USF Tampa Graduate Theses and Dissertations   USF Graduate Theses and Dissertations

November 2022

Data-Driven State Estimation for Improved Wide Area Situational Awareness in Smart Grids

Md Jakir Hossain
University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd

Part of the Engineering Commons, and the Oil, Gas, and Energy Commons

Data-Driven State Estimation for Improved Wide Area Situational Awareness in Smart Grids

by

Md Jakir Hossain

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Mia Naeini, Ph.D.
Yasin Yilmaz, Ph.D.
Nasir Ghani, Ph.D.
Xinming Ou, Ph.D.
Kaiqi Xiong, Ph.D.

Date of Approval:
November 22, 2022

Keywords: Multi-Region, Kalman Filters, Information Sharing, Cyber-physical Stresses

Dedication

To my family and friends all who inspired and supported me in this journey.

Acknowledgments

It might be challenging to articulate how some people affect one's life in specific ways. This may explain why I have put off writing this section of my dissertation, in addition to the fact that I am naturally lazy. But now that the time has come, let's begin this challenging endeavor.

I must start by saying that Dr. Mia Naeini, my advisor, deserves all the credit for making this journey both successful and fun for me. I count myself fortunate to have seen firsthand Mia's passion for study, her unwavering commitment to her work, her lofty objectives, her sharp perceptions, and most importantly, her aptitude for deconstructing challenging issues. She has taught me a variety of things throughout our weekly interactions, and I sincerely appreciate how calm and patient she has been. Her prompt comments have helped me acquire confidence in my skills and have kept my research from deviating over the years. She has spent endless hours working on my writing and presentation skills, and I am appreciative of that. To put it briefly, I couldn't have asked for a better doctorate advisor.

I want to express my gratitude to everyone on my committee, especially Professors Yasin Yilmaz, Nasir Ghani, Xinming Ou, and Kaiqi Xiong, for their service and insightful contributions. I would like to extend a sincere appreciation to the outstanding personnel at the department of electrical engineering who worked quietly in the background to effectively filter out all the administrative turbulence so that I could focus solely on my studies. I would like to thank the members of my lab, Upama, Hasnat, and Shuvo, for contributing to the pleasant research environment. I have gained a lot from our conversations and will certainly miss our time together.

And last, Nafisa is the kind of spouse I could never have imagined having. Graduate school can be difficult and upsetting at times, but I can't even begin to count the number of times Nafisa has calmed all my concerns with her patient words of solace and logic, intermingled with her humorous antics. I cherish every second I spend with you, Nafisa. From our leisurely strolls in

the neighborhood to our adventures in Georgia to our fascination with the falls in Philadelphia, we still have a long way to go. Thank you for being my confidante and best friend, for never failing to make me laugh with your wit, humor, and sarcasm, for showing me your love, and for staying by my side.

Table of Contents

List of Tables

List of Figures

Abstract

Wide area situational awareness (WASA) in smart grids includes automatic monitoring, perception and detection of anomalies in these systems. The goal of WASA is to make smart grids aware of their physical and operational state for more effective operational decisions and control. As such, tracking the system's state or state estimation is one of the key objectives of WASA. The extensive integration of cyber elements into smart grids, such as large deployment of various monitoring and measurement devices, provides new opportunities to improve WASA. However, the tight coupling of power grids with cyber components introduces vulnerabilities to cyber and physical stresses.

State estimation is one of the key functions in WASA. The conventional state estimators have been widely deployed in utility control centers to help with monitoring the state of the system. However, traditional model-based state estimation methods do not adequately meet the real-time monitoring and accuracy requirements for smart grids. Many of the model-based state estimation techniques are based on steady-state analysis, which cannot be accurate for modern power systems due to highly dynamic and stochastic variations introduced by, for instance, distributed energy generations and fast-changing loads. The availability of large volume of measurement data in smart grids have opened new directions to complement the traditional state estimation techniques using data-driven state estimation methods. In this dissertation, data-driven state estimation techniques are developed to support the WASA functions, such as monitoring the state of the system and detecting cyber and physical stresses in the system. The presented data-driven and machine learning models include linear Minimum Mean Square Error (MMSE) estimation, Bayesian Multivariate Linear Regression (BMLR) combined with Auto-Regressive AR(p) process, and Kalman filters and Temporal Graph Convolution Neural Networks (T-GCNNs). In addition to the measurement data, the T-GCNN can learn the features in the non-Euclidean domain of the system's topology, which can capture the structures and interactions among the components of power grids. The performance

of the proposed techniques are evaluated using simulated power system measurement data under various normal and stressed scenarios.

Moreover, low latency data processing is important for real-time WASA in smart grids. Distributed and local processing of data is a promising strategy that can improve system monitoring tasks, as it satisfies the low latency requirements while avoiding the enormous overhead of transferring a huge volume of time-sensitive data to central processing units. Distributed data processing may improve the efficiency of many tasks and one such task is state estimation in power systems. In this dissertation, multi-region distributed state estimation is modeled and analyzed under various information sharing techniques among the regions. The regions in the system are defined based on physical distance and the correlation among the state of the components. Several data-driven and machine learning models for centralized and distributed state estimation are evaluated for the system with respect to the various ways of information sharing techniques. It is discussed that the multi-region distributed state estimation can achieve comparable performance to centralized techniques with reduce communication and computation cost.

Chapter 1: Introduction

The age of the internet and data has revolutionized the way people interact with information. Likewise, cyber-physical systems (CPS) are transforming the way people interact with engineered systems. Composed of sensor networks, computational hardware, and communication technologies, cyber-physical systems are platforms that monitor and control the interaction among physical and cyber systems. Smart grids are examples of such cyber-physical systems. Figure 1.1 represents the cyber and physical layers of a smart grid.
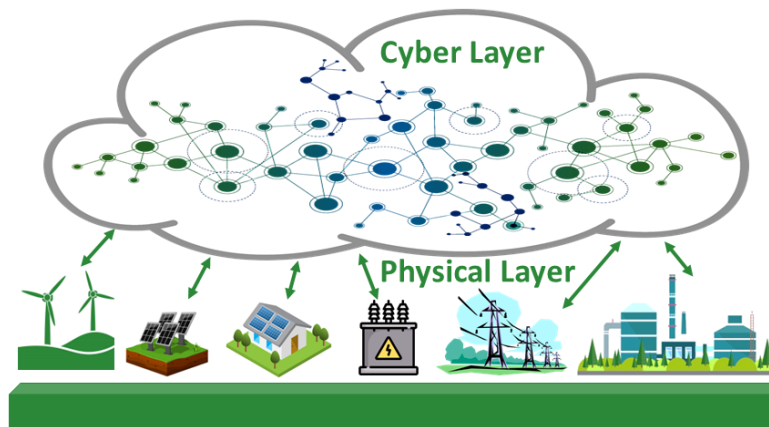
Figure 1.1: Cyber and physical layers of a smart grid.

As one of the most important cyber-physical systems, smart grids use advanced monitoring, control, and communication technologies to provide a dependable and secure energy supply while improving operational efficiency for electricity generation, transmission, and distribution [1]. The state of the physical system (the power network infrastructure) is controlled and monitored by the cyber system (including sensors, measurement devices, and communication elements) and fed to the control centers and functions as input to help in more efficient and reliable operation of the power systems. Moreover, timely responses are essential in the dynamic environment of smart

grids. Thus, real-time distributed computation and information processing of data streams in power systems are essential to help deliver timely decisions for operating these critical infrastructures.

Due to the increasing integration of various distributed components, including distributed energy resources (DERs) and electric vehicles (EVs), there is a growing demand to improve the reliability, energy efficiency, management, and security of the future smart grids considering the dynamic and stochastic nature of these systems. Moreover, smart grids are being equipped with various cyber components, including communication and computation elements, such as advanced metering infrastructure (AMIs) and measurement devices, including phasor measurement units (PMUs). These components are becoming more and more tightly integrated with power systems, forming an Internet of things (IoT) framework for data collection and processing. The sensing and measurement components are expected to generate a vast amount of data to support various applications in smart grids, such as distributed energy management, generation forecasting, grid health monitoring, fault detection, home energy management, wide area situational awareness (WASA) functions, state estimation, etc. With the vast amount of data from these new components, data-driven, machine learning, and artificial intelligence techniques are gaining more attention to automate and further improve the performance of smart grids [2].

In the light of the above, the focus of this dissertation is on developing and evaluating data-driven state estimation techniques both in centralized and distributed forms to support key functions in WASA for smart grids. In this chapter, a short review of the WASA and its main objectives are first discussed and then an overview of research challenges and contributions of this dissertation are presented.

1.1 Wide Area Situational Awareness

Situational awareness is important in a variety of fields, including business, military, aviation, and distributed engineered systems, such as smart grids. Situational awareness consists of three key components: perception, response, and restoration. The perception component is focused on data collection, sensing, monitoring, and descriptive analytics of data to track and monitor the state of

the system and also to identify vulnerabilities in the system. The response component is focused on functions that support the reliable and secure operation of the system, including cyber and physical stress detection and mitigation and demand and generation management supporting functions. The restoration component supports system improvements and operations after stresses, system planning and hardening and more. Providing the essential level of situational awareness in power systems is challenging and has been the focus of a large body of research [3, 4]. The fundamentals of situational awareness and failures and issues in control centers caused by insufficient situational awareness in power systems are discussed in more detail in [3].

As the name suggests, WASA is situational awareness on a comprehensive and large-scale implementation that makes smart grids aware of the situations in their geographically distributed physical components and systems to enable efficient decision-making, operation, and control of the system [5]. The information provided by WASA to operators and other control and management functions is essential for well-informed decision-making and the reliable and efficient operation of power systems. State estimation in smart grids is regarded as an essential part of the energy management system (EMS) that enables WASA. The performance of the state estimation affects many critical functions of smart grids.

1.2   Smart Grid State Estimation

State estimation is a numerical process to infer the state of system parameters from limited measurement data. This is a vital function to monitor the status of the system's components and to ensure efficient and reliable operation of the system. Power system state estimation exploits measurement data collected at Supervisory Control and Data Acquisition (SCADA) systems or through PMUs and smart meters to produce an estimation of the power system's status [6, 7].

Some of the primary functions of state estimators include topology processing, bad data detection, observability analysis, and state forecasting. State estimation takes the unprocessed measurements from meters and transforms them into state information for the system and its downstream applications, making it one of the most significant functions in EMS. Since the cyber and phys-

ical sides of contemporary power systems may not always be consistent with one another, state estimation also functions as a crucial intermediary that can filter inaccurate data and information produced by either side.

Historically, monitoring the power grid was challenging due to the small number of measurement devices compared to the size of the power grid. Even with the increased deployment of monitoring and sensing devices in today's power grids, traditional state estimation techniques using system models have to tackle challenges such as inaccurate system models, unbalanced operation, renewable energy integration, and cyber-security issues. As such, data-driven state estimation techniques have been pursued to complement traditional state estimation techniques. The data-driven state estimation techniques also need to address challenges such as communication and computation limitations (e.g., bandwidth, latency, security, accuracy, processing power, and data storage capacity) given the large scale of the smart grids. Moreover, state estimation algorithms aid in wide-area monitoring, detection, and locating anomalies at various levels of cyber and physical components of the smart grids. Thus smart grid state estimation is a critical function in enabling the situational awareness functionalities.

1.3 Research Challenge and Motivation

This section discusses the research scope and challenges in situational awareness and particularly the state estimation in smart grids considering data-driven approaches and cyber and physical stresses in the system.

1.3.1 Data-Driven State Estimation for Wide Area Situational Awareness

As discussed earlier, state estimation is one of the main functions of the WASA. The conventional state estimators have been widely deployed in utility control centers to help with monitoring the state of the system. However, traditional state estimation methods do not adequately meet the real-time monitoring and accuracy requirements for smart grids. Traditionally, power grids were not intended to house variable (renewable) distributed generations. However, new smart grid technologies have

made it conceivable. Specifically, utility-scale energy storage and new sensors and actuators that could observe and control the power grid at points ranging from generation through the transmission to distribution are examples of such technologies supporting distributed generations.

Many of the model-based state estimation techniques are based on steady-state analysis, which cannot be accurate for modern power systems due to highly dynamic and stochastic variations introduced by, for instance, distributed energy generations and fast-changing loads. Besides, the deployment of PMUs and the availability of a large volume of measurement data, introduce new opportunities for improving and complementing the conventional model-based state estimation in power systems. As such, in addition to conventional model-based state estimation, various data-driven state estimation methods have been developed and studied recently. Along these lines, the goal of the work in this dissertation is to evaluate new data-driven state estimation techniques that can enable tracking the state of the system with the required precision using the collected data from the system without the system model. Such data-driven approaches are particularly important when there are inaccuracies or missing information in the system model.

1.3.2   Distributed State Estimation

Many of the data processing and computations related to monitoring the power systems have been traditionally performed centrally on utility-owned servers or cloud platforms. Communicating the large volume of data and its processing in central units inevitably adds delay and inaccuracies in the state estimation and monitoring of the system. For certain time-sensitive functions, such delays could affect the reliability and stability of the system. To address such issues, distributed and local processing of data can provide a good solution, especially for time-sensitive functions. Data-driven state estimation is one of such function that can improve the response time of the system, particularly, to critical conditions, such as failures or cyber-stresses. The data-driven distributed state estimation can provide a faster and more accurate estimation of current conditions in the local regions as well as predicting future trends in state changes to identify abnormal conditions. New

technologies, such as Edge or Fog computing, can provide a platform to enable these functions by local and distributed processing of data to enhance system monitoring in power grids [8, 9].

While various data-driven state estimation approaches for smart grids, have been proposed and studied in the literature [10, 11, 12, 13, 14, 15], most of them are centralized techniques [10, 11, 13, 14, 15]. Examples of distributed state estimation studies include [12, 16, 17], where the state estimation has been implemented using various machine learning-based frameworks. To address this challenge, in this dissertation, various distributed data-driven state estimation techniques under various scenarios of information sharing among distributed regions are developed and studied. It is demonstrated that multi-region distributed state estimations can speed up the system's reaction time, especially in response to urgent situations like failures or cyber stressors [12]. The focus of this dissertation is specifically on interpretable and generalizable data-driven approaches for state estimation including Kalman filters with information sharing in multi-region settings.

1.3.3 Cyber Physical Stress Detection and Mitigation

Due to increased dependency on cyber components, smart grids exhibit new vulnerabilities to cyber threats. When cyber attacks occur jointly with physical attacks or failures in the power grid, they could have even more serious impacts and cause large-scale blackouts with severe societal and economic consequences [14].

Notable real-world incident

Date / Sector	Incident
2017, Business	Qatar news agency hacked
2017, Energy	Widespread electricity cuts across Istanbul, Turkey
2016, Business	The operation Ghoul in UAE
2016, Business	Shamoon 2, Saudi
2016, Defense	Police data licked, Turkey
2015, Energy	Ukrainian Power grid attack
2015, Energy	UAE, attacks on Energy companies
2015, Transport	Attack on Istanbul Airport passport control system
2013, Defense	Saudi Arabian Defense Ministry System
2012, Defense	Syrian Ministry of Foreign Affairs
2012, Energy	RasGas Attack, Qatar
2012, Transport	Maritime transport sector, Egypt
2012, Energy	Aramco, RasGas
2011, Defense	US drone captured, Iran
2011, Defense	Drone keylogging
2011, Water	Illinois Water Plant
2010, Transport	Web-based immobilization
2010, Water	South Houston water treatment
2010, Energy	San Bruno pipeline
2009, Defense	Drone video feed interception, Iraq
2009, Health	Arlington clinic HVAC
2009, Health	Worm on MRIs
2009, Energy	En. Fut. Holdings
2009, Energy	STUXNET
2008, Transport	Lodz Trams
2008, Energy	Hatch Nuclear Plant
2007, Transport	Alaska ATC
2006, Health	Epilepsy website
2006, Water	Sacramento River
2006, Transport	L.A. Traffic Engineers Strikes
2006, Water	Harrisburg water filtering Plant
2005, Transport	Railcorp Sydney
2002, Energy	Davis-Besse
2002, Transport	CSX Washington DC
2001, Transport	Port of Houston Dos
2001, Water	Maroochy, SCADA
1999, Energy	Bellingham Pipeline
1997, Transport	Worcester Airport
1994, Water	Salt River Project
1982, Energy	Siberian pipeline

Figure 1.2: Historical time-line of reported cyber and physical attacks on various infrastructures (energy infrastructure is indicated by red).

In the case of physical attacks or failures, the system's stability can be maintained if the Supervisory Control and Data Acquisition (SCADA) receives precise information about the status of the components and take proper action accordingly. If however, the flow of information is obstructed by a cyber attack, the status of the components will be unobservable to the SCADA, which prevents the control center from taking necessary and appropriate actions in a timely manner. Figure 1.2 presents the historical timeline of reported cyber-physical attacks, which is a clear

indication of ever-increasing threats and concerns on cyber-physical system's security, such as the smart grid security.

In this dissertation, a scenario of joint cyber and physical attack on the smart grid is considered. It is discuss how a data-driven method based on PMU data can help in recovering the status information of the components. Specifically, scenarios in which an attacker conducts a physical attack on the power system by disconnecting few transmission lines and simultaneously launches a cyber attack on the communication system and prevents the flow of information from the region around the physically attacked area or other regions of the system to the control center are considered. This joint cyber attack leads to unobservability on a portion of the power system, which has experienced line outages. The goal of the presented work is to use the PMU data from outside the attacked zone (observable parts of the system) to estimate the state of the lines in the attacked zone using a data-driven technique.

1.4   Key Contributions of this Dissertation

In this section, a summary of the key contributions of this dissertation is presented. A more detailed discussion of each of the following is presented in the following chapters. The contributions are categorized into three classes centralized data-driven state estimation, distributed state estimation, and state estimation for cyber-physical stress mitigation, following the discussions in the previous section.

1.4.1   Centralized Data-Driven State Estimation

- Interpretable and generalizable data-driven approaches for state estimation based on Kalman filters are presented. Both linear and non-linear Kalman filters, particularly extended and unscented Kalman filters, with data-driven models to approximate the power system dynamics are considered. These data-driven models make the approach independent of the system models, including the topology, admittance matrix of the system, and the physics of the electricity. Moreover, a Bayesian Multivariate Linear Regression (BMLR) approach

combined with the auto-regressive AR(p) process, and a Minimum Mean Squared Error (MMSE) estimator are also considered and evaluated for state estimation in smart grids. The performance of these models is compared with other data-driven models from the literature.

- A data-driven approach based on Graph Convolution Neural Networks (G-CNNs) is presented for state estimation in smart grids. The G-CNN can learn the features in the non-Euclidean domain of graphs, which can capture the structures and interactions among the components of power grids. By integrating the temporal dependencies in the time-series data, a temporal G-CNN (T-GCNN) is adopted for the data-driven state estimation problem. Specifically, a message-passing G-CNN is used to capture the topological structure of the smart grid and the gated recurrent units (GRU) are used to capture the dynamic variation of state information for temporal dependencies. The performance of this method is evaluated for cases of full measurement availability and availability of a subset of measurements.

1.4.2   Distributed Data-Driven State Estimation

- A wide range of distributed scenarios, from multi-region to fully distributed data-driven state estimation settings is considered, where the regions are defined based on the geographical distribution of the buses as well as their state correlations. The centralized data-driven techniques based on Kalman filters and BMLR with AR(p) process are adjusted and extended to be applied to the distributed scenarios. The performance of the presented data-driven distributed techniques is studied under various region definitions on the system.

- Various information-sharing techniques among distributed regions of state estimation are considered to improve the performance of the state estimation. These include defining regions with overlapped buses and incorporating auxiliary buses in each region for information sharing based on various statistical features of the neighboring regions. The performance of the presented data-driven distributed techniques is studied under various message-passing

settings and compared with various other data-driven and machine-learning-based state estimation techniques from the literature.

- The performance of the BMLR with AR(p) estimation model has been evaluated under normal operating conditions and partially unobservable scenarios. The partially unobservable scenarios can model cases in which the data from a subset of PMUs become unavailable due to, for example, cyber stresses such as the denial of service (DoS) attack or physical failures of the PMUs.

1.4.3   State Estimation for Cyber Physical Stress Mitigation

- A joint cyber-physical attack is considered in which an adversary damages some lines physically (physical attack) and prevents the information flow from the attacked zone to the control center to tamper the observability of the grid and mask the physical failure (cyber attack). The PMU data available from outside of the attacked zone is used to estimate the state of the components in the attacked zone. Specifically, a linear MMSE estimation is developed to estimate the state of the components in the attack zone in a centralized approach. An idea to extend the MMSE to an iterative process with feedback to improve the performance of estimation is also presented. The effects of various scenarios of attacks on the state estimation performance are evaluated.

1.5   Structure of this Dissertation

This dissertation is organized as following. Chapter 1 presents a general overview of the situational awareness, state estimation in smart grids. Chapter 2, presents an overview of related work in state estimation in smart grids from centralized to distributed state estimation as well as examples of work considering joint cyber physical stresses in smart grids. In chapter 3, centralized data-driven state estimation techniques have been discussed including centralized linear and non-linear Kalman filter models, MMSE, BMLR with AR(p) model, and T-GCNN model. Chapter 4

discusses the grid partitioning technique and multi-region distributed state estimation techniques along with information sharing strategies. A fully distributed Kalman filter have also been discussed as a special case of multi-region state estimation. Chapter 5 presents the state estimation for recovery of component states under a joint cyber and physical stress scenario. In Chapter 6 numerical results and performance evaluation of various proposed techniques have been presented under various scenarios of distributed estimation, information sharing, and normal operating conditions and partially unobservable scenarios. The performance of the techniques in comparison with existing work in the literature is also presented and discussed in this chapter. Chapter 7 presents an overall summary and future direction of the research.

1.6 Publications Resulting from this Dissertation

[1] Hossain, Md Jakir, and Mia Naeini, "Multi-Area Distributed State Estimation in Smart Grids Using Data-Driven Kalman Filters" Energies 15, no. 19: 7105, 2022.

[2] M. J. Hossain and M. Rahnamay–Naeini, "State Estimation in Smart Grids Using Temporal Graph Convolution Networks," 2021 North American Power Symposium (NAPS), pp. 01-05, 2021.

[3] M. J. Hossain and M. Rahnamay-Naeini, "Data-Driven, Multi-Region Distributed State Estimation for Smart Grids," 2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), pp. 1-6, 2021.

[4] M. J. Hossain and M. Rahnamay-Naeini, "Line Failure Detection from PMU Data after a Joint Cyber-Physical Attack," 2019 IEEE Power & Energy Society General Meeting (PESGM), pp. 1-5, 2019.

1.7 Summary

State estimation in smart grids is regarded as an essential part of the EMS that enables wide-area monitoring of the system. The performance of the state estimation affects many critical functions of smart grids. However, inaccurate or unavailable system models can challenge conventional power system state estimation. As such, data-driven state estimation techniques have been pursued to

complement the conventional approaches by exploiting the large volume of available energy data. The data-driven approaches can also face challenges due to data communication and processing delays as well as missing or inaccurate data due to, for instance, measurement/sensor failures and cyber threats. Hence, developing effective mechanisms to improve the state estimation and the related communication and computation delays and enabling the recovery of missing information is crucial for enhancing the power system's reliability and security. As such, the focus of this dissertation is on developing and evaluating data-driven state estimation to improve the WASA. The key contributions of this work are briefly discussed and more details on each will be discussed in the following chapters.

Chapter 2: A Review of the State of the Art in State Estimation in Power Systems

2.1 Introduction

State estimation is a crucial component of the EMS that enables wide-area system monitoring in smart grids. The performance of the state estimation affects a variety of essential smart grid functions. Inaccurate or unavailable system models can, however, provide a challenge to conventional power system state estimation. Consequently, data-driven state estimate strategies have been developed to supplement conventional approaches by utilizing the vast amount of energy data that is readily available. Data-driven state estimation may also face challenges, such as data transfer and processing delays and missing or erroneous data due to, for example, measurement sensor failures and cyber stresses. Developing effective ways to enhance state estimation and the associated communication and computation delays, as well as permitting the recovery of missing information, is vital for boosting the reliability and security of the power systems. In this chapter, a review of the state-of-the-art techniques in state estimation in power systems, ranging from conventional techniques to data-driven approaches in centralized and distributed frameworks, is presented. Moreover, state estimation under cyber and physical stresses has also been discussed briefly.

Various data-driven state estimation methodologies for smart grids have been developed and explored in the literature [10, 11, 12, 13, 14, 15, 18], the majority of which are centralized techniques [10, 11, 13, 14, 15]. However, real-time, wide-area smart grid monitoring requires low-latency data processing. Historically, a significant amount of data processing and computations related to power system monitoring have been performed centrally on utility-owned servers or in the cloud. Smart grid state estimation and monitoring can be hampered by delays and inaccuracies due to the huge amount of data that needs to be delivered and processed in centralized units. Such delays may affect the system's stability and reliability for some of the time-sensitive operations.

Distributed and local data processing solutions are new approaches that are being pursued to address the aforementioned challenges. They can improve system monitoring capabilities by satisfying low-latency requirements and reducing the cost of transferring a huge volume of time-sensitive data to central processing units.

Examples of distributed state estimation studies include [12, 16, 17, 18]. This chapter discusses several distributed data-driven state estimate strategies in a variety of scenarios, including the sharing of information between distributed regions. It is shown that multi-region distributed state estimations can improve the system's response time, particularly in response to severe scenarios, such as system failures or cyber stressors [12]. New technologies, such as edge or fog computing, can enable these functions with the aid of local and distributed data processing in power grids [8, 19]. In the following subsections, more detailed discussions of the state estimation techniques in power systems are presented. The reviews are organized under three main subsections, including centralized techniques, distributed techniques, and state estimation under cyber and physical stresses.

2.2   Centralized State Estimation Approaches

In centralized state estimation techniques, it is generally assumed that the state information and data from the measurement sensors (such as PMUs) are being processed in a centralized model. The majority of the conventional and data-driven state estimation for power systems are centralized, which will be discussed next.

2.2.1   Conventional Centralized State Estimation

Conventional power system state estimation techniques have been in use for a long time in power systems. Such techniques rely heavily on power system models, including the connectivity, attributes, and operating conditions of the components. A detailed review of various methods for model-based power system state estimation can be found in [20, 21]. Power system state estimation has traditionally been carried out using model-based static approaches, such as the weighted least

square (WLS)-based method [22, 23]. As the least squares approaches have analytical solutions for linear systems they were widely used for this purpose. Specifically, for the ease of use and quick convergence, WLS-based algorithms are most commonly used state estimation technique in power systems. Also, possess features that are optimal for certain kinds of error distributions, such as Gaussian distribution, is one further significant reason why they are so widely used. Apart from the WLS algorithm, other state estimation techniques, such as Least Absolute Value (LAV) [24, 25], Weighted Least Absolute Value (WLAV) [26], and Schweppe Hubber Generalized M (SHGM) [27] and their variations have been introduced in the literature.

The power grid's state vector can be defined as a set of variables, which describe the state of its components. By tracking the state variables other electric power quantities can be monitored and tracked as well. Many attributes of power grid's components, such as bus voltage magnitudes and phase angles, real and reactive power injections and current attributes can be used to define the state vector. The state estimation problem in general can be formulated as an overdetermined system of nonlinear equations to be solves as an unconstrained optimization problem. Let $\underline{z} \in \mathbb{R}^{m \times 1}$ and $\underline{x} \in \mathbb{R}^{n \times 1}$ denote the vector of meter measurements and state vector, and m is the number of samples, n is the number of states, respectively. The measurement equation is $\underline{z} = h(\underline{x}) + \underline{e}$, where $h(\underline{x}) \in \mathbb{R}^{m \times 1}$ is the nonlinear function relating the system states to the measurements and $\underline{e}$ is the vector of the measurement errors. The residual can be defined as $\underline{r} = \mathbf{w}(h(\underline{x}) - \underline{z})$, where $\mathbf{w} \in \mathbb{R}^{m \times n}$ is the diagonal weight matrix. The objective function of the general state estimation formulation, $J(\underline{x})$, for some of the popular techniques are summarized in Table 2.1.

In a comparative study [28], the accuracy of WLS, WLAV, and SHGM estimators were assessed based on three criteria: bias, consistency, and quality. It has been demonstrated that the WLS algorithm has superior performance in these three criteria. The popular WLS state estimation in power systems is an iterative normal equations (NE) method. The NE method can handle zero injections by assigning the zero injection equations larger weights. However, the weights might not be well-conditioned, which would hurt the convergence. Many alternative solution methods have been proposed to improve the accuracy and convergence of this technique. For example, in [29], an

Table 2.1: The general formulation of conventional state estimation optimization.

Method	General Formulation
WLS	$\text{minimize}_{\mathbf{x}} \ J(\underline{x}) = \sum_{i=1}^{m} r_i^2$
LAV	$\text{minimize}_{\mathbf{x}} \ J(\underline{x}) = \sum_{i=1}^{m} \lvert r_i \rvert,$ here $\lvert . \rvert$ is the L_1 norm
WLAV	$\text{minimize}_{\mathbf{x}} \ J(\underline{x}) = \sum_{i=1}^{m} w_i \lvert r_i \rvert$
SHGM	$\text{minimize}_{\mathbf{x}} \ J(\underline{x}) = \sum_{i=1}^{m} \lvert s_i \rvert$ $s_i = \begin{cases} \frac{r_i^2}{2} & if \ \lvert r_i \rvert \leq \alpha w_i \\ \alpha w_i \lvert r_i \rvert - \frac{1}{2}\alpha^2 w_i^2 & otherwise \end{cases}$ where α is the tuning parameter

alternative solution approach is proposed based on orthogonal transformations to alleviate the ill-conditioning problems. Moreover, the formulation of WLS with equality constraints [30] has been suggested to handle zero-injections. A method based on the direct elimination of variables using the equalities is also suggested in [31]. Hachtel's augmented matrix method for solving least squares problems is used in [32]. When Hachtel's method is compared to two different NE approaches, it is found to significantly improve numerical stability, which is crucial for ill-conditioned systems. Hachtel's method takes less computing time than the standard NE method but requires slightly more memory. In another comparative study, the NE methods, the orthogonal transformation method, the NE with added constraints techniques, and Hachtel's augmented matrix technique have been studied and their performance have been compared [33]. Comparisons are performed based on three factors: (i) computational effectiveness, (ii) implementation complexity, and (iii) numerical stability. According to the theoretical study, the orthogonal transformation technique is numerically the most stable. However, the effective fast decoupled version does not support the orthogonal transformation method. It is demonstrated that both Hachtel's method and the hybrid NE methods offer good trade-offs between computational effectiveness and numerical stability.

Dynamic model-based state prediction and state forecasting have also been in use in power system state estimation. Family of Gaussian filtering techniques, such as variations of Kalman filters (KF) are extensively studied for power system state estimation problems. The popular

variations are linear-KF [34], extended Kalman filter (EKF) [35, 36, 37, 35], unscented Kalman filter (UKF) [38, 39], and Particle Filter (PF) [40]. In these approaches, power system dynamics are modeled as a system of nonlinear equations as

$$
\begin{aligned}
\underline{x}_{t+1} &= f(\underline{x}_t, \underline{\omega}_t) \\
\underline{z}_t &= h(\underline{x}_t, \underline{\delta}_t),
\end{aligned}
\tag{2.1}
$$

where $\underline{x}_t$ is the state vector, $\underline{z}_t$ is the measurement vector, $\underline{\omega}_t$ is the vector of white Gaussian noise, $\underline{\delta}_t$ is the vector of measurement noise at time instance t, and $f(.)$ and $h(.)$ are vector-valued nonlinear functions describing the system and state equations. The size of the aforementioned vectors are $1 \times N$, where N is the number of buses in the system. The state of each bus at time t, $x_{i,t}$, can be represented by its attributes such as $v_{i,t}$ or $\theta_{i,t}$, where $v_{i,t}$ and $\theta_{i,t}$ are the voltage and phase angle at bus i at time t, respectively. The measurements at bus i at time t, $z_{i,t}$, can represent the attributes, such as $\mathbf{p}_{i,t}, \mathbf{q}_{i,t}, v_{i,t}, \theta_{i,t}$, where $\mathbf{p}_{i,t}$ and $\mathbf{q}_{i,t}$ are real and reactive power injections, respectively. The goal of the state estimation here is to estimate the vector $\underline{x}_{t+1}$ given the measurement $\underline{z}_t$.

As discussed in [41], traditionally, the state estimation has been solved using dynamic complex power-flow equations, such as

$$
\begin{aligned}
\mathbf{p}_{i,t} &= \sum_{j=1}^{N} |v_{i,t}||v_{j,t}||Y_{ij}| \cos(\theta_{i,t} - \theta_{j,t} - \Delta_{ij}) \\
\mathbf{q}_{i,t} &= \sum_{j=1}^{N} |v_{i,t}||v_{j,t}||Y_{ij}| \sin(\theta_{i,t} - \theta_{j,t} - \Delta_{ij}).
\end{aligned}
\tag{2.2}
$$

Here, i and j are indices of buses in the system, Δ_{ij} is the complex power angle, and Y_{ij} is the ijth entry of the admittance matrix $\mathbf{Y}$ ($|.|$ represents the absolute value). One of the challenges of this approach for state estimation is the need for the system model, for instance, $\mathbf{Y}$. The feasibility of applying KF-based techniques in dynamic state estimation problems has been extensively studied in [42]. A comparative study on the variations of the KF-based dynamic state estimation algorithm has also been studied in [43, 44].

2.2.2 Data-Driven Centralized State Estimation

The conventional and static state estimators need accurate system models and, in some cases, may not allow for predicting the future operating points of the system. With the rapid deployment of the PMUs, tracking the real-time dynamics of power system states is becoming possible. Thus, power system dynamic state estimation has become a scorching topic in recent years [45, 21, 46]. The accessibility of a significant number of measurement data opens up new possibilities for enhancing and complementing the traditional model-based state estimate in power systems. In addition, the robustness of data-driven approaches to system and topology changes, as well as missing or inaccuracies in system information, has made them the focus of many researchers recently [10, 11, 12, 13, 14, 15]. Many of the centralized state estimation techniques are focused on addressing the challenges regarding data quality issues, such as non-Gaussian measurement noise, bad data, cyber stresses, missing data, and frequent topology changes. Examples of such techniques based on forecasting-aided and predictive information filters are [14, 15, 47, 48, 49, 50].

Examples of data-driven and machine-learning-based state estimation techniques in power systems include techniques based on neural networks [51, 13], the Bayesian approach [12, 52], minimum mean squared error (MMSE) [14, 50], and the Kalman filter (KF) [11, 53]. Since KF and its variations, including the extended Kalman filter (EKF) and the unscented Kalman filter (UKF), have a great ability to model and express dynamic systems, many dynamic state estimation algorithms for power systems are developed based on these techniques [11, 54, 46, 55, 56].

Specifically, EKF [57], UKF [58], Cubature Kalman Filter (CKF) [59], Particle Filter [60], and Gaussian Mixture Filter [61] are examples KF variations that have been applied to the state estimation problem in power systems. The state estimation methods based on regression-based optimization using past measurements [14] and instantaneous correlations [47] of the measurements have also been proposed in the literature.

With the availability of a vast amount of data from smart grid sensors, neural network-based techniques are becoming increasingly popular in solving critical operations of smart grids [2].

18

For instance, variants of artificial neural networks (ANN) [62], such as recurrent neural network (RNN) [63], Long short-term memory (LSTM) [64], Residual Neural Network (ResNet) [65], and Convolutional Neural Network (CNN) [66] have been adopted for solving data-driven forecasting-aided state estimation in smart grids. In the later techniques, measurement data are fed as the input to the models either as multivariate time series or images in Euclidean space. The aforementioned techniques do not explicitly consider the information about the underlying graph structures in the system data. Particularly, CNN generally handles data in Euclidean space and fails to address non-Euclidean spaces created by graphs; especially when there is no spatial locality due to the arbitrary structure of the graph [67].

To incorporate graph information, such as information on the topological structure of the power system in the model, graph-aware machine-learning approaches have also been considered in the literature for state estimation in power systems. Next, we briefly review state estimation techniques that consider graph or topological information of the power system in their model. G-CNN, first introduced in [68], extends the existing neural network methods for processing the data represented in graph domains. The underlying structures and graph of connections among the components of power grids suggest the potential of G-CNN to capture and use such information. Applications of G-CNN to smart grid problems are emerging [69]. For instance, in [70], the authors have modeled the fault localization problem as a graph search approach using G-CNN in the distribution grid. The authors in [71] proposed a quasi-Monte-Carlo method based on G-CNN to calculate the distribution characteristics of system power flow. The problem of intentional islanding considering load-generation balance has also been addressed using G-CNN in [72]. A physics-aware graph-pruned neural network model for distribution grid state estimation has been proposed in [73]. Joint detection of false data injection attacks in smart grids using auto-regressive moving average (ARMA) graph convolutional filters on G-CNN has been proposed in [74]. In addition, in the G-CNN domain, both graph spectral filtering CNN and message-passing neural networks are developed [75]. In this dissertation, both data-driven and machine-learning models, with and without graph consideration, have been studied.

2.3 Distributed State Estimation Approaches

Smart grids are large-scale infrastructures that are distributed over large geographical regions. Transmitting the large volume of data collected by thousands of sensors to a centralized unit for processing imposes new challenges for the communication and computation systems as well as satisfying time-critical operations. To address these challenges and to improve wide-area situational awareness, distributed processing of data and computation of the system's state are promising solutions. Distributed solutions were pursued since the statistical formulation of state estimation [6, 76, 77], when it was realized that for a chain of serially interconnected regions, distributed updates could be invoked incrementally in space. In distributed state estimation, each subsystem tracks the state of its components using local measurements and limited communications with its neighboring subsystems or a central unit. Distributed state estimation may also serve the regulations in information sharing in each region since it limits the information sharing among the regions. Distributed state estimation is becoming increasingly popular, and a large body of work is emerging in this direction.

2.3.1 Conventional Distributed State Estimation

Traditionally, state estimation in power grids has been done in a centralized framework. With grid deregulation, increased awareness of information privacy and security, and more strict response time requirements, more attention has been given to multi-area state estimation. A detailed review of conventional techniques for multi-region distributed state estimation can be found in [78].

In this direction, hierarchical, multi-level approaches with global coordinators have been proposed in the literature. These approaches involve breaking down the overall system into subsystems and performing a two-level or multi-level state estimation computation in each level. Specifically, state estimation is first performed in parallel for all subsystems in the lower levels of the hierarchy. The coordination of these local estimates is then realized at an upper level of the hierarchy. This way the higher levels receive and process a smaller number of variables to form a global state estima-

tion [79]. In [80, 81], a comprehensive review of hierarchical distributed estimation techniques have been presented, and several renditions of this approach can also be found in [80, 81, 82, 83, 84, 85].

The goal of the hierarchical scheme is to extend the conventional techniques in a multi-region setting so that they can be applied to large-scale power systems. For instance, in [82], a fast rectangular-coordinate method is proposed, which involves an extension of a fast second-order load-flow method, and allows a fixed Jacobian matrix to be used in the hierarchical algorithm for state estimation. In [83], a stochastic fractal search technique (SFS) is used to perform the local/region-level state estimation and a simulated annealing technique is used at the global coordination level for state estimation. KF-based state estimation techniques are also proposed in a hierarchical fashion. For example, EKF for dynamic state prediction and hierarchical filtering for power systems have been proposed in [86]. A comprehensive analysis for the experimental implementation of KF-based hierarchical state estimation has also been performed in terms of computation power and communication bandwidth requirements in [87]. The author has also compared central, distributed, and hierarchical state estimation algorithms in terms of time complexity and concluded that hierarchical Kalman filtering (HKF), needs about 34% communication bandwidth and $O(1/N^3)$ computation power in subsystems compared to central state estimation, while giving approximately the same level of estimation precision. Moreover, most of these hierarchical techniques assume local observability, meaning that local state estimates, excluding the boundary bus measurements, are uniquely identifiable. Such an assumption may not hold in a real-world scenario due to missing or inaccurate information in the system. Moreover, the need for a central coordinator can also hinder the system's reliability.

To ease the computational and communication burden, another category of distributed state estimation techniques are the ones, which do not assume any centralized coordinator. Examples of such techniques include decentralized state estimation solutions, such as block Jacobian iterations [88, 89], gradient-based WLS solutions [90], and approximate algorithm modifying the conventional state estimation algorithms and applying a coupling constraints optimization technique [91]. These methods generally assume local observability; however, convergence is not always guaran-

teed in these techniques. A novel decentralized parallel optimization algorithm called auxiliary problem principle (APP) is proposed in [92]; however, it requires tuning of several parameters for optimal performance. In [93], a spectral graph theory-based algorithm is proposed for this purpose. The requirement for local observability is relaxed and the communication structure can be different from the physical topology of the power network. However, each region is envisioned to maintain a copy of the entire high-dimensional state vector.

Apart from popular distributed KF-based methods [94, 95] and distributed block Jacobian approach, distributed and parallel optimization-based method [96, 97, 98, 99, 100] have also been proposed for distributed state estimation. The key advantages of such approaches are the decomposition and parallelization of the original problem and the absence of a central computation unit. In [98], the authors proposed a bi-linear state estimation technique. In both linear stages, the state estimation problem in each area is solved locally, with minimal data exchange with its neighbors. The intermediate nonlinear transformation can be performed by all areas in parallel without any need for inter-regional communication. In [99], an MMSE-based approach has been proposed and a finite-time average consensus algorithm is utilized in conjunction with an influence function to realize the proposed distributed approach within a totally distributed framework. A distributed incremental Quais-Newton (D-IQN) algorithm has also been proposed for multi-region distributed state estimation in [100], where the maximum correntropy criterion (MCC) is used in the objective function in order to address the non-Gaussian noise scenario. In [101], a distributed innovation-based model is proposed. The authors have approached the distributed model solution of the state-of-the-art WLS through an Alternating Direction Method of Multipliers (ADMM), which requires minimal information communication. Though some of the conventional approaches are computationally efficient, some criteria, such as local observability, an accurate system model, the need for global coordination, and stochastic and non-gaussian noise, may limit the performance and applicability of such methods. Data-driven, distributed solutions may become a remedy to address some of the issues faced by conventional approaches.

2.3.2 Data-Driven Distributed State Estimation

Due to the constantly increasing complexity of power systems and higher penetration of stochastic renewable generation, as well as increased deployment of sensing and monitoring devices, distributed state estimation is gaining more attention in these systems. As such, the development of novel techniques to exploit the dynamic and large volume of energy data, along with increasing their robustness, are important research problems to tackle. The focus of some of the distributed data-driven and machine learning-based models for state estimation is to approximate the power system dynamics in a region and use such approximation for state estimation. Moreover, distributed state estimation in power systems is considered in a wide range of scenarios, including fully distributed, multi-region distributed, and multi-region distributed with collaborations among the regions. Regions in such techniques are usually defined to consist of buses and their connections based on specific properties, such as their geographic proximity, state correlations, or other data-driven or physics-based metrics of interactions among the buses [102]. A fully-distributed scenario is a special case of a multi-region scenario, where each area has a single bus as the member. A fully-distributed state estimation without information sharing/collaboration would mean a univariate estimation.

An example of a fully distributed state estimation can be found in [103], which uses predictive information filtering. Moreover, the work in [104] uses diffusion-based KF for fully distributed state estimation in which a selected set of nodes in the system are allowed to share a subset of intermediate estimates with their neighbors using information propagation strategies. A Gaussian process-based distributed Bayes filter has been proposed in [105] for the dynamic distributed state estimation. Examples of techniques that propose state estimation over regions in power systems include [12, 106, 107, 108, 109, 18, 110]. In these examples, regions have been defined according to a number of criteria, including geographic distance, operational similarity, and communication resources [106]. The work in [8], for instance, addresses an early stress detection and locating method based on a linear predictive filter that may be used over an edge computing platform for regions determined based on geographical distances. The researchers in [108] offer a multi-region

23

distributed state estimation solution that incorporates edge computing and uses local estimates that are calculated using the belief propagation algorithm over spatially defined regions. In a multi-region distributed setting, ANNs have also been taken into consideration for state estimation problems in [109, 110]. In [110], to reduce the computational cost of the SE problem with distributed computation, multiple ANN three-phase estimators are executed in parallel and integrated into a multi-region state estimation architecture. A stochastic optimization mechanism-based differential evolution algorithm (DEA) has been proposed in [111]. The algorithm is highly scalable. However, a very robust and fast communication link between parallel processors is a fundamental requirement in this algorithm. Emerging ANN techniques such as GCNN have also been proposed for multi-region settings in [112].

In the multi-region framework for large-scale power systems, distributed cubature KF has also been investigated for the state estimation problem in [113]. Special cases of multi-region distributed state estimation, in which regions are defined based on each bus and its one-hop neighborhood, which we term *1-hop neighborhood* state estimation, have also been considered in the literature. Examples of 1-hop neighborhood distributed state estimation techniques in power systems include techniques based on neural networks [114], belief propagation [115], and KF [116, 117, 18]. For instance, the work in [116] uses diffusion-based KF for 1-hop neighborhood distributed state estimation. Specifically, a selected set of buses in the system share a subset of intermediate estimates with their neighbors. The proposed technique in [117] utilizes the information diffusion strategy in modeling 1-hop neighborhood distributed KF to detect cyber-physical stresses. In this dissertation data-driven state estimation methods that use the local information in the defined regions over the system to provide fast and effective state estimates have been studied.

2.4 State Estimation Under Cyber and Physical Stresses

Security of the cyber-physical systems, including smart grids, has been the focus of much research. Studying and mitigating the effects of joint cyber and physical attacks in CPSs are categories of such research that have gained lots of attention recently [118, 119, 120, 121, 122]. For

instance, Sultan et. al in [118, 119, 120] exploited a joint graph-based and power analysis approach for state estimation and line failure detection. In [121], the authors considered coordinated cyber-physical attacks that can lead to line outages. In the latter work, the goal is to identify the most damaging and undetectable line outages using power system analysis and an optimization framework. In [122], an in-depth review of the smart grid security from a CPS perspective is presented, and prominent cyber-physical attack schemes with significant impact on the smart grid operation and corresponding defense solutions have also been discussed. In [70], the authors have modeled the fault localization problem as a graph search approach using G-CNN in the distribution grid.

Other than the power-based and graph-based analyses of the security threats, many researchers have used PMU data for detecting the line outage in power grids in case of failures or attacks. For instance, in [123], the authors use PMU data along with the topology information to detect line outages. State estimation of power systems using PMU data has also been extensively studied [124, 125, 126, 127, 128, 129]. For instance, in [125], a two-step hybrid state estimation combines both the conventional WLS method and linear estimation that utilizes PMU measurements. In [128], real-time fault detection and faulted line identification functionality is proposed based on computing parallel synchrophasor-based state estimators. Based on the optimal filter and graph theory, a robust distributed smart grid state estimation algorithm is proposed in [129] for defense against false data injection attacks. A deep auto-encoding Gaussian mixture model (DAGMM) has been proposed as a data-driven detection mechanism of stealthy false data injection attack against power system state estimation is proposed in [130]. For instance, the work in [8] addresses an early stress detection and locating method based on a linear predictive filter that may be used over an edge computing platform for regions determined based on geographical distances.

In addition to new techniques based on graph and PMU data analyses, power system security has been extensively studied using traditional state estimation methods [20, 45, 131, 132, 133] in which accurate knowledge of the system model is required. The work in [131] provides a survey discussing state of the art in electric power system state estimation. A review of power

system dynamic state estimation techniques using conventional methods have also been discussed in [45, 20]. Strategies for malicious data manipulation and data integrity attacks against state estimation and countermeasures for such attacks have been reviewed in [133, 132]. Although many powerful techniques have been developed in state estimation for power systems, the availability of large volumes of data and data analytics techniques can provide new opportunities to help with state estimation in special situations, for example, when the system model is not available or accurate (such as in the cases of joint cyber attacks). The presented work in this dissertation is focused on a data-driven approach for state estimation using PMU data for transmission line state estimation. A state estimation method in case of joint cyber and physical attacks has also been discussed in Chapter 6.

2.5 Summary

A review of conventional and data-driven state estimation strategies in power systems was presented in this chapter. It was discussed that state estimation has been conducted in a centralized manner in power systems, traditionally. However, the communication and computation burden of processing the measurements from the vast deployed sensors in the system suggest the need for distributed state estimation to reduce the communication and processing cost and to satisfy the fast response time requirements of critical functions in power systems that rely on state estimation. While some of the conventional distributed approaches are computationally efficient, criteria, such as local observability, need for an accurate system model, the need for a global coordinator, and stochastic and non-Gaussian noise, may limit their performance and applicability. Data-driven, distributed approaches can be a solution to some of such challenges encountered by traditional approaches. A brief review of central and distributed state estimation for both conventional and data-driven approaches were presented in this chapter. Moreover, state estimation under cyber and physical stresses in power systems was discussed. It was particularly discussed that the state estimation can enable observability and recovery of the state of the components under cyber stresses.

Chapter 3: Data-Driven Centralized State Estimation

3.1 Introduction

An important function of the wide-area monitoring systems in power grids is monitoring the state of operational conditions of the system. The information provided by these systems to operators and other control and management functions is essential for well-informed decision-making and reliable and efficient operation of power systems. The conventional state estimators, which have been widely deployed in utility control centers, have been in use to help with monitoring the state of the system. As discussed in Chapter 2, conventional power system state estimation relies heavily on the power system models, including the connectivity, attributes, and operating conditions of the components. The steady-state analysis serves as the foundation for a good portion of the model-based state estimation approaches. However, due to extremely dynamic and stochastic changes induced by, for example, distributed energy generations (DEGs) and rapidly changing loads, steady-state studies cannot be effective for current power systems. In addition, the widespread use of PMUs and the availability of a significant amount of measurement data have created new possibilities for enhancing and supplementing the traditional model-based state estimation used in power systems. As a consequence of this, data-driven and machine learning-based approaches to state estimation are receiving more attention. In this chapter, we present novel data-driven techniques for state estimation in power systems. These techniques are discussed under a centralized setting in this chapter, while some of these techniques have been studied under distributed setting in Chapter 4. Specifically, techniques based on KF, Bayesian regression with auto-regressive approach, minimum mean squares error, and graph neural network have been presented in this chapter. These techniques are discussed under two categories of linear and non-linear techniques.

3.2 Linear Data-Driven State Estimation

The techniques discussed in this category adopt a linear approximation of the system model and its dynamics in developing the models. As data-driven linear state estimation methods, the MMSE estimation, Bayesian linear regression with the auto-regressive process, and linear data-driven KF have been discussed in this section. The MMSE method has the ability to estimate state of certain components using the information from the rest of the system, which is important in mitigating and recovering from the effects of cyber stresses. However, the presented MMSE technique does not consider the temporal information from the measurement time-series. On the other hand, Baysian regression combined with the auto-regressive process captures the temporal aspects of the measurements and The data-driven KF model is also suitable to estimate the state of dynamic and stochastic systems, while using a data-driven linear model to approximate the system model and as such, the model can handle missing and inaccuracies in the system model information.

3.2.1 Minimum Mean Square Error Method

The MMSE method presented in this section enables estimating the state of certain components (which can represent unobservable components in the system due to, for instance, lack of measurement devices or cyber and physical stresses) using the information from the rest of the system. In this model, a linear approximation of the nonlinear dynamics of the power system can be written as $\underline{x} = G(\underline{z}) + \mathbf{B}$, where $\underline{x} = [x_1, x_2, \dots, x_n]^T$ is the state vector, $\underline{z} = [z_1, z_2, \dots, z_m]^T$ is the measurement vector, and $G(\underline{z})$ is the linear function relating the state vectors to the measurements and n and m are the number of states and measurements, respectively and $\mathbf{B}$ is the vector of intercepts. It is assumed that the individual state variables are random variables and the statistical relationships among the state variables can be characterized by specifying the mean and covariance. Let us assume the state variable x_i needs to be estimated in terms of the measurements $\underline{z} \in \mathbb{R}^m$. The linear estimate of state variable x_i can be expressed as $\hat{x}_i = a_{1i}z_1 + a_{2i}z_2, \dots, a_{mi}z_m + b_i$ and the estimation error can be expressed as $e_{ms} = E[(x_i - \hat{x}_i)^2]$. The problem is then identifying $\underline{a}_i = [a_{1i}, a_{2i}, \dots, a_{mi}]$ and b_i, such that error e_{ms} is minimized. Denoting these parameters in a matrix form, such that $a_{ij} \in \mathbf{A}^{n \times m}$ and

$b_i \in \mathbf{B}^{n \times 1}$, the co-efficient matrix can be learnt as:

$$\mathbf{A} = \mathbf{R_{zx}}\mathbf{R_z}^{-1} \tag{3.1}$$

$$\mathbf{B} = \underline{\mu_x} - \mathbf{A}\underline{\mu_z}. \tag{3.2}$$

Here, $\mathbf{A}$ is the coefficient matrix, $\mathbf{B}$ is the vector of intercepts and $\mathbf{R_{zx}}$ and $\mathbf{R_z}$ are the cross-correlation and auto-correlation matrices. Finally, $\underline{\mu_z}$ and $\underline{\mu_x}$ are the vector of mean value of the measurements and the state variables. The overall MMSE state estimation model can be summarized as:

$$\underline{x} = \mathbf{A}\underline{z} + \mathbf{B} \tag{3.3}$$

$$\mathbf{A} = \begin{bmatrix} a_{1i}, a_{2i}, \dots, a_{mi} \\ \vdots, \vdots, \dots, \vdots \\ a_{1n}, a_{2n}, \dots, a_{mn} \end{bmatrix}_{n \times m} \tag{3.4}$$

$$\mathbf{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}_{n \times 1} \tag{3.5}$$

This method exploits the statistical relation among the variables in the system for state estimation; however, it does not consider the temporal information in the measurement data. This approach is specifically helpful in situations where a subset of the system states are unobservable due to failure or cyber stresses or lack of mounted measurement devices to track the state of the component. In such cases, using the available measurement and observable components, MMSE can recover the unobservable states. This techniques and the scenarios that it is applied to are discussed further in Chapter 5.

3.2.2 Bayesian Regression Combined with Auto-Regressive Approach

As discussed in the previous section, the presented MMSE method only considers the statistical relation among the state of individual component and do not consider the temporal information in the measurement. The collected measurement from power systems are generally in the form of time-series and contain important temporal information that can help in the estimation process. In this section, to capture the temporal information in the data, a Bayesian approach has been presented.

Specifically, the PMUs in power system sample the state of the system's components and provide a sequence of phasor measurement observations in a form of time series. We denote the phase angles of buses in the system at time t by $\underline{\theta}_t = [\theta_{1t}, \theta_{2t}, .., \theta_{Nt}]$ and bus voltages by $\underline{V}_t = [V_{1t}, V_{2t}, .., V_{Nt}]$. In this method, we will focus on phase angle time series, while the study can be applied to other collected attributes from the system.

Measurements from the PMUs are multivariate time series while the consecutive measurements have high auto-correlation among themselves. To capture this property of time series, we have expanded our feature space with an AR(p) process such that $\underline{\theta}'_t = f(\underline{\theta}_{t-1}, .., \underline{\theta}_{t-P})$, where $\underline{\theta}_{t-i}$ is the past observation at time $t - i$ in the time series. Then the state estimation problem may be described by linear regression as $\hat{\theta}_{it} = \beta^T \underline{\theta}'_t + e_t$ with coefficients as β, where the goal is to find maximum likelihood (ML) estimate at node i, $\hat{\theta}_{it}$, for θ_{it} while minimizing frobenius norm $\| \beta^T \underline{\theta}'_t - \hat{\theta}_{it} \|$.

However, the computed weight vector β in the linear regression may not be able to capture all uncertainties, especially in the case of noisy measurements, since it only gives the ML estimate. One solution to address this issue is to adopt the Bayesian approach to regression, which will learn the probability distribution of all possible β values that describe the relations between $\underline{\theta}'_t$ and $\hat{\theta}_{it}$. The Bayesian linear regression with AR(p) can be realized as $\hat{\theta}_{it} \sim \mathcal{N}(\beta^T \underline{\theta}'_t, \sigma^2)$. Unlike linear regression, which finds the ML estimate for coefficients β that describe the relationship between the inputs and the outputs, we are interested in computing a probability distribution for β values that describe this relationship. This can be calculated by defining prior distributions for β, and later applying the Bayes rule to calculate the posterior distribution of β. There are several choices

for the prior distribution of the coefficients. In this section, we follow the Bayesian regression model described in [134]. The authors in [134] assumed conjugate normal inverse-gamma prior for β, σ^2, and used variational inference that makes the computations faster. Under this assumption, the likelihood of the $\hat{\theta}_{it}$ can be written as in Equation 3.6.

$$\rho(\theta_{it}|\underline{\theta}'_t, \beta, \sigma^2) = \mathcal{N}(\theta_{it}|\underline{\theta}'_t, \beta, \sigma^2) \tag{3.6}$$

Note that in the Bayesian approach a general assumption is that the individual data streams from the PMUs have Gaussian distribution, which may not be accurate and valid in all cases. Modeling individual time series with their true distribution may improve the estimation performance. However, in this case, the mathematical model may become intractable and will have a higher computational cost.

3.2.3 Kalman Filter with Linear System Identification

The data-driven KF approach tries to develop models to track the dynamics of the power system; specifically by approximating functions $f(\underline{x})$ and $h(\underline{x})$ in equation Equation 2.2 using PMU measurement data. In this section, linear KF and linear data-driven models to approximate functions $f(\underline{x})$ and $h(\underline{x})$ are considered. This technique is named multivariate linear regression KF (MLR-KF). Specifically, for the multivariate time-series modeling, a Multiple Linear Auto-regressive (MLAR) combined with Least Absolute Shrinkage and Selection Operator (LASSO) algorithm is utilized. The multivariate time-series approximation, in this case, can be formulated as:

$$\min_{\beta_1,\dots,\beta_{N\times p}} \left(\frac{1}{2M} \sum_{l=1}^{M} \left(\underline{x}_{i,l} - \sum_{j=1}^{N} \sum_{\tau=1}^{p} \underline{x}_{j,l-\tau} \beta_{j,\tau}^i \right)^2 + \lambda \sum_{j=1}^{N} \sum_{\tau=1}^{p} \left| \beta_{j,\tau}^i \right| \right) \tag{3.7}$$

where M is the number of observations, N is the number of buses, $\beta_{j,\tau}^i$ is the regularized coefficient at time-lag τ for the jth bus when modeling bus i, $|.|$ is the ℓ_1 norm, and λ is the ℓ_1 regularization parameter controlling the boundary of the coefficients, and p represents the number of time-lag steps. Once the coefficients of the MLAR model are learned, the data-driven KF model preliminaries can

be defined as:

$$\mathbf{X} = [\underline{x}_t, \underline{x}_{t-1}, \dots, \underline{x}_{t-p}]^T \tag{3.8}$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{B}_{N \times (N \times p)} \\ \mathbf{W}_{((N-1) \times p) \times (N \times p)} \end{bmatrix}_{(N \times p) \times (N \times p)} \tag{3.9}$$

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}_{N \times (N \times p)} \tag{3.10}$$

where $\underline{x}_t$ is the vector of state attributes of all the buses in the system, $\mathbf{F}$ is the state transition matrix in which $\mathbf{B}$ has $\beta_{j,\tau}^i$ in its ith row and $j\tau$th column, and $\mathbf{W}$ consists of an identity matrix of $(N \times p - N) \times (N \times p - N)$ and a $(N \times p - N) \times N$ vector of zeros in the last column of the matrix. Matrix $\mathbf{H}$ is the observation matrix consisting of an $N \times N$ identity matrix and a matrix of $N \times (N \times p - N)$ zeros. Now, the multivariate KF model can be presented as:

(1) Prediction step:

$$\mathbf{X}_{k+1|k} = \mathbf{F}\mathbf{X}_{k|k-1} \tag{3.11}$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}\mathbf{P}_{k|k}\mathbf{F}^T + \mathbf{Q} \tag{3.12}$$

(2) Correction step:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}^T[\mathbf{R} + \mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T]^{-1} \tag{3.13}$$

$$\mathbf{X}_{k+1|k+1} = \mathbf{X}_{k+1|k} + \mathbf{K}_{k+1}[\mathbf{Z}_k - \mathbf{H}\mathbf{X}_{k+1|k}] \tag{3.14}$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1}[\mathbf{R} + \mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T]\mathbf{K}_{k+1}^{-1} \tag{3.15}$$

where $\mathbf{X}_{k+1|k}$ and $\mathbf{X}_{k+1|k+1}$ are the predicted and corrected state vectors; $\mathbf{P}_{k+1|k}$ and $\mathbf{P}_{k+1|k+1}$ are the predicted and corrected state covariance matrices. $\mathbf{K}_{k+1}$ is the Kalman gain matrix, $\mathbf{Q}_{Np \times Np}$ is the modeling error matrix, and $\mathbf{R}_{N \times N}$ is the measurement error matrix. In the next section, the linearity assumption of the model is relaxed, and a non-linear data-driven KF model is presented.

3.3 Nonlinear Data-Driven State Estimation

In this section, data-driven nonlinear approximation of system dynamics are derived from measurement data to estimate the state of the system. These models are expected to track the state more closely due to the nonlinear dynamics of the power systems. In this section, extended KF (EKF), unscented KF (UKF), and the temporal graph convolutional neural network (T-GCNN) models have been presented. As discussed in Chapter 2, EKF and UKF are among the most popular dynamic state estimation algorithms. The nonlinear estimation problem can be solved using an EKF by first linearizing the state and/or measurement equations and then applying a standard KF to the linearized estimation problem. However, the linearization process results in approximation errors that can result in underestimating state uncertainties. On the other hand, the UKF selects what is called sigma point samples from the filtering distribution and then propagates or updates those samples all the way through the (nonlinear) state and measurement models. The updated filtering distribution is represented by the weighted set of sigma points that were produced as a result, and this distribution is then approximated as a moment-matched Gaussian distribution. This leads to state estimates that can better capture the uncertainty in the states compared to the estimates derived from the EKF. However, UKF has a higher computational cost compared to EKF. In addition to KF-based model, we have adopted a neural network-based approach, T-GCNN, which enables building nonlinear models while capturing both the temporal and topological information among the components. In this model, the power system topology will be modeled as graph structure and captured the underlying inter-relations among the variables. T-GCNN extracts the topological information using graph convolution and temporal aspects of the measurements using gated recurrent units. The details of these models are presented next.

3.3.1 Extended Kalman Filter with Nonlinear System Identification

In this section, the non-linear KF approach has been considered to better track the non-linear dynamics of power systems. In the general non-linear system of equations, presented in Equation 2.1, functions $f(.)$ and $h(.)$ represent the non-linear measurement and system equations, respectively,

which can be learned using data-driven techniques to track the system's non-linear dynamics. To this end, the data-driven multivariate polynomial regression (MPR) model [135] has been utilized for non-linear system identification. Specifically, to have comparable linear and non-linear KF models, the non-linear auto-regressive exogenous model (NARX) has been considered in this section as follows:

$$
x_{i,t} = \sum_{\kappa=1}^{n_v} a_{i,\kappa} \gamma_\kappa + \xi_i
$$

$$
\gamma_\kappa = \prod_{j=1}^{N} \prod_{\tau=1}^{p} x_{j,t-\tau}^{\sigma_{j,\tau}}
$$

$$
\sum_{j=1}^{N} \sum_{\tau=1}^{p} \sigma_{j,\tau} \le \eta
$$

(3.16)

where n_v is the total number of unique combinations of $N \times p$ independent variables and p time-lagged measurements. If the order of the polynomial is considered to be η, and γ_κ represents the κ-th regressor.

Extended KF is a non-linear version of KF (here, named MPR-EKF), which specifies the non-linear systems and state equations at step k as:

$$
\mathbf{F}_k = \left. \frac{\delta f}{\delta \mathbf{X}} \right|_{\mathbf{X}_k}
$$

(3.17)

$$
\mathbf{H}_k = \left. \frac{\delta h}{\delta \mathbf{X}} \right|_{\mathbf{X}_k}
$$

(3.18)

Then, the main steps in EKF estimation are as follows:

(1) Prediction step:

$$
\mathbf{X}_{k+1|k} = f(\mathbf{X}_{k|k-1})
$$

(3.19)

$$
\mathbf{P}_{k+1|k} = \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{Q}
$$

(3.20)

(2) Correction step:

$$
\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_k^T [\mathbf{R} + \mathbf{H}_k \mathbf{P}_{k+1|k} \mathbf{H}_k^T]^{-1}
$$

(3.21)

$$\mathbf{X}_{k+1|k+1} = \mathbf{X}_{k+1|k} + \mathbf{K}_{k+1}[\mathbf{Z}_k - h(\mathbf{X}_{k+1|k})] \tag{3.22}$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1}[\mathbf{R} + \mathbf{H}_k\mathbf{P}_{k+1|k}\mathbf{H}_k^T]\mathbf{K}_{k+1}^{-1} \tag{3.23}$$

3.3.2 Unscented Kalman Filter with Nonlinear System Identification

The propagation of error in covariance in Equation 3.20 through the linearization of the underlying non-linear model in EKF, as shown in Equation 3.17, can lead to poor performance in highly dynamic systems. In such circumstances, the unscented KF can be applied, which generates a minimal set of sample points (known as sigma points) around the mean using the deterministic sampling technique known as the unscented transformation. The estimations of the new mean and error covariance are then generated from these altered sigma points using non-linear functions. The computation is performed in the following phases. First, sigma points are selected to propagate from k to $k + 1$ time step using the best guess of $\mathbf{p_k}$ and x_k as follows, where $\mathbf{p_k}$ is the diagonal element of matrix $\mathbf{P}_k$, and x_k is an element of vector $\mathbf{X}_k$.

$$x_k^{(i)} = x_k + \tilde{x}^{(i)} \quad \text{for } i = 1, \dots, 2N \tag{3.24}$$

$$\tilde{x}^{(i)} = \left(\sqrt{N\mathbf{p}_k}\right)_i^T \quad \text{for } i = 1, \dots, N \tag{3.25}$$

$$\tilde{x}^{(i+N)} = -\left(\sqrt{N\mathbf{p}_k}\right)_i^T \quad \text{for } i = 1, \dots, N, \tag{3.26}$$

where $x_k^{(i)}$ represents the sigma point, i is the index of the sigma point, and N is the number of buses or states to be estimated. The variable $\tilde{x}^{(i)}$ denotes the change in sigma point, and both negative and positive changes are considered in defining sigma points. Next, the non-linear function $f(.)$ is used to transform all the sigma points $x_k^{(i)}$ into $x_{k+1}^{(i)}$ as $x_{k+1}^{(i)} = f(x_k^{(i)})$.

The initial state estimate will then be obtained by combining $x_{k+1}^{(i)}$ values as follows:

$$\hat{x}_{k+1|k} = \frac{1}{2N} \sum_{i=1}^{2N} x_{k+1}^{(i)} \tag{3.27}$$

$$\hat{\mathbf{p}}_{k+1|k} = \frac{1}{2N} \sum_{i=1}^{2N} (x_{k+1}^{(i)} - \hat{x}_{k+1|k})^2 + \mathfrak{q}, \tag{3.28}$$

where $\mathfrak{q}$ is the modeling error and an element of $\mathbf{Q}$. Using the updated estimate $\hat{x}_{k+1}$ and $\hat{p}_{k+1}$, new sigma points $\hat{x}_{k+1}^{(i)}$ will be computed as in Equations 3.24 to 3.26 to update the measurements $z_{k+1}^{(i)}$. Then, the updated estimate $\hat{z}_{k+1}$ can be calculated similarly to that of $\hat{x}_{k+1}$ in Equation 3.27. The variance, $\mathbf{p}_z$ (an element of $\mathbf{P}_Z$), and covariance, $\mathbf{p}_{xz}$ (an element of $\mathbf{P}_{XZ}$), of the predicted measurements are calculated as follows:

$$\mathbf{p}_z = \frac{1}{2N} \sum_{i=1}^{2N} (z_{k+1}^{(i)} - \hat{z}_{k+1|k})^2 + \rho \tag{3.29}$$

$$\mathbf{p}_{xz} = \frac{1}{2N} \sum_{i=1}^{2N} (x_{k+1}^{(i)} - \hat{x}_{k+1|k})(z_{k+1}^{(i)} - \hat{z}_{k+1|k}), \tag{3.30}$$

where ρ is the measurement error and an element of $\mathbf{R}$. The correction state can then be calculated as follows (denoted in vector and matrix notation):

$$\mathbf{K}_{k+1|k+1} = \mathbf{P}_{XZ}\mathbf{P}_Z^{-1} \tag{3.31}$$

$$\mathbf{X}_{k+1|k+1} = \mathbf{X}_{k+1|k} + \mathbf{K}_{k+1|k+1}(\mathbf{Z}_{k+1|k} - \hat{\mathbf{Z}}_{k+1|k}) \tag{3.32}$$

$$\mathbf{P}_{k+1|k+1} = \hat{\mathbf{P}}_{k+1|k} - \mathbf{K}_{k+1|k+1}\mathbf{P}_Z\mathbf{K}_{k+1|k+1}^T. \tag{3.33}$$

One benefit of UKF over EKF is that it can reduce the computational complexity since it does not require calculating the Jacobin matrix in the estimation process. This approach is referred to as MPR-UKF in this method.

3.3.3 Temporal Graph Convolutional Neural Network Approach

Graph has an excellent ability to capture structural information of a network. Power systems are complex interconnected networks where the physical topology also carries important spatial information which may further improve the data-driven state estimation. In an effort to combine

spatial and temporal information in state estimation process, a Spatial temporal graph convolutional neural network based approach is proposed as follows.

In general, G-CNN combines the feature information and the graph structure to learn better representations on graphs, for instance, via feature propagation and aggregation. In this method, given the input measurement features Z_t, the state estimation process is modeled in a message-passing framework of Graph Neural Network [68]. In this framework, the state x_1 of the node n_1 depends on the information contained in its neighborhood (see Figure 3.1). In the example in Figure 3.1, the state of node n_1 can be learnt by aggregating the information of its neighbours using function f as $x_1 = f(x_2, x_3, x_4, x_5)$, where $f(.)$ is a nonlinear function modeled in the neural network. The same goes for each node of the graph. In this framework, message h (hidden node state information) will be passed among neighbors for node-level state prediction.
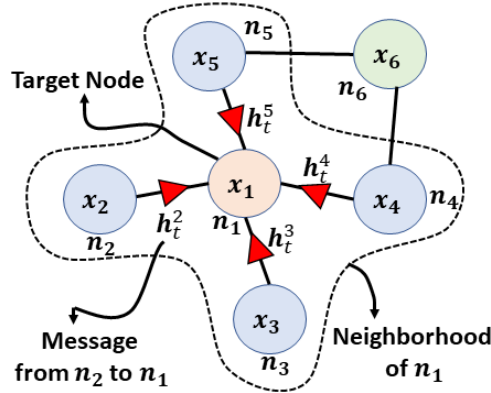
Figure 3.1: An example of the message passing process in a neighborhood for G-CNN.

This step along with integrating the temporal dependencies in the data for one-step ahead prediction of the system's state can be modeled as

$$Z_t = f_1(.)X_t, \tag{3.34}$$

$$X_{t+1} = f_2(.)X_t, \tag{3.35}$$

Here, $f_1(.)$ and $f_2(.)$ are nonlinear functions of state variables X_t that will be learnt from the measurement data. Particularly, the function $f_1(.)$ maps the measurements to the system state, and

$f_2(.)$ does the one-step ahead forecasting. Instead of learning two separate nonlinear relationships, both functions can be combined into one mapping function $F(.)$ as $X_{t+1} = F(Z_t, X_t)$. In this section, a two layer spatio-temporal G-CNN is used to learn the mapping function F. Since power system measurements are highly correlated multivariate time-series, to capture the spatial dependencies, a graph convolution layer with message passing is used as the first layer of the network. The second layer is a GRU layer, which is responsible for capturing the temporal dependencies of the measurements.

The traditional CNN can obtain local spatial features in the Euclidean data space. Power system graphs are complex networks and as such, CNN cannot accurately capture the embedded spatial dependencies. G-CNN can learn the spatial features of complex graph structured data based on the neighborhood aggregation (for message passing framework), given the adjacency matrix $\mathcal{A} := \{0, 1\} \in \mathbb{R}^{N \times N}$, and feature matrix Z. A typical G-CNN layer [136] can be expressed as follows:

$$\mathcal{H}^{l+1} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \mathcal{H}^l \mathcal{W}_l). \tag{3.36}$$

Here, $\tilde{A} := \mathcal{A} + I_N$ (where I_N is the identity matrix of size N) and $\tilde{D} := I_N \sum_j \tilde{A}_{i,j}$ are the adjacency and degree matrix, respectively. Moreover, $\sigma(.)$ is the sigmoid activation function, l is the layer number, $\mathcal{W}_l$ holds the weights of layer l, and $\mathcal{H}^l$ is the output of layer l. Multiple such layers can be added on top of one another to create a multi-layer G-CNN model.

To perform forecasting-aided estimation on multivariate time-series, a GRU model is used. The basic principle of GRU and LSTM models are roughly the same [137]. However, LSTM has a comparatively complex structure and longer training time. Stacking the GRU model with G-CNN creates a temporal G-CNN (T-GCNN), which was first proposed in [138]. The T-GCNN can be described as follows

$$F(Z_t, \mathcal{A}) = \sigma(\hat{A} ReLU(\hat{A} Z_t \mathcal{W}_0) \mathcal{W}_1) \tag{3.37}$$

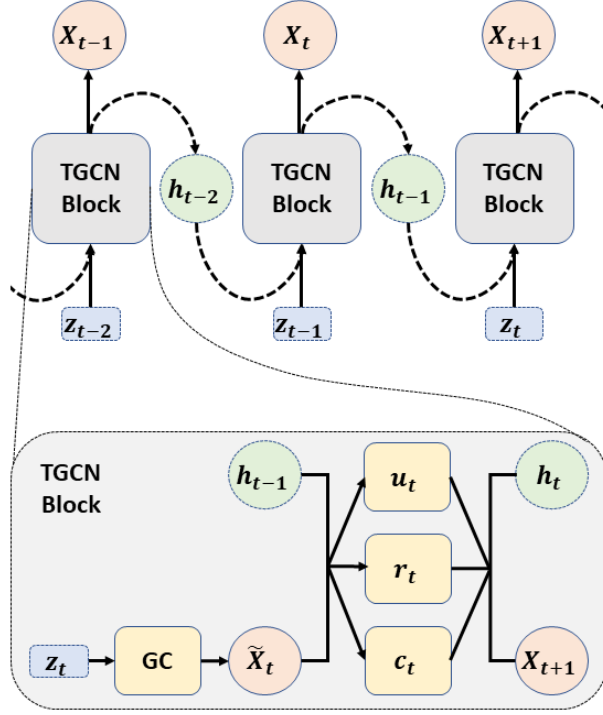$$u_t = \sigma(W_u[F(Z_t, \mathcal{A}), h_{t-1}] + b_u) \tag{3.38}$$

Figure 3.2: Schematics of the temporal G-CNN model for state estimation in the smart grid.

$$r_t = \sigma(W_r[F(Z_t, \mathcal{A}), h_{t-1}] + b_u) \tag{3.39}$$

$$c_t = \tanh(W_c[F(Z_t, \mathcal{A}), (r_t * h_{t-1})] + b_c) \tag{3.40}$$

$$h_t = u_t * h_{t-1} + (1 - u_t) * c_t \tag{3.41}$$

Here, $\hat{A} := \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$ is the pre-processing of graph convolution layer. $\mathcal{W}_0 \in \mathbb{R}^{\beta \times \delta}$ and $\mathcal{W}_1 \in \mathbb{R}^{\delta \times \tau}$ are the model weights, where the parameters β, δ, and τ denote the batch size, hidden unit, and prediction length, respectively. In this model, r_t is the reset gate, which is used to decide how much of the past information to forget and u_t is the update gate, which helps the model to determine how much of the past information needs to be passed to the future. Moreover, c_t is the memory unit, which calculates information stored at time t and h_t is the hidden state at time t. The parameter b denotes the bias parameter at respective levels. Over the model training iteration h_t will slowly converge into model prediction X_{t+1}. Figure 3.2 shows the overall schematics and information flow of T-GCNN model.

3.4 Summary

In this chapter, several centralized data-driven state estimation approaches have been discussed. We started the discussion by introducing few data-driven linear models to approximate the dynamics of the power system including MMSE, Bayesian Multiple Regression combined with an Auto-regressive approach, and linear KF. As nonlinear state estimation techniques, nonlinear KF-base techniques, such as the Extended KF and the Unscented KF with nonlinear system identification have been presented. We discussed that the linear MMSE is particularly helpful in scenarios in which a sub set of components are unobservable due to, for instance, cyber and physical stresses. The Bayesian approach considers the measurements as multi-variate time series to capture the temporal information in the data in the estimation process. The data-driven system identification techniques have been implemented to approximate the underlying power system dynamics in both linear and nonlinear fashion in KF-based state estimation. These data-driven system models are used specifically to develop linear, extended, and unscented KF models for state estimation in the system. All these approaches are completely data-driven and do not require system topological information. However, a data-driven approach based on G-CNN is presented for the state estimation problem in smart grids. Since G-CNNs are deep learning-based methods that operate on the graph domain, they lend themselves well to the problem of state estimation in smart grids with underlying graph-based structures and interactions. A modified variant of G-CNN, namely temporal G-CNN (T-GCNN), suitable for analyses of graph time series, is presented for one-step ahead state prediction in smart grids. The T-GCNN model can deal with complex spatial dependencies over graphs as well as temporal dynamics in the measurements. Specifically, a message passing G-CNN is used to capture the topological structure of the smart grid network in the spatial dependency analyses, and the gated recurrent units are used to capture the dynamic variation of state information for obtaining the temporal dependencies. These data-driven models are particularly helpful in supporting state estimation under conditions of missing or limited measurements, such as partial unobservability due to failures or attacks in the sensing and monitoring system or the limited availability of PMUs.

Chapter 4: Distributed Multi-Region State Estimation

4.1 Introduction

Deregulation of energy markets, penetration of renewables, advanced metering capabilities, and the urge for situational awareness, all call for system-wide power system state estimation. Implementing a centralized state estimator though is practically challenging due to the large scale of these systems, the communication bottleneck in real-time monitoring, regional disclosure policies, and reliability issues.

Many of the data processing and computations related to monitoring the power systems have been traditionally performed centrally in utility-owned servers or cloud platforms. Communicating the large volume of data and their processing in central units, inevitably adds delay and inaccuracies in the state estimation and monitoring of the system. For certain time-sensitive functions, such delays could cost the reliability and stability of the system. To address this issue, distributed and local processing of data can provide a good solution, especially for time-sensitive functions. Data-driven state estimation is one of such functions that can improve the response time of the system, particularly, to critical conditions, such as failures or cyber-stresses. The data-driven distributed state estimation can provide a faster and more accurate estimation of current conditions in the local regions as well as predicting future trends in state changes to identify abnormal conditions. New technologies, such as Edge or Fog computing, can provide a platform to enable these functions by local and distributed processing of data to enhance system monitoring in power grids [8, 9].

To enable local processing of data for distributed state estimation, local neighborhoods or regions need to be defined over the physical layout of the power system. This will also allow for provisioning the computational resources required for each region and their placement (for example in an edge computing platform). While the partitioning of the power system to regions

can be dictated by physical, geographical, or economical constraints due to the layout of the power system and the communication and computing systems' resources, the physics of the system and relations and interactions among the components can also be considered in defining the regions as they can affect the accuracy of the estimations. In this chapter, we consider the geographical properties and the correlations among the PMU time series to define regions over the power system for state estimation purposes. Additionally, a multi-region distributed state estimation framework have been discussed. We will discuss some of the state estimation techniques presented in Chapter 3 under this distributed framework. In this framework, each local region will work independently or in collaboration with other regions to compute the local estimates using the local measurements and limited information from neighboring regions. This chapter discusses the information sharing strategies among the distributed regions, that can improve overall estimation. Additionally, as a special case of multi-region, a fully distributed state estimator based on KF has also been presented. To easily keep track of the various state estimation techniques, the fully distributed state estimation is denoted as Case-I, centralized estimation processes as Case-II, multi-region state estimation without information sharing/ collaboration as Case-III, and with information sharing as Case-IV.

4.2   Power Grid Partitioning

We have considered multi-region distributed state estimation for smart grids. As such, the first step is to define and identify partitions in the system that facilitate the distributed state estimation using the local information within each region. We assume that the system consists of N buses connected using transmission lines.

As computations are supposed to be local, the geographical proximity of the PMUs is one criterion in defining the regions. To incorporate the dynamics and properties of the power system, which can affect the state estimation performance, we consider the cross-correlations of PMU time series at buses as the second criterion in defining the regions. This criterion will result in more inter-related feature space for the estimation model, which can improve the estimation performance.

In partitioning power systems, various criteria have been considered in the literature [139, 140]. Specifically, in this section, for multi-region state estimation, as the computations are performed locally, the geographical proximity of the PMUs can be one criterion in defining the regions. Further, cross-correlations among PMU time-series at buses can serve as the second criterion in defining the regions as they enable capturing information on the dynamics and interaction properties of the power system components. The latter criterion will result in more inter-related feature space for the estimation model, which can improve the performance.

The objective is to divide the grid into J regions, where S_i represents the collection of nodes that make up a region, and $\mathcal{S} = \{S_1, S_2, \dots, S_J\}$ signifies the set of regions. Given the total number of regions J, the partitioning problem has been solved using a density-based clustering technique, specifically k-means, to identify the non-overlapping regions. The grid will be partitioned using k-means clustering while minimizing the sum of squared distances of the nodes within a region and maximizing the spatial correlation. In this case, the cross-correlation values between the PMU time-series of the buses and their geographical distance from one another are employed in the feature space for the k-means algorithm. Figure 4.1 shows an example of a partitioned grid over IEEE 118.
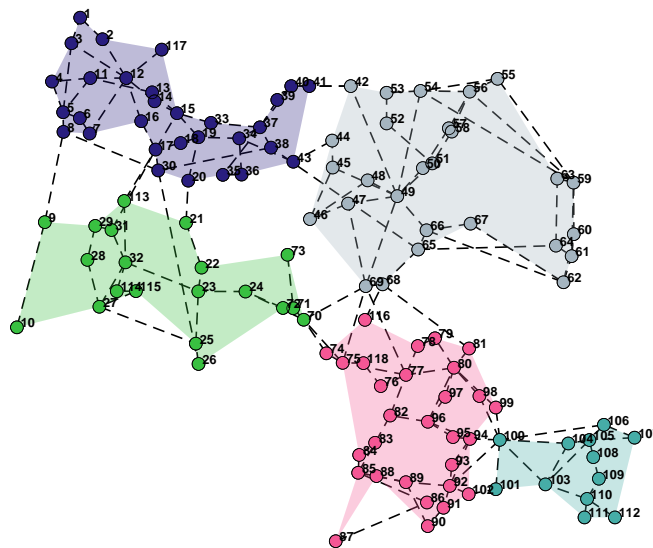
Figure 4.1: An example of a partition over the IEEE 118 bus system using the modified k-means partitioning discussed in Case-III for five regions.

Figure 4.2a shows that non-homogeneous partitions will be formed by this strategy. However, because of the limitations of the local computing and communication capabilities (e.g., communication bandwidth, storage, and computational power), substantial inequalities in the sizes of the regions might not be appropriate. In order to overcome this issue, the aforementioned k-means method is modified to develop a size-constrained k-means algorithm that can produce homogeneous regions in terms of size. First, the grid will be divided into J non-homogeneous regions using the k-means algorithm, as mentioned above (assumed $N \gg J$). Then, the approximate ideal size of the homogeneous partitions is defined by the expression $\lfloor \frac{N}{J} \rfloor$. In the following stage, the largest partition will be determined, and $\lfloor \frac{N}{J} \rfloor$ PMUs that are closest to the region's center will be chosen and assigned to it. This area will be designated as finished and temporarily removed from the graph. For the reduced graph, the same procedure will be repeated. To be more precise, the assignments for the largest partition will be finalized when k-means is used to find $J - 1$ non-homogeneous partitions on the reduced graph in the following step. Up till J homogeneous regions are obtained, this process will be repeated. If there are PMUs left out in the last step (due to $(N \mod J) \neq 0$), they will be assigned to their closest region. An example distribution of homogeneous region sizes are presented in Figure 4.2b. This process is applied once before the state estimation and thus does not affect the computational complexity or run-time of the state estimation techniques.

4.3   Muti-Region State Estimation Framework

For the local processing of data in distributed computing units, we consider edge computing as an example supporting platform of multi-region distributed state estimation. First, the smart grid is divided into regions and the center of each region μ_r (based on its geographical coordinates and average cross correlation) is adjusted to its nearest PMU for optimal placement of edge computing unit. Each edge computing unit is connected via wireless links to PMUs in the region and will run the local state estimation in parallel. It collects the associated regional PMU measurements (such as V, θ) and processes the data for one step ahead state estimation using the estimation
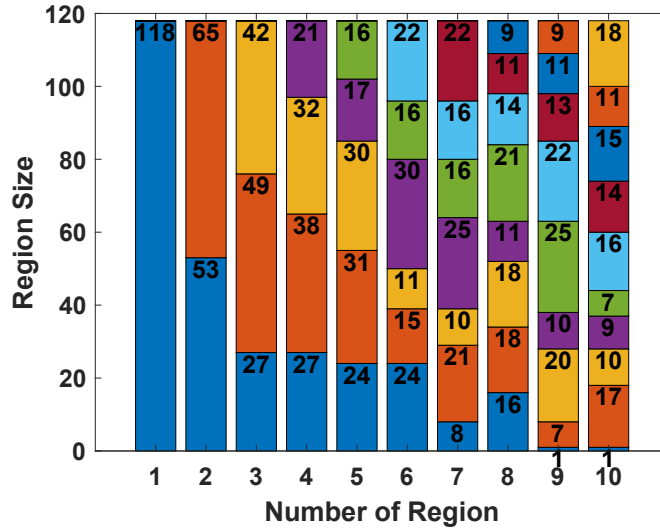
(a) Non-homogeneous
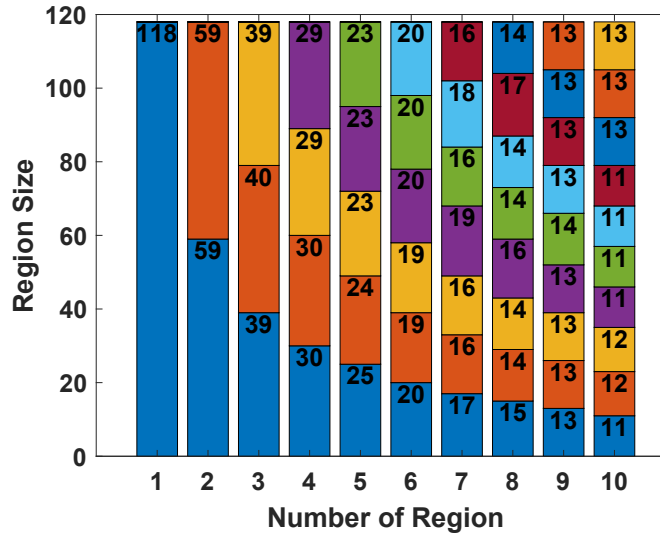
(b) Homogeneous

Figure 4.2: Stacked representation of different region sizes for (a) non-homogeneous partitioning and (b) homogeneous partitioning discussed in Case-III. Each color represents a different region.

model (e.g., BMLR, Linear-KF, EKF, and UKF). Since data is being processed by local edge nodes instead of being transferred to the central cloud servers, the overall communication latency will be reduced and faster decision making can be achieved from the local estimations [8]. Finally, local estimates can be used locally for responding to a situation identified or can be communicated to the central systems and also be combined to achieve the overall system estimates. Figure 4.3 shows the general schematics of the multi-region state estimation on the IEEE 118 bus system. In this figure, each colored node of the graph represents a bus in the system and the black solid lines represent a transmission line. Different colors assigned to the nodes specify the regions defined over the system. Note that in this figure, we do not consider the coordinated and collaborative state estimation among the edge servers of different regions. We also assume that there is no central coordinator to support the distributed estimation. However, collaborative state estimation ad considering communication among the edge servers of different regions in a distributed manner can also be considered, as is considered in [18, 110, 113] as well as later in this chapter. KF-based linear and nonlinear methods (e.g., Linear-KF, EKF, and UKF) have been evaluated in Chapter 6 considering various method of collaboration in the same multi-region framework.
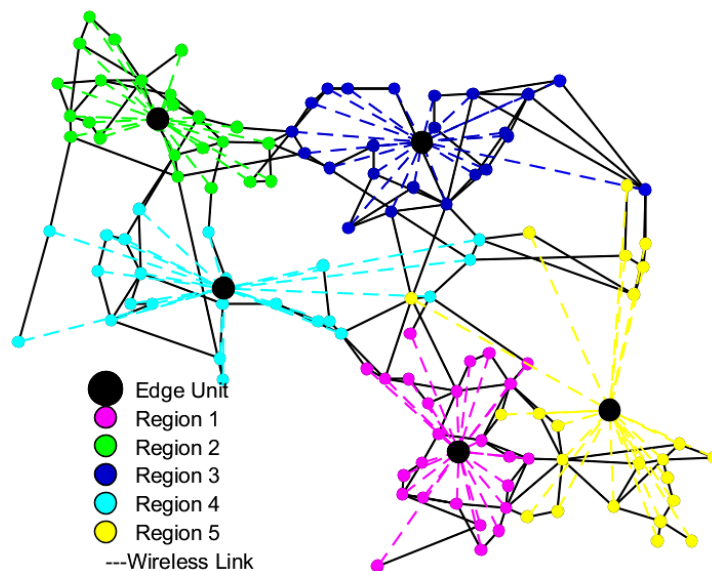
Figure 4.3: Schematics of the proposed multi-region DSE framework enabled by distributed edge computing over the IEEE 118 bus system.

4.4 Multi-Region State Estimation without Information Sharing

In this scenario, the distributed region will only utilize the local information for computing the local state. No external information will shared among the region. The algorithms that have been discussed in Chapter 3 in both linear and non-linear fashion will be evaluated in this setting. The estimation algorithm will be applied to each region as if it is a centralized estimation process. Although, this scenario may not be practical, we will utilize the results from this evaluation as benchmark for other collaborative techniques. We will name this scenario as Case-III for referring to when we discuss evaluations.

4.5 Multi-Region State Estimation with Information Sharing

In multi-region state estimation, allowing limited information sharing among the regions can improve the estimation accuracy by making the estimation process aware of the overall state of other regions and adaptive to changes in them. This case is an extension of non-collaborating state estimation (Case-III), in which regions share selected data or features with their neighboring regions to be used in the local estimation process (Case-IV). In this section, various information-sharing schemes have been considered.

4.5.1 Information Sharing through Overlapped Regions

In this scenario, common buses among regions are considered to enable information sharing among regions for the state estimation process. As such, instead of the partitioning mechanism presented in section 4.2, a soft clustering technique, specifically, fuzzy c-means (FCM) clustering [141, 142], has been utilized to allow overlapping regions. The partitioning mechanism in this case is similar to the one presented in Case-III, except that buses can be assigned to multiple regions with varying degrees of membership belonging, where the membership can be tuned through specified thresholds (Case-IVa). The process for keeping the region sizes homogeneous is the same as the hard partitioning mechanism presented in Case-III. An example of overlapped partition scenario

has shown in Figure 4.4. The overlapping regions expand the feature space of individual regions by capturing partial information about the dynamics of neighboring regions through common buses.
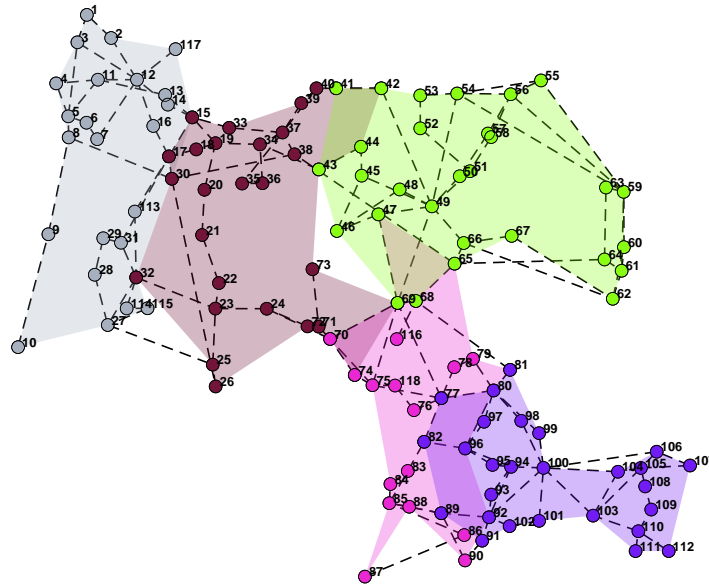

Figure 4.4: An example of five overlapped regions discussed in Case-IVa.

4.5.2   Information Sharing through Auxiliary Buses

In this information-sharing technique, an auxiliary bus will be associated with each region, which is responsible to summarize and share the information about that region with the rest of the regions. As such, considering J regions, each region, say ith, will use data from its N_i buses plus the data from $J - 1$ auxiliary buses representing the rest of the regions.

The next step in information sharing through auxiliary buses is to assign features to them, such that the features embed the state information of the region that they are representing. The feature associated with the auxiliary bus of the region i at a specific time instance is denoted by m_i. In this section, the following representing features are considered for m_i:

- Mean of the state values of the buses of region i;

- State data of the bus with the largest variance in the region i;

- State data of the bus in the region *i* with maximum average correlation with the rest of the buses in the region *i*;

- State data of the bus in the region *i* with maximum average correlation with the buses in the region *j*;

- As the first principal component of states in region *i*.

4.6 Fully Distributed State Estimation

A fully-distributed scenario is a special case of multi-region scenario, where each area has a single bus as the member. A fully-distributed state estimation without information sharing/collaboration would mean a univariate estimation. The assumption in the fully distributed case is that the full set of measurements are available on all the buses in the system, and the goal is to perform a 1-step-ahead prediction for the state of each bus using only the information from the same bus. We will denote the fully distributed scenario as Case-I for simulation and comparison purpose. In this case, the individual bus measurement time-series are modeled as an auto regressive process as follows:

$$x_{i,t} = \sum_{\tau=1}^{p} \phi_{i,\tau} x_{i,t-\tau} + \xi_i, \tag{4.1}$$

Here, p represents the number of time-lag steps, and ξ_i is the modeling error, which is independent of measurements.

Now, the model parameter $\underline{\Phi}_i = [\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,p}]^T$, where $(.)^T$ denotes the transpose operation, is defined as $\underline{\Phi}_i = \underline{\mathbb{R}}_i^{-1} \mathbf{r}_i$. Here, $r_{i,\tau}$, the τth element of vector $\mathbf{r}_i$, is the auto-correlation of bus x_i considering τ time-lagged samples, and $\underline{\mathbb{R}}_i$ is defined as the following:

$$\underline{\mathbb{R}}_i = \begin{bmatrix} 1 & r_{i,1} & ... & r_{i,t-p} \\ \vdots & \ddots & ... & \vdots \\ r_{i,p-1} & r_{i,p-2} & ... & 1 \end{bmatrix}. \tag{4.2}$$

This AR model is used to design the $f(.)$ and $h(.)$ functions in Equation (1) as follows:

$$\mathbf{F}_i = \begin{bmatrix} \underline{\Phi}_i \\ \mathbf{W} \end{bmatrix}_{p \times p} \tag{4.3}$$

$$\mathbf{H}_i = \begin{bmatrix} 1 & \underline{\mathbf{0}} \end{bmatrix}_{1 \times p} \tag{4.4}$$

where $\mathbf{F}_i$ is the state transition matrix with matrix $\mathbf{W}_{p-1 \times p}$ that consists of an identity matrix of $1 - p \times 1 - p$ and a $p - 1 \times 1$ vector of zeros in the last column of the matrix. The matrix $\mathbf{H}_i$ is the observation matrix with a 1 in its first entity and a vector of $1 \times p - 1$ in the rest to capture that the measurements, and the states are the same in this case. For simplicity of notations, hereafter, the time index is dropped, and index k is used to keep track of KF iterations. Vector $\underline{x}_{i,k}$ is the current state of bus x_i along with the past $p - 1$ time-lagged states. The $\underline{x}_{i,k}$ can be considered as the current state in KF, and $\underline{x}_{i,k+1|k}$ and $\underline{x}_{i,k+1|k+1}$ are the predicted and corrected state vector, respectively. Now, the KF model can be presented as follows:

(1) Prediction step:

$$\underline{x}_{i,k+1|k} = \mathbf{F}_i \underline{x}_{i,k|k-1} \tag{4.5}$$

$$\mathbf{P}_{i,k+1|k} = \mathbf{F}_i \mathbf{P}_{i,k|k} F_i^T + \mathbf{Q}_i \tag{4.6}$$

(2) Correction step:

$$\mathbf{K}_{i,k+1} = \mathbf{P}_{i,k+1|k} \mathbf{H}^T [\rho_i + \mathbf{H}_i \mathbf{P}_{i,k+1|k} \mathbf{H}_i^T]^{-1} \tag{4.7}$$

$$\underline{x}_{i,k+1|k+1} = \underline{x}_{i,k+1|k} + \mathbf{K}_{i,k+1} [\underline{z}_{i,k} - \mathbf{H}_i \underline{x}_{i,k+1|k}] \tag{4.8}$$

$$\mathbf{P}_{i,k+1|k+1} = \mathbf{P}_{i,k+1|k} - \mathbf{K}_{i,k+1} [\rho_i + \mathbf{H}_i \mathbf{P}_{i,k+1|k} \mathbf{H}_i^T] \mathbf{K}_{i,k+1}^{-1}, \tag{4.9}$$

where $\mathbf{P}_{i,k+1|k}$ and $\mathbf{P}_{i,k+1|k+1}$ are the predicted and corrected state covariance matrix, $\mathbf{K}_{i,k+1}$ is the Kalman gain matrix, $\mathbf{Q}_i$ is an all-zero matrix except for the modeling error ξ_i at the first element, and ρ_i is the measurement error.

4.7   Summary

In this chapter, multiple regions or areas are considered over the power grid to implement distributed state estimation. The distributed state estimation will support wide-area situational awareness in power systems with reduced data transmission and processing delay by processing the measurements locally in regions. Specifically, in multi-region distributed state estimation, the models presented in Chapter 3 was considered in a distributed multi-region state estimation framework. We partitioning of the power system into regions with geographical and power system considerations (such as correlation among the PMU time series). We considered both homogeneous and non-homogeneous region sizes in our study. We also discussed that the proposed method can be implemented over a distributed computing platform, such as edge computing. The proposed framework can lead to low latency and faster data processing, which results in improved wide-area monitoring for smart grids. Various modes of information sharing among the regions have also been discussed to allow a collaborative multi-region state estimation framework, which can also improve the performance of distributed state estimation. The presented techniques and scenarios for multi-region distributed state estimation with and without information sharing are evaluated in Chapter 6.

Chapter 5: State Estimation Under Cyber and Physical Stresses

5.1 Introduction

Modern power grids are becoming more and more equipped with cyber elements for sensing, monitoring, communication, computation, and control, which make them exemplary complex cyber-physical systems. Due to such increased dependency on cyber components, these systems exhibit new vulnerabilities to cyber threats. When cyber attacks occur jointly with physical attacks or failures in the power grid, they could have even more serious impacts and cause large-scale blackouts with severe societal and economic consequences [118]. In the case of physical attacks or failures, the system's stability can be maintained if the SCADA receives precise information about the status of the components and take proper action accordingly. If however, the flow of information is obstructed by a cyber attack, the status of the components will be unobservable to the SCADA, which prevents the control center from taking necessary and appropriate actions in a timely manner.

On the other hand, cyber components provide invaluable opportunities for a more secure and reliable operation of smart grids. For instance, the immense volume of energy data collected by various sensors, such as PMUs, provide new opportunities for detecting, estimating and predicting various events in the system using big data analytics techniques. In this chapter, we consider a scenario of joint cyber and physical attack on the smart grid and discuss how a data-driven method based on PMU data can help in recovering the status information of the components. Similar to the work in [118, 119, 120], we consider the scenario in which an attacker conducts a physical attack on the power system by disconnecting few transmission lines and simultaneously launches a cyber attack on the communication system and prevents the flow of information from the region around the physically attacked area or other regions of the system to the control center. This joint

cyber attack leads to unobservability on a portion of the power system, which has experienced line outages. The goal of the presented work is to use the PMU data from outside the attacked zone (observable parts of the system) to estimate the state of the lines in the attacked zone using a data-driven technique. The availability of large volumes of PMU data in future smart grids and limitations of the traditional power system state estimation due to dependency on accurate power system models, make the data-driven approaches more appealing than before as a complement to the traditional state estimation or individually. In this section, we have specifically used a linear minimum mean square error estimator for recovering the status information of components in the attacked zone. We have evaluated various scenarios and observed that recovering the status information of certain power components are more difficult than others and thus, we have proposed an extension to the linear MMSE estimator by adding iterative feedback to the estimator, which can improve the estimation performance. We have evaluated these data-driven estimation methods in Chapter 6 on various scenarios of joint-cyber attacks including scattered and localized attacks. The results show that the data-driven approaches can be promising approaches for state estimation, particularly during cyber-physical attacks.

5.2   Stress Model

We consider joint cyber and physical attacks and thus the attack definition has two parts. Specifically, to model the *cyber attack*, we assume that the attacker randomly selects a subset A_c of transmission lines (i.e., $A_c \in L$) and masks the flow of information from them to the control center. We call the set A_c the *cyber attack zone* or *attack zone* for short. Further, to model the *physical attack*, we assume that a subset A_p of lines from the attack zone (i.e., $A_p \in A_c$) experiences physical attack or failure.

Further, we consider two scenarios for the attack zone:

5.2.1 Randomly Scattered Attacks

In case of randomly scattered attack scenario, the set A_c of transmission lines is geographically scattered on the system. Figure 5.1a depicts one example of a scattered attack on the IEEE 118 test case topology.

5.2.2 Localized Attacks

In case of localized attack scenario, the set A_c of transmission lines are all adjacent to each other (i.e., have physical connection in the topology of the system). Figure 5.1b represents an example of a localized attack scenario on IEEE 118 test case.

5.3 MMSE Under Joint Cyber and Physical Stresses

After a joint cyber and physical attack occurs in the system, we use the collected PMU dataset and apply MMSE to estimate the status of the transmission lines (branches) in the unobservable portion of the grid. We specifically use a linear MMSE estimation model, where the unobservable portion of the grid is the estimation target and is denoted by $\underline{Y}$. The size of vector $\underline{Y}$ is equal to $|A_c|$, where $|.|$ represents the cardinality of the set. The elements Y_is of $\underline{Y}$ represent the power flow through the unobservable transmission lines in the attack zone. In this section, we use the real power flow through the lines to identify the physically attacked/failed lines. The rest of the information outside the attacked zone provided by the PMUs are considered as the estimation features $\underline{X}$, where the size of vector $\underline{X}$ is given by $(|L| - |A_c|) * f$ and f is the number of feature parameters to be used. Specifically, in this section we consider three possible feature parameters including real and reactive power flow and the phase angel. We can use a single feature parameter or a combination of them as well as certain lines or all the lines as a part of our estimation features.

The linear MMSE model suggests that our estimation of $\underline{Y}$ is related to features through $\hat{\underline{Y}} = \mathbf{A}\underline{X} + \mathbf{B}$, where matrix $\mathbf{A}$ and column matrix $\mathbf{B}$ can be characterized based on the data such that estimation error is minimized. Specifically, for linear MMSE we have matrix $\mathbf{A} = \mathbf{R}_{XY}\mathbf{R}_X^{-1}$, where
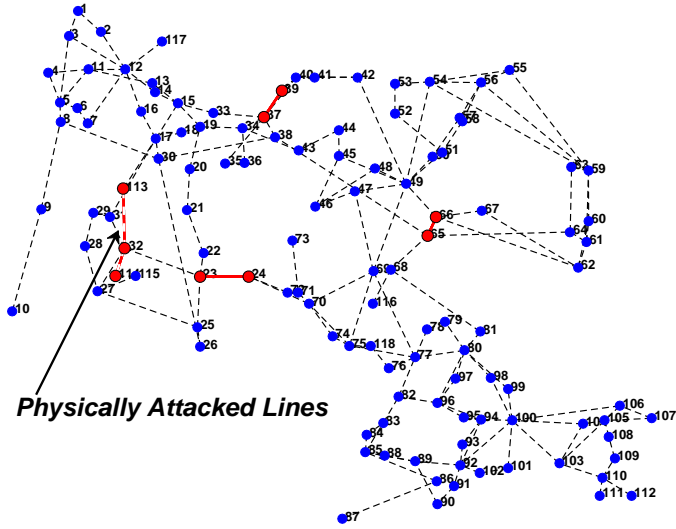
(a) Scattered Attacks
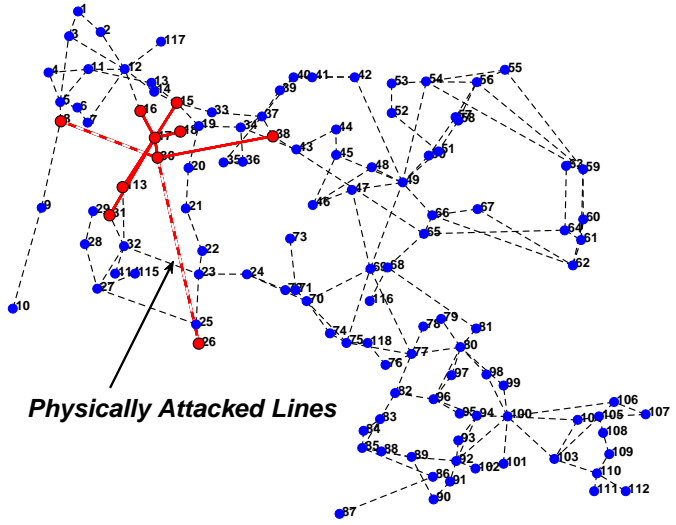
(b) Localized Attacks

Figure 5.1: Example of (a) scattered attack scenario, and (b) a localized attack scenario. The red marked branches have experienced cyber attacks and became unobservable and the red dashed lines indicate branches, which are physically attacked.

the matrix $\mathbf{R}_{XY}$ and $\mathbf{R}_X$ are the cross-correlation and auto-correlation matrices and $\mathbf{B} = \underline{\mu}_Y - \mathbf{A}\underline{\mu}_X$, where $\underline{\mu}_X$ and $\underline{\mu}_Y$ are the mean of the variables $\underline{X}$ and $\underline{Y}$.

An important observation based on our simulations is that when a subset of the grid branches changes their status (e.g., fail), not all the other lines will be effected equally due to such changes. For example in Figure 5.1b, changes inside the red portion of the grid (e.g., failure in the attack

zone) does not equally affect other branches outside the attack zone. Figure 5.2 shows a heatmap of the real power flow changes in all transmission lines due to the changes in the status of components inside the attack zone. This result is obtained based on 250 different scenarios with multiple combinations of failed transmission lines inside the attack zone. Based on this observation we can conclude that to determine the status of the unobservable components, one does not need data on all other branches outside this zone. Thus we can use a feature selection mechanism based on such analyses, which will allow selecting features with the most information to ease the computational complexity.
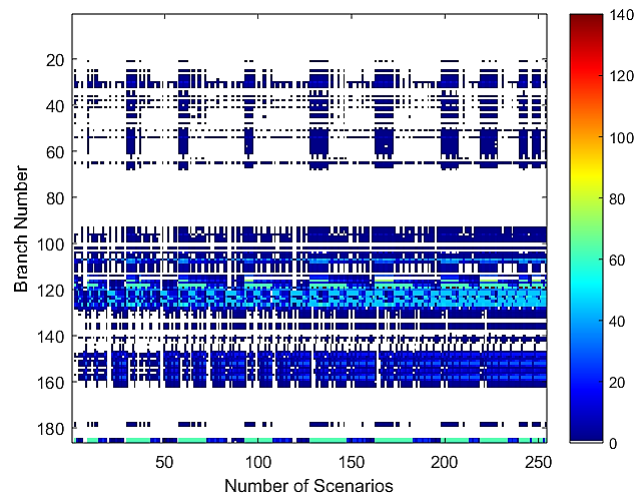
Figure 5.2: Power flow changes in IEEE 118 branches due to state changes in attack zone shown in Figure 5.1b. The state changes include failure of different combinations of lines inside the attack zone.

5.4 Iterative MMSE with Feedback

As we will show in Chapter 6, when multiple lines becomes unobservable, it gets more difficult to estimate the status of the line with an acceptable confidence level. To improve the performance of the estimation, we can use a feedback mechanism in the linear MMSE to use the components that are easier to estimate their states as features for estimating the rest of the components. Note that this requires a pre-assessment of estimation capabilities for various components of the system,

which can be a cumbersome task. In this section, we assume such information is available and has been pre-evaluated for the components and thus the focus is on the concept of the feedback MMSE.

In this approach, we assume that if the status of a subset of unobservable branches can be estimated with 90% confidence level (this level can be adjusted) then this subset will be used in the next iteration of the estimation as additional estimation feature (as a part of vector $\underline{X}$). This means that the estimated components's states with a predefined confidence level will no longer be a part of the attack zone and thus the attack zone shrinks to a smaller size, which we again assume that we know the estimation capability for the components inside the new attack zone. This feedback process can continue until the status of the whole unobservable portion is estimated or no further improvements are feasible.

5.5   Summary

In this chapter, we considered the effects of a combined cyber and physical stress scenario on the estimation of the state of power system components. We considered scenarios, which can lead to the unobservability of a portion of the grid and causes failures in transmission lines. In order to estimate the state of the unobservable component of the grid when it was subjected to cyber stressors, we used a data-driven technique using the data obtained from PMUs located outside of the attack region. In particular, the linear MMSE method introduced in Chapter 3 was adopted. To further enhance the performance of the estimation, another strategy that we suggested was iterative estimation combined with feedback. In addition, we took into consideration two other kinds of stress situations, which were localized attacks and scattered attacks. This study demonstrates the significance of data as well as the efficacy of data analytics tools in the context of defending smart grids against combined cyber and physical attacks.

Chapter 6: Numerical Analysis and Performance Evaluation

6.1   Introduction

Empirical evaluation is important to understand the effectiveness of any theoretical methodologies. This chapter evaluates the data-driven approaches for state estimation described in Chapters 3 and 4 under centralized and multi-region settings. Both linear and non-linear state estimation techniques are considered in a wide range of scenarios with various number of regions and information sharing methodologies. Under centralized techniques MMSE, BMLAR, T-GCNN, KF, EKF, and UKF are evaluated and under distributed state estimation fully distributed KF, multi-region distributed BMLAR, KF, EKF, and UKF, and multi-area distributed with collaborations among the areas for KF, EKF, and UKF have been evaluated. Regions are defined to consist of buses and their connections based on specific properties, such as their geographic proximity, state correlations, or other data-driven or physics-based metrics of interactions among the buses. The collaborative distributed state estimation described in Chapter 4 has been evaluated through overlapped regions and information sharing using auxiliary nodes with statistical attributes reflecting the neighboring region's states, such as mean, variance, and correlation among the states and their principal components. The performance of the presented models under these scenarios and collaboration settings has been evaluated using the IEEE 118 bus system and is compared with other data-driven and machine-learning-based methods in the literature.

In all the evaluations, the IEEE 118 bus system has been used. We have simulated a large dataset of PMU time series in both normal and also under partially unobservable scenarios using MATPOWER [143] simulation toolbox. The considered unobservable scenarios (due to physical failures or DoS stresses) are described details in later sections. We have used real load profiles from the New York Independent System Operator (NYISO) and sampled at 30*Hz* to generate

PMU time series by solving power flow at each sample. From the simulation, the bus phase angle time series, voltage magnitude, real, and reactive power flow have been recorded as the measurements. The actual grid topology also presented in the form graph data, where each bus have been presented as graph nodes and each transmission line as graph edges. To easily keep track of the various state estimation techniques, the fully distributed state estimation is denoted as Case-I, centralized estimation processes as Case-II, multi-region state estimation without information sharing/ collaboration as Case-III, and with information sharing as Case-IV.

6.2 Centralized MMSE for State Estimation Under Cyber and Physical Stresses

The MMSE method described in Chapter 3, and MMSE under joint cyber and physical stresses in Chapter 5 has the trainable parameter A and B that can be learned using the simulated data as described above. The joint cyber and physical stress is a scenario when a portion of the grid is unobserbable duo to cyber attack or sensor failure and among that unobservable set of buses, a subset of buses could actually be physically failed. Two such scenario such as such as (i) unobservable buses are randomly scattered throughout the network, (ii) unobservable buses are confined in a geographical location have been evaluated here. The MMSE method with various combination measurements features have been applied to unmask the unobservable and detect the physically failed lines in the process. The iterative extension to the MMSE model have also been evaluated and case studies from each scenario have been presented.

6.2.1 Randomly Scattered Attacks

To evaluate the performance of our trained estimator under the scattered attack scenarios, we create randomly scattered attacked zones (where the lines under cyber attack are geographically distant as shown in Figure 5.1a. We specifically create attack zones of size one to seven (while larger attack zones are possible, we assume that attackers have limited resources and the size of the attack zones are relatively small compared to the size of the grid.) We represent the attack zones with size i by F_i, representing the unobservable components under cyber attack. In each of the

randomly generated attack zones, there might be any number ($\leq i$) of physically failed lines. For each size of attack zone, we have generated 250 random attack zones. The average estimation error for each size of the attack zone is presented in Figure 6.1 when different features are used in the estimation. We observe that the estimation error increases with attack size and combined features gives the best estimate for the power flow status of branches.
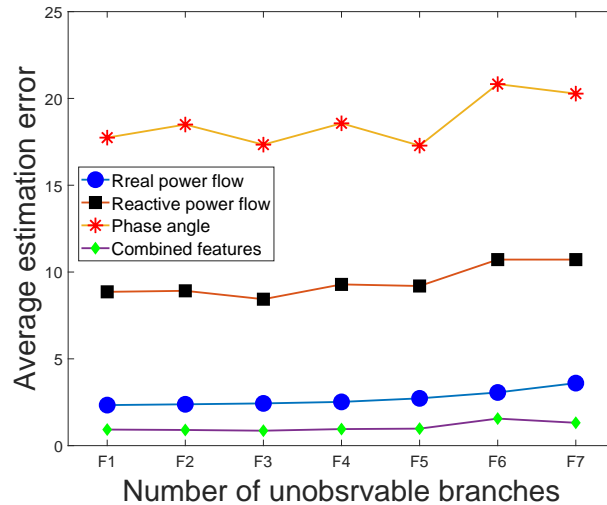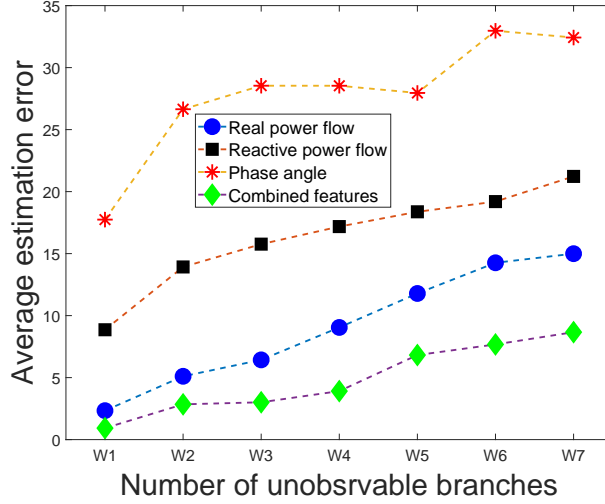

Figure 6.1: Average estimation error using different features for randomly scattered attack

6.2.2 Localized Attacks

To evaluate the performance of our trained estimator under the localized attack scenarios (as shown in Figure 5.1b), we generate attack zones with topologically adjacent lines under cyber attack. We call these attack zones, windows and consider sizes of one to seven for the attack zone. In this case, we represent the attack zones with size i by W_i. Similar to the scattered attacks, we generated 250 random scenarios of localized attacks for each window size. The average estimation error for the localized attacks is shown in Figure 6.2, when different features are used in the estimation. From the results, we observe that the estimation error increase with the attack size and the combined features give the best estimate for the power flow of the branches.

Figure 6.2: Average estimation error using different features for localized attack.

6.2.3 Detection of Physical Stresses

To evaluate the performance of the estimator in detecting the failed or physically attacked components in the attack zone, we have evaluated the average detection rate for both scattered and localized scenarios, where the failure is identified when the power flow through the line is estimated to be below certain threshold. The results are shown in Figure 6.3. From the results, we observe that the detection rate is lower for localized attacks (dashed lines) than the scattered attacks (solid lines). This is because when a transmission line is affected by a physical attack or failure, usually the adjacent lines will bear the most impact (as shown in Figure 5.1b) and thus the most information to help with the estimation. In the localized attack scenarios, since information from a portion of the locally adjacent lines are unavailable (due to the cyber attack), estimating the state of components in the attack zone is more difficult.

Note that one of the key observations that we obtained from our estimation results is that the estimation performance is different for various transmission lines. The results in Figures 6.1, Figure 6.2 and Figure 6.3 show the average performance over all transmission lines, while Figure 6.4 shows the average performance of estimation for the individual lines in an attack zone. In the notations used in this results F_{ij} and W_{ij} denote the jth transmission line in attack zone/window F_i and W_i). The results are shown for attack size seven.
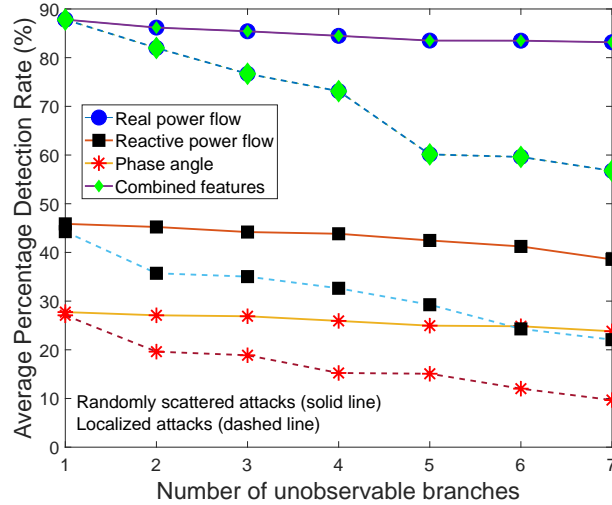
Figure 6.3: Average failure detection rate for different attack sizes using different features.
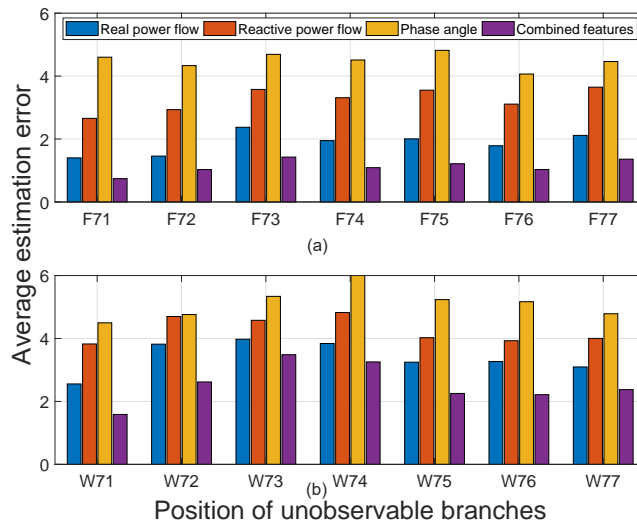
Figure 6.4: Average estimation error for individual lines in attack zone of size seven for (a) a randomly scattered attack, (b) a localized attack using different features.

6.2.4 Iterative Estimator with Feedback

The results in Figure 6.4, suggests that due to the power system attributes and topological location of the lines, it is easier to recover the unobservable information on the state of some lines. Identifying such components using similar studies can help in the iterative estimator with feedback discussed in section 5.4. In this section, we present two examples of attack zones with

such components that can help improving the estimation of the rest of the components in the attack zone. Although the results for the iterative estimator with feedback are very dependent on the attack zone and the pre-assessment of our estimation capability for lines, these examples show how the approach can help the state information recovery with such information in an iterative process. In these examples, we use the linear MMSE estimator to find the status of the lines that we know they can be estimated with 90% confidence rate. We will then update the attack zone size and use the estimated states in the previous step as new features for estimation. The iterative process will go on until all components are estimated with 90% confidence rate or we cannot improve the estimation confidence for the remaining components. The steps of the process for a scattered and localized attack are presented in Figure 6.5. In addition, feature selection using maximum variance in the data (as shown in Figure 5.2) is also applied to eliminate the unnecessary PMU data for the lines that were not impacted by the changes in the state of the components in the attack zone to ease the computational complexity. These examples show that the iterative estimation with feedback can improve the estimation in the attack zone. However, due to high computational complexity of pre-assessments it requires further studies to alleviate this limitation.

6.3   Centralized and Multi-Region BMLAR Process for State Estimation

The previous state estimation method did not consider the measurement as time series. Whereas the real world measurements are highly correlated multivariate time series data. To incorporate these features into estimation process (as described in section 3.3.3) Baysian regression combined with auto-regressive processes haev been introduced. Using the simulated data the spatial cross-correlations of state variables have been calculated. Using the partitioning strategy discussed in Chapter 4, the grid topology is divided into non-overlapping regions. We have solved the partitioning in both non-homogeneous and homogeneous partition sizes using 1) geographical distance (G), 2) geographical distance and cross-correlation (GC), 3) geographical distance in homogeneous

	Feedback	Unobservable Lines	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration
Randomly Scattered Attacks	Before	F7 – [B28, B49, B102, B171, B175, B177, B183]	(B28, 100%), ***(B49, 1.75%)***, (B102, 100%), ***(B171, 0%)***, ***(B175, 0%)***, ***(B177, 0%)***, (B183, 100%)			
		F6 – [B26, B74, B92, B145, B149, B166]	(B26, 100%), ***(B74, 60%)***, (B92, 100%), (B145, 100%), (B149, 100%), ***(B166, 74%)***			
	After	F7 – [B28, B49, B102, B171, B175, B177, B183]	(B28, 100%), ***(B49, 1.75%)***, (B102, 100%), ***(B171, 0%)***, ***(B175, 0%)***, ***(B177, 0%)***, (B183, 100%)	***(B49, 1.75%), (B171, 0%), (B175, 0%)***, (B177, 95.75%)	(B49, 100%), (B171, 100%), (B175, 100%)	0
		F6 – [B26, B74, B92, B145, B149, B166]	(B26, 100%), ***(B74, 60%)***, (B92, 100%), (B145, 100%), (B149, 100%), ***(B166, 74%)***	(B74, 100%), (B166, 100%)	0	0
Localized Attacks	Before	W7 - [B21, B22, B23, B36, B39, B40, B42]	***(B21, 0%), (B22, 0%), (B23, 0%)***, (B36, 100%), ***(B39, 0%), (B40, 0%), (B42, 0%)***			
		W6 – [B45, B47, B48, B49, B50, B51]	***(B45, 25%), (B47, 0%), (B48, 0%), (B49, 0%), (B50, 0%)***, (B51, 100%)			
	After	W7 - [B21, B22, B23, B36, B39, B40, B42]	***(B21, 0%), (B22, 0%), (B23, 0%)***, (B36, 100%), ***(B39, 0%), (B40, 0%), (B42, 0%)***	(B21, 100%), ***(B22, 0%)***, (B23, 100%), ***(B39, 0%), (B40, 0%), (B42, 0%)***	(B22, 100%), ***(B39, 74%), (B40, 87%), (B42, 49%)***	(B39, 100%), ***(B40, 87%), (B42, 54%)***
		W6 – [B45, B47, B48, B49, B50, B51]	***(B45, 25%), (B47, 0%), (B48, 0%), (B49, 0%), (B50, 0%)***, (B51, 100%)	***(B45, 25%), (B47, 0%), (B48, 0%), (B49, 0%)***, (B50, 100%)	***(B45, 25%)***, (B47, 100%), (B48, 98%), (B49, 100%)	***(B45, 25%)***

Figure 6.5: Examples from localized attack and a scattered attack scenario with and without the iterative estimation with feedback. The data pairs shown in each column represent the line number and their detection rate. Bold-underlined values show the components with low detection rate in each iteration.

partitioning (GS), and 4) geographical distance and cross-correlation in homogeneous partitioning (GCS). We have compared the performance of estimation for these partitioning scenarios.

The performance of DSE in terms of average mean square error (MSE) is depicted in Figure 6.6 as a function of the number of regions for different partitioning strategies in both homogeneous and non-homogeneous partition sizes. The average MSE is taken over all the buses of the system. In this figure, the dashed line is the estimation error if the model was to be applied centrally using the data from the whole system. From the figure, we see that partitioning the grid can improve the overall estimation performance depending on the number/size of partitions. As the defined regions consider the cross-correlation among the PMUs, in the smaller region sizes the feature space for the model is more densely correlated thus improving the overall average estimation accuracy. However, as the number of regions increases the model will have access to less information compared to the
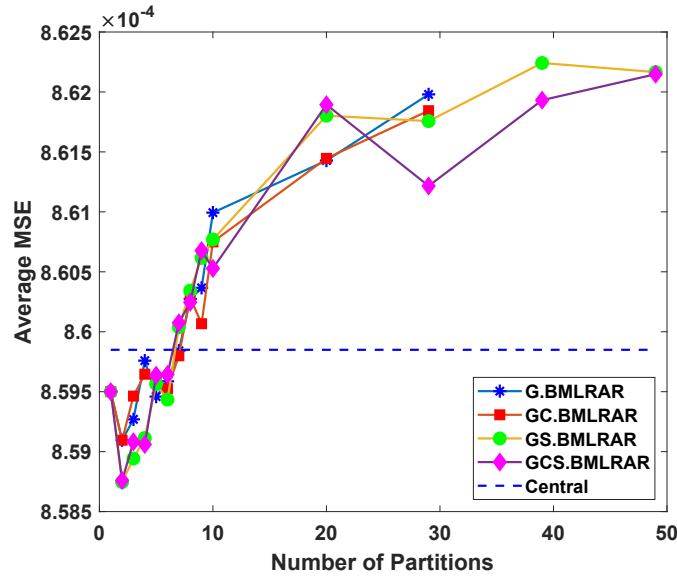
Figure 6.6: Average MSE over all the buses compared for different partitioning strategies and for different number of regions.

larger partition sizes thus the overall estimation error increases again. The variation of estimation error among different partitioning is small but the results show that incorporating correlation into the partitioning process slightly improves the overall estimation accuracy. The best estimation accuracy occurs when the number of partitions is homogeneous and small (less than ten). As such, we will focus on a small number of GCS-based partitions for the rest of the analyses.
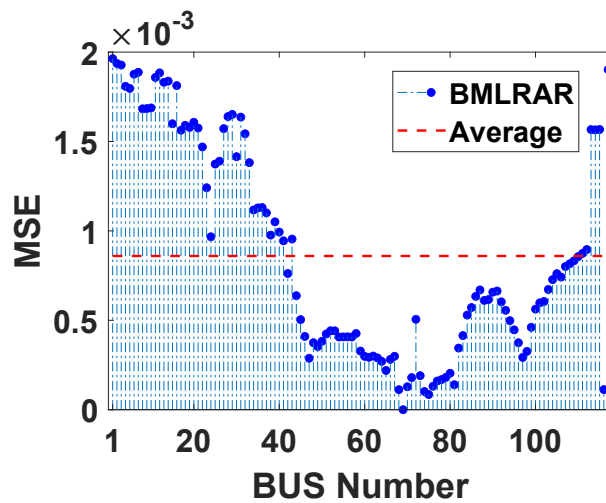
Figure 6.7: MSE at each bus for five regions, $R = 5$ and GCS partitioning technique.
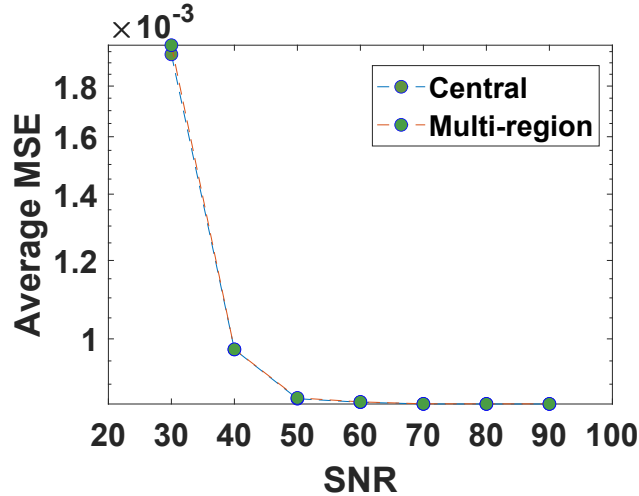
Figure 6.8: Average MSE as a function of added noise (SNR) for central estimation and multi-region estimation for *R* in the range of 2 to 10 and GCS partitioning technique.

Note that the estimation accuracy values are different at various buses of the system as shown in Figure 6.7 Specifically, the state of some buses is difficult to estimate, which can be due to their complex interaction dynamics with other components. The dotted line shows the average performance for all buses. Figure 6.8 shows the average estimation performance as a function of added Gaussian white noise to the PMU time series data. This added noise represents measurement noise or communication channel noise. As expected the performance increases with the increase in Signal to Noise Ration (SNR). The results also show that the estimator maintains good performance for SNR values larger than 40*db*. Since our results show small variations of estimation error under noise for different partition sizes, Figure 6.8 only depicts the MSE for the central case and for averaged over the number of regions in the range 2 to 10.

To evaluate the performance of the DSE under partial unobservability, we have considered scenarios that can be resulted from cyber stresses (such as DoS stresses) or the physical failure of PMU devices or disconnections in their communication links. In such cases, the local and the central servers will not receive any data or only receive the channel noise instead of the actual measurements from the PMU. We have considered two different scenarios of partial unobservability: a) when the unobservable PMUs are scattered throughout the grid randomly, and b) when the unobservable
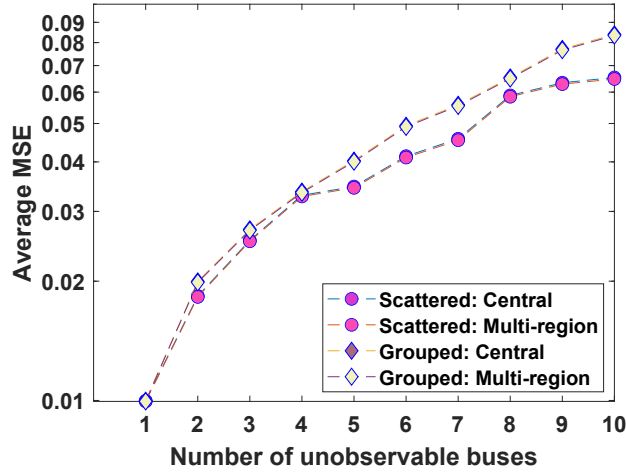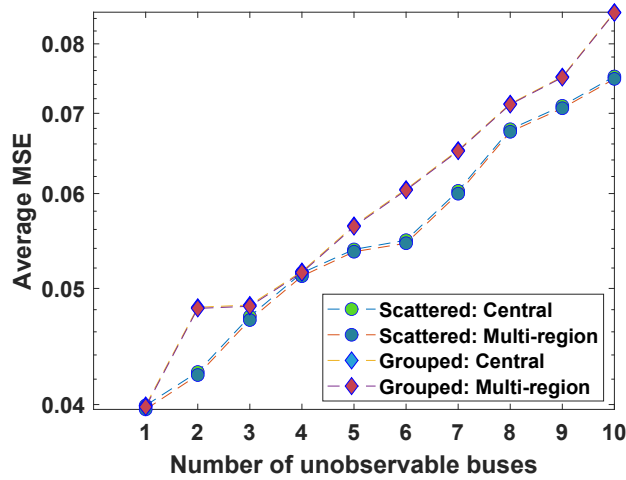
Figure 6.9: Average MSE for different stress sizes (number of unobservable buses) in two different stress scenarios (scattered and grouped). The results are shown for the central estimation and averaged multi-region estimation for R from 2 to 10 and GCS partitioning technique.

PMUs are localized geographically (for instance due to localized events such as earthquakes or attacks). We have simulated 100 stress scenarios for each case and each stress size (i.e., number of unobservable PMUs). The average MSE over all the buses under different stress scenarios for different stress sizes is represented in Figure 6.9. As the variations of estimation error are very small for different numbers of partitions, in this figure, we have presented MSE for the central estimator and for the multi-region estimator averaged over the number of regions in the range 2 to 10. It can be observed that the estimation error rises as the number of unobservable buses increases in both stress scenarios. However, the grouped stress cases impose more strain on the estimator as the number of correlated features becoming unavailable increases in a region. In the simulations, we have considered less than ten percent of the buses may become unobservable as a result of the stress; however, these results can be extended to any number of stressed buses. Overall the estimator retains promising performance under a small number of unobservable buses in both scenarios. Figure 6.10 shows similar results for the MSE of recovering the state of unobservable buses from other PMU data streams in both stress scenarios.

The estimation error of our proposed approach in this section, which is $\cong 8.6 \times 10^{-4}$ (averaged over all buses for the number of partition < 10), is comparable to the state of the art physics-

Figure 6.10: Average MSE of unobservable buses for different stress sizes (number of unobservable buses) in two different stress scenarios (scattered and grouped). The results are shown for the central estimation and averaged multi-region estimation for R from 2 to 10 and GCS partitioning technique.

driven estimation techniques. For instance, the performance (the average estimation error over all the buses) of the adaptive multi-region distributed Quasi-newton (A-DQN) algorithm presented in [144] is $\cong 5.8 \times 10^{-4}$, which is slightly better than the proposed technique in the current method due to collaboration consideration between the regions. However, the proposed technique in the current method outperforms the physics-driven hybrid linear multi-region state estimation techniques [145], which is based on the traditional weighted least square approach, with a performance of $\cong 1.38 \times 10^{-2}$ for average estimation error over all the buses.

6.4 Centralized and Multi-Region Data-Driven KF-based State Estimation

In this section centralized and multi-region distributed state estimation based on variations of KF along with linear and non-linear system identifications have been evaluated. The fully distributed KF implementation is discussed as a special case of mult-region where each region has only one element. The measurement data $x(V, \theta, \mathbf{p}, \mathbf{q})$, are utilized in the form of multivariate time-series. The cross-correlations of state variables among the buses have been calculated and used in the partitioning process (as discussed in Chapter 4) to divide the system topology into multiple regions.
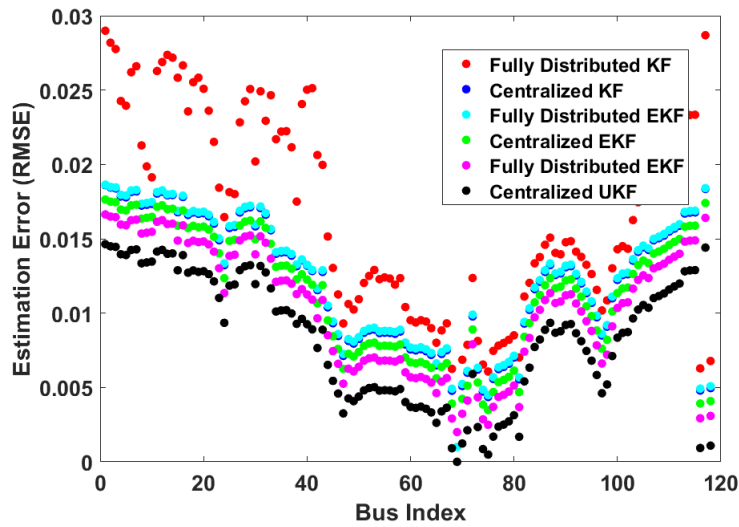
Figure 6.11: The state estimation error at each bus for fully distributed and centralized estimation using MLR-KF, MPR-EKF, and MPR-UKF.

Following the discussion in Chapter 3 (Linear and Non-linear KF) and Chapter 4, the state estimation problem has been considered for a range of scenarios from fully centralized to multi-region and fully distributed cases (labeled, Case-I to Case-IV). Moreover, various information-sharing mechanisms for collaborative state estimation among multi-region (as discussed in Case-IV) have been considered for the performance evaluation studies. These scenarios have been considered for both linear and non-linear KF models. The results of the first study presented in Figure 6.11, show the estimation error for the two extreme cases where the state estimation is performed fully centralized or fully distributed. As expected, the estimation error is higher for the fully distributed estimation case (Case-I) both for linear and non-linear KF models, which is due to limited information used for estimation at each individual bus. On the other hand, the fully centralized cases (Case-II) perform better due to full observability and full-information access at all the buses. However, in the centralized cases, it is expected that the measurements from the geographically distributed components of the system are communicated to a central system for building models and developing functions such as state estimators. Communicating those measurements to a remote central system naturally imposes higher communication costs than localized models, which only use measurements from a local region in a distributed fashion.
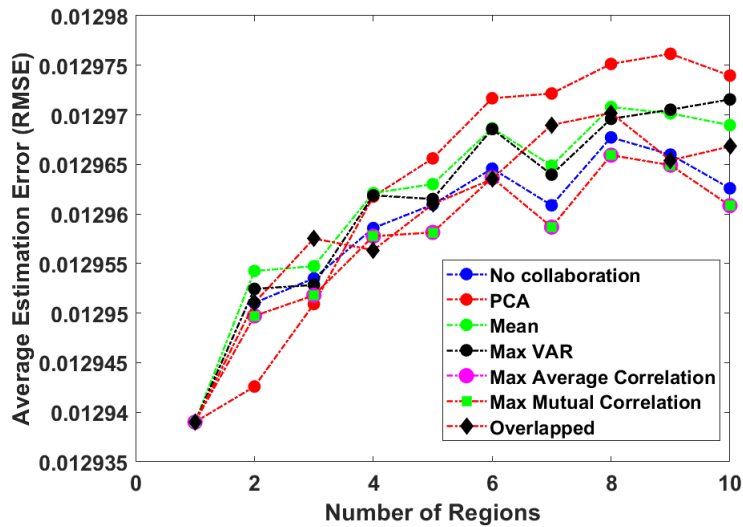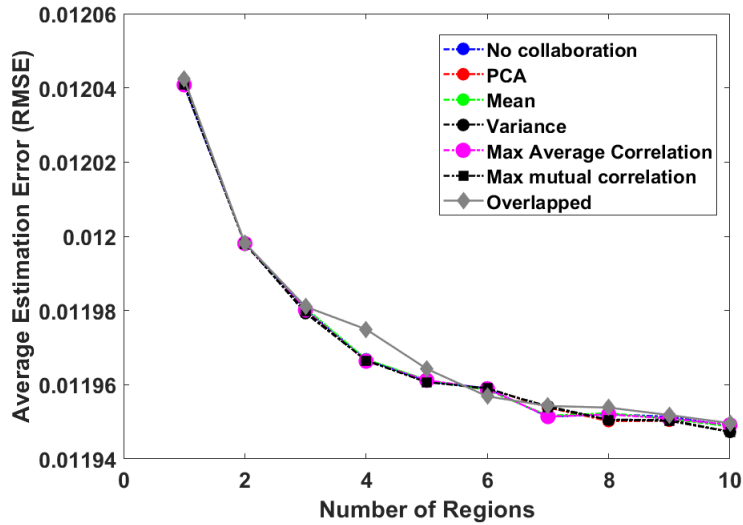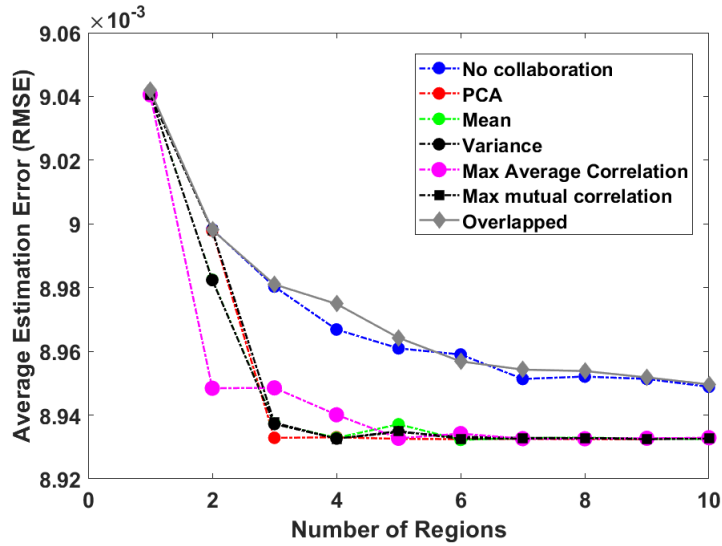
Figure 6.12: Average estimation error over a varying number of regions for a different mode of information sharing among the regions for MLR-KF.

Studies have shown that lower dimensionality in the distributed estimation process results in low communication cost [146]. In [8, 19], it is also discussed that distributed state estimation enabled by new technologies, such as edge computing, can reduce computational cost over centralized processing of data. Moreover, it can be observed that non-linear KF models, particularly UKF, perform better in estimation by better enabling the capture of non-linear dynamics of the system.

Next, the state estimation performance for multi-region setting is considered both under a non-collaborative setting (Case-III) and various information-sharing mechanisms under Case-IV for the linear KF model. From Figure 6.12, it can be observed that as the number of regions increases, the performance of the state estimation decreases. Note that here the performance is calculated by taking the average estimation error over all the buses of the system. Moreover, it can be observed that allowing information sharing by overlapping a small subset of buses among the regions in Case-IV(a) overall results in better accuracy than in Case-III, which does not consider information sharing. Another approach of information sharing discussed in Case-IV(b) is to allow sharing of a few resenting features from each region to help with state estimation. Among the various features discussed in Case-IV(b), sharing the state data of the bus in the region *i* with a maximum average

correlation with the buses in the region *j* (named, Max Mutual Correlation case) results in better performance on average and particularly for the number of regions higher than four.

Figure 6.13: Average estimation error over varying number of regions for different mode of information sharing among the regions for MPR-EKF.

Figures 6.13 and Figure 6.14 also depict the performance of various modes of information sharing over multi-region distributed state estimation for EKF and UKF, respectively. A key observation here is that in the case of EKF and UKF with non-linear system approximations, the estimation error is decreasing with the increase in the number of regions. This can be explained through the characteristics of the average approximate entropy for different number of regions as shown in Figure 6.15. Specifically, it has been discussed that entropy can be an indicator of the level of non-linearity in the system [147, 148]. As can be observed from Figure 6.15, the average entropy over regions increases with the increase in the number of regions. Since, large number of regions are not generally practical to consider for power systems, the trend observed for small number of regions (up to ten) is considered in the studies in this method. The result in Figure 6.15 suggests that as the number of regions increase the non-linearity increase and as such EKF and UKF performance show slight improvements compared to the opposite trend observed in Figure 6.12 for the linear KF model. Table 6.1 presents the average performance of various information sharing

Figure 6.14: Average estimation error over varying number of regions for different mode of information sharing among the regions for MPR-UKF.

approaches (i.e., average over various number of regions ranging from 1 to 10). From the table, it can be observed that maximum mutual correlation based information sharing approach results in the lowest estimation error for KF, EKF, and UKF. This result suggests that in the multi-region state estimation setting, sharing the state information of the bus with the maximum correlation with the buses of the other region can help improving the performance of the model, while still only using local information.

While Table 6.1 shows the average performance with small differences, if Figures 6.12, 6.13, and 6.14 are considered, it can be observed that for different number of regions, the performance gain from different information sharing approach are different. Specifically, a comparatively wider range of variations in performance for different information sharing approaches for different regions can be observed. From Figure 6.16 we can also see that specific information sharing approaches show different performance for different number of regions.

Next, the performance of the proposed techniques in this section are compared with some of the existing methods in the literature including historical averaging (HA) [149], support vector regression (SVR) [150], MMSE state estimation method [14], BMLAR [12], temporal graph T-
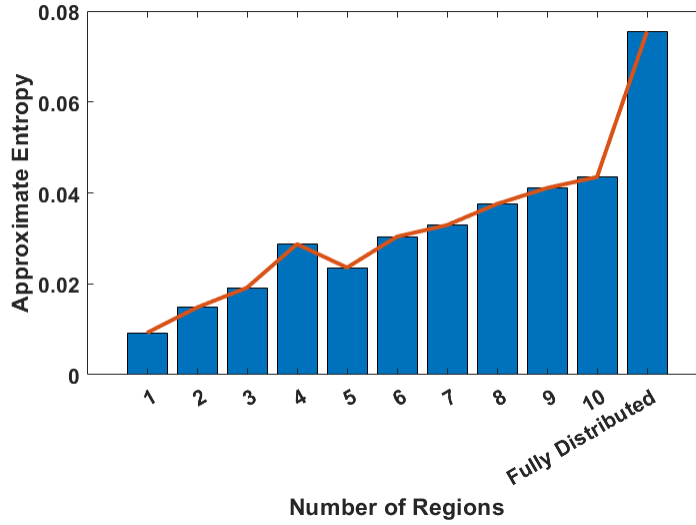
Figure 6.15: Average approximate entropy as a measure of system non-linearity for different number of regions.

Table 6.1: The average RMSE for state estimation using MLR-KF, MPR-EKF, and MPR-UKF for various information-sharing techniques.

Methods	MLR-KF	MPR-EKF	MPR-UKF
No Collaboration	0.012963	0.011971	0.00897
PCA	0.0129616	0.0119706	0.00895
Mean	0.0129612	0.0119709	0.0089499
Variance	0.012961	0.0119707	0.0089493
Self Correlation	0.0129572	0.0119708	0.0089491
Mutual Correlation	**0.012957**	**0.01197**	**0.0089475**
Overlapped	0.012959	0.011972	0.008972

GCNN [13]. For comparison purposes, these techniques are applied to two scenarios, including centralized state estimation and multi-region state estimation with four regions. The number of regions are selected as four as it shows a good balance of state estimation performance based on the results presented in previous figures and also from practicality perspective. Also, no information sharing is considered in this comparison as the techniques from the literature are mainly developed for centralized setting with no information sharing mechanism in their original form. The average estimation error in the form of RMSE for Linear-KF, EKF, and UKF are presented in Table 6.2 as 0.01295, 0.01197, and 0.00897, respectively, with UKF showing leading
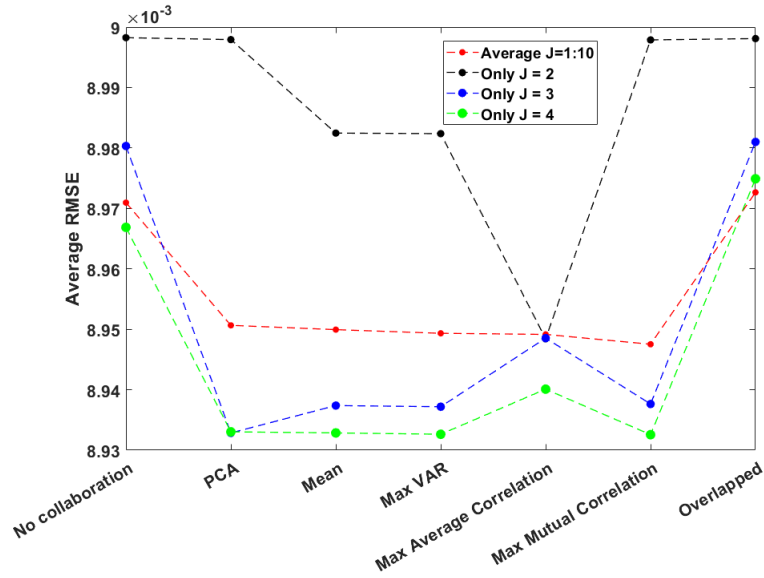
Figure 6.16: Performance of various information sharing mechanisms for various regions in comparison with average performance as shown in Table 6.1 for MPR-UKF.

performance. While T-GCNN shows slightly better performance compared to MPR-UKF, it uses topological information (system's adjacency matrix) as an input. However, the availability of accurate topological information in certain cases cannot be considered. Additionally, the sensitivity to noise has been evaluated for SNR 90dB, 60dB and 30dB and the results in Table 6.3 show that linear-KF is more susceptible to noise compared to the non-linear KF models. However, the proposed methods retain acceptable performance up to 60dB of SNR.

Table 6.2: The average RMSE for various state estimation techniques for the centralized and multi-region state estimation with four regions and no information sharing.

Models	Centralized	Multi-region with No Information Sharing
HA	0.7950	0.0791
SVR	0.04290	0.0435
MMSE	1.2530	1.3394
BMLAR	0.04000	0.0348
T-GCNN	0.00534	0.009
MLR-KF	**0.0129**	**0.01295**
MPR-EKF	**0.01206**	**0.01197**
MPR-UKF	**0.0091**	**0.00897**

Table 6.3: The average RMSE for various noise levels for the proposed state estimation techniques with four regions and no information sharing.

Models	Multi-region with No Information Sharing	Multi-region with No Information Sharing and 90dB SNR	Multi-region with No Information Sharing and 60dB SNR	Multi-region with No Information Sharing and 30dB SNR
MLR-KF	0.01295	0.0183	0.01834	0.03645
MPR-EKF	0.01197	0.01197	0.011976	0.011977
MPR-UKF	0.00897	0.00897	0.008976	0.008978

6.5 Centralized TGCN for State Estimation

State estimation algorithms that have been discussed in previous section did not consider any spatial information such as the network topology of the smart grid. Whereas incorporating the spatio information may greatly improve over the data-driven state estimation techniques. In this section the T-GCNN discussed in Chapter 3 will be evaluated. From the simulation, time-series measurements of real power flow, reactive power flow, voltage angle, and voltage magnitude have been recorded. We have tested the T-GCNN algorithm for two scenarios as follows.

6.5.1 Scenario-I: Full Set of Measurements Are Available

In this scenario, the assumption is that measurements are available at all the buses in the system. As such, this case is a multivariate time-series forecasting problem described as

$$X_{t+1} = F(Z_{t-p}, A,). \ [p \in \mathbb{N}] \tag{6.1}$$

6.5.2 Scenario-II: A Subset of Measurements Are Available

In this scenario, the assumption is that measurements are available only at a subset of buses (i.e., at $n \in \mathcal{N}^o$), which can be identified, for instance, using a PMU placement strategy to ensure full

observability of the system. Three different PMU placement strategies for the IEEE 118 bus system have been adopted and considered from [151] as shown in Table 6.4 to evaluate the performance of the presented state estimation technique for different availability of the measurements. Note that the measurement at the rest of the buses are modeled as white Gaussian noise, which can, for instance, represent the channel noise. The state estimation process in this case will estimate the state of all the nodes from the available measurements of $\mathcal{N}^o$ buses along with one-step ahead state prediction.

Table 6.4: Three PMU placement strategies for the IEEE 118 test case system

PMU Placement Strategy	BUS index
O_1	2, 5, 10, 11, 12, 17, 20, 23, 25, 29, 34, 37, 40, 45, 49, 50, 51, 52, 59, 65, 66, 71, 75, 77, 80, 85, 87, 91, 94, 101, 105, 110, 114, 116
O_2	1, 5, 10, 12, 13, 17, 21, 25, 28, 34, 37, 40, 45, 49, 52, 56, 62, 63, 68, 70, 71, 75, 77, 80, 85, 87, 90, 94, 102, 105, 110, 114
O_3	1, 4, 5, 6, 8, 9, 10, 11, 12, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 30, 32, 34, 37, 40, 43, 45, 49, 50, 56, 59, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 72, 73, 75, 77, 79, 80, 83, 85, 87, 89, 90, 92, 94, 96, 100, 101, 105, 106, 108, 110, 111, 112, 114, 116, 117, 118

Table 6.5: T-GCNN model simulation parameters

Key	Value
GRU units	64
p	7
learning rate	0.001
epoch	250
training ratio	0.8
batch size	32

The model parameters used for the evaluations are presented in Table 6.5. Specifically, for the hyper-parameter p, its effects on the performance of the model have been evaluated for up to 12 sequence lengths. Since the lowest RMSE is observed at $p = 7$ (as shown in Figure 6.17), and a large value of p will increase the model complexity and thus the execution time, this value is considered for the rest of the evaluations.
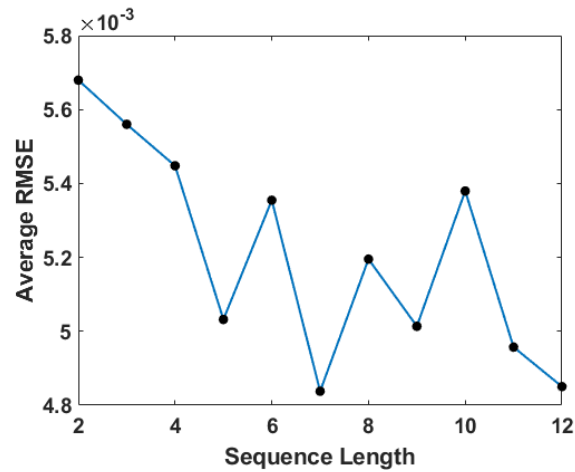
(a)
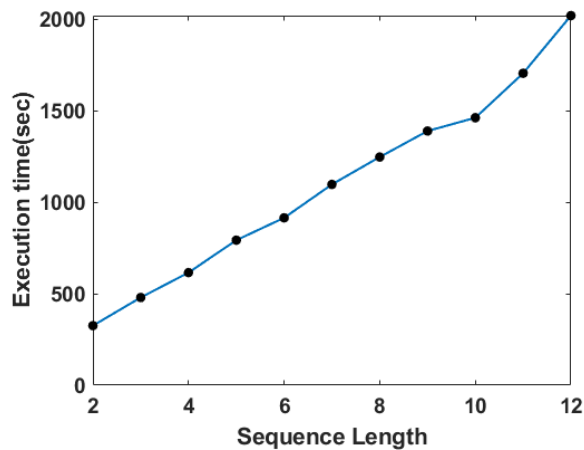
(b)

Figure 6.17: The evaluation of the impacts of sequence length, p, on the performance of the model based on (a) the average RMSE, and (b) the execution time (training and testing).

The performance of the T-GCNN for state estimation in the smart grid is compared with some baseline forecasting models for state estimation including History Average (HA) model, Support

Vector Regression (SVR) model, Minimum Mean Square Error optimization-based model, Bayesian Multivariate Regression combined with Auto-Regressive model, and LSTM. As can be observed from the results presented in Table 6.6, the presented T-GCNN shows superior performance in state estimation compared to the aforementioned models.

Table 6.6: The average RMSE for various state estimation techniques for the defined Scenario I and II for the availability of measurements.

Models	Scenario-I	Scenario-II(O_1)	Scenario-II(O_2)	Scenario-II(O_3)
HA	0.7950	0.89675	0.91307	0.70449
SVR	0.04290	0.84007	0.85523	0.66068
MMSE	1.2530	1.4510	1.53070	1.33950
BMLAR	0.04000	0.72340	0.74300	0.45660
LSTM	0.04473	0.99445	0.98533	0.53607
T-GCNN	**0.00533**	**0.01452**	**0.01278**	**0.01215**

Moreover, according to [63], the average RMSE for the state estimation in the IEEE 118 case based on Gauss-Newton, 6-layer Feed-forward Neural Network (FNN) and 8-layer FNN, and Prox-linear net are 4.71×10^{-2}, 1.645×10^{-3}, 2.366×10^{-3}, and 2.97×10^{-4}, respectively. Here, the two layer T-GCNN, which considers both the spatial (in the form of a graph) and temporal information shows a competitive performance of 5.33×10^{-3} for the scenario when full measurements are available. G-CNN also improves performance for Scenario-II, where only a subset of PMU measurements are available. For different subsets of available PMUs, performance slightly improves with larger number of available PMUs (as in Scenario:II-O_3).

6.6   Summary

In this chapter, the proposed data-driven, centralized, multi-region distributed, linear and non-linear state estimation frameworks for smart grids have been evaluated. These data-driven models are particularly helpful in supporting state estimation under missing or limited measurements, such as in cases of partial unobservability due to failures or attacks in the sensing and monitoring system

or limited availability of PMUs. The presented models are also evaluated under various distributed settings, which can support low-latency requirements of critical functions and reduce the communication overhead through state estimation using local data from each region. Various modes of information sharing among the regions have also been studied to improve the performance of distributed state estimation. It has been shown through numerical evaluations that the distributed state estimation using data-driven linear and nonlinear Kalman filters with selective feature sharing among the regions, particularly sharing the state of a node with the highest correlation with the other region, can lead to comparable results to centralized state estimation. The presented models also show a leading performance compared to some of the other data-driven and machine-learning models in the literature. Another key observation from Figures 6.7 and Figure 6.11 is that regardless of the type of algorithm being used, due to the dynamic nature, individual buses have different estimation accuracy. This could be due to the load variations or the non-Gaussian noise statistics present in the system. Among all the centralized estimation techniques discussed, T-GCNN led to the best performance, an RMSE of 0.0053, followed by the MPR-UKF at 0.0091, which indicates that the nonlinear algorithms have the ability to approximate actual system dynamics with more precision compared to the linear approximations. Additionally, the T-GCNN algorithm also has an added advantage as it can also incorporate spatial information such as grid adjacency matrix in the estimation process. This could also become a limitation if accurate spatial information is not available. However, modifications can be made to the T-GCNN algorithm to use an interaction-driven virtual topology rather than the actual grid topology, which needs further investigation in the future.

Chapter 7: Conclusion and Future Direction

7.1   Introduction

Wide-area situational awareness (WASA) is essential in operating the large scale power systems under normal and stresses situations. Tracking the state of the system to provide the necessary awareness about the operating conditions of the power systems is essential in efficient and reliable operation of power systems. State estimation is one of the key functions for this purpose, which has been studied in this dissertation.

Due to substantial penetration of distributed energy resources, such as distributed generation, demand-responsive loads, storage devices, and renewable energy generation, power systems have recently seen higher level of stochasticness, fluctuation and complexity. On the other hand, vast deployment of monitoring, sensing, communication and computation devices have produced a powerful platform to collect and process data from these systems. To improve situational awareness in these highly dynamics systems with tightly coupled cyber elements, a first step can be applying advanced real-time data processing techniques from machine learning and artificial intelligence domain to handle the large volume of the data for better tracking of the state of the system. As such, the focus of the work presented in this dissertation is on developing and studying data-driven state estimation techniques under various scenarios, ranging from central, multi-region, and fully distributed settings to support improved state estimation over large scale power systems both under normal and cyber and physical stresses scenarios.

7.2 Concluding Remarks

In this dissertation, we presented a review of existing conventional and data-driven state estimation techniques for power systems. We discussed that one of the key limitations of the conventional techniques is that they heavily rely on system models and in case of inaccuracies or missing information and highly dynamic systems, for instance due to renewable resources, these model may not produce accurate estimations of the system states. As such, we focused on data-driven techniques, both in the form of linear and non-linear models, which use the measurement data to track or recover the state of the system by approximating the non-linear dynamics of the power system without the need for system model.

We first discussed the data-driven models in a central setting, in which we assumed all the measurement from the components of the system are available in a central unit for processing and developing the state estimation models. Later, we presented a distributed multi-region state estimation framework that allows generating local state estimates under information sharing and without information sharing among the regions.

We started by discussing linear models, such as, MMSE-based state estimation, which allows recovering unobservable or partially observable states of the component from the measurements of the rest of the system. To capture the time dependent aspects of the measurements, a Bayesian technique combining Bayesian multiple linear regression with an auto-regressive process was presented next. Furthermore, given the KF family's reputation and demonstrated ability to express linear and nonlinear system dynamics, linear KF, extended KF, and unscented KF-based state estimation techniques have been proposed for use in conjunction with linear and nonlinear system identification techniques. The data-driven system identification approximated the power system dynamics in a data-driven manner, removing system model reliance for the state estimation process. Finally, a spatial feature, such as topological information, was considered that enables capturing the underlying relationships among the variables. A temporal graph convolutional network has been developed, which uses the graph topology of the system combined with temporal characteristics in the measurement time-series in its model.

In the distributed multi-region framework, we also discussed that the proposed framework can be implemented over a distributed computing platform, such as edge computing. The proposed framework can lead to low latency and faster data processing, which results in improved wide-area monitoring for smart grids. We also considered partially unobservable scenarios that can result from cyber or physical stresses on PMUs and showed that the distributed estimation approach can handle the estimations under such scenarios well. Further, power grid partitioning into regions with geographical and power system considerations (such as correlation among the PMU time series) was discussed. We considered both homogeneous and non-homogeneous region sizes in our study. We showed that the distributed state estimation can achieve better estimates compared to its central counterpart when the grid is partitioned into few regions.

In the distributed framework, in order to boost collaboration across surrounding regions, several information sharing methodologies were developed, which can limit data sharing among regions while improving accuracy of state estimation over no-collaboration. Specifically, it was shown through numerical evaluations that the distributed state estimation using data-driven linear and non-linear Kalman filters with selective feature sharing among the regions; particularly, sharing the state of a node with highest correlation with the other region, can lead to comparable results to centralized state estimation. The presented models also show a leading performance compared to some of the other data-driven and machine learning models in the literature as were discussed in the evaluation of the techniques.

Finally, a joint cyber and physical attack on smart grids was considered, which results in unobservablity of a portion of the grid while causing transmission lines failures. We used a data-driven approach to estimate the state of the unobservable portion of the grid under cyber attack from the PMU data available outside the attack area. Specifically, a linear MMSE approach was used and was trained based on the simulated PMU data. We also proposed the idea of iterative estimation with feedback to improve the estimation performance. Further, we considered two different types of attack scenarios including localized and scattered attacks and showed that estimating the state of components in a scattered attack is easier compared to localized attacks. This study shows the

importance and the power of data and data analytics methods in addressing joint cyber and physical attacks on smart grids.

7.3   Future Research Directions

The dynamics of power systems are governed by various physical and operational attributes of these systems. Modern power grids' dynamics are more complex and exhibit stochastic properties due to increasing deployment of new technologies for renewable resources and energy storage systems. As such, state estimation techniques that can capture such complex dynamics and stochastic behavior are yet under study. Moreover, distributed forms of state estimations are gaining more attention as their supporting technologies emerging and the need for them growing. On the other hand, the underlying interactions and interconnections among the components of power systems are important attributes that their effects are reflected in the measurement data from these systems; however, they are not yet captured enough in the state estimation models. Moreover, the form of underlying structure to be considered for power system structural interactions has been limited to its physical topology. Improving each of the aforementioned areas require research investments in these supporting domains. In this direction, we have identified few open research problems that can be the direction of future studies.

7.3.1   State Estimation in the Presence of Non-Gaussian Noise

Renewable energy and smart grid advancements are altering the generation and usage of electricity. This new combination begins to alter the character of electric grid dynamics. According to data, the electrical grid is suffering more frequent and severe frequency fluctuations during normal operation, which vary further from the steady-state assumption in classic static state estimation [152]. The stochastic variation introduced by distributed and renewable energy sources results in highly nonlinear and non-Gaussian system dynamics. Many dynamic state estimations are based on Kalman filters or a set of Gaussian filters that provide estimates of the state in terms of minimum mean squared error. However, modern power grids are becoming increasingly nonlinear

and non-Gaussian. Finding a closed-form solution for the state estimate under non-Gaussian noise is quite challenging, hence approximation techniques have been suggestion in such cases. Several studies have attempted to extend these models to non-Gaussian noise environments. For example, the researchers in [153] have shown that one strategy to handle such scenarios with non-Gaussian noise statistics is to decompose (or approximate) the noise using the sum of many Gaussian processes and then apply the Kalman filter model. In another technique, researchers [58] recommended creating redundant measurements from projected states using linearized batch mode regression and then filtering out data with non-Gaussian noise statistics using the generalized maximum likelihood (GM)-estimator. Chapter 3 presents nonlinear dynamic state estimation strategies; however, the robustness of these techniques in the presence of non-Gaussian noise requires further exploration.

7.3.2 State Estimation Over Data-Driven Power Grid's Graph of Interactions

The interconnected components with complex interactions within power systems make them complex networks that can be represented by graphs. In addition to the physical topology of the power grid, various data-driven and power physics-based methods have modeled and revealed the underlying graph of interconnections in smart grids [154]. In Chapter 3, a data-driven nonlinear graph-based state estimation method (T-GCNN) was presented, which allowed capturing the topological structure of the component interactions based on the physical structure of the power system as well as temporal information from the measurements. Though incorporating grid topology within the model can enhance the state estimation accuracy, accurate system topology may not always be available or there may be other interactions among the components that are not well-represented in the physical topology. To address this challenge, one solution could be to use a data-driven approach to learn the graph of interactions among the components of the system. In [154], it was demonstrated that multiple virtual typologies can be constructed while taking into account varying levels of interaction among the components. The advantage of interaction graphs is that interactions among components are topologically local, which simplifies the analysis, particularly during anomalous events. Inferring and learning the underlying graph of interactions

that best supports the state estimation process is another research direction that can improve state estimation.

7.3.3 Edge Computing Platform for Distributed State Estimation

One of the key research direction to support fast and effective state estimation in power systems is developing distributed computational platforms to support distributed state estimation. Edge computing can be a candidate platform for critical and time-sensitive applications in the monitoring and operation of power systems. In a collaborative work presented in [8], we explored one of the important operations for smart grid reliability, situational awareness, and discussed the role that edge computing can play to enhance this operation by enabling state estimation locally at the edge nodes. We notably focused on the network of PMUs as an example of the industrial internet of things in smart grids and discussed the edge-computing platform architecture to enable data analytics for state estimation using the PMU time-series. Furthermore, in [19], we have highlighted the promise of edge computing in supporting privacy conscious federated learning for distributed non-intrusive load monitoring function. In [108], researchers has also highlighted the significance of high speed communication technologies, such as 5G technology, to support distributed state estimation. Multi-region distributed state estimation techniques presented in Chapter 4 may benefit considerably from edge computing and sophisticated communication network technologies. As such, to make these techniques ready for implementation more research is needed in the direction of their enabling technologies.

7.3.4 Improved Information Sharing for Multi-Region State Estimation

Privacy and security has became increasingly important to maintain security from adversarial attacks. For example, in [106] researchers intended to protect the network database and the network communication channels against attacks and data manipulations via a blockchain (BC), an encryption-based system design. Additionally, the volume of shared information among multi-region is also important considering the limitation of distributed computational hardware or data

communication channel as discussed in Chapter 4. Our main goal in this dissertation was to show that the data-driven distributed state estimation approach with limited information sharing can achieve competitive performance compared to the central techniques. We used Kalman filter algorithms with simple statistical messages to evaluate the objective. Thus construction and selection of information to be shared with neighboring regions that can improve the distributed state estimation while maintaining privacy and security of the system suggest more investigation.

References

[1] X. Yu and Y. Xue. Smart grids: A cyber–physical systems perspective. *Proceedings of the IEEE*, 104(5):1058–1070, 2016.

[2] S. S. Ali and B. J. Choi. State-of-the-art artificial intelligence techniques for distributed smart grids: A review. *Electronics*, 9(6), 2020.

[3] M. Panteli and D. S. Kirschen. Situation awareness in power systems: Theory, challenges and applications. *Electric Power Systems Research*, 122:140–151, 2015.

[4] M. Panteli, P. A. Crossley, D. S. Kirschen, and D. J. Sobajic. Assessing the impact of insufficient situation awareness on power system operation. *IEEE Transactions on Power Systems*, 28(3):2967–2977, 2013.

[5] C. Alcaraz and J. Lopez. Wasam: A dynamic wide-area situational awareness model for critical domains in smart grids. *Future Generation Computer Systems*, 30:146–154, 2014.

[6] F. C. Schweppe and J. Wildes. Power system static-state estimation, part i: Exact model. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(1):120–125, 1970.

[7] F. C. Schweppe and Douglas B. Rom. Power system static-state estimation, part ii: Approximate model. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(1):125–130, 1970.

[8] M. A. Hasnat, M. J. Hossain, A. Adeniran, M. Rahnamay-Naeini, and H. Khamfroush. Situational awareness using edge-computing enabled internet of things for smart grids. In *2019 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2019.

[9] A. Adeniran, M. A. Hasnat, M. Hosseinzadeh, H. Khamfroush, and M. Rahnamay-Naeini. Edge layer design and optimization for smart grids. In *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGrid-Comm)*, pages 1–6, 2020.

[10] Y. Weng, R. Negi, C. Faloutsos, and M. D. Ilić. Robust data-driven state estimation for smart grid. *IEEE Transactions on Smart Grid*, 8(4):1956–1967, 2017.

[11] M. Netto and L. Mili. A robust data-driven koopman kalman filter for power systems dynamic state estimation. *IEEE Transactions on Power Systems*, 33(6):7228–7237, 2018.

[12] M. J. Hossain and M. Rahnamay-Naeini. Data-driven, multi-region distributed state estimation for smart grids. In *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pages 1–6, 2021.

[13] M. J. Hossain and M. Rahnamay–Naeini. State estimation in smart grids using temporal graph convolution networks. In *2021 North American Power Symposium (NAPS)*, pages 01–05, 2021.

[14] M. J. Hossain and M. Rahnamay-Naeini. Line failure detection from pmu data after a joint cyber-physical attack. In *2019 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5, 2019.

[15] X. Ji, Z. Yin, Y. Zhang, M. Wang, X. Zhang, C. Zhang, and D. Wang. Real-time robust forecasting-aided state estimation of power system based on data-driven models. *International Journal of Electrical Power Energy Systems*, 125:106412, 2021.

[16] A. Tajer, S. Kar, and H. V. Poor. *Distributed state estimation: a learning-based framework*, page 191–202. Cambridge University Press, 2012.

[17] H. Ma, Y. Yang, Y. Chen, and K. J. R. Liu. Distributed state estimation in smart grid with communication constraints. In *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, pages 1–4, 2012.

[18] M. J. Hossain and M. Naeini. Multi-area distributed state estimation in smart grids using data-driven kalman filters. *Energies*, 15(19), 2022.

[19] N. Hudson, M. J. Hossain, M. Hosseinzadeh, H. Khamfroush, M. Rahnamay-Naeini, and N. Ghani. A framework for edge intelligent smart distribution grids via federated learning. In *2021 International Conference on Computer Communications and Networks (ICCCN)*, pages 1–9, 2021.

[20] M. R. Karamta and J. G. Jamnani. A review of power system state estimation: Techniques, state-of-the-art and inclusion of facts controllers. In *2016 International Conference on Electrical Power and Energy Systems (ICEPES)*, pages 533–538, 2016.

[21] J. Zhao, A. Gómez-Expósito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang, and A. P. S. Meliopoulos. Power system dynamic state estimation: Motivations, definitions, methodologies, and future work. *IEEE Transactions on Power Systems*, 34(4):3188–3198, 2019.

[22] A. Abur and M. K. Celik. A fast algorithm for the weighted least absolute value state estimation. *IEEE Power Engineering Review*, 11:41–, 1991.

[23] Felix F. W. Power system state estimation: a survey. *International Journal of Electrical Power Energy Systems*, 12(2):80–87, 1990.

[24] G.S. Christensen, S.A. Soliman, and M.Y. Mohamed. Power systems state estimation based on least absolute value (lav). In C.T. LEONDES, editor, *Analysis and Control System Techniques for Electric Power Systems, Part 4 of 4*, volume 44 of *Control and Dynamic Systems*, pages 345–487. Academic Press, 1991.

[25] A. Abur and M.K. Celik. Least absolute value state estimation with equality and inequality constraints. *IEEE Transactions on Power Systems*, 8(2):680–686, 1993.

[26] H. Singh and F.L. Alvarado. Weighted least absolute value state estimation using interior point methods. *IEEE Transactions on Power Systems*, 9(3):1478–1484, 1994.

[27] R.C. Pires, A. Simoes Costa, and L. Mili. Iteratively reweighted least-squares state estimation through givens rotations. *IEEE Transactions on Power Systems*, 14(4):1499–1507, 1999.

[28] R. Singh, B. Chandra Pal, and R. A. Jabr. Choice of estimator for distribution system state estimation. *Iet Generation Transmission & Distribution*, 3:666–678, 2009.

[29] A. Monticelli, C.A.F. Murari, and Felix F. Wu. A hybrid state estimator: Solving normal equations by orthogonal transformations. *IEEE Transactions on Power Apparatus and Systems*, PAS-104(12):3460–3468, 1985.

[30] G.N. Korres. A robust method for equality constrained state estimation. *IEEE Transactions on Power Systems*, 17(2):305–314, 2002.

[31] A. S. Costa, S. Seleme Jr, and R. Salgado. Equality constraints in power system state estimation via orthogonal row-processing techniques. *IFAC Proceedings Volumes*, 18(7):43–49, 1985.

[32] A. Gjelsvik, S. Aam, and L. Holten. Hachtel's augmented matrix method - a rapid method improving numerical stability in power system static state estimation. *IEEE Transactions on Power Apparatus and Systems*, PAS-104(11):2987–2993, 1985.

[33] F.C. Schweppe and E.J. Handschin. Static state estimation in electric power systems. *Proceedings of the IEEE*, 62(7):972–982, 1974.

[34] E. A. Blood, B. H. Krogh, and M. D. Ilic. Electric power system static state estimation through kalman filtering and load forecasting. In *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–6, 2008.

[35] J. Zhao, M. Netto, and L. Mili. A robust iterated extended kalman filter for power system dynamic state estimation. *IEEE Transactions on Power Systems*, 32(4):3205–3216, 2017.

[36] E. Ghahremani and I. Kamwa. Dynamic state estimation in power system by applying the extended kalman filter with unknown inputs to phasor measurements. *IEEE Transactions on Power Systems*, 26(4):2556–2566, 2011.

[37] M. Netto, J. Zhao, and L. Mili. A robust extended kalman filter for power system dynamic state estimation using pmu measurements. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, 2016.

[38] G. Valverde. Unscented kalman filter for power system dynamic state estimation. *IET Generation, Transmission Distribution*, 5:29–37(8), January 2011.

[39] E. Ghahremani and I. Kamwa. Online state estimation of a synchronous generator using unscented kalman filter from phasor measurements units. *IEEE Transactions on Energy Conversion*, 26(4):1099–1108, 2011.

[40] C. Huanyuan, L. Xindong, S. Caiqi, and Y. Cheng. Power system dynamic state estimation based on a new particle filter. *Procedia Environmental Sciences*, 11:655–661, 2011. 2011 2nd International Conference on Challenges in Environmental Science and Computer Engineering (CESCE 2011).

[41] S. Weng, D. Yue, C. Dou, J. Shi, and C. Huang. Distributed event-triggered cooperative control for frequency and voltage stability and power sharing in isolated inverter-based microgrid. *IEEE Transactions on Cybernetics*, 49(4):1427–1439, 2019.

[42] H. Zhenyu, K. Schneider, and J. Nieplocha. Feasibility studies of applying kalman filter techniques to power system dynamic state estimation. In *2007 International Power Engineering Conference (IPEC 2007)*, pages 376–382, 2007.

[43] H. Liu, F. Hu, J. Su, X. Wei, and R. Qin. Comparisons on kalman-filter-based dynamic state estimation algorithms of power systems. *IEEE Access*, 8:51035–51043, 2020.

[44] A. Khandelwal and A. Tondan. Power system state estimation comparison of kalman filters with a new approach. In *2016 7th India International Conference on Power Electronics (IICPE)*, pages 1–6, 2016.

[45] N. R. Shivakumar and A. Jain. A review of power system dynamic state estimation techniques. In *2008 Joint International Conference on Power System Technology and IEEE Power India Conference*, pages 1–6, 2008.

[46] N. Zhou, D. Meng, Z. Huang, and G. Welch. Dynamic state estimation of a synchronous machine using pmu data: A comparative study. *IEEE Transactions on Smart Grid*, 6(1):450–460, 2015.

[47] M. A. Hasnat and M. Rahnamay-Naeini. A data-driven dynamic state estimation for smart grids under dos attack using state correlations. In *2019 North American Power Symposium (NAPS)*, pages 1–6, 2019.

[48] S. Soltan, P. Mittal, and H. V. Poor. Bayesian regression for robust power grid state estimation following a cyber-physical attack. In *2018 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5, 2018.

[49] J. Zhao, G. Zhang, Z. Y. Dong, and M. La Scala. Robust forecasting aided power system state estimation considering state correlations. *IEEE Transactions on Smart Grid*, 9(4):2658–2666, 2018.

[50] H. Bilil and H. Gharavi. Mmse-based analytical estimator for uncertain power system with limited number of measurements. *IEEE Transactions on Power Systems*, 33(5):5236–5247, 2018.

[51] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis. Physics-informed neural networks for power systems. In *2020 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5, 2020.

[52] W. Zhou, Y. Wu, X. Huang, R. Lu, and H. Zhang. A group sparse bayesian learning algorithm for harmonic state estimation in power systems. *Applied Energy*, 306:118063, 2022.

[53] J. Zhang, G. Welch, G. Bishop, and Z. Huang. A two-stage kalman filter approach for robust and real-time power system state estimation. *IEEE Transactions on Sustainable Energy*, 5(2):629–636, 2014.

[54] D. Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[55] Z. Yang, R. Gao, and W. He. A review of the research on kalman filtering in power system dynamic state estimation. In *2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, volume 4, pages 856–861, 2021.

[56] X. Jin, R. Jonhson R. Jeremiah, T. Su, Y. Bai, and J. Kong. The new trend of state estimation: From model-driven to hybrid-driven methods. *Sensors*, 21(6), 2021.

[57] S. V. S. Chauhan and G. X. Gao. Spoofing resilient state estimation for the power grid using an extended kalman filter. *IEEE Transactions on Smart Grid*, 12(4):3404–3414, 2021.

[58] J. Zhao and L. Mili. Robust unscented kalman filter for power system dynamic state estimation with unknown noise statistics. *IEEE Transactions on Smart Grid*, 10(2):1215–1224, 2019.

[59] S. Li, Z. Li, J. Li, Q. Wang, Z. Song, Z. Chen, Y. Sheng, X. Liu, and Y. Liu. Event-based cubature kalman filter for smart grid subject to communication constraint. *IFAC-PapersOnLine*, 50(1):49–54, 2017. 20th IFAC World Congress.

[60] X. Liu, L. Li, Z. Li, X. Chen, T. Fernando, H. H. Iu, and G. He. Event-trigger particle filter for smart grids with limited communication bandwidth infrastructure. *IEEE Transactions on Smart Grid*, 9(6):6918–6928, 2018.

[61] Z. Guo, D. Shi, D. E. Quevedo, and L. Shi. Secure state estimation against integrity attacks: A gaussian mixture model approach. *IEEE Transactions on Signal Processing*, 67(1):194–207, 2019.

[62] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar. Deep learning in power systems research: A review. *CSEE Journal of Power and Energy Systems*, 7(2):209–220, 2021.

[63] L. Zhang, G. Wang, and G. B. Giannakis. Real-time power system state estimation and forecasting via deep unrolled neural networks. *IEEE Transactions on Signal Processing*, 67(15):4069–4077, 2019.

[64] Z. Cao, Y. Wang, C. Chu, and R. Gadh. Scalable distribution systems state estimation using long short-term memory networks as surrogates. *IEEE Access*, 8:23359–23368, 2020.

[65] N. Bhusal, R. M. Shukla, M. Gautam, M. Benidris, and S. Sengupta. Deep ensemble learning-based approach to real-time power system state estimation. *International Journal of Electrical Power Energy Systems*, 129:106806, 2021.

[66] X. Niu, J. Li, J. Sun, and K. Tomsovic. Dynamic detection of false data injection attack in smart grid using deep learning. In *Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–6, 2019.

[67] J. Zhou, G. Cui, Z. Zhang, C. Yang, Z. Liu, and M. Sun. Graph neural networks: A review of methods and applications. *CoRR*, abs/1812.08434, 2018.

[68] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2009.

[69] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang. A review of graph neural networks and their applications in power systems, 2021.

[70] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He. Fault location in power distribution systems via deep graph convolutional networks. *IEEE Journal on Selected Areas in Communications*, 38(1):119–131, 2020.

[71] D. Wang, K. Zheng, Q. Chen, G. Luo, and X. Zhang. Probabilistic power flow solution with graph convolutional network. In *Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 650–654, 2020.

[72] Z. Sun, Y. Spyridis, T. Lagkas, A. Sesis, G. Efstathopoulos, and P. Sarigiannidis. End-to-end deep graph convolutional neural network approach for intentional islanding in power systems considering load-generation balance. *Sensors*, 21(5), 2021.

[73] A. S. Zamzam and N. D. Sidiropoulos. Physics-aware neural networks for distribution system state estimation, 2019.

[74] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin. Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks. *CoRR*, abs/2104.11846, 2021.

[75] H. Zhu and P. Koniusz. Simple spectral graph convolution. In *International Conference on Learning Representations*, 2021.

[76] F. C. Schweppe and D. B. Rom. Power system static-state estimation, part ii: Approximate model. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(1):125–130, 1970.

[77] F. C. Schweppe. Power system static-state estimation, part iii: Implementation. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(1):130–135, 1970.

[78] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem. A taxonomy of multi-area state estimation methods. *Electric Power Systems Research*, 81(4):1060–1069, 2011.

[79] T. van Cutsem, J.L. Horward, and M. Ribbens-Pavella. Hierarchical state estimator: A new way for large power systems estimation. *Mathematics and Computers in Simulation*, 22(2):133–140, 1980.

[80] T. V. Cutsem and M. Ribbens-Pavella. Critical survey of hierarchical methods for state estimation of electric power systems. *IEEE Transactions on Power Apparatus and Systems*, PAS-102(10):3415–3424, 1983.

[81] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles. A multilevel state estimation paradigm for smart grids. *Proceedings of the IEEE*, 99(6):952–976, 2011.

[82] S. Iwamoto, M. Kusano, and V.H. Quintana. Hierarchical state estimation using a fast rectangular-coordinate method (power system analysis computing). *IEEE Transactions on Power Systems*, 4(3):870–880, 1989.

[83] G. N. Korres. A distributed multiarea state estimation. *IEEE Transactions on Power Systems*, 26(1):73–84, 2011.

[84] Liang Z. and A. Abur. Multi area state estimation using synchronized phasor measurements. *IEEE Transactions on Power Systems*, 20(2):611–617, 2005.

[85] G. N. Korres. A distributed multiarea state estimation. *IEEE Transactions on Power Systems*, 26(1):73–84, 2011.

[86] P. Rousseaux, T. Van Cutsem, D. Mallieu, and M. Ribbens-Pavella. Dynamic state prediction and hierarchical filtering for power system estimation. *IFAC Proceedings Volumes*, 20(9):721–726, 1987. 4th IFAC/IFORS Symposium on Large Scale Systems: Theory and Applications 1986, Zurich, Switzerland, 26-29 August 1986.

[87] S. A. Zonouz and W. H. Sanders. A kalman-based coordination for hierarchical state estimation: Agorithm and analysis. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 187–187, 2008.

[88] S.-Y. Bin and C.-H. Lin. An implementable distributed state estimator and distributed bad data processing schemes for electric power systems. *IEEE Transactions on Power Systems*, 9(3):1277–1284, 1994.

[89] A. J. Conejo, S. de la Torre, and M. Canas. An optimization approach to multiarea state estimation. *IEEE Transactions on Power Systems*, 22(1):213–221, 2007.

[90] X. Zhou, Z. Liu, Y. Guo, C. Zhao, J. Huang, and L. Chen. Gradient-based multi-area distribution system state estimation. *IEEE Transactions on Smart Grid*, 11(6):5325–5338, 2020.

[91] D.M. Falcao, F.F. Wu, and L. Murphy. Parallel and distributed state estimation. *IEEE Transactions on Power Systems*, 10(2):724–730, 1995.

[92] R. Ebrahimian and R. Baldick. State estimation distributed processing [for power systems]. *IEEE Transactions on Power Systems*, 15(4):1240–1246, 2000.

[93] L. Xie, D. Choi, and S. Kar. Cooperative distributed state estimation: Local observability relaxed. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–11, 2011.

[94] P. Du, Z. Huang, Y. Sun, R. Diao, K. Kalsi, K. K. Anderson, Y. Li, and B. Lee. Distributed dynamic state estimation with extended kalman filter. In *2011 North American Power Symposium*, pages 1–6, 2011.

[95] F. Wen and Z. Wang. Distributed kalman filtering for robust state estimation over wireless sensor networks under malicious cyber attacks. *Digital Signal Processing*, 78:92–97, 2018.

[96] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. 2011.

[97] S. Parsegov, S. Kubentayeva, E. Gryazina, A. Gasnikov, and F. Ibáñez. Admm-based distributed state estimation for power systems: Evaluation of performance∗∗the research of a. gasnikov is partially supported by rfbr project 19-31-51001 and by the ministry of science and higher education of the russian federation (goszadaniye) no. 075-00337-20-03, project no. 0714-2020-0005. *IFAC-PapersOnLine*, 53(5):182–188, 2020. 3rd IFAC Workshop on Cyber-Physical Human Systems CPHS 2020.

[98] W. Zheng, W. Wu, A. Gomez-Exposito, B. Zhang, and Y. Guo. Distributed robust bilinear state estimation for power systems with nonlinear measurements. *IEEE Transactions on Power Systems*, 32(1):499–509, 2017.

[99] T. Chen, Y. Cao, X. Chen, L. Sun, J. Zhang, and G. A. J. Amaratunga. A distributed maximum-likelihood-based state estimation approach for power systems. *IEEE Transactions on Instrumentation and Measurement*, 70:1–10, 2021.

[100] Y. Bai, W. Li, and B. Zhang. Distributed incremental quasi-newton algorithm for power system state estimation. In *2021 8th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, pages 369–374, 2021.

[101] N. Aljohani, T. Zou, A. S. Bretas, and N. G. Bretas. Multi-area state estimation: A distributed quasi-static innovation-based model with an alternative direction method of multipliers. *Applied Sciences*, 11(10), 2021.

[102] U. Nakarmi, M. Rahnamay Naeini, M. J. Hossain, and M. A. Hasnat. Interaction graphs for cascading failure analysis in power grids: A survey. *Energies*, 13(9):1–25, 2020.

[103] J. Yang, W. Zhang, and F. Guo. Dynamic state estimation for power networks by distributed unscented information filter. *IEEE Transactions on Smart Grid*, 11(3):2162–2171, 2020.

[104] V. Vahidpour, A. Rastegarnia, A. Khalili, and S. Sanei. Partial diffusion kalman filtering for distributed state estimation in multiagent networks. *IEEE Transactions on Neural Networks and Learning Systems*, 30(12):3839–3846, 2019.

[105] R. Yu, Z. Yuan, M. Zhu, and Z. Zhou. Data-driven distributed state estimation and behavior modeling in sensor networks. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8192–8199, 2020.

[106] M. N. Kurt, Y. Yılmaz, and X. Wang. Secure distributed dynamic state estimation in wide-area smart grids. *IEEE Transactions on Information Forensics and Security*, 15:800–815, 2020.

[107] C. Muscas, M. Pau, P. A. Pegoraro, S. Sulis, F. Ponci, and A. Monti. Multiarea distribution system state estimation. *IEEE Transactions on Instrumentation and Measurement*, 64(5):1140–1148, 2015.

[108] M. Cosovic, A. Tsitsimelis, D. Vukobratovic, J. Matamoros, and C. Anton-Haro. 5g mobile cellular networks: Enabling distributed state estimation for smart grids. *IEEE Communications Magazine*, 55(10):62–69, 2017.

[109] M. Tran, A. S. Zamzam, P. H. Nguyen, and G. Pemen. Multi-area distribution system state estimation using decentralized physics-aware neural networks. *Energies*, 14(11), 2021.

[110] B. Zargar, A. Angioni, F. Ponci, and A. Monti. Multiarea parallel data-driven three-phase distribution system state estimation using synchrophasor measurements. *IEEE Transactions on Instrumentation and Measurement*, 69(9):6186–6202, 2020.

[111] N. Nusrat, M. Irving, and G. Taylor. Development of distributed state estimation methods to enable smart distribution management systems. In *2011 IEEE International Symposium on Industrial Electronics*, pages 1691–1696, 2011.

[112] S. Park, F. Gama, J. Lavaei, and S. Sojoudi. Distributed power system state estimation using graph convolutional neural networks. 2021.

[113] Y. Sun, Y.and Zhao. Distributed cubature kalman filter with performance comparison for large-scale power systems. *International Journal of Control, Automation and Systems*, 19:1319–1327, 2021.

[114] F. M. Zegers, R. Sun, G. Chowdhary, and W. E. Dixon. Distributed state estimation with deep neural networks for uncertain nonlinear systems under event-triggered communication, 2022.

[115] O. Vuković and G. Dán. Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks. *IEEE Journal on Selected Areas in Communications*, 32(7):1500–1508, 2014.

[116] X. Zhang and Y. Shen. Distributed kalman filtering based on the non-repeated diffusion strategy. *Sensors*, 20(23), 2020.

[117] Y. Jiang and Q. Hui. Kalman filter with diffusion strategies for detecting power grid false data injection attacks. In *2017 IEEE International Conference on Electro Information Technology (EIT)*, pages 254–259, 2017.

[118] S. Soltan and G. Zussman. Power grid state estimation after a cyber-physical attack under the ac power flow model. In *2017 IEEE Power Energy Society General Meeting*, pages 1–5, 2017.

[119] S. Soltan, M. Yannakakis, and G. Zussman. Power grid state estimation following a joint cyber and physical attack. *IEEE Transactions on Control of Network Systems*, 5(1):499–512, 2018.

[120] S. Soltan and G. Zussman. Quantifying the effect of k-line failures in power grids. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, 2016.

[121] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah. Bilevel model for analyzing coordinated cyber-physical attacks on power systems. *IEEE Transactions on Smart Grid*, 7(5):2260–2272, 2016.

[122] H. He and J. Yan. Cyber-physical attacks and defences in the smart grid: a survey. *IET Cyber-Physical Systems: Theory & Applications*, 1(1):13–27, 2016.

[123] J. E. Tate and T. J. Overbye. Line outage detection using phasor angle measurements. *IEEE Transactions on Power Systems*, 23(4):1644–1652, 2008.

[124] H. Zhao. A new state estimation model of utilizing pmu measurements. In *2006 International Conference on Power System Technology*, pages 1–5, 2006.

[125] J. James and Bindu S. Hybrid state estimation including pmu measurements. In *2015 International Conference on Control Communication Computing India (ICCC)*, pages 309–313, 2015.

[126] I. Kolosok, E. Korkina, and E. Buchinsky. The test equation method for linear state estimation based on pmu data. In *2014 Power Systems Computation Conference*, pages 1–7, 2014.

[127] S. Hou, Z. Xu, H. Lv, Z. Jiang, and W. Lingyi. Research into harmonic state estimation in power system based on pmu and svd. In *2006 International Conference on Power System Technology*, pages 1–6, 2006.

[128] M. Pignati, L. Zanni, P. Romano, R. Cherkaoui, and M. Paolone. Fault detection and faulted line identification in active distribution networks using synchrophasors-based real-time state estimation. *IEEE Transactions on Power Delivery*, 32(1):381–392, 2017.

[129] M. M. Rana, R. Bo, and A. Abdelhadi. Distributed grid state estimation under cyber attacks using optimal filter and bayesian approach. *IEEE Systems Journal*, 15(2):1970–1978, 2021.

[130] C. Chen, Y. Wang, M. Cui, J. Zhao, W. Bi, Y. Chen, and X. Zhang. Data-driven detection of stealthy false data injection attack against power system state estimation. *IEEE Transactions on Industrial Informatics*, 18(12):8467–8476, 2022.

[131] A. Monticelli. Electric power system state estimation. *Proceedings of the IEEE*, 88(2):262–282, 2000.

[132] O. Kosut, L. Jia, R. J. Thomas, and L. Tong. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 220–225, 2010.

[133] A. Gul and S. D. Wolthusen. *A Review on Attacks and Their Countermeasures in Power System State Estimation*, pages 9–28. Springer International Publishing, Cham, 2018.

[134] J. Drugowitsch. Variational bayesian inference for linear and logistic regression, 2013.

[135] D. A. Vaccari and H. K. Wang. Multivariate polynomial regression for identification of chaotic time series. *Mathematical and Computer Modelling of Dynamical Systems*, 13(4):395–412, 2007.

[136] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks, 2017.

[137] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling, 2014.

[138] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li. T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, 21(9):3848–3858, 2020.

[139] C. Zhao, J. Zhao, C. Wu, X. Wang, F. Xue, and S. Lu. Power grid partitioning based on functional community structure. *IEEE Access*, 7:152624–152634, 2019.

[140] Miao Zhang, Zhixin Miao, and Lingling Fan. Power grid partitioning: Static and dynamic approaches. In *2018 North American Power Symposium (NAPS)*, pages 1–6, 2018.

[141] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Springer New York, NY, Springer Science+Business Media, 1981.

[142] J. Mezquita, D. Asber, S. Lefebvre, M. Saad, and P. J. Lagacé. Power network partitioning with a fuzzy c-means. 2011.

[143] 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*, 26(1):12–19, 2011.

[144] W. Zheng and W. Wu. An adaptive distributed quasi-newton method for power system state estimation. *IEEE Transactions on Smart Grid*, 10(5):5114–5124, 2019.

[145] N. A. Rathod, H. H. Patel, and S. R. Joshi. Distributed hybrid state estimation with ill conditioning of sub area. In *2019 8th International Conference on Power Systems (ICPS)*, pages 1–6, 2019.

[146] A. Garg, T. Ma, and H. Nguyen. On communication cost of distributed statistical estimation and dimensionality. *Advances in Neural Information Processing Systems*, 27, 2014.

[147] A. Delgado-Bonal and A. Marshak. Approximate entropy and sample entropy: A comprehensive tutorial. *MDPI: Entropy (Basel)*, 21(6):541, 2019.

[148] S. Pincus. Approximate entropy (apen) as a complexity measure. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):110–117, 1995.

[149] Y. Weng, R. Negi, and M. D. Ilić. Historical data-driven state estimation for electric power systems. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 97–102, Oct 2013.

[150] V. Kirinčić, E. Čeperić, S. Vlahinić, and J. Lerga. Support vector machine state estimation. *Applied Artificial Intelligence*, 33(6):517–530, 2019.

[151] N. M. Manousakis and G. N. Korres. Optimal allocation of phasor measurement units considering various contingencies and measurement redundancy. *IEEE Transactions on Instrumentation and Measurement*, 69(6):3403–3411, 2020.

[152] Z. Huang, N. Zhou, R. Diao, S. Wang, S. Elbert, D. Meng, and S. Lu. Capturing real-time power system dynamics: Opportunities and challenges. In *2015 IEEE Power Energy Society General Meeting*, pages 1–5, 2015.

[153] X. Sun, J. Duan, X. Li, and X. Wang. State estimation under non-gaussian levy noise: A modified kalman filtering method, 2013.

[154] U. Nakarmi, M. Rahnamay-Naeini, M. J. Hossain, and M. A. Hasnat. Interaction graphs for cascading failure analysis in power grids: A survey. *Energies*, 13(9), 2020.

Appendix A: Copyright Permissions

The permission to reuse the materials from the published papers are below:

CCC | **RightsLink**

IEEE
Requesting permission to reuse content from an IEEE publication

Line Failure Detection from PMU Data after a Joint Cyber-Physical Attack

Conference Proceedings: 2019 IEEE Power & Energy Society General Meeting (PESGM)
Author: MD Jakir Hossain
Publisher: IEEE
Date: August 2019

Copyright © 2019, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK   CLOSE WINDOW

The permission to reuse the materials from the published papers are below:

The permission to reuse the materials from the published papers are below:

The permission to reuse the materials from the published papers are below::