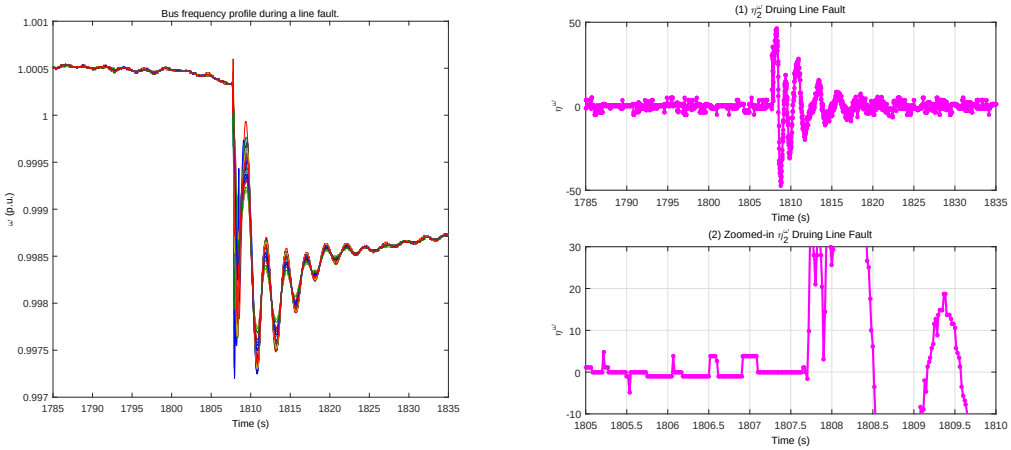


(a) ω profile. (b) η_{13}^{ω}.

Figure 3.20: Oscillation in Western U.S..

3.4.2.3 Online Detection of Line Fault

Figure 3.21(a) illustrates the bus frequency profile during a line fault, while the anomaly indicator of PMU #2 is shown in Figure 3.21(b). As can be observed from Figure 3.21(b.2), with a threshold of $\gamma = 10$, the anomaly can be detected at time $t = 1807.72$ s. However, at the same time, from Figure 3.21(a), the frequency deviation from the normal operating condition is only 0.0002 p.u., which is too small to be considered as an anomaly. Again, the benefit of utilizing the proposed indicator is shown for an effective anomaly detection.



(a) ω profile. (b) η_2^ω.

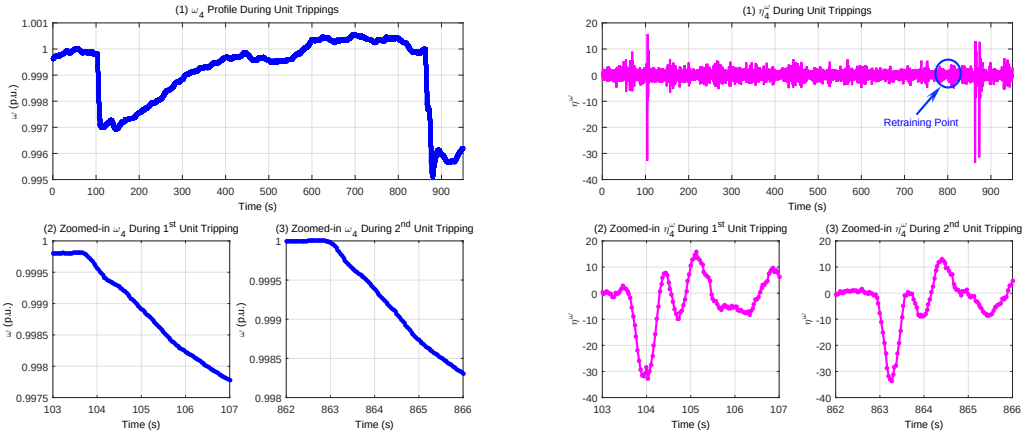
Figure 3.21: Line fault in Western U.S..

3.4.2.4 Online Detection of Unit Tripping

As can be observed from Figure 3.22(a), there are two unit tripping occurring around $t = 104$ s and $t = 863$ s. We demonstrate the adaptive training scheme in this case.

After the first anomaly, it takes about 300 s for the system to recover to normal operating conditions. In this case, assume the updating period is $T_{up} = 100$ s and the retraining period is $T_{retrn} = 250$ s. Therefore, according to the early anomaly detection algorithm, the adaptive training results in Section 3.4.1 will work only for the first anomaly. The latest training model before the second anomaly will be updated at time $t = 800$ s with the latest 250 s data. The detection of the second anomaly will be achieved using the latest training model. In this case, the retraining data do not contain any anomalies and therefore can better demonstrate the efficacy of the adaptive training. The cumulative variances preserved by PCs for the two times of training are shown in Figures 3.14 and 3.15.

The anomaly indicator of PMU #4, η_4^{ω}, is shown in Figure 3.22(b). In the zoomed-in Figures 3.22(b.2) and 3.22(b.3), the changes of system operating conditions can be detected at $t = 103.7$ s and $t = 863$ s, respectively with a threshold of $\gamma = 10$. However, from the bus frequency profile in Figure 3.22(a.2), the bus frequency deviation $\Delta\omega_4 \approx 0.00001$ p.u. at $t = 103.7$ s is too small to be detected early or accurately. Similar conclusions can be drawn for $\Delta\omega_4 \approx 0.00001$ p.u. at $t = 863$ s in Figure 3.22(a.3).



(a) ω_4 profile. (b) η_4^{ω}.

Figure 3.22: Unit trippings in TX.

Table 3.3 summarizes the comparisons of detection time for the simulated cases. From the above analysis, the efficacy of the proposed algorithm is demonstrated for the considered anomalies. Further studies are conducted to test the robustness and accuracy of the detection.

Table 3.3: Comparison of detection time for bus frequency cases

Case	Detection Time from Direct Profile	Detection Time from Proposed Algorithm
PSS/E Line 152-202 Tripping	10.25	**10.033**
PSS/E Unit 3011 Tripping	40.5	**40.033**
PSS/E V_{ref}^{211} Input Change	11	**10.1**
Oscillation	2799.2	**2798.52**
Line Fault	1807.83	**1807.72**
Unit Tripping 1	104.5	**104**
Unit Tripping 2	864	**863**

4. PMU-BASED ROBUST MONITORING OF LOW-FREQUENCY OSCILLATIONS*

4.1 Introduction

The impact of power system low-frequency oscillations has been discussed in Section 1. The severe system separation and blackout bring special attentions to the analysis of power system low-frequency oscillations [4, 5, 6, 7]. However, traditional approaches requires a detailed dynamical representation of the system to conduct modal analysis. The increasing penetration of spatially dispersed and temporally variable resources into power grids introduces difficulties in achieving an accurate dynamic model of the system. This further deteriorates the accuracy of model-based oscillation analysis, and therefore a data-based analytical approach is of great need.

Synchrophasor measurement, with the high sampling rate and the synchronization provided by GPS, becomes more and more attractive in monitoring and control of low-frequency oscillations in power grids [10, 11, 12]. Three data-driven methods, subspace identification, spectral independent component analysis, and wavelet transform, are compared in [44] for damping estimation of electromechanical oscillations. [45] utilizes Fourier spectral analysis on synchrophasor data to estimate the eigenvalues for monitoring inter-area oscillation.

Although the existing methods perform well in mode estimation for power system low-frequency oscillations, there are still several gaps between low-frequency

*This section is in part a reprint of the material in the papers: (1) Reprinted with permission from Yang Chen, Le Xie, and P. R. Kumar, "Power System Event Classification via Dimensionality Reduction of Synchrophasor Data," *Sensor Array and Multichannel Signal Processing Workshop, 2014. SAM 2014. 8th IEEE*, pp. 57-60, 2014. Copyright 2014, IEEE. (2) Reprinted with permission from Yang Chen, Harish Chintakunta, Le Xie, Yuliy M. Baryshnikov, and P. R. Kumar, "Robust Detection and Mode Estimation of Power System Low-frequency Oscillations using Synchrophasor Data," *IEEE Transactions on Power Systems*, to be submitted.

oscillation monitoring and its online applications:

- Robust Early Detection of Oscillation. Most existing approaches recursively estimate the modes using PMU measurements. However, recursive estimation by use of uneventful ambient data will increase the computational burden.

- Online Efficient Estimation of Oscillation Parameters. The increasing deployment of synchrophasors has resulted in high-volume PMU data injecting into power grid. Such data will also increase the computational burden. The high measurement noise in such data may also deteriorate the estimation performance.

With such gaps, an effective approach to monitoring power system low-frequency oscillations with robustness against noise using large-scale synchrophasor data is of great need.

In this section, by continuing exploring the low-dimensional characteristics of high-volume synchrophasor data, we propose a robust data-driven framework to enhance the detection and mode estimation of power system low-frequency oscillations, without requiring any knowledge of system model/topology. Scatter plot is first applied to visualize oscillations using the topological patterns. Secondly, PCA is applied to extract the pre-PCA features from the high-volume raw measurements. Then a persistent-homology-based cyclicity response is proposed by using the pre-PCA features for oscillation detection. A pre-defined threshold can be determined by historical eventful PMU data. Cyclicity response with a value exceeding the threshold indicates the occurrence of an oscillation. After an oscillation is detected, PCA is applied again to extract the post-PCA features from the transient oscillatory PMU data. Theoretical justification shows that the post-PCA features retain the modal information that is present in the raw measurements. FFT and Prony analysis are

finally applied to the post-PCA features for mode estimation.

The main contributions of this section are as follows.

- Visualization of power system low-frequency oscillations is presented to differentiate oscillations from other types of anomalies by use of topological patterns.
- The cyclicity response provides high accuracy in detecting low-frequency oscillations, and benefits real-time operations in power systems.
- The mode estimation with post-PCA features preserves robustness of mode estimation against high measurement noise.
- The proposed robust monitoring algorithm is purely online data-driven, requiring no knowledge of system model or topology.

4.2 Scatter Plot based Visualization of Power System Oscillations

In this section, a scatter-plot-based approach is presented for visualization of power system oscillations using the trajectory of measurements from PMUs. Scatter plot is briefly introduced in Figures 3.3(b) and 3.6(b). In this section, we will systematically propose an oscillation visualization algorithm.

The definitions below follow those in Section 3.2. Define the measurement matrix $\mathbf{Y_{n\times N}} = [\mathbf{y^{(1)}}, \ \ldots, \ \mathbf{y^{(N)}}]$ with N measurements of ambient PMU data. Each measurement has n samples constituting a time history, i.e., $\mathbf{y^{(i)}} = [y_1^{(i)}, \ \ldots, \ y_n^{(i)}]^T$, $i = 1, \ \ldots, \ N$.

The proposed oscillation visualization algorithm is described as the following two parts, with the flowchart shown in Figure 4.1.

4.2.1 PCA based Adaptive Training

As the first part, the PCA-based dimensionality reduction is utilized for the adaptive training. The adaptiveness refers to the update of measurement matrix $\mathbf{Y}$

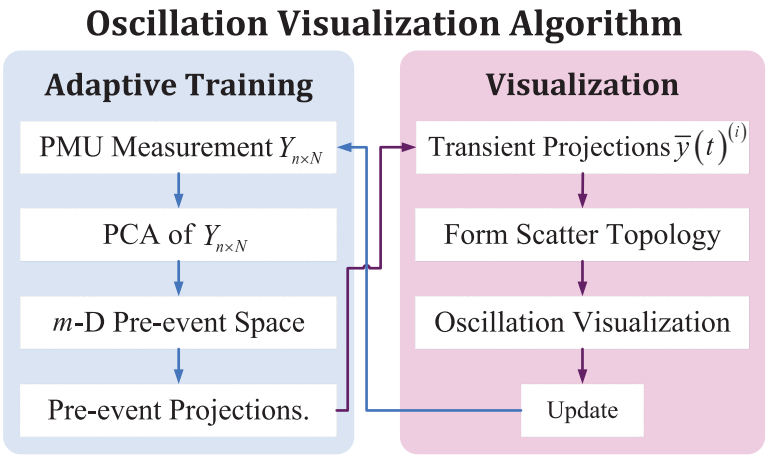


Figure 4.1: Flowchart of oscillation visualization algorithm.

after an oscillation is visualized as shown in Figure 4.1.

1) **Form the PMU measurement matrix Y with ambient data**.

2) **Perform PCA on Y**.

3) **Construct the m-dimensional (m-D) pre-event space**.

Select the top m out of N PCs, which preserve a cumulative variance satisfying $\sum_{i=1}^{m} var_i \geq \tau$. τ is a pre-defined variance threshold and $m \ll N$. From synthetic and realistic PMU data analysis, the value of τ usually satisfies $\tau \geq 99.9\%$ for bus frequency. For the purpose of visualization, m is selected as 2 or 3 in this dissertation.

Then, construct an m-D pre-event space with the bases being the m selected PCs.

4) **Find pre-event projections**.

At each of the n time steps, the original N-dimensional (N-D) sample can be reduced to m-D by a projection from the raw data space to the pre-event space. These n projections are part of the pre-event projections.

Three types of projections onto the pre-event space are defined as follows:

(1) *Pre-event projections*: projections of pre-event data, including training data, and any real-time ambient data before the occurrence of an anomaly.

(2) *Transient projections*: projections of transient data, sampled right after the occurrence of an anomaly, usually within several seconds.

(3) *Post-event projections*: projections of data when the system reaches a new operating condition after the occurrence of an anomaly.

4.2.2 Scatter Plot based Visualization

The pre-event space from the adaptive training will be utilized as prerequisites for the online oscillation visualization using the scatter plot.

1) **Calculate transient projections**.

For the real-time measurement $y(t)^{(i)}$ of the i^{th} PMU, calculate the transient projection $\bar{y}(t)^{(i)}$.

2) **Construct the scatter-plot topology**.

Project all calculated transient projections into the pre-event space to construct the scatter-plot topology.

3) **Visualize the oscillation based on the topological pattern.**

4) **Adaptive update**.

The update signal is sent from the second stage of visualization.

Denote the *update period* as T_{up}. T_{up} is a system-dependent variable, and usually can be chosen as 3-5 minutes. The adaptive update mechanism is designed as follows:

- If there is no oscillation detected[†] during a period of T_{up}, the training procedure is adaptively updated every T_{up} time units with new ambient data.
- If an oscillation is visualized within T_{up}, the training procedure is updated immediately with new ambient data after the system operates at a new condition.

[†]The oscillation detection will be described in Section 4.3 using cyclicity response.

4.3 Cyclicity Response based Oscillation Detection

In this section, we describe *cyclicity response*, which is used to quantify the oscillatory behavior of the power grid. The oscillation detection procedure takes the pre-PCA features as inputs, and calculates the corresponding cyclicity responses as outputs. Cyclicity response is based on:

1. Takens' delay embedding theorem [83], which implies that a mapping $g : X \to \mathbb{R}^2$ from the phase space [84] of an oscillatory system to $\mathbb{R}^2$ contains a cyclic structure, and

2. Persistent homology [85], which quantifies the cyclic structure of a point cloud.

In the context of the application considered here, X represents the phase space of the power grid, and the mapping g represents the projection of PMU measurements on to the two significant PCs, which yields a scatter plot. The premise here is that when the power grid shifts into an oscillatory condition, the projection of the PMU measurements will start to behave cyclically (either circularly or elliptically), which can be detected using persistent homology computation as described in Section 4.3.2.

4.3.1 Delay Embedding Theorem

Consider a discrete-time dynamical system $f : X \to X$, which has a strange attractor [86]. Then for sufficiently large k, and for any generic function $g : X \to \mathbb{R}^k$, the delay embedding theorem states that the map

$$G(x) = \left(g(x), g\left(\phi(x)\right), g\left(\phi^2(x)\right), \ldots, g\left(\phi^{k-1}(x)\right)\right) \tag{4.1}$$

is an embedding. The dimension k is determined by the structure of the attractor. For an attractor with dimension d, $k > 2d$ is usually sufficient. Readers may refer to [83, 87] for a formal treatment of the delay embedding theorem.

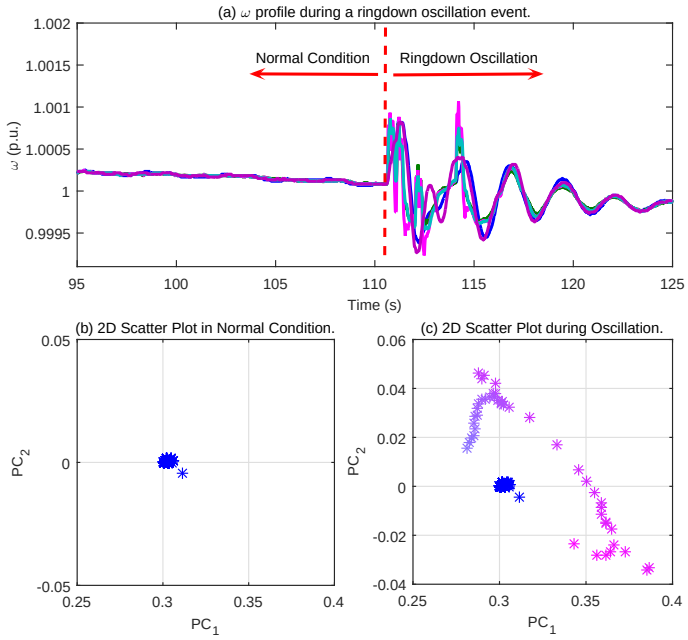


Figure 4.2: Bus frequency profile and scatter plot during a ringdown oscillation.

We model the power grid as a discrete-time dynamical system f, where the time points are determined by the sampling rate of the PMUs. When the power grid is in an oscillatory state, the phase of the system exhibits cyclic behavior, and by the delay embedding theorem, so does the mapping g. Figure 4.2 illustrates this phenomenon in the power grid. Figure 4.2(b) shows the scatter plot when the power grid is not undergoing oscillations, and Figure 4.2(c) shows the case of an oscillation. The task then is to distinguish between these two cases automatically, which is what we aim to accomplish using persistent homology computation.

4.3.2 Persistent Homology and Cyclicity Response

We now give a brief introduction to persistent homology. The purpose here is to provide the reader an intuition for what the tool computes. A formal introduction to the theory is beyond the scope of this dissertation, and we refer the readers to [85] for

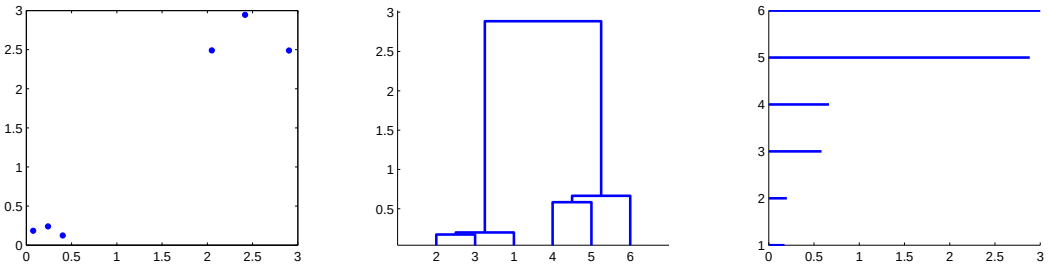
a thorough exposition. *Persistent homology* may be viewed as a generalization of traditional hierarchical clustering to other topological features. We first briefly describe hierarchical clustering, followed by how this is paralleled by persistent homology to cyclic structures.

4.3.2.1 Hierarchical Clustering

Given a point cloud, hierarchical clustering is a traditional tool for obtaining a multi-scale summary of how the points are clustered. Figure 4.3 illustrates single linkage clustering where the output, either a "dendrogram" or a "barcode" as described below, provides a multi-scale summary of the clustering. The algorithm is as follows: for each value ϵ, draw an edge between any two points (v_i, v_j), where the distance between the two points $d(v_i, v_j) \leq \epsilon$. The vertices corresponding to connected components in the resulting graph are the clusters. One can then follow how these connected components merge as we increase the value of ϵ. This pattern in which the connected components merge can be described by either a dendrogram (as in Figure 4.3(b)) or by a barcode (as in Figure 4.3(c)). Each vertical line segment in the dendrogram corresponds to a connected component, and these segments are joined as the connected components merge. Equivalently, we may view the points as connected components "born" at $\epsilon = 0$, and whenever two components merge together, we may say one of the connected components has "died" at that ϵ value. This birth-death pattern is represented in the "barcode" where each bar represents the birth and death of a connected component. For example point cloud in Figure 4.3(a), it is quite apparent from the output that there are two clusters.

4.3.2.2 Persistent Homology

We now describe how to generalize the above hierarchical clustering to cyclic structures. For this purpose, we use the example of the point cloud shown in Figure



(a) Point cloud. (b) Dendrogram. (c) Barcode.

Figure 4.3: Demonstration of hierarchical clustering.

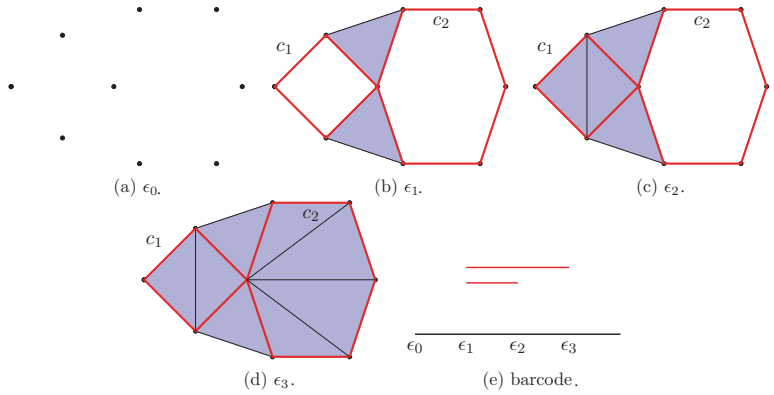


Figure 4.4: Demonstration for the computation of persistent homology.

4.4(a). In the case of clustering described above, we have followed the "birth" and "death" of connected components in the graph. Here, we instead follow the birth and death of cycles.

As in the case of clustering, we increase the value of ϵ, and for each ϵ value, we add all the edges in the Delaunay triangulation [88] whose length is less than or equal to ϵ.

At ϵ_1, we see the birth of two cycles in red: c_1 and c_2. If there is a triangle in the resulting graph, we consider that triangle to be "filled in." Therefore, a cycle is considered "filled in", and therefore dead, if it bounds a set of triangles. We see that c_1 dies at ϵ_2, and c_2 dies at ϵ_3, resulting in the barcode shown in Figure 4.4(e). It is straight forward to infer from the barcode that the point cloud exhibits two cyclic structures, one of which is larger than the other.

Given a point cloud, the *cyclicity response* is defined as the length of the largest bar in the barcode. As evidenced in the above example, the longest bar represents the biggest cyclic structure in the point cloud, and we see a cyclic structure when the system is undergoing oscillations. For the sake of concreteness, we will now provide some algebraic background which gives rise to these bars.

For a triangulation K_ϵ, as in Figure 4.4, one can define an abstract vector space $H_1(K_\epsilon)$, called the first homology space, where the basis elements are equivalence classes of cycles in K_ϵ. All cycles which bound triangles are considered to be equivalent to 0. For example, the equivalence class of $[c_1] \neq 0$ in K_{ϵ_1}, where as $[c_1] = 0$ in K_{ϵ_2}. Also, since ϵ_2 is the smallest value for which $[c_1] = 0$, we say that $[c_1]$ persists in the interval $[\epsilon_1, \epsilon_2)$, and this is represented as bar from ϵ_1 to ϵ_2 in the output. More generally, the homology which persists in the interval $[a, b]$ is given by $Img\left(H_1(K_a) \xrightarrow{i_*} H_1(K_b)\right)$, where i_* is the map induced by the inclusion $i : K_a \to K_b$. In other words, the number of bars in any interval $[a, b]$ in the barcode is equal to the dimension of $Img\left(H_1(K_a) \xrightarrow{i_*} H_1(K_b)\right)$.

From a data analysis perspective, each bar of interval $[a, b]$ in the barcode corresponds to a cyclic structure which is present in the triangulations corresponding to thresholds in that interval. The length of the longest bar corresponds to the largest cyclic structure present in the point cloud, which is the basis of cyclicity response.

For further technical details behind the computation of the barcode, we refer the

readers to [85].

4.4 PCA-enabled Mode Estimation Approach

Based on the proposed cyclicity response in Section 4.3, oscillations can be effectively detected. An efficient approach is of great necessity to further estimate the mode and damping ratio of an oscillation. In this section, a novel mode estimation approach is developed upon the frequency-domain analysis of the post-PCA features extracted from large-scale transient oscillatory data.

4.4.1 Post-PCA on Oscillatory Data

With the increasing deployment of PMUs in the power grid with the high sampling rate, the resulted large amount of PMU data brings new challenges to the online applications in terms of the processing and computational burden.

Considering the issues from the high-volume PMU data, in the proposed mode estimation approach, PCA is applied to the oscillatory PMU measurements after an oscillation is detected using cyclicity response.

Mathematically, PCA aims at finding a low-dimensional embedding from a high-dimensional space by preserving the most variance. Therefore, the modal properties of the low-frequency oscillation do not change through the PCA process. This property of PCA is the reason for the feasibility of the proposed framework that reduces the dimensionality before extracting the modal information. Besides, by conducting singular value decomposition on a zero-mean measurement matrix, PCA has a denoising functionality, which further improves the robustness of the proposed approach with regard to the high measurement noise.

Assume an oscillation occurs at time t_d. A measurement matrix is formulated afterwards with the oscillatory data as $\mathbf{Y_e} := [\mathbf{y^{(1)}}, \ \ldots, \ \mathbf{y^{(N)}}] \in \mathbb{R}^{n \times N}$, which includes a total number of N measurements. Each measurement has n samples

constituting a time history, i.e., $\mathbf{y^{(i)}} := [y_1^{(i)},\ \ldots,\ y_n^{(i)}]^T$, $i = 1,\ \ldots,\ N$.

One can apply PCA on $\mathbf{Y_e}$ to extract the first two PCs as

$$\mathbf{F_1} = (\mathbf{Y_e}(t) - \mu)\,\mathbf{u_1},\ \mathbf{F_2} = (\mathbf{Y_e}(t) - \mu)\,\mathbf{u_2}, \tag{4.2}$$

where the $\mathbf{u_i}$'s, $i = 1, \ldots, N$, are the orthonormal eigenvectors for the covariance matrix of $\mathbf{Y_e}$ corresponding to the nonnegative eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. $\mu = E[\mathbf{Y_e}]$. $\mathbf{F_1}$ is PC_1, having the maximal variance in the direction of $\mathbf{u_1}$. PC_2, $\mathbf{F_2}$, has the maximal variance in a direction perpendicular to $\mathbf{u_1}$.

4.4.2 Mode Estimation with Post-PCA Feature

Since PCA retains the modal property of the raw PMU measurements, the *post-PCA features* $\mathbf{F_1}$ and $\mathbf{F_2}$ will be utilized in the mode estimation. Fourier analysis is employed as a primary tool to estimate the mode of an oscillation, while Prony analysis, as a supplementary of the mode estimation, provides an extra estimate of damping ratio.

4.4.2.1 Fourier Analysis

Suppose that an oscillation is detected at time t_d. Then we construct a measurement matrix $\mathbf{Y_e}(t)$ with the latest $n = 2M$ samples starting from t_d. Then the i^{th} measurement vector $\mathbf{y^{(i)}}(t)$ in $\mathbf{Y_e}(t)$, which contains n samples with sampling interval of $t = k\Delta t,\ k = 0,\ 1,\ \ldots,\ n-1$, can be represented by a finite Fourier series as

$$\mathbf{y^{(i)}}(k\Delta t) = \frac{a_0^{(i)}}{2} + \sum_{j=1}^{M-1}\left(a_j^{(i)}\cos\frac{\pi k}{M}j + b_j^{(i)}\sin\frac{\pi k}{M}j\right) + \frac{a_M}{2}\cos\pi k, \tag{4.3}$$

where

$$a_j^{(i)} = \frac{1}{M}\sum_{k=0}^{n-1} \mathbf{y^{(i)}}\left(k\Delta t\right)\cos\left(\frac{\pi j}{M}k\right),\ j=0,\ 1,\ \ldots,\ M,$$
$$b_j^{(i)} = \frac{1}{M}\sum_{k=0}^{n-1} \mathbf{y^{(i)}}\left(k\Delta t\right)\sin\left(\frac{\pi j}{M}k\right),\ j=1,\ 2,\ \ldots, M-1.$$

We can convert the trigonometric Fourier series in (4.3) into complex Fourier series as

$$\mathbf{y^{(i)}}\left(k\Delta t\right) = \sum_{j=0}^{M} c_j^{(i)}\cos\left(\frac{\pi k}{M}j + \psi_j^{(i)}\right), \tag{4.4}$$

where

$$c_0^{(i)} = \frac{a_0^{(i)}}{2},\ c_M^{(i)} = \frac{a_M^{(i)}}{2},\ \psi_0^{(i)} = \psi_M^{(i)} = 0,$$

and for $j = 1,\ 2,\ \ldots,\ M-1$,

$$c_j^{(i)} = \sqrt{\left(a_j^{(i)}\right)^2 + \left(b_j^{(i)}\right)^2},\ \psi_j^{(i)} = \arctan\left(-\frac{b_j^{(i)}}{a_j^{(i)}}\right).$$

Equivalently, (4.4) can be rewritten as

$$\mathbf{y^{(i)}}\left(k\Delta t\right) = \sum_{j=0}^{M} c_j^{(i)}\cos\left(\frac{\pi j}{M\Delta t}k\Delta t + \psi_j^{(i)}\right), \tag{4.5}$$

or equivalently,

$$\mathbf{y^{(i)}}\left(t\right) = \sum_{j=0}^{M} c_j^{(i)}\cos\left(\omega_j^{(i)}t + \psi_j^{(i)}\right), \tag{4.6}$$

where Δt is the reciprocal of the sampling frequency f_s. The quantities $c_j^{(i)}$, $\omega_j^{(i)}$, and $\psi_j^{(i)}$ are the amplitude, frequency, and phase, respectively, with

$$\omega_j^{(i)} = \frac{2\pi}{n\Delta t} j. \tag{4.7}$$

Equations (4.6) and (4.7), together with FFT, are widely used to efficiently extract the modal information from synchrophasor data in order for oscillation analysis [45].

Next, we provide theoretical justification to show that the modal information is preserved by post-PCA features. This further establishes the feasibility of using the post-PCA features for mode estimation.

Proposition 2 (Preservation of Modal Information by Post-PCA Features)**.** *Given measurement matrix* $\mathbf{Y_e}(t)$ *constructed with oscillatory data, assume a set of modes* $\omega^{(i)}$ *can be identified from every measurement vector* $\mathbf{y^{(i)}}(t)$ *in* $\mathbf{Y_e}(t)$ *by single-channel Fourier analysis as in (4.7). Then these modes can also be identified by Fourier analysis from any of the N PCs as shown in (4.2) up to resolution on n-samples.*

Proof

Assume there is a set of modes $\omega^{(i)} := [\omega_1^{(i)},\ \ldots,\ \omega_M^{(i)}]$ that can be identified from each measurement vector by Fourier analysis as

$$\mathbf{y^{(i)}}(\mathbf{t}) = \sum_{j=1}^{M} c_j^{(i)} \cos\left(\omega_j^{(i)} t + \psi_j^{(i)}\right), \tag{4.8}$$

where $i = 1,\ 2,\ \ldots,\ N$, and $M = n/2$.

Let $\mathbf{F_i}$ be the i^{th} PC of $\mathbf{Y_e}$ after applying PCA, i.e.,

$$\mathbf{F_i} = (\mathbf{Y_e}(t) - \mu)\,\mathbf{u_i}, \tag{4.9}$$

according to (4.2). $\mathbf{u_i}$ satisfies

$$\mathbf{\Sigma u_i} = \lambda_i \mathbf{u_i}, \tag{4.10}$$

where

$$\mathbf{\Sigma} = E\left[\left(\mathbf{Y_e}(t) - \mu\right)^T \left(\mathbf{Y_e}(t) - \mu\right)\right] \in \mathbb{R}^{N \times N} \tag{4.11}$$

is the covariance matrix of $\mathbf{Y_e}(t)$.

From (4.9), the Fourier transform of $\mathbf{F_i}$ can be represented as

$$\mathfrak{F}\left(\mathbf{F_i}\right) = \mathfrak{F}\left(\left(\mathbf{Y_e}\left(t\right) - \mu\right)\mathbf{u_i}\right). \tag{4.12}$$

The expectation in (4.11) eliminates the time dependence of the covariance matrix $\mathbf{\Sigma}$. As a result, its i^{th} eigenvector $\mathbf{u_i}$ is also time independent, i.e., $\mathfrak{F}\left(\mathbf{u_i}\right) = \mathbf{u_i}$. Therefore, (4.12) becomes

$$\mathfrak{F}\left(\mathbf{F_i}\right) = \mathfrak{F}\left(\mathbf{Y_e}\left(t\right) - \mu\right)\mathbf{u_i} = \left[\mathfrak{F}\left(\mathbf{Y_e}\left(t\right)\right) - \mathfrak{F}\left(\mu\right)\right]\mathbf{u_i}. \tag{4.13}$$

The Fourier transform of the constant value of the mean μ is $\mathfrak{F}\left(\mu\right) = \delta\left(\omega\right)$, which has a value of ∞ only at $\omega = 0$. In other words, for any $\omega \neq 0$, $\mathfrak{F}\left(\mu\right) = 0$, yielding (4.13)

$$\mathfrak{F}\left(\mathbf{F_i}\right) = \mathfrak{F}\left(\mathbf{Y_e}\left(t\right)\right)\mathbf{u_i}. \tag{4.14}$$

Assume that a single-channel Fourier analysis is applied to the i^{th} measurement vector to identify the set of modes, as shown in (4.8). Then (4.14) indicates that the same set of modes can also be identified from the i^{th} PC of the measurement matrix. □

4.4.2.2 Prony Analysis

Compared to Fourier analysis, Prony analysis has the advantage of estimating the damping ratio in addition to the mode, phase, and amplitude, and therefore is becoming more attractive in power system oscillation monitoring [51]. Mathematically, Prony analysis fits a time series signal $y(t)$ with a sum of *damped* complex exponentials:

$$\hat{y}(t) = \sum_{i=1}^{L} A_i e^{j\phi_i} e^{(\sigma_i + j2\pi f_i)t}, \tag{4.15}$$

where the results represent the identifications of the amplitude A_i, the phase ϕ_i, the mode f_i, and the damping $\frac{-\sigma_i}{\sqrt{(\sigma_i)^2+(2\pi f_i)^2}}$.

Similar to Fourier analysis, Prony analysis applied to any PC is also capable of identifying the same set of modes that are present in the raw measurements.

4.5 Robust Monitoring Framework of Low-frequency Oscillation

In this section, by integrating the cyclicity response based oscillation detection in Section 4.3, and the PCA-based mode estimation in Section 4.4, we propose a robust monitoring algorithm for low-frequency oscillations in power systems. The implementation flowchart is presented in Figure 4.5. The issue of missing data is not considered in this dissertation. The algorithm is described as follows:

0) **Offline training**.

Calculate the threshold for oscillation detection using historical eventful PMU data.

1) **Perform PCA to extract the pre-PCA features.**

2) **Construct cyclicity response using pre-PCA features**.

3) **Issue oscillation detection command based on comparison with threshold**.

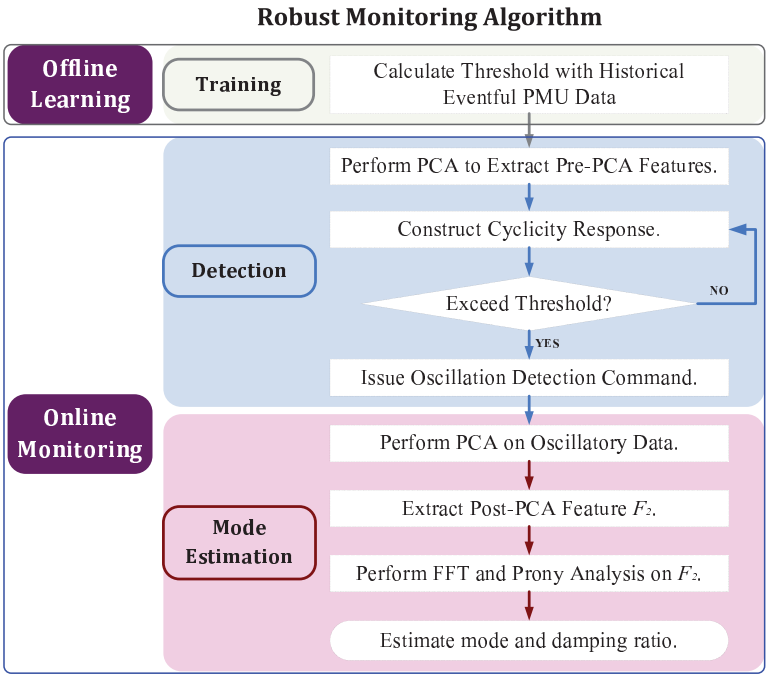


Figure 4.5: Implementation of the robust monitoring algorithm.

If the resulting cyclicity response has a value exceeding the threshold, then issue an oscillation detection command to **Step 4**. Otherwise, return to **Step 2**.

4) **Perform PCA on oscillatory data to extract post-PCA feature**.

Only upon a command from **Step 3**, perform PCA on the oscillatory signals to extract the post-PCA feature vector $\mathbf{F_2}$. The reason for choosing $\mathbf{F_2}$ is discussed in Remark 3.

5) **Mode estimation**.

Perform frequency-domain analysis, FFT and Prony analysis, on $\mathbf{F_2}$ to estimate the mode and damping ratio.

Remark 2. *Steps 1-5 correspond to online monitoring. Steps 1-3 are for oscillation detection, while mode estimation is conducted in Steps 4-5.*

Next, some practical issues for implementing the algorithm are discussed.

4.5.1 Choice of Measurements

Similarly as in Section 3.3.4, this part discusses how to choose appropriate measurements for the proposed algorithm.

Considering the deterioration of mode estimation using the PCA-based approach on multiple categories, again in this dissertation, we apply measurements from a single category for the proposed algorithm. Specifically, voltage magnitude measurements are utilized for pre-PCA feature extraction, and bus frequency measurements are employed to obtain the post-PCA feature. With such choices, the PCA retains the original variance, and therefore the modal information. This will further benefit the mode estimation, as shown in Proposition 2.

Remark 3 (Choice of PC_2 in Mode Estimation)**.** *The global property of bus frequency is reflected by the similar values of PC_1 in PCA. The similarities among PC_1 will not contain much useful modal information. Therefore, PC_2 is employed instead for mode estimation.*

4.5.2 Accuracy of Oscillation Detection

In this part, we discuss the robustness of the threshold for an accurate oscillation detection.

The oscillation detection is based on a threshold calculated from historical eventful PMU data. Currently, with real eventful PMU data from the Western and Texas Interconnection, 30 anomalies are analyzed with a threshold of 2. With such a threshold, the apparent error of detection is 13.33%, including a mis-detection rate of 10%, and a false-alarm rate of 3%.

With the availability of more eventful data, the detection threshold can be refined to achieve higher accuracy of detection.

4.6 Numerical Examples

This section illustrates the efficacy of the PMU-based approaches, including oscillation visualization, detection, and mode estimation. Synthetic PMU data generated from PSS/E and MATLAB [89] are utilized for the demonstration together with real data from the Western U.S. and Texas.

4.6.1 Scatter-plot-based Oscillation Visualization

In this part, the oscillation visualization will be demonstrated using the scatter topological patterns with PCA. Only bus frequency measurements are utilized.

4.6.1.1 Oscillation Visualization using Synthetic PSS/E Data

The 23-bus system in PSS/E [82] is utilized to generate synthetic PMU data with added noise of 92dB SNR. The topology is shown in Figure 3.10. Assumed each bus has one PMU installed, with a sampling rate of 30 Hz. The total length of training data is 250s. For the purpose of visualization, a 2D pre-event space is employed for the scatter plot.

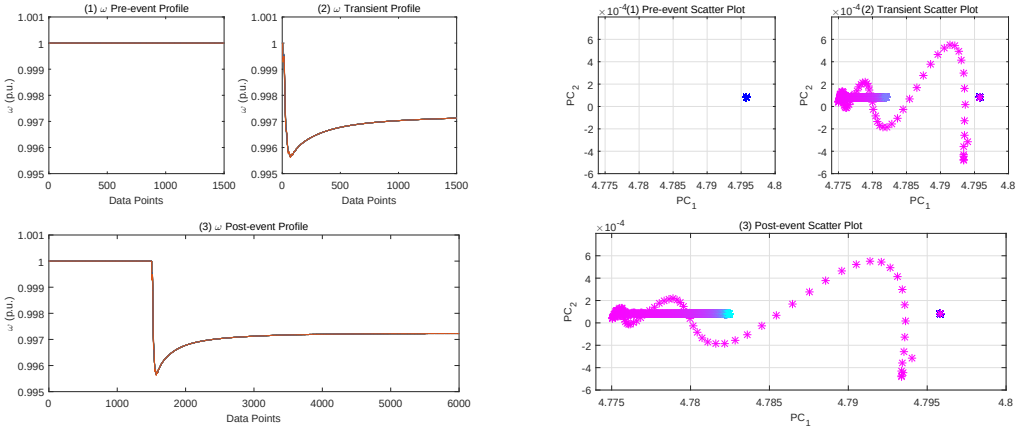
In the rest of this section, we will use non-oscillatory and oscillatory anomalies to demonstrate two different scatter topological patterns. The probability of false alarms is also discussed.

4.6.1.1.1 Visualization of Non-Oscillatory Anomalies

Figure 4.6 illustrates a case where unit 3011 in Figure 3.10 is tripped. Figure 4.6(a) shows the bus frequency profile before, during and after the unit tripping, while Figure 4.6(b) presents the scatter plot accordingly. In the scatter plot, the blue dots in Figure 4.6(b.1) represent the pre-event projections, the color-changing dots in Figure 4.6(b.2) from pink to purple illustrate the transient projections, and the cyan dots in Figure 4.6(b.3) correspond to the post-event projections. Comparing Figures

4.6(b.1) and 4.6(b.2), the deviation of the transient projections from the pre-event ones indicates the occurrence of the anomaly. As can be observed, this non-oscillatory anomaly presents an open pattern with fast convergence in the transient projections. This open pattern can be classified as a simple line in the topology theory.

Another non-oscillatory anomaly is shown in Figure 4.7, corresponding to the control input change of unit 101 in the topology of Figure 3.10. Similarly as Figure 4.6, the scatter plot in Figure 4.7(b) indicates an open-line pattern for this non-oscillatory anomaly.



(a) Bus frequency profile.

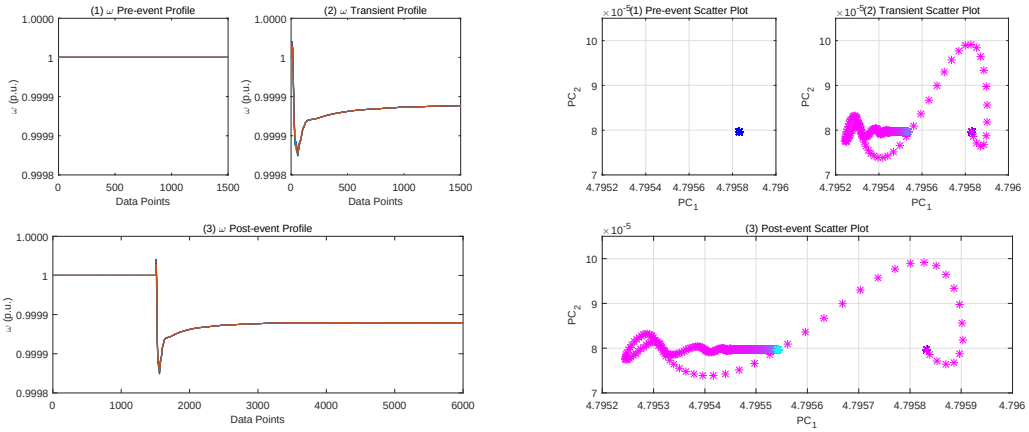
(b) Scatter plot.

Figure 4.6: Scatter plot visualization of unit 3011 tripping.

4.6.1.1.2 Visualization of Oscillatory Anomalies

In this section, we illustrate the visualization of oscillatory anomalies using scatter plot.

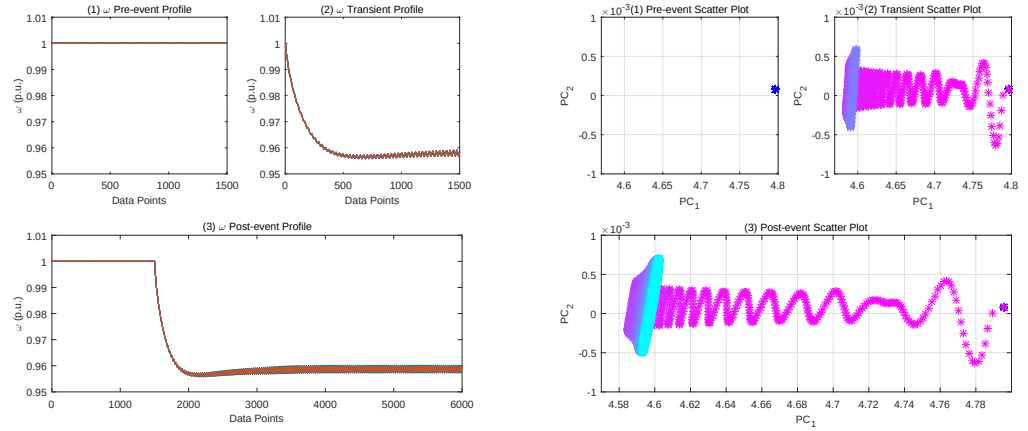
Figure 4.8 corresponds to a tripping of unit 211. As can be observed from Figure 4.8(a), the bus frequency starts the oscillatory behavior following the unit tripping.



(a) Bus frequency profile. (b) Scatter plot.

Figure 4.7: Scatter plot visualization of control input change of unit 101.

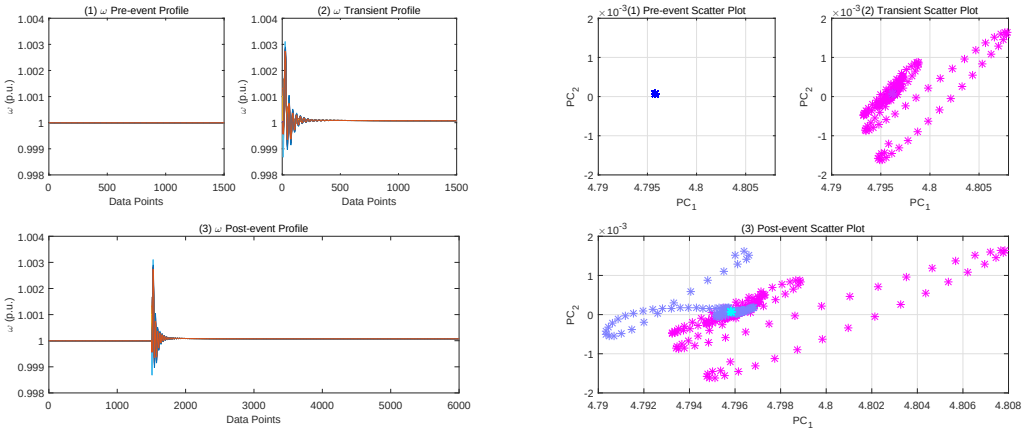
Consequently, in the scatter plot, the transient projections form a circular pattern.



(a) Bus frequency profile. (b) Scatter plot.

Figure 4.8: Scatter plot visualization of unit 211 tripping.

Similarly, the transmission line connecting units 151 and 201 is tripped with the visualization presented in Figure 4.9. Again, a circular pattern can be observed from



(a) Bus frequency profile.

(b) Scatter plot.

Figure 4.9: Scatter plot visualization of line 151-201 tripping.

the scatter plot in Figure 4.9(b).

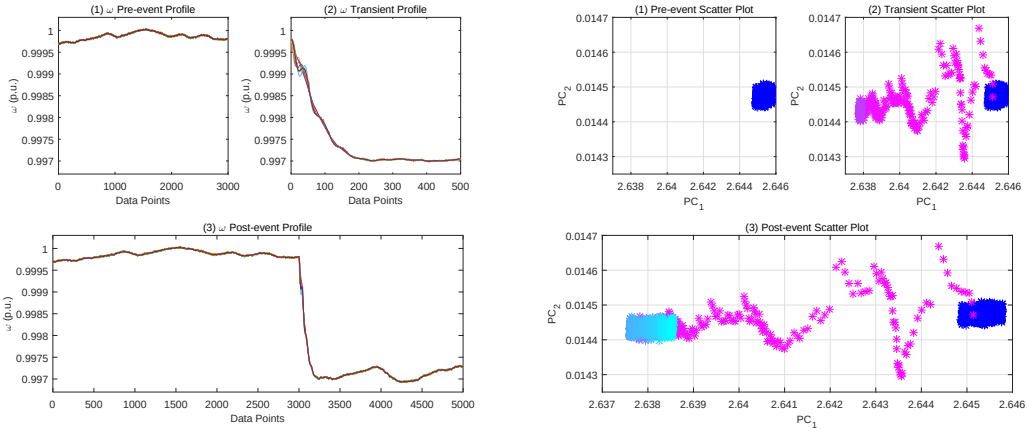
4.6.1.2 Oscillation Visualization using Real Data

In this part, the visualization is demonstrated by use of the real PMU data. Similarly, both non-oscillatory and oscillatory anomalies are employed to illustrate the scatter topological pattern summarized in Section 4.6.1.1.

4.6.1.2.1 Visualization of Non-Oscillatory Anomalies

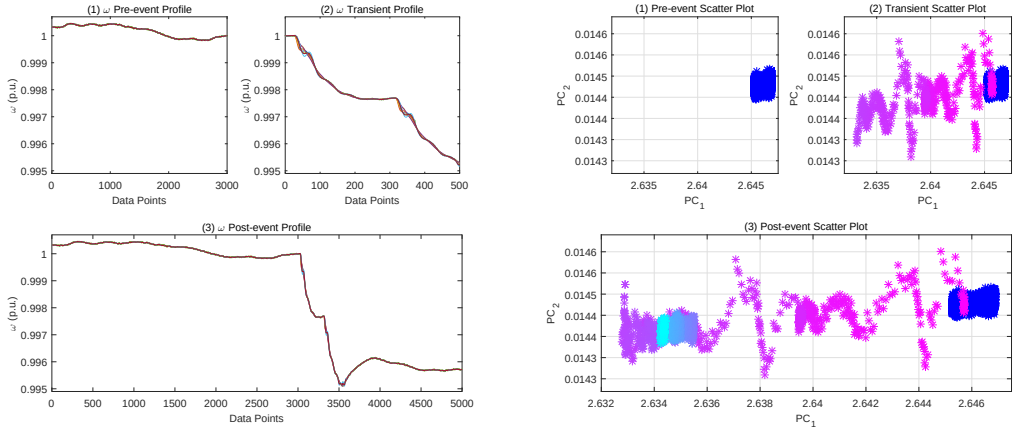
Two unit trippings are recorded by the PMUs from Texas Interconnection. Figures 4.10(a) and 4.11(a) present the bus frequency profiles for these two anomalies before, during and after the unit trippings. Correspondingly, Figures 4.10(b) and 4.11(b) show the scatter plots. As can be observed from both cases, the transition from the pre-event projections in blue to the transient projections in pink indicates the occurrences of such anomalies. The one-directional transient projections suggest the open-line pattern of these two non-oscillatory anomalies.

4.6.1.2.2 Visualization of Oscillatory Anomalies



(a) Bus frequency profile. (b) Scatter plot.

Figure 4.10: Scatter plot visualization of a unit tripping.

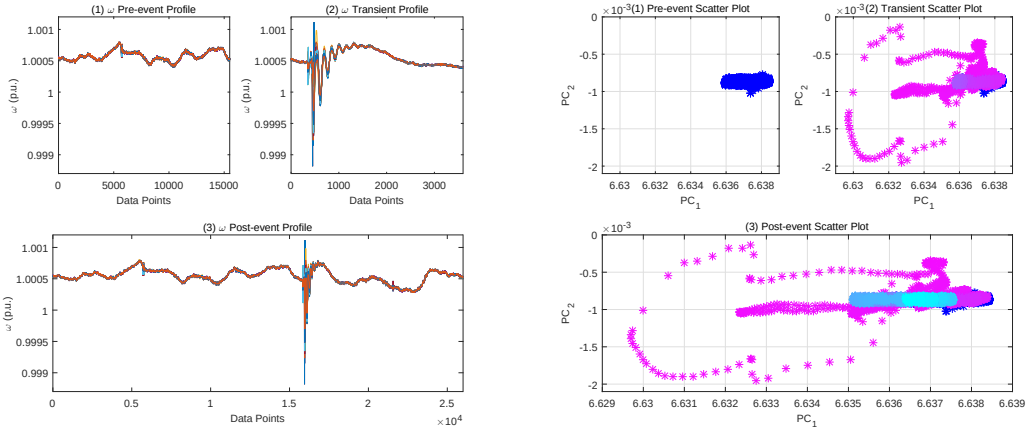


(a) Bus frequency profile. (b) Scatter plot.

Figure 4.11: Scatter plot visualization of another unit tripping.

Figure 4.12 corresponds to a ringdown oscillation recorded from the Western Interconnection, with Figure 4.12(a.2) clearly showing the oscillatory behaviors. As indicated in the scatter plot in Figure 4.12(b), the transient projections form a closed circle pattern.

Similarly, in Figure 4.13, a brake insertion anomaly is presented with PMU data from the Western Interconnection. A similar closed-circle pattern is indicated in Figure 4.13(b).



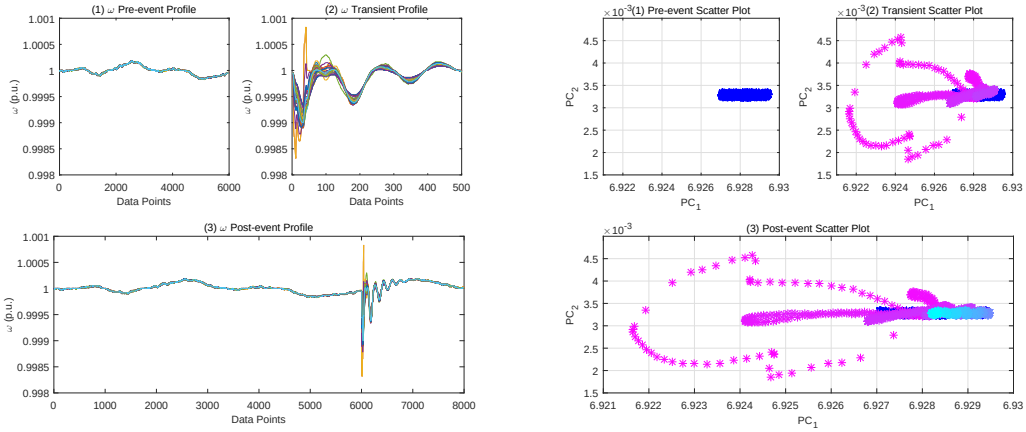
(a) Bus frequency profile. (b) Scatter plot.

Figure 4.12: Scatter plot visualization of a ringdown oscillation.

4.6.2 Cyclicity Response based Oscillation Detection

The visualization of the oscillatory behaviors has already been illustrated in Section 4.6.1.

In this part, we illustrate how the cyclicity response detects the oscillations. Figure 4.14 demonstrates the case for a sustained oscillation in (a.1) and several ringdown oscillations in (b.1). As a comparison, in Figure 4.15, the cyclicity responses



(a) Bus frequency profile. (b) Scatter plot.

Figure 4.13: Scatter plot visualization of a braking anomaly.

for two non-oscillatory anomalies are presented.

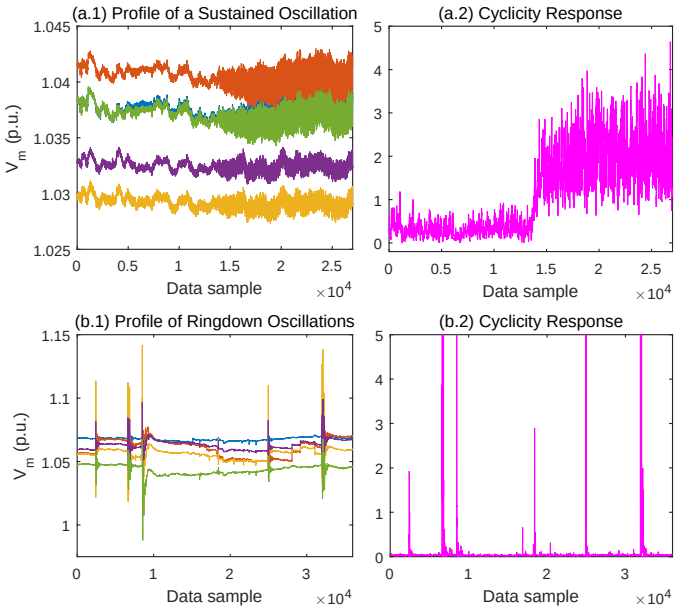


Figure 4.14: Classification of oscillations using cyclicity response.

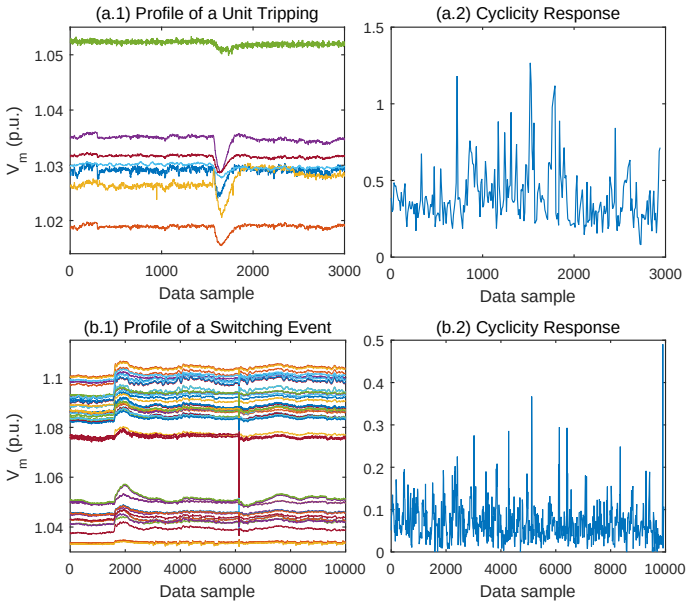


Figure 4.15: Classification of non-oscillatory anomalies using cyclicity response.

The threshold for detection is selected as 2. It can be observed from Figures 4.14 and 4.15 that the occurrence of an oscillation is indicated by the cyclicity response with a value exceeding the threshold.

4.6.3 PCA-based Mode Estimation

This section illustrates the efficacy of the PCA-based mode estimation approach. Both real PMU data [90] and synthetic PMU data generated from MATLAB are utilized.

4.6.3.1 Early Mode Estimation

The sustained oscillation in Figure 4.14(a.1) is utilized to demonstrate the capability of early mode estimation for the robust monitoring algorithm. The bus frequency measurements of this oscillation are shown in Figure 4.16, with a 60 Hz PMU sampling rate. Observing the raw measurements, the anomaly starts around

time $t = 220$ s. The official results published by North American SynchroPhasor Initiative (NASPI) are listed in TABLE 4.1, which indicates that this oscillation can be detected at time $t = 245$ s, with a mode of 1.25 Hz and a damping ratio of 0.

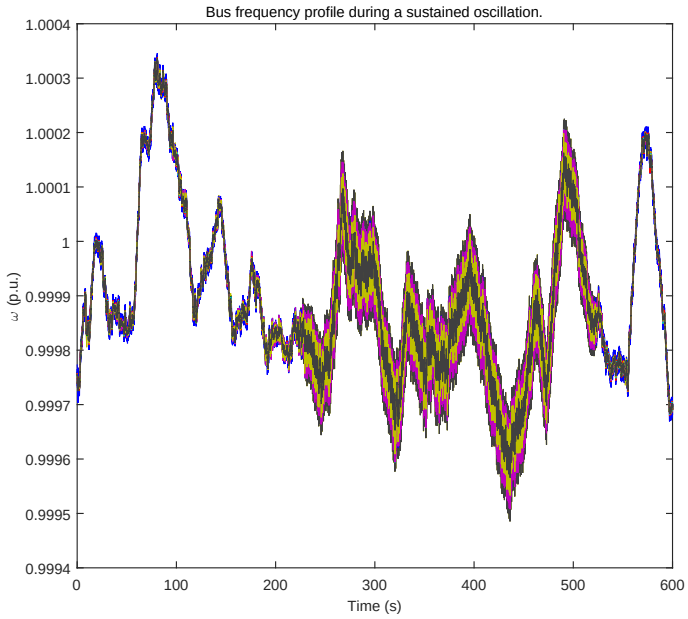


Figure 4.16: Bus frequency measurements during a sustained oscillation.

Table 4.1: Official NASPI results of the sustained oscillation

	NASPI Results
Time When Oscillation Detected (s)	245
Mode (Hz)	1.25
Damping Ratio	0
Signal Used	Frequency, Real Power

TABLE 4.2 presents the mode estimation results using the proposed algorithm, with NASPI "benchmark" results shown in the last column. The first row presents the detection time using Step 2 in Figure 4.5, which is at $t = 228.3833$ s. Then for the FFT-based and Prony-based approaches, we define window sizes of 10 s and 4 s respectively for mode estimation.

Table 4.2: Results comparisons of the sustained oscillation

	FFT	**Prony**	***Benchmark***
Detection Time (s)	228.3833	228.3833	*NA*
Window Size (s)	10	4	*NA*
Estimation Time (s)	238.3833	232.3833	*245*
Time Saving (s)	**6.6167**	**12.6167**	*0*
Estimated Mode (Hz)	**1.23**	**1.2457**	*1.25*
Estimated Damping Ratio	NA	**0.64%**	*0*

The FFT-based approach is capable of estimating a mode of 1.23 Hz at time $t = 238.3833$ s. The total time saved is 6.6167 s in this case compared with the benchmark. The Prony-based approach estimates a mode of 1.2457 Hz, and a damping ratio of 0.64% at time $t = 232.3833$ s, with time saving of 12.6167 s. These time savings indicate that the mode estimation can be achieved by the proposed algorithm at an earlier stage.

The estimated damping ratio from the Prony-based approach is not exactly zero. This is potentially due to the vulnerability of Prony analysis to the high measurement noise. Our future work will explore more robust and accurate approaches to improve mode estimation.

4.6.3.2 Robust Mode Estimation

This part illustrates the robustness of the proposed algorithm on mode estimation.

4.6.3.2.1 Case 1: Real Data of Sustained Oscillation

The sustained oscillation shown in Figure 4.16 is utilized in this case.

A single-channel FFT is first applied on the raw PMU measurement. The results shown in Figure 4.17(a) indicate that the benchmark mode of 1.25 Hz is mis-detected. Potential reasons could be: 1) the high measurement noise, and 2) an incorrect choice of the channel. To handle these problems, we instead adopt the proposed FFT-based approach on the post-PCA feature. As indicated in Figure 4.17(b), a 1.23 Hz mode can be estimated directly from the frequency spectrum. Therefore, it can be concluded that the proposed approach retains more robustness in the presence of high measurement noise than the traditional single-channel FFT, and in the meantime avoids the channel-selection issue.

4.6.3.2.2 Case 2: Real Data of Ringdown Oscillation

Figure 4.18 illustrates the bus frequency profile during a ringdown oscillation. The data from 5 PMUs are provided by NASPI, with benchmark information shown in TABLE 4.3.

Figure 4.19 demonstrates the results of applying single-channel FFT on the 5 PMU measurements, respectively. As can be observed, the first mode estimated from the spectrum is 0.3516 Hz for all the 5 measurements. In other words, the benchmark mode of 0.233 Hz is mis-detected, no matter which measurement channel is selected. This indicates the invalidity of the single-channel FFT in face of high measurement noise.

Instead in Figure 4.20, a mode of 0.2344 Hz can be estimated from the spectrum by applying FFT to the post-PCA feature. This again demonstrates the robustness

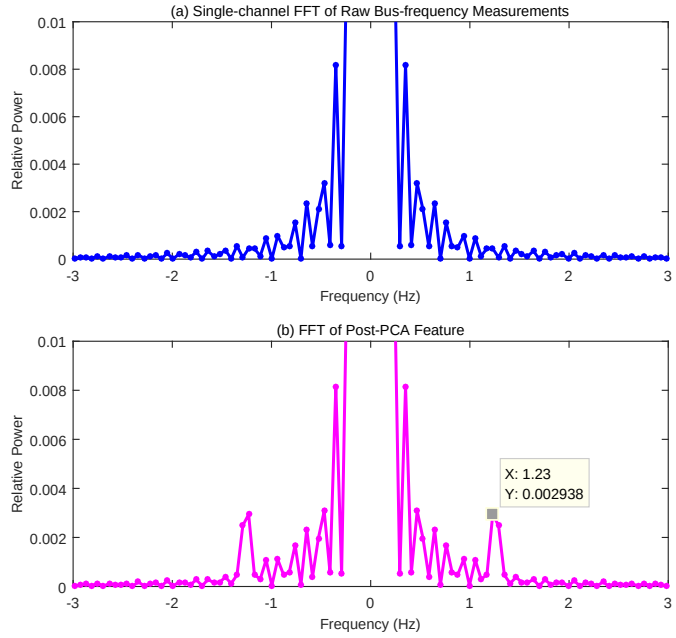


Figure 4.17: FFT results for the sustained oscillation on: (a) raw PMU measurements, and (b) post-PCA feature.

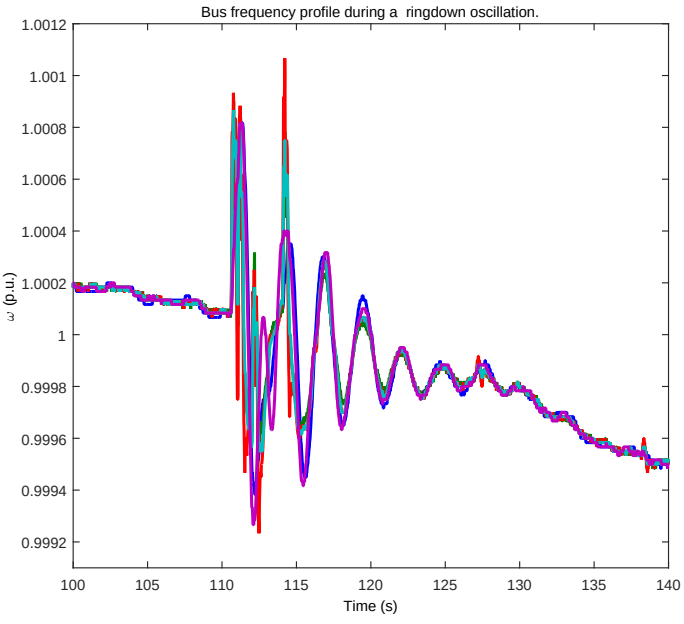


Figure 4.18: Bus frequency measurements during a ringdown oscillation.

Table 4.3: Official NASPI results of the ringdown oscillation

	NASPI Results
Time Start (s)	114.6
Signal Used	Power
Mode 1 Frequency (Hz)	0.233
Mode 1 Damping	11.05%
Mode 2 Frequency (Hz)	0.386
Mode 2 Damping	10.9%

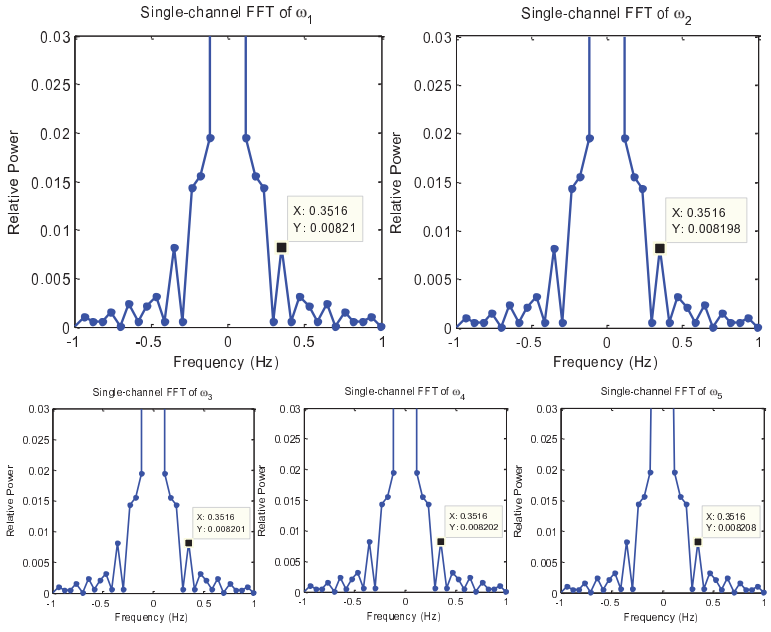


Figure 4.19: FFT results of raw PMU measurements for ringdown oscillation.

of the proposed algorithm against high measurement noises.

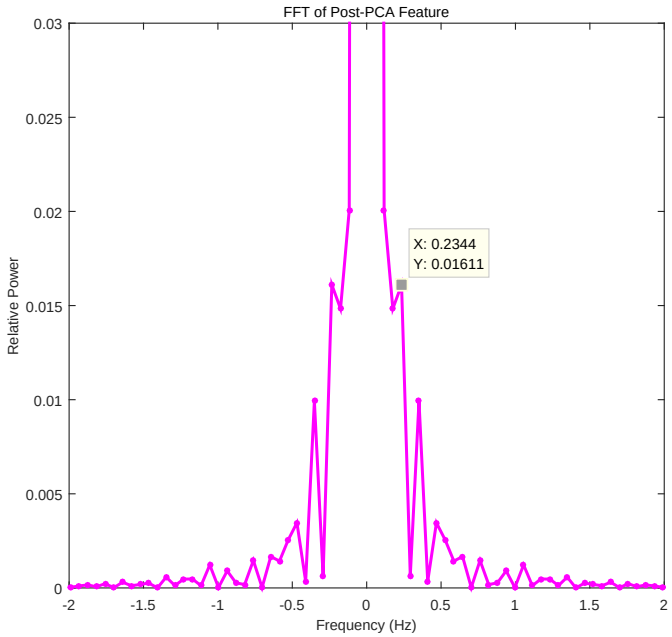


Figure 4.20: FFT of post-PCA feature for ringdown oscillation.

4.6.3.2.3 Case 3: Synthetic Data of Ringdown Oscillation

In this case, synthetic PMU data is generated from the 4-generator system in MATLAB [5], with topology shown in Figure 4.21. Assume each generator is equipped with 1 PMU and the sampling rate is 30 Hz.

Figure 4.22 presents the frequency deviations of the four generators in response to an impulse signal. The four measurements are then analyzed by single-channel FFT, respectively, with the corresponding frequency spectrums shown in Figure 4.23. As can be observed, an oscillatory mode of 1.221 Hz can only be estimated from FFT on $\Delta\omega_1$, while the spectrums of the other three measurements lead to mis-detection

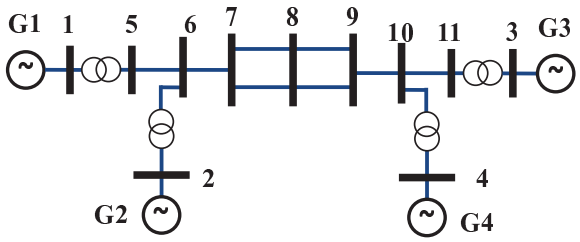


Figure 4.21: Topology of 2-area system [5].

of such a mode. This indicates a problem in that the mis-detection probability using single-channel FFT is 75% even for such a small system with 4 PMUs. Given the large amount of PMUs implemented in real power systems, the risk of mis-detection will be even higher if the choice of measurement channel is incorrectly selected.

By using the proposed algorithm, we can effectively avoid the channel selection problem. The frequency spectrum shown in Figure 4.24 clearly demonstrates the

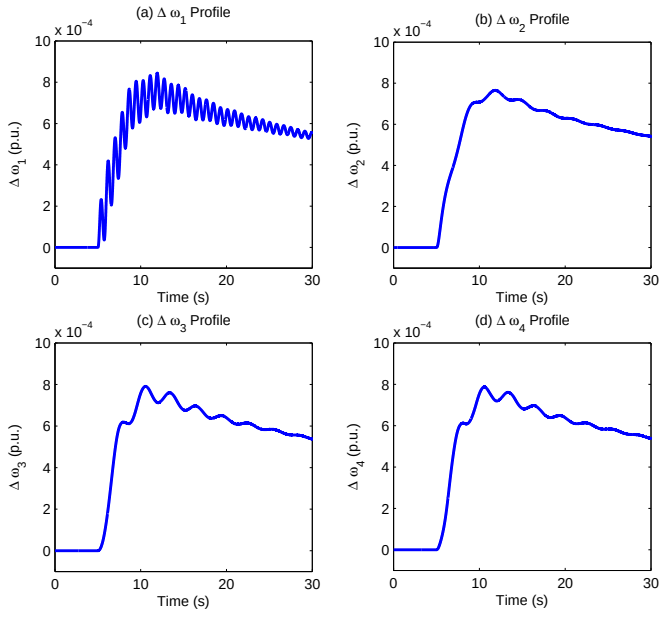


Figure 4.22: Bus frequency deviation for synthetic case.

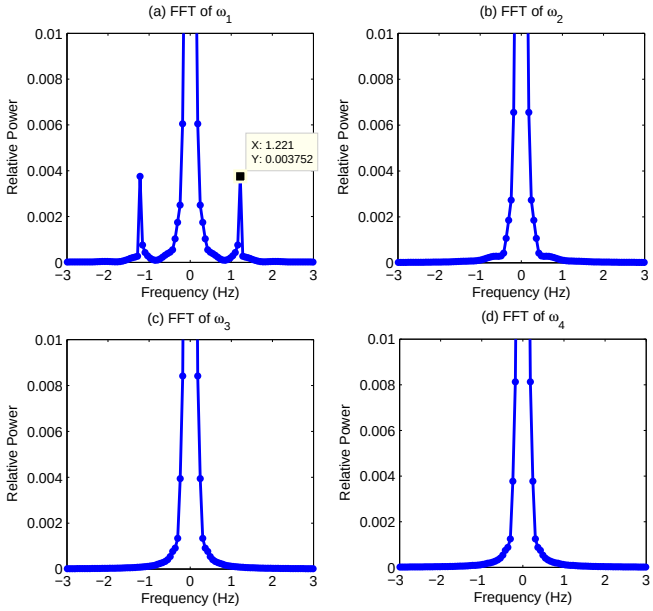


Figure 4.23: FFT of raw PMU measurements for synthetic case.

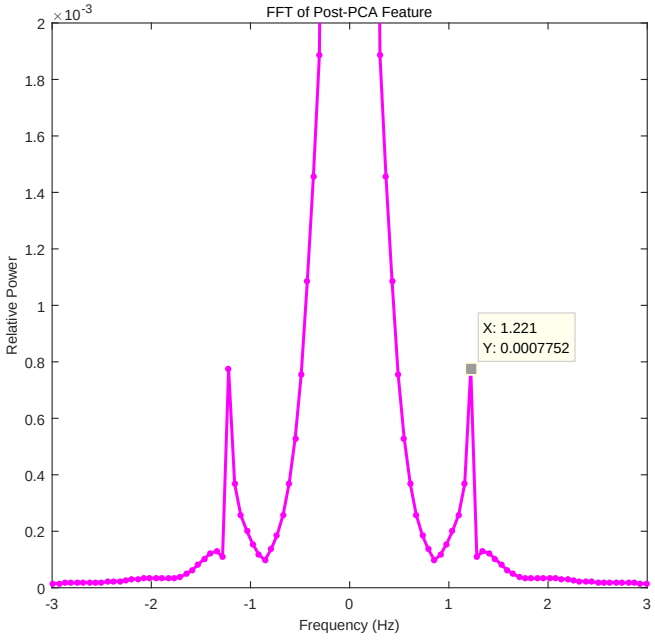


Figure 4.24: FFT of post-PCA data for synthetic case.

detection of the oscillatory mode at 1.221 Hz.

Comparing Figures 4.23 and 4.24, it can be observed that: 1) the modal information is retained in the post-PCA features; 2) the proposed algorithm is capable of effectively estimating the oscillatory mode by avoiding the issue from measurement channel selection. With such a capability, the risk of mis-detection is therefore reduced.

5. CONCLUSIONS

This dissertation introduces a dimensionality-reduction-based framework for early detection and robust classification of power system anomalies by use of synchrophasor data. PCA plays the key role in reducing the dimensionality of the large-scale synchrophasor data.

In Section 2, we introduce background on synchrophasor technology, dimensionality reduction, and power system anomaly detection and classification. In Section 3, we present a PMU-based early anomaly detection algorithm with theoretical justification using linear dynamical system theory. In Section 4, we discuss the PMU-based visualization, followed by robust monitoring of low-frequency oscillations. The conclusions and proposed future work for both of these problems are discussed in the following two subsections.

5.1 PMU-based Early Anomaly Detection

5.1.1 Summary

In Section 3, by exploring the dimensionality reduction of large-scale PMU data, we propose an early anomaly detection algorithm, which lends itself to early detection of general anomalies in power systems, including oscillations, outages, and control input changes, etc.

Dimensionality reduction techniques, including PCA and Isomap, are utilized to examine the underlying dimensionality of large-scale PMU data. With tests using both synthetic and real PMU data, the cumulative variance preserved by the top several PCs yield a high percentage. This indicates an extremely low dimensionality of high-volume PMU data when power system is under either normal operating conditions or contingencies.

Based on the results from dimensionality reduction, a PCA-based dimensionality reduction method with PMU data is implemented with an adaptive training procedure. A basis matrix, consisting only of the pilot PMUs, can be employed to linearly approximate the non-pilot PMUs. The value of the approximation error is utilized to form an anomaly indicator (3.6), which is designed for the robust data-driven early anomaly detection in online monitoring. With the assumption of no consideration of missing data issues, the proposed early anomaly detection algorithm in Figure 3.9 yields high accuracy in early detection of general anomalies. Theoretical justification based on linear dynamical system theory is provided to show the capability of the proposed algorithm in detecting any system anomaly.

Both synthetic PMU data generated from PSS/E and real PMU data collected from the Western, Eastern, and Texas Interconnections suggest the efficacy of the early anomaly detection algorithm in detecting events in an online setting. Such detection is much faster than would be possible by monitoring the raw measurements.

5.1.2 Future Work

Section 3 is only a first step towards understanding and utilizing dimensionality reduction of online PMU data for real-time monitoring. Much more research could be done along this direction. First, with the accumulation of more real eventful data, we plan to continue investigating the efficacy of the proposed algorithm. Considering the false alarms from system high-stressed conditions, another research direction is to improve the robustness of the early anomaly detection algorithm. Given the fact that PCA is capable of reducing the dimensionality and detecting anomalies, other linear dimensionality reduction techniques, such as linear discriminant analysis (LDA) and independent component analysis (ICA), will be the candidates in our future research in improving the robustness. Last but not least, given the fact that some types of

anomalies may have much severer impacts on power system operations than others, our future research avenue will focus on the PMU-based monitoring and analysis of specific anomalies, such as low-frequency oscillation, sub-synchronous resonance, etc.

5.2 PMU-based Robust Monitoring of Low-frequency Oscillations

5.2.1 Summary

In Section 4, a scatter-plot-based visualization approach is first presented to illustrate the characteristics of the topological patterns for the oscillations and non-oscillatory anomalies. From the scatter-plot visualization, the non-oscillatory anomalies behave in an open-line pattern, while oscillations present a closed-loop pattern.

After oscillation visualization, we propose a PMU-based framework for online robust detection and mode estimation of low-frequency oscillations. This framework consists of two stages, where in the first stage, oscillation detection is achieved by use of the persistent-homology-based cyclicity response. By comparing the cyclicity response value calculated from the pre-PCA features in a sliding window with a pre-defined threshold, the low-frequency oscillation can be effectively detected much faster than conventional approaches. After the detection, PCA is applied to extract the post-PCA features from multi-channel oscillatory PMU data. With theoretical justification of Proposition 2, the post-PCA features are shown to preserve the modal information in a more robust way comparing with raw PMU measurements. FFT and Prony analysis are applied on the post-PCA features for mode estimation. Numerical simulations based on both real and synthetic PMU data suggest that the proposed framework is capable in early oscillation detection and robust mode estimation, and at the same time it avoids the channel selection issues of traditional Fourier-based analysis methods.

5.2.2 Future Work

Built upon our preliminary work in Section 4, one of our future research avenues plans to test the oscillation detection approach and train a more robust threshold with more eventful PMU data from the real world. Besides, we plan to continue exploring persistent homology for the application of general anomaly classification in power systems. Considering the oscillation detection errors discussed in Section 4.5.2, another research direction would be to improve the accuracy and robustness of mode estimation by considering advanced approaches in frequency-domain analysis.

Besides the extension of the work in Section 4, we also plan to integrate the work in Sections 3 and 4 in the following aspect:

- **Development of online PMU-based automated anomaly detection, classification, and control scheme.**

With effective detection from Section 3 and robust classification from Section 4, our next research direction would be the development of a data-driven control and mitigation scheme for low-frequency oscillations to close the loop. In order to effectively mitigate oscillations with certain modes, an accurate estimate of the modes will be required. Two aspects will be pursued along this direction:

- More advanced approaches of frequency-domain analysis will be applied to directly estimate the mode and/or damping ratio from real PMU measurements. In the presence of high measurement noise, modified Prony algorithm [91] and matrix pencil algorithm [92] will be good candidate approaches to achieve a more robust estimate of mode and damping ratio.

- System identification techniques [93, 94] will be utilized to estimate the dynamic model of a power system. A preliminary work has been done in [95]

using system identification with available system-wide modal information. We plan to continue this work in developing piecewise linear dynamic models within certain frequency ranges of interest. With prior knowledge of system operating conditions, parameter estimation and tuning can also be achieved from the identified system model.

REFERENCES

[1] B. Pal and B. Chaudhuri, *Robust control in power systems*. Springer Science & Business Media, 2006.

[2] N. Power and C. Council, "1996 system disturbances," August 2002. Available: http://www.nerc.com/pa/rrm/ea/System%20Disturbance%20Reports%20DL/1996SystemDisturbance.pdf.

[3] K. Prasertwong, N. Mithulananthan, and D. Thakur, "Understanding low-frequency oscillation in power systems," *International Journal of Electrical Engineering Education*, vol. 47, no. 3, pp. 248–262, 2010.

[4] M. Klein, G. Rogers, P. Kundur, *et al.*, "A fundamental study of inter-area oscillations in power systems," *Power Systems, IEEE Transactions on*, vol. 6, no. 3, pp. 914–921, 1991.

[5] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*, vol. 4. McGraw-hill New York, 1994.

[6] A. R. Messina, J. Ramirez, C. Canedo, *et al.*, "An investigation on the use of power system stabilizers for damping inter-area oscillations in longitudinal power systems," *Power Systems, IEEE Transactions on*, vol. 13, no. 2, pp. 552–559, 1998.

[7] H. Vu and J. Agee, "Comparison of power system stabilizers for damping local mode oscillations," *Energy Conversion, IEEE Transactions on*, vol. 8, no. 3, pp. 533–538, 1993.

[8] A. Phadke and R. M. de Moraes, "The wide world of wide-area measurement," *Power and Energy Magazine, IEEE*, vol. 6, no. 5, pp. 52–65, 2008.

[9] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 80–93, 2011.

[10] D. J. Trudnowski, J. W. Pierre, N. Zhou, J. F. Hauer, and M. Parashar, "Performance of three mode-meter block-processing algorithms for automated dynamic stability assessment," *Power Systems, IEEE Transactions on*, vol. 23, no. 2, pp. 680–690, 2008.

[11] J. Ma, P. Zhang, H.-j. Fu, B. Bo, and Z.-y. Dong, "Application of phasor measurement unit on locating disturbance source for low-frequency oscillation," *Smart Grid, IEEE Transactions on*, vol. 1, no. 3, pp. 340–346, 2010.

[12] J. Ma, T. Wang, Z. Wang, and J. S. Thorp, "Adaptive damping control of inter-area oscillations based on federated Kalman filter using wide area signals," *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1627–1635, 2013.

[13] Z. Zhong, C. Xu, B. Billian, L. Zhang, S. Tsai, R. Conners, V. Centeno, A. Phadke, and Y. Liu, "Power system frequency monitoring network (FNET) implementation," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1914–1921, 2005.

[14] L. Wang, J. Burgett, J. Zuo, C. Xu, B. Billian, R. Conners, and Y. Liu, "Frequency disturbance recorder design and developments," in *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1–7, IEEE, 2007.

[15] M. Patel, S. Aivaliotis, E. Ellen, *et al.*, "Real-time application of synchrophasors for improving reliability," *North American Elecricity Reliability Corporation, Princeton, Princeton, NJ, Tech. Rep*, 2010.

[16] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early anomaly detection: Linearized analysis," *Power Systems, IEEE Transactions on*, vol. 29, no. 4, pp. 2784–2794, 2014.

[17] Y. Chen, L. Xie, and P. R. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *Power and Energy Society General Meeting, 2013. IEEE*, pp. 1–5, IEEE, 2013.

[18] Y. Chen, L. Xie, and P. R. Kumar, "Power system event classification via dimensionality reduction of synchrophasor data," in *Sensor Array and Multichannel Signal Processing Workshop, 2014. SAM 2014. 8th IEEE*, pp. 57–60, IEEE, 2014.

[19] I. Fodor, "A survey of dimension reduction techniques," *Center for Applied Scientific Computing, Lawrence Livermore National Laboratory*, vol. 9, pp. 1–18, 2002.

[20] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[21] L. Van der Maaten, E. Postma, and H. Van Den Herik, "Dimensionality reduction: A comparative review," *Journal of Machine Learning Research (submited)*, 2009.

[22] K. Anaparthi, B. Chaudhuri, N. Thornhill, and B. Pal, "Coherency identification in power systems through principal component analysis," *Power Systems, IEEE Transactions on*, vol. 20, no. 3, pp. 1658–1660, 2005.

[23] Y. Zhang, Z. Wang, J. Zhang, and J. Ma, "PCA fault feature extraction in complex electric power systems," *Advances in Electrical and Computer Engineering*, vol. 10, no. 3, pp. 102–107, 2010.

[24] Z. Wang, Y. Zhang, and J. Zhang, "Principal components fault location based on WAMS/PMU measure system," in *Power and Energy Society General Meeting, 2011. IEEE*, pp. 1–5, IEEE, 2011.

[25] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *Smart Grid, IEEE Transactions on*, vol. 6, no. 6, pp. 3104–3113, 2015.

[26] S. Pan, T. Morris, and U. Adhikari, "Classification of disturbances and cyber-attacks in power systems using heterogeneous time-synchronized data," *Industrial Informatics, IEEE Transactions on*, vol. 11, no. 3, pp. 650–662, 2015.

[27] J. Barros and E. Pérez, "Automatic detection and analysis of voltage events in power systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 55, no. 5, pp. 1487–1493, 2006.

[28] S.-J. Huang and J.-M. Lin, "Enhancement of anomalous data mining in power system predicting-aided state estimation," *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 610–619, 2004.

[29] S. M. Strachan, S. D. McArthur, B. Stephen, J. R. McDonald, and A. Campbell, "Providing decision support for the condition-based maintenance of circuit breakers through data mining of trip coil current signatures," *Power Delivery, IEEE Transactions on*, vol. 22, no. 1, pp. 178–186, 2007.

[30] S. Tso, J. Lin, H. Ho, C. Mak, K. Yung, and Y. Ho, "Data mining for detection of sensitive buses and influential buses in a power system subjected to disturbances," *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 563–568, 2004.

[31] H. Mori, "State-of-the-art overview on data mining in power systems," in *Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES*, pp. 33–34, IEEE, 2006.

[32] M. Kazerooni, H. Zhu, and T. J. Overbye, "Literature review on the applications of data mining in power systems," in *Power and Energy Conference at Illinois (PECI), 2014*, pp. 1–8, IEEE, 2014.

[33] A. Allen, S. Santoso, and E. Muljadi, "Algorithm for screening phasor measurement unit data for power system events and categories and common characteristics for events seen in phasor measurement unit relative phase-angle differences and frequency signals," *Report*, 2013.

[34] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," *Power Systems, IEEE Transactions on*, no. 99, pp. 1–10, 2013.

[35] J. Jiang, J. Yang, Y. Lin, C. Liu, and J. Ma, "An adaptive PMU based fault detection/location technique for transmission lines. I. theory and algorithms," *Power Delivery, IEEE Transactions on*, vol. 15, no. 2, pp. 486–493, 2000.

[36] J. Jiang, Y. Lin, J. Yang, T. Too, and C. Liu, "An adaptive PMU based fault detection/location technique for transmission lines. II. PMU implementation and performance evaluation," *Power Delivery, IEEE Transactions on*, vol. 15, no. 4, pp. 1136–1146, 2000.

[37] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *Power Systems, IEEE Transactions on*, vol. 23, no. 4, pp. 1644–1652, 2008.

[38] Z. Tashman, H. Khalilinia, and V. Venkatasubramanian, "Multi-dimensional fourier ringdown analysis for power systems using synchrophasors," *Power Systems, IEEE Transactions on*, vol. 29, no. 2, pp. 731–741, 2014.

[39] A. H. Al-Mohammed and M. A. Abido, "A fully adaptive PMU-based fault location algorithm for series-compensated lines," *Power Systems, IEEE Transactions on*, vol. 29, no. 5, pp. 2129–2137, 2014.

[40] N. Dahal, R. King, and V. Madani, "Online dimension reduction of synchrophasor data," in *Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES*, pp. 1–7, IEEE, 2012.

[41] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "Power system real-time event detection and associated data archival reduction based on synchrophasors," *Smart Grid, IEEE Transactions on*, vol. 6, no. 4, pp. 2088–2097, 2015.

[42] J. J. Sanchez-Gasca and J. H. Chow, "Performance comparison of three identification methods for the analysis of electromechanical oscillations," *Power Systems, IEEE Transactions on*, vol. 14, no. 3, pp. 995–1002, 1999.

[43] D. Ke, C. Chung, and Y. Xue, "An eigenstructure-based performance index and its application to control design for damping inter-area oscillations in power systems," *Power Systems, IEEE Transactions on*, vol. 26, no. 4, pp. 2371–2380, 2011.

[44] J. Turunen, J. Thambirajah, M. Larsson, B. C. Pal, N. F. Thornhill, L. C. Haarla, W. W. Hung, A. M. Carter, and T. Rauhala, "Comparison of three electromechanical oscillation damping estimation methods," *Power Systems, IEEE Transactions on*, vol. 26, no. 4, pp. 2398–2407, 2011.

[45] N. Kakimoto, M. Sugumi, T. Makino, and K. Tomiyama, "Monitoring of inter-area oscillation mode by synchronized phasor measurement," *Power Systems, IEEE Transactions on*, vol. 21, no. 1, pp. 260–268, 2006.

[46] S. Nezam Sarmadi and V. Venkatasubramanian, "Electromechanical mode estimation using recursive adaptive stochastic subspace identification," *Power Systems, IEEE Transactions on*, vol. 29, no. 1, pp. 349–358, 2014.

[47] S. Nabavi, J. Zhang, and A. Chakrabortty, "Distributed optimization algorithms for wide-area oscillation monitoring in power systems using interregional PMU-PDC architectures," *Smart Grid, IEEE Transactions on*, vol. 6, no. 5, pp. 2529–2538, 2015.

[48] G. Liu, J. Quintero, and V. Venkatasubramanian, "Oscillation monitoring system based on wide area synchrophasors in power systems," in *Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, 2007 iREP Symposium*, pp. 1–13, IEEE, 2007.

[49] G. Liu, J. Ning, Z. Tashman, V. M. Venkatasubramanian, and P. Trachian, "Oscillation monitoring system using synchrophasors," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, IEEE, 2012.

[50] J. Haver and F. Vakili, "An oscillation detector used in the BPA power system disturbance monitor," *Power Systems, IEEE Transactions on*, vol. 5, no. 1, 1990.

[51] Z. Huang, N. Zhou, F. K. Tuffner, Y. Chen, D. J. Trudnowski, R. Diao, J. C. Fuller, W. A. Mittelstadt, J. F. Hauer, and J. E. Dagle, *MANGO: Modal Analysis for Grid Operation: a Method for Damping Improvement Through Operating Point Adjustment.* Pacific Northwest National Laboratory, 2010.

[52] J. Giri, "Synchrophasor PMU data analysis for enhanced control center operations," 2011.

[53] K. E. Martin, "Synchrophasor standards development-IEEE C37. 118 & IEC 61850," in *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, pp. 1–8, IEEE, 2011.

[54] K. Martin, "Synchrophasor measurements under the IEEE standard C37. 118.1-2011 with amendment C37. 118.1 a," *Power Delivery, IEEE Transactions on*, vol. 30, no. 3, pp. 1514–1522, 2015.

[55] "Phasor advanced concepts," Available: http://www.phasor-rtdms.com/phaserconcepts/phasor_adv_faq.html.

[56] "Low-cost microcontroller-based phasor measurement units improve smart grid reliability," Available: http://www.digikey.co.il/en/articles/techzone/2014/jan/low-cost-microcontroller-based-phasor-measurement-units-improve-smart-grid-reliability.

[57] K. Martin, G. Brunello, M. Adamiak, G. Antonova, M. Begovic, G. Benmouyal, P. Bui, H. Falk, V. Gharpure, A. Goldstein, *et al.*, "An overview of the IEEE standard C37. 118.2-synchrophasor data transfer for power systems," *Smart Grid, IEEE Transactions on*, vol. 5, no. 4, pp. 1980–1984, 2014.

[58] North American SynchroPhasor Initiative, "Synchrophasor update success story," June 2008. Available: https://www.naspi.org/meetingarchives.

[59] J. F. Hauer, D. Trudnowski, and J. G. DeSteese, "A perspective on WAMS analysis tools for tracking of oscillatory dynamics," in *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1–10, IEEE, 2007.

[60] ABB, "Wide area monitoring systems: Portfolio, applications and experiences," April 2012. Available: http://www05.abb.com/global/scot/scot221.nsf/veritydisplay/3d85757b8c7f3bb6c125784d0056a586/$file/1KHL501042%20PSGuard%20WAMS%20Overview%202012-04.pdf.

[61] North American SynchroPhasor Initiative, "Wide area monitoring in the middle of the central European system," September 2007. Available: https://www.naspi.org/meetingarchives.

[62] US Department of Energy, "Model validation using synchrophasors," October 2013. Available: https://smartgrid.gov/sites/default/files/doc/files/NASPI_model_validation_workshop.pdf.

[63] US Department of Energy, "Recovery act smart grid investments – synchrophasor project status," June 2014. Available: https://smartgrid.gov/sites/default/files/synchrophasor_project_status_061014-1.pdf.

[64] US Department of Energy, "Deployment of synchrophasor systems: Decisions and cost impacts," September 2014. Available: www.smartgrid.gov/sites/default/files/doc/files/PMU_cost_study_Final_09232014.pdf.

[65] "Investment in synchrophasors and wide area situational awareness systems for the grid is expected to total nearly $107 billion from 2014 to 2023," Available: http://www.navigantresearch.com/newsroom/investment-in-synchrophasors-and-wide-area-situational-awareness-systems-for-the-grid-is-expected-to-total-nearly-107-billion-from-2014-to-2023.

[66] North American SynchroPhasor Initiative, "Representative data flows from transmission operators to regional hubs, between reliability coordinators, and between transmission operators," October 2014. Available: https://www.naspi.org/documents.

[67] L. Van der Maaten, "An introduction to dimensionality reduction using MATLAB," *Report*, vol. 1201, pp. 07–07, 2007.

[68] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic press, 2013.

[69] M. Partridge and R. Calvo, "Fast dimensionality reduction and simple PCA," *Intelligent data analysis*, vol. 2, no. 3, pp. 292–298, 1997.

[70] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[71] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[72] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[73] W. Rebizant, J. Szafran, and A. Wiszniewski, *Digital signal processing in power system protection and control*. Springer Science & Business Media, 2011.

[74] L. L. Grigsby *et al.*, *The electric power engineering handbook*. CRC Press, 2001.

[75] US Department of Energy, "Smart grid investment grant program-progress report," July 2012. Available: http://energy.gov/sites/prod/files/Smart%20Grid%20Investment%20Grant%20Program%20-%20Progress%20Report%20July%202012.pdf.

[76] "New technology can improve electric power system efficiency and reliability," Available: http://www.eia.gov/todayinenergy/detail.cfm?id=5630.

[77] "Principal component analysis." Available: https://en.wikipedia.org/wiki/Principal_component_analysis#Compute_the_cumulative_energy_content_for_each_eigenvector.

[78] S. Van De Geer, "Estimating a regression function," *The Annals of Statistics*, pp. 907–924, 1990.

[79] M. D. Ilic and J. Zaborszky, *Dynamics and control of large electric power systems*. Wiley, 2000.

[80] Y. Zhang, Y. Chen, and L. Xie, "Multi-scale integration and aggregation of power system modules for dynamic security assessment," in *Power and Energy Society General Meeting, 2013. IEEE*, pp. 1–5, IEEE, 2013.

[81] C.-T. Chen, *Linear system theory and design*. Oxford University Press, Inc., 1998.

[82] Siemens, "PSSE 30.2 program operational manual," 2009.

[83] L. Noakes, "The takens embedding theorem," *International Journal of Bifurcation and Chaos*, vol. 1, no. 04, pp. 867–872, 1991.

[84] "Phase space." Available: https://en.wikipedia.org/wiki/Phase_space.

[85] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2005.

[86] "Attractor." Available: https://en.wikipedia.org/wiki/Attractor.

[87] F. Takens, *Detecting strange attractors in turbulence*. Springer, 1981.

[88] P.-L. George and H. Borouchaki, "Delaunay triangulation and meshing," 1998.

[89] M. U. Guide, "The mathworks," *Inc., Natick, MA*, vol. 5, p. 333, 1998.

[90] Available: https://www.naspi.org/site/Module/Meeting/Forms/General.aspx?m__ID=MEETING&meetingid=347.

[91] M. Osborne and G. K. Smyth, "A modified prony algorithm for fitting functions defined by difference equations," *SIAM journal on scientific and statistical computing*, vol. 12, no. 2, pp. 362–382, 1991.

[92] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 5, pp. 814–824, 1990.

[93] D. J. Trudnowski, *Power system identification toolbox: phase two progress*. PNL-9456, Pacific Northwest Laboratory, Richard, Washington, August 1994.

[94] L. Ljung, "System identification toolbox for use with MATLAB," 2007.

[95] B. P. Wiseman, Y. Chen, L. Xie, and P. R. Kumar, "PMU-based reduced-order modeling of power system dynamics via selective modal analysis," in *Transmission and Distribution Conference and Exposition (T&D), 2016 IEEE PES*, accepted.