Synchrophasor Data-Analytics for a More Resilient Electric Power System

Final Project Report

S-74

Power Systems Engineering Research Center
Empowering Minds to Engineer
the Future Electric Energy System

Synchrophasor Data-Analytics for a More Resilient Electric Power System

Final Project Report

Project Team
Anamitra Pal, Project Leader
Lalitha Sankar
Arizona State University

Christopher DeMarco
University of Wisconsin-Madison

Graduate Students
Reetam Sen Biswas
Zhigang Chu
Andrea Pinceti
Arizona State University

Sowmya Acharya
Jong Min Lim
University of Wisconsin-Madison

PSERC Publication 19-02

September 2019

For information about this project, contact:

Anamitra Pal
Arizona State University
School of Electrical, Computer, and Energy Engineering
P.O. BOX 875706
Tempe, AZ 85287-5706
Phone: 480-965-2882
Fax: 480-727-2052
Email: anamitra.pal@asu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Acknowledgements

This is the final report for the Power Systems Engineering Research Center (PSERC) research project titled "Synchrophasor data-analytics for a more resilient electric power systems" (project S-74). We express our appreciation for the support provided by PSERC's industry members. In particular we wish to thank industry advisors Alan Engelmann (ComED), Blake Buescher (MISO), Zhongyu Wu (MISO), Di Shi (GEIRI North America), Jay Giri (GE Energy), Reynaldo Nuqui (ABB), Evangelos Farantatos (EPRI), Mahendra Patel (EPRI), Dejan Sobajic (NYISO), Curtis Roe (ATC), Emanuel Bernabeu (PJM), Jianzhong Tong (PJM), George Stefopoulos (NYPA), Gordon Matthews (BPA), Qiang Zhang (ISONE), Orlando Ciniglio (Idaho Power Company), Brian Keel (SRP), Atena Darvishi (NYPA), Harvey Scribner (SPP), and Michael Nugent (SPP).

Executive Summary

Deployment of synchrophasor infrastructure is occurring at an exceptionally fast rate in the US power grid; especially at the transmission and the sub-transmission networks. The world's first three-phase phasor measurement unit (PMU)-only linear state estimator has been developed and is running successfully at Dominion Virginia Power. However, the data obtained from PMUs has been primarily used for *forensics* analysis in the past; i.e., *after* an undesirable event has occurred. This PSERC S-74 project is a step towards the utilization of PMU data in *near-real-time* environment. The main focus of this project is to develop algorithms that can distinguish normal system operations from anomalous system behavior using synchrophasor data; and consequently, enhance situational awareness for operational decision making. In this research, following applications of PMU data have been considered:

(a) *Power system monitoring application*: faster islanding detection and robust power system asset health monitoring;
(b) *Power system cyber-protection application*: evaluating the efficacy of PMUs to combat cyber-attacks on the SCADA system and developing data analytics algorithm using synchrophasor data to enhance resiliency against cyber-attacks;
(c) *Power system control application*: Predicting system stability in presence of renewable generation.

To be able to produce the research deliverables with respect to the three above-mentioned applications of synchrophasor technology, the tasks were distributed among the three researchers in the following way: *Task 1*- power system monitoring application (led by Anamitra Pal and his students, and supported by Lalitha Sankar and her students), *Task 2*- power system cyber-protection application (led by Lalitha Sankar and her students, and supported by Anamitra Pal and his students), and *Task 3*- power system control application (led by Christopher DeMarco and his students, and supported by Anamitra Pal and his students).

The first sub-task (Task 1.1) of the monitoring application of this project was power system islanding detection. Synchrophasor measurement based wide-area power system islanding detection has mostly relied on voltage phase angle differences between two buses across the islanded systems. However, *noise due to instrument transformers* can severely degrade the measurement quality and in turn *alter the accuracy of the detection technique*. The errors in the voltage angles could be as high as $\pm 4°$ with respect to existing standards. Such high errors in the PMU data due to the instrumentation channel errors, could result in considerable misclassification in islanding detection. Therefore, a new PMU-based passive islanding detection technique is proposed *which is immune to instrumentation channel errors*. The proposed islanding detection technique monitors the *voltage phase angle difference trajectory* measured from the same PMU as a means to counter the instrumentation channel errors. The voltage phase angle difference is accumulated over a window of PMU samples to minimize misclassification. This approach is termed "cumulated sum of voltage phase angle difference (CUSPAD)".

The second sub-task (Task 1.2) of the monitoring application of this project was power system asset health monitoring. One of the biggest challenges faced by the electric power industry is the successful management of its aging infrastructure. Untimely loss of a power system's critical equipment, e.g., large power transformer (LPT) could be catastrophic to the grid operations. Power system equipment provide information about their health through the sensors that monitor them.

The data captured by these sensors is a treasure-house of knowledge because it contains information about actual as well as potential failures. The sensors for LPTs include online dissolved gas analyzers (DGAs), power quality (PQ) meters, potential discharge (PD) testers, and bushing monitors, amongst others. *However, the output generated by many of these sensors are not monitored continuously*. It is only when they generate an alarm that their outputs are considered for decision-making. Now, it is possible that the sensors generate alarms when the device is very close to an imminent failure, and no possible intervention (at that stage) can prevent the failure and/or subsequent disruption from occurring. PMUs provide time-synchronized measurements of voltage and current phasors at the timescale of 30 to 60 samples per second. The main research question being explored in this sub-task is as follows: *could PMU measurements capture the deteriorating health of an LPT?* The research done in the course of this sub-task has found that the *signal-to-noise ratio* (SNR) of PMU measurements is a robust metric that can quantify transformer health in real-time. SNR is a statistical measure of the strength of the desirable components to that of undesirable components present in a signal. When an equipment is malfunctioning and is close to failure, noise component in the signal tends to increase, resulting in a wider SNR bandwidth. The asset health monitoring scheme proposed in this research utilizes data-driven methods to monitor the SNR bandwidth obtained from PMU measurements for real-time assessment of equipment health.

One of the sub-tasks (Task 2.1) of the power system cyber-protection application was to create a realistic synthetic test system that can be used to verify the performance of different PMU-based applications as well as to test cyber-attacks and countermeasures. A crucial aspect in designing a testbed that can be used to observe system behaviors at PMU sampling speeds is to accurately model the behavior of the system loads. In fact, not considering faults or other such unpredictable events, the dynamics of the system are mostly governed by the variation of loads over time and how the generators respond to such changes. The power system cyber-protection application of this project proposed a new data-driven algorithm for the generation of synthetic bus-level time series load data at 30 samples per second that can be applied to any system model. The proposed data-driven algorithm is unique because it can learn the spatial and temporal correlation from a dataset of real system loads and use the learnt model to generate new synthetic data that retains the same characteristics. A utility in the Western Interconnection (WI), which is also a part of PSERC provided the data that was used to investigate the spatio-temporal correlation in the utility scale PMU data. We have used singular value decomposition (SVD) to screen out the dominant load patterns in the real PMU data and proposed a generalized scheme to create synthetic data in a test system that retains the spatio-temporal attributes of real PMU data.

The second sub-task (Task 2.2) of the power system cyber-protection application was to investigate the vulnerability of PMUs to cyber-attacks. One of the simplest ways in which PMU data can be compromised by a cyber-attacker is via false data injection (FDI). FDI attacks involve an intelligent attacker who replaces a subset of measurements with counterfeits. Prior research had shown that a sub-class of cyber-attacks can bypass the conventional bad data detector, that does not consider the *temporal correlation* in PMU measurements to detect an anomaly. This PSERC project was the first effort to investigate if *predictive-filters* could be used to identify a cyber-attack. Predictive filters study the temporal correlations in PMU measurements from prior data to predict the future measurements. If the predicted measurements do not correspond to the actual measurements, it indicates an anomaly. Two types of cyber-attacks have been investigated in this research: *sudden attack* and *ramping attack*. A *sudden cyber-attack* refers to the situation when an attacker injects false measurements suddenly at a specific time. A *ramping cyber-attack* refers to the scenario, when the attacker injects the false measurements slowly over a period of time. Our research findings have shown that the sudden *cyber-attacks* could be detected by predictive filters. However, it is more challenging to detect an intelligently designed *ramping cyber-attack*.

The power system control application (Task 3) of this project involved analysis of power system voltage stability using synchrophasor data. Major power system outages take place when a range of different phenomena occur in quick succession. However, it has often been found that the loss of voltage stability, and ultimately voltage collapse are the immediate precursors of the outage. In current utility practice, operational measures of vulnerability to voltage instability are based on the state estimator that uses a network model to compute the steady state operating point of the grid, with typical update rates in the order of several minutes. The dependence on accurate knowledge of network parameters and topology, and relatively infrequent update rate may be viewed as shortcomings of the existing practice. Among advances that can support new approaches has been a proliferation of vastly improved measurement technology in the grid. In the bulk transmission system, these improved measurements have predominantly taken the form of synchrophasor measurements via PMUs. Typical reporting rates for such measurements are 30 or 60 samples per second (25 or 50 for 50 Hz-based networks). The much higher reporting rate from PMUs suggests the value of developing efficient PMU-based metrics of system performance, which may be computed in near real-time. The metric employed here was based on SVD, alternately known as Karhunen-Loeve decomposition, principal component analysis (PCA), or proper orthogonal decomposition (POD). In power system engineering, SVD has been employed in the context of "full-model-based" analysis to assess voltage stability by examination of the smallest singular value of the power flow Jacobian. Most of the research on voltage stability had relied on full dynamic models. On the contrary, the work presented here could be viewed as an evolution of a model-free approach for voltage stability assessment; which pre-dominantly relies on PMU data. The proposed work also involved precise identification of "noise dominated" measurement channels that contributes no useful information to the SVD calculation and are therefore considered good candidates for removal from the measurements set.

Project Publications:

[1] R. Sen Biswas, and A. Pal, "A robust techno-economic analysis of PMU-based islanding detection schemes," in *Proc. IEEE Texas Power and Energy Conf. (TPEC)*, College Station, TX, pp. 1-6, 9-10 Feb. 2017 [**Third Best Paper Award**].

[2] M. Barkakati, R. Sen Biswas, and A. Pal, "A PMU based islanding detection scheme immune to additive instrumentation channel errors", *accepted 2019 North American Power Symposium (NAPS)*, Wichita, Kansas, USA.

[3] K. Basu, M. Padhee, S. Roy, A. Pal, A. Sen, M. Rhodes, and B. Keel, "Health monitoring of critical power system equipment using identifying codes," in *Proc. CRITIS 2018 Conf.* [**PI Pal received the 2018 Young CRITIS Award for this research**].

[4] M. Padhee, R. Sen Biswas, A. Pal, K. Basu, and A. Sen, "Identifying unique power system signatures for determining vulnerability of critical power system assets," *submitted to ACM sigmetrics performance evaluation review (PER)*.

[5] A. Pinceti, O. Kosut, and L. Sankar, "Data-driven generation of synthetic load datasets preserving spatio-temporal features," accepted *2019 IEEE Power & Energy Society General Meeting (PESGM),* Atlanta, GA.

[6] Z. Chu, J. Zhang, O. Kosut, and L. Sankar, “Unobservable false data injection attacks against PMUs: feasible conditions and multiplicative attacks,” in *Proc. IEEE SmartGridComm* 2018, Aalborg, Denmark, Oct. 2018.

[7] Z. Chu, A. Pinceti, R. Sen Biswas, O. Kosut, A. Pal, and L. Sankar, “Can predictive filters detect gradually ramping false data injection attacks against PMUs?” accepted IEEE *SmartGridComm* 2019, China.

[8] J. Zhang, Z. Chu, L. Sankar and O. Kosut, "False data injection attacks on phasor measurements that bypass low-rank decomposition," in *Proc. 2017 IEEE Intl. Conf. Smart Grid Comm. (SmartGridComm),* Dresden, pp. 96-101, 2017.

[9] S. Acharya, and C.L. DeMarco, “Exploiting Network-induced Correlation for Efficient Compression of PMU Data,” in *Proc. 2018 North American Power Symposium* (NAPS), Fargo, ND, 2018.

[10] M. Lim, and C.L. DeMarco, “SVD-based voltage stability assessment from phasor measurement unit data,” IEEE Trans. Power Syst., vol. 31, no. 4, pp. 2557-2565, Jul. 2016.

Student Theses:

[1] M. Barkakati, “Transmission system reliability: monitoring and analysis”, M.S. Thesis, Arizona State University, 2018.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Potential benefits

The reliability of the electric power system often depends on the *presence-of-mind* of the operator; a correct decision made by the operator at the time of need can be crucial for the survival of the system. The proposed work is intended to enhance the system's resiliency by providing appropriate tools to operators so that they can make judicious decisions. At the same time, modern technology is often thrust upon operators without taking their apprehensions into considerations. Since misunderstanding of a technology may have seriously negative outcomes, operator-industry acknowledgement is very important during the technology development and transfer process. This PSERC S-74 project is an effort to aid the operators in operational decision making during critical situations. The research pursued in this project demonstrates how to take decisions using real-time phasor measurement unit (PMU) data.

Benefits to RTOs/ISOs: RTOs and ISOs have to integrate a diverse mix of energy resources into the electric grid in a reliable manner to match generation and demand. In order to do this, they also have to analyze a variety of contingency scenarios. The results of this research can help the RTOs/ISOs to perform enhanced power system reliability studies to facilitate integration of renewable generation in a judicious manner that will not incur violations with regards to power system security and resiliency.

Benefits to vendors: The complexity and the cost of updating energy management systems (EMSs) make it essential to explore methods that evaluate and improve power system resiliency without interfering with existing architectures. The proposed work can aid vendors by working in parallel with existing EMS software to provide information regarding equipment health and knowledge of impending failures.

Benefits to power utilities/non-market entities: A PMU-based online asset health monitoring tool realized as an outcome of this project will be very useful for power utilities. Another outcome of this project will be a synchrophasor data-based cyber-attack-resilient detection and control algorithm that can be implemented in real-time. Since the proposed tools will minimize the susceptibility of the electric power system to component failures and cyber-attacks, it will be of significant benefit to power utilities as well as non-market entities.

To summarize, the potential benefits of this PSERC S-74 project are as follows:

- A robust tool for detecting island formation as well as monitoring health of critical power system assets in real-time
- A cyber-attack-resilient detection and control algorithm that overlays existing EMS architectures
- A robust model-free approach for voltage stability assessment using synchrophasor measurements considering system uncertainty

1.2 Key challenges addressed in different tasks

As described in the Executive Summary, this PSERC S-74 project encompasses (i) power system monitoring application, (ii) power system cyber-protection application, and (iii) power system control application. Fig. 1.1 provides a schematic overview of this project. The power system monitoring application is aimed towards best utilization of synchrophasor measurements for robust islanding detection and online asset health monitoring. The power system cyber-protection application is aimed towards improving the power system resiliency against cyber-attacks. Finally, the power system control application makes use of PMU data for real-time prediction of voltage stability in stochastic systems.

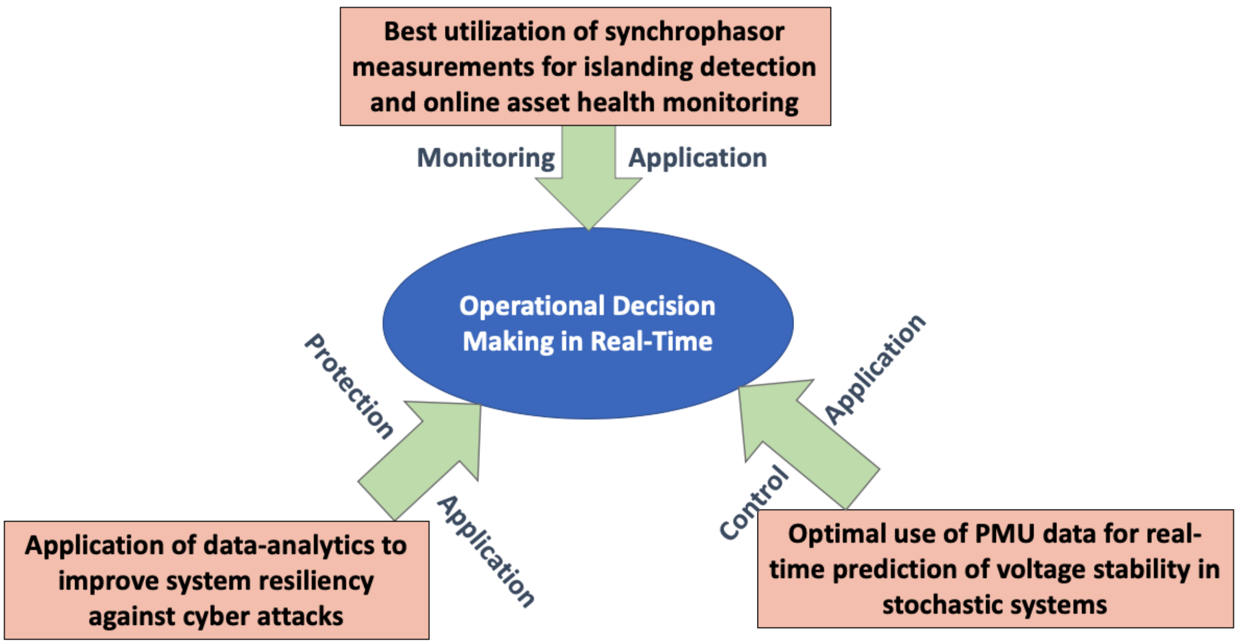


Fig. 1.1: Different components of the PSERC project S-74

The key challenges that this project addressed with regards to each of these three applications are enumerated below:

(i) *Power system monitoring application*:

a. *Synchrophasor-based power system islanding detection*: During un-intentional power system islanding the voltage angle spread across different islands of the power system becomes very large. PMUs installed at different locations of the transmission network could be utilized to observe the relative voltage angle spread across different islands to detect un-intentional separation of the system. However, noise content in the PMU data due to the instrumentation channel errors can severely degrade the measurement quality and in turn affect the accuracy of the islanding detection technique. As per the existing synchrophasor standards [1], due to instrumentation channel errors, the errors in voltage angles can be as high as $\pm 4°$. *The key challenge here is to perform robust islanding detection even in the presence of such high errors.* The proposed research circumvented this problem by proposing a new islanding detection technique that is immune to the instrumentation errors present in the synchrophasor measurements.

b. *Power system asset health monitoring*: Successful maintenance of the aging infrastructure of the electric power industry is a challenging task. Untimely loss of a power system's critical infrastructure; e.g., large power transformer (LPT) could be catastrophic to grid operations. The sensors that monitor an LPT are online dissolved gas analyzers (DGAs), power quality (PQ) meters, potential discharge (PD) testers, bushing monitors, etc. *In practice, the outputs generated by many of these sensors are not monitored continuously.* Only when such sensors generate an alarm, their outputs are considered for decision making. It may happen that when the sensors generate an alarm, it is already too late for the corrective actions to be initiated. The research conducted in this project proposed a new data-driven analysis based on PMU data which can assess the health of the equipment in real-time and generate warnings before the "point of no-return" is reached.

(ii) *Power system cyber-protection application*:

a. *Real-time load monitoring using PMUs*: It has been shown in prior research [2] that when cyber attackers have knowledge of a small sub-network, they can change supervisory control and data acquisition (SCADA) measurements for that sub-network in a way that causes physical damage to the system. For example, a cyber-attacker can create an apparent overload in a line (*that was previously congested*). Such types of cyber-attacks are achieved by manipulating the SCADA data to a credible state, *so as to make it appear to the security constrained economic dispatch (SCED)* that there is a change in *load distribution* in the network. It would lead to a new dispatch and therefore result in an overload of a transmission asset. An enhanced real-time load monitoring system would have the capability to identify anomalous load changes. *However, the key challenge here is that such an option is not feasible in present generation energy management systems (EMSs).* Therefore, this PSERC project exploited the finer granularity of PMU data to learn load patterns at PMU timescales for enhanced real-time load monitoring algorithms.

b. *Vulnerability of PMUs to cyber-attacks*: With the largescale deployment of PMUs at the transmission level, it is also important to investigate the vulnerability of PMUs to cyber-attacks. Prior research [2] has shown that FDIs can be mounted on PMU measurements. However, temporal correlations in the PMU measurements have not been investigated in the past to detect cyber-attacks in PMU measurements. In this project we exploited the use of *predictive filters* to analyze the temporal correlation in synchrophasor measurements to detect an anomaly.

(iii) *Power system control application*:

Voltage stability and ultimately voltage collapse played a major role in multiple power system outages [3]. In current utility practice, operational measures of vulnerability to voltage instability are based on the state estimator that uses a network model to compute the steady state operating point of the grid, with typical update rates in the order of several minutes. Most of the prior work [4], [5] on voltage stability had primarily relied on full dynamic models of the power system. *The dependence on accurate knowledge of network parameters and topology, and relatively infrequent update rate may be viewed as a practical limitation*. To circumvent this limitation, this PSERC project aimed for a measurement-based metric for power system voltage stability assessment. The work presented here is an evolution of a model-free approach for power system voltage stability assessment.

1.3 Report organization

This report is structured as follows.

Chapter 2 presents the research conducted in the synchrophasor based power system monitoring applications. Research findings in the context of power system islanding detection and power system asset health monitoring are documented in Chapter 2.

Chapter 3 documents the research outcomes for the power system cyber-protection application. The algorithm for real-time power system load monitoring and the research findings in the context of vulnerability of PMUs to cyber-attacks are documented in Chapter 3.

Chapter 4 presents the research conducted in the domain of synchrophasor-based power system control applications; the focus was on assessing voltage stability. It describes a measurement-based approach for power system voltage stability assessment.

Chapter 5 summarizes the outcomes of this project and describes the scope of future work.

2. PMU based power system monitoring applications

2.1 Task 1.1: Power system islanding detection

Un-intentional power system islanding refers to an uncontrolled separation of a portion of the electrical network from the rest of the system. It can occur due to power system disturbances (such as faults), natural events (such as hurricanes), or human mis-operation [6]-[8]. Fast and accurate detection of an island when it has formed is essential for the prompt restoration of the system. The role of PMUs in detecting, identifying, maintaining, and eventually restoring the system after the 2008 Hurricane Gustav has been described in [7], [8]. Frequency measurements from PMUs obtained during Hurricane Gustav helped operators monitor the island's load generation balance by adjusting governor controls, which prevented system collapse.

PMUs provide time-synchronized information of complex voltage and current phasors, frequency, and rate-of-change-of-frequency. References [9] and [10] have used frequency differences and voltage phase angle differences for islanding detection, respectively. Principal component analysis (PCA) on voltage magnitudes, phase angles, and frequency measurements were investigated for reliable islanding detection in [10], [11]. Data mining techniques such as support vector machine (SVM) and decision trees (DTs) were applied for islanding detection in [12] and [13], respectively. In [14], a comparative study revealed that DT based classifiers were most dependable for passive islanding detection. Additionally, Tokyo Electric Power Company (TEPCO) Inc. compared different PMU attributes for islanding detection and acknowledged that phase angle difference was the most reliable method for detecting un-intentional islanding [15].

However, PMU measurements are susceptible to both device errors as well as instrumentation channel errors. Fig. 2.1 shows a schematic representation of how the voltage and current transformers are connected with a PMU inside a substation. The voltage and current transformers measure the bus voltage and current phasors, which are then passed through a burden and an attenuator, before the input signals are sent to the PMU device. As per [16], PMU device error expressed as a total vector error (TVE) is typically less than 1%. However, the errors introduced by the instrumentation channel may cause a phase-shift that can be as high as $\pm 4°$ [1]. Thus, reliable and fast detection of un-intentional islanding in the presence of instrumentation channel errors in PMU measurements can be a major challenge [17]. Considering the recent advancements made in renewable energy generation technology, the contribution of inverter-based generation (IBG) such as wind and solar in the transmission network is expected to increase significantly in the near-future. Additionally, during transient phenomena, high renewable energy penetration may have a substantial impact on system stability [18]. Prior research on islanding detection considering renewable energy penetration has primarily focused on the distribution grid [19], [20]. Islanding detection with high penetration of IBG is important because when a renewable rich sub-system gets isolated from the bulk power system, power quality issues such as frequency deviation, voltage fluctuation, and power system harmonics manifest as critical problems in the power system. In addition, if the islanded operation is continued there could be serious concerns about physical injuries because of inspection or people coming in contact with *live* parts [20]. Therefore, it is important to detect un-intentional islanding quickly and initiate immediate corrective actions (such as fast tripping of isolated IBG). The research problem explored in this project in the domain of power system islanding detection is stated as follows: *accurately detect un-intentional islanding in the transmission grid in presence of IBG and additive instrumentation channel errors in the PMU measurements*.

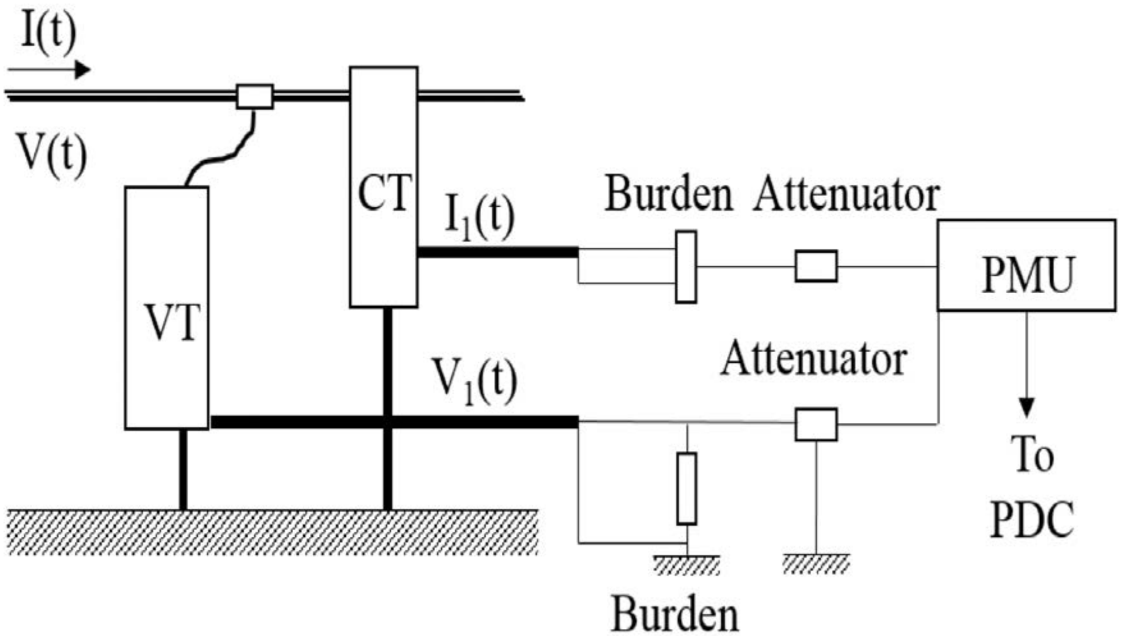


Fig. 2.1: PMU installation depicting locations of instrument transformers

2.1.1 Proposed islanding detection methodology

This section introduces the need for a new PMU-based islanding detection scheme. The reason why the proposed technique is immune to additive instrumentation channel errors is explained next. The methodology followed for modeling wind energy penetration using a positive sequence simulation software is described afterwards. Finally, this section concludes by describing a supervised learning scheme using DTs as well as the methodology that was employed for placing the PMUs.

2.1.1.1 Need for a new islanding detection scheme

Let the true bus voltage angles at any two buses i and j at time instant t be given by θ_i^t and θ_j^t, respectively. The traditional angle difference (AD) approach for islanding detection computes the difference between θ_i^t and θ_j^t [21] as shown below:

$$\boldsymbol{\Delta\theta^t = \theta_i^t - \theta_j^t} \tag{1}$$

When the calculated voltage angle difference, $\Delta\theta^t$, exceeds a pre-determined threshold, τ, the AD approach concludes that an island has formed. It is worth mentioning here that τ is often obtained from offline analyses which do not account for the actual errors present in the system. Consider that the PMU errors associated with the PMUs at buses i and j are e_i^P and e_j^P, respectively, and the instrumentation channel errors associated with the PMUs at buses i and j are e_i^I and e_j^I, respectively. Therefore, the total error at the PMU at bus i is $e_i = e_i^P + e_i^I$ and the total error at the PMU at bus j is $e_j = e_j^P + e_j^I$. Now, the PMU errors (e_i^P and e_j^P) are typically within a total vector error (TVE) of 1% [16], while the instrumentation channel errors (e_i^I and e_j^I) may introduce a phase-shift as high as $\pm 4°$ [1]. Since, the PMU errors are tiny compared to the instrumentation channel errors, the PMU errors can be ignored with respect to the instrumentation errors; i.e., $e_i \approx e_i^I$ and $e_j \approx e_j^I$.

Therefore, the measured voltage phase angle differences at buses i and j are given by $\theta_i^m = \theta_i^t + e_i$ and $\theta_j^m = \theta_j^t + e_j$, respectively. Consequently, the measured voltage phase angle difference between buses i and j, can be written as:

$$\Delta\theta^m = \theta_i^m - \theta_j^m = (\theta_i^t + e_i) - (\theta_j^t + e_j) = \Delta\theta^t + (e_i - e_j) \qquad (2)$$

Due to the error, $e_i - e_j$, in the measured voltage angle $\Delta\theta^m$, the following situations may occur:

a) $\boldsymbol{\Delta\theta^t > \tau}$, but $\boldsymbol{\Delta\theta^m < \tau}$: In this scenario, an un-intentional islanding may not be detected.
b) $\boldsymbol{\Delta\theta^t < \tau}$, but $\boldsymbol{\Delta\theta^m > \tau}$: In this scenario, a non-islanding contingency may be misclassified as an un-intentional islanding.

In light of the two scenarios mentioned above, it is clear that the accuracy of the conventional AD approach would decrease in presence of large instrumentation channel error.

2.1.1.2 Input feature for islanding detection

As described in Section 2.1.1.1, conventional AD approach for islanding detection may not be reliable for detecting un-intentional islanding in presence of large instrumentation channel errors. The major contribution of this research is the development of a pre-processing technique on the input feature set that makes the islanding detection methodology immune to fixed additive instrumentation errors. In our case, the input features are the voltage phase angles obtained from PMUs. We do this by first stating (and proving) the following lemma.

Lemma 1: *Cumulative sum of change in voltage phase angle computed with respect to a pre-contingency reference angle obtained from the same PMU device over a given time-period is immune to instrumentation channel errors.*

Proof: Let θ_x^t and θ_x^m denote the true voltage angle and the measured voltage angle, respectively, i.e. $\theta_x^m = \theta_x^t + e_x$ holds true for every time instant, where e_x denotes the fixed but unknown instrumentation error [22]. Now, let a contingency occurs at $t = t_c$ that causes the voltage angles to change in the manner shown in Fig. 2.2. Note that the pre-contingency voltage angle is the reference voltage angle, denoted by $\theta_x^t(t_c^-)$ for the true angle and $\theta_x^m(t_c^-)$ for the measured angle, respectively, where $\theta_x^m(t_c^-) = \theta_x^t(t_c^-) + e_x$. The cumulative sum of change in voltage phase angles for the true angle and the measured angle are denoted by the green and the blue shaded regions (see Fig. 2.2) and can be mathematically written as:

$$\left.\begin{aligned} S_x^t &= \sum_{n=1}^{w} \left|\theta_x^t(t_c + n) - \theta_x^t(t_c^-)\right| \\ S_x^m &= \sum_{n=1}^{w} \left|\theta_x^m(t_c + n) - \theta_x^m(t_c^-)\right| \end{aligned}\right\} \qquad (3)$$

Now, as the additive instrumentation error is an unknown but fixed quantity, they will cancel out at every time instant of S_x^m making it equal to S_x^t. Therefore, although the true and the measured voltage phase angles are numerically different (the blue and green curves have different intercepts on the Y-axis), the area under the curve between the post-contingency voltage angle and the reference voltage angle (over a window of w samples) for both true and measured voltage angles will be the same. In other words, the following holds true:

$$\left.\begin{array}{l}\theta_x^m \neq \theta_x^t \\ S_x^m = S_x^t\end{array}\right\} \tag{4}$$

From (4), it can be concluded that the cumulated sum of change in voltage phase angle obtained from a specific PMU over a time trajectory does not get affected by instrumentation channel errors. This proves Lemma 1.

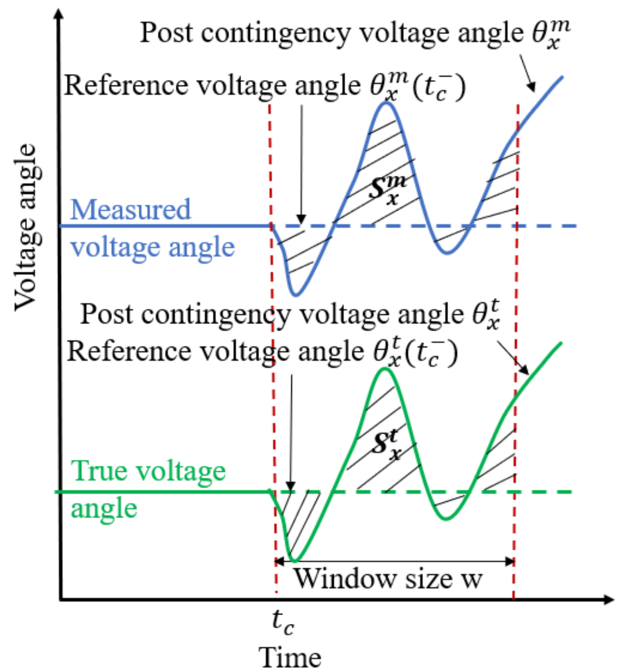


Fig. 2.2: Schematic diagram depicting immunity of CUSPAD to additive instrumentation errors

The traditional approach for detecting islands used the raw angle differences between two buses, say, x and y, given by $\theta_x^m - \theta_y^m$, as input feature for decision-making. Considering Lemma 1, in this research, the following methodology is devised for selecting the input feature. If a contingency occurs at $t = t_c$, the cumulative sum of change in voltage phase angles for buses x and y, denoted by S_x^m and S_y^m is computed based on the relation shown in (3). Since S_x^m and S_y^m are immune to instrumentation channel errors, the input feature for islanding detection is chosen to be the cumulative sum of change in voltage phase angle difference (CUSPAD) between buses x and y, which is mathematically described by:

$$CUSPAD_{xy} = S_x^m - S_y^m \tag{5}$$

For real-time applications, the *determination of the pre-contingency reference voltage angle* $\theta_x(t_c^-)$ *and* $\theta_y(t_c^-)$ *in real-time* is a major concern. This problem can be resolved by using the three-sample based quadratic prediction algorithm (TSQPA) proposed by Gao et al. in [23], and extended to multiple load models in [24]. TSQPA states that for a linear change in load (which is a valid assumption to make considering the fast output rates of PMUs), the relationship between successive voltages is given by:

$$V(n|n-1) = 3V(n-1) - 3V(n-2) + V(n-3) \qquad (6)$$

where, $V(n|n-1)$ denotes the predicted value of complex voltage at time instant n, when the voltages at time instants $n-3$ through $n-1$ are known. From the predicted value of the complex voltage, $V(n|n-1)$, the predicted voltage phase angle, $\theta(n|n-1)$, can be obtained. Knowing the predicted phase angle, $\theta(n|n-1)$, and the measured phase angle, $\theta(n)$, an observation residual, $r(n)$, can be computed as follows:

$$r(n) = \theta(n|n-1) - \theta(n) \qquad (7)$$

When the observation residual, $r(n)$, manifests a sudden change, it means a contingency has occurred at time instant n and the reference voltage angle for CUSPAD calculation must be the angle just before that time instant, i.e., $\theta(n-1)$. Based on the analysis done above, the main conclusion is described by the following theorem.

Theorem 1: *For islanding detection in presence of additive instrumentation channel errors, a CUSPAD-based approach has higher accuracy than the conventional angle difference (AD)-based approach.*

Proof: Section 2.1.1.1 demonstrates how islanding detection accuracy of the conventional AD approach would deteriorate in presence of additive instrumentation channel errors. Lemma 1 proves how CUSPAD computed with respect to a pre-contingency reference angle becomes immune to additive instrumentation channel errors. By combining the two arguments it can be concluded that CUSPAD will provide better performance in comparison to the conventional AD approach for islanding detection in presence of large instrumentation channel errors. This proves Theorem 1.

2.1.1.3 Wind energy modeling

A wind farm is a collective group of interconnected wind turbines that are tied to a point of common coupling (PCC) before the power is fed to the grid. In accordance with the WECC Wind Plant Power Flow Modeling Guide, wind power plants must be represented by an equivalent generator, generator transformer, collector system, and substation transformer [25]. The characteristic features of the wind farm used in this study are described below.

A wind farm containing several wind turbines is modeled as an equivalent generator as depicted in Fig. 2.3. An individual wind turbine is typically rated for capacities 1-4 MW at around 690 V. A pad mounted generator step-up transformer usually steps up the generation voltage of 600-690 V to 34.5 kV by the transformer between buses 4 and 5. Multiple wind turbine models are connected at the 34.5 kV collector bus between buses 3 and 4. The operating voltage at the collector bus is further stepped up at the interconnection to the transmission voltage level at 132 kV or 230 kV via a substation transformer between buses 2 and 3. The representation in Fig. 2.3 is considered adequate for positive sequence dynamic simulations [25]. Type 4 wind energy generator (Wt4g), turbine (Wt4t) and exciter models (Wt4e) are used to represent the wind energy penetration. The power system simulator used to carry out dynamic simulations is GE PSLF.

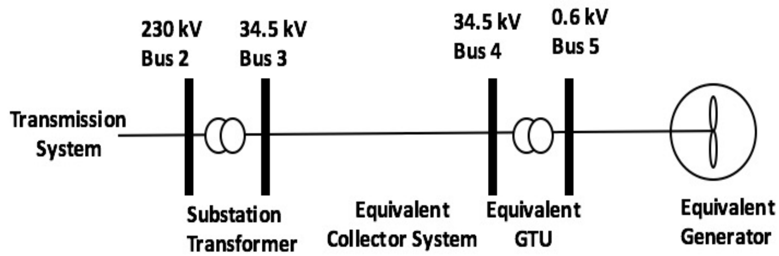


Fig. 2.3: Single line diagram of wind turbine [25]

2.1.1.4 Supervised learning for islanding detection

Supervised learning techniques such as DTs and Random Forest (RFs) have often been used for islanding detection. DT is a supervised learning-based data mining technique which infers hidden relationships from the data and classifies it based on binary partitioning through if-else statements [26]. RF fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve predictive accuracy and control overfitting [27]. In this research, a Classification and Regression Tree (CART)-based DT and RF is trained offline with the help of a training database and a mapping is developed by finding correlations between the input and the output. In [14], [26]-[29] it is observed that DT based classifiers detect island formation accurately and reliably. As such, DTs and RFs were used to evaluate the performance of the proposed methodology.

To create a robust dataset for accurate islanding detection, islanding and non-islanding scenarios were created and simulated in accordance with the following methodology.

1. *Generation of simulation cases*: For non-islanding scenarios, some extreme cases such as line trips, faults, and generator trips were simulated, and the measurement of voltage phase angle for these cases recorded from GE PSLF. For creating island in large test systems $i, 1 \leq i \leq 5$ transmission lines were removed at different instants of time using the community-based partitioning scheme developed in [30].

2. *Measurement of voltage phase angle*: For each case, the voltage phase angle measurements required for calculating CUSPAD values are obtained using the model *ametr* in GE PSLF. It is assumed that PMUs are installed on multiple locations in the system under study and the bus voltage angle measurements are provided by them; see Section 2.1.1.5 for the PMU placement methodology that was employed in this research.

3. *Calculation of CUSPAD*: Dynamic simulations were run in GE PSLF to record the phase angle measurements at the rate of 30 samples per second to emulate PMU data reporting rate. CUSPAD is computed based on the methodology described in Section 2.1.1.2.

4. *Training Data*: After the CUSPAD values for every simulation is obtained, they are fed as inputs to CART and RF. Every case in the training dataset is identified as an islanding case or a non-islanding case by labeling it as 0 or 1 [26]. This serves as the training database for the DT and RF.

5. *Testing Data*: To test the supervised learning model built in the previous step, realistic measurements are replicated through introduction of measurement errors in the training database. The error model used is additive and includes both PMU and instrumentation channel errors:

   i. PMU errors in voltage phase angles are assumed to be a Gaussian distribution with zero mean and standard deviation of 0.104° [1].

   ii. Instrumentation channel errors in voltage phase angle are assumed to follow a uniform distribution that lies in the range of ±1°, ±2°, or ± 4° for the different case studies considered. Good quality measurements (for example, revenue quality instrument transformers) are also considered for testing purpose. They are assumed to introduce an angle error of the order of 0.1° [1].

The resultant voltage phase angles after incorporation of additive PMU and instrumentation channel errors is given by [31]:

$$\theta_V^m = \theta_V^t + \alpha_{VT}^{error} + \alpha_{PMU}^{error} \tag{8}$$

where θ_V^m is the measured voltage phase angle and θ_V^t is the true voltage phase angle. The instrumentation channel errors are denoted by α_{VT}^{error} while the PMU errors are denoted by α_{PMU}^{error}. A schematic diagram describing the different steps that were followed for training and testing the DT-based islanding detection classifier is shown in Fig. 2.4.

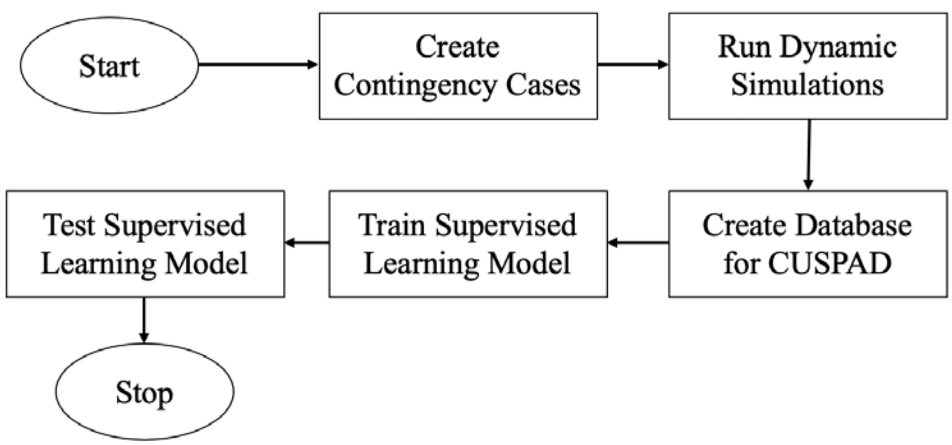


Fig. 2.4: Flowchart for the proposed CUSPAD approach

2.1.1.5 PMU placement

When PMUs are placed in a network, the primary objective is to ensure observability, i.e. the PMUs should have the ability to directly or indirectly observe all the bus voltages of the network. In addition to ensuring topological observability, the PMU placement scheme proposed in [32] takes into consideration PMU redundancy for critical buses as well as the cost of disrupting a substation for PMU installation. Accordingly, the core concept of [32] was employed here for determining the locations where PMUs must be placed.

Let the power network be denoted by an undirected graph $G(V, E)$ such that V is the set of nodes (buses) and E is the set of edges (transmission lines or transformers). The buses are grouped into substations, S, using the rationale that buses connected by transformers will lie inside the same substation. It is assumed in this study that all PMUs are of the dual-use line relay (DULR)-type. For each substation $S_i \in S$, a binary variable x_i is used such that the following holds true:

$$x_i = \begin{cases} 1, & \text{if Substation } \mathrm{S_i} \text{ is distrupted} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Each edge $e \in E$ is associated with two binary variables w_e^l and w_e^h such that following holds true:

$$w_e^l = \begin{cases} 1, & \text{if DULR is placed at the low end of edge e} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

$$w_e^h = \begin{cases} 1, & \text{if DULR is placed at the high end of edge e} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

The objective is to minimize the total cost of PMU installations which involve cost of PMU devices as well as the cost of disrupting a substation. This objective function is mathematically described by:

$$Minimize \left(\sum_{i=1}^{k} c_i x_i + \Delta \sum_{e \in E} \{w_e^h + w_e^l\} \right) \tag{12}$$

where, c_i is the cost of disrupting a substation, Δ is the cost of a DULR, and $k = |S|$. If E_v denotes all outgoing phases from a vertex v, the constraint for phase observability is given by:

$$\sum_{e \in E_v} \{w_e^h + w_e^l\} \geq 1 \tag{13}$$

2.1.2 Simulation results

In this section, the efficiency of CUSPAD in islanding detection is compared with that of the conventional AD approach. For the AD approach, pairs of voltage phase angle differences are calculated through instantaneous combinations of PMU measurements. The test systems comprised of a modified version of the 18-bus system available in the GE PSLF library and the IEEE 118-bus system. Measurement errors consisting of both PMU and instrumentation channel errors were included in the test data. The error model used for the two error types can be found in Section 2.1.1.4. The simulations were repeated 50 times and accuracy with a 95% confidence interval was computed for the test data.

2.1.2.1 Modified 18-bus test case

The original 18-bus system is modified to include wind energy penetration in the network by replacing one of the conventional generators with an equivalent capacity wind farm connected at the 230-kV voltage level. The number of PMUs required for complete observability of the 18-bus system was 5 and they were located at buses 1, 11, 14, 23, and 31. Total cases simulated were 467, out of which 200 were islanding cases and 267 were non-islanding cases. For the 18-bus test case, we have used DT based supervised learning scheme for islanding detection. To determine a suitable window length for calculating CUSPAD, the DT accuracies obtained with various window sizes are presented in Fig. 2.5. We observe that the relative increase in DT accuracy corresponding to the window size between 30 and 40 samples is lesser as compared to that obtained for window sizes between 20 and 30 samples. A larger window size would however negatively influence the detection time (by adding more delay). As with any islanding detection algorithm, a lower detection time is preferred and therefore a compromise between DT accuracy and window size must be made. In the literature, a time delay of 100-150 ms was considered in [17] to prevent misclassifications. Islanding detection time as high as 2-3 seconds is discussed in [9]. Taking all this into account, we believe that for the proposed study, a window size of 30 samples would be appropriate. Comparing accuracies in Table 2.1, it can be concluded that for the 18-bus system, for a window-size of 30 samples, the CUSPAD approach was not affected by increasing amounts of additive measurement errors while the performance of the conventional AD approach deteriorated considerably as the errors increase in the measurements.

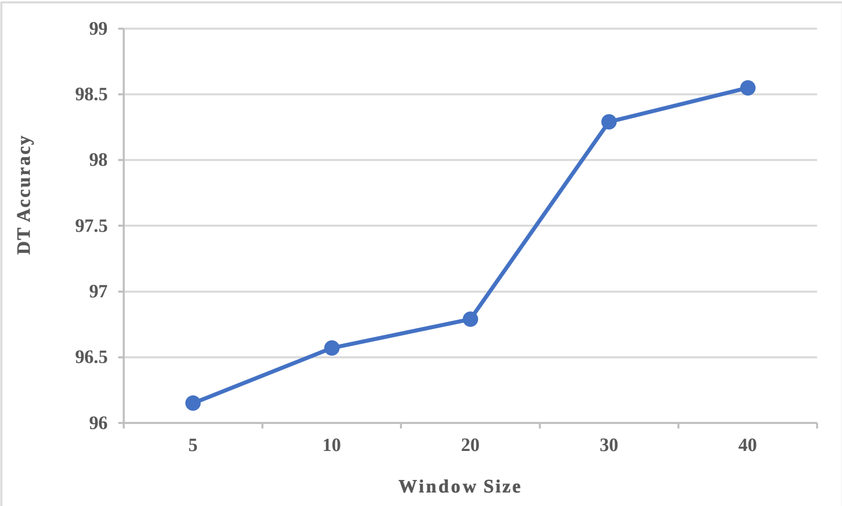


Fig. 2.5: Selection of window size for CUSPAD calculation

Table 2.1: Accuracy comparison of DT models for modified 18-bus system (16% wind penetration)

Error		**AD**		**CUSPAD**	
Instrumentation Channel Error	**PMU Error**	**Accuracy (95%)**	**Depth**	**Accuracy (95%)**	**Depth**
0°	0 Mean ±0.104° SD	99.79	5	98.29	5
-0.1° $\leq \alpha_{VT}^{error}$ ≤0.1°		99.37	5	98.01	5
-1° $\leq \alpha_{VT}^{error}$ ≤1°		96.85	5	98.06	5
-2° $\leq \alpha_{VT}^{error}$ ≤2°		94.66	5	98.08	5
-4° $\leq \alpha_{VT}^{error}$ ≤4°		93.04	5	97.97	5

2.1.2.2 IEEE 118-bus test case

The original IEEE 118-bus system was modified to include variable percentages of wind energy penetration in the network by replacing some of the conventional generators with equivalent capacity wind farms connected at the 132kV voltage level. The number of PMUs required for complete observability of this system was 38. They were placed on buses 3, 5, 8, 9, 12, 15, 17, 21, 23, 28, 30, 36, 40, 43, 45, 49, 52, 56, 59, 63, 65, 66, 68, 69, 71, 75, 77, 80, 85, 86, 84, 91, 94, 101, 105, 110, 114, and 116. To create islands in the IEEE 118-bus system, the community-based partitioning logic developed in [30] was used. It identifies the minimum number of edges that must be lost for islands of a given size to form. In total, 2,000 cases were simulated for three levels of wind energy penetration, namely, 10%, 20%, and 30%. Of these 2,000 cases, 1,000 were islanding cases and 1,000 were non-islanding cases. The 30-sample window size was again selected for computing the CUSPAD accuracy for the 118-bus system. For the IEEE 118-bus test system two supervised learning techniques were tried for islanding detection, namely, CART-based DT and RF. Table 2.2 compares the accuracy for AD and CUSPAD for the IEEE 118-bus test system using DTs. The most interesting observation here is that with increased instrumentation channel errors the accuracy for islanding detection with simple angle-difference as an attribute monotonically decreases. However, the proposed CUSPAD accuracy remains approximately constant at 85%, thereby proving that it is immune to the percentage of instrumentation channel errors added to the measurements. For a fair comparison all the DTs were pruned till a depth level of 6.

Table 2.2: Accuracy comparison with Decision Trees (DTs) for 118-bus system

Error		AD		CUSPAD	
Instrumentation Channel Error	**PMU Error**	**AD Accuracy (95%)**	**Depth**	**CUSPAD Accuracy (95%)**	**Depth**
0	0 Mean ±0.104 SD	85.50	6	85.50	6
-1		80.50	6	83.50	6
-2		73.00	6	84.00	6
-4		55.50	6	85.00	6

Table 2.3: Accuracy comparison with Random Forest (RF) for 118-bus system

Error		AD		CUSPAD	
Instrumentation Channel Error	**PMU Error**	**AD Accuracy (95%)**	**Depth**	**CUSPAD Accuracy (95%)**	**Depth**
0	0 Mean ±0.104 SD	93.02	6	91.44	6
-1		85.05	6	91.03	6
-2		82.36	6	91.98	6
-4		78.99	6	92.04	6

Table 2.3 compares the islanding detection accuracy for AD and CUSPAD using RF as the supervised learning technique. From the RF results, we observe that the CUSPAD accuracies are immune to the instrumentation channel errors, while the AD accuracy with DTs monotonically decreases. Another observation from Table 2.2 and Table 2.3 is that the accuracy of islanding

detection using RF is relatively more compared to DTs. This is expected, because of the following: DTs use a single tree to train the data, while RFs operate by training a multitude of trees during training time and outputs a mean prediction by taking attributes from multiple trees during the testing [26], [27]. Therefore, RFs are observed to perform better against DTs on the same dataset.

2.1.2.3 Summary of the findings

A PMU-based passive islanding detection technique was proposed in this sub-task, which is immune to additive instrumentation channel errors present in the PMU measurements. The cumulated sum of voltage phase angle difference (CUSPAD) obtained from a specific PMU device over a given time-period successfully cancels the effect of instrumentation channel errors present in the PMU measurements. This is the underlying reason behind the increased islanding detection accuracy and its immunity to instrumentation channel errors. The proposed approach was tested on an 18-bus system using DT based CART classifier. The results indicate that in the presence of instrumentation channel errors the proposed CUSPAD technique was superior to conventional AD approach. The performance of this technique is further evaluated for the IEEE 118-bus system where 10%, 20%, and 30% wind penetration was modeled by replacing corresponding conventional generation. The performance of the CUSPAD approach was also found to be superior for the 118-bus system in presence of increasing amounts of instrumentation channel errors when compared with that of the AD approach using both DTs and RFs. We can therefore conclude that islanding detection in renewable rich systems using CUSPAD is more reliable than AD in presence of additive instrumentation channel errors.

2.2 Task 1.2: Power system online-asset health monitoring

The power grid is the most important of all critical infrastructures as it has the highest degree of influence over other critical infrastructures [33]. Some of the most important equipment of the bulk power system (BPS) are generators, transmission lines and large power transformers (LPTs). LPTs are typically located in open-air switchyards, where they are at the mercy of elements of nature, and more recently, trigger-happy humans [34]. An untimely loss of an LPT can be catastrophic for not only the electric power grid, but also the other critical infrastructures that depend on it for normal operations [33].

One of the biggest challenges faced by the electric power industry is the successful management of its aging infrastructure [33]. Some of the equipment being worth millions of dollars, power utilities want to achieve maximum return of investment (ROI) on the purchase of their equipment. At the same time, a failure of the power system's critical assets at a crucial time-period may be catastrophic for the reliable operation of the grid. Therefore, the way forward is to create an *asset health monitoring* and *management system* that can not only continuously track the condition of critical power system equipment, but also generate alarms sufficiently in advance so that necessary interventions can be made. The motivation for doing this research came from an actual LPT failure that took place at the Avondale substation of Salt River Project (SRP), a power utility located in Arizona, USA on June 1, 2016. During the early morning hours of that day, a large power transformer at the 500/230 kV substation in Avondale, Arizona caught fire, leading to the burning of 27,000 gallons of mineral oil [35]. Such an untimely loss of an LPT could be catastrophic for the electric power grid [36].

Different types of sensors are placed in the power network to monitor the health of its critical equipment. The sensors for LPTs include online DGAs, PQ meters, PD testers and bushing monitors. However, the output generated by most of these sensors are not monitored continuously. It is only when they generate an alarm that their outputs are considered for decision-making. Now, it is quite possible that these sensors generate alarms when the device is very close to an imminent failure, and no possible intervention (at that stage) can prevent the failure and/or subsequent disruption from occurring. PMUs are becoming popular in the US power grid. As per the NASPI report [37], 2,800 PMUs have already been installed in North America by 2014. The numbers will be much higher now. PMUs provide time-synchronized measurements of voltage and current phasors at the time-scale of 30 to 60 samples per second [38]. The use of intelligent electronic devices (IEDs) in monitoring the health of power system equipment was first demonstrated by Jones et al. in [39]. The IEDs used for that analysis were PMUs, which provide voltage and current measurements at the locations where they are placed. Most prior research work on health monitoring carried out using PMUs has been directed towards improving the security/stability of the BPS [40]-[42]. A sudden failure of LPTs may considerably worsen power system security/stability; therefore, real-time health monitoring of LPTs is a task worth undertaking. The main question that is asked in this task of the PSERC project is as follows: *can PMU measurements capture the deteriorating health of an LPT in advance?* To answer this question, we need to first address the following issues:

- Since PMUs provide both magnitude and angle information, it is not clear how inferences drawn from voltage magnitudes in Volts/kilo Volts or current magnitudes in Amperes/kilo

Amperes can be compared with those drawn from angles in degrees/radians. Thus, there is a need of a *metric which is independent of the unit of the measured quantity*.

- PMU data quality can be affected by equipment lying in the vicinity of the PMU. Therefore, one needs to identify the equipment which is the *unique* cause for PMU-data quality degradation. To circumvent this problem, we have used a mathematical technique called *Discriminating Code*, which is a modification of the more well-known Identifying Code concept [43], [44].

The proposed research addressed all the above-mentioned needs. The salient contributions of this research are: (1) first documented research exploiting role of PMU measurements in LPT health assessment; (2) first time use of *Discriminating Code* to solve a power system engineering problem; (3) validating proposed algorithm with field data.

2.2.1 Background of Avondale LPT failure

During the early morning hours of June 1, 2016, a large power transformer at the 500/230 kV substation in Avondale, Arizona, caught fire, leading to the burning of 27,000-gallons of mineral oil [35]. Due to system redundancy as well as low load conditions, no power outage occurred. For the analysis done here, SRP provided *real PMU data for a two-year period* for the 500-kV and 230-kV substations located close to the Avondale substation. The primary objective of the analysis was to investigate whether PMU measurements obtained from locations near to a power transformer could be utilized to make accurate predictions about any degradation in the health of the transformer *before* the failure actually happened.

2.2.2 Robust metric for asset health indicator

Most equipment failures build-up slowly over time; hence signs of an impending failure may be observable days before the actual failure occurs [39]. PMUs can act as sensors for LPT health monitoring, only if a suitable indicator/metric is found. Signal-to-noise (SNR) ratio is a classical statistical measure of the strength of the desirable components to the undesirable components present in a signal. Mathematically, SNR can be expressed as shown below.

$$SNR = 10log\left(\frac{\mu}{\sigma}\right) \quad (13)$$

where, μ represents the mean and σ denotes the standard deviation of the signal. It must be noted that SNR is a relative metric. This is the reason why SNR can be applied to magnitude and angle signals independently. It was also observed in [39], that the variations in SNRs were better indicators of deteriorating equipment health than actual PMU measurements. Thus, in this research SNR of PMU measurements was considered for LPT health assessment and not raw measurements.

The analysis done in this research is based on the real PMU data provided by SRP. PMU measurements from 10 substations located (S1-S10) up to an electrical distance of six hops from Z0 (Avondale substation) were collected. Three scenarios were considered to validate our hypothesis *that the SNR of PMU measurements is a good indicator of equipment health*, as described below: Scenario I: One year away from failure, Scenario II: One month away from failure, and Scenario III: On the day of failure. For each of the above scenarios, the SNR of the

PMU measurements from a substation that was close to the LPT of the Avondale substation was obtained. In our study, we refer that specific substation as S_2. Fig. 2.6, Fig. 2.7, and Fig. 2.8 depict the variation of SNRs of the voltage magnitude obtained from the PMU at substation S_2, for scenarios I, II, and III respectively.

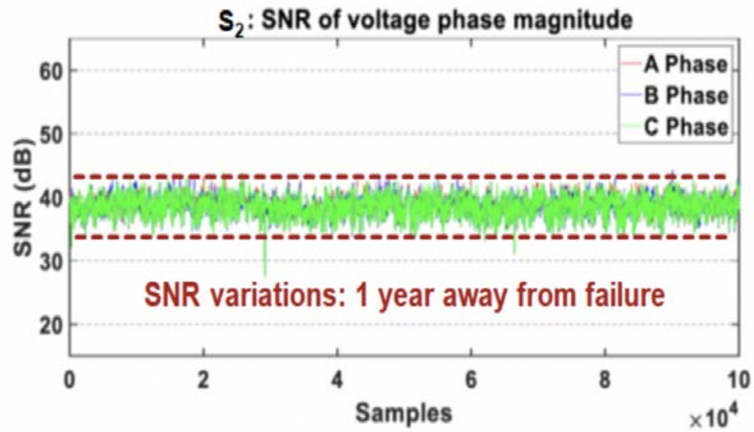


Fig. 2.6: SNR variations of PMU measurements (1 year away from failure)

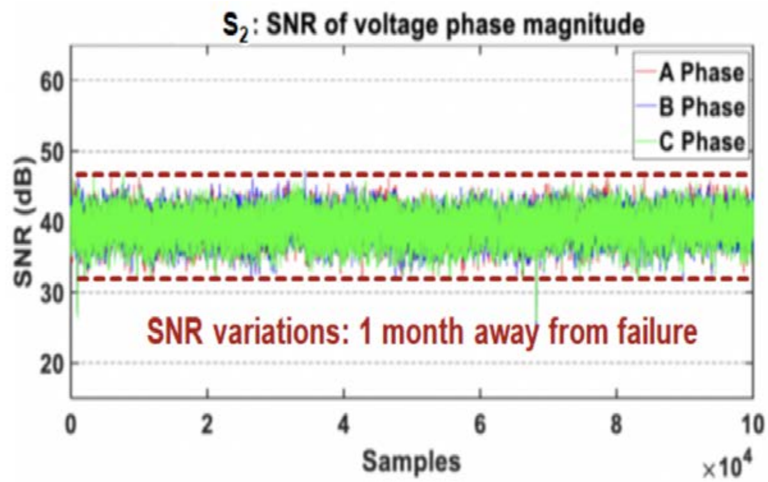


Fig. 2.7: SNR variations of PMU measurements (1 month away from failure)

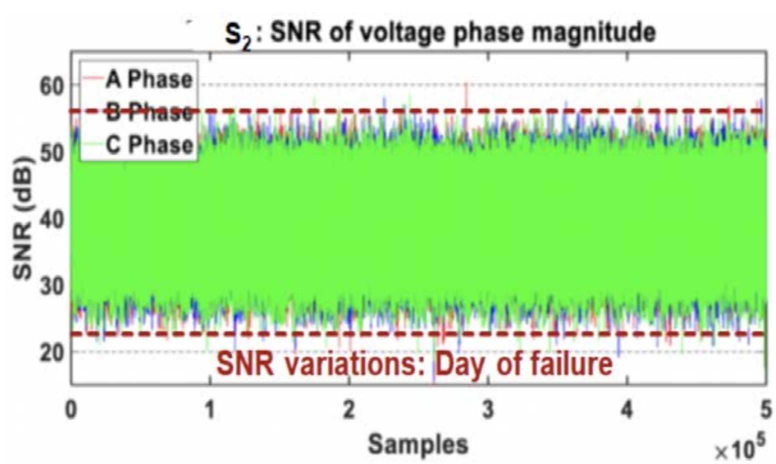


Fig. 2.8: SNR variations of PMU measurements (on the day of failure)

To analyze the rate of increase of the SNR bandwidth obtained from PMU measurements, it is important to quantify and compare the SNR bandwidth at different time-periods before the actual failure. In this research, *we calculated the standard deviation of SNR over a given time window*, for investigating how the SNR profile changes as an equipment moves toward an impending failure. Fig. 2.9 indicates that the standard deviation starts to increase from approximately 75 hours before failure. The standard deviation increases from 5 dB to 25 dB by the time of failure. In the figure, "Vm_a", "Vm_b" and "Vm_c" refers to the voltage magnitude for phases A, B, and C, respectively. The consistent increase of the SNR is an indication of the degradation of the equipment health. Moreover, (a) *as all the three phases show a similar trend in their SNRs, the phenomenon is a three-phase event and not a single-phase event*, and (b) *as the SNR trends are consistent over a time-period ranging in hours (and not seconds or minutes), the phenomenon being captured is not a transient event*.

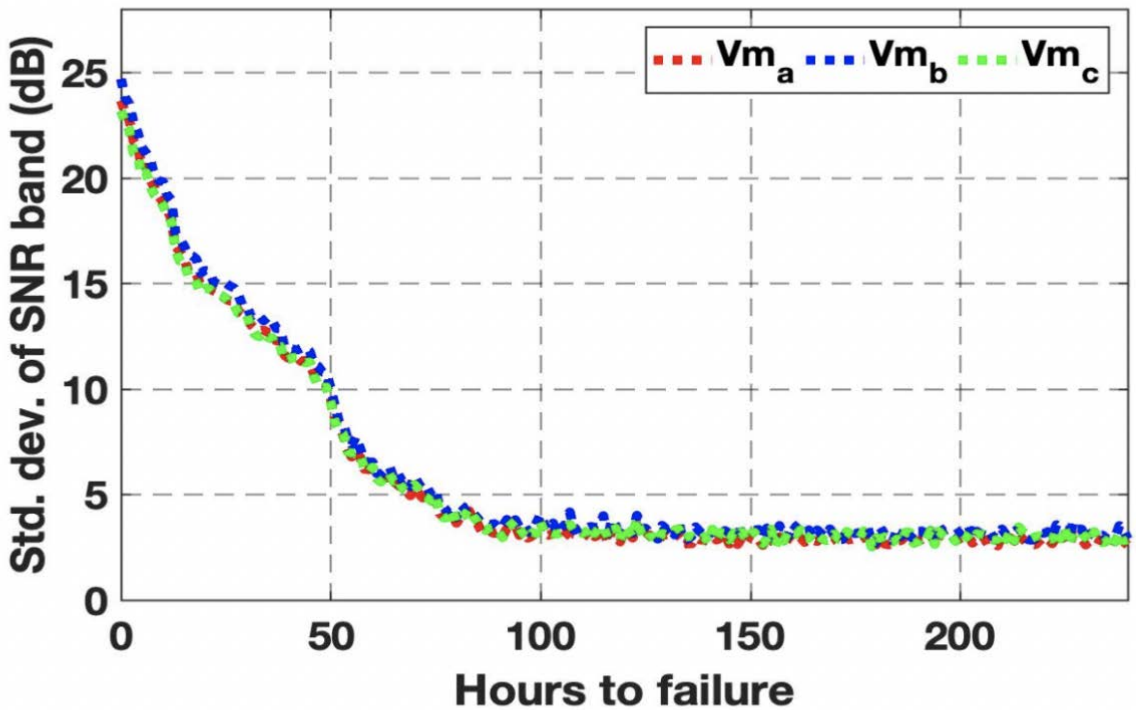


Fig. 2.9: Variations in the standard deviation of SNR (before a failure) at a substation which is two hops away from Avondale

Fig. 2.10 plots the standard deviation of the SNR band after the failure had occurred. It is interesting to observe that as soon as the transformer at the Avondale substation had tripped, the standard deviation of the SNR band fell to its normal operating range. This observation provides additional justification that the SNR computed from the PMU measurements is a robust metric that can identify malfunctioning of a power system equipment.

Next, it was investigated how the increase in the electrical distance (termed hop in this research) between a substation and the monitored equipment affects the standard deviation of the SNR variations. For this study, a time-period between 12:00 AM and 5 AM on June 1, 2016 was selected. It can be observed from Fig. 2.11 that for all three phases, a *monotonic decrease* in the standard deviation of SNR band of real component of current occurred with increasing distance from the transformer that failed. Fig. 2.11 also indicates that the measurements obtained within three hops of the failing transformer (Z0) were better indicators than the substations that are further away. This is an important observation, because it confirms that the problem is located in the vicinity of Z0 and is not a system-wide event.

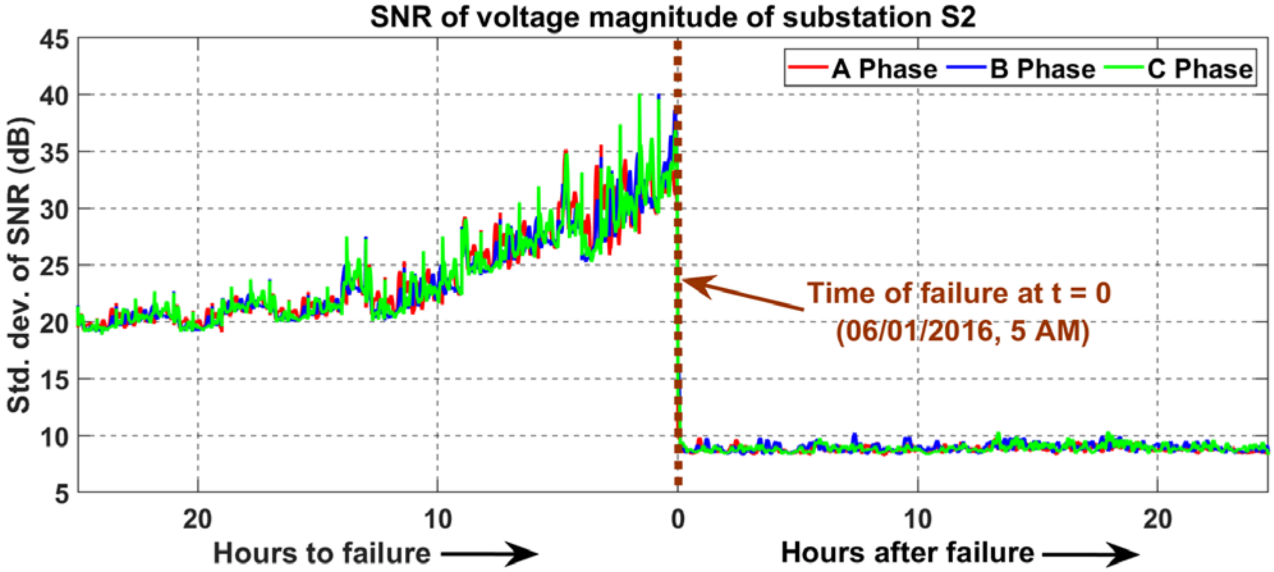


Fig. 2.10: Variations in standard deviations of SNR (after failure) at a substation which is two hops away from Avondale substation

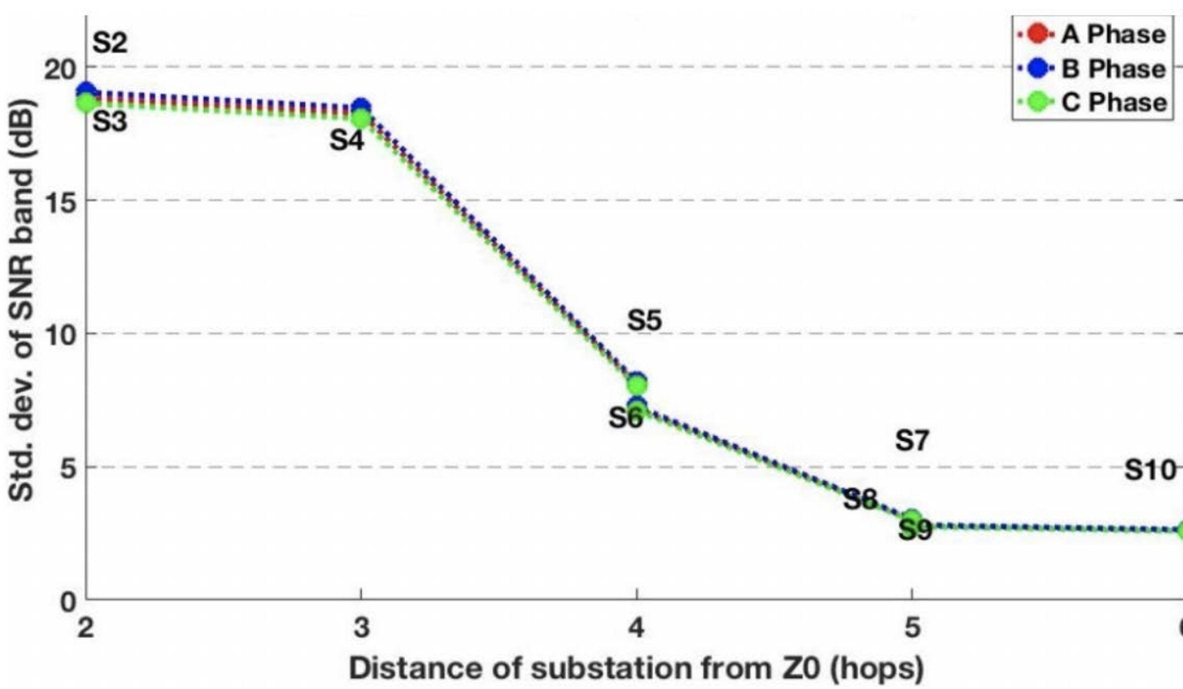


Fig. 2.11: Variation in standard deviation of SNR band with electrical distance for real component of current on the day of transformer failure

2.2.3 Optimal sensor selection using Discriminating Code

2.2.3.1 Theoretical background

A given PMU device may be affected by different equipment lying in its vicinity. Therefore, research must be done to uniquely identify the equipment which is the primary cause of degradation in the quality of the PMU measurements. In other words, all the important power system equipment must be associated with a unique signature. *Identifying code* is a mathematical abstraction that enables the unique detection of one or more objects of interest, by generating exclusive signatures for those objects [43]. The concept of *Identifying Code* would enable us to uniquely monitor every LPT.

The simplest definition of the *Identifying Code set problem* is as follows: *for any graph $G = (\boldsymbol{V}, \boldsymbol{E})$, a vertex set $\boldsymbol{V}' \subseteq \boldsymbol{V}$ is defined as the Identifying Code set (ICS) for the vertex set $\boldsymbol{V}$, if $\forall\, v \in \boldsymbol{V}, \boldsymbol{N}^{+}(v) \cap \boldsymbol{V}'$, is unique where, $\boldsymbol{N}^{+}(v) = v \cup \boldsymbol{N}(v)$, and $\boldsymbol{N}(v)$ represents the set of nodes adjacent to v in $G = (\boldsymbol{V}, \boldsymbol{E})$*. The *minimum Identifying Code set* (MICS) problem finds the identifying code set of smallest cardinality. The *Identifying Code* set becomes useful if the objective is to uniquely monitor all the nodes in a graph. However, as our goal here is to uniquely monitor the health of LPTs only, the formulation of Identifying Code set problem must be modified. We now provide a definition of the modified version of the *Identifying Code* that is known as the *Discriminating Code*. *Given a bipartite graph $G = (\boldsymbol{V}_1 \cup \boldsymbol{V}_2, \boldsymbol{E})$, for any vertex $v \in \boldsymbol{V}_1$, a subset $\boldsymbol{V}_2' \subseteq \boldsymbol{V}_2$ is called the Discriminating Code of G, if $\forall v \in \boldsymbol{V}_1, \boldsymbol{N}(v) \cap \boldsymbol{V}_2'$ is unique*. This is explained with an example. For the power system we consider that the set $\boldsymbol{V}_1$ contain the transformers and the set $\boldsymbol{V}_2$ contain the buses. $\boldsymbol{V}_2'$ is the subset of the set $\boldsymbol{V}_2$ where the sensors are to be placed, such that all the transformers contained in the set $\boldsymbol{V}_1$ are uniquely observed. The set $\boldsymbol{N}(v) \cap \boldsymbol{V}_2'$ refer to the set of PMUs observing any transformer $v \in \boldsymbol{V}_1$. The objective of the Discriminating Code problem is to find the minimum set of vertices $\boldsymbol{V}_2'$ for which $\boldsymbol{N}(v) \cap \boldsymbol{V}_2'$ is unique for any transformer v contained in the set $\boldsymbol{V}_1$. The mathematical formulation of the Discriminating Code set problem and its application to the LPT health monitoring is described as follows.

2.2.3.2 Mathematical formulation

Monitoring of critical equipment (MCE) problem: For the power system graph, let $\boldsymbol{V_1}$ denotes the set of critical equipment (namely, LPTs) and $\boldsymbol{V_2}$ denotes the set of buses. $\boldsymbol{V_2'}$ is the subset of set $\boldsymbol{V_2}$, where sensors must be placed, such that all the critical equipment contained in $\boldsymbol{V_1}$ are uniquely observed. Further, $\boldsymbol{N}(\boldsymbol{v}) \cap \boldsymbol{V_2'}$ refers to the set of PMUs observing any critical equipment $\boldsymbol{v} \in \boldsymbol{V_1}$. Then, the objective of the monitoring of critical equipment (MCE) problem is to find out the minimum set of vertices, $\boldsymbol{V_2'}$, for which $\boldsymbol{N}(\boldsymbol{v}) \cap \boldsymbol{V_2'}$ is unique for any critical equipment $\boldsymbol{v} \in \boldsymbol{V_1}$. An ILP formulation to solve the MCE problem is now described.

Let a binary variable x_i be associated with every node $v_i \in \boldsymbol{V}_2$, such that,

$$x_i = \begin{cases} 1, & \text{if a PMU is placed at node } \mathrm{v_i} \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

The objective function is formulated as shown below:

$$Minimize \sum_{\forall v_i \in \boldsymbol{V}_2} x_i \tag{15}$$

The observability and the unique observability constraints are obtained from the two equations shown below respectively.

$$\sum_{v_i \in \boldsymbol{N}^{\boldsymbol{k}}(v_j)} x_i \geq 1 \ \forall v_j \in \boldsymbol{V}_1 \tag{16}$$

$$\sum_{v_i \in \{\boldsymbol{N}^{\boldsymbol{k}}(v_j) \oplus \boldsymbol{N}^{\boldsymbol{k}}(v_l)\}} x_i \geq 1 \qquad \forall v_j \neq v_l \ \in \boldsymbol{V}_1 \tag{17}$$

In equations (16) and (17), $\boldsymbol{N}^{\boldsymbol{k}}(v_j)$ denotes the neighborhood of bus v_j within an electrical distance of k hops. In (17), $\boldsymbol{N}^{\boldsymbol{k}}(v_j) \oplus \boldsymbol{N}^{\boldsymbol{k}}(v_l)$ denotes the symmetric difference operation of the node sets $\boldsymbol{N}^{\boldsymbol{k}}(v_j)$ and $\boldsymbol{N}^{\boldsymbol{k}}(v_l)$. It may be noted that the objective function ensures that the fewest number of nodes in $\boldsymbol{V}_2$ are installed with a PMU. The observability constraint ensures that every node in $\boldsymbol{V}_1$ receives at least one signature via the PMUs installed in $\boldsymbol{V}_2$. A consequence of the observability constraint is that a node in $\boldsymbol{V}_1$ may receive more than one signature from the PMUs installed at the nodes in $\boldsymbol{V}_2$. The unique observability constraint ensures that, for every pair of nodes (v_j, v_l) in $\boldsymbol{V}_1$, at least one node in the node set $\boldsymbol{N}^{\boldsymbol{k}}(v_j) \oplus \boldsymbol{N}^{\boldsymbol{k}}(v_l) \subseteq \boldsymbol{V}_2$ is associated with a PMU. This guarantees that v_j and v_l will not have same *identifying signature*.

Augmented monitoring of critical equipment (AMCE) problem: It must be noted that the MCE problem determined the locations where the sensors can be placed for unique monitoring of the LPTs, *assuming that no sensors were initially present in the system*. We now propose an enhanced variant of the MCE problem, called the augmented monitoring of critical equipment (AMCE) problem, which accounts for the presence of pre-existing sensors (or PMUs) in the network. It can be formally stated as follows. *Given a bipartite graph $G = (\boldsymbol{V}_1 \cup \boldsymbol{V}_2, \boldsymbol{E})$ and a set $\boldsymbol{V}_2' \subseteq \boldsymbol{V}_2$, determine the smallest subset $\boldsymbol{V}_2'' \subseteq \boldsymbol{V}_2$, such that, $\forall v \in \boldsymbol{V}_1$, $N^k(v) \cap \boldsymbol{V}_2'''$ is unique; where $\boldsymbol{N}^k(v)$ represents the k-hop neighbors of v, and $\boldsymbol{V}_2''' = \boldsymbol{V}_2' \cup \boldsymbol{V}_2''$.* In the AMCE problem, we consider that a certain number of PMUs, present in $\boldsymbol{V}_2'$, have been pre-installed at some of the buses in the system. Our goal is to determine the smallest subset of $\boldsymbol{V}_2$, which when augmented with $\boldsymbol{V}_2'$, can uniquely monitor the nodes in $\boldsymbol{V}_1$. An ILP formulation to solve the AMCE problem is now described.

Let a binary variable x_i be associated with every node $v_i \in \boldsymbol{V}_2$, such that,

$$x_i = \begin{cases} 1, & \text{if a PMU is placed at node } \mathrm{v_i} \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

The objective function is stated as follows:

$$Minimize \sum_{\forall v_i \in \boldsymbol{V}_2} x_i \tag{19}$$

The observability, unique-observability and the pre-installed sensor location constraints are obtained from (20), (21) and (22).

$$\sum_{v_i \in \boldsymbol{N}^k(v_j)} x_i \geq 1 \;\; \forall v_j \in \boldsymbol{V}_1 \tag{20}$$

$$\sum_{v_i \in \{\boldsymbol{N}^k(v_j) \oplus \boldsymbol{N}^k(v_l)\}} x_i \geq 1 \qquad \forall v_j \neq v_l \;\in \boldsymbol{V}_1 \tag{21}$$

$$\forall x_i \in \boldsymbol{V}_2', x_i = 1 \tag{22}$$

The constraints with (20) and (21) are the same as those in the MCE problem. The constraint specified in (22) accounts for the sensors already placed in the network.

2.2.3.3 Performance evaluation of Discriminating Code

This section presents the results of the Discriminating Code set problem on the IEEE standard test systems (IEEE 14, 30 and 118-bus systems), and very large power systems (2,383-bus Polish, 2,603-bus SRP, and 22,978-bus Western Electricity Coordinating Council (WECC) systems). Studies were conducted for $k = 1,2,3$. The application of the Discriminating Code set problem to the power system is described in detail using the IEEE 14-bus test system shown in Fig. 2.12. There were 5 LPTs and 40 potential sensor locations. This is based on the assumption that the PMUs are of the dual-use line relay (DULR) type, which are placed on buses and monitor either ends of the branches of the power network [32]. It is found that for $k = 3$, the 5 LPTs can be monitored by 3 sensors. In Fig. 2.12, for $k = 3$ if the 3 selected sensors are located at nodes B6, B8, and B11 (or, three colors A, B and C are injected at these nodes), transformers T1 through T5 will receive unique signatures AB, ABC, AC, B, and BC, respectively. The fact that *every transformer* of the IEEE 14-bus system receives a *unique signature*, is of *great practical significance*, because in such a case the degrading quality of *every LPT* could be *uniquely determined* by the sensors monitoring them.

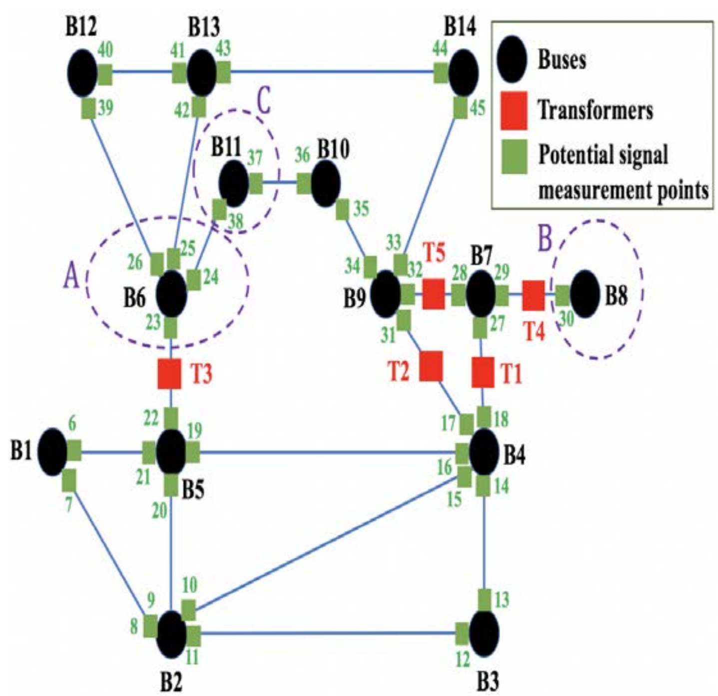


Fig. 2.12: Discriminating code result of the IEEE 14 bus system

The sensor selection results for the IEEE test-systems and the three large power systems are shown in Table 3.1, for the MCE problem. For all the systems, it is observed that the number of sensors (S) required to monitor all the LPTs is less than or equal to the number of LPTs. Fig. 2.11 showed that there is minimal difference between the nature of variations of standard deviations of the SNR bands at substations S3 and S4, which are two and three-hops away, respectively from Z0. Therefore, it is suitable to consider $k = 3$ in the proposed formulation for the selection of sensors, which results in a significant reduction in the number of sensors required. For example, with $k = 3$, for the WECC system, the number of sensors required to monitor all the LPTs is approximately 40% lesser than that required for $k = 1$. GUROBI for python was used to solve the mathematical formulation described in Section 3 for the six power networks. An Intel Core i5-6300HQ CPU with 2.30 GHz and 32 GB RAM was used for performing the different simulations. The time taken to solve the MCE problem was 0.17 seconds for the smallest test systems (14-bus system, $k = 1$) and 392 seconds for the largest test system (22,978-bus system, $k = 3$). The results indicate that the proposed formulation can be successfully applied to real-world power systems.

Table 2.5 shows the results obtained for the AMCE problem. Pre-existing PMUs were assumed to be placed on the highest voltage buses of the system that were also close to the large generators. The results can be explained as follows: for the 22,978-bus system, for $k = 3$, the minimum number of additional sensors (AS) required to uniquely monitor all the LPTs is 4,397; the corresponding CPU time was also reasonable (=457 seconds).

Table 2.4: Results for the MCE problem

System	#Transformer	$\boldsymbol{S_{k=1}}$	$\boldsymbol{S_{k=2}}$	$\boldsymbol{S_{k=3}}$
IEEE 14-bus	5	4	3	3
IEEE 30-bus	7	6	4	4
IEEE 118-bus	9	9	5	5
Polish 2383-bus	155	155	106	76
SRP 2603-bus	1145	1145	756	595
WECC 22978-bus	8999	8999	6020	5127

Table 2.5: Results for the AMCE problem

System	#Transformer	$\boldsymbol{AS_{k=1}}$	$\boldsymbol{AS_{k=2}}$	$\boldsymbol{AS_{k=3}}$
IEEE 14-bus	5	4	2	2
IEEE 30-bus	7	6	4	2
IEEE 118-bus	9	6	2	3
Polish 2383-bus	155	139	95	66
SRP 2603-bus	1145	664	478	393
WECC 22978-bus	8999	7189	5289	4397

2.2.3.4 Summary of the findings

A novel approach for monitoring and predicting the health of LPTs by utilizing PMU measurements from substations was proposed. The studies were based on an actual power transformer failure that occurred in the transmission network of a large power utility in the US Southwest. By enhancing an existing sensor selection approach, a new technique called the *Discriminating Code* technique was developed to appropriately select and augment the sensors which might be already deployed in the network to uniquely monitor the health of LPTs. In the future, we plan to investigate whether, (i) the proposed approach can be used to classify the health status of critical equipment into normal, alert, or alarm categories, (ii) one can perform identification of events when one or more sensors are malfunctioning, and (iii) inferences obtained from PMU measurements compare favorably with those obtained from other sensors.

3. PMU based power system cyber-protection applications

3.1 Task 2.1: PMU-based load prediction and monitoring

One of the goals of this project is to create a realistic synthetic test system that can be used to verify the performance of the proposed PMU-based applications as well as to test cyber-attacks and countermeasures. One of the crucial aspects in designing a testbed that can be used to observe system behaviors at PMU sampling speeds is to accurately model the behavior of the system loads. In fact, not considering faults or other such unpredictable events, the dynamics of a system are governed by the variation of loads over time and how the generators respond to such changes. In this section we describe a data-driven algorithm for the generation of synthetic, bus-level, time-series load data at 30 samples per second that can be used on any system model. The approach we adopted in this section is mainly based on the work described in [44]. In that paper, we have presented a data-driven algorithm that can be used to learn the spatial and temporal correlation from a dataset of real system loads and use the learnt model to generate new synthetic data that retains the same characteristics. Fig. 3.1 presents a diagram illustrating the main blocks of the algorithm and the required input data. Given real time-series load data, a matrix factorization technique called singular value decomposition (SVD) is used to identify and extract typical load patterns from the data, each describing different load behaviors over time. Moreover, based on the system topology, it is possible to study how the relationship between different loads is influenced by their geographical location. After learning these characteristics from the real data, synthetic but realistic load profiles can be generated for a new system taking into consideration its topology. The time series data for each new bus is created by taking combinations of the typical patterns learnt from the real loads and adjusting them to reflect the same spatial correlation observed in the real system.

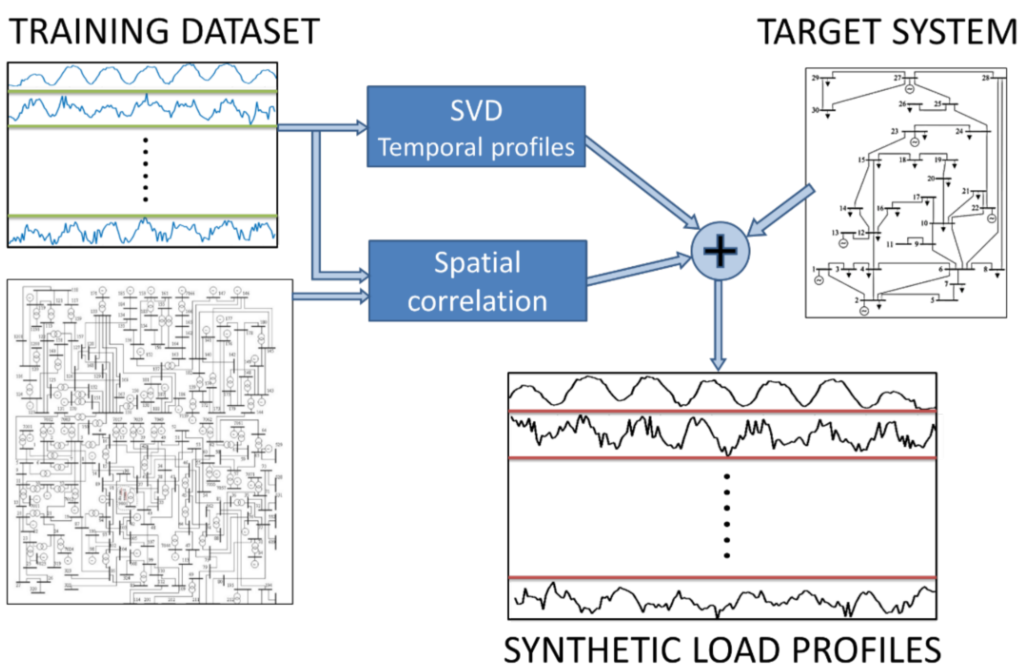


Fig. 3.1: Synthetic load generation scheme

In [44], this approach is demonstrated on SCADA based hourly load data. In this report, we provide the details on how this technique was used on PMU data at 30 samples per second. Salt River Project (SRP), a power utility in Arizona, provided us with one-week worth of PMU data for a group of neighboring substations. As illustrated in Fig. 3.2, from the voltage and current measurements of each bus and line, it is possible to compute the net load of a substation. From the data available to us, we were able to calculate the loads of two different substations, one at the 500-kV level and one at 230-kV level.

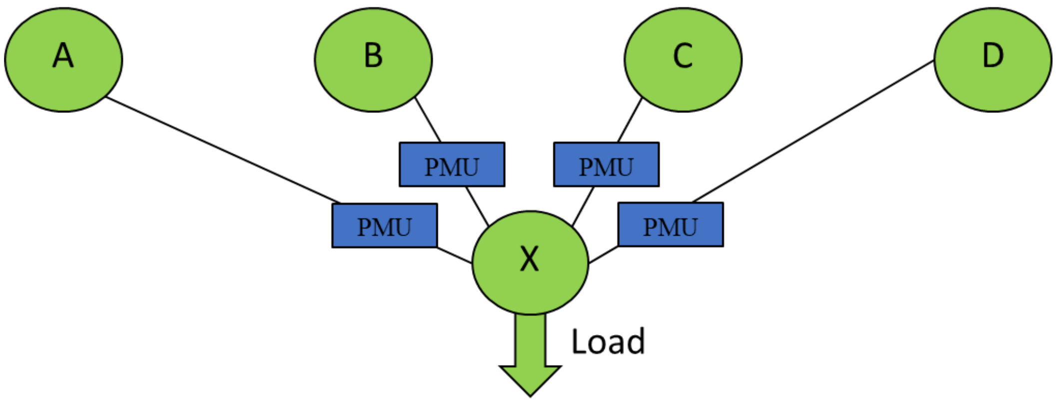


Fig. 3.2: Estimating a load from a PMU

Each of the two data streams (230-kV load and 500-kV load) were analyzed independently in accordance with the following six-step procedure:

1) The time-series load data for one consecutive week is broken into segments of length of 10 minutes; this results in around 1,008 segments. Each segment contains: 30 samples/sec×60 sec/min×10 min=18,000 samples. The segments are then stacked to form the load matrix $\boldsymbol{P} \in \mathbb{R}^{\mathbf{1008\times18000}}$.

2) The load matrix $\boldsymbol{P}$ is factorized using singular value decomposition: $\boldsymbol{P} = \boldsymbol{U\Sigma V^T}$. The rows of $\boldsymbol{V^T}$, which are vector of size 1×18,000 samples, correspond to archetypal *temporal profiles*. Each element of the diagonal matrix $\boldsymbol{\Sigma}$, called a *singular value*, represents a scale factor which multiplies each corresponding temporal profile. Moreover, because the singular values are ordered by magnitude, they give an indication of the relative importance of each temporal profile. The synthetic loads are generated by taking linear combinations of the first rows of $\boldsymbol{V^T}$ (load basis).

3) To determine the number of basis (or temporal profiles) to be used in the generative model it is useful to look at approximations of $\boldsymbol{P}$, defined as $\widehat{\boldsymbol{P}} = \boldsymbol{U^f}\ \boldsymbol{\Sigma^f}\ \boldsymbol{V^f}^{\boldsymbol{T}}$, where $\boldsymbol{U^f}$ indicates the first $\boldsymbol{f}$ columns of $\boldsymbol{U}$, $\boldsymbol{\Sigma^f}$ first $\boldsymbol{f}$ columns and $\boldsymbol{f}$ rows of $\boldsymbol{\Sigma}$, and $\boldsymbol{V^f}$ first $\boldsymbol{f}$ columns of $\boldsymbol{V}$. By varying the value of $\boldsymbol{f}$ (corresponding to the number of basis to be used) in steps of 1 and measuring the root mean square error (RMSE) between $\boldsymbol{P}$ and $\widehat{\boldsymbol{P}}$ we have determined an appropriate number of base temporal profiles to be used by the generative model. In Fig. 3.3 the error is plotted as a function of the number of basis used. It can be seen that the error decreases rapidly up to $\boldsymbol{f}$=10 and then it slowly reaches zero when all the basis are used. For this reason, we used the first 10 temporal profiles in our generative model.

4) Having identified some typical temporal load patterns, a new load profile can be created by generating a vector of coefficients and multiplying it by the set of base profiles contained in $\boldsymbol{V}$. To compute these new coefficients, we need to learn the distribution of the coefficients in the original data (e.g. the first 10 columns of $\boldsymbol{U}$). Using the MATLAB distribution fitter app, a Gaussian distribution if fitted to each column of $\boldsymbol{U}$.

5) A new matrix of load profiles for $\boldsymbol{n}$ buses is generated as: $\boldsymbol{P_{new}} = \boldsymbol{U_{new}^{10}}\, \boldsymbol{\Sigma^{10}}\, \boldsymbol{V^{10^T}}$, where $\boldsymbol{P}_{\mathbf{new}}$ is a n×18,000, $\boldsymbol{U_{new}^{10}}$ is a n×10 matrix of coefficients randomly sampled from the distributions learnt in step 4, $\boldsymbol{\Sigma^{10}}$ and $\boldsymbol{V^{10^T}}$ represent the first 10 singular values and first 10 temporal profiles respectively.

6) To account for the spatial correlation which exists between neighboring loads, the model is modified as $\boldsymbol{P_{new}} = (\boldsymbol{D}\boldsymbol{U_{new}})\boldsymbol{\Sigma}\boldsymbol{V^T} = \boldsymbol{U'_{\mathbf{new}}}\boldsymbol{\Sigma}\boldsymbol{V^T}$, where each entry $\boldsymbol{d_{i,j}}$ of $\boldsymbol{D}$ is as follows:

$$\boldsymbol{d_{i,j}} = \begin{cases} \mathbf{1}, & \mathbf{if\ \ i = j} \\ \boldsymbol{e^{-2dist_{i,j}}}, & \mathbf{if\ dist_{i,j} \leq 3\ and\ \ i \neq j} \\ \mathbf{0}, & \mathbf{otherwise} \end{cases} \qquad (\mathbf{23})$$

Where, $dist_{i,j}$ is the distance in hops between bus i and bus $\boldsymbol{j}$. Overall, this relation was experimentally determined in [44] and was adapted to the system for which we designed the synthetic loads. This step ensures that neighboring buses have higher correlation compared to loads which are further apart. Fig. 3.3 shows the results of applying this correction factor to the generative model. In particular, these plots represent the average and percentiles of the correlation coefficient as a function of the distance between buses. The plot on the left shows these statistics for the real data: as expected, the closer two buses (small distance) the more similar the load profiles are (correlation coefficient close to 1). The center and right plots represent the correlation in the synthetic data we generated with and without applying the correction factor $\boldsymbol{D}$; when the spatial correlation is not considered, the distance-dependence is lost.

The process described above was used to generate individual load profiles for 10 minutes for the loads in the IEEE 118-bus system. Fig. 3.4 depicts the synthetic load profiles generated for two adjacent loads (at buses 93 and 94). It can be seen that they show a similar pattern over a period of 10 minutes, due to strong spatial correlation. Fig. 3.5 depicts a sample load profile at bus 15 which is far away from the buses 93 and 94. As expected, due to weaker spatial correlation between far away buses, the load profile at bus 15 is very different from the load profiles at buses 93 or 94.

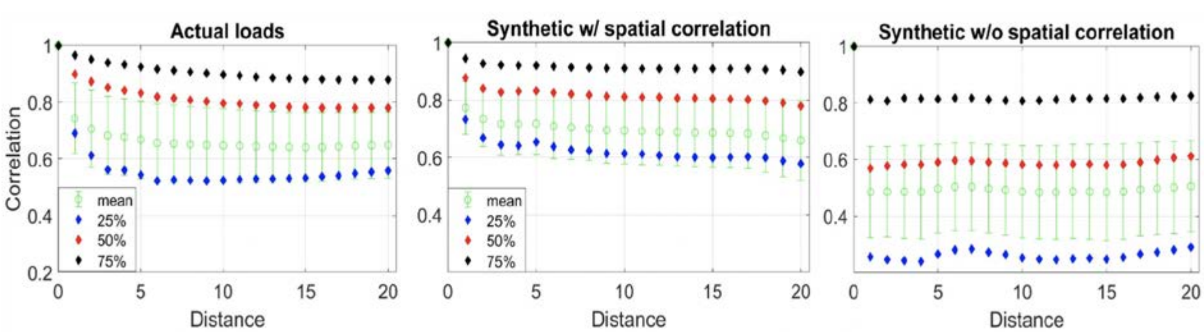


Fig. 3.3: Statistics of the correlation coefficients between load profiles as a function of the distance between buses

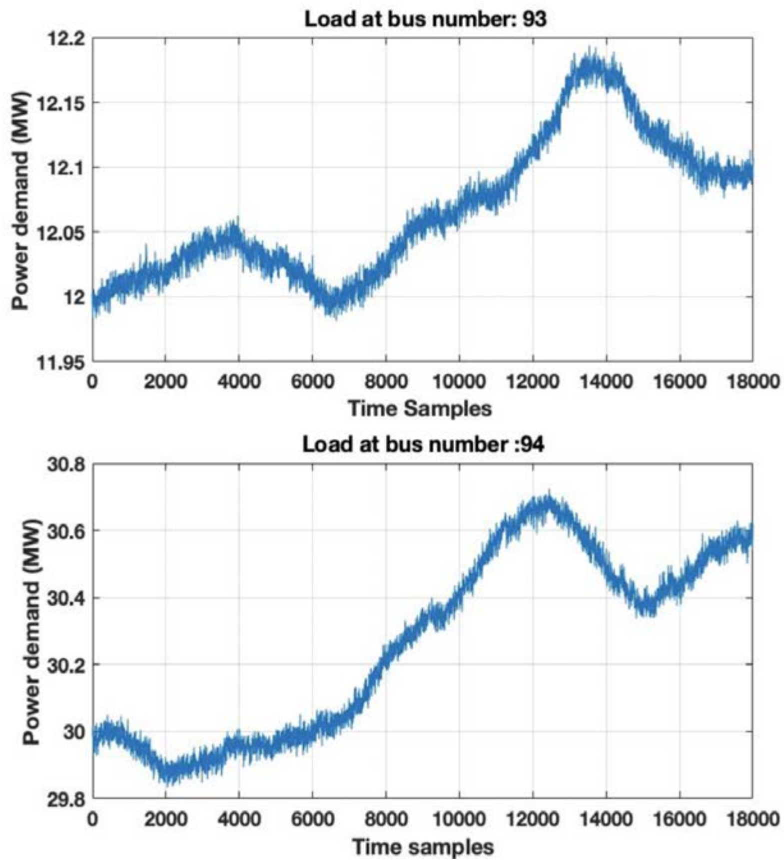


Fig. 3.4: Example load profiles for two neighboring buses

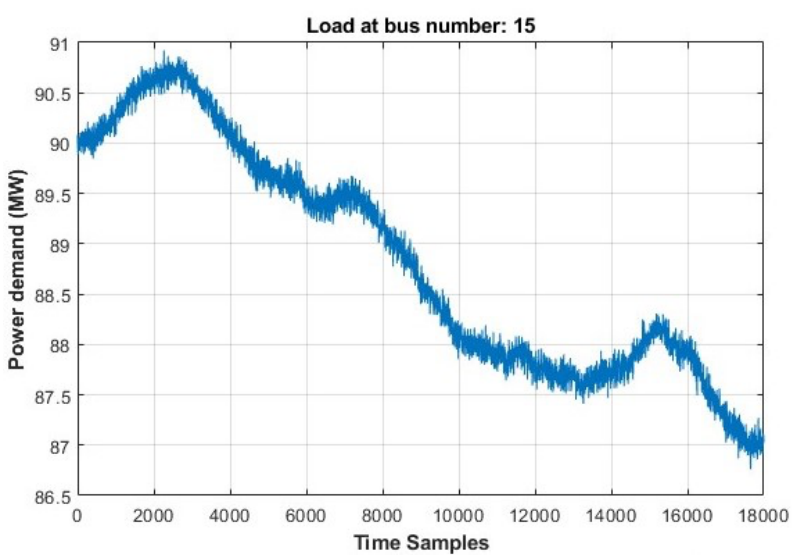


Fig. 3.5: Example of load profile at a bus which is far away from buses 93 and 94

3.2 Assessing vulnerability of PMUs to cyber-attacks

3.2.1 False data injection (FDI) attacks and low rank detector

PMUs have been widely deployed in the electric power system to directly measure the bus voltages and phase angles. Due to their high sampling rate and accuracy, PMUs have the potential to play a significant role in real-time power system state estimation (SE), dynamic security assessment, system protection, and system awareness. Several incidents have demonstrated that the cyber layer of power system is vulnerable to cyberattacks that impact the system operation status and lead to serious physical consequences. As increasingly important monitoring devices, PMUs are also prone to cyber-attacks. Therefore, it is crucial to evaluate the vulnerability of PMUs to potential cyber-attacks.

False data injection (FDI) attacks are a type of cyber-attacks which involve an intelligent attacker who replaces a subset of measurements with counterfeits. We focus on the sub-class of unobservable attacks which render the false data unobservable to the operator. PMU data collected at each time instance is given by $\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}$, where $\boldsymbol{z}$ is the PMU measurement vector consisting of the complex voltage measurements at PMU buses and current measurements on all branches connected to these buses, $\boldsymbol{H}$ is the measurement state dependency matrix, and $\boldsymbol{e}$ is the PMU measurement error vector. As the PMU measurements are all linearly related to the states, the least square solution to this problem is given by $\widehat{\boldsymbol{x}} = \boldsymbol{H}^{+}z$, where $\boldsymbol{H}^{+}$ is the pseudo-inverse of $\boldsymbol{H}$. A conventional residual-based bad data detector (BDD) performs χ^2-test on the measurement residual $\boldsymbol{r} = \boldsymbol{z} - \boldsymbol{H}\widehat{\boldsymbol{x}}$ to detect bad data. In an FDI attack, an attacker may replace $\boldsymbol{z}$ with $\bar{\boldsymbol{z}} = \boldsymbol{z} + \boldsymbol{a} = \boldsymbol{H}(\widehat{\boldsymbol{x}} + \boldsymbol{c})$, where $\boldsymbol{a}$ and $\boldsymbol{c}$ are measurement and state attack vectors, respectively, so that the resulting residual matrix remains unchanged, and the attack is undetected.

However, by representing PMU time-series data as matrix $\boldsymbol{Z}$ whose rows are PMU measurements (voltage and current) at each time instant and given the high data rate of PMUs (typically 30 samples/sec), $\boldsymbol{Z}$ is low rank. The low-rank nature of such a measurement matrix allows for a new detection mechanism. The authors in [46]-[48] propose a new attack detection scheme based on low-rank decomposition (LD) to detect and identify column sparse FDI attacks on PMU data. Assuming an attacker can change the output of a subset of all PMUs in the system, it can launch an attack $\bar{\boldsymbol{Z}} = \boldsymbol{Z} + \boldsymbol{D}$ where $\boldsymbol{D}$ is column sparse. One natural way to form a column sparse measurement matrix is through a column sparse state (complex voltage) matrix $\boldsymbol{C}$ such that $\bar{\boldsymbol{Z}} = \boldsymbol{Z} + \boldsymbol{D} = (\boldsymbol{X} + \boldsymbol{C})\boldsymbol{H}^{\boldsymbol{T}}$, where $\boldsymbol{X}$ is the state matrix.

The proposed LD detector uses an optimization problem to identify the attack matrix as shown in Fig. 3.6. In particular, given a measurement matrix $\bar{\boldsymbol{Z}}^{(LD)}$ and the attack-free measurement matrix $\boldsymbol{Z}^{(LD)}$, the attack matrix $\boldsymbol{C}^{(LD)}$ can be found by solving the following convex optimization problem:

$$\min_{\boldsymbol{Z}^{(LD)} \in \mathbb{C}^{N\times n_z},\, \boldsymbol{C}^{(LD)} \in \mathbb{C}^{N\times n_b}} \left\|\boldsymbol{Z}^{(LD)}\right\|_* + \lambda\left\|\boldsymbol{C}^{(LD)}\right\|_{1,2} \tag{24}$$

$$\text{subject to } \bar{\boldsymbol{Z}}^{(LD)} = \boldsymbol{Z}^{(LD)} + \boldsymbol{C}^{(LD)}\bar{\boldsymbol{H}}^{T} \tag{25}$$

where $\left\|\boldsymbol{Z}^{(LD)}\right\|_*$ is the nuclear norm of $\boldsymbol{Z}^{(LD)}$; $\left\|\boldsymbol{C}^{(LD)}\right\|_{1,2}$ is the $l_{1,2}$-norm of $\boldsymbol{C}^{(LD)}$, i.e., the sum of l_2-norm of columns of $\boldsymbol{C}^{(LD)}$; λ is a weight factor; and $\bar{\boldsymbol{H}}$ is the normalized dependency matrix, where for each row vector $\boldsymbol{H}_i$, $\bar{\boldsymbol{H}}_i = \boldsymbol{H}_i/\|\boldsymbol{H}_i\|$. The objective is to minimize the rank of $\boldsymbol{Z}^{*(LD)}$ (captured by its nuclear norm) and the column sparsity of $\boldsymbol{C}^{*(LD)}$ (captured by its $l_{1,2}$-norm). After obtaining the optimal solution, $\left(\boldsymbol{Z}^{*(LD)},\ \boldsymbol{C}^{*(LD)}\right)$, the set of attacked measurements and states, $supp\left(\boldsymbol{C}^{*(LD)}\overline{\boldsymbol{H}}^T\right)$ and $supp\left(\boldsymbol{C}^{*(LD)}\right)$, respectively, can be identified as the column support of $\boldsymbol{C}^{*(LD)}\overline{\boldsymbol{H}}^T$ and $\boldsymbol{C}^{*(LD)}$. Assume there exists unobservable attacks in $\overline{\boldsymbol{Z}}^{(LD)}$, such that $\overline{\boldsymbol{Z}}^{(LD)} = \boldsymbol{Z} + \boldsymbol{C}\overline{\boldsymbol{H}}^T$. The authors prove that for a specific range of λ, the optimization in (24), (25) can successfully identify $supp(\boldsymbol{C})$, i.e., $supp\left(\boldsymbol{C}^{*(LD)}\right) = supp(\boldsymbol{C})$, under the assumption that every nonzero column of $\boldsymbol{C}\overline{\boldsymbol{H}}^T$ does not lie in the column space of $\boldsymbol{Z}$. In the next sections we introduce two new classes of FDI attacks that cannot be detected by the LD detector and we prove that the LD detector can either detect no attack, or incorrectly identify attacked states.

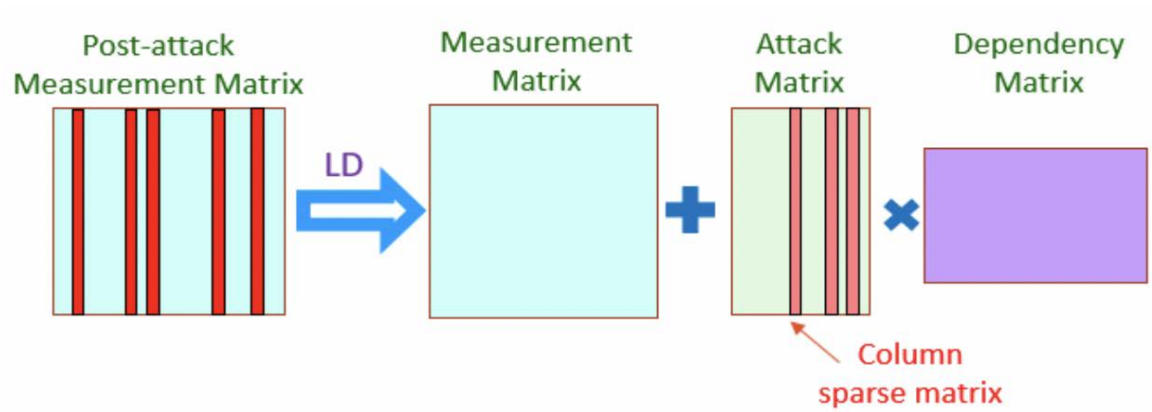


Fig. 3.6: The low-rank decomposition detector

3.2.2 FDI attacks exploiting low-rank property of PMU measurement matrix

In this section, we describe a convex optimization problem that allows an attacker with knowledge of the time correlation of the PMU data to design FDI attacks that can bypass the LD detector. We assume that the attacker has the following knowledge and capabilities:

1) The attacker has full system topology information
2) The attacker can perfectly predict the measurements in the following N instances
3) The attacker has control of the measurements in a subset S of the network

Given a PMU measurement matrix $\boldsymbol{Z}$ and the potential attacked states $\boldsymbol{\mathcal{J}}$, we propose the following optimization problem to design FDI attacks:

$$\text{subject to } supp(\boldsymbol{C}) \subseteq \boldsymbol{\mathcal{J}} \qquad (26)$$

where $\|\cdot\|_*$ denotes the nuclear norm. For optimal solution $\boldsymbol{C}^*$, the optimal post-attack measurement matrix denoted as $\overline{\boldsymbol{Z}}^*$ can be written as

$$\overline{\boldsymbol{Z}}^* = \boldsymbol{Z} + \boldsymbol{C}^*\overline{\boldsymbol{H}}^T \qquad (27)$$

The goal of the attacker is to ensure that the attacked measurement matrix $\overline{\boldsymbol{Z}}^*$ is low-rank when $\boldsymbol{Z}$ is low-rank. This can be approximated by minimizing the nuclear norm of $\overline{\boldsymbol{Z}}^*$ as in (26). Constraint (27) ensures that the attacker can only attack states in $\boldsymbol{\mathcal{J}}$, i.e., $\boldsymbol{C}^*$ is a column-sparse matrix. Moreover, we prove that either $\overline{\boldsymbol{Z}}^*$ bypasses the LD detector, or the LD detector identifies at least one measurement as corrupted which is actually not corrupt. This can be described by the following theorem.

Theorem 2: *Assume the attack-free measurement matrix* $\boldsymbol{Z}$ *can bypass the LD detector, i.e., for* $\overline{\boldsymbol{Z}}^{(LD)} = \boldsymbol{Z}, \left(\boldsymbol{Z}^{*(LD)},\ \boldsymbol{C}^{*(LD)}\right) = (\boldsymbol{Z}, 0)$. *Assume the solution* $\boldsymbol{C}^*$ *of* (26)-(27) *is non-zero. Then, using* $\overline{\boldsymbol{Z}}^*$ *in the LD detector, the resulting* $\boldsymbol{C}^{*(LD)}$ *satisfies that either* $\boldsymbol{C}^{*(LD)} = 0$, *or* $supp\left(\boldsymbol{C}^{*(LD)}\right) \nsubseteq supp(\boldsymbol{C}^*)$.

The efficacy of this attack model is verified by designing unobservable FDI attacks using the proposed optimization problem and verifying that they are not detected by the LD detector. The test systems used are the IEEE 24-bus reliability test system (RTS) and the IEEE 118-bus system. For testing purposes, we generated synthetic PMU-data over 5 seconds in each test system and modeled a sudden load increase to simulate a disturbance. The resulting synthetic measurements are illustrated in Fig. 3.7.

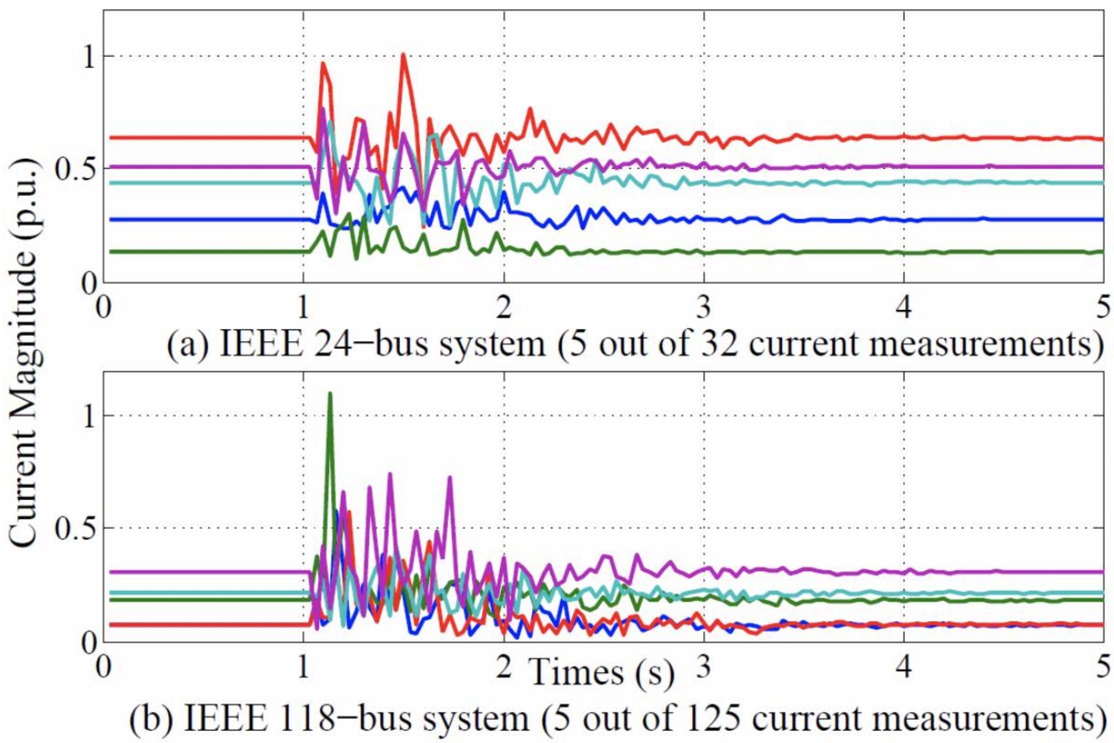
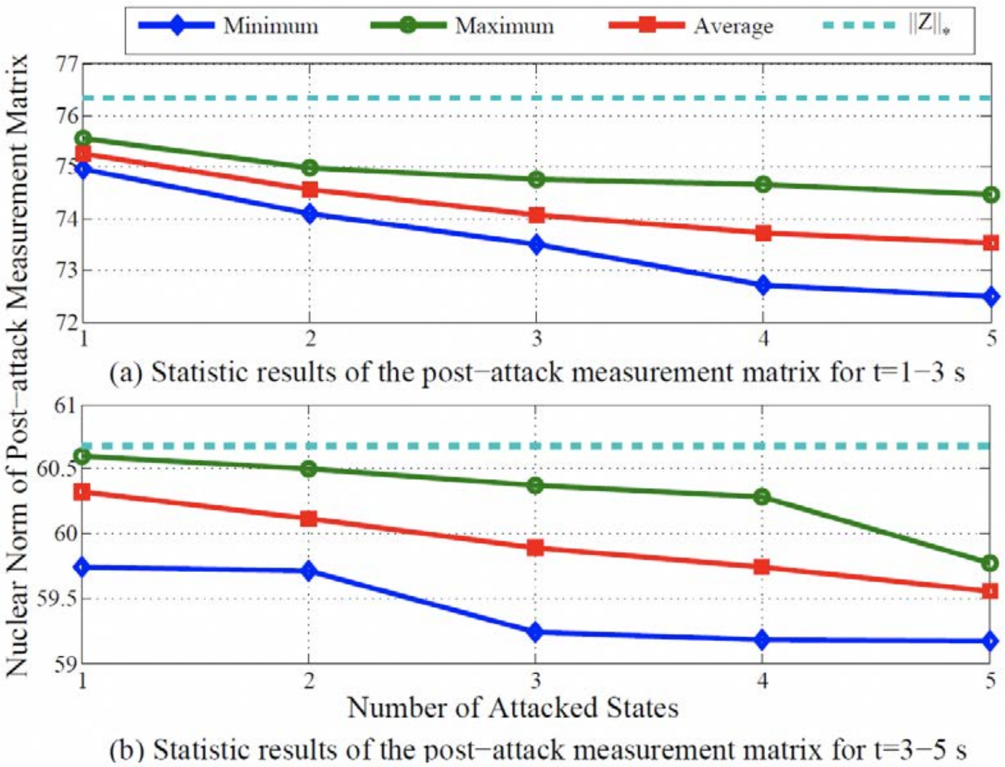


Fig. 3.7 Current magnitudes of synthetic PMU data

For every attack that we tested, the LD detector is completely bypassed; the statistic results are summarized in Fig. 3.8 and Table 3.1 for the IEEE 24-bus system. From these results, it can be seen that for every attack we tested, $\|\overline{\boldsymbol{Z}}^*\|_* \leq \|\boldsymbol{Z}\|_*$ always holds. Additionally, in Fig. 3.8 we also find that for the IEEE 24-bus system, $\|\overline{\boldsymbol{Z}}^*\|_*$ gradually decreases as the number of attacked states increases.

In conclusion, we showed that an intelligently designed attack can bypass the LD detector, if the attacker captures the temporal correlation of the measurement matrix $\boldsymbol{Z}$. Assuming an attacker can predict the measurements in a certain length of time, it can design an attack optimization problem that minimizes the rank of the post-attack measurement matrix, while fixing the column support of the attack matrix $\boldsymbol{C}$.



(a) Statistic results of the post−attack measurement matrix for t=1−3 s

(b) Statistic results of the post−attack measurement matrix for t=3−5 s

Fig. 3.8: Statistic results of $||\overline{\boldsymbol{Z}}^{*}||_{*}$ in the IEEE 24-bus system

Table 3.1: Statistic results of $||\overline{\boldsymbol{Z}}^{*}||_{*}$ in the IEEE 24-bus system

Time Period	$\|\bar{Z}^*\|$			$\|Z\|$
	Min	Max	Ave	
1–3 second	116.1	116.7	116.5	116.8
3–5 second	56.9	57.1	57.0	57.1

3.2.3 Rank preserving multiplicative attacks that can bypass the LD detector

The underlying assumption for LD detector to work is that an attack will violate the low-rank nature of the PMU measurement matrix $\boldsymbol{Z}$. Therefore, attacks that preserves the rank of $\boldsymbol{Z}$ can potentially avoid detection by the LD detector. To this end, we introduce a class of rank preserving multiplicative FDI attack. The false measurement matrix resulting from this class of attack is given by

$$\overline{\boldsymbol{Z}} = \boldsymbol{X}\boldsymbol{F}\boldsymbol{H}^{\boldsymbol{T}} = \boldsymbol{X}\boldsymbol{H}^{\boldsymbol{T}} + \boldsymbol{C}\boldsymbol{H}^{\boldsymbol{T}} \tag{28}$$

and the resulting additive attack matrix is given by

$$\boldsymbol{C} = \boldsymbol{X}(\boldsymbol{F} - \boldsymbol{I}) \tag{29}$$

Note that multiplicative attacks do not change the rank of $\boldsymbol{Z}$, but the resulting nuclear norm may change. Thus, theoretically LD detector is still possible to detect these attacks as it uses the nuclear norm as a proxy for rank.

We now illustrate the efficacy of the multiplicative FDI attacks by applying the LD detector on the false measurement matrix $\overline{\boldsymbol{Z}}$. We assume that the LD detector selects 5 seconds worth of PMU measurement data, while the attacker continuously injects bad data. λ is chosen to be 1.05 in the LD detector. The test system is the IEEE RTS-24-bus system. An optimal PMU placement problem as introduced in [49] is solved to ensure the system is fully observable with PMUs. The details of the PMU placement scheme and available measurements are illustrated in Fig. 3.9. Buses in red are buses with PMUs. We generate synthetic PMU data over 5 seconds in the test system. A base case of the system operating status is obtained by solving an AC optimal power flow problem. To model realistic data with a disturbance, at the first time instant t after 1 second, we change the load at each bus by adding a random value d to the base load. We then solve an AC power flow to obtain the measured phasors of bus voltage and branch current as measurements at time instant t. The singular values for the synthetic measurement matrices are illustrated in Fig. 3.10. It can be seen that these synthetic measurements have the same low-rank property as the actual PMU data as illustrated in [47]. Furthermore, we assume noiseless measurements.

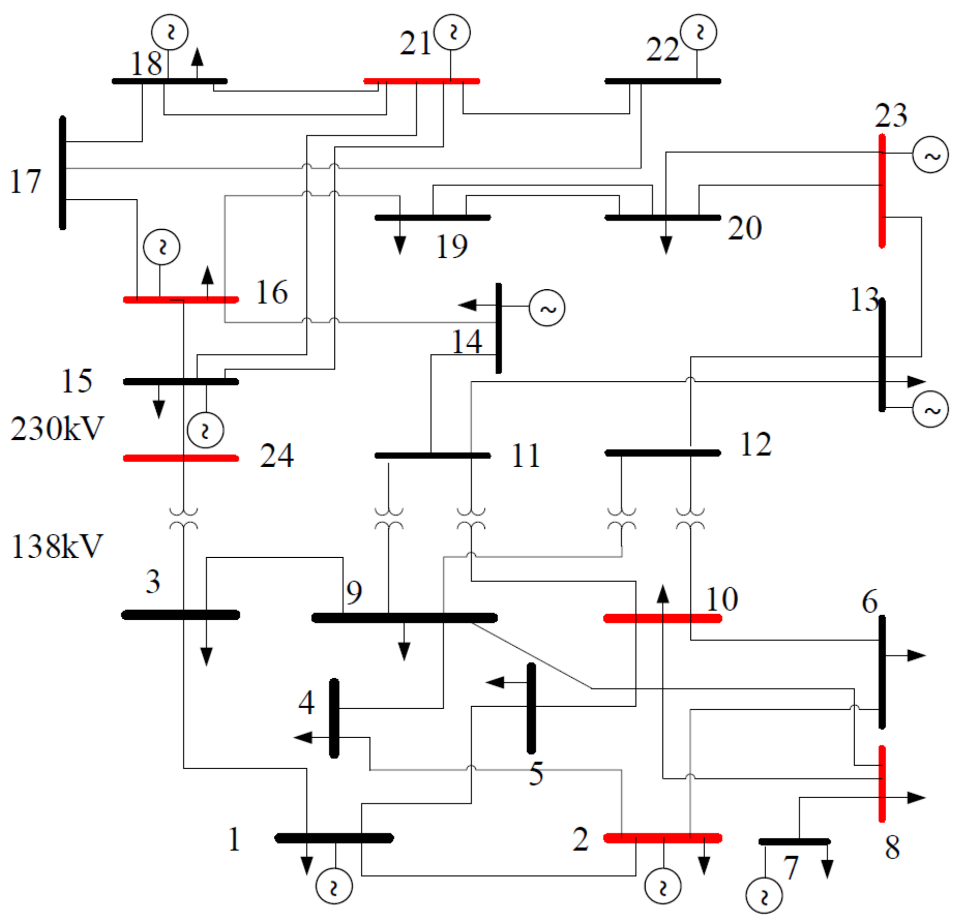


Fig. 3.9: PMU placement scheme in IEEE RTS 24-bus system

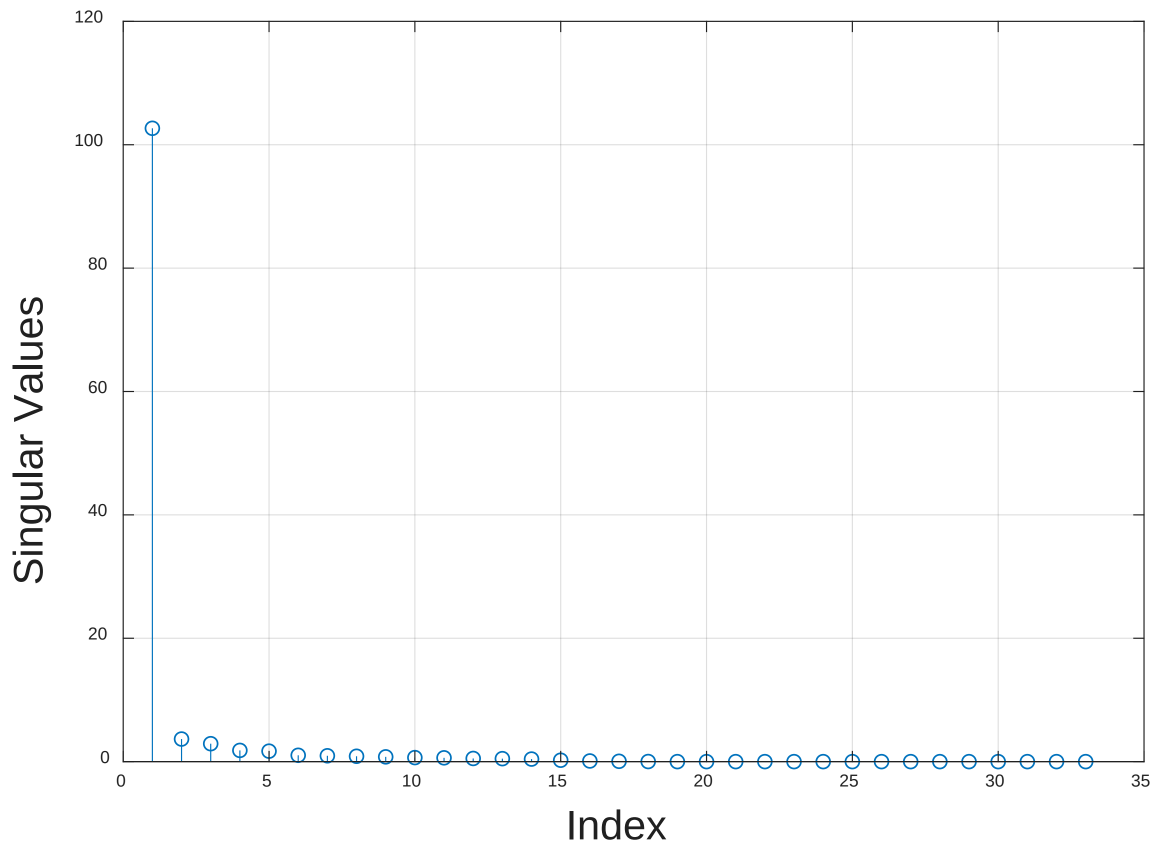


Fig. 3.10: Singular values of the synthetic PMU data matrix in decreasing order

We focus on unobservable 1-sparse multiplicative attacks and illustrate our results using two specific cases. One case illustrates that no attacks are detected, and the other case illustrates that attacks are detected at incorrect buses. To change the state at bus b, the entries of the attack matrix $\boldsymbol{F}$ are set to be,

$$F_{ij} = \begin{cases} 1, & i = j \neq b \\ c, & i = j = b \\ 0, & i \neq j \end{cases} \tag{30}$$

Clearly, $\boldsymbol{F}$ is a diagonal matrix, and hence has full rank.

Fig. 3.11(a) illustrates the detection result when there is no attack. Fig. 3.11(b) illustrates the attack detection result for an attack on bus 4. $\boldsymbol{F}$ is constructed with $b = 4$ and $c = e^{j0.2}$. Compare these two subfigures, we conclude that the LD detector fails to detect the attack at bus 4. Fig. 3.11(c) illustrates the attack detection result for an attack on bus 16. $\boldsymbol{F}$ is constructed with $b = 16$ and $c = e^{j0.3}$. The LD detector incorrectly detects that buses 18, 21, and 22 are under attack.

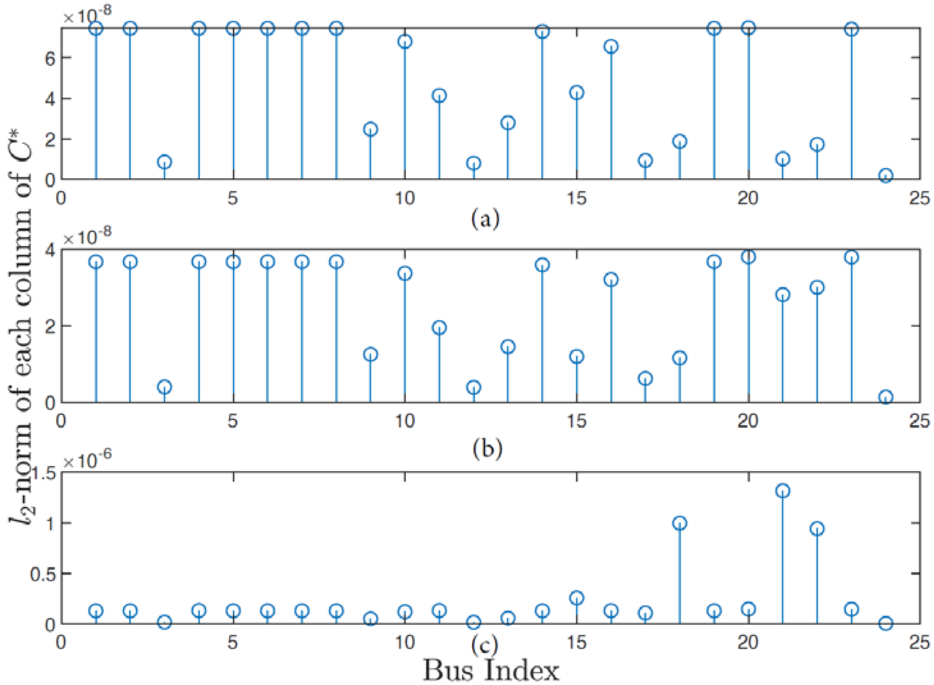


Fig. 3.11: Normalized l_2-norm of each column of $\hat{C}^*$ under (a) no attack; (b) attack at bus 4; (c) attack at bus 16

A comparison of the $l_{1,2}$-norm of $\widehat{\boldsymbol{C}}^*$ with no attack and with the attack at bus 4 for $\lambda = [1.05, 1.5]$ is illustrated in Fig. 3.12. It can be seen that the $||\widehat{\boldsymbol{C}}^*||_{1,2}$ with and without this attack are very similar. Intuitively, this attack cannot be detected by the LD detector, and our result in Fig. 3.11(b) supports such intuition.

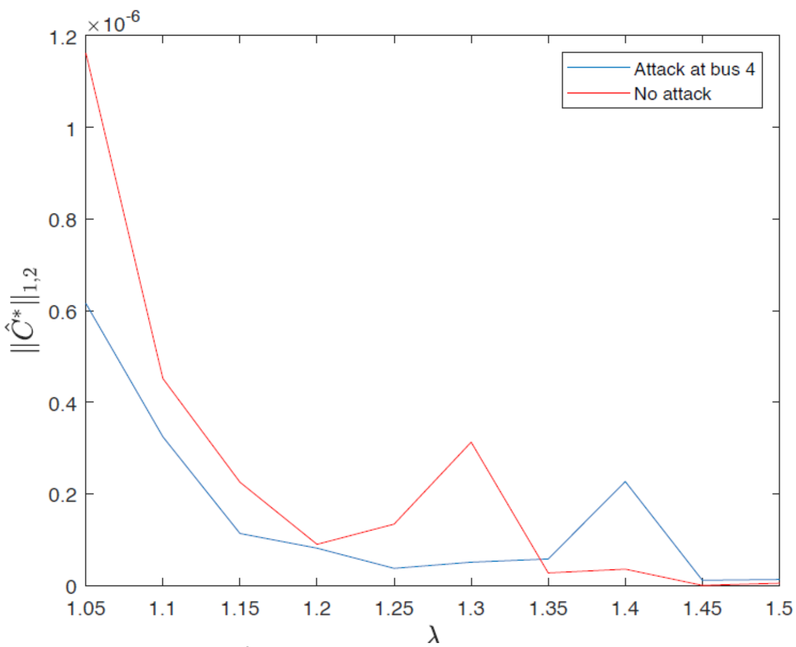


Fig. 3.12: $l_{1,2}$-norm of $\hat{C}^*$ under no attack and attack at bus 4 for different λ

3.2.4 Predictive filters to capture temporal correlation of the PMU measurements

As shown in the previous section, FDI attacks can be created to be unobservable to the residual based BDD. However, given the high sampling rate of PMUs, one would expect the PMU measurements to be highly correlated in time, because the power system is unlikely to have dramatic changes in such short time period under normal operating conditions. Therefore, a detector that takes into consideration the temporal correlations of the PMU measurements may be able to detect such "unobservable" FDI attacks. One way to do this is to have a predictive filter that accurately predicts the PMU measurements. This filter flags anomaly if the difference between measured value and predicted value is larger than a threshold, which is often the case when an FDI attack is launched. The authors of [23] investigated the temporal correlation in the PMU data to find the relationship between consecutive measurements. They proved that for loads changing at a constant power factor, the real and imaginary components of the voltage phasor follow a quadratic trajectory. Under this condition, voltage at a future step can be predicted using the present and the past states as shown in (31).

$$V'_{x,y}(n+1) = 3V_{x,y}(n) - 3V_{x,y}(n-1) + V_{x,y}(n-2) \tag{31}$$

In (31), $V_{x,y}$ denotes the actual value of complex voltage, $V'_{x,y}$ denotes the predicted value of the complex voltage at the future time instant $(n+1)$. The present and past states are denoted by $n - i, \forall i \in \{0,1,2\}$. Since the future state is predicted by utilizing the knowledge of the three prior states, the algorithm was named "three-sample quadratic prediction algorithm (TSQPA)". Considering that $V'_{x,y}(n+1)$ and $V_{x,y}(n+1)$ are the predicted and actual value of the voltage phasor at the $(n+1)^{th}$ time instant, an observation residual $(R_{x,y})$ in real and imaginary components can be obtained as follows:

$$R_{x,y}(n+1) = V'_{x,y}(n+1) - V_{x,y}(n+1) \tag{32}$$

A change in the observation residual R would be considered an anomaly due to a cyber-attack. This prediction technique is actually a third order FIR filter and can be used as an anomaly detector to detect unobservable FDI attacks that are suddenly injected in PMU measurements.

Data-driven five-sample predictive (FSP) filter: Based on the real PMU measurements, a moving window linear regression is performed to learn the best coefficients of a five-sample predictive filter. This predictive filter is given by

$$\begin{aligned} V'_{x,y}(n+1) = {} & 0.9186V_{x,y}(n) + 0.0196V_{x,y}(n-1) + 0.0438V_{x,y}(n-2) \\ & +0.0058V_{x,y}(n-3) + 0.0122V_{x,y}(n-4) \end{aligned} \tag{33}$$

3.2.5 Gradually ramping unobservable FDI attacks

Sudden attacks may be easily detected by predictive filters because the attack magnitude often has to be sufficiently large, in order to cause severe consequences on the system. Thus, to avoid detection by predictive filters, the attacker may gradually increase the attack magnitude, so that at each time step, the increase in PMU measurements is sufficiently small. Consequently, the differences between the predicted measurements and the attacked measurements are also small.

Assume that the power system performs DC optimal power flow (OPF) every five minutes, using the states estimated by PMU-based linear state estimation (LSE) at that time. A malicious attacker can inject intelligently designed unobservable FDI attacks into the system, to spoof the LSE to estimate fake states, and subsequently, fake loads. The generation re-dispatch (which might occur in 5 minutes intervals) based on the fake loads can cause overflows in the physical system. The attacker can either inject the false measurements at $t = 5$ minutes, or gradually increase the attack magnitude from $t = 0$ to $t = 5$ minutes, so that the false measurements at $t = 5$ minutes are the same under these two scenarios.

The aim of the attacker is to cause physical overflows in the system. A bi-level optimization problem can be formulated to maximize the physical power flow on a target line, wherein the first level models the attacker's capabilities and limitations, while the second level models the system response to the attack via DC OPF [50]. The output of this attack optimization problem is the state attack vector $\boldsymbol{c}$. To create false measurements, it is assumed that the attacker has control of PMUs in a subgraph of the whole system, which can be constructed as described in [51]. The attacker first estimates the states in the subgraph and adds the attack vector $\boldsymbol{c}$ to the voltage angles. Then, to ensure the estimated loads are zero at non-load buses, the attacker solves for the false states in the subgraph via Newton-Raphson method. Finally, the attacker computes the measurement attack vector $\boldsymbol{a}$ in the subgraph based-on the false states and adds $\boldsymbol{a}$ to the true measurements.

In a sudden attack, the attacker injects the attack vector $\boldsymbol{a}$ at any time instance between $t = 0$ and $t = 5$ minutes. Such an attack may be easily detected by the aforementioned predictive filter, as there will be a big change between the two measurement values, so that the difference between predicted value and measurement value is large. Alternatively, the attacker may gradually ramp up the attack magnitude to avoid detection by predictive filters. Assume that at $t = 0$, the attacker computes $\boldsymbol{a}$, and it keeps increasing the measurement attack vector injected at each sample to reach the desired attack magnitude at $t = 5$ minutes (sample number 9,000). Mathematically, the measurement attack vector at the PMU time sample n is given by $a_n = \frac{n}{9000} \times a, n = 1,2,3, \ldots, 9000$. Since the change caused by the attack between two consecutive samples are small, this attack can avoid detection by a predictive filter.

3.2.6 Attack detection using predictive filters

In this research, two types of cyber-attacks were designed: "sudden attacks" and "ramping attacks". In the "sudden attack" situation, the voltage measurements were changed suddenly at the fifth minute. On the other hand, in the "ramping attack" case, the PMU measurements were changed gradually over a period of time. In this section, we investigate if such type of attacks can be detected by predictive filters, which investigates the temporal correlation of the PMU data to identify an anomaly.

False measurements resulting from sudden and ramping attack, as well as attack-free measurements at two buses of the IEEE 118-bus system are illustrated in Fig. 3.13. It can be seen that the measurements of both attack strategies were identical after 5 minutes (9,000 samples). Fig. 3.13(a) shows a relatively large attack, where the attack magnitude on the real part of the voltage at bus 8 at the fifth minute was 0.0141 per unit, while Fig. 3.13(b) shows a small attack at bus 40 where the attack magnitude to the real part of the voltage was merely 0.0017 per unit.

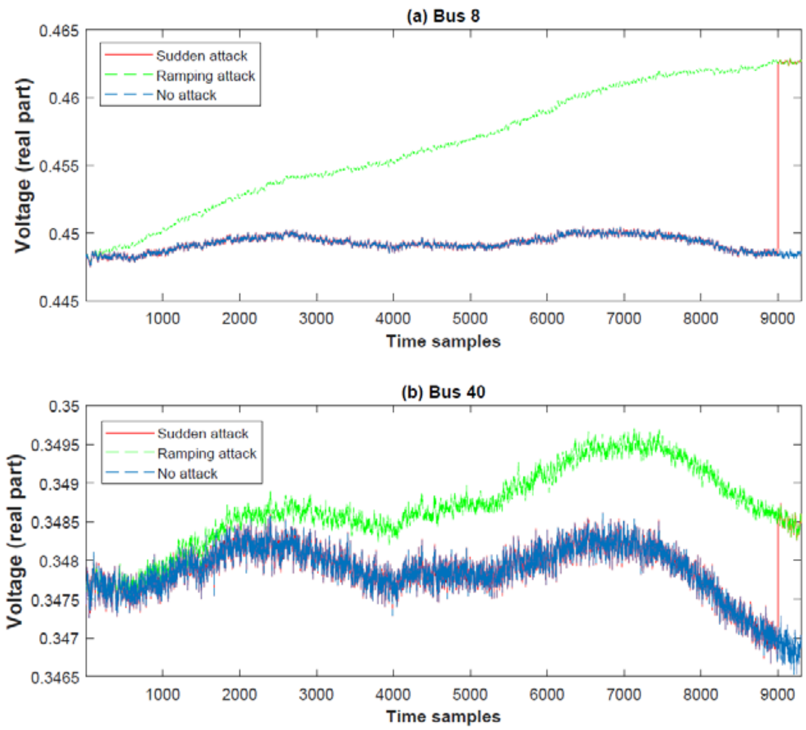


Fig. 3.13: Examples of false measurements at (a) bus 8, and (b) bus 40

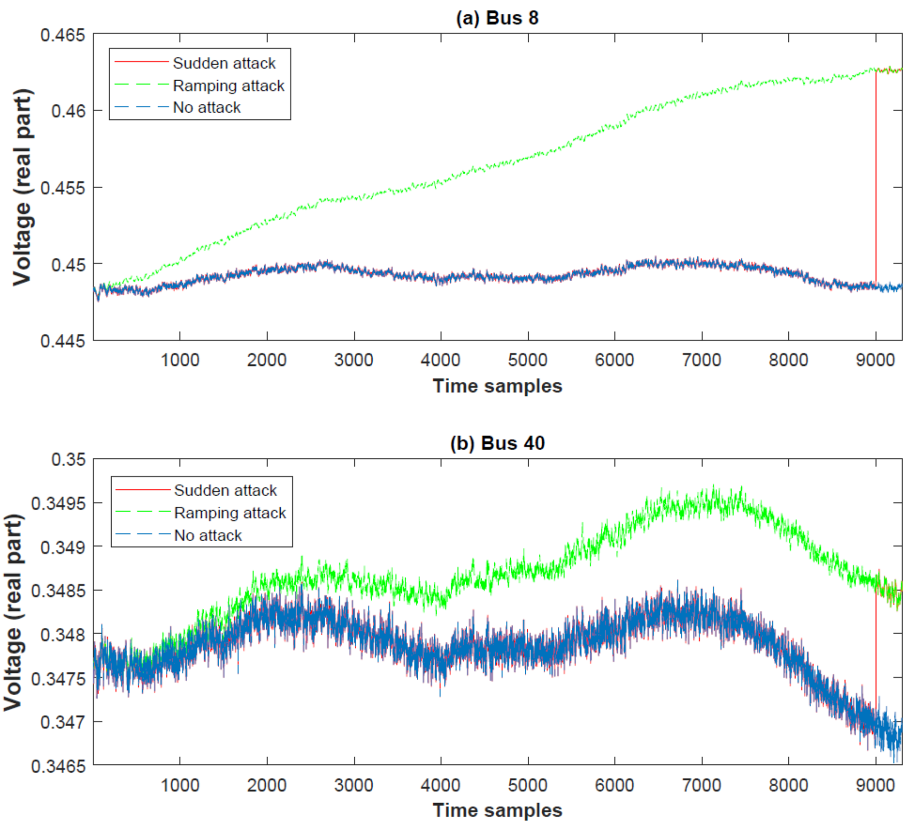


Fig. 3.14: Examples of false measurements at (a) bus 8, and (b) bus 40

Fig. 3.14 demonstrates the observation residues when applying the predictive filters on measurements with sudden attack. Both TSQPA and FSP gave a large residue at the fifth minute when the attack is injected, indicating that they were both able to detect sudden attacks. Moreover, they could detect both the attacks at bus 8 and at bus 40, even though the attack magnitude at bus 40 was much smaller. Fig. 3.15 illustrates the observation residues obtained by applying predictive filters on measurements with ramping attack. The residues did not increase because the attack magnitude at each time instant was too small. These observations indicate that gradually ramping attacks can avoid detection by the selected predictive filters. Fig. 3.16 illustrates the observation residues obtained by applying predictive filters on measurements with ramping attack. The residues did not increase because the attack magnitude at each time instant was too small. These observations indicate that gradually ramping attacks can avoid detection by the selected predictive filter.

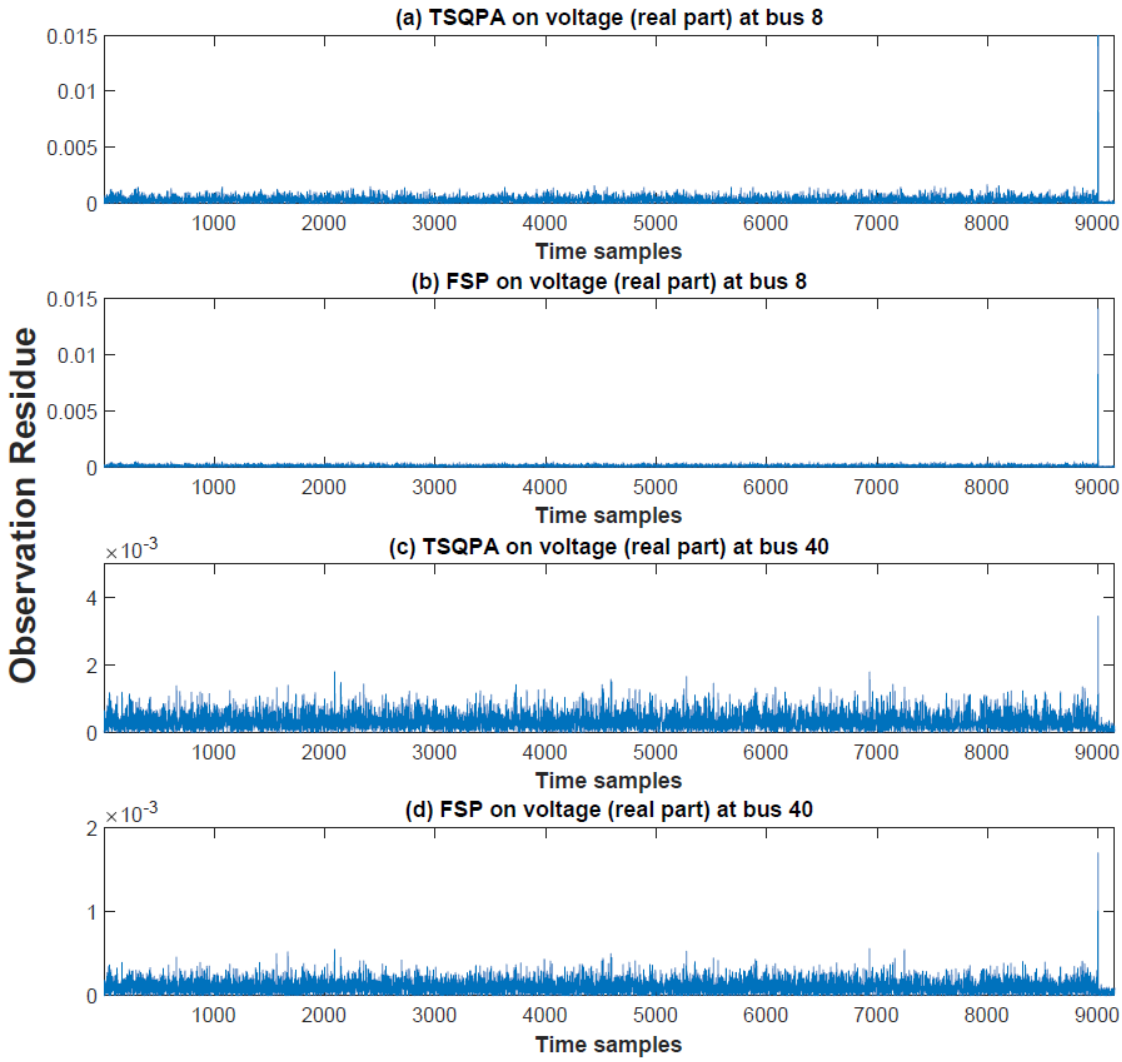


Fig. 3.15: Sudden attack detected by predictive filters

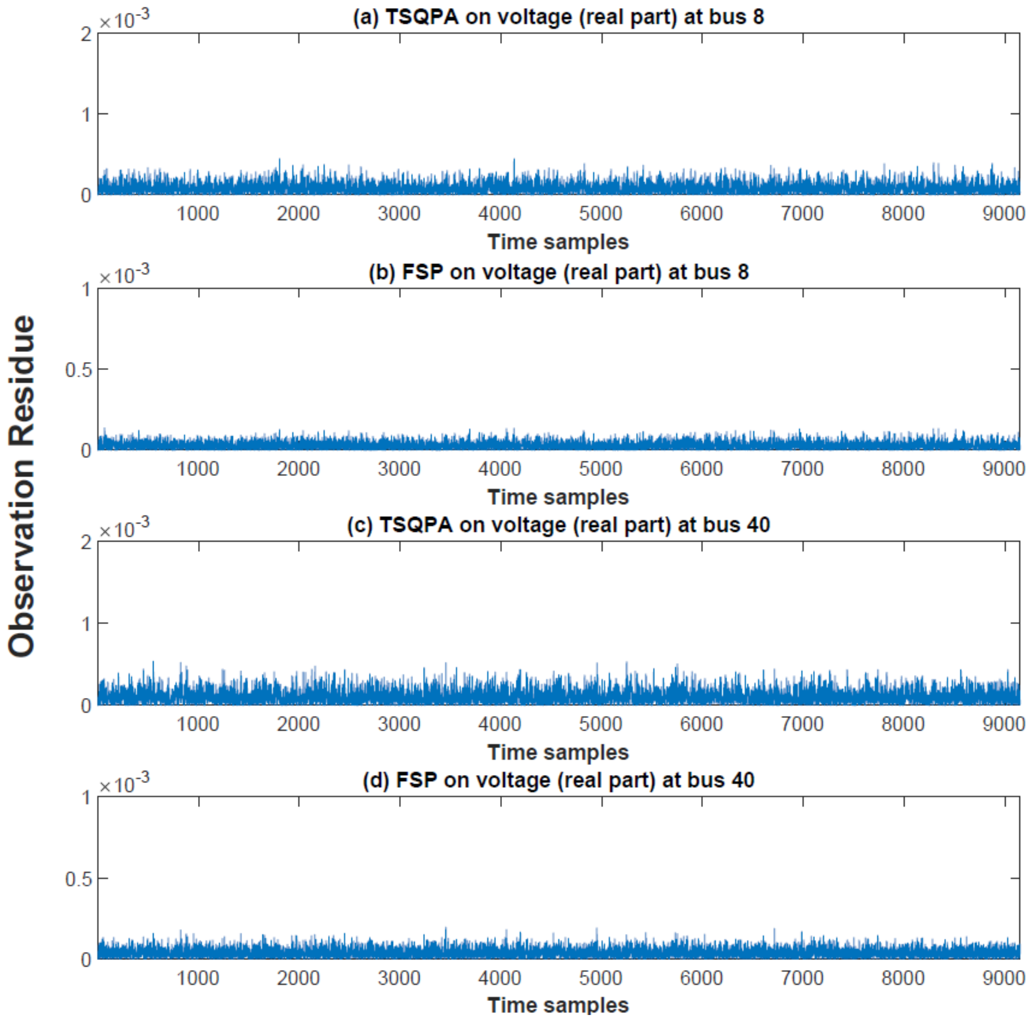


Fig. 3.16: Ramping attack undetected by predictive filters

4. Task 3: PMU based power system control application

4.1 Background: PMU based voltage stability assessment for stochastic systems

The power flow equations define the instantaneous, equilibrium operating condition for a synchronous electric power grid, in terms of sinusoidal bus (node) voltages in phasor form. Derivation of these equations begins from basic, linear KCL current balance constraints imposed at each bus. However, they become nonlinear when current conservation is modified to power conservation, and generation and load demands are modeled as fixed power injections into or withdrawals from the network. Looking at the evolution of the equilibrium operating point, time varying changes in power demand and generator injections drive the system; the bus voltages evolve in response to these changes. Other electrical quantities such as branch currents are then simple linear functions of the bus voltages. Provided the quasi-static, near-equilibrium assumption remains valid, the power system can be viewed abstractly as a power flow solution engine, taking power injection variations as inputs, and "computing" bus voltage phasors as outputs.

In its standard formulation, the power flow is a square mapping, with an equal number of input arguments (the power injections) and output results (in polar form, the magnitudes and angles of bus voltage phasors). A key assumption in our formulation is a separation of time scales in the driving inputs to this system. We view the load power variation at any bus as being decomposable into a sum of two parts. The first component represents the slowly varying bulk consumption, evolving on a time scale of 10's of minutes to hours, and typically displaying nearly periodic behavior on the daily 24-hour cycle of human behavior. This relatively slowly varying component is dominant and has long been the focus of utility studies of time behavior of load. Indeed, in most utility operations, this slower evolution of load is very accurately predicted hours or days in advance of real-time operations and is referred to as the "load cycle." However, both first principles and more detailed electrical measurements suggest that load demand must inevitably display faster time scale behavior also. Load demand at a bulk distribution bus aggregates the individual behavior of hundreds of thousands of individual power consuming devices, most of which display individual on-off behavior governed by human users or by control systems responding to very local environments. As a result, one may expect the fast behavior of load (time scale seconds or less) to display small magnitude random jump behavior, largely uncorrelated between locations, and smoothed by filtering inherently present in the electrical characteristics of transformers and other equipment in the distribution system. This random component of load is typically small in magnitude, with variance at a given bus no more than a few percent of the total bulk load at that location. However, a key premise in this project's work emerges from this viewpoint: *that the vector of driving inputs contains both a large signal component, that slowly move the operating point, and small-signal, randomly varying components, that persistently excite the system about its operating point*.

As is traditional in many branches of circuit analysis, this split of large-signal component and small signal component suggests the usefulness of linearized approximations. The nominal operating point is set by the large signal component, and the impact of the small signal component is analyzed via linearization about this (slowly varying) operating point. The power flow conditioning to be evaluated can then be interpreted as the conditioning of the linearization about the operating point, and this will in turn be slowly varying as the operating point evolves in time. In this context, the linearization of interest will be represented by the familiar Jacobian of the power flow. To be precise, the proposed approach here will not seek to estimate the condition number of the power flow Jacobian, but rather a measure of its nearness to singularity, as reflected in the largest singular value of its inverse.

Goal of the proposed method will be to consider the impact of such random load variations as driving terms in power balance equations for the electric grid. The "forward" power flow equations, linearized about an operating point can be written as shown in (34). Viewing the physical power system as a power flow solver, it is useful to invert this forward form, treating loads and power injections as inputs, and the output response being phasor angles and voltage magnitudes (as measured by PMUs). Rearranging (34) leads us to (35), which provides the desired input-output relationship. If the smallest singular value approaches zero, small variations in power have the potential to yield large response in bus voltage magnitude and angle variations; as noted earlier, such high sensitivity behavior is recognized as a precursor to voltage instability problems. Indeed, the smallest singular value of Jacobian matrix has been specifically proposed as an index of vulnerability voltage collapse [4], [5], [52], [53], [54], [55], [56].

$$\begin{bmatrix} \Delta \boldsymbol{P} \\ \Delta \boldsymbol{Q} \end{bmatrix} = \begin{bmatrix} \dfrac{\boldsymbol{\partial P}(\boldsymbol{\delta}, |\boldsymbol{V}|)}{\boldsymbol{\partial \delta}} & \dfrac{\boldsymbol{\partial P}(\boldsymbol{\delta}, |\boldsymbol{V}|)}{\boldsymbol{\partial} |\boldsymbol{V}|} \\ \dfrac{\boldsymbol{\partial Q}(\boldsymbol{\delta}, |\boldsymbol{V}|)}{\boldsymbol{\partial \delta}} & \dfrac{\boldsymbol{\partial Q}(\boldsymbol{\delta}, |\boldsymbol{V}|)}{\boldsymbol{\partial} |\boldsymbol{V}|} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Delta \delta} \\ \Delta |\boldsymbol{V}| \end{bmatrix} \tag{34}$$

$$\begin{bmatrix} \boldsymbol{\Delta \delta} \\ \Delta |\boldsymbol{V}| \end{bmatrix} = \boldsymbol{J^{-1}} \begin{bmatrix} \boldsymbol{\Delta P} \\ \boldsymbol{\Delta Q} \end{bmatrix} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \frac{\boldsymbol{u_i^T v_i}}{\boldsymbol{s_{ii}}} \begin{bmatrix} \boldsymbol{\Delta P} \\ \boldsymbol{\Delta Q} \end{bmatrix} \tag{35}$$

Traditional methods of power flow analysis often seek to exploit approximate decoupling between angle-active power, versus voltage magnitude-reactive power. Such decoupling approximation might suggest that with a goal of considering voltage stability, one might focus exclusively on the block of the Jacobian that characterizes sensitivity of reactive power to voltage magnitude. However, such a decoupling approximation has not been considered here, in part because systems under highly stressed operating conditions, for which risk of voltage instability may be most severe, are precisely those for which decoupling approximations may be least accurate.

4.2 Model-free estimation of the power flow Jacobian's smallest singular value

In this project, a "model-free" method was considered, and practical aspects that impact the method were considered, such as base state selection, PMU data scaling, time window, and measurement noise removal and filtering using different algorithms. The fundamental concept of the proposed approach is to identify ill-conditioned operating conditions, which may serve as an indicator of vulnerability to voltage collapse. Since proposed algorithm seeks a model-free analysis, the available information will primarily be limited to voltage magnitudes and phase angles as could be measured via PMUs from the buses in the power system. Practical PMU measurement sets may also include bus power injections and demands, which can possibly be included. Under the simplified assumption of measurements, the goal of the proposed algorithm is simple: *Estimate the major axis of the "ellipse" formed by the set of measurements and track how this quantity evolves in response to variations in grid operating conditions or network topology changes*. The algorithm employed to estimate the major axis is also quite simple, and is closely analogous to the use of SVD tools in other streaming data applications [56]: *after subtracting a base state from the measured PMU data, and possible filtering/bad-data-correction, one constructs a sliding windowed array of the streaming data, up to the most recent measurement*. For this array, one computes the largest (or several largest) singular value(s) and left singular vector(s) and assess voltage stability of the system and detect topology change in the system with this information. Investigation of base state, PMU data scaling, time window and computation efficiency are essential factors to make the proposed method feasible to our application.

4.3 Jacobian conditioning and voltage stability assessment via PMU data

The concepts described above suggest desirable features in an algorithm to identify ill-conditioned operating conditions, which in turn may serve as an indicator of vulnerability to voltage collapse. First, because we seek a "model free" analysis, we will limit the information available to the algorithm to be that of the output measurements only; i.e., the measured bus voltage phasor magnitudes and angles. In practice, this is likely restrictive, because practical PMU measurement sets may include some bus power injections and demands (inputs) also. These are not considered in the project work here. With the simplifying assumption of output measurements only, our conceptual goal is straightforward: *estimate the length of the major axis of the "output" quantities of the inverse power flow (i.e., the voltage magnitudes and angles, in response to the input of active and reactive load variation), and track how this quantity changes in time*. However, a number of implementation questions follow: (i) what is the "base case," nominal point about which variation is being measured? (ii) how many sample output points are necessary to form an accurate estimate of the ellipse's major axis? (iii) given a desired number/density of measurement points, how are they sampled in time? (iv) if one has the flexibility to select different possible measurement locations in the network, how should the measurement points be selected in "space" (i.e., which bus locations)?

The algorithm developed here makes practical choices in regard to the questions above. First, with regard to selection of a base case about which variation is measured, our choice is dictated in part by the synthetic load cycle constructed for study. For each of the standard test power systems to follow, publicly available datasets specify a single vector of generation and load injections. Our approach was to treat such values as one "snapshot" in time along a 24-hour interval, and to synthesize for every bus a plausible 24-hour curve of power injection/load behavior. Clearly, the family of load curves that might be judged "plausible" is very large, and our selection reflects a highly subjective judgment as to what represented an interesting study case. However, once this vector valued time function (over the 24-hour interval) is selected, our choice of base case was quite simple: *it was simply the power flow solution associated with the time point at which total load demand was minimum for power flow conditioning estimation, while it was the power flow solution with the predicted load demands at a given time point for topology change detection*.

In regard to questions (ii) and (iii) above, the choices made for the algorithm here are also influenced by specifics of the synthetic datasets generated to test the algorithm. Because of our focus on the evolution of the system toward operating conditions in which it may be vulnerable to voltage instability, rather than the final "collapse" of system with system state rapidly diverging from acceptable operation, the time scale of interest starts at seconds, rather than the 30 or 60 Hz sampling rate of the PMUs. Therefore, for the synthetic computational studies to follow, we considered a number of scenarios in which measurements are down sampled to once per 10 seconds (i.e., a $1/10^{th}$ Hz rate). The output array of interest is then constructed as a sliding window, with each column representing a vector of reported measurements, running backward from the present (or from the "clock time" in the simulation), over an 8 minute and 40 second window (8 minutes and 40 seconds representing 52 samples at 10 second intervals). Each one of these measurement vectors represents a set of deviations in bus voltage phase angles (in radians) and voltage magnitude (in per unit). As described above, the deviations are taken relative to a nominal set of angles and voltage magnitudes associated with a low load operating point for power flow conditioning estimation and predicted operating point for topology detection. Reflected in choice of the length of time interval, and correspondingly the number of sample measurements, are two competing requirements. From the geometric standpoint, one might like a very large number of measurements, to densely "fill" the output space ellipse whose major axis we are trying to estimate. However, weighing against too large a number of samples is the issue of time scale separation. One wants the time window interval to be sufficiently short, relative to the underlying 24-hour, "slow" variation of average load. Different samples (columns) within the window should reflect only random variation about an operating point, and hence, that operating point should remain nearly constant throughout the window interval. Finally, the point (iv) above represents, in practical terms, the selection of substation locations from which measurements are collected, in order to construct the sliding window data array described above. In the illustrative test cases to follow, we make a range of possible selections, from full measurement availability at every bus (unlikely in present-day implementations), to a more realistic selections of subsets representing much smaller percentage penetration of measurements.

The proposed time varying, scalar measure of power flow conditioning is then obtained in a very simple fashion: *it is computed as the largest singular value of the windowed array*. For each step forward in time, the array is updated with one new column of measurements, the oldest column discarded, and the new largest singular value computed. Note that this computation was implemented by a power-method-like Lanzcos algorithm [57], [58], [59], [60]. Further computational efficiency may be gained by exploiting the rank one nature of the update at each time-step. The figure below shows the graphical description of sliding windowed PMU data matrix. The quantities $\Delta\delta$ and $\Delta|V|$ denote the voltage phase angles and magnitudes after subtracting the base state and the subscript k correspond to the k^{th} time point. The largest singular value of the PMU data array with time points from 1 to k^{th} instant is proposed as a means to estimate the largest singular value of the inverse power flow Jacobian corresponding to the operating point at the k^{th} time instant, and is therefore, also a measure of vulnerability to voltage instability at that time point. A premise of this work is that the time series of the largest singular values from PMU data matrix should be closely related to the time series of the largest singular value of the power flow Jacobian inverse matrix, as loads and injections change in time. As mentioned previously, the smallest singular value of the power flow Jacobian has been utilized as an indicator of vulnerability of a system to voltage collapse. If the smallest singular value of the power flow Jacobian can be reliably estimated from the PMU data matrix, this provides a real-time indicator of system stress without the need of network parameter values or state estimator results.

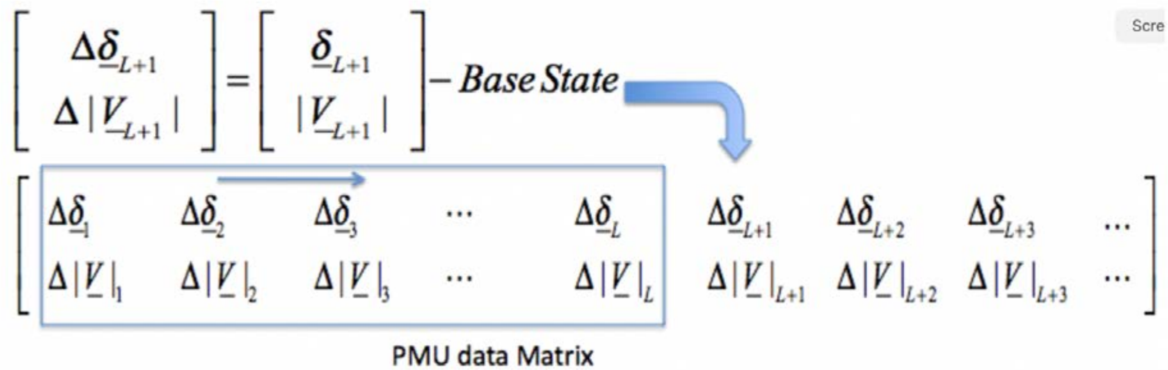


Fig. 4.1: Construction of PMU Data Matrix (from which singular values of interest are computed)

To verify the hypothesized relationship of largest singular value of the analytic inverse Jacobian matrix (as might be computed in off-line studies) and that of PMU data matrix, a simple least square error is used. In test cases to follow, for which system data is available to compute the operating point and the power flow Jacobian, its inverse can be computed with full knowledge of power system model, load demands, and power injections. One can then compute the classic voltage stability index of smallest singular value of the power flow Jacobian directly. This can then be benchmarked against the measurement-based estimate computed "model free," from the PMU data matrix. For this comparison, system states at each bus are also computed with the full knowledge of power system model, load demands, and power injections and the computed states can be used as synthetic PMU measurements. The new voltage stability index, the largest singular value of PMU data array, computed from only the synthetic PMU measurements, can then be compared to the voltage stability index from the analytically computed power flow Jacobian inverse.

PMU data matrix should be closely related to the time series of the largest singular value of the power flow Jacobian inverse matrix, as loads and injections change in time. As mentioned previously, the smallest singular value of the power flow Jacobian has been utilized as an indicator of vulnerability of a system to voltage collapse. If the smallest singular value of the power flow Jacobian can be reliably estimated from the PMU data matrix, this provides a real-time indicator of system stress without the need of network parameter values or state estimator results.

To quantify the relationship of largest singular value of the analytic inverse Jacobian matrix (as might be computed in off-line studies) and that of PMU data matrix, a simple least square error LSE over an affine fit is used. In test cases for which system data is available to compute the operating point and the power flow Jacobian, its inverse can be computed with full knowledge of power system model, load demands, and power injections. The Jacobian matrix is able to provide voltage stability index as discussed previously. We then can benchmark the largest singular value of power flow Jacobian inverse and singular vectors associated with the largest singular values as the accurate voltage stability index under the full knowledge of power system model, load demands, and power injections. For this comparison, system states at each bus are also computed with the full knowledge of power system model, load demands, and power injections and the computed states can be used as "pseudo-PMU measurements". The new voltage stability index, which is the largest singular value of PMU data array, will be compared with the voltage stability index from the analytically computed power flow Jacobian inverse.

Fig. 4.2 and Fig. 4.3 below provide this benchmark comparison for numerical computations performed using data from the IEEE 14-bus and IEEE 300-bus test systems. To understand the plots provided, it is first important to understand that a synthetic 24-hour load curve is considered for each test system, with load and operating point updated every 10 seconds across the 24-hour study period (hence the horizontal time axis represents 8,640 operating points). For IEEE 14-bus system, system loads were varied from the operating point provided as the standard base case, using a uniform scaling to create a peak load 1.6 times the base. Generation is likewise uniformly scaled to follow load, with a distributed slack allocating increase in system losses across multiple generators. For the IEEE 300-bus system, where the base case load was judged to represent a relatively more heavily loaded condition, the synthetic load curve uniformly varied system loads and generation over a range from a minimum of 0.25 times base-case, to a peak of 1.05 times the base case. We wish to stress our judgment that the exact degree of "realism" in the synthetic load curve is not critical in this test, as long as the system is exercised across a significant range of system "stress" at each operating point. The "true" largest singular value of the power flow Jacobian inverse is computed next. The measurement-based voltage stability metric, which is the largest singular values computed from the synthetic PMU data matrix, is allowed two degrees of freedom for data fit. In particular, we use a small number of training points, out of the 8,640 points of the day, the reader may think of it as a short training period shortly after 12 am to identify two parameters. These are a dc-offset, and a normalization/scaling factor, between the true Jacobian based largest singular value, and the measurement based largest singular value. The graphs illustrate the two quantities over the 8,640 operating points of the 24-hour day. The "full system information" Jacobian-based singular values, and the PMU measurement-based singular values behave nearly identically in the 14-bus test system (see Fig. 4.2) and show very good agreement in the 300-bus test system (see Fig. 4.3).

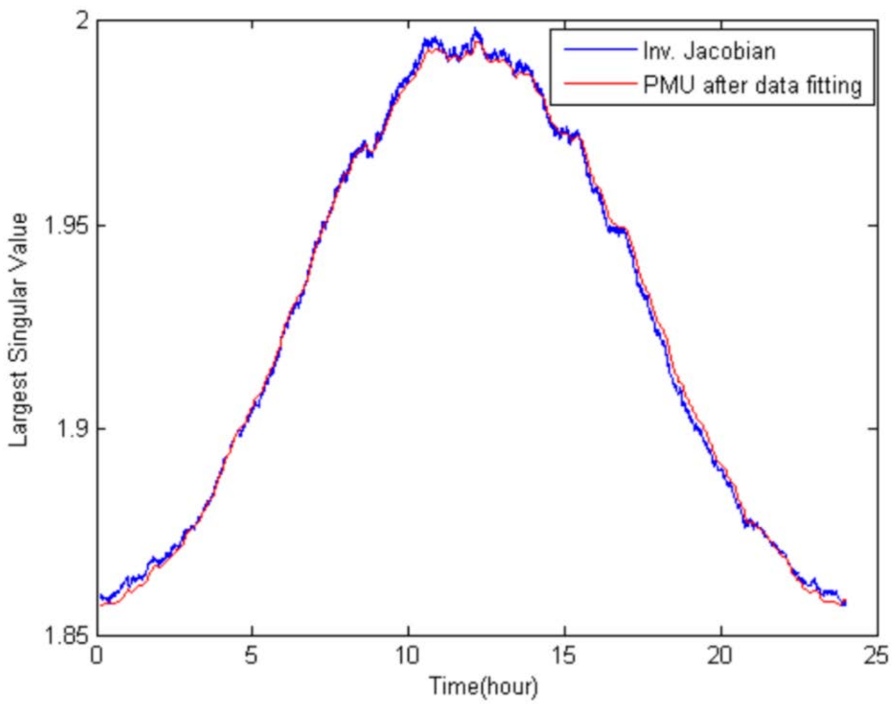


Fig. 4.2: Quality of fit between Inverse Jacobian-based largest singular value versus PMU measurement-based largest singular value, IEEE 14-bus test case

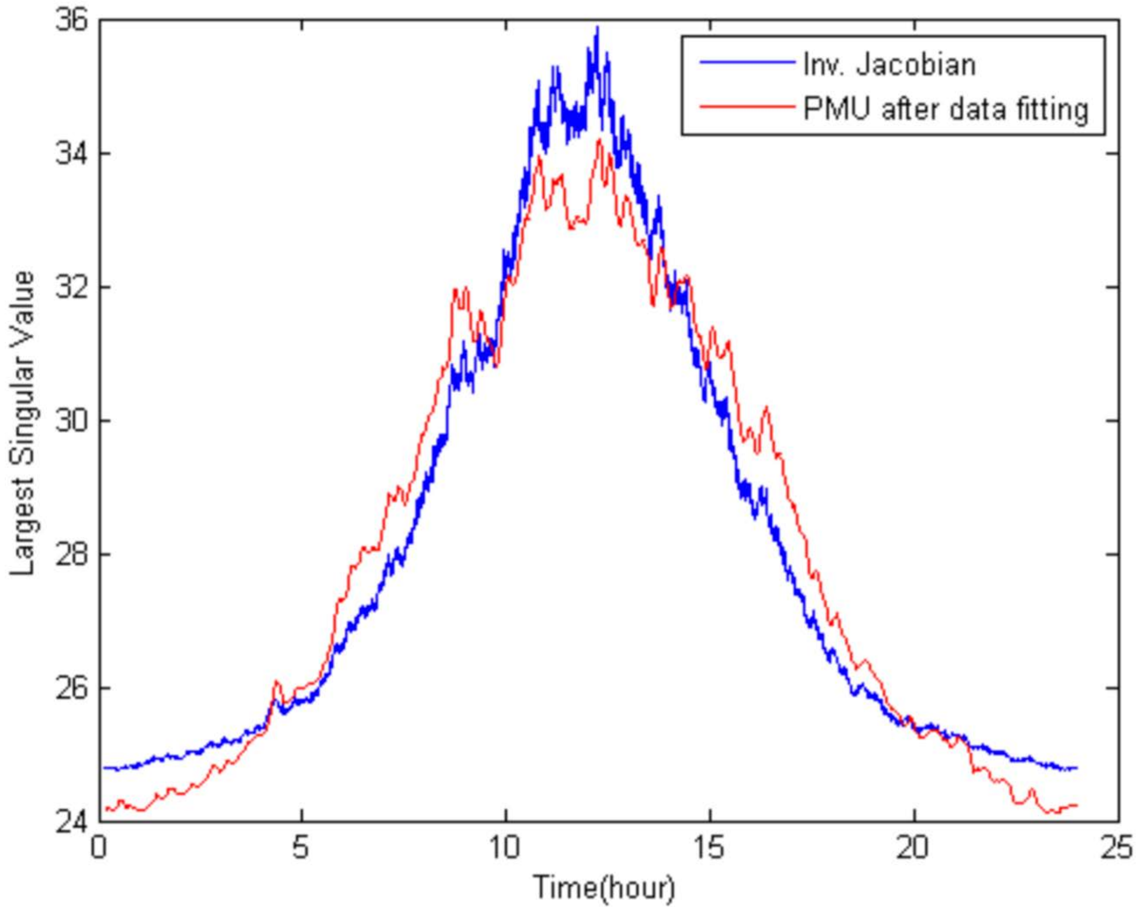


Fig. 4.3: Quality of fit between inverse Jacobian-based largest singular value versus PMU measurement-based largest singular value, IEEE 300 bus test case

4.4 Selection of window length for the PMU data matrix

The length of time window employed in the SVD calculation here may be considered from two perspectives, if one allows for the possibility of different PMU reporting rates (or down sampling, as was done for our prior numerical examples): one may consider either the "absolute" time-period spanned by the window, in physical units of seconds, or simply the number of time points in the window; i.e. the number of samples. Choice of the time window length for the SVD calculation presents natural trade-offs. A long window length may be expected to provide more dependable estimation of the "major axis" of the PMU measurements, when these are viewed as output data from the power systems mapping from input of load variation. However, a long window, with commensurate larger matrix dimension, imposes higher computational cost for the SVD. For sufficiently large matrices, with limited computational resources, this computation time delay might impede near-real-time display of system vulnerability to voltage instability (which is, of course, the whole point of this method). However, the method assumes that the load "inputs" randomly vary around an operating point to excite the system. Therefore, too small a window length may fail to excite the system response adequately, over a range of input directions, and thereby fail to capture the maximum singular value direction.

An attractive method to consider the issue of “sufficient excitation” is to consider the dimension of the space spanned by the columns of the PMU data matrix; i.e., the PMU data matrix rank. This rank must ultimately saturate once a sufficient number of columns are included. One may hypothesize that for any dominant phenomenon of interest the rank of the PMU-data matrix is reasonably low. This low-rank property of the PMU-data matrix would saturate the inclusion of a modest number of columns, yielding a computationally tractable SVD computation. From the perspective of the voltage instability measure, our premise would be that after some number of measurement samples are included, additional samples do not add significant new information to estimate the voltage instability metric. This class of problem is widely recognized in the literature, and determining threshold of interest (here, the number of columns that impact the rank) is often termed the “main dimension.” One approach to characterizing the main dimension computes the least number of components of normalized left singular vector such that the sum of squares of these components is greater than a predetermined threshold value, say μ. In keeping with our hypothesis above, for the numerical examples examined here, the main dimension proves to be relatively small relative to the overall dimension of the measurement vector; i.e., many components of the normalized left singular vectors are essentially zero. Fig. 4.4, Fig. 4.5, and Fig. 4.6 below display numeric results of the PMU-data matrix rank versus length of time window, plotted with four different choices of numeric threshold in the rank test (the horizontal axes show number of time sample points, and hence number of columns in the data matrix).

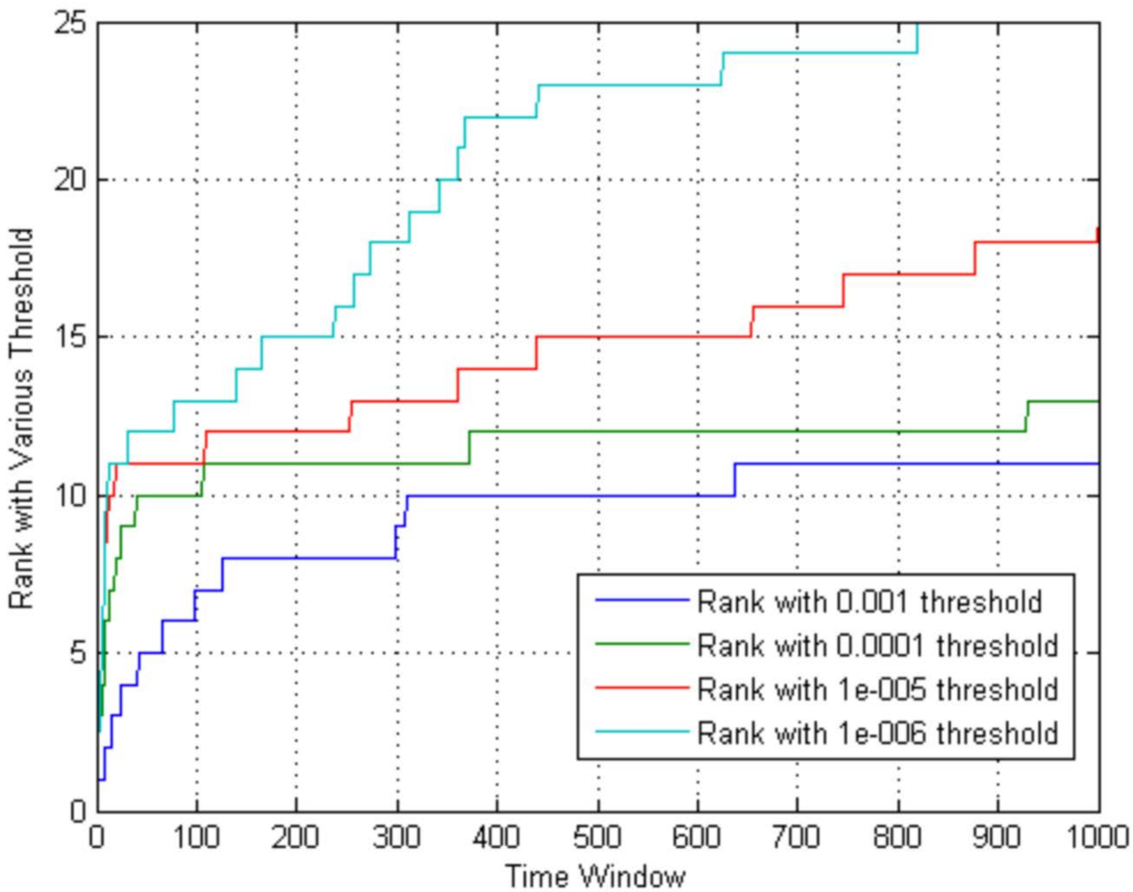


Fig. 4.4: Synthetic PMU Data Matrix rank versus window length, IEEE 14 bus example (note: 26 measurements considered, and hence rank is upper bounded by 26)

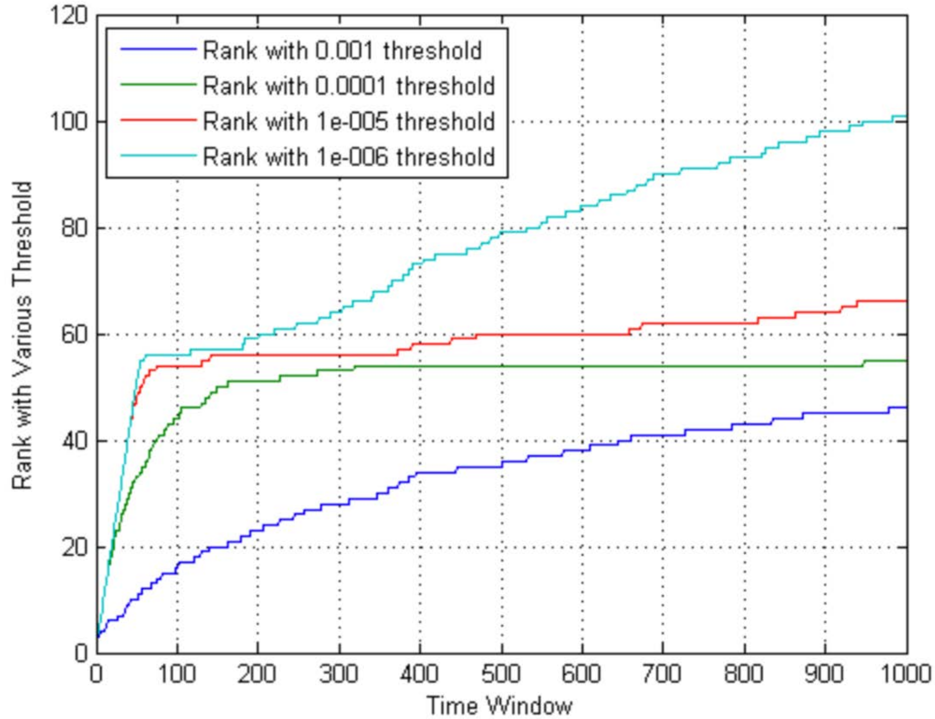


Fig. 4.5: Synthetic PMU Data Matrix rank versus window length, IEEE 118 bus example (note: 234 measurements considered, and hence rank is upper bounded by 234)

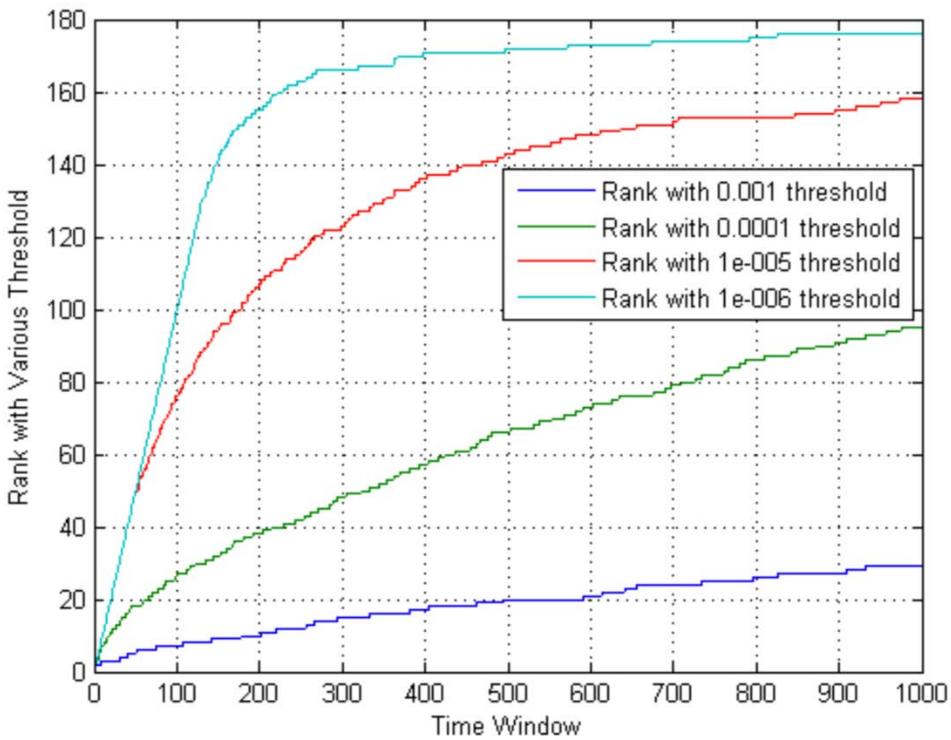


Fig. 4.6: Synthetic PMU Data Matrix rank versus window length, IEEE 300-bus example (note: 598 measurements considered, and hence rank is upper bounded by 598)

Interestingly, in our numerical experiments in this project using data from the IEEE 14, 118, and 300-bus systems, a larger number of components of the singular vector associated with voltage phase-angle is intended to play a role in the main dimension, relative to those components associated with voltage magnitude. We hypothesize that this is in keeping with voltage-reactive behavior being more localized in power systems, and hence fewer voltage magnitude components of the singular vector contribute to the main dimension. Also, our numerical experience indicates that the main dimension remains largely invariant over families of operating points, with load varying over a wide range. This observation further reinforces the premise that the main dimension identifies a useful intrinsic property in PMU datasets.

The main dimension computation is closely related to a problem that has been considered in other researchers work on PMU datasets, that of identifying core subspaces. In [60] the authors examined the $\boldsymbol{m}$ principal components of covariance matrix of PMU data, where $\boldsymbol{m}$ preserves the cumulative covariance. Both the work of [60] and the research conducted in this project indicate that the significant information content in PMU-data is relatively sparse. The threshold measuring significant information-content is set relatively "loosely" (i.e., one does not set the numeric threshold for rank too small).

4.5 Computational experiments using a measurement-based voltage stability metric

As mentioned above, the benchmark against which the method is compared is the established voltage stability metric of the largest singular value of the power flow Jacobian inverse, or equivalently, the inverse of the smallest singular value of the Jacobian. The numerical experiments reported below provide further experience using the comparisons described earlier; i.e., we employ synthetic examples in which the full network information is known and compute a large number of quasi-steady state operating point over some time window using synthetic load curves. The quality of the PMU measurement-based computation is judged by the accuracy with which it matches the singular value computed from full system information, using the power-flow Jacobian.

The load curves for studies here are constructed as the sum of two processes: *slowly varying deterministic process* and *fast stochastic process*. Even though the slowly varying load demand is not truly deterministic, utilities are typically able to accurately estimate the slowly varying load demands due to daily, seasonal, or annual cycles of human activity. Fast stochastic process in the load demands is from the millions of customer devices, switching on and off. Here, Ornstein-Uhlenbeck process has been used to model the load demand over 24- hour period; slowly varying deterministic process in the load demand is simply represented as peak at noon and lowest at midnight. For the stochastic process, the parameters, mean reverting rate and volatility, from the literature [61],[62],[63] are used with some modifications according to time-scale of interest.

The IEEE 118-bus system is considered first; from the base case data associated with this test case, variations in its load curve are hypothesized for each bus over a 24-hour period, with corresponding generation redispatch, and the pseudo-PMU data is generated over this time interval. As discussed previously, the exact hours of the day at which the load curve reaches minimum and maximum are not relevant for the test being performed here; hence for simplicity, the slowly varying deterministic process is assumed to have its peak at noon, and its minimum at midnight. The random process parameters, the mean reverting rate and the volatility, in Ornstein-Uhlenbeck process that represent stochastic process follow [61],[64]. The window size for the PMU data array in IEEE 118-bus system is 52 data samples, whose time interval is 8 minutes 40 seconds using a 0.1 Hz data sampling rate. From the pseudo-PMU data, the largest singular value of a windowed array of the data is computed and compared to the largest singular value of the inverse power flow Jacobian at each time step/operating point. The first set of results are displayed in Fig. 4.7 below, for a load curve representing a relatively lightly loaded case. Again, applying a fixed offset and scaling, the data derived measure and the power flow Jacobian inverse-based measure show very close agreement.

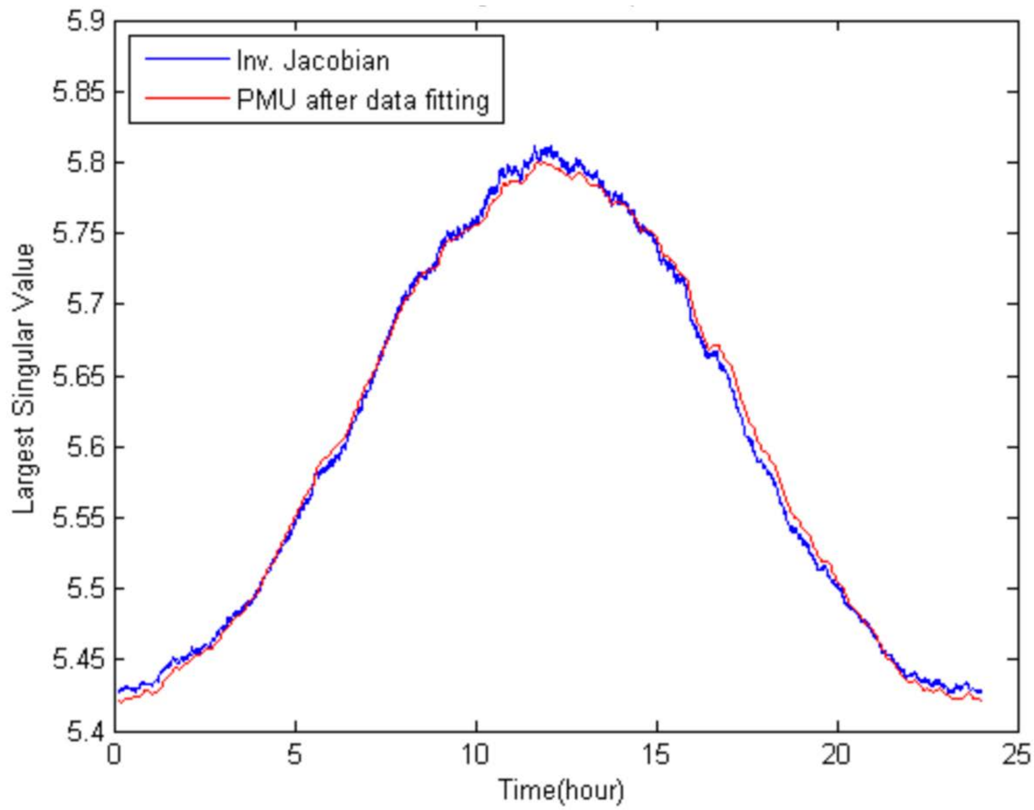


Fig. 4.7: Largest Singular Value: Inverse Jacobian-based versus PMU Measurement-Based IEEE 118 Bus System, Lightly Loaded Case

A second analysis is performed in the IEEE 118-bus example, but with a load curve that represents considerably more heavily loaded operating conditions. When system is very heavily loaded, the correlation between largest singular values from PMU data array and power flow Jacobian inverse does degrade slightly. Fig. 4.8 below shows that the accuracy of the proposed measure is slightly degraded in heavily loaded IEEE 118-bus system (i.e., the load curve employed for this example has higher values at each hour than that employed in the prior example). Although the correlation is not as exact as the system with lightly loaded condition, there still exist very strong correlations between the two measures. Because of the strong correlation of largest singular value of the Jacobian inverse, and that of the measurement array, both off-line system study and simpler engineering judgment can offer insight into selection of a voltage instability threshold. For IEEE 118-bus system case shown in Fig. 4.8, the system loses its steady-state solution (to within the computational accuracy of a full Newton-Raphson power flow computation) at a loading level one step beyond which the largest singular value of the power flow Jacobian inverse is equal to 7.402.

The corresponding largest singular value of PMU data array is 48.11, and the loading level is 2.514 times that of the nominal operating point. This type of off-line study provides a method for selecting an appropriate threshold for our proposed measure. This observation confirms that even the method here will not be completely “model free” some limited amount of off-line, model-based study to perform the affine fit between the largest singular value of the power flow Jacobian inverse, and the largest singular value of the PMU data matrix. However, once that affine fit is established, on-line computation of the largest singular value of the PMU data matrix then provides an estimate for the largest singular value for the power flow Jacobian inverse, without any on-line calculation of that Jacobian.

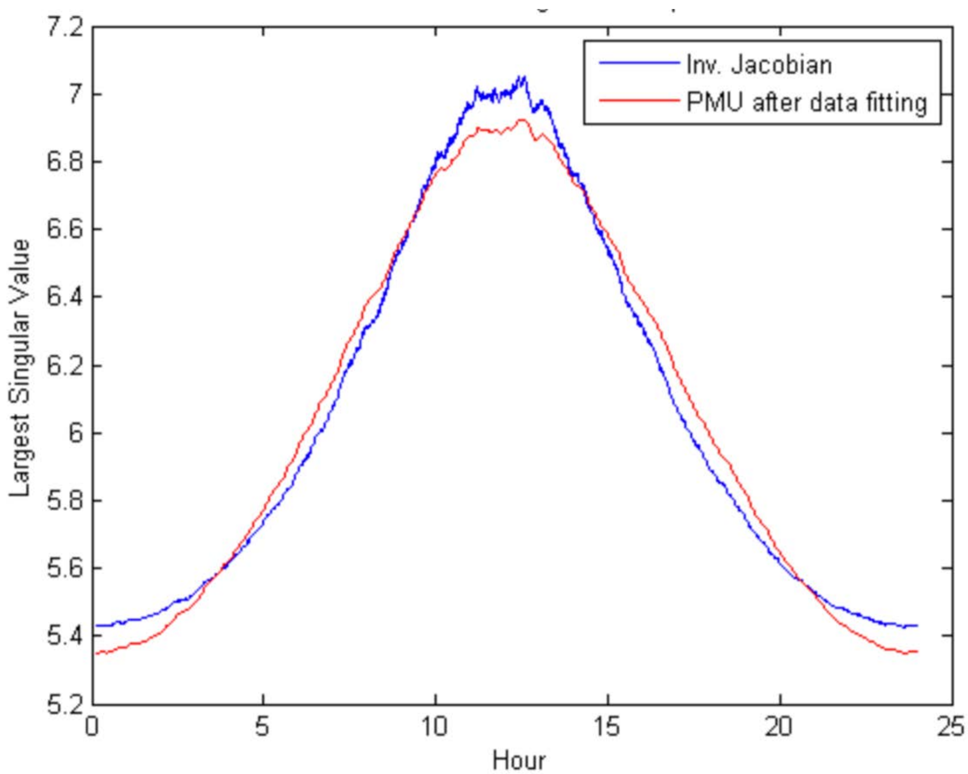


Fig. 4.8: Largest Singular Value: Inverse Jacobian-based versus PMU Measurement-Based IEEE 118 Bus System, Heavily Loaded Case

As an alternative method to select a maximum acceptable threshold below which the system is to be judged secure against voltage instability, recall the interpretation of largest singular value as approximating the maximum 2-norm gain of the power flow solution operator. Suppose one identifies a maximum credible deviation of load (ΔP, ΔQ) (as measured in 2-norm), and a maximum acceptable deviation in the power flow solution quantities ($\Delta\delta$, ΔV) (again, as measured in 2-norm). The ratio between these determines a worst-case allowable sensitivity of the power flow solution to changes in load; the largest singular value of the power flow Jacobian inverse is precisely this worst-case sensitivity. For the illustrative example in the IEEE 118-bus system here, our selection of a threshold of 6 has the engineering interpretation as a bound on the maximum ratio allowed between the load variation (in per unit) and power solution variation (in radian angle and per unit voltage magnitude). Again, the reader is reminded that the actual quantity calculated on-line will be the “model-free” PMU data matrix’s largest singular value, which in this example would have a corresponding threshold of 40.

As a somewhat larger study case, analogous computations were applied to a time sequence of operating points generated over hypothetical 24-hour pattern of load and generation variation for the IEEE 300-bus system. The window size for the PMU data array is 52 data samples, whose time interval is 8 minutes 40 seconds using 0.1 Hz data sampling rate. The outcome is displayed in the Fig. 4.9 below. In IEEE 300-bus system, the data fit is less exact that in the previous cases. We hypothesize that some of the differences is inherent to comparing the linearized information of the Jacobian inverse to the full nonlinear behavior of the power flow. In particular, the 300-bus test case had the property that a number of generators reach reactive power limits at high load levels. In typical power flow studies, this represents a change in the constitutive relations for the generator's reactive power/voltage magnitude behavior, which would imply a discontinuity in some terms of the power flow Jacobian as discussed previously. For the studies here, we adopted a smoothed approximation to avoid the discontinuity in the Jacobian. None-the-less, in the vicinity of operating points at which some generators encounter reactive limits, it is reasonable to assume that the linearized approximation to the power flow becomes less accurate and may account for some of the discrepancy between the pseudo-measurement-based singular value result, versus that based on the Jacobian inverse. While the researchers do not claim that the results here are comprehensive enough to verify this premise, one might argue that the measurement-based measure could prove a better indication of the conditioning of an operating point.

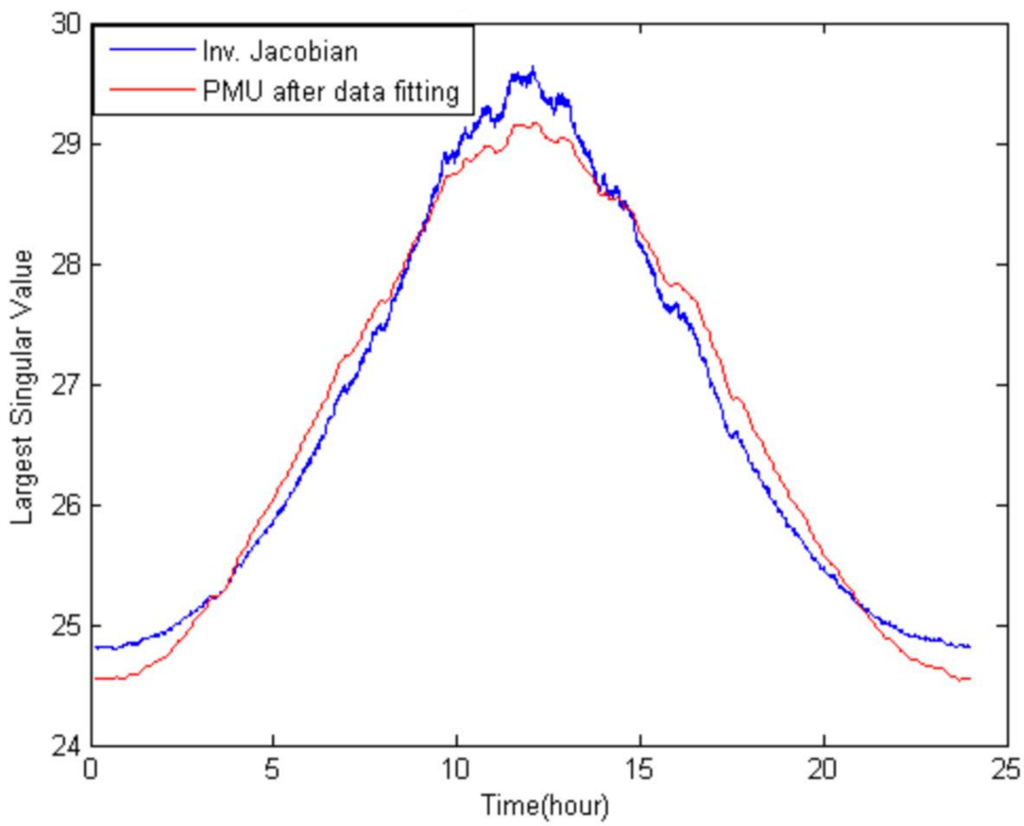


Fig. 4.9: Largest Singular Value: Inverse Jacobian-based versus PMU Measurement-Based IEEE 300 Bus System

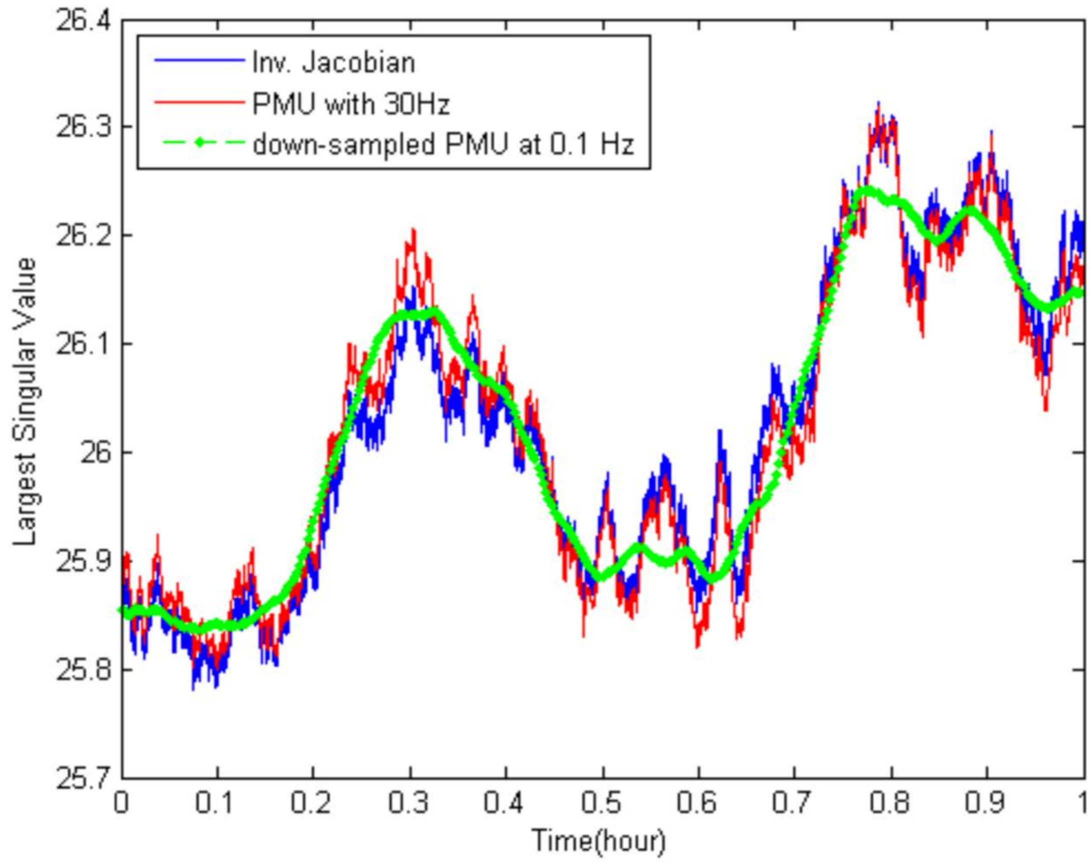


Fig. 4.10: Impact of PMU Data Down-Sampling on Singular Value Estimate

4.6 Cleaning PMU measurements for voltage stability applications

One drawback of the studies above, because they are performed on synthetic system data, is their inability to consider non-ideal, real-world effects of PMU data quality. The quality of PMU measurement data is an important factor in the performance of various applications relying on the data and may have significant impact on resulting stability measures such as the voltage stability index considered here. Raw PMU data may contain high-frequency noise and other artifacts due to bad data and data dropouts that can adversely affect the performance of compression algorithms. An in-depth look at the possible sources of data corruption and their implications can be found in [4]. It has been widely reported in literature that filtering out high frequency noise and correcting for bad and missing data is either a necessity or a preferred step in many applications. The specific quality requirements and the approaches for achieving those will vary based on the application. For the voltage stability analysis proposed in this work, the following schemes are being employed: (a) eliminating high frequency noise using a low-pass filter, and (b) discarding measurements with very high noise content. These schemes are designed by first analyzing the spectral properties of noise present in real PMU data.

4.7 Characterizing noise in PMU measurement data

To effectively identify and eliminate noise, it is important to recognize the nature and quantum of noise in a given measurement. While majority of the literature assumes noise in PMU data to be

Gaussian distributed [52], the authors of [5] argue by analyzing real PMU measurements that the noise is unlikely to follow Gaussian distribution. Regardless of its distribution, it is useful to characterize noise in terms of its spectral content. shows the spectral characteristics of a noise signal using a periodogram. The noise signal is extracted from a real PMU voltage measurement by subtracting a filtered version of itself. From Fig. 4.11, it can be noted that the noise is characterized by a nearly flat profile in the power spectral density diagram. This knowledge is used to identify the range of noise frequencies while designing the low-pass filter.

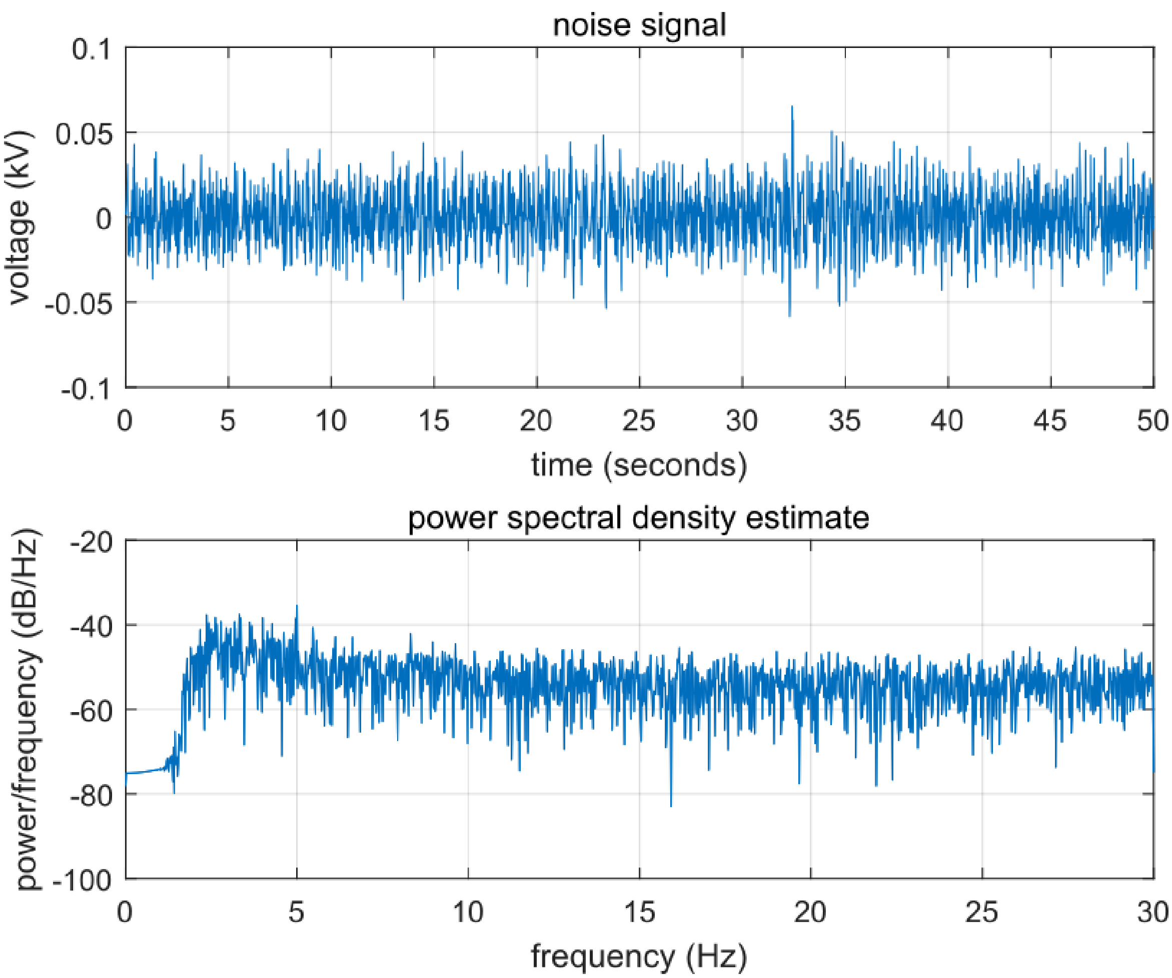


Fig. 4.11: Periodogram of noise in a voltage measurement

4.8 Low-pass filtering of PMU data

Designing a low-pass noise removal filter for voltage stability applications is an important task. Adequate care should be exercised to ensure that important information required for accurate detection by the algorithm are not eliminated or distorted by filtering. PMU measurements may contain a wide range of frequency information pertaining to disturbances caused by events in the power system. Fig. 4.12 shows the frequency ranges of different events and disturbances that can affect measurements.

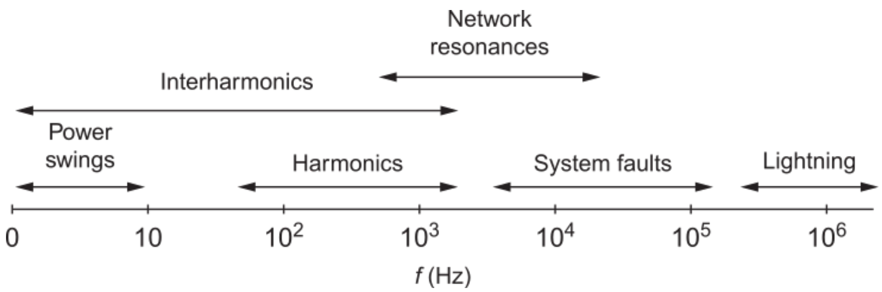


Fig. 4.12: Frequencies of events and disturbances [53]

High-frequency non-sinusoidal events are removed in the PMU itself by the anti-aliasing filter which limits the bandwidth of input signal. Furthermore, the PMU cannot accurately track the signals during the high-frequency transients. This is because the phasor estimation is performed on band-limited signals under the assumption of a quasi-steady-state, wherein the frequency is assumed to be at its nominal value. This non-trivial assumption significantly affects the accuracy of the PMU measurements during transients. The events that can be faithfully captured by the PMU are of fairly low frequency, like power swings. Inspecting the spectral properties of real PMU measurements also confirm that the noise content is of significantly higher frequency as compared to the phenomena of interest. This observation allows one to choose the cut-off frequency (f_c) to be in the range of 2 - 10 Hz based on the nature of the data and the applications.

In this work, the low-pass filter is implemented as a 30th order Hamming window-based filter with $f_c = 2$ Hz, after examining the spectral properties of real PMU data used in the study. The information about important system dynamics are preserved while the noise is eliminated. Fig. 4.13 shows the frequency response of the filter. The frequency response is characterized by reduced side-lobe magnitudes, which is a desirable feature for frequency-selective analysis. Fig. 4.14 shows the application of the low-pass filter to a typical voltage measurement with noise. The spectral content of these signals is quantified using a periodogram. An inspection of the filtered signal and its periodogram confirms that the filtering has not altered dynamics of interest either in time-domain or frequency-domain. The filtering process may introduce additional artefacts like delays and spikes. It is assumed that their effects on compression are negligible and are not corrected for in this work.

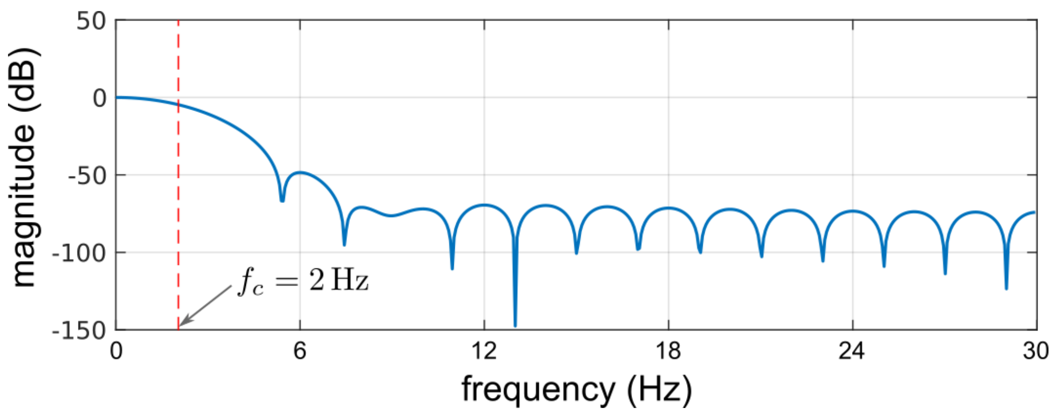


Fig. 4.13: Frequency response of a Hamming filter with $f_c = 2$ Hz

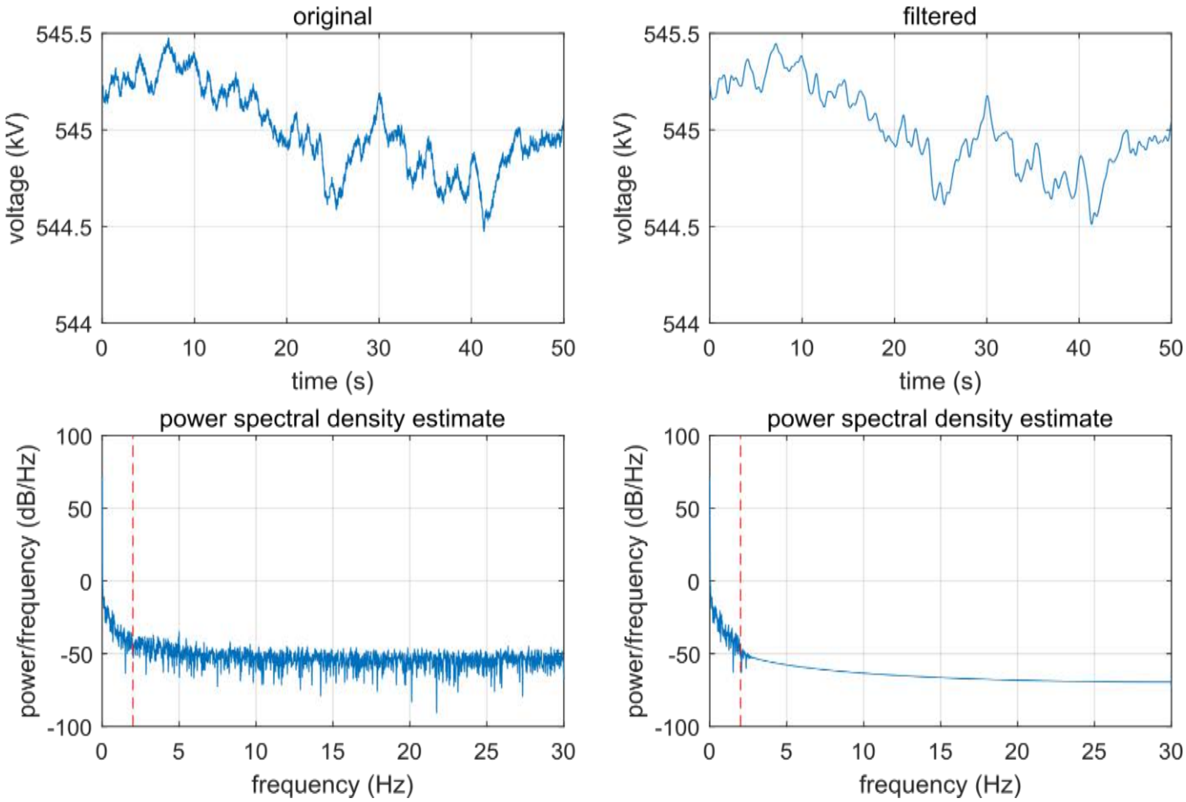


Fig. 4.14: Filtering the noisy signal with low-pass filter ($f_c = 2$ Hz). Each periodogram corresponds to the signal above it

4.9 Removing measurements with high noise content

Real PMU measurements can sometimes be highly corrupted and noisy. This can occur due to faulty measuring apparatus, interference in the communication channel, etc. When a measured quantity has magnitude near zero, the noise can dominate the signal of interest. For example, active and reactive power measurements from a PMU monitoring a line carrying negligible power may be dominated by noise. In such cases, the filtering approach described above cannot satisfactorily remove the noise content. Tests performed using real PMU data show that the presence of such measurements adversely affect the SVD-based tests for voltage stability margin and may raise "false alarms". Since the magnitude of the measurement is zero (or near zero) for the period of interest, the measurement can be temporarily removed from the dataset while performing the calculations. For this, it is required to quantify the noise in the signals and identify those measurements which are highly corrupted by noise.

A long-standing approach for quantifying noise in signals examines the autocorrelation in the noisy signal. Autocorrelation of a random process quantifies the correlation between its values at different points in time as a function of the time difference. For a discrete signal x of finite length T, the autocorrelation coefficient $R(\tau)$ can be defined as a function of the delay τ as

$$R(\tau) = \frac{\frac{1}{T}\sum_{t=1}^{T-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x})}{\sigma_x^2} \tag{36}$$

where $\bar{x}$ and σ_x^2 are the sample mean and sample variance of x, respectively. The autocorrelation coefficients have the following important properties:

1. The autocorrelation function in (36) is an even function
2. A maximum value of $R(\tau)$ occurs at $\tau = 0$
3. If x is non-periodic and has zero mean, then $\lim_{\tau\to\infty} R(\tau) = 0$

The last property can be exploited to quantify the noise content in signals. It can be observed that the rate at which $R(\tau)$ decays as $\tau \to \infty$ increases with higher noise content in the signal. This rate may be characterized by defining the *half-life* of autocorrelation as follows:

$$half - life \triangleq the\ time\ corresponding\ to\ the\ least\ value\ of\ \tau\ such\ that\ R(\tau) \leq R(0) \tag{37}$$

Fig. 4.15 shows an active power measurement for 10 minutes and its autocorrelation coefficients. The coefficients decay as τ increases. For this signal, the half-life can be identified as 297.95 seconds (corresponding to $\tau = 17877$). Fig. 4.16 shows an example of how the half-life can be used to identify measurements with high noise content. In the first signal, the magnitude of the measurement is nearly zero for the entire window considered. The autocorrelation coefficients reduce to small values very quickly, yielding a half-life of 0.05 seconds (corresponding to $\tau = 3$). In the other two signals, despite the presence of noise, the half-life is much larger as compared to 0.05 seconds. This significant separation in the half-life of signals can be used to identify and remove measurements with high noise content.

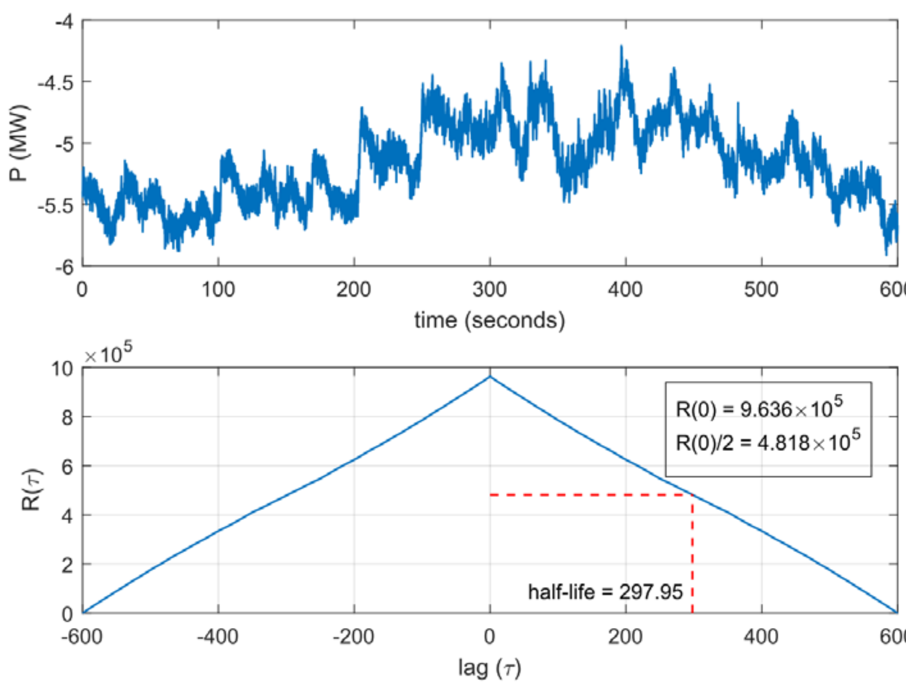


Fig. 4.15: Autocorrelation coefficients of an active power measurement from real PMU data. The dashed lines identify the half-life

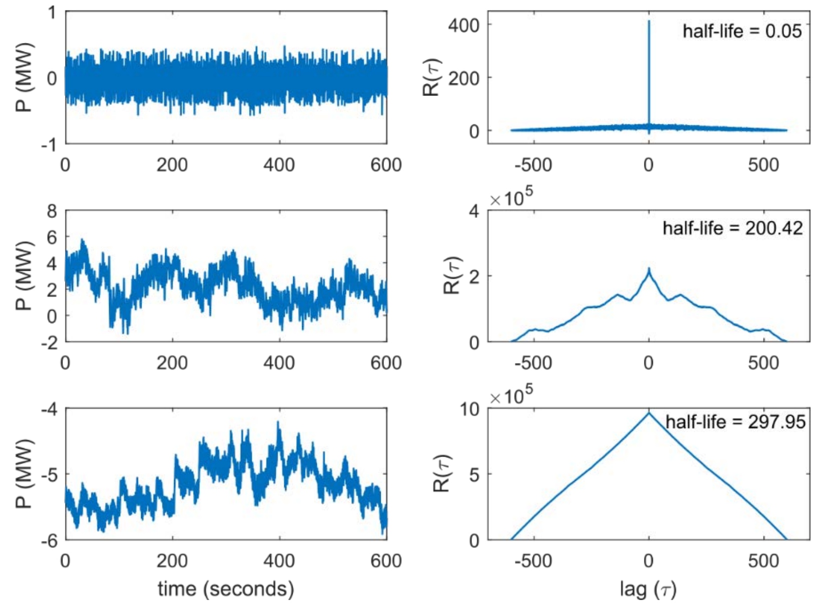


Fig. 4.16: Power measurements from different locations and the corresponding autocorrelation coefficients. The measurements are from real PMU data.

A possible concern regarding the noise quantification procedure described in this manner is of ensuring that signals with electro-mechanical oscillations of interest are not being wrongly classified as noise. Dynamic data generated by simulating various transient events like series and shunt faults, and load drops has been examined and the half-life of their autocorrelations computed. The decay of $R(\tau)$ observed was very slow as compared to noisy signals in all studied cases. This is illustrated in Fig. 4.17 for the IEEE 39-bus test system. From the analysis performed, it is seen that the half-life test is able to discriminate between signals with high noise content and measurements corresponding to transients with large excursions in signal magnitude.

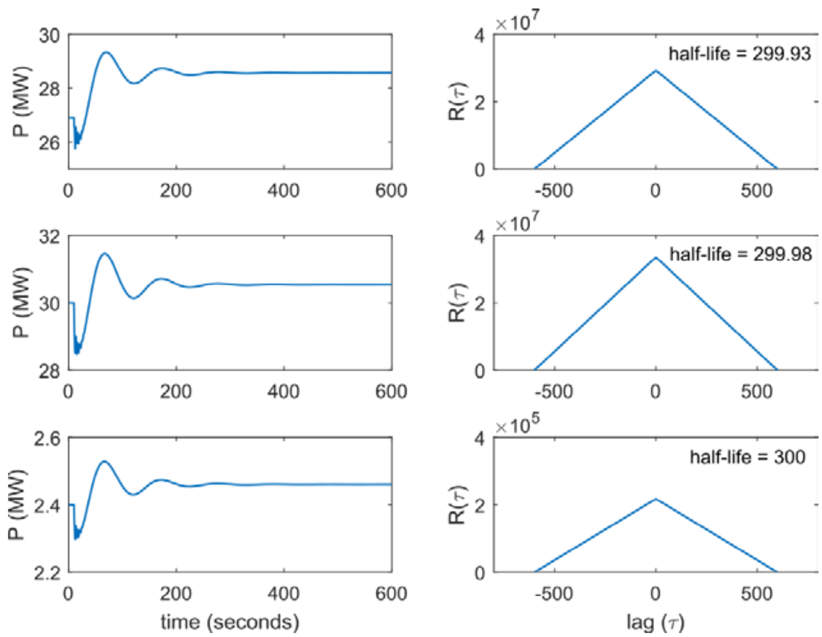


Fig. 4.17: Autocorrelation coefficients of active power measurements from different buses. The measurements are synthesized by simulating a line outage in the IEEE 39-bus system.

5. Conclusions

5.1 Research outcomes

With respect to the synchrophasor based power system monitoring applications, this research has made unique contributions to *power system islanding detection* and *online power system asset health monitoring* scheme. The instrumentation channel errors present in PMU data can deteriorate the islanding detection accuracy significantly. The research has proposed a novel wide-area measurement-based power-system islanding detection scheme which is immune to instrumentation channel errors contained in practical PMU measurements. The unique research finding in this context is as follows: *the cumulated sum of voltage phase angle difference (CUSPAD) obtained from a specific PMU device over a given time period cancels the effect of instrumentation channel errors present in PMU measurements*. In the context of *online-power system asset health monitoring scheme*, this project has proposed a new PMU based real-time power system asset health monitoring algorithm. The research has found out that the *signal-to-noise* (SNR) of PMU measurements is a robust metric to quantify equipment health. The bigger challenge was to identify the equipment (in the neighborhood of a PMU device) which is the cause of degradation in the quality of PMU measurements. Therefore, this research has also proposed an algorithm called the *Discriminating Code* to identify the equipment which is the unique cause of degradation of the measurement quality.

The power system cyber-protection application of this project proposed a new data-driven algorithm for the generation of synthetic bus-level time series load data at 30 samples per second that can be used on any system model. The proposed data-driven model is unique in its way because it can be used to learn spatial and temporal correlations from a dataset of real system loads and use the learnt models to generate new synthetic data that retains the same characteristics. In addition, the vulnerability of PMUs to cyber-attacks has been thoroughly investigated in this project. It has been observed that predictive filters that exploit the temporal correlations in the PMU measurements can be used to detect a sudden cyber-attack launched on PMU data, that would previously remain undetected by conventional bad data detectors. However, a more intelligently designed ramping cyber-attack is more challenging to detect.

With respect to the power system-based control application, the research conducted in this PSERC project proposed a measurement-based approach for power system voltage stability assessment. The metric employed was based on singular value decomposition (SVD). The research reviewed the underlying modeling assumptions for SVD-based voltage stability metric and described the data structure employed to organize the PMU measurements for computation of this metric. However, such measurement-based techniques can be largely affected by noise present in PMU data. Therefore, the research also focused on measurement filtering techniques adapted to this application. It described a novel method to identify “noise dominated” measurement channels that contributed no useful information to the SVD calculation and are therefore candidates for removal from the measurement set.

5.2 Future scope of work

The following two future scopes of work have been identified for the research done in the course of this project.

- PMU data is at times characterized by bad data (data dropouts, stale data) due to loss of GPS synchronization. In the presence of such biased PMU measurements, it becomes more challenging to detect the malfunctioning of an equipment from the measurements. As a scope of future work, such practical constraints will be addressed. It is known that the PMU data has high spatio-temporal correlation. Machine learning techniques will be explored in the future research to learn the spatial and temporal correlation in PMU measurements for robust health prediction of power system equipment even in the presence of biased measurements.

- This project has demonstrated that PMU devices are susceptible to cyber-attacks, especially when the attack is an intelligently ramping cyber-attack. Therefore, it extremely important to detect such type of cyber-attacks to improve the resiliency of the PMUs. Future research will be directed towards a robust detection algorithm for such type of cyber-attacks on PMU measurements. Application of machine learning techniques for detecting anomalies in PMU measurements will also be explored as a future scope of research.

6. Appendix

6.1 Dynamic data of the Type-IV wind turbine generator in GE-PSLF

The dynamic data that was used for the Type-IV wind turbine generator used in the GE-PSLF software is given as follows:

- wt4g Bus no. "Bus name" 0.6 "1 " : #9 MVA 1.0000 10.000 0.9000/ 0.4000 1.2200 1.2000 0.8000 0.4000 -1.300 0.7000
- wt4e Bus no. "Bus name" 0.6 "1 " : #9 1.0000 0.1000 20.00 1.1000 /

  0.900 4000 -0.4000 0.0200 0.150 18.000 5.000 1.0000 0.0500 0.0500 1.2400 /

  0.901 1.2500 0.0000 1.700 1.600
- wt4t Bus no. "Bus name" 0.6 "1 " : #9 0.0500 0.0800 0.1000 0.0800 /

  0.0 0.1000 -0.1000

The controller gains for the wind turbine models are as below:

- Controller gains for Wt4g:

  Kpp 0.08 PI controller proportional gain, p.u.

  Kip 0.10 PI controller integral gain, p.u
- Controller gains for Wt4e:

  Kqi 0.1 Q control integral gain

  Kvi 12. V control integral gain

6.2 Modified 118-bus system with 10% wind penetration

Bus data

Bus	Type	V (p.u.)	V(deg.)	Bus	Type	V (p.u.)	V(deg.)
1	2	0.9705	-17.31	37	1	0.9919	-16.21
2	1	0.9773	-16.64	38	1	0.9619	-11.05
3	1	0.979	-16.37	39	1	0.9704	-19.59
4	2	1.01	-12.7	40	2	0.97	-20.67
5	1	1.0101	-12.21	41	1	0.9668	-21.12
6	2	0.99	-14.79	42	2	0.985	-19.54
7	1	0.9893	-15.23	43	1	0.9785	-16.73
8	2	1.015	-7.2	44	1	0.985	-14.27
9	1	1.0049	0.49	45	1	0.9867	-12.45
10	2	0.9763	9.11	46	2	1.005	-9.65
11	1	0.9887	-15.12	47	1	1.017	-7.43
12	2	0.99	-15.58	48	1	1.0206	-8.21
13	1	0.9711	-16.5	49	2	1.025	-7.21
14	1	0.9836	-16.32	50	1	1.0213	-9.54
15	2	0.97	-16.69	51	1	1.0153	-12.44
16	1	0.9839	-15.92	52	1	1.0137	-13.46
17	1	0.9951	-14.19	53	1	1.0253	-14.6
18	2	0.973	-16.4	54	2	1.049	-14.01
19	2	0.9631	-16.88	55	2	1.0346	-14.09
20	1	0.956	-16.04	56	2	1.0368	-13.95
21	1	0.9553	-14.47	57	1	1.0273	-12.48
22	1	0.965	-11.93	58	1	1.0223	-13.36
23	1	0.9927	-7	59	2	0.9862	-8.79
24	2	0.992	-7.24	60	1	0.9932	-4.97
25	2	1.027	0.23	61	2	0.995	-4.07
26	2	1.015	1.97	62	2	0.998	-4.7
27	2	0.968	-12.77	63	1	0.9692	-5.37
28	1	0.9616	-14.46	64	1	0.9839	-3.58
29	1	0.9632	-15.43	65	2	1.005	-0.38
30	1	0.9854	-9.13	66	2	1.039	-0.51
31	2	0.967	-15.3	67	1	1.0137	-3.23
32	2	0.963	-13.28	68	1	1.0005	-0.47
33	1	0.9715	-17.31	69	0	1.035	1.8
34	2	0.9858	-16.67	70	2	1.0115	-5.94
35	1	0.9807	-17.1	71	1	1.0008	-6.16
36	2	0.98	-17.09	72	2	0.98	-7.14

Bus	Type	V (p.u.)	V(deg.)	Bus	Type	V (p.u.)	V(deg.)
73	2	0.991	-6.22	113	2	0.993	-14.21
74	2	1.0322	-7.65	114	1	0.9601	-13.63
75	1	1.0072	-5.96	115	1	0.96	-13.64
76	2	0.9701	-6.95	116	2	1.005	-0.92
77	2	0.9853	-1.46	117	1	0.9738	-17.12
78	1	0.9783	-1.71	118	1	0.9836	-6.85
79	1	0.9757	-1.29				
80	1	0.9841	1.44				
280	1	0.9841	14.75				
380	2	1.04	27.34				
81	1	0.9749	0.31				
82	1	0.9709	-0.93				
83	1	0.972	0.21				
84	1	0.9757	2.63				
85	2	0.985	4.11				
86	1	0.9867	2.74				
87	2	1.015	3				
88	1	0.9874	7.25				
89	2	1.005	11.32				
90	2	0.985	4.91				
91	2	0.98	4.92				
92	2	0.99	5.46				
93	1	0.9803	2.51				
94	1	0.9802	0.4				
95	1	0.9647	-0.5				
96	1	0.9696	-0.54				
97	1	0.9719	0.05				
98	1	0.988	-0.55				
99	2	1.01	-1.35				
100	2	1.017	-0.38				
101	1	0.9914	1.22				
102	1	0.9891	3.94				
103	2	1.0007	-3.97				
104	2	0.971	-6.72				
105	2	0.966	-7.84				
106	1	0.9618	-8.09				
107	2	0.952	-10.88				
108	1	0.9668	-9.03				
109	1	0.9675	-9.48				
110	2	0.973	-10.32				
111	2	0.98	-8.67				
112	2	0.975	-13.42				

Line Data

From	To	R(p.u.)	X(p.u.)	B(p.u.)
1	2	0.0303	0.0999	0.0254
1	3	0.0129	0.0424	0.0108
2	12	0.0187	0.0616	0.0157
3	5	0.0241	0.108	0.0284
3	12	0.0484	0.16	0.0406
4	5	0.0018	0.008	0.0021
4	11	0.0209	0.0688	0.0175
5	6	0.0119	0.054	0.0143
5	11	0.0203	0.0682	0.0174
6	7	0.0046	0.0208	0.0055
7	12	0.0086	0.034	0.0087
8	9	0.0024	0.0305	1.162
8	30	0.0043	0.0504	0.514
9	10	0.0026	0.0322	1.23
11	12	0.006	0.0196	0.005
11	13	0.0223	0.0731	0.0188
12	14	0.0215	0.0707	0.0182
12	16	0.0212	0.0834	0.0214
12	117	0.0329	0.14	0.0358
13	15	0.0744	0.2444	0.0627
14	15	0.0595	0.195	0.0502
15	17	0.0132	0.0437	0.0444
15	19	0.012	0.0394	0.0101
15	33	0.038	0.1244	0.0319
16	17	0.0454	0.1801	0.0466
17	18	0.0123	0.0505	0.013
17	31	0.0474	0.1563	0.0399
17	113	0.0091	0.0301	0.0077
18	19	0.0112	0.0493	0.0114
19	20	0.0252	0.117	0.0298
19	34	0.0752	0.247	0.0632
20	21	0.0183	0.0849	0.0216
21	22	0.0209	0.097	0.0246
22	23	0.0342	0.159	0.0404
23	24	0.0135	0.0492	0.0498
23	25	0.0156	0.08	0.0864
23	32	0.0317	0.1153	0.1173
24	70	0.0022	0.4115	0.102
24	72	0.0488	0.196	0.0488
25	27	0.0318	0.163	0.1764

From	To	R(p.u.)	X(p.u.)	B(p.u.)
26	30	0.008	0.086	0.908
27	28	0.0191	0.0855	0.0216
27	32	0.0229	0.0755	0.0193
27	115	0.0164	0.0741	0.0197
28	29	0.0237	0.0943	0.0238
29	31	0.0108	0.0331	0.0083
30	38	0.0046	0.054	0.422
31	32	0.0298	0.0985	0.0251
32	113	0.0615	0.203	0.0518
32	114	0.0135	0.0612	0.0163
33	37	0.0415	0.142	0.0366
34	36	0.0087	0.0268	0.0057
34	37	0.0026	0.0094	0.0098
34	43	0.0413	0.1681	0.0423
35	36	0.0022	0.0102	0.0027
35	37	0.011	0.0497	0.0132
37	39	0.0321	0.106	0.027
37	40	0.0593	0.168	0.042
38	65	0.009	0.0986	1.046
39	40	0.0184	0.0605	0.0155
40	41	0.0145	0.0487	0.0122
40	42	0.0555	0.183	0.0466
41	42	0.041	0.135	0.0344
42	49	0.0715	0.323	0.086
42	49	0.0715	0.323	0.086
43	44	0.0608	0.2454	0.0607
44	45	0.0224	0.0901	0.0224
45	46	0.04	0.1356	0.0332
45	49	0.0684	0.186	0.0444
46	47	0.038	0.127	0.0316
46	48	0.0601	0.189	0.0472
47	49	0.0191	0.0625	0.016
47	69	0.0844	0.2778	0.0709
48	49	0.0179	0.0505	0.0126
49	50	0.0267	0.0752	0.0187
49	51	0.0486	0.137	0.0342
49	54	0.0869	0.291	0.073
49	54	0.073	0.289	0.0738
49	66	0.018	0.0919	0.0248
49	66	0.018	0.0919	0.0248

From	To	R(p.u.)	X(p.u.)	B(p.u.)
49	69	0.0985	0.324	0.0828
50	57	0.0474	0.134	0.0332
51	52	0.0203	0.0588	0.014
51	58	0.0255	0.0719	0.0179
52	53	0.0405	0.1635	0.0406
53	54	0.0263	0.122	0.031
54	55	0.0169	0.0707	0.0202
54	56	0.0027	0.0095	0.0073
54	59	0.0503	0.2293	0.0598
55	56	0.0049	0.0151	0.0037
55	59	0.0474	0.2158	0.0565
56	57	0.0343	0.0966	0.0242
56	58	0.0343	0.0966	0.0242
56	59	0.0803	0.239	0.0536
56	59	0.0825	0.251	0.0569
59	60	0.0317	0.145	0.0376
59	61	0.0328	0.15	0.0388
60	61	0.0026	0.0135	0.0146
60	62	0.0123	0.0561	0.0147
61	62	0.0082	0.0376	0.0098
62	66	0.0482	0.218	0.0578
62	67	0.0258	0.117	0.031
63	64	0.0017	0.02	0.216
64	65	0.0027	0.0302	0.38
65	68	0.0014	0.016	0.638
66	67	0.0224	0.1015	0.0268
68	81	0.0018	0.0202	0.808
68	116	0.0003	0.0041	0.164
69	70	0.03	0.127	0.122
69	75	0.0405	0.122	0.124
69	77	0.0309	0.101	0.1038
70	71	0.0088	0.0355	0.0088
70	74	0.0401	0.1323	0.0337
70	75	0.0428	0.141	0.036
71	72	0.0446	0.18	0.0444
71	73	0.0087	0.0454	0.0118
74	75	0.0123	0.0406	0.0103
75	77	0.0601	0.1999	0.0498
75	118	0.0145	0.0481	0.012
76	77	0.0444	0.148	0.0368
76	118	0.0164	0.0544	0.0136
77	78	0.0038	0.0124	0.0126
77	80	0.0294	0.105	0.0228

From	To	R(p.u.)	X(p.u.)	B(p.u.)
77	80	0.017	0.0485	0.0472
77	82	0.0298	0.0853	0.0817
78	79	0.0055	0.0244	0.0065
79	80	0.0156	0.0704	0.0187
80	96	0.0356	0.182	0.0494
80	97	0.0183	0.0934	0.0254
80	98	0.0238	0.108	0.0286
80	99	0.0454	0.206	0.0546
82	83	0.0112	0.0366	0.038
82	96	0.0162	0.053	0.0544
83	84	0.0625	0.132	0.0258
83	85	0.043	0.148	0.0348
84	85	0.0302	0.0641	0.0123
85	86	0.035	0.123	0.0276
85	88	0.02	0.102	0.0276
85	89	0.0239	0.173	0.047
86	87	0.0283	0.2074	0.0445
88	89	0.0139	0.0712	0.0193
89	90	0.0518	0.188	0.0528
89	90	0.0238	0.0997	0.106
89	92	0.0393	0.1581	0.0414
89	92	0.0099	0.0505	0.0548
90	91	0.0254	0.0836	0.0214
91	92	0.0387	0.1272	0.0327
92	93	0.0258	0.0848	0.0218
92	94	0.0481	0.158	0.0406
92	100	0.0648	0.295	0.0472
92	102	0.0123	0.0559	0.0146
93	94	0.0223	0.0732	0.0188
94	95	0.0132	0.0434	0.0111
94	96	0.0269	0.0869	0.023
94	100	0.0178	0.058	0.0604
95	96	0.0171	0.0547	0.0147
96	97	0.0173	0.0885	0.024
98	100	0.0397	0.179	0.0476
99	100	0.018	0.0813	0.0216
100	101	0.0277	0.1262	0.0328
100	103	0.016	0.0525	0.0536
100	104	0.0451	0.204	0.0541
100	106	0.0605	0.229	0.062
101	102	0.0246	0.112	0.0294
103	104	0.0466	0.1584	0.0407
103	105	0.0535	0.1625	0.0408

From	To	R(p.u.)	X(p.u.)	B(p.u.)
103	110	0.0391	0.1813	0.0461
104	105	0.0099	0.0378	0.0099
105	106	0.014	0.0547	0.0143
105	107	0.053	0.183	0.0472
105	108	0.0261	0.0703	0.0184
106	107	0.053	0.183	0.0472
108	109	0.0105	0.0288	0.0076
109	110	0.0278	0.0762	0.0202
110	111	0.022	0.0755	0.02
110	112	0.0247	0.064	0.062
114	115	0.0023	0.0104	0.0028

Transformer Data

From	To	MVA	R(p.u.)	X(p.u.)
8	5	100	0	0.0267
26	25	100	0	0.0382
30	17	100	0	0.0388
38	37	100	0	0.0375
63	59	100	0	0.0386
64	61	100	0	0.0268
65	66	100	0	0.037
68	69	100	0	0.037
81	80	100	0	0.037
80	280	100	0	0.05
280	380	100	0	0.05

Generator Data

Bus	Pg(MW)	Qg(MVar)	Bus	Pg(MW)	Qg(MVar)
1	0	15	36	0	-1.1
4	0	46.4	40	0	27
6	0	0.9	42	0	41
8	0	160.6	46	19	-5.2
10	450	-147	49	204	13.8
12	85	56.8	54	48	262.8
15	0	2.7	55	0	23
18	0	26.2	56	0	-8
19	0	-8	59	155	-60
24	0	-7.8	61	160	-42.6
25	220	-47	62	0	11.3
26	314	74.1	65	391	130
27	0	18.4	66	392	-67
31	7	32.6	69	518.2	-110.4
32	0	-10.5	70	0	32
34	0	-8	72	0	-18.7

Bus	Pg(MW)	Qg(MVar)
73	0	-20.8
74	0	9
76	0	23
77	0	70
380	477	26.6
85	0	8.4
87	4	11
89	607	-5.9
90	0	59.3
91	0	-13.1
92	0	-2.1
99	0	10.2
100	252	150
103	40	40
104	0	5.7
105	0	-8
107	0	5.7
110	0	4.9
111	36	-1.8
112	0	41.5
113	0	6.6
116	0	120.8

Shunt Data

Bus	G(pu)	B(pu)
5	0	-0.4
34	0	0.14
37	0	-0.25
44	0	0.1
45	0	0.1
46	0	0.1
48	0	0.15
74	0	1.1
79	0	0.2
82	0	0.2
83	0	0.1
105	0	0.2
107	0	0.06
110	0	0.06

Load Data

Bus	Pload	Qload	Bus	Pload	Qload	Bus	Pload	Qload
1	51	27	49	87	30	99	42	0
2	20	9	50	17	4	100	37	18
3	39	10	51	17	8	101	22	15
4	39	12	52	18	5	102	5	3
6	52	22	53	23	11	103	23	16
7	19	2	54	113	32	104	38	25
8	28	0	55	63	22	105	31	26
11	70	23	56	84	18	106	43	16
12	47	10	57	12	3	107	50	12
13	34	16	58	12	3	108	2	1
14	14	1	59	277	113	109	8	3
15	90	30	60	78	3	110	39	30
16	25	10	62	77	14	112	68	13
17	11	3	66	39	18	113	6	0
18	60	34	67	28	7	114	8	3
19	45	25	70	66	20	115	22	7
20	18	3	72	12	0	116	184	0
21	14	8	73	6	0	117	20	8
22	10	5	74	68	27	118	33	15
23	7	3	75	47	11			
24	13	0	76	68	36			
27	71	13	77	61	28			
28	17	7	78	71	26			
29	24	4	79	39	32			
31	43	27	80	130	26			
32	59	23	82	54	27			
33	23	9	83	20	10			
34	59	26	84	11	7			
35	33	9	85	24	15			
36	31	17	86	21	10			
39	27	11	88	48	10			
40	66	23	90	163	42			
41	37	10	91	10	0			
42	96	23	92	65	10			
43	18	7	93	12	7			
44	16	8	94	30	16			
45	53	22	95	42	31			
46	28	10	96	38	15			
47	34	0	97	15	9			
48	20	11	98	34	8			

6.3 Modified 118-bus system with 20% wind penetration

Bus data

Bus	Type	V (p.u.)	V(deg.)	Bus	Type	V (p.u.)	V(deg.)
1	2	0.9561	-17.3	35	1	0.9807	-17.2
2	1	0.9587	-16.53	36	2	0.98	-17.2
3	1	0.9666	-16.36	37	1	0.9919	-16.31
4	2	1.01	-12.8	38	1	0.9617	-11.14
5	1	1.0086	-12.29	39	1	0.9704	-19.71
6	2	0.99	-14.89	40	2	0.97	-20.8
7	1	0.9814	-15.22	41	1	0.9668	-21.26
8	2	1.015	-7.28	42	2	0.985	-19.69
9	1	1.0049	0.41	43	1	0.9785	-16.86
10	2	0.9763	9.04	44	1	0.9851	-14.44
11	1	0.9756	-15.04	45	1	0.9867	-12.63
12	1	0.9691	-15.38	46	2	1.005	-9.83
212	1	1.0166	-13.07	47	1	1.017	-7.6
312	2	1.0702	-10.97	48	1	1.0206	-8.41
13	1	0.9608	-16.49	49	2	1.025	-7.42
14	1	0.9681	-16.23	50	1	1.0208	-9.79
15	2	0.97	-16.81	51	1	1.0145	-12.75
16	1	0.9692	-15.83	52	1	1.013	-13.78
17	1	0.9943	-14.28	53	1	1.025	-14.97
18	2	0.973	-16.51	54	2	1.049	-14.41
19	2	0.9631	-17	55	2	1.0321	-14.46
20	1	0.956	-16.15	56	2	1.035	-14.33
21	1	0.9553	-14.57	57	1	1.026	-12.81
22	1	0.9649	-12.02	58	1	1.0211	-13.7
23	1	0.9927	-7.08	59	1	0.966	-8.91
24	2	0.992	-7.31	259	1	0.9981	-4.6
25	2	1.027	0.15	359	2	1.04	-0.6
26	2	1.015	1.89	60	1	0.9651	-4.93
27	2	0.968	-12.86	61	1	0.965	-3.97
28	1	0.9616	-14.56	261	1	0.9748	0.59
29	1	0.9632	-15.52	361	2	0.995	5.02
30	1	0.9851	-9.21	62	2	0.9747	-4.7
31	2	0.967	-15.39	63	1	0.9515	-5.36
32	2	0.963	-13.37	64	1	0.9673	-3.51
33	1	0.9715	-17.43	65	2	1.005	-0.41
34	2	0.9858	-16.77	66	2	1.0354	-0.6

Bus	Type	V (p.u.)	V(deg.)
67	1	1.0008	-3.28
68	1	1.0005	-0.48
69	0	1.035	1.8
70	2	1.0115	-5.95
71	1	1.0008	-6.18
72	2	0.98	-7.18
73	2	0.991	-6.24
74	2	1.0321	-7.66
75	1	1.0072	-5.97
76	2	0.9701	-6.96
77	2	0.9853	-1.47
78	1	0.9783	-1.71
79	1	0.9757	-1.3
80	1	0.984	1.43
280	1	0.9841	14.74
380	2	1.04	27.33
81	1	0.9749	0.3
82	1	0.9709	-0.94
83	1	0.972	0.2
84	1	0.9757	2.62
85	2	0.985	4.1
86	1	0.9867	2.73
87	2	1.015	2.99
88	1	0.9874	7.24
89	2	1.005	11.31
90	2	0.985	4.9
91	2	0.98	4.91
92	2	0.99	5.45
93	1	0.9803	2.5
94	1	0.9802	0.39
95	1	0.9647	-0.51
96	1	0.9695	-0.55
97	1	0.9719	0.04
98	1	0.988	-0.56
99	2	1.01	-1.37
100	2	1.017	-0.39

Bus	Type	V (p.u.)	V(deg.)
101	1	0.9914	1.21
102	1	0.9891	3.93
103	2	1.0007	-3.98
104	2	0.971	-6.73
105	2	0.966	-7.85
106	1	0.9618	-8.1
107	2	0.952	-10.89
108	1	0.9668	-9.04
109	1	0.9675	-9.49
110	2	0.973	-10.33
111	2	0.98	-8.68
112	2	0.975	-13.43
113	2	0.993	-14.32
114	1	0.9601	-13.72
115	1	0.96	-13.73
116	2	1.005	-0.93
117	1	0.9524	-16.99
118	1	0.9836	-6.86

Line Data

From	To	R(p.u.)	X(p.u.)	B(p.u.)
1	2	0.0303	0.0999	0.0254
1	3	0.0129	0.0424	0.0108
2	12	0.0187	0.0616	0.0157
3	5	0.0241	0.108	0.0284
3	12	0.0484	0.16	0.0406
4	5	0.0018	0.008	0.0021
4	11	0.0209	0.0688	0.0175
5	6	0.0119	0.054	0.0143
5	11	0.0203	0.0682	0.0174
6	7	0.0046	0.0208	0.0055
7	12	0.0086	0.034	0.0087
8	9	0.0024	0.0305	1.162
8	30	0.0043	0.0504	0.514
9	10	0.0026	0.0322	1.23
11	12	0.006	0.0196	0.005
11	13	0.0223	0.0731	0.0188
12	14	0.0215	0.0707	0.0182
12	16	0.0212	0.0834	0.0214
12	117	0.0329	0.14	0.0358
13	15	0.0744	0.2444	0.0627
14	15	0.0595	0.195	0.0502
15	17	0.0132	0.0437	0.0444
15	19	0.012	0.0394	0.0101
15	33	0.038	0.1244	0.0319
16	17	0.0454	0.1801	0.0466
17	18	0.0123	0.0505	0.013
17	31	0.0474	0.1563	0.0399
17	113	0.0091	0.0301	0.0077
18	19	0.0112	0.0493	0.0114
19	20	0.0252	0.117	0.0298
19	34	0.0752	0.247	0.0632
20	21	0.0183	0.0849	0.0216
21	22	0.0209	0.097	0.0246
22	23	0.0342	0.159	0.0404
23	24	0.0135	0.0492	0.0498
23	25	0.0156	0.08	0.0864
23	32	0.0317	0.1153	0.1173
24	70	0.0022	0.4115	0.102
24	72	0.0488	0.196	0.0488
25	27	0.0318	0.163	0.1764

From	To	R(p.u.)	X(p.u.)	B(p.u.)
26	30	0.008	0.086	0.908
27	28	0.0191	0.0855	0.0216
27	32	0.0229	0.0755	0.0193
27	115	0.0164	0.0741	0.0197
28	29	0.0237	0.0943	0.0238
29	31	0.0108	0.0331	0.0083
30	38	0.0046	0.054	0.422
31	32	0.0298	0.0985	0.0251
32	113	0.0615	0.203	0.0518
32	114	0.0135	0.0612	0.0163
33	37	0.0415	0.142	0.0366
34	36	0.0087	0.0268	0.0057
34	37	0.0026	0.0094	0.0098
34	43	0.0413	0.1681	0.0423
35	36	0.0022	0.0102	0.0027
35	37	0.011	0.0497	0.0132
37	39	0.0321	0.106	0.027
37	40	0.0593	0.168	0.042
38	65	0.009	0.0986	1.046
39	40	0.0184	0.0605	0.0155
40	41	0.0145	0.0487	0.0122
40	42	0.0555	0.183	0.0466
41	42	0.041	0.135	0.0344
42	49	0.0715	0.323	0.086
42	49	0.0715	0.323	0.086
43	44	0.0608	0.2454	0.0607
44	45	0.0224	0.0901	0.0224
45	46	0.04	0.1356	0.0332
45	49	0.0684	0.186	0.0444
46	47	0.038	0.127	0.0316
46	48	0.0601	0.189	0.0472
47	49	0.0191	0.0625	0.016
47	69	0.0844	0.2778	0.0709
48	49	0.0179	0.0505	0.0126
49	50	0.0267	0.0752	0.0187
49	51	0.0486	0.137	0.0342
49	54	0.0869	0.291	0.073
49	54	0.073	0.289	0.0738
49	66	0.018	0.0919	0.0248
49	66	0.018	0.0919	0.0248

From	To	R(p.u.)	X(p.u.)	B(p.u.)
49	69	0.0985	0.324	0.0828
50	57	0.0474	0.134	0.0332
51	52	0.0203	0.0588	0.014
51	58	0.0255	0.0719	0.0179
52	53	0.0405	0.1635	0.0406
53	54	0.0263	0.122	0.031
54	55	0.0169	0.0707	0.0202
54	56	0.0027	0.0095	0.0073
54	59	0.0503	0.2293	0.0598
55	56	0.0049	0.0151	0.0037
55	59	0.0474	0.2158	0.0565
56	57	0.0343	0.0966	0.0242
56	58	0.0343	0.0966	0.0242
56	59	0.0803	0.239	0.0536
56	59	0.0825	0.251	0.0569
59	60	0.0317	0.145	0.0376
59	61	0.0328	0.15	0.0388
60	61	0.0026	0.0135	0.0146
60	62	0.0123	0.0561	0.0147
61	62	0.0082	0.0376	0.0098
62	66	0.0482	0.218	0.0578
62	67	0.0258	0.117	0.031
63	64	0.0017	0.02	0.216
64	65	0.0027	0.0302	0.38
65	68	0.0014	0.016	0.638
66	67	0.0224	0.1015	0.0268
68	81	0.0018	0.0202	0.808
68	116	0.0003	0.0041	0.164
69	70	0.03	0.127	0.122
69	75	0.0405	0.122	0.124
69	77	0.0309	0.101	0.1038
70	71	0.0088	0.0355	0.0088
70	74	0.0401	0.1323	0.0337
70	75	0.0428	0.141	0.036
71	72	0.0446	0.18	0.0444
71	73	0.0087	0.0454	0.0118
74	75	0.0123	0.0406	0.0103
75	77	0.0601	0.1999	0.0498
75	118	0.0145	0.0481	0.012
76	77	0.0444	0.148	0.0368
76	118	0.0164	0.0544	0.0136
77	78	0.0038	0.0124	0.0126
77	80	0.0294	0.105	0.0228
77	80	0.017	0.0485	0.0472

From	To	R(p.u.)	X(p.u.)	B(p.u.)
77	82	0.0298	0.0853	0.0817
78	79	0.0055	0.0244	0.0065
79	80	0.0156	0.0704	0.0187
80	96	0.0356	0.182	0.0494
80	97	0.0183	0.0934	0.0254
80	98	0.0238	0.108	0.0286
80	99	0.0454	0.206	0.0546
82	83	0.0112	0.0366	0.038
82	96	0.0162	0.053	0.0544
83	84	0.0625	0.132	0.0258
83	85	0.043	0.148	0.0348
84	85	0.0302	0.0641	0.0123
85	86	0.035	0.123	0.0276
85	88	0.02	0.102	0.0276
85	89	0.0239	0.173	0.047
86	87	0.0283	0.2074	0.0445
88	89	0.0139	0.0712	0.0193
89	90	0.0518	0.188	0.0528
89	90	0.0238	0.0997	0.106
89	92	0.0393	0.1581	0.0414
89	92	0.0099	0.0505	0.0548
90	91	0.0254	0.0836	0.0214
91	92	0.0387	0.1272	0.0327
92	93	0.0258	0.0848	0.0218
92	94	0.0481	0.158	0.0406
92	100	0.0648	0.295	0.0472
92	102	0.0123	0.0559	0.0146
93	94	0.0223	0.0732	0.0188
94	95	0.0132	0.0434	0.0111
94	96	0.0269	0.0869	0.023
94	100	0.0178	0.058	0.0604
95	96	0.0171	0.0547	0.0147
96	97	0.0173	0.0885	0.024
98	100	0.0397	0.179	0.0476
99	100	0.018	0.0813	0.0216
100	101	0.0277	0.1262	0.0328
100	103	0.016	0.0525	0.0536
100	104	0.0451	0.204	0.0541
100	106	0.0605	0.229	0.062
101	102	0.0246	0.112	0.0294
103	104	0.0466	0.1584	0.0407
103	105	0.0535	0.1625	0.0408
103	110	0.0391	0.1813	0.0461
104	105	0.0099	0.0378	0.0099

From	To	R(p.u.)	X(p.u.)	B(p.u.)
105	106	0.014	0.0547	0.0143
105	107	0.053	0.183	0.0472
105	108	0.0261	0.0703	0.0184
106	107	0.053	0.183	0.0472
108	109	0.0105	0.0288	0.0076
109	110	0.0278	0.0762	0.0202
110	111	0.022	0.0755	0.02
110	112	0.0247	0.064	0.062
114	115	0.0023	0.0104	0.0028

Transformer Data

From	To	MVA	R(p.u.)	X(p.u.)
8	5	100	0	0.0267
26	25	100	0	0.0382
30	17	100	0	0.0388
38	37	100	0	0.0375
63	59	100	0	0.0386
64	61	100	0	0.0268
65	66	100	0	0.037
68	69	100	0	0.037
81	80	100	0	0.037
12	212	100	0	0.05
59	259	100	0	0.05
61	261	100	0	0.05
80	280	100	0	0.05
212	312	100	0	0.05
259	359	100	0	0.05
261	361	100	0	0.05
280	380	100	0	0.05

Generator Data

Bus	Pg(MW)	Qg(MVar)	Bus	Pg(MW)	Qg(MVar)
1	0	15	65	391	130
4	0	46.4	66	392	-67
6	0	0.9	69	518.2	-110.4
8	0	160.6	70	0	32
10	450	-147	72	0	-18.7
12	85	56.8	73	0	-20.8
15	0	2.7	74	0	9
18	0	26.2	76	0	23
19	0	-8	77	0	70
24	0	-7.8	380	477	26.6
25	220	-47	85	0	8.4

Bus	Pg(MW)	Qg(MVar)	Bus	Pg(MW)	Qg(MVar)
27	0	18.4	89	607	-5.9
31	7	32.6	90	0	59.3
32	0	-10.5	91	0	-13.1
34	0	-8	92	0	-2.1
36	0	-1.1	99	0	10.2
40	0	27	100	252	150
42	0	41	103	40	40
46	19	-5.2	104	0	5.7
49	204	13.8	105	0	-8
54	48	262.8	107	0	5.7
55	0	23	110	0	4.9
56	0	-8	111	36	-1.8
59	155	-60	112	0	41.5
61	160	-42.6	113	0	6.6
62	0	11.3	116	0	120.8

Load Data

Bus	Pload	Qload	Bus	Pload	Qload	Bus	Pload	Qload
1	51	27	49	87	30	99	42	0
2	20	9	50	17	4	100	37	18
3	39	10	51	17	8	101	22	15
4	39	12	52	18	5	102	5	3
6	52	22	53	23	11	103	23	16
7	19	2	54	113	32	104	38	25
8	28	0	55	63	22	105	31	26
11	70	23	56	84	18	106	43	16
12	47	10	57	12	3	107	50	12
13	34	16	58	12	3	108	2	1
14	14	1	59	277	113	109	8	3
15	90	30	60	78	3	110	39	30
16	25	10	62	77	14	112	68	13
17	11	3	66	39	18	113	6	0
18	60	34	67	28	7	114	8	3
19	45	25	70	66	20	115	22	7
20	18	3	72	12	0	116	184	0
21	14	8	73	6	0	117	20	8
22	10	5	74	68	27	118	33	15
23	7	3	75	47	11			
24	13	0	76	68	36			
27	71	13	77	61	28			
28	17	7	78	71	26			
29	24	4	79	39	32			

Bus	Pload	Qload	Bus	Pload	Qload
32	59	23	82	54	27
33	23	9	83	20	10
34	59	26	84	11	7
35	33	9	85	24	15
36	31	17	86	21	10
39	27	11	88	48	10
40	66	23	90	163	42
41	37	10	91	10	0
42	96	23	92	65	10
43	18	7	93	12	7
44	16	8	94	30	16
45	53	22	95	42	31
46	28	10	96	38	15
47	34	0	97	15	9
48	20	11	98	34	8

Shunt Data

Bus	G(pu)	B(pu)
5	0	-0.4
34	0	0.14
37	0	-0.25
44	0	0.1
45	0	0.1
46	0	0.1
48	0	0.15
74	0	1.1
79	0	0.2
82	0	0.2
83	0	0.1
105	0	0.2
107	0	0.06
110	0	0.06

6.4 Modified 118-bus system with 30% wind penetration

Bus data

Bus	Type	V (p.u.)	V(deg.)	Bus	Type	V (p.u.)	V(deg.)
1	2	0.9561	-17.78	35	1	0.9806	-17.67
2	1	0.9587	-17.01	36	2	0.98	-17.67
3	1	0.9666	-16.85	37	1	0.9917	-16.78
4	2	1.01	-13.28	38	1	0.9612	-11.59
5	1	1.0086	-12.77	39	1	0.9703	-20.18
6	2	0.99	-15.37	40	2	0.97	-21.28
7	1	0.9814	-15.71	41	1	0.9668	-21.74
8	2	1.015	-7.76	42	2	0.985	-20.18
9	1	1.0049	-0.08	43	1	0.9784	-17.35
10	2	0.9763	8.55	44	1	0.9851	-14.95
11	1	0.9756	-15.52	45	1	0.9867	-13.15
12	1	0.9691	-15.87	46	2	1.005	-10.37
212	1	1.0166	-13.55	47	1	1.0124	-8.09
312	2	1.0702	-11.46	48	1	1.0206	-8.92
13	1	0.9608	-16.97	49	2	1.025	-7.92
14	1	0.9681	-16.71	50	1	1.0208	-10.28
15	2	0.97	-17.3	51	1	1.0145	-13.23
16	1	0.9692	-16.32	52	1	1.013	-14.26
17	1	0.9943	-14.77	53	1	1.025	-15.44
18	2	0.973	-17	54	2	1.049	-14.87
19	2	0.9631	-17.49	55	2	1.032	-14.92
20	1	0.956	-16.65	56	2	1.035	-14.79
21	1	0.9554	-15.09	57	1	1.0259	-13.28
22	1	0.965	-12.55	58	1	1.0211	-14.17
23	1	0.9927	-7.63	59	1	0.9655	-9.33
24	2	0.992	-7.89	259	1	0.9978	-5.01
25	2	1.027	-0.37	359	2	1.04	-1.01
26	2	1.015	1.38	60	1	0.9642	-5.32
27	2	0.968	-13.37	61	1	0.9642	-4.35
28	1	0.9616	-15.07	261	1	0.9744	0.21
29	1	0.9632	-16.03	361	2	0.995	4.64
30	1	0.9849	-9.7	62	2	0.9738	-5.09
31	2	0.967	-15.9	63	1	0.9505	-5.75
32	2	0.963	-13.89	64	1	0.9661	-3.88
33	1	0.9715	-17.9	65	2	1.0031	-0.75
34	2	0.9857	-17.24	66	2	1.0343	-0.99

Bus	Type	V (p.u.)	V(deg.)	Bus	Type	V (p.u.)	V(deg.)
67	1	0.9998	-3.68	99	2	1.01	-1.89
68	1	0.9971	-0.76	100	2	1.0168	-0.92
69	1	1.0017	1.8	101	1	0.9913	0.67
269	1	0.9855	16.14	102	1	0.9891	3.39
369	0	1.035	30	103	2	1.0006	-4.51
70	2	0.9967	-6.55	104	2	0.971	-7.26
71	1	0.9933	-6.87	105	2	0.966	-8.38
72	2	0.98	-7.88	106	1	0.9617	-8.63
73	2	0.991	-7	107	2	0.952	-11.42
74	2	1.0117	-8.23	108	1	0.9668	-9.57
75	1	0.9863	-6.49	109	1	0.9674	-10.02
76	2	0.9518	-7.52	110	2	0.973	-10.86
77	2	0.9726	-1.86	111	2	0.98	-9.22
78	1	0.9661	-2.12	112	2	0.975	-13.96
79	1	0.9647	-1.7	113	2	0.993	-14.81
80	1	0.9769	1.04	114	1	0.9601	-14.24
280	1	0.9802	14.51	115	1	0.96	-14.25
380	2	1.04	27.15	116	2	1.005	-1.23
81	1	0.9703	-0.02	117	1	0.9524	-17.47
82	1	0.9646	-1.41	118	1	0.9638	-7.41
83	1	0.9676	-0.28				
84	1	0.9743	2.09				
85	2	0.985	3.55				
86	1	0.9867	2.18				
87	2	1.015	2.44				
88	1	0.9874	6.7				
89	2	1.005	10.77				
90	2	0.985	4.36				
91	2	0.98	4.37				
92	2	0.99	4.91				
93	1	0.9792	1.98				
94	1	0.9781	-0.12				
95	1	0.9614	-1				
96	1	0.9648	-1.02				
97	1	0.966	-0.39				
98	1	0.9834	-1.02				

Line Data

From	To	R(p.u.)	X(p.u.)	B(p.u.)
1	2	0.0303	0.0999	0.0254
1	3	0.0129	0.0424	0.0108
2	12	0.0187	0.0616	0.0157
3	5	0.0241	0.108	0.0284
3	12	0.0484	0.16	0.0406
4	5	0.0018	0.008	0.0021
4	11	0.0209	0.0688	0.0175
5	6	0.0119	0.054	0.0143
5	11	0.0203	0.0682	0.0174
6	7	0.0046	0.0208	0.0055
7	12	0.0086	0.034	0.0087
8	9	0.0024	0.0305	1.162
8	30	0.0043	0.0504	0.514
9	10	0.0026	0.0322	1.23
11	12	0.006	0.0196	0.005
11	13	0.0223	0.0731	0.0188
12	14	0.0215	0.0707	0.0182
12	16	0.0212	0.0834	0.0214
12	117	0.0329	0.14	0.0358
13	15	0.0744	0.2444	0.0627
14	15	0.0595	0.195	0.0502
15	17	0.0132	0.0437	0.0444
15	19	0.012	0.0394	0.0101
15	33	0.038	0.1244	0.0319
16	17	0.0454	0.1801	0.0466
17	18	0.0123	0.0505	0.013
17	31	0.0474	0.1563	0.0399
17	113	0.0091	0.0301	0.0077
18	19	0.0112	0.0493	0.0114
19	20	0.0252	0.117	0.0298
19	34	0.0752	0.247	0.0632
20	21	0.0183	0.0849	0.0216
21	22	0.0209	0.097	0.0246
22	23	0.0342	0.159	0.0404
23	24	0.0135	0.0492	0.0498
23	25	0.0156	0.08	0.0864
23	32	0.0317	0.1153	0.1173
24	70	0.0022	0.4115	0.102
24	72	0.0488	0.196	0.0488
25	27	0.0318	0.163	0.1764

From	To	R(p.u.)	X(p.u.)	B(p.u.)
26	30	0.008	0.086	0.908
27	28	0.0191	0.0855	0.0216
27	32	0.0229	0.0755	0.0193
27	115	0.0164	0.0741	0.0197
28	29	0.0237	0.0943	0.0238
29	31	0.0108	0.0331	0.0083
30	38	0.0046	0.054	0.422
31	32	0.0298	0.0985	0.0251
32	113	0.0615	0.203	0.0518
32	114	0.0135	0.0612	0.0163
33	37	0.0415	0.142	0.0366
34	36	0.0087	0.0268	0.0057
34	37	0.0026	0.0094	0.0098
34	43	0.0413	0.1681	0.0423
35	36	0.0022	0.0102	0.0027
35	37	0.011	0.0497	0.0132
37	39	0.0321	0.106	0.027
37	40	0.0593	0.168	0.042
38	65	0.009	0.0986	1.046
39	40	0.0184	0.0605	0.0155
40	41	0.0145	0.0487	0.0122
40	42	0.0555	0.183	0.0466
41	42	0.041	0.135	0.0344
42	49	0.0715	0.323	0.086
42	49	0.0715	0.323	0.086
43	44	0.0608	0.2454	0.0607
44	45	0.0224	0.0901	0.0224
45	46	0.04	0.1356	0.0332
45	49	0.0684	0.186	0.0444
46	47	0.038	0.127	0.0316
46	48	0.0601	0.189	0.0472
47	49	0.0191	0.0625	0.016
47	69	0.0844	0.2778	0.0709
48	49	0.0179	0.0505	0.0126
49	50	0.0267	0.0752	0.0187
49	51	0.0486	0.137	0.0342
49	54	0.0869	0.291	0.073
49	54	0.073	0.289	0.0738
49	66	0.018	0.0919	0.0248
49	66	0.018	0.0919	0.0248

From	To	R(p.u.)	X(p.u.)	B(p.u.)
49	69	0.0985	0.324	0.0828
50	57	0.0474	0.134	0.0332
51	52	0.0203	0.0588	0.014
51	58	0.0255	0.0719	0.0179
52	53	0.0405	0.1635	0.0406
53	54	0.0263	0.122	0.031
54	55	0.0169	0.0707	0.0202
54	56	0.0027	0.0095	0.0073
54	59	0.0503	0.2293	0.0598
55	56	0.0049	0.0151	0.0037
55	59	0.0474	0.2158	0.0565
56	57	0.0343	0.0966	0.0242
56	58	0.0343	0.0966	0.0242
56	59	0.0803	0.239	0.0536
56	59	0.0825	0.251	0.0569
59	60	0.0317	0.145	0.0376
59	61	0.0328	0.15	0.0388
60	61	0.0026	0.0135	0.0146
60	62	0.0123	0.0561	0.0147
61	62	0.0082	0.0376	0.0098
62	66	0.0482	0.218	0.0578
62	67	0.0258	0.117	0.031
63	64	0.0017	0.02	0.216
64	65	0.0027	0.0302	0.38
65	68	0.0014	0.016	0.638
66	67	0.0224	0.1015	0.0268
68	81	0.0018	0.0202	0.808
68	116	0.0003	0.0041	0.164
69	70	0.03	0.127	0.122
69	75	0.0405	0.122	0.124
69	77	0.0309	0.101	0.1038
70	71	0.0088	0.0355	0.0088
70	74	0.0401	0.1323	0.0337
70	75	0.0428	0.141	0.036
71	72	0.0446	0.18	0.0444
71	73	0.0087	0.0454	0.0118
74	75	0.0123	0.0406	0.0103
75	77	0.0601	0.1999	0.0498
75	118	0.0145	0.0481	0.012
76	77	0.0444	0.148	0.0368
76	118	0.0164	0.0544	0.0136

From	To	R(p.u.)	X(p.u.)	B(p.u.)
76	118	0.0164	0.0544	0.0136
77	78	0.0038	0.0124	0.0126
77	80	0.0294	0.105	0.0228
77	80	0.017	0.0485	0.0472
77	82	0.0298	0.0853	0.0817
78	79	0.0055	0.0244	0.0065
79	80	0.0156	0.0704	0.0187
80	96	0.0356	0.182	0.0494
80	97	0.0183	0.0934	0.0254
80	98	0.0238	0.108	0.0286
80	99	0.0454	0.206	0.0546
82	83	0.0112	0.0366	0.038
82	96	0.0162	0.053	0.0544
83	84	0.0625	0.132	0.0258
83	85	0.043	0.148	0.0348
84	85	0.0302	0.0641	0.0123
85	86	0.035	0.123	0.0276
85	88	0.02	0.102	0.0276
85	89	0.0239	0.173	0.047
86	87	0.0283	0.2074	0.0445
88	89	0.0139	0.0712	0.0193
89	90	0.0518	0.188	0.0528
89	90	0.0238	0.0997	0.106
89	92	0.0393	0.1581	0.0414
89	92	0.0099	0.0505	0.0548
90	91	0.0254	0.0836	0.0214
91	92	0.0387	0.1272	0.0327
92	93	0.0258	0.0848	0.0218
92	94	0.0481	0.158	0.0406
92	100	0.0648	0.295	0.0472
92	102	0.0123	0.0559	0.0146
93	94	0.0223	0.0732	0.0188
94	95	0.0132	0.0434	0.0111
94	96	0.0269	0.0869	0.023
94	100	0.0178	0.058	0.0604
95	96	0.0171	0.0547	0.0147
96	97	0.0173	0.0885	0.024
98	100	0.0397	0.179	0.0476
99	100	0.018	0.0813	0.0216
100	101	0.0277	0.1262	0.0328

From	To	R(p.u.)	X(p.u.)	B(p.u.)
100	103	0.016	0.0525	0.0536
100	104	0.0451	0.204	0.0541
100	106	0.0605	0.229	0.062
101	102	0.0246	0.112	0.0294
103	104	0.0466	0.1584	0.0407
103	105	0.0535	0.1625	0.0408
103	110	0.0391	0.1813	0.0461
104	105	0.0099	0.0378	0.0099
105	106	0.014	0.0547	0.0143
105	107	0.053	0.183	0.0472
105	108	0.0261	0.0703	0.0184
106	107	0.053	0.183	0.0472
108	109	0.0105	0.0288	0.0076
109	110	0.0278	0.0762	0.0202
110	111	0.022	0.0755	0.02
110	112	0.0247	0.064	0.062
114	115	0.0023	0.0104	0.0028

Transformer Data

8	5	100	0	0.0267
26	25	100	0	0.0382
30	17	100	0	0.0388
38	37	100	0	0.0375
63	59	100	0	0.0386
64	61	100	0	0.0268
65	66	100	0	0.037
68	69	100	0	0.037
81	80	100	0	0.037
12	212	100	0	0.05
59	259	100	0	0.05
61	261	100	0	0.05
69	269	100	0	0.05
80	280	100	0	0.05
212	312	100	0	0.05
259	359	100	0	0.05
261	361	100	0	0.05
269	369	100	0	0.05
280	380	100	0	0.05

Generator Data

Bus	Pg(MW)	Qg(MVar)	Bus	Pg(MW)	Qg(MVar)
1	0	15	15	0	16.5
4	0	84.2	18	0	27.8
6	0	41.3	19	0	-8
8	0	167.2	24	0	-4.2
10	450	-147	25	220	-47
312	85	-35	26	314	74.6

Bus	Pg(MW)	Qg(MVar)
27	0	18.4
31	7	33.1
32	0	-10.5
34	0	-8
36	0	-0.9
40	0	27
42	0	40.9
46	19	-5.2
49	204	24.1
54	48	296.5
55	0	23
56	0	-8
359	155	-51.8
361	160	-88.5
62	0	20
65	391	196.3
66	392	-67
369	522.7	24.3
70	0	32
72	0	-14.6
73	0	-4.4
74	0	9
76	0	23
77	0	70
380	477	35.4
85	0	13.4
87	4	11
89	607	-5.9
90	0	59.3
91	0	-13.1
92	0	0.5
99	0	14

Bus	Pg(MW)	Qg(MVar)
103	40	40
104	0	5.9
105	0	-8
107	0	5.7
110	0	4.9
111	36	-1.8
112	0	41.5
113	0	9.3
116	0	204.2

Shunt Data

Bus	G(pu)	B(pu)
5	0	-0.4
34	0	0.14
37	0	-0.25
44	0	0.1
45	0	0.1
46	0	0.1
48	0	0.15
74	0	1.1
79	0	0.2
82	0	0.2
83	0	0.1
105	0	0.2
107	0	0.06
110	0	0.06

Load Data

Bus	Pload	Qload	Bus	Pload	Qload
1	51	27	43	18	7
2	20	9	44	16	8
3	39	10	45	53	22
4	39	12	46	28	10
6	52	22	47	34	0
7	19	2	48	20	11
8	28	0	49	87	30
11	70	23	50	17	4
12	47	10	51	17	8
13	34	16	52	18	5
14	14	1	53	23	11
15	90	30	54	113	32
16	25	10	55	63	22
17	11	3	56	84	18
18	60	34	57	12	3
19	45	25	58	12	3
20	18	3	59	277	113
21	14	8	60	78	3
22	10	5	62	77	14
23	7	3	66	39	18
24	13	0	67	28	7
27	71	13	70	66	20
28	17	7	72	12	0
29	24	4	73	6	0
31	43	27	74	68	27
32	59	23	75	47	11
33	23	9	76	68	36
34	59	26	77	61	28
35	33	9	78	71	26
36	31	17	79	39	32
39	27	11	80	130	26
40	66	23	82	54	27
41	37	10	83	20	10
42	96	23	84	11	7

Bus	Pload	Qload
85	24	15
86	21	10
88	48	10
90	163	42
91	10	0
92	65	10
93	12	7
94	30	16
95	42	31
96	38	15
97	15	9
98	34	8
99	42	0
100	37	18
101	22	15
102	5	3
103	23	16
104	38	25
105	31	26
106	43	16
107	50	12
108	2	1
109	8	3
110	39	30
112	68	13
113	6	0
114	8	3
115	22	7
116	184	0
117	20	8
118	33	15

References

[1] A. Pal, P. Chatterjee, J. S. Thorp, and V. A. Centeno, "On-line calibration of voltage transformers using synchrophasor measurements," *IEEE Trans. Power Del.*, vol. 31, no. 1, pp. 370-380, Feb. 2016.

[2] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Trans. Power Syst.*, 31(5), pp. 3864-3872, Sep.2016.

[3] P. Kaur, M. Jaiswal, and P. Jaiswal, "Review and analysis of voltage collapse in power system", *Int. J. Scientific and Research Publications*, vol. 2, iss. 1, Jan. 2012.

[4] C. W. Taylor, "Power system voltage stability", New York, NY, USA: McGraw-Hill, 1994.

[5] P.-A. Lof, G. Anderson, and D. Hill, "Voltage stability indices for stressed power systems," *IEEE Trans. Power Syst.*, vol. 8, no. 1, pp. 326-335, Feb. 1993.

[6] FERC and NERC staff "Arizona-Southern California outages on September 8, 2011, causes and recommendations," *Federal Energy Regulatory Commission* and *North American Electric Reliability Corporation*, April 2012. [Online] Available: https://www.ferc.gov/legal/staff-reports/04-27-2012-ferc-nerc-report.pdf

[7] F Galvan, S Mandal, and M Thomas, "Phasor Measurement Units (PMU) instrumental in detecting and managing the electrical island created in the aftermath of Hurricane Gustav", *Power Syst. Conf. Expo.*, pp. 1 – 4, Mar. 15-18, 2009.

[8] F. Galvan C. H. Wells "Detecting and managing the electrical island created in the aftermath of Hurricane Gustav using Phasor Measurement Units (PMUs)", *in Proc. Transm. Distrib. Conf. Expo*. pp. 1-5, 2010.

[9] Z. Lin, et al., "Application of wide area measurement systems to islanding detection of bulk power systems", *IEEE Trans. Power Syst.*, vol. 28, no. 2 pp. 2006-2015, 2013.

[10] X. Liu, D. M. Laverty, R. J. Best, K. Li, D. J. Morrow, and S. S. McLoone, "Principal compoenent analysis of wide-area phasor measurements for islanding detection a geometric view," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 976-985, Apr. 2015.

[11] Y. Guo, K. Li, D. M. Laverty, and Y. Xue, "Synchrophasor based island detection for distributed generation systems using systematic principal component analysis," *IEEE Trans. Power Del.*, vol. 30, no. 6, pp. 2544-2552, Dec. 2015.

[12] Y. Zhu, R. Azim, H. A. Saleem, K. Sun, D. Shi, and R. Sharma, "Microgrid security assessment and islanding control by Support Vector Machine," in *Proc. IEEE Power & Energy Soc. Gen. Meeting*, Denver, CO, pp. 1-5, 2015.

[13] R. Azim, Y. Zhu, H. A. Saleem, K. Sun, F. Li, D. Shi and R. Sharma, "A decision tree based approach for microgrid island detection," in *Proc. Innov. Smart Grid Technol. Conf.*, Washington, DC, pp. 1-5, Feb. 2015.

[14] R. Azim *et al*., "A comparative analysis of intelligent classifiers for passive islanding detection in microgrids," in *Proc. IEEE Eindhoven PowerTech*, Eindhoven, pp. 1-5, Sep. 2015.

[15] T. Ohno, T. Yasuda, O. Takahashi, M. Kaminaga, and S. Imai, "Islanding protection system based on synchronized phasor measurements and its operational experiences", in *Proc. IEEE Power Energy Soc. Gen. Meeting*, pp. 1-5, Aug. 2008.

[16] "IEEE standard for synchrophasor measurements for power systems-amendment 1: Modifications of selected performance requirements," *IEEE Std C37.118.1a-2014 (Amendment to IEEE Std C37.118.1-2011)*, pp. 1-25, Apr. 2014.

[17] J. Zhao, et al., "Impact of measurement errors on synchrophasor applications", in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Chicago, 2017.

[18] S. Wang, E. Farantatos, and K. Tomsovic, "Wind turbine generator modeling considerations for stability studies of weak systems", North American Power Symposium (NAPS), 2017

[19] N.W. A. Lidula, and A. D. Rajapakse, "A pattern recognition approach for detecting power islands using transient signals-Part I:design and implementation," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 3070-3077, Oct. 2010.

[20] T. Funabashi, K. Koyanagi, and R. Yokoyama, "A review of islanding detection methods for distributed resources," in *Proc. IEEE Bologna Power Tech Conf.*, vol. 2, pp. 6–11, 2003.

[21] "Using Synchrophasor Data for Phase Angle Monitoring", North American SynchroPhasor Initiative, May 2016. [Online] Available: https://www.naspi.org/sites/default/files/reference_documents/0.pdf

[22] P. Chatterjee, A. Pal, J. S. Thorp, J. De La Ree, and V. A. Centeno, "Error reduction of phasor measurement unit (PMU) data considering practical constraints," *IET Gener., Transm. Distrib*., vol. 12, no. 10, pp. 2332-2339, May 2018.

[23] F. Gao, J. S. Thorp, A. Pal, and S. Gao, "Dynamic state prediction based on Auto-Regressive (AR) model using PMU data," in *Proc. IEEE Power Energy Conf. Illinois (PECI)*, Champaign, IL, pp. 1-5, 24-25 Feb. 2012.

[24] A. Pal, "Effect of different load models on the three-sample based quadratic prediction algorithm," in *Proc. IEEE Power Eng. Soc. Conf. Innov. Smart Grid Technol.*, Washington D.C, pp. 1-5, 18-20 Feb. 2015.

[25] "WECC Wind Power Plant Power Flow Modeling Guide", WECC Modeling and Validation Work Group, May 2008. [Online] Available: https://www.wecc.biz/Reliability/WECC%20Wind%20Plant%20Power%20Flow%20Modeling%20Guide.pdf

[26] R. Sen Biswas, and A. Pal, "A robust techno-economic analysis of PMU-based islanding detection schemes," in *Proc. 2017 IEEE Texas Power Energy Conf. (TPEC)*, College Station, TX, pp. 1-6, 2017.

[27] S. Adari, and B. R. Bhalja, "Islanding detection of Distributed Generation using Random Forest technique," in *Proc. 2016 IEEE 6th Int. Conf. Power Syst. (ICPS)*, New Delhi, pp. 1-6, 2016.

[28] R. Sun, and V. A. Centeno, "Wide area system islanding contingency detecting and warning scheme," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2581-2589, Nov. 2014.

[29] J. D. L. Ree, V. Centeno, J. S. Thorp, and A. G. Phadke "Synchronized phasor measurement applications in power systems", *IEEE Trans. Smart Grid*, vol. 1, no. 1 pp. 20-27 ,Jun. 2010.

[30] A. Pal, G. A. Sanchez-Ayala, J. S. Thorp, and V. A. Centeno, "A community based partitioning approach for phasor measurement unit placement in large systems," *Electr. Power Compon. Syst.*, vol. 44, no. 12, pp. 1317-1329, Jun. 2016.

[31] M. Asprou, and E. Kyriakides, “Identification and Estimation of Erroneous Transmission Line Parameters Using PMU Measurements”, *IEEE Trans. on Power Del.*, vol. 32, no. 6, pp. 2510-2519, Dec. 2017.

[32] A. Pal, A. K. S. Vullikanti, and S. S. Ravi, “A PMU placement scheme considering realistic costs and modern trends in relaying,” *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 552-561, Jan. 2017.

[33] R. Setola, S. D. Porcellinis, and M. Sforna, “Critical infrastructure dependency assessment using the input–output inoperability model,” *Int. J. Crit. Infr. Prot.*, vol. 2, no. 4, pp. 170-178, Dec. 2009.

[34] CNN, “Sniper attack on California’s power grid may have been an insider,” DHS says (CNN Tech) [Online]. Available: https://money.cnn.com/2015/10/16/technology/sniper-power-grid/index.html

[35] Transformer explodes into fireball at SRP substation in Avondale (azfamily) [Online]. Available: http://www.azfamily.com/story/32111291/transformer-explodes-into-fireball-at-srp-substation-in-avondale

[36] T. Hong, D. Deswal, and F. de León, “An online data-driven technique for the detection of transformer winding deformations,” *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 600-609, Apr. 2018.

[37] NASPI, “Synchrophasor fact sheet”, Oct. 2014 [online] Available: https://www.naspi.org/sites/default/files/reference_documents/33.pdf?fileID=1326A

[38] F. Aminifar, M. Fotuhi-Firuzabad, A. Safdarian, A. Davoudi and M. Shahidehpour, “Synchrophasor measurement technology in power systems: panorama and state-of-the-art,” *IEEE Access*, vol. 2, pp. 1607-1628, 2014.

[39] K. D. Jones, A. Pal, and J. S. Thorp, “Methodology for performing synchrophasor data conditioning and validation,” *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1121-1130, May 2015.

[40] J. M. Lim, and C. L. DeMarco, “SVD-based voltage stability assessment from phasor measurement unit data,” *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2063-2072, Jul. 2017.

[41] T. Wang, A. Pal, J. S. Thorp, Z. Wang, J. Liu, and Y. Yang, “Multi-polytope based adaptive robust damping control in power systems using CART,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2063-2072, Jul. 2015.

[42] T. Wang, A. Pal, J. S. Thorp, and Y. Yang, “Use of polytopic convexity in developing an adaptive inter-area oscillation damping scheme,” *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2509-2520, Jul. 2017.

[43] M. Laifenfeld, and A. Trachtenberg, “Identifying codes and covering problems,” *IEEE Trans. Inform. Theory*, vol. 54, no. 9, pp. 3929-3950, Sep. 2008.

[44] Y. Xu and R. Xiao, “Solving the Identifying Code Problem by a Genetic Algorithm,” *IEEE Trans. Power Syst., Man, and Cybernetics - Part A: Systems and Humans*, vol. 37, no. 1, pp. 41-46, Jan. 2007.

[45] A. Pinceti, O. Kosut, and L. Sankar, “Data-Driven Generation of Synthetic Load Datasets Preserving Spatio-Temporal Features”, accepted in *Proc. 2019 IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Atlanta, GA.

[46] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, “Detecting false data injection attacks on power grid by sparse optimization,” *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 612–621, Mar. 2014.

[47] M. Wang, P. Gao, S. G. Ghiocel, J. H. Chow, B. Fardanesh, G. Stefopoulos, and M. P.

Razanousky, “Identification of "unobservable" cyber data attacks on power grids,” in *Proc. 2014 IEEE Int. Conf. on Smart Grid Comm. (SmartGridComm)*, pp. 830–835, Nov. 2014.
[48] P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. P. Razanousky, “Identification of successive "unobservable" cyber data attacks in power systems through matrix decomposition,” *IEEE Trans. Signal Process.*, vol. 64, no. 21, pp. 5557–5570, Nov 2016.
[49] B. Xu, and A. Abur, “Observability analysis and measurement placement for systems with PMUs,” in *Proc. IEEE PES Power Syst. Conf. Expos.*, pp. 943–946 vol. 2, Oct. 2004.
[50] J. Liang, L. Sankar, and O. Kosut, “Vulnerability analysis and consequences of false data injection attack on power system state estimation,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3864–3872, 2016.
[51] G. Hug, and J. Giampapa, “Vulnerability Assessment of AC State Estimation With Respect to False Data Injection Cyber-Attacks,” *IEEE Trans. Smart Grid*, pp. 1362-1370, 2012.
[52] A. Tiranuchit, and R. Thomas, “A posturing strategy against voltage instabilities in electric power systems,” *IEEE Trans. Power Syst.*, vol. 3, no. 1, pp. 87–93, 1988.
[53] Y. Wang, L. C. P. Da Silva, W. Xu, and Y. Zhang, “Analysis of ill-conditioned power-flow problems using voltage stability methodology,” in *Proc. IEEE Proc. Gener., Transm. Distrib.*, vol. 148, no. 5, pp. 384–390, 2001.
[54] A. Tiranuchit, L. M. Ewerbring, R. A. Duryea, R. Thomas, and F. Luk, “Towards a computationally feasible on-line voltage instability index,” *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 669–675, 1988.
[55] “Voltage stability assessment: Concepts, practices and tools,” in *IEEE-PES Power System Dynamic Performance Committee and Power Systems Stability Subcommittee*, Special Publication, Piscataway, NJ, USA, Aug. 2002.
[56] L. Balzano, R. Nowak, and B. Recht, “Online identification and tracking of subspaces from highly incomplete information,” in *Proc. 2010 48th Annual Allerton Conf. Comm., Control, and Computing*, pp. 704–711, 2010.
[57] Y. Saad, “Numerical Methods for Large Eigenvalue Problems,” 2011. [Online]. Available: http://www.eia.gov/electricity/annual/pdf/tablees1.pdf
[58] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. Mckenney, S. Ostrouchov, and D. Sorensen, “LAPACK user’s guide,” 1992.
[59] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, “An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices,” *SIAM J. Sci. Computing*, vol. 14, no. 1, pp. 137–158, Jan. 1993. [Online]. Available: http://dx.doi.org/10.1137/0914009
[60] L. Xie, Y. Chen, and P. Kumar, “Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis,” *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, 2014.
[61] M. Perninge, V. Knazkins, M Amelin, and L. Söder, “Modeling the electric power consumption in a multi-area system”, *European Trans. Electric Power*, pp. 413-423, Jan. 2011.
[62] J. C. Hull, “Options, Futures, and Other Derivatives”, *5th Ed. Prentice Hall*, 2003.
[63] P. J. Brockwell, and R. A. Davis, “Introduction to Time series and forecasting”, *2nd Ed. Springer*, 2002.

[64] B. Bibby, and M. Sorensen, “Martingale estimation functions for discretely observed diffusion processes,” *Int. J. Statistical Institute. (ISI) and Bernoulli Society for Mathematical Statistics and Probability*, vol. 1, no. 1/2, pp. 17–39, 1995.