Sparse Sensing Methods for Model-Free Sensitivity Estimation and Topology Change Detection using Synchro-Phasor Measurements

Final Project Report

S-59

Power Systems Engineering Research Center

Empowering Minds to Engineer
the Future Electric Energy System

Sparse Sensing Methods for Model-Free Sensitivity Estimation and Topology Change Detection using Synchro-Phasor Measurements (S-59)

Final Project Report

Project Team

Alejandro D. Domínguez-García, Project Leader
Peter W. Sauer
Yu Christine Chen, Graduate Student*
Xichen Jiang, Graduate Student**
University of Illinois at Urbana-Champaign

Christopher L. DeMarco
Stephen J. Wright
Sowmya Acharya, Graduate Student
Taedong Kim, Graduate Student
University of Wisconsin–Madison

PSERC Publication 16-09

December 2016

* Currently an Assistant Professor at the University of British Columbia.
**Currently an Assistant Professor at Western Washington University

For information about this project, contact

Alejandro D. Domínguez-García
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
306 N. Wright Street,
Urbana, IL 61801
Phone: (217) 333-0394
Email: aledan@illinois.edu

Power Systems Engineering Research Center

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

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
Engineering Research Center #527
551 E. Tyler Mall
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Acknowledgments

We would like to express our appreciation for the continued support provided by PSERC member companies. In particular we would like to thank the industry advisors of this project: Mirrasoul Mousavi (ABB); Prashant Kansal (AEP); Jim Kleitsch (ATC); Jim Gronquist (BPA); Evangelos Farantatos (EPRI); Alan Engelmann (Exelon-ComEd); Slava Maslennikov (ISO-NE); George Stefopoulos (NYPA); Angel A. Aquino-Lugo (PowerWorld); Mahendra Patel (PJM); Jay Caspary (SPP).

Executive Summary

State estimation, and health monitoring for general situational awareness at a power system control center, inevitably employs a mix of a priori information, based on system model and parameter data, and observations in the form of measurement data. In traditional power systems applications, when SCADA systems provided only low-bandwidth, unsynchronized measurement data to a control center, observations and a priori information typically contributed on nearly equal footing in applications such as state estimation and topology error identification. Today, the availability of high-bandwidth, time-synchronized PMU data shifts this balance, creating a much larger role for observations, and reducing the need for full model information, thereby opening the door to much faster time-scale estimation. However, rather than replacing existing state estimation tools and the applications that depend on them, the techniques developed in this project seek to provide much higher bandwidth estimation, to "fill in" situational awareness between slower, periodic state estimator updates. The particular algorithms developed focus on two applications: (A1) fast time-scale, dynamic detection and identification of network topology changes; and (A2) PMU measurement-based, (nearly) "model-free" estimation of the power flow Jacobian. The report is divided into four parts: Parts I-III cover the work conducted under Application A1, whereas Part IV covers the work conducted under Application A2.

Part I: Quickest Change Line Outage Detection and Identification
In this part, we propose a framework for detection and identification of system topological changes in near real-time that utilizes the statistical properties of electricity generation and demand, which are assumed to be known. Instead of relying on offline models as with traditional methods, the proposed method is model-free, and exploits the high-speed synchronized measurements provided by phasor measurement units (PMUs). In this framework, a statistical quickest change algorithm is applied to the voltage phase angle measurements collected from PMUs to detect the change-point that corresponds to the system topology change instant. An advantage of this algorithm is that the operator also has full control over the tradeoff between detection delay and false alarm rate. In order to lend support for the work conducted, case studies are done through simulations on standard IEEE test systems

Part II: PMU Placement via Multinomial Logistic Regression
In this part, we also consider the problem of identifying a single line outage in a power grid by using data from PMUs. When a line outage occurs, the voltage phasor of each bus node changes in response to the change in network topology. Each individual line outage has a distinctive "signature," and a multinomial logistic regression (MLR) classifier can be trained to distinguish between these signatures reliably. We consider first the ideal case in which PMUs are attached to every bus. We then describe techniques from regularized optimization for placing PMUs selectively on a subset of buses, with the subset being chosen to allow discrimination between as many outage events as possible. Experimental results with synthetic 24-hour demand profile data generated for several IEEE test systems.

Part III: An Observer-Based Approach of Topology Change Estimation
In this part, the goal here is identification of topology errors from PMU data, exploiting dynamic characteristics of the power systems response, while being able to discriminate topology errors from bad data. Types of topology errors common in Power Systems include incorrect status of line switches, transformer tap settings, shunt compensation, etc. The focus here is on line switching events. In this work, the estimation of line switching events from measured errors is sought by concentrating the disturbance input due to the line outage(s) into an invariant subspace of the state matrix of the Luenberger observer. This formulation has the advantage of allowing operation of multiple observers in parallel, facilitating simultaneous monitoring of multiple line switching events.

Part IV: Measurement-Based Estimation of the Power Flow Jacobian
In this part, we propose a measurement-based method to compute the power flow Jacobian matrix, from which we can infer pertinent information about the system topology in near real-time. A salient feature of our approach is that it readily adapts to changes in system operating point and topology; this is desirable as it provides power system operators with a way to update, as the system evolves, the models used in many reliability analysis tools. The method uses high-speed synchronized voltage and current phasor data collected from phasor measurement units to estimate entries of the Jacobian matrix through linear total least-squares (TLS) estimation. In addition to centralized TLS-based algorithms, we provide distributed alternatives aimed at reducing computational burden. Through numerical case studies involving standard IEEE test systems, we illustrate the effectiveness of our proposed Jacobian-matrix estimation approach as compared to the conventional model-based one.

Project Publications:
The work reported in the publications below was supported in part by project S-59.

Journal Papers:

[J1] T. Kim, and S. J. Wright, "PMU Placement for Line Outage Identification via Multiclass Logistic Regression," *IEEE Transactions on Smart Grid*, to appear.

[J2] Y. C. Chen, J. Wang, A. D. Domínguez-Garcia, "Measurement-Based Estimation of the Power Flow Jacobian Matrix," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, 2507-2515, September 2016.

[J3] Y. C. Chen, T. Banerjee, A. D. Domínguez-Garcia, and V. V. Veeravalli, "Quickest Line Outage Detection and Identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749-758, January 2016.

Conference Papers:

[C1] G. Rovatsos, X. Jiang, A. D. Domínguez-Garcia, and V. V. Veeravalli, "Comparison of Statistical Algorithms for Power System Line Outage Detection," *in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Shanghai, China, March 2016.

Student Theses:

[T1] X. Jiang, *Real-Time Power System Topology Change Detection and Identification*, Ph.D. Thesis, University of Illinois at Urbana-Champaign, August 2016.

[T2] S. Acharya, *Topology Error Estimation in Power System Dynamic Models*, MS Thesis, University of Wisconsin-Madison, December 2015.

[T3] T. Kim, *Optimization Approaches for Reliable Operation of Electric Power Systems*, Ph.D. Thesis, University of Wisconsin–Madison, July 2015.

[T4] Y. C. Chen, *Measurement-Based Tools for Power System Monitoring and Operations*, Ph.D. Thesis, University of Illinois at Urbana-Champaign, December~2014.

Part I

Quickest Line Outage Detection and Identification

Alejandro D. Domínguez-García

Xichen Jiang, Graduate Student*

University of Illinois at Urbana-Champaign

*Currently an Assistant Professor at Western Washington University.

For information about this part of the project report, contact:

Alejandro D. Domínguez-García
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
306 N. Wright Street,
Urbana, IL 61801
Phone: (217) 333-0394
Email: aledan@illinois.edu

Power Systems Engineering Research Center

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

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Timely line outage detection for power systems is crucial for maintaining operational reliability. Currently, many of the methods for online power system monitoring rely on a system model that is obtained offline, which can be inaccurate due to bad historical or telemetry data; such inaccuracies have been a contributing factor in many recent blackouts. For example, in the 2011 San Diego blackout, operators were unable to determine overloaded lines because the network model was not up to date [1]. This lack of situational awareness limited the operators' ability to identify and prevent the next critical contingency, leading to instability and cascading failures. Similarly, during the 2003 Northeast blackout, operators failed to initiate the correct control schemes because they had an inaccurate model of the system, and could not identify the loss of key transmission elements [2]. These blackouts highlight the importance of developing online techniques to detect and identify system topological changes. This work addresses the problems discussed above by establishing a framework for quickly detecting system topological changes. Specifically, we focus on the problem of line outage detection in power systems, and exploit fast measurements provided by PMUs to develop a statistical method that allows for quick detection of changes in the power system network topology.

Early approaches for topological change detection and identification include algorithms based on state estimation [3]–[5], and rule-based algorithms that mimic system operator decisions [6]. The issue of external system topology error detection was explored in [7]. More recent proposed methods exploit the fast sampling of voltage magnitudes and phases provided by PMUs to detect events in a power system in near real-time [8]–[11]. While these works allow for improved situational awareness of the power system, they do have shortcomings. Mainly, they do not exploit the fact that the line outage is persistent; i.e., once a line outage occurs, it persists until it is detected and brought back into service. Instead, only the most recent PMU measurement is used to determine if an outage has occurred. The authors of [12] proposed a method to detect line outages using statistical classifiers where a maximum likelihood estimation is performed on the PMU data. The authors also considered the transient response of the system after a line outage by comparing synthesized data against actual data. However, their method requires the exact instant the line outage occurs to be known before applying the algorithm, whereas the method we propose in this work does not have this restriction.

In [13], [14], the authors proposed a statistical method based on the theory of quickest change detection (QCD) for line outage detection and identification. This method observes a sequence of measured voltage phase angles provided by PMUs and exploits the fact that their statistics change following a line outage. The objective is to detect this change in distribution quickly while subject to a fixed false alarm rate. The statistics of the measured voltage angles pre- and post line outage are related to the known distributions in the real power injections through a linear mapping involving a linearized power flow model. For this method, the incremental changes in real power injections are modeled as independent random variables. Then, the probability distribution of such incremental changes is mapped to that of the incremental changes in voltage phase angles via a linear transformation obtained from the power flow equations. The PMUs provide a random sequence of voltage phase angle measurements in real-time; when a line outage occurs, the probability distribution of the incremental changes in the voltage phase angles changes abruptly. The objective is to detect a change in this probability distribution after the occurrence of a line outage as quickly as possible while maintaining a desired false alarm rate. In the previous work in [13], [14], the Cumulative Sum (CuSum) algorithm was proposed to solve this problem. For this algorithm, a sequence of CuSum statistics is computed, one for each line in the system. An outage is declared when any one of the statistics crosses a prespecified threshold for the first time. The performance of this algorithm is characterized by a parameter known as the Kullback-Leibler (KL) divergence (see, e.g., [15]), which is a distance measure between the pre- and post-outage voltage phase angle distributions.

The remainder of this document is organized as follows. In Chapter 2, we provide the preliminary background of this research along with the statement of the problem to be addressed. Specifically, we introduce the power system model and the assumptions we adopt, the pre- and post-outage statistical model of the voltage phase angles, and the problem statement of quickest change detection. Starting from the nonlinear power balance equations, we derive the linearized incremental model of the system and then apply the DC power flow assumptions. The resulting proposed model captures the transient dynamics following a line outage. Finally, the statistical model for the incremental power injections and how they relate to the voltage phase angle statistics are also introduced. Chapter 3 outlines the QCD-based line outage identification algorithms, the CuSum algorithm and the Generalized Likelihood Ratio Test algorithm. Small examples are provided to illustrate this process. We also introduce the KL divergence, which is an important measure for characterizing the performance of these detection algorithms. Finally, various other algorithms for line outage detection that exist in the literature are discussed. These algorithms are "one-shot" detection schemes that do not exploit the persistence of the line outage. We compare the performance of our proposed method against those of others and show that the CuSum-based method is better. Chapter 4 summarizes what has been done as part of this research along with additional insights and remarks; future research directions are also provided in this chapter.

Chapter 2

Preliminaries

In this chapter, we present the power system model adopted in this research. A linearized small-signal power system model is used in conjunction with synchronized voltage phase angle measurements obtained from phasor measurement units. We provide a general framework where the system transient dynamics after a line outage are captured in the model. We then establish a statistical framework for both the pre- and post-outage scenarios that is used in the line detection algorithm. This chapter concludes with the statement of the line outage detection problem.

2.1 Power System Model

We represent the power system network by a graph consisting of N nodes and L edges, corresponding to buses and transmission lines, respectively. The set of buses is denoted by $\mathcal{V} = \{1, \ldots, N\}$, and the set of transmission lines is denoted by $\mathcal{E}$, where for $m, n \in \mathcal{V}$, $(m, n) \in \mathcal{E}$ if there exists a transmission line between buses m and n. At time t, let $V_i(t)$ and $\theta_i(t)$ denote the voltage magnitude and phase angle at bus i, and let $P_i(t)$ and $Q_i(t)$ denote the net active and reactive power injection at bus i. Then, the quasi-steady-state behavior of the system can be described by the power flow equations, which for bus i can be written as:

$$\begin{aligned} P_i(t) &= p_i(\theta_1(t), \ldots, \theta_N(t), V_1(t), \ldots, V_N(t)), \\ Q_i(t) &= q_i(\theta_1(t), \ldots, \theta_N(t), V_1(t), \ldots, V_N(t)), \end{aligned} \tag{2.1}$$

where the dependence on the system network parameters is implicitly captured by $p_i(\cdot)$ and $q_i(\cdot)$ (see e.g., [16]).

2.1.1 Pre-outage Incremental Power Flow Model

Let $P_i[k] \coloneqq P_i(k\Delta t)$ and $Q_i[k] \coloneqq Q_i(k\Delta t)$, $\Delta t > 0$, $k = 0, 1, 2, \ldots$, denote the k^{th} measurement sample of active and reactive power injections into bus i. Similarly, let $V_i[k]$ and $\theta_i[k]$ denote bus i's k^{th} voltage magnitude and angle measurement sample at $t = k\Delta t$, $k = 0, 1, 2, \ldots$. Furthermore, define variations in voltage magnitudes and phase angles between consecutive sampling times $k\Delta t$ and $(k+1)\Delta t$ as $\Delta V_i[k] \coloneqq V_i[k+1] - V_i[k]$

and $\Delta\theta_i[k] \coloneqq \theta_i[k+1] - \theta_i[k]$, respectively. Similarly, variations in the active and reactive power injections at bus i between two consecutive sampling times are defined as $\Delta P_i[k] = P_i[k+1] - P_i[k]$ and $\Delta Q_i[k] = Q_i[k+1] - Q_i[k]$.

Suppose a solution to the power flow equations exists at $(\theta_i[k], V_i[k], P_i[k], Q_i[k])$, $i = 1, \ldots, N$, such that $p_i(\cdot)$ and $q_i(\cdot)$ in (2.1) are continuously differentiable with respect to all θ_i and V_i at $\theta_i[k]$ and $V_i[k]$, $i = 1, \ldots, N$. Then, assuming that $\Delta\theta_i[k]$ and $\Delta V_i[k]$ are sufficiently small, we can approximate $\Delta P_i[k]$ and $\Delta Q_i[k]$ as

$$\Delta P_i[k] \approx \sum_{j=1}^{N} a_{ij}[k]\Delta\theta_j + \sum_{j=1}^{N} b_{ij}[k]\Delta V_j, \tag{2.2}$$

$$\Delta Q_i[k] \approx \sum_{j=1}^{N} c_{ij}[k]\Delta\theta_j + \sum_{j=1}^{N} d_{ij}[k]\Delta V_j, \tag{2.3}$$

where

$$a_{ij}[k] = \frac{\partial p_i}{\partial \theta_j},\ b_{ij}[k] = \frac{\partial p_i}{\partial V_j},\ c_{ij}[k] = \frac{\partial q_i}{\partial \theta_j},\ d_{ij}[k] = \frac{\partial q_i}{\partial V_j},$$

for each bus $i = 1, \ldots, N$, all evaluated at $(\theta_i[k], V_i[k], P_i[k], Q_i[k])$.

Under standard assumptions used in power system analysis, we assume that $b_{ij} << a_{ij}$ and $c_{ij} << d_{ij}$ in (2.2) and (2.3) [16]. This allows for the decoupling of (2.2) and (2.3) as the variations in the active power injections primarily affect the bus voltage angles; therefore, we can write $\Delta P_i[k] \approx \sum_{j=1}^{N} a_{ij}[k]\Delta\theta_j[k]$. Furthermore, under the so-called DC assumptions, namely (i) the system is lossless, (ii) $V_i[k] = 1$ per unit (p.u.) for all $i \in \mathcal{V}$, k, and (iii) $\theta_m[k] - \theta_n[k] << 1$ for all k and for $m, n \in \mathcal{V}$, $a_{ij}[k]$ simply becomes the negative of the imaginary part of the $(i,j)^{th}$ entry of the network admittance matrix and independent of k [16]. One of the buses must be designated as reference (i.e., $\theta = 0$) for the other buses in the system. Therefore, since the reference bus angle is assumed to be known, the equation for the reference bus is omitted from (2.2). We can express the variations in the voltage phase angles due to the variations in the real power flow in matrix form by

$$\Delta P[k] \approx H_0 \Delta\theta[k], \tag{2.4}$$

where $\Delta P[k]$, $\Delta\theta[k] \in \mathbb{R}^{(N-1)}$ and $H_0 \in \mathbb{R}^{(N-1)\times(N-1)}$. Note that the $N-1$ dimension of the vectors is the result of omitting the reference bus equation.

In an actual power system, random fluctuations in the load drive the generator response. Therefore, in this work, we use the so-called governor power flow model (see e.g., [17]), which is more realistic than the conventional power flow model, where the slack bus picks up any changes in the load power demand. In the governor power flow model, at time instant k, the relation between changes in the load demand vector, $\Delta P^d[k] \in \mathbb{R}^{N_d}$, and changes in the power generation vector, $\Delta P^g[k] \in \mathbb{R}^{N_g}$, is described by

$$\Delta P^g[k] = B[k]\Delta P^d[k], \tag{2.5}$$

where $B[k]$ is a time dependent matrix of participation factors. We approximate $B[k]$ by quantizing it to take values B_i, $i = 0, 1, \ldots, T$, where i denotes the time period of interest.

Let $B[k] = B_0$ and $M_0 \coloneqq H_0^{-1}$ during the pre-outage period. Then, we can substitute (2.5) into (2.4) to obtain a pre-outage relation between the changes in the voltage angles and the real power demand at the load buses as follows:

$$\begin{aligned}
\Delta\theta[k] &\approx M_0 \Delta P[k] \\
&= M_0 \begin{bmatrix} \Delta P^g[k] \\ \Delta P^d[k] \end{bmatrix} \\
&= [M_0^1 \ M_0^2] \begin{bmatrix} B_0 \Delta P^d[k] \\ \Delta P^d[k] \end{bmatrix} \\
&= (M_0^1 B_0 + M_0^2) \Delta P^d[k] \\
&= \tilde{M}_0 \Delta P^d[k],
\end{aligned} \tag{2.6}$$

where $\tilde{M}_0 = M_0^1 B_0 + M_0^2$.

2.1.2 Post-outage Incremental Power Flow Model

Suppose an outage occurs for the line (m, n) at time $t = t_f$, where $\gamma_0 \Delta t \leq t_f < (\gamma_0 + 1)\Delta t$. In addition, assume that the loss of line (m, n) does not cause islands to form in the post-event system (i.e., the underlying graph representing the internal power system remains connected).

Following a line outage, the power system undergoes a transient response governed by B_i, $i = 1, 2, \ldots, T-1$ until quasi steady state is reached, in which $B[k]$ settles to a constant B_T. For example, immediately after the outage occurs, the power system is dominated by the inertial response of the generators, which is then followed by the governor response. As a result of the line outage, the system topology changes, which manifests itself in the matrix H_0. This change in the matrix H_0 resulting from the outage can be expressed as the sum of the pre-outage matrix and a perturbation matrix, $\Delta H_{(m,n)}$, i.e., $H_{(m,n)} = H_0 + \Delta H_{(m,n)}$. Since H_0 has the same sparsity structure as the graph Laplacian of the internal system network, we conclude that the only non-zero terms in the matrix $\Delta H_{(m,n)}$ are $\Delta H_{(m,n)}[n, n] = -\Delta H_{(m,n)}[m, n] = -\Delta H_{(m,n)}[n, m] = \Delta H_{(m,n)}[m, m] = -1/X_{(m,n)}$, where $X_{(m,n)}$ is the imaginary part of the impedance of the outaged line. Thus, the matrix $\Delta H_{(m,n)}$ is a rank-one matrix and can be expressed as

$$\Delta H_{(m,n)} = -\frac{1}{X_{(m,n)}} r_{(m,n)} r_{(m,n)}^T, \tag{2.7}$$

where $r_{(m,n)} \in \mathbb{R}^{N-1}$ is a vector with the m^{th} entry equal to 1, the n^{th} entry equal to -1, and all other entries equal to 0.

We can use the matrix inversion lemma [18] to obtain an expression for $M_{(m,n)} \coloneqq H_{(m,n)}^{-1}$ as follows:

$$M_{(m,n)} = M_0 + \beta_{(m,n)} \ s_{(m,n)} \ s_{(m,n)}^T, \tag{2.8}$$

where $\beta_{(m,n)} = 1/(X_{(m,n)} - r_{(m,n)}^T H_0^{-1} r_{(m,n)})$ and $s_{(m,n)} = H_0^{-1} r_{(m,n)}$.

Then, by letting $M_{(m,n)} \coloneqq H^{-1}_{(m,n)} = [M^1_{(m,n)}\ M^2_{(m,n)}]$, and deriving in the same manner as the pre-outage model of (2.6), we obtain the post-outage relation between the changes in the voltage angles and the real power demand as:

$$\Delta\theta[k] \approx \tilde{M}_{(m,n),i}\Delta P^d[k], \quad \gamma_{i-1} \leq k < \gamma_i, \tag{2.9}$$

where $\tilde{M}_{(m,n),i} = M^1_{(m,n)}B_i + M^2_{(m,n)}$, $i = 1, 2, \ldots, T$.

2.1.3 Instantaneous Change During Outage

At the time of outage, $t = t_f$, there is an instantaneous change in the mean of the voltage phase angle measurements that affects only one incremental sample, namely, $\Delta\theta[\gamma_0] = \theta[\gamma_0 + 1] - \theta[\gamma_0]$. The measurement $\theta[\gamma_0]$ is taken immediately prior to the outage, whereas $\theta[\gamma_0 + 1]$ is the measurement taken immediately after the outage. Then, the effect of an outage in line (m, n) can be modeled with a power injection of $P_{(m,n)}[\gamma_0]$ at bus m and $-P_{(m,n)}[\gamma_0]$ at bus n, where $P_{(m,n)}[\gamma_0]$ is the pre-outage line flow across line (m, n) from m to n. Following a similar approach as the one in [13], the relation between the incremental voltage phase angle at the instant of outage, $\Delta\theta[\gamma_0]$, and the variations in the real power flow can be expressed as:

$$\Delta\theta[\gamma_0] \approx M_0\Delta P[\gamma_0] - P_{(m,n)}[\gamma_0 + 1]M_0 r_{(m,n)}, \tag{2.10}$$

where $r_{(m,n)} \in \mathbb{R}^{N-1}$ is a vector with the $(m-1)^{\text{th}}$ entry equal to 1, the $(n-1)^{\text{th}}$ entry equal to -1, and all other entries equal to 0. Furthermore, by using the governor power flow model of (2.5) and substituting into (2.10), and simplifying, we obtain:

$$\Delta\theta[\gamma_0] \approx \tilde{M}_0\Delta P^d[\gamma_0] - P_{(m,n)}[\gamma_0 + 1]M_0 r_{(m,n)}. \tag{2.11}$$

Example 1 (Three-Bus System)**.** Consider the lossless three-bus system shown in Fig. 2.1. The parameters for this system are listed in Table 2.1 and all quantities are in per unit. The nonlinear real power balance equations of (2.2) for this system are

$$\begin{aligned}
P_1 &= \frac{V_1V_2}{X_{1,2}}\sin(\theta_1 - \theta_2) + \frac{V_1V_3}{X_{1,3}}\sin(\theta_1 - \theta_3),\\
P_2 &= \frac{V_2V_1}{X_{1,2}}\sin(\theta_2 - \theta_1) + \frac{V_2V_3}{X_{2,3}}\sin(\theta_2 - \theta_3),\\
P_3 &= \frac{V_3V_1}{X_{1,3}}\sin(\theta_3 - \theta_1) + \frac{V_3V_2}{X_{2,3}}\sin(\theta_3 - \theta_2).
\end{aligned} \tag{2.12}$$

The first equation with P_1 is removed from (2.12) since bus 1 is the reference bus. Then, using the DC assumptions, the model of (2.12) can be approximated by a small-signal linear incremental model of the form in (2.4), where

$$\begin{aligned}
H_0 &= \begin{bmatrix} \frac{1}{X_{1,2}} + \frac{1}{X_{2,3}} & -\frac{1}{X_{2,3}} \\ -\frac{1}{X_{2,3}} & \frac{1}{X_{1,3}} + \frac{1}{X_{2,3}} \end{bmatrix}\\
&= \begin{bmatrix} 46.72 & -26.88 \\ -26.88 & 42.60 \end{bmatrix}.
\end{aligned}$$

Table 2.1: Parameter values for 3-bus system shown in Fig. 2.1.

P_2	P_3	$X_{1,2}$	$X_{2,3}$	$X_{1,3}$
-1	-0.9	0.0504	0.0372	0.0636

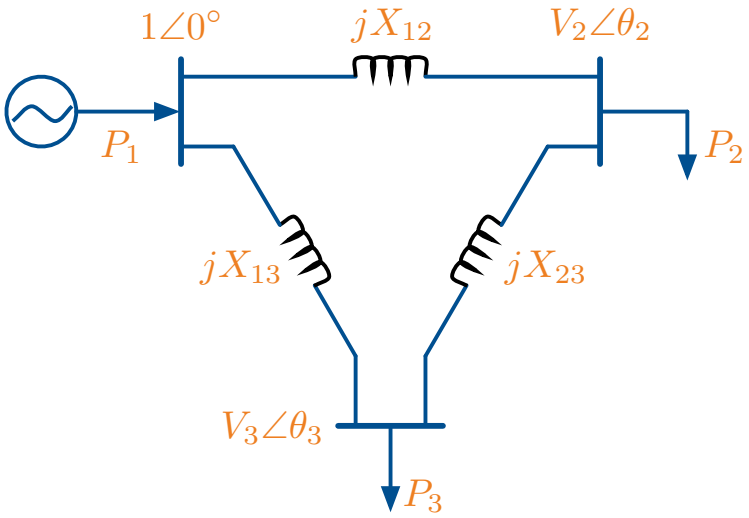


Figure 2.1: Network topology for 3-bus system.

Accordingly, $M_0 = H_0^{-1}$ is computed as:

$$M_0 = \begin{bmatrix} 0.033 & 0.021 \\ 0.021 & 0.037 \end{bmatrix}.$$

Suppose an outage occurs on line $(1, 2)$. Then, according to (2.7) and (2.8),

$$\begin{aligned} \Delta H_{(1,2)} &= -\frac{1}{0.0504}[-1, 0]^T[-1, 0] \\ &= \begin{bmatrix} 19.84 & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned} \Delta M_{(1,2)} &= 59.53[-0.033, -0.021]^T[-0.033, -0.021] \\ &= \begin{bmatrix} 0.067 & 0.042 \\ 0.042 & 0.026 \end{bmatrix}. \end{aligned}$$

□

2.2 Measurement Model

Since the voltage phase angles, $\theta[k]$, are assumed to be measured by PMUs, we allow for the scenario where the angles are measured at only a subset of the load buses, and denote

this reduced measurement set by $\hat{\theta}[k]$. Suppose that there are N_d load buses and we select $p \leq N_d$ locations to deploy the PMUs. Then, there are $\binom{N_d}{p}$ possible locations to place the PMUs. Let

$$\tilde{M} = \begin{cases} \tilde{M}_0, \text{ if } 1 \leq k < \gamma_0, \\ \vdots \\ \tilde{M}_{(m,n),T}, \text{ if } k \geq \gamma_T. \end{cases} \tag{2.13}$$

Then, the absence of a PMU at bus i corresponds to removing the i^{th} row of $\tilde{M}$. Thus, let $\hat{M} \in \mathbb{R}^{p \times N_d}$ be the matrix obtained by removing $N - p - 1$ rows from $\tilde{M}$. Therefore, we can relate $\hat{M}$ to $\tilde{M}$ in (2.13) as follows:

$$\hat{M} = C\tilde{M}, \tag{2.14}$$

where $C \in \mathbb{R}^{p \times (N-1)}$ is a matrix of 1's and 0's that appropriately selects the rows of $\tilde{M}$. Accordingly, the increments in the phase angle can be expressed as follows:

$$\Delta\hat{\theta}[k] \approx \hat{M}\Delta P^d[k]. \tag{2.15}$$

The small variations in the real power injections at the load buses, $\Delta P^d[k]$, can be attributed to random fluctuations in electricity consumption. In this regard, we may model the $\Delta P^d[k]$'s as independent and identically distributed (i.i.d.) random vectors. By the Central Limit Theorem [19], it can be shown that each $\Delta P^d[k]$ is a Gaussian vector, i.e., $\Delta P^d[k] \sim \mathcal{N}(0, \Lambda)$, where Λ is the covariance matrix. Note that the elements $\Delta P^d[k]$ need not be independent. Since $\Delta\hat{\theta}[k]$ depends on $\Delta P^d[k]$ through the linear relationship given in (2.15), we have that:

$$\Delta\hat{\theta}[k] \sim \begin{cases} f_0 \coloneqq \mathcal{N}(0, \hat{M}_0\Lambda\hat{M}_0^T), \text{ if } 1 \leq k < \gamma_0, \\ f_{(m,n)}^{(0)} \coloneqq \mathcal{N}(-P_{(m,n)}[\gamma_0 + 1]CM_0 r_{(m,n)}, \\ \qquad\qquad \hat{M}_0\Lambda\hat{M}_0^T), \text{ if } k = \gamma_0, \\ \vdots \\ f_{(m,n)}^{(T)} \coloneqq \mathcal{N}(0, \hat{M}_{(m,n),T}\Lambda\hat{M}_{(m,n),T}^T), \text{ if } k \geq \gamma_T, \end{cases} \tag{2.16}$$

It is important to note that for $\mathcal{N}\left(0, \hat{M}\Lambda\hat{M}^T\right)$ to be a nondegenerate p.d.f., its covariance matrix, $\hat{M}\Lambda\hat{M}^T$, must be full rank. We enforce this by ensuring that the number of PMUs allocated, p, is less than or equal to the number of load buses, N_d.

2.3 Problem Statement

Our goal is to detect the change in the probability distribution of the sequence $\{\Delta\hat{\theta}[k]\}_{k \geq 1}$ (that results from the line outage) as quickly as possible while maintaining a certain level of detection accuracy, which is related to, e.g., the false alarm rate. This problem, referred to as quickest change detection (QCD), has been well studied in the statistical signal processing literature. Next, we provide a precise mathematical description of the QCD problem and an algorithm that we will use to detect a line outage; we refer the reader to [20]-[21] for a survey of the theory of QCD and algorithms.

We assume that the sequence $\{\Delta\hat{\theta}[k]\}_{k\geq 1}$ of random vectors is available from PMU measurements. For the base case where no line outages are present, we have that $\Delta\hat{\theta}[k] \sim f_0$. At some random time, t_f, an outage occurs on line (m,n). Then, the pdf of the sequence $\{\Delta\hat{\theta}[k]\}$ changes from f_0 to $f^{(0)}_{(m,n)}$. Then, the system undergoes a series of transient responses which corresponds to the distribution of $\Delta\hat{\theta}[k]$ evolving from $f^{(0)}_{(m,n)}$ to $f^{(T)}_{(m,n)}$. First, a meanshift takes place during the instant of change t_f, where the pdf is $f^{(0)}_{(m,n)}$. Then, the statistical behavior of the process is characterized by a series of changes only in the covariance matrix of the measurements. The objective is to detect this transition in the pdf of $\{\Delta\hat{\theta}[k]\}$ as quickly as possible. Mathematically, when a line outage occurs, the objective is to find the optimal stopping time τ on the sequence of observations for $\Delta\hat{\theta}$. In the absence of a change, the expectation of τ, $\mathbb{E}[\tau]$, should be maximized so as to avoid false alarms. On the other hand, once a line outage occurs, we expect $\mathbb{E}[\tau]$ to be as small as possible. A formulation that captures this trade-off is as follows [22]:

$$\begin{aligned} &\min_{\tau} \ \sup_{\gamma_0 \geq 1} \mathbb{E}_{\gamma_0}[\tau - \gamma_0 | \tau \geq \gamma_0] \\ &\text{subject to } \ \mathbb{E}_{\infty}[\tau] \geq \beta, \end{aligned} \tag{2.17}$$

where $\mathbb{E}_{\gamma_0}$ denotes the expectation with respect to probability measure when a change occurs at time sample γ_0, $\mathbb{E}_{\infty}$ denotes the corresponding expectation when the change never occurs, and $\beta > 0$ is the given constraint on the mean time to false alarm.

2.4 Summary

This chapter set up the framework for the proposed line outage detection and identification method that is to be developed in this work. An incremental DC-like power flow model was derived along with the statistical model for the voltage phase angles, the measurements of which are assumed to be provided by PMUs. In the derivation of the statistical model, we assumed that the incremental variations in the active power injections at each load bus are independent random variables and that the generators react to the changes in these load demands.

Chapter 3

Line Outage Identification

This chapter begins by introducing the CuSum algorithm for change detection for the case where only the meanshift and the quasi steady state period following a line outage is considered. This is followed by the presentation of the Kullback-Leibler (KL) divergence, which is an important measure characterizing the distance between two probability distributions; the detection algorithms we present are based on this measure. Then, we introduce the Generalized Likelihood Ratio Test (GLRT) algorithm, which serves as a basis for our line-outage detection method. The performance of the proposed algorithm is compared against other existing line outage detection algorithms in the literature. We show that our algorithm performs better as it exploits the statistical properties of the measured voltage phase angles before, during, and after a line outage, whereas other methods in the literature only utilize the change in statistics that occurs at the instant of outage.

3.1 CuSum Algorithm

Recall the measurement model in (2.16) for the case where $T = 1$:

$$\Delta\hat{\theta}[k] \sim \begin{cases} f_0 \coloneqq \mathcal{N}(0, \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } 1 \leq k < \gamma_0, \\ f_{(m,n)}^{(0)} \coloneqq \mathcal{N}(-P_{(m,n)}[\gamma_0 + 1] C M_0 r_{(m,n)}, \\ \qquad\qquad \hat{M}_0 \Lambda \hat{M}_0^T), \text{ if } k = \gamma_0, \\ f_{(m,n)}^{(1)} \coloneqq \mathcal{N}(0, \hat{M}_{(m,n),T} \Lambda \hat{M}_{(m,n),T}^T), \text{ if } k \geq \gamma_1. \end{cases} \tag{3.1}$$

Suppose that the p.d.f.'s f_0, $f_{(m,n)}^{(0)}$, and $f_{(m,n)}^{(1)}$, are known. Then, one particular algorithm that possesses the optimality properties with respect to the formulation in (2.17) is the Cumulative Sum (CuSum) algorithm [23]. From the sequence of phase angle measurements, the CuSum algorithm computes a sequence of statistics recursively so that for $k \geq 0$, the statistic $W_{(m,n)}^{\text{CU}}[k+1]$ is computed as

$$W^{\mathrm{CU}}_{(m,n)}[k+1] = \max\left\{ W^{\mathrm{CU}}_{(m,n)}[k] + \log\frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[k+1])}{f_0(\Delta\hat{\theta}[k+1])}, \log\frac{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k+1])}{f_0(\Delta\hat{\theta}[k+1])}, 0 \right\}, \tag{3.2}$$

where $W^{\mathrm{CU}}_{(m,n)}[0] = 0$ for all $(m,n) \in \mathcal{E}$. Denote τ_C to be the time at which the CuSum algorithm declares a line outage; then,

$$\tau_{\mathrm{C}} = \inf\{k \geq 1 : W^{\mathrm{CU}}_{(m,n)}[k] > A^{\mathrm{CU}}_{(m,n)}\}, \tag{3.3}$$

where $A^{\mathrm{CU}}_{(m,n)}$ is a threshold selected for the corresponding $W^{\mathrm{CU}}_{(m,n)}[k]$ statistic. An optimal method to select this threshold will be presented in Section 3.3.

3.2 Generalized Likelihood Ratio Test Algorithm

In the setting we consider in this work, since the line for which an outage occurs is unknown, the post-change pdf of $\Delta\hat{\theta}$ is also unknown. However, since the single-line outage can occur in at most L ways, the post-change distribution is known to belong to the finite set $\{f_{(m,n)}, (m,n) \in \mathcal{E}\}$. In this context, we can apply the Generalized Likelihood Ratio Test (GLRT) algorithm where we compute L CuSum statistics in parallel, one for each post-change scenario, and declare an outage the first time a change is detected in any one of the parallel CuSum tests. In other words, we compute (3.2) for each line (m,n) in the system, with $W^{\mathrm{CU}}_{(m,n)}[0] = 0$, and stop at

$$\tau^{\mathrm{CU}} = \inf\left\{ k \geq 1 : \max_{(m,n)\in\mathcal{E}} W^{\mathrm{CU}}_{(m,n)}[k] > A^{\mathrm{CU}}_{(m,n)} \right\}. \tag{3.4}$$

In [13], a single threshold $A^{\mathrm{CU}}_{(m,n)}$ was chosen for all line outage streams $W^{\mathrm{CU}}_{(m,n)}$. However, faster detection can be achieved by choosing an individual threshold $A^{\mathrm{CU}}_{(m,n)}$ for each $W^{\mathrm{CU}}_{(m,n)}$ that is proportional to its corresponding KL divergence. The threshold $A^{\mathrm{CU}}_{(m,n)}$ can be chosen to control the mean time to false alarm; if a larger mean time to false alarm is required, then $A^{\mathrm{CU}}_{(m,n)}$ is set to a larger value, and vice-versa. Finally, this algorithm also identifies the line that is outaged at τ^{CU} to be

$$(\hat{m}, \hat{n}) = \arg\max_{(m,n)\in\mathcal{E}} W^{\mathrm{CU}}_{(m,n)}[\tau^{\mathrm{CU}}]. \tag{3.5}$$

3.3 Threshold Selection

We now present ways of choosing the thresholds for the CuSum test. It can be shown (see, e.g., [24]) that by choosing

$$A^{\mathrm{CU}}_{(m,n)} = \log\gamma - \log\beta_{(m,n)}, \tag{3.6}$$

with $\beta_{(m,n)}$ being a positive constant independent of γ, the expected delay for each possible outage differs from the corresponding minimum delay among the class of stopping times $\mathcal{C}_\gamma = \{\tau : \mathbb{E}_0(\tau) \geq \gamma\}$, as $\gamma \to \infty$, by a bounded constant.

A choice of thresholds for the CuSum algorithm is obtained by setting $\beta_{(m,n)} = \frac{1}{L}$ for all $(m,n) \in \mathcal{E}$. This way we get a common threshold, i.e., $A^{\mathrm{CU}}_{(m,n)} = A^{\mathrm{CU}} = \log(\gamma L)$ for all $(m,n) \in \mathcal{E}$. It can be shown (see, e.g., [25]) that by choosing the thresholds this way, we can guarantee that $\mathbb{E}_0[\tau^{\mathrm{CU}}] \geq \gamma$.

Using the results in [24], another choice of the thresholds could be based on a relative performance loss criterion, i.e.,

$$\beta_{(m,n)} = \frac{1}{D(f^{(1)}_{(m,n)} \parallel f_0) L (\zeta_{(m,n)})^2)}, \tag{3.7}$$

where

$$\zeta_{(m,n)} = \lim_{b \to \infty} \mathbb{E}^{(1)}_{(m,n)} \left[e^{\{-(S_{(m,n)}[\tau^{\mathrm{b}}_{(m,n)}] - b)\}} \right], \tag{3.8}$$

with

$$\tau^{\mathrm{b}}_{(m,n)} = \inf\{k \geq 1 : S_{(m,n)}[k] \geq b\}, \tag{3.9}$$

and

$$S_{(m,n)}[k] = \sum_{l=1}^{k} \log \frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[l])}{f_0(\Delta\hat{\theta}[l])}. \tag{3.10}$$

This choice of threshold depends on the asymptotic overshoot of an Sequential Probability Ratio Test (SPRT)-based test, which is often used in hypothesis testing [15]. As we show through case studies, these thresholds result in performance gains.

3.4 Intuition Behind the Operation of the GLRT Algorithm

The algorithm we presented in (3.2) for line outage detection is based on the Kullback-Leibler (KL) divergence, which for any two probability densities f and g is defined as follows:

$$D(f \parallel g) := \int f(x) \log \frac{f(x)}{g(x)} dx \geq 0, \tag{3.11}$$

with equality if and only if $f = g$ almost surely. In the context of the line outage detection problem, for an outage of line (m,n), the KL divergence is

$$D(f^{(1)}_{(m,n)} \parallel f_0) = \mathbb{E}\left[\log \left(\frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])} \right) \middle| (m,n) \text{ outage} \right], \tag{3.12}$$

which provides a bound on the delay for detecting an outage in line (m,n); a larger KL divergence results in lower detection delay and vice versa. Prior to any changes, the mean of the log likelihood ratio is negative due to (3.11). Therefore, $W^{\mathrm{CU}}_{(m,n)}[k]$ would remain close

to or at 0 prior to a line outage. On the other hand, after an outage occurs, the mean of the log-likelihood ratio is positive. As a result, $W^{\text{CU}}_{(m,n)}[k]$ increases unboundedly after the outage in line (m,n), and the CuSum algorithm in (3.2) declares the occurrence of an outage in line (m,n) the first time that $W^{\text{CU}}_{(m,n)}[k]$ reaches $A^{\text{CU}}_{(m,n)}$.

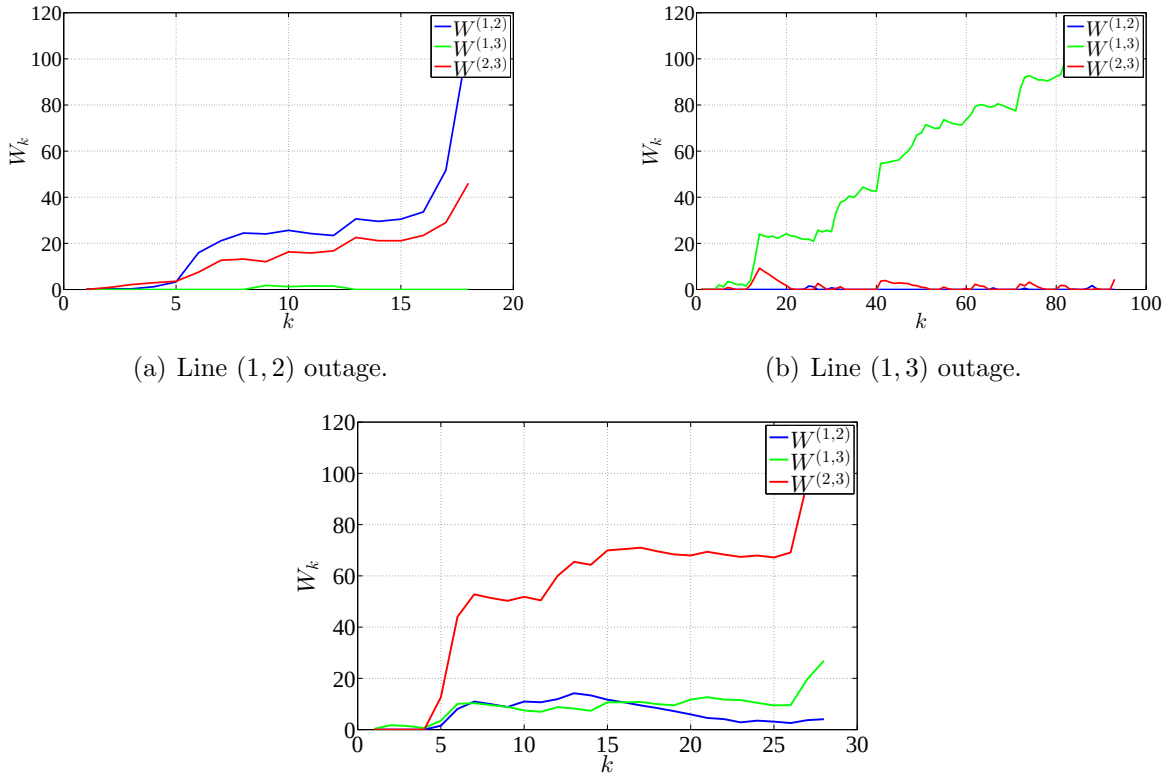
In addition, We can use (3.11) to obtain bounds on the false isolation rates. Consider an outage of line (m,n); if $\mathbb{E}\left[\log\left(\frac{f^{(1)}_{(k,l)}}{f_0}\right)\right]$ is positive, then $W^{\text{CU}}_{(k,l)}[k]$ would increase despite no outage in line (k,l). Hence, we would like $\mathbb{E}\left[\log\left(\frac{f^{(1)}_{(m,n)}}{f^{(1)}_{(k,l)}}\right)\right]$ to be maximized so that the false isolation rate for line (m,n) outage is minimized. For example, we can compute $D(f^{(1)}_{(m,n)} \parallel f^{(1)}_{(k,l)})$ to estimate a bound on the false isolation rate, where a small value indicates that an outage in line (k,l) causes a statistical change in the voltage phase angles that is very similar to that corresponding to an outage in line (m,n). An extreme case of $D(f^{(1)}_{(m,n)} \parallel f^{(1)}_{(k,l)}) = 0$ occurs when lines (k,l) and (m,n) have the same impedance and share the same terminal buses, i.e., $k = m$, $l = n$. In this case, our algorithm cannot distinguish between the occurrence of an outage on either of the two lines. Next, we illustrate these ideas in a small power system example.

Example 2 (Three-Bus System)**.** Consider the 3-bus system shown in Fig. 2.1. We apply the GLRT algorithm to detect and identify a line $(2,3)$ outage at $\gamma_0 = 1$. The PMU measurements are simulated by creating an active power injection time-series data for each load bus i with

$$P_i[k] = P_i[k-1] + v_i, \tag{3.13}$$

where $P_i[0]$ is the nominal power injection at load bus i at instant k, and $v_i \sim \mathcal{N}(0, 0.5)$ is a pseudorandom value representing random fluctuations in electricity consumption. For each set of bus injection data, we solve the nonlinear power flow equations in (2.1) to obtain the sequence of phase angle "measurements" $\{\hat{\theta}[k]\}$. In this example, we assume that bus 1 is the reference bus and the random fluctuations at buses 2 and 3 are uncorrelated, so Λ is a diagonal matrix.

Using the GLRT-based algorithm, we execute three CuSum tests in parallel and compute each $W^{\text{CU}}_{(m,n)}[k]$ defined in (3.2), one for each line of the system. Figure 3.1 shows the typical progressions of $W^{\text{CU}}_{(m,n)}[k]$ for each line outage of the 3-bus system. In Fig. 3.1(a), $W^{\text{CU}}_{(1,2)}[k]$ crosses the threshold of $A = 100$ first, while $W^{\text{CU}}_{(1,3)}[k]$ and $W^{\text{CU}}_{(2,3)}[k]$ remain close to 0. Therefore, the algorithm was able to correctly detect and identify the line outage after $k = 25$ samples. Similar behavior was observed for outages of line (1,3) and (2,3). In addition to the plots, Table 3.1 shows the computed KL divergences of all line outages according to (3.11). For example, a positive $\mathbb{E}\left[\log\left(\frac{f_{(1,2)}}{f_0}\right)\middle|\,(1,2)\text{ outage}\right]$ value of 3.69 means that in the long run, if the outage were indeed on line $(1,2)$, $W^{\text{CU}}_{(1,2)}[k]$ will increase; the rate of this increase depends on the magnitude of the KL divergence. Since $\mathbb{E}\left[\log\left(\frac{f_{(1,2)}}{f_0}\right)\middle|\,(1,2)\text{ outage}\right] > \mathbb{E}\left[\log\left(\frac{f_{(1,3)}}{f_0}\right)\middle|\,(1,3)\text{ outage}\right]$, on average, $W^{\text{CU}}_{(1,2)}[k]$ reaches the threshold A in less samples than $W^{\text{CU}}_{(1,3)}[k]$ for an outage of their respective lines. On the other hand, the negative value of $\mathbb{E}\left[\log\left(\frac{f_{(1,2)}}{f_0}\right)\middle|\,(1,3)\text{ outage}\right]$ means that $W^{\text{CU}}_{(1,2)}[k]$ tends to stay near 0 for a line $(1,3)$ outage,



(a) Line $(1, 2)$ outage.

(b) Line $(1, 3)$ outage.

(c) Line $(2, 3)$ outage.

Figure 3.1: Realizations of $W^{\text{CU}}_{(m,n)}[k]$ for each line outage of 3-bus system with PMU at buses 2 and 3.

which is observed in Fig. 3.1(b). □

3.5 Other Statistical Algorithms for Power System Line Outage Detection

This section introduces some of the other algorithms that are used for change detection. Specifically, we introduce the meanshift and Shewhart tests, which only consider the latest sample of the voltage phase angles instead of using all of the samples. For example, the line outage detection algorithm proposed in [8] can be shown to be equivalent to a log-likelihood ratio test that only uses the most recent measurements.

3.5.1 Meanshift Test

The meanshift test is a "one-shot" detection scheme in that the algorithm uses only the most recent observation to decide whether a change in the mean has occurred and ignores all past observations. The meanshift statistic corresponding to line (m, n) is defined as follows:

Table 3.1: 3-bus system KL Div.

KL Div. \ Line Outage (m,n)	$(1,2)$	$(1,3)$	$(2,3)$
$\mathbb{E}\left[\log\left(\frac{f_{(1,2)}}{f_0}\right) \mid (m,n) \text{ outage}\right]$	3.69	-1.09	0.75
$\mathbb{E}\left[\log\left(\frac{f_{(1,3)}}{f_0}\right) \mid (m,n) \text{ outage}\right]$	-0.86	1.77	1.34
$\mathbb{E}\left[\log\left(\frac{f_{(2,3)}}{f_0}\right) \mid (m,n) \text{ outage}\right]$	0.59	-0.02	6.42

$$W^{\text{MS}}_{(m,n)}[k] = \log \frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}. \tag{3.14}$$

The decision maker declares a change when one of the $|\mathcal{E}|$ statistics crosses a corresponding threshold, $A^{\text{MS}}_{(m,n)}$. The stopping time for this algorithm is defined as:

$$\tau^{\text{MS}} = \inf_{(m,n)\in\mathcal{E}} \left\{ \inf\{k \geq 1 : W^{\text{MS}}_{(m,n)}[k] > A^{\text{MS}}_{(m,n)}\} \right\}. \tag{3.15}$$

The meanshift test ignores the persistent covariance change that occurs after the outage. In particular, note that the meanshift test is using the likelihood ratio between the distribution of the observations before and at the changepoint. More specifically, assuming that an outage occurs in line (m,n), the expected value of the statistic at the changepoint is given by

$$\mathbb{E}^{(0)}_{(m,n)}\left[\log \frac{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] = D(f^{(0)}_{(m,n)} \parallel f_0) > 0. \tag{3.16}$$

On the other hand, after the changepoint ($k > \lambda_0$), the expected value of the statistic is given by

$$\begin{aligned} \mathbb{E}^{(1)}_{(m,n)}\left[\log \frac{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right] &= \\ &= D(f^{(1)}_{(m,n)} \parallel f_0) - D(f^{(1)}_{(m,n)} \parallel f^{(0)}_{(m,n)}), \end{aligned} \tag{3.17}$$

which could be either positive or negative.

3.5.2 Shewhart Test

Similar to the meanshift test, the Shewhart test is also a "one-shot" detection scheme. This test attempts to detect a change on the observation sequence through the meanshift and the change in the covariance matrix of $\Delta\hat{\theta}[k]$. The Shewhart test statistic for line (m,n) outage is defined as:

$$W^{\mathrm{SH}}_{(m,n)}[k] = \max\left\{\log\frac{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}, \log\frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right\}, \tag{3.18}$$

where the first log-likelihood ratio is used to detect the meanshift, while the second log-likelihood ratio is used to detect the persistent change in the covariance. The stopping time is:

$$\tau^{\mathrm{SH}} = \inf_{(m,n)\in\mathcal{E}}\left\{\inf\{k \geq 1 : W^{\mathrm{SH}}_{(m,n)}[k] > A^{\mathrm{SH}}_{(m,n)}\}\right\}. \tag{3.19}$$

Since the Shewhart test exploits the covariance change in addition to the meanshift statistic, it should perform better than the meanshift test, at least as the meantime to false alarm goes to infinity, which is verified in the case studies.

3.6 Case Studies

This section provides a case study of the concepts introduced in this chapter on the IEEE 14-bus and 118-bus systems [26]. The importance sampling technique for rare events is also presented. An outage is simulated and the proposed algorithm is used to detect this outage. In addition, we demonstrate the effectiveness of our proposed line outage detection algorithm by comparing against other line outage detection algorithms on the IEEE 14-bus system.

3.6.1 Importance Sampling

Since the meanshift in the voltage phase angles occurs between the sample immediately prior to and after the line outage, it is not persistent. If the meanshift test presented in Section 3.5.1 does not capture the outage exactly when it occurs, then the likelihood of correctly identifying the outage would be a rare event. This is because the log-likelihood ratio used in the meanshift statistic of (3.14) matches only the meanshift but not the covariance shift that is persistent after a line outage. Therefore, in order to simulate detection delays of the meanshift test, the technique of importance sampling is used.

The usual Monte Carlo method of estimating the average detection delay of the meanshift test is

$$\hat{\tau}^{\mathrm{MS}} := \mathbb{E}[\tau^{\mathrm{MS}}] \approx \frac{1}{N}\sum_{i=1}^{N}\tau^{\mathrm{MS}}_i, \tag{3.20}$$

where N is a large sample size and τ^{MS}_i is the detection delay for the i^{th} sample run. For line (m,n) outage simulation, starting with the second sample after the line outage, the voltage phase angles samples are generated from the probability distribution $f^{(1)}_{(m,n)}$. Therefore, the numerator of (3.14) does not match the distribution from which the samples are generated, making the crossing of threshold rare. In order to use importance sampling, we sample from the distribution $f^{(0)}_{(m,n)}$ instead of $f^{(1)}_{(m,n)}$ for all samples after the outage. We use the fact that

$$\mathbb{E}_{f^{(1)}_{(m,n)}}\left[\log\frac{f^{(0)}_{(m,n)}}{f_0}\right] = \mathbb{E}_{f^{(0)}_{(m,n)}}\left[\log\left(\frac{f^{(0)}_{(m,n)}}{f_0}\right)\frac{f^{(1)}_{(m,n)}}{f^{(0)}_{(m,n)}}\right]. \tag{3.21}$$

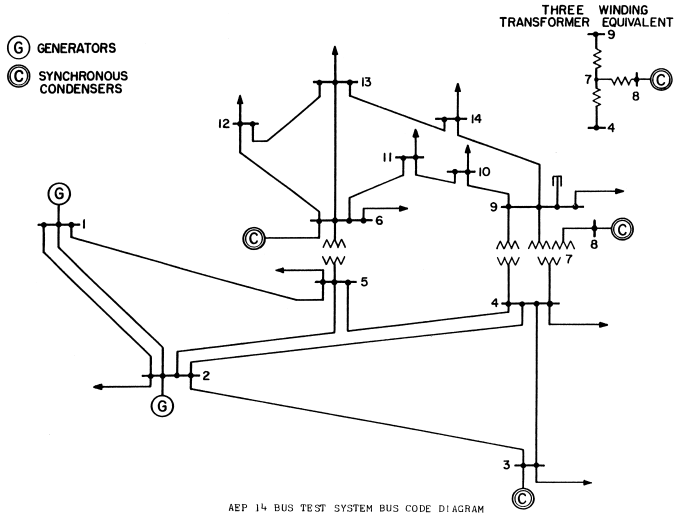


Figure 3.2: Network topology for 14-bus system.

We modify the meanshift statistic of (3.14) to

$$W^{\mathrm{MS}}_{(m,n)}[k] = \log\left(\frac{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k])}{f_0(\Delta\hat{\theta}[k])}\right)\frac{f^{(1)}_{(m,n)}(\Delta\hat{\theta}[k])}{f^{(0)}_{(m,n)}(\Delta\hat{\theta}[k])} \tag{3.22}$$

but with sampling of $\Delta\hat{\theta}[k]$ from the distribution $f^{(0)}_{(m,n)}$ after the line outage instead of $f^{(1)}_{(m,n)}$. We declare the detection of line outage when the statistic crosses the threshold. This method greatly reduces the number of sample paths that must be simulated to estimate the detection delay, resulting in greater efficiency.

3.6.2 14-Bus System

The one-line diagram of the IEEE 14-bus test system is shown in Fig. 3.2. We simulate a line $(2,5)$ outage at $k = 10$ and apply the CuSum tests of (3.2) to the voltage phase angle measurements. The random fluctuations in the active power injections are considered independent Gaussian random variables with a variance of 0.03. Therefore, Λ is a diagonal matrix. Figure 3.3 shows that the $W_{(2,5)}[k]$ stream crosses the threshold of 100 before all the other streams at $k = 48$, resulting in a detection delay of 38 samples. Assuming that the PMUs sample voltage angles at a rate of 30 samples per second, the detection delay in this case is about 2 seconds. Again, we see that the streams for the other lines either remain close to 0 or grow at a slower rate than $W_{(2,5)}[k]$.

Next, we perform Monte Carlo simulations for the Shewhart, meanshift, and CuSum algorithms to obtain plots of average detection delay versus mean time to false alarm. The values for the average detection delay are obtained by simulating an outage in line $(4,5)$ and running the corresponding detection algorithms for different thresholds until a detection of the outaged line is declared. For computing the mean time to false alarm, the detection algorithms are executed for the power system under normal operation until a false alarm occurs. Since false alarm events are in general rare, averaging many sample runs would

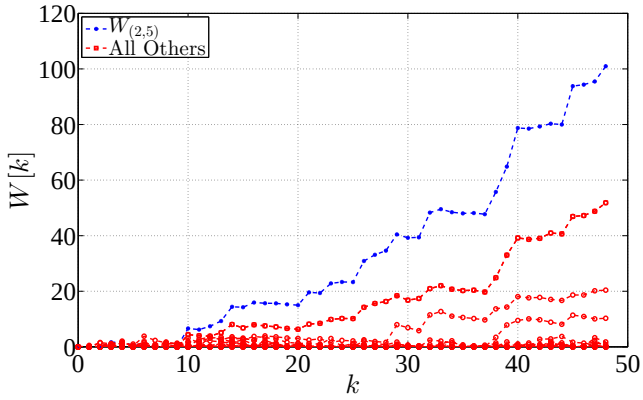


Figure 3.3: Sample run for 14-bus system.

incur significant computation time. In order to reduce the simulation time, importance sampling is used for the meanshift and Shewhart tests. For our simulations, we found that the error bounds for all the simulated values are within 5% of the means. Figure 3.4 shows the average detection delay versus mean time to false alarm for all of the detection methods mentioned in this paper. As evidenced by the crossing of the Shewhart and meanshift plots, for small values of mean time to false alarm, the meanshift test performs better than the Shewhart test. It can be verified from QCD theory that the slope of Delay versus log(mean time to false alarm) for the Shewhart and CuSum tests is given by $\frac{1}{D(f^{(1)}_{(m,n)} \parallel f_0)}$ for large mean time to false alarm [20].

From the plots, we conclude that for the same value of mean time to false alarm, both CuSum-based algorithms have a much lower average detection delay compared to the Shewhart and meanshift algorithms. In addition, the figure shows that when we use varied thresholds for the CuSum test as opposed to a fixed threshold, even lower detection delay can be achieved for the same mean time to false alarm. This illustrates that our algorithm is an improvement over that of [13]. Lastly, simulation results demonstrate that the detection delay scales exponentially with the selected thresholds for both the meanshift and Shewhart tests, and linearly for the CuSum-based tests.

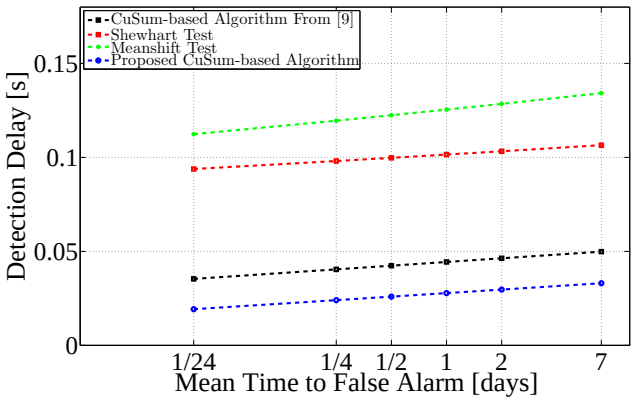


Figure 3.4: Detection delay vs. mean time to false alarm.

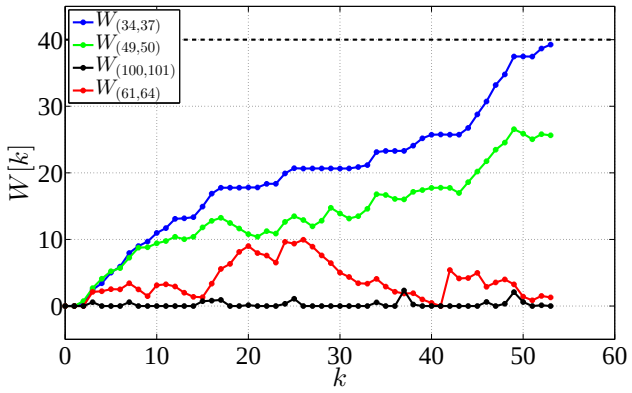


Figure 3.5: Sample run of 118-bus system.

3.6.3 118-Bus System

Next, we illustrate the scalability of the proposed line outage detection algorithm on the IEEE 118-bus test system. The simulation tool MATPOWER [27] is used to compute the voltage angles by repeatedly solving AC power flow solutions of the system. The real power injection is generated using (3.13) with $\sigma = 0.03$. We also assume these random fluctuations are uncorrelated; thus, Λ is a diagonal matrix.

An outage in line $(34, 37)$ starting at $k = 1$ is simulated and the algorithm of (3.2) is applied. Some sample test statistics are shown in Fig. 3.5. With a threshold of 40, the line outage is declared 53 samples after the outage.

Chapter 4

Conclusions

This work provided a framework for line outage detection in power systems, which is crucial for maintaining operational reliability. Many of the current methods for online power system monitoring rely on a system model that can be inaccurate due to bad historical or telemetry data. These inaccuracies were a major factor in many blackouts. Therefore, there is a significant need for developing online techniques to detect and identify system topological changes. The algorithm proposed in this work exploits fast measurements provided by PMUs and uses a statistical method to quickly detect network topological changes. The results of the proposed method are compared against the other line outage detection algorithms in literature.

There are several extensions to this current work that are left for future work. The current method is not capable of detecting double line outage; new techniques that allow for quick detection of double line outages would be beneficial. Additionally, the algorithms presented in this work could be applied to other event detection problems in power systems such as switching of capacitors and transformers.

References

[1] FERC and NERC. (2012, Apr.) Arizona-southern California outages on September 8, 2011: Causes and recommendations. [Online]. Available: http://www.ferc.gov.

[2] U.S.-Canada Power System Outage Task Force. (2004, Apr.) Final report on the August 14th blackout in the United States and Canada: causes and recommendations. [Online]. Available: http://energy.gov.

[3] K. A. Clements and P. W. Davis, "Detection and identification of topology errors in electric power systems," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1748–1753, 1988.

[4] F. F. Wu and W. E. Liu, "Detection of topology errors by state estimation," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 176–183, 1989.

[5] L. Zhao and W. Song, "Distributed power-line outage detection based on wide area measurement system," *Sensors*, vol. 14, no. 7, p. 13114, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/7/13114

[6] N. Singh and H. Glavitsch, "Detection and identification of topological errors in online power system analysis," *IEEE Transactions on Power Systems*, vol. 6, no. 1, pp. 324–331, 1991.

[7] F. Alvarado, "Determination of external system topology errors," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 11, pp. 4553–4561, 1981.

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

[9] H. Zhu and G. B. Giannakis, "Sparse overcomplete representations for efficient identification of power line outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, 2012.

[10] H. Sehwail and I. Dobson, "Locating line outages in a specific area of a power system with synchrophasors," in *North American Power Symposium*, Sept 2012, pp. 1–6.

[11] R. Emami and A. Abur, "External system line outage identification using phasor measurement units," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1035–1040, 2013.

[12] M. Garcia, T. Catanach, S. V. Wiel, R. Bent, and E. Lawrence, "Line outage localization using phasor measurement data in transient state," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3019–3027, July 2016.

[13] Y. C. Chen, T. Banerjee, A. D. Domínguez-García, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, 2016.

[14] T. Banerjee, Y. C. Chen, A. D. Domínguez-García, and V. V. Veeravalli, "Power system line outage detection and identification—a quickest change detection approach," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2014.

[15] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, 2009.

[16] A. R. Bergen and V. Vittal, *Power Systems Analysis*. Prentice Hall, 2000.

[17] M. Lotfalian, R. Schlueter, D. Idizior, P. Rusche, S. Tedeschi, L. Shu, and A. Yazdankhah, "Inertial, governor, and AGC/economic dispatch load flow simulations of loss of generation contingencies," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 11, pp. 3020 –3028, Nov. 1985.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[19] B. Hajek, *Random Processes for Engineers*. Cambridge University Press, 2015.

[20] V. V. Veeravalli and T. Banerjee, *Quickest Change Detection*. Elsevier: E-reference Signal Processing, 2013.

[21] A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Change-Point Detection*, ser. Statistics. CRC Press, 2014.

[22] M. Pollak, "Optimal detection of a change in distribution," *Ann. Statist.*, vol. 13, no. 1, pp. 206–227, Mar. 1985.

[23] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[24] G. Fellouris and G. Sokolov. (2014, October) Second-order asymptotic optimality in multisensor sequential change detection. http://arxiv.org/abs/1410.3815v2. [Online]. Available: http://arxiv.org/abs/1410.3815v2

[25] A. G. Tartakovsky and A. S. Polunchenko, "Quickest changepoint detection in distributed multisensor systems under unknown parameters," in *Proc. of IEEE International Conference on Information Fusion*, Jul. 2008, pp. 1–8.

[26] "Power system test case archive," Oct. 2012. [Online]. Available: http://www.ee.washington.edu/research/pstca.

[27] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

Part II

PMU Placement via Multinomial Logistic Regression

Stephen J. Wright

Taedong Kim, Graduate Student

University of Wisconsin–Madison

For information about this part of the project report, contact:

Stephen J. Wright
Computer Sciences Department
University of Wisconsin
1210 West Dayton Street
Madison, WI 53706
Phone: (608) 262-4838
Email: swright@cs.wisc.edu

Power Systems Engineering Research Center

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

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, phasor measurement units (PMUs) have been introduced as a way to monitor power system networks. Unlike the more conventional Supervisory Control and Data Acquisitions (SCADA) system, whose measurements include active and reactive power and voltage magnitude, PMUs can provide accurate, high-sampling-rate, synchronized measurements of voltage phasor. There has been much recent research on how the real-time measurement information gathered from PMUs can be exploited in many areas of power system studies, including system control and state estimation.

In this work, we study the use of PMU data in detecting topological network changes caused by single-line outages, and propose techniques for determining optimal placement of a limited number of PMU devices in a grid, so as to maximize the capability for detecting such outages. Our PMU placement approach can also be used as a tie-breaker for the other types of strategies that have multiple optimal solutions (for example, maximum observability problems). It can also be used to enhance a PMU placement scheme that satisfies a property such as topological observability, by identifying PMU locations that can be added to the basic placement in a way that optimizes outage detection capabilities.

We note that modern power grids contain intelligent electronic devices (IEDs) that can identify outages directly. In this context, PMUs can provide a backup detection capability that is based on an alternative set of observations of the state of the grid.

Knowledge of topological changes as a result of line failure (outage) can be critical in deciding how to respond to a blackout. Rapid detection of such changes can enable actions to be taken that reduce risks of cascading failures that lead to large-scale blackouts. One of the main causes of the catastrophic Northeast blackout of 2003 was faulty topological information about the grid following the initial failures; see [1].

Numerous approaches have been proposed for identifying line outages using PMU device measurements. In [2, 3], phasor angle changes are measured and compared with expected phasor angle variations for all single- or double-line outage scenarios. Support vector machines (SVM) were proposed for identification of single-line outages in [4]. A compressed-sensing approach was applied to DC power balance equations to find sparse topological changes in [5], while a cross-entropy optimization technique was considered in [6]. Since the approaches in [5] or [6] use the linearized DC power flow models to represent a power system, their line outage identification strategies rely only on changes to phase angles, ignoring the voltage magnitude measurements from PMUs. Our use of the AC power flow model

allow phase angles and voltage magnitudes to be used in outage detection, leading to more complete exploitation of the available data. In [7], a distributed framework is proposed to avoid privacy issues which can be caused by sharing raw measurement data. In the model, each phasor data concentrator (PDC) performs line outage identification locally, then the processed data (instead of the raw measurement data) is collected at a central system to identify line outage in the entire system. The alternating direction method of multipliers is used to resolve the solution mismatches at shared resources between PDCs.

The key feature that makes line outage identification possible is that voltage phasor measurements reported by PMUs are different for different line-outage scenarios. Our approach aims to distinguish between these different "signatures" by using a multinomial logistic regression (MLR) model [8]. The model can be trained by a convex optimization approach, using standard techniques. The coefficients learned during training can be applied during grid operation to detect outage scenarios. Our approach could in principle be applied to multiline outages too. One practical way to extend our approach to multiline outages without exploding the number of scenarios is to include only those multiline outages that could plausibly occur, that is, sets of lines that are geographically close or are strongly coupled and that could thus fail simultaneously. We note too that our single-line outage classifier could be useful in multiline outage situations when the coupling between the lines is weak (as discussed in [9, Section 2.2]). In other words, some multiline outage cases can be decomposed into single-line outage events on different parts of the grid.

Our approach has several advantages over the other methods that have been proposed previously. First, training of the MLR model, which is the most compute-intensive part of the approach, can be done *offline*; it is not a time-critical task. The *online* portion of the line outage identification process requires only a number of vector-vector inner products (one for each possible outage), so it can be done in real-time. This situation contrasts with alternative approaches for which the "online" computational requirements can be quite heavy. Second, the use of a fairly standard data analysis approach—the MLR classifier—allows integration of full modeling and engineering information, in the form of AC power flow modeling, measurements of phase angle *and* voltage magnitude changes, knowledge of changes to the voltage phasors under a variety of demand conditions. The generic nature of the MLR framework makes it flexible and extendable to detection of other kinds of fault conditions, provided that these conditions yield distinctive signatures. Moreover, the model can be retrained and tuned to changes in demand conditions at different times of the day, or for unusual demand situations.

Because of the expense of installation and maintenance, PMUs are usually installed on just a subset of buses in a grid. We therefore need to formulate an *optimal placement problem* to determine the choice of PMU locations that gives the best information about system state. Several different criteria have been proposed to measure quality of a given choice of PMU locations. We mention below the approaches that use line outage identification as a criterion for PMU placement.

In [10], the authors use pre-computed phase angles as outage signatures and attempt to find the optimal PMU locations by identifying a projection (by setting to zero the entries which are not selected as PMU locations) that maximizes the minimum distance in the ℓ_p-norm of the projected signatures. The problem is formulated as an integer program (IP) and a greedy heuristic and branch-and-bound approaches are proposed. These algo-

rithms are used again in [11] to seek optimal PMU location for line outage identification with uncertain states. In the model, the power injections are considered as network states and provided as a form of prior distribution. Three metrics are proposed to evaluate the PMU locations and the effect of prior accuracy for line outage identification is also studied. PMU placement for the line outage identification method discussed in [5] is studied in [9]. A non-convex mixed-integer nonlinear program (MINLP) is formulated, leading to a linear programming convex relaxation. Again, a greedy heuristic and a branch-and-bound algorithm are suggested as a solution methodology. In [12], the PMU placement is optimized to maximize the average identification capability (AIC) of multiple line outages. The AIC is defined using the dissimilarity distance, which is inspired by Kullback-Leibler distance, between the probability densities of pre-outage and post-outage measurements from PMU locations. Exhaustive search optimal method is proposed as a solution methodology.

We build our optimal PMU placement formulation on our MLR model for single-line outage detection, by adding nonsmooth "Group LASSO" regularizers to the MLR objective and "greedy" heuristics. Our approach is flexible, and can also be used to decide additional PMU locations to maximize detection performance for given existing PMU locations. Moreover, we aim directly to optimize detection performance in our choice of PMU locations, not a surrogate objective such as observability.

The rest of this document is organized as follows: In Chapter 2, the line outage identification problem is described along with the multinomial logistic regression (MLR) formulation. The problem of PMU placement to identify a line outage is described in Chapter 3: We describe a group-sparse heuristic and its greedy variant used to formulate and solve this problem. Numerical results on synthetically generated data are presented in Chapter 4 and a conclusion appears in Chapter 5.

Chapter 2

Line Outage Identification

We describe an approach that uses changes in voltage phasor measurements at PMUs to detect single-line outages in the power grid. As in [2, 5], we assume that the fast dynamics of the system are well damped and voltage measurements reflect the quasi-static equilibrium that is reached after the disruption. We use a quasi-steady state AC power flow model (see, for example, [13, Chapter 10]) as a mapping from time varying load variation (and line outage events) to the polar coordinate "outputs" of voltage magnitude and angle.

PMUs report phasor measurements with high frequency, and changes in voltage due to topology changes of the power grid tend to be larger than the variation of voltage phasor during normal operation (for example, demand fluctuation that occurs during the sampling time period). We construct signature vectors from these voltage changes, under the various single-line outage situations, and use them to train a classifier for line outage identification. We now describe the multinomial logistic regression model for determining the outage scenario.

2.1 Multinomial Logistic Regression Model

Multinomial logistic regression (MLR) is a machine-learning approach for classifying a vector of features as belonging to one of several possible classes. A set of functions is constructed from the feature vector, each function giving the odds of a feature vector belonging to a particular class. The parameters that define these functions are obtained from a "training" process, in which a large collection of feature vectors are presented along with their *known* classes. The parameters are adjusted so that when presented with a new feature vector that is similar to several others that have been encountered in training, the MLR classifier will output a high probability that the new vector has the same class as the training vectors that are similar to it.

MLR generalizes the well-known two-class logistic regression procedure (see [8]) to multiple classes. Each class has its own vector of parameters, which is applied to the feature vector to determine the odds of the feature vector belonging to that class. In our application, the feature vectors consist of voltage phasor changes at the PMUs, differences of the complex voltage measurements before and after an outage event.

The training process consists of choosing the parameters that define the function so as

to maximize an a posteriori likelihood function. Suppose that there are K possible outcomes (classes) labeled as $k \in \{1, 2, \ldots, K\}$, and let X denote a feature vector of length n. In the MLR model, the probability that X belongs to class Y (which is one of the K possible classes) is given by the following formula:

$$\Pr(Y = k|X) := \frac{e^{\langle \beta_k, X \rangle}}{\sum_{i=1}^{K} e^{\langle \beta_i, X \rangle}} \text{ for } k = 1, 2, \ldots, K, \tag{2.1}$$

where $\langle \cdot, \cdot \rangle$ is the inner product of vectors and $\beta_1, \beta_2, \ldots, \beta_K$ are regression coefficients—the parameter vectors that define each odds function. Once values of the coefficients β_k, $k \in \{1, 2, \ldots, K\}$ have been obtained from the training process, we can predict the outcome associated with a given feature vector X by evaluating

$$k^* \in \underset{k \in \{1,2,\ldots,K\}}{\arg\max} \Pr(Y = k|X), \tag{2.2}$$

or equivalently,

$$k^* \in \underset{k \in \{1,2,\ldots,K\}}{\arg\max} \langle \beta_k, X \rangle . \tag{2.3}$$

From the engineering perspective, a more useful form of output would be to report all outcomes k that are assigned by the model a probability above a certain threshold (say .05 or .1) and/or to report the top-2 or top-3 most likely outcomes. The results described in Chapter 4 are based on this more nuanced interpretation of the output.

Training of the regression coefficients $\beta_1, \beta_2, \ldots, \beta_K$ can be performed by maximum likelihood estimation. The training data consists of M pairs $(X_1, Y_1), (X_2, Y_2), \ldots, (X_M, Y_M)$, each consisting of a feature vector X_m and its corresponding outcome Y_m for a line outage scenario, that is, X_m represents the shift of voltage phasors when line Y_m is removed from the system for $m = 1, 2, \cdots, M$. Note that $M \geq K$ since we need at least one training sample for each line outage scenario. Given formula (2.1), a posteriori likelihood of observing $Y_1, Y_2, \ldots, Y_M$ given the events $X_1, X_2, \ldots, X_M$ is

$$\prod_{m=1}^{M} P(Y = Y_m|X_m) = \prod_{m=1}^{M} \left(\frac{e^{\langle \beta_{Y_m}, X_m \rangle}}{\sum_{k=1}^{K} e^{\langle \beta_k, X_m \rangle}} \right). \tag{2.4}$$

By taking log of (2.4), we have log-likelihood function

$$f(\beta) := \sum_{m=1}^{M} \left(\langle \beta_{Y_m}, X_m \rangle - \log \sum_{k=1}^{K} e^{\langle \beta_k, X_m \rangle} \right), \tag{2.5}$$

where the matrix β (of dimensions $n \times K$) is obtained by arranging the coefficient vectors as $\begin{bmatrix} \beta_1 & \beta_2 & \ldots & \beta_K \end{bmatrix}$. Note that $f(\beta) \leq 0$ for all β, since the value of (2.4) is in $(0, 1]$. The maximum likelihood estimate β^* of regression coefficients is obtained by solving the following optimization problem:

$$\beta^* = \underset{\beta}{\arg\max}\ f(\beta). \tag{2.6}$$

This is a smooth convex problem [8, Section 4.3.4] that can be solved by fairly standard techniques for smooth nonlinear optimization, such as limited-memory BFGS (L-BFGS) algorithm [14].

If the training data is separable (that is, there exists β such that $\langle \beta_k, X_m \rangle \geq 0$ for $Y_m = k$ and $\langle \beta_k, X_m \rangle \leq 0$ for $Y_m \neq k$ for all $k = 1, 2, \cdots, K$), the value of $f(\beta)$ can be made to approach zero arbitrarily closely by multiplying β by a large positive value (see [15]). To maintain finiteness of the recovered solution β, we can solve instead the following regularized form of (2.6):

$$\beta^* = \arg\max_{\beta} \; f(\beta) - \tau w(\beta) \tag{2.7}$$

where $\tau > 0$ is a penalty parameter and the penalty function $w(\beta)$ is typically convex and nonsmooth. The penalized form can also be used to promote structure in the solution β^*, such as sparsity or group-sparsity. This property is key to our PMU placement formulation, and we discuss it further in Chapter 3.

Training of the MLR model, via solution of (2.6) or (2.7), can be done offline, as described in the next section. Once the model is trained (that is, the coefficients β_k, $k = 1, 2, \ldots, K$ have been calculated), classification can be done via (2.3), at the cost of multiplying the matrix β by the observed feature vector X, an operation that can be done in real time.

2.2 Training Data: Observation Vectors and Outcomes

In our MLR model for line outage identification problems, the observation vector X_m is constructed from the change of voltage phasor at each bus, under a particular outage scenario. The corresponding outcome is the index of the failed line.

Suppose that a power system consists of N buses, all equipped with PMUs that report the voltage values periodically. Let (V_i, θ_i) and (V_i', θ_i'), $i = 1, 2, \ldots, N$, be two phasor measurements obtained from PMU devices, one taken before a possible failure scenario and one afterward. The observation vector X which describes the voltage phasor difference is defined to be

$$X = \begin{bmatrix} \Delta V_1 & \ldots & \Delta V_N & \Delta\theta_1 & \ldots & \Delta\theta_N \end{bmatrix}^T \tag{2.8}$$

where $\Delta V_i = V_i' - V_i$ and $\Delta\theta_i = \theta_i' - \theta_i$, for $i = 1, 2, \ldots, N$. If we assume that the measurement interval is small enough that loads and demands on the grid do not change significantly between measurements, we would expect the entries of X to be small, unless an outage scenario (leading to a topological change to the grid) occurred. Some such outages would lead to failure of the grid. More often, feasible operation can continue, but with significant changes in the voltage phasors, indicated by large components of X.

The training data (X_m, Y_m) can be assembled by considering a variety of realistic demand scenarios for the grid, solving the AC power flow equations for each possible outage scenario (setting the value of Y_m according to the index of that failure), then setting X_m to be the shift in voltage phasor that corresponds to that scenario. The phasor shifts for a particular scenario change somewhat as the pattern of loads and generations changes, so it is important to train the model using a sample of phasor changes under different realistic patterns of supply and demand.

The observation vector can be extended to include additional information beyond the voltage phasor information from the PMUs, if such information can be gathered easily and exploited to improve the performance of the MLR approach. For example, the system operator may be able to monitor the power generation level G (expressed as a fraction of the long-term average generation) that is injected to the system at the same time points at which the voltage phasor measurements are reported. If included in the observation vector, this quantity might need to be scaled so that it does not dominate the phasor difference information. Also, a constant entry can be added to the observation vector to allow for an "intercept" term (as is usually done in all forms of regression). The extended observation vector thus has the form

$$\overline{X} = \begin{bmatrix} \Delta V_1 & \dots & \Delta V_N & \Delta\theta_1 & \dots & \Delta\theta_N & \rho G & \rho \end{bmatrix}^T \tag{2.9}$$

where ρ is a scaling factor that approximately balances the magnitudes of all entries in the vector. (Note that since G is not too far from 1, it is appropriate to use the same scaling factor for the last two terms.) The numerical experiments in Chapter 4 make use of this extended observation vector.

Chapter 3

PMU Placement

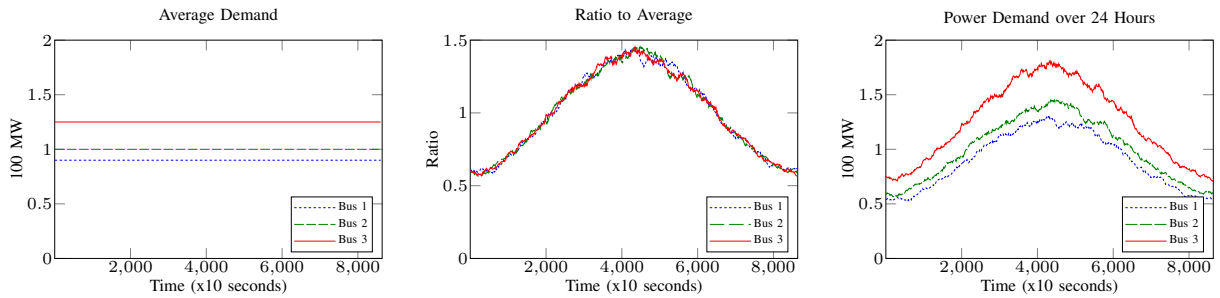
As we mentioned in Chapter 1, installation of PMUs at all buses is impractical. Indeed, if it were possible to do so, single-line outage detection would become a trivial problem, as each outage could be observed directly by PMU measurements of line current flows in or out of a bus; there would be no need to use the "indirect" evidence provided by voltage phasor changes. In this Chapter we address the problem of placing a limited number of PMUs around the grid, with the locations chosen in a fashion that maximizes the system's ability to detect single-line outages. This PMU placement problem selects a subset of *buses* for PMU placement, and assumes that PMUs are placed to monitor voltage phasors at the selected buses.

A naive approach is simply to declare a "budget" of the number of buses at which PMU placement can take place, and consider all possible choices that satisfy this budget. This approach is of course computationally intractable, except when the budget is very small (at most two or three locations). Another approach is a mixed-integer nonlinear programming formulation [10, 9], but this formulation is also quite difficult to solve since its computational complexity is generally quite high. In this work, we use a regularizer function $w(\beta)$ in (2.7) to promote a particular kind of sparsity structure in the coefficient matrix β. Specifically, a *group ℓ_1-regularizer* (GroupLASSO [16]) is used to impose a common sparsity pattern on the rows of the coefficient matrix β, allowing nonzero in β to occur only those rows corresponding to the voltage magnitude and phase angle changes at a particular subset of buses. The numerical results show that approaches based on this regularizer give good results in selection of PMU placements.

3.1 Group-Sparse Heuristic

Let $\mathcal{P}$ be the set of indices in the vector of features $X \in \mathbb{R}^n$, that is $\mathcal{P} = \{1, 2, \ldots, n\}$ where $n = 2N$ (or $n = 2(N+1)$ for $\overline{X}$. See (2.8) and (2.9).) Consider S mutually disjoint subsets of $\mathcal{P}$, denoted by $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_S$. For each $s \in \mathcal{S} := \{1, 2, \ldots, S\}$, define $q_s([\beta]_{\mathcal{P}_s})$ as follows:

$$q_s([\beta]_{\mathcal{P}_s}) = \|[\beta]_{\mathcal{P}_s}\|_F \,, \tag{3.1}$$



(a) Base Case Demand (Considered as Average) (b) The Ratio Generated by Stochastic Process (c) Generated 24-Hour Demand Profile

Figure 3.1: Generating Synthetic Demand Data by A Stochastic Process

where $[\beta]_{\mathcal{P}_s}$ is the submatrix of β constructed by choosing the rows whose indices are in $\mathcal{P}_s$, $\|\cdot\|_F$ is the Frobenius norm[1], and β_{ik} is the (i,k) entry of matrix β. (Thus, β_{ik} is the ith entry of the coefficient vector β_k in (2.3) and (2.4).) The value of $q_s([\beta]_{\mathcal{P}_s})$ is the ℓ_2-norm over the entries of matrix β which are involved in group s. For our observation vectors X (2.8) and $\overline{X}$ (2.9), we can choose the number of groups $|\mathcal{S}|$ equal to the number of buses N, and set

$$\mathcal{P}_s = \{s, s+N\}, \quad s = 1, 2, \ldots, N. \tag{3.2}$$

Thus, if bus s is "selected" in the placement problem, the components of β that are associated with phasor changes (ΔV_s and $\Delta\theta_s$) on bus s are allowed to be nonzeros. Buses that are *not* selected need not be instrumented with PMUs, because the components of β that correspond to these buses are all zero. Note that for the extended vector $\overline{X}$, we do not place the last two entries (the constant and the total generation quantities) into any group, as we assume that these are always "selected" for use in the classification process.

For any subset $\mathcal{R}$ of $\mathcal{S}$, we define a group-ℓ_1-regularizer $w_{\mathcal{R}}(\beta)$ to be the sum of $q_s([\beta]_{\mathcal{P}_s})$ for $s \in \mathcal{R}$, that is,

$$w_{\mathcal{R}}(\beta) = \sum_{s\in\mathcal{R}} q_s([\beta]_{\mathcal{P}_s}). \tag{3.3}$$

Setting $\mathcal{R} = \mathcal{S}$, the penalized form (2.7) with $w = w_{\mathcal{S}}$ can be solved to identify a group-sparse solution:

$$\max_{\beta} \; f(\beta) - \tau w_{\mathcal{S}}(\beta). \tag{3.4}$$

With an appropriate choice of the parameter τ, the solution β^* of (3.4) will be group-row-sparse, that is, the set $\{s \in \mathcal{S} \,|\, q_s([\beta^*]_{\mathcal{P}_s}) \neq 0\}$ will have significantly fewer than S elements. Given a solution β^* of (3.4) for some value of τ, we could define the r-sparse solution as follows (for a given value of r, and assuming that the solution of (3.4) has at least r nonzero values of $q_s([\beta^*]_{\mathcal{P}_s})$):

$$\mathcal{R}^* := \operatorname*{arg\,max}_{\mathcal{R}:|\mathcal{R}|=r, \mathcal{R}\subset\mathcal{S}} \; w_{\mathcal{R}}(\beta^*). \tag{3.5}$$

Since the minimizer β^* of (3.4) is biased due to the presence of the penalty term, we should not use the submatrix extracted from β^* according to the selected group $\mathcal{R}^*$ as the regression

[1]The Frobenius norm of a matrix A is the square root of the sum of squares of the elements of A.

Algorithm 1 Group-Sparse Heuristic

Input:

Parameter τ, r.

Output:

$\mathcal{R}^r$: Set of groups after selecting r groups.

$\tilde{\beta}^r$: Maximum likelihood estimate for r-group observation.

1: $\beta^* \leftarrow \arg\max_\beta \ f(\beta) - \tau w_{\mathcal{S}}(\beta)$

2: $\mathcal{R}^r \leftarrow \arg\max_{\mathcal{R}:|\mathcal{R}|=r,\mathcal{R}\subseteq\mathcal{S}} \sum_{s\in\mathcal{S}} q_s(\beta^*)$

3: Solve (3.6) with $\mathcal{R}^* = \mathcal{R}^r$ to obtain $\tilde{\beta}^r$. (▷) *debiasing*

coefficients for purposes of multinomial classification. Rather, we should solve a reduced, unpenalized version of the problem in which the coefficients from sets $\mathcal{P}_s$ that were *not* selected are fixed at zero. That is, we define a *debiased* solution $\tilde{\beta}^*$ corresponding to $\mathcal{R}^*$ as follows:

$$\max_\beta \ f(\beta) \ \text{ subject to } \beta_{ik} = 0 \text{ for all } (i,k) \text{ with} \tag{3.6}$$
$$k = 1, 2, \ldots, K \text{ and } i \in \mathcal{P}_s \text{ for each } s \notin \mathcal{R}^*.$$

The GroupLASSO approach is shown in Algorithm 1.

3.2 Greedy Heuristic

The regularization approach can be combined with a greedy strategy, in which groups are selected one at a time, with each selection made by solving a regularized problem. Suppose that $\mathcal{R}^{l-1}$ is set of selected groups after $l-1$ iterations of the selection heuristic. The problem solved at iteration l of the heuristic to choose the next group is

$$\hat{\beta}^l = \arg\max_\beta \ f(\beta) - \tau w_{\mathcal{S}\setminus\mathcal{R}^{l-1}}(\beta). \tag{3.7}$$

The next group s^l is obtained from $\hat{\beta}^l$ as follows:

$$s^l = \arg\max_{s\in\mathcal{S}\setminus\mathcal{R}^{l-1}} \ q_s([\hat{\beta}^l]_{\mathcal{P}_s}), \tag{3.8}$$

and we set $\mathcal{R}^l = \mathcal{R}^{l-1} \cup \{s^l\}$. Note that we do *not* penalize groups in $\mathcal{R}^{l-1}$ that have been selected already, in deciding on the next group s^l. After choosing r groups by this process, the debiasing step is performed to find the best maximum likelihood estimate for the sparse observation. Algorithm 2 describes this greedy approach. Note that the initial set of groups $\mathcal{R}^0$ might not be empty since we can use additional information that is independent from the PMU measurement. The initial set $\mathcal{R}^0$ may also contain the locations of PMUs that are already installed in the grid; our problem in this case is to decide locations for additional PMUs to improve the outage detection performance of the PMU network.

The major advantage of this approach is that redundant observations are suppressed by already-selected, non-penalized observations at each iteration. We give more details in Chapter 4.

Algorithm 2 Greedy Heuristic

Input:

Choose an initial set of groups: $\mathcal{R}^0$.

Parameter τ, r.

Output:

$\mathcal{R}^r$: Set of groups after selecting r groups.

$\tilde{\beta}^r$: Maximum likelihood estimate for r-group observation.

1: **for** $l = 1, 2, \cdots, r$ **do**

2: Solve (3.7) with $\mathcal{R}^{l-1}$ for $\hat{\beta}^l$.

3: $s^l \leftarrow \arg\max_{s \in \mathcal{S} \setminus \mathcal{R}^{l-1}} q_s([\beta^l]_{\mathcal{P}_s})$

4: $\mathcal{R}^l \leftarrow \mathcal{R}^{l-1} \cup \{s^l\}$

5: **end for**

6: Solve (3.6) with $\mathcal{R}^* = \mathcal{R}^r$ to obtain $\tilde{\beta}^r$. (▷) *debiasing*

Chapter 4

Numerical Experiments

Here we present experimental results for the approaches proposed above. The test sets considered here are based on the power system test cases from MATPOWER [17] (originally from [18]), with demands altered to generate training and test sets for the MLR approach.

4.1 Synthetic Data Generation

Since the data provided from IEEE test case archive [18] is a single snapshot of the states of power systems, we extend them to a synthetic 24-hour demand data cycle by using a stochastic process, as follows.

a. Take the demand values given by the IEEE test case archive as the average load demand over 24-hours.

b. Generate the demand variation profile by using an additive Ornstein-Uhlenbeck process as described in [19], separately and independently on each demand bus.

c. Combine the average demand and the variation ratio to obtain the 24-hour load demand profile for the system.

Figure 3.1 shows demand data generated by this procedure at three demand buses in the 9-Bus system (`case9.m`) from MATPOWER. Figure a shows the data drawn from the data file, now taken to be a 24-hour average. Figure b shows the ratio generated by the additive

Table 4.1: Test Cases.

System	Filename in MATPOWER	# of Lines		Train (5)	Test (50)
		Feas.	Infeas./Dup.		
14-Bus	case14.m	18	2	90	900
30-Bus	case_ieee30.m	37	4	185	1850
57-Bus	case57.m	67	13	335	3350
118-Bus	case118.m	170	16	850	8500

Table 4.2: Line Outage Detection Error Rates on Test Set with PMUs on All Buses.

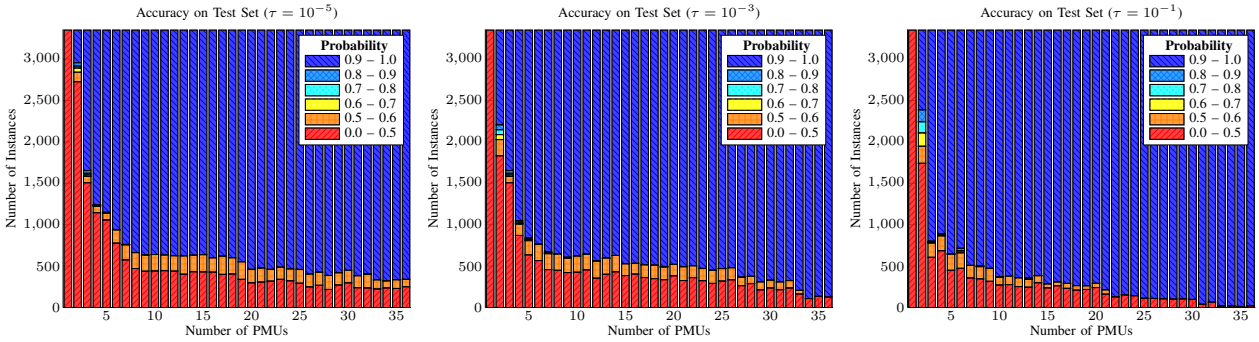
System	Probability						Ranking						Using DC Model (NAD)
	Using X			Using $\overline{X}$			Using X			Using $\overline{X}$			
	≥ 0.9	≥ 0.7	≥ 0.5	≥ 0.9	≥ 0.7	≥ 0.5	1	≤ 2	≤ 3	1	≤ 2	≤ 3	
14-Bus	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
30-Bus	0.32%	0.32%	0.32%	0%	0%	0%	0.32%	0%	0%	0%	0%	0%	1.30%
57-Bus	0.48%	0.27%	0.18%	0.48%	0.33%	0.21%	0.18%	0.06%	0.06%	0.21%	0.03%	0%	4.42%
118-Bus	0.53%	0.51%	0.47%	0.24%	0.21%	0.19%	0.47%	0.32%	0.31%	0.19%	0.08%	0.02%	3.11%

- The values in each cell are the percentage of incorrectly predicted line outages among the test instances.
- "Probability" measure indicates statistics for the probability assigned by the MLR classifier to the actual outage event.
- "Ranking" measure indicates whether the actual event was ranked in the top 1, 2, or 3 of probable outage events by the MLR classifier.
- Normalized angle distances (NAD) [2] are obtained by solving DC power flow models for test instances using MATPOWER [17].

Ornstein-Uhlenbeck process, and Figure c shows the products of the average and ratio. Since the power injected to the system needs to increase proportionally to the total demands, all power generation is multiplied by the average of the demand ratios. This average of ratios is used as the generation level G for the observation vector $\overline{X}$ defined by (2.9). The data assumes a 10-second interval between the measurements, so the total number of time points in the generated data is $24 \times 60 \times 6 = 8640$.

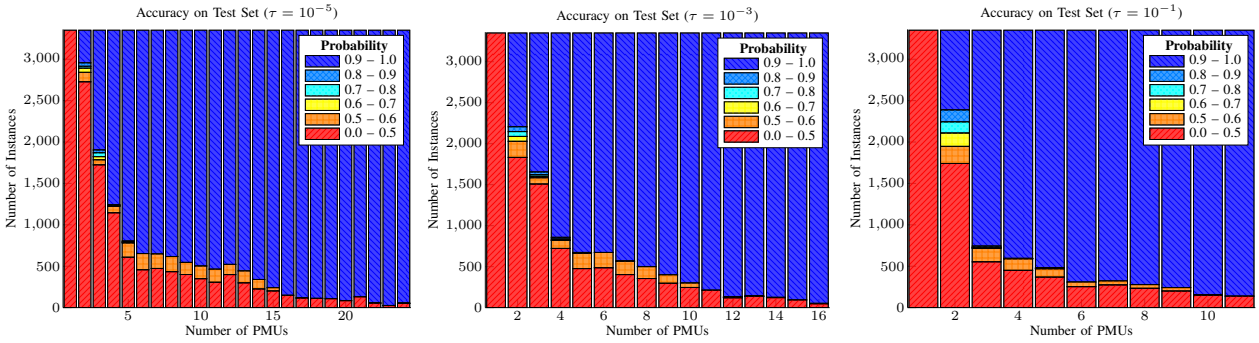
Once the 24-hour load demand profile is obtained, the AC-power equations are solved using MATPOWER to calculate the voltage phasor values at each time point. These phasor values are taken to be the PMU measurements for a normal operation cycle over a 24-hour period. MATPOWER's AC power flow equations solver is also used to evaluate voltage phasors for each single-line outage scenario that does not lead to an infeasible system. (During this process, if there exist duplicated lines that connect the same pair of buses, they are considered as a single line, that is, we do not allow only a fraction of multiple lines that connect the same set of buses to be failed.) Simulation of single-line failures to generate training data is necessary because there are typically few instances of actual outages available for study. The voltage variation for each line outage at time t is calculated by subtracting these normal-operation voltages at timepoint $t-1$ from line outage voltages at time point t. (The 10-second interval between measurements is usually sufficient time to allow transient fluctuations in phasor values to settle down; see [2].) This process leads to a number of labeled data pairs (X, Y) (or $(\overline{X}, Y)$) which we can use to train or tune the MLR classifier.

Table 4.1 provides basic information on the power systems used for the experiments. The number of lines whose removal does not prevent feasible operation of the grid is given in the column "Feas.,", while the number of lines that are duplicated or whose removal leads to an infeasible power flow problem is shown in the column "Infeas./Dup.". For each *feasible* line outage, five equally spaced samples are selected from the first half (that is, the first 12-hour period) of voltage variation data as training instances. Fifty samples are selected



(a) $\tau = 10^{-5}$ (b) $\tau = 10^{-3}$ (c) $\tau = 10^{-1}$

Figure 4.1: Accuracy on Test Set of IEEE 57-Bus System for different values of τ: Group-Sparse Heuristic



(a) $\tau = 10^{-5}$ (b) $\tau = 10^{-3}$ (c) $\tau = 10^{-1}$

Figure 4.2: Accuracy on Test Set of IEEE 57-Bus System for different values of τ: Greedy Heuristic

randomly from the second half of voltage variation data as test instances. The numbers of training and test instances are shown in the last two columns of the table. We note that the number of training instances is quite modest; just five samples per outage scenario are needed to yield reliable identification, as we show next.

4.2 PMUs on All Buses

We present results for line outage detection when phasor measurement data from all buses is used. The maximum likelihood estimation problem (2.4) with these observation vectors is solved by L-BFGS algorithm [14], coded in MATLAB. The training time to obtain the coefficient β for 118-Bus system is about 9 seconds on a personal laptop computer (2 GHz Intel Core i7 processor with 8GB RAM), with an additional 6 seconds required to solve AC power flow problems for $5 \times (170 + 1) = 855$ instances for training set generation.[1] We measure performance of the identification procedure in two ways. The first measure is based

[1]We need to solve one AC power flow problem to obtain the pre-outage base state, and one additional AC power flow problem for each of the 170 possible outages. This entire process is carried out once at each of the five training times, each of which has a different pattern of demands.

Table 4.3: Error Rates on Test Set for GroupLASSO and Greedy Heuristic Selections on the 57-Bus System.

# of PMUs	Strategy	τ	PMU Locations	Probability			Ranking		
				≥ 0.9	≥ 0.7	≥ 0.5	1	≤ 2	≤ 3
10	GroupLASSO	1.1	1* 8 17 27 28 51 52 53 54 55	27.2%	26.9%	21.2%	21.1%	7.3%	4.3%
	Greedy	1.2×10^{-1}	1* 2 17 19 26 39 40 45 46 57	7.4%	7.3%	5.7%	5.7%	0.3%	0.1%
15	GroupLASSO	8.0×10^{-1}	1* 2 4 17 23 27 28 43 46 47 51 52 53 54 55	17.2%	17.2%	11.7%	11.7%	4.3%	4.2%
	Greedy	1.7×10^{-3}	1* 2 5 12 17 20 21 26 39 40 43 45 46 54 57	1.7%	1.7%	1.7%	1.7%	0%	0%

* indicates the reference bus.

Table 4.4: Line Outage Detection Error Rates on Test Set with PMUs on About 25% of Buses.

System	τ	# of PMUs	PMU Locations	Probability			Ranking			DC Model (NAD)
				≥ 0.9	≥ 0.7	≥ 0.5	1	≤ 2	≤ 3	
14-Bus	5×10^{-2}	3	1* 7 12	0.4%	0.3%	0.2%	0.2%	0%	0%	30.9%
	5×10^{-3}	3	1* 11 12	0%	0%	0%	0%	0%	0%	43.0%
30-Bus	5×10^{-2}	4	1* 3 23 30	0.4%	0.4%	0.4%	0.4%	0%	0%	57.9%
	5×10^{-3}	5	1* 3 14 22 29	0%	0%	0%	0%	0%	0%	45.9%
57-Bus	5×10^{-2}	12	1* 2 5 17 21 26 39 40 45 46 54 57	2.9%	2.9%	2.9%	2.9%	0.2%	0.2%	28.0%
	5×10^{-3}	14	1* 2 5 17 20 21 26 39 40 41 45 46 54 57	1.5%	1.5%	1.5%	1.5%	0.1%	0.1%	25.0%
118-Bus	5×10^{-2}	15	2 22 29 36 48 58 62 63 69* 81 91 95 106 108 115	5.8%	5.8%	5.8%	5.8%	3.8%	3.8%	30.3%
	5×10^{-3}	21	3 13 29 35 43 47 55 58 62 63 69* 75 81 82 91 93 104 106 107 113 115	0.7%	0.5%	0.4%	0.4%	0.1%	0.1%	28.2%

* indicates the reference bus.

on the probability assigned by the MLR model to the actual line outage. The "probability" column on Table 4.2 shows failure rates for the classifiers according to this measure, for both the original phasor difference vector X (2.8) and the extended vector $\overline{X}$ (2.9). Each column shows the percentage of testing samples for which the probability assigned to the correct outage *does not* exceed 0.9, 0.7, and 0.5, respectively. The result shows that the performance of line outage identification is very good, even for the original observation vector X. For both X and $\overline{X}$, the actual line outage is assigned a probability of greater than 0.5 in at least 99% of test cases.

The second measure is obtained by ranking the probabilities assigned to each line outage on the test datum, and score a positive mark if the correct outage is one of the top one, two, or three cases in the ranking. We see in the "ranking" column on Table 4.2 that the actual case appears in the top two in almost every case.

For comparison, the last column of Table 4.2 shows the error rates of line outage identification which uses only *phase angles* from a DC model. The pre-outage phase angles and post-outage phase angles are computed using MATPOWER's DC power flow problem solver, and the line whose pre- and post-outage angle difference is the closest to the observed angle difference (obtained from AC power flow equations) using the normalized-angle-distance (NAD) metric [2] is identified as the failed line. The identification performance is still good, but the error rates (as high as 4.4%) are significantly higher than our MLR approach. Even in the case in which PMUs are installed on all buses, there seems to be a significant price to pay in terms of degraded identification performance, when only phase angle data is used.

4.3 PMU Placement

In this section we only consider the extended observation vector $\overline{X}$ defined by (2.9). We assume too that a PMU is installed on the reference bus, for purposes of maintaining consistency in phase angle measurement. We describe in some detail the performance of the proposed algorithm on the IEEE 57-Bus system, showing that line-outage identification performance when PMUs are placed judiciously almost matches performance in the fully-instrumented case. We then summarize our computational experience on IEEE 14-, 30-, 57-, and 118-Bus systems.

For our regularization schemes, we used groups $\mathcal{P}_s$, $s = 1, 2, \ldots, N$, defined as in (3.2). The final two entries in the extended observation vectors (the average-generation and constant terms) are not included in any group, so there two features are used by all classifiers.

4.3.1 IEEE 57-Bus System

We describe here results obtained on the IEEE 57-Bus system with two heuristics discussed in Chapter 3: The GroupLASSO and Greedy Heuristics.

In Figure 4.1, results for the GroupLASSO heuristic are displayed for different values of τ. The horizontal axis indicates the number of PMUs selected by this heuristic. The vertical axis indicates the number of test cases for which the true outage was classified by the heuristic. Each bar is divided into sections according to the probability assigned to the true outage by the MLR classifier. The top partitions indicate cases in which high probability is assigned (that is, the outage was identified correctly) while the bottom section of each column indicates cases in which the probability assigned to the true outage scenario is less than 0.5. For example, the second bar from the left in Figure c, which corresponds to two installed PMUs, corresponds to the following distribution of probabilities assigned to the correct outage scenario, among the 3350 test instances.

Probability	$[.9, 1]$	$[.8, .9]$	$[.7, .8]$	$[.6, .7]$	$[.5, .6]$	$[0, .5]$
# of Instances	963	144	135	161	206	1741

Note that the area for probability $[.9, 1]$ (top-most area) occupies a fraction $963/3350$ of the bar, the area for probability $[.8, .9]$ (second from top) occupies $144/3350$, and so on.

When only one PMU is installed, that bus naturally serves as the angle reference, so no phase angle difference information is available, and identification cannot be performed. As expected, identification becomes more reliable as PMUs are installed on more buses. The value $\tau = 0.1$ (Figure c) appears to select locations better than the smaller choices of regularization parameter. For this value, about 10 buses are sufficient to assign a probability of greater than 90% to the correct outage event for more than 90% of the test cases, while near-perfect identification occurs when 30 PMUs are installed. Note that for $\tau = .1$, there is only slow marginal improvement after 10 buses; we see a similar pattern for the other values of τ. The locations added after the initial selection are being chosen on the basis of information from the single regularized problem (3.4), so locations added later may be providing only redundant information over locations selected earlier.

Figure 4.2 shows performance of the Greedy Heuristic, plotted in the same fashion as in Figure 4.1. For each value of τ, Algorithm 2 is performed with $\mathcal{R}^0 = \emptyset$, with iterations

continuing until there is no group $s \in \mathcal{S} \setminus \mathcal{R}^{l-1}$ such that $q_s([\beta^l]_{\mathcal{P}_s}) > 0$. Termination occurs at 24, 16, and 11 PMU locations for the values $\tau = 10^{-5}, 10^{-3}$, and 10^{-1}, respectively. As the value of τ increases, the number of PMUs which are selected for line outage identification decreases. We can see by comparing Figures 4.1 and 4.2 that classification performance improves more rapidly as new locations are added for the Greedy Heuristic than for the GroupLASSO Heuristic. Larger values of τ give slightly better results. We note Figure c that almost perfect identification occurs with only 11 PMU locations, while only 6 locations suffice to identify 90% of outage events with high confidence.

Although we can manipulate the GroupLASSO technique to achieve sparsity equivalent to the Greedy Heuristic (by choosing a larger value of τ), the PMUs selected by the latter give much better identification performance on this test set. In Table 4.3, the parameter τ in the GroupLASSO heuristic is chosen manually, to find the solutions with 10 PMUs and 15 PMUs for the 57-Bus system. Performance is compared to that obtained from the Greedy Heuristic, with a much smaller value of τ. Results for the Greedy Heuristic are clearly superior.

4.3.2 Greedy Heuristic on 14-, 30-, 57- and 118-Bus System

We applied the Greedy Heuristic to 14-, 30-, 57-, and 118-Bus Systems with two values of $\tau = 5 \times 10^{-2}$ and $\tau = 5 \times 10^{-3}$, and found that the phasor measurements from the small set of buses are enough to have the similar line outage identification performance to the full measurement cases. Table 4.4 shows the PMU locations selected for each case, and line outage identification performance. Identification performance is hardly degraded from the fully instrumented case, even when phasor measurements are available from only about 25% of buses. The location of PMUs for IEEE 30-Bus and IEEE 57-Bus systems are displayed in Figure 4.3, with instrumented buses indicated by circles.

If, however, only phase angles are used for outage identification, the error rates increase dramatically; see the column "DC Model (NAD)" in Table 4.4. These results show that the advantages of using full phasor measurements (voltage magnitudes as well as phase angles) are significant when PMU measurements are available from limited number of buses. We note that the PMU locations were chosen optimally for our MLR classifier, and that the optimal locations may be different for the NAD metric. However, we expect that the chosen PMU locations would be broadly similar, so we believe that the comparison is worth reporting.

4.3.3 Combining PMU Placement with Retraining

The optimal placement of PMUs in a network will depend on network conditions, including typical demand and generation profiles. However, PMU placement is a design decision made for the medium-long term, and this decision should take account of a wide variety of demand profiles that may be encountered during the decision period. When the grid contains renewable energy sources, PMU placement decisions should also account for the wide range of output from these sources.

Having decided on PMU placement, however, we have the flexibility to *retrain* our identification algorithm periodically, determining each time a new value of the coefficients β

Table 4.5: Error Rates Obtained for General PMU Placement plus Retraining.

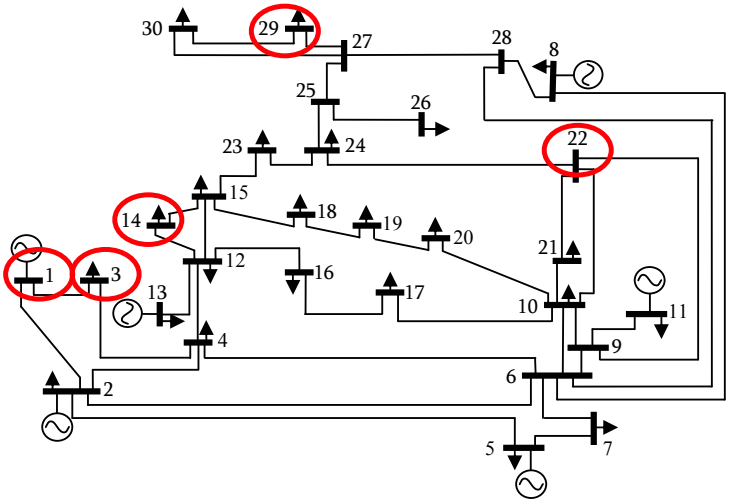
PMU Placement		Performance Test on Different Load Profiles						
# of PMUs	PMU Locations	Load Profile (κ) for Testing	Probability			Ranking		
			≥ 0.9	≥ 0.7	≥ 0.5	1	≤ 2	≤ 3
15 ($\tau = 2.8 \times 10^{-1}$)	2 22 29 36 48 58 63 67 69* 81 84 95 102 107 108	0.85	5.6%	5.5%	5.4%	5.4%	2.5%	1.8%
		1.00	6.2%	6.2%	6.2%	6.2%	4.8%	4.7%
		1.15	11.4%	11.1%	11.0%	11.0%	3.6%	2.1%
21 ($\tau = 2.25 \times 10^{-2}$)	2 13 14 27 35 43 47 55 58 62 63 69* 75 81 86 91 95 104 106 108 114	0.85	1.9%	1.8%	1.8%	1.8%	0.1%	0.1%
		1.00	0.4%	0.3%	0.2%	0.2%	0.1%	0%
		1.15	1.1%	1.0%	1.0%	1.0%	0.2%	0.1%

* indicates the reference bus.

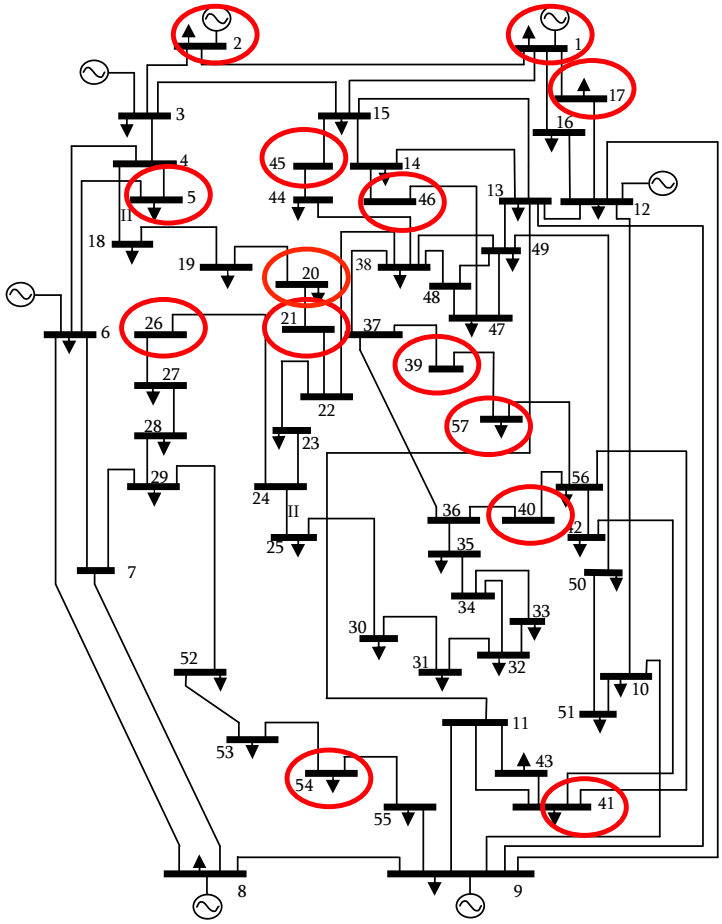
that is customized to prevailing grid conditions but that makes use *only of the selected PMU locations*.

We demonstrate the effects of this procedure on out 118-Bus test set as follows. First, we took the standard demand profiles used as the basis for our experiments above, and added extra training points to the data set in which (a) all demands were scaled by .85 and (b) all demands were scaled by 1.15. The set of training data used to select the PMUs was thus tripled in size. Second, we obtained optimal PMU locations, for varying numbers of PMUs, using the greedy heuristic described above. Third, we retraining the classifier to obtain a different set of coefficients β for each of the three scenarios — standard demand data, demands scaled by .85, and demands scaled by 1.15 — requiring the classifier to use only data from the selected PMU locations. (That is, the coefficients of β at non-selected locations was fixed at zero.) We then tested the performance of these three classifiers on test data for each of the three scenarios.

Results are shown in Table 4.5. To allow direct comparison with Table 4.4, we chose two values for the number of PMUs: 15 and 21. (Note that the values of τ used in the greedy heuristic in Table 4.5 differ from the values used in Table 4.4. This is in part because the number of terms in the $f(\beta)$ term in the objective (3.7) is tripled, so that τ should be approximately tripled to compensate.) Table 4.5 shows locations for the PMUs using the unified data set, with the final columns showing results on the test data for each of the three scaling factors, which are represented by the parameter κ. In comparing the results for 15 PMU locations with those chosen for the single value $\kappa = 1$ in Table 4.4, we see that 4 out of 15 PMUs were placed in different locations. A comparison of identification performance shows that the error rates go up slightly for $\kappa = 1$ (the "top-3" error rate increases from 3.8% to 4.7%), while the error rates for the new values of κ (.85 and 1.15) are good (top-3 rates of 1.8% and 2.1%, respectively). For the 21-location results, 7 PMUs are placed in different locations from Table 4.4. Identification performance is strong, with top-2 and top-3 error rates near zero, essentially matching the results of Table 4.4.



(a) IEEE 30-Bus System. ($\tau = 5 \times 10^{-3}$, 5 PMUs)



(b) IEEE 57-Bus System. ($\tau = 5 \times 10^{-3}$, 14 PMUs)

$*$ System diagrams are taken from [20, 21].

Figure 4.3: PMU Placement for IEEE 30-Bus and IEEE 57-Bus Systems.

Chapter 5

Conclusions

We have presented a novel approach for identifying single-line outages in a power grid from data supplied from PMUs, using a multinomial regression classifier. The model employs historical load demand data and simulated output of an AC power-flow model to train the classifier, then uses the trained classifier to identify outages in real time from streaming PMU data. Numerical results obtained on IEEE 14-, 30-, 57-, and 118-Bus systems prove that the approach can identify outages reliably.

Building on this line outage identification framework, we study the problem of placing PMUs devices so as to optimize performance of the classification framework. Heuristics are proposed to decide which buses should be instrumented with PMUs. In our test cases, the performance of the classifier when PMUs are installed on just 25% of the buses is almost as good as when PMUs are attached to all buses.

Several extensions to this work are possible. We have found that detection performance becomes even better if we take account of the fact that line outages can be detected *directly* for lines to which a PMU is attached; there is no need to rely on the indirect evidence of voltage phasor changes, for these lines. We have also found that multiple-line outages can be detected reliably (at least for a small grid), despite the much greater number of classes in the MLR classifier.

Bibliography

[1] U.S.-Canada Power System Outage Task Force, "Final Report on the August 14, 2003 Blackout in the United States and Canada," Apr. 2004. [Online]. Available: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf

[2] J. E. Tate and T. J. Overbye, "Line Outage Detection Using Phasor Angle Measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, Nov. 2008.

[3] ——, "Double line outage detection using phasor angle measurements," in *2009 IEEE Power & Energy Society General Meeting*, Calgary, AB, Jul. 2009, pp. 1–5.

[4] A. Y. Abdelaziz, S. F. Mekhamer, M. Ezzat, and E. F. El-Saadany, "Line outage detection using Support Vector Machine (SVM) based on the Phasor Measurement Units (PMUs) technology," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[5] H. Zhu and G. B. Giannakis, "Sparse Overcomplete Representations for Efficient Identification of Power Line Outages," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2215–2224, Nov. 2012.

[6] J.-C. Chen, W.-T. Li, C.-K. Wen, J.-H. Teng, and P. Ting, "Efficient Identification Method for Power Line Outages in the Smart Power Grid," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1788–1800, Jul. 2014.

[7] L. Zhao and W.-Z. Song, "Distributed power-line outage detection based on wide area measurement system." *Sensors*, vol. 14, no. 7, pp. 13 114–13 133, Jan. 2014.

[8] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, Aug. 2006.

[9] H. Zhu and Y. Shi, "Phasor Measurement Unit Placement for Identifying Power Line Outages in Wide-Area Transmission System Monitoring," in *2014 47th Hawaii International Conference on System Sciences*, no. Section 2, Waikoloa, HI, Jan. 2014, pp. 2483–2492.

[10] Y. Zhao, A. Goldsmith, and H. V. Poor, "On PMU location selection for line outage detection in wide-area transmission networks," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[11] Y. Zhao, J. Chen, A. Goldsmith, and H. V. Poor, "Identification of Outages in Power Systems with Uncertain States and Optimal Sensor Locations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1140–1153, Dec. 2014.

[12] J. Wu, J. Xiong, P. Shil, and Y. Shi, "Optimal PMU placement for identification of multiple power line outages in smart grids," in *2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, College Station, TX, Aug. 2014, pp. 354–357.

[13] A. R. Bergen and V. Vittal, *Power systems analysis*, 2nd ed. Prentice Hall, 1999.

[14] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, pp. 503–528, 1989.

[15] B. Krishnapuram, L. Carin, M. A. T. Figueiredo, and A. J. Hartemink, "Sparse multinomial logistic regression: fast algorithms and generalization bounds." *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 6, pp. 957–68, Jun. 2005.

[16] L. Meier, S. van de Geer, and P. Bühlmann, "The group LASSO for logistic regression," *Journal of the Royal Statistical Society B*, vol. 70, no. 1, pp. 53–71, Jan. 2008.

[17] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[18] University of Washington, Electrical Engineering, "Power Systems Test Case Archive," 2014. [Online]. Available: http://www.ee.washington.edu/research/pstca/

[19] M. Perninge, V. Knazkins, M. Amelin, and L. Söder, "Modeling the electric power consumption in a multiarea system," *European Transactions on Electrical Power*, no. 21, pp. 413–423, Jan. 2011.

[20] B. Allaoua and A. Laoufi, "Collective intelligence for optimal power flow solution using ant colony optimization," *Leonardo Electronic Journal of Practices and Technologies*, no. 13, pp. 88–104, 2008.

[21] B. Gasbaoui and B. Allaoua, "Ant Colony Optimization Applied on Combinatorial Problem for Optimal Power Flow Solution," *Leonardo Journal of Sciences*, no. 14, pp. 1–16, 2009.

Part III

An Observer-Based Approach to Topology Change Estimation

Christopher L. DeMarco
Sowmya Acharya, Graduate Student

University of Wisconsin–Madison

For information about this part of the project report, contact:

Christopher L. DeMarco
Department of Electrical and Computer Engineering
University of Wisconsin–Madison
1415 Engineering Drive Madison, WI 53706
Phone: 608 262-5546
Email: cdemarco@wisc.edu

Power Systems Engineering Research Center

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

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

Identification of power line outages and, generally, changes in the open/close status of transmission line breakers is particularly critical for a number of tasks, including state estimation, optimal power flow, real-time contingency analysis, and thus, security assessment of power systems.

Phasor measurement units (PMUs) provide voltage and power data per local system in real time. Likewise, real-time data are telemetered internally to offer topology-bearing information on the connectivity status of local circuit breakers and switches. Although PMUs have become increasingly widespread throughout power networks, the buses monitored by PMUs still constitute a very small percentage of the total number of system buses. Thus, useful phasor measurement data is available only at a subset of buses. The system data exchange (SDX) module of the North American Electric Reliability Corporation (NERC) can provide grid-wide interarea (i.e. basecase) topology information on an hourly basis [7], but this may prove inadequate in near real-time monitoring of transmission lines. Following a line switching event, evaluation of the power system uses real time data from PMUs, but may mistakenly continue using the original basecase topology information. This may lead to incorrect values for the power flows on lines and voltages, which in turn could lead to risk of voltage collapse and blackouts if left unchecked.

In a nutshell, the need arises for each internal system to identify, in a computationally efficient manner, line outages (and, generally, line status changes) in its external counterpart relying only on basecase topology information and local PMU data. The aim of this research is to derive useful information from PMU data in spite of its limited coverage, and use it to detect system line outages using known system basecase topology information.

1.2 Literature Review

Existing approaches are either challenged by the combinatorial complexity issues involved and are thus computationally tractable for identifying single and double line-outages or they require less pragmatic assumptions such as conditionally independent phasor angle measurements available across the grid.

The work on detection of topology errors by state estimation in [5] involves using measurement data collected by the Supervisory Control and Data Acquisition (SCADA) system, namely, status data of breakers and switches, and analog data of real and reactive power flows, injections, and bus voltages to determine real-time topology of the network. State estimation is used to process this data globally, and certain conditions for detectability of topology errors are presented. Apart from line outages, this paper also considers detection of bus splits and shunt capacitor/reactor switching. The residual-based r^N- test for bad data in state estimation is used for the detection of topology errors from the measurement data of the status. However, this approach may fail to detect the location of the topology error, as well as multiple line outages. Another major disadvantage is that the error is detectable only if true state of the system is known.

The problem addressed in [2] is the detection of single line outages using only PMU data, transmission line and transformer parameter data, and system topology information. Following a system event, the phasor angle differences at buses with PMUs (measured buses) in the system with respect to their pre-event values can be determined as $\Delta\theta_{observed}$. Once the changes in angles at each bus have been determined (denoted as a K -dimensional vector, where K is the number of phasor angles measured using PMUs), the following optimization problem is solved:

$$E^* = \arg\min_{E\in\varepsilon} ||\Delta\theta_{observed} - f(E)||$$

where ε is the set of events to be checked for occurrence, and $f(E)$ is a function which relates an event E to the changes in angles caused by the event. Although it is not feasible to include every possible event on the system, one useful application of the algorithm developed in this paper would be the extension of the event set to include a wider array of potential events such as multiple line outages and single generator outages. The results shown in [3] indicate that the methods used to detect single line outages [2] can be extended to detect double line outages on the system. However, defining the external line identification problem in this fashion, as a combinatorially complex (integer programming) problem, would be computationally demanding.

One of the recent papers published on power systems topology error estimation, perhaps closest to the work to be presented here, is by Zhu and Giannakis entitled 'Sparse Overcomplete Representation for Efficient Identification of Power Line Outages' [4]. This paper develops a near-real time algorithm for identifying multiple line outages by solving a sparse signal reconstruction problem via

(a) Greedy Orthogonal Matching Pursuit;

(b) Lasso Technique using coordinate descent (CD) iterations;

using only a subset of voltage phasor angle data from PMUs and hourly basecase topology information. Representing the power system in a linear DC power flow model, the paper constructs the pre- and post-event phase differences of the internal system's PMU data as a sparse overcomplete representation:

$$\tilde{\theta}_I = [B^{-1}]_I(\sum_{l\in\tilde{\varepsilon}} s_l m_l) + [B^{-1}]_I\eta$$

where B is the admittance matrix used to represent the DC power flow model, m_l's are columns of the incidence matrix, η accounts for small perturbations between the pre- and post- power injections (which is independent of the line outage event), and the subscript I represents the measured subset of buses in the system. The outaged lines are collected in the subset $\tilde{\varepsilon}$, and the s_l's are the coefficients of this sparse overcomplete representation.

In order to overcome the combinatorial complexity of previous exhaustive search algorithms, this work reformulates the problem into a sparse linear regression model capturing all possible line outages. Using the Singular Value Decomposition (SVD) of the fat matrix $[B^{-1}]_I = U_I \Sigma_I V_I^T$, and transforming $\tilde{\theta}_I$ to $y = \Sigma_I^{-1} U_I^T \tilde{\theta}_I$:

$$\begin{aligned} y &= V_I^T (\sum_{l \in \tilde{\varepsilon}} s_l m_l) + V_I^T \eta \\ &= V_I^T M s + V_I^T \eta \\ &= As + V_I^T \eta \end{aligned}$$

This is a sparse linear regression model with A as the transformed incidence matrix (or regression matrix) and unknown regression coefficients in the sparse $L \times 1$ vector $\mathbf{s}$ with non-zero entries only in the places corresponding to line outages i.e. whose l_{th} entry is non-zero if $l \in \tilde{\varepsilon}$ and 0 otherwise. The problem now involves estimating the sparse vector $\mathbf{s}$ to recover the line outage set $\tilde{\varepsilon}$. The line outage set $\tilde{\varepsilon}$ is no longer present in the model which eliminates the need for exhaustive search.

This sparse signal reconstruction is approached in two ways. Using the Greedy Orthogonal Matching Pursuit (OMP) algorithm, a least squares fit to y is found using the columns of the regression matrix A. If the number of line outages L_o is known, the OMP algorithm is shown to find the solution in a single step. If only the maximum number of line outages prescribed is given, say k_{max}, the OMP algorithm outputs a string of greedy solutions for different number of possible line outages. One can adopt the Minimum Description Length (MDL) type test to find the optimum of these solutions. The signal reconstruction could also be carried out by the method of Lassoing Line Outages via CD iterations. Here, the most sparse solution of the over-determined system $y = As + \eta$ is found by formulating the problem as optimizing the l_0 norm of the sparse vector $\mathbf{s}$:

$$\min ||s||_0 \;\; s.t. \;\; ||y - As||_2^2 \leq \delta \tag{1.1}$$

This non-convex problem, being NP-hard, is relaxed to a related convex optimization problem instead, where the l_1 norm is used instead of the l_0 norm (known as convex relaxation). The convex quadratic program can be solved to find the global minimizer, using a CD iterative solver. Tests run on the IEEE 118- and 300- bus systems show that both algorithms perform well for upto double line outages.

The premise of the work to follow here is that greater information content for topology estimation is present in power system models that represents dynamic evolution of system states (i.e. an ODE or DAE model), as opposed to the steady-state, purely algebraic representations that dominate current literature.

Chapter 2

Luenberger Observers and the Structure of State Matrix

2.1 Review of Luenberger State Observer Theory

State feedback and many other situations rely on the availability of the internal states. However, accessibility to the internal states is often impractical, or even impossible. An alternate approach is to develop a replica of the dynamic system that provides an "estimate" of the system states based on the measured input and output of the physical system [1]. Assume that the linearized steady state system model is of the form:

$$\dot{\underline{x}} = A_{actual}\underline{x}(t) + B\underline{u}(t) \;\; ; \;\; \underline{x}(0) = \underline{x}_0 \tag{2.1}$$

$$\underline{y}(t) = C\underline{x}(t)$$

where $\underline{x} \in \mathbb{R}^{n\times 1}$ are the state variables, $\underline{y} \in \mathbb{R}^{r\times 1}$ are the measured outputs with $r \leq n$, and $\underline{u} \in \mathbb{R}^{m\times 1}$ are the inputs to the system. The system state matrix $A_0 \in \mathbb{R}^{n\times n}$, the input matrix $B \in \mathbb{R}^{n\times m}$ and the output matrix $C \in \mathbb{R}^{r\times n}$ are known. In the application here, the state variables $\underline{x}$ would be the generator angles and frequencies at the buses of the system, the inputs $\underline{u}$ would predominantly be load demands, generator active power set points, and the measured outputs $\underline{y}$ will be primarily the PMU measurements from a small subset of buses in the system.

The observer then creates a parallel simulation of the system in a model:

$$\dot{\hat{\underline{x}}} = A_{model}\hat{\underline{x}}(t) + B_{model}\underline{u}(t) \;\; ; \;\; \hat{\underline{x}}(0) = \hat{\underline{x}}_0 \tag{2.2}$$

$$\hat{\underline{y}}(t) = C_{model}\hat{\underline{x}}(t)$$

The Luenberger observer attempts to use the output error :

$$\underline{y}(t) - \hat{\underline{y}}(t)$$

as a weighted input to the model to try to correct the estimated states and cause $\hat{\underline{x}}(t) \rightarrow \underline{x}(t)$

The observer model with the output error feedback can be expressed as:

$$\dot{\hat{\underline{x}}} = A_{model}\hat{\underline{x}}(t) + B\underline{u}(t) - L(\underline{y}(t) - C\hat{\underline{x}}(t)) \tag{2.3}$$

$$= (A_{model} + LC)\hat{\underline{x}}(t) + B\underline{u}(t) - L\underline{y}(t) \tag{2.4}$$

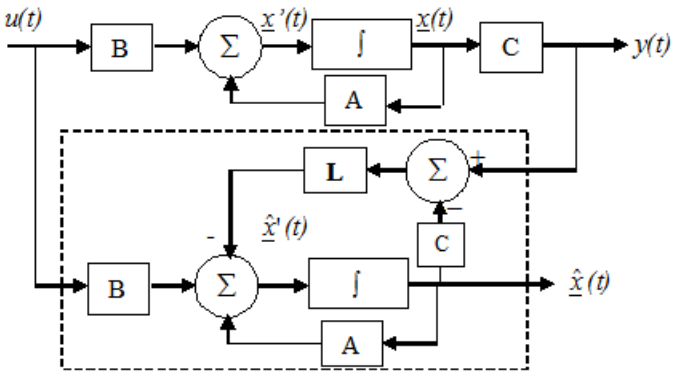


Figure 2.1: Closed Loop State Estimator

where $L \in \mathbb{R}^{n \times r}$ is the observer gain matrix.

The estimation error dynamics of the system can be defined as the difference in the actual state variables and the estimate variables, i.e. $\underline{e}(t) = \underline{x}(t) - \underline{\hat{x}}(t)$. Hence, from (2.1) and (2.4) :

$$\begin{aligned} \dot{\underline{e}} &= \dot{\underline{x}}(t) - \dot{\underline{\hat{x}}}(t) \\ &= A_{actual}\underline{x}(t) - (A_{model} + LC)\underline{\hat{x}}(t) + L\underline{y}(t) \end{aligned}$$

Defining the difference between the actual system matrix A_{actual} and the estimator model A_{model} as :

$$A_{actual} = A_{model} + \Delta A \tag{2.5}$$

the error dynamics can be simplified as:

$$\dot{\underline{e}}(t) = (A_{model} + LC)\ \underline{e}(t) + \Delta A \cdot \underline{\hat{x}}(t) \tag{2.6}$$

If the model is accurate, i.e. $A_{model} = A_{actual}$, the disturbance term $\Delta A \cdot \underline{\hat{x}}(t)$ disappears, and the Luenberger observer theory shows that for an appropriate choice of the gain matrix L, the error will asymptotically approach zero regardless of the initial conditions. Choosing the gain matrix L such that $(A_{model} + LC)$ has stable eigenvalues faster than the eigenvalues of the physical system ensures rapid conversion of the estimates $\underline{\hat{x}}(t)$ to the actual states $\underline{x}(t)$.

2.2 Line Outage Problem

In this paper, considering topology errors in the system (for e.g. lines switching out, etc.), the model system matrix A_{model} (or A_0) may not be true to the actual system matrix A_{actual} leading to a non-zero difference matrix ΔA defined above in (2.5). The error dynamics (2.6) is no longer guaranteed to asymptotically converge to zero by assigning the appropriate

eigenvalues to the feedback system $(A_0 + LC)$ due to the additional term $\Delta A \cdot \underline{\hat{x}}(t)$, that can be thought of as a disturbance input.

As explained later in this chapter in section 2.2.1, due to the structure of the system matrix A_0 in a power system, a system line outage creates a ΔA that would be a low rank matrix. The hypothesis in this research is that if one picks an appropriate gain matrix L to assign a subset of desired eigenvalues *and* eigenvectors to the feedback system $(A_0 + LC)$, the estimate variables could be made to align with the low-dimensional range space of this difference matrix ΔA and thus lead to a way to "concentrate" the error in a particular, known subspace when a line outage occurs.

2.2.1 State Matrix Formation

Consider an n bus system with $m \leq n$ generators and l transmission lines. Without loss of generality, assume the m generator buses are the first m buses, and the remaining $n - m$ buses are load buses. Select one bus, say bus 1, as the reference or slack bus that is assumed to have a generator phase angle equal to zero ($\delta_1 = 0$), and the phase angles at all the other buses are measured relative to this. The state variables are the generator phase angles and generator frequencies:

$$\underline{x} = \begin{bmatrix} \delta_2 \\ \delta_3 \\ \vdots \\ \delta_m \\ - - - \\ \omega_2 \\ \omega_3 \\ \vdots \\ \omega_m \end{bmatrix}$$

Writing the swing equations for the generator buses:

$$\begin{aligned} \dot{\delta} &= \omega(t) - \omega_{syn} \\ M\dot{\omega} &= P_m - P_e - \frac{D}{\omega_{syn}}(\omega(t) - \omega_{syn}) \end{aligned}$$

where ω_{syn} is the generator synchronous frequency, P_m is the mechanical input and P_e is the electrical output to the generator, $M \in \mathbb{R}^{m \times m}$ and $D \in \mathbb{R}^{m \times m}$ are diagonal matrices with diagonal entries corresponding to the generator inertias and damping coefficients respectively.

Assuming an unforced system and linearizing about the steady state operating point, we get

$$\Delta\dot{\delta} = I_{m \times m} \Delta\omega \tag{2.7}$$

$$\begin{aligned} \Delta\dot{\omega} &= -M^{-1}D \cdot \Delta\omega - M^{-1}\{\frac{\partial}{\partial \delta_g}(P_g - P_D) \cdot \Delta\delta\} \\ &= -M^{-1}D \cdot \Delta\omega - M^{-1}\{\frac{\partial P_g}{\partial \delta_g} - \begin{bmatrix} \frac{\partial P_g}{\partial \delta_l} & \frac{\partial P_g}{\partial V} \end{bmatrix} \begin{bmatrix} \frac{\partial P_l}{\partial \delta_l} & \frac{\partial P_l}{\partial V} \\ \frac{\partial Q_l}{\partial \delta_l} & \frac{\partial Q_l}{\partial V} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial P_l}{\partial \delta_g} \\ \frac{\partial Q_l}{\partial \delta_g} \end{bmatrix}\} \end{aligned}$$

In matrix form:

$$\begin{bmatrix} \Delta\dot{\delta}_2 \\ \vdots \\ \Delta\dot{\delta}_m \\ - - - \\ \Delta\dot{\omega}_2 \\ \vdots \\ \Delta\dot{\omega}_m \end{bmatrix} = \begin{bmatrix} 0_{m-1\times m-1} & | & I_{m-1\times m-1} \\ - - - - - - - - - - - - - & - - - & - - - - \\ -M^{-1}\{\frac{\partial}{\partial\delta_g}(P_g - P_D)\cdot\Delta\delta\} & | & -M^{-1}D \end{bmatrix} \cdot \begin{bmatrix} \Delta\delta_2 \\ \vdots \\ \Delta\delta_m \\ - - - \\ \Delta\omega_2 \\ \vdots \\ \Delta\omega_m \end{bmatrix} \tag{2.8}$$

Further simplifying the network model by assuming all buses to have generators attached, we recall the DC power flow approximation. The $\frac{\partial \underline{P}_e}{\partial \underline{\delta}}$ is a symmetric, positive semi-definite matrix with the form of a weigthed Laplacian matrix:

$$\frac{\partial \underline{P}_e}{\partial \underline{\delta}} \cong Z \cdot diag\{b_k|V_i^0||V_j^0|\cos(\delta_i^0 - \delta_j^0)\}Z^T$$

where Z is the incidence matrix defined in section 2.2.2, b_k's are the imaginary part of the admittance of the lines, $\underline{V}^0$ and $\underline{\delta}^0$ are the steady state operating points of the system.

In the DC power flow, one further simplifies to

$$\frac{\partial \underline{P}_e}{\partial \underline{\delta}} \cong [B]$$

where B is the imaginary part of the admittance matrix Y_{bus} of the system. The simplified unforced electromechanical state equations for the power system are then:

$$\begin{bmatrix} \Delta\dot{\underline{\delta}} \\ \Delta\dot{\underline{\omega}} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -M^{-1}\underbrace{\frac{\partial P_e}{\partial \delta}} & -M^{-1}D \end{bmatrix} \cdot \begin{bmatrix} \Delta\underline{\delta} \\ \Delta\underline{\omega} \end{bmatrix} \tag{2.9}$$

$$= \underbrace{\begin{bmatrix} 0 & I \\ -M^{-1}B & -M^{-1}D \end{bmatrix}}_{system\ state\ matrix A_0} \cdot \begin{bmatrix} \Delta\underline{\delta} \\ \Delta\underline{\omega} \end{bmatrix} \tag{2.10}$$

2.2.2 Low rank change due to loss of line

Given the node vs. branch incidence matrix of the system represented as a graph: $Z = [z_1\ z_2\ \dots\ z_l] \in \mathbb{R}^{n\times l}$ where a column z_k represents the line k, from say bus p to q, has zeroes in all the rows except a 1 in the pth row and a -1 in the qth row. The Y_{bus} of the system can be represented as:

$$Y_{bus} = Z \cdot diag\{y_1\ y_2\ \dots\ y_l\} \cdot Z^T$$

where y_i's are the series admittance of each line with shunt effects neglected. Switching out one line, say the kth line, we have the new Y_{bus} expressed as:

$$Y_{bus}^{new} = Y_{bus}^{old} + \underbrace{Z \cdot diag\{0\ 0\ \dots\ -y_k\ \dots\ 0\} \cdot Z^T}_{\Delta Y \to rank1}$$

Thus, a line outage shows up as only a rank 1 change in the admittance matrix of the system. Similarly, in the DC Power Flow approximation,

$$\frac{\partial \underline{P}_e}{\partial \underline{\delta}} \cong Z \cdot diag\{\underline{b}\} \cdot Z^T$$

where b's are the imaginary part of the admittance of a line. Translating it into the system matrix (2.10), it follows that the difference in the post-event and pre-event state matrices ΔA in (2.5) would also be only a rank 1 matrix.

2.2.3 Change of Coordinates

In an attempt to get a block symmetrical system state matrix, we propose a change of coordinates.

Let

$$\begin{aligned} \Delta\underline{\tilde{\delta}} &= M^{1/2}\Delta\underline{\delta} \\ \Delta\underline{\tilde{\omega}} &= M^{1/2}\Delta\underline{\omega} \end{aligned} \tag{2.11}$$

then, from (2.10) and (2.11),

$$\begin{bmatrix} \Delta\dot{\underline{\tilde{\delta}}} \\ \Delta\dot{\underline{\tilde{\omega}}} \end{bmatrix} = \begin{bmatrix} 0 & | & I \\ - - - - & - - - & - - - - \\ -M^{-1/2}[B]M^{-1/2} & | & -M^{-1/2}[D]M^{-1/2} \end{bmatrix} \cdot \begin{bmatrix} \Delta\underline{\tilde{\delta}} \\ \Delta\underline{\tilde{\omega}} \end{bmatrix}$$

i.e.

$$\begin{bmatrix} \Delta\dot{\underline{\tilde{\delta}}} \\ \Delta\dot{\underline{\tilde{\omega}}} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & | & I \\ - - - - & - - - & - - - - \\ -\tilde{B} & | & -\tilde{D} \end{bmatrix}}_{\tilde{A}} \cdot \begin{bmatrix} \Delta\underline{\tilde{\delta}} \\ \Delta\underline{\tilde{\omega}} \end{bmatrix} \tag{2.12}$$

where $\tilde{B} = -M^{-1/2}[B]M^{-1/2}$ and $\tilde{D} = -M^{-1/2}[D]M^{-1/2}$ are symmetric positive semi-definite matrices and $\tilde{A}$ is block symmetrical.

The system dynamics can now be expressed as:

Physical System:

$$\dot{\underline{\tilde{x}}} = (\tilde{A}_0 + \tilde{A})\underline{\tilde{x}} \quad ; \ \underline{\tilde{x}}(t=0) = \underline{\tilde{x}}^0$$

Observer:

$$\dot{\underline{\hat{\tilde{x}}}} = \tilde{A}_0\underline{\hat{\tilde{x}}} + L(C\underline{\tilde{x}} - C\underline{\hat{\tilde{x}}}) \quad ; \ \underline{\hat{\tilde{x}}}(t=0) = \underline{\hat{\tilde{x}}}^0$$

and the error dynamics given as:

$$\dot{\underline{\tilde{e}}} = (\tilde{A}_0 + LC)\underline{\tilde{e}} + \Delta\tilde{A}\underline{\tilde{x}}(t)$$

For ease of notation, we drop the tildes from now on.

Chapter 3

Eigenvector-Eigenvalue Placement

Consider the case of a single perturbation of the state matrix (e.g. line failure event(s)) characterized by

$$A^{actual} = A_0 + \Delta A \tag{3.1}$$

where $A^{actual} \in \mathbb{R}^{n \times n}$ is the actual linearized state matrix, $A_0 \in \mathbb{R}^{n \times n}$ is the state matrix employed by the observer (base case state matrix) and ΔA is the error.

Let the range space of ΔA be spanned by the vectors $w_1, w_2 \dots w_s$ where $n > s \geq 1$ and $W = [\, w_1 w_2 \dots w_s]$

The error dynamics of the Luenberger observer will take the form

$$\dot{e} = (A_0 + LC) \cdot e(t) + \Delta A \cdot x(t) \tag{3.2}$$

Then $\exists d : \mathbb{R} \to \mathbb{R}^s$ such that $\Delta A \cdot x(t) = W \cdot d(t)$; hence we treat $d(t)$ as an unknown disturbance input in

$$\dot{e} = (A_0 + LC) \cdot e + W \cdot d \tag{3.3}$$

Our goal is to concentrate the effect of the disturbance input into a small number (here illustrated as just one) of invariant subspaces of $(A_0 + LC)$, and minimize its effect in all other invariant subspaces.

Based on the classic work of [10], [11] investigated eigenvector-eigenvalue placement for power system controller design. However, adapting these results to observer design requires more degrees of freedom than available. We consider solving an optimization problem instead.

3.1 The Optimization Problem

To quantify the goal of minimizing the impact of disturbances in directions spanned by columns of W on the invariant subspace determined by an eigenvector v_k of $(A_0 + LC)$, we seek to minimize $v_k^T W W^T v_k$, with normalization $||v_k|| = 1$.

So we consider framing our problem of selecting $L \in \mathbb{R}^{n \times r}$ (where r is the number of measurements) as:

Given : $A_0 \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{r\times n}, W \in \mathbb{R}^{n\times s}$
Find : $L \in \mathbb{R}^{n\times r}$ to

$$minimize \quad \sum_{k=1}^{n-1} v_k^T W W^T v_k \tag{3.4}$$

$$subject\ to\ (A_0 + LC)\, V - V\Lambda = 0_{n\times n}$$

with $V = [v_1\ v_2 \dots v_n]$, $||v_k|| = 1$ and $\Lambda = diag\{\lambda_1, \lambda_2, \dots \lambda_n\}$ where $\lambda_i < 0$ for all $i = 1, 2, \dots, n$
Additionally, we include a constraint that ensures that the set of eigenvectors V spans the whole of $\mathbb{R}^n$. This can be done by adding a constraint to make sure that the matrix composed of eigenvectors V is full rank, here n. A constraint such as $\det(V) \geq \epsilon$ or the smallest singlular value of the the matrix V from the singular value decomposition of the matrix $\sigma_{min} \geq \epsilon$, where ϵ is a sufficiently small positive number, would serve the purpose.

For the structure of the system matrix in power systems, it is natural to get complex conjugate eigenvalue pairs along with real eigenvalues. We consider a transformation to avoid working in the complex plane [12]. Each complex conjugate pair of eigenvalues are represented by a 2×2 block as:

$$\begin{bmatrix} \mu_1 & \omega_1 \\ -\omega_1 & \mu_1 \end{bmatrix}$$

where $\lambda_1 = \mu_1 + j\omega_1$ and $\lambda_2 = \lambda_1^* = \mu_1 - j\omega_1$.

Thus, the diagonal matrix of eigenvalues, i.e.

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & & \lambda_n \end{bmatrix}$$

is transformed into the block diagonal form:

$$\Lambda_{block} = \begin{bmatrix} \mu_1 & \omega_1 & | & & & & & & \\ -\omega_1 & \mu_1 & | & & & & & & \\ - & - & - & - & - & - & & & \\ & & | & \mu_3 & \omega_3 & | & & & \\ & & | & -\omega_3 & \mu_3 & | & & & \\ & & - & - & - & - & & & \\ & & & & & & \ddots & & \\ & & & & & & - & - & - \\ & & & & & & | & \mu_{n-1} & \omega_{n-1} \\ & & & & & & | & -\omega_{n-1} & \mu_{n-1} \end{bmatrix} \tag{3.5}$$

The corresponding transformation matrix for the eigenvectors picks out the real and imaginary parts of the complex conjugate pair of eigenvectors corresponding to these complex

conjugate pair of eigenvalues. That is, the pair of eigenvectors $\underline{v}_1$ and $\underline{v}_2 = \underline{v}_1^*$ would be expressed as

$$\begin{bmatrix} \underline{w}_1 & \underline{w}_1 \end{bmatrix}$$

where $\underline{w}_1 = \mathrm{Re}(\underline{v}_1)$ and $\underline{w}_2 = \mathrm{Im}(\underline{v}_1)$.

The matrix composed of complex eigenvectors

$$V = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

is then transformed to the real matrix

$$V_{block} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathrm{Re}(\underline{v}_1) & \mathrm{Im}(\underline{v}_1) & \cdots & \mathrm{Im}(\underline{v}_{n-1}) \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \underline{w}_1 & \underline{w}_2 & \cdots & \underline{w}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{3.6}$$

The transformed eigenvalue-eigenvector equation

$$(A_0 + LC) \cdot V_{block} = V_{block} \cdot \Lambda_{block} \tag{3.7}$$

from (3.4), (3.5) and (3.6) is then completely real.

Finally, the constraints to the optimization problem are :

a. $(A_0 + LC) \cdot V_{block} = V_{block} \cdot \Lambda_{block}$, i.e. eigenvalue-eigenvector relation,

b. $||\underline{v}_k|| = 1$, i.e. typical normalization,

c. $rank(V) = rank(V_{block}) = n$, ensuring non-trivial Jordon blocks,

d. $\mathrm{Re}(\lambda_k) < 0$, alternatively $\mu_k < 0$,

e. $\mathrm{Im}(\lambda_k) \neq 0$, alternatively $\omega_k > 0$

While the first three constraints are nonlinear, the rest of the constraints are linear constraints. The last constraint is an optional constraint included initially for ease to have *all* complex conjugate pairs of eigenvalues.

The above nonlinear optimization problem can be solved by various algorithms including the interior-point method, which is used in this work.

3.2 5-bus test case

Considering a 5-bus test case, with generators assumed to be connected at every bus, and line data as shown below. Bus 1 is selected as the reference bus. The state variables are the generator phase angles and frequencies, and PMU measurement data is available from bus 2 and 3. At time t_0, line 4 between bus 4 and bus 5 is assumed to go out of service.

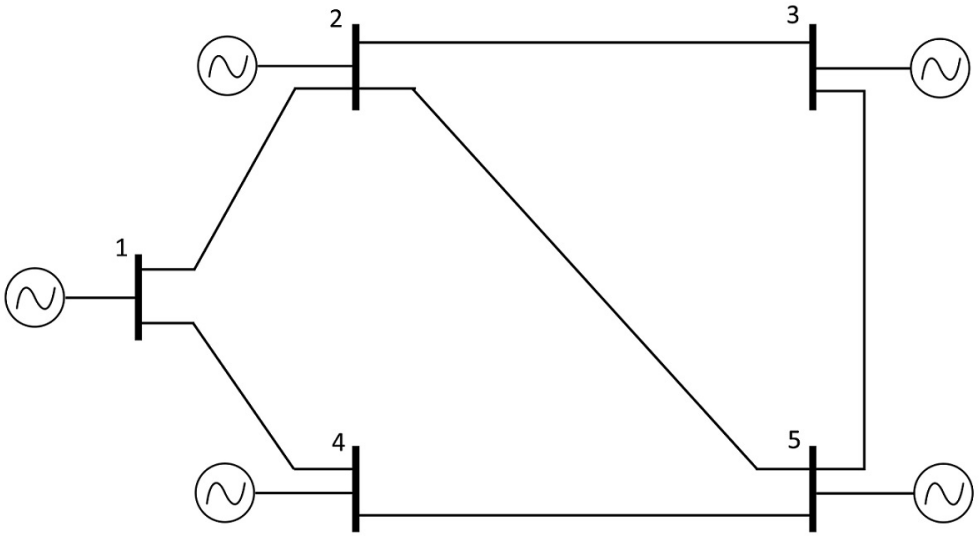


Figure 3.1: 5-bus test case

Line	'from' bus	'to' bus	susceptence b_l
1	2	1	2
2	2	3	3
3	1	4	5
4	4	5	4
5	5	2	9
6	3	5	2

Table 3.1: Line Data for 5-bus test case

The complex conjugate eigenvalue pairs of the original basecase matrix turn out to be:

Choosing a gain matrix L to ensure that the disturbance input is concentrated mostly in one of the invariant subspaces of the observer error system matrix $(A_0 + LC)$:

$$L = \begin{bmatrix} 2.0326 & -13.5736 \\ 22.0239 & -2.6280 \\ 98.5044 & 56.9155 \\ 21.9845 & -16.3897 \\ -135.6580 & -22.8317 \\ 83.1067 & 6.4995 \\ 20.2698 & -5.5386 \\ 187.4140 & 61.5838 \end{bmatrix}$$

Eigenvalues of A_0
$-2.0220 \pm j19.6301$
$-1.1912 \pm j12.6861$
$-1.3237 \pm j9.5855$
$-1.5401 \pm j3.6459$

Table 3.2: Eigenvalues of the basecase system matrix

The eigenvalues of the observer control system are then:

Eigenvalues of $A_0 + LC$
$-1.7323 \pm j22.7467$
$-1.7283 \pm j22.7434$
$-2.5905 \pm j6.3219$
$-0.3236 \pm j1.6028$

Table 3.3: Eigenvalues of the observer control system matrix

Running the composite dynamics of the system given by:

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} (A_0 + \Delta A) & 0 \\ -LC & (A_0 + LC) \end{bmatrix} \cdot \begin{bmatrix} x \\ \hat{x} \end{bmatrix} \tag{3.8}$$

from the initial conditions

$$\underline{x} \neq \underline{0}$$

$$\underline{\hat{x}} = \underline{0}$$

we see that the error dynamics of the system asymptotically converge to zero even with initial conditions of the estimates removed away from that of the actual states.

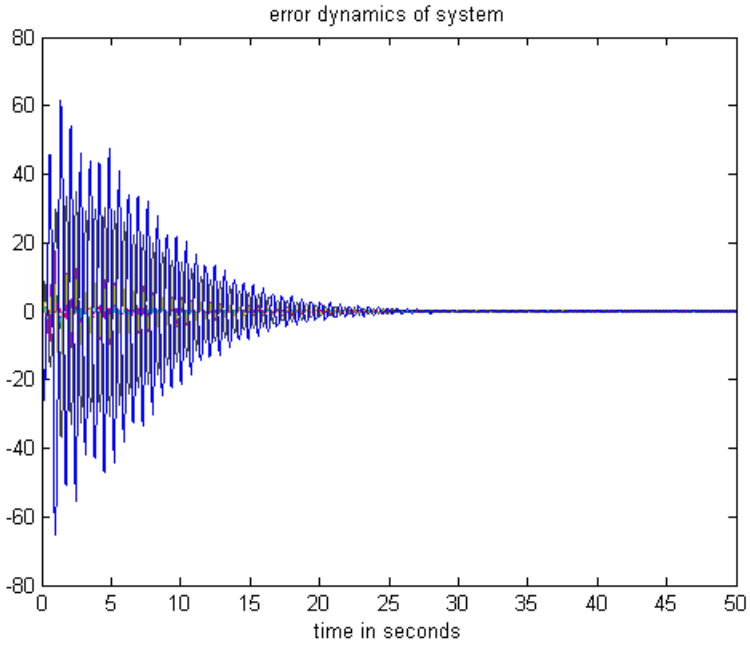


Figure 3.2: Error dynamics of the system

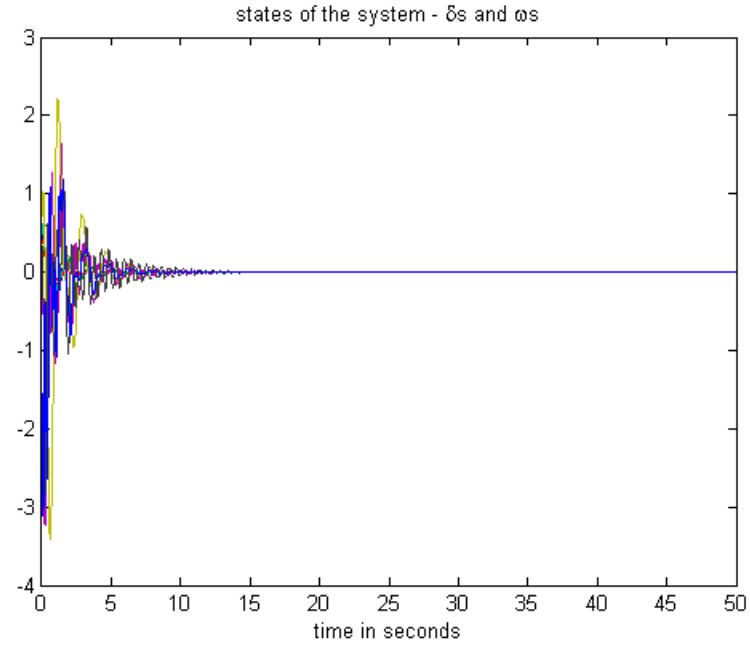


Figure 3.3: States of the system

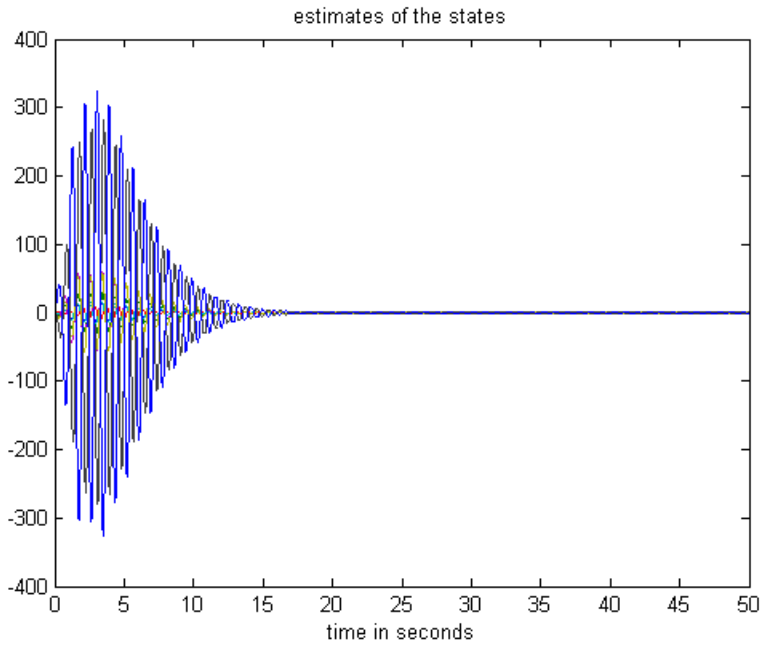


Figure 3.4: Estimates of the states of the system

3.3 Projection of the state estimates on the subspace spanned by eigenvectors

In order to observe the projection of the state estimates $\hat{x}$ onto the subspace spanned by the eignvectors of the observer state matrix $A_0 + LC$, it is required to obtain an orthonormal basis of this subspace. Consider the eigenvector $\underline{v}_i = \underline{v}_{R_i} + j\underline{v}_{I_i}$ where $\underline{v}_{R_i}$ and $\underline{v}_{I_i}$ are the real and imaginary parts of the complex eigenvector. The subspace spanned by this eigenvector is the same as the subspace spanned by the vectors $\underline{v}_{R_i}$ and $\underline{v}_{I_i}$. To get an orthonormal basis for this subspace, the Gram-Schmidt orthogonalization algorithm is used

$\underline{v}_{R_{\perp i}} = \underline{v}_{I_i} - \frac{\underline{v}_{I_i}^T \underline{v}_{R_i}}{||\underline{v}_{R_i}||^2}$ giving the normalized basis vectors to be $\hat{v}_{R_i}$ and $\hat{v}_{R_{\perp i}}$.

Observing the projection of $\hat{x}$ onto the eigenvectors of $A_0 + LC$, one can see that the estimates $\hat{x}$ tends to align with the subspace spanned by the eigenvector corresponding to the smallest magnitude eigenvalue even though this may not be the last eigenvalue (recall optimization problem is set up to try to ensure $\hat{x}$ aligns with the space spanned by eigenvector corresponding to the last eigenvalue).

Consider additional constraints to ensure that the smallest (i.e. slowest) eigenvalue is sufficiently removed away from all other eigenvalues. That is,

$$|\lambda_{n-1}| \ < \ 2|\lambda_i| \ \ for \ \ i = 1, 3, 5, ..n-3$$

with $\lambda_{i+1} = \lambda_i^*$ for DC power flow assumption. Alternately, constraints to impose a descending order to the set of eigenvalues could be introduced, but its immediate benefits are

minimal. One can now observe that, when the algorithm converges, the estimates $\underline{\hat{x}}$ tend to align with the subspace associated with the eigenvector corresponding to the last complex conjugate eigenvalue pair (which also turns out to be the smallest magnitude eigenvalue pair removed away from other eigenvalue pairs).

Testing the algorithm with 10 random initial conditions, the percentage of times the algorithm converged and gave the desired results for different line outage locations is displayed below. For the initial condition on the L matrix, picking the r columns to be normally distributed random numbers:

Location of line outage	% of success
line 5	100%
line 6	80%

Table 3.4: % of success of optimization algorithm for 5-bus 6-line test case with 10 trials

Chapter 4

Conclusions

This section has presented a method for line outage identification in a power system using the basecase topology information and PMU measurement data from a limited number of buses in the system, in which the disturbance input to the Luenberger observer is concentrated into an invariant subspace of the observer error system matrix. Further, a test of the projection of the estimated states on this invariant subspace would lead to identification of lines switched out of service.

Part 1 contains an overview of the line outage estimation problem and the motivation behind addressing this issue. Previous works in this area have been presented, highlighting the combinatorial complexity issues and limitations to single or at most double line outage identification.

Part 2 reviews the Luenberger Observer Control Theory including the conditions for appropriate choice of the gain matrix to guarantee a stable system with the error asymptotically approaching zero. The line outage identification problem is defined in terms of the Luenberger observer, and the structure of the state matrix of the power system is described. As a preliminary work, the linear DC approximation is used, and all buses are assumed to have generators connected. Some properties of the basecase state matrix are detailed as well as some useful transformations. It is important to note that a line switching event leads to a low rank change difference matrix between the pre and post event state matrices.

Part 3 outlines the method used in this work for the eigenvector-eigenvalue placement of the state feedback matrix in order to get the desired gain matrix that ensures that the disturbance input concentrates in a single subspace of the feedback matrix. Transformations to the eigenvector and eigenvalue matrices are considered while working with complex conjugate pairs of eigenvalues. The method is tested initially on a 5 bus test case, and the composite system simulation with the choice of gain matrix from the optimization problem defined is shown to stabilize. A further test would be a check of the estimate variables projecting mainly onto the chosen invariant subspace of the feedback system matrix.

Future work in this area will involve testing the method on larger systems, and then relaxing the assumptions made on the power system. Much greater information could be got out of the system by representing it as a ODE or a DAE model instead of DC power flow assumption. It is also hoped to achieve multiple line outage identifications by possibly running multiple observers in parallel checking for errors in all or a clever subset of lines.

References

[1] C.T. Chen, *Linear System Theory and Design*, 3rd ed. Oxford University Press 1999

[2] J. E. Tate, T. J. Overbye (2008). Line outage detection using phasor angle measurements. *IEEE Trans. Power Syst, vol. 23, no. 4*, 1644-1652

[3] J. E. Tate, T. J. Overbye(2009). Double line outage detection using phasor angle measurements. *Proc. IEEE PES Gen. Meeting*. 1-5.

[4] H. Zhu, G.B. Giannakis (2012). Sparse overcomplete representation for efficient identification of power line outages. *IEEE Transactions on Power Systems*, vol. 27, no. 4, 2215-2224

[5] F. W. Wu et. al(1989). Detection of topology errors by state estimation. *IEEE Transactions on Power Systems*, vol. 4, no. 1, 176-183

[6] J.D. Glover, M.Sarma, T.J.Overbye. *Power System Analysis and Design*, 5th ed. 2012

[7] "Final report on the implementation of the task force recommendation," U.S.-Canada Power System Outage Task Force, Sept. 2006 [Online: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/Outage_Task_Force_-_DRAFT_Report_on_Implementation.pdf]

[8] C.L.DeMarco, G.C.Verghese. Bringing Phasor Dynamics into the Power System Load Flow. *IEEE Power Engineering Society*, 1993

[9] IEEE standard for synchrophasor measurements for power systems, C37.118.1-2011, Dec 2011. [Online: http://ieeexplore.ieee.org/servlet/opac?punumber=6111217]

[10] B.C. Moore(1976). On the flexibility offered by state feedback in multivariable systems beyond closed loop eigenvalue assignment. *IEEE Transactions on Automatic Control*, vol. 21(5), pp. 689692

[11] C.L.DeMarco (1998). "The threat of predatory generator control to the U.S. electric power infrastructure". Report of studies performed under sandia national laboratories professional service agreement

[12] D.F. Delchamps, *State Space and Input-Output Linear Systems*, Springer-Verlag New York, 1988

[13] C.L.DeMarco, Eigenvector assignment in power system controller design: illustration through predatory control, *IEEE Power Engineering Society Summer Meeting*, vol. 2, pp. 830-834, 2000

[14] C.L.DeMarco, J.Wassner, A generalized eigenvalue perturbation approach to coherency, *Wisconsin. University - Madison. Department of Electrical and Computer Engineering. [Papers]*, University of Wisconsin, Engineering Experiment Station, ECE-95-5, pp. 36

[15] G. Klein, B.C. Moore(1977). Eigenvalue-Generalize Eigenvector Assignment with State Feedback. *IEEE Transactions on Automatic Control*, pp. 140-141

Part IV

Measurement-Based Estimation of the Power Flow Jacobian

Alejandro D. Domínguez-García
Peter W. Sauer
Yu Christine Chen, Graduate Student

University of Illinois at Urbana-Champaign

* Currently an Assistant Professor at the University of British Columbia.

For information about this part of the project report, contact:

Alejandro D. Domínguez-García
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
306 N. Wright Street,
Urbana, IL 61801
Phone: (217) 333-0394
Email: aledan@illinois.edu

Power Systems Engineering Research Center

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

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Our approach to online Jacobian matrix estimation builds upon our previous work in [1, 2]. In [1], by relying on active power bus injection and line flow data obtained from PMUs, linear sensitivity distribution factors are computed via the solution of a linear least-squares errors (LSE) estimation problem. In this work, by exploiting slight fluctuations in measurements of bus voltage magnitudes and phase angles, as well as those of net active and reactive power injections obtained from PMUs, we construct an overdetermined set of linear equations, and solve it via total least-squares (TLS) estimation; the solution to the problem provides the entries of the Jacobian matrix. In this regard, in [1], even though the regressor matrix is constructed from PMU measurements, it is assumed to be error-free as per the LSE estimation framework. In contrast, the TLS-based estimation method proposed here to compute the Jacobian matrix accounts for errors present in both the regressor matrix and the observation vector. Furthermore, we improve the adaptability of the proposed method by formulating a weighted TLS (WTLS) problem, in which recent measurements are weighted more favorably than past ones.

We illustrate the effectiveness of the proposed measurement-based Jacobian estimation method by comparing its results to benchmark values obtained via direct linearization of the power flow equations at a particular operating point. The estimated Jacobian matrix is quite accurate and can therefore be used in studies that rely on the power flow model. Moreover, the network topology can be inferred in order to facilitate standard power system analyses that are heavily dependent upon an accurate system model.

The remainder of this document is organized as follows. In Chapter 2, we define the elements of the power flow Jacobian matrix as partial derivatives of the power flow equations, describe the conventional model-based approach to compute them, and formulate the measurement-based counterpart proposed in this chapter. In Chapter 3, we formulate two algorithms to solve the Jacobian matrix estimation problem via TLS and WTLS estimation. Chapter 4 modifies the centralized problem formulation in Chapter 2 and offers a distributed scheme to estimate the Jacobian matrix. Finally, in Chapter 5, we illustrate the proposed ideas via case studies involving the IEEE 118-bus system.

Chapter 2

Preliminaries

Let $\mathcal{V}$ denote the set of N buses in the system. Let V_i and θ_i, respectively, denote the voltage magnitude and phase angle at bus i; additionally, let P_i and Q_i, respectively, denote the net active and reactive power injections at bus i. The entries of the power flow Jacobian matrix are composed of partial derivatives of P_i with respect to θ_j and V_j, which we denote by Ψ_i^j and Φ_i^j, respectively, and partial derivatives of Q_i with respect to θ_j and V_j, which we denote by Γ_i^j and Λ_i^j, respectively. Suppose θ_j varies by a small amount, denoted by $\Delta\theta_j$. Also denote by $\Delta P_i^{\theta_j}$ the change in active power injection at bus i, resulting from $\Delta\theta_j$, with all other system quantities held constant. Then, it follows that

$$\Psi_i^j := \frac{\partial P_i}{\partial \theta_j} \approx \frac{\Delta P_i^{\theta_j}}{\Delta \theta_j}. \tag{2.1}$$

On the other hand, suppose V_j varies by a small amount, denoted by ΔV_j. Also denote by $\Delta P_i^{V_j}$ the change in active power injection at bus i, resulting from ΔV_j, with all other system quantities held constant. Then, it follows that

$$\Phi_i^j := \frac{\partial P_i}{\partial V_j} \approx \frac{\Delta P_i^{V_j}}{\Delta V_j}. \tag{2.2}$$

Similarly, we define the analogue of (2.1)–(2.2) for reactive power as follows:

$$\Gamma_i^j := \frac{\partial Q_i}{\partial \theta_j} \approx \frac{\Delta Q_i^{\theta_j}}{\Delta \theta_j}, \tag{2.3}$$

where $\Delta Q_i^{\theta_j}$ denotes the change in reactive power injection at bus i, resulting from $\Delta\theta_j$, with all other quantities held constant; and

$$\Lambda_i^j := \frac{\partial Q_i}{\partial V_j} \approx \frac{\Delta Q_i^{V_j}}{\Delta V_j}, \tag{2.4}$$

where $\Delta Q_i^{V_j}$ denotes the change in reactive power injection at bus i, resulting from ΔV_j. Traditionally, the sensitivity factors in (2.1)–(2.4) have been computed offline based on a model of the power system, including its topology and pertinent parameters. Next, we describe this traditional model-based approach.

2.1 Model-Based Approach to Jacobian Computation

Consider a power system with N buses, each of which is categorized into one of the following: (i) slack bus, for which the voltage magnitude is fixed and with respect to which the phase angles of all other buses are measured, (ii) voltage-controlled bus, for which the voltage magnitude is fixed, or (iii) load bus, for which neither voltage magnitude nor phase angle are fixed (see, e.g., [**?**]). Let $\mathcal{V}_L$ ($\mathcal{V}_G$) denote the set of N_L load (N_G voltage-controlled) buses. Furthermore, without loss of generality, in subsequent developments, we assume that bus 1 is designated as the slack bus. Then, the static behavior of the power system can be described by the power flow equations:

$$P_i = p_i(\theta_1, \ldots, \theta_N, V_1, \ldots, V_N),\ i \in \mathcal{V}_G \cup \mathcal{V}_L, \tag{2.5}$$

and

$$Q_i = q_i(\theta_1, \ldots, \theta_N, V_1, \ldots, V_N),\ i \in \mathcal{V}_L. \tag{2.6}$$

In (2.5)–(2.6), the dependence on network parameters is implicitly considered in $p_i(\cdot)$ and $q_i(\cdot)$. Suppose a solution to (2.5)–(2.6) exists at $(\theta_i^0, V_i^0, P_i^0, Q_i^0)$, $i = 1, \ldots, N$. Further, assume $p_i(\cdot)$, for all $i \in \mathcal{V}_G \cup \mathcal{V}_L$, and $q_i(\cdot)$, for all $i \in \mathcal{V}_L$, are continuously differentiable with respect to θ_i and V_i, for all $i = 1, \ldots, N$, at $(\theta_i^0, V_i^0, P_i^0, Q_i^0)$, $i = 1, \ldots, N$. For each i, let $\theta_i = \theta_i^0 + \Delta\theta_i$, $V_i = V_i^0 + \Delta V_i$, $P_i = P_i^0 + \Delta P_i$, and $Q_i = Q_i^0 + \Delta Q_i$. Then, assuming $\Delta\theta_i$, ΔV_i, ΔP_i, and ΔQ_i are sufficiently small, we can approximate (2.5) as

$$\begin{aligned} P_i^0 + \Delta P_i \approx\ & p_i(\theta_1^0, \ldots, \theta_N^0, V_1^0, \ldots, V_N^0) \\ & + \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Psi_i^j \Delta\theta_j + \sum_{j \in \mathcal{V}_L} \Phi_i^j \Delta V_j, \end{aligned} \tag{2.7}$$

for each $i \in \mathcal{V}_G \cup \mathcal{V}_L$, and (2.6) as

$$\begin{aligned} Q_i^0 + \Delta Q_i \approx\ & q_i(\theta_1^0, \ldots, \theta_N^0, V_1^0, \ldots, V_N^0) \\ & + \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Gamma_i^j \Delta\theta_j + \sum_{j \in \mathcal{V}_L} \Lambda_i^j \Delta V_j, \end{aligned} \tag{2.8}$$

for each $i \in \mathcal{V}_L$, where

$$\Psi_i^j = \frac{\partial p_i}{\partial \theta_j},\ \Phi_i^j = \frac{\partial p_i}{\partial V_j},\ \Gamma_i^j = \frac{\partial q_i}{\partial \theta_j},\ \text{and}\ \Lambda_i^j = \frac{\partial q_i}{\partial V_j},$$

all of which are evaluated at the nominal operating point $(\theta_i^0, V_i^0, P_i^0, Q_i^0)$, $i = 1, \ldots, N$. Note that in (2.7)–(2.8), we have accounted for the fact that the voltages at the slack bus and the voltage-controlled buses are fixed. Next, we illustrate the ideas presented above with an example.

Example 1 (3-Machine 9-Bus System) *Here, we consider the WECC 3-machine, 9-bus system model (see, e.g., [3]), the topology of which is shown in Fig. 2.1. In this system, bus 1 is designated as the slack bus; there are $N_G = 2$ voltage-controlled buses, consisting of $\mathcal{V}_G = \{2, 3\}$; and there are $N_L = 6$ load buses, consisting of $\mathcal{V}_L = \{4, 5, \ldots, 9\}$. In this example, we compute model-based sensitivity factors by linearizing the power flow equations in* (2.7)–(2.8)*. In Table 2.1, we report the sensitivities of the active and reactive power injections at bus 4 with respect to voltage magnitudes and phase angles at all other buses.* ■

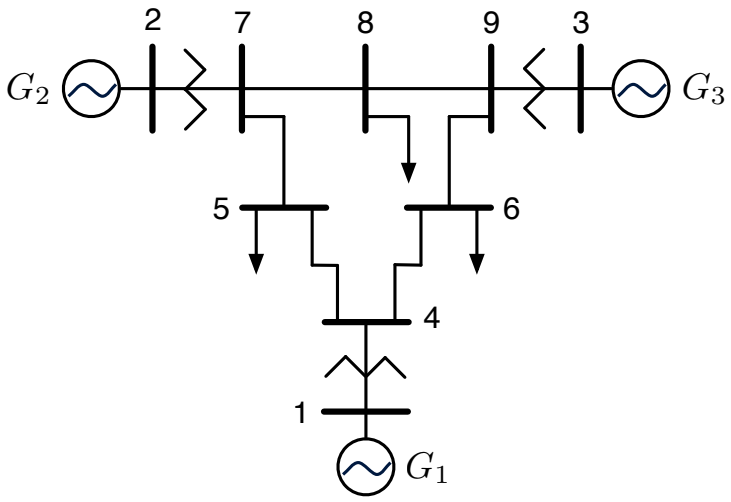


Figure 2.1: Network topology for WECC 3-machine 9-bus system.

Table 2.1: WECC 3-machine 9-bus systems—model-based sensitivity factors obtained in Examples 1.

Ψ_5^2	Ψ_5^3	Ψ_5^4	Ψ_5^5	Ψ_5^6	Ψ_5^7	Ψ_5^8
0	0	-10.86	0	16.54	0	0
Ψ_5^9	Φ_5^4	Φ_5^5	Φ_5^6	Φ_5^7	Φ_5^8	Φ_5^9
-5.6816	-2.239	0	2.3762	0	0	-1.8495
Γ_5^2	Γ_5^3	Γ_5^4	Γ_5^5	Γ_5^6	Γ_5^7	Γ_5^8
0	0	0	0	0	2.4279	-3.861
Γ_5^9	Λ_5^4	Λ_5^5	Λ_5^6	Λ_5^7	Λ_5^8	Λ_5^9
1.433	0	0	0	-13.81	23.33	-9.912

The traditional model-based approach described above is not ideal since accurate and up-to-date network topology, parameters, and operating point are required. In this work, we aim to eradicate the reliance on system models in the computation of the sensitivities defined in (2.1)–(2.4), and improve adaptability to changes occurring in the system. With regard to this, we propose a method to estimate these sensitivities using only PMU measurements obtained in near real-time without relying on the full nonlinear power flow model of the system.

2.2 Measurement-Based Approach to Jacobian Computation

Denote the voltage phase angle at bus j at times t and $t + \Delta t$, $\Delta t > 0$ and small, as $\theta_j(t)$ and $\theta_j(t + \Delta t)$, respectively. Also denote the voltage magnitude at bus j at times t and $t + \Delta t$, as $V_j(t)$ and $V_j(t + \Delta t)$, respectively. Define $\Delta\theta_j(t) = \theta_j(t + \Delta t) - \theta_j(t)$ and $\Delta V_j(t) = V_j(t + \Delta t) - V_j(t)$; then, according to the approximations of Ψ_i^j, Φ_i^j, Γ_i^j, and Λ_i^j

in (2.1)–(2.4), we have that, at time t,

$$\Psi_i^j \approx \frac{\Delta P_i^{\theta_j}(t)}{\Delta \theta_j(t)},\ \Phi_i^j \approx \frac{\Delta P_i^{V_j}(t)}{\Delta V_j(t)}, \tag{2.9}$$

$$\Gamma_i^j \approx \frac{\Delta Q_i^{\theta_j}(t)}{\Delta \theta_j(t)},\ \text{and}\ \Lambda_i^j \approx \frac{\Delta Q_i^{V_j}(t)}{\Delta V_j(t)}. \tag{2.10}$$

We assume $\theta_j(t)$, $V_j(t)$, $\theta_j(t+\Delta t)$, and $V_j(t+\Delta t)$ are measurements available from PMUs. As evidenced in (2.9), in order to compute Ψ_i^j and Φ_i^j, we also need $\Delta P_i^{\theta_j}(t)$ and $\Delta P_i^{V_j}(t)$, which are not readily available from PMU measurements. However, we assume that the net variation in net active power injection at bus i is available from PMU measurements. We express this net variation at the sum of active power injection variations at bus $i \in \mathcal{V}_G \cup \mathcal{V}_L$ due to variations in voltage phase angle $j \in \mathcal{V}_G \cup \mathcal{V}_L$ and magnitude at each bus $j \in \mathcal{V}_L$:

$$\Delta P_i(t) \approx \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Delta P_i^{\theta_j}(t) + \sum_{j \in \mathcal{V}_L} \Delta P_i^{V_j}(t). \tag{2.11}$$

Similarly, from (2.10), we note that in order to compute Γ_i^j and Λ_i^j, we need $\Delta Q_i^{\theta_j}(t)$ and $\Delta Q_i^{V_j}(t)$, which are not readily available from PMU measurements. By making similar assumptions to the ones used in the derivation of (2.11), we express the net variation in net reactive power injection at bus i as

$$\Delta Q_i(t) \approx \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Delta Q_i^{\theta_j}(t) + \sum_{j \in \mathcal{V}_L} \Delta Q_i^{V_j}(t). \tag{2.12}$$

Now, by substituting (2.9) into (2.11), we can express (2.11) as

$$\Delta P_i(t) \approx \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Delta \theta_j(t) \Psi_i^j + \sum_{j \in \mathcal{V}_L} \Delta V_j(t) \Phi_i^j,$$

where $\Psi_i^j \approx \frac{\Delta P_i^{\theta_j}}{\Delta \theta_j}$ and $\Phi_i^j \approx \frac{\Delta P_i^{V_j}}{\Delta V_j}$. Analogously, by substituting (2.10) into (2.12), we can express (2.12) as

$$\Delta Q_i(t) \approx \sum_{j \in \mathcal{V}_G \cup \mathcal{V}_L} \Delta \theta_j(t) \Gamma_i^j + \sum_{j \in \mathcal{V}_L} \Delta V_j(t) \Lambda_i^j,$$

where $\Gamma_i^j \approx \frac{\Delta Q_i^{\theta_j}}{\Delta \theta_j}$ and $\Lambda_i^j \approx \frac{\Delta Q_i^{V_j}}{\Delta V_j}$.

Suppose $M+1$ sets of synchronized measurements are available. Let

$$\begin{aligned}
\Delta P_i[k] &= P_i((k+1)\Delta t) - P_i(k\Delta t), \\
\Delta Q_i[k] &= Q_i((k+1)\Delta t) - Q_i(k\Delta t), \\
\Delta \theta_i[k] &= \theta_i((k+1)\Delta t) - \theta_i(k\Delta t), \\
\Delta V_i[k] &= V_i((k+1)\Delta t) - V_i(k\Delta t),
\end{aligned}$$

$k = 1, \ldots, M$. Next, define $\Delta P_i = [\Delta P_i[1], \ldots, \Delta P_i[M]]^T$ and $\Delta Q_i = [\Delta Q_i[1], \ldots, \Delta Q_i[M]]^T$; similarly, define $\Delta\theta_i = [\Delta\theta_i[1], \ldots, \Delta\theta_i[M]]^T$ and $\Delta V_i = [\Delta V_i[1], \ldots, \Delta V_i[M]]^T$. Then, we obtain the following systems of equations:

$$\Delta P_i \approx \begin{bmatrix} (\Delta\theta_j)_{j\in\mathcal{V}_G\cup\mathcal{V}_L} & (\Delta V_j)_{j\in\mathcal{V}_L} \end{bmatrix} \begin{bmatrix} \Psi_i \\ \Phi_i \end{bmatrix}, \tag{2.13}$$

where

$$\Psi_i = \left[\left(\Psi_i^j\right)_{j\in\mathcal{V}_G\cup\mathcal{V}_L} \right] \text{ and } \Phi_i = \left[\left(\Phi_i^j\right)_{j\in\mathcal{V}_L} \right],$$

and

$$\Delta Q_i \approx \begin{bmatrix} (\Delta\theta_j)_{j\in\mathcal{V}_G\cup\mathcal{V}_L} & (\Delta V_j)_{j\in\mathcal{V}_L} \end{bmatrix} \begin{bmatrix} \Gamma_i \\ \Lambda_i \end{bmatrix}, \tag{2.14}$$

where

$$\Gamma_i = \left[\left(\Gamma_i^j\right)_{j\in\mathcal{V}_G\cup\mathcal{V}_L} \right] \text{ and } \Lambda_i = \left[\left(\Lambda_i^j\right)_{j\in\mathcal{V}_L} \right].$$

In (2.13)–(2.14), we assume that the relationship between ΔP_i and $[\Psi_i^T, \Phi_i^T]^T$ and the one between ΔQ_i and $[\Gamma_i^T, \Lambda_i^T]^T$ are approximately linear. Under this assumption, we seek the best estimate for $[\Psi_i^T, \Phi_i^T]^T$ and $[\Gamma_i^T, \Lambda_i^T]^T$ given the measured observations.

2.3 Problem Statement

Suppose the systems in (2.13)–(2.14) are overdetermined, i.e., $M > \bar{N} = N_G + 2N_L$. Then, a natural solution approach is to obtain Ψ_i, Φ_i, Γ_i, and Λ_i via LSE estimation. In ordinary LSE estimation, the regressor matrix is assumed to be free of error; hence all errors are confined to the observation vector (in our setting, ΔP_i or ΔQ_i). This assumption, however, is not entirely appropriate in our problem setting, since ΔP_i, ΔQ_i, $\Delta\theta_j$, and ΔV_j are all constructed from PMU measurements obtained in real-time. In such a case where modeling and measurement errors are associated with both the observation vectors and the regressor matrix, total least-squares (TLS) estimation is one appropriate method for fitting [4]. Next, we describe a TLS-based estimation algorithm as it applies to the solution of (2.13)–(2.14).

Chapter 3

Total Least-Squares Approach to Jacobian Estimation

In our setting, as described in Section 2.2, measurement and modeling errors enter into both the regressor matrix and the observation vectors in (2.13)–(2.14). In this section, we formulate the TLS estimation problem and its solution with respect to the system in (2.13) (the formulation with respect to the system in (2.14) is analogous). Further, for ease of notation, let

$$A = \begin{bmatrix} (\Delta\theta_j)_{j\in\mathcal{V}_G\cup\mathcal{V}_L} & (\Delta V_j)_{j\in\mathcal{V}_L} \end{bmatrix},$$

and also let $b_i = \Delta P_i$. Based on the expression above, we can rewrite (2.13) as

$$b_i \approx A \begin{bmatrix} \Psi_i^T & \Phi_i^T \end{bmatrix}^T. \tag{3.1}$$

Since (3.1) is an overdetermined system of equations, in the remainder of this section, we formulate the problem of computing $[\Psi_i^T, \Phi_i^T]^T$ in (3.1) as a TLS estimation problem. We note, however, that the ideas presented in this section are immediately applicable to estimate the unknown vectors in both systems described in (2.13)–(2.14).

3.1 Basic Total Least-Squares Approach

Before delving into the TLS estimation problem formulation and associated solution, we briefly describe the ordinary LSE problem formulation and its solution, as it applies to our setting. In ordinary LSE, since the regressor matrix is assumed to be error free, the rationale behind this estimation method is to correct the observations b_i as little as possible under the Euclidean norm metric; this can be formulated as an optimization program as follows (see, e.g., [4]):

$$\begin{aligned} &\min_{\hat{b}_i\in\mathbb{R}^M} ||\Delta b_i||_2, \\ &\text{s.t. } \hat{b}_i = A \begin{bmatrix} \Psi_i^T & \Phi_i^T \end{bmatrix}^T, \end{aligned} \tag{3.2}$$

where $\Delta b_i = b_i - \hat{b}_i$. Once a minimizer, $\hat{b}_i$, is found, then any $[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$ satisfying $\hat{b}_i = A[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$ is a LSE solution to (3.1). We assume A has full column rank; under this

Algorithm 1

Input: $b_i \in \mathbb{R}^M$, $A \in \mathbb{R}^{M \times \bar{N}}$.
Output: A vector $x \in \mathbb{R}^{\bar{N}}$

1: **Compute the SVD.**
$[A\ b_i] = [u_1, \ldots, u_M]\Sigma[v_1, \ldots, v_{\bar{N}+1}]^T$
2: **if** $v_{\bar{N}+1}^{\bar{N}+1} \neq 0$ **then**
3: **Set.** $x = -\frac{1}{v_{\bar{N}+1}^{\bar{N}+1}}[v_{\bar{N}+1}^1, \ldots, v_{\bar{N}+1}^{\bar{N}}]^T$
4: **else**
5: **Output.** Problem in (3.4) has no solution.
6: **Stop.**
7: **end if**

condition, the closed-form unique solution to (3.2) is (see, e.g., [5])

$$\begin{bmatrix} \hat{\Psi}_i \\ \hat{\Phi}_i \end{bmatrix} = \left(A^T A\right)^{-1} A^T b_i. \tag{3.3}$$

In contrast to the LSE problem formulation in (3.2), since TLS estimation accounts for errors in A as well, analogous to the vector Euclidean norm, its problem formulation seeks to minimize the matrix Frobenius norm, as follows:

$$\begin{aligned} \min_{[\hat{A}\ \hat{b}_i] \in \mathbb{R}^{M \times (\bar{N}+1)}} \quad & \left|\left|[\Delta A\ \Delta b_i]\right|\right|_F, \\ \text{s.t.} \quad & \hat{b}_i = \hat{A}\left[\Psi_i^T \quad \Phi_i^T\right]^T, \end{aligned} \tag{3.4}$$

where $\Delta A = A - \hat{A}$, $\Delta b_i = b_i - \hat{b}_i$, and $\bar{N} = N_G + 2N_L$ [4]. Then, once a minimizing $[\hat{A}\ \hat{b}_i]$ is found, then any $[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$ satisfying $\hat{b}_i = \hat{A}[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$ is a TLS solution to (3.1).

The solution to the TLS estimation problem in (3.4) relies heavily on the singular value decomposition (SVD) (see, e.g., [6]); below, we describe the procedure (its pseudocode is provided in Algorithm 1 (see, e.g., [7])). To obtain the solution to (3.4), we rewrite (3.1) as (see, e.g., [4])

$$\begin{bmatrix} A & b_i \end{bmatrix}\left[\Psi_i^T \quad \Phi_i^T \quad -1\right]^T \approx 0. \tag{3.5}$$

By using the SVD, we can write

$$\begin{bmatrix} A & b_i \end{bmatrix} = U\Sigma V^T, \tag{3.6}$$

where $U = [u_1, \ldots, u_M]$ and $V = [v_1, \ldots, v_{\bar{N}+1}]$ are unitary matrices, Σ is a diagonal matrix in which the diagonal elements σ_i are the singular values of $[A\ b_i]$ (see, e.g., [6]). If $\sigma_{\bar{N}+1} \neq 0$, then $[A\ b_i]$ has rank $\bar{N}+1$ and the unique solution to (3.5) is the zero vector. In order to obtain a nonzero solution to (3.5), the rank of $[A\ b_i]$ must be reduced to $\bar{N}$. According to the Eckart-Young-Mirsky low-rank matrix approximation theorem [8], the rank $\bar{N}$ approximation of $[A\ b_i]$, which minimizes the objective function in (3.4), is

$$\begin{bmatrix} \hat{A} & \hat{b}_i \end{bmatrix} = U\hat{\Sigma}V^T, \tag{3.7}$$

Table 3.1: WECC 3-machine 9-bus systems—measurement-based sensitivity factors obtained in Examples 2.

$\hat{\Psi}_5^2$	$\hat{\Psi}_5^3$	$\hat{\Psi}_5^4$	$\hat{\Psi}_5^5$	$\hat{\Psi}_5^6$	$\hat{\Psi}_5^7$	$\hat{\Psi}_5^8$
0.07631	0.05691	-11.03	0.03587	16.64	-0.06743	-0.02094
$\hat{\Psi}_5^9$	$\hat{\Phi}_5^4$	$\hat{\Phi}_5^5$	$\hat{\Phi}_5^6$	$\hat{\Phi}_5^7$	$\hat{\Phi}_5^8$	$\hat{\Phi}_5^9$
-5.7703	-2.236	-0.02054	2.557	-0.08047	0.03681	-2.017
$\hat{\Gamma}_5^2$	$\hat{\Gamma}_5^3$	$\hat{\Gamma}_5^4$	$\hat{\Gamma}_5^5$	$\hat{\Gamma}_5^6$	$\hat{\Gamma}_5^7$	$\hat{\Gamma}_5^8$
-0.05818	0.02315	-0.1354	0.02593	0.06445	2.519	-3.855
$\hat{\Gamma}_5^9$	$\hat{\Lambda}_5^4$	$\hat{\Lambda}_5^5$	$\hat{\Lambda}_5^6$	$\hat{\Lambda}_5^7$	$\hat{\Lambda}_5^8$	$\hat{\Lambda}_5^9$
1.363	-0.04853	-0.006372	0.02142	-13.75	23.33	-9.890

where $\hat{\Sigma}$ is a diagonal matrix in which the diagonal elements $\hat{\sigma}_i = \sigma_i$, if $i < \bar{N}+1$, and $\hat{\sigma}_i = 0$, otherwise. Since the approximate matrix $[\hat{A}\ \hat{b}_i]$ has rank $\bar{N}$, (3.5) has a nonzero solution. Based on properties of the SVD, $v_{\bar{N}+1}$ is the only vector that belongs to the null space of $[\hat{A}\ \hat{b}_i]$. Then, the TLS solution is obtained by scaling the vector $v_{\bar{N}+1}$ until its last component is equal to -1, namely,

$$\begin{bmatrix}\hat{\Psi}_i^T & \hat{\Phi}_i^T & -1\end{bmatrix}^T = -\frac{1}{v_{\bar{N}+1}^{\bar{N}+1}} v_{\bar{N}+1},$$

where $v_{\bar{N}+1}^{\bar{N}+1}$ denotes the $(\bar{N}+1)^{\text{th}}$ element of $v_{\bar{N}+1}$. Thus, the unique TLS solution to (3.1) is

$$\begin{bmatrix}\hat{\Psi}_i^T & \hat{\Phi}_i^T\end{bmatrix}^T = -\frac{1}{v_{\bar{N}+1}^{\bar{N}+1}} \begin{bmatrix}v_{\bar{N}+1}^1 & \cdots & v_{\bar{N}+1}^{\bar{N}}\end{bmatrix}^T. \tag{3.8}$$

Next, we illustrate the concepts introduced above.

Example 2 (3-Machine 9-Bus System) *In this example, we consider the same system as in Example 1. Here, we use Algorithm 1 to estimate the entries in each row of the power flow Jacobian matrix and compare the results to the benchmark values recorded in Table 2.1. In order to simulate PMU measurements of slight fluctuations in active and reactive power generated and consumed at each bus, we generate power injection (positive or negative) time-series data. To this end, we assume the active power injection at bus i at time instant k, denoted by $P_i[k]$, can be modeled as*

$$P_i[k] = P_i^0[k] + P_i^0[k]\nu_1^P + \nu_2^P, \tag{3.9}$$

where $P_i^0[k]$ is the nominal active power injection at time instant k, and ν_1^P and ν_2^P are pseudorandom values drawn from standard normal distributions with zero mean and standard deviations $\sigma_1^P = 0.1$ and $\sigma_2^P = 0.1$, respectively. Similarly, we assume the reactive power injection at bus i at time instant k, denoted by $Q_i[k]$, can be modeled as

$$Q_i[k] = Q_i^0[k] + Q_i^0[k]\nu_1^Q + \nu_2^Q, \tag{3.10}$$

where $Q_i^0[k]$ is the nominal reactive power injection, and ν_1^Q and ν_2^Q are pseudorandom values drawn from standard normal distributions with zero mean and standard deviations $\sigma_1^Q = 0.1$ and $\sigma_2^Q = 0.1$, respectively. In both (3.9) and (3.10), there are two random components added

Table 3.2: WECC 3-machine 9-bus systems—MSE of sensitivity factors obtained via Algorithm 1 compared to corresponding model-based benchmark.

$\hat{\Psi}_2$	$\hat{\Psi}_3$	$\hat{\Psi}_4$	$\hat{\Psi}_5$
9.289×10^{-5}	1.357×10^{-4}	0.01373	8.616×10^{-3}
$\hat{\Psi}_6$	$\hat{\Psi}_7$	$\hat{\Psi}_8$	$\hat{\Psi}_9$
7.446×10^{-3}	2.792×10^{-3}	0.01026	2.987×10^{-3}
$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\hat{\Phi}_4$	$\hat{\Phi}_5$
6.921×10^{-4}	9.754×10^{-4}	0.02228	0.04182
$\hat{\Phi}_6$	$\hat{\Phi}_7$	$\hat{\Phi}_8$	$\hat{\Phi}_9$
0.01153	6.452×10^{-3}	1.054×10^{-3}	7.142×10^{-3}
$\hat{\Gamma}_4$	$\hat{\Gamma}_5$	$\hat{\Gamma}_6$	$\hat{\Gamma}_7$
0.04260	0.01655	4.911×10^{-3}	1.857×10^{-3}
	$\hat{\Gamma}_8$	$\hat{\Gamma}_9$	
	5.044×10^{-3}	3.730×10^{-3}	
$\hat{\Lambda}_4$	$\hat{\Lambda}_5$	$\hat{\Lambda}_6$	$\hat{\Lambda}_7$
0.05839	5.933×10^{-3}	0.02212	0.01184
	$\hat{\Lambda}_8$	$\hat{\Lambda}_9$	
	1.219×10^{-3}	2.831×10^{-3}	

to the deterministic nominal quantities. The first component, $P_i^0[k]\nu_1^P$ in (3.9) ($Q_i^0[k]\nu_1^Q$ in (3.10)), represents the inherent fluctuations in active (reactive) power generation and load. The second component, ν_2^P in (3.9) (ν_2^Q in (3.10)), represents random measurement noise, which is independent of the nominal active (reactive) power injection values. For each set of bus injection data, we solve the power flow equations, with the slack bus absorbing all power imbalances, to obtain the voltage magnitude and phase "measurements".

In this example, we simulate 100 sets of power injection and voltage measurements with the same network topology and operating point. In Table 3.1, we report TLS estimates corresponding to the entries in Table 2.1. By visually comparing the model-based sensitivities and measurement-based estimates in Tables 2.1 and 3.1, we note that the measurement-based TLS estimation achieves values that are very close to the model-based benchmark values obtained by directly linearizing the power flow equations. Furthermore, we compute the mean-squared error (MSE) of each sensitivity vector and report them in Table 3.2; in this case, the average MSE is 0.01169. ■

3.2 Weighted Total Least-Squares Approach

One of the assumptions we make in (3.4) is that the Jacobian matrix sensitivity factors are approximately constant across the estimation time window. One way to eliminate this restriction and to obtain an estimator that is more adaptive to changes in operating point is to place more importance on recent measurements and less on earlier ones, which may be out of date. Again, before we delve into the WTLS estimation problem formulation, we briefly describe the ordinary weighted least-squares (WLS) estimation problem setting in which the objective function in (3.2) becomes

$$\min_{\hat{b}_i \in \mathbb{R}^M} \quad \left|\left| \sqrt{W} \Delta b_i \right|\right|_2, \tag{3.11}$$

where W is a positive definite symmetric matrix. The solution to (3.11) is given by (see, e.g., [5])

$$\begin{bmatrix} \hat{\Psi}_i \\ \hat{\Phi}_i \end{bmatrix} = \left(A^T W A\right)^{-1} A^T W b_i. \tag{3.12}$$

The idea is to choose appropriate values for W so that more recent measurements are weighted preferentially over past ones. If the elements of the error vector Δb_i are uncorrelated, then W is a diagonal matrix. The WLS estimation problem is often formulated using an exponential forgetting factor [9], in which the more recent measurements are preferentially weighted by setting $W[i,i] = f^{M-i}$ for some fixed $f \in (0,1]$, where f is called a "forgetting" factor.

In the WTLS estimation problem setting, the optimization in (3.4) becomes

$$\begin{aligned} \min_{[\hat{A}\ \hat{b}_i] \in \mathbb{R}^{M \times (\bar{N}+1)}} & F_0\left(\Delta A, \Delta b_i\right), \\ \text{s.t.} \qquad & \hat{b}_i = \hat{A}\begin{bmatrix} \Psi_i^T & \Phi_i^T \end{bmatrix}^T, \end{aligned} \tag{3.13}$$

with

$$F_0(\cdot) = \sum_{k=1}^{M} \Delta a[k] W_k \Delta a[k]^T + w_k \Delta b_i[k]^2, \tag{3.14}$$

where $\Delta a[k]$ denotes the k^{th} row of ΔA, $\Delta b_i[k]$ is the k^{th} element of Δb_i, and matrix W_k and scalar w_k represent weighting factors for elements in $\Delta a[k]$ and $\Delta b_i[k]$, respectively. Next, we discuss the selection of these weighting factors.

Choice of Weighting Factors. Inspired by ordinary WLS estimation, we set $w_k = f^{M-k}$, so as to weigh the more recent elements in the observation vector, b_i, more heavily. With regard to the choice of W_k's, first, we assume that the elements of the error vector $\Delta a[k]$ are uncorrelated; therefore the matrix W_k is diagonal. Furthermore, if measurements obtained at each bus are equally reliable, then the elements of $\Delta a[k]$ are equally weighted. Then, by employing the exponential forgetting factor, we set $W_k[i,i] = w_k$, for all i. With the above choices for w_k and W_k, if $f = 1$, then all measurements are given equal weighting, and the WTLS formulation in (3.13) is equivalent to the TLS one in (3.4). On the other hand, if $f < 1$, then earlier measurements would not contribute as much to the final estimate $[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$ as more recent ones. In this way, the WTLS formulation is useful if the system experiences a change in operating point during the measurement acquisition time window. With the weighting factors chosen as described above, we next describe the solution to the optimization problem in (3.13).

WTLS Problem Solution. Note that if $W_k = \mathbb{I}_{\bar{N}}$, where $\mathbb{I}_{\bar{N}}$ denotes an $\bar{N} \times \bar{N}$ identity matrix, and $w_k = 1$, for all $k = 1, \ldots, M$, then the formulation in (3.13) is equivalent to that in (3.4). Unlike the basic TLS problem, however, the WTLS problem does not have a SVD-based closed-form solution. In order to solve (3.13), we follow the development described in [10], which is summarized below. We first note that the equality constraint in (3.13) is equivalent to $b_i - \Delta b_i = (A - \Delta A)[\hat{\Psi}_i^T, \hat{\Phi}_i^T]^T$, i.e.,

$$\Delta b_i[k] = \begin{bmatrix} \Psi_i^T & \Phi_i^T \end{bmatrix} \left(\Delta a[k]^T - a[k]^T\right) + b_i[k], \tag{3.15}$$

Algorithm 2

Input: $b_i \in \mathbb{R}^M$, $A \in \mathbb{R}^{M\times\bar{N}}$; $W_k \in \mathbb{R}^{\bar{N}\times\bar{N}}$, $w_k \in \mathbb{R}$, $k = 1, \ldots, M$.
Output: A vector $x \in \mathbb{R}^{\bar{N}}$

1: **Initialize.** Set x_0 to previously known value, tolerance ϵ, and counter $p = 1$
2: **while** $F(x_p) > \epsilon$ **do**
3: **Iterate.** Compute new iterate x_p based on x_{p-1}
4: **Update.** $x = x_p$
5: **Set.** $p \leftarrow p + 1$
6: **end while**

for each $k = 1, \ldots, M$, where $a[k]$ denotes the k^{th} row of A. Substituting (3.15) into (3.14), we obtain the following unconstrained optimization problem:

$$\min_{\hat{A},[\Psi_i^T,\Phi_i^T]^T} F_u\left(\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix}^T, \Delta A\right), \tag{3.16}$$

where

$$F_u(\cdot) = \sum_{k=1}^{M} \Delta a[k] W_k \Delta a[k]^T + w_k \left(\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} \Delta a[k]^T - \begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} a[k]^T + b_i[k]\right)^2. \tag{3.17}$$

We note that $F_u(\cdot)$ is differentiable with respect to $\Delta a[k]$, for each $k = 1, \ldots, M$. Suppose ΔA^* is a local minimizer of (3.16). Then, according to first-order necessary conditions of optimality, at ΔA^* (see, e.g., [11, Chap. 11]), we have that

$$0 = \left.\frac{dF_u}{d\Delta a[k]}\right|_{\Delta a[k]=\Delta a^*[k]}, \ k = 1, \ldots, M,$$

from which we obtain

$$\Delta a^*[k]^T = \left[W_k + \begin{bmatrix}\Psi_i \\ \Phi_i\end{bmatrix}\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} w_k\right]^{-1} \times w_k \left(\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} a[k]^T - b_i[k]\right)\begin{bmatrix}\Psi_i \\ \Phi_i\end{bmatrix}, \tag{3.18}$$

for each $k = 1, \ldots, M$. By invoking the matrix inversion lemma (see, e.g., [6]), (3.18) simplifies to

$$\Delta a^*[k]^T = \frac{\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} a[k]^T - b_i[k]}{w_k^{-1} + \begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} W_k^{-1} \begin{bmatrix}\Psi_i \\ \Phi_i\end{bmatrix}} W_k^{-1} \begin{bmatrix}\Psi_i \\ \Phi_i\end{bmatrix}. \tag{3.19}$$

Finally, we substitute each optimal $\Delta a^*[k]$ as given in (3.19) into (3.17), from which we reformulate the optimization problem in (3.16) as

$$\min_{[\Psi_i^T,\Phi_i^T]^T} F\left(\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix}^T\right), \tag{3.20}$$

where

$$F(\cdot) = \sum_{k=1}^{M} \frac{\left(\begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} a[k]^T - b_i[k]\right)^2}{w_k^{-1} + \begin{bmatrix}\Psi_i^T & \Phi_i^T\end{bmatrix} W_k^{-1} \begin{bmatrix}\Psi_i \\ \Phi_i\end{bmatrix}}. \tag{3.21}$$

Through the development above (readers may refer to Appendix A for a full derivation), we convert the original constrained WTLS problem in (3.13) into the less troublesome unconstrained minimization problem in (3.20). The optimization problem in (3.13) (or (3.20)) is nonconvex; therefore, numerical solution methods do not guarantee convergence to a global minimum. Many numerical algorithms, most of which are iterative, have been proposed to solve (3.13) or (3.20) (see, e.g., [7], for an overview). For our case studies, we find that built-in optimization routines in MATLAB are sufficient as proof-of-concept to demonstrate the feasibility of the proposed power flow Jacobian estimation framework. In a commercial implementation of the proposed framework, it may be prudent to investigate convergence properties of various solution methods. We refrain from further discussion on this topic here as it is beyond the scope of the present work. The WTLS solution algorithm is summarized in Algorithm 2 for a generic iterative optimization scheme. Next, we illustrate the ideas presented above with an example.

Example 3 (3-Machine 9-Bus System) *We consider the same system as in Example 1 and simulate 200 sets of PMU measurements of slight fluctuations. In order to simulate an undetected change in operating point, without updating the model, the active load at bus 6 linearly increases by 1.6 p.u. over the span of 20 measurements beginning at $k = 80$, with the generation at bus 2 also increasing commensurately at each time step.*

As in Example 2, we compute the power flow, with the slack bus absorbing all power imbalances for each particular time k. Table 3.3 shows a comparison between benchmark sensitivity factors obtained via direct linearization of the power flow equations around the operating point (both before and after the change), and those obtained via the proposed WTLS framework with forgetting factors $f = 0.96$ and $f = 1$. Both measurement-based estimations are executed at $k = 200$ with the previous $M = 200$ measurements. Since the operating point is undetected by operators, under the pre-change system model, the power flow Jacobian matrix (some entries of which are shown in column 3 of Table 3.3) results in an average MSE of 0.2956. In column 5 of Table 3.3, we record results for WTLS with $f = 1$ (or,

Table 3.3: WECC 3-machine 9-bus systems—model- and measurement-based sensitivity factors obtained in Example 3.

	Model-based		Measurement-based	
	Post-change	Pre-change	$f = 0.96$	$f = 1$
Ψ_{52}^{2}	0	0	-0.5254	-0.2637
Ψ_{53}^{3}	0	0	0.08467	0.9688
Ψ_{54}^{5}	-9.685	-10.86	-9.599	-10.66
Ψ_{55}^{6}	0	0	-0.2559	0.1843
Ψ_{56}^{7}	14.60	16.54	14.97	15.52
Ψ_{57}^{8}	0	0	0.6563	0.2575
Ψ_{58}^{9}	0	0	-0.1905	0.4154
Ψ_{59}^{5}	-4.911	-5.6816	-5.137	-6.780
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Average MSE		0.2956	0.1979	0.4501

equivalently, TLS). From the average MSE metric, as well as a survey of the individual values, reported in this column, we note that the basic TLS scheme is unable to estimate the updated Jacobian matrix elements. On the other hand, the WTLS method with $f = 0.96$ *is able to track elements in the Jacobian matrix, as shown in column 4 of Table 3.3. However, compared to the SVD computation in basic TLS, the WTLS optimization incurs much higher computational burden; thus, the cost of better tracking is longer computation time and lack of optimality guarantee.*

By observing estimation results in both Tables 2.1 and 3.3, we note that while the TLS-based schemes are able to track the nonzero terms with sufficient accuracy, the resulting estimated signal is quite noisy, with many near-zero terms. ■

Chapter 4

Estimation with a Subset of Measurements

In the proposed Jacobian matrix estimation framework presented thus far, to estimate the unknown sensitivity factors, with respect to bus i, voltage magnitude and phase angle measurements are required from all buses. In other words, the framework necessitates a central data collector to whom all measurements are passed. Moreover, it requires at least as many time-sampled sets of measurements as the number of columns of the Jacobian matrix. Both of these restrictions become unwieldy for large-scale power systems. First, since power systems are constantly undergoing changes and operators often need to quickly determine the current system state, it would be ideal to obtain accurate estimates using fewer data sets. Furthermore, in a practical setting, the entire set of measurements may not be available for transmission to a central data collector.

In order to relax the restrictions described above, we note that, due to the structure of the power flow equations, the sensitivity factors Ψ_i^j, Φ_i^j, Γ_i^j, and Λ_i^j are only nonzero if $i = j$, or if there exists a transmission line connecting buses i and j. With this in mind, we define $\mathcal{N}_i$ as the set of buses that are connected to bus i, including bus i itself. Then, based on the full systems of equations in (2.13)–(2.14), we can obtain the following reduced systems of equations:

$$\Delta P_i \approx \begin{bmatrix} (\Delta\theta_j)_{j\in(\mathcal{V}_G\cup\mathcal{V}_L)\cap\mathcal{N}_i} & (\Delta V_j)_{j\in\mathcal{V}_L\cap\mathcal{N}_i} \end{bmatrix} \begin{bmatrix} \Psi_i^{\mathcal{N}_i} \\ \Phi_i^{\mathcal{N}_i} \end{bmatrix}, \tag{4.1}$$

where

$$\Psi_i^{\mathcal{N}_i} = \left[\left(\Psi_i^j\right)_{j\in(\mathcal{V}_G\cup\mathcal{V}_L)\cap\mathcal{N}_i} \right] \text{ and } \Phi_i^{\mathcal{N}_i} = \left[\left(\Phi_i^j\right)_{j\in\mathcal{V}_L\cap\mathcal{N}_i} \right]$$

are reduced sensitivity vectors that contain only the nonzero entries of Ψ_i and Φ_i, respectively; and

$$\Delta Q_i \approx \begin{bmatrix} (\Delta\theta_j)_{j\in(\mathcal{V}_G\cup\mathcal{V}_L)\cap\mathcal{N}_i} & (\Delta V_j)_{j\in\mathcal{V}_L\cap\mathcal{N}_i} \end{bmatrix} \begin{bmatrix} \Gamma_i^{\mathcal{N}_i} \\ \Lambda_i^{\mathcal{N}_i} \end{bmatrix}, \tag{4.2}$$

where

$$\Gamma_i^{\mathcal{N}_i} = \left[\left(\Gamma_i^j\right)_{j\in(\mathcal{V}_G\cup\mathcal{V}_L)\cap\mathcal{N}_i} \right] \text{ and } \Lambda_i^{\mathcal{N}_i} = \left[\left(\Lambda_i^j\right)_{j\in\mathcal{V}_L\cap\mathcal{N}_i} \right]$$

contain only the nonzero entries of Γ_i and Λ_i, respectively.

Similar to the full formulation in (2.13)–(2.14), we can obtain estimates of the reduced sensitivity vectors in (4.1)–(4.2) via Algorithms 1 and 2. Unlike the full formulation, however, to obtain estimates of these reduced sensitivity factors with respect to bus i, it suffices to acquire $M > 2(\#\mathcal{N}_i)$ sets of synchronized measurements,[1] thus reducing the computational burden involved. As a direct consequence of conducting computations at each bus, parallel processing can be utilized so that the full system topology and relevant parameters can be obtained quickly. The local topology information can be transmitted to a central controller periodically, or when the resulting estimates indicate an update is required.

[1] $\#\mathcal{A}$ denotes the cardinality of set $\mathcal{A}$.

Chapter 5

Case Studies

We use the proposed measurement-based approach to estimate the Jacobian matrix in the IEEE 118-bus system. The simulation tool MATPOWER [3] is used throughout to solve the power flow and generate voltage magnitude and phase angle measurements from pseudo-random bus injections generated using (3.9) and (3.10).

5.1 Base Case

We consider the base case model for the IEEE 118-bus system and assess the effectiveness of the proposed measurement-based method to estimate the power flow Jacobian matrix under constant nominal operating point. As in Example 2, we simulate bus injection data by adding noise to the nominal injections, as given in (3.9)–(3.10), with $\sigma_1^P = \sigma_1^Q = 0.03$ and $\sigma_2^P = \sigma_2^Q = 0.01$. For comparison, we obtain benchmark values by linearizing the power flow equations around the nominal operating point.

Measurements from All Buses. We utilize data from all buses and compute estimates for the elements of the power flow Jacobian by solving the full problems in (2.13)–(2.14). We assume the time window under consideration contains $M = 1000$ sets of synchronized measurements. Using Algorithm 1, in conjunction with simulated measurements from all buses, we obtain estimates of Ψ_i and Φ_i, for $i \in \mathcal{V}_G \cup \mathcal{V}_L$, as well as Γ_i and Λ_i, for $i \in \mathcal{V}_L$. When comparing these estimated vectors to their corresponding model-based benchmark values, we find that the mean MSE for all estimated vectors is 0.00497, with the maximum being 0.5090.

Measurements from a Subset of Buses. Suppose each bus is equipped with the computational capability required to conduct its own sensitivity estimation. Then, as described in Section 4, each estimation problem solves fewer unknown sensitivity factors and requires fewer sets of synchronized measurements. Therefore, we use the first $M = 40$ sets of measurements from the full-system Jacobian matrix estimation from above. We assume that each bus is able to attain voltage magnitude and phase measurements from its immediate neighbors. Using Algorithm 1, we solve for the unknown vectors in (4.1)–(4.2) for each $i = 1, \ldots, N$, and further compare them to corresponding model-based benchmark values. We find the mean MSE to be 0.001523, with the maximum being 0.1936.

5.2 Change in Topology

Under the reduced formulation presented in Section 4, we assess the performance of the proposed WTLS framework, as described in Section 3.2, to update the entries of the Jacobian matrix after a topology change. With respect to this, we simulate $M = 100$ sets of synchronous measurements by computing the power flow solution using power injection data generated via (3.9)–(3.10) with $\sigma_1^P = \sigma_1^Q = \sigma_2^P = \sigma_2^Q = 0.1$. To simulate a topology change, we introduce a credible line outage (i.e., one that does not island the system) at time step $k = 30$. As in Example 3, we use a forgetting factor of $f = 0.96$. We repeatedly simulate random sample paths with random line outages.

Overall, the proposed WTLS estimation method is able to adapt and obtain accurate estimates for 63.84% of the affected Jacobian matrix entries. Since the optimization problem in WTLS is nonconvex, iterative numerical solution methods may only attain a local minimum, as evidenced by the low estimation accuracy. In contrast, for the same random sample paths and forgetting factor, the WLS estimates, obtained via (3.12), are accurate for 84.72% of the affected entries.

Chapter 6

Conclusions

In this work, we presented a measurement-based method to estimate the power flow Jacobian matrix without relying on the system power flow model. The proposed method relies on the solution of an overdetermined set of linear equations constructed from real-time measurements obtained with PMUs installed throughout the system.

Via TLS estimation, we account for measurement errors in both the observation vector as well as the regressor matrix. We showed that the proposed method provides accurate estimates of the Jacobian matrix entries. Furthermore, we improve the adaptability of the proposed method by employing WTLS and WLS estimation.

Appendix A

Weighted Total Least-Squares Estimation

Consider an overdetermined system

$$y \approx Ax, \tag{A.1}$$

where $x \in \mathbb{R}^N$ is a vector of unknowns to be estimated, $y \in \mathbb{R}^M$ is a vector of measurements. In the classical least-squares estimation approach, the entries of the regressor matrix A are assumed to be free of error, i.e., all errors are confined to the observation vector y. This assumption, however, is frequently inaccurate. In Chapter 3, both the regressor matrix and observation vector are constructed from real-time measurements. Hence, we would like to account for sampling and modeling errors in both A and y. One way to do so is via the total least-squares (TLS) estimation framework, in which the following optimization problem is solved:

$$\begin{aligned} \min_{[\hat{A}\ \hat{y}] \in \mathbb{R}^{M \times (N+1)}} \quad & \left|\left| \begin{bmatrix} \Delta A & \Delta y \end{bmatrix} \right|\right|_F, \\ \text{s.t.} \quad & \hat{y} = \hat{A}x, \end{aligned} \tag{A.2}$$

where

$$\Delta A = A - \hat{A},\ \Delta y = y - \hat{y}. \tag{A.3}$$

Using the definition of the Frobenius norm, we can rewrite (A.2) as

$$\begin{aligned} \min_{[\hat{A}\ \hat{y}] \in \mathbb{R}^{M \times (N+1)}} \quad & \sum_{k=1}^{M} \Delta a[k] \Delta a[k]^T + \Delta y[k]^2, \\ \text{s.t.} \quad & \hat{y} = \hat{A}x, \end{aligned} \tag{A.4}$$

where $\Delta a[k]$ denotes the k^{th} row of ΔA, $\Delta y[k]$ is the k^{th} element of Δy. In (A.4), each set of measurement errors, which consists of $\Delta a[k]$ and $\Delta y[k]$, is weighted equally and assumed to be uncorrelated. This assumption may not be valid in all circumstances: sensor errors may be correlated, some sensors may be more accurate, and past measurements may be less reliable. One way to account for this is via a weighted total least-squares (WTLS) estimation framework. In this appendix, we derive the WTLS estimation optimization problem in (3.20)–(3.21).

In the WTLS estimation problem setting, the optimization in (A.4) becomes

$$\min_{[\hat{A}\ \hat{y}]\in\mathbb{R}^{M\times(N+1)}} \quad F_0\left(\Delta A, \Delta y\right) = \sum_{k=1}^{M} \Delta a[k] W_k \Delta a[k]^T + w_k \Delta y[k]^2, \tag{A.5}$$

$$\text{s.t.} \quad \hat{y} = \hat{A}x, \tag{A.6}$$

where matrix W_k and scalar w_k represent weighting factors for elements in $\Delta a[k]$ and $\Delta y[k]$, respectively.

Substituting (A.3) into the equality constraint in (A.6), we obtain

$$\begin{aligned} y - \Delta y &= (A - \Delta A)x, \\ \Delta y &= -(A - \Delta A)x + y, \end{aligned}$$

which leads to

$$\Delta y[k] = -(a[k] - \Delta a[k])x + y[k], \tag{A.7}$$

for each $k = 1, \ldots, M$, and where $a[k]$ denotes the k^{th} row of A. Rewriting (A.7), we obtain

$$\Delta y[k] = x^T \Delta a[k]^T - x^T a[k]^T + y[k]. \tag{A.8}$$

Next, we substitute (A.8) into (A.5) to obtain the following unconstrained optimization problem:

$$\min_{\hat{A},x} F_u\left(x, \Delta A\right) = \sum_{k=1}^{M} \Delta a[k] W_k \Delta a[k]^T + w_k \left(x^T \Delta a[k]^T - x^T a[k]^T + y[k]\right)^2. \tag{A.9}$$

We note that $F_u(\cdot)$ is differentiable with respect to $\Delta a[k]$, for each $k = 1, \ldots, M$. Suppose ΔA^* is a local minimizer of (A.9). Then, according to first-order necessary conditions, at ΔA^*,

$$0 = \left.\frac{dF_u}{d\Delta a[k]}\right|_{\Delta a[k]=\Delta a^*[k]}, \ k = 1, \ldots, M,$$

from which we obtain, for each $k = 1, \ldots, M$,

$$\begin{aligned} 0 &= 2\left[W_k \Delta a^*[k]^T + \left(x^T \Delta a^*[k]^T - x^T a[k]^T + y[k]\right) w_k x\right], \\ &= W_k \Delta a^*[k]^T + w_k \left(x^T \Delta a^*[k]^T\right) x - w_k \left(x^T a[k]^T - y[k]\right) x, \\ &= \left(W_k + x w_k x^T\right) \Delta a^*[k]^T - w_k \left(x^T a[k]^T - y[k]\right) x. \end{aligned}$$

Finally, after rearranging, we obtain

$$\Delta a^*[k]^T = \left[W_k + x w_k x^T\right]^{-1} w_k \left(x^T a[k]^T - y[k]\right) x, \tag{A.10}$$

for each $k = 1, \ldots, M$. By invoking the matrix inversion lemma, (A.10) becomes

$$\begin{aligned}
\Delta a^*[k]^T &= \left(W_k^{-1} - \frac{W_k^{-1} x x^T W_k^{-1}}{w_k^{-1} + x^T W_k^{-1} x} \right) w_k \left(x^T a[k]^T - y[k] \right) x, \\
&= \frac{1}{w_k^{-1} + x^T W_k^{-1} x} \left[W_k^{-1} \left(w_k^{-1} + x^T W_k^{-1} x \right) - W_k^{-1} x x^T W_k^{-1} \right] \\
&\qquad\qquad w_k \left(x^T a[k]^T - y[k] \right) x \\
&= \frac{1}{w_k^{-1} + x^T W_k^{-1} x} \left[W_k^{-1} \left(w_k^{-1} + x^T W_k^{-1} x \right) w_k \left(x^T a[k]^T - y[k] \right) x \right. \\
&\qquad\qquad \left. - W_k^{-1} x x^T W_k^{-1} w_k \left(x^T a[k]^T - y[k] \right) x \right] \\
&= \frac{1}{w_k^{-1} + x^T W_k^{-1} x} \left\{ W_k^{-1} \left(x^T a[k]^T - y[k] \right) x \right. \\
&\qquad \left. + \left[W_k^{-1} \left(x^T W_k^{-1} x \right) - W_k^{-1} x x^T W_k^{-1} \right] w_k \left(x^T a[k]^T - y[k] \right) x \right\}. \\
&= \frac{1}{w_k^{-1} + x^T W_k^{-1} x} \left\{ W_k^{-1} \left(x^T a[k]^T - y[k] \right) x \right. \\
&\qquad \left. + w_k \left(x^T a[k]^T - y[k] \right) \left[W_k^{-1} \left(x^T W_k^{-1} x \right) x - W_k^{-1} x \left(x^T W_k^{-1} x \right) \right] \right\}.
\end{aligned}$$

Noting that $x^T W_k^{-1} x$ is a scalar and therefore $W_k^{-1} x \left(x^T W_k^{-1} x \right) = W_k^{-1} \left(x^T W_k^{-1} x \right) x$, we obtain

$$\Delta a^*[k]^T = \frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} W_k^{-1} x. \tag{A.11}$$

Evaluating $F_u(\cdot)$ in (A.9) at $\Delta a[k] = \Delta a^*[k]$, we obtain the following optimization problem:

$$\min_x F(x) = \sum_{k=1}^{M} \Delta a^*[k] W_k \Delta a^*[k]^T + w_k \left(x^T \Delta a^*[k]^T - x^T a[k]^T + y[k] \right)^2. \tag{A.12}$$

Substituting (A.11) into (A.12), we simplify $F(x)$ as follows:

$$\begin{aligned}
F(x) &= \sum_{k=1}^{M} \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 (W_k^{-1}x)^T W_k W_k^{-1} x \\
&\qquad + w_k \left[x^T \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right) W_k^{-1} x - x^T a[k]^T + y[k] \right]^2, \\
&= \sum_{k=1}^{M} \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 x^T W_k^{-1} x \\
&\qquad + \frac{w_k}{\left(w_k^{-1} + x^T W_k^{-1} x \right)^2} \left[\left(x^T a[k]^T - y[k] \right) x^T W_k^{-1} x \right. \\
&\qquad \left. - \left(x^T a[k]^T - y[k] \right) \left(w_k^{-1} + x^T W_k^{-1} x \right) \right]^2, \\
&= \sum_{k=1}^{M} \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 x^T W_k^{-1} x \\
&\qquad + \frac{w_k}{\left(w_k^{-1} + x^T W_k^{-1} x \right)^2} \left[- \left(x^T a[k]^T - y[k] \right) w_k^{-1} \right]^2, \\
&= \sum_{k=1}^{M} \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 x^T W_k^{-1} x + \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 w_k^{-1}, \\
&= \sum_{k=1}^{M} \left(\frac{x^T a[k]^T - y[k]}{w_k^{-1} + x^T W_k^{-1} x} \right)^2 \left(w_k^{-1} + x^T W_k^{-1} x \right), \\
F(x) &= \sum_{k=1}^{M} \frac{\left(x^T a[k]^T - y[k] \right)^2}{w_k^{-1} + x^T W_k^{-1} x}.
\end{aligned} \tag{A.13}$$

Finally, replacing $F(x)$ in (A.12) with (A.13), the WTLS optimization problem reduces to the following unconstrained minimization:

$$\min_x F(x) = \sum_{k=1}^{M} \frac{\left(x^T a[k]^T - y[k] \right)^2}{w_k^{-1} + x^T W_k^{-1} x}. \tag{A.14}$$

References

[1] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, "Measurement-based estimation of linear sensitivity distribution factors and applications," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1372 – 1382, 2014.

[2] ——, "A sparse representation approach to online estimation of power system distribution factors," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1727–1738, July 2015.

[3] R. D. Zimmerman, C. E. Murillo-Snchez, and R. J. Thomas, "Matpower: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12 – 19, Feb. 2011.

[4] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*. The Society for Industrial and Applied Mathematics, 1991.

[5] F. Schweppe, *Uncertain Dynamic Systems*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1973.

[6] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1985.

[7] I. Markovsky and S. V. Huffel, "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283–2302, 2007.

[8] G. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrica*, pp. 211–218, 1936.

[9] L. Ljung and T. Söderström, *Theory and practice of recursive identification*. MIT Press, 1983.

[10] A. Premoli and M. Rastello, "The parametric quadratic form method for solving tls problems with elementwise weighting," in *Total Least Squares and Errors-in-Variables Modeling*, S. Van Huffel and P. Lemmerling, Eds. Springer Netherlands, 2002, pp. 67–76.

[11] I. Griva, S. G. Nash, and A. Sofer, *Linear and Nonlinear Optimization*. Society for Industrial and Applied Mathematics, 2009.