# Security Analysis of Interdependent Critical Infrastructures: Power, Cyber and Gas

by

Mahdi Jamei

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved August 2018 by the
Graduate Supervisory Committee:

Anna Scaglione, Chair
Raja Ayyanar
Kory W. Hedman
Oliver Kosut

ARIZONA STATE UNIVERSITY

December 2018

ABSTRACT

Our daily life is becoming more and more reliant on services provided by the infrastructures power, gas , communication networks. Ensuring the security of these infrastructures is of utmost importance. This task becomes ever more challenging as the inter-dependence among these infrastructures growsand a security breach in one infrastructure can spill over to the others. The implication is that the security practices/analysis recommended for these infrastructures should be done in coordination. This thesis, focusing on the power grid, explores strategies to secure the system that look into the coupling of the power grid to the cyber infrastructure, used to manage and control it, and to the gas grid, that supplies an increasing amount of reserves to overcome contingencies.

The first part (Part I) of the thesis, including chapters 2 through 4, focuses on the coupling of the power and the cyber infrastructure that is used for its control and operations. The goal is to detect malicious attacks gaining information about the operation of the power grid to later attack the system. In chapter 2, we propose a hierarchical architecture that correlates the analysis of high resolution Micro-Phasor Measurement Unit ($\mu$PMU) data and traffic analysis on the Supervisory Control and Data Acquisition (SCADA) packets, to infer the security status of the grid and detect the presence of possible intruders. An essential part of this architecture is tied to the analysis on the $\mu$PMU data. In chapter 3 we establish a set of anomaly detection rules on $\mu$PMU data that flag "abnormal behavior". A placement strategy of $\mu$PMU sensors is also proposed to maximize the sensitivity in detecting anomalies. In chapter 4, we focus on developing rules that can localize the source of an events using $\mu$PMU to further check whether a cyber attack is causing the anomaly, by correlating SCADA traffic with the $\mu$PMU data analysis results. The thread that unifies the data analysis in this chapter is the fact that decision are made without

fully estimating the state of the system; on the contrary, decisions are made using a set of physical measurements that falls short by orders of magnitude to meet the needs for observability. More specifically, in the first part of this chapter (sections 4.1-4.2), using $\mu$PMU data in the substation, methodologies for online identification of the source Thevenin parameters are presented. This methodology is used to identify reconnaissance activity on the normally-open switches in the substation, initiated by attackers to gauge its controllability over the cyber network. The applications of this methodology in monitoring the voltage stability of the grid is also discussed. In the second part of this chapter (sections 4.3-4.5), we investigate the localization of faults. Since the number of PMU sensors available to carry out the inference is insufficient to ensure observability, the problem can be viewed as that of under-sampling a "graph signal"; the analysis leads to a PMU placement strategy that can achieve the highest resolution in localizing the fault, for a given number of sensors. In both cases, the results of the analysis are leveraged in the detection of cyber-physical attacks, where $\mu$PMU data and relevant SCADA network traffic information are compared to determine if a network breach has affected the integrity of the system information and/or operations.

In second part of this thesis (Part II), the security analysis considers the adequacy and reliability of schedules for the gas and power network. The motivation for scheduling jointly supply in gas and power networks is motivated by the increasing reliance of power grids on natural gas generators (and, indirectly, on gas pipelines) as providing critical reserves. Chapter 5 focuses on unveiling the challenges and providing solution to this problem.

DEDICATION

To my beloved parents, Gholamreza and Fatemeh

ACKNOWLEDGMENTS

Now that my journey at ASU has come to an end, I cannot express enough how grateful I am to be around people who mean a lot to me. I owe a big thank to my advisor Prof. Anna Scaglione who tirelessly helped me to find my path in research and life. You have always believed in me and taught me how to never give up and to do my best to achieve my goals. I would like to thank my collaborators Dr. Sean Peisert, Dr. Reinhard Gentz, Mr. Ciaran Roberts, Dr. Emma Stewart, Mr. Alex McEachern, Mr. Chuck McParland, Mr. Eran Schweitzer, Dr. Teklemariam Tesfay for your great work and helpful comments. I would also like to thank my committee members Dr. Raja Ayyanar, Dr. Kory Hedman, and Dr. Oliver Kosut for all the valuable comments and guidance you gave me throughout my PhD. Life as a PhD student has a lot of ups and downs. Being away from my family and country for a long time did not make it any easier but I have known that I have friends who are always there for me. Thank you Lorenzo, Kari, Eran, Reinhard, To, Raksha, Bita, Teklemariam, Nurullah, Nikhil. Also, thank to my friends Soroush, Hadi, Amin, Saman, Arman, Behdad, and Sadra. You all defined for me what a true friendship means. A special thank goes to my sisters Shabnam and Maryam and my brother-in-law Saeed who have always supported me in every aspect of my life. I am very blessed to have you in my life. I dedicated this thesis to my lovely parents Gholamreza Jamei and Fatemeh Tazareh. I do not think words can describe how I feel about you but you are the reason of everything good in my life. Finally, to my love Tina who has been a great support for me and has brought so much happiness and joy into my life. Thank you all.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF ALGORITHMS

# Nomenclature

### Parameters

$\mathbf{C}_g^x$ Production cost vector of generator $g$ corresponding to cost function's segments.

$\mathbf{C}_g^r$ Generation reserve cost vector of generator $g$.

$c_g^{\text{nl}}$ No-load cost of generator $g$.

$c_g^{\text{st}}$ Start-up cost of generator $g$.

$\mathbf{S}_g$ Ramp-rate limit vector of generator $g$.

$\underline{\mathbf{X}}_g$ Generator $g$ minimum generation limit vector corresponding to cost function's segments.

$\overline{\mathbf{X}}_g$ Generator $g$ maximum generation limit vector corresponding to cost function's segments.

$UT_g$ Minimum up-time of generator $g$.

$DT_g$ Minimum down-time of generator $g$.

$P_g^{max/min}$ Generator $g$ maximum/minimum generation limit.

$Pl_b^{max,a/c}$ Rate-A/Rate-C power flow limit of line $b$.

$PF_{b,n}^R$ Power transfer distribution factor on line $b$ from bus $n$, with reference bus $R$.

$n(c)$ Bus number of the generator $c$.

$d_n(t)$ Electric load at time $t$ of bus $n$.

$\beta_{el/gas}$ Electric power/Gas reserve factor.

$d_w$ Natural-gas production cost at well/storage $w$.

$d_w^{res}$ Natural-gas reserve cost at well/storage $w$.

$G_w^{max/min}$ Maximum/minimum gas well production limit.

$\pi_i^{max/min}$ Node $i$ maximum/minimum gas pressure limit.

$\pi_{\text{pipe}}^{max}$ Maximum tolerable pressure by pipeline.

$\rho_k(t)$ Base-case compressor ratio at time $t$ of pipeline $k$.

$L_i(t)$ Natural-gas load at time $t$ of node $i$.

$h_k$ Gas pipeline $k$ constant.

$\eta_g$ Conversion rate for gas-fired generator, $g$, from electric output to natural-gas input.

**Sets and Indices**

$\mathcal{V}_e$ Set of electric grid buses.

$\mathcal{E}_e$ Set of electric grid lines.

$\mathcal{G}$ Set of generators.

$\mathcal{G}^a$ Generators subset with automatic generation control.

$\mathcal{G}^f$ Subset of fast-start generators.

$\mathcal{G}(n)$ Subset of generators at bus $n$.

$\mathcal{Q}$ Set of segments for generators' production cost curve.

$\mathcal{S}(i)$ Subset of gas-fired generators at gas node $i$.

$\mathcal{V}_g$ Set of gas grid nodes.

$\mathcal{E}_g$ Set of gas pipelines.

$\mathcal{E}_g^a$ Subset of active gas pipelines.

$\mu_i^{+/-}$ Subset of pipelines with node $i$ as their "to"/"from" node.

$\mathcal{W}$ Set of natural gas wells/storage.

$\mathcal{W}(i)$ Subset of natural gas wells/storage at node $i$.

$\mathcal{W}^s$ Subset of natural gas slack node.

$\mathcal{T}$ Set of time epochs.

**Variables**

$\mathbf{X}_g(t)$ Power production vector at time $t$ of generator $g$ for different segments.

$\mathbf{R}_g(t)$ Generation reserve vector at time $t$ of generator $g$.

$u_g(t)$ Commitment status at time $t$ of generator $g$.

$v_g(t)$ Start-up status at time $t$ of generator $g$.

$r_g^c(t)$ Reserve at time $t$ of generator $g$ for generator $c$ contingency.

$\pi_i(t)$ Gas pressure at time $t$ of node $i$.

$G_w(t)$ Gas production at time $t$ of gas well/storage $w$.

$rg_w(t)$ Gas reserve at time $t$ of gas well/storage $w$.

$f_k(t)$ Gas flow at time $t$ of pipeline $k$.

$z_k(t)$ Gas flow direction indicator at time $t$ of pipeline $k$.

$\pi_i^c(t)$ Gas pressure at time $t$ of node $i$ for generator $c$ contingency.

$rg_w^c(t)$ Reserve at time $t$ of gas well/storage $w$ for generator $c$ contingency.

$f_k^c(t)$ Gas flow at time $t$ of pipeline $k$ for generator $c$ contingency.

$z_k^c(t)$ Gas flow direction indicator at time $t$ of pipeline $k$ for generator $c$ contingency.

# CHAPTER 1

# INTRODUCTION

Designing secure practices and schedules for the operation of the critical infrastructures cannot be achieved if one looks at these infrastructures as independent entities. In fact, as the inter-dependency between these infrastructures increases, any practice and schedule should be in coordination with the coupled entities. We focus on the power system as the main critical infrastructure of interest and study the security issues that might arise due the coupling of this grid with the cyber network and the gas grid.

Power grid operation has heavily become reliant on the communication network for information exchange and control of the assets. Adding this cyber layer on the top of the physical power grid has introduced vulnerabilities in the grid since the common industrial protocols that are used are not secure in design. Therefore, breaking into this network by hackers means that the underlying physical grid is at a high risk to be compromised or to be disrupted from doing its normal operation. For example, the results of a successful cyber attack on the Ukraine power grid that occurred in 2015 were:

- 30 substations were switched off,
- about 230,000 people did not have electricity for a period from 1 to 6 hours.

Part I of this thesis that includes chapters 2 through 4 investigates the security of the power distribution grid due to the coupling with the cyber network. More specifically, an architecture is proposed in chapter 2 that receives physical data from Micro Phasor Measurement Units ($\mu$PMUs) and cyber packets sniffed by the packet

sniffers from the Supervisory Control and Data Acquisition (SCADA) to analyze and cross-correlate the data based on a set of defined rules and decide whether an intrusion has happened. Since this architecture heavily relies on the analysis of the data from $\mu$PMUs for intrusion detection, a set of anomaly detection rules for distribution grid are proposed in 3. Having an event detected is necessary but it is not sufficient most of the times. In chapter 4, we focus on methods to localize the source of an event. The results of the $\mu$PMU-based event source identification can then be compared to the corresponding mirrored SCADA to check whether the two sources of information are in agreement. In particular, we first discuss a "reconnaissance activity" identification method for normally-open switches in the substation (sections 4.1-4.2). In the next part of this chapter (sections 4.3-4.5), we investigate the localization of faults using $\mu$PMU data, in an operating regime where the number of $\mu$PMU sensors available to carry out the inference is insufficient to have observability.

The second part of this thesis investigates the coupling between the power and gas grid and analyzes the security challenges that might arise due to this coupling. The share of the gas-fired generators to produce electricity has been growing over the past few years. Figure. 1.0.1 shows the share of each of the resources in producing electricity in the US. It can be observed that natural-gas fired generators produce a significant portion of electricity in the US. This observation has been quantified in Figure. 1.0.2 and proves the point. It implies that the power grid would heavily rely on the gas network to be able to produce enough electricity to serve the customers. Therefore, a secure operation of the power network in terms of adequacy of generation and reliability of the service cannot be guaranteed unless the schedule is in coordination with the constraints of the gas network. Part II of this thesis that includes chapter 5 is dedicated to investigate the impact of uncoordinated scheduling of power and gas grid on security of the power grid and to propose a methodology

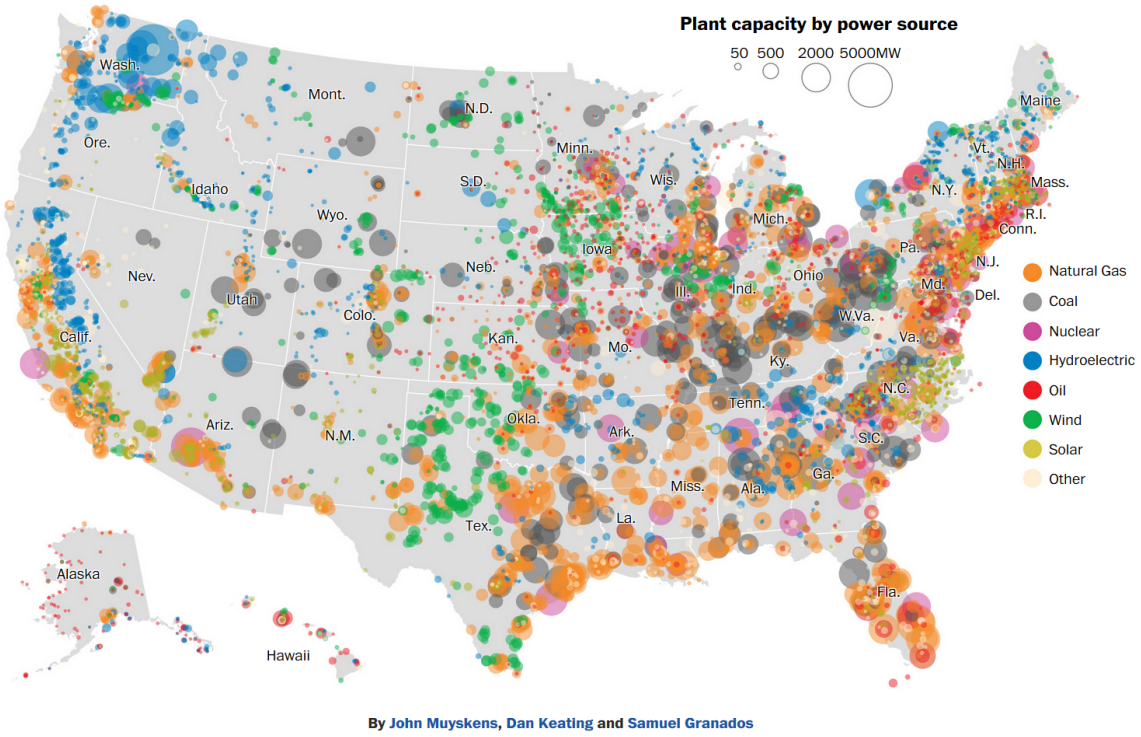



Figure 1.0.1: How the US Generates its Electricity?

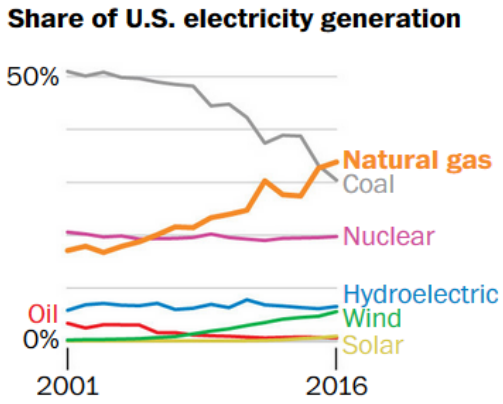



Figure 1.0.2: Share of U.S. Electricity Generation. source: U.S. Energy Information Agency.

for security-constrained scheduling of power and gas grid.

The breakdown of introduction of each of these two parts are given in the following:

## 1.1 Part I-Cyber-Physical Security of Power Distribution Grid

According to the U.S. Department of Homeland Security (DHS) reports, the sophistication and frequency of the attacks on the power grid are increasing [tec(2015)]. For example, a recent report on the Ukraine attack [cbc news Technology and Science(2016)] showed how a failure in the communication network security resulted in significant power outages. The Stuxnet malware or the sewage spill incident at the Maroochy Water Station, due to the wireless attack, are other examples showcasing how a misinformed control system can lead to catastrophic consequences.

Cyber security for energy delivery systems has, until now, focused primarily on the transmission grid and on securely transferring bits of information about the condition of power-grid elements (e.g., "Is this switch open or closed?", "Which tap is selected on this transformer?") and preventing unauthorized access to sensor and control packets. Once that access has been gained, there is little remediation action for the power grid, other than a communications blackout and manual fieldwork. The industry is seeking new approaches to this problem, also focusing on understanding security at the distribution level, in anticipation to a growth in automation.

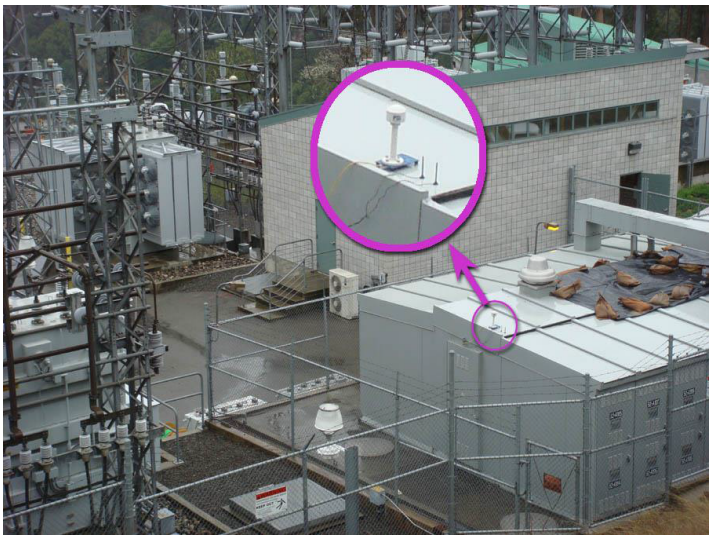
While transmission grids states have been tightly monitored and their behavior at the physical level is reasonably well understood, the operators have been largely blind towards the real time condition of the distribution grid. Hence, in tandem with the effort of gaining situational awareness on the security of the system there is a growing need and interest in the deployment of sensors, like the $\mu$PMUs that can capture the state at the distribution level [Eto *et al.*(2015)]. These devices, recently developed by PSL [Power Standards Lab(2016)], address both the technical and economic barriers limiting the deployment of conventional PMUs, which are aimed at the transmission grid, for the distribution level [von Meier *et al.*(2014)].

Fig. 1.1.1a shows a sample $\mu$PMU device installed in the partner utility grid. These devices sample at a rate of 120 samples/sec the three phase voltage and current phasors. In comparison to SCADA that samples power flow and power injections every 3-5 minutes, $\mu$PMU provides significantly more information, and often nuances which are missed in SCADA data, as illustrated in the example in Fig. 1.1.1b. In this example, the magnitude of the current measured by the SCADA meter is missing an overcurrent event that $\mu$PMU could capture. This information may prove critical in identifying cyber-attacks.

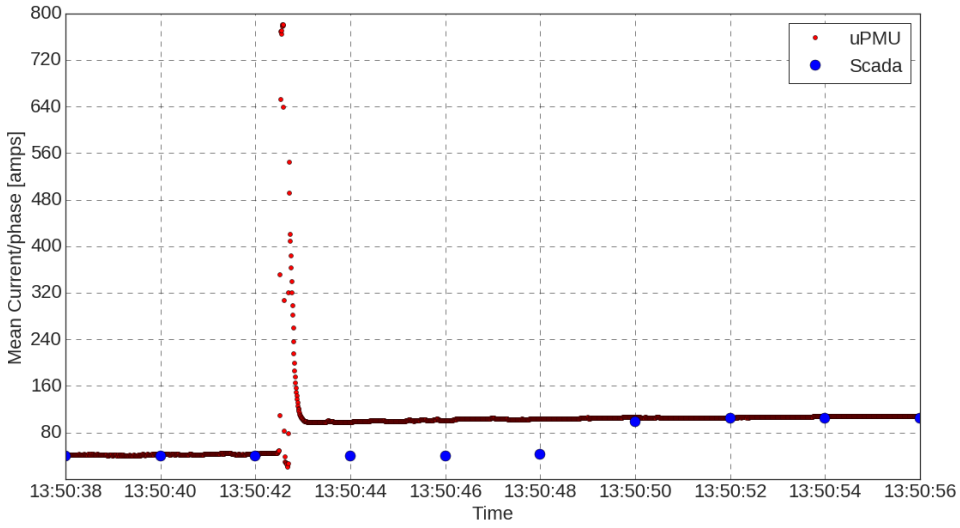
However, visualizing and interpreting raw sensor streams can be overwhelming for DSOs, considering the large quantity of data that could flow in from different parts of the grid [Kezunovic *et al.*(2013)]. Therefore, it is essential to mine the data collected with analytic tools that can derive informative measurements and form automated reports.

In chapter 2 of this dissertation, we propose an architecture for analysis of the $\mu$PMU data and sniffed SCADA packets in a hierarchical manner to alarm the operator about a possible cyber attack on the grid [Jamei *et al.*(2016)]. Due to the critical role of $\mu$PMUs in this framework, we establish a set of rules in chapter 3 that are implemented in this architecture and look at the $\mu$PMU data at different levels of aggregation to flag any anomalies [Jamei *et al.*(2017b), Jamei *et al.*(2017c)]. These anamolous segments of data are filtered out for further co-analysis with the cyber packets to distinguish between a natural anomaly and a malicious one.

In chapter 4, a set of rules are developed to localize an event and be able to verify whether the SCADA-reported data is in agreement with the $\mu$PMU-based data analysis. In particular, in sections 4.1-4.2 we fist propose a methodology for online estimation of the Thevenin source parameters using $\mu$PMU data in balanced and unbalanced grid [Jamei *et al.*(2017a)]. An application that relates this part of the



(a)




(b)

Figure 1.1.1: (a) $\mu$PMU Instrument from Power Sensors Ltd, (b) Measurement Comparison of $\mu$PMU and SCADA

chapter to the cyber-physical security of distribution grid is when these parameters are used to infer the status of the normally-open switch at the substation. The point here is that an attacker may launch "reconnaissance activities" in the grid to gauge his controllability over the assets. A point of interest for this type of attack

would be the manipulation of the normally-open switch at a substation between two parallel transformers since it does not interrupt/change the delivery of the power to downstream feeder. Using $\mu$PMU data to estimate the Thevenin source parameters, we can infer about the status of these switches and compare it to what SCADA data represent to understand a possible breach in the system. However, the contribution of this work goes beyond that and includes the application of these parameters in monitoring the voltage stability index.

In the next part of this chapter (sections 4.3-4.5), we investigate the localization of faults using $\mu$PMU data, in an operating regime where the number of sensors available to carry out the inference is insufficient to have observability [Jamei *et al.*(2018a), Jamei *et al.*(2018c)]. The problem can be viewed as that of sampling a graph signal and the analysis leads to a $\mu$PMU placement strategy that can achieve the highest resolution in localizing the fault, for a given number of sensors. In addition to verifying our theory, in our numerical analysis our strategy is leveraged in the detection of cyber-physical attacks, where $\mu$PMU data and relevant SCADA network traffic information are compared to determine if a network breach has affected the integrity of the system information and/or operations.

### *1.1.1 Prior Work and Contributions*

**Intrusion Detection Architecture:** The first steps in adding security in operational environments are typically to deploy firewalls and device-level authentication. Encryption is often also added to enhance confidentiality and integrity of the message content. Another common security mechanism on computer networks is intrusion detection, in which network traffic is monitored and analyzed to detect activities that either fit into a "known bad" category or deviate in a statistically significant way from "normal." The Tofino Security Appliance, the Digital Bond Quickdraw SCADA

intrusion detection system, the Radiflow Secure Gateway, the Bro Security Framework [Berthier *et al.*(2010)] are all examples of network intrusion detection systems (IDSs) that can be applied to control systems.

Numerous examples have shown that all of these methods leave significant gaps in security and safety [Slay and Miller(2008)]. It has been recognized that one of the reasons for this is that most of these security methods are divorced from the knowledge of the physics of the system, its safe operations and limits, and its current physical operating point. This gap was recognized early on by e.g. [Cárdenas *et al.*(2011)]. Some of our own previous work for monitoring SCADA traffic expanded the notion of intrusion detection by leveraging the laws of physics governing the grid and imposing them as security constraints [McParland *et al.*(2014), Koutsandria *et al.*(2015)]. Nonetheless, these methods also remain blind to more sophisticated attacks. One reason is that the data coming from SCADA systems are not updated with high frequency, so events causing many changes in a short period of time can be missed. In addition, attackers can inject false data at the device level, thus evading detection by the IDS.

An underpinning of the activity described in our work is to design a security framework in anticipation of the impending move towards Advanced Distribution Management System (ADMS). ADMS limits the need for direct human intervention, and when working properly, its functionalities enhance the reliability and safety of the system. While ADMSs are developed with careful consideration for safe physical operation, a number of their features make them uniquely vulnerable to cyber-attacks [Tesfay *et al.*(2014)]. In ADMS, a SCADA network is responsible to collect the information from field devices (e.g. switches, meters ...), and send back the according control commands. The presence of such network opens up a large attack surface. What makes the case even more challenging is that ADMS is an integrated network,

so failures in one section could cascade into a large and widespread series of events. In particular, communication that lacks end-to-end security can permit difficult-to-detect interference between sub-systems that could cause them to function in ways that threaten the safety and reliability of the power grid. Additionally, unlike modern computer systems that are upgraded every three to six years, many of the CPSs, such as electric power system equipped with ADMS are amalgam of decades-old and very new components, operating side by side often with inconsistent operating controls, algorithms and guidelines.

The work in [Mitchell and Chen(2014)] gives a detailed survey on the IDSs that are designed to enhance the security of CPSs. Figure 1.1.2 obtained from this work categorizes different approaches of designing an IDS for a CPSs. Our approach here

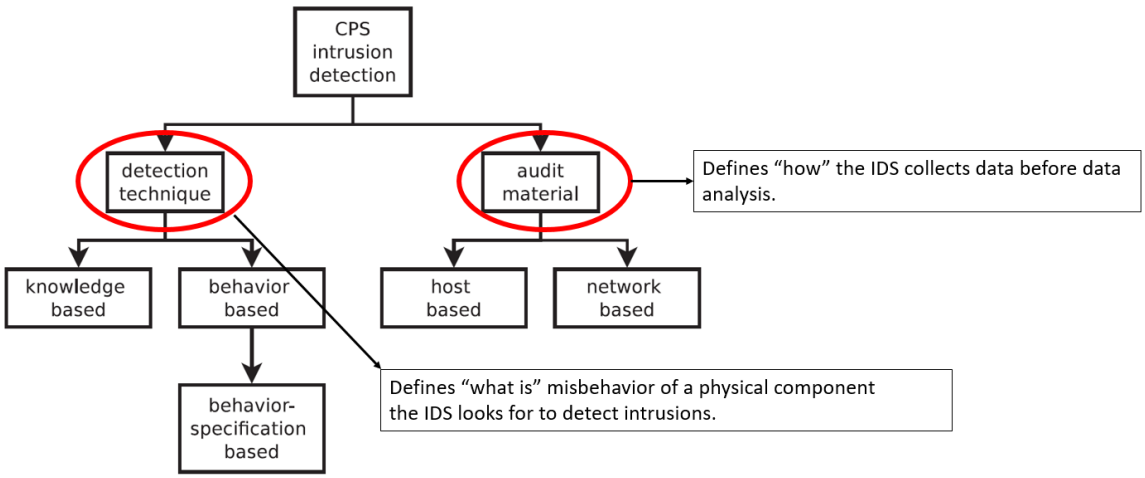



Figure 1.1.2: CPS Intrusion Detection Tree [Mitchell and Chen(2014)]

is a mix of auditing materials at the network level for the SCADA traffic anomaly detection and behavior-specification based techniques, where the legitimate behavior is defined and intrusion is detected when the system departs from this model.

To ensure that ADMS operates and fails in well-understood and controlled ways, one needs to tightly monitor the parts most exposed to an attack. The level of

monitoring in CPSs is often a compromise between two competing design mandates: least function (design systems as simple as possible to perform their grid management functions); and robust monitoring (incorporate high-fidelity system status indicators to enable detection of and response to cyber-security events). Our framework leans towards the second predicament through the integration of $\mu$PMUs information in the security architecture, which provides a clear image of the physical trail left by cyber-physical-attacks.

**$\mu$PMU Data Anomaly Detection:** In the context of gaining awareness using synchrophaor data, recent work has focused on the PMU data utilization at the transmission level to improve the Wide-Area Monitoring, Protection, and Control [Phadke and de Moraes(2008), Terzija *et al.*(2011)]. Distribution grids, however, are still lagging in that respect, since tools for the transmission grid may not be directly applicable to the distribution grid due to a different operation environment, such as load imbalances, untransposed lines, and the existence of single-phase and two-phase laterals. Transmission operations and system wide analysis are concerned with large imbalances in load and generation, and as a result frequency, whereas distribution operations are concerned with localized, but frequent events such as voltage imbalance, overloading, and outage management. Most of the prior research on sensor data analytics (including SCADA and PMU measurements) is concerned with detecting events on the grid transmission level. For example, Pan et al., [Pan *et al.*(2015)] use data mining techniques on PMU measurements and audit logs for event classification. A linear basis expansion for the PMU data is described by Chen et al., [Chen *et al.*(2013), Xie *et al.*(2014)] for event detection application. A similar approach, based on Principal Component Analysis (PCA), is used in [Valenzuela *et al.*(2013), Ge *et al.*(2015)] for event detection and data archival. Allen et al., [Allen *et al.*(2014)] describe the use of voltage phasor angle differences between different

PMU readings to detect events. Biswal et al., [Biswal *et al.*(2016)] use the strongest signatures of event in PMU data for situational awareness enhancement. A study by Brahma et al., [Brahma *et al.*(2016)] describes the real-time dynamic event identification in power system using PMU data based on a data-driven and also physics-based method.

Our approach for event detection using $\mu$PMU data is a combination of the data-driven methods, as well as criteria resulting from analyzing the underlying physical model of the system. Unlike model agnostic machine learning methods that look for statistical anomalies in a feature space that is often heuristic, the *anomalies* that algorithms in this chapter identify are defined in the context of power quality and protection, in addition to what is imposed by the grid governing physical equations (i.e. Kirchhoff Voltage and Current Laws). This, in turn, gives a DSO much greater insight and help in the forensic analysis. Because of the important role played by $\mu$PMU data in the framework, a $\mu$PMU placement optimization with limited number of the sensors is described, to achieve the maximum sensitivity in detecting a change[1].

**Event Localization:** The contributions and the review of the related work of this chapter is divided into two sections, where the first section discusses Thevenin estimation and substation normally-open switch monitoring and the second section discusses the related works in the fault localization domain and cyber-security of the fault detectors.

*Thevenin Estimation:* Many studies focus on estimating Thevenin parameters from field sensors, such as PMUs. Most of these studies are at the transmission level, where the grid is typically balanced. Vu et al. [Vu *et al.*(1999)] used the linear, least-square method to solve for the Thevenin parameters and employed the

[1]Note that PMU placement studies for state estimation are not applicable here since the state is not directly observable by solely depending on the $\mu$PMUs due to the scarcity of sensors.

Thevenin impedance to apparent load impedance ratio (well-known L-index) to monitor the voltage stability margin. Using two consecutive phasor measurements, Smon et al. [Smon *et al.*(2006)] employ Tellegen's theorem and the notion of adjoint networks to estimate the Thevenin circuit and apply this analytic to monitor grid stability. Tsai et al. [Tsai and Wong(2008)] use recursive least-square to estimate Thevenin parameters and track the voltage stability margin. Corsi et al. [Corsi and Taranto(2008)] propose adaptive identification of Thevenin voltage and impedance equivalents with real-time voltage instability monitoring purposes. Wang et al. [Wang *et al.*(2013)] estimate the equivalent Thevenin circuit of generators and incorporate the estimated model to extend the traditional L-index. Yuan et al. [Yuan and Li(2014)] conduct a comparative study between different PMU-based methods of Thevenin equivalents identification in terms of how fast and accurate each method is. Others [Abdelkader and Morrow(2012), Alinejad and Karegar(2016)] identify the adverse effect of quasi steady-state condition in the grid on Thevenin parameter estimation, and propose to estimate the off-nominal frequency and compensate for this phenomena to obtain accurate solutions. The authors in [Parniani *et al.*(2006), Arefifar and Xu(2009)] estimate the Thevenin parameters in a non-linear least-square setting considering the variations of the Thevenin voltage source angle. Using measurements from a load busbar in an unbalanced distribution grid, Hart [Hart(1986)] proposes a method that requires only RMS values of the voltage to estimate the Thevenin circuit in a three-phase, 3- or 4-wire system.

In this chapter, using the synchrophasor voltage and current measurements at the substation, we propose methods for estimating the Thevenin source impedance and voltage corresponding to substations delivering power to an unbalanced and balanced feeder. The advent of $\mu$PMU sensors, which measure the three-phase voltage and current phasor at different points over the distribution grid [von Meier *et al.*(2014)],

enables analytics that are similar to the ones at the transmission level. However, the extension to the distribution grid requires care, as distribution feeders are typically unbalanced. In the unbalanced case, we take advantage of having non-zero negative and zero sequence to estimate not only the voltage and positive sequence impedance but also the zero sequence source impedance. It should be noted that the use of synchrophasor data for Thevenin estimation is a side benefit of PMUs but it does not mean that obtaining required input data for Thevenin estimation is not possible with current facilities other than PMUs.

In the case of balanced grid, we propose a new method that is effective even when the system of equations are close to be rank-deficient. In our proposed approach, only the magnitude of the Thevenin voltage and Thevenin impedance are assumed to be constant over a short window, and no assumption is made about the angle of the Thevenin voltage source. This in fact improves the results as compared to the methods that require the Thevenin voltage angle to be constant over a window of samples; an assumption that is not valid in practice. In this sense, our work is similar to the proposed methods in [Parniani *et al.*(2006), Arefifar and Xu(2009)]. We show how our method is successful in tracking the voltage stability index on a load point at the transmission level. In addition, given growing concerns about the security vulnerabilities of network-connected control systems in the distribution grid, we show, through an example, how our Thevenin estimation can be useful to detect specific types of cyber-attacks in a distribution substation. Specifically, in this chapter, we discuss the detection of certain types of *reconnaissance* activity over SCADA network by attackers—that is, subtle probes and manipulation of network-connected substation components by an attacker to test degree of access and possible broader physical impact on the distribution grid.

*Fault Localization:* PMU-based fault detection and localization is an active area

of research. For example, Ardekanian et al., [Ardakanian *et al.*(2017)] exploit PMU data to detect and localize a change in the admittance matrix of a power grid. Zhou et al., [Zhou *et al.*(2017)] use PMU measurements for event detection in the distribution grid with partial information. Our previous work [Jamei *et al.*(2017b)] proposes a hierarchical architecture as the host for event detection rules in distribution system using PMU data when scarce measurements are available. Farajollahi et al., [Farajollahi *et al.*(2018)] use distribution level PMUs for general event detection and localization by generating an equivalent circuit so that the event can be represented by the voltage and current phasors.

Event detection by itself is necessary but insufficient most of the times and it is important for the grid operators to have knowledge about the source of the event, to be able to locate a faulty section for isolation and service restoration. Tackling the localization problem becomes challenging when one considers in a regime where the sensors are insufficient to have observability, which is usually the case in the distribution feeders. This regime is important in general in an adversarial setting where data injection attacks can leave operators with few trustworthy sensors [Kim and Poor(2011)]. In this case, gaining information from a subset of sensors on what is plausible can be very useful in identifying data stealth attack since all subsets of sensors must agree on a common version of the facts.

Zhu et al., [Zhu *et al.*(1997)] propose an automated fault localization and diagnosis for distribution grids. Using measurements from the substation, the method first find a set of plausible locations for the fault. During the diagnosis, the set of possiblities are ranked for the operator. The effect of the DC componenet in the phsor data and how it can affect momentary fault localization process is invetsigate in the work by Min and Santoso [Min and Santoso(2017)]. Kahyap et al., [Kashyap *et al.*(2015)] implement a fault location and isolation algorithm in a distributed fashion, where the focus is put

on the communication network. Lee [Lee(2014)] uses PMU voltage data to search for a fault in a radial network in a timely manner. A graph marking approach is taken by Dzafic et al., [Dzafic *et al.*(2018)] to spot the location of a fault. There are also a number of non-parametric methods for fault localization that use spatio-temporal patterns of the measurements. For example, Jian et al., [Jiang *et al.*(2014)] extract the time-frequency features of frequency and voltage from a dictionary by *matching pursuit* [Mallat and Zhang(1993)], which is then followed by a clustering method for fault detection. Borghetti et al., [Borghetti *et al.*(2008)] perform wavelet analysis on the voltage waveform generated during a fault-induced transients to obtain the location of a fault in the distribution network. While the exploitation of temporal patterns helps in the localization, they do not provide an understanding on how the performance is affected by the grid parameters and the sensor deployment. In this work we ignore temporal features (which can be always included) and take inspiration from Brahma's work [Brahma(2011)] in using pre and post fault measurements along with the admittance/impedance matrix for fault localization.

Using a statistical approach, we first improve upon Brahma's method and study the fault localization problem with measurements that are too scarce for observability (we refer to it as the *under-sampled grid* regime). Our main contribution is to show that in an *under-sampled grid* faults can be resolved reliably at level of *clusters* that typically form connected sub-graphs. These clusters depend both on the properties of grid admittance matrix as well as the placement of the sensors. As a figure of merit, the *cluster level fault localization* resolution can measured in terms of the maximum size of all clusters in which the graph is divided. Leveraging the insights from our analysis we propose a placement strategy that achieves the highest localization resolution over the grid, by clustering the graph in a number of connected components equal to the number of measurements, with clusters sizes as even as possible. In this

section of the chapter we also compare our method with [Brahma(2011)] to better highlight the benefits of our statistical approach, and showcase its application in the forensic analysis of cyber-physical attacks to a distribution Fault Location, Isolation, and Service Restoration (FLISR) system [Uluski(2012)]. To the best of our knowledge, the connection between the graph clustering and fault localization resolution is new. However, we would like to acknowledge the graph clustering works in the transmission grid (see e.g., [Wang *et al.*(2010)], [Sanchez-Garcia *et al.*(2014)]), in a different context (not for fault localization). The results of this research is presented in:

## 1.2 Part II-Coordinated Security-Constrained Scheduling of Power and Gas Grid:

There has been a significant growth in the share of gas-fired generators (especially combined-cycle plants) for electricity production. The result of this trend is a strong coupling between two of the largest infrastructures in a way that operation of one can directly affect the other [Liu *et al.*(2009), Urbina and Li(2007)]. This becomes even more important in the future, considering that expected gas-fired generation is expected to increase by 230% by 2030 [Gruenspecht(2010)].

This issue is not just confined to the normal operation of the two grids. In fact, coupling becomes more problematic when a contingency occurs in one, and affects the other. Therefore, any security-constrained analysis should consider the interdependency between the coupled systems. Chapter 5 has been dedicated to address these issues and to present a coordinated methodology for secure scheduling of the power and the gas grid [Jamei *et al.*(2018b)].

### *1.2.1 Prior Work and Contributions*

Interest in the coupling between the natural-gas and electric grids is relatively recent. The aspects of this interdependence that have been explored in the literature are briefly reviewed next. An early contribution is [Liu *et al.*(2009)] in which the authors have formulated an electric grid Unit Commitment (UC) problem including the natural gas grid constraints. The optimization is solved from an electric utility perspective, but neither the co-optimization aiming to reach an Optimal Gas Flow (OGF) and Optimal Power Flow (OPF) was considered nor the coordinated N-1 contingency analysis.

The focus of [Damavandi *et al.*(2011)] is incorporating a quasi dynamic model for the gas grid in the UC problem. In [Qadrdan *et al.*(2014)], deterministic and stochastic operating strategies are proposed to deal with an integrated gas and electricity network under wind power forecast uncertainties. The authors in [Correa-Posada and Sanchez-Martin(2014)] present a security-constrained optimal power and natural gas flow. However, the spinning and non-spinning reserves, that are critical in studies concerning gas-fired generators, are not considered. Also, the contingency analysis is limited to the N-1 gas pipeline contingency. The authors in the paper simulate only the contingencies for the pipelines in a loop, so that the grid stays connected even after losing a pipeline. The issue is that the gas grid is largely radial, with very few loops, and clearly the standard N-1 analysis is not well matched to radial pipelines systems. The cascading effects of the gas-fired generators on the pressure fluctuations in the gas network is investigated in [Chertkov *et al.*(2015b), Chertkov *et al.*(2015a)]. Different coordination scenarios of gas and electric grid are proposed in [Zlotnik *et al.*(2017)], using new methods to control gas flow dynamics to schedule the day-ahead generators dispatch and compressors operation.

The contributions in this work are as follows: First, a unified formulation for Gas and Electric Grid Coordinated Unit Commitment (GECUC) is presented that includes the spinning and non-spinning reserves, and deals with the non-linearity of the gas grid constraints to transform the problem into a Mixed-Integer Linear Program (MILP). Also, a coordinated N-1 generator contingency analysis is presented that finds the contingency dispatch for the gas wells, storage, and generators when a generator is lost.

# Part I

# Cyber-Physical Security of Power Distribution Grid

# CHAPTER 2

# MICRO SYNCHROPHASOR-BASED INTRUSION DETECTION IN AUTOMATED DISTRIBUTION SYSTEMS: TOWARDS CRITICAL INFRASTRUCTURE SECURITY

Electric power distribution systems are undergoing many technological changes and concerns are surfacing on possible additional vulnerabilities. Resilient cyber-physical systems (CPSs) in general must leverage state measures and operational models that interlink the physical and the cyber assets that compose them, to assess the global state. In this chapter we describe a viable process of abstraction to obtain this holistic system state exploration tool, through the analysis of data from Micro Phasor Measurement Units ($\mu$PMUs) combined with the monitoring of Supervisory Control and Data Acquisition (SCADA) traffic, and using semantics to interpret these data that expresses the specific system physical and operational constraints in both cyber and physical realms. We then formulate anomaly detection rules for this architecture that looks into the $\mu$PMU data at different levels of aggregation to filter out anamolous chunks of data for further analysis in order to distinguish between a normal or a malicious anomaly.

## 2.1 Micro Synchrophasor Data: A Game Changer?

We believe deploying $\mu$PMUs can significantly increase the detection and classification capabilities of distribution operators. Many of the cyber-attacks aiming to cause changes in the physical layer leave footprints or anomalies in the $\mu$PMU measurements, such as voltage sags and swells, change of power flow direction, and electric current events. Our basic idea is utilizing the $\mu$PMU measurements to correlate the observed state of the system and the set of detected events through $\mu$PMU to form

building blocks for the estimation of the grid security status. The knowledge about the system topology and operation provides the rules to check the compliance of the events seen in the $\mu$PMU measurements and in the network traffic, with the normal behavior of the system, with some level of certainty.

While ADMS and in particular SCADA have potential existing security flaws due to the use of traditional and outdated security measured, $\mu$PMUs, as a new measurement device, are designed having modernized and advanced security practices in mind. As the first step, they are placed on a separate network from SCADA, and are designed to be read-only devices, and to communicate over secure protocols. However, even if some of the $\mu$PMUs are compromised, since they only provide measurements (in spite of SCADA, which also controls the devices and switches), many of the bad data detection techniques (e.g. [Kim and Poor(2011)]) can be used to remove the false data unless the number of compromised devices is large enough that data injection attack is lost in the noise. In this regard, the optimally-placed $\mu$PMUs can not only detect the bad data injection in SCADA meters but also can be used to identify the bad data injection attack on a subset of $\mu$PMUs. To illustrate the use of $\mu$PMU data in event detection and classification, we offer next an example based on the real data. Specifically, on April-16-2015 a power quality event was captured by the $\mu$PMUs installed at the partner utility grid shown in Fig. 2.1.1. The $\mu$PMUs data showed that a voltage sag occurred, impacting all the $\mu$PMUs placed on two separate feeders. The voltage and current phasor profiles during the event can be seen in Fig. 2.1.2a, 2.1.2b, respectively. Different hypotheses can be formulated about what caused the voltage sag to happen, for example a local or remote transmission or distribution level fault, with a possible protection operation ensuing. Given the brevity of the event, it is extremely unlikely that SCADA data would have captured the sags. But the enormous potential benefit of $\mu$PMU data in assessing security threats is best

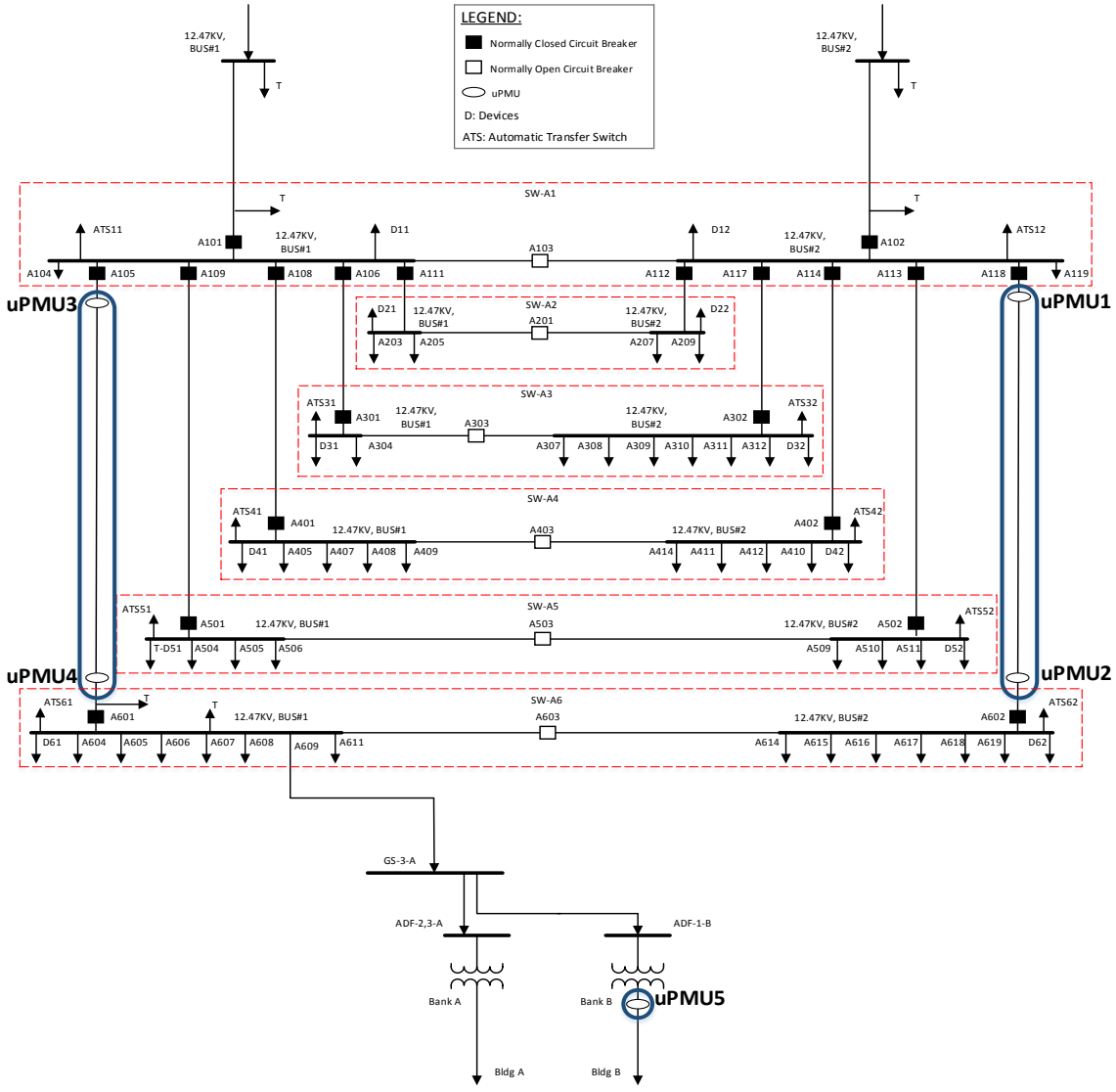
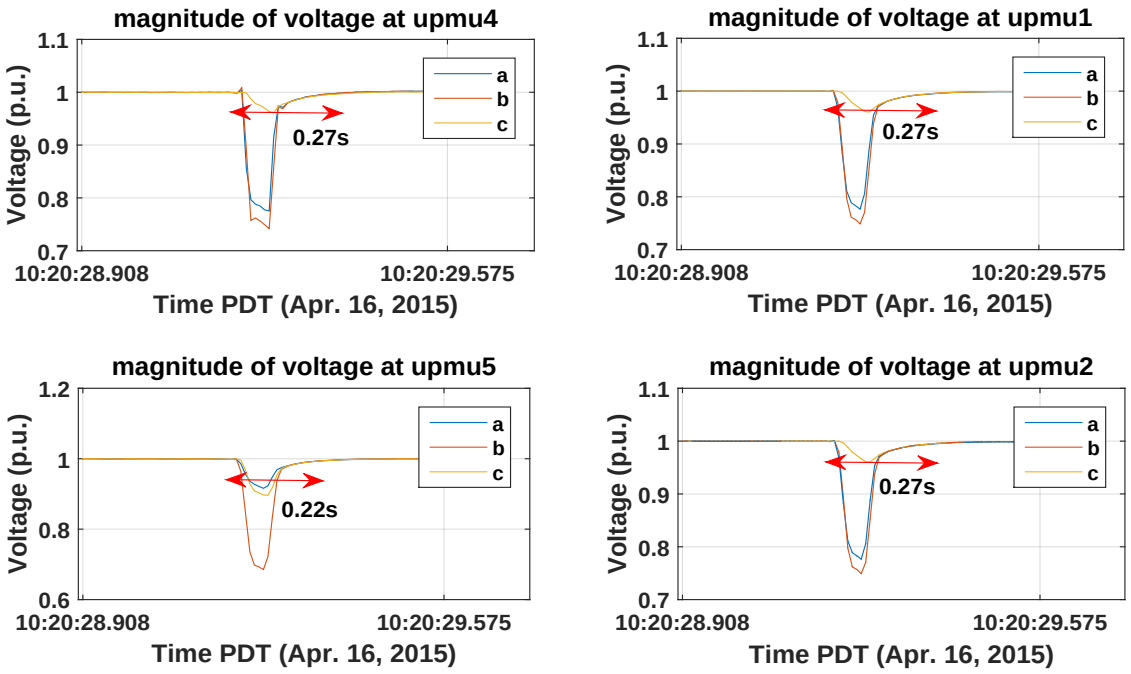



Figure 2.1.1: Partner Utility Distribution Grid One-Line Diagram

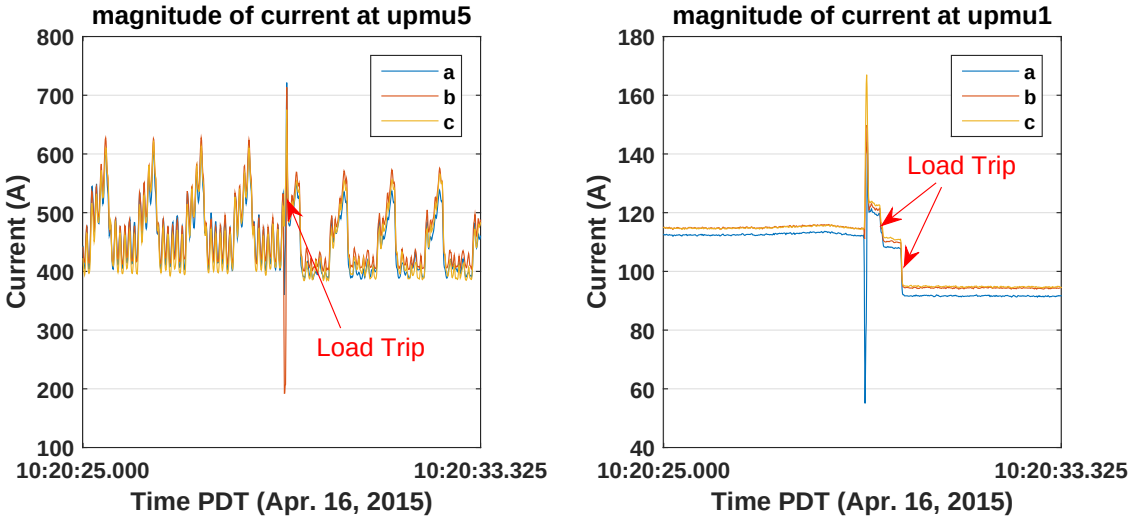
illustrated by the ability they offered to identify the likely source of the problem.

From Fig. 2.1.2a, 2.1.2b it is apparent that the severity of voltage sag is similar for the $\mu$PMUs on both circuits at the same voltage level. In addition, all the $\mu$PMUs captured the voltage sag simultaneously.

A distribution level fault at one feeder causing the simultaneous transients that is transferred through sub-transmission to the other feeder is plausible only if the




(a)




(b)

Figure 2.1.2: (a) Captured Voltage Sag by $\mu$PMUs, (b) Captured Current Phasor during Voltage Sag by $\mu$PMUs

transmission grid is not stiff with respect to transients happening at the distribution feeders, which is usually not the case. Even if this is the case, the captured severity would be more significant on the feeder that the fault happened compared to the other that is not confirmed by the data. Another hypothesis is spreading of voltage sag through the closed Normally Open (N.O.) breakers to the other feeder. This is not corroborated by the data because the N.O. breaker between left and right side are secondary action, which means another breaker should first clear the fault and then this switch is closed to feed the healthy part of the grid. In that case, the sag is already over when the switch gets closed. Even if the attacker tries to close the switch before the fault clearance, the sag is transferred with a delay and different severity and shape to the other side. The transmission level event is the most plausible scenario, as it was visualized concurrently at all the two separate feeders, and is consistent with the $\mu$PMU data.

## 2.2 All-embracing IDS Framework: How to Utilize All the Resources?

The analysis of this event revealed the ability of the $\mu$PMUs to capture the footprints of a grid anomaly that led to physical impact. Based on this analysis we believe that this new rich source of data, combined with knowledge of the grid configuration and operations, allows to reason about different hypotheses and establish the likely cause of an event in a way that would not have been possible using SCADA data or network traffic alone. In addition, it is worth mentioning that, depending on the type of events, some signatures would be more indicative than others of the situation. In the example we offered what the $\mu$PMU data cannot do is to clarify further what happened at transmission level, where we have neither observations nor detailed knowledge of the configuration and operations.

The abstraction of our $\mu$PMU-Based Intrusion Detection System ($\mu$PMU-IDS)

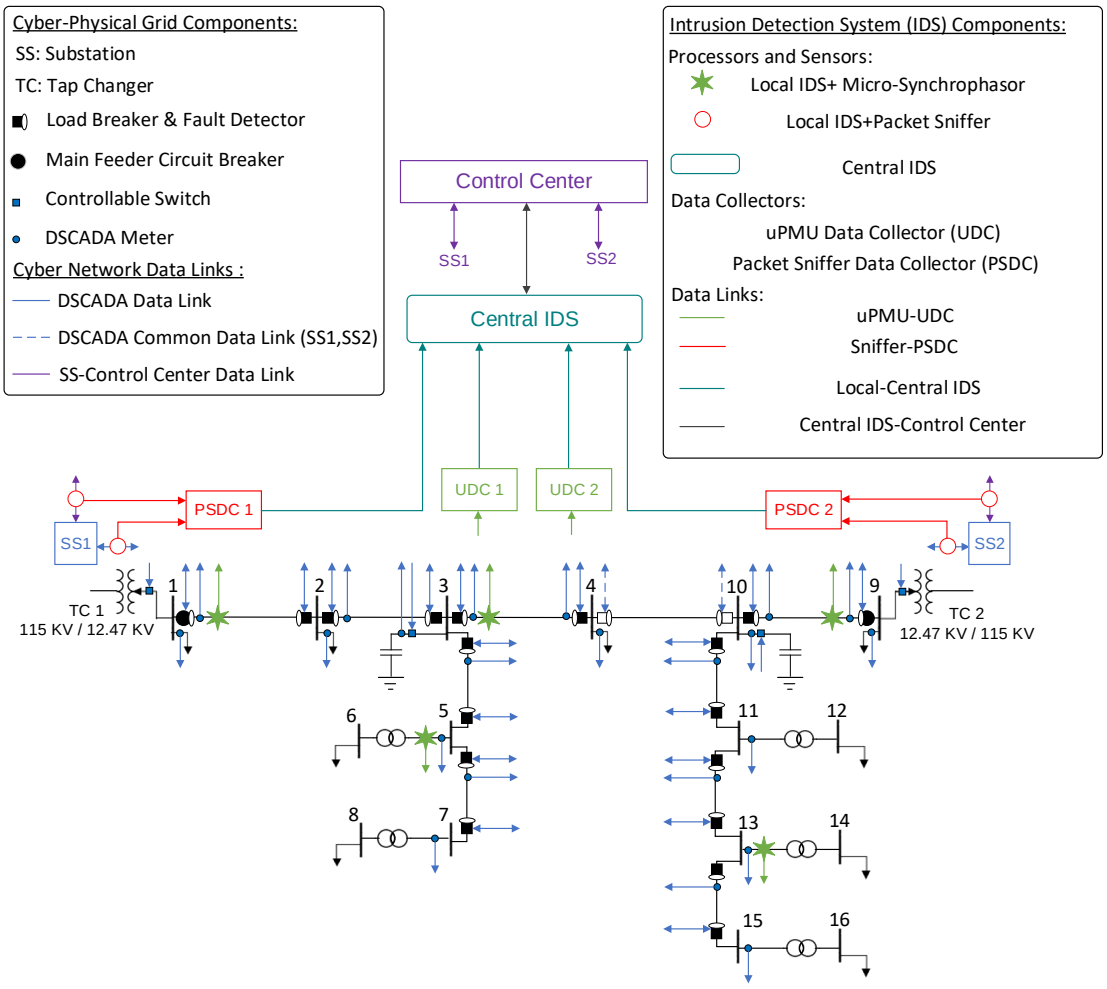



Figure 2.2.1: The Hierarchical IDS Framework Includes: Local (Node Next to Each $\mu$PMU ($\star$) and Packet Sniffer), and Central IDS.

architecture is shown in Fig. 2.2.1. In this figure, local and central IDS form the two levels of IDS data processing, respectively. The correlation of the different data sources including the real-time $\mu$PMU measurements at multiple sites, and monitored SCADA traffic are checked at the central level constantly to draw conclusions about the security state of the grid, while the algorithms at the local level focus on the data coming from each individual sensor irrespective to the others to find anomalies. In Section 2.2, we provide an example that illustrates how our data analytics differ from

the standard network intrusion detection system.

The $\mu$PMU-IDS is designed to be scalable by partitioning the security rules hierarchically. The rules are established based on the physical constraints implied by the Physics of the grid in addition to the common cyber inspection in the computer networks security. The filters used in $\mu$PMU-IDS generalize and automate the process of hypothesis testing that we illustrated in the example we offered before, also utilize the SCADA packets, and are based on encoding the semantics of the rules in a decision tree that can be inspected automatically by the $\mu$PMU-IDS components.

The $\mu$PMU-IDS is an incarnation of the Bro Network Security Monitor framework [Paxson(1999)]. Functionally, the Bro Network Security Monitor is the "glue" that binds passive SCADA system state observations, results leveraged from $\mu$PMU data archiving and analysis tools, and results obtained from circuit analysis activities. Output from the Bro framework will be in the form of predefined software events that can be customized to interact with commercial substation control systems.

Interestingly, the $\mu$PMU-IDS rules pertaining the physical state emulate the behavior of an expert in the field looking at the logged data. Even without cyber security concerns this effort is important to address the *big data* issue, arising from the large amount of sensors and controllers placed on the grid, which would overwhelm the operators.

*What Happens at Local and Central IDS Nodes?*

Prior to forwarding the information for its aggregation, that data from each $\mu$PMU (marked together with green star in Fig. 2.2.1) and SCADA packet sniffer are analyzed by an IDS node that executes local rules. Such rules inspect for the signatures of anomalies in the phasor data streams of the corresponding $\mu$PMU and cyber packets. The rules on the $\mu$PMU data inspect the anomalies in the voltage magnitude,

estimated grid frequency, current magnitude, active and reactive power. In addition, the rules utilize the deviation from the steady-state Kirchhoff and Ohm's law as an indicator of transient behavior and possible changes in the physical parameters of the grid. The formal description of these rules are presented in the following sections. Their specific characteristic is that the analysis is agnostic of the cyber-physical infrastructure configuration, which means that these particular data analytics can be applied broadly.

As an example, in the left side of Fig. 2.2.2 we show how specific rules on the voltage magnitude can convert the data into inferences on various possible hypotheses. On the left side of Fig. 2.2.2, the data are first classified depending on the deviation from the nominal voltage and event time duration values. The way that some of the analysis on the $\mu$PMU data lead to the selection of different hypotheses on some of the voltage events in the local IDS is shown in the right side of Fig. 2.2.2.

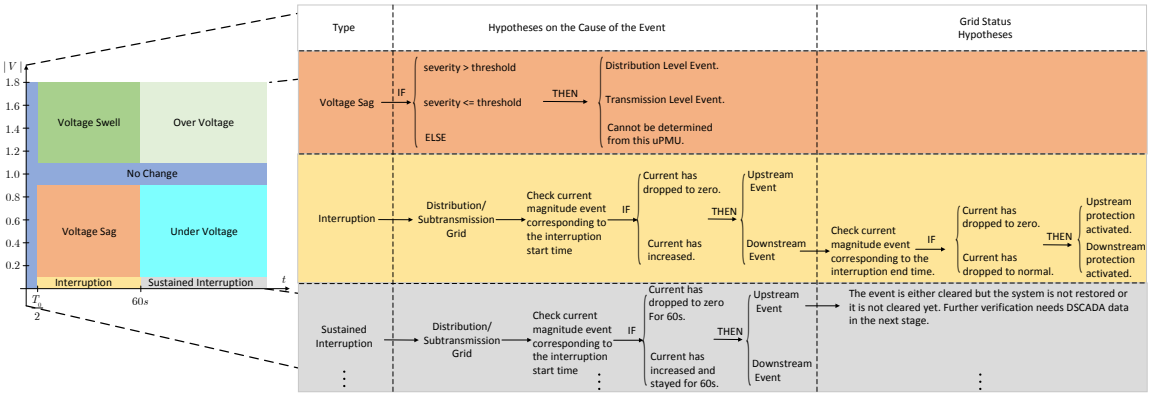



Figure 2.2.2: Decision Region and Hypotheses of the Voltage Magnitude Events in the First Stage of the Local IDS.

The chunks of data containing the event, along with a description of the analysis performed at the local level, are then collected via $\mu$PMU Data Collectors (UDCs) and Packet Sniffer Data Collectors (PSDCs) and reported to the central nodes. The analysis performed by these nodes that are higher in the hierarchy fuses the $\mu$PMUs

measurements and the monitored SCADA traffic to infer the status of the system, taking into account the cyber-physical infrastructure configuration. For example, if the interruption is detected, the bidirectional fault detectors that should see the fault must be checked to unveil a possible spoofing attack. The duration of the interruption can also be compared with the protective load breaker time of operation to determine if the breaker has tripped on time.

An example of an attack scenario is now outlined to demonstrate the hypotheses and process formulated in this work, and tested using the $\mu$PMUs and SCADA packets. This example also clarifies how the SCADA commands, along with the $\mu$PMU data are leveraged in the anomaly detection. In the test case shown in Fig. 2.2.1, a short circuit fault happens on the line connecting bus 5 to 7. In its normal operation, a protection algorithm in substation 1 will detect the fault and use the relay on bus-1 breaker to deenergize the left feeder, at which point the load breakers placed on line 5-7 will receive a command to isolate the fault and finally energy will be restored to the healthy part of the feeder, by closing the circuit breaker at bus 1. Assume that a knowledgeable attacker has gained access to the network and the IP address of the substation controller. For instance, in a first scenario, the attacker could stage a Man-in-the-Middle attack jamming the command of the controller to the relay intended to open the circuit breaker. A second possible scenario is that the attacker changes the firmware of the relay at bus 1 (as in the Ukraine attack case [cbc news Technology and Science(2016)]) and prevent it from tripping. The local IDSs, monitoring for anomalies in the data from the $\mu$PMUs on the left feeder, will detect a transient and alert the central IDS by sending the data through the UDC1 (see Fig. 2.2.1). The packets sniffed by network taps placed on the links that connect the substation to the relays operating the switches at bus 1 and line 5-7, are also sent to central IDS from the sniffers through the PSDC1. Depending on the location of the sniffers, the

analysis of the packets could reveal the man-in-the-Middle attack, corroborating the anomaly detected from the $\mu$PMU. If not, the $\mu$PMU will still indicate that the fault is not cleared, in spite of the opening command having been issued, revealing an attack either to the relay in bus 1 firmware, like in the aforementioned second scenario, or the other possible attack mentioned as the first scenario that is launched after the sniffer. Notice that the latter would not be detectable from the packet analysis only.

# CHAPTER 3

# ANOMALY DETECTION USING OPTIMALLY-PLACED $\mu$PMU SENSORS IN DISTRIBUTION GRIDS

In this chapter, focusing on $\mu$PMU data, we propose a set of analytics and sensor fusion primitives for the detection of abnormal behavior in the control perimeter. Due to the key role of the $\mu$PMU devices in our architecture, an optimal $\mu$PMU placement with limited number of sensors is also described that finds the best location of the devices with respect to our rules. The effectiveness of the proposed methods are tested through the synthetic and real $\mu$PMU data. The following notations are used throughout the rest of the chapter:

| | |
|---|---|
| $j$ | Imaginary unit. |
| $\mathcal{I}_N$ | $N \times N$ identity matrix. |
| $\mathbf{A}^T, \mathbf{A}^*,$ | Transpose, conjugate, |
| $\mathbf{A}^H$ | and conjugate transpose of matrix $\mathbf{A}$. |
| $\|\|\mathbf{A}\|\|, \|\|\mathbf{A}\|\|_F$ | 2-norm and F-norm of matrix $\mathbf{A}$. |
| $\mathbf{1}_{m\times n}$ | $m \times n$ size matrix with entries 1. |
| $\mathbf{A}^\dagger$ | Pseudo-inverse of matrix $\mathbf{A}$. |
| $\otimes$ | Kronecker product. |

The algorithms for anomaly detection applied next to each $\mu$PMU sensor are referred to as *"local rules"* and those that aggregate readings of multiple $\mu$PMUs are referred as *"central rules"*. The design of these rules are in compliance with the architecture presented in Figure. 2.2.1 as the local rules would sit in the local IDSs and central rules would be hosted at the central IDS. For large grid sizes, the

aggregation can occur in multiple steps, where mid-level stages analyze part of the data and forward them upward.

The suite of algorithms that are proposed for the $\mu$PMU-IDS layers have the following advantages in comparison to the present state of the art: (1) due to the near-real-time analysis, analytic results can be used to prioritize the traffic flow from the lower to higher layer, thus pushing forward reports of anomalies faster than data that do not raise a flag and need to simply be accrued for historical purposes; (2) it employs three-phase distribution grid equations rather than the more commonly-used positive sequence solution, thus avoiding the errors arising due to poor modeling; (3) a quasi steady-state condition is considered as the *normal regime* of operation rather than the idealistic steady-state, which assumes there is no frequency drift. These modeling aspects are clarified in our discussion next.

## 3.1 The $\mu$PMU Data in a Distribution Grid

Fig. 3.1.1 shows the $\pi$ model of a distribution line that connects bus $m$ to $n$. Assuming normal conditions, the $\mu$PMUs are designed to capture samples of the

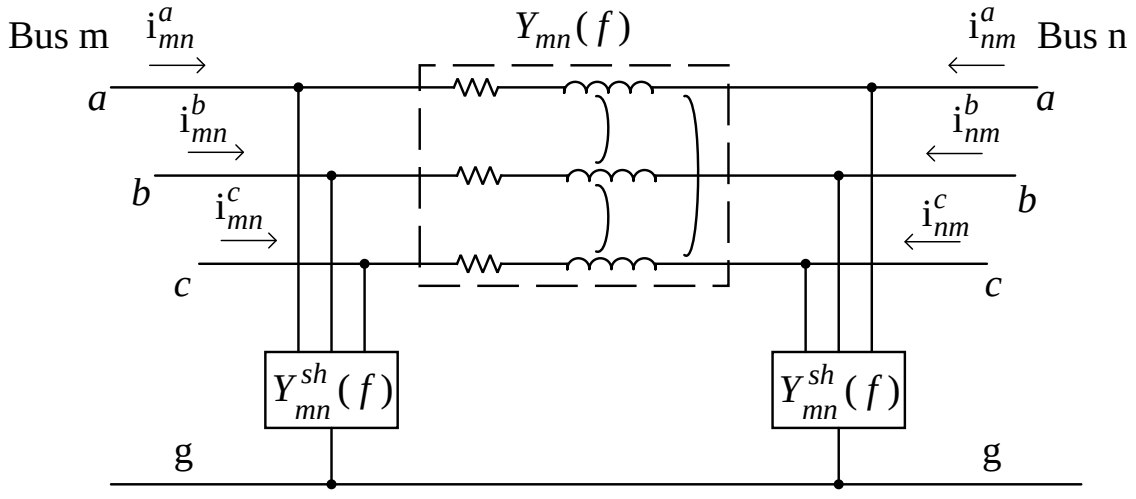



Figure 3.1.1: $\pi$ Model of a Distribution Line

three phase voltage phasor, which is denoted by $\mathbf{v}[k] \in \mathbb{C}^{3\times 1}$, and of the current phasor $\mathbf{i}[k] \in \mathbb{C}^{3\times 1}$ in specific sites of a distribution grid. Next we apply Kirchhoff's and Ohm's law for a three-phase line in a quasi steady-state condition.

A well-known fact from signals and systems theory is that the relationship between voltage and current through a passive circuit with a certain admittance matrix can be represented as a multiplication in the frequency domain and as a convolution in the time domain. Because the circuit is three phase, these will be represented by a Multi-Input Multi-Output system. This fact also holds for the phasor of the signals. We first define $\boldsymbol{y}_{mn}(t)$ and $\boldsymbol{y}^{sh}_{mn}(t)$ to denote the time domain equivalents of the matrices $\boldsymbol{Y}_{mn}(f+f_0)$ and $\boldsymbol{Y}^{sh}_{mn}(f+f_0)$, respectively, where $f_0$ is the main frequency of the grid. In discrete time, the convolution relationship is as follows:

$$\mathbf{i}_{mn}[k] = \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]\mathbf{v}_m[k-r] - \boldsymbol{y}_{mn}[r]\mathbf{v}_n[k-r] \tag{3.1.1}$$

where $\overline{\boldsymbol{y}}_{mn}[r] = \boldsymbol{y}_{mn}[r] + \boldsymbol{y}^{sh}_{mn}[r]$ and it is assumed that $\boldsymbol{y}^{sh}_{mn}[r]$ and $\boldsymbol{y}_{mn}[r]$ are the samples of $\boldsymbol{y}^{sh}_{mn}(t) \star h(t)$ and $\boldsymbol{y}_{mn}(t) \star h(t)$, respectively, and have finite support $N$ and are causal[1].

Next, the form of this relationship during the quasi steady-state is shown, since the steady-state in reality never happens, which in turn represents the governing equations during a regime of operation that is referred to as *"normal"*.

The fundamental frequency of the voltage and current signals are always varying, even in a normal state (although slowly and over a very small range), because of load-

[1] $h(t)$ denotes the low-pass filter implemented in the $\mu$PMU to extract the base-band signal and $H(f)$ is its corresponding frequency response. It should be noted that the outputs of $\mu$PMU, and their corresponding functions, are not the exact phasors if the bandwidth of the signal (voltage/current) is greater than $2f_0$, and are aliased. However, what we are interested in is to track any data abnormality (instead of the exact grid state during the event). Therefore, as long as it is not suppressing the anomaly, having aliasing is not an issue for our rules.

generation imbalances, active power demand interactions, large generators inertia, and the automatic speed controllers of the generators [Phadke and Thorp(2008)]. The effect observable is a change of grid operation regime from steady-state to quasi steady-state. The off-nominal frequency therefore affects the phase angle captured by phasor measurement devices. To highlight that mathematically, the phasor readings $\mathbf{v}_m[k]$ and $\mathbf{i}_{mn}[k]$ are decomposed as follows:

$$\mathbf{v}_m[k] = \hat{\mathbf{v}}_m[k]e^{j\beta_m[k]k},\ \mathbf{i}_{mn}[k] = \hat{\mathbf{i}}_{mn}[k]e^{j\beta_m[k]k} \tag{3.1.2}$$

where $\hat{\mathbf{v}}_m[k]$ is the voltage phasor that is observed at nominal frequency, and $\hat{\mathbf{i}}_{mn}[k]$ is the current phasor after removing the exponential term due to the off-nominal frequency at bus $m$, $\beta_m[k]$, that represents the (time-varying) drift in the frequency induced by the above-mentioned reasons.

Considering the Discrete Time Fourier Transform, we have:

$$\boldsymbol{Y}_{mn}(f + f_0)H(f) = rect(T_s f)T_s \sum_{r=0}^{N-1} \boldsymbol{y}_{mn}[r]e^{-j2\pi r T_s f} \tag{3.1.3}$$

where $\frac{1}{T_s} = 120$Hz is the $\mu$PMU output rate, and $H(f)$ is the frequency response of the filter $h(t)$. Introducing:

$$\boldsymbol{Y}_{mn}(f_0, k) \ \triangleq\ \frac{1}{T_s}\boldsymbol{Y}_{mn}\left(f_0 + \frac{\beta_n[k]}{2\pi T_s}\right) H\left(\frac{\beta_n[k]}{2\pi T_s}\right), \tag{3.1.4}$$

we have that

$$\boldsymbol{Y}_{mn}(f_0, k) = \sum_{r=0}^{N-1} \boldsymbol{y}_{mn}[r]e^{-j\beta_n[k]r}, \tag{3.1.5}$$

and similarly $\overline{\boldsymbol{Y}}_{mn}(f_0, k) = \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]e^{-j\beta_m[k]r}$.

During the quasi steady-state the variations in $\hat{\mathbf{v}}_m[k]$ and $\beta_m[k]$ are negligible over a window of $N$ samples, which means $\hat{\mathbf{v}}_m[k-r] \approx \hat{\mathbf{v}}_m[k]$ and $\beta_m[k-r] \approx \beta_m[k]$. Using

this approximation, (3.1.1) can be re-written as:

$$\begin{aligned}\mathbf{i}_{mn}[k] &\approx \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]\hat{\mathbf{v}}_m[k]e^{j\beta_m[k](k-r)} \\ &\quad - \boldsymbol{y}_{mn}[r]\hat{\mathbf{v}}_n[k]e^{j\beta_n[k](k-r)} \\ &= \overline{\boldsymbol{Y}}_{mn}(f_0,k)\mathbf{v}_m[k] - \boldsymbol{Y}_{mn}(f_0,k)\mathbf{v}_n[k]\end{aligned} \tag{3.1.6}$$

Equation (3.1.6) is Ohm's law in the phasor domain and is the cornerstone for our transients detection algorithm derived in the following section. The analysis above explains through (3.1.6) why in the phasor domain, the equivalent effect of the quasi-steady state conditions is that the admittances (3.1.4) fluctuate. The effect is usually modest, because $\beta[k]$ is small. However, during a severe transient or frequency in the order of 10Hz the relationship (3.1.6) with the matrices in (3.1.4) does not hold, due to the manifestation of the full dynamic behavior in (3.1.1).

In the following, the *"local"* and *"central"* rules are defined leveraging these insights and the knowledge about power system operation. Note that our rules are set up in a way that all the local engines are agnostic about $\overline{\boldsymbol{Y}}_{mn}(f_0,k)$ and $\boldsymbol{Y}_{mn}(f_0,k)$, and the sensor siting. However, it is assumed that the central engine knows $\boldsymbol{Y}^{sh}_{mn}(f_0,k)|_{\beta[k]=0}$ and $\boldsymbol{Y}_{mn}(f_0,k)|_{\beta[k]=0}$ for the lines within its monitoring region and the difference from (3.1.4) is treated as equivalent to noise in the observation model. Therefore, when dealing with central rules, we will simply use $\boldsymbol{Y}^{sh}_{mn}$ and $\boldsymbol{Y}_{mn}$ to refer to $\boldsymbol{Y}^{sh}_{mn}(f_0,k)|_{\beta[k]=0}$ and $\boldsymbol{Y}_{mn}(f_0,k)|_{\beta[k]=0}$.

## 3.2 Data Analysis

Our rules monitor for abnormalities in the following quantities 1) voltage magnitude, 2) current magnitude, 3) active power, 4) reactive power, 5) instantaneous frequency drift, and 6) the validity of quasi steady-state equations. The *"local rules"* just require the stream of phasors from a single $\mu$PMU, while the *"central rules"*

combine multiple streams across $\mu$PMUs.

### *3.2.1 Local Rules*

The *"local rules"* are applied at the lowest layer of the *"anomaly detection architecture"*, and run on systems adjacent to the $\mu$PMUs. Their common feature is that they require no specific prior knowledge of the grid network parameters.

Table 3.1: Voltage Magnitude Anomalies

| anomaly | signature [2] |
|---|---|
| voltage sag | $0.1 \leq \|\mathrm{v}\| \leq 0.9,\ T_0/2 \leq \tau \leq 60s$ |
| voltage swell | $1.1 \leq \|\mathrm{v}\| \leq 1.8,\ T_0/2 \leq \tau \leq 60s$ |
| interruption | $\|\mathrm{v}\| < 0.1,\ T_0/2 \leq \tau \leq 60s$ |
| sustained interruption | $\|\mathrm{v}\| < 0.1,\ \tau > 60s$ |
| undervoltage | $0.1 \leq \|\mathrm{v}\| \leq 0.9,\ \tau > 60s$ |
| overvoltage | $1.1 \leq \|\mathrm{v}\| \leq 1.8,\ \tau > 60s$ |

#### 3.2.1.1 Voltage Magnitude Changes

The magnitude of the voltage varies within a small range that power quality standards enforce during the normal operations [515(2009)]. Therefore, any large deviation from that range indicates an abnormal condition. Table. 3.1 lists the anomalies that can be observed in the voltage magnitude labeled by their severity and duration, denoted by $|\mathrm{v}|$ and $\tau$, respectively.

[2]The voltage magnitude is in p.u.

#### 3.2.1.2 Current Magnitude, Active, and Reactive Power Changes

Even when the voltage magnitude is within the safe range discussed previously, changes in active and reactive power can still happen due to the change of the load, affecting current magnitude and the phase angle between current and voltage phasors. Therefore, it is also of interest to track fast-changes of these quantities using the method described in Section. 3.2.3.

The three phase apparent power can simply be computed as:

$$\mathbf{S}_{mn}[k] = \mathbf{P}_{mn}[k] + j\mathbf{Q}_{mn}[k] = \text{diag}(\mathbf{v}_m[k])\mathbf{i}^*_{mn}[k] \tag{3.2.1}$$

where $\mathbf{P}_{mn}[k]$ and $\mathbf{Q}_{mn}[k]$ are the three-phase active and reactive powers, respectively. Note that for the sake of tracking a power flow change in the distribution grid, tracking the active power and reactive power is preferable over monitoring the phase angle difference, since the resistance of the lines is not negligible, and therefore the angle difference does not necessarily indicate the direction of the power flow. We observed this fact in our partner utility grid, when the voltage phase angle at one end of the line was less than the angle at the other end, though the direction of the power flow was not from the higher angle to the lower one.

Because physical changes in this class of data on the distribution grid can be potentially indicative of negative behavior, it is important to determine the direction (upward trend, downward trend or oscillation) of the change.

The anomalies related to fast changes are labeled in this class of data with *surge*, *drop*, and *oscillation* for increasing, decreasing, and swinging trends respectively by estimating the slope of the signal around the time of change.

While we have primarily introduced fast state changing events in both the dynamic and transient realms, we must also consider events in the steady state time frame, slower changing yet also potentially critical. An example of this could include a

line rating or transformer load being slowly but consistently exceeded leading to accelerated failure. During the quasi steady-state, the three phase current phasor magnitude flowing in each line, i.e., $|\mathbf{i}_{mn}[k]|$, should be less than or equal the line rated current, $|\mathbf{i}_{mn}|_{\max}$. This constraint is imposed as feeder limit, and the violation is flagged as *overcurrent*.

#### 3.2.1.3 Instantaneous Frequency Changes

For a $\mu$PMU at bus $m$, we propose to estimate adaptively the instantaneous local frequency deviation from the nominal frequency during the quasi-steady state, using the approach for instance in [Xia and Mandic(2012)] that is tailored to three-phase distribution lines, to isolate abnormal changes in the estimated frequency.

#### 3.2.1.4 Quasi Steady-State Regime Validity

As previously noted, when the grid is not in the normal quasi steady-state conditions, the relationship between voltage and the current phasors represents its full dynamic behavior, i.e. the grid is no longer well-approximated by the set of memory-less algebraic equations. Therefore, it is proposed to check the validity of the quasi steady-state regime to flag the presence of transients in the grid. At the local engines, (3.1.6) provides the basis for our rule. For the line in Fig. 3.1.1, assuming that a $\mu$PMU is installed at bus $m$ means that $\mathbf{i}_{mn}[k]$, and $\mathbf{v}_m[k]$ are both available.

Let $\boldsymbol{\alpha}[k]$ be the diagonal matrix such that voltage phasors of bus $m$ and bus $n$, connected via a power line, are related through:

$$\mathbf{v}_n[k] = \boldsymbol{\alpha}[k]\mathbf{v}_m[k], \tag{3.2.2}$$

Defining:

$$\boldsymbol{R}_{\text{iv}}^{(mn)}[k] = \frac{1}{M-1}\sum_{r=0}^{M-1} \mathbf{i}_{mn}[k-r]\mathbf{v}_m^H[k-r], \tag{3.2.3}$$

$$\boldsymbol{R}_{\text{vv}}^{(nm)}[k] = \frac{1}{M-1}\sum_{r=0}^{M-1}\mathbf{v}_n[k-r]\mathbf{v}_m^H[k-r], \tag{3.2.4}$$

Assuming that $\boldsymbol{\alpha}[k]$ remains constant over a window of $M$ samples in the quasi-steady state, one can write:

$$\boldsymbol{R}_{\text{vv}}^{(nm)}[k] \approx \boldsymbol{\alpha}[k]\boldsymbol{R}_{\text{vv}}^{(mm)}[k] \tag{3.2.5}$$

It can be assumed that the variation of $\overline{\boldsymbol{Y}}_{mn}(f_0,k)$ is negligible over $M$ samples in normal operation and use (3.1.6) to write:

$$\left(\begin{array}{c:c} \mathcal{I}_3 & -\overline{\boldsymbol{Y}}_{mn}(f_0,k)+\boldsymbol{Y}_{mn}(f_0,k)\boldsymbol{\alpha}[k] \end{array}\right) \underbrace{\begin{pmatrix} \boldsymbol{R}_{\text{iv}}^{(mn)}[k] \\ \boldsymbol{R}_{\text{vv}}^{(mm)}[k] \end{pmatrix}}_{\boldsymbol{R}_k^{(mn)}} \approx \mathbf{0}$$

**Proposition 1.** *Correlation matrix $\boldsymbol{R}_k^{(mn)}$ is approximately rank-1 during the quasi steady-state.*

*Proof.* During the quasi-steady state along a distribution line, the following assumptions hold with a very good approximation for $r = 0, 1, ..., M-1$:

$$\hat{\mathbf{v}}_m[k-r] \approx \hat{\mathbf{v}}_m[k], \quad \beta_m[k-r] \approx \beta_m[k] \tag{3.2.6}$$

Therefore, we can write:

$$\begin{aligned} \boldsymbol{R}_{\text{vv}}^{(mm)}[k] &= \frac{1}{M-1}(\mathbf{v}_m[k]\otimes\mathbf{E}_m[k])(\mathbf{v}_m^H[k]\otimes\mathbf{E}_m^H[k]) \\ &= \frac{1}{M-1}(\mathbf{v}_m[k]\mathbf{v}_m^H[k])\otimes(\mathbf{E}_m[k]\mathbf{E}_m^H[k]) \end{aligned} \tag{3.2.7}$$

where $\mathbf{E}_m[k]$ is defined as follows and represents the variations due to the off-nominal frequency:

$$\mathbf{E}_m[k] = \mathbf{1}_{3\times 1}\otimes\begin{pmatrix} e^{-j\beta_m[k](M-1)} & \dots & e^{-j\beta_m[k]} & 1 \end{pmatrix} \tag{3.2.8}$$

which we can then write:

$$\mathbf{E}_m[k]\mathbf{E}_m^H[k] = (\mathbf{1}_{3\times1}\mathbf{1}_{1\times3}) \otimes (M) = M\mathbf{1}_{3\times3} \tag{3.2.9}$$

and therefore:

$$\boldsymbol{R}_{\text{vv}}^{(mm)}[k] = \frac{M}{M-1}(\mathbf{v}_m[k]\mathbf{v}_m^H[k]) \otimes (\mathbf{1}_{3\times3}) \tag{3.2.10}$$

which accordingly means that:

$$rank(\boldsymbol{R}_{\text{vv}}^{(mm)}[k]) = rank(\mathbf{v}_m[k]\mathbf{v}_m^H[k]) \times rank(\mathbf{1}_{3\times3}) = 1$$

Because:

$$rank(\boldsymbol{R}_k^{(mn)}) = rank((\boldsymbol{R}_k^{(mn)})^H\boldsymbol{R}_k^{(mn)})$$

we analyze the rank of $(\boldsymbol{R}_k^{(mn)})^H\boldsymbol{R}_k^{(mn)}$ here, where:

$$\begin{aligned}(\boldsymbol{R}_k^{(mn)})^H\boldsymbol{R}_k^{(mn)} =&(\boldsymbol{R}_{\text{iv}}^{(mn)}[k])^H\boldsymbol{R}_{\text{iv}}^{(mn)}[k]+\\ &(\boldsymbol{R}_{\text{vv}}^{(mm)}[k])^H\boldsymbol{R}_{\text{vv}}^{(mm)}[k]\end{aligned} \tag{3.2.11}$$

From the structure of (3.2.6) during the quasi-steady state, we have:

$$\begin{aligned}\boldsymbol{R}_{\text{iv}}^{(mn)}[k] &= \tilde{\mathbf{Y}}_{mn}(f_0,k)\boldsymbol{R}_{\text{vv}}^{(mm)}[k]\\ \tilde{\mathbf{Y}}_{mn}(f_0,k) &= \overline{\boldsymbol{Y}}_{mn}(f_0,k) - \boldsymbol{Y}_{mn}(f_0,k)\text{diag}(\boldsymbol{\alpha}[k])\end{aligned} \tag{3.2.12}$$

Substituting (3.2.12) in (3.2.11), we have:

$$(\boldsymbol{R}_k^{(mn)})^H\boldsymbol{R}_k^{(mn)} = (\boldsymbol{R}_{\text{vv}}^{(mm)}[k])^H\boldsymbol{\mathcal{Y}}_{mn}(f_0,k)\boldsymbol{R}_{\text{vv}}^{(mm)}[k] \tag{3.2.13}$$

where:

$$\boldsymbol{\mathcal{Y}}_{mn}(f_0,k) = \tilde{\mathbf{Y}}_{mn}^H(f_0,k)\tilde{\mathbf{Y}}_{mn}(f_0,k) + \mathcal{I}$$

Since the linear transformation of $\boldsymbol{R}_{\text{vv}}^{(mm)}[k]$ does not increase its rank and since it has already been shown that $\boldsymbol{R}_{\text{vv}}^{(mm)}[k]$ is of rank-1 during the quasi-steady state, one can conclude that:

$$\begin{aligned} &rank(\boldsymbol{R}_k^{(mn)}) = rank((\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)}) \leq \\ &rank(\boldsymbol{R}_{\text{vv}}^{(mm)}[k]) \rightarrow rank(\boldsymbol{R}_k^{(mn)}) = 1 \end{aligned} \tag{3.2.14}$$

□

Hence:

$$\begin{aligned} &\boldsymbol{R}_k^{(mn)} \approx \sigma_1^{(mn)}[k]\boldsymbol{u}_1^{(mn)}[k](\boldsymbol{\nu}_1^{(mn)}[k])^H \rightarrow \\ &\boldsymbol{R}_k^{(mn)}(\boldsymbol{R}_k^{(mn)})^H \approx (\sigma_1^{(mn)}[k])^2\boldsymbol{u}_1^{(mn)}[k](\boldsymbol{u}_1^{(mn)}[k])^H \end{aligned} \tag{3.2.15}$$

where $\sigma_1^{(mn)}$ is the largest singular value of $\boldsymbol{R}_k^{(mn)}$, and $\boldsymbol{u}_1^{(mn)}$ and $\boldsymbol{\nu}_1^{(mn)}$ are the corresponding left and right singular vectors to that, respectively. Deviation from this subspace structure can indicate that the line is experiencing a transient. We can automate the detection of anomaly using this criterion by computing the following cost minimization and tracking the fast changes in $x[k]$ for each incidental line to a bus with $\mu$PMU:

$$\begin{aligned} x[k] = &\min_{\boldsymbol{u}} ||(\mathcal{I}_6 - \boldsymbol{u}\boldsymbol{u}^H)\boldsymbol{R}_k^{(mn)}(\boldsymbol{R}_k^{(mn)})^H||_F \\ &\text{s.t. } ||\boldsymbol{u}|| = 1 \end{aligned} \tag{3.2.16}$$

In other words, $x[k]$ measures the size of the residual that $\boldsymbol{R}_k^{(mn)}$ has in the space orthogonal to the optimal $\boldsymbol{u}$, which should be zero in the quasi steady-state and non-zero otherwise. Since it has already been shown that $\boldsymbol{R}_k^{(mn)}$ is of rank-1 during the quasi steady-state, the left singular vector corresponding to the largest singular value of $\boldsymbol{R}_k^{(mn)}$ is the only quantity of interest to compute the metric, instead of computing all the singular vectors. Therefore, also owing it to the small size of the matrix, the method is not computationally very expensive.

To conclude the local rules, a flowchart of the analysis performed at the local engine next to $\mu$PMU at bus $m$ is presented in Fig. 3.2.1. At each instant of time, the phasor readings are received by the local engine and the introduced metrics above are calculated. The pre-processed data are then passed to the local rules to check for any violation. A violation could be trespassing pre-defined limits (e.g., the voltage magnitude rule, or maximum current magnitude limit) or fast changes in a metric with smooth behavior during the normal condition (e.g., quasi steady-state validity rule). Once a violation is found in one of the metrics, the start time is recorded. The search for anomaly continues until no new violation is found for a certain window of time ( *"Count1* $> T_1$*"*), and that specifies the end time of the anomaly. The type of anomaly is then determined based on the behavior of the data between the start time and the end time (e.g., active power surge, voltage interruption,...). The start time, end time and the anomaly label is then sent upstream for further analysis/visualization, and the parameters are reset for next event. If the number of detected violations related to a certain event passes a pre-defined threshold ( *"Count2* $> T_2$*"*), the end time is replaced with a *"Persistent"* label, and the results are sent to the central engine, without waiting for the end of the event to arrive. The reason is to be able to inform the operator about the anomaly in time, and not waiting too long before something more damaging happens.

### *3.2.2 Central Rule*

At higher levels of aggregation, the central engine in our case, the natural way to relate the measurements is through the grid interconnection. For a grid with $B$ buses, let $\mathbf{I}[k]$ denote the vector of three-phase current injection phasors with size $3B$, and $\mathbf{V}[k]$ represent the vector of three-phase voltage phasor at all the buses, which contains $3B$ elements. We define the measurement vector $\mathbf{d}[k] = (\mathbf{I}^T[k], \mathbf{V}^T[k])^T$.

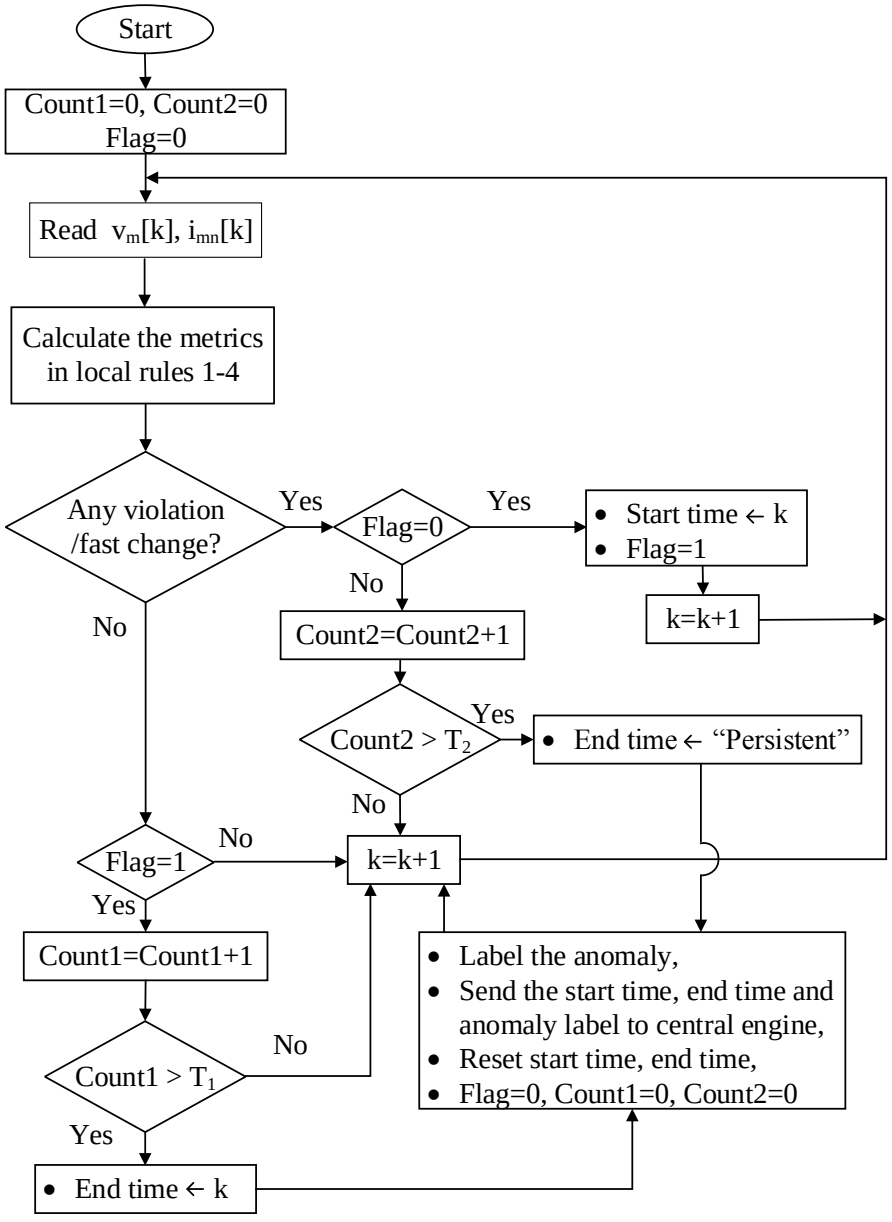



Figure 3.2.1: Local Engine Analysis Flowchart Next to $\mu$PMU at Bus $m$.

During the steady-state, the following algebraic equation should hold:

$$\mathbf{H}\mathbf{d}[k] = \mathbf{0}, \quad \mathbf{H} = \left( \mathcal{I}_{3B} \,\middle|\, -\boldsymbol{Y}_{3(B\times B)} \right) \tag{3.2.17}$$

where $\boldsymbol{Y}$ is the admittance matrix of the grid that connects the current injection to the bus voltages. During the quasi steady-state these equations are close to be homogeneous. Since the distribution grid is unbalanced, and the lines are not transposed the set of equations that should be dealt with are three phase instead of working with positive sequence [Kersting(2012)]. In this framework, we are also able to include the laterals in the admittance matrix by putting the entries corresponding to the phases that do not exist equal to zero.

It is assumed that $\mathbf{H}$ is known, $K$ denotes the number of $\mu$PMUs that are available and that each device has enough channels to measure the voltage and all incidental current measurements of the bus on which it is placed. Therefore, having a $\mu$PMU at bus $m$ means that the entries in $\mathbf{V}[k]$ corresponding to $\mathbf{v}_m[k]$ and entries in $\mathbf{I}[k]$ corresponding to $\sum_{n:m\sim n} \mathbf{i}_{mn}[k]$ are both available, where $m \sim n$ denotes that bus $m$ and $n$ are connected through a line. Let $\mathbf{T}$ denote a matrix parsing the vector $\mathbf{d}[k]$ into two parts corresponding to the unavailable measurements, $\mathbf{d}_u[k]$, and the available measurements, $\mathbf{d}_a[k]$:

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_u \\ \mathbf{T}_a \end{pmatrix} \;\rightarrow\; \mathbf{T}\mathbf{d} = \begin{pmatrix} \mathbf{d}_u \\ \mathbf{d}_a \end{pmatrix}, \; \mathbf{H}\mathbf{T}^T = \left( \mathbf{H}_u \,\middle|\, \mathbf{H}_a \right) \tag{3.2.18}$$

where $K' = B - K$ and

$$\mathbf{T}_u \in \{0,1\}^{6(K'\times B)}, \quad \mathbf{T}_a \in \{0,1\}^{6(K\times B)} \tag{3.2.19}$$

$\mathbf{T}_u$ and $\mathbf{T}_a$ here are block diagonal matrices of size $6(K' \times B)$ and $6(K \times B)$ with entries equal to 0 or 1. The former is supposed to select the unavailable current and voltage phasors by having entries equal to 1 at corresponding locations, and the latter

is supposed to pick the available phasors. Since $\mathbf{T}^T\mathbf{T} = \mathcal{I}$, we can rewrite (3.2.17) as follows:

$$\mathbf{H}_u\mathbf{d}_u[k] + \mathbf{H}_a\mathbf{d}_a[k] = \mathbf{0} \rightarrow \tag{3.2.20a}$$

$$\mathbf{H}_a\mathbf{d}_a[k] = -\mathbf{H}_u\mathbf{d}_u[k]. \tag{3.2.20b}$$

Let first assume that we have enough $\mu$PMUs that satisfy $K > \frac{B}{2}$. Therefore, the matrix $\mathbf{H}_u$ would be a tall matrix and has a left null-space. Premultiplying both sides of (3.2.20b) by the projector on the left null-space of $\mathbf{H}_u$, we have:

$$(\mathcal{I} - \mathbf{H}_u\mathbf{H}_u^\dagger)\mathbf{H}_a\mathbf{d}_a[k] = \mathbf{0} \tag{3.2.21}$$

because $(\mathcal{I} - \mathbf{H}_u\mathbf{H}_u^\dagger)\mathbf{H}_u$=$\mathbf{0}$. The equality in (3.2.20b) only holds during the steady-state (and for quasi steady-state with a good approximation), which means that equation (3.2.21) is homogeneous only in the steady-state (nearly-homogeneous in the quasi steady-state), and non-homogeneous otherwise. The following metric should, therefore, be close to zero only during normal operation and if $\mathbf{H}$ is unchanged:

$$x[k] = \frac{||(\mathcal{I} - \mathbf{H}_u\mathbf{H}_u^\dagger)\mathbf{H}_a\mathbf{d}_a[k]||_2^2}{||\mathbf{d}_a[k]||_2^2} \tag{3.2.22}$$

However, in reality, the number of available $\mu$PMUs, $K \ll \frac{B}{2}$. In this case, $\mathbf{H}_u$ is a fat matrix and, in general, is of full row rank, which in turn means that $(\mathcal{I} - \mathbf{H}_u\mathbf{H}_u^\dagger)$=$\mathbf{0}$, and our criterion becomes trivial. However, $\mathbf{H}_u\mathbf{H}_u^H$ has generally a high condition number, due to the weak grid connectivity. Therefore, both sides of (3.2.20b) can be projected on the subspace spanned by the left singular vector, denoted by $\mathbf{u}_{u,s}$ corresponding to the smallest singular value of $\mathbf{H}_u$ expecting that $\mathbf{u}_{u,s}^H\mathbf{H}_u\mathbf{d}_u[k]$ to be small. Accordingly, when the quasi steady-state is the regime of operation, i.e., when the equality in (3.2.20b) holds, it is expected that the following function to be small and to vary smoothly:

$$x[k] = \frac{|\mathbf{u}_{u,s}^H\mathbf{H}_a\mathbf{d}_a[k]|^2}{||\mathbf{d}_a[k]||^2}. \tag{3.2.23}$$

The exit from this behavior is then marked as an anomaly in the central engine and $x[k]$ is the quantity that is proposed to be tracked for fast changes for this purpose.

The flowchart for the central engine is similar to the one for local engine to some extent except that all the $\mu$PMU readings are required to form the central metric, and the results from the local analysis are received by the central engine, and the analysis results from this stage are not shipped anywhere, and are ready to be displayed for the operator.

### *3.2.3 Fast Change Detection Method*

As explained above, some of the criteria defined in the local and central rules require tracking fast changes in the quantities that are defined, because severe variations in $x[k]$ are signatures of an anomaly. From real data and simulations we have verified that variations in the mean value for these quantities during the quasi steady-state regime are extremely smooth. This observation prompted us to consider changes in their mean value as the common statistical trade-mark of anomalies in all of these quantities and to use the sequential two-sided Cumulative Sum (CUSUM) algorithm [Page(1954), Basseville *et al.*(1993)] as a heuristic. The use of this algorithm amounts to approximating the samples, for all the aforementioned quantities, as outcomes of a Gaussian non-zero mean process with independent observations. Although the observations are in fact temporally correlated, our objective (i.e. fast change detection in mean) justifies the relaxation that the random process has independent observation samples. The algorithm decides between two hypotheses $\mathcal{H}_0$: *no change in the mean*, or $\mathcal{H}_1$: *change in the mean*, at time $k$.

A change is detected when the decision function of CUSUM crosses a pre-defined threshold. Therefore, choosing a proper threshold would be essential for practical applications. In this sense, our intrusion detection architecture falls into "semi-

supervised behavior-based detection" category, where a collection of real data points containing a variety of anomalies are required to train the detector and find the proper threshold in the learning phase. Several metrics are proposed to evaluate the performance of the CUSUM. The *"average run length"* proposed in [Page(1954)] and the *"worst case detection delay"* proposed in [Lorden(1971)] are the two most popular ones. Having the approximation of a Gaussian distribution, these metrics can be computed accordingly.

We expect to see multiple change points during an event. Detection of multiple change points is achieved by resetting the decision functions and cumulative sums to zero after the change is detected, and continuing the inspection of upcoming samples. The fast change anomaly is completed if no new changes are detected for a defined window of time.

## 3.3 Optimal $\mu$PMU Placement

In tandem with the anomaly detection, the criterion that was described for the central engine can be the basis to determine an optimal placement for the $\mu$PMUs. As it was mentioned, the challenge here is that we cannot feasibly nor practically deploy $\mu$PMUs at all nodes in the distribution system, therefore we consider a limited deployment, where $K \ll \frac{B}{2}$.

Ideally, we want the matrix $\mathbf{H}_u$ to have a left null-space for the criterion in (3.2.23) to be zero in the steady-sate. Considering the high condition number of $\mathbf{H}_u\mathbf{H}_u^{\dagger}$, a reasonable approach to find the optimal configuration is to minimize the norm in (3.2.23) over all the possible placement configurations. Let first rewrite the defined metric in (3.2.23):

$$x[k] = \frac{\mathbf{d}_a^H[k]\mathbf{H}_a^H\mathbf{u}_{u,s}\mathbf{u}_{u,s}^H\mathbf{H}_a\mathbf{d}_a[k]}{||\mathbf{d}_a||^2}. \tag{3.3.1}$$

We desire our formulation to be only topology-dependent and not to be a function of measurements. Therefore, the optimal placement problem is formulated as a min-max optimization with the following structure:

$$\begin{aligned}
\mathbf{\Pi}^{\text{opt}} = \underset{\mathbf{\Pi}}{\text{argmin}} \quad & \lambda_{\max}(\mathbf{W}) \\
\text{s.t.} \quad & \left( \mathbf{H}_u \;\middle|\; \mathbf{H}_a \right) = \mathbf{H} \left( \mathbf{T}_u^T \;\middle|\; \mathbf{T}_a^T \right) \\
& \mathbf{T} = \mathcal{I}_2 \otimes (\mathbf{\Pi} \otimes \mathcal{I}_3), \\
& \mathbf{W} = \mathbf{H}_a^H \mathbf{u}_{u,s} \mathbf{u}_{u,s}^H \mathbf{H}_a, \quad [\mathbf{\Pi}]_{i,j} \in \{0,1\} \\
& \sum_j [\mathbf{\Pi}]_{i,j} = 1, \quad \sum_i [\mathbf{\Pi}]_{i,j} = 1
\end{aligned} \tag{3.3.2}$$

Since $\lambda_{\max}(\mathbf{W}) = \max_{\mathbf{d}_a[k]} \frac{\mathbf{d}_a^H[k]\mathbf{W}\mathbf{d}_a[k]}{||\mathbf{d}_a[k]||^2}$, essentially we are choosing a placement that minimizes the maximum value that our objective function can take over possible set of available measurement vectors $\mathbf{d}_a[k]$.

An exhaustive search is required to find the global optimum of the optimization problem in (3.3.2), which is exponentially complex, and therefore does not scale well. This becomes a barrier when the size of the grid is large, i.e., in most of the real grids. Therefore, we propose to employ a *"Greedy Search"* as an alternative to reduce the time complexity to *polylog*, while accepting to be near-optimal. The pseudo-code of the employed greedy search is illustrated in Algorithm. 3.1.

## 3.4 Numerical Results

In this section, we first find the optimal placement for our $\mu$PMUs based on (3.3.2) and then test the effectiveness of the proposed anomaly detection criteria through simulated data and real data, provided from the $\mu$PMUs that are installed in our partner utility medium voltage (12.47 kV) grid.

**Algorithm 3.1:** Greedy Search Pseudo-Code for Anomaly Detection Optimal $\mu$PMU Placement in Application.

**Initialization**

$K :=$ Number of $\mu$PMUs;

$\mathcal{P} := \emptyset$, // Set of selected placement locations;

$\mathcal{L} :=$ Set of candidate placement buses;

**begin**

**for** *n=1..K* **do**

$Cost \leftarrow$ inf;

**for** *each* $l \in \mathcal{L}$ **do**

$\mathcal{P} := \mathcal{P} \cup \{l\}$;

given $\mathcal{P}$, calculate $\lambda_{\max}(\mathbf{W})$;

**if** $\lambda_{\max}(\mathbf{W}) < Cost$ **then**

$l_{opt} \leftarrow l$;

$Cost \leftarrow \lambda_{\max}(\mathbf{W})$;

$\mathcal{P} := \mathcal{P} \setminus \{l\}$;

$\mathcal{P} := \mathcal{P} \cup \{l_{opt}\}$;

$\mathcal{L} := \mathcal{L} \setminus \{l_{opt}\}$;

### 3.4.1 Synthetic Data

The IEEE-34 bus test case [Group(2010)] is simulated using the time-domain simulation environment of DIgSILENT [Manual and PowerFactory(2009)] that deals with differential equations rather than memory-less equations. The sampling rate is selected to be equal to the sampling rate of the Analog-to-Digital Converter (ADC) in a real $\mu$PMU, which is 512 $\times$ 60 Hz =30720 samples per sec.. We then processed these time-domain data through our phasor estimation algorithm, that emulates a two-cycle, P-class algorithm based on the IEEE C37.118.1 [for Power Systems Working Group(2011)] producing phasor samples at a rate of 120 Hz. The single-line diagram of the test case is shown in Fig. 3.4.1. This case includes single-phase laterals, voltage regulators, and untransposed lines, which all are modeled exactly in our admittance matrix.

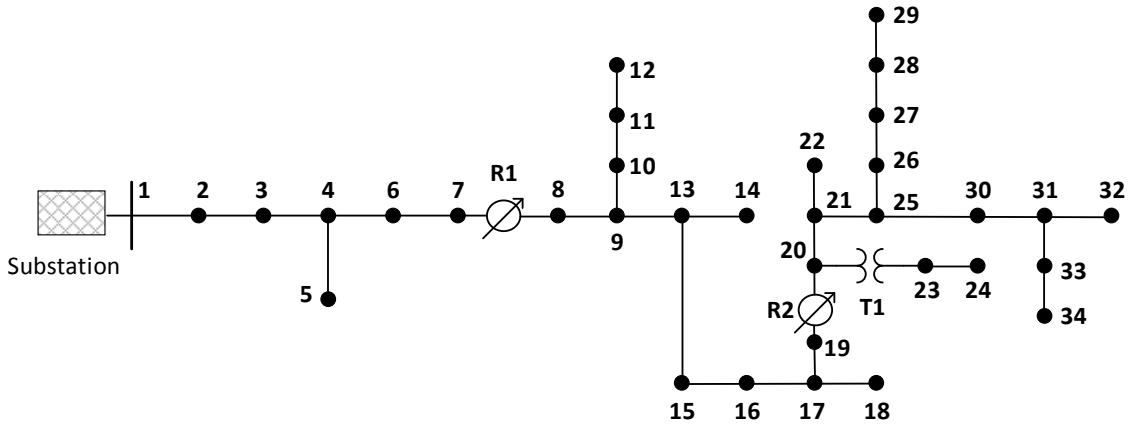



Figure 3.4.1: IEEE 34-Bus Test Feeder Single-Line Diagram

Table. 3.2 compares the objective value for a random placement, *"Greedy Search"* and *"Exhaustive Search"*, and the time complexity of each solver, assuming that $K = 3$ $\mu$PMUs are available. The objective value of the *"Greedy Search"* and the *"Exhaustive Search"*, and the set of the selected buses are close to each other, while

Table 3.2: IEEE-34 Case Optimal $\mu$PMU Placement Result for $K = 3$

| | **Random** | **Greedy** | **Exhaustive** |
|---|---|---|---|
| **Optimum Cost** | 1.7085 | 0.51477 | 0.51477 |
| **Buses with $\mu$PMUs** | {1,3,9} | {7,19,31} | {9,19,31} |
| **Run Time** | – | 2.84 s | 290.266 s |

the run time of the *"Greedy Search"* is 102.206 times faster. Hence, the *"Greedy Search"* can be a very good choice to solve our optimal placement problem.

As expected, the placement rule tries to scatter the available $\mu$PMUs all over the grid, in order to achieve the maximum possible coverage.

In order to investigate how the time complexity grows, and also analyze the results of the placement criterion, the 123 standard test case in [Group(2010)] was used, considering 20 $\mu$PMUs available (i.e., $K = 20$). Without loss of generality, the 123 test feeder was reduced to 70 buses to only include three-phase lines, and roll up all the laterals (single-phase and two-phase lines). The reasoning behind this reduction is that visibility on the main feeder is more important for us than the the visibility on the laterals, considering the limited number of $\mu$PMUs. Using a machine with 60 Intel(R) Xeon(R) CPU E7-4870 v2 @ 2.30GHz cores, it took 11.56 Sec for the algorithm to place 20 $\mu$PMUs over 70 buses. Fig. 3.4.2 illustrates the location of the optimally-placed $\mu$PMUs in the reduced grid. It can be observed from the figure that the $\mu$PMUs are scattered over the grid to achieve the maximum sensitivity with respect to different locations of anomalies.

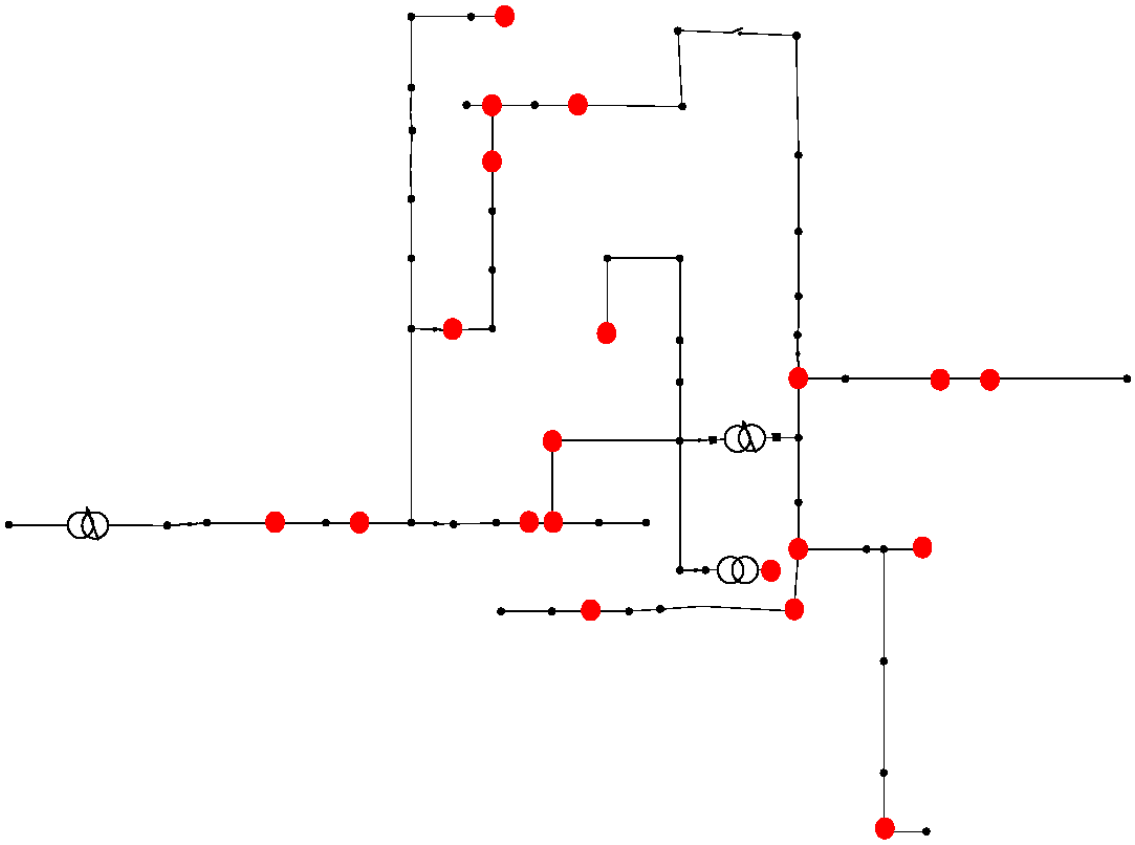


Figure 3.4.2: Optimally-Placed $\mu$PMUs on Reduced IEEE 123 Test Feeder.

#### 3.4.1.1 Single-Line to Ground Fault-IEEE 34 Bus

In this simulation, the rules are tested for detecting anomaly with respect to a Single-Line to Ground Fault (SLGF), which is a very common type of short-circuit fault in the distribution grid. A SLGF was introduced on "Phase a" of line (25,26), which then caused the fuse placed on the phase a of this line near bus 25 to melt down. Our three $\mu$PMUs are placed on buses, 7, 19, and 31 based on the "greedy search" result for the optimal placement criterion.

The results of the *"voltage magnitude change"* rule are shown in Fig. 3.4.3 for $\mu$PMUs 7, and 19 for instance. Simply, the rule inspects the data for large deviations, label them accordingly, and find the start time and the end time of the event, marked

with blue and red stars.

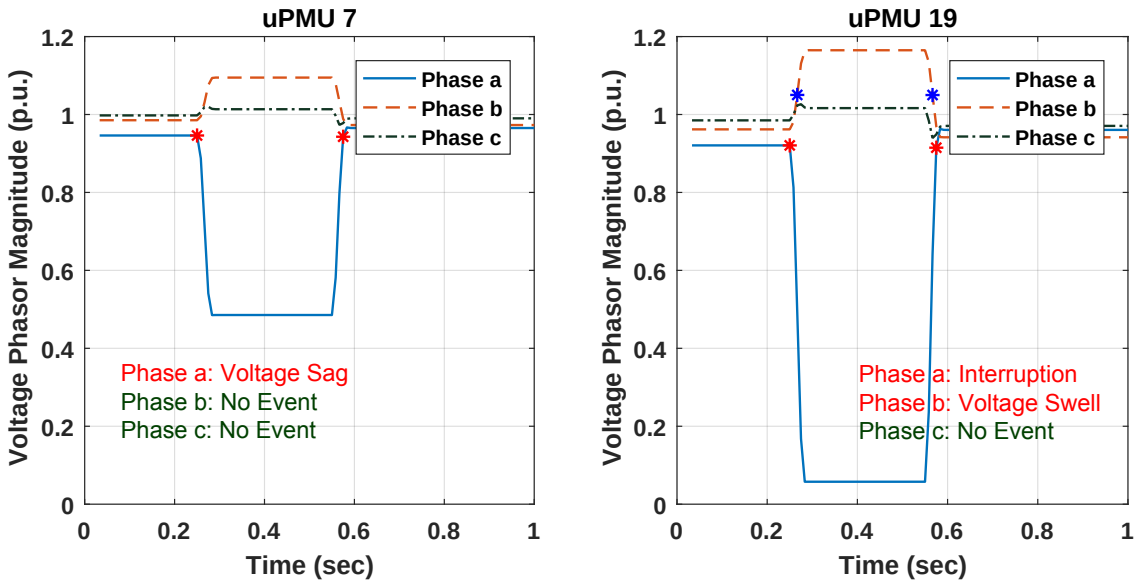



Figure 3.4.3: Voltage Magnitude Local Rule Result for SLGF.

Fig. 3.4.4 illustrates the metric value in (3.2.16) for the specified lines that is inspected with $M = 12$ to check whether the grid is in the quasi steady-state or not. The voltage and current data are first converted to per-unit system assuming $S_b = 1$ MVA. The start time of the detected changes are also marked, setting the CUSUM detector parameters fixed in all the three local engines corresponding to each $\mu$PMU. As it can be observed, there are two periods in which the grid manifests its dynamic: the first one corresponds to the occurrence of the fault and the second matches with the fuse meltdown. In addition, based on the severity of the transient that each $\mu$PMU measures, the number of detected changes via CUSUM varies. In this case, considering the location of the $\mu$PMUs and the location and type of the fault, the most severe change appears in the metric corresponding to measurements from line (19,20), while the changes in metric for line (31,32) is very small. Therefore, based on the defined parameters on the detector, CUSUM finds quite a large number of change

points in the former, while it is not set to be sensitive to the changes in the order that appears in the latter. In fact, if the detector is set to be too sensitive, it can increase "*false alarms*" in the system. Also note that due to the two-cycle calculation of the phasor, and use of $M$ samples to calculate the correlation matrix, the event appears and disappears with a systematic delay.

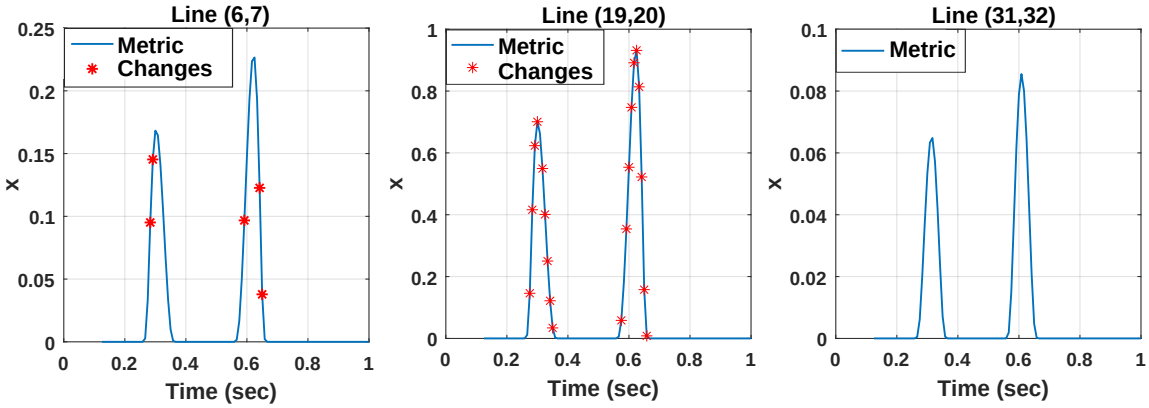



Figure 3.4.4: Quasi Steady-State Validity Checking for SLGF.

The other local rules also capture the anomaly, though with different severity and behavior (we just show some of the results here for lack of space). In fact, many of the local rules may detect the same event, though some rules are more informative than others depending on the cause. Each triggered rule reports a start and an end time for the event. Storing these time-tags for eventful segments of data helps understanding their relationship.

The metric defined for our central engine is also illustrated in Fig. 3.4.5. The delay in appearing and disappearing of the event in here is solely due to the two-cycle phasor calculation.

It should be noted that the value of the metric in (3.2.23) during the quasi steady-state is highly dependent on the number of $\mu$PMUs and the topology of the grid that together determine the structure of $\mathbf{H}_u$ and $\mathbf{H}_a$. However, the best placement of the

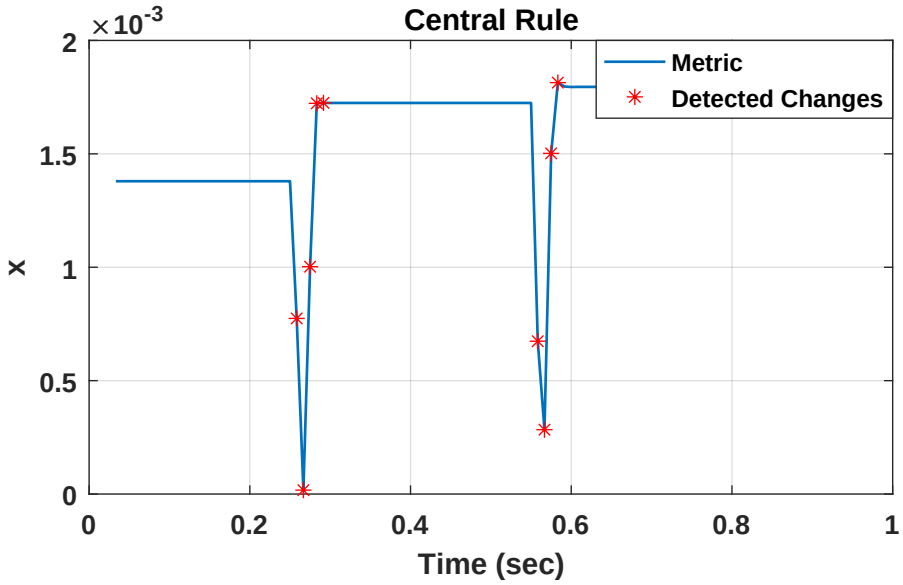



Figure 3.4.5: Central Rule Inspection for SLGF.

$\mu$PMUs for a certain topology and number of $\mu$PMUs that makes the metric to be as close as possible to zero, returns a certain objective value that can be used as the baseline to determine what is a normal value for $x[k]$ and what would be an anomaly.

#### 3.4.1.2 Feeder Reconfiguration-IEEE 34 Bus

In this case, it is assumed that there is a normally-open line that connects node 13 to 21. This line is closed as part of a hypothetical feeder reconfiguration scheme when the line 13-15 is opened. We run the simulation for 1.5 sec, and the feeder reconfiguration happens at t = 0.75 sec. Fig. 3.4.6 shows the voltage phasor magnitude and the metric related to the quasi steady-state validity check at the local level for the reconfiguration event. As it can be seen, the event does not trigger any anomaly in the voltage phasor magnitude. For example, if we wanted to solely rely on the raw phasor data coming from $\mu$PMU 7 and not checking other metrics, this event would have been missed because the current phasors measured by this $\mu$PMU also do not see significant change, mainly because the downstream load of bus 7 has not changed. However,

checking the validity of quasi steady-state for line 6-7 using the measured phasor data by $\mu$PMU 7 clearly reveals the existence of an anomaly in the system.

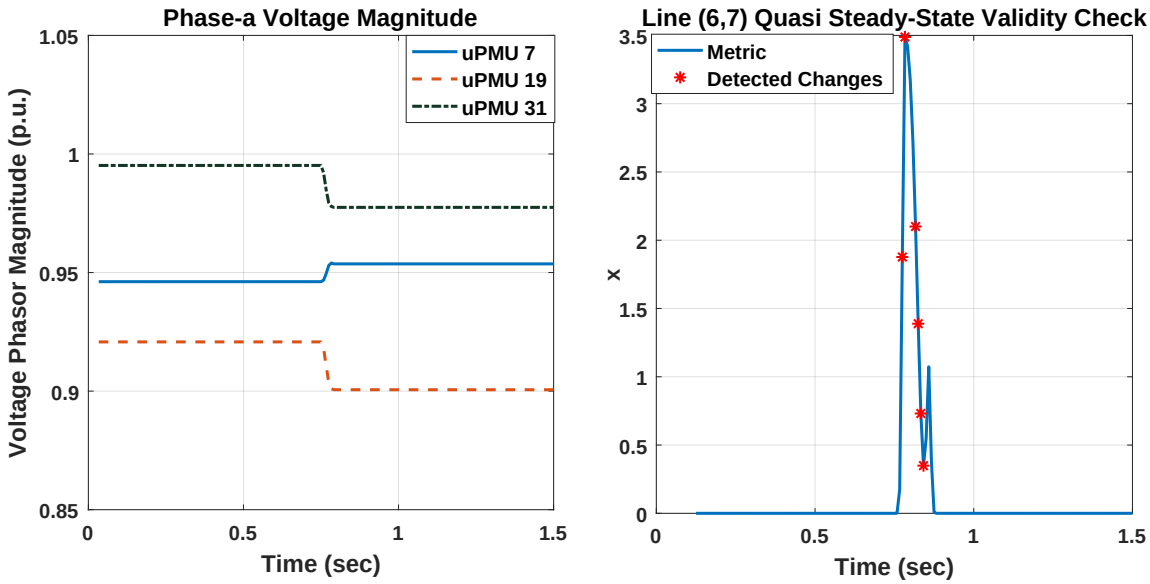



Figure 3.4.6: Voltage Magnitude and Quasi Steady-State Validity Check Rule for Feeder Reconfiguration.

### 3.4.1.3 Case Study on Data Injection Attack

We illustrate here the resilience of the architecture to data injection attacks. For this purpose, three data attack scenarios are investigated happening concurrently with the SLGF event discussed previously. We consider the case of the attacker manipulating the data of $\mu$PMU 7 in the first case, and $\mu$PMU 19 in the second case, and finally $\mu$PMU 7 and 19 at the same time in the third case on their way to the central engine. In all cases the data injected are a *replay* of the last available data set before the anomaly starts. Setting the change detector parameters fixed for all the three cases, Fig. 3.4.7 shows, for each case, the central rule and the start time of the detected changes. As can be seen, since $\mu$PMU 19 is playing an important role for this

event, having $\mu$PMU 7 compromised will not affect our central rule significantly (case-1). However, when the $\mu$PMU 19 is compromised, the number of detected changes reduce significantly (case-2), and when both $\mu$PMU 7 and 19 are compromised and the only healthy data is coming from $\mu$PMU 31, the detector does not pick any fast changes based on the set parameters. This also reveals the importance of tuning the detector thresholds to have a certain "false alarm," while maximizing the "detection" probability. We wish to remark that in all the three cases, the local analytics that

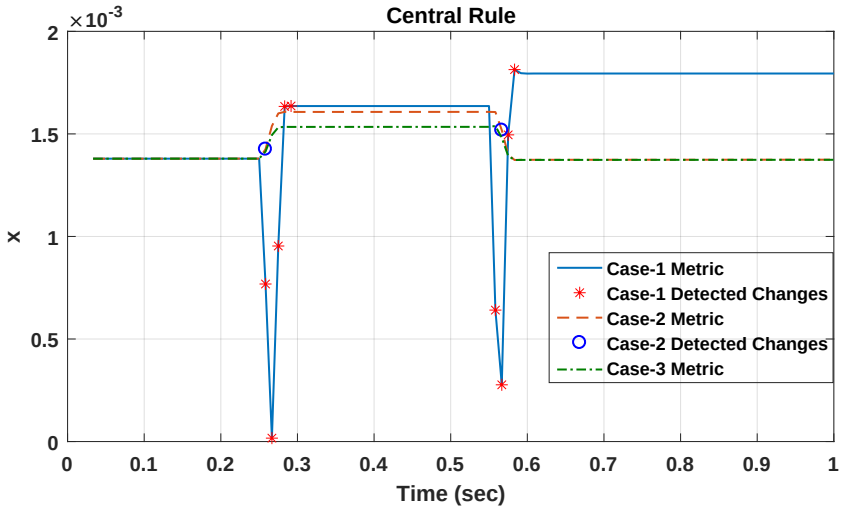



Figure 3.4.7: Central Rule for SLGF with Manipulated $\mu$PMU Data.

directly draw data from $\mu$PMUs will still flag the alarm, so buffering locally at the site of the event these data can be an important way of helping understand what communications were compromised in an ex-post analysis.

#### 3.4.1.4 Optimal versus Non-Optimal Placement

As it was mentioned, the placement criterion tries to scatter the available $\mu$PMUs over the grid to achieve the maximum coverage, and therefore make the central rule more sensitive to anomalies.

In order to compare the performance of an optimal versus non-optimal placement, a load loss event is created on bus 24 of IEEE-34 test case at $t = 0.4$s. The $\mu$PMUs are placed based on the random placement and greedy search result given in Table. 3.2, corresponding to non-optimal and optimal placement, respectively. Fig. 3.4.8 shows the central metric for these two cases. Since the relative change is what matters to our detector, the metrics corresponding to two placements are brought on the same scale. As it can be observed, since the $\mu$PMUs in the random placement are concentrated at a certain area, events like this would not be very pronounceable in the central metric, which could possibly lead to a "false-negative".

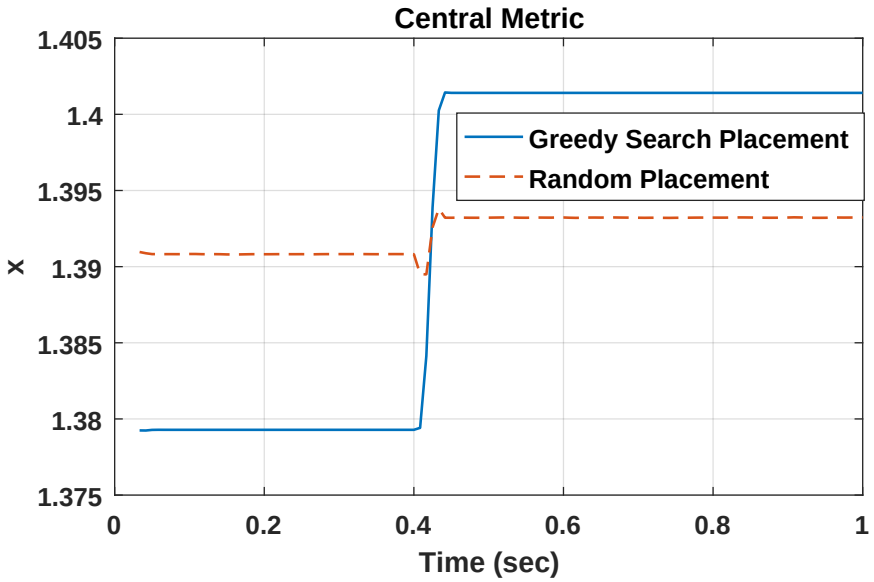



Figure 3.4.8: Central Metric Change of a Load Loss Event; Optimal versus Non-optimal Placement.

### *3.4.2 Real Data*

Fig. 3.4.9 shows the abstract one-line diagram of the partner utility grid and the location of the installed $\mu$PMUs. The installed $\mu$PMUs sample the voltage and current with a rate of $512 \times 60$ Hz, and output the estimated phasors at 120 Hz rate.

These devices achieve an accuracy of 0.001 deg resolution for phasor angle, 0.0002% for phasor magnitude, and 0.01% for Total Vector of Error (TVE) [Power Standards Lab(2016)]. The two feeders here are connected through the subtransmission grid.

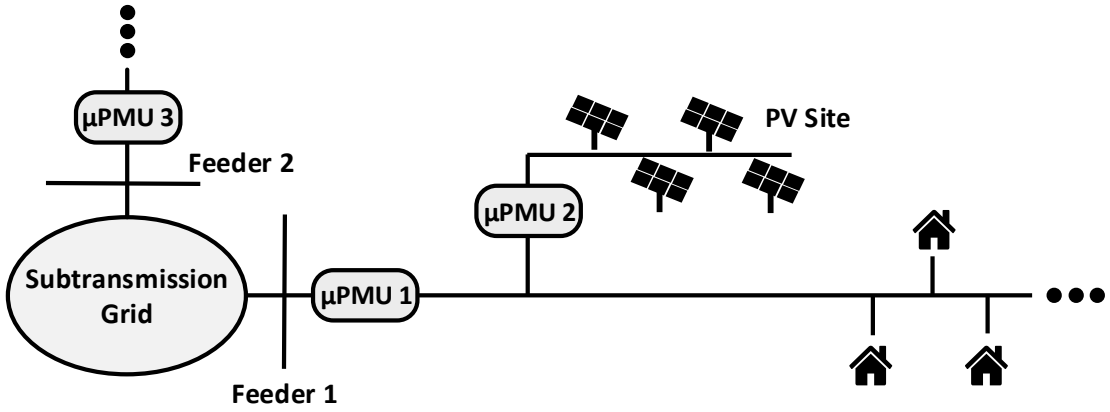



Figure 3.4.9: Location of Installed $\mu$PMUs in Our Partner Utility Grid.

Fig. 3.4.10 shows the voltage magnitude change rule inspected on the data from these $\mu$PMUs over a certain time. The results of the fast change inspection on the current magnitude of phase a and instantaneous local frequency at the substation bus of feeder 2 are also illustrated in Fig. 3.4.11(a) and Fig. 3.4.11(b), respectively. Observing the results of the analysis, the DSO can deduce that the cause of the event is most probably located on feeder 2. Also, from the pre- and post-anomaly value of the current magnitude, the DSO can conclude that some of the loads on this feeder tripped due to the voltage sag. We just showed the results from some of the rules due to the space limit but other rules can also flag the existence of anomaly on feeder 2.

All the metrics introduced and tested above are designed considering the specifications of a distribution grid. It should be noted that not all the proposed methods in the literature for transmission grid are applicable in the distribution side. For example, the phase angle difference in the distribution grid is known to be much smaller than that in the transmission level. Therefore, as opposed to transmission grid that

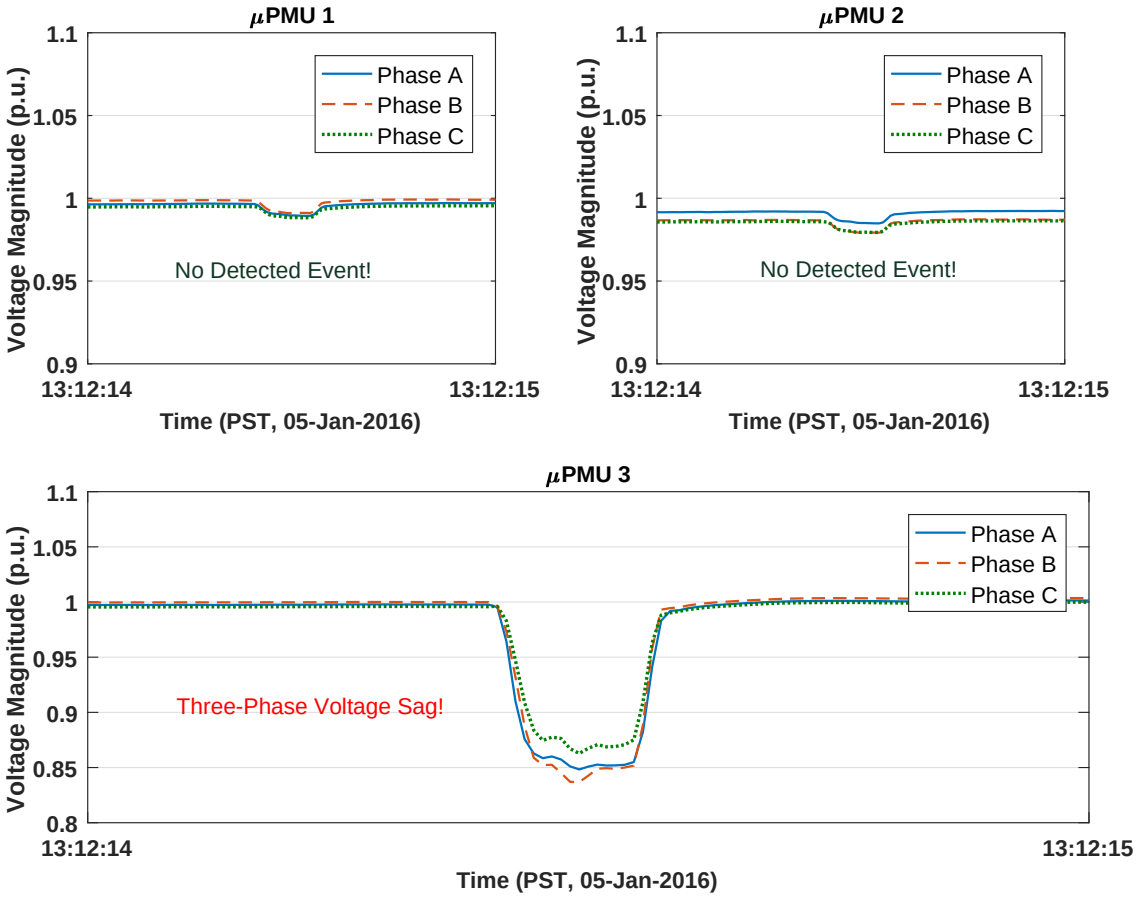



Figure 3.4.10: Voltage Magnitude Change Rule for Real Data.

this metric would work well for event detection [Allen *et al.*(2014)], it might not be a proper metric to look at in the distribution grid, since the signal to noise ratio might be small. The example next illustrates how the voltage angle difference metric could have failed if it was used as a "local rule". Using the data from two $\mu$PMUs installed at two ends of a line at a second utility grid (not the one in Fig. 3.4.9), the voltage magnitude captured by the $\mu$PMU at one end of the line is shown in Fig. 3.4.12(a) and the voltage phasor angle difference between the two $\mu$PMUs at two ends of the line is shown in Fig. 3.4.12(b). As it can be observed, the important event in this period corresponds to the two voltage sags. However, the angle difference shows a significant number of spikes without clearly marking these two events with the same

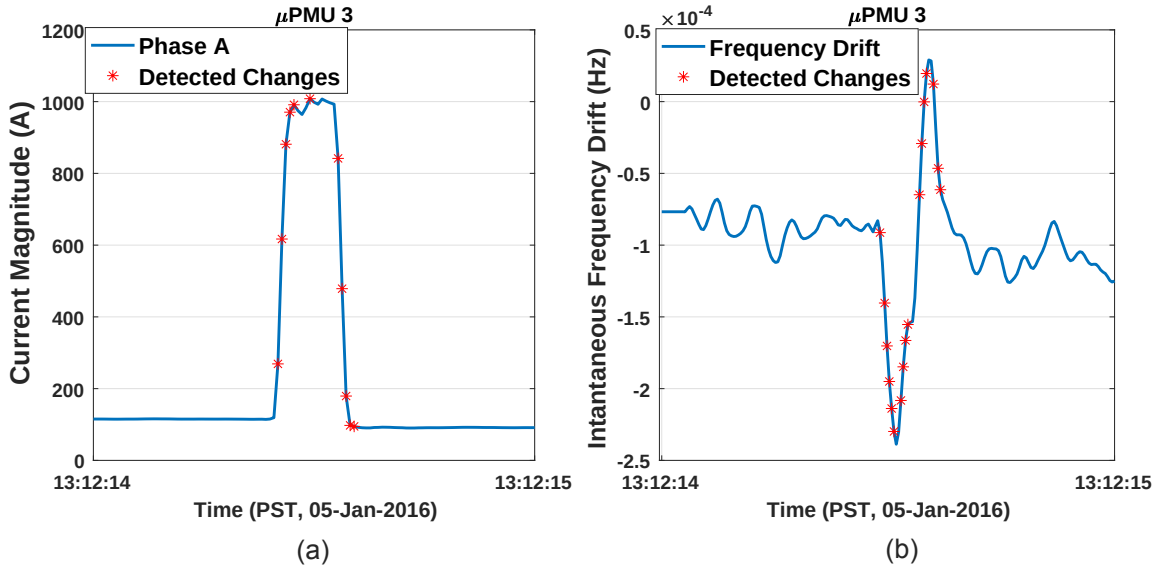



Figure 3.4.11: Fast Change Tracking of Current Phasor Magnitude and Bus Instantaneous Frequency Drift for Real Data.

significance. Passing this metric to our detector, we would pick too many fast changes that do not represent any specific event of interest, which therefore means an increase in the number of "false positives". All streams of interest were examined during this period, including both the active and reactive power, and all anomalies detected were in agreement with those visible in Fig. 3.4.12(a).

#### 3.4.2.1 Case Study-Robustness Against Volatility

The increasing presence of renewable resources, like wind and solar, in the distribution grid may raise some concerns about the robustness of the proposed rules to "false alarms" due to their inherent volatility. However, our method is based on an adaptive estimation of the data mean using an exponential window and, as shown in the following numerical example, this makes the algorithm capable to only pick the fast changes over a very short period of time, while being insensitive when the rate of change is in the same order of that of renewable resources fluctuations. To show it

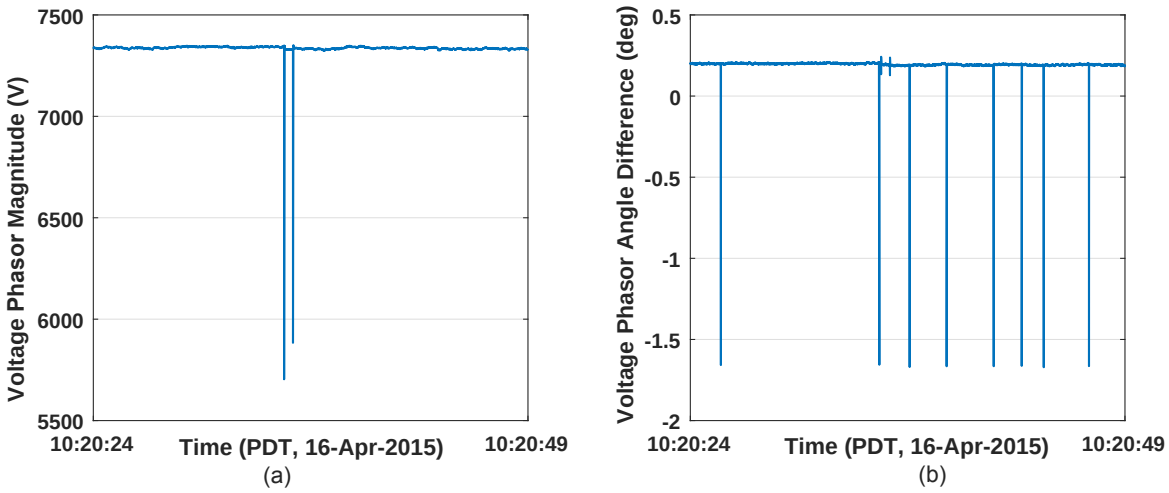



Figure 3.4.12: Voltage Phasor Angle Difference Between Two $\mu$PMUs at Distribution Grid.

numerically, we refer to two real events captured by $\mu$PMU 2 in the grid in Fig.3.4.9. Fig. 3.4.13(a) shows a dramatic change of about -18.25% over 9 seconds in the PV site active power injection due to the cloud effects, and Fig. 3.4.13(b) shows a step change in the active power after the PV site went out of service. Setting the detection parameters to be the same for both events, it can be seen that the detector flags no event for the first case and finds multiple of fast changes in the second. This is ideal since the events of our interest in the this context are of the second type, which should be discriminated from the normal quasi steady-state behavior in the first case.

Considering the existence of different types of loads in the distribution grid, the detector should also be able to differentiate between possible normal fluctuations in the load profile and those that are caused by a rare event. To show the performance of CUSUM for this case, real data of a $\mu$PMU installed behind a building with a non-linear load is used. As it can be seen in Fig. 3.4.14, the fast change tracker of the current magnitude does not pick the fluctuations due to the non-linear load behavior as an event and only flags a segment of data that corresponds to a voltage sag in the

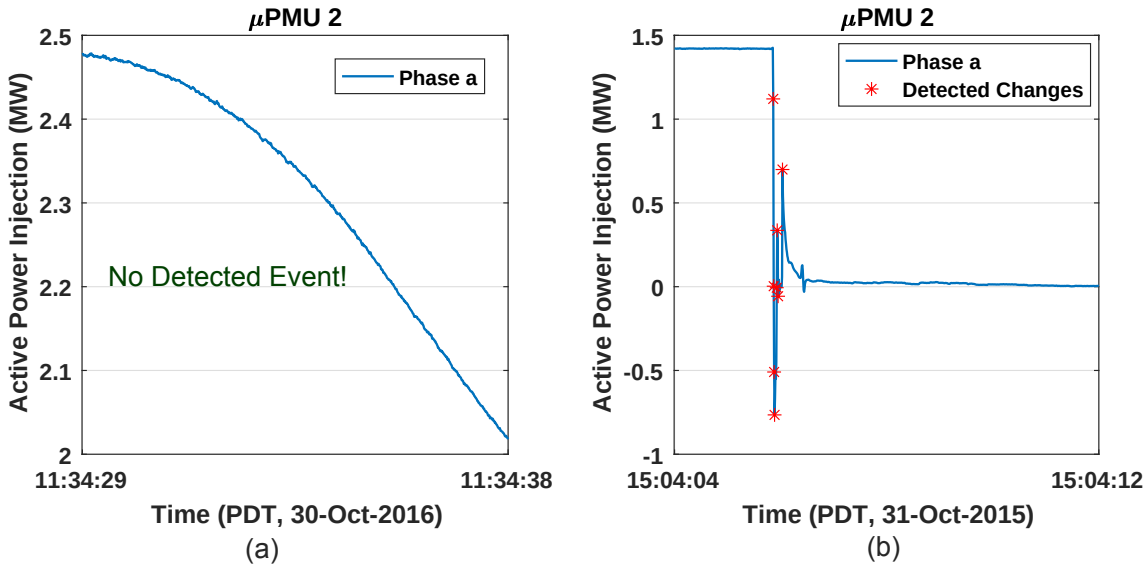



Figure 3.4.13: Fast Change Tracking of Injected Active Power in the PV Site.

grid as an eventful segment.

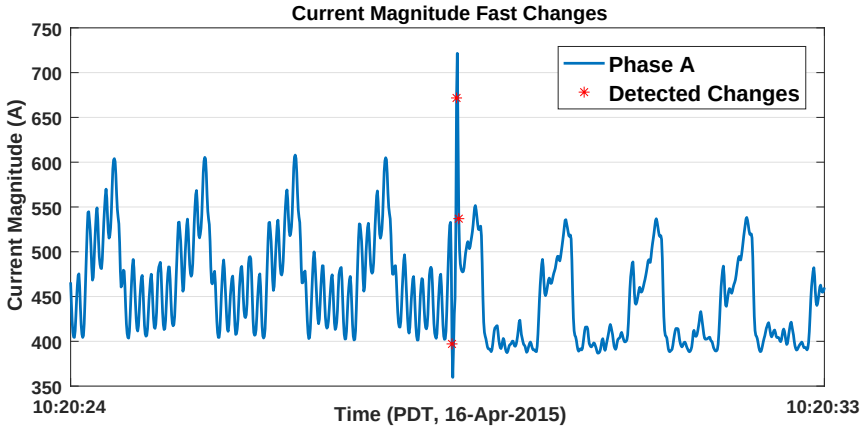



Figure 3.4.14: Fast Change Tracking of Current Magnitude for a Non-Linear Load.

# CHAPTER 4

# EVENT LOCALIZATION FOR CYBER INTRUSION DETECTION

Even though event detection is an integral part of our cyber-security framework, it is not sufficient for intrusion detection most of the times. To be able to correlate the SCADA data with the analytics form $\mu$PMUs, one needs to localize the source of the event to make sure which SCADA packets need to be checked. Inspired by this requirement, we focus on localizing the source of two important types of events.

The first part of the chapter (sections 4.1-4.2) investigates methodologies to localize a "reconnaissance activity" on the normally-open switch in the substation. In this section, we propose methods using phasor measurement data to track Thevenin parameters at substations delivering power to both an unbalanced and balanced feeder. We then show how tracking Thevenin source impedance can be employed in identifying cyber attackers performing "reconnaissance" in a distribution substation.

With the goal to detect cyber intrusion on the fault detectors that are part of a Fault Location, Isolation, and Service Restoration (FLISR) system, we develop fault localization method, where a fault can be localized up to a cluster of neighboring nodes when there is very few sensors in the grid. We prove that these clusters can be inferred from the properties of the network and where the sensors are placed. Having that identified, we propose an optimal PMU placement to achieve fault localization with highest resolution over the grid. The method is integrated in our cyber-physical security framework, where the results of PMU-based fault localization is correlated with the SCADA-reported fault location to check the integrity of the SCADA data.

## 4.1 Thevenin Equivalent Estimation

Fig. 4.1.1 shows the Thevenin equivalent circuit as seen from bus 1. This could be a subtransmission or distribution substation. Due to the integration of PMUs at the transmission level and $\mu$PMUs at the distribution level, measurements of the three-phase voltage phasor, $\mathbf{v}[k]$, and the three-phase current phasor, $\mathbf{i}[k]$, are now available at bus 1 at a very high rate. In steady-state, the following equation holds

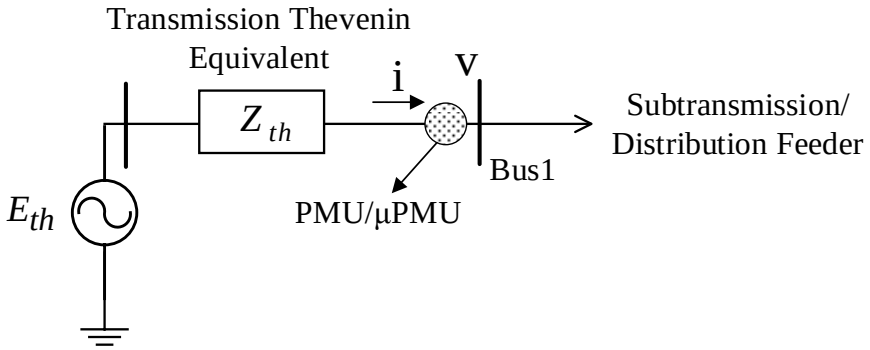



Figure 4.1.1: Transmission Grid Thevenin Equivalent Seen from Substation

for the circuit:

$$\mathbf{v}[k] = \mathbf{E}_{th}[k] - \mathbf{Z}_{th}[k]\mathbf{i}[k] \tag{4.1.1}$$

Equivalently, in the sequence domain, we have:

$$\mathbf{v}_s[k] = \mathbf{E}_s[k] - \mathbf{Z}_s[k]\mathbf{i}_s[k] \tag{4.1.2}$$

Assuming transposed lines implies that the mutual coupling between the zero, positive and negative sequence impedance at transmission level is zero. Hence, $\mathbf{Z}_s$ has the following form:

$$\mathbf{Z}_s = \begin{bmatrix} Z_0 & 0 & 0 \\ 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \end{bmatrix} \tag{4.1.3}$$

In addition, since $\mathbf{E}_{th}$ is representing the generators voltage at the transmission, it can be considered to be balanced, and therefore can be written in the sequence domain as:

$$\mathbf{E}_s[k] = \begin{bmatrix} 0 \\ E_1[k] \\ 0 \end{bmatrix} \tag{4.1.4}$$

Therefore, we obtain the well-known decoupled set of equations as follows:

$$\begin{bmatrix} v_0[k] \\ v_1[k] \\ v_2[k] \end{bmatrix} = \begin{bmatrix} 0 \\ E_1[k] \\ 0 \end{bmatrix} - \begin{bmatrix} Z_0[k] & 0 & 0 \\ 0 & Z_1[k] & 0 \\ 0 & 0 & Z_2[k] \end{bmatrix} \begin{bmatrix} i_0[k] \\ i_1[k] \\ i_2[k] \end{bmatrix}. \tag{4.1.5}$$

### *4.1.1 Unbalanced Grid*

Clearly, from (4.1.5), if the feeder connected to the substation is unbalanced (that is the case most of the times in the distribution feeders), all the Thevenin parameters can be estimated at each instant of time using the fact that $Z_1[k] \approx Z_2[k]$:

$$\begin{aligned} Z_0[k] &= -\frac{v_0[k]}{i_0[k]}, \quad Z_1[k] = -\frac{v_2[k]}{i_2[k]}, \\ E_1[k] &= \frac{v_1[k]i_2[k] - v_2[k]i_1[k]}{i_2[k]}. \end{aligned} \tag{4.1.6}$$

In this formulation we are taking full advantage of the inherent unbalanced behavior of the distribution grid that is usually considered a complication for many other applications. In this context, this feature allows us to update our estimation with the same frequency as the $\mu$PMU output rate. In addition, we are not only estimating the positive sequence impedance but also the zero sequence impedance. It should be noted that in some cases such as when the $\mu$PMU is placed after a Wye-Delta distribution transformer, the zero sequence current, $i_0$ is zero at point of measurement, which means that the zero sequence source impedance is infinite as expected.

### 4.1.2 Balanced Grid

When the grid is balanced, which is usually the case at the substations connecting the transmission grid to a subtransmision circuit, the method discussed above does not work. In this case, the only non-trivial equation in (4.1.5) is:

$$v_1[k] = E_1[k] - Z_1[k]i_1[k] \tag{4.1.7}$$

Clearly, what we have in (4.1.7) is a single equation with two unknowns, $E_1[k]$ and $Z_1[k]$, which does not have a unique solution. Therefore, we need to make appropriate approximations. Without loss of generality, we set the voltage phasor angle in (4.1.7) to zero and accordingly all the angles are reported relative to the voltage phasor angle. We drop the subscripts in (4.1.7) for simplicity of notation in the rest of this section. They will be reintroduced in the next sections.

Replacing the load in Fig. 4.1.1 with an equivalent apparent load impedance $Z_L = R_L + jX_L$ and neglecting the Thevenin resistance of the transmission grid in comparison with the Thevenin reactance, the corresponding phasor diagram is illustrated in Fig. 4.1.2. It is worth to mention here that the effect of off-nominal

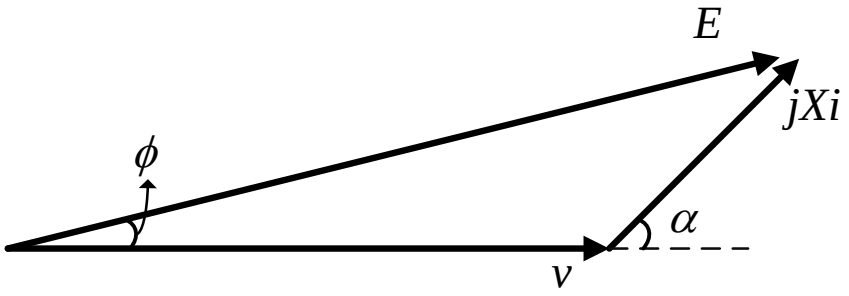



Figure 4.1.2: Phasor Diagram of the Equivalent Thevenin Circuit in Fig. 4.1.1 for Balanced Grid.

frequency is an equal rotation in the vectors of the phasor diagram. Since we report the phasor angles relative to the voltage, this modulation term is accordingly

removed from the voltage, current and the modeled Thevenin voltage source. We propose to use the following relationship based on the phasor diagram in Fig. 4.1.2. Let $A[k] = |E[k]|$ and $i_{im}$ be the imaginary component of the current, then we have:

$$\underbrace{A^2[k] - v^2[k] - X^2[k]|i[k]|^2 + 2i_{im}[k]X[k]v[k]}_{r(A,X;k)} = 0 \tag{4.1.8}$$

At each instant of time, there are infinite solutions for $A$ and $X$ as zeros of the "residual function" $r$. Assuming that $A$ and $X$ are constant over a window of $M$ samples, and there is sufficient change in the feeder over this window, we form the following over-determined homogeneous set of equations:

$$\underbrace{\begin{pmatrix} r(A,X;k-M+1) \\ r(A,X;k-M+2) \\ \vdots \\ r(A,X;k) \end{pmatrix}}_{\vec{\mathbf{r}}(A,X;k)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4.1.9}$$

Note that we do not assume the angle $\phi$ is constant, and the assumptions made pertain exclusively to the Thevenin voltage magnitude and source reactance, which are considered constant over the examined window. Having constant voltage magnitude is reasonable since the Thevenin voltage source is representing the bulk grid, whose automatic voltage regulators operate on time-scales longer than those examined here. Our problem now translates into solving $M$ non-linear quadratic equations. Taking the squared-norm of the residual vector, $\vec{\mathbf{r}}(A,X;k)$, in (4.1.9), we form the following objective function:

$$f(A,X;k) = \frac{1}{2}||\vec{\mathbf{r}}(A,X;k)||^2 \tag{4.1.10}$$

which is aimed to be minimized:

$$\min_{\boldsymbol{\theta}[k]} f(A,X;k) \tag{4.1.11}$$

where $\boldsymbol{\theta}[k] = [A[k], X[k]]^T$. The Levenberg-Marquardt Algorithm (LMA) is used to solve the non-linear least square problem. This method replaces the line search method in Gauss-Newton with a trust region strategy. The LMA overcomes one of the weaknesses of Gauss-Newton, i.e., the behavior when the Jacobian matrix of the residual vector is rank-deficient, or is close to be rank-deficient [Nocedal and Wright(2006)]. This is important in tracking the Thevenin parameter online since the change in the data during some intervals may return a Jacobian matrix which is ill-conditioned. The other advantage of the LMA as opposed to the Gauss-Newton is when the initial guess is far from the minimizer. In this case, Gauss-Newton converges very slowly or may not converge at all but LMA is proved to perform better under certain conditions. The implementation of the LMA in our system is shown in Algorithm. 4.1.

Note that in calculating $\mathbf{P}_{LM}$, we have used an ellipsoidal trust region in place of a spherical trust region by using the term $\lambda \text{diag}(\mathbf{J}^T\mathbf{J})$ instead of $\lambda\mathcal{I}$ as suggested in [Seber and Wild(1989)], where $\mathcal{I}$ is the identity matrix. The reason is that our problem is poorly scaled, or in other words the range of $A$ and $X$ are so different, and therefore we may face numerical difficulties or produce solutions with poor quality. The initial guess at time instant $k$ can be provided by the estimated parameters at time instant k-1, i.e., $\boldsymbol{\theta}_0[k] \leftarrow \boldsymbol{\theta}[k-1]$.

Compared to the methods in [Parniani *et al.*(2006), Arefifar and Xu(2009)], our formulation leads to a smaller size of the Jacobian matrix, and therefore is computationally more efficient. Also, the mentioned advantages of LMA over the Gauss-Newton method used in [Arefifar and Xu(2009)] still remains.

**Algorithm 4.1:** Levenberg-Marquardt Algorithm (LMA) at time instant $k$

**Input:** $\vec{\mathbf{r}}(A, X; k)$, and an initial guess $\boldsymbol{\theta}_0[k]$

**Output:** Thevenin parameters at time $k$

**begin**

flag=1;

initialize $\rho < 1$, $\lambda$, and $\epsilon$ ;

$\boldsymbol{\theta}[k] \leftarrow \boldsymbol{\theta}_0[k]$;

**while** *flag==1* **do**

$\mathbf{J} = \nabla \vec{\mathbf{r}}(\boldsymbol{\theta}; k)$;

$\mathbf{P}_{LM} = -(\mathbf{J}^T\mathbf{J} + \lambda \text{diag}(\mathbf{J}^T\mathbf{J}))^{-1}\mathbf{J}^T\vec{\mathbf{r}}(\boldsymbol{\theta}; k)$;

$\boldsymbol{\theta}_{new}[k] \leftarrow \boldsymbol{\theta}[k] + \mathbf{P}_{LM}$;

**if** $f(\boldsymbol{\theta}_{new}; k) < f(\boldsymbol{\theta}; k)$ **then**

$\lambda \leftarrow \rho\lambda$;

$\boldsymbol{\theta}[k] \leftarrow \boldsymbol{\theta}_{new}[k]$;

**else**

$\lambda \leftarrow \frac{\lambda}{\rho}$;

**if** $f(\boldsymbol{\theta}; k) < \epsilon$ **then**

flag $\leftarrow 0$;

$\phi[k] = \sin^{-1}(X[k]i_r[k]/A[k])$

**return** $E[k]$, $X[k]$;

## 4.2 Thevenin Estimation Results

In this section, we present the performance of the proposed Thevenin estimation through the numerical results for both cases of unbalanced and balanced grid.

### *4.2.1 Unbalanced Grid Thevenin Parameter Estimation*

We modified IEEE-34 test case for our simulation by connecting it to a voltage source through a lumped transmission line representing the Thevenin impedance as shown in Fig. 4.2.1. The simulation is done in the DIgSILENT software [Manual and PowerFactory(2009)] with phasor report rate of 120 Hz. We intentionally imposed

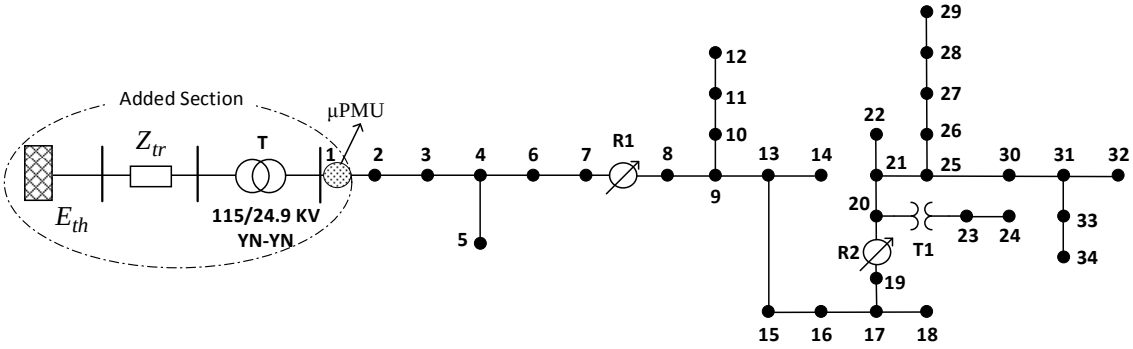



Figure 4.2.1: Modified IEEE-34 Bus Test Case One-Line Diagram.

the grid to work at off-nominal frequency of 59.9 Hz to simultaneously examine the effect of the off-nominal frequency on our estimation performance. We expect that our methodology for unbalanced grid will not be affected by off-nominal frequency because the estimation at each instant of time only uses the samples at that time. Table. 4.1 shows the estimated values for positive and zero sequence versus the actual ones. The estimated values are very close to the actual, which indicates the successful performance of the estimation. Fig. 4.2.2 shows the magnitude and the angle of the estimated Thevenin voltage versus the actual values captured from the simulation. As can be seen, the estimation of the Thevenin voltage is successful, and our algorithm tracks the parameters even when the off-nominal frequency is non-negligible, which

Table 4.1: Actual vs Estimated Thevenin Sequence Impedance

| | Estimated | Actual |
|---|---|---|
| $Z_0$ | $2.5533 + j9.4392$ | $2.5716 + j9.4320$ |
| $Z_1$ | $2.9922 + j10.92$ | $2.99 + j10.8901$ |

is clear from the voltage angle.

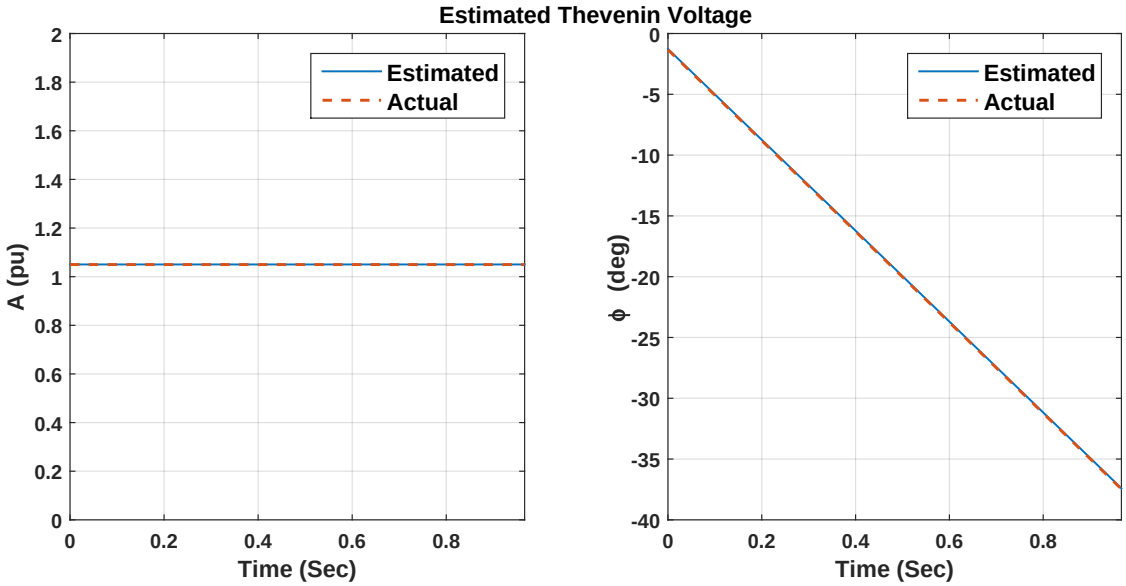



Figure 4.2.2: Estimated Thevenin Voltage for Modified IEEE-34 Bus Under Off-Nominal Frequency.

### *4.2.2 Balanced Grid Thevenin Parameter Estimation*

In balanced case, we first simulated the 39-Bus New-England dynamic model [Pai(2012)] using DIgSILENT software with a phasor report rate of 120 Hz. A very slow load ramping event of +2%/sec is defined at load 16, while other loads are kept constant. The reason of the slow ramping event is to show how our method is suitable even under slow changes that is necessary for online application since large changes may not always happen over a short window. Taking a sliding window

of 30 samples, i.e., $M = 30$, Fig. 4.2.3 shows the estimated versus actual value of the Thevenin impedance seen from bus 16. This window corresponds to 0.25 sec, where the assumption of constant Thevenin voltage magnitude is valid with a very good approximation. As can be seen, the estimates are very close to the actual values,

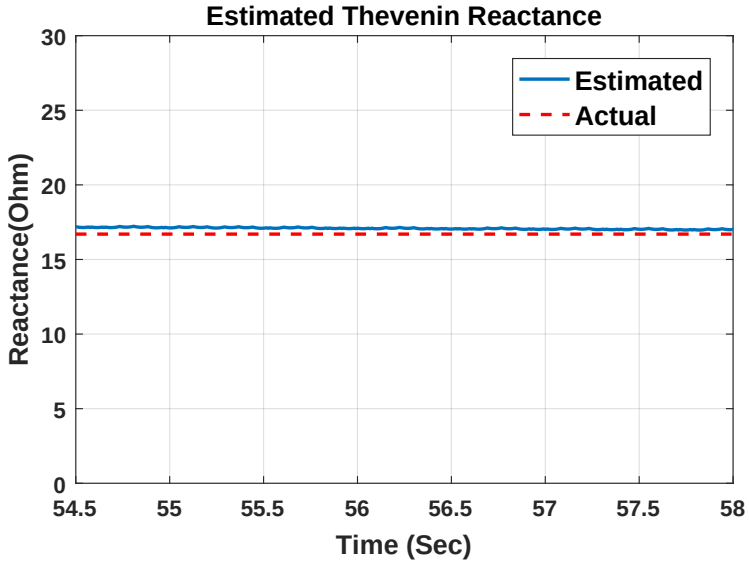



Figure 4.2.3: Estimated Thevenin Reactance Seen from Bus 16 of New England Test Case Using LMA Method.

which indicate that the LMA method can successfully track the Thevenin parameters. We attribute the small residual error to the fact that we neglect the resistance.

Using the same sliding window size, the performance of the linear least-square method under this slow rate of change is illustrated in Fig. 4.2.4. As shown, this method completely fails to track the actual value of the Thevenin reactance in this scenario. The reason is most probably due to the fact that the phase angle is assumed to be constant over the window of samples.

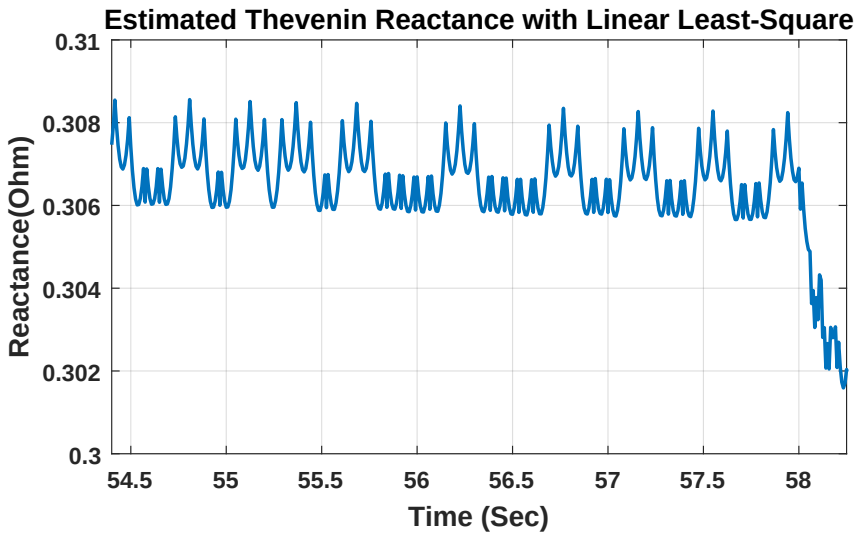



Figure 4.2.4: Estimated Thevenin Reactance Seen from Bus 16 of New England Test Case Using Linear Least Square Method.

### *4.2.3 Voltage Stability Index Monitoring*

An important application of the online Thevenin estimation algorithm is to monitor the voltage stability index at a load point in the transmission level. We use the voltage stability index in (4.2.1) that is employed in [Vu *et al.*(1999)] to test the effectiveness of our Thevenin tracking method for a balanced grid. The index is defined as follows:

$$L[k] = \frac{X_1[k]}{|Z_L[k]|} = \frac{X_1[k]|i_1[k]|}{v_1[k]} \tag{4.2.1}$$

where $Z_L[k]$ is the apparent load impedance at time $k$. Based on (4.2.1), the stable and unstable regions are defined as follows:

$$\begin{cases} \text{Stable Region} & \text{if } L[k] \leq 1 \\ \text{Unstable Region} & \text{if } L[k] > 1 \end{cases} \tag{4.2.2}$$

For the numerical result, a ramping load event with load steps of +18%/sec is introduced at all the load buses of the New-England 39 Bus test case until the grid

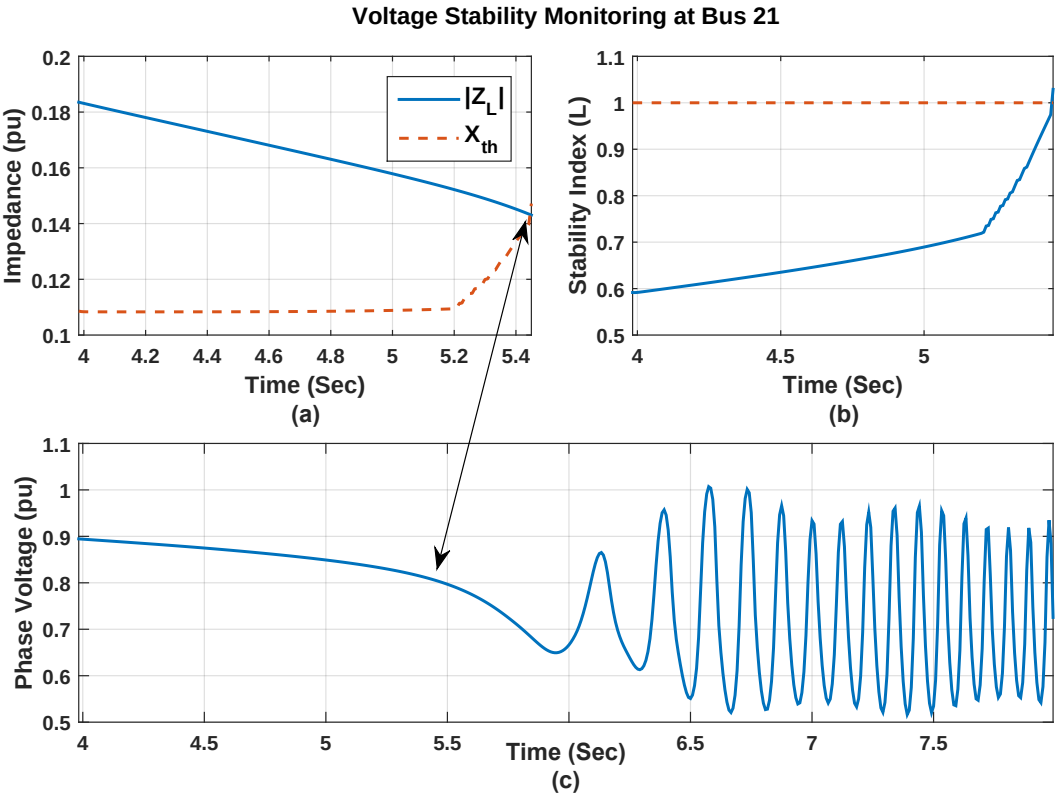



Figure 4.2.5: Voltage Stability Index Monitoring at Bus 21.

goes unstable. Fig. 4.2.5(a) shows the estimated Thevenin impedance and the magnitude of the apparent impedance. The voltage stability index (L) is also shown in Fig. 4.2.5(b). From the figure it is evident that the oscillation in the voltage starts when the magnitude of the apparent impedance meets the estimated Thevenin impedance or, equivalently, when the voltage stability index (L) crosses 1. This point corresponds to the maximum loadability of the system and the grid goes unstable beyond that. This is another indication that the Thevenin parameter is being tracked correctly, due to $L$ correctly determining the point of maximum loadability.

### *4.2.4 Reconnaissance Activity Identification*

In this use-case we assume that the substation connected to the distribution feeder is equipped with a normally-open spare transformer parallel to the main transformer, to restore the delivery of the power to the feeder when the main transformer fails or

requires maintenance. Inserting the spare transformer into service, while the main transformer is also in service, has no adverse effect on the delivery of the power to the feeder downstream. It is therefore appealing for an attacker to try and test its ability to control the network through a compromised SCADA network, by changing the service status of the backup transformer switch when the main transformer is in service. The attacker may then mask this control action by spoofing the SCADA packages sent to the control center. The question is "can we detect such an activity with $\mu$PMU data?"

Suppose that we have a $\mu$PMU placed at the head of distribution feeder as shown in Fig. 4.2.6. In this circuit, the dominant term in the Thevenin source impedance would be due to the transformer impedance. Therefore, when the spare transformer switch is toggled maliciously, while the main transformer is in service, the source impedance will roughly reduces to half of its normal value. We propose to track the fast changes in the estimated positive sequence of the Thevenin source impedance. When a fast change is found, the magnitude of the change can be extracted to determine whether it can be attributed to the spare transformer switch or not. From Fig. 4.2.7, when

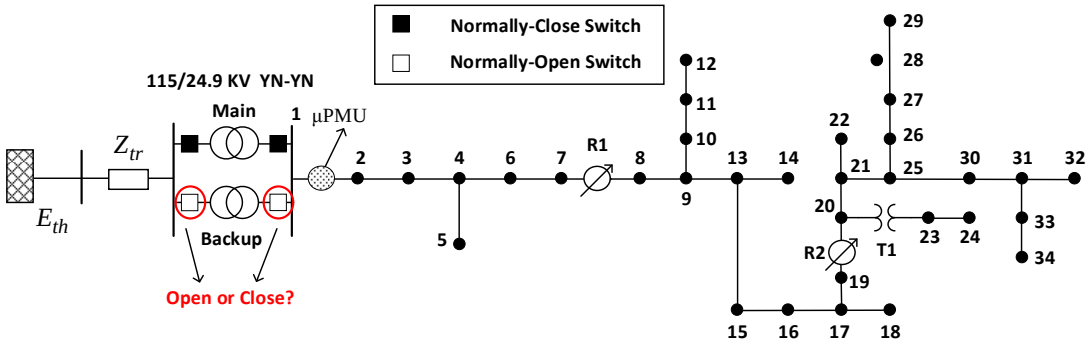



Figure 4.2.6: Modified IEEE-34 Test Case for Reconnaissance Attack Identification

the switch gets closed the estimated impedance is almost half of what it was before. Considering the fact that the dominant term is due to the transformer and such a change cannot be the consequence of change in the transmission grid topology,

the most plausible cause is the event that the spare transformer switch was closed. Once such a change has been detected, the control center is notified to determine if the change was a scheduled switch or an unexpected switch, potentially indicating malicious behavior.

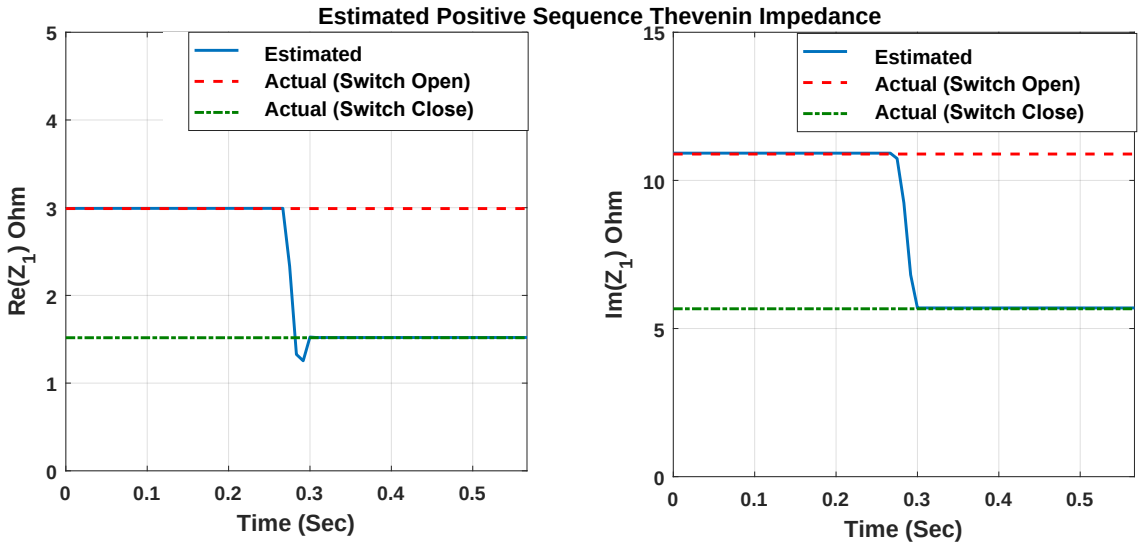



Figure 4.2.7: Online Tracking of the Thevenin Source Impedance using $\mu$PMU Measurements

## 4.3 The Fault Localization Problem

*Notation:* In this part of the chapter: 1) $\mathcal{I}_N$ $N \times N$ identity matrix. 2) $\mathbf{A}^T, \mathbf{A}^H$ transpose and conjugate transpose of $\mathbf{A}$; 3) $||\mathbf{a}||$ is the 2-norm of vector $\mathbf{a}$; 4) $||\mathbf{A}||_F$ is the Frobenius norm of matrix $\mathbf{A}$; 5) $|\mathbf{A}|, tr(\mathbf{A})$ are, respectively, determinant and trace of $\mathbf{A}$. 6) $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for $\mathbf{x}$ multivariate normal with expected value $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

Fault localization is a two-step process that first detects a fault and then seek to determine the location of said fault. This part of the cahpter focuses on the latter, when a fault has already been detected (for instance using the method we presented in [Jamei *et al.*(2017b)], that also provides an estimate of the beginning and end of the transient event associated with the fault). To keep the formulation general for

balanced and unbalanced grids with transposed/untransposed lines, we use the three phase formulation of the Ohm's law (instead of the sequence domain) so that single and two phase lines can be considered. For a network of size $N$, let $\mathbf{V}$ and $\mathbf{I}$ denote the nodal voltage and injection current $M \times 1$ vectors:

$$\mathbf{V} = [V_1^T, V_2^T, ..., V_N^T]^T, \quad \mathbf{I} = [I_1^T, I_2^T, ..., I_N^T]^T$$

where $V_i$ and $I_i$ are column vectors of size 1, 2 or 3, depending on the number of phases connected to bus $i$. From Ohm's law:

$$\mathbf{I} = \mathbf{Y}\mathbf{V} \tag{4.3.1}$$

where $\mathbf{Y}$ is the admittance matrix. The sources in the grid are modeled with their Norton equivalent and their internal admittances are included in the $\mathbf{Y}$ matrix as well. We wish to remark that even though the formulation is general, all the numerical analyses are done using distribution grid test cases, and the algorithm has not been tested against transmission case in our work.

Let us denote the pre-fault voltage and current as $\mathbf{V}_0$ and $\mathbf{I}_0$. If we assume that a fault happens at bus $j$, we can decompose the nodal injection current post-fault as $\mathbf{I}_F + \mathbf{I}_E$, where $\mathbf{I}_E$ is a sparse vector containing the injected fault current at bus $j$ in its non-zero entries. The localization problem amounts to finding the most likely set of indexes for these non-zero entries, given the observed measurements.

From (4.3.1), $\mathbf{I}_0 = \mathbf{Y}\mathbf{V}_0$ and when the measurements have stabilized during fault, $\mathbf{I}_F + \mathbf{I}_E = \mathbf{Y}\mathbf{V}_F$ where $\mathbf{V}_F$ is the post-fault voltage[1]. Subtracting the pre-fault from the post-fault equation:

$$(\mathbf{I}_F - \mathbf{I}_0) + \mathbf{I}_E = \boldsymbol{\delta}\mathbf{I} + \mathbf{I}_E = \mathbf{Y}(\mathbf{V}_F - \mathbf{V}_0) = \mathbf{Y}\boldsymbol{\delta}\mathbf{V} \tag{4.3.2}$$

[1]The assumption is that the post-fault data is recorded before any corrective protection measure is taken so that the admittance matrix stays the same.

### 4.3.1 The Statistical Measurements Model

Let us parse the equations (4.3.2) into those that are associated with available PMU measurements and those that are not. Let $K$ denote the total number of phases, for which the nodal voltage and injection current are measured with PMUs. The matrix $\mathbf{\Pi}_a \in \{0,1\}^{K\times M}$ is a selection matrix that picks the available measurements in the voltage/current vector and $\mathbf{\Pi}_u \in \{0,1\}^{(M-K)\times M}$ selects the unavailable measurements. Pre-multiplying both sides of (4.3.2) by $\mathbf{\Pi} = (\mathbf{\Pi}_a^T \mid \mathbf{\Pi}_u^T)^T$, and replacing $\mathbf{Y}$ with $\mathbf{Y}\mathbf{\Pi}^{-1}\mathbf{\Pi}$ (noting that $\mathbf{\Pi}^{-1} = \mathbf{\Pi}^T$) we have:

$$\begin{aligned}
\begin{pmatrix}\mathbf{\Pi}_a\\ \mathbf{\Pi}_u\end{pmatrix}(\boldsymbol{\delta}\mathbf{I} + \mathbf{I}_E) &= \begin{pmatrix}\mathbf{\Pi}_a\\ \mathbf{\Pi}_u\end{pmatrix}\mathbf{Y}\begin{pmatrix}\mathbf{\Pi}_a^T & \mathbf{\Pi}_u^T\end{pmatrix}\begin{pmatrix}\mathbf{\Pi}_a\\ \mathbf{\Pi}_u\end{pmatrix}\boldsymbol{\delta}\mathbf{V}\\
\begin{pmatrix}\boldsymbol{\delta}\mathbf{I}_a\\ \boldsymbol{\delta}\mathbf{I}_u\end{pmatrix} + \begin{pmatrix}\mathbf{I}_E^a\\ \mathbf{I}_E^u\end{pmatrix} &= \begin{pmatrix}\mathbf{Y}_{aa} & \mathbf{Y}_{au}\\ \mathbf{Y}_{au}^T & \mathbf{Y}_{uu}\end{pmatrix}\begin{pmatrix}\boldsymbol{\delta}\mathbf{V}_a\\ \boldsymbol{\delta}\mathbf{V}_u\end{pmatrix}.
\end{aligned} \tag{4.3.3}$$

From (4.3.3), we obtain the two following set of equations:

$$\boldsymbol{\delta}\mathbf{I}_a + \mathbf{I}_E^a = \mathbf{Y}_{aa}\boldsymbol{\delta}\mathbf{V}_a + \mathbf{Y}_{au}\boldsymbol{\delta}\mathbf{V}_u \tag{4.3.4}$$

$$\boldsymbol{\delta}\mathbf{I}_u + \mathbf{I}_E^u = \mathbf{Y}_{au}^T\boldsymbol{\delta}\mathbf{V}_a + \mathbf{Y}_{uu}\boldsymbol{\delta}\mathbf{V}_u \tag{4.3.5}$$

Substituting $\boldsymbol{\delta}\mathbf{V}_u$ in (4.3.4) using (4.3.5) and reordering some terms:

$$\begin{aligned}
\mathbf{z} &= \overbrace{\left(\mathcal{I} \mid -(\mathbf{Y}_{aa} - \mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}\mathbf{Y}_{au}^T)\right)}^{\mathbf{H}}\overbrace{\begin{pmatrix}\boldsymbol{\delta}\mathbf{I}_a\\ \boldsymbol{\delta}\mathbf{V}_a\end{pmatrix}}^{\mathbf{s}}\\
&= \overbrace{(-\mathcal{I} \mid \mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1})}^{\mathbf{D}}\overbrace{\begin{pmatrix}\mathbf{I}_E^a\\ \mathbf{I}_E^u\end{pmatrix}}^{\mathbf{x}} + \overbrace{\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}}^{\mathbf{C}}\overbrace{\boldsymbol{\delta}\mathbf{I}_u}^{\boldsymbol{\epsilon}}
\end{aligned} \tag{4.3.6}$$

The term $\mathbf{C}\boldsymbol{\epsilon}$ can be viewed as a noise term, while the term $\mathbf{D}\mathbf{x}$, in which $\mathbf{x}$ is a sparse vector (i.e. the vector $\mathbf{I}_E$ reordered) whose non-zero entries point to the sites of the fault.

**Remark.** *Generally speaking, the noise term* $\mathbf{C}\boldsymbol{\epsilon}$ *will be relatively small, if the constant impedance loads and capacitors/reactors are also included in the* $\mathbf{Y}$ *matrix, instead of being modeled in the current injection vector. For constant power loads, their equivalent admittance at the nominal voltage can be included in the bus admittance matrix, and the deviation of the actual consumed power from the nominal can be included in the nodal injection vector, so that a similar model can be adopted when the assumption of constant power loads is more appropriate. This modeling implies that the vectors* $\mathbf{I}_F$ *and* $\mathbf{I}_0$ *are small and, accordingly, their difference* $\boldsymbol{\delta}\mathbf{I}$ *is small as well.*

Assuming that[2] $\boldsymbol{\epsilon} \sim \mathcal{CN}(\mathbf{0}, \sigma_\epsilon^2 \mathcal{I})$, whitening the noise term is appropriate. Let the singular value decomposition of $\mathbf{C}$ be:

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{W}^H \tag{4.3.7}$$

where $\mathbf{U}$ is a $(K \times \tilde{K})$, $\tilde{K} \leq K$ containing the columns spanning the column space of $\mathbf{C}$, $\boldsymbol{\Sigma}$ is a diagonal matrix with non-zero singular values of $\mathbf{C}$ on its diagonal entries and $\mathbf{W}$ is of size $(M - K) \times \tilde{K}$ containing the columns spanning the row space of $\mathbf{C}$. Pre-multiplying $\mathbf{z}$ by $\boldsymbol{\Sigma}^{-1}\mathbf{U}^H$, we have:

$$\mathbf{y} = \boldsymbol{\Sigma}^{-1}\mathbf{U}^H\mathbf{z} = \boldsymbol{\varepsilon} + \mathbf{F}\mathbf{x} \tag{4.3.8}$$

where $\mathbf{F} = \boldsymbol{\Sigma}^{-1}\mathbf{U}^H\mathbf{D}$ and the noise $\boldsymbol{\varepsilon} = \mathbf{W}^H\boldsymbol{\epsilon}$ and, since $\mathbf{W}$ is semi-unitary, $\boldsymbol{\varepsilon} \sim \mathcal{CN}(\mathbf{0}, \sigma_\epsilon^2 \mathcal{I})$.

We formulate the fault location hypotheses testing problem based on (4.3.8). Let $\mathcal{H}_\ell$ denote the hypothesis that a fault has occurred at location $\ell \in \mathcal{F}_t$, where $\mathcal{F}_t$ is a set of candidate locations for a detected fault $t$. To better understand the structure of $\mathcal{F}_t$,

[2]The pseudo-covariance matrix is assumed to be zero for all complex normally distributed vectors so it is omitted from the definition for simplicity.

consider the three-phase diagram in Fig. 4.3.1 of a sample 5-bus system including 3 three-phase nodes and two single phase nodes: suppose that detected fault $t$ indicates

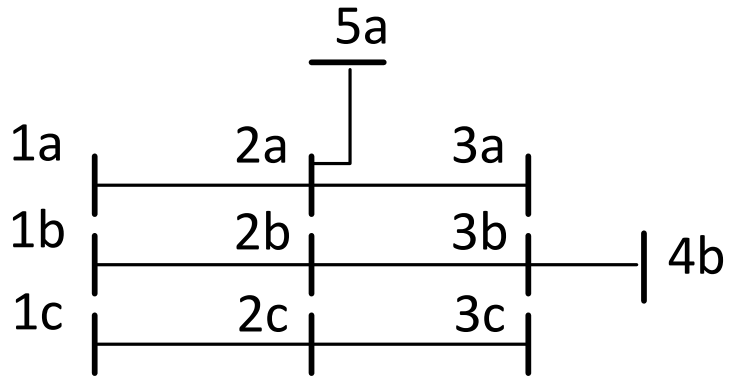



Figure 4.3.1: Sample Three-Phase Line Diagram.

a three phase fault. The candidate fault locations are:

$$\mathcal{F}_t = \big\{\{1a, 1b, 1c\}, \{2a, 2b, 2c\}, \{3a, 3b, 3c\}\big\}$$

and buses 4 and 5 are excluded since they are single phase.

Let us assume that the noise variance $\sigma_\epsilon^2$ is known. We also assume that: 1) conditioned on the hypothesis $\mathcal{H}_\ell$, the sparse vector $\mathbf{x}$ has its non-zero entries normally distributed. This can be captured by $\mathbf{x}|\mathcal{H}_\ell \sim \mathcal{CN}(\boldsymbol{\mu}_\ell, \boldsymbol{\Phi}_\ell)$, where $\boldsymbol{\mu}_\ell$ is a sparse vector containing the expected value for the fault current under hypothesis $\mathcal{H}_\ell$ at its non-zero entries and $\boldsymbol{\Phi}_l = \mathbb{E}[\mathbf{x}\mathbf{x}^H|\mathcal{H}_\ell]$ is a diagonal covariance matrix, which is only non-zero at entries corresponding to the candidate location $\ell$. Thus:

$$\mathbf{F}\mathbf{x}|\mathcal{H}_\ell \sim \mathcal{CN}(\mathbf{F}\boldsymbol{\mu}_\ell, \mathbf{F}\boldsymbol{\Phi}_\ell\mathbf{F}^H)$$

Assuming that $\mathbf{F}\mathbf{x}|\mathcal{H}_\ell$ and $\boldsymbol{\varepsilon}|\mathcal{H}_\ell$ are independent, we have:

$$\mathbf{y}|\mathcal{H}_\ell \sim \mathcal{CN}(\overbrace{\mathbf{F}\boldsymbol{\mu}_\ell}^{\mathbf{m}_\ell}, \overbrace{\sigma_\epsilon^2\mathcal{I} + \mathbf{F}\boldsymbol{\Phi}_\ell\mathbf{F}^H}^{\boldsymbol{\Psi}_\ell}) \tag{4.3.9}$$

Denoting by $f(\mathbf{y}|\mathcal{H}_\ell)$ the probability density function of $\mathbf{y}|\mathcal{H}_\ell$ and by $\lambda_\ell(\mathbf{y}) = \ln f(\mathbf{y}|\mathcal{H}_\ell)$ the log-likelihood function, the maximum likelihood (ML) detector of the fault location is:

$$\begin{aligned}\ell^* &= \underset{\ell\in\mathcal{F}_t}{\text{argmax}}\ \lambda_\ell(\mathbf{y}) \\ &= \underset{\ell\in\mathcal{F}_t}{\text{argmin}}\ (\mathbf{y}-\mathbf{m}_\ell)^H\mathbf{\Psi}_\ell^{-1}(\mathbf{y}-\mathbf{m}_\ell) + \ln(\pi^K|\mathbf{\Psi}_\ell|). \end{aligned} \tag{4.3.10}$$

In this case, the values of these non-zero entries in the mean and covariance are not known but can be approximated. Suppose that there is a PMU installed at each source node in the grid (e.g., substation, generators,...). Let the vectors $I_{s,0}$ and $I_{s,F}$ denote the sum of the current injected by each source for each faulty phase into the grid during pre and post-fault condition, respectively. Fig. 4.3.2 shows one-line diagram of a sample test case after a fault, which is supplied through sources 1 and 2. There is a PMU installed next to each one measuring the three-phase line pre-fault

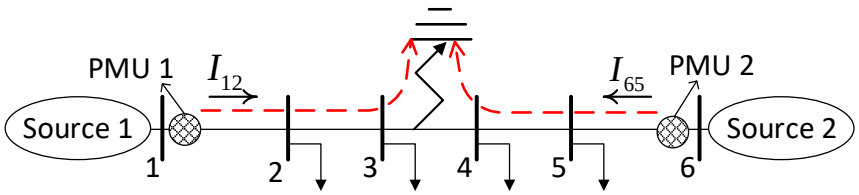



Figure 4.3.2: Composition of Vectors $I_{s,0}$ and $I_{s,F}$.

currents $I_{12,0} = (i^a_{12,0}, i^b_{12,0}, i^c_{12,0})^T$ and $I_{65,0} = (i^a_{65,0}, i^b_{65,0}, i^c_{65,0})^T$ and post-fault currents $I_{12,F}$ and $I_{65,F}$. Assume that the fault detectors indicate that a fault is on phase b and c. The vectors, $I_{s,0}$ and $I_{s,F}$ are formed as follows for this fault:

$$I_{s,0} = [i^b_{12,0}, i^c_{12,0}]^T + [i^b_{65,0}, i^c_{65,0}]^T$$

$$I_{s,F} = [i^b_{12,F}, i^c_{12,F}]^T + [i^b_{65,F}, i^c_{65,F}]^T$$

Forming the vectors $I_{s,0}$ and $I_{s,F}$ for other types of fault follows a similar approach. Note that if $P$ denotes the number of faulty phases, the vectors $I_{s,0}$ and $I_{s,F}$ are of

size $P \times 1$. We can use these vectors to estimate the values of the mean and the covariance of the non-zero entries of $\mathbf{I}_E$. We first define matrix $\mathbf{A}_l \in \{0,1\}^{M \times P}$ that is zero everywhere except for a block equal to the identity matrix $\mathcal{I}_P$ corresponding to the fault location indexed by $l$. For a given fault at node $\ell^*$, the value of the non-zero entries in $\boldsymbol{\mu}_\ell$ equal approximately to the sum of the currents flowing from the sources into the grid minus the current flowing into the grid before the fault occurred on the corresponding faulty phases. Denoting by $\boldsymbol{\delta} I_s = (I_{s,F} - I_{s,0})$, no matter where the fault is, we can use the following approximation:

$$\boldsymbol{\mu}_\ell = \mathbb{E}[\mathbf{x}|\mathcal{H}_l] = \mathbf{A}_l \boldsymbol{\delta} I_s \tag{4.3.11}$$

Also, the variance can be set to be a certain percentage of the value $|[\boldsymbol{\delta} I_s]_j|^2$, i.e. $[\boldsymbol{\Gamma}_s]_{jj} = \rho |[\boldsymbol{\delta} I_s]_j|^2$, where $\rho$ accounts for the percentage approximation error made in assuming that the fault current is completely accounted for by the source currents change. Hence:

$$\boldsymbol{\Phi}_\ell = \mathbf{A}_l \boldsymbol{\Gamma}_s \mathbf{A}_l^T \;\rightarrow\; \boldsymbol{\Psi}_\ell = \sigma_\epsilon^2 \mathcal{I} + \mathbf{F} \mathbf{A}_l \boldsymbol{\Gamma}_s \mathbf{A}_l^T \mathbf{F}^H \tag{4.3.12}$$

Before going through the discussion of the fault identifiability, it should be noted that we have introduced faults at nodes, to have finite and countable number of hypotheses. In reality, a fault occurs on a line. If the method was able to locate the fault at full resolution, then the two nodes with the highest $\lambda_l(\mathbf{y})$ would determine the faulty line. However, as we discuss in the next section, the fault localization method here can locate a fault with low-resolution up to a certain neighborhood of the actual location, and therefore putting a fault on a line or on its closest node would not change the end result.

If for a given fault at location $\ell^*$, the value of the metric $\lambda_\ell(\mathbf{y})$ in (4.3.10) is close to $\lambda_k(\mathbf{y})$ $(k \neq \ell^*, \ell = \ell^*)$, the location of the fault can be mis-identified due to different sources of error such as approximation errors, noisy measurements, etc.

**Definition.** *For each possible $\ell^*$ a cluster is a set of nodes for which $\lambda_\ell(\mathbf{y}) \approx \lambda_{\ell^*}(\mathbf{y})$ under $\mathbf{y}|\mathcal{H}_{\ell^*}$.*

We wish to investigate if the properties of the grid and the connection of the nodes affect the closeness of the log-likelihood value in (4.3.10), using the Kullback-Leibler (KL) divergence [Kullback and Leibler(1951)] of the two probability density functions $f(\mathbf{y}|\mathcal{H}_\ell)$ and $f(\mathbf{y}|\mathcal{H}_k)$. The KL divergence is:

$$\begin{aligned} D_{KL}(f(\mathbf{y}|\mathcal{H}_\ell)\|f(\mathbf{y}|\mathcal{H}_k)) &= \ln\frac{|\mathbf{\Psi}_k|}{|\mathbf{\Psi}_\ell|} + tr\big(\mathbf{\Psi}_k^{-1}\mathbf{\Psi}_\ell\big) - K \\ &+ \big(\mathbf{F}(\mathbf{A}_k - \mathbf{A}_\ell)\boldsymbol{\delta}\mathrm{I}_s\big)^H \mathbf{\Psi}_k^{-1}\big(\mathbf{F}(\mathbf{A}_k - \mathbf{A}_\ell)\boldsymbol{\delta}\mathrm{I}_s\big) \end{aligned} \tag{4.4.1}$$

In this application the weighted mean of the observations is very large compared to the noise $\sigma_\epsilon^2$. This leads to the following proposition, which is valid in our application:

**Proposition 2.** *Let $\tilde{\mathbf{F}}_k = \mathbf{F}\mathbf{A}_k\mathbf{\Gamma}_s^{1/2}$, $\boldsymbol{\xi} = \mathbf{\Gamma}_s^{-1/2}\boldsymbol{\delta}I_s$ and assume: $\forall k$ $\frac{\|\tilde{\mathbf{F}}_k\|_F^2}{\sigma_\epsilon^2} \gg 1$. Then, the following holds:*

$$D_{KL}(f(\mathbf{y}|\mathcal{H}_\ell)\|f(\mathbf{y}|\mathcal{H}_k)) \approx \frac{\boldsymbol{\xi}^H\tilde{\mathbf{F}}_\ell^H\tilde{\mathbf{\Pi}}_k^\perp\tilde{\mathbf{F}}_\ell\boldsymbol{\xi}}{\sigma_\epsilon^2} \tag{4.4.2}$$

*where $\tilde{\mathbf{\Pi}}_k^\perp = (\mathcal{I} - \tilde{\mathbf{F}}_k(\tilde{\mathbf{F}}_k^H\tilde{\mathbf{F}}_k)^{-1}\tilde{\mathbf{F}}_k^H)$ is the projector onto the subspace orthogonal to the columns of $\tilde{\mathbf{F}}_k$.*

*Proof.* First, the first and the second terms in the KL divergence expression are just a function of the covariance matrices. Therefore, the following approximation holds:

$$D_{KL}(f(\mathbf{y}|\mathcal{H}_\ell)\|f(\mathbf{y}|\mathcal{H}_k)) \approx \|\mathbf{F}(\mathbf{A}_k - \mathbf{A}_\ell)\boldsymbol{\delta}\mathrm{I}_s\|^2_{\mathbf{\Psi}_k^{-1}} \tag{4.4.3}$$

Second, we note that $\mathbf{F}(\mathbf{A}_k - \mathbf{A}_\ell)\boldsymbol{\delta}\mathrm{I}_s = (\tilde{\mathbf{F}}_k - \tilde{\mathbf{F}}_\ell)\boldsymbol{\xi}$ and applying the matrix inversion lemma and using the assumption $\frac{\|\tilde{\mathbf{F}}_k\|_F^2}{\sigma_\epsilon^2} \gg 1$:

$$\boldsymbol{\Psi}_k^{-1} = \frac{1}{\sigma_\epsilon^2}(\mathcal{I} - \tilde{\mathbf{F}}_k(\sigma_\epsilon^2\mathcal{I} + \tilde{\mathbf{F}}_k^H\tilde{\mathbf{F}}_k)^{-1}\tilde{\mathbf{F}}_k^H) \approx \frac{1}{\sigma_\epsilon^2}\tilde{\boldsymbol{\Pi}}_k^\perp \tag{4.4.4}$$

Noting that by definition $\tilde{\boldsymbol{\Pi}}_k^\perp\tilde{\mathbf{F}}_k = \mathbf{0}$, the statement follows. □

Note that $\tilde{\mathbf{F}}_k$ include a weighted subset of columns of the matrix $\mathbf{F} = \boldsymbol{\Sigma}^{-1}\mathbf{U}^H\mathbf{D}$ defined after (4.3.8). Let $\mathbf{f}_k$ denote a column of the matrix $\mathbf{F}$. Their correlation is the absolute value of the cosine of the angle between them:

$$r_{\ell k} = \frac{|\mathbf{f}_\ell^H\mathbf{f}_k|}{||\mathbf{f}_\ell||\ ||\mathbf{f}_k||} \tag{4.4.5}$$

From what has been shown, as $r_{\ell k}$ increases for two locations $\ell$ and $k$, the divergence of their corresponding hypotheses reduces and therefore it is more likely that they are mistaken with each other. Locations with high correlation coefficient belong to the same cluster.

What we wish for the design is to have the sizes of the classes of vectors with high mutual correlation to be as even as possible, so that we can have the finest localization resolution. The consequence that we draw from this observation is that, it is desirable to have the matrix $\mathbf{F}$ with the flattest possible spectrum. It is worth noting that the rows of $\mathbf{F}$ are orthogonal and its spectrum is easily relatable to the placement. In fact, $\mathbf{D}\mathbf{D}^H = (\mathcal{I} + \mathbf{C}\mathbf{C}^H)$ since $\mathbf{D} = (-\mathcal{I} \mid \mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}) = (-\mathcal{I} \mid \mathbf{C})$ and, because of (4.3.7):

$$\mathbf{F}\mathbf{F}^H = \boldsymbol{\Sigma}^{-1}\mathbf{U}^H(\mathcal{I} + \mathbf{C}\mathbf{C}^H)\mathbf{U}\boldsymbol{\Sigma}^{-1} = (\boldsymbol{\Sigma}^{-2} + \mathcal{I}) \tag{4.4.6}$$

which clearly indicates that in designing the placement we want as flat a spectrum as possible for $\mathbf{C}\mathbf{C}^H = \mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}\mathbf{Y}_{uu}^{-H}\mathbf{Y}_{au}^H$, since $\boldsymbol{\Sigma}^{-2}$ are the eigenvalues of its inverse. This suggests to minimize the so called *Shatten* infinity norm of the matrix $\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}$,

which is the infinity norm $\|\boldsymbol{\sigma}\|_\infty$ of the vector $\boldsymbol{\sigma}$ containing the non-zero singular values of $\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}$.

### *4.4.1 Optimal Phasor Measurement Unit Placement*

Let $\mathfrak{P}$ denote the set of all permutation matrices. We present the following placement strategy:

$$\begin{aligned} \boldsymbol{\Pi}^{\text{opt}} = \ & \underset{\boldsymbol{\Pi}\in\mathfrak{P}}{\text{argmin}} \quad \|\boldsymbol{\sigma}(\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1})\|_\infty \\ & \text{s.t.} \quad \boldsymbol{\Pi}\mathbf{Y}\boldsymbol{\Pi}^T = \begin{pmatrix} \mathbf{Y}_{aa} & \mathbf{Y}_{au} \\ \mathbf{Y}_{au}^T & \mathbf{Y}_{uu} \end{pmatrix} \end{aligned} \tag{4.4.7}$$

From (4.4.7), the placement problem is combinatorial and exponentially complex so it does not scale well. We use a *greedy search* as an alternative to reduce the time complexity to *polylog*. The pseudo-code of the used greedy search is shown in Algorithm. 4.2. In spite of the complexity of (4.4.7) it is possible to predict how to obtain good placements looking at the structure of the graph, and particularly of its natural clustering in sub-graphs with higher connectivity within themselves (in network science, these sub-graphs are often referred to as *communities*). This is because the spectral properties we are seeking can be tied to selecting nodes in such a way that the sparsity patterns of the rows of $\mathbf{Y}_{au}$ and of $\mathbf{Y}_{uu}$ separate in clusters. Since there are few observable nodes, the rank of the matrix $\mathbf{Y}_{au}$ is limited by its number of rows. The algorithm performs best if the rows of $\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}$ are as uncorrelated from each other as possible. Given the similarity of line parameter values, low correlation is mostly attained by having non overlapping support among of the rows in $\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1}$. As the next example illustrates, this can be attained by having $\mathbf{Y}_{uu}$ as close as possible to a block diagonal matrix, where the diagonal blocks as matched to the non zero portions of rows $\mathbf{Y}_{au}$. This is possible if the neighborhoods

**Algorithm 4.2:** Greedy Search Pseudo-Code for PMU Placement.

**Initialization**

$P :=$ Number of available PMUs;

$S :=$ Number of source nodes;

$\mathcal{P} :=$ {Source Nodes} //Nodes with PMU;

$\mathcal{M} :=$ Set of candidate placement buses;

**begin**

**for** *p=1..P-S* **do**

$Cost \leftarrow \infty$;

**for** *each* $m \in \mathcal{M}$ **do**

$\mathcal{P} := \mathcal{P} \cup \{m\}$;

given $\mathcal{P}$, calculate $\|\boldsymbol{\sigma}(\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1})\|_{\infty}$;

**if** $\|\boldsymbol{\sigma}(\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1})\|_{\infty} < Cost$ **then**

$m_{opt} \leftarrow m$;

$Cost \leftarrow \|\boldsymbol{\sigma}(\mathbf{Y}_{au}\mathbf{Y}_{uu}^{-1})\|_{\infty}$;

$\mathcal{P} := \mathcal{P} \setminus \{m\}$;

$\mathcal{P} := \mathcal{P} \cup \{m_{opt}\}$;

$\mathcal{M} := \mathcal{M} \setminus \{m_{opt}\}$;

of the buses where the sensors are located, have the smallest intersection possible. For example, for the system in Fig. 4.4.1, if the PMUs are placed at bus 2 and 5, the

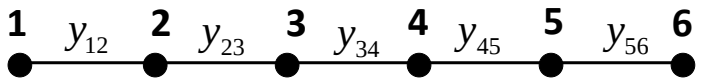



Figure 4.4.1: One-Line Diagram of a Sample Radial Network

admittance matrix is partitioned in the following desirable way:

$$\mathbf{Y} = \left(\begin{array}{cc:cc:cc} y_2 & 0 & -y_{12} & -y_{23} & 0 & 0 \\ 0 & y_5 & 0 & 0 & -y_{45} & -y_{56} \\ \hdashline -y_{12} & 0 & y_1 & 0 & 0 & 0 \\ -y_{23} & 0 & 0 & y_3 & -y_{34} & 0 \\ \hdashline 0 & -y_{45} & 0 & -y_{34} & y_4 & 0 \\ 0 & -y_{56} & 0 & 0 & 0 & y_6 \end{array}\right)$$

This design is clearly distancing the PMUs in the graph and dividing the network in neighborhoods, each associated with one of the sensors immediate and two hop neighbors where measurements are unavailable. The ambiguity that remains is confined to a connected set of buses that are topologically close to the PMU sensors. Hence, the optimum resolution for a certain grid is tied to the same intrinsic topological properties that are studied in graph clustering.

## 4.5 Numerical results and applications

In this section, we first corroborate our technical discussion and then showcase its application in the forensic analysis of cyber attacks to FLISR systems. In the numerical simulations we use the IEEE-34 bus and 123 test feeders (c.f. [Kersting(2001)]).

### 4.5.1 Sensor Placement and Identified Clusters

Before analyzing the results of the fault identification algorithm based on (4.3.10), we identify the *greedy-based* placement for the PMUs based on (4.4.7). We highlight the clustering by showing the values of $r_{\ell k}$ in (4.4.5) that exceed a certain threshold $\tau$ through images (all $r_{\ell k} < \tau$ are set to zero).

#### 4.5.1.1 IEEE-34 Bus Test Feeder

The one-line diagram of the test-case is shown in Fig. 4.5.1, where a 100 kW generator is added at bus 848. This grid is unbalanced and radial and has untransposed lines and one phase laterals. Also, without loss of generality, only the nodes at the same voltage level have been studied. As did the placement on the reduced network, where

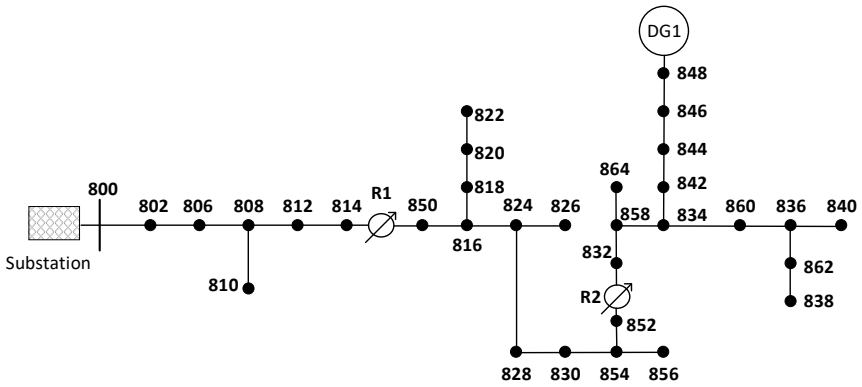



Figure 4.5.1: Reduced IEEE-34 Test Case with Added Generator.

the laterals are rolled up to avoid putting a PMU on a single phase node. Assuming that there are 4 PMUs available i.e., $P = 4$, Table. 4.2 shows the greedy-based location of the sensors and the random placement used as a comparison. Since this is a small case, we are able to do the exhaustive search to understand how well the greedy search has performed. For this purpose, we calculated the relative optimality gap that shows a value of 0.75% for the greedy-based search versus 7.61% for the

Table 4.2: PMU Locations for IEEE-34 Bus

| Test Case | #PMUs | Greedy-based sites | Random sites |
|---|---|---|---|
| IEEE-34 | 4 | 800-830-848-840 | 800-814-816-848 |

random case that puts the greedy-based to fall below 1% optimality gap.

It is clear from the locations in Fig. 4.5.1 according to Table. 4.2 that the sensors are placed to cover the grid. Fig. 4.5.2(a) shows the thresholded correlation (4.4.5) of the columns of **F** corresponding to phase-A for the greedy-based placement in Table 4.2 and Fig. 4.5.2(b) is the same quantity for the **F** that corresponds to the random choice. Since high correlation leads to low KL divergence, the clusters of light gray values effectively represent fault locations that are hard to discriminate. The larger the cluster, the lower the resolution. Hence, Fig. 4.5.2(b) clearly illustrates the impact on the resolution of a bad sensor siting, when compared to the clustering that emerges in Fig. 4.5.2(a). The correlation coefficients heat-map for the columns

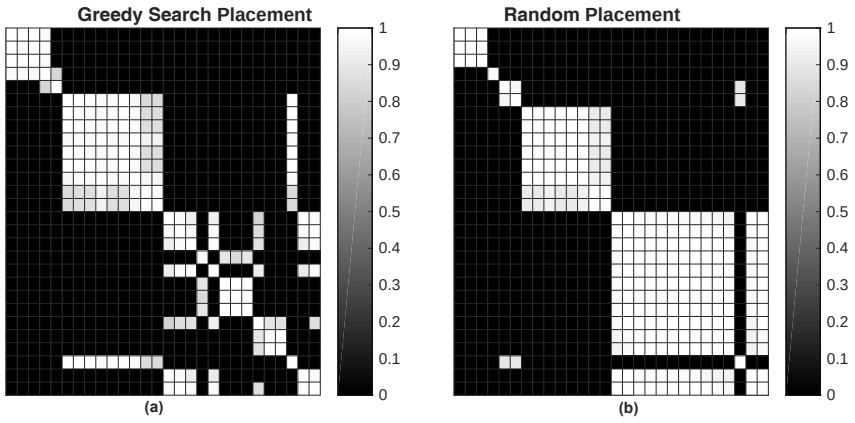



Figure 4.5.2: Correlation of Columns of **F** for IEEE-34 Bus Case Related to Phase-A for a) Greedy-based Location b) Random Location with $\tau = 0.814$

corresponding to phase-B and phase-C in the matrix **F** have a similar pattern as in Fig. 4.5.2. To corroborate our conjecture that the placement is affected by the

intrinsic topology of the graph and its clusters, in Fig. 4.5.3 we highlight the set of buses that exhibit high correlation based on the greedy-based placement, by building a graph adjacency matrix such that its $\ell k$ entry $[\mathbf{A}]_{\ell k} = u(r_{\ell k} - \tau)$[3] using the correlation coefficients of columns of $\mathbf{F}$ corresponding to phase-A and overlaying it on the IEEE-34 test case topology. As we predicted, the highly correlated nodes are those that are

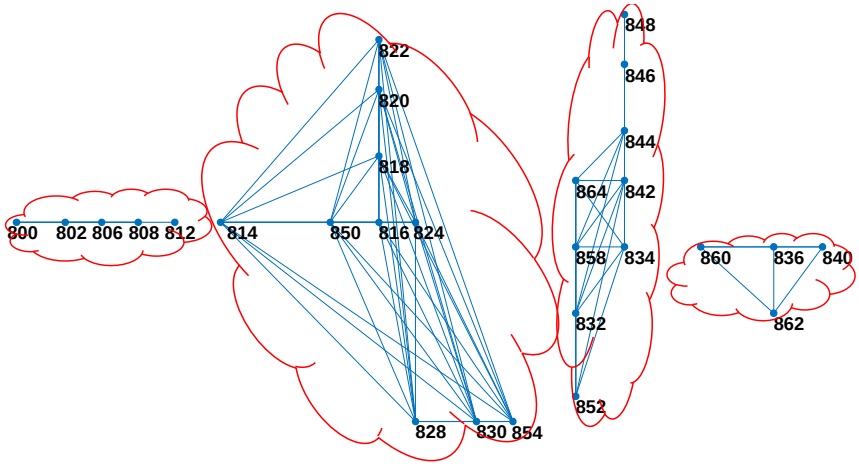



Figure 4.5.3: Adjacency Matrix Graph for Correlation Coefficients of Columns of $\mathbf{F}$ with $\tau = 0.814$ Corresponding to Phase-A.

located in a neighborhood of each other.

#### 4.5.1.2 IEEE-123 Bus Test Feeder

The analysis of the larger IEEE-123 test case, with a greedy-based assignment of 10 PMUs, returns the results in Table 4.3. Once again, we did the placement after rolling up the single and two-phase laterals and only nodes at the same voltage level are analyzed. The one-line diagram and node numbering of this case can be found in [Kersting(2001)]. The PMUs in this case also are spread over different areas of the grid topology to form communities of neighboring nodes of comparable sizes.

[3] $u(x)$ is the step function $u(x) = 1$ if $x > 0$ and $u(x) = 0$ else.

Table 4.3: Greedy-based Sensor Locations for IEEE-123 Bus

| Test Case | #PMUs | Greedy-based Location |
|---|---|---|
| IEEE-123 | 10 | 149-81-61-56-105-250-86-151-72-57 |

The thresholded heat-map given in Fig. 4.5.4 based on the placement of Table 4.3, highlights the balance across clusters.

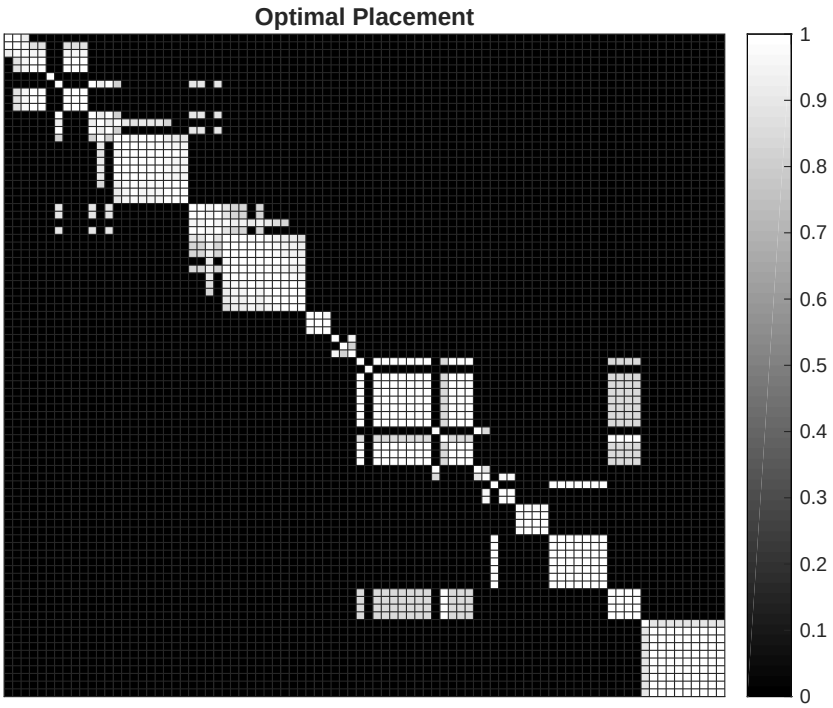



Figure 4.5.4: Correlation of Columns of $\mathbf{F}$ for IEEE-123 Bus Case Related to Phase-A for Greedy-based Location with $\tau = 0.814$

### *4.5.2 Fault Localization Algorithm*

The synthetic PMU data for this experiment were generated using the OpenDSS software [Dugan(2012)]. We first record the pre-fault data were using a power-flow snapshot and then we introduce a fault in the dynamic mode to represent the behavior

of the grid after a fault occurs. Note that the tap changers usually have a delay for 15-30 seconds in order to respond to a change so the voltage and current data should be recorded before the tap value changes, so that the admittance matrix stays the same before and after the fault.

#### 4.5.2.1 IEEE-34 Bus Test Feeder

Fig. 4.5.5 shows the value of the log-likelihood function $\lambda_\ell(\mathbf{y})$ for different hypotheses when a single-phase to ground fault occurs at phase-A of bus 820. As our analysis

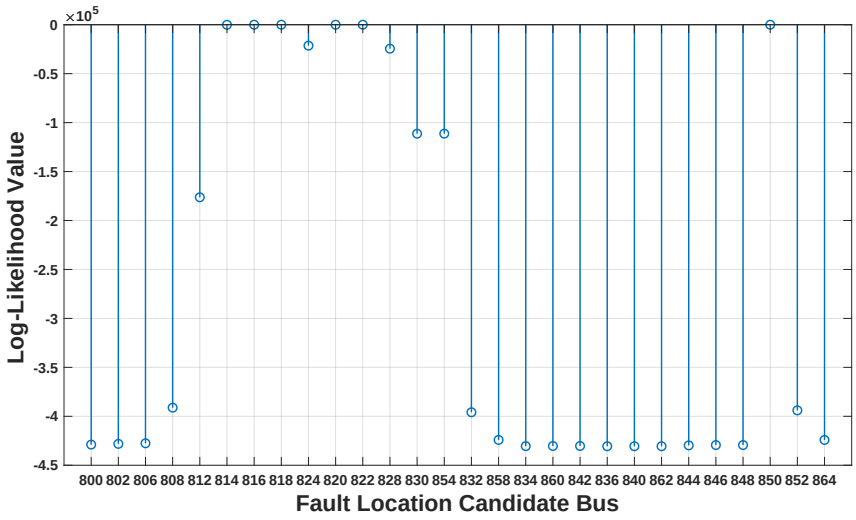



Figure 4.5.5: Log-Likelihood Value for Fault at Bus 820-Phase-A

indicated, using the metric in (4.3.10) shown in Fig. 4.5.3 it becomes readily apparent that one can narrow down the location of the fault up to a certain resolution, since a cluster of nodes return very close log-likelihood value. These are also the nodes that correspond to high values of $r_{\ell k}$ with respect to the actual location of the fault. To further test and verify the fault localization method, we introduce different types of fault and show the results of the fault localization in Table 4.4 under the greedy-based placement.

Table 4.4: Identified Fault Locations for IEEE-34 Bus Case

| Fault Type | Exact Fault Location | Locations with Highest $\lambda_\ell(\mathbf{y})$ |
|---|---|---|
| LLL | 816 | 814,816,850 |
| A-G | 822 | 814,816,818,820,822,850 |
| BC-G | 852 | 832,852,858 |
| AC-G | 836 | 836,840,862 |
| AB-G | 808 | 800,802,806,808 |

#### 4.5.2.2 IEEE-123 Bus Test Feeder

We extend the analysis of the fault localization algorithm to IEEE-123 test case for different types of fault after the placement of the sensors on the locations in Table 4.3. Fig. 4.5.6(a) shows the log-likelihood value of different hypotheses for a three-phase fault at bus 160. Part of the correlation coefficients heat-map that illustrates the correlation of the columns that have high correlation with bus 160 is snipped out of Fig. 4.5.4 and is illustrated in Fig. 4.5.6(b). As expected, the neighboring nodes that also have high correlation with bus 160 are those which have the closest log-likelihood value in our metric and can be mis-identified as the fault location.

Other types of faults and the potential locations with highest $\lambda_\ell(\mathbf{y})$ are summarized in Table 4.5.

We introduce faults at different locations of the grid and the results show that the fault is identifiable up to the resolution of the clusters of neighboring nodes identified with the correlation coefficients. An interesting case is the two-phase fault at node 57 (last row in Table 4.5), where the fault is identifiable with a high resolution. This is consistent with our claims, since the coefficients $r_{\ell k}$ reveal that there is no other

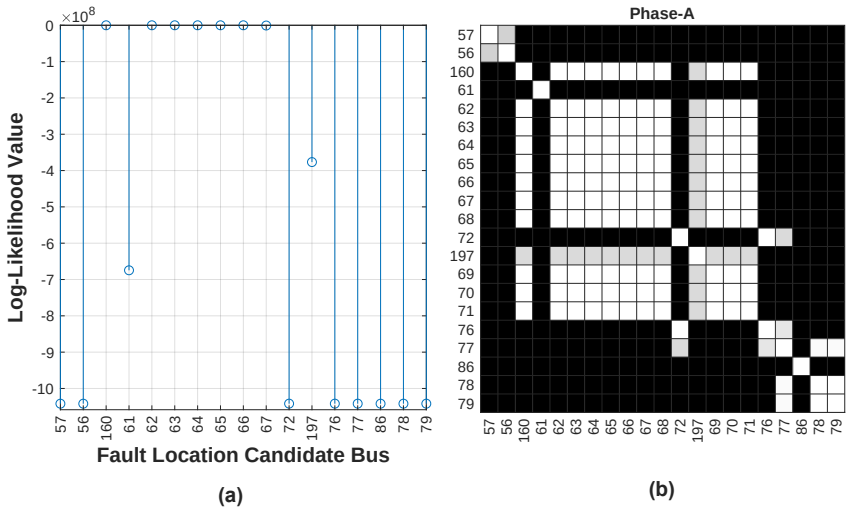



Figure 4.5.6: a) Log-Likelihood Value for Three-Phase Fault at Bus 160 and b) Correlation of Corresponding Columns of **F**-phase A.

node with very high correlation with this specific node.

#### 4.5.2.3 Comparison with the State of the Art

Our approach in using the pre and post-fault samples and treating the fault as a current injection improves upon the work in [Brahma(2011)] through its statistical underpinning. To show this, we introduce a two-phase fault at bus 836-A-C and the results are shown in Fig. 4.5.7. The original method in [Brahma(2011)] is designed as a minimization problem to find the fault location so we changed it to a maximization by adding a negative sign to make a better visual comparison with our method. The results clearly show that our described method can locate the fault more accurately than the algorithm in [Brahma(2011)].

Besides this improvement, investigating the reason behind this ambiguity is missing in [Brahma(2011)], whereas we have investigates this matter theoretically and analytically in our work. In addition, we have presented a sensor placement strategy here to place the sensors at nodes that would return the highest fault localization

Table 4.5: Identified Fault Locations for IEEE-123 Bus Case

| Fault Type | Exact Fault Location | Locations with Highest $\lambda_\ell(\mathbf{y})$ |
|---|---|---|
| LLL | 42 | 40,42,44,47,48 |
| A-G | 108 | 105,108,109,300,110, 111,112,113,114 |
| BC-G | 89 | 86,87,89,91,93,95 |
| AC-G | 50 | 47,48,49,50,51,151 |
| AB-G | 57 | 57 |

resolution.

### *4.5.3 Application to Cyber-Physical Intrusion Detection*

We leveraged the insights from this work to enhance our network intrusion detection system testbed hosted at Lawrence Berkeley National Lab (LBNL) [Gentz(2017)] and incorporate additional rules to monitor for cyber attacks that interfere with the normal operations of FLISR systems [Uluski(2012)]. Our testbed (see Fig. 4.5.8) is defined as a hierarchical architecture to fast-detect the presence of attacks against networked devices controlling physical systems by correlating analytic results from SCADA traffic (which includes traces from fault detector communications) with PMUs deployed sparsely over a distribution system [Gentz(2017)].

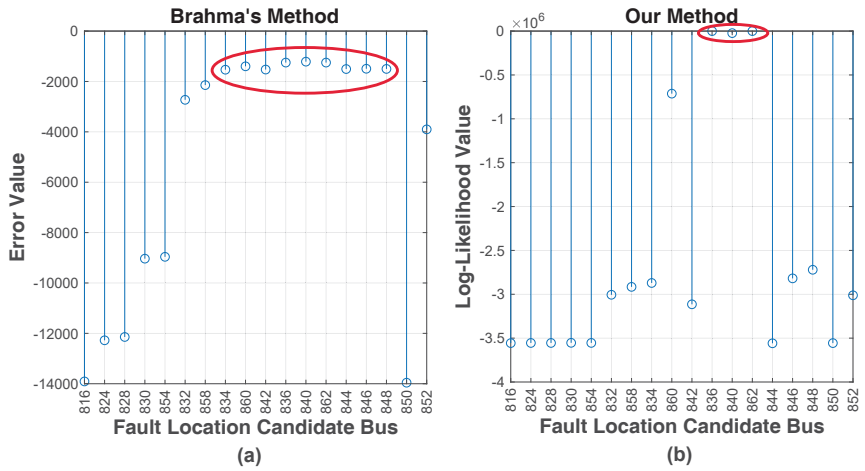



Figure 4.5.7: Fault Localization for Fault at Bus 836-A-C Using a) Method in [Brahma(2011)] b) Described Method.

#### 4.5.3.1 The FLISR System Operations and Vulnerabilities

FLISR systems detect the location of a permanent fault [4] in a feeder and automatically restore service to customers in the healthy section of the feeder. To do so they employ directional fault detectors, installed at every line in a distribution feeder, that communicate the occurrence of a fault and its direction to a distribution management system (DMS). The DMS analyzes these data and, once the faulted section is identified, issues commands to a predetermined set of switches to first isolate the faulted section and then restore service to the non-faulted areas [Uluski(2012)], so as to minimize the service disruption. Consider the one-line diagram of a sample radial network shown in Fig. 4.5.9 to describe how FLISR works. The network is connected to a substation at bus 1 and a distributed generation is connected to bus 6. Fault detectors and normally-closed switches are connected at the end of every line. When a fault happens at line $3-4$, all the fault detectors detect the presence of this fault

[4]Temporary faults that clear themselves below a certain time threshold do not trigger the FLISR operation.

such that fault detectors at buses $1, 2$ and 3 indicate that the fault is in their right side, whereas fault detectors at buses $4, 5$ and 6 indicate the fault is to their left. Fault localization analysis at the DMS using this information identifies the faulted section to be line $3 - 4$. The DMS, therefore, sends control signals to automatically open switches to the right of bus 3 and to the left of bus 4 splitting the network into two parts and restoring service to customers on both sections.

Fault detector data is transmitted to the DMS through a wide area network using industrial protocols such as DNP3 [IEEE(2012)] or Modbus [Modbus(2004)] which are not designed with security concerns in mind. Therefore, a dedicated attacker can tamper with fault detector data either remotely, through a compromised network device (e.g., a router), or by physically connecting herself to an exposed section of the communication network. For instance, the tampering can lead the DMS analysis to pick the wrong location for an actual fault or hide the presence of a fault.

Although such attacks can be partially prevented by enabling photographically authenticated communication on top of the stated communication protocols, such a measure does not prevent the whole spectrum of possible cyber attacks. Therefore, a more robust approach, that does not merely rely on cryptographic solutions, is desired that is capable of detecting the presence of an attacker. The analysis we carried out in this part of the chapter shows that, even with a very limited number of sensors scattered in the systems, the PMU localization performed using (4.3.10) enhances an operator's confidence about the fault localization information (or the lack of thereof) extracted from fault detectors data, by giving an additional means to verify the trustworthiness of the SCADA messaging, albeit at a lower spatial resolution. In the next section we illustrate the effectiveness of the forensic analysis on FLISR attacks carried out using a experiment, with simulated streaming SCADA and PMU data, to our LBNL testbed in Fig. 4.5.8.

#### 4.5.3.2 Intrusion Detection on Fault Detectors

In our simulation to demonstrate intrusion detection on fault detectors, we use the modified IEEE 34-bus system shown in Fig. 4.5.10. We introduce a two-phase AC-G fault at line $860 - 836$. Our PMU data analytic results correctly indicate the fault to be in Zone D. We generate simulated SCADA data for each fault detector in the network using the openDNP3 library [Open DNP3 Group and others(2012)] that implements the DNP3 communication protocol. In our simulation, we modify the packets from some of the fault detectors (those at buses 834 and 860) so that the SCADA data analytics indicate the fault is at line $834 - 842$, which is in zone C. The inconsistent results from the PMU data analytics and the SCADA data analytics raise an alarm about a possible cyber-attack on the fault detectors.

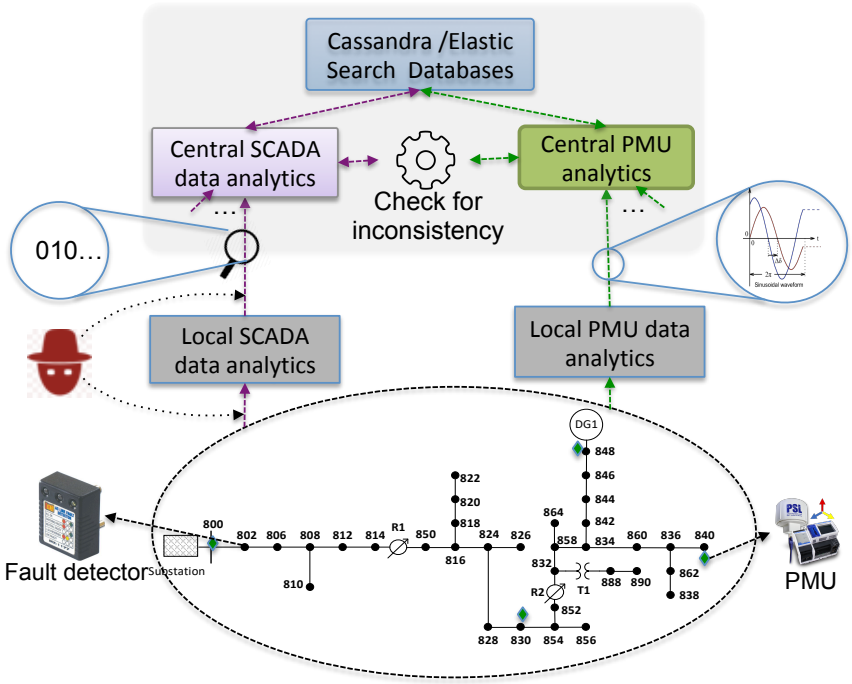



Figure 4.5.8: LBNL Stream-Processing Architecture for Real-time Cyber-physical Security (SPARCS) used in our experiments [Gentz(2017)], [Peisert *et al.*(2017)], [Peisert *et al.*(2018)]. Its components are: 1) The Bro IDS, the industry-standard, network monitoring framework; 2) a publish-subscribe messaging system called RabbitMQ used to transfer PMU and SCADA data and analysis results; 3) the Cassandra database, used for permanent archiving of all data; 4) The Elasticsearch database, which collects analysis results and receives events notifications in real time; 5) BeagleBone Black (BBB) devices that receive synchrophasor data at a rate of 120 samples/sec, and analyze 1 sec. worth of data at a time to rapidly detect a cyber-physical event [Jamei *et al.*(2016)].

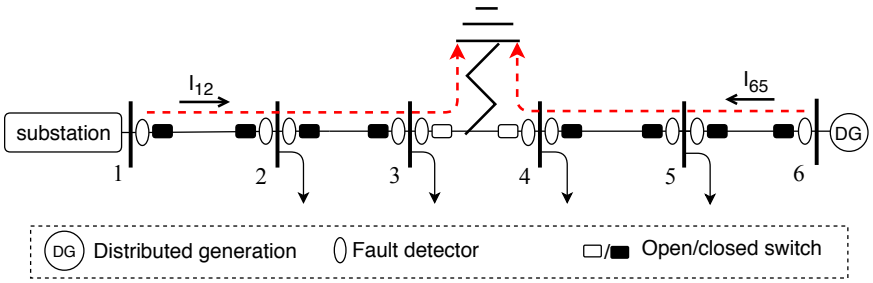



Figure 4.5.9: Diagram of a Sample Radial Network with a Fault in Line $3-4$.

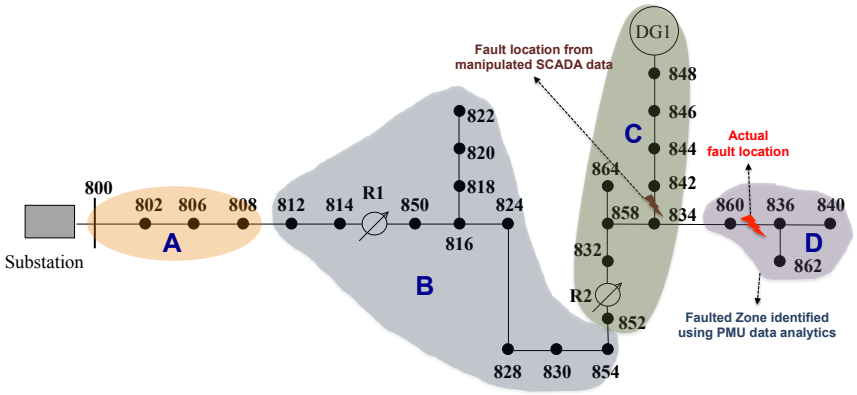



Figure 4.5.10: Output Mismatch as a Result of an Attack on SCADA data

# Part II

# Coordinated Security-Constrained Scheduling of Power and Gas Grid

# CHAPTER 5

# GAS AND ELECTRIC GRID UNIT COMMITMENT WITH COORDINATED N-1 GENERATOR CONTINGENCY ANALYSIS

The inter-dependency between the gas and electric grid is growing as the reliance on the gas-fired generators increases. Strong coupling between two large infrastructures requires a coordinated planning and operational framework. In fact, the constraints of both grids must be respected, while their objective functions are co-optimized. In this chapter, we present a new formulation for Gas and Electric Grid Coordinated Unit Commitment (GECUC). We apply the standard DC power flow approximation and propose a new formulation to relax the non-linear gas grid constraints so that the GECUC can be mapped into a Mixed-Integer Linear Program (MILP). In addition, the N-1 generator contingency security constraints are extended to include the gas grid constraints and find the required reserve from gas wells and generators after the loss of a generating unit.

Throughout the chapter, for simplicity of notation, the following convention is made unless explicitly otherwise mentioned: $\forall t \equiv \forall t \in \mathcal{T}, \forall w \equiv \forall w \in \mathcal{W}, \forall g \equiv \forall g \in \mathcal{G}, \forall i \equiv \forall i \in \mathcal{V}_g, \forall k \equiv \forall k \in \mathcal{E}_g, \forall b \equiv \forall b \in \mathcal{E}_e, \forall q \equiv \forall q \in \mathcal{Q}$.
$\mathbf{A}^T$ and $[\mathbf{A}]_m$ denotes the transpose and m-th element of $\mathbf{A}$, respectively, and finally $T = |\mathcal{T}|$.

## 5.1 Gas Network

The following components comprise the natural gas grid.

***Suppliers and Loads:*** Gas is injected or withdrawn at different points in the gas grid. Most of the gas is supplied by the wells. The gas loads could be residential, commercial, or industrial (gas-fired generators are in the third category). Unlike for

electric power systems, gas can be stored in some locations during the off-peak hours and withdrawn during the peak hours. These locations therefore play the role of supplier during some hours and of load during the other hours.

***Pipelines:*** Similar to power grid lines, the pipelines in a natural gas grid are designed to deliver gas from suppliers to loads. Generally speaking, gas grids are radial or very weakly-meshed.

***Compressors:*** Since the gas pressure drops over the gas pipeline, compressors are placed at different points of the grid (usually every 50-100 miles) [Misra *et al.*(2015)] to increase the gas pressure. Some compressors use part of the natural gas to run their engines, and others use electricity. Gas pipelines with compressor are referred to as active pipelines, which also form the set $\mathcal{E}_g^a \subseteq \mathcal{E}_g$, while the pipelines without a compressor are referred to as passive pipelines.

### *5.1.1 Steady-State Gas Flow Equations*

For a passive pipeline $k$ that is defined "from" bus $i$ "to" bus $j$, the relationship between the gas flow over the line and the nodal pressure at two ends is as follows [Misra *et al.*(2015)]:

$$\pi_i^2 - \pi_j^2 = h_k f_k |f_k|, \quad \forall k \in \mathcal{E}_g \backslash \mathcal{E}_g^a. \tag{5.1.1}$$

where the constant $h_k$ depends on the pipeline friction factor, diameter, length, gas compressibility factor, gas constant, and gas temperature. The relationship in (5.1.1) can be extended to include active pipelines, i.e.,

$$\rho_k^2 \pi_i^2 - \pi_j^2 = h_k f_k |f_k|, \quad \forall k \in \mathcal{E}_g. \tag{5.1.2}$$

where $\rho_k$ is the compression ratio of line $k$, that is the ratio of the compressor outlet and inlet square pressure. Clearly, for passive pipelines, $\rho_k = 1$, and therefore (5.1.2)

reduces to (5.1.1). In eq. (5.1.2), it is assumed that the compressors are placed at the "from" node of the pipelines for simplicity of notation. However, the model is simply extendable to more general cases.

### *5.1.2 Optimal Gas Flow (OGF) Problem*

Similar to the Optimal Power Flow (OPF) in the electric grid, the objective of the OGF is to find the economic hourly schedule of the gas wells/storage, while respecting the constraints of the gas grid. Throughout the chapter, it is assumed that the gas wells are always committed. For simplicity, the gas storage is modeled similar to a gas well, which allows for zero production unlike the gas well , which the minimum production is constrained based on a contract, but the inter-temporal constraints of the storage have not been considered. The problem is then formulated as:

$$\text{Minimize: } \sum_{t\in\mathcal{T}}\sum_{w\in\mathcal{W}} d_w G_w(t) + d_w^{res} rg_w(t) \tag{5.1.3}$$

Subject to:

$$G_w(t) \geq G_w^{min}, \qquad \forall t, w \tag{5.1.4a}$$

$$G_w(t) + rg_w(t) \leq G_w^{max}, \qquad \forall t, w \tag{5.1.4b}$$

$$rg_w(t) \geq 0, \qquad \forall t, w \tag{5.1.4c}$$

$$\sum_{w\in\mathcal{W}} rg_w(t) \geq \beta_{gas} \sum_{i\in\mathcal{V}_g} L_i(t), \qquad \forall t \tag{5.1.4d}$$

$$\sum_{w\in\mathcal{W}(i)} G_w(t) + \sum_{k\in\mu_i^+} f_k(t) - \sum_{k\in\mu_i^-} f_k(t) = L_i(t) \qquad \forall t, i \tag{5.1.4e}$$

$$\pi_i^{min} \leq \pi_i(t) \leq \pi_i^{max}, \qquad \forall t, i \tag{5.1.4f}$$

$$\rho_k(t)\pi_i(t) \leq \pi_{\text{pipe}}^{max} \qquad \forall t, k \in \mathcal{E}_g^a \tag{5.1.4g}$$

$$\pi_i^2(t) = \pi_s^2, \qquad \forall t, i \in \mathcal{W}^s \tag{5.1.4h}$$

$$\rho_k^2(t)\pi_i^2(t) - \pi_j^2(t) = h_k f_k(t)|f_k(t)|, \qquad \forall t, k \tag{5.1.4i}$$

The gas production at each well/storage is limited by the contracts and the capacity of the production, which is imposed by (5.1.4a) and (5.1.4b). For a storage, the minimum limit would be zero. The gas nodal balance constraint is imposed by (5.1.4e). Similar to the electric grid, we define a variable $rg_w(t)$ to account for an increase in the gas demand, when a contingency of a base-load generator in the electric grid happens, and as a result some of the gas generators are re-dispatched/committed to make up for the lost generator, especially in the form of non-spinning reserve. It is assumed that the summation of the potential reserve over all the wells/storage at each hour should be greater than $\beta_{gas} \in (0, 1)$ times of the net gas base-case load at that hour. The pressure limit at each node and pressure limit of the pipelines are imposed by Eq. (5.1.4f),(5.1.4g). Finally, the pressure at the slack node is fixed.

The only non-linearity in the OGF appears in Eq. (5.1.4i) due to the quadratic terms and the absolute value. In order to deal with the absolute value term, we first define the following auxiliary variables:

$$f_k(t) = f_k^+(t) - f_k^-(t), \quad f_k^+(t) \geq 0, \quad f_k^-(t) \geq 0 \ \ \forall t, k \tag{5.1.5}$$

where $f_k^+(t)$ should be non-zero only when the flow is in the positive direction, otherwise $f_k^-(t)$ should be non-zero. To enforce this, we define:

$$z_k(t) = \begin{cases} 1 & \text{if } f_k(t) > 0 \\ 0 & \text{else} \end{cases} \tag{5.1.6}$$

Therefore, after adding the following set of constraints to the problem:

$$0 \leq f_k^+(t) \leq M_1 Z_k(t), \qquad \forall t, k \tag{5.1.7a}$$

$$0 \leq f_k^-(t) \leq M_1(1 - Z_k(t)), \qquad \forall t, k \notin \mathcal{E}_g^a \qquad (5.1.7b)$$

$$f_k^-(t) = 0, \qquad \forall t, k \in \mathcal{E}_g^a \qquad (5.1.7c)$$

we can rewrite Eq. (5.1.4i), (5.1.4e) as:

$$\sum_{w \in \mathcal{W}(i)} G_w(t) + \sum_{k \in \mu_i^+} \big(f_k^+(t) - f_k^-(t)\big) - \sum_{k \in \mu_i^-} \big(f_k^+(t) - f_k^-(t)\big) = L_i(t) \qquad \forall t, i \qquad (5.1.8a)$$

$$\rho_k^2(t)\pi_i^2(t) - \pi_j^2(t) = h_k\big(f_k^{+^2}(t) - f_k^{-^2}(t)\big), \qquad \forall t, k \qquad (5.1.8b)$$

In the added constraints above, $M_1$ is a parameter and is chosen to be a big enough value for the flow. Also, constraint Eq. (5.1.7c) is added to account for the fact that the active pipelines only pass the flow in one direction. The advantage of this reformulation is that the flow direction has been implicitly considered using linear constraints. However, new variables including new binary variables have been introduced that can add to the complexity of the problem. Also, the number of constraints of the problem has increased compared to the non-linear formulation.

The only remaining non-linearity is due to the quadratic terms. For the pressure, it is easy to define the auxiliary variables as follows:

$$\pi_i^2(t) = \tilde{\pi}_i(t) \Rightarrow \tilde{\pi}_i^{min} \leq \tilde{\pi}_i(t) \leq \tilde{\pi}_i^{max}, \qquad \forall t, i \qquad (5.1.9)$$

$$\rho_k^2(t)\tilde{\pi}_i(t) \leq \tilde{\pi}_{\text{pipe}}^{max} \qquad \forall t, k \in \mathcal{E}_g^a \qquad (5.1.10)$$

$$\tilde{\pi}_i(t) = \pi_s^2, \qquad \forall t, i \in \mathcal{W}^s \qquad (5.1.11)$$

However, the same approach cannot be used for the flow terms as the flow itself (and not the squared flow) appears in the gas nodal balance constraint. We use a piecewise linearization formulation of the flows $f_k^{+^2}$ and $f_k^{-^2}$ over a range $[0, M_2]$ in this case. The following constraints are therefore added to the problem:

$$\gamma_k^+(t) \geq \alpha[n] f_k^+(t) + \beta[n], \qquad \forall t, k, n \qquad (5.1.12)$$

$$\gamma_k^-(t) \geq \alpha[n] f_k^-(t) + \beta[n], \qquad \forall t, k, n \tag{5.1.13}$$

where $\gamma_k^+(t)$ and $\gamma_k^-(t)$ are approximations of $f_k^{+^2}(t)$ and $f_k^{-^2}(t)$ over a range $M_2$, respectively. $\alpha[n]$ and $\beta[n]$ are the coefficients of the piece-wise linearized quadratic curve. The constraint Eq. (5.1.8b) is then rewritten as:

$$\rho_k^2(t)\pi_i^2(t) - \pi_j^2(t) = h_k\big(\gamma_k^+(t) - \gamma_k^-(t)\big), \qquad \forall t, k \tag{5.1.14}$$

The objective function in Eq. (5.1.3) should also change accordingly in order to make sure $\gamma_k^+(t)$ and $\gamma_k^-(t)$ does not sit above the second order curve as follows:

$$\begin{aligned} \text{Min} \sum_{t\in\mathcal{T}} \big( & \sum_{w\in\mathcal{W}} d_w G_w(t) + d_w^{res} rg_w(t) \\ & + \sum_{k\in\mathcal{E}_g} a_1\gamma_k^+(t) + a_2\gamma_k^-(t)\big) \end{aligned} \tag{5.1.15}$$

where $a_1$ and $a_2$ are the weights of the added terms in the objective function.

## 5.2 GECUC Formulation

The GECUC problem formulation brings the OGF and UC problem under a unified framework. In the numerical results, we show how solving a UC irrespective to the gas grid constraints can result in infeasible scheduled grid operation. The optimization problem can then be formulated as follows[1]:

$$\begin{aligned} \text{Min} \sum_{t\in\mathcal{T}} \Bigg( & \sum_{g\in\mathcal{G}} \big(\mathbf{C}_g^{x^T}\mathbf{X}_g(t) + \mathbf{C}_g^{r^T}\mathbf{R}_g(t) + c_g^{\text{nl}} u_g(t) \\ & + c_g^{\text{st}} v_g(t)\big) + \sum_{w\in\mathcal{W}} \big(d_w G_w(t) + d_w^{res} rg_w(t)\big) \\ & + \big(\sum_{k\in\mathcal{E}_g} a_1\gamma_k^+(t) + a_2\gamma_k^-(t)\big)\Bigg) \end{aligned} \tag{5.2.1}$$

where the q-th element of vector $\mathbf{X}_g(t)$ with size $|\mathcal{Q}|$, corresponds to the power production at time $t$ of generator $g$ for segment $q$, $P_{g,q}(t)$ and the corresponding element

[1] $\mathbf{A}^T$ is the transpose of vector $\mathbf{A}$.

in vector $\mathbf{C}_g^x$ is the cost related to it. For generator $g$ at time $t$, the generation reserve vector is defined as:

$$\mathbf{R}_g(t) = [r_g^{ru}(t), r_g^{rd}(t), r_g^s(t), r_g^{ns}(t)]^T$$

where the elements correspond to the regulation-up, regulation-down, spinning, and non-spinning reserve, respectively. Accordingly, the cost related to each reserve is defined in the vector $\mathbf{C}_g^r$.

The minimization is subject to the power grid and gas grid constraints that are presented in the following sections.

### *5.2.1 Power Grid Constraints*

The power grid constraints are split into three parts. The first part goes over the production and reserve level requirements. The second part describes the ramp-rate limits, and finally the third part explains the constraints on commitment variables.

*1) Production and Reserve Limits*

$$P_g(t) = \mathbf{1}^T.\mathbf{X}_g(t), \qquad \forall t, g \tag{5.2.2a}$$

$$[\underline{\mathbf{X}}_g]_q.u_g(t) \leq [\mathbf{X}_g(t)]_q \leq [\overline{\mathbf{X}}_g]_q.u_g(t), \quad \forall t, g, q \tag{5.2.2b}$$

$$P_g(t) + [1, 0, 1, 0].\mathbf{R}_g(t) \leq P_g^{max}.u_g(t), \ \forall t, g \tag{5.2.2c}$$

$$P_g^{min}.u_g(t) \leq P_g(t) - [\mathbf{R}_g(t)]_2, \qquad \forall t, g \tag{5.2.2d}$$

$$\sum_{g\in\mathcal{G}}[\mathbf{R}_g(t)]_m \geq \beta_{el} \sum_{n\in\mathcal{V}_e} d_n(t), \qquad \forall t, m = 1, 2 \tag{5.2.2e}$$

$$\sum_{g\in\mathcal{G}}[\mathbf{R}_g(t)]_3 \geq \frac{1}{2}\max_g(P_g^{max}), \qquad \forall t \tag{5.2.2f}$$

$$[0, 0, 1, 1].\sum_{g\in\mathcal{G}}\mathbf{R}_g(t) \geq \max_g(P_g^{max}), \qquad \forall t \tag{5.2.2g}$$

$$\left|\sum_{n\in\mathcal{V}_e} PF_{b,n}^R \left(\sum_{g\in\mathcal{G}(n)} P_g(t) - d_n(t)\right)\right| \leq Pl_b^{max,a}, \forall t, b \tag{5.2.2h}$$

$$\sum_{g\in\mathcal{G}} P_g(t) = \sum_{n\in\mathcal{V}_e} d_n(t). \qquad \forall t \tag{5.2.2i}$$

The vectors $\underline{\mathbf{X}}_g$ and $\overline{\mathbf{X}}_g$ are defined as follows:

$$\underline{\mathbf{X}}_g = [P_g^{min}, 0, 0, ..., 0]^T,$$

$$\overline{\mathbf{X}}_g = [P_g^{min}, P_{g,2}^{max}, P_{g,3}^{max}, ..., P_{g,q}^{max}, ..., P_{g,|\mathcal{Q}|}^{max}]^T$$

where $P_{g,q}^{max}$ is generator $g$ production upper-limit corresponding to segment $q$ of the piece-wise constant cost-function. Note that in this constraint, the generation for segment 1 is forced to be equal to $P_g^{min}$.

Constraints (5.2.2b)-(5.2.2d) ensures that generator $g$ does not violate its generation limits. Constraints (5.2.2e)-(5.2.2g) define the requirements on the amount of spinning and non-spinning reserves, and constraints (5.2.2h)-(5.2.2i) are the power-flow model.

*2) Ramp-Rate Limits*

Provided generation reserve is limited by the ramp-rate limits of the generators, and are enforced through constraints (5.2.3a)-(5.2.3f). Constraints (5.2.3g) and (5.2.3h) on the other hand impose the hourly inter-temporal constraint on the generators.

$$0 \leq [\mathbf{R}_g(t)]_m \leq [\mathbf{S}_g]_1.u_g(t), \quad \forall t, g \in \mathcal{G}^a, m = 1, 2 \tag{5.2.3a}$$

$$[\mathbf{R}_g(t)]_m = 0 \qquad \forall t, g \notin \mathcal{G}^a, m = 1, 2 \tag{5.2.3b}$$

$$[1, 0, 1, 0].\mathbf{R}_g(t) \leq [\mathbf{S}_g]_2.u_g(t), \qquad \forall t, g \tag{5.2.3c}$$

$$0 \leq [\mathbf{R}_g(t)]_4 \leq [\mathbf{S}_g]_2.(1 - u_g(t)), \qquad \forall t, g \in \mathcal{G}^f \tag{5.2.3d}$$

$$[\mathbf{R}_g(t)]_4 = 0, \qquad \forall t, g \notin \mathcal{G}^f \tag{5.2.3e}$$

$$0 \leq [\mathbf{R}_g(t)]_3, \qquad \forall t, g \tag{5.2.3f}$$

$$P_g(t) - P_g(t-1) \leq [\mathbf{S}_g]_3.u_g(t-1) + [\mathbf{S}_g]_5.v_g(t), \tag{5.2.3g}$$

$$P_g(t-1) - P_g(t) \leq [\mathbf{S}_g]_3.u_g(t) + [\mathbf{S}_g]_4.(v_g(t) - u_g(t) + u_g(t-1)). \qquad \forall t, g \tag{5.2.3h}$$

For generator $g$, the ramp-rate limit vector is defined as:

$$\mathbf{S}_g = [s_g^5, s_g^{10}, s_g^{60}, s_g^{SD}, s_g^{SU}]^T$$

where the elements correspond to the generator 5-minute, 10-minute, 60-minute, shut-down, and start-up ramp-rate limit, respectively.

In these constraints, the time horizon over which the commitment problem is solved is treated cyclically, i.e.,

$$t - s \equiv T - (t - s) \;\; \forall \; s \geq t. \tag{5.2.4}$$

*3) Commitment Variable Constraints*

The explicit constraints on the unit commitment variables are presented in eqs. (5.2.5a)-(5.2.5d). Similar to the previous part, the time horizon is cyclical in these equations.

$$u_g(t) \in \{0, 1\}, \; 0 \leq v_g(t) \leq 1, \qquad \forall t, g \tag{5.2.5a}$$

$$v_g(t) \geq u_g(t) - u_g(t-1), \qquad \forall t, g \tag{5.2.5b}$$

$$\sum_{s=t-UT_g+1}^{t} v_g(s) \leq u_g(t), \qquad \forall t, g \tag{5.2.5c}$$

$$\sum_{s=t-DT_g+1}^{t} v_g(s) \leq 1 - u_g(t - DT_g), \quad \forall t, g \tag{5.2.5d}$$

### *5.2.2 Gas Grid Constraints*

Constraints (5.1.4a)-(5.1.4d),

$$\sum_{w \in \mathcal{W}(i)} G_w(t) + \sum_{k \in \mu_i^+} \left(f_k^+(t) - f_k^-(t)\right) - \sum_{k \in \mu_i^-} \left(f_k^+(t) - f_k^-(t)\right) - \sum_{g \in \mathcal{S}(i)} P_g(t)\eta_g = L_i(t), \qquad \forall t, i \tag{5.2.6}$$

Constraints (5.1.7a)-(5.1.7b), and (5.1.9)-(5.1.14).

Compared to the nodal balance constraint in (5.1.4e), the new nodal balance constraint of the gas grid in (5.2.6) includes the demand of the gas-fired generators supplied from the gas grid, where $P_g(t)$ is also a variable for the electric grid. This equation is, in fact, what couples the two infrastructures.

Based on the current common practice, it is assumed that the compressor ratio is set by the operator heuristically rather than being optimally chosen, and is imported in the formulation as a parameter over the normal operation.

## 5.3 Coordinated N-1 Generator Contingency Analysis

Coordinated scheduling of the two infrastructures over the normal operation is necessary but not sufficient. This becomes more evident when we look at the problem in the context of N-1 generator contingency analysis, where losing a large generator usually means a sudden increase in natural-gas demand. The increase is due to the re-dispatch/commitment of the gas-fired generators to support the grid with spinning and, more critically, non-spinning reserves. Therefore, it is important to make sure that the planned contingency dispatch respects the constraints of the gas and electric grid using a coordinated N-1 generator contingency analysis.

The following constraints are then added to the GECUC formulation to account for N-1 generator contingency analysis.

### *5.3.1 Power Grid Constraints*

Constraint (5.3.1a) enforces the power flow in the lines to stay within the rate-c limit after a generator contingency occurs. The required contingency reserve and the limits on the available contingency reserve are imposed through constraints (5.3.1b)-

(5.3.1e).

$$\left| \sum_{n \in \mathcal{V}_e} PF^R_{b,n} \left( \sum_{g \in \mathcal{G}(n)} (P_g(t) + r^c_g(t)) - d_n(t) \right) - PF^R_{b,n(c)} P_{g=c}(t) \right| \leq Pl^{max,c}_b, \quad \forall t, b, c \tag{5.3.1a}$$

$$\sum_{g \in \mathcal{G}} r^c_g(t) = P_{g=c}(t), \quad \forall t, c \tag{5.3.1b}$$

$$r^c_{g=c}(t) = 0, \quad \forall t, c \tag{5.3.1c}$$

$$0 \leq r^c_g(t) \leq [0, 0, 1, 1].\mathbf{R}_g(t), \quad \forall t, g, c \tag{5.3.1d}$$

$$P^{min}_g.u_g(t) \leq P_g(t) + r^c_g(t), \quad \forall t, g, c \tag{5.3.1e}$$

### *5.3.2 Gas Grid Constraints*

For each contingency , $c$, and time $t$, the required contingency gas reserve after loss of a generator and the limits on the gas pressure and the available gas reserve during the contingency are imposed through constraints presented in (5.3.2a)-(5.3.2k).

$$0 \leq rg^c_w(t) \leq rg_w(t), \quad \forall w \tag{5.3.2a}$$

$$\sum_{w \in \mathcal{W}} rg^c_w(t) = \sum_{g \in \mathcal{G}_w} r^c_g(t)\eta_g, \tag{5.3.2b}$$

$$\sum_{w \in \mathcal{W}(i)} (G_w(t) + rg^c_w(t)) + \sum_{k \in \mu^+(i)} \left( f^{c^+}_k(t) - f^{c^-}_k(t) \right) - \sum_{k \in \mu^-(i)} \left( f^{c^+}_k(t) - f^{c^-}_k(t) \right) - \sum_{g \in \mathcal{S}(i)} (P_g(t) + r^c_g(t))\eta_g = GL_i(t), \quad \forall i \tag{5.3.2c}$$

$$\rho^2_k(t)\tilde{\pi}^c_i(t) - \tilde{\pi}^c_j(t) = h(k)(\gamma^{c^+}_k(t) - \gamma^{c^-}_k(t)), \quad \forall k \tag{5.3.2d}$$

$$\gamma^{c^+}_k(t) >= \alpha[n] f^{c^+}_k(t) + \beta[n], \quad \forall k, n \tag{5.3.2e}$$

$$\gamma^{c^-}_k(t) >= \alpha[n] f^{c^-}_k(t) + \beta[n], \quad \forall k, n \tag{5.3.2f}$$

$$f_k^{c^-}(t) = 0, \qquad \forall k \in \mathcal{E}_g^a \qquad (5.3.2g)$$

$$0 \leq f_k^{c^+}(t) \leq M_1.Z_k^c(t), \qquad \forall k \qquad (5.3.2h)$$

$$0 \leq f_k^{c^-}(t) \leq M_1.(1 - Z_k^c(t)), \qquad \forall k \notin \mathcal{E}_g^a \qquad (5.3.2i)$$

$$\tilde{\pi}_i^{min} \leq \tilde{\pi}_i^c(t) \leq \tilde{\pi}_i^{max}, \qquad \forall i \qquad (5.3.2j)$$

$$\rho_k^2(t)\tilde{\pi}_i^c(t) \leq \tilde{\pi}_{\text{pipe}}^{max}, \qquad \forall k \in \mathcal{E}_g^a \qquad (5.3.2k)$$

## 5.4 Numerical Results

In this section, we first show how an uncoordinated UC and OGF can lead to an infeasible solution. We then show how solving GECUC for the base-case would mitigate the problem. Finally, we include the coordinated contingency analysis in the formulation of the GECUC.

In the simulation, the RTS-96 test case is considered as the power grid [Force(1999)]. The modified Belgium high-calorific gas grid in Fig. 5.4.1 is considered in this study, with original grid data found in [De Wolf and Smeers(2000)]. The following modifications/assumptions are made in the gas grid data-set:

- All the gas wells/storage maximum production limits are increased by 75% to account for the increased demand due to gas-fired generators,
- The given gas load data is the net demand for a day that is first divided by 24 to obtain average hourly load. This hourly value is considered as the mean of a Gaussian random demand with standard deviation equal to $10^{-3}$ of the mean value to consider for daily variations.
- The average hourly load of node 20 is assumed to be $1.5 \times 10^4 m^3/h$ higher than the given value.
- It is assumed that node 18 has a load with mean value $10^5 m^3/h$, and therefore

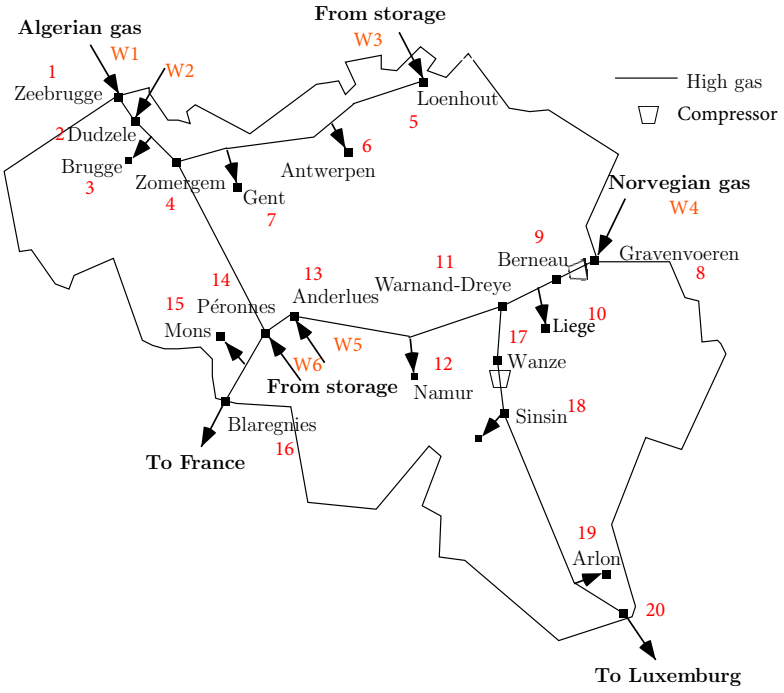



Figure 5.4.1: Belgium High-Calorific Natural Gas Grid.

the minimum pressure is changed from 0 in the main data-set to be equal to minimum pressure of node 10.

- The values for $h_{ij}$ are converted to hourly basis from daily basis.
- In the original data-set, there are a few parallel pipelines that are removed in the modified case. For parallel pipelines with different properties, the one with smaller diameter is removed.
- 50 MW gas-fired generators in RTS are modeled as combined cycle with efficiency 50%.
- 20 MW gas-fired generators in RTS are modeled as only gas turbines with efficiency 35%.
- It is assumed that 1 $ft^3$ of natural-gas can produce energy equal to 1109 $BTU$.

### 5.4.1 Uncoordinated UC and OGF Results

In this case, the UC is first solved and the gas-fired generators dispatch is sent to the OGF as parameters from the UC. The location of the gas fired generators that are assumed to be supplied from the Belgium grid are listed below:

| Gas Node # | RTS Bus # | Generators IDs |
|---|---|---|
| 12 | 101 | 1,2 (of 4) |
| 18 | 102 | 1,2 (of 4) |
| 20 | 122 | 1,2 (of 6) |
| 6 | 322 | 3,4,5 (of 6) |

Also, the compressor ratios are assumed to be: $\rho_{8-9}(t) = 1, \rho_{17-18}(t) = 1.25$.

The UC is solved with MIP Gap=0.1%, using GUROBI solver. In Fig. 5.4.2 top, the dispatch of some of the gas generators supplied from the studied gas grid is presented. From Fig. 5.4.2 bottom, it can be observed that if this schedule were to be implemented, the gas-flow limit on pipeline 17-18 would be violated. This flow limit has been obtained assuming one end has the maximum pressure and the other end has the minimum pressure, and therefore the flow in he pipeline is at its maximum. This situation may not be attained because of the limits that other nodes put on the pressure. For this uncoordinated scenario, what would happen in practice is a shortage in gas-delivery due to the pressure drop. This example highlights the fact that uncoordinated scheduling of the gas and electric can lead to infeasibility.

### 5.4.2 GECUC Results

In this case, with the same generator locations and assumption as the uncoordinated case, we run the optimization using the unified formulation with MIP Gap=0.1%. Comparing the generators dispatch in this case to the uncoordinated

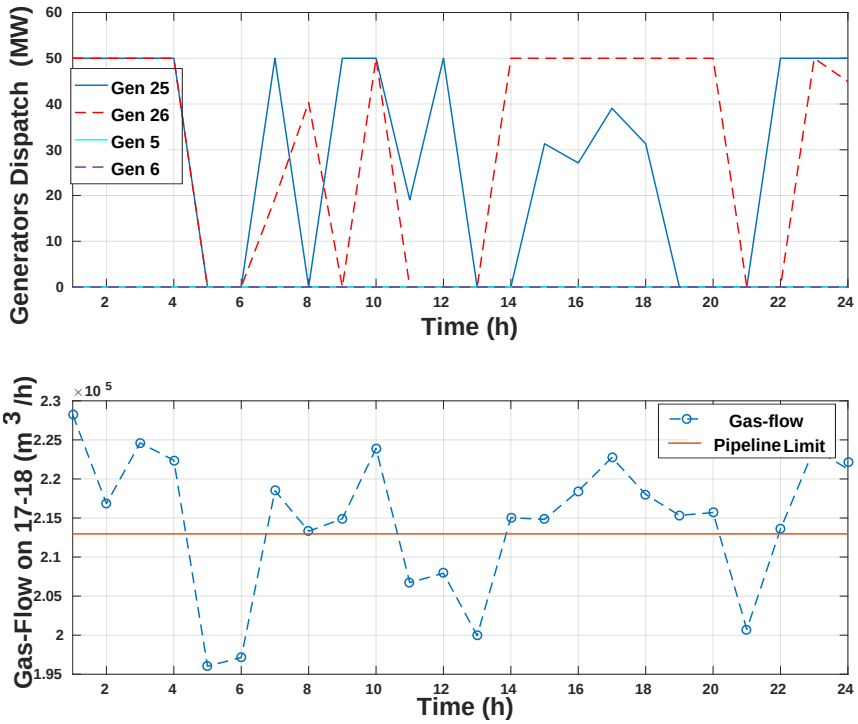



Figure 5.4.2: Uncoordinated UC and OGF.

Table 5.1: Electric Grid UC Cost

| | |
|---|---|
| Uncoordinated | $2.9484M |
| Coordinated | $2.9676M |

case, it can be seen that the uncoordinated case changes the gas-fired generators profile. Sometimes they are not committed to respect the gas grid constraints. Intuitively, this should have come at the expense of higher electric grid UC cost. As expected, the cost in the coordinated case is higher than that of uncoordinated (c.f. Table. 5.1) but the results are feasible for both gas and electric infrastructures.

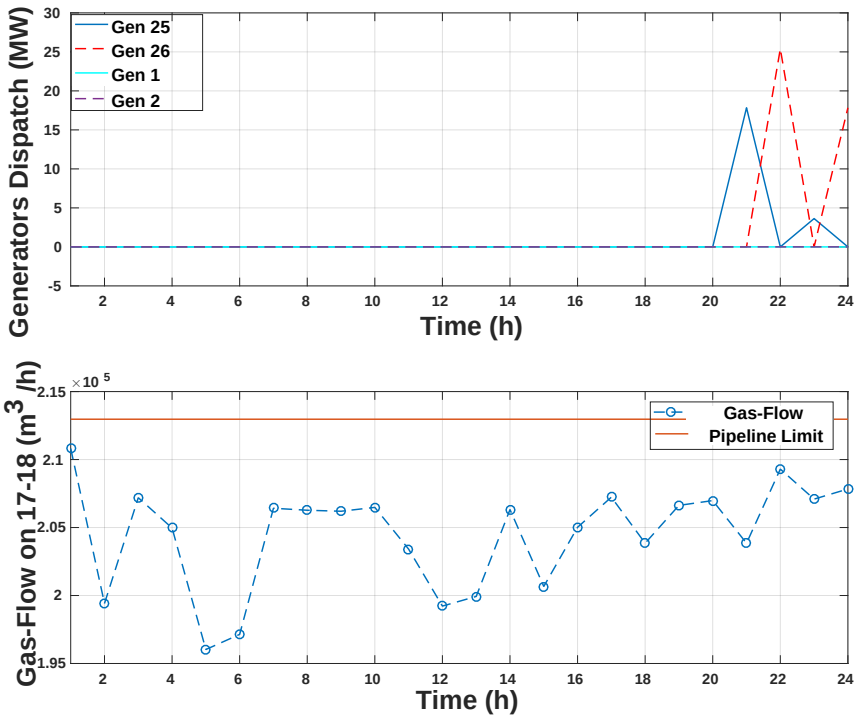



Figure 5.4.3: Dispatch results for GECUC.

### *5.4.3 GECUC with coordinated N-1 Generator Contingency Analysis Results*

In the scenario, the GECUC is solve along with the coordinated N-1 generator contingency analysis. The following gas-fired generators in RTS are assumed to be supplied from the Belgium Grid:

| Gas Node # | RTS Bus # | Generators IDs |
|---|---|---|
| 20 | 101 | 1,2 (of 4) |
| 18 | 102 | 1,2 (0f 4) |
| 12 | 122 | 1,2,3 (of 6) |
| 6 | 322 | 3,4,5 (of 6) |

Also, the compressor ratios are assumed to be: $\rho_{8-9}(t) = 1, \rho_{17-18}(t) = 1.25$ To illustrate a case, where coordinated N-1 generator contingency analysis is necessary, the assumed list of gas-fired generators supplied by the Belgium Grid is different in

here compared to part A.

Fig. 5.4.4 illustrates the reserves that need to be activated in the gas wells/storage and gas-fired generators when generator 90 is out of service. It can be noted that the gas reserve pattern successfully tracks the pattern of the gas-fired generators activated reserve to meet the required demand.

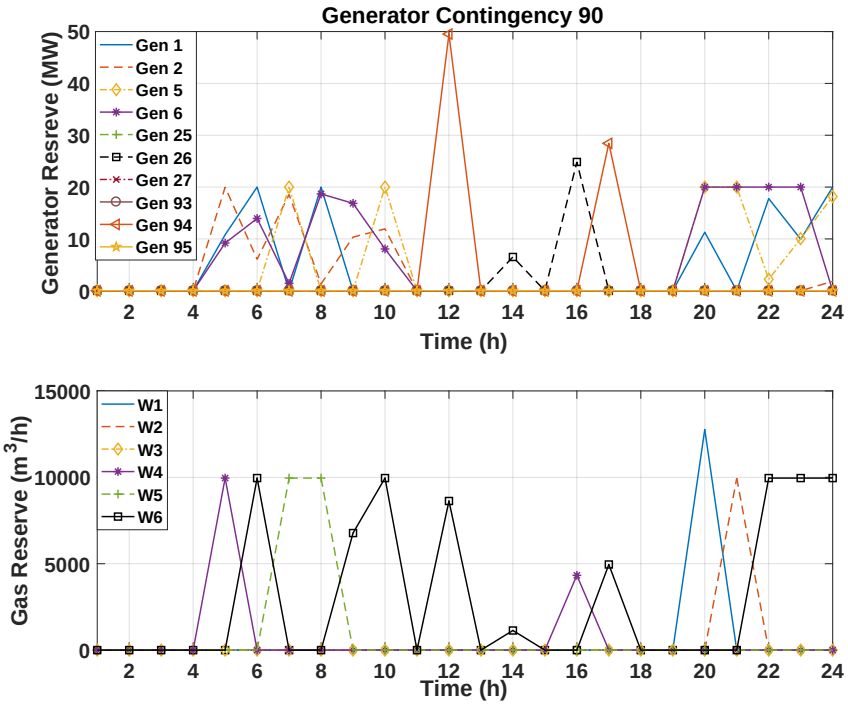



Figure 5.4.4: Activated Gas Generators and Gas Wells Reserve for Gen-90 Contingency.

[515(2009)] "IEEE Recommended Practice for Monitoring Electric Power Quality", IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995) pp. c1–81 (2009).

[tec(2015)] *Enabling Modernization of the Electric Power System* (U.S. Department of Energy, Quadrennial Technology Review 2015).

[Abdelkader and Morrow(2012)] Abdelkader, S. M. and D. J. Morrow, "Online tracking of Thévenin equivalent parameters using PMU measurements", IEEE Transactions on Power Systems **27**, 2, 975–983 (2012).

[Alinejad and Karegar(2016)] Alinejad, B. and H. K. Karegar, "On-line thevenin impedance estimation based on pmu data and phase drift correction", IEEE Transactions on Smart Grid (2016).

[Allen *et al.*(2014)] Allen, A., M. Singh, E. Muljadi and S. Santoso, *PMU Data Event Detection: A User Guide for Power Engineers* (National Renewable Energy Laboratory, 2014).

[Ardakanian *et al.*(2017)] Ardakanian, O., Y. Yuan, R. Dobbe, A. von Meier, S. Low and C. Tomlin, "Event detection and localization in distribution grids with phasor measurement units", in "2017 IEEE Power Energy Society General Meeting", pp. 1–5 (2017).

[Arefifar and Xu(2009)] Arefifar, S. A. and W. Xu, "Online tracking of power system impedance parameters and field experiences", IEEE Transactions on Power Delivery **24**, 4, 1781–1788 (2009).

[Basseville *et al.*(1993)] Basseville, M., I. V. Nikiforov *et al.*, *Detection of abrupt changes: theory and application*, vol. 104 (Prentice Hall Englewood Cliffs, 1993).

[Berthier *et al.*(2010)] Berthier, R., W. H. Sanders and H. Khurana, "Intrusion detection for advanced metering infrastructures: Requirements and architectural directions", in "Proc. SmartGridComm", pp. 350–355 (IEEE, 2010).

[Biswal *et al.*(2016)] Biswal, M., S. M. Brahma and H. Cao, "Supervisory protection and automated event diagnosis using pmu data", IEEE Transactions on Power Delivery **31**, 4, 1855–1863 (2016).

[Borghetti *et al.*(2008)] Borghetti, A., M. Bosetti, M. Di Silvestro, C. A. Nucci and M. Paolone, "Continuous-wavelet transform for fault location in distribution power networks: Definition of mother wavelets inferred from fault originated transients", IEEE Transactions on Power Systems **23**, 2, 380–388 (2008).

[Brahma *et al.*(2016)] Brahma, S., R. Kavasseri, H. Cao, N. R. Chaudhuri, T. Alexopoulos and Y. Cui, "Real time identification of dynamic events in power systems using pmu data, and potential applications - models, promises, and challenges", IEEE Transactions on Power Delivery **PP**, 99, 1–1 (2016).

[Brahma(2011)] Brahma, S. M., "Fault location in power distribution system with penetration of distributed generation", IEEE transactions on power delivery **26**, 3, 1545–1553 (2011).

[Cárdenas *et al.*(2011)] Cárdenas, A. A., S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang and S. Sastry, "Attacks against process control systems: risk assessment, detection, and response", in "Proc. 6th ACM Symposium on Information, Computer and Communications Security", pp. 355–366 (2011).

[cbc news Technology and Science(2016)] cbc news Technology and Science, "cyberattack that crippled ukrainian power grid", http://www.cbc.ca/news/technology/ukraine-cyberattack-1.3398492 (2016).

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

[Chertkov *et al.*(2015a)] Chertkov, M., S. Backhaus and V. Lebedev, "Cascading of fluctuations in interdependent energy infrastructures: Gas-grid coupling", Applied Energy **160**, 541–551 (2015a).

[Chertkov *et al.*(2015b)] Chertkov, M., M. Fisher, S. Backhaus, R. Bent and S. Misra, "Pressure fluctuations in natural gas networks caused by gas-electric coupling", in "System Sciences (HICSS), 2015 48th Hawaii International Conference on", pp. 2738–2747 (IEEE, 2015b).

[Correa-Posada and Sanchez-Martin(2014)] Correa-Posada, C. M. and P. Sanchez-Martin, "Security-constrained optimal power and natural-gas flow", IEEE Transactions on Power Systems **29**, 4, 1780–1787 (2014).

[Corsi and Taranto(2008)] Corsi, S. and G. N. Taranto, "A real-time voltage instability identification algorithm based on local phasor measurements", IEEE Transactions on Power Systems **23**, 3, 1271–1279 (2008).

[Damavandi *et al.*(2011)] Damavandi, M. Y., I. Kiaei, M. K. Sheikh-El-Eslami and H. Seifi, "New approach to gas network modeling in unit commitment", Energy **36**, 10, 6243–6250 (2011).

[De Wolf and Smeers(2000)] De Wolf, D. and Y. Smeers, "The gas transmission problem solved by an extension of the simplex algorithm", Management Science **46**, 11, 1454–1465 (2000).

[Dugan(2012)] Dugan, R. C., "Reference guide: The open distribution system simulator (opendss)", Electric Power Research Institute, Inc **7** (2012).

[Dzafic *et al.*(2018)] Dzafic, I., R. A. Jabr, S. Henselmeyer and T. Donlagic, "Fault location in distribution networks through graph marking", IEEE Transactions on Smart Grid **9**, 2, 1345–1353 (2018).

[Eto *et al.*(2015)] Eto, J. H., E. M. Stewart, T. Smith, M. Buckner, H. Kirkham, F. Tuffner and D. Schoenwald, "Scoping study on research and priorities for distribution-system phasor measurement units", (2015).

[Farajollahi *et al.*(2018)] Farajollahi, M., A. Shahsavari, E. Stewart and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using Micro-PMU data", IEEE Transactions on Power Systems (2018).

[for Power Systems Working Group(2011)] for Power Systems Working Group, S. M., "IEEE Standard for Synchrophasor Measurements for Power Systems", IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005) pp. 1–61 (2011).

[Force(1999)] Force, R. T., "The ieee reliability test system-1996", IEEE Trans. Power Syst **14**, 3, 1010–1020 (1999).

[Ge *et al.*(2015)] Ge, Y., A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee and D.-Y. Kwon, "Power system real-time event detection and associated data archival reduction based on synchrophasors", IEEE Transactions on Smart Grid **6**, 4, 2088–2097 (2015).

[Gentz(2017)] Gentz, R., *Wireless Sensor Data Transport, Aggregation and Security*, Ph.D. thesis, Arizona State University (2017).

[Group(2010)] Group, D. T. F. W., http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html (2010).

[Gruenspecht(2010)] Gruenspecht, H., "International energy outlook 2011", Center for Strategic and International Studies (2010).

[Hart(1986)] Hart, P., "Characterising the power system at a load busbar by measurement", in "IEE Proceedings C-Generation, Transmission and Distribution", vol. 133, pp. 87–94 (IET, 1986).

[IEEE(2012)] IEEE, "IEEE Std 1815-2012 (Revision of IEEE Std 1815-2010) - IEEE Standard for Electric Power Systems Communications-Distributed Network Protocol (DNP3)", URL `http://standards.ieee.org/findstds/standard/1815-2012.html` (2012).

[Jamei *et al.*(2018a)] Jamei, M., A. Scaglione and S. Peisert, "Low-resolution fault localization using phasor measurement units with community detection", in "Accepted in 2018 IEEE International Conference on Smart Grid Communications (SmartGridComm)", (2018a).

[Jamei *et al.*(2017a)] Jamei, M., A. Scaglione, C. Roberts, A. McEachern, E. Stewart, S. Peisert and C. McParland, "Online thevenin parameter tracking using synchrophasor data", in "Power & Energy Society General Meeting, 2017 IEEE", pp. 1–5 (IEEE, 2017a).

[Jamei *et al.*(2017b)] Jamei, M., A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland and A. McEachern, "Anomaly detection using optimally-placed $\mu$PMU sensors in distribution grids", IEEE Transactions on Power Systems (2017b).

[Jamei *et al.*(2017c)] Jamei, M., A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland and A. McEachern, "Automated anomaly detection in distribution grids using upmu measurements", in "Proceedings of the 50th Hawaii International Conference on System Sciences", (2017c).

[Jamei *et al.*(2018b)] Jamei, M., E. Schweitzer, A. Scaglione and K. Hedman, "Gas and electric grid unit commitment with coordinated n-1 generator contingency analysis", in "2018 Power System Computation Conference", (2018b).

[Jamei *et al.*(2016)] Jamei, M., E. Stewart, S. Peisert, A. Scaglione, C. McParland, C. Roberts and A. McEachern, "Micro synchrophasor-based intrusion detection in automated distribution systems: Toward critical infrastructure security", IEEE Internet Computing **20**, 5, 18–27 (2016).

[Jamei *et al.*(2018c)] Jamei, M., T. Tesfay, R. Gentz, C. Roberts, A. Scaglione and S. Peisert, "On localizing faults without observability using phasor measurement units", Submitted to IEEE Transactions on Smart Grid (2018c).

[Jiang *et al.*(2014)] Jiang, H., J. J. Zhang, W. Gao and Z. Wu, "Fault detection, identification, and location in smart grid based on data-driven computational methods", IEEE Transactions on Smart Grid **5**, 6, 2947–2956 (2014).

[Kashyap *et al.*(2015)] Kashyap, N., C.-W. Yang, S. Sierla and P. G. Flikkema, "Automated fault location and isolation in distribution grids with distributed control and unreliable communication", IEEE Transactions on Industrial Electronics **62**, 4, 2612–2619 (2015).

[Kersting(2001)] Kersting, W. H., "Radial distribution test feeders", in "Power Engineering Society Winter Meeting, 2001. IEEE", vol. 2, pp. 908–912 (IEEE, 2001).

[Kersting(2012)] Kersting, W. H., *Distribution system modeling and analysis* (CRC press, 2012).

[Kezunovic *et al.*(2013)] Kezunovic, M., L. Xie and S. Grijalva, "The role of big data in improving power system operation and protection", in "2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid", pp. 1–9 (2013).

[Kim and Poor(2011)] Kim, T. T. and H. V. Poor, "Strategic protection against data injection attacks on power grids", Smart Grid, IEEE Transactions on **2**, 2, 326–333 (2011).

[Koutsandria *et al.*(2015)] Koutsandria, G., R. Gentz, M. Jamei, A. Scaglione, S. Peisert and C. McParland, "A real-time testbed environment for cyber-physical security on the power grid", in "Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy", pp. 67–78 (ACM, 2015).

[Kullback and Leibler(1951)] Kullback, S. and R. A. Leibler, "On Information and Sufficiency", Ann. Math. Statist. **22**, 1, 79–86 (1951).

[Lee(2014)] Lee, J., "Automatic fault location on distribution networks using synchronized voltage phasor measurement units", in "ASME 2014 Power Conference", pp. V002T14A008–V002T14A008 (American Society of Mechanical Engineers, 2014).

[Liu *et al.*(2009)] Liu, C., M. Shahidehpour, Y. Fu and Z. Li, "Security-constrained unit commitment with natural gas transmission constraints", IEEE Transactions on Power Systems **24**, 3, 1523–1536 (2009).

[Lorden(1971)] Lorden, G., "Procedures for reacting to a change in distribution", The Annals of Mathematical Statistics pp. 1897–1908 (1971).

[Mallat and Zhang(1993)] Mallat, S. G. and Z. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Transactions on signal processing **41**, 12, 3397–3415 (1993).

[Manual and PowerFactory(2009)] Manual, D. P. F. and D. PowerFactory, "Version 14.0", DIgSILENT GmbH, Gomaringen, Germany (2009).

[McParland *et al.*(2014)] McParland, C., S. Peisert and A. Scaglione, "Monitoring security of networked control systems: It's the physics", Security & Privacy, IEEE **12**, 6, 32–39 (2014).

[Min and Santoso(2017)] Min, K. W. and S. Santoso, "DC offset removal algorithm for improving location estimates of momentary faults", IEEE Transactions on Smart Grid (2017).

[Misra *et al.*(2015)] Misra, S., M. W. Fisher, S. Backhaus, R. Bent, M. Chertkov and F. Pan, "Optimal compression in natural gas networks: A geometric programming approach", IEEE Transactions on Control of Network Systems **2**, 1, 47–56 (2015).

[Mitchell and Chen(2014)] Mitchell, R. and I.-R. Chen, "A survey of intrusion detection techniques for cyber-physical systems", ACM Computing Surveys (CSUR) **46**, 4, 55 (2014).

[Modbus(2004)] Modbus, I., "Modbus messaging on tcp", IP Implementation Guide v1. 0a, North Grafton, Massachusetts (www. modbus. org/specs. php) (2004).

[Nocedal and Wright(2006)] Nocedal, J. and S. Wright, *Numerical optimization* (Springer Science & Business Media, 2006).

[Open DNP3 Group and others(2012)] Open DNP3 Group and others, "DNP3–distributed network protocol 3.0–google project hosting", URL `https://www.automatak.com/opendnp3/` (2012).

[Page(1954)] Page, E., "Continuous inspection schemes", Biometrika pp. 100–115 (1954).

[Pai(2012)] Pai, A., *Energy function analysis for power system stability* (Springer Science & Business Media, 2012).

[Pan *et al.*(2015)] Pan, S., T. Morris and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems", Smart Grid, IEEE Transactions on **6**, 6, 3104–3113 (2015).

[Parniani *et al.*(2006)] Parniani, M., J. H. Chow, L. Vanfretti, B. Bhargava and A. Salazar, "Voltage stability analysis of a multiple-infeed load center using phasor measurement data", in "Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES", pp. 1299–1305 (IEEE, 2006).

[Paxson(1999)] Paxson, V., "Bro: a system for detecting network intruders in real-time", Computer networks **31**, 23, 2435–2463 (1999).

[Peisert *et al.*(2017)] Peisert, S., R. Gentz, J. Boverhof, C. McParland, S. Engle, A. Elbashandy and D. Gunter, "LBNL Open Power Data", `http://powerdata-explore.lbl.gov` (2017).

[Peisert *et al.*(2018)] Peisert, S., R. Gentz, M. Jamei, J. Boverhof, C. McParland, S. Engle, A. Elbashandy and D. Gunter, "LBNL Stream-Processing Architecture for Real-time Cyber-physical Security (SPARCS)", `https://github.com/lbnl-cybersecurity/sparcs` (2018).

[Phadke and de Moraes(2008)] Phadke, A. and R. M. de Moraes, "The wide world of wide-area measurement", Power and Energy Magazine, IEEE **6**, 5, 52–65 (2008).

[Phadke and Thorp(2008)] Phadke, A. G. and J. S. Thorp, *Synchronized phasor measurements and their applications* (Springer Science & Business Media, 2008).

[Power Standards Lab(2016)] Power Standards Lab, "PQube3 information", http://www.powersensorsltd.com/PQube3.php, accessed: 2016-01-15 (2016).

[Qadrdan *et al.*(2014)] Qadrdan, M., J. Wu, N. Jenkins and J. Ekanayake, "Operating strategies for a gb integrated gas and electricity network considering the uncertainty in wind power forecasts", IEEE Transactions on Sustainable Energy **5**, 1, 128–138 (2014).

[Sanchez-Garcia *et al.*(2014)] Sanchez-Garcia, R. J., M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki and J. W. Bialek, "Hierarchical spectral clustering of power grids", IEEE Transactions on Power Systems **29**, 5, 2229–2237 (2014).

[Seber and Wild(1989)] Seber, G. and C. Wild, *Nonlinear regression* (Wiley, New York, 1989).

[Slay and Miller(2008)] Slay, J. and M. Miller, *Lessons learned from the maroochy water breach* (Springer, 2008).

[Smon *et al.*(2006)] Smon, I., G. Verbic and F. Gubina, "Local voltage-stability index using Tellegen's theorem", IEEE Transactions on Power Systems **21**, 3, 1267–1275 (2006).

[Terzija *et al.*(2011)] Terzija, V., G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks", Proceedings of the IEEE **99**, 1, 80–93 (2011).

[Tesfay *et al.*(2014)] Tesfay, T. T., J.-P. Hubaux, J.-Y. Le Boudec and P. Oechslin, "Cyber-secure communication architecture for active power distribution networks", in "Proceedings of the 29th Annual ACM Symposium on Applied Computing", pp. 545–552 (ACM, 2014).

[Tsai and Wong(2008)] Tsai, S.-J. S. and K.-H. Wong, "On-line estimation of Thevenin equivalent with varying system states", in "Proc. Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century", pp. 1–7 (IEEE, 2008).

[Uluski(2012)] Uluski, R., "Using distribution automation for a self-healing grid", in "Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES", pp. 1–5 (IEEE, 2012).

[Urbina and Li(2007)] Urbina, M. and Z. Li, "A combined model for analyzing the interdependency of electrical and gas systems", in "Power Symposium, 2007. NAPS'07. 39th North American", pp. 468–472 (IEEE, 2007).

[Valenzuela *et al.*(2013)] Valenzuela, J., J. Wang and N. Bissinger, "Real-time intrusion detection in power system operations", Power Systems, IEEE Transactions on **28**, 2, 1052–1062 (2013).

[von Meier *et al.*(2014)] von Meier, A., D. Culler, A. McEachern and R. Arghandeh, "Micro-synchrophasors for distribution systems", in "Proc. IEEE PES Innovative Smart Grid Technologies Conference (ISGT)", pp. 1–5 (2014).

[Vu *et al.*(1999)] Vu, K., M. M. Begovic, D. Novosel and M. M. Saha, "Use of local measurements to estimate voltage-stability margin", IEEE Transactions on Power Systems **14**, 3, 1029–1035 (1999).

[Wang *et al.*(2013)] Wang, Y., C. Wang, F. Lin, W. Li, L. Y. Wang and J. Zhao, "Incorporating generator equivalent model into voltage stability analysis", IEEE Transactions on Power Systems **28**, 4, 4857–4866 (2013).

[Wang *et al.*(2010)] Wang, Z., A. Scaglione and R. J. Thomas, "Compressing electrical power grids", in "Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on", pp. 13–18 (IEEE, 2010).

[Xia and Mandic(2012)] Xia, Y. and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems", IEEE Transactions on Instrumentation and Measurement **61**, 1, 74–83 (2012).

[Xie *et al.*(2014)] Xie, L., Y. Chen and P. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis", IEEE Transactions on Power Systems **29**, 6, 2784–2794 (2014).

[Yuan and Li(2014)] Yuan, H. and F. Li, "A comparative study of measurement-based thevenin equivalents identification methods", in "Proc. North American Power Symposium (NAPS)", pp. 1–6 (IEEE, 2014).

[Zhou *et al.*(2017)] Zhou, Y., R. Arghandeh and C. J. Spanos, "Partial knowledge data-driven event detection for power distribution networks", IEEE Transactions on Smart Grid (2017).

[Zhu *et al.*(1997)] Zhu, J., D. L. Lubkeman and A. A. Girgis, "Automated fault location and diagnosis on electric power distribution feeders", IEEE Transactions on Power Delivery **12**, 2, 801–809 (1997).

[Zlotnik *et al.*(2017)] Zlotnik, A., L. Roald, S. Backhaus, M. Chertkov and G. Andersson, "Coordinated scheduling for interdependent electric power and natural gas infrastructures", IEEE Transactions on Power Systems **32**, 1, 600–610 (2017).