Summer 8-4-2020

Model-Based and Data-driven Situational Awareness for Distribution System Monitoring and Control

Ying Zhang
Southern Methodist University, imyingzhang@ieee.org

MODEL-BASED AND DATA-DRIVEN
SITUATIONAL AWARENESS FOR
DISTRIBUTION SYSTEM MONITORING AND CONTROL

Approved by:

Dr. Jianhui Wang
Associate Professor

Dr. Behrouz Peikari
Professor

Dr. Khaled Abdelghany
Professor

Dr. Mohammad Khodayar
Associate Professor

Dr. Harsha Gangammanavar
Assistant Professor

MODEL-BASED AND DATA-DRIVEN SITUATIONAL AWARENESS FOR DISTRIBUTION SYSTEM MONITORING AND CONTROL

A Dissertation Presented to the Graduate Faculty of the

Bobby B. Lyle School of Engineering

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Electrical Engineering

by

Ying Zhang

Bachelor of Science, Electrical Engineering, Shandong University, 2014
Master of Science, Electrical Engineering, Shandong University, 2017

04 August, 2020

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor and mentor, Prof. Jianhui Wang, for offering me an opportunity to pursue the Ph.D. degree and continuous supports of my research in the past three years. I am grateful for his patience, wisdom, enthusiasm, and suggestions, and also for motivating me farther than I thought I could go.

My graduation committee members, Prof. Behrouz Peikari, Prof. Mohammad Khodayar, Prof. Harsha Gangammanavar, and Prof. Khaled Abdelghany, have provided valuable advice and guidance. Thanks for their unwavering supports.

Besides, I would like to acknowledge Dr. Yingchen Zhang, Dr. Bo Chen, Dr. Zhenghsuo Li, Dr. Mingjian Cui, and Dr. Jianzhe Liu. The dissertation obtains their constructive technical inputs. Meanwhile, sincere thanks to my colleagues and school fellows, Xinan Wang, Shengfei Yin, Mahdi Khodayar, You Lin, Yanling Lin, Xinyun Lu, Junyang Mi, Tao Wu, Bin Huang, Saeed D Manshadi, and Juwariyah Khalid, who have supported and accompanied me over years in SMU. In the end, I thank to all my family and friends for their consistent helps and supports on my educational path, and they keep me going.

This work is supported in part by the U.S. National Science Foundation under Grant ECCS-1800716, Office of Electricity Delivery and Energy Reliability in the U.S. Department of Energy (DOE), and Argonne National Laboratory.

Zhang, Ying Bachelor of Science, Electrical Engineering, Shandong University, 2014
Master of Science, Electrical Engineering, Shandong University, 2017

Model-based and Data-driven
Situational Awareness for
Distribution System Monitoring and Control

Advisor: Dr. Jianhui Wang
Doctor of Philosophy degree conferred 04 August, 2020
Dissertation completed July, 2020

Electric power systems are undergoing a dramatic change. The penetration of distributed energy resources (DERs) such as wind turbine generators and photovoltaic panels is turning a traditional power system into the active distribution network. Power system situational awareness, which provides critical information for system monitoring and control, is being challenged by multiple sources of uncertainties such as random meter errors, stochastic power output of DERs, and imprecise network parameters. On the other hand, cyber-physical power system operation is vulnerable to cyberattacks against effective state estimation, such as false data injection attacks (FDIAs). To construct next-generation smart grids, this dissertation develops a comprehensive situational awareness framework for distribution system monitoring and control via optimization, machine learning, and artificial intelligence. Specifically, this dissertation explores advanced model-based and data-driven methodologies in this framework, including state estimation, cyberattack detection, fault location, and voltage control.

Firstly, we propose a highly efficient distribution system state estimation (DSSE) method using supervisory control and data acquisition (SCADA) systems. The proposed method uses the Taylor series of voltages for constructing a linear DSSE model in the interval form and then solves this model by interval arithmetic, which avoids high calculation costs that result from a nonlinear iterative process in traditional weighted least square (WLS)-based methods. Subsequently, a general interval state estimation (ISE) model is developed to

formulate multiple uncertainties. A modified Krawczyk operator in interval arithmetic is proposed to solve the general ISE model efficiently and provides upper and lower bounds of state variables. Case studies the developed DSSE method overcomes the traditional challenges from poor observability due to lack of meters, stochastic outputs of renewables, and imprecise network parameters, and works conveniently with limited phasor measurement units (PMUs) in three-phase unbalanced distribution systems. Meanwhile, we seek to develop fault location via DSSE integrating PMU data in a distributed manner. Based on the graph model of the feeder, we further perform the decentralized DSSE algorithm in a hierarchical structure and identify the location of the fault source. The proposed approach captures the impact of distributed generations on distribution system operation and works with high-level noises in measurements.

We further apply DSSE to data-driven cyberattack detection and voltage control via machine learning and artificial intelligence (AI). In the propose semi-supervised learning-based FDIA detection algorithm, we use autoencoders for efficient dimension reduction and feature extraction of measurement datasets and integrate them into advanced generative adversarial networks (GANs) to detect anomalies by capturing the unconformity between abnormal and secure measurements. Finally, a model-free volt-VAR optimization (VVO) algorithm via multi-agent deep reinforcement learning (MADRL) is developed. Numerical tests show the proposed situational awareness technologies, combining sophistication and flexibility, are highly applicable to practical distribution systems and have good potential to improve operation and control efficiency in real-world power grids.

TABLE OF CONTENTS

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER

1. **Introduction** 1

    1.1. Power System Situational Awareness 1

    1.2. Research Objectives and Achievements 3

        1.2.1. Innovation: Preparing for DER Uncertainty 4

        1.2.2. Integration: Transforming the Grid with Big Data, Machine Learning, and Artificial Intelligence 5

    1.3. Chapter Organization 7

2. **Towards Highly Efficient State Estimation with SCADA Measurements in Distribution Systems** 8

    2.1. Motivation 8

    2.2. Conventional DSSE 9

        2.2.1. State Estimation Theory 9

        2.2.2. Formulation of DSSE 10

    2.3. Proposed Methodology 12

        2.3.1. DSSE Modeling in Interval Arithmatic 12

        2.3.2. Solution in Interval Arithmetic 15

    2.4. Numerical Test 16

        2.4.1. Estimation Performance 17

        2.4.2. Robustness Analysis 19

    2.5. Conclusion and Discussion 20

3. **Interval State Estimation with Uncertainty of Distributed Generation and Line Parameters Integrating PMU Data** 22

3.1. Literature Review 22

3.2. Hybrid DSSE Algorithm and Interval Arithmetic 24

3.2.1. DSSE with Hybrid Measurements 25

3.2.2. Interval Arithmetic and Interval Prediction of DG 29

3.3. General ISE Framework and Proposed Algorithm 29

3.3.1. ISE Model with Multiple Uncertainties 30

3.3.2. General ISE Framework 31

3.3.3. MKO Algorithm for Solving ISE Models 33

3.4. Simulation Result 37

3.4.1. Result Analysis 39

3.4.2. Effect of Parameter Uncertainty 41

3.4.3. Robustness Tests 43

3.4.4. Computational Efficiency 44

3.4.5. Results of Model I and Model III 46

3.5. Conclusion 47

4. **Graph-based Faulted Line Identification Using PMU Data in Distribution Networks** 48

4.1. Introduction 48

4.2. Theoretical Basis 50

4.3. Graph-Based Faulted Line Identification Method 54

4.3.1. Step One: Identifying the Faulted Subgraph 54

4.3.2. Step Two: Locating the Faulted Line 58

4.3.3. PMU Placement 61

4.4. Case Study 63

4.4.1. Faulted-subgraph Identification . . . 64
4.4.2. Faulted-line Location . . . 65
4.4.3. Robustness and Sensitivity Analysis . . . 68
4.4.4. Impact of Line Parameters . . . 70
4.4.5. Computational Efficiency . . . 71
4.5. Conclusion . . . 71
5. **Semi-supervised Deep Learning for False Data Injection Attack Detection** . . . 73
5.1. Introduction . . . 73
5.2. State Estimation and FDIA . . . 75
5.2.1. BDD based on Estimation Residuals . . . 75
5.2.2. On the Existence of Linear Approximation of DSSE . . . 77
5.2.3. Unobservable FDIA . . . 78
5.3. Construction of FDIA . . . 79
5.4. Proposed Detection Mechanism . . . 81
5.4.1. Autoencoders and GAN . . . 83
5.4.2. Adversarial Autoencoder . . . 84
5.5. Case Study . . . 88
5.5.1. Unobservable FDIA . . . 90
5.5.2. Detection Performance . . . 90
5.5.3. Other Semi-supervised Techniques . . . 93
5.5.4. Impacts of Measurement Noises . . . 93
5.5.5. Sensitivity Analysis . . . 94
5.6. Conclusion . . . 97
6. **Deep Reinforcement Learning Based Volt-VAR Control in Smart Grids** 98
6.1. Introduction . . . 98

6.2. System Model of Unbalanced Distribution Networks 101
6.3. Proposed VVO Algorithm 104
6.3.1. Voltage Regulation Devices 105
6.3.2. MADRL Design for VVO 108
6.3.3. Multi-agent DQN Algorithm 110
6.4. Case Study 114
6.4.1. Learning Performance 116
6.4.2. VVO Performance 118
6.4.3. Computation Time 121
6.4.4. Comparison with Existing Reward Design 122
6.5. Conclusion 124
7. **Conclusion and Future Work** 125
7.1. Summary 125
7.2. Technical Prospect and Outlook 126
8. **Publications** 129
APPENDIX
A. PMU Latency for Faulted Line Identification 131
BIBLIOGRAPHY 132

LIST OF FIGURES

Figure Page

1.1 Research scheme of this dissertation for distribution system monitoring and control. . . . 4

2.1 Schematic diagram of the proposed linearization. (a) The complex plane of V_k. (b) Approximation loss $F(V_k)$ . . . 13

2.2 IEEE 34-bus distribution system . . . 17

2.3 Comparison in the estimation errors of the proposed method and the MAEs that may happen in a random sampling of measurements. . . . 18

3.1 Three-phase line model in unbalanced distribution systems . . . 28

3.2 Uncertainties of DSSE in interval arithmetic. The model complexity increases from Model I to Model IV. . . . 33

3.3 Solution hulls of a 2-dimensional interval equation . . . 35

3.4 One-line diagram of unbalanced IEEE 13-bus system . . . 38

3.5 DG profiles during one day (a) Wind speed (b) Solar radiation . . . 39

3.6 Voltage magnitude results of the proposed algorithm . . . 40

3.7 Voltage magnitude results in the 123-bus system. (a) Case 1: Without line uncertainty (b) Case 2: With line uncertainty. Under multiple uncertainties, although a linear approximation technique is adopted, the interval results of our algorithm are similar to the ones of MC simulations without this approximation. . . . 42

4.1 A sample of radial distribution networks with three PMUs. The dotted lines with arrows at nodes denote the laterals (if present) . . . 55

4.2 Subgraph K (a) Embedded in the whole feeder (b) Decoupled with other subgraphs. The reference directions of branch currents measured by two PMUs are shown. . . . 57

4.3 Sample network of a 5-node subgraph, and a lateral is connected to node 3 shown as a dotted line. (a) A fault occurs at branch 3-5. (b) A fault occurs at branch 3-4. . . . 60

4.4 Single-line diagram for the three-phase 34-node test feeder 63

4.5 Identification results in different faulted subgraphs, where we set LG faults on phase A with 100 Ω impedance in G_K. 65

4.6 Identification results in different faulted subgraphs. LL faults on phases B and C with 50 Ω fault impedance. 66

4.7 Location results in G_1, when LL faults with 10Ω fault impedance occur at branch 3-4. The secondary y-axis shows the WMRs at the corrupted path in 100 Monte Carlo simulations. 66

4.8 Performance for high-impedance LG faults at branch 3-4. 69

5.1 An overview of the proposed FDIA detection mechanism 82

5.2 Structure of autoencoders with three fully connected hidden layers 82

5.3 The learning process of GAN 85

5.4 Semi-supervised AAE architecture: (a) in reconstruction phase, (b) in adversarial phase, and (c) in supervised phase. 86

5.5 The LNR results under unobservable FDIAs and no attacks 91

5.6 Estimation results of voltages under no attacks and an unobservable attack 91

5.7 Detection accuracy with maximum measurement errors ranging from 4% to 12%. 95

6.1 Flowchart of the forward-backward sweep method 104

6.2 The NN structure for agent j. 111

6.3 Multi-agent DQN training loop 112

6.4 Three-phase unbalanced 13-bus distribution system 115

6.5 Multi-agent DQN training process in the IEEE 123-bus system 118

6.6 Three-phase voltage magnitude comparison before and after VVO control in the 13-bus system. 120

6.7 Power loss comparison before and after control in the 13-bus system. 121

6.8 The comparison of the learning performance for VVO 123

A.1 The diagram of a waveform of fault current and the latency of the proposed method in identifying a fault. 131

LIST OF TABLES

Table		Page
2.1	Measurement arrangement in test systems	17
2.2	Comparison in accuracy and computation time	19
2.3	Estimation performance in robustness analysis	19
3.1	DG placements in 123-bus system	38
3.2	Measurement locations in test systems	40
3.3	Detailed comparison in voltage results	43
3.4	Estimation accuracy in robustness tests	44
3.5	Comparison in balanced and unbalanced 123-bus system	44
3.6	CPU time in different cases	45
3.7	Computation time in 123-bus system	45
3.8	Estimation accuracy in robustness tests	46
3.9	CPU time in 123-bus system	47
4.1	State variables and measurements in a 5-node subgraph	60
4.2	Measurement information of subgraphs	64
4.3	Performance with different fault types (50 Ω impedance)	67
4.4	Performance with different fault impedances (LL faults)	68
4.5	Performance with higher measurement errors (LG, 50 Ω impedance)	69
4.6	Impact of identification thresholds	70
4.7	Location accuracy with uncertainty in line parameters	70
4.8	CPU time in test system	71
5.1	Measurement locations in test system	89

5.2 Definitions of performance indices 92
5.3 Detection performance of proposed algorithm 93
5.4 Comparison with other semi-supervised methods 94
5.5 FDIA detection with fewer or more labeled data 95
5.6 Data structure in sensitivity analysis 96
5.7 Comparison of detection performance in two cases 96
5.8 Confusion matrix of FDIA detection with fewer labeled data 97
6.1 Literature review of VVO methods in distribution systems 99
6.2 Installation settings of control devices 116
6.3 Parameter settings of multi-agent DQN 117
6.4 Performance of voltage regulation 118
6.5 Voltage comparison in the IEEE 123-bus system 121
6.6 Performance of power loss reduction 122
6.7 Comparison of voltage regulation in the 13-bus system 123

Chapter 1

Introduction

1.1. Power System Situational Awareness

High penetration of distributed energy resources (DERs) is driving current distribution systems toward smart next-generation systems [1]. DERs are defined as small power sources on the distribution systems that produce electricity and are not otherwise included in the formal definition of the bulk electric system by the North American Electric Reliability Corporation (NERC) [2]. There are various types of DERs, such as photovoltaic (PV), electric vehicle (EV), and wind turbine generators (WTGs). DERs have advantages of low investment costs, low power losses, and flexible operation by unlocking new or existing energy resources to be developed near customers. Motivated by these benefits, the installation of DERs have increased significantly worldwide in recent years. As a good example of such an increment, DERs in the United States have grown almost three times faster than net total generation capacity (168 GW vs. 57 GW) on a 5-year basis (2015-2019) [3].

This increase brings huge challenges to practical power system and electricity market operations. The traditional "passive" distribution systems are transforming into an active network, which requires operators to revise their operation strategies to accommodate the transition from the traditional top-down power flow to a bottom-up paradigm [4] and update almost all of their control applications. Moreover, at high levels of DER penetration, electric power is found to flow back towards the upstream transformer, causing nodal voltages to rise. This phenomenon leads to an undesirable cycling of voltage regulators and voltage quality degradation [5]. Besides, the presence of DERs reduces the fault current for downstream faults and increases the fault current for upstream faults. Therefore, either of them may exceed the existing current setting, resulting in mal-trip or fail-to-trip of protection devices

at a certain penetration level [6].

On the other hand, highly relying on the communication infrastructure in the cyber-physical power systems, the risk and vulnerability of cyber-attacks against power grid operation increase. For example, in the 2015 Ukrainian power grid cyber-attack event, attackers intruded the supervisory control and data acquisition (SCADA) system of distribution networks, which leads to a power outage involving 225,000 customers [7]. Liu et al. presented an false data injection attack (FDIA) algorithm against dc state estimation to avoid being detected by conventional bad data detection (BDD) methods [8]. Since then, researchers devote significant efforts to investigating the FDIA detection methods. These attacks that can successfully circumvent the BDD methods are defined as unobservable cyber-attacks. Distribution systems are also vulnerable to cyber-attacks due to their direct connections to customer loads and emerging distributed generators (DGs) [9].

As a typical component for distribution automation, power system situational awareness, represented by state estimation, enables continuous and reliable monitoring and control of distribution system operation with DER penetration. This technology is paramount in practice as it allows distribution system operators (DSOs) to perceive the operating status of the system and make decisions. Various control-room applications contain state estimation, topology identification, fault management, cyberattack mitigation, and voltage control, etc. However, the existing situational awareness technologies in transmission systems cannot be trivially applied to typical distribution systems. The reason originates from several considerable differences between electric transmission and distribution systems:

- **Radial or weakly-meshed networks**: The vast majority of distribution systems operate in radial configurations for practical engineering concerns, which distinguishes distribution networks from transmission systems that generally have multiple loops. The weakly-meshed structure guarantees the system reliability against contingencies or scheduled maintenance.

- **Poor Observability**: Distribution systems may be unobservable due to data missing and lack of real-time meters, especially phasor measurement units (PMUs), which

means that the number of metering instruments in a network is generally limited. In contrast, transmission systems have a high measurement redundancy due to prevalent PMU installation.

- **Low x/r ratio**: Low reactance to resistance ratio in distribution lines make conventional Newton-Raphson power flow and state estimation algorithms in transmission systems fail at a distribution level [10].

- **Unbalanced operation**: Distribution systems are highly unbalanced in practice, resulting in higher complexity in the formulation of power flow calculation and distribution system state estimation (DSSE). This unbalanced property is reflected not only in the different network parameters (or loads) on each phase but also in the existences of single-phase and two-phase branches.

- **"Invisible" network configuration**: The complete data of network topology are not commonly stored [11]. Moreover, the topology information in distribution networks is often not available at the low-voltage level due to uninformed changes that frequently happen [12].

These characteristics of distribution systems increase the challenges and difficulties of system operation and control. As a result, accurate and effective modeling and methodology of power system situational awareness call for holistic improvement.

1.2. Research Objectives and Achievements

Active distribution systems emerge with extensive DERs aggregating on the demand side. Quantitatively assessing these impacts of DER uncertainty on these active distribution systems via DSSE becomes more necessary. On the other hand, data-driven approaches, which exploit valuable information by abundant real-time and historical data from periodic patterns of power systems, hold the promise to enhance operational efficiency of distribution system operation. In pursuit of these goals, we have pioneered a general theoretical framework that advances statistical theory, optimization, and machine learning for distribution system monitoring and control.

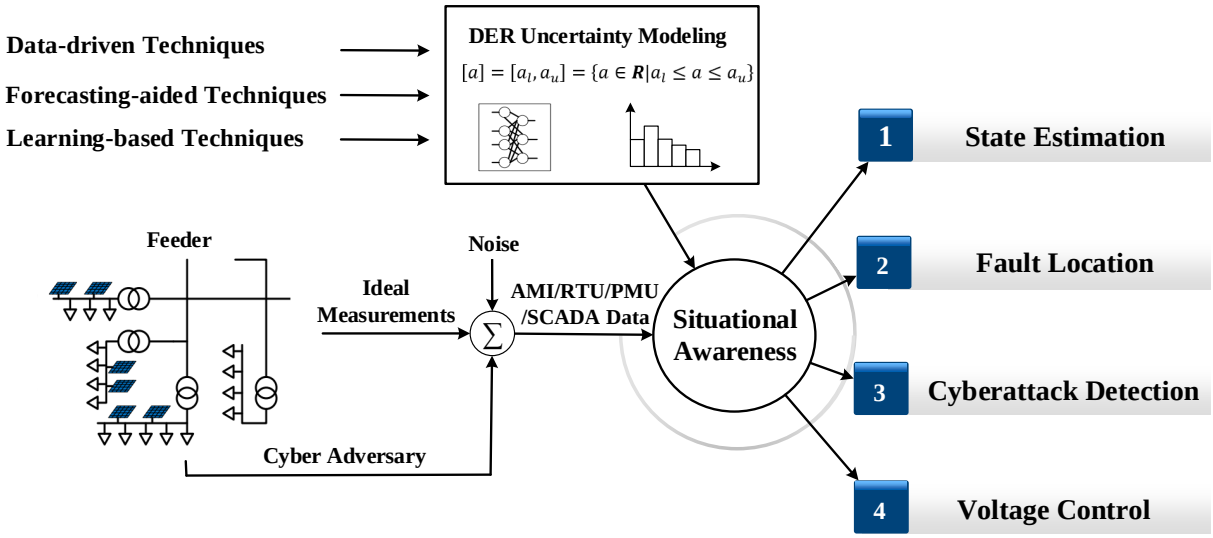

Figure 1.1. Research scheme of this dissertation for distribution system monitoring and control.

Fig. 1.1 depicts the general schematic of a model-based and data-driven situational awareness framework. In the framework, the uncertain power outputs of DER are modeled by diverse techniques and then deemed as the inputs of DSSE along with measurements to depict the system states. Meanwhile, the random noises from meters and malicious cyber-attacks against these measurements degrade the estimation accuracy. Further DSSE-based applications appear across voltage control, topology identification, and cyberattack mitigation. Specifically, the main contributions and achievements of this dissertation are outlined below.

1.2.1. Innovation: Preparing for DER Uncertainty

The variability and intermittency of DER outputs pose significant uncertainty to system operation; besides, measurements with noises and imprecise line parameters because of aging also result in uncertain inputs to DSSE. In the meantime, the reliability events represented

by faults and cyberattacks bring immense uncertainty factors to normal system operation. All these issues result in stringent requirements for system modeling and operation practices.

DSSE & Renewable Integration. To capture the impacts of these uncertainties from DER outputs, imprecise line parameters, and measurements with noises, we develop a general framework of interval state estimation (ISE) leveraging the heterogeneous measurements from SCADA systems and the emerging PMUs. Our proposed algorithms obtain the upper and lower bounds of state variables for better monitoring power grids under the coordinated impacts of multiple uncertainties. In brief, we see through these uncertainties by ISE and realize 100x acceleration compared with the conventional weighted least square (WLS)-based methods based on Monte Carlo simulation in this field. Towards a 100% renewable power grid in the near future, where significant uncertainty produces a far-reaching influence on system operation, the proposed ISE algorithm will be a promising candidate for effective solutions to uncertainty analysis.

DSSE & Fault Location. Fast and accurate fault location helps the utilities to clear the faults and accelerate system restoration; however, this is a challenging task as the bidirectional power flow due to the increasing penetration of DERs leads to the mal-trip or fail-to-trip of conventional protection devices and imprecise fault location. To mitigate the impacts, a graph-based fault location algorithm in distribution networks is proposed by advanced DSSE techniques integrating PMU data. This faulted-line location method running in a decentralized manner has a lower computation cost and enables fast fault location within 15 milliseconds, which shows the application potential in larger-scale systems. The proposed approach captures the impact of DER penetration on distribution system operation, and its location performance is independent of fault types and fault impedances. Furthermore, the proposed algorithm is robust against high-level noises in measurements.

1.2.2. Integration: Transforming the Grid with Big Data, Machine Learning, and Artificial Intelligence

With the fast development of advanced metering infrastructure (AMI) that collects a

massive volume of data, machine learning (ML) and artificial intelligence (AI) techniques are widely applied to power system operation due to their powerful capability of extracting useful information and flexible extensibility. Based on DSSE, this dissertation further focus on developing data-driven AI-based methods for upgrading the monitoring and control of electric distribution systems with high penetration of DERs.

DSSE & Cyberattack Detection. The high dependence of cyber-physical power systems on information technology increases vulnerability from malicious cyberattacks. Considering the high dimensionality and correlated nature of power system measurements, the advanced ML technique using PMU data is applied to cyberattack detection and mitigation. Autoencoders are integrated into an advanced generative adversarial network (GAN) framework, which detects abnormal measurements under the unobservable cyberattacks by capturing the unconformity between anomalies and secure measurements. Also, because of the expensive labeling costs and potential missing labeled data in practical systems, this method only requires unlabeled data and a few labeled data from measuring instruments by leveraging a powerful generation capability of GAN and thus is semi-supervised learning. Specifically, the proposed method with as few as 1000 labeled training data self-learns with an accurate detection ability of higher than 95%.

DSSE & Voltage Control. Encouraged by AlphaGo's big hit, a real-time volt-VAR optimization (VVO) paradigm is developed to support distribution system operators in effective voltage control via deep reinforcement learning (DRL). This work is the first attempt to apply such an innovative AI technique to VVO control on voltage regulating devices in distribution systems. This method realizes power loss reduction and voltage regulation concurrently. In brief, under time-varying operating conditions, this work enables a power distribution grid to self-learn with the "cognitive" function of system operation by mimicking the human mind. Furthermore, by assigning the global control variables to multiple DRL agents with effective information exchange, the proposed method addresses the scalability issue and raise computational efficiency in larger-scale distribution systems, compared with single-agent DRL-based algorithms.

1.3. Chapter Organization

The remainder of this dissertation is organized as follows.

Chapter 2 proposes a novel and highly efficient DSSE algorithm with measurements from SCADA systems. The proposed method uses the Taylor series of voltages for constructing a linear DSSE model in the interval form and then solves this model by interval arithmetic.

Chapter 3 extends the work proposed in Chapter 2 and proposes a general ISE framework in DSSE integrating PMU data in unbalanced distribution systems. The proposed model simultaneously formulates multiple uncertainties in unbalanced distribution systems by interval arithmetic. Moreover, this model can accommodate partially available measurements of DG outputs and inaccurate line parameters. A modified Krawczyk-operator algorithm is proposed to solve the general ISE model, and effectively provides the upper and lower bounds of state variables under coordinated impacts of these uncertainties.

Chapter 4 proposes a graph-based faulted line identification algorithm using a limited number of PMUs. The proposed method first applies a distributed DSSE algorithm to efficiently restrict the searching region for the fault source in the feeder between two adjacent PMUs. Based on the graph model of the feeder in the reduced searching region, this method performed in a hierarchical structure further identifies the location of the fault source.

Chapter 5 explores the advanced semi-supervised learning technique for detecting cyberattack against DSSE. Autoencoders are integrated into a GAN framework, which detects abnormal measurements under the unobservable cyberattacks by capturing the unconformity between anomalies and secure measurements.

Chapter 6 develops a model-free VVO algorithm via multi-agent DRL, where the forward-backward sweep method for DSSE provides power flow results within a few iterations as the DRL environment. The VVO problem in unbalanced distribution networks is cast to an intelligent deep Q-network framework, which avoids solving a specific optimization model directly when facing time-varying system operating conditions.

Finally, a conclusion of this dissertation and future works unveiled during the research efforts are discussed in Chapter 7.

Chapter 2

Towards Highly Efficient State Estimation with SCADA Measurements in Distribution Systems

2.1. Motivation

Distribution systems are undergoing radical changes in operation and control, which calls for effective situation awareness techniques [13]. Distribution system state estimation (DSSE) converts redundant meter readings and other available information into an estimate of system states and thus develops into a subject of active research [14]. The measurements can be the voltage magnitudes, power injections, and power flows from supervisory control and data acquisition (SCADA) systems, or voltage and current phasor recorded by phasor measurement units (PMUs) [15]. Monte Carlo simulations (MCSs) [16] are widely used to obtain measurements samplings including noises owing to the assumption that these noises follow Gaussian distributions. Also, traditional nonlinear DSSE methods adopt the Gauss-Newton method based on the WLS criterion to perform the iterative estimation process. Besides, to correctly evaluate the estimation performance, the number of the required samples in MCSs is tremendous, which brings a heavy computation load to the WLS-based methods [13].

To mitigate the deficiency of MCSs, analytical methods such as nonlinear programming [17, 18] are proposed to provide the upper and lower bounds of all possible state variables that meet all constraints from measurements. For instance, the authors of [17] used a constrained nonlinear programming approach to obtain the ranges of states in transmission systems, while [18] extends this boundary optimization method to distribution systems with PMU installation. However, the method in [18] requires the installation of PMUs in distribution systems. Due to lack of PMUs in some distribution systems, the metering data

recorded by SCADA systems and pseudo-measurements collected at loads or distributed generators (DGs) are widely used in the existing DSSE methods [15]. Moreover, these measurements lead to nonlinear DSSE models that are iteratively solved, and thus the process is time-consuming. This chapter presents a highly efficient DSSE method to handle the uncertainty of random measurement noises. The main contributions of the proposed method are concluded as (i) constructing a novel DSSE model to avoid multiple runs of the WLS-based DSSE procedure and (ii) accelerating accurate state estimates by interval arithmetic with no requirements of PMU data.

2.2. Conventional DSSE

2.2.1. State Estimation Theory

In classical state estimators, the relationship between redundant measurements and state variables is depicted as:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e} \tag{2.1}$$

where $\boldsymbol{x}$ is an n-dimension state vector, and $\boldsymbol{z}$ is an m-dimension measurement vector; $\boldsymbol{h}(\boldsymbol{x})$ is the measurement function about $\boldsymbol{x}$; the measurement noise vector $\boldsymbol{e}$ obeys a Gaussian distribution $\boldsymbol{e} \sim N(0, \boldsymbol{R})$ is a covariance matrix and is usually considered diagonal (for instance, see [19]), $\boldsymbol{R} = \text{diag}\,[\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2]$, and σ_i^2 denotes the variance of the $\boldsymbol{i}$th measurement error, $i = 1, 2, \ldots, m$.

The state variables are obtained via a WLS criterion that minimizes the sum of weighted measurement residuals J as:

$$\hat{\boldsymbol{x}} = \arg\min J = \arg\min[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})]^T \boldsymbol{W} [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] \tag{2.2}$$

where $\boldsymbol{W}$ is a weight matrix of measurements to quantify the trust levels of diverse measurements, and $\boldsymbol{W} = \boldsymbol{R}^{-1}$.

Optimal estimated states are solved iteratively by the Gauss-Newton method until each component of the vector $\Delta \boldsymbol{x}$ at each iteration is sufficiently small:

$$\partial J/\partial \boldsymbol{x} = \boldsymbol{H}(\boldsymbol{x})^T \boldsymbol{W}[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] = 0 \tag{2.3}$$

$$\boldsymbol{H}(\boldsymbol{x})^T \boldsymbol{W} \boldsymbol{H}(\boldsymbol{x}) \Delta \boldsymbol{x} = \boldsymbol{H}(\boldsymbol{x})^T \boldsymbol{W}[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] \tag{2.4}$$

$$\boldsymbol{x}^{(t+1)} = \boldsymbol{x}^{(t)} + \Delta \boldsymbol{x} \tag{2.5}$$

where $\boldsymbol{H}(\boldsymbol{x})$ is the Jacobian matrix of the measurement function with respect to $\boldsymbol{x}$, and $\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{\partial h}(\boldsymbol{x})/\boldsymbol{\partial x}$.

2.2.2. Formulation of DSSE

In the conventional DSSE methods, SCADA systems provide the metering data of voltage magnitudes and powers, and pseudo-measurements are also used to achieve the system observability. Also, the substation acts as a slack bus [15].

The system states are usually chosen as the voltage phasors at all buses and expressed as

$$\boldsymbol{x} = [v_1, v_1, \cdots, v_n] \tag{2.6}$$

where v_j denotes the voltage phasor at bus j and $j = 1, \ \ldots, n$; n is the number of buses in the system.

The vector of measurements, $\boldsymbol{z}$, in DSSE generally includes 1) voltage and current phasors from distribution-level PMUs, 2) power flows recorded by supervisory control and data acquisition systems, and 3) power injections from smart meters or pseudo-measurements, including load consumption and DG [20]. The DSSE model is nonlinear since PMUs are not installed at each node in a practical distribution system [10].

Denote the voltage at node k as $\boldsymbol{V}_k$ and the current at branch i- j as $\boldsymbol{I}_{ij}$, $k \in \psi_V$ and $\{i, j\} \in \psi_I$; ψ_V and ψ_I are the sets of nodes and branches with voltage/current measurements

from limited PMUs installed in the distribution system. Power measurements exist at node k or branch i- j, $\{i,j\}$ *or* $k \in \psi_S$ is the set of load/DG nodes or the branches installed with a meter.

The measurement function of the three-phase voltage measurement at node k, $k \in \psi_V$, can be depicted as

$$\boldsymbol{V}_k = [V_k^a, V_k^b, V_k^c]^T = \boldsymbol{v}_k \tag{2.7}$$

The relationship between the current measurement at branch i- j and the state can be expressed as:

$$\boldsymbol{I}_{ij} = [I_{ij}^a, I_{ij}^b, I_{ij}^c]^T = \boldsymbol{Y}_{ij}(\boldsymbol{v}_i - \boldsymbol{v}_j) \tag{2.8}$$

where $\boldsymbol{Y}_{ij}$ denotes the line admittance at this branch.

The complex power measurements at node k or at branch i- j can be expressed as a nonlinear relationship about the states as

$$\boldsymbol{S}_k = \boldsymbol{v}_k \cdot (\boldsymbol{I}_k)^* \tag{2.9}$$

$$\boldsymbol{S}_{ij} = \boldsymbol{v}_i \cdot (\boldsymbol{I}_{ij})^* \tag{2.10}$$

where the power flow measurement $\boldsymbol{S}_{ij} \in \mathbb{C}^{3\times 1}$ and the power injection measurement $\boldsymbol{S}_k \in \mathbb{C}^{3\times 1}$; $\boldsymbol{v}_k$ and $\boldsymbol{v}_i$ denote the estimated voltages at node k and i, respectively, and come from the corresponding elements in $\boldsymbol{x}$; Further, $\boldsymbol{I}_{ij}$ can be obtained by (2.8), and the current injection at node k can be expressed as $\boldsymbol{I}_k = [I_k^a, I_k^b, I_k^c]^T = \boldsymbol{Y}_k \boldsymbol{v}_k$, where $\boldsymbol{Y}_k \in \mathbb{C}^{3\times 3}$ denotes the nodal admittance.

The DSSE model in the complex form is expressed as

$$\boldsymbol{z} = [\boldsymbol{V}_k, \boldsymbol{I}_{ij}, \boldsymbol{S}_k, \boldsymbol{S}_{ij}]^T = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e} \tag{2.11}$$

Due to the nonlinear relationships between the voltages and the power measurements, the model (2.11) is nonlinear. The DSSE process in the three-phase distribution system is

iteratively implemented in the following steps [21]:

1) *Backward Sweep*: Get initial values of branch currents by a backward approach. An initial voltage at each node is set as the substation voltage, *i.e.*, $\widetilde{\boldsymbol{v}}_i = \boldsymbol{V}_{slack}$, and (2.12) is used to calculate the current injections is calculated as through nodal power injections:

$$\boldsymbol{I}_{k\_eq} = (\boldsymbol{S}_k/\widetilde{\boldsymbol{v}}_i)^* \tag{2.12}$$

Next, these injections are used to obtain branch currents.

2) *Forward Sweep*: The branch currents in step 1) and the substation voltage are used to calculate initial nodal voltages, $\boldsymbol{x}^0$.

3) Obtain $\boldsymbol{h}(\boldsymbol{x})$ by (2.7)-(2.10) with the latest states $\boldsymbol{x}^t$, and then update system state variables as

$$\boldsymbol{\Delta x}^t = (\boldsymbol{H}(\boldsymbol{x}^t)^T \boldsymbol{W} \boldsymbol{H}(\boldsymbol{x}^t))^{-1} \boldsymbol{H}(\boldsymbol{x}^t)^T \boldsymbol{W} \left[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}^t)\right] \tag{2.13}$$

4) Update the nodal voltages by $\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \boldsymbol{\Delta x}^t$.

5) If $\boldsymbol{\Delta x}^t$ is less than a pre-set tolerance, stop the iterative process. Otherwise, go to step 3).

2.3. Proposed Methodology

2.3.1. DSSE Modeling in Interval Arithmatic

A linear approximation is developed in this section. Define $\Delta V_k = 1 - V_k$, where V_k denotes the voltage phasor at bus k. Considering the small voltage drops along the distribution lines and the normal voltage limits ($0.95 \sim 1.05$ p.u.) in practical systems [22], apply the Taylor series of ΔV_k around zero as $1(1 - \Delta V_k) = \sum_{n=0}^{\infty} (\Delta V_k)^n$. Further, the following equation is obtained by ignoring the high order terms [23]:

$$1/(1 - \Delta V_k) \approx 1 + \Delta V_k = 2 - V_k \tag{2.14}$$

Figure 2.1. Schematic diagram of the proposed linearization. (a) The complex plane of V_k. (b) Approximation loss $F(V_k)$

Fig.2.1 depicts the accuracy loss introduced by (2.14) as $F(V_k) = \frac{1}{V_k} - (2 - V_k)$, and for instance, the error for $V_k = 0.1$ is around 0.01. The approximation relationship $\frac{1}{V_k} = 2 - V_k$ is used for constructing the measurement functions in the proposed DSSE model, which are shown below.

1) Power Flow Measurements From SCADA System

The power flow measurements P_{ik} and Q_{ik} at branch i–k are expressed as

$$S_{ik} = P_{ik} + jQ_{ik} = V_i[y_{ik}(V_i - V_k)]^* \tag{2.15}$$

where P_{ik} and Q_{ik} denote the real and reactive powers at this branch; y_ik denotes the nodal admittance between buses i and k, and the function $[\cdot]^*$ represents the complex conjugate; V_i and V_k denote the voltage phasors at these two buses.

Apply (2.14) to (2.15), and a closed-form expression is obtained.

$$S_{ik}(2 - V_i) = V_i[y_{ik}(V_i - V_k)]^* \tag{2.16}$$

Formula (2.15) can be linearly expressed as $S_{ik}V_i + y_{ik}^* V_i^* - y_{ik}^* V_k^* = 2S_ik$.

2) Power Injection Measurements From SCADA System or Pseudo-measurements

The measurement function of the power injection at bus k holds below, and similar to (2.15), further expressed as a linear one.

$$S_k = P_k + jQ_k = V_k \sum_{l \in N(k)} [y_{lk}(V_l - V_k)]^* \tag{2.17}$$

$$S_k(2 - V_k) = \sum_{l \in N(k)} [y_{lk}(V_l - V_k)]^* \tag{2.18}$$

where P_k and Q_k denote the real and reactive powers at bus k, and $N(k)$ is the set of all buses connected to bus k.

3) Voltage Magnitudes From SCADA System

The measurement function of voltage magnitudes at bus k is approximated as

$$V_k = \sqrt{(V_{k,r}^2 + V_{k,x}^2} \approx V_{k,r} \tag{2.19}$$

where $V_{k,r}$ and $V_{k,x}$ represent the real and imaginary parts of voltages at bus k, and the small angle differences of distribution lines are considered, e.g., 0.1 degrees per mile [2].

Reorganize the linear equations (2.16), (2.18), and (2.19) as

$$\boldsymbol{BV} + \boldsymbol{DV}^* = \boldsymbol{E} \tag{2.20}$$

where $\boldsymbol{V}$ denotes the vector of the nodal voltage phasors, and $\boldsymbol{B}$, $\boldsymbol{D}$, and $\boldsymbol{E}$ are the corresponding coefficient matrices and suppressed here due to the limited space. Further, express (2.20) in rectangular coordinates as

$$\begin{bmatrix} \boldsymbol{B}_r + \boldsymbol{D}_r & -\boldsymbol{B}_x + \boldsymbol{D}_x \\ \boldsymbol{B}_x + \boldsymbol{D}_x & \boldsymbol{B}_r - \boldsymbol{D}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_r \\ \boldsymbol{V}_x \end{bmatrix} = \begin{bmatrix} \boldsymbol{E}_r \\ \boldsymbol{E}_x \end{bmatrix} \tag{2.21}$$

where the subscripts r and x denote the real and imaginary parts of complex numbers. For simplicity, (2.21) is expressed as

$$\boldsymbol{Ax} = \boldsymbol{b} \tag{2.22}$$

where $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{V}_r \\ \boldsymbol{V}_x \end{bmatrix}$ denotes the state vector; $A = \begin{bmatrix} \boldsymbol{B}_r + \boldsymbol{D}_r & -\boldsymbol{B}_x + \boldsymbol{D}_x \\ \boldsymbol{B}_x + \boldsymbol{D}_x & \boldsymbol{B}_r - \boldsymbol{D}_r \end{bmatrix}$, and $\boldsymbol{A} \in \mathbb{R}^{m \times n}$; rank($\boldsymbol{A}$)=n, *i.e.*, a full rank; $\boldsymbol{b} = \begin{bmatrix} \boldsymbol{E_r} \\ \boldsymbol{E_x} \end{bmatrix}$ and $\boldsymbol{b} \in \mathbb{R}^{m \times 1}$.

Note that (2.22) does not consider measurement noises and involves various levels of approximation on the voltage magnitudes and powers. Next, based on (2.22), we use interval arithmetic to handle these accuracy losses in DSSE. We consider the measurement noises by updating (2.22) to an interval equation, where an interval number is defined as $[a] = [a_l, a_u] = \{a \in \mathbb{R} \mid a_l \leq a \leq a_u\}$, and interval vectors and matrices are constructed similarly [24]. According to the 3σ rule of a Gaussian distribution, where σ denotes the standard deviation, 99.73% of values from the distribution are within three times of standard deviations [2]. Hence, the maximum measurement errors (*i.e.*,$\mp 3\sigma$) are superposed onto the corresponding measurements to obtain the lower and upper bounds of $\boldsymbol{A}$ and $\boldsymbol{b}$, *i.e.*, $[\boldsymbol{A}]$ and $[\boldsymbol{b}]$. By this relaxation, these measurement intervals enclose their true values.

2.3.2. Solution in Interval Arithmetic

This section presents a novel solving method towards the linear DSSE model by further taking into account measurement noises based on (2.22). Also, interval arithmetic [24] is used to handle the accuracy loss introduced by the linear approximation and these measurement noises in the DSSE process.

The proposed method considers the measurement noises by updating (2.22) to an interval equation, where an interval number is defined as $[a] = [a_l, a_u] = \{a \in \mathbb{R} \mid a_l \leq a \leq a_u\}$, and an interval vector and matrix are constructed similarly. The maximum measurement errors ($\pm 3\sigma$, and σ is the standard deviation of a Gaussian distribution) are superposed onto the corresponding measurements to obtain the lower and upper bounds of $\boldsymbol{A}$ and $\boldsymbol{b}$, *i.e.*, $[\boldsymbol{A}]$ and $[\boldsymbol{b}]$. By this relaxation, these measurement intervals enclose the true values of measurements

$$[\boldsymbol{A}][\boldsymbol{x}] = [\boldsymbol{b}] \Rightarrow \begin{bmatrix} [\boldsymbol{A}] & -\boldsymbol{I} \\ \boldsymbol{0} & [\boldsymbol{A}]^T \end{bmatrix} \begin{bmatrix} [\boldsymbol{x}] \\ [\boldsymbol{y}] \end{bmatrix} = \begin{bmatrix} [\boldsymbol{b}] \\ \boldsymbol{0} \end{bmatrix} \tag{2.23}$$

where $\boldsymbol{I} \in \mathbb{R}^{m\times m}$ is an identity matrix, and $[\boldsymbol{y}] \in \mathbb{R}^{m\times 1}$.

Compactly express (2.23) as $[\boldsymbol{A}][\boldsymbol{X}] = [\boldsymbol{B}]$, and the interval symbol $[\cdot]$ is omitted below, e.g., the augmented state vector $\boldsymbol{X} = \begin{bmatrix} [\boldsymbol{x}] \\ [\boldsymbol{y}] \end{bmatrix}$. Equation (2.23) is solved by a Krawczyk-operator algorithm. An initial interval solution $\boldsymbol{X}^{(0)}$ is calculated by

$$\boldsymbol{X}^{(0)} = ([-\alpha, \alpha], \ldots, [-\alpha, \alpha])^T \tag{2.24}$$

where $\alpha = \|\boldsymbol{CB}\|_\infty/(1-\beta)$ and $\beta = \|\boldsymbol{I} - \boldsymbol{CA}\|_\infty$; $\boldsymbol{C}$ is a preconditioning point matrix, $\boldsymbol{C}^{-1}$=$Mid[\boldsymbol{A}]$, and $Mid[\cdot]$ is the medium of intervals; $\|\cdot\|_\infty$ denotes the infinite norm of a vector, and this initial solution $\boldsymbol{X}^{(0)}$ contains the final solution.

The following process at iteration j is used to gradually approach the final solution hull until $\|\boldsymbol{X}^{(j+1)} - \boldsymbol{X}^{(j)}\|_\infty \leq \epsilon$:

$$\boldsymbol{X}^{(j+1)} = (\boldsymbol{CB} + (\boldsymbol{I} - \boldsymbol{CA})\boldsymbol{X}^{(j)}) \cap \boldsymbol{X}^{(j)} \tag{2.25}$$

where $\epsilon = 10^{-4}$. It is proved that starting from $\boldsymbol{X}^{(0)}$, the iterative process rapidly converges if $\|\boldsymbol{I} - \boldsymbol{CA}\| < 1$, and $\|\cdot\|$ is any norm [24].

The final state estimate of $\boldsymbol{x}$ in the proposed DSSE model is obtained by Mid$[\boldsymbol{X}^{(j+1)}]$ considering that the measurement noises follow symmetric Gaussian distributions about zero means.

2.4. Numerical Test

We test the proposed algorithm on the IEEE 34-bus distribution system [25]. The system is modified by adding four DGs , shown as Fig. 2.2:, and the installed capacity of each DG

Table 2.1. Measurement arrangement in test systems

Measurements		Location
SCADA	$\|V\|$	1, 11, 20, 25
	P, Q	1-2, 4-6, 13-15, 20-23,30-31
Pseudo-meas.	P, Q	All load nodes and DG nodes

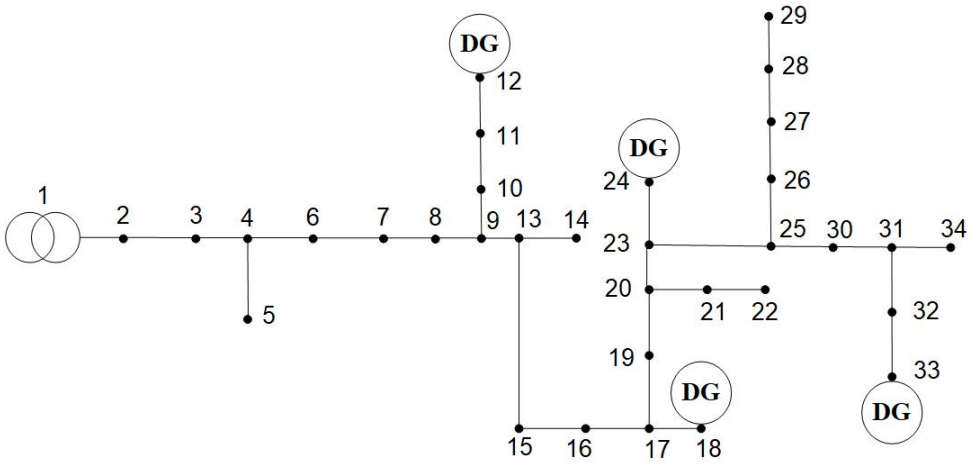

Figure 2.2. IEEE 34-bus distribution system

is 200 kVA. The maximum errors of measurements are set as 1% of these true values for the voltage magnitudes and powers from SCADA systems and 20% for pseudo-measurements. Table 2.1 displays the measurement placement scheme in the test system.

2.4.1. Estimation Performance

To evaluate the performance of the proposed method, we adopt the nonlinear WLS-based method in (2.1)-(2.5) as the baseline. In Fig. 2.3, the maximum absolute errors (MAEs) of the real and imaginary parts of voltages in 1000 MCSs are used to evaluate the estimation accuracy of the nonlinear WLS-based method in [15]. Also, using measurements from one of these MCSs, we calculate the errors of the estimated voltages by the proposed method. Fig. 2.3 implies that the WLS-based method in MCSs may produce the MAEs that reach up to around 5.34×10^{-3} p.u., while the proposed method obtains accurate estimates with the

maximum error 1.54×10^{-3} p.u. The root mean square errors (RMSEs) of these estimated voltages in all MCSs are used to evaluate the overall estimation performance of this nonlinear DSSE method.

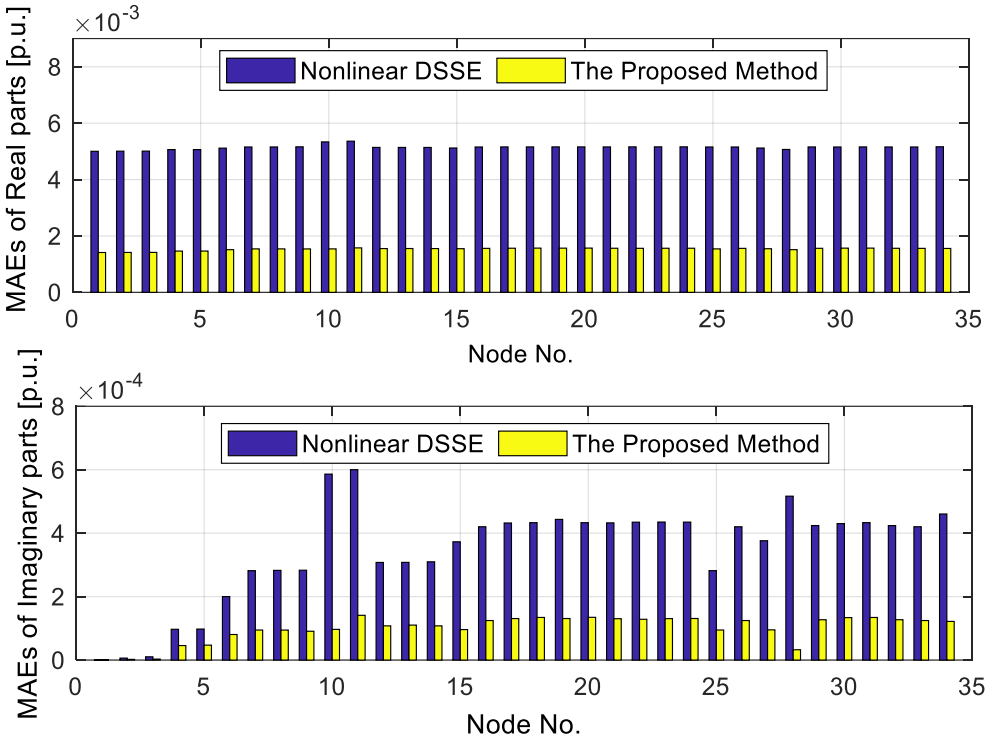

Figure 2.3. Comparison in the estimation errors of the proposed method and the MAEs that may happen in a random sampling of measurements.

Moreover, the maximums of the RMSEs are compared with the MAEs of the proposed method at all buses, shown in Table 2.2. In contrast to the WLS-based method in a Monte Carlo trial, the proposed method obtains the states more accurately. Also, the estimation accuracy of our method is close to that of the baseline method in 1000 MCSs, and however, the latter requires considerable sets of samplings.

Table 2.2 compares the computational efficiency of the proposed method with that of the WLS-based method. The CPU time of the proposed method accounts for about 12% of the average one that this nonlinear method takes in a single Monte Carlo trial, *i.e.*, 20.30 vs 174.79 milliseconds.

Table 2.2. Comparison in accuracy and computation time

Algorithm Type	Maximum Errors or RMSEs [p.u.]		CPU Time
	Real Part	Imaginary Part	
Proposed Method	1.541×10^{-3}	1.451×10^{-4}	20.30 [ms]
WLS-based DSSE	1.681×10^{-3}	1.842×10^{-4}	174.7 [s] for 1000 trials

2.4.2. Robustness Analysis

This section discusses influence factors, such as operating condition and measurement redundancy, to illustrate the robustness of the proposed method.

Considering the impacts of DG penetration on voltage profile, we investigate the estimation results of the proposed method in various operation ranges, *i.e.*, 0.90∼0.95 p.u. and 0.9∼1.1 p.u., shown in Table 2.3. The comparison between these cases illustrates that a narrower voltage range around 1.0 leads to higher estimation accuracy and computational efficiency of the proposed method. Table 2.3 validates the robustness of this method for various operating conditions.

Table 2.3. Estimation performance in robustness analysis

Robustness Analysis		Maximum Errors or RMSEs [p.u.]		CPU Time
		Real Part	Imaginary Part	[ms]
Voltage Profile	0.90∼0.95	1.652×10^{-3}	1.607×10^{-3}	20.43
	0.90∼1.1	1.831×10^{-3}	1.793×10^{-3}	21.41
Measurement Redun.	1.221	1.330×10^{3}	1.693×10^{-4}	18.30
	1.265	1.029×10^{3}	1.642×10^{-4}	17.11

The measurement redundancy shown in Table 2.1 is 1.176, and we further test the proposed algorithm with other measurement redundancies by adjusting the number of measurements and their locations. Table 2.3 gives the estimation performance of these tests, which shows that the efficacy of this method does not depend on the measurement arrangements. Also, the higher measurement redundancy leads to the overall improvement in estimation accuracy and computational efficiency.

In Fig. 2.2, the maximum absolute errors (MAEs) of the real and imaginary parts of voltages in 1000 times of MCSs are used to evaluate the estimation accuracy of the nonlinear DSSE method in [15]; Fig. 3 also lists the errors of the estimated voltages in the proposed method. It implies in Fig.3 that a single Monte Carlo trial may produce significant estimation errors that reach up to around 5.34×10^3 p.u., while the proposed method obtains accurate estimates with the maximum error 1.54×10^3 p.u. The root mean square errors (RMSEs) of these estimated voltages in all MCSs are also used to evaluate the overall estimation performance of this nonlinear DSSE method. Moreover, the maximums of the RMSEs at all nodes are compared with those of the errors of the proposed method, shown in Table 2.3. It concludes that in contrast to a Monte Carlo trial, the proposed method estimates the states via a single run accurately and robustly; the estimation accuracy of this method is close to that of 1000 MCSs, which require considerable sets of samplings and high computational cost.

2.5. Conclusion and Discussion

This chapter presents a highly efficient DSSE algorithm using the Taylor series of complex numbers and interval arithmetic techniques. The proposed method provides a highly efficient substitute of nonlinear DSSE methods without the use of PMU measurements. Numerical simulations illustrate the accuracy and efficiency of the proposed method in tackling nonlinear measurements with Gaussian noises.

Owing to the characteristics of distribution systems, several other factors associated with distributed energy resource (DER) penetrations further increase the challenges and difficulties of DSSE, which are delineated as follows:

Accurate point forecasts for DER outputs are difficult. The power outputs of DERs are mostly dependent upon the weather conditions and multiple ambient factors. Wind speed is regarded as a typical deciding factor of wind power, while the power generation of a photovoltaic cell mainly depends on solar radiation and ambient temperature [26]. Since accurate quantification of wind speed or solar radiation is difficult to obtain, their forecasting

errors inevitably lead to considerable forecasting errors of DER outputs [27].

DER operates with unbalanced and bidirectional power flows. Conventional distribution power flows in one direction, *i.e.*, from central power plants to customers, while the reverse flow from customers to the grid occurs with DERs integrated. In addition, practical power systems at the distribution level operate with unbalanced power flow, which is usually the case where multiple DERs are installed in distribution systems [26].

Lack of meters at DER installations. Due to currently limited metering and communication infrastructure, distribution systems are generally underdetermined with poor observability. Besides, the assumption that measurement devices are mounted at DER installations might be impractical owing to the lack of specific agreements between DER operators and distribution system operators (DSOs) [28].

Spatial correlations of DER outputs. DERs that belong to the same geographical area present similarity in power outputs. Weather conditions and field operations (e.g., active power curtailment or reactive generation control) affect the correlation between DERs at different nodes of the system [29].

All the above challenges deserve further investigation. Therefore, by updating the proposed DSSE algorithm in this chapter, Chapter 3 will formulate the uncertainty from DER integration to develop high-performance DSSE in three-phase unbalanced distribution systems.

Chapter 3

Interval State Estimation with Uncertainty of Distributed Generation and Line Parameters Integrating PMU Data

3.1. Literature Review

Distribution system state estimation (DSSE) is largely driven by the diffusion of distributed generation (DG) such as wind turbine generators (WTGs) and photovoltaic (PV) panels [30]. DG has advantages of low investment costs, flexible and eco-friendly operations, and low power losses. However, the variability and intermittency of DG pose significant uncertainty to DSSE [31]. Apart from the uncertainty from these emerging DG units, inputs to DSSE also contain measurements with noises and imprecise line parameters. For example, the uncertainty of line parameters originates from varying field ambient conditions and aging wirings. While the uncertainty of measurements is commonplace and their impacts on the DSSE are investigated in [32], multiple other uncertainties besides DG uncertainty in DSSE call for innovative solutions.

Uncertainty studies that account for the variable and stochastic nature of input data in conventional DSSE are pursued by Monte Carlo (MC) simulation, as reviewed in [13]. In these studies, DG power outputs are assumed available in real time via measuring instruments installed by distribution system operators (DSOs) [33] or in the form of pseudo-measurements that follow Gaussian distributions [34]. Nevertheless, these assumptions might be impractical due to 1) currently limited metering and communication infrastructure, and 2) lack of specific agreements between DG operators and DSOs [28]. Moreover, the stochastic nature of DG outputs weakens the assumption that the DG outputs follow a known family of parametric distributions [15]. As discussed in [17, 26, 35–40], the statistical data of DG outputs are a prerequisite in MC simulation, however, such information may not be available

in practice. In addition, these methods based on MC simulation require plenty of runs for various combinations of measurement samplings and/or DG outputs and thus are generally applied for evaluating the overall accuracy of state estimators [32].

Motivated by the deficiencies of these methods, interval state estimation (ISE) is proposed to obtain the boundaries of state variables, which provide more intuitive information such as the upper and lower bounds of these states [17]. In ISE, all data with uncertainty are modeled as inputs in the interval form, since the upper and lower limits are available in most practical cases. For example, the range of line parameters can be specified (e.g., within $\pm$ 5% of their nominal values) [40].

Boundary optimization methods such as [17, 35–37] are proposed to address the ISE problem by maximizing and minimizing the state variables that meet all constraints from measurements. For instance, in [17], a constrained nonlinear programming approach using the measurements from supervisory control and data acquisition (SCADA) systems is used to obtain the ranges of states in transmission systems. The solving method is applied to distribution systems in [37]. However, with more PMUs or micro-PMUs emerging at the distribution level, this approach cannot deal with hybrid measurements including PMU data. Moreover, the authors of [37] did not consider the uncertainty of line parameters. In addition, the limitation of these optimization-based methods is that lower and upper bounds of each state variable need to be computed separately, and thus the total number of the optimization models for all states proliferates with the scale of the distribution systems. This leads to their low efficiency in the solving process.

Recent efforts to apply interval arithmetic to study uncertainties in power system operations are noteworthy, such as power flow calculation in [38], reliability evaluation in [39], and ISE in [40] and [41]. In ISE, interval arithmetic deals with the uncertain inputs that lie within a certain interval and enables the direct computation towards the bounds of state variables. For instance, focusing on transmission systems, an ISE model with only PMU measurements is formulated as interval linear equations in [40], and the ranges of states with the line parameter uncertainty are solved. However, such a high PMU deployment is not available at the

distribution level, and the impacts of DG uncertainty are not considered. Further research is conducted in active distribution systems. In [41], an iterative Krawczyk-operator algorithm is used to obtain interval states, which takes the solution solved by interval Gaussian elimination (IGE) as initial values of the states. Nevertheless, IGE presents the drawback of "wrapping effect", where the widths of intervals expand since each interval variable is treated independently. When IGE is applied to distribution systems with high uncertainty, this over-conservatism is further intensified. As a result, the IGE-based Krawczyk operator (IKO) is computationally expensive to obtain final states since the initial states are far from them [37]. In addition, all DG outputs in [41] are assumed to obey Gaussian distributions. As mentioned above, this assumption may not be practical.

To sum up, the existing studies still lack generality in modeling to formulate multiple uncertainties, and most of them focus on the modeling for a single type of uncertainty in DSSE. Moreover, the direct impacts of uncertain DG outputs on DSSE are not fully addressed in unbalanced distribution systems. The existing ISE methods are established on the strong assumptions that the probability information or real-time measurements of DG outputs are available. In addition, the limitations of the solution strategies for the existing ISE models, including conservative estimation results and time-consuming solving process, persist.

We propose a novel and fast ISE algorithm considering multiple uncertainties of DG outputs and line parameters in unbalanced distribution systems. As a solid reference to DSOs, the upper and lower bounds of state variables are provided by the proposed algorithm with hybrid SCADA and PMU measurements. First, based on the interval prediction for DG power outputs, an ISE model is formulated in interval arithmetic. Moreover, the model consider the line parameter uncertainty, and a weighted least square (WLS) criterion is integrated to deal with these hybrid measurements. Finally, a modified Krawczyk-operator (MKO) algorithm, which enables fast and accurate computation towards the bounds of state variables, is presented to obtain the interval solution of the proposed ISE model.

3.2. Hybrid DSSE Algorithm and Interval Arithmetic

3.2.1. DSSE with Hybrid Measurements

Recent research interests focus on the applications of PMUs to DSSE, since PMUs measure voltage and current phasors with high sampling precision and short update cycles [28]. Considering that a limited number of PMUs are installed in distribution systems due to their high technical and financial costs, hybrid state estimators incorporate conventional SCADA data with PMU data to improve estimation accuracy. Moreover, by adopting the state variables in the rectangular form, hybrid DSSE integrating PMU data results in a linear estimator, while conventional estimators with SCADA data are nonlinear [14].

Based on [13] and [42], the hybrid estimator in [43] is used owing to its improved and recognized performance, where the voltage at a slack node (*i.e.*, a substation) and branch currents are chosen as state variables. In three-phase distribution systems, the state vector is expressed in rectangular coordinates as

$$\boldsymbol{x} = \left[v_{slack,r}^{a,b,c}, v_{slack,x}^{a,b,c}, i_{1r}^{a,b,c}, \ldots, i_{Nr}^{a,b,c}, i_{1x}^{a,b,c}, \cdots, i_{Nx}^{a,b,c}\right] \tag{3.1}$$

where $v_{slack,r}^{a,b,c}$ and $v_{slack,x}^{a,b,c}$ are the real and imaginary parts of the three-phase substation voltage, and $i_{lr}^{a,b,c}$ and $i_{lx}^{a,b,c}$ are the real and imaginary parts of the three-phase current at branch l, $l = 1,\ 2, \ldots, N$. The superscripts a, b, and c denote the phase indices.

In the hybrid estimator, power flows and power injections at loads are obtained by SCADA systems or pseudo-measurements, while PMUs provide the magnitude and phase angle measurements of voltages and currents. Moreover, the three-phase power measurements are converted into equivalent currents by

$$z_{I_{kr}} + j z_{I_{kx}} = \left[\frac{z_{P_k} + j z_{Q_k}}{V_k}\right]^* \tag{3.2}$$

$$z_{I_{pr}} + j z_{I_{px}} = \left[\frac{z_{P_p} + j z_{Q_p}}{V_p}\right]^* \tag{3.3}$$

where $z_{I_{kr}}$ and $z_{I_{kx}}$ (or $z_{I_{pr}}$ and $z_{I_{px}}$) are the real and imaginary parts of the current at node k (or at branch p), and V_k and V_p are the voltage phasors at node k and connected to branch

p, respectively; z_{P_k} and z_{Q_k} (or z_{P_p} and z_{Q_p}) denote the active and reactive powers at node k (or at branch p). $[\cdot]^*$ denotes the complex conjugate.

The hybrid DSSE process in an unbalanced distribution system is iteratively implemented in the following steps [42]:

1. **Backward Sweep**: Get initial values of branch currents by a backward approach. An initial voltage at each node is set as the substation voltage V_{slack}, and (3.2) is used to calculate current injections through nodal power injections as follows:

$$z_{I_{kr}} + j z_{I_{kx}} = \left[\frac{z_{P_k} + j z_{Q_k}}{V_{slack}} \right]^* \tag{3.4}$$

   Next, these injections are used to obtain branch currents.

2. **Forward Sweep**: The branch currents in step 1 and the substation voltage are used to calculate initial nodal voltages.

3. Calculate $\boldsymbol{h}(\boldsymbol{x})$, and then update system state variables as

$$\Delta \boldsymbol{x}^k = (\boldsymbol{H}^T(\boldsymbol{x}^k) \boldsymbol{W} \boldsymbol{H}(\boldsymbol{x}^k))^{-1} \boldsymbol{H}^T(\boldsymbol{x}^k) \boldsymbol{W} [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}^k)] \tag{3.5}$$

4. Update the branch currents by $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{\Delta x}^k$, then calculate nodal voltages by forward sweep in step 2.

5. If $\Delta \boldsymbol{x}^k$ is less than a pre-set tolerance, stop the iterative process. Otherwise, use these updated voltages to calculate the equivalent currents by (3.2) and (3.3), then go to step 3.

To improve the computational efficiency in the above iterative process, a linear approximation technique in [44] is used, where each nodal voltage is fixed as V_{slack} to calculate equivalent currents in (3.2) and (3.3). This approximation is based on two observations for practical distribution systems, that 1) the voltage drops along feeders do not exceed 5%, since voltage phase differences are very small [22], and 2) nodal voltages are usually kept

within normal operation limits (0.95 to 1.05 p.u.) [45]. Then, the Jacobian matrix is independent of $\boldsymbol{x}$ and highly sparse, and the general formulas (2.3) and (2.4) are updated with $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{H}\boldsymbol{x}$ as:

$$\partial \boldsymbol{J}/\partial \boldsymbol{x} = \boldsymbol{H}^T \boldsymbol{W} \boldsymbol{z} - \boldsymbol{H}^T \boldsymbol{W} \boldsymbol{H} \boldsymbol{x} = 0 \tag{3.6}$$

$$\boldsymbol{x} = (\boldsymbol{H}^T \boldsymbol{W} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{W} \boldsymbol{z} \tag{3.7}$$

The constant Jacobian elements of this estimator are briefly listed below, and more details about $\boldsymbol{H}$ can be referred to [43].

1) PMU voltage measurements

For the PMU voltage at node k, the measurement function is shown as

$$h(x) = h_{V_{kr}}^{a,b,c} + j h_{V_{kx}}^{a,b,c} = (v_{slack,r}^{a,b,c} + j v_{slack,x}^{a,b,c}) - \sum\nolimits_{p \in \mathfrak{I}_k} (i_{pr}^{a,b,c} + j i_{px}^{a,b,c}) z_p \tag{3.8}$$

where $\mathfrak{I}_k$ denotes a set of line segments from the slack node to node k, and $p \in \mathfrak{I}_k$; $\boldsymbol{z_p}$ is the 3×3 impedance matrix of branch p, and $v_{slack,r}^{a,b,c}$, $v_{slack,x}^{a,b,c}$, $i_{pr}^{a,b,c}$, and $i_{px}^{a,b,c}$ as state variables are defined in (3.1). Also, the off-diagonal elements with non-zero values in z_p reflect the coupling effect among three-phase lines, and z_p can be shown as:

$$\boldsymbol{z}_p = \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} = \begin{bmatrix} r_p^{aa} & r_p^{ab} & r_p^{ac} \\ r_p^{ba} & r_p^{bb} & r_p^{bc} \\ r_p^{ca} & r_p^{cb} & r_p^{cc} \end{bmatrix} + j \begin{bmatrix} x_p^{aa} & x_p^{ab} & x_p^{ac} \\ x_p^{ba} & x_p^{bb} & x_p^{bc} \\ x_p^{ca} & x_p^{cb} & x_p^{cc} \end{bmatrix}$$

where the diagonal and off-diagonal elements such as Z_{aa} and Z_{ab} denote the self-impedances and mutual impedances between two phases, respectively, shown as Fig. 3.1. Besides, the mutual impedances between any two phases may not be the same each other due to the unbalanced nature of the system.

Take the A-phase voltage at node k as an example, the non-zero Jacobian elements of

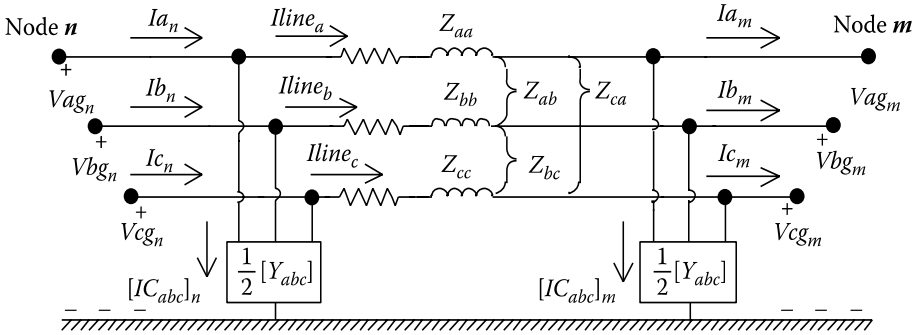

Figure 3.1. Three-phase line model in unbalanced distribution systems

the function (3.8) are expressed as:

$$\frac{\partial h^a_{V_{kr}}}{\partial v^a_{slack,r}} = 1 \qquad \frac{\partial h^a_{V_{kx}}}{\partial v^a_{slack,x}} = 1$$

$$\frac{\partial h^a_{V_{kr}}}{\partial i^a_{pr}} = -r^{aa}_p \quad \frac{\partial h^a_{V_{kr}}}{\partial i^a_{px}} = x^{aa}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^a_{pr}} = -x^{aa}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^a_{px}} = -r^{aa}_p$$

$$\frac{\partial h^a_{V_{kr}}}{\partial i^b_{pr}} = -r^{ab}_p \quad \frac{\partial h^a_{V_{kr}}}{\partial i^b_{px}} = x^{ab}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^b_{pr}} = -x^{ab}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^b_{px}} = -r^{ab}_p$$

$$\frac{\partial h^a_{V_{kr}}}{\partial i^c_{pr}} = -r^{ac}_p \quad \frac{\partial h^a_{V_{kr}}}{\partial i^c_{px}} = x^{ac}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^c_{pr}} = -x^{ac}_p \quad \frac{\partial h^a_{V_{kx}}}{\partial i^c_{px}} = -r^{ac}_p$$

where $r^{a\varphi}_p$ denotes the mutual or self-resistance at branch p, $p \in \mathfrak{I}_k$, and $x^{a\varphi}_p$ denotes the mutual or self-reactance, $\varphi = a, b, c$. These phase indices are omitted for simplicity below.

2) PMU current measurements

For the PMU current at branch p,$h(x) = h_{I_{pr}} + jh_{I_{px}} = i_{pr} + ji_{px}$, and the only nonzero Jacobian elements are shown as:

$$\frac{\partial h_{I_{pr}}}{\partial i_{lr}} = \begin{cases} 1, & when\ p = l \\ 0, & elsewhere \end{cases} \qquad \frac{\partial h_{I_{px}}}{\partial i_{lx}} = \begin{cases} 1, & when\ p = l \\ 0, & elsewhere \end{cases}$$

where l denotes the index of all branches.

3) Power measurements including line flows and power injections

For the power injections at node k, the Jacobian elements only has nonzero values of $+1$ and -1, since the measurement function holds based on Kirchhoff's current law:

$$h(x) = h_{I_{kr}} + jh_{I_{kx}} = i_{in,r} + ji_{in,x} - \sum (i_{out,r} + ji_{out,x}) \tag{3.9}$$

where $i_{in,r} + ji_{in,x}$ and $i_{out,r} + ji_{out,x}$as state variables denote the input and output currents at node k, respectively.

Besides, the Jacobian elements of the equivalent currents at branch p from the corresponding line flows only have nonzero values of $+1$, similar to PMU current measurements.

3.2.2. Interval Arithmetic and Interval Prediction of DG

An interval number is defined as a compact set $[a] = [a_l, a_u] = \{a \in \mathbb{R} \mid a_l \leq a \leq a_u\}$, and similarly, an interval vector is defined as a vector with interval elements [26].

When meters are not available at DG locations, effective forecasting techniques are utilized to obtain DG power outputs as pseudo-measurements for achieving system observability [1]. Due to the difficulty in accurate forecasts for instantaneous wind speeds or solar radiations, their forecast errors inevitably result in considerable forecast errors of DG outputs [46]. Hence, DG outputs in the interval form are modeled to quantify the uncertainty levels in interval predictions, which is more feasible in practice [26]. Conventional pseudo-measurements originate from the historical or forecast data on generator production and load consumption. Moreover, they obey Gaussian distributions with high-level noises as in [34] or are represented as other known distribution information as in [28]. In the proposed algorithm, we relax these assumptions and use the interval prediction of DG outputs. Further, the interval DG outputs are deemed as another form of pseudo-measurements in DSSE to obtain the interval estimate of states in the subsequent section.

3.3. General ISE Framework and Proposed Algorithm

3.3.1. ISE Model with Multiple Uncertainties

In this section, an ISE model with multiple uncertainties in DSSE is proposed, where measurements with noise, uncertain DG outputs, and imprecise line parameters are considered. The above highly efficient estimator is used to achieve fast monitoring of distribution networks with these uncertainties.

The impacts of the uncertainty sources on the deterministic estimation model (3.6) are analyzed and then updated into an interval estimation model as

$$\begin{bmatrix} [\boldsymbol{H}_1] \\ \boldsymbol{H}_2 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{W}_1 & 0 \\ 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} [\boldsymbol{H}_1] \\ \boldsymbol{H}_2 \end{bmatrix} [\boldsymbol{x}] = \begin{bmatrix} [\boldsymbol{H}_1] \\ \boldsymbol{H}_2 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{W}_1 & 0 \\ 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} [\boldsymbol{z}_1] \\ [\boldsymbol{z}_2] \end{bmatrix} \tag{3.10}$$

where $[\boldsymbol{x}]$ is an interval state vector, and $[\boldsymbol{x}] \in \mathbb{R}^{n\times 1}$; $[\boldsymbol{z}_1]$ denotes an interval measurement vector, and $[\boldsymbol{z}_1] \in \mathbb{R}^{m_1\times 1}$, while $[\boldsymbol{z}_2]$ denotes an interval vector of DG power outputs, and $[\boldsymbol{z}_2] \in \mathbb{R}^{m_2\times 1}$; $[\boldsymbol{H}_1]$ and $\boldsymbol{H}_2$ are the Jacobian matrices of the measurements and the DG outputs, and $[\boldsymbol{H}_1] \in \mathbb{R}^{m_1\times n}$, $\boldsymbol{H}_2 \in \mathbb{R}^{m_2\times n}$; $\boldsymbol{W}_1$ is the weighted matrix of the measurements, and $\boldsymbol{W}_1 \in \mathbb{R}^{m_1\times m_1}$; $\boldsymbol{I}$ is an identity matrix, and $\boldsymbol{I} \in \mathbb{R}^{m_2\times m_2}$.

The top-row equation describes the relationship between the measurements and $[\boldsymbol{x}]$; the bottom-row equation provides the constraints for the states related to DG outputs. The uncertain outputs of DG are modeled as pseudo-measurements of the system according to their interval predictions mentioned in Section 3.2.2. Note that only solving the equation set at the top row in (3.10) may not obtain the solutions of these state variables. This is owing to practical engineering concerns that the measurements of DG outputs are not available at all DG locations. As a result, the top-row system with measurements may be unobservable by DSSE due to the lack of necessary measurements [33]. The formula (3.10) ensures that an interval solution not only meets a WLS criterion for all measurements but also follows the relationship with these interval DG outputs.

According to different uncertainty sources, the details of (3.10) are discussed for clarity.

1) **Measurements**. Power measurements are converted to equivalent currents by (3.4),

then $[\boldsymbol{z}_1]$ is expressed as

$$[z_1] = \begin{bmatrix} [\boldsymbol{U}_l, \boldsymbol{U}_u] \\ [\boldsymbol{I}_l, \boldsymbol{I}_u] \\ [\boldsymbol{I}_{eq,l}, \boldsymbol{I}_{eq,u}] \end{bmatrix} \tag{3.11}$$

where $\boldsymbol{U}_l$ and $\boldsymbol{I}_l$ represent the lower bounds of voltage and current vectors from PMU measurements, and $\boldsymbol{I}_{eq,l}$ is the lower bound of equivalent current measurements, while $\boldsymbol{U}_u$, $\boldsymbol{I}_u$, and $\boldsymbol{I}_{eq,u}$ represent the corresponding upper bounds.

The 3σ deviation criterion about the mean in a Gaussian distribution covers more than 99.7% of the area of the distribution is used to obtain $[\boldsymbol{z}_1]$ based on measurements with noises [34]. For any measurement Z_m with a random noise $\boldsymbol{e}\sim N(0,\sigma^2)$, $Z_m \in [Z_0-3\sigma, Z_0+3\sigma]$, where Z_0 represents the true value. Hence, $Z_0 \in [Z_m - 3\sigma, Z_m + 3\sigma]$, and this measurement is modeled as an interval enclosing the corresponding true value.

2)**DG outputs**. The upper and lower bounds of DG outputs as pseudo-measurements are transformed to equivalent currents by (3.4), respectively. $[\boldsymbol{z}_2] = [\boldsymbol{I}_{DG,l}, \boldsymbol{I}_{DG,u}]$, where $\boldsymbol{I}_{DG,l}$ and $\boldsymbol{I}_{DG,u}$ represent the lower and upper bounds of the equivalent currents. As discussed in (3.9), the Jacobian matrix $\boldsymbol{H}_2$ related with DG outputs only has nonzero elements of $+1$ and -1. Besides, we relax the strong assumption that the statistical information of these DG outputs is known, *i.e.*, there is no requirement of the knowledge of the mean and covariance of DG outputs. Such correlations can be translated into respective DG output intervals and then the proposed method can be still applicable. The detailed consideration of the DG correlation will be left for our future work.

3) **Line Parameters**. The uncertainty of line parameters is evaluated in a range based on their nameplate values [40]. In the adopted estimator, the uncertain line parameters are only present at the locations corresponding to PMU voltage measurements, *i.e.*, in (3.8) and in the matrix $[\boldsymbol{H}_1]$ of (3.10). Hence, with the line parameter uncertainty, $\boldsymbol{H}_1$ is updated into an interval matrix $[\boldsymbol{H}_1]$ in (3.10).

3.3.2. General ISE Framework

The impacts of multiple uncertainties on ISE are decoupled in (3.10), and thus an ISE framework is proposed to deal with different combinations of multiple uncertainties to obtain the bounds of state variables. The model (3.10) is expressed in a compact form:

$$[\boldsymbol{H}]^T \boldsymbol{W} [\boldsymbol{H}] [\boldsymbol{x}] = [\boldsymbol{H}]^T \boldsymbol{W} [\boldsymbol{z}] \tag{3.12}$$

where $[\boldsymbol{H}] = \begin{bmatrix} [\boldsymbol{H}_1] \\ \boldsymbol{H}_2 \end{bmatrix}$, and $[\boldsymbol{H}] \in \mathbb{R}^{m \times n}$; $\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$, and $\boldsymbol{W} \in \mathbb{R}^{m \times m}$; $[\boldsymbol{z}] = \begin{bmatrix} [\boldsymbol{z}_1] \\ [\boldsymbol{z}_2] \end{bmatrix}$, and $[\boldsymbol{z}] \in \mathbb{R}^{m \times 1}$; m is the total number of the measurements and DG prediction intervals, and $m = m_1 + m_2$.

In order to avoid computing the interval multiplication in $[\boldsymbol{H}]^T \boldsymbol{W} [\boldsymbol{H}]$, which is computationally expensive, a dummy interval vector $[\boldsymbol{y}]$ is introduced into (3.12) as suggested in [47]. Then, an equivalent equation is obtained:

$$\begin{bmatrix} [\boldsymbol{H}] & -\boldsymbol{I} \\ 0 & [\boldsymbol{H}]^T \boldsymbol{W} \end{bmatrix} \begin{bmatrix} [\boldsymbol{x}] \\ [\boldsymbol{y}] \end{bmatrix} = \begin{bmatrix} [\boldsymbol{z}] \\ 0 \end{bmatrix} \tag{3.13}$$

where $\boldsymbol{I}$ is an identity matrix, and $\boldsymbol{I} \in \mathbb{R}^{m \times m}$; $[\boldsymbol{y}] \in \mathbb{R}^{m \times 1}$.

The formula (3.13) is further expressed below for brevity:

$$[\boldsymbol{A}] [\boldsymbol{X}] = [\boldsymbol{B}] \tag{3.14}$$

where $[\boldsymbol{A}] = \begin{bmatrix} [\boldsymbol{H}] & -\boldsymbol{I} \\ 0 & [\boldsymbol{H}]^T \boldsymbol{W} \end{bmatrix}$, and $[\boldsymbol{A}] \in \mathbb{R}^{(m+n) \times (m+n)}$; $[\boldsymbol{B}] = \begin{bmatrix} [\boldsymbol{z}] \\ 0 \end{bmatrix}$, and $[\boldsymbol{B}] \in \mathbb{R}^{(m+n) \times 1}$; $[\boldsymbol{X}] = \begin{bmatrix} [\boldsymbol{x}] \\ [\boldsymbol{y}] \end{bmatrix}$, and $[\boldsymbol{X}] \in \mathbb{R}^{(m+n) \times 1}$.

The model (3.14) is straightforward to realize the mutual transformation between differ-

ent combinations of uncertainties. This transformation is shown in Fig. 3.2. A general ISE model considering three types of uncertainties is formulated as (3.14) and named *Model* IV, while *Model* I is a basic ISE model only considering the measurements with noises. When meters are available at all DG units, with the DG measurements merged into $[\boldsymbol{z}_1]$, *Model* IV is simplified into *Model* III, where $[\boldsymbol{A}_1] = \begin{bmatrix} [\boldsymbol{H}_1] & -\boldsymbol{I} \\ 0 & [\boldsymbol{H}_1]^T \boldsymbol{W}_1 \end{bmatrix}$, and $[\boldsymbol{B}_1] = \begin{bmatrix} [\boldsymbol{z}_1] \\ 0 \end{bmatrix}$. In another case, *Model* IV is simplified into *Model* II, when parameter identification techniques or the assumption with accurate line parameters are adopted, *i.e.*, $[\boldsymbol{H}_1]$ is fixed as $\boldsymbol{H}_1$.

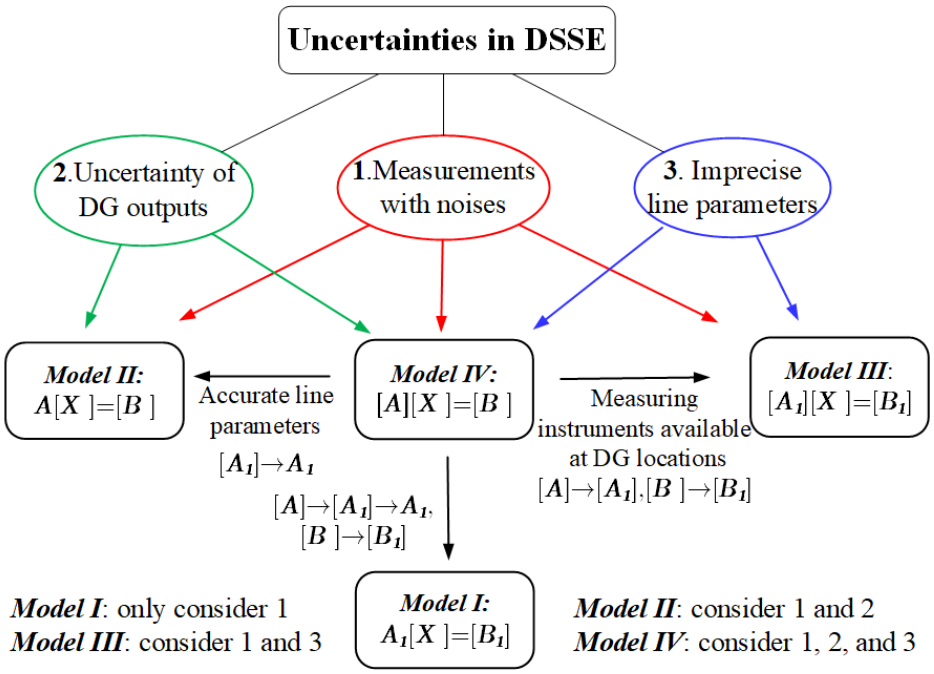

Figure 3.2. Uncertainties of DSSE in interval arithmetic. The model complexity increases from Model I to Model IV.

3.3.3. MKO Algorithm for Solving ISE Models

In this section, an MKO algorithm in interval arithmetic is proposed to efficiently solve the compact ISE model in (3.14). The interval symbols $[\cdot]$ are omitted here for clarity.

For the interval system $\boldsymbol{AX} = \boldsymbol{B}$, its solution set in the interval form is expressed as $\sum(\boldsymbol{A}, \boldsymbol{B}) = \{\widetilde{\boldsymbol{X}} : \ \boldsymbol{a}\widetilde{\boldsymbol{X}} = \boldsymbol{b}, \forall \boldsymbol{a} \in \boldsymbol{A} \ and \ \forall \boldsymbol{b} \in \boldsymbol{B}\}$, where $\widetilde{\boldsymbol{X}}$, $\boldsymbol{a}$, and $\boldsymbol{b}$ are the corresponding deterministic vectors or matrices. Moreover, its interval solution hull $\boldsymbol{X}$ is the interval vector with the smallest radius containing $\sum(\boldsymbol{A}, \boldsymbol{B})$.

A Krawczyk operator is widely employed as a solver for linear interval equations [48]. The core of this operator is to utilize an approximate interval solution $\boldsymbol{X}^{(0)}$ that contains the final solution hull as an initial value, then gradually approach the final solution hull by the following iterative process:

$$\boldsymbol{X}^{(i+1)} = \left(\boldsymbol{CB} + (\boldsymbol{I} - \boldsymbol{CA})\boldsymbol{X}^{(i)}\right) \cap \boldsymbol{X}^{(i)} \tag{3.15}$$

where $\boldsymbol{C}$ is a preconditioning point matrix, and the inverse of $\boldsymbol{C}$ is the midpoint matrix of $\boldsymbol{A}$, expressed as

$$\boldsymbol{C^{-1}} = Mid\,[\boldsymbol{A}] = \begin{bmatrix} Mid\,[\boldsymbol{a_{11}}] & \cdots & Mid\,[\boldsymbol{a_{1,m+n}}] \\ \vdots & \ddots & \vdots \\ Mid\,[\boldsymbol{a_{m+n,1}}] & \cdots & Mid\,[\boldsymbol{a_{m+n,m+n}}] \end{bmatrix}$$

where $Mid\,[\cdot]$ is the median function of an interval variable.

It is deduced that with this matrix $\boldsymbol{C}$ that satisfies $\|\boldsymbol{I} - \boldsymbol{CA}\| < 1$, where $\|\cdot\|$ is any subordinate norm, (3.15) converges according to the fixed point theorem [49]. The iterative process runs until $\|\boldsymbol{X}^{(i+1)} - \boldsymbol{X}^{(i)}\|_{\infty} \leq \varepsilon$, and we set $\varepsilon = 10^{-4}$. Hence, the Krawczyk operator avoids the issue of interval extension when the tolerance of variables in the iterative process is sufficiently small, and the interested readers can refer to the proof in [49].

Next, the above Krawczyk operator is modified to improve algorithmic performance in both accuracy and efficiency. In the proposed MKO algorithm, two computational strategies, Strategy One and Strategy Two, are jointly used to solve (3.14) quickly and accurately:

1) **Strategy One**: Start from an initial solution $\boldsymbol{X}^{(0)}$, which is closer to the final solution

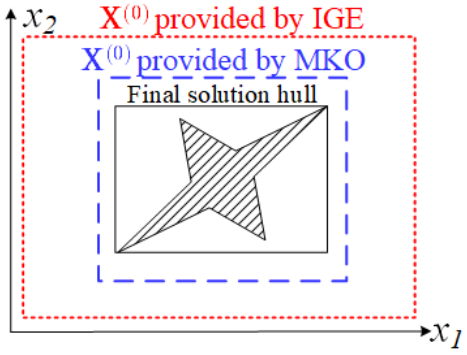

Figure 3.3. Solution hulls of a 2-dimensional interval equation

hull compared to the one that IGE produces.

In the Krawczyk operator, an initial solution $\boldsymbol{X}^{(0)}$, which meets $\widetilde{\boldsymbol{X}} \in \boldsymbol{X}^{(0)}$ and $\sum(\boldsymbol{A}, \boldsymbol{B}) \subseteq \boldsymbol{X}^{(0)}$, is required to start the iterative process. A straightforward approach to getting $\boldsymbol{X}^{(0)}$ is IGE, which is used in [41] as an extension of Gaussian elimination in interval arithmetic. However, IGE largely expands the widths of interval solutions due to its over-conservatism [50]. In addition, IGE is expensive to compute since its forward elimination and back substitution procedure cannot be parallelized.

To address the limitations of IGE, a tighter $\boldsymbol{X}^{(0)}$ is obtained by the following theorem [24].

Theorem 1. If $\boldsymbol{C}$ satisfies $\|\boldsymbol{I} - \boldsymbol{C}\boldsymbol{a}\| = \beta < 1$, $\widetilde{\boldsymbol{X}} = \boldsymbol{a}^{-1}\boldsymbol{b}$ and $\|\cdot\|$ is any subordinate norm, then $\left\|\widetilde{\boldsymbol{X}}\right\| \leq \frac{\|\boldsymbol{C}\boldsymbol{b}\|}{1-\beta}$.

Proof. From $\boldsymbol{a}\widetilde{\boldsymbol{X}} = \boldsymbol{b}$, we have $\widetilde{\boldsymbol{X}} = \boldsymbol{C}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{a})\widetilde{\boldsymbol{X}}$, and hence

$$\left\|\widetilde{\boldsymbol{X}}\right\| \leq \|\boldsymbol{C}\boldsymbol{b}\| + \|\boldsymbol{I} - \boldsymbol{C}\boldsymbol{a}\| \left\|\widetilde{\boldsymbol{X}}\right\| \rightarrow (1-\beta)\left\|\widetilde{\boldsymbol{X}}\right\| \leq \|\boldsymbol{C}\boldsymbol{b}\|$$

where $\beta < 1$ exists for $\boldsymbol{C}$, which is the inverse of the midpoint matrix of $\boldsymbol{A}$.

Since $\|\boldsymbol{C}\boldsymbol{b}\|_\infty \leq \|\boldsymbol{C}\boldsymbol{B}\|_\infty$ and $\|\boldsymbol{I} - \boldsymbol{C}\boldsymbol{a}\|_\infty \leq \|\boldsymbol{I} - \boldsymbol{C}\boldsymbol{A}\|_\infty$, an initial vector $\boldsymbol{X}^{(0)}$ is defined as

$$\boldsymbol{X}^{(0)} = ([-\alpha, \alpha], \ldots, [-\alpha, \alpha])^T \tag{3.16}$$

where $\alpha = \frac{\|\boldsymbol{CB}\|_\infty}{1-\beta}$ and $\beta = \|\boldsymbol{I} - \boldsymbol{CA}\|_\infty$. $\|\cdot\|_\infty$ is the infinite norm of a vector.

The tighter initial solution in (3.16), which is closer to the final solution hull compared to the IGE solution, largely improves the efficiency of the iterative process. A schematic diagram in Fig.3 illustrates the phenomenon, where the accurate solution set of the interval equation constitutes a star-shaped area [48].

2) **Strategy Two**: Modify an enclosure of the difference between the solution at the ith iteration $\boldsymbol{X}^{(i)}$ and an approximate solution, rather than $\boldsymbol{X}^{(i)}$. This modification, combined with Strategy One, guarantees that the proposed algorithm produces the final solution at least as tight as the original Krawczyk operator.

First, the approximate solution $\boldsymbol{X}_s$, which is a point solution located at the center of the solution space, is calculated by multiplying $\boldsymbol{C}$ by the midpoint vector of $\boldsymbol{B}$:

$$\boldsymbol{X}_s = \boldsymbol{C}(Mid\,[\boldsymbol{B}]) \tag{3.17}$$

Let $\boldsymbol{d} = \boldsymbol{X} - \boldsymbol{X}_s$, and get an equivalent interval equation $\boldsymbol{Ad} = \boldsymbol{B} - \boldsymbol{AX}_s$. The initial solution for this equation is $\boldsymbol{d}^{(0)} = \boldsymbol{X}^{(0)} - \boldsymbol{X}_s$, where $\boldsymbol{X}^{(0)}$ is calculated by (3.16).

Applying (3.16) to the enclosure $\boldsymbol{d}^{(i+1)}$ gives a modified residual Krawczyk iterative process

$$\boldsymbol{d}^{(i+1)} = \left(\boldsymbol{C}\,(\boldsymbol{B} - \boldsymbol{AX}_s) + (\boldsymbol{I} - \boldsymbol{CA})\boldsymbol{d}^{(i)}\right) \cap \boldsymbol{d}^{(i)} \tag{3.18}$$

until $\|\boldsymbol{d}^{(i+1)} - \boldsymbol{d}^{(i)}\|_\infty \leq \varepsilon$.

A final solution is computed by $\boldsymbol{X}^{(i+1)} = \boldsymbol{X}_s + \boldsymbol{d}^{(i+1)}$. The final solution produced by (3.18) is at least as tight as the original Krawczyk operator in (3.15), which is verified through the sub-distributive law for interval arithmetic.

Theorem 2. Sub-distributive Law [50]. For interval variables x, y, and z, the law holds

$$x(y+z) \subseteq xy + xz \tag{3.19}$$

Apply Theorem 2 to (3.18), and we have

$$\boldsymbol{X}_s+\boldsymbol{C}\left(\boldsymbol{B}-\boldsymbol{A}\boldsymbol{X}_s\right)+(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{A})\boldsymbol{d}^{(i)} \supseteq \boldsymbol{C}\boldsymbol{B}+(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{A})\left(\boldsymbol{X}_s+\boldsymbol{d}^{(i)}\right)=\boldsymbol{C}\boldsymbol{B}+(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{A})\boldsymbol{X}^{(i)} \tag{3.20}$$

where $\boldsymbol{X}^{(i)}=\boldsymbol{X}_s+\boldsymbol{d}^{(i)}$.

The formula (3.20) implies that $\boldsymbol{X}^{(i+1)}=\boldsymbol{X}_s+\boldsymbol{d}^{(i+1)} \supseteq (\boldsymbol{C}\boldsymbol{B}+(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{A})\boldsymbol{X}^{(i+1)})\cap\boldsymbol{X}^{(i+1)}$, *i.e.*, the final solution hull provided by the original Krawczyk operator contains the one that the MKO algorithm solves. Hence, the proposed algorithm obtains a tighter boundary than the original Krawczyk operator. Moreover, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are thin (an interval with zero radius is defined as a thin interval), the residual $\boldsymbol{B}-\boldsymbol{A}\boldsymbol{X}_s$ is enclosed with fewer rounding errors than $\boldsymbol{B}$ [24]. These features lead to the higher accuracy and less memory space of the proposed algorithm, since many thin interval elements exist due to the highly sparse $\boldsymbol{A}$ and $\boldsymbol{B}$ shown in (3.13), i.e., $0=[0,0]$.

3.4. Simulation Result

The proposed algorithm is tested on unbalanced IEEE 13-bus and 123-bus distribution systems [25]. The 13-bus system is modified by adding two single-phase PV units at buses 675 and 684, and a three-phase wind farm at bus 680, shown in Fig. 3.4. The 123-bus system is modified by adding six DG units, and the installation details of these DG units are listed in Table 3.1. Based on the weather data [51] shown in Fig. 3.5, at 12:00 am, the wind speed interval is [8.886, 9.805] m/s, and the solar radiation interval is [191.246, 286.870] W/m^2. The interval outputs of PVs and WTGs are obtained by the method in [26], and constant power factors are used [15]: [106.72, 149.53] kW for PVs, 0.95 lagging and [84.52, 103.31] kW for WTGs, 0.85 lagging. For these DGs, the upper and lower bounds of reactive power are calculated by $Q_{DG,u}=P_{DG,u}\tan\phi$ and $Q_{DG,l}=P_{DG,l}\tan\phi$, where $P_{DG,u}$ and $P_{DG,l}$ are the upper and lower bounds of power outputs, and ϕ is the power factor angle.

For simulation purposes, the deterministic DG outputs and constant line parameters that fall in the corresponding intervals are used to obtain true values of voltages, currents, and

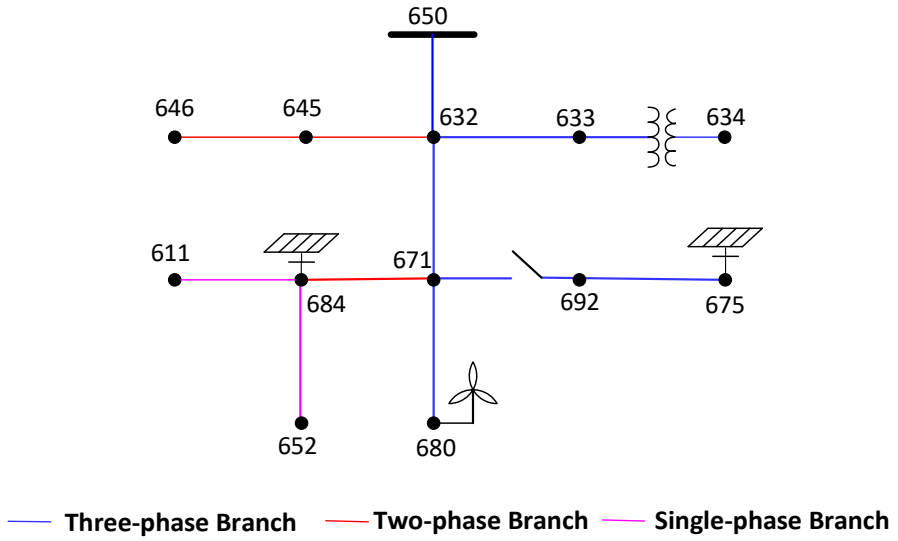

Figure 3.4. One-line diagram of unbalanced IEEE 13-bus system

Table 3.1. DG placements in 123-bus system

#	No. node	Type	Phase	#	No.node	Type	Phase
1	14	PV	C	4	250	WTG	A, B, C
2	61	WTG	A, B, C	5	300	PV	A
3	151	WTG	A, B, C	6	450	PV	A

powers by the power flow program. Measurements are obtained by adding Gaussian noises to these true values. The following conditions are applied to maximum percentage errors of hybrid measurements in Table 3.2: 0.7% for magnitudes and 0.7 crad (10^{-2} rad) for phase angles in PMU data [52], 2% for active and reactive powers of SCADA data, and 10% for active and reactive powers at load nodes as pseudo-measurements [37]. Besides, these hybrid measurements with different sampling rates can be pre-processed for synchronization by the method in [53]. Note that meters and statistical information of power outputs at some DGs are not available, *e.g.*, DGs at buses 680 and 675 in the 13-bus system. Conventional DSSE defines such systems as unobservable, *i.e.*, DSSE fails due to lack of key measurements. In these cases, the proposed algorithm provides the ranges of state variables with the aid of

prediction intervals of DG outputs via *Model* II and *Model* IV.

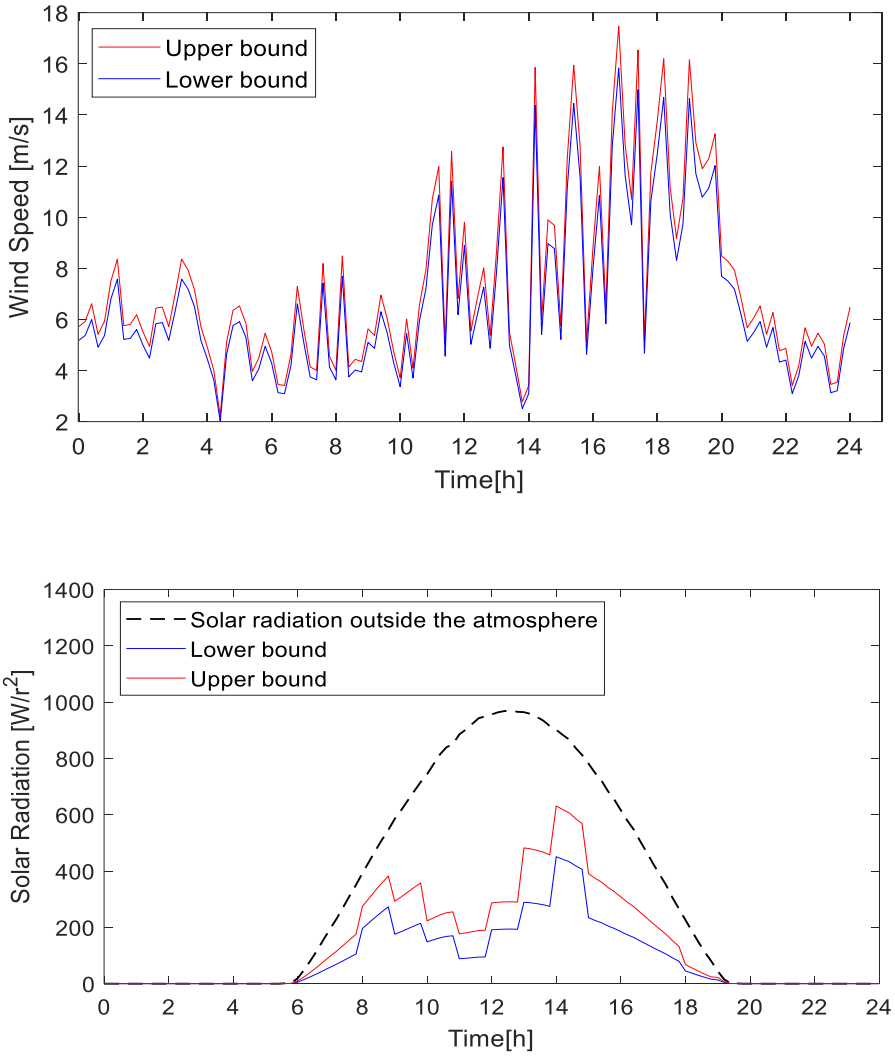

Figure 3.5. DG profiles during one day (a) Wind speed (b) Solar radiation

3.4.1. Result Analysis

The proposed algorithm is tested with *Model* II on the 13-bus system, provided that accurate line parameters are known. In this section, we display the voltage magnitude results of different methods considering that they are more concerned by DSOs in practice, and these voltage magnitudes are calculated from state variables.

Table 3.2. Measurement locations in test systems

Measurement Type	Placement Location	
	13-bus System	123-bus System
SCADA	632-633, 645-646, 684-652	1-7, 9-14, 15-16, 13-52, 18-35, 44-45, 57-60, 76-77, 86-87, 99-100, 110-112
PMU	650, 671	149, 8, 25, 54, 97, 108

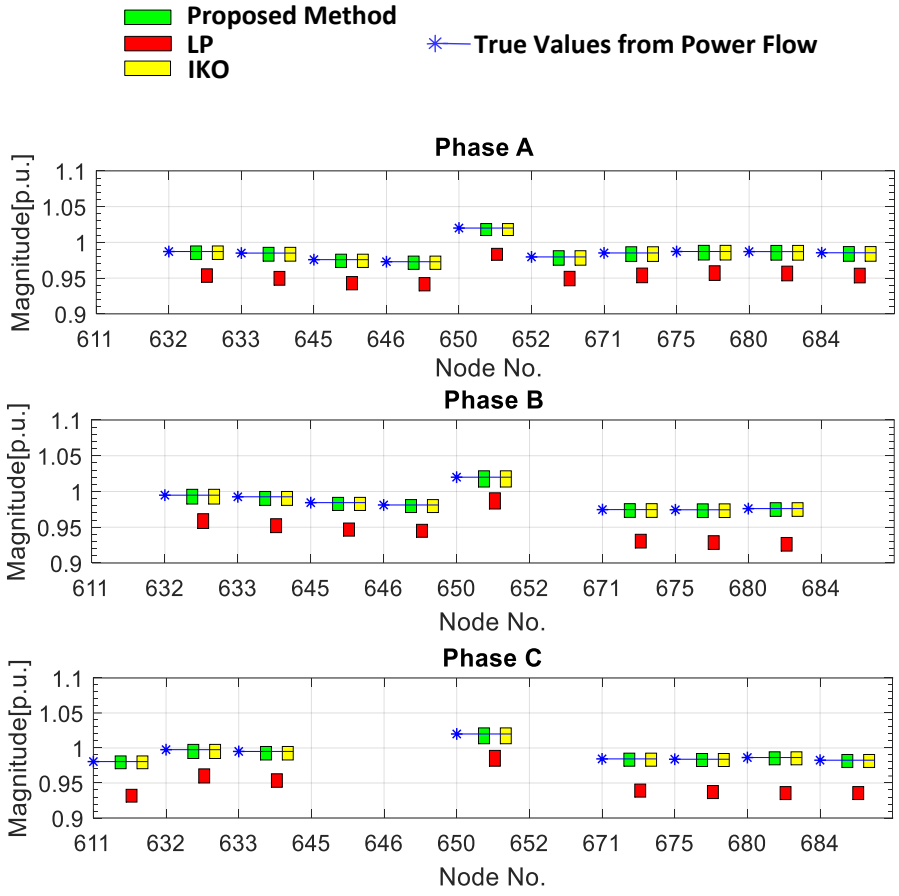

Figure 3.6. Voltage magnitude results of the proposed algorithm

The estimation results of the proposed algorithm, a linear programming (LP)-based method, and the IKO method in [41] are compared with true values of voltages, and the three-phase voltage magnitudes in these methods are intuitively displayed in Fig. 3.6. The proposed method provides the tight bounds that all possible state variables fall in under multiple uncertainties, and these ranges contain the true values of voltages. However, the voltages solved by the LP-based method deviate a lot from the true values at some buses and even exceed the normal operation voltage range ($0.95 \sim 1.05$ p.u.). This is because this method does not consider the various weights of hybrid measurements in the test system.

Moreover, the sums of these voltage widths between the proposed method and the IGE-based method are numerically shown in Table 3.3. The widths of voltages for the proposed algorithm are narrower than the IGE-based method.

3.4.2. Effect of Parameter Uncertainty

The proposed algorithm is applied to two situations in which the line parameters are determined or in certain ranges, via *Model* II and *Model* IV, respectively. To investigate the influence of uncertain line parameters, two cases are designed:

Case 1: Constant line parameter vector, $\boldsymbol{P}_0$.

Case 2: Line parameters change in $[0.95\boldsymbol{P}_0, 1.05\boldsymbol{P}_0]$.

Two indices Q_1 and Q_2 are used to evaluate the precision of interval estimation:

$$Q_1 = \frac{1}{n}\sum_{i=1}^{n}(\overline{x}_i - \underline{x}_i) \tag{3.21}$$

$$Q_2 = \max(|\overline{x}_i - \tilde{x}_i|, |\tilde{x}_i - \underline{x}_i|) \tag{3.22}$$

where $\overline{x}_i$ and $\underline{x}_i$ are the upper and lower bounds of the ith interval variable, and $\tilde{x}_i$ denotes the true value of the ith state variable. Q_1 is the average value of interval widths, and Q_2 is the maximum deviation relative to the true values. The interval estimation with smaller Q_1 and Q_2 has better accuracy.

To verify the effectiveness of the proposed algorithm, in Case 1 and Case 2, the deter-

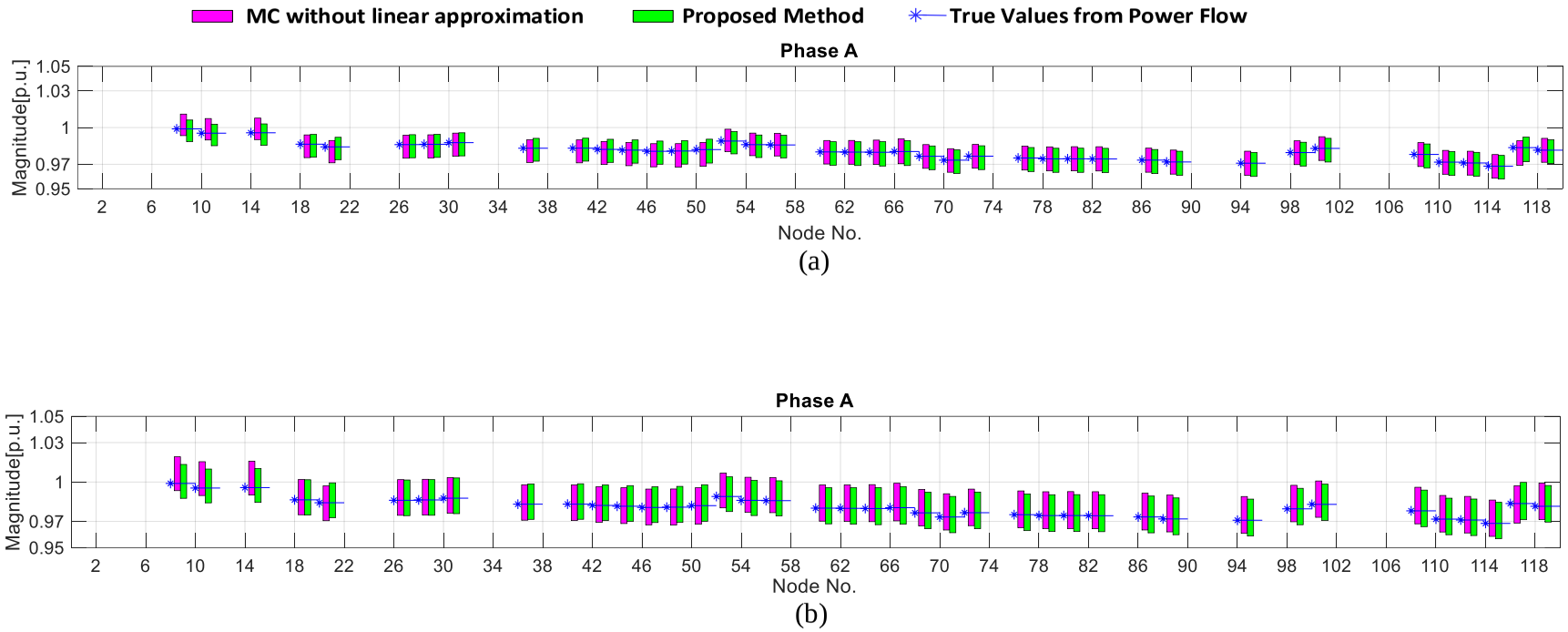

Figure 3.7. Voltage magnitude results in the 123-bus system. (a) Case 1: Without line uncertainty (b) Case 2: With line uncertainty. Under multiple uncertainties, although a linear approximation technique is adopted, the interval results of our algorithm are similar to the ones of MC simulations without this approximation.

Table 3.3. Detailed comparison in voltage results

Sum of width	Proposed Method		IKO	
[V]	Re. Part	Im. Part	Re. Part	Im. Part
Phase A	471.05	260.47	471.08	260.49
Phase B	365.64	214.67	365.65	214.69
Phase C	393.31	215.95	393.34	215.97

ministic DSSE algorithm in [43] without linear approximation runs for 3000 times of MC trials. In these MC trials, random DG outputs in their predication intervals are regarded as the inputs to the DSSE method in Case 1, while the random combinations of deterministic DG outputs and line parameters that fall in the corresponding intervals are used in Case 2. The minimum and maximum values of these state variables in all MC trials are collected and compared with the interval estimation results of the proposed algorithm.

Limited by space, the voltage results at the even-numbered nodes of the 123-bus system are depicted in Fig. 3.7, where the true values of these states are also marked. The interval results of both methods are shown as similar in the two cases. Concretely, the accuracy indices Q_1 and Q_2 in Fig. 3.7(b) are 0.0196 and 0.0163 in the proposed algorithm, while they are 0.0182 and 0.0171 in MC simulations. These results illustrate that under multiple uncertainties, the proposed algorithm obtains the tight boundaries of state variables via a single run, compared with thousands of times of MC runs. Also, the comparison between Fig. 3.7(a) and Fig. 3.7(b) demonstrates that the line parameter uncertainty further intensifies the variability of state variables. With these imprecise line parameters, the proposed method provides the ranges that all possible values of states fall in, as a reliable reference to system operators, shown as Fig. 3.7(b).

3.4.3. Robustness Tests

The robustness of the proposed algorithm is tested on the 123-bus system. Based on Case 2, three cases are established, where the true values of DG outputs lie on the edge of interval predictions, *i.e.*, asymmetric intervals. These cases are designed below.

Table 3.4. Estimation accuracy in robustness tests

Accuracy Indices [p.u.]		Case 3		Case 4		Case 5	
		Q_1	Q_2	Q_1	Q_2	Q_1	Q_2
Phase A	Re. Part	0.020	0.018	0.020	0.017	0.019	0.017
	Im. Part	0.014	0.013	0.014	0.012	0.013	0.014
Phase B	Re. Part	0.017	0.017	0.017	0.015	0.016	0.016
	Im. Part	0.012	0.015	0.012	0.013	0.012	0.012
Phase C	Re. Part	0.020	0.019	0.020	0.016	0.020	0.017
	Im. Part	0.012	0.012	0.012	0.013	0.012	0.012

Case 3: The intervals of PV and WTG outputs are still [106.72, 149.53] kW and [84.52, 103.31] kW. Their true values are fixed at 107 kW and 85 kW in power flow calculation, respectively, to generate the measurements for ISE.

Case 4: True values of PV and WTG outputs are 149 kW and 85 kW, respectively. Other settings are the same as the ones in Case 3.

Case 5: True values of PV and WTG outputs are 149 kW and 103 kW, respectively. Other settings are the same as the ones in Case 3.

Table 3.4 summarizes the accuracy indices Q_1 and Q_2 on three phases in these cases. It is concluded that these estimation results are not significantly affected by the extent of the deviation of DG interval predictions relative to their true values. In contrast, the true values of DG outputs in [37, 41] are always assumed in the center of their intervals, which may not be robust due to the variability of DG outputs.

Table 3.5. Comparison in balanced and unbalanced 123-bus system

Dimensional Analysis	Balanced	Unbalanced
State Variables:$[\boldsymbol{x}]$	238×1	714×1
Measurements and DG Outputs: $[\boldsymbol{z}]$	296×1	888×1
Augmented Variables: $[\boldsymbol{X}]$	534×1	1602×1
Coefficient Matrix: $[\boldsymbol{A}]$	534×534	1602×1602

3.4.4. Computational Efficiency

Numerical experiments are carried out to investigate the computational efficiency of the proposed algorithm. All the tests are performed in MATLAB with the INTLAB toolbox using a 2.5 GHz, 8 GB of RAM, Intel Core i5 computer.

Table 3.6. CPU time in different cases

CPU time [s]	Case 1		Case 2	
	13-bus	123-bus	13-bus	123-bus
Proposed Method	0.0099	0.847	0.012	0.965
MC (3000 trials)	89.52	5655.5	89.78	6209.7
LP	2.560	164.68	-	-
IKO	8.994	795.57	9.957	911.50

The dimensional analysis towards the ISE model in the balanced and unbalanced 123-bus systems is given in Table 3.5. The comparison shows that the unbalanced system leads to a higher-dimensional interval equation. Also, the complexity in unbalanced systems intensifies low efficiency of the existing methods such as [41], which is proposed in balanced systems. Further, in both unbalanced systems, the CPU times of the proposed algorithm, MC simulations, the LP-based method, and the IKO method are summarized in Table 3.6. It shows that the proposed algorithm solves the ISE model in less than one-hundredth amount of time, compared with other methods.

Table 3.7. Computation time in 123-bus system

CPU Time [s]	Time for Initial Solution [s]	Iteration Number	Average Time in Single Iteration [s]
Proposed Method	0.876	2	0.0445
IKO	908.30	8	0.400

Algorithmically, the LP-based method requires solving the equal-scale minimum and maximum problems for n times, where n is the total number of state variables. It should be noted that the LP-based method cannot deal with the uncertain line parameters as in Case

2. Also, as discussed in Section 3.3.3, the initial intervals provided by IGE are much wider than the final solution hull, and the iterative process of the IKO method is time-consuming. The comparative analysis between the proposed algorithm and the IKO method is shown in Table 3.7. It is concluded that the high computational efficiency of the proposed algorithm firmly holds in the 123-bus unbalanced system.

3.4.5. Results of Model I and Model III

We test the proposed algorithm via *Model* I and *Model* III on the 123-bus system, *i.e.*, DG outputs can be obtained by meters or pseudo-measurements. Two cases are designed below.

Table 3.8. Estimation accuracy in robustness tests

Accuracy Indices [p.u.]		Case 6		Case 7	
		Q_1	Q_2	Q_1	Q_2
Phase A	Re. Part	0.0135	0.0107	0.0136	0.0124
	Im. Part	0.0078	0.0090	0.0080	0.0122
Phase B	Re. Part	0.0116	0.0106	0.0120	0.0121
	Im. Part	0.0068	0.0093	0.0070	0.0117
Phase C	Re. Part	0.0138	0.0116	0.0139	0.0133
	Im. Part	0.0080	0.0084	0.0081	0.0118

Case 6: Constant line parameter vector, $\boldsymbol{P}_0$. Also, the DG outputs are modeled as pseudo-measurements following known Gaussian distributions, and their maximum errors are 10% of active and reactive powers [37].

Case 7: Line parameters change in $[0.95\boldsymbol{P}_0, 1.05\boldsymbol{P}_0]$. Other settings are the same as the ones in Case 6.

The indices Q_1 and Q_2 in both cases are shown in Table 3.8. Besides, we list the computation time of these cases in Table 3.9. The comparison between these two cases illustrates that less uncertainty leads to tighter interval results and higher computational efficiency.

Table 3.9. CPU time in 123-bus system

CPU time [s]	Case 6	Case 7
Proposed Method	0.633	0.720
MC (3000 trials)	5178.2	5750.3
LP	135.57	-
IKO	300.44	316.36

3.5. Conclusion

This chapter proposes a novel and fast ISE algorithm with multiple uncertainties in unbalanced distribution systems. We establish a general ISE framework that simultaneously considers imprecise line parameters, measurements with noises, and uncertain DG outputs. An MKO algorithm is proposed to solve these ISE models and obtain the upper and lower bounds of state variables for better monitoring systems under the coordinated impacts of these multiple uncertainties. The proposed algorithm is tested on unbalanced 13-bus and 123-bus distribution systems. In contrast to MC simulations and the existing alternatives, the proposed algorithm encloses tighter boundaries of state variables in a faster manner. This dissertation will focus on the applications of the proposed DSSE algorithm to bad data or cyberattack detection in active distribution systems later.

Chapter 4

Graph-based Faulted Line Identification Using PMU Data in Distribution Networks

4.1. Introduction

Faults are regarded as an important type of reliability events, which may immensely affect normal system operation. In the past decade, 22.2 million customers in California experienced about 6000 outage hours resulted from sustained faults [54]. Extensive studies on fault location are developed in meshed transmission systems (*e.g.*, [55–57]). However, distribution systems are largely different from transmission systems due to their radial topology and limited real-time meters. Consequently, these existing fault location methods in transmission systems cannot be applied to distribution systems. On the other hand, increasingly pervasive installation of distribution-level phasor measurement units (PMUs), *i.e.*, micro-PMUs, improves the system monitoring significantly. Compared with conventional meters, PMUs provide more accurate measurements of voltage and current phasors at a high resolution. Several emerging applications of PMUs include distribution system state estimation (DSSE), fault detection, and faulted line location [58]. For instance, the authors of [59] applied data-driven techniques with PMU data to detect the presence of a fault in distribution systems; however, these detection algorithms cannot identify the location of the faulted line.

Quick and accurate location of faults in distribution systems helps the utilities to clear the faults and accelerate the system restoration; however, this is a challenging task as the mal-trip or fail-to-trip of the protection devices may lead to inaccurate location of the fault. The chances of such unfavorable events grow with the bidirectional power flow and the increasing penetration of distributed generators (DGs). The authors of [60] pointed out that conventional protection devices such as fault indicators may fail to clear a fault under the

bidirectional current flow conditions. Also, the overcurrent-based protection devices may not be able to locate high-impedance faults in distribution systems since it is difficult to identify the small fault currents [61].

The existing fault location methods are classified into three main types: 1) traveling wave-based, 2) training-based, and 3) impedance-based. The traveling wave-based algorithms (*e.g.*, [62, 63]) locate a fault by utilizing the arrival time of the original and reflected waves generated by the fault. These methods require high-speed communication and high sampling rate measurements that may not be prevalent in distribution networks. Besides, the training-based fault location methods, such as artificial neural networks (ANN) [64] and support vector machine (SVM) [65], require a large number of high-quality measurements as training datasets and thus suffer from a high computational burden in a training process.

Recent efforts are devoted to proposing impedance-based location methods in distribution systems [66–70]. For instance, the authors of [67] proposed the fault location methods focusing on single-phase to ground faults. However, [67] can only localize a faulted area, rather than yielding an exact faulted line. Emerging works are applying PMU measurements to fault location by constructing generalized impedance-based location methods [60, 68–70]. We conclude that the search strategy in these works is to select each bus or each line as the candidate fault source and then calculate the values of the self-defined objective function for each candidate. Then, the fault location is determined by minimizing or maximizing these function values. Specifically, the authors of [68] used a state estimation technique with sufficient PMU data to identify the fault at a distribution line. However, these PMUs are assumed available at each bus, which is impractical due to economic and technical restrictions in distribution systems. Further efforts are put into fault location with a fewer number of PMUs, such as [69, 70]. The approach in [70] requires equipping with PMUs at all DGs. This arrangement may not be practical due to a limited number of available PMUs. Also, this method does not consider high-impedance faults, which are regarded as an untraceable fault type in system operation. The authors of [70] defined a fault as a generalized reliability event and presented an optimization model to locate the event bus by PMU data and pseudo-

measurements recorded at load/DG buses. Also, due to the presence of local minimums in the objective function, the method needs to compare all local minimums to obtain a global one. Further, the global minimum of this function points to the final faulted bus. However, this process may increase computational complexity due to this traversal search strategy.

Various influence factors, such as fault types, fault impedances, DG penetration, and measurement errors, may degrade the effectiveness of the existing fault location methods. To mitigate these impacts, this chapter proposes a graph-based fault location method using advanced DSSE techniques with PMU data in unbalanced distribution systems. The core of the proposed method is to determine the faulted line by comparing the weighted measurement residuals (WMRs) of DSSE in different topologies/graphs. This idea, as a typical application of state estimation, is proposed in [57, 68], where the power systems are observable by an adequate number of PMUs. In comparison, the proposed method only requires a limited number of PMUs in distribution networks for such an application. Specifically, we present an efficient distributed DSSE algorithm to restrict the search region in a shorter feeder between two adjacent PMUs. Further, in the shorter feeder, the fault source is identified at the exact line by applying the DSSE methods to a hierarchical structure. The hierarchical structure built on the graph theory is presented in Section 4.3 and captures the graphs, subgraphs, and paths in the network.

4.2. Theoretical Basis

This section describes the theoretical basis for applying DSSE to fault location. We introduce a classical state estimator and further extend it to an advanced DSSE method using measurements from a limited number of PMUs.

Developed from this classical estimator, the branch current based DSSE method integrating PMU data is regarded as a computationally efficient method due to its constant and sparse-structured Jacobian matrix, as reviewed in [14]. Therefore, this section uses the branch current based DSSE method proposed in [43] for the fault location task. Also, the voltage at the slack node and branch currents are chosen as state variables, and we express

these states in a three-phase network as

$$\boldsymbol{x} = \left[v^a_{slack,r}, v^a_{slack,x}, \ldots, v^c_{slack,x}, i^a_{1r}, i^a_{1x}, \cdots, i^c_{Lx}\right] \tag{4.1}$$

where $v^{\varphi}_{slack,r}$ and $v^{\varphi}_{slack,x}$ denote the real and imaginary parts of the φ -phase slack node's voltage, and $\varphi \in \{a, b, c\}$; i^{φ}_{lr} and i^{φ}_{lx} denote the real and imaginary parts of the branch current at branch l, $l = 1, \ldots, L$, and L is the number of branches. In the following, the phase index φ is suppressed for simplicity.

Here, the measurement vector includes the PMUs' recorded magnitudes and phase angles of voltages and currents as well as power measurements from pseudo-measurements, and the latter provides the historical or forecasting data with a low-level accuracy of power consumption/production at loads/DGs [70]. We list the measurement functions for voltages, currents, and powers in this estimator as follows:

$$\begin{cases} h_{V_{kr}}(x) = z_j = z_{V_{kr}} \ , & k \in \psi_V \\ h_{V_{kx}}(x) = z_j = z_{V_{kx}} \ , & k \in \psi_V \end{cases} \tag{4.2}$$

$$\begin{cases} h_{I_{pr}}(x) = z_j = z_{I_{pr}}, & p \in \psi_I \\ h_{I_{px}}(x) = z_j = z_{I_{px}}, & p \in \psi_I \end{cases} \tag{4.3}$$

$$\begin{cases} h_{P_k}(x) = z_j = z_{P_k}, & k \in \psi_S \\ h_{Q_k}(x) = z_j = z_{Q_k}, & k \in \psi_S \end{cases} \tag{4.4}$$

where z_j denotes measurement j and is expressed as 1) the real and imaginary parts of voltages, $z_{V_{kr}}$ and $z_{V_{kx}}$, 2) the real and imaginary parts of currents, $z_{I_{pr}}$ and $z_{I_{px}}$, or 3) the real and reactive powers, z_{P_k} and z_{Q_k}; $h_{V_{kr}}(x)$, $h_{V_{kx}}(x)$, $h_{I_{pr}}(x)$, $h_{I_{px}}(x)$, $h_{P_k}(x)$, and $h_{Q_k}(x)$ denote the corresponding measurement functions; ψ_V and ψ_I are the sets of nodes and branches with voltage/current measurements from limited PMUs installed in the distribution

system, and ψ_S is the set of load/DG nodes; k and p are the indices of nodes and branches, respectively.

For $k \in \psi_S$, the pseudo-measurements at node k are further converted into equivalent currents in (4.4) by

$$z_{I_{kr}}^{eq} + j z_{I_{kx}}^{eq} = \left[\frac{z_{P_k} + j z_{Q_k}}{V_k} \right]^* \tag{4.5}$$

where $z_{I_{kr}}^{eq}$ and $z_{I_{kx}}^{eq}$ are the real and imaginary parts of the equivalent injection current at node k; V_k as the voltage phasor at the node is updated during the DSSE procedure, since the PMU measurement of V_k is not available at each node; $[\cdot]^*$ denotes the complex conjugate.

By the processing in (4.5), the Jacobian matrix is independent of $\boldsymbol{x}^t$, *i.e.*, $\boldsymbol{H}(\boldsymbol{x}^t) = \boldsymbol{H}$ [43]. The measurement functions of (4.2), (4.3), (4.5), and Jacobian elements of $\boldsymbol{H}$ are listed below.

1) Voltages

The voltage measurement function of the PMU at node $k \in \psi_V$ is expressed as:

$$h_{V_{kr}} + j h_{V_{kx}} = v_{slack} - \sum\nolimits_{p \in \Im_k} (R_p + j X_p) i_p \tag{4.6}$$

where $\Im_k$ denotes a set of line segments from the slack node to node k, and $p \in \Im_k$. R_p and X_p denote the 3×3 resistance and reactance matrices of branch p. Also, the complex variables $i_p = i_{pr} + j i_{px}$ and $v_{slack} = v_{slack,\,r} + j v_{slack,x}$ are the voltage phasor at the slack node and the current phasor at branch p. The Jacobian elements of (4.6) for $p \in \Im_k$ are constant, expressed as:

$$\frac{\partial h_{V_{kr}}}{\partial v_{slack,r}} = 1 \quad \frac{\partial h_{V_{kr}}}{\partial v_{slack,x}} = 0 \quad \frac{\partial h_{V_{kx}}}{\partial v_{slack,r}} = 0 \quad \frac{\partial h_{V_{kx}}}{\partial v_{slack,x}} = 1$$

$$\frac{\partial h_{V_{kr}}}{\partial i_{pr}} = -R_p \quad \frac{\partial h_{V_{kr}}}{\partial i_{px}} = X_p \quad \frac{\partial h_{V_{kx}}}{\partial i_{pr}} = -X_p \quad \frac{\partial h_{V_{kx}}}{\partial i_{px}} = -R_p$$

2) Currents

The current measurement function of the PMU at branch $p \in \psi_I$ is shown as

$$h_{I_{pr}} + j h_{I_{px}} = i_{pr} + j i_{px} \tag{4.7}$$

where i_{pr} and i_{px} denote the real and imaginary parts of the current states, and thus the Jacobian elements are present at:

$$\frac{\partial h_{I_{pr}}}{\partial i_{lr}} = \begin{cases} 1, & when\ p = l \\ 0, & elsewhere \end{cases} \qquad \frac{\partial h_{I_{pr}}}{\partial i_{lx}} = 0$$

$$\frac{\partial h_{I_{px}}}{\partial i_{lr}} = 0 \qquad \frac{\partial h_{I_{px}}}{\partial i_{lx}} = \begin{cases} 1, & when\ p = l \\ 0, & elsewhere \end{cases}$$

where l denotes the branch index, and $l = 1, \ldots, L$.

3) Power Injections

The power measurements presenting at node $k \in \psi_S$ are converted into equivalent current injection by (4.5), and then the current measurement function is expressed as

$$h^{eq}_{I_{kr}} + j h^{eq}_{I_{kx}} = \sum\nolimits_{l \in \Lambda'_k} (i_{lr} + j i_{lx}) - \sum\nolimits_{l \in \Lambda_k} (i_{lr} + j i_{lx}) \tag{4.8}$$

where i_{lr} and i_{lx} as state variables denote the real and imaginary inflow and outflowing currents at node k , and Λ'_k and Λ_k denote the set of branches with the inflow and outflow currents at node k, respectively. The Jacobian elements of (4.8) are calculated by

$$\frac{\partial h^{eq}_{I_{kr}}}{\partial i_{lr}} = \begin{cases} 1, & when\ l \in \Lambda'_k \\ -1, & when\ l \in \Lambda_k \\ 0, & elsewhere \end{cases} \qquad \frac{\partial h^{eq}_{I_{kr}}}{\partial i_{lx}} = 0$$

$$\frac{\partial h^{eq}_{I_{kx}}}{\partial i_{lr}} = 0 \qquad \frac{\partial h^{eq}_{I_{kx}}}{\partial i_{lx}} = \begin{cases} 1, & when\ l \in \Lambda'_k \\ -1, & when\ l \in \Lambda_k \\ 0, & elsewhere \end{cases}$$

The complete DSSE procedure can be found in [43], and the next section gives the details of the modified DSSE method for faulted line identification.

4.3. Graph-Based Faulted Line Identification Method

This section proposes a graph-based fault location method that leverages the above DSSE method to narrow down the searching area and then locate the faulted line.

We consider a distribution network as a graph $G(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the sets of vertices (nodes) and the edges (branches), respectively. A PMU is installed at the substation, and other PMUs are installed at a limited number of nodes along the feeder. Each of these PMUs measures the nodal voltage and the currents on the branches connected to that node [43]. We define a subgraph $G_K(\mathcal{V}_K, \mathcal{E}_K)$ as the subset of $G(\mathcal{V}, \mathcal{E})$ that connects two adjacent PMUs, PMUs K and $K+1$, where $K = 1, \ldots, M-1$ and $M \geq 2$. Here, M is the number of PMUs installed in the network. Fig.4.1 shows the schematic diagram of the subgraphs. In the figure, G_1 is a subgraph that includes the branches and nodes between PMUs 1 and 2, while G_2 contains those between PMUs 2 and 3.

We briefly describe the proposed fault location method:

1. Step One: Using a distributed DSSE algorithm, the searching area for the fault is restricted to the feeder between two adjacent PMUs, i.e., a certain subgraph.

2. Step Two: The location of the fault is further identified as the faulted line.

4.3.1. Step One: Identifying the Faulted Subgraph

This step proposes an efficient DSSE algorithm in G_K to identify the subgraph that contains the faulted line, *i.e.*, the faulted subgraph. By the graph partition and the subsequent

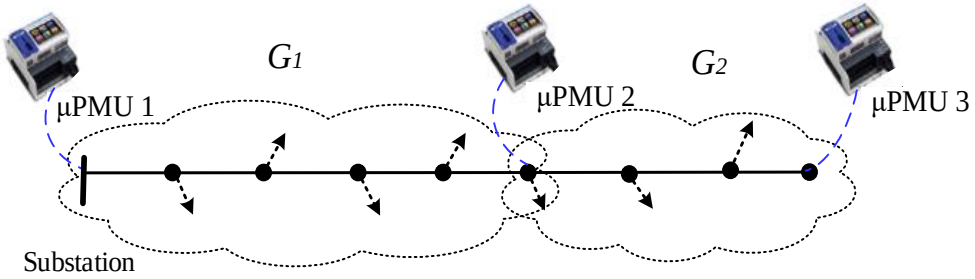

Figure 4.1. A sample of radial distribution networks with three PMUs. The dotted lines with arrows at nodes denote the laterals (if present)

network equivalence, the DSSE method leverages the PMU and pseudo-measurement data in G_K and runs in parallel for these subgraphs with shorter feeders, *i.e.*, distributed DSSE [71].

1) Network Equivalence

In each subgraph, we suppose that the vertex of G_K acts as the root node of this subgraph. The lateral connected to PMU K+1 is also included in G_K, while the lateral at the root node of G_K is included in the last subgraph, *i.e.*, G_{K-1}. Fig. 4.2 shows the schematic diagram of G_K in this design. At node $k \in \mathcal{V}_K$, one type of the following measurements exist and $\mathcal{V}_K = \mathcal{V}_{K1} + \mathcal{V}_{K2} + \mathcal{V}_{K3}$: 1) Only PMU data (*i.e.*, the measurements of PMU K at the root node), and let $k \in \mathcal{V}_{K1}$; 2) PMU data (*i.e.*, measurements of PMU $K+1$) and pseudo-measurements, and $k \in \mathcal{V}_{K2}$; 3) Only pseudo-measurements, and $k \in \mathcal{V}_{K3}$.

To reduce the impact of the graph partition on the power flow in the original network shown as Fig. 4.2(a), we do the equivalent current calculation at node $k \in \mathcal{V}_{K2}$. Specifically, the real and imaginary parts of the injected current at node k in G_K, $z_{I_{kr}}^{sub}$ and $z_{I_{kx}}^{sub}$, are equivalently calculated by

$$z_{I_{kr}}^{sub} + j z_{I_{kx}}^{sub} = (z_{I_{kr}}^{eq} + j z_{I_{kx}}^{eq}) + (z_{I_{kr}} + j z_{I_{kx}}) \tag{4.9}$$

where $z_{I_{kr}}^{eq}$ and $z_{I_{kx}}^{eq}$ denote the real and imaginary parts of the injection currents of pseudo-

measurements obtained by (4.5), and $z_{I_{kr}}$ and $z_{I_{kx}}$ are the real and imaginary parts of the current to the downstream network measured by PMU $K+1$, shown in Fig. 4.2(b). For simplicity, (4.9) does not show the measurement noises.

At node $k \in \mathcal{V}_{K1} \cup \mathcal{V}_{K3}$, the measurement functions (4.5) – (4.8) hold.

2) Identification Metric

We use the WMR in DSSE as the metric to determine the faulted subgraph. In normal operation, assume measurement noises follow Gaussian distribution, WMRs obey a Chi-square distribution with at most $m-n$ degrees of freedom [19]. With a limited number of PMUs installed in distribution systems, the degree of freedom is low and equal to the number of these PMUs. Therefore, the values of a WMR in each subgraph fluctuate within a limited range under the impact of measurement noises, when no faults occur.

On the other hand, according to [57], a fault introduces an additional unknown fault current I_F injected to the ground or other phases, while the DSSE equations are built on the precondition $I_F = 0$. If a fault occurs in G_K, the presence of the fault violates the state estimation relationship and leads to a high WMR in the faulted subgraph; The DSSE in normal subgraphs have low WMRs even under the impact of measurement noises [68]. Hence, the faulted subgraph is determined by selecting the maximum of WMRs:

$$K^* = \arg\max_{K} \ J_K \qquad K = 1, \ldots, M-1 \tag{4.10}$$

where J_K denotes the WMR in subgraph K calculated by (2.2).

Based on the state estimator presented in (2.1) – (2.5), we conclude the procedure for identifying the faulted subgraph below:

1) Considering $G_K(\mathcal{V}_K, \mathcal{E}_K)$, the measurements in each subgraph are collected to form the Jacobian and weight matrices, *i.e.*, $\boldsymbol{H}_K$ and $\boldsymbol{W}_K$.

2) For $K = 1, \ldots, M-1$, the DSSE process in subgraph K is shown in the following steps:

a. *Initialization–forward-backward sweep* [72]: Set the initial voltage at each node as the

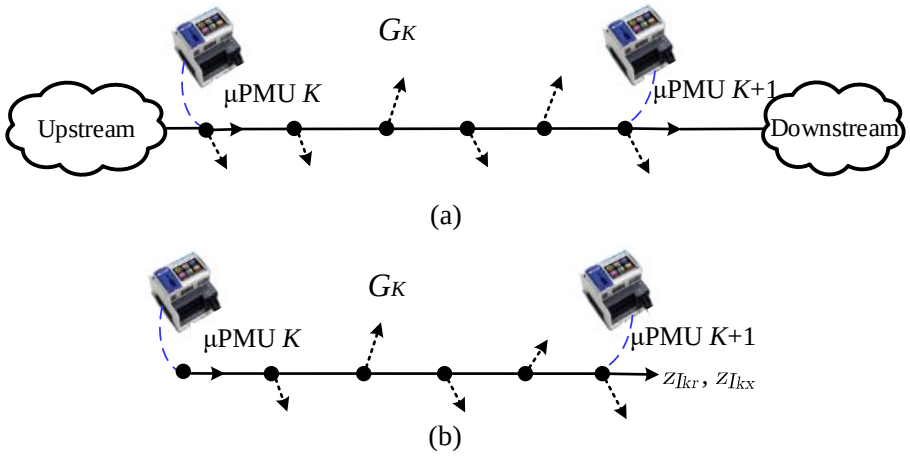

Figure 4.2. Subgraph K (a) Embedded in the whole feeder (b) Decoupled with other subgraphs. The reference directions of branch currents measured by two PMUs are shown.

voltage of the root node V_{root}, and calculate the current injections of power measurements by

$$z_{I_{kr}}^{eq} + jz_{I_{kx}}^{eq} = \left[\frac{z_{P_k} + jz_{Q_k}}{V_{root}}\right]^* \tag{4.11}$$

$$z_{I_{kr}}^{sub} + jz_{I_{kx}}^{sub} = \left[\frac{z_{P_k} + jz_{Q_k}}{V_{root}}\right]^* + (z_{I_{kr}} + jz_{I_{kx}}) \tag{4.12}$$

where (4.11) is used for $k \in \mathcal{V}_{K3}$, and (4.12) holds at $k \in \mathcal{V}_{K2}$; V_{root} comes from the voltage measurement from the PMU at the root node.

Then, obtain the initial branch currents $\boldsymbol{x}^0$ by a backward sweep method. Use $\boldsymbol{x}^0$ and V_{root} to calculate initial nodal voltages V_k by a forward sweep method.

b. Obtain $\boldsymbol{h}(\boldsymbol{x})$ using (4.6) – (4.8), and calculate $\Delta\boldsymbol{x}^t$ and update the new state variables by $\boldsymbol{x}^{t+1} = \boldsymbol{x}^t + \Delta\boldsymbol{x}^t$. Calculate the latest voltages V_k based on the new states by the forward sweep.

c. If $\Delta\boldsymbol{x}^t$ is less than a pre-set tolerance or t reaches the maximum iteration number, yield J_K using (4.1) as the WMR of G_K; otherwise, use the latest V_k to calculate injection currents by (4.5) or (4.9), then go to step b.

3) Procure the faulted subgraph K^* using (4.10).

Finding the faulted subgraph at this stage reduces the computation burden associated with locating the faulted line in subgraphs without faults.

4.3.2. Step Two: Locating the Faulted Line

Once we obtain the faulted subgraph G_{K^*} by Step One, a similar WMR metric based on the DSSE technique is developed to identify the exact line that a fault lies at. Also, we use the following definitions to present Step Two in G_{K^*}.

Definition 1 (Paths in a subgraph). A path in a subgraph is a set of interconnected edges that begins with the root node of the subgraph. A path that a fault is located in is a corrupted path.

Definition 2 (Adjacent Paths and Boundary Edge). Two paths denoted by $\mathcal{P}_{s-1}$ and $\mathcal{P}_s$, $s = 1,\ldots, S$, are defined as adjacent paths, and if $\mathcal{P}_{s-1} \subseteq \mathcal{P}_s$ and $\mathcal{P}_s = \mathcal{P}_{s-1} \cup \{\varepsilon\}$, where ε is the boundary edge that connects two vertices ν and μ, $\nu \in \mathcal{P}_{s-1}$ and $\mu \in \mathcal{P}_s$.

All paths in a faulted subgraph share a starting vertex (root), and different paths are formed by radially expanding the topology of G_{K^*}. The paths $\mathcal{P}_s$ in each subgraph are sorted by their depth. The shortest path in the subgraph K^* only includes one edge, while the longest path is the whole subgraph G_{K^*}.

In theory, the WMRs in two neighboring paths without fault current injections should be close to each other; the WMR of DSSE in a path is low if there is no fault in the path, while WMR is significantly high once faults occur in the path. Therefore, we convert the fault location problem into a problem of searching for the corrupted path that includes the fault, and this corrupted path is characterized by abnormally high WMR in DSSE. To find the corrupted path, DSSE runs for each path in G_{K^*}, and the sending-end branch currents in the corresponding path are chosen as state variables shown in Fig. 4.3. We apply the

DSSE algorithm in Step One for path s and calculate the WMR by

$$WMR_s = [\boldsymbol{z}_s - \boldsymbol{h}_s(\boldsymbol{x})]^T \boldsymbol{W}_s [\boldsymbol{z}_s - \boldsymbol{h}_s(\boldsymbol{x})] \tag{4.13}$$

where $\boldsymbol{z}_s$ and $\boldsymbol{h}_s(\boldsymbol{x})$ denote the measurement vector and measurement functions for path s, and $s = 1,\ldots, S$; $\boldsymbol{W}_s$ denotes the diagonal weight matrix for this path.

According to Definition 2, if a fault occurs at the boundary edge ε, we have

$$WMR_s \gg WMR_{s-1} \tag{4.14}$$

where WMR_s and WMR_{s-1} are the WMRs in paths $\mathcal{P}_{s-1}$ and $\mathcal{P}_s$ calculated by (4.1), respectively. To find the faulted boundary edge, set the user-defined identification thresholds to quantize the relationship in (4.14):

$$\begin{cases} WMR_{s-1} \leq \epsilon \\ WMR_s > \epsilon \end{cases} \tag{4.15}$$

where ϵ denotes the identification threshold for evaluating the abnormally high WMR.

We consider that various fault conditions may occur, and they are unpredictable for system operators. As a result, although a proper identification threshold is beneficial for fault location, the specific value of this threshold is difficult to determine when the fault location, fault impedance, and fault type are unknown. Similar to [55] and [73], the identification threshold ϵ could be properly selected by using historical or simulation data of different faults to enforce (4.15). Also, it is efficient to run the efficient DSSE method for verifying the relationship in (4.15), since the Jacobian matrix $\boldsymbol{H}_s$ for path s is sparse and independent of state variables.

Illustrative Example: To clarify the procedure of the proposed method, let us consider a 5-node subgraph shown in Fig. 4.3, where $l = \{1, 2, 3, 4, 5\}$, and a lateral is connected to node 3. There are three paths: 1-2, 1-3, and 1-5 in this subgraph. Path 1-3 is the set of

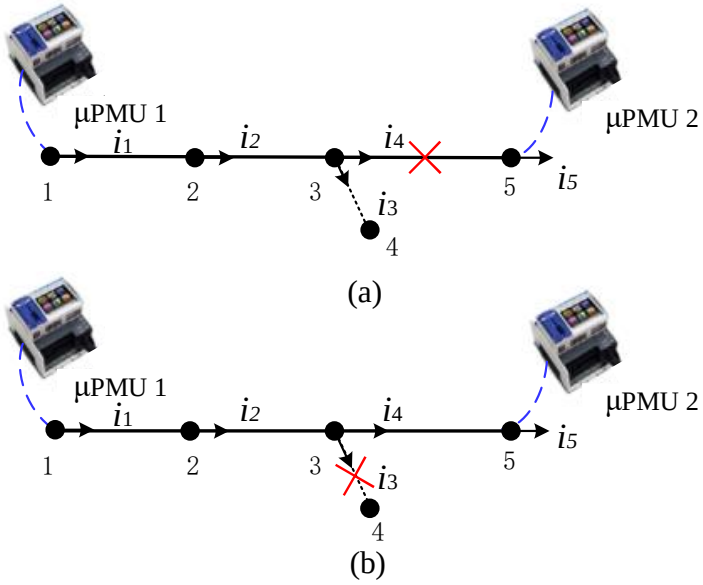

Figure 4.3. Sample network of a 5-node subgraph, and a lateral is connected to node 3 shown as a dotted line. (a) A fault occurs at branch 3-5. (b) A fault occurs at branch 3-4.

Table 4.1. State variables and measurements in a 5-node subgraph

s	$\mathcal{P}_s$	State Variables x	Measurements z_s
1	1-2	i_1, i_2	$z_{V_1}, z_{i_1}, z_{i_2}^{eq}$
2	1-3	i_1, i_2, i_3, i_4	$z_{V_1}, z_{i_1}, z_{i_2}^{eq}, z_{i_3}^{eq}, z_{i_4}^{eq}$
3	1-5	i_1, i_2, i_3, i_4, i_5	$z_{V_1}, z_{I_1}, z_{i_2}, z_{i_2}^{eq}, z_{i_3}^{eq}, z_{i_4}^{eq}$

branches from node 1 to node 3, including the lateral 3-4. Table 4.1 lists the state variables and measurements used in these paths, and we show $\boldsymbol{H}_s$ for path $s = 1$, 2, and 3, which are marked by three block matrices, respectively. In Fig. 4.3(a), the fault is located at the boundary edge between paths 2 and 3 by the proposed method, *i.e.*, branch 3-5; the source of the fault in Fig. 4.3(b) is located at the boundary edges between paths 1 and 2, *i.e.*, branches 2-3 and 3-4, and there are two boundary edges due to the existence of a lateral. In the case that a fault occurs at a lateral as Fig. 4.3(b), the fault source could be located at the lateral or the only upstream branch connected to it, even when there is no PMU installed

at the lateral. Granting complete observability on laterals may not be of economic interest, and therefore, in many practical cases, it is sufficient to identify the faulted laterals [70].

$$
\boldsymbol{H}_s = \begin{array}{c}
\begin{array}{cccccccccc} i_{1r} & i_{1x} & i_{2r} & i_{2x} & i_{3r} & i_{3x} & i_{4r} & i_{4x} & i_{5r} & i_{5x} \end{array} \\
\begin{bmatrix}
1 & & & & & & & & & \\
 & 1 & & & & & & & & \\
1 & & -1 & & & & & & & \\
 & 1 & & -1 & & & & & & \\
 & & 1 & & -1 & & -1 & & & \\
 & & & 1 & & -1 & & -1 & & \\
 & & & & & & 1 & & & \\
 & & & & & & & 1 & & \\
 & & & & & & 1 & & -1 & \\
 & & & & & & & 1 & & -1 \\
 & & & & & & & & 1 & \\
 & & & & & & & & & 1
\end{bmatrix}
\end{array}
\begin{array}{l}
z_{I_{1r}} \\ z_{I_{1x}} \\ z^{eq}_{I_{2r}} \\ z^{eq}_{I_{2x}} \\ z^{eq}_{I_{3r}} \\ z^{eq}_{I_{3x}} \\ z^{eq}_{I_{4r}} \\ z^{eq}_{I_{4x}} \\ z^{eq}_{I_{5r}} \\ z^{eq}_{I_{5x}} \\ z_{I_{5r}} \\ z_{I_{5x}}
\end{array}
$$

We summarize the proposed algorithm in the pseudo-code. Owing to the hierarchical graph-subgraph-path structure in the proposed method and the adaptation of the advanced DSSE method, the search along the faulted subgraph is more efficient, compared with the ones considering the whole distribution feeder.

4.3.3. PMU Placement

The identification accuracy of the proposed method relies on the number and locations of PMUs. As the number of PMUs increases, the size of a partitioned graph in the distribution network will be smaller. This would increase the measurement redundancy defined as the ratio of the number of measurements to that of states. The minimum number of PMUs to be installed in a distribution feeder is two. Moreover, with more PMUs installed, the location performance and computational efficiency of the proposed method can be improved.

Algorithm 1 Graph-based Faulted Line Identification Algorithm

Input: System model and measurement data.

While: The presence of a fault is detected, and its location is unknown.

Step One:

Run the distributed DSSE algorithm for G_K in parallel, and obtain G_{K^*} by (4.10).

Step Two:

Let $s = 1$, and obtain $\boldsymbol{z}_s$ and $h_s(x)$. Then, calculate WMR_s by (4.13)

if $WMR_s > \epsilon$ **then**

The faulted line is identified as the first branch in G_{K^*}.

else

  for $s = 2$ to S **do**

    if $WMR_{s-1} \leq \epsilon$ and $WMR_s > \epsilon$ **then**

      The boundary edge between paths $s-1$ and s is located as the faulted line.

    end if

  end for

end if

Output: the faulted line.

To guarantee the observability for faulted lines, the following conditions presented in earlier research about meter placement are considered:

1) PMU measurements are available at a substation and nodes in the main feeders that have many downstream nodes. Such design is suggested in [71] and [72] to provide improved observability with a limited number of PMU measurements.

2) PMUs can be installed at the ends of feeders or long laterals to ensure the observability and identify the faulted lateral if necessary [69, 70].

To maximize the location accuracy using a certain number of PMUs, an optimal PMU placement method presented in [74] can be implemented by considering the probability of various fault types' occurrence. However, the optimal meter placement is a complicated multi-objective optimization problem, involving multiple impact factors, such as the installation cost, estimation accuracy, and faulted line observability, etc. We adopt a meter

placement scheme with a low measurement redundancy following the above-mentioned conditions, which illustrates the potential of the proposed method for faulted line identification with limited PMU installations.

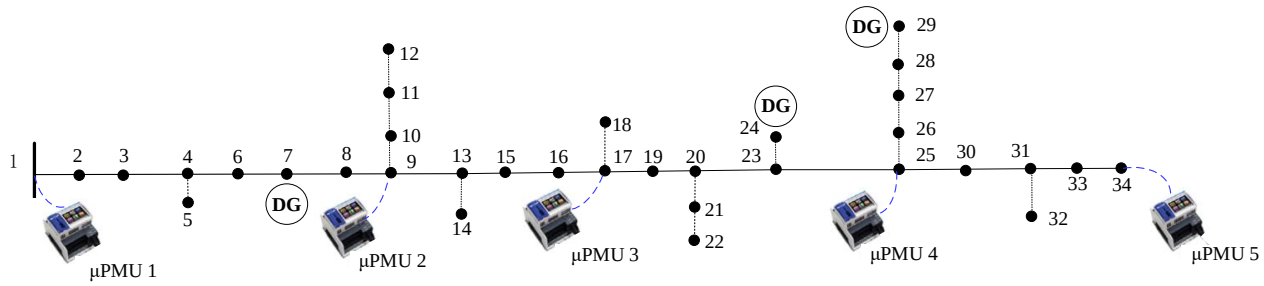

Figure 4.4. Single-line diagram for the three-phase 34-node test feeder

4.4. Case Study

We test the proposed algorithm on a three-phase unbalanced 24.9kV, IEEE 34-node test feeder. The test system is modified by adding three DGs (two synchronous generators and one PV [69]) into the original systems and simulated in PSCAD, and the proposed location method runs in MATLAB. The detailed models of these DGs can be found in [75]. Fig. 4.4 shows the PMU and DG placement in the system, where five PMUs are installed and the nameplate capacity of DGs is 500kVA. The procedure adopted by PMUs to obtain voltage or current phasor measurements is described in Appendix A. Illustrated as Table 4.2, the propose graph partition in Section 4.3 divides the system into four subgraphs, *i.e.*, $G_K \subseteq G$ and $K = 1, 2, 3, 4$.

Assume that measurement noises obey Gaussian distributions. Moreover, the maximum errors of magnitudes and phase angles for PMU data are 1% of the true values and 0.01 rads, respectively, while the maximum errors for the powers recorded by pseudo-measurements at load/DG nodes are 20% of the true values. Also, smart meters can be installed at DGs for accurately monitoring power outputs, and the maximum errors of these outputs are 3%.

Table 4.2. Measurement information of subgraphs

Subgraph G_K	PMU K	PMU $K+1$	Nodes in G_K
G_1	1	2	1-12
G_2	2	3	9-18
G_3	3	4	17-29
G_4	3	4	25-34

By collecting measurements at the DG nodes, distribution system operators (DSOs) do not require the specific DG models. Moreover, DG operators may not share these detailed models and control policies with DSOs due to a lack of agreements between them. However, DSOs can still monitor their power dispatch by the measurement data [75].

4.4.1. Faulted-subgraph Identification

This section shows the identification performance of the proposed method in Step One for faulted subgraphs. We test the proposed algorithm with single-phase LG faults, which are set at three branches in each subgraph, *e.g.*, branches 3-4, 7-8, and 10-11 in G_1. Moreover, these faults are placed at the beginning ($0.25L_l$), in the middle ($0.5L_l$), and at the end ($0.75L_l$) of the lines, and L_l denotes the corresponding line length. In each fault location, fifty sets of measurements are generated by Monte Carlo simulations. Also, considering nine fault locations for each subgraph, 9×50=450 fault scenarios for two influence factors (fault locations and measurement noises) in each subgraph are tested. In all tests, the subgraphs with the highest values of the identification function J_K correctly point to those faulted subgraphs. The values of J_K across $K = 1, 2, 3, 4$ for these faults are shown in Fig. 4.5, where we average the WMRs of each subgraph for conciseness. As discussed in Section ó, WMR greatly increases in the faulted subgraph, indicating that the fault is located at that subgraph. For example, when a fault occurs in G_1, J_1 is abnormally higher than J_2, J_3, and J_4. This leads to an immediate conclusion that the fault is located in G_1.

Furthermore, we test the two-phase line-to-line faults with 50 Ω fault impedance in G_K, and Fig. 4.6 depicts J_K in all subgraphs. We observe that the maximum of J_K correctly

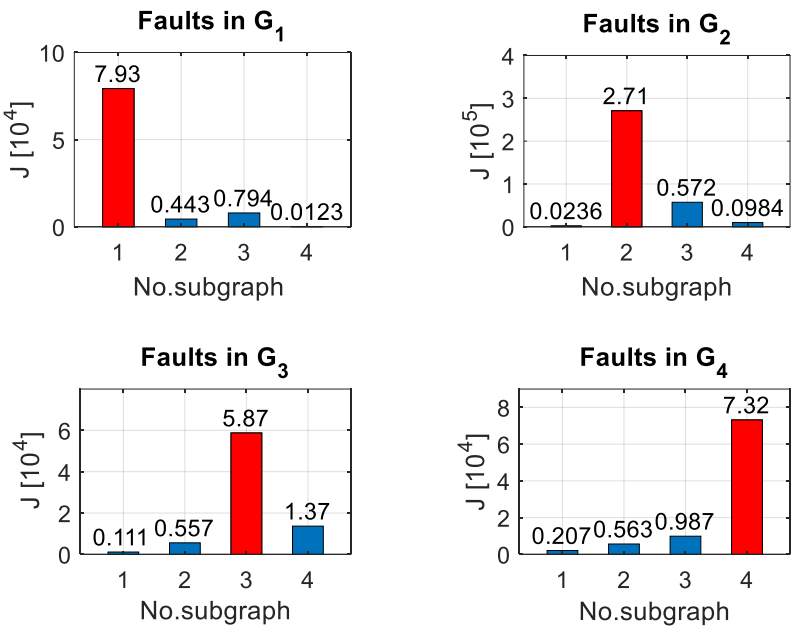

Figure 4.5. Identification results in different faulted subgraphs, where we set LG faults on phase A with 100 Ω impedance in G_K.

indicates the location of the faulted subgraph. Also, the identification performance for the faulted subgraph is not influenced by fault types and fault impedances.

4.4.2. Faulted-line Location

We test various fault scenarios to evaluate the location performance of the proposed method. Fig.4.7 shows the WMRs for different paths in the faulted subgraph for LL faults on phases B and C on branch 3-4 in G_1, where we run 100 Monte Carlo simulations for random combinations of measurement noises. In this figure, the WMRs of the normal paths are much lower than those for the corrupted paths. Also, with the radial expansion of paths, the WMRs of the corrupted path that the boundary edge 3-4 lies in have high values. Consequently, branch 3-4 is identified as the faulted line.

We further evaluate the location accuracy of the proposed method by calculating the probabilities of two types of test results: 1) the faulted branch is correctly located and 2) an

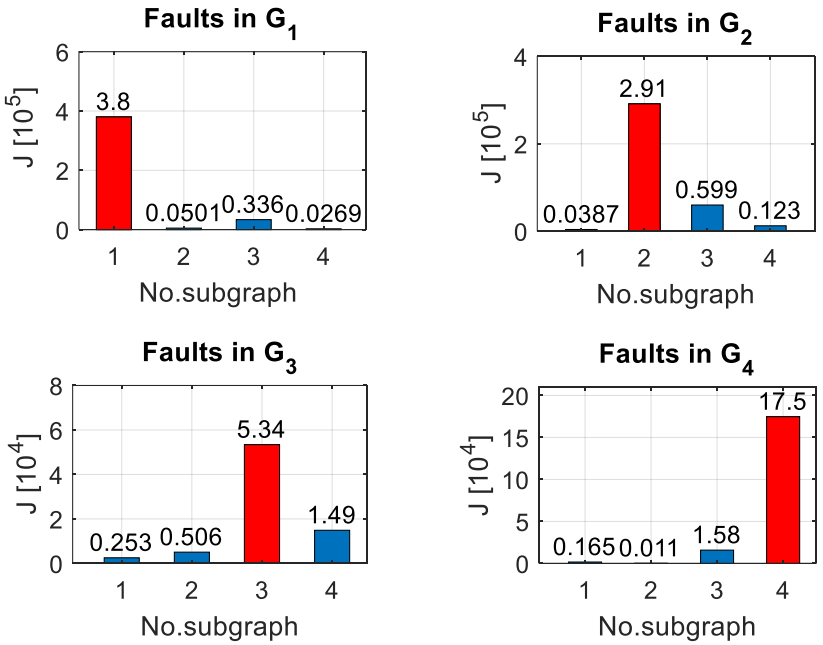

Figure 4.6. Identification results in different faulted subgraphs. LL faults on phases B and C with 50 Ω fault impedance.

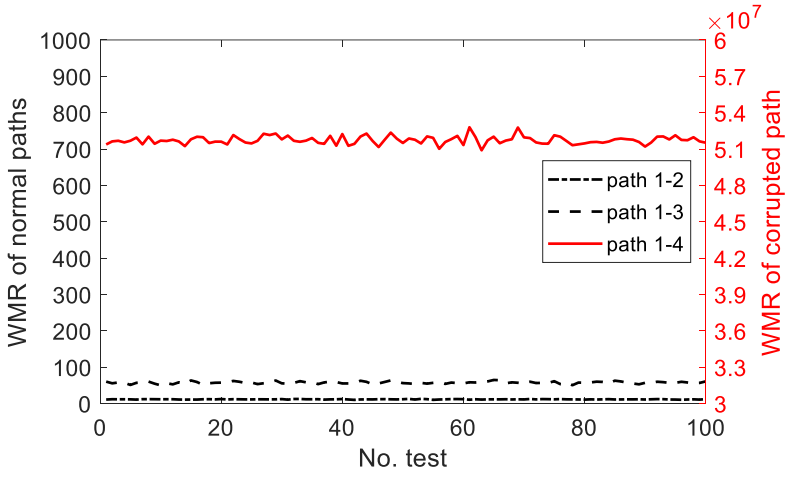

Figure 4.7. Location results in G_1, when LL faults with 10Ω fault impedance occur at branch 3-4. The secondary y-axis shows the WMRs at the corrupted path in 100 Monte Carlo simulations.

Table 4.3. Performance with different fault types (50 Ω impedance)

Fault Type	α	β	$1-\alpha-\beta$	Max Error
G_1	94.50%	4.67%	0.83%	2 branches
G_2	95.75%	4.25%	0%	1 branch
G_3	96.83%	3.17%	0%	1 branch
G_4	95.67%	4.33%	0%	1 branch

immediate neighboring branch of the faulted branch is determined as a faulted one [70]:

$$\alpha = N_0/N_t \tag{4.16}$$

$$\beta = N_1/N_t \tag{4.17}$$

where N_0 and N_1 denote the number of the tests in these two cases, respectively, and N_t is the total number of the tests; also, $1-\alpha-\beta$ is used to calculate the probability of other results, *i.e.*, other branches are determined as a faulted line.

We calculate these accuracy indices α and β in scenarios with various fault types and fault impedances, where $N_t = 1200$ is set to obtain statistical results in each scenario, and here the identification threshold $\epsilon = 500$.

1) Fault Type

The impacts of various fault types on the location accuracy of the proposed algorithm are investigated. Four fault types denoted as LG, LL, LLG, and LLL, are tested. We list the location results of these fault types in Table 4.3. It is shown that the proposed method enables correct faulted line location with various fault types and reaches 94% and higher accuracy.

2) Fault Impedance

We test the impacts of fault impedances on the accuracy of the proposed algorithm. We set different fault impedances at each branch of G_2, and Table 4.4 shows the accuracies of this method to locate faults with these impedances. Especially, the proposed method enables

accurate locations of bolted faults, owing to the existence of the fault injection currents with high magnitudes. The results in Table 4.4 show that the proposed method enables correct fault-line location with multiple fault impedances.

Table 4.4. Performance with different fault impedances (LL faults)

Fault Type	α	β	Max Error
0Ω	100%	0%	0 branches
10Ω	94.67%	5.33%	1 branch
50Ω	94.83%	5.17%	1 branch
100Ω	100%	0%	0 branch
200Ω	95.08%	4.92%	1 branch

The proposed method is tested with high-impedance LG faults (100, 200, 500, 800, and 1000 Ω) at branch 3-4 to show the sensitivity towards magnitudes of fault currents. Fig.4.8 shows the probabilities of the correct location of the faulted branch, where Monte Carlo simulations with 400 samples of measurements are used. The location probabilities are higher than 88% under these various current injections. The reason is that the measurement errors of voltages and currents are proportional to the measurement values, while the measurement weights are inversely proportional to them. While smaller fault current injections occur, as the weights of measurements are higher in this case, the WMR will be high. We conclude that the proposed method works effectively when the fault impedance is not higher than 1000 Ω in the test system. Once the fault impedance exceeds about 2000 Ω, the proposed approach may not observe the small fault injection at long branches in the 34-node system.

4.4.3. Robustness and Sensitivity Analysis

We investigate the robustness and sensitivity of the proposed method against various measurement noises and identification thresholds.

1) Measurement Errors

We conduct robustness analysis concerning higher measurement noises. We set the measurement noises of PMUs as 2% in magnitudes and 0.02 rads in phase angles, while consider-

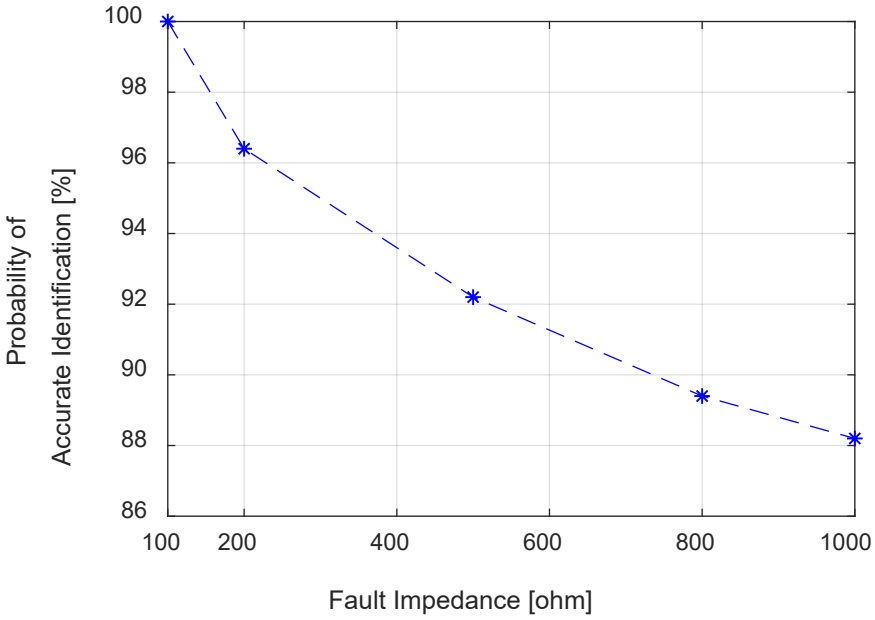

Figure 4.8. Performance for high-impedance LG faults at branch 3-4.

ing the maximum errors of pseudo-measurements as 10%, 30%, and 50%. Table 4.5 lists the accuracy of the proposed algorithm with these measurement noises. As shown, even with high pseudo-measurement errors up to 50%, either the correct line or its immediately neighboring line is identified. It implies that such high-level noises do not degrade the location performance since DSSE takes the weights of measurement noises into full account. Also, the location performance of this algorithm is robust against measurement errors.

Table 4.5. Performance with higher measurement errors (LG, 50 Ω impedance)

Max Error of PMU Data	Max Error of Pseudo-meas.	α	Max Error
2%, 0.02 rad	10%	94.67%	1 branch
	30%	94.50%	1 branch
	50%	94.42%	1 branch

2) Identification Threshold

We test the location performance of the proposed method with various identification

thresholds ϵ. Fig. 4.7 shows that the WMRs in the corrupted paths are much higher than those for the normal paths. Further, different thresholds are set in the cases of Section 4.4.2, and the location accuracy with these thresholds is calculated and listed in Table 4.6. We conclude that the identification threshold could be properly selected to maintain a desirable identification sensitivity.

Table 4.6. Impact of identification thresholds

Threshold ϵ	α	β	$1-\alpha-\beta$	Max Error
100	88.69%	10.35%	0.96%	2 branches
500	95.92%	3.91%	0.17%	2 branches
1000	97.08%	2.92%	0%	1 branch
2000	97.25%	2.75%	0%	1 branch

4.4.4. Impact of Line Parameters

Line parameters in distribution systems are subject to changes with environmental conditions. Considering this uncertainty, the range of line parameters is generally set within $\pm 5\%$ of their nominal values [20]. Therefore, we consider the variation in line parameters to evaluate the accuracy of the proposed method.

Table 4.7. Location accuracy with uncertainty in line parameters

Maximum Errors of Line Parameters	α	β	Max Error
2%	100%	0%	0 branch
5%	97.5%	2.5%	1 branch
10%	95.5%	3.25%	2 branches

We use Monte Carlo simulations to generate 400 test scenarios, where imprecise line parameters are assumed to obey Gaussian distribution with various maximum deviations and zero means. Table 4.7 lists the location accuracy of the proposed algorithm, and the maximum errors are 2%, 5%, and 10% of true values of line parameters. Also, LG faults

with 100 Ω are set on different branches in G_2. With 5% deviation in the line parameters, either the correct faulted line or its immediate neighboring branch is identified.

We conclude that inaccurate line parameters degrade the location accuracy of the proposed method, and hence line calibration in power systems is necessary periodically.

4.4.5. Computational Efficiency

Numerical experiments for different faults are performed to demonstrate the computational efficiency of the proposed algorithm. We run this method on a PC with 2.6 GHz i5, and 8GB RAM using MATLAB 2017b.

Table 4.8. CPU time in test system

Fault Subgraph	Average CPU Time [ms]
G_1	14.28
G_2	13.50
G_3	14.27
G_4	9.98

Table 4.8 lists the average CPU time of the proposed method, including two steps, for faulted line identification in these test systems. It shows that once the measurement data are collected, this algorithm locates the faults within 15 milliseconds in the 34-node distribution system. Compared to the traversal search strategy for a whole feeder, the proposed method runs in parallel for feeders with a reduced size, which improves the computational efficiency for application to the larger-scale networks. It should be noted that owing to the increase in the number of nodes in subgraphs of this larger-scale system, the proposed method takes a longer CPU time, *i.e.*, about 20 ms, for faulted line identification. With more PMUs installed, the number of nodes in a subgraph decrease, and the computational efficiency of this method can be further improved.

4.5. Conclusion

This chapter proposes a graph-based faulted line identification algorithm using PMU data in distribution systems. We present a distributed DSSE algorithm to identify the faulted subgraph efficiently, and this method significantly reduces the searching scale and speeds up the subsequent fault location procedure. Further, we conveniently determine a faulted line by applying a hierarchical graph-subgraph-path structure to the DSSE method. The proposed method is suitable for radial distribution systems and can be updated by incorporating the linear measurement functions proposed in [42] for weakly meshed distribution systems. In the case of inadequate PMU and DGs installed at the distribution level, the proposed method enables accurate faulted-line location. Extensive simulations verify the accuracy and efficiency of this method under various fault scenarios.

Chapter 5

Semi-supervised Deep Learning for False Data Injection Attack Detection

5.1. Introduction

Power distribution systems are transforming into smart grids with the development of advanced communicating devices, such as phasor measurement units (PMUs) and smart meters, which facilitate the system monitoring and control. However, the high dependence on information technology also increases vulnerability from malicious cyber-attacks [76]. Among common attacks in cyber-physical systems, false data injection attacks (FDIAs) are regarded as one of the most challenging threats against secure system operation. An unobservable FDIA can circumvent the conventional bad data detection (BDD) mechanism based on measurement residuals of state estimation [8]. Without the aids of the effective detection mechanism, attackers can stealthily launch the FDIA multiple times, which degrades the performance of the state estimation algorithm and may render a significant threat to the grids [30].

Some research on FDIA construction is reported in different application scenarios in DC power systems, while recent work in AC transmission systems emerges due to their reactively accurate analytical models [77, 78]. Liang *et al.* [79] conducted a comprehensive survey on construction methods for FDIAs. On the other hand, many results using various statistical and probabilistic techniques are reported to defend against FDIA in DC system state estimation, such as sparse optimization [80] and Kalman filter [81]. However, these methods require information on measurement data distributions and system operation states, and once these perquisites change, detection for FDIAs may become ineffective and outdated.

Recently, with the fast development of advanced metering infrastructure that collects a massive volume of data, machine-learning and data-driven techniques are being widely

applied to power system operation because of their powerful capability of extracting useful information and flexible extensibility [82]. Also, various learning-based techniques for detecting FDIAs in transmission systems have emerged, including deep belief networks (DBN) [83], support vector machines (SVM) [84,85], and deep neural networks (DNN) [86]. For instance, He *et al.* [83] proposed a conditional DBN-based method with a restricted Boltzmann machine for detecting compromised data in DC power systems. The authors of [84] proposed a semi-supervised learning approach based on mixture Gaussian distribution and SVM for detecting FDIAs against state estimation in DC transmission systems, and since a linear system model is used to generate the measurement data, principal component analysis (PCA) is used for extracting the data feature. However, real-world utilities widely employ AC power system models, and these algorithms, such as [80, 83–85], performed on DC power systems, ignore the complexity of power systems or the sophistication of unobservable attacks by the conventional BDD mechanism. To overcome this deficiency when dealing with unobservable attacks in AC transmission systems, [86] uses wavelet transform and DNN techniques to capture the inconsistency of abnormal and normal measurements by analyzing the state dynamics. Nevertheless, the method in [86] requires measurements with labels from continuous samplings that may be unavailable in practical operation and leads to a high computational burden. Note that most existing machine-learning algorithms for detecting FDIAs such as [83] and [86] are supervised and test the abnormal data that differ in some manner from the labeled data available during training. However, the datasets collected from practical cyber-physical systems are partially labeled due to expensive labeling costs [87]. Moreover, the scale of unlabeled data is usually much larger than that of the labeled data in practice, and these extensive unlabeled data seldom take part in the supervised learning process. This absence leads to the loss of useful information, even the failure in this process.

In contrast to the work in transmission systems, there is a handful of research related to FDIAs at the distribution level, although the vulnerability of distribution systems has been discovered over the years [79]. For instance, Dai *et al.* [88] presented two simple yet powerful cyber-attack methods targeting feeder automation and introduced a search

theory-based method for modeling the probability of feeders being attacked. Deng *et al.* [89] proposed an FDIA model with limited knowledge of system states, which exposes the feasibility of attacks without being detected by the current BDD mechanism. Then, they extended this work focusing on balanced networks to unbalanced distribution systems in [44]; these systems are more consistent with practical models. Motivated by these studies on constructing unobservable FDIAs in distribution systems, reliable system operation demands countermeasures against these FDIAs urgently.

This chapter proposes an adversarial autoencoder (AAE)-based detection algorithm for unobservable FDIAs in distribution systems. Considering the high dimensionality and nonlinear correlated nature of measurements, we apply autoencoders to dimension reduction and feature extraction of measurement datasets in the three-phase unbalanced networks. Further, we integrate the autoencoders into an advanced generative adversarial networks (GAN) framework [90], which successfully detects abnormal measurements under FDIAs by capturing the unconformity between anomalies and secure measurements. Also, because of the expensive labeling costs and potential missing labeled data in practical systems, this method only requires unlabeled data and a few labeled data from measuring instruments by leveraging the powerful generation capability of GAN and thus is semi-supervised learning.

5.2. State Estimation and FDIA

5.2.1. BDD based on Estimation Residuals

Section 2.2.1 introduces classical state estimation theory [10]. The relationship between redundant measurements and state variables is depicted as:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e} \tag{5.1}$$

where $\boldsymbol{x}$ is an n-dimension state vector, and $\boldsymbol{z}$ is an m-dimension measurement vector; $\boldsymbol{h}(\boldsymbol{x})$ is the measurement function about $\boldsymbol{x}$; the measurement noise vector $\boldsymbol{e}$ obeys a Gaus-

sian distribution $\boldsymbol{e} \sim N(0, \boldsymbol{R})$ is a covariance matrix and is usually considered diagonal, $\boldsymbol{R} = \text{diag}\,[\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2]$, and σ_i^2 denotes the variance of the $\boldsymbol{i}$th measurement error, $i = 1, 2, \ldots, m$.

The state variables are obtained via a WLS criterion that minimizes the sum of weighted measurement residuals J as:

$$\hat{\boldsymbol{x}} = \arg\min J = \arg\min[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})]^T \boldsymbol{W} [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] \tag{5.2}$$

where $\boldsymbol{W}$ is a weight matrix of measurements to quantify the trust levels of diverse measurements, and $\boldsymbol{W} = \boldsymbol{R}^{-1}$.

In three-phase unbalanced distribution systems, the system states are chosen as the voltage phasors at all buses and expressed as

$$\boldsymbol{x} = [v_1^a, v_1^b, \cdots, v_n^c] \tag{5.3}$$

where v_j^φ denotes the φ-phase voltage phasor at bus j, $\varphi = \{a, b, c\}$ and $j = 1, 2, \ \ldots, n$; n is the number of the buses in the system.

The measurement vector $\boldsymbol{z}$ in DSSE generally includes 1) voltage and current phasors from distribution-level PMUs, 2) power flows recorded by supervisory control and data acquisition systems, and 3) power injections from smart meters or pseudo-measurements, including load consumption and DG [20]. The detailed formulation of the measurement functions $\boldsymbol{h}\,(\boldsymbol{x})$ can be found in Section 2.2.2. The DSSE model is nonlinear since PMUs are not installed at each node in a practical distribution system [10]. Then, the solution of the DSSE model is solved iteratively based on the Newton-Gauss process (2.3)-(2.5).

Considering the sampling errors of various measuring instruments and potential malicious cyberattacks, the conventional BDD mechanism of state estimation applies the LNR test [79]. The LNR test calculates the measurement residual $\boldsymbol{r}$ and the normalized measurement

residual vector $\boldsymbol{r}^N$ to test the existence of bad measurements:

$$\boldsymbol{r}^N = \frac{|\boldsymbol{r}|}{\sqrt{diag(\boldsymbol{SR})}} \tag{5.4}$$

where $\boldsymbol{r} = \boldsymbol{z} - \boldsymbol{h}(\widehat{\boldsymbol{x}})$; $\boldsymbol{S}$ represents the measurement sensitivity matrix of this estimator, and $\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{W}$.

The LNR test is used to detect the presence of bad data due to malicious cyber-attacks, faulty sensors, or topological errors, expressed as:

$$\max_i \boldsymbol{r}_i^N \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{5.5}$$

where $\boldsymbol{r}_i^N$ denotes the ith element of $\boldsymbol{r}^N$, and $i = 1, 2, \ldots, M$; $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the hypotheses without and with bad data or false data injection, respectively, and the threshold λ is set to some predetermined significant level.

5.2.2. On the Existence of Linear Approximation of DSSE

The power flow equations already contain the linear relationships (2.7) and (2.8) between $\boldsymbol{x}$ and PMU measurements, together with nonlinear relationships between $\boldsymbol{x}$ and power measurements. The authors of [91] proposed linear approximation theorems and the error analysis to establish a linear model between the voltages and powers. Based on [91], [44] proposes an FDIA construction model in the distribution system.

In [91], the complex power measurements at node k or at branch i-j can be converted to the equivalent currents as

$$\boldsymbol{I}_{k\_eq} = (\boldsymbol{S}_k/\boldsymbol{V}_{slack})^* \tag{5.6}$$

$$\boldsymbol{I}_{ij\_eq} = (\boldsymbol{S}_{ij}/\boldsymbol{V}_{slack})^* \tag{5.7}$$

where $\boldsymbol{V}_{slack}$ denotes the voltage at the substation.

The DSSE model with this approximation is expressed as

$$\boldsymbol{z} = [\boldsymbol{V}_k, \boldsymbol{I}_{ij}, \boldsymbol{I}_{k\_eq}, \boldsymbol{I}_{ij\_eq}]^T = [\boldsymbol{I}, \boldsymbol{Y}_{br}, \boldsymbol{Y}_{bus}, \boldsymbol{Y}_{br\_eq}]^T \boldsymbol{x} + \boldsymbol{e} \tag{5.8}$$

where the Jacobian matrix $\boldsymbol{H} = [\boldsymbol{I}, \boldsymbol{Y}_{br}, \boldsymbol{Y}_{bus}, \boldsymbol{Y}_{br\_eq}]^T$ and $\boldsymbol{I}$ denotes the identity matrix; $\boldsymbol{Y}_{br}$ is a matrix composed of all the $\boldsymbol{Y}_{ij}$ at $\{i,j\} \in \psi_I$ and zero elements, $\boldsymbol{Y}_{bus}$ (or $\boldsymbol{Y}_{br\_eq}$) is a matrix composed of all the $\boldsymbol{Y}_k$ (or $\boldsymbol{Y}_{ij}$) at $k \in \psi_S$ (or $\{i,j\} \in \psi_S$) and zero elements.

This linear approximation solution is closer to the nonlinear solution provided by (2.11), compared with a linear solution by simplifying the AC distribution system as a DC model. The conclusion is validated by the case study in [44,91]. Based on this linear approximation, an unobservable FDIA in three-phase distribution systems can be constructed by the method presented in the next section.

5.2.3. Unobservable FDIA

The objective of FDIAs is to mislead system operators to consider $\boldsymbol{x}_a = \boldsymbol{x} + \boldsymbol{c}$ as the estimated state vector, where $\boldsymbol{c}$ is the deviation of normal system states $\boldsymbol{x}$ [8]. The attackers can manipulate the received measurements at a control center into $\boldsymbol{z}_a = \boldsymbol{z} + \boldsymbol{a}$, and $\boldsymbol{a}$ is an injected attack vector. Also, the measurement residual vector of $\boldsymbol{z}_a$ is expressed as

$$\boldsymbol{r}_a = \boldsymbol{z}_a - \boldsymbol{h}(\boldsymbol{x} + \boldsymbol{c}) = \boldsymbol{z} + \boldsymbol{a} - \boldsymbol{h}(\boldsymbol{x} + \boldsymbol{c}). \tag{5.9}$$

To circumvent the detection method in (5.4), the attack vector should be constructed as

$$\boldsymbol{a} = \boldsymbol{h}(\boldsymbol{x} + \boldsymbol{c}) - \boldsymbol{h}(\boldsymbol{x}). \tag{5.10}$$

The attack can bypass the residual-based detection, if the normalized residual vector $\boldsymbol{r}_a^N$ satisfies the condition

$$\max_i \boldsymbol{r}_{a,i}^N \leq \lambda. \tag{5.11}$$

Here, we refer to the attack as an unobservable FDIA with a well-structured attack vector $\boldsymbol{a}$ that enforces (5.11). By launching such an FDIA, the attacker can inject estimation errors without being detected by the conventional LNR test.

5.3. Construction of FDIA

To evaluate the FDIA detection capability, this section introduces an unobservable attack construction method in distribution systems. Assuming that an attacker has limited ability to hack into meters, they can use the attack method proposed in [44] for constructing FDIAs without paying high calculation costs.

Emerging research indicates that approximately linear models for state estimation in distribution systems exist and provide convenience for hackers to launch unobservable FDIAs, *e.g.*, [44,89]. Specifically, the linear formulas in these DSSE methods approximate the nodal voltages in equivalent current measurements as the substation voltage $\boldsymbol{V}_{slack}$. This approximation originates from two observations for distribution systems: 1) the voltage magnitudes are close to each other, *i.e.*, 0.95 $\sim$ 1.05 p.u., and 2) the voltage phase angle differences are very small, such as 0.1 degrees per mile. Take the three-phase power flow measurement at branch i-j, $\boldsymbol{S}_{ij}$, as an example, convert the power measurement into an equivalent current measurement by $\boldsymbol{I}_{ij\_eq} \approx (\boldsymbol{S}_{ij}/\boldsymbol{V}_{slack})^*$, and then the measurement function can be expressed as

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{Y}_{ij}(\boldsymbol{v}_i - \boldsymbol{v}_j) \approx \boldsymbol{I}_{ij\_eq} \tag{5.12}$$

where $\boldsymbol{Y}_{ij} \in \mathbb{C}^{3\times3}$ denotes the nodal admittance between nodes i and j, which is constant, and hence, $\boldsymbol{h}(\boldsymbol{x})$ is linear about the state variables $\boldsymbol{x} = \{\boldsymbol{v}_i, \boldsymbol{v}_j\}$; $[\cdot]^*$ denotes the complex

conjugate.

More details of the linear approximation can be found in Section 5.2.2. Then, due to the existence of the nearly linear relationship in DSSE, the estimator (5.1) can be expressed in a linear form as

$$\widetilde{\boldsymbol{z}} = \boldsymbol{H}\widetilde{\boldsymbol{x}} + \boldsymbol{e} \tag{5.13}$$

where $\widetilde{\boldsymbol{z}}$ and $\widetilde{\boldsymbol{x}}$ denote the measurement and the closed-form estimated vector.

When the attack vector $\boldsymbol{a}$ is injected, the compromised measurement residual $\boldsymbol{r}_a$ can be expressed as [44]

$$\boldsymbol{r}_a = \widetilde{\boldsymbol{z}}_a - \boldsymbol{H}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{z}} + \boldsymbol{a} - \boldsymbol{H}[\widehat{\boldsymbol{x}} + (\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{a}] \tag{5.14}$$

If $\boldsymbol{a} = \boldsymbol{H}\boldsymbol{c}$, the compromised measurement residual $\boldsymbol{r}_a$ after the attack is the same as the measurement residual $\boldsymbol{r}$ before the attack, as follows.

$$\boldsymbol{r}_a = \widetilde{\boldsymbol{z}} - \boldsymbol{H}\widehat{\boldsymbol{x}} + \boldsymbol{H}\boldsymbol{c} - \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H}\boldsymbol{c} = \widetilde{\boldsymbol{z}} - \boldsymbol{H}\widehat{\boldsymbol{x}} = \boldsymbol{r} \tag{5.15}$$

If the residual $\boldsymbol{r}$ can bypass the LNR test, the compromised residual $\boldsymbol{r}_a$ with malicious data can also bypass this test by (5.11). Furthermore, to construct an attack vector that meets $\boldsymbol{a} = \boldsymbol{H}\boldsymbol{c}$, let $\boldsymbol{A} = \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T$ and $\boldsymbol{A} \in \mathbb{R}^{M\times M}$, and solve $\boldsymbol{a}$ by the following equation:

$$\boldsymbol{A}\boldsymbol{a} = \boldsymbol{A}\boldsymbol{H}\boldsymbol{c} \leftrightarrow \boldsymbol{A}\boldsymbol{a} = \boldsymbol{a} \leftrightarrow \boldsymbol{B}\boldsymbol{a} = \boldsymbol{0} \tag{5.16}$$

where $\boldsymbol{B} = \boldsymbol{A} - \boldsymbol{I}$ and $\boldsymbol{I}$ is an identity matrix. Assume that the maximum number of measurements that hackers can compromise is K, and express the attack vector as $\boldsymbol{a} = [0, a_{i_1}, \ldots, 0, a_{i_k}, \ldots, 0]^T$ and $0 < k \leq K$. Also, the elements a_i with $i \in \{i_1, \ldots, i_k\}$ are the unknown variables to solve. Then $\boldsymbol{B}\boldsymbol{a} = \boldsymbol{0}$ is equivalent to $\boldsymbol{B}'\boldsymbol{a}' = \boldsymbol{0}$, where $\boldsymbol{a}'$ is

the k-dimension vector that removes zero-value elements in $\boldsymbol{a}$, and $\boldsymbol{B}'$ is the matrix that removes the corresponding columns at the locations related to these zero-value elements in $\boldsymbol{a}$, $\boldsymbol{B}' \in \mathbb{R}^{M\times k}$. If $\text{rank}(\boldsymbol{B}') < k$, there is at least one non-zero solution in $\boldsymbol{B}'\boldsymbol{a}' = \boldsymbol{0}$. Moreover, $\boldsymbol{a}'$ can be obtained by $\boldsymbol{a}' = (\boldsymbol{I} - \boldsymbol{B}'^{+}\boldsymbol{B}')\boldsymbol{d}$, where $\boldsymbol{B}'^{+}$ is the pseudo inverse of $\boldsymbol{B}'$, and $\boldsymbol{d}$ is an arbitrary non-zero vector, $\boldsymbol{d} \in \mathbb{R}^{k\times 1}$.

Attackers can successfully launch unobservable attacks by the attack model (5.16). Therefore, the mechanism for detecting the unobservable FDIAs demands an effective solution.

5.4. Proposed Detection Mechanism

This section proposes a deep learning mechanism for detecting unobservable FDIAs in three-phase distribution systems; Fig. 5.1 provides an overview of the proposed AAE-based detection mechanism. We define the detection problem for unobservable FDIAs or outliers represented by the attack vector $\boldsymbol{a}$ as a binary classification problem with the detection indicator α:

$$\alpha = \begin{cases} 0 & if\ \boldsymbol{a} = 0 \\ 1 & if\ \boldsymbol{a} \neq 0 \end{cases} \tag{5.17}$$

In the AAE-based detection algorithm, the measurement vector $\boldsymbol{z}$ including three-phase voltages, currents, and powers are collected as the inputs of AAE, where only a limited number of them is labeled with $\alpha = 0$ or 1. Compared with the conventional LNR test, the learning-based detection method is fully data-driven and does not require the knowledge of system knowledge of topology and parameters, *i.e.*, $\boldsymbol{h}(\boldsymbol{x})$. Furthermore, we train the AAE network to extract the node-to-node and phase-to-phase features of normal and abnormal data, and then detect the presence of an unobservable FDIA. To evaluate the detection performance, we generate unobservable attacks by the method introduced in Section 5.3, while the normal dataset comes from the nonlinear distribution system model in Section 2.2.2.

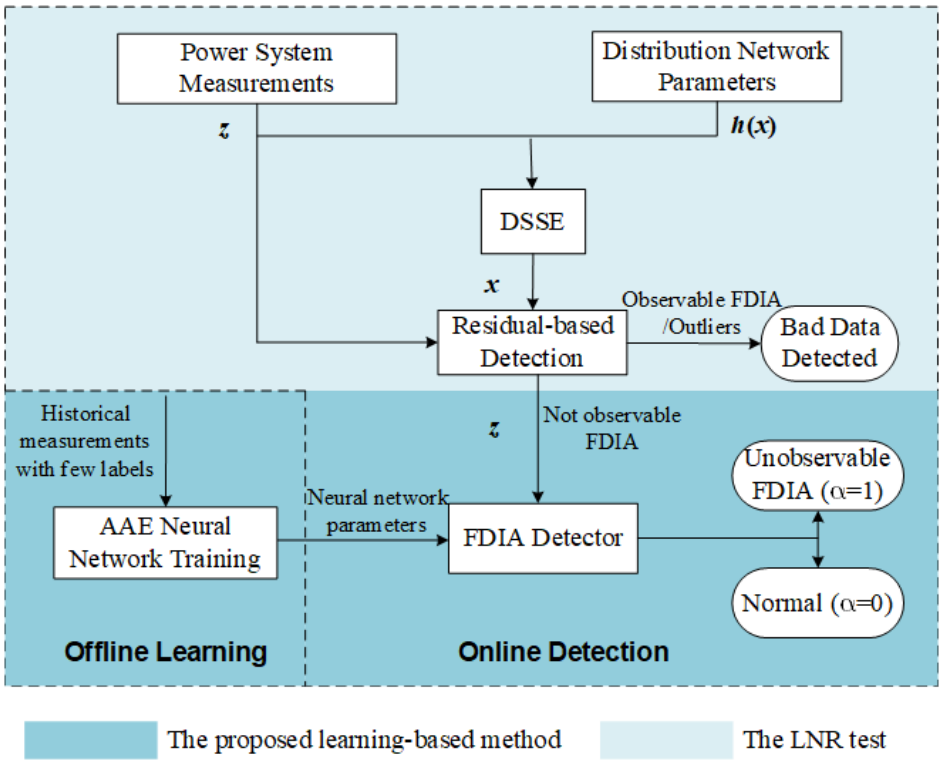

Figure 5.1. An overview of the proposed FDIA detection mechanism

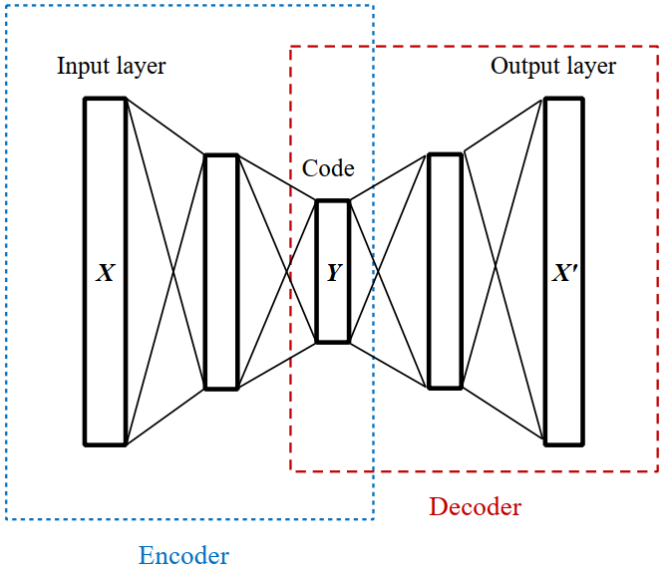

Figure 5.2. Structure of autoencoders with three fully connected hidden layers

5.4.1. Autoencoders and GAN

1) Autoencoders

Autoencoders are widely used for dimension reduction and feature extraction of highly dimensional and correlated data [92]. Fig. 5.2 shows a classical structure of autoencoders, including an encoder and a decoder. Specifically, the autoencoders learn a mapping from an input $\boldsymbol{X}$ to a hidden code $\boldsymbol{Y}$ and the mapping is parameterized as $q(\boldsymbol{Y}|\boldsymbol{X}) = q(\boldsymbol{Y}|\boldsymbol{X};\boldsymbol{\theta})$ with the parameters $\boldsymbol{\theta}$ that we want to learn.

Define $p(\boldsymbol{X})$ as the prior distribution and $p(\boldsymbol{Y})$ as the prior distribution that we want to impose on the code. The encoding function of the autoencoders $q(\boldsymbol{Y}|\boldsymbol{X})$ defines an aggregated posterior distribution $q(\boldsymbol{Y})$ on the hidden code vector as

$$q(\boldsymbol{Y}) = \int q(\boldsymbol{Y}|\boldsymbol{X})p_d(\boldsymbol{X})d\boldsymbol{X} \tag{5.18}$$

where $p_d(\boldsymbol{X})$ denotes the data distribution of $\boldsymbol{X}$.

Encoder: The mapping f_θ transforms the input $\boldsymbol{X}$ into $\boldsymbol{Y}$ and is expressed as

$$f_\theta(\boldsymbol{X}) = s(\boldsymbol{w}\boldsymbol{X}_j + \boldsymbol{b}) \tag{5.19}$$

where $\boldsymbol{w}$ and $\boldsymbol{b}$ denote the weight matrix and offset vector, respectively, and $\boldsymbol{\theta} = \{\boldsymbol{w}, \boldsymbol{b}\}$, $\boldsymbol{w} \in \mathbb{R}^{h\times M}$, and $\boldsymbol{b} \in \mathbb{R}^{h\times 1}$; h denotes the number of the hidden units; $\boldsymbol{X}_j \in \mathbb{R}^{M\times 1}$ is the jth vector of the input samples $\boldsymbol{X}$, $j \in \{1, 2, \ldots, N_d\}$, and $\boldsymbol{X} \in \mathbb{R}^{M\times N_d}$, where N_d is the number of these samples; $s(\cdot)$ denotes the squashing nonlinearity of the neural network.

Decoder: The hidden code $\boldsymbol{Y}$ is then mapped back to a reconstruction $\boldsymbol{X}'$ in the input space, *i.e.*, $\boldsymbol{X}' = g_{\theta'}(\boldsymbol{Y})$. This mapping $g_{\theta'}$ is called the decoder, and based on the parameters $\boldsymbol{\theta}' = \{\boldsymbol{w}', \boldsymbol{b}'\}$, it is shown as

$$g_{\theta'}(\boldsymbol{Y}) = s(\boldsymbol{w}'\boldsymbol{Y} + \boldsymbol{b}'). \tag{5.20}$$

The autoencoders optimize the network parameters by minimizing the mean square error

between $\boldsymbol{X}$ and $\boldsymbol{X}'$ as the reconstruction cost, L_R [92]:

$$arg \min_{\boldsymbol{\theta},\boldsymbol{\theta}'} L_R = \frac{1}{N_d}\|\boldsymbol{X} - \boldsymbol{X}'\|^2 \tag{5.21}$$

The parameters $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ are usually backpropagated by stochastic gradient descent (SGD) in the training process. The convergence proof of autoencoders can be found in [93].

2) GAN

GAN establishes a min-max adversarial game between two neural networks, a generator, G, and a discriminator, D, shown as Fig. 5.3. The generator produces the measurement data samples (fake samples) that follow the distribution of the original training data (real samples), while the discriminator distinguishes between the generated data samples and these real samples. In a nutshell, GAN is alternatively trained in two stages: a) update the discriminator with fixed generator parameters to distinguish the real samples from the generated samples, and b) update the generator with the fixed discriminator parameters to fool the discriminator with its generated fake samples. The solution of the two-player game is globally optimal, and [90] provides the proof of this optimality and the convergence analysis.

Because of limited labeled measurements available for training in practical power systems, we use GAN to aid the autoencoders in shaping the hidden code for accurately detecting FDIAs, which is detailed later.

5.4.2. Adversarial Autoencoder

This section introduces the structure of AAE and its training procedure for FDIA detection. Here, the input $\boldsymbol{X}$ is a measurement dataset with P labeled samples $\{(\boldsymbol{z}_1, \alpha_1), (\boldsymbol{z}_2, \alpha_2), \ldots, (\boldsymbol{z}_P, \alpha_P)\}$ and Q unlabeled samples $\{\boldsymbol{z}_{P+1}, \boldsymbol{z}_{P+2}, \ldots, \boldsymbol{z}_{P+Q}\}$, where $\alpha_p = 0$ or 1 denotes the label of the pth set of measurements, $p = 1, \ldots, P$ and $P \ll Q$. As the inputs of AAE, $\boldsymbol{X} \in \mathbb{R}^{M \times N_d}$, where M is the number of measurements in (5.1) and here the number of the samples $N_d = P + Q$. Each sample of $\boldsymbol{z}$ is further represented in the neural

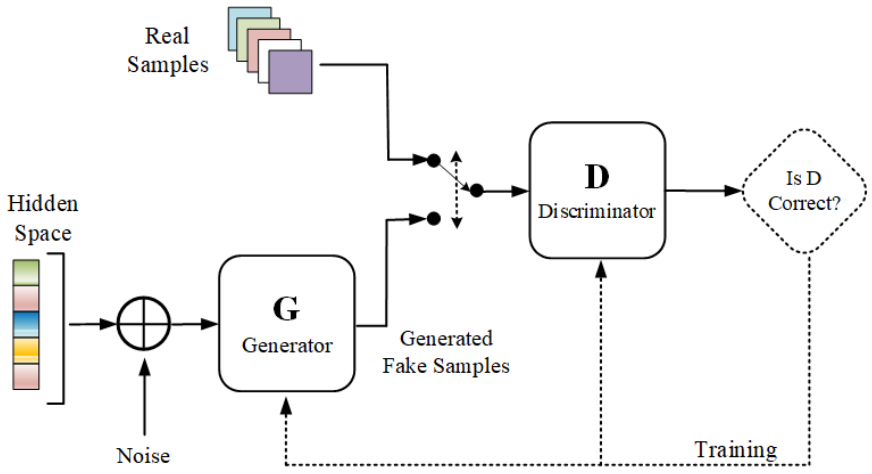

Figure 5.3. The learning process of GAN

networks by $\boldsymbol{X}_j$, and $j \in \{1, 2, \ldots, N_d\}$, shown in (5.19).

In AAE, the encoder learns how to encode a given data into a prior distribution, while the decoder learns a deep generative model that matches the aggregated posterior distribution of the hidden representation from the encoder to an arbitrary prior distribution. Fig. 4 shows the combination and division of work of the autoencoders and GAN in the attack detection task.

In Fig. 5.4, $\boldsymbol{Y}_l$ and $\boldsymbol{Y}_u$ denote the hidden representation for the labeled and unlabeled inputs, and $q_\theta(\boldsymbol{Y}_u, \boldsymbol{Y}_l \mid \boldsymbol{X})$ and $p_{\theta'}(\boldsymbol{X}|\boldsymbol{Y}_u, \boldsymbol{Y}_l)$ denote the encoder and decoder in this semi-supervised learning, respectively. Assume the data generated by $\boldsymbol{Y}_l$ and $\boldsymbol{Y}_u$, named $\boldsymbol{Y}'_l$ and $\boldsymbol{Y}'_u$, follow a two-dimensional categorical distribution for the binary classification problem in (5.17) and Gaussian distributions, respectively, *i.e.*,$\boldsymbol{Y}'_l \sim Cat(2)$ and $\boldsymbol{Y}'_u \sim N(\boldsymbol{0}, \boldsymbol{I})$. Here, adding the Gaussian noises is to stabilize the GAN training [90]. Also, we assume that the aggregated posteriors $p(\boldsymbol{Y}_l)$ and $p(\boldsymbol{Y}_u)$ obey Gaussian distributions. To match the aggregated posterior to the prior distributions of the mixture data, the encoder $q_\theta(\boldsymbol{Y}_u, \boldsymbol{Y}_l|\boldsymbol{X})$ works as the generator of GAN. In the meantime, the adversarial network has two discriminators, D_{cat} and D_{gauss}, for the labelled and unlabeled inputs, respectively.

We train the AAE network at three stages: the reconstruction phase, the adversarial

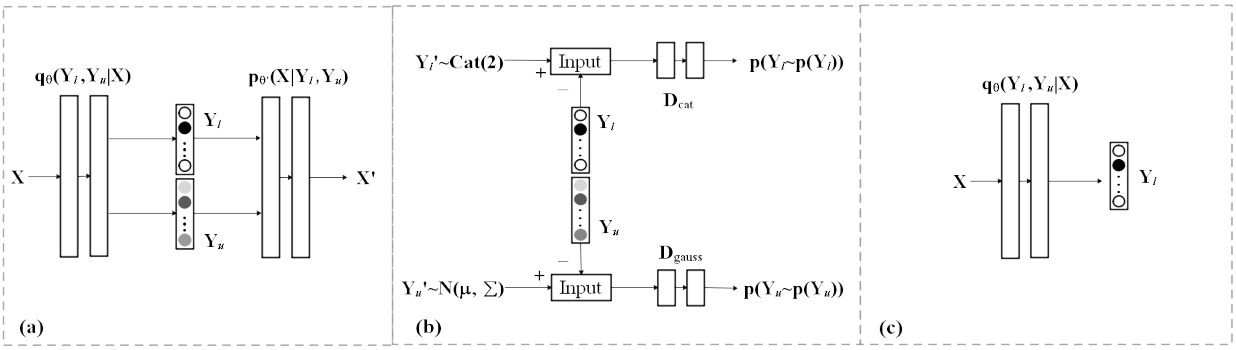

Figure 5.4. Semi-supervised AAE architecture: (a) in reconstruction phase, (b) in adversarial phase, and (c) in supervised phase.

phase, and the supervised phase [94]. The batch normalization technique is used to improve the training speed, performance, and stability of neural networks. We introduce the training procedure of AAE below and show it in pseudo-code.

1) Reconstruction Phase

The AAE detector first works as traditional autoencoders in this phase, shown in Fig. 5.4(a), and both the encoder and decoder are trained to minimize the total reconstruction loss, L_R, for the labeled and unlabeled inputs $\boldsymbol{X}$ as in (5.21).

2) Adversarial Phase

In this phase, the encoder $q_\theta\left(\boldsymbol{Y}_u, \boldsymbol{Y}_l|\boldsymbol{X}\right)$ is reserved for training the discriminators and generator, illustrated in Fig. 5.4(b) the discriminators D_{cat} and D_{gauss} to distinguish the true samples of the categorical from Gaussian priors from the generated samples. Here, the goal of a discriminator is to maximize the probability that $\boldsymbol{Y}_l$ or $\boldsymbol{Y}_u$ comes from the generated data rather than from the true sample distribution, *i.e.*, its confidence score. Hence, we formulate the loss function of D_{cat} for the labeled data as

$$\max D_{cat} V\left(D_{cat}, G\right) = E_{\boldsymbol{Y}'_l \sim \mathrm{Cat}(2)}\left[\log D_{cat}\left(\boldsymbol{Y}'_l\right)\right] + E_{Y_l \sim p(\boldsymbol{Y}_l)}\left[\log\left(1 - D_{cat}\left(G\left(\boldsymbol{Y}_l\right)\right)\right)\right] \quad (5.22)$$

where $E_{\boldsymbol{Y}'_l}$ and $E_{\boldsymbol{Y}_l}$ denote the expectations under the corresponding distributions.

Then, we express the loss function of the generator G as $\max_G V\left(D_{cat}, G\right) = E_{\boldsymbol{Y}_l \sim p(\boldsymbol{Y}_l)}$ $\left[\log D_{cat}\left(G\left(\boldsymbol{Y}_l\right)\right)\right]$, which is equivalently written as

$$\min_G V(D_{cat}, G) = E_{\boldsymbol{Y}_l \sim p(\boldsymbol{Y}_l)}\left[\log\left(1 - D_{cat}\left(G(\boldsymbol{Y}_l)\right)\right)\right]. \tag{5.23}$$

A two-player minimax game for the labeled data is formulated by combining (5.23) with (5.22) as

$$\min_G \max_{D_{cat}} E_{\boldsymbol{Y}'_l \sim \mathrm{Cat}(2)}\left[\log D_{cat}\left(\boldsymbol{Y}'_l\right)\right] + E_{Y_l \sim p(\boldsymbol{Y}_l)}[\log(1 - D_{cat}\left(G\left(\boldsymbol{Y}_l\right)\right))] \tag{5.24}$$

Similarly, the objective function for the unlabeled data is expressed as follows:

$$\min_G \max_{D_{gauss}} E_{\boldsymbol{Y}'_u \sim N(\mu, \sum)}\left[\log D_{gauss}\left(\boldsymbol{Y}'_u\right)\right] + E_{\boldsymbol{Y}_u \sim p(\boldsymbol{Y}_u)}\left[\log\left(1 - D_{gauss}\left(G(\boldsymbol{Y}_u)\right)\right)\right] \tag{5.25}$$

3) Supervised Phase

Using only the labeled data, the autoencoders continue to update the encoder network, shown in Fig. 5.4(c). Train the encoder for the labeled data by minimizing the cross-entropy as the supervised cost by

$$\min_\theta L_S = E_{q(\boldsymbol{Y}_l)}[-\log\, p(\boldsymbol{Y}_l)] \tag{5.26}$$

where the aggregated posterior distribution $q(\boldsymbol{Y}_l)$ is calculated by (5.18), and $p(\boldsymbol{Y}_l)$ is the distribution of $\boldsymbol{Y}_l$ inherited from the results in the adversarial phase. During this training, we use the *Adam* optimization technique to computes adaptive learning rates for each parameter. *Adam* is straightforward to implement and computationally efficient and has little memory requirements [95]. This technique is widely used as a replacement of SGD in the application research of GAN [90], especially for the optimization of objective functions with high-dimensional parameters spaces. Specifically, *Adam* computes bias-corrected first and second moment estimates. We provide the pseudo-code of Adam, where the default

hyperparameters in [95] are used.

Algorithm 2 AAE Training Process

Input:Learning rate γ, mini-batch size, the number of epochs.

for $t = 1$ to N_{ep} **do**

Sample mini-batch.

Reconstruction Phase

Calculate L_R by (5.21) and update $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ by descending the gradients:

$f_\theta \leftarrow \nabla_\theta L_R, \boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \gamma \operatorname{Adam}(f_\theta)$

$g_{\theta'} \leftarrow \nabla_{\theta'} L_R, \boldsymbol{\theta}' \leftarrow \boldsymbol{\theta}' - \gamma \operatorname{Adam}(g_{\theta'})$

Reconstruction Phase

Obtain the hidden representation of the encoder $q_\theta(\boldsymbol{Y}_u, \boldsymbol{Y}_l|\boldsymbol{X})$ and sample from the prior distributions, and calculate the confidence scores of D_{gauss} and D_{cat}.

Discriminator: Train the discriminators to update their network parameters when fixing the generator parameters

Generator: Update $q_\theta(\boldsymbol{Y}_u, \boldsymbol{Y}_l|\boldsymbol{X})$ as the generator with the fixed discriminator parameters.

Supervised Phase: Using only the labeled data, update the encoder to minimize L_S by (5.26).

end for

Output:the encoder $q_\theta(\cdot)$.

5.5. Case Study

We test the proposed AAE-based algorithm on three-phase unbalanced benchmarks: IEEE 13-bus and 123-bus distribution systems [25]. These systems are modified by adding DG units; more details about the location and types of these DG units are provided in [20]. The measurement arrangement of these systems is listed in Table 5.1. The true values of measurements and states are obtained by running power flow program and DSSE in MATLAB, and the proposed AAE-based algorithm runs in Python. Measurements with noises consist of voltage phasors, current phasors, and complex powers, and these measurement noises obey

Table 5.1. Measurement locations in test system

Measurement Types	Placement Locations	
	13-bus System	123-bus system
Power Flow	650-632,645-646, 684-652	149-1,9-14,15-16,18-35, 44-45, 76-77, 86-87, 99-100, 110-112
Power Injection	8 load/DG nodes	91 load/DG nodes
Voltage/currents from PMUs	650, 671	149, 14, 18, 25, 60, 76, 97, 108

Gaussian distributions [96]. Specifically, the maximum meter noises of PMUs [52] are 1% of the true values for voltage/current magnitudes and 0.01 rads for the phase angles, and assume that a PMU measures the nodal voltage and the currents at the branches connected to this bus; the measurement errors of power data at limited branches and all load/DG nodes from smart meters are 3% of the true values [97].

Dataset Structure: The input of the proposed AAE detector is the collection of the measurement vector $\boldsymbol{z}$. In the modified 13-bus system, there are 11 nodes by closing the switch installed at the branch 671-692; the state vector, $\boldsymbol{x}\in\mathbb{R}^{66\times1}$, is composed of the three-phase voltage magnitudes and voltage phase angles, and 17 measurement phasors in Table 5.1 produce a measurement vector, $\boldsymbol{z}\in\mathbb{R}^{102\times1}$. In the modified 123-bus system, there are 119 nodes due to three normally closed switches and one normally open switch; $\boldsymbol{x}\in\mathbb{R}^{714\times1}$ and $\boldsymbol{z}\in\mathbb{R}^{870\times1}$. We record 5,000 sets of measurements from Monte Carlo simulations and generate other 5,000 sets of measurements under unobservable FDIAs by the attack construction methods proposed in Section 5.3 and [8]. In the training process, 80% of measurements are chosen as the training dataset and the rest are used for evaluating detection performance. Further, for semi-supervised learning, we label 1,000 sets of measurements with a ratio of 1:1 as the secure and attacked data. These secure and attacked data are fed to the proposed AAE detector for offline training.

Neural Network Specification: The learning rate is chosen as 0.0001, and the number of epochs is 400. The encoder, decoder, and discriminators have two layers of 1,000 hidden units

with a ReLU activation function, and a sigmoid activation function is used in the output layer of the autoencoder. $\boldsymbol{Y}_l$ and $\boldsymbol{Y}_u$ are two-dimensional for the binary classification problem. We use the Adam optimization method to train these neural networks with mini-batches of 64 samples for optimizing all the loss functions.

5.5.1. Unobservable FDIA

We investigate the detection performance of the conventional BDD method under unobservable attacks to show its insufficiency. Assuming that an attacker has access only to at most K measurements from half the number of all meters, we randomly choose k in $(0, K]$ to generate a k-sparse attack vector $\boldsymbol{a}$ in these systems. Also, the DSSE method in Section 2.2.2 runs 5,000 Monte Carlo simulations under no attacks, and we choose the maximum of $\boldsymbol{r}^N$ in all trials as the detection threshold, λ_0.

Fig.5.5 lists the results of the LNR test under these FDIAs in the 13-bus system, and we compare these results with those with no attacks. We find that these FDIAs are unobservable by the LNR test, since all the residuals are located under the detection threshold. For instance, we construct FDIAs targeting on the A-phase and C-phase voltages at buses 611, 671, and 680 by constructing a sparse attack vector with $k = 4$, and Fig.5.6 shows the estimated states under this unobservable attack. The unobservable FDIA can stealthily compromise the state estimation for voltage magnitudes to make them violate the operation ranges, *e.g.*, below 0.95 p.u. at some buses in Fig.5.6. Estimated states with such significant biases may mislead the decisions made by system operators for voltage regulation.

5.5.2. Detection Performance

Targeting at finding these FDIAs, we evaluate the detection performance of the proposed method in the 13-bus and 123-bus distribution systems. All simulation and training are conducted on a computer with a 2.5 GHz Intel Core i5 CPU and 8 GB of RAM.

We calculate the true positive (TP), the true negative (TN), the false positive (FP), and the false negative (FN) rates, which are defined in Table 5.2. For instance, TP denotes the

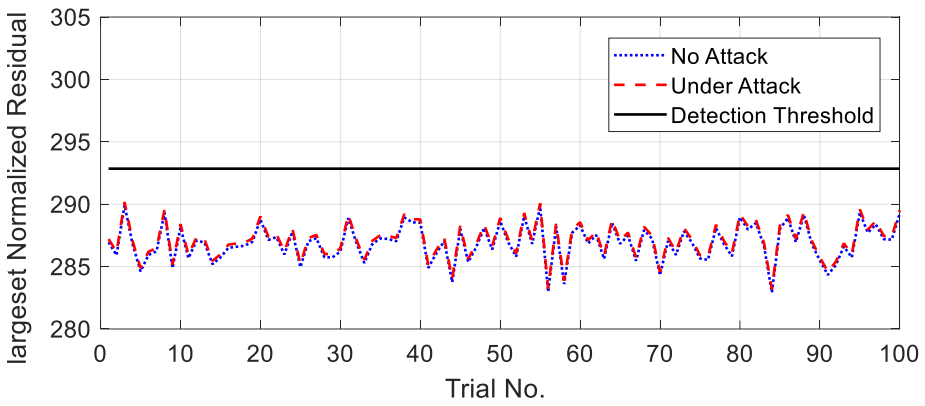

Figure 5.5. The LNR results under unobservable FDIAs and no attacks

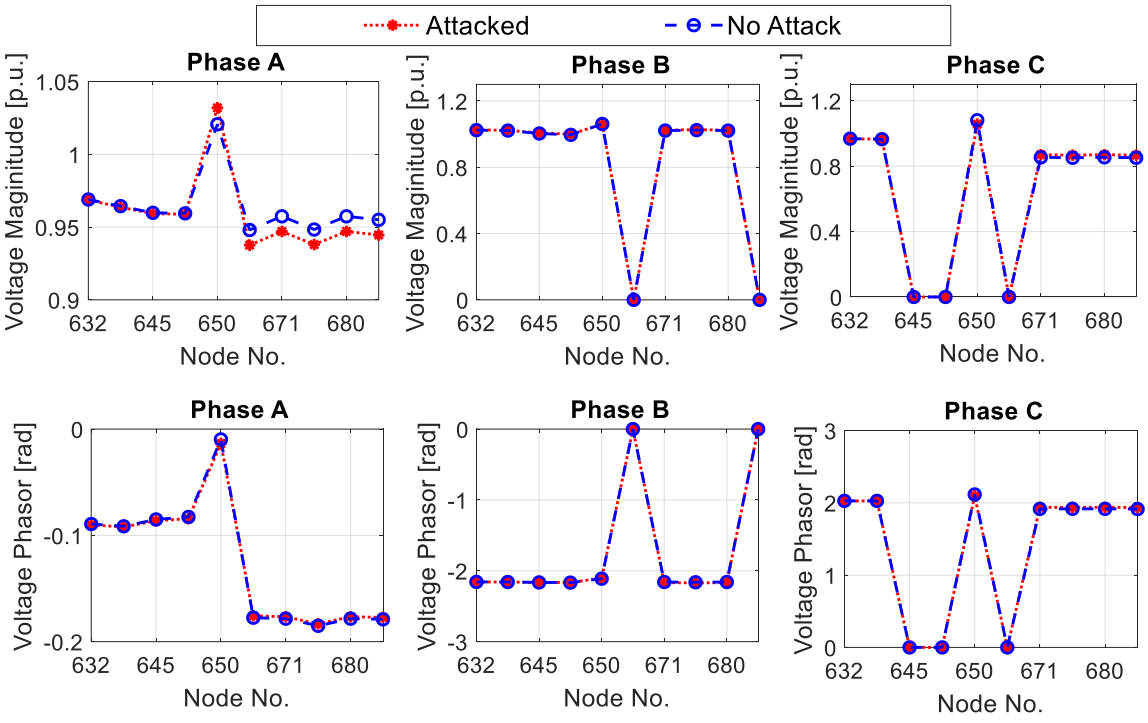

Figure 5.6. Estimation results of voltages under no attacks and an unobservable attack

Table 5.2. Definitions of performance indices

	Attacked	Secure
Classified as attacked	TP	FP
Classified as secure	FN	TN

probability that a measurement classified as attacked is actually exposed to an attack. We evaluate the learning ability of the proposed method by the values of precision (Prec), recall (Rec), and accuracy (Acc) [85]. The precision values are used to evaluate the classification performance for the attacked measurements, while the recall values measure the probability that the secure measurements are not misclassified as attacked. The overall detection performance is measured by the index Acc. Furthermore, we calculate these three indices by

$$\text{Prec} = \frac{TP}{TP + FP} \tag{5.27}$$

$$\text{Rec} = \frac{TP}{TP + FN} \tag{5.28}$$

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.29}$$

Table 5.3 shows the detection accuracy, training time, and detection time of the proposed method in two test systems. The proposed algorithm used for detecting the attacked metering data has a detection error of 3.75% in the 13-bus system, while the detection error is 2.15% in the 123-bus system. Also, the average computation time of the proposed method is 9.30 and 14.81 milliseconds in the 13-bus and 123-bus systems, which can satisfy the requirement of online detection. Without the use of graphics processing unit (GPU), the training time reaches about four hours in the 123-bus test system. Moreover, we further try the fine-tuning technique [98] for neural network training based on the original hardware configuration, the

Table 5.3. Detection performance of proposed algorithm

Performance		13-bus System	123-bus system
Training Accuracy	Labeled	100%	100%
	Unlabeled	96.35%	98.95%
Detection Accuracy		96.25%	97.85%
Average Detection Time		9.30 ms	14.81 ms
Training Time		2.5 hours	4 hours

new training process shows that this time (four hours) is shortened to less than two hours.

5.5.3. Other Semi-supervised Techniques

To evaluate the detection performance of the proposed AAE method, we compare the proposed method with other data-driven detection techniques. We employ the semi-supervised support vector machine (S3VM) proposed in [85] for attack detection as the baseline. Moreover, for a fair comparison, we adopt the semi-supervised autoencoders (SS-AE) and update the k-nearest neighbor (k-NN) method [99] into a semi-supervised version.

Table 5.4 lists the detection performance of these methods. With limited labeled data for training, our approach has a higher detection accuracy owing to the powerful combination of autoencoders and GAN. For instance, the proposed algorithm achieves a high detection accuracy of up to 95%, while the S3VM-based scheme has a worse performance with a detection accuracy of less than 80%. Our conclusion is similar to that of Ozay *et al.* [85] that extensive unlabeled samples in the training dataset largely degrade the classification performance of the SVM method. Moreover, owing to the adoption of the advanced generative models in AAE, the detection accuracy of the proposed method is higher than that of the individual SS-AE algorithm without the use of generative models. Therefore, the proposed method is more competitive.

5.5.4. Impacts of Measurement Noises

To test the robustness against measurement noises in the proposed detection method, we

Table 5.4. Comparison with other semi-supervised methods

Semi-supervised Algorithm	Detection Accuracy [%]	
	13-bus System	123-bus system
Proposed Method	96.25	97.85
S3VM [85]	73.60	76.65
k-NN [99]	84.90	86.55
Autoencoders [92]	91.85	92.40

investigate the impact of various noise levels of measurement data on the detection accuracy in the 123-bus distribution system. In the case with no installation or malfunction of smart meters at load/DG nodes, we use pseudo-measurements with higher errors (*e.g.*, 10% of the true values), which are obtained from the historical or forecasting data of customer loads and DG production, to realize the system observability. In experiments, the maximum errors of the power measurements are set to vary from 4% to 12%. Fig. 5.7 shows the detection results of the proposed method, SS-AE, and S3VM under these noise levels.

It implies in Fig.5.7 that the proposed algorithm achieves a detection accuracy of more than 94% even with the maximum measurement errors of up to 12% of the true values; this finding illustrates the robustness of the proposed detection method against measurement noises. In comparison, the detection accuracy of the SS-AE and S3VM approaches decreases when dealing with higher measurement errors. We conclude that the detection accuracy of the proposed method still remains high when the noise level of the test dataset increases in distribution systems. This is because the adopted generative model has the capability of better shaping the distribution of the hidden code of autoencoders to make the measurement data distinguishable.

5.5.5. Sensitivity Analysis

1) Impact of Amount of Labeled Data

We investigate the detection performance when using relatively fewer labeled data during the training. We set the different amount of the labeled data, ranging from 500 to 1250 in

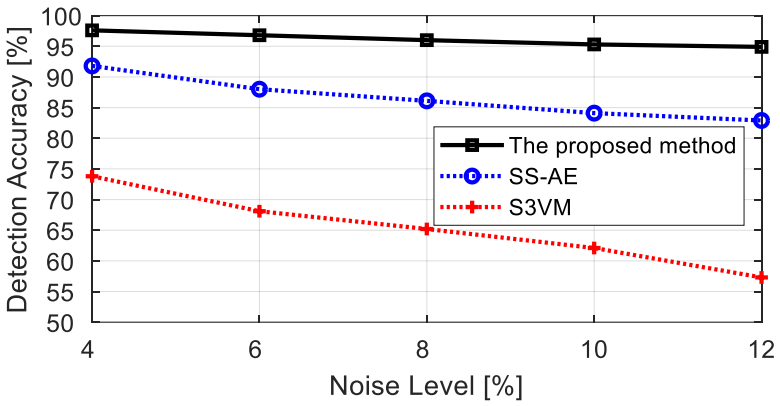

Figure 5.7. Detection accuracy with maximum measurement errors ranging from 4% to 12%.

Table 5.5. FDIA detection with fewer or more labeled data

Number of labeled data	Prec [%]	Rec [%]	Acc [%]
500	91.17	92.26	91.70
750	93.18	94.98	94.15
1000	94.16	96.78	96.25
1250	95.47	97.83	96.70

the 13-bus test system. Table 5.5 provides the confusion matrix and evaluation indices of the proposed algorithm, in which the number of training samples is 8,000. More labeled data during the training leads to a more accurate detection performance. However, with 500 sets of data labeled, the proposed method detects the unobservable FDIAs with precision, recall, and accuracy values of about 91.17%, 92.26%, and 91.70%, respectively, which illustrates the detection effectiveness of this method.

2) Impact of Different Attacks

Considering that there are some potential FDIAs that are not fully investigated and thus are not labeled in the training stage, we test the detection performance towards new attacked samples. New attacked samples here are defined as those that are not labeled in the training stage and produced by different attack construction methods from that of those historical known FDIAs. This case study can be summarized as "using few attacked samples to detect

Table 5.6. Data structure in sensitivity analysis

Data in stage of		Number of normal data	Number of attacked data	
			From [5]	From [20]
Training (8000)	Labeled	8%	0	2%
	Unlabeled	78%	2%	10%
Test (2000)		50%	40%	10%

* The numbers in the brackets denote the size of the samples.

Table 5.7. Comparison of detection performance in two cases

Performance with 400 training epochs		Number of labeled attacked data	
		400	160
Training accuracy	Labeled (800)	100%	100%
	Unlabled (7200)	95.75%	93.15%
Detection accuracy		94.95%	91.60%

* The numbers in the brackets denote the size of the samples.

more new samples" by adopting the generative models.

Specifically, the attacked samples with labels in the training dataset are only from the construction method in [44], and the ratio of these samples to all the training samples is low, *i.e.*, 2%, shown in Table 5.6. Furthermore, we use the method in [8] to constructs different attacks from those labeled samples, and these attacked samples without attaching labels are randomly chosen and put in the training and test dataset as new attacks. These details of the adopted training and test datasets are shown in Table 5.6, and here only 160 attacks are labeled. Other settings are the same as Section 5.5.1.

Shown in Table 5.7, the detection accuracy of the unlabeled data in the training decreases to 93.15%, compared with 95.75% in the case study where 400 attacks are labeled in the total 800 labeled data. In the test stage, the proposed method detects the unobservable FDIAs with an accuracy of 91.60% in the 123-bus system. We conclude that the limited attacked data that are labeled in the training process degrade the detection performance of the semi-supervised learning.

Table 5.8. Confusion matrix of FDIA detection with fewer labeled data

Actual label	Classified as attacked	Classified as secure	Prec [%]	Rec [%]
Attacked (994)	923	71	92.86	90.49
Secure (1006)	97	909		

* The numbers in the brackets denote the size of the samples.

More details about the test performance can be found in Table 5.8. These "new attack" influence the recall value more obviously, and this index is 90.49%. However, the detection performance of the proposed algorithm might be acceptable, since only 2% of the attacked data is labeled in the training stage.

5.6. Conclusion

This chapter proposes a semi-supervised AAE-based algorithm for detecting FDIAs in smart distribution systems. In the case of only a small fraction of labeled measurement data, the proposed method leverages a state-of-the-art GAN framework to realize the effective detection of unobservable FDIAs that bypass the conventional BDD method. Compared with other semi-supervised learning techniques, the proposed algorithm has a high and robust detection accuracy owing to the powerful combination of autoencoders and GAN. The proposed detection method is fully data-driven and does not depend the specific estimation methods and system knowledge. Numerical simulations validate the detection performance of this method.

Chapter 6

Deep Reinforcement Learning Based Volt-VAR Control in Smart Grids

6.1. Introduction

Electric power systems currently adopt volt-VAR optimization (VVO) to improve operational efficiency and reduce power losses in distribution systems. About 10% of the energy losses occur during transmission and distribution to customers, while 40% of the total losses occur at the distribution side, according to the U.S. Energy Information Administration [100]. Research shows that effective VVO control on various regulating devices, such as automatic voltage regulators (AVRs) and switchable capacitors, can realize voltage regulation as well as loss reduction. As a typical tool in the distribution management system (DMS), the primary goal of VVO is to keep voltages at all buses within a normal operation range, *e.g.*, 0.95∼1.05 p.u., according to ANSI C84.1 standard. This topic is further motivated by the penetration of distributed generation (DG), since bidirectional power flow in active distribution systems raises the risk of voltage violation [101]. The DG units equipped with smart inverters have the flexible capability of absorbing or providing reactive power [102]. Thus, a VVO tool with an effective control strategy on these smart inverters can enhance the operational performance of distribution systems with DG penetration [103].

Traditionally, VVO is modeled as a mixed-integer nonlinear programming (MINLP) problem established on optimal power flow (OPF) [104]. Due to the existence of integer variables and nonlinear voltage-dependent load models in systems, the VVO formulation is *nonconvex* and *NP-hard* [105]. More research converts this problem to various optimization problems, namely, mixed-integer quadratically constrained quadratic programming (MIQCQP) and mixed-integer quadratic programming (MIQP), etc. [106, 107]. However, the iteration process in these methods is time-consuming. This low computational efficiency originates from

two reasons: 1) the comprehensive modeling of various control devices largely increases the complexity of these optimization models, and 2) the combination of action variables from multiple control devices results in a huge searching space [108]. On the other hand, these studies run in a centralized manner and adopt linear or nonlinear power flow formulation for single-phase distribution systems to simplify the modeling complexity, such as [104–107] and [109–112]. However, the three-phase unbalanced operation of distribution systems is more consistent with practice, and these approaches can be computationally demanding and do not guarantee optimal performance.

Table 6.1. Literature review of VVO methods in distribution systems

Method	Reference	Model practical consideration			Single/multiple agent(s)
		ZIP load	Accurate three-phase power flow	Smart inverters	
Model-based	[108]	✓	×	×	-
	[113]	×	✓	×	Multiple agents
	[114]	✓	×	✓	-
DRL-based	[115, 116]	×	×	×	Single agent
	This Chapter	✓	✓	✓	Multiple agents

To reduce the computational burden of these centralized algorithms, the decentralized or hierarchical methods, such as [108], [113], and [114], are used for VVO in unbalanced distribution systems. For instance, dividing the feeder into several regulating zones, the authors of [108] formulated a linearized power system model to solve a zone-based optimization problem in each stage via MIQP and then performed a multi-stage coordinated operation to achieve the overall voltage regulation. Unfortunately, the iterations recorded in [108] reach up to thousands and take hundreds of seconds due to this multi-stage operation. Also, these approximation techniques may cause accuracy losses in power flow calculation and lead to suboptimal control strategies. Recently, [114] develops a bi-level VVO formulation, and the lower level models a MILP problem using a nearly linear power flow, while the upper level solves a quadratically constrained nonlinear programming (QCNP) problem based on

nonlinear power flow approximation.

The physical model-based methods mentioned above highly depend on specific optimization models and have limited capability in rapidly adapting to time-varying DG/loads in distribution systems. To address the limitations of these model-based approaches, recent effort applies reinforcement learning (RL) to power system operation, such as [117] and [118]. Furthermore, deep reinforcement learning (DRL) combining deep learning with RL is regarded as valuable alternatives to model-based methods, due to its strong exploration capability of neural networks (NNs) towards nonlinear high-dimensional searching spaces. For example, the DRL-based methods proposed in [119] adaptively provide the voltage setpoints for generators in transmission systems. However, the existing voltage control methods via DRL, such as [119] and [120], only focus on adjusting voltage profiles without looking into the potential of VVO in power loss reduction.

Recently, the authors of [116] extended their work in [115] and solved a constrained Markov decision process for VVO in the DistFlow environment via a constrained soft actor-critic algorithm. However, [116] and [115] do not consider voltage-dependent loads and smart inverters installed at DG units, both of which are widely used in practical distribution systems with renewable integration [104, 114]. On the other hand, the DistFlow equations may suffer from high errors in power flow, loss, and voltage calculations in power networks with heavy load and high renewable penetration [121]. Such inaccuracies are more pronounced when operating power factors deviate from unity, *e.g.,* in systems with DG participating in voltage regulation or high r/x ratios. Also, coordinated VVO control on various regulating devices has not been fully investigated in three-phase unbalanced distribution systems. A literature review of the VVO methods in such systems is shown in Table 6.1. Most of the existing RL-based VVO methods adopt a single agent, and if directly applied to larger-scale systems with multiple regulation devices, they have a slower learning speed due to a huge searching space of variables [122].

Targeting at auto-adaptive voltage control under time-varying operating conditions, we propose a data-driven and model-free VVO approach via multi-agent DRL (MADRL) in

unbalanced distribution systems. The proposed method is novel since we cast the multi-objective VVO problem for distribution systems to an intelligent DQN framework. In this framework, we consider the statuses/ratios of capacitors, AVRs, and smart inverters as action variables. These actions are determined via the agents that are trained by interacting with their environment, *i.e.*, the distribution system. Moreover, by customizing a reward function that effectively guides the DQN training process, this method realizes dual goals on power loss reduction and voltage regulation simultaneously.

6.2. System Model of Unbalanced Distribution Networks

Unlike transmission systems, distribution networks have radial or weakly meshed topologies with lines of a high r/x ratio, which may make the traditional Newton-Raphson power flow method fail to converge [21]. The efficient forward-backward sweep method provides exactly accurate power flow results without any approximation, even in large-scale unbalanced distribution systems. Hence, this method is widely used in power flow calculation for radial distribution systems.

We adopt the ZIP model in distribution systems, which is a voltage-dependent load model widely used in related research, such as [112, 114]. The ZIP load models for active and reactive powers at bus i are depicted as

$$P_i^{\varphi} = P_{i,0}^{\varphi}\left[k_{p,1}\left(\frac{|U_i^{\varphi}|}{U_0}\right)^2 + k_{p,2}\left(\frac{|U_i^{\varphi}|}{U_0}\right) + k_{p,3}\right] \tag{6.1}$$

$$Q_i^{\varphi} = Q_{i,0}^{\varphi}\left[k_{q,1}\left(\frac{|U_i^{\varphi}|}{U_0}\right)^2 + k_{q,2}\left(\frac{|U_i^{\varphi}|}{U_0}\right) + k_{q,3}\right] \tag{6.2}$$

where P_i^{φ} and Q_i^{φ} denote the φ -phase active and reactive powers at bus i, respectively, and $\varphi = \{a, b, c\}$; $k_{p,1} + k_{p,2} + k_{p,3} = 1$, and $k_{q,1} + k_{q,2} + k_{q,3} = 1$; $P_{i,0}^{\varphi}$ and $Q_{i,0}^{\varphi}$ denote the φ -phase active and reactive powers at the nominal voltage U_0; $|U_i^{\varphi}|$ represents the φ -phase voltage magnitude at bus i.

We briefly introduce the procedure of the forward-backward sweep method for power flow calculation in unbalanced distribution systems as follows:

1. *Current Injection Calculation.* Initialize three-phase voltages at all buses as the values of nominal voltages. In each iteration, the three-phase current injections at bus k are calculated by

$$\begin{bmatrix} I^a_{k,in} \\ I^b_{k,in} \\ I^c_{k,in} \end{bmatrix} = \begin{bmatrix} (S^a_k/U^a_k)^* \\ (S^b_k/U^b_k)^* \\ (S^c_k/U^c_k)^* \end{bmatrix} \tag{6.3}$$

where $S^\varphi_k = P^\varphi_k + jQ^\varphi_k$ denotes the φ –phase complex power of load consumption or DG production at bus k and is considered as the ZIP model in (6.1) and (6.2) ; U^φ_k denotes the φ -phase voltage at bus k in the current iteration, and $k = 1,\ 2, \ldots, N$; $[\cdot]^*$ represents the complex conjugate.

2. *Backward Sweep.* Starting from the end bus of the feeder, we calculate the receiving-end current $\boldsymbol{I}_m$ at branch n - m via the Kirchhoff's current law:

$$\boldsymbol{I}_m = \begin{bmatrix} I^a_m \\ I^b_m \\ I^c_m \end{bmatrix} = \begin{bmatrix} I^a_{m,in} \\ I^b_{m,in} \\ I^c_{m,in} \end{bmatrix} + \sum\nolimits_{l\in\mathcal{N}(m)} \begin{bmatrix} I^a_l \\ I^b_l \\ I^c_l \end{bmatrix} \tag{6.4}$$

where $\mathcal{N}(m)$ denotes the set of the downstream branches connected to bus m, and I^φ_l denotes the sending-end current on phase φ at branch l.

The sending-end current $\boldsymbol{I}_n$ at branch n - m is calculated based on (6.4) by

$$\boldsymbol{I}_n = \boldsymbol{c}_{nm} \begin{bmatrix} U^a_m \\ U^b_m \\ U^c_m \end{bmatrix} + \boldsymbol{d}_{nm} \begin{bmatrix} I^a_m \\ I^b_m \\ I^c_m \end{bmatrix} = \boldsymbol{c}_{nm}\boldsymbol{U}_m + \boldsymbol{d}_{nm}\boldsymbol{I}_m \tag{6.5}$$

where $\boldsymbol{c}_{nm}$ and $\boldsymbol{d}_{nm}$ are determined by the line parameters at this branch, $\boldsymbol{c}_{nm} = \boldsymbol{Y}_{abc} + \frac{1}{4}\boldsymbol{Y}_{abc}\boldsymbol{Z}_{abc}\boldsymbol{Y}_{abc}$ and $\boldsymbol{d}_{nm} = \boldsymbol{I} + \frac{1}{2}\boldsymbol{Z}_{abc}\boldsymbol{Y}_{abc}$; $\boldsymbol{Z}_{abc} \in \mathbb{C}^{3\times 3}$ denotes the three-phase line impedance matrix at branch n- m, and $\boldsymbol{Y}_{abc} \in \mathbb{C}^{3\times 3}$ is the shunt capacitance matrix; $\boldsymbol{I}$ denotes an identity matrix.

3. *Forward Sweep.* Starting from the root bus and moving towards the end bus of the feeder, the nodal voltage at bus m is calculated from the voltage at bus n and the sending-end current by

$$\boldsymbol{U}_m = \boldsymbol{a}_{nm}\begin{bmatrix} U_n^a \\ U_n^b \\ U_n^c \end{bmatrix} - \boldsymbol{b}_{nm}\begin{bmatrix} I_n^a \\ I_n^b \\ I_n^c \end{bmatrix} = \boldsymbol{a}_{nm}\boldsymbol{U}_n - \boldsymbol{b}_{nm}\boldsymbol{I}_n \tag{6.6}$$

where $\boldsymbol{a}_{nm} = \boldsymbol{I} + \frac{1}{2}\boldsymbol{Z}_{abc}\boldsymbol{Y}_{abc}$ and $\boldsymbol{b}_{nm} = \boldsymbol{Z}_{abc}$.

The iterative procedure continues until the voltage differences at all nodes in two successive iterations satisfy

$$\Delta\boldsymbol{U}_k = \left\|\boldsymbol{U}_k^{(t+1)} - \boldsymbol{U}_k^{(t)}\right\|_\infty < \varepsilon \quad k \in \{1,\ 2, \ldots, N\} \tag{6.7}$$

where $\boldsymbol{U}_k^{(t)}$ and $\boldsymbol{U}_k^{(t+1)}$ denote the three-phase voltages at bus k at iterations t and $t+1$, and ε denotes the iteration tolerance.

The flowchart of power flow calculation is shown in Fig. 6.1. Furthermore, we calculate the total active power loss in the whole system based on the nodal voltages at all buses by:

$$P_{loss} = \sum\nolimits_{m:n\to m} \text{real}(\boldsymbol{U}_{nm}^T \boldsymbol{I}_{nm}) \tag{6.8}$$

where $\boldsymbol{U}_{nm}$ denotes the three-phase voltage drop at branch n-m, and $\boldsymbol{U}_{nm} = \boldsymbol{U}_n - \boldsymbol{U}_m$; $\boldsymbol{I}_{nm}$ is the three-phase current through the line impedance, and $\boldsymbol{I}_{nm} = \boldsymbol{Z}_{abc}^{-1}\boldsymbol{U}_{nm}$; the function real$(\cdot)$ takes the real part of the complex number, and $(\cdot)^T$ denotes the transpose of a vector.

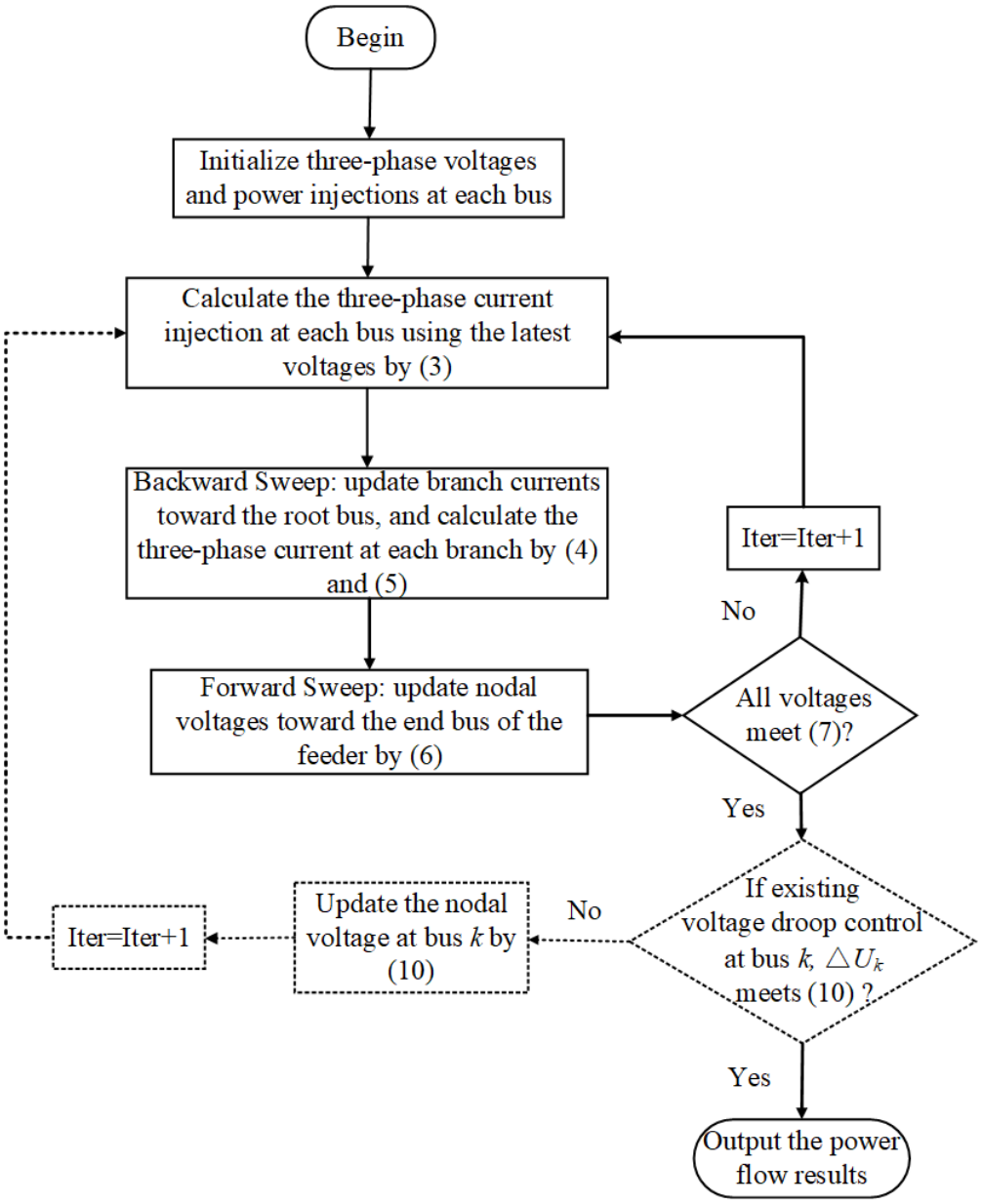

Figure 6.1. Flowchart of the forward-backward sweep method

Power flow calculation requires prior knowledge of the network topology. For distribution networks with an unknown topology, efficient topology identification methods, such as [123], should be adopted for further system monitoring and control.

6.3. Proposed VVO Algorithm

This section integrates the models of switchable capacitor banks, AVRs, and smart inverters into the power flow calculation to evaluate the impacts of the status changes of these regulating devices on system operation. Leveraging these changes as control actions, we propose a multi-agent DQN-based VVO method to realize effective voltage regulation and power loss reduction.

6.3.1. Voltage Regulation Devices

1) DG With Smart Inverters

We develop a per-phase model of smart inverters installed at DGs. Various voltage control strategies can be used in these smart inverters. Here, we assume these smart inverters adopt the reactive power control (RPC) strategy, and voltage droop control can also be used by updating the backward-forward sweep method.

i. RPC

Assume a DG unit with the smart inverter installed at bus k and the active power of DG outputs is known, and its nominal per-phase capacity is $S^{\varphi}_{dg,k}$. In RPC, the φ - phase reactive power provided or absorbed by this DG unit at bus k can be expressed as the following box constraint [114]:

$$\underline{Q}^{\varphi}_{dg,k} \leq Q^{\varphi}_{dg,k} \leq \overline{Q}^{\varphi}_{dg,k} \tag{6.9}$$

where $\overline{Q}^{\varphi}_{dg,k}$ denotes the maximum reactive power of this DG unit installed at bus k, and $\underline{Q}^{\varphi}_{dg,k} = -\overline{Q}^{\varphi}_{dg,k}$; $\overline{Q}^{\varphi}_{dg,k} = \sqrt{\left(S^{\varphi}_{dg,k}\right)^2 - \left(P^{\varphi}_{dg,k}\right)^2}$, and $P^{\varphi}_{dg,k}$ denotes the φ -phase active power; here we define the control variable as $a^{\varphi}_{dg} \in [-1, 1]$, and $Q^{\varphi}_{dg,k} = a^{\varphi}_{dg}\overline{Q}^{\varphi}_{dg,k}$. The dispatchable range of $Q^{\varphi}_{dg,k}$ is relatively narrow since a high power factor (*e.g.*, 0.95) is preferable during DG operation.

ii. Droop Control

In voltage droop control at bus k, the inverter output voltage $\boldsymbol{U}_k$ are controlled by the droop characteristics defined by the following equation [124]:

$$\boldsymbol{U}_k - \boldsymbol{U}_{k,0} = k_q(\boldsymbol{Q}_{dg,k} - \boldsymbol{Q}_{k,0}) \tag{6.10}$$

where $\boldsymbol{U}_k - \boldsymbol{U}_{k,0}$ denotes the three-phase voltage deviations from their rated values, $\boldsymbol{U}_{k,0}$; $\boldsymbol{Q}_{dg,k} - \boldsymbol{Q}_{k,0}$ is the variation of the three-phase reactive powers delivered by the power converter to compensate such deviations; the parameter k_q represents the slope of the voltage droop characteristic.

For each $\boldsymbol{Q}_{dg,k}$ computed by the action that the proposed DRL agent gives, the nodal voltage can be obtained by (6.10). The backward-forward sweep process can be further modified for power flow calculation in droop-regulated distribution systems or microgrids by using the method proposed in [125]. This process can be found in the dotted part of Fig. 6.1.

2) Capacitor Bank

We adopt the three-phase model of capacitor banks. Specifically, we define the reactive power provided by the capacitor installed on phase φ as a function of the control variable, $a_c^\varphi \in \{0, 1\}$, which indicates the status (on/off) of this capacitor. The capacitor provides reactive power when it is connected, *i.e.*, $a_c^\varphi = 1$, and the reactive power at bus k is calculated by the following nonlinear function of U_k^φ.

$$Q_c^\varphi = a_c^\varphi (U_k^\varphi)^2 B_k^\varphi \tag{6.11}$$

where U_k^φ denotes the φ -phase voltage of the capacitor installed at bus k, and B_k^φ denotes the susceptance of the capacitor on phase φ.

3) Voltage Regulator

A voltage regulator with a regulating range of $\pm 10\%$ is used in distribution systems, and the series and shunt impedance of the voltage regulators are neglected since their values can be regarded as extremely small [114]. Define a_r^φ as the step for the voltage regulator on phase φ. For example, in a 32-step regulator, a_r^φ takes values between 0.9 and 1.1 at a step of 0.00625 p.u. The control variable for the regulator is defined by

$$a_r^\varphi = \sum_{j=1}^{33} b_j \alpha_j \tag{6.12}$$

where α_j denotes the binary variable for the jth regulator step position, and $\sum_{j=1}^{33} \alpha_j = 1$; $b_j \in \{0.9,\ 0.90625, \ldots,\ 1.1\}$ and $\varphi \in \{a, b, c\}$.

For branch n - m that with the regulator installation, an additional bus n' is introduced between buses n and m. The impacts of the regulators installed at this branch on the voltage $\boldsymbol{U}_n$ and current $\boldsymbol{I}_n$ are quantified by [21]:

$$\boldsymbol{U}'_n = \boldsymbol{A}_r \boldsymbol{U}_n \tag{6.13}$$

$$\boldsymbol{I}'_n = \boldsymbol{D}_r \boldsymbol{I}_n \tag{6.14}$$

where $\boldsymbol{U}'_n$ denotes the three-phase voltage at bus n', and $\boldsymbol{I}'_n$ denotes the three-phase current that flows out from this regulator; $\boldsymbol{A}_r = diag\{a_r^a, a_r^b, a_r^c\}$, and $\boldsymbol{D}_r = \boldsymbol{A}_r^{-1}$.

In the power flow calculation, we replace $\boldsymbol{U}_n$ and $\boldsymbol{I}_n$ with $\boldsymbol{U}'_n$ and $\boldsymbol{I}'_n$ at the locations of regulators to run the forward and backward sweeps in (6.5) and (6.6).

To introduce the control strategies of smart inverters into the DQN-based VVO framework, we discretize the action space of a_{dg}^{φ} to handle the performance of these actions in RL. This processing method is widely accepted and used for flexible Q-learning or DQN applications, such as [126, 127]. Here, we suppose that DG operators have a certain number of strategies for each control interval (6.9) in practice [128], and a_{dg}^{φ} can take values between -1 and 1 at a step of 0.1.

Fully considering all setting changes of the capacitor banks, smart inverters, and AVRs, the power flow calculation process in (6.3)-(6.7) is updated by integrating (6.9)-(6.14), along with the voltage-dependent loads modeled by (6.1) and (6.2).

The objective of conventional VVO is to keep the nodal voltages in a normal range (*e.g.*, 0.95 to 1.05 p.u.) and concurrently minimize total active power loss [109, 110]. Hence, the

optimization model for this purpose can be formulated as

$$\boldsymbol{a} = \arg \min \ P_{loss} \tag{6.15a}$$

$$\text{s.t.} \quad \text{unbalanced power flow equations in (5.1)-(5.7)} \tag{6.15b}$$

$$0.95 \leq |U_k^{\varphi}| \leq 1.05 \tag{6.15c}$$

where $|U_k^{\varphi}|$ denotes the voltage magnitude on phase φ at bus k, and $\varphi = \{a, b, c\}$; P_{loss} is calculated by (6.8).

Under time-varying operating conditions, the effective power flow calculation acts as the environment for DQN agents, and the details of the proposed VVO algorithm are shown in the next section.

6.3.2. MADRL Design for VVO

In an RL process, a NN is defined as an agent, and the part where the agent takes control actions is the environment. Massive episodes of training are applied to the agent, and in the environment, the load consumption and DG production in a distribution system vary in each episode. The agent is required to take control actions with respect to the given operating condition to achieve VVO. The dimension of the action space increases explosively with the number of controllable devices installed in the three-phase distribution system. Also, a single-agent DQN is challenging to efficiently provide actions due to the extremely high dimension of the joint action space [122]. To improve computational efficiency and ensure scalability for VVO, we propose a multi-agent DQN-based algorithm. The interaction between multiple agents and the environment is depicted via three elements: state $\boldsymbol{s}$, action $\boldsymbol{a}$, and reward r^t at episode t. We describe these elements for VVO below.

1) State and Action

The action vector for the VVO problem decided by all agents is expressed as

$$\boldsymbol{a} = [a_1, a_2, \ldots, a_{Ni}]^T \tag{6.16}$$

where a_i denotes the control action that originates from the statuses of three-phase smart inverters, regulators, or capacitors, *i.e.*, a_r^φ, a_c^φ, and a_{dg}^φ, and $a_i \in \mathcal{A}_i$; $\mathcal{A}_i$ represents the searching space of the ith action, and $i \in \{1, 2, \ldots, Ni\}$.

For the action vector $\boldsymbol{a}$ provided by the agents, the environment provides three-phase voltages at all buses in distribution systems, which act as the states of the DQN. Moreover, these states are expressed as $\boldsymbol{s} = [\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_N]$, where $\boldsymbol{U}_k = [U_k^a, U_k^b, U_k^c]$ and $k = \{1, 2, \ldots, N\}$.

2) Reward

To solve the optimization model (6.15), we construct a reward interpreter of DQN by putting the voltage constraints into the objective function. The proposed reward interpreter has the following characteristics:

- If the constraint (6.15c) is violated, a significant penalty M is imposed on the reward of the DQN, which moves the voltages into the normal range [109]. To accurately quantize the voltage deviation degree in the whole distribution system, the reward function at all the time steps is calculated by

$$r^t = -M \sum_{\varphi} \sum_{k} [\max(|U_k^\varphi| - 1.05, 0) \ + \max(0.95 - |U_k^\varphi|, 0) \ ] \tag{6.17}$$

  where the more significant the degree of voltage violation is, the more negative reward the DQN agents obtain.

Remark. The primary goal of VVO is to make all voltage magnitudes within a normal operating range. Otherwise, the VVO problem has no feasible solution. Therefore, before calculating the power loss reduction, the proposed reward function firstly examines whether the current actions lead to a voltage violation. The agents must avoid providing the actions that result in the voltage violation as much as possible.

- When the voltage constraint (6.15c) is not violated, we calculate the reward value at episode t by

$$r^t = M^+(P_{loss,0}^t - P_{loss}^t) \tag{6.18}$$

where P_{loss}^{t} denotes the active power loss based on the current action variables, and $P_{loss,0}^{t}$ denotes the one that takes the default actions at episode t, both of which are calculated according to the corresponding states; $M^{+} > 0$ denotes an incentive factor that motivates the agents to obtain a greater positive reward; $r^{t} > 0$ implies that the proposed DQN further reduces the power loss after conducting the new action given by the agents at episode t.

Remark. In the case of no voltage violation, we set the reduction of power losses compared with the default settings of the regulation devices as the reward, rather than the absolute power loss at the current episode. Doing so is to guarantee that the agents always attempt to improve their policy based on the default actions, when facing new operational conditions. Moreover, this design coordinates with voltage constraints in (6.15c) and make agents learn in the direction of obtaining a greater positive reward. It is because either the voltage violation or increasing power loss has a negative reward, which informs the agents that the actions being taken are undesirable.

6.3.3. Multi-agent DQN Algorithm

This section describes the learning process of the proposed MADRL algorithm. The NN structure of agent j is depicted in Fig. 6.2, where the input is state $\boldsymbol{s}$, and the output is the action to be taken by this agent. The NN-based agent is trained to learn the mapping between the state and action, *i.e.*, an action-reward Q function, by maximizing the possible reward.

1) Multi-agent Operation

MADRL is capable of solving more complex tasks through the cooperation of individual agents effectively. In our VVO framework, the agents' training process can be summarized as "centralized learning, decentralized execution", which belongs to the concurrent learning category in multi-agent training schemes and is computationally efficient [122]. Moreover, the multi-agent DQN distributes the global control actions to each agent that performs coordinated RL by exploring the shared environment. Specifically, each agent takes actions

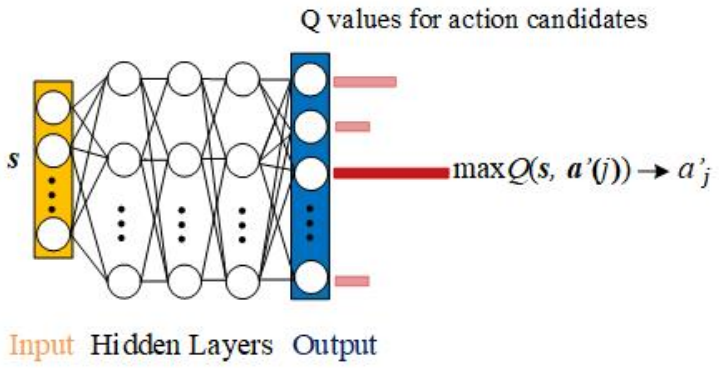

Figure 6.2. The NN structure for agent j.

based on their local observations, and the action choice is evaluated by the overall Q value of the combined actions from all agents under a specific state.

Here, let j denote the serial number of an agent, and $j \in \{1, 2, \ldots, n\}$, where n is the number of agents. Considering various regulating devices, the actions of all the devices are assigned to multiple agents, shown in Fig. 6.3; assume there are N_j actions in agent j. Due to the existence of heterogeneous regulation devices, all agents can perform in parallel by distributing the actions' searching space as uniformly as possible, *i.e.*, making the dimension of the joint space in each agent close to each other. For example, we can assign one action of a per-phase 32-step regulator to agent 1 and five actions from multiple capacitors to agent 2, and then $\dim(\mathcal{A}_1) = 33$ and $\dim\left(\prod_{i=2}^{6} \mathcal{A}_i\right) = 2^5 = 32$. It is noted that the agents require no physical information of these actions, such as their locations and phases in the system. When sent to the environment, these actions are executed at the corresponding locations for effective power flow calculation.

During the learning process, efficient communication among these agents is conducted to select the optimal actions via their shared observation of the current state $\boldsymbol{s}$ and the latest action $\boldsymbol{a}$, shown as Fig. 6.3. Moreover, the information exchange among the agents is the current actions that these agents jointly take, $\boldsymbol{a}$. In each training episode, based on the current state $\boldsymbol{s}$, the agents provide the new control actions to the environment. Specifically, in agent j, the new action participating in the local learning process is $\boldsymbol{a}'(j) =$

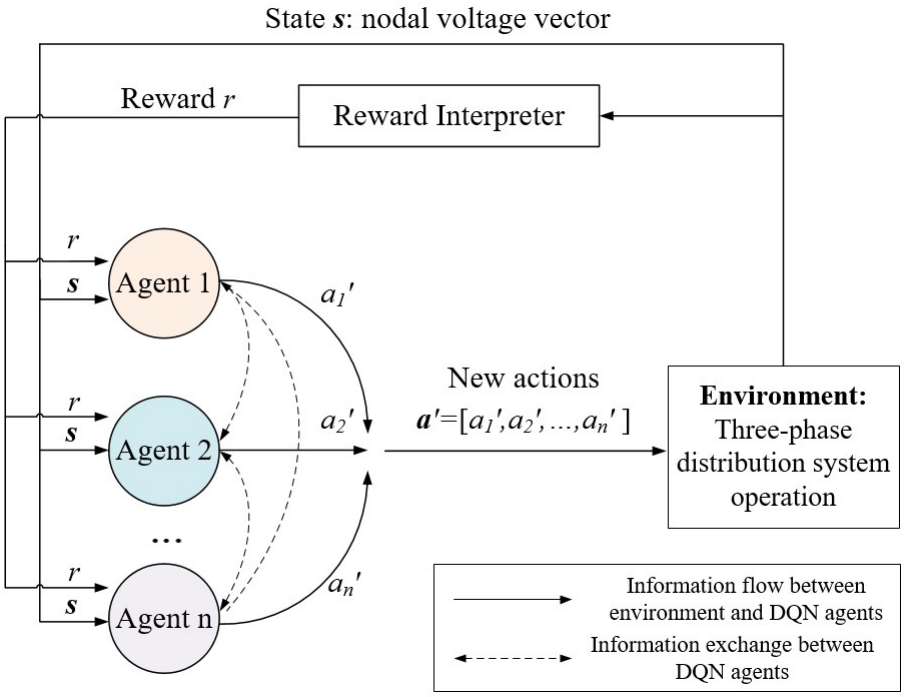

Figure 6.3. Multi-agent DQN training loop

$[a_1, \ldots, a_{j-1}, a'_j, a_{j+1}, \ldots, a_n]^T$, where a'_j is the latest control policy from this agent, and the rest of the actions a_i, $i \in \{1, 2, \ldots, n\} \backslash j$, are reserved from the last episode. Then, collect the control action a'_j from $\boldsymbol{a}'(j)$ in all these agents, and form the new action vector for the whole distribution system, $\boldsymbol{a}' = [a'_1, a'_2, \ldots, a'_n]^T$.

The environment, *i.e.*, the power flow calculation procedure, then implements the joint action $\boldsymbol{a}'$ and get a new reward and a new state $\boldsymbol{s}'$, until the training process terminates.

2) Offline Training and Online Test

According to our customized reward function (6.17) and (6.18), the new state $\boldsymbol{s}'$ and the corresponding system power loss are interpreted into the immediate reward r after taking action $\boldsymbol{a}$ at state $\boldsymbol{s}$. For $t = 1, 2, \ldots, N_{ep}$, agent j updates the action-reward Q function at episode t via the following Bellman equation:

$$Q_j^{t+1}(\boldsymbol{s}, \boldsymbol{a}) = Q_j^t(\boldsymbol{s}, \boldsymbol{a}) + \alpha(r^t + \gamma \max Q_j^t(\boldsymbol{s}', \boldsymbol{a}') - Q_j^t(\boldsymbol{s}, \boldsymbol{a})) \tag{6.19}$$

where $\gamma \in [0,1]$ is a discount rate, and α denotes the learning rate of the DQN.

An *experience replay* technique is used to store the latest N_b sets of the agents' experience in episode t, *i.e.*, the transition tuple $(\boldsymbol{s}, \boldsymbol{a}, r^t, \boldsymbol{s'})$, to a replay buffer $\mathcal{M}$. We sample a mini-batch memory $\mathcal{D}$ from the replay buffer to improve the generality of the agents towards diverse states. The agent j is trained by $\mathcal{D}$ together with the current transition tuple. Moreover, the stochastic gradient descent on NN parameters $\boldsymbol{\theta}_j$ for the agent is conducted using the following loss function $\mathcal{L}_j(\boldsymbol{\theta}_j)$, which enforces the Bellman equation (6.19):

$$\mathcal{L}_j\left(\boldsymbol{\theta}_j\right) = \mathbb{E}\left[\left(r^t + \gamma \text{max} Q_j^t(\boldsymbol{s'}, \boldsymbol{a'}) - Q_j^t(\boldsymbol{s}, \boldsymbol{a})\right)^2\right] \tag{6.20}$$

where we define the target Q function as $y{=}r^t + \gamma \cdot \text{max} Q_j^t(\boldsymbol{s'}, \boldsymbol{a'})$, and $\mathbb{E}(\cdot)$ denotes the expectation function.

The Q function iteratively updates following (6.19), shown as $\boldsymbol{Q}_j^{t+1}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}(r^t + \gamma \text{max} Q_j^t(\boldsymbol{s'}, \boldsymbol{a'}))$. Such iterations converge to the optimal action-value function, $Q_j^t \rightarrow Q_j^*$ as $t \rightarrow \infty$ [129].

During the training process, we apply the ε-greedy policy [120] to select the actions efficiently, as it encourages each agent to fully explore the corresponding action space. Specifically, as the training continues, the action selection relies more on the action policy from $Q_j^t(\boldsymbol{s'}, \boldsymbol{a'})$, shown as:

$$a_j = \pi_j(\boldsymbol{s}) = \begin{cases} \text{random action from } \mathcal{A}_j, & \text{if } \xi < \varepsilon_t \\ \arg\max_{a'_j \in \mathcal{A}_j} Q_j^t(\boldsymbol{s'}, \boldsymbol{a'}), \text{otherwise} \end{cases} \tag{6.21}$$

where π_j denotes the action selection policy for agent j, and $0 < \xi < 1$ is a random number; the searching criteria ε_t is updated based on the last episode by a decay factor η, *i.e.*, $\varepsilon_t = \varepsilon_{t-1}\eta$.

The pseudo-code summarizes the offline training process of the proposed MADRL algorithm. When the training process terminates, the agent j with parameter $\boldsymbol{\theta}_j^*$ is applied to

the test cases, where new operating conditions in the distribution system are fed. For each test case, these well-trained agents provide the action policy by

$$a_j = Q_j^* \left(\boldsymbol{s}, \boldsymbol{a}; \boldsymbol{\theta}_j^* \right) \qquad j \in \{1, 2, \ldots, n\} \tag{6.22}$$

These actions from all agents are combined by (6.16) and given to the environment as the solution of the model (6.15) for online system VVO control.

Algorithm 3 Multi-agent DQN Training Process

Input:Distribution system model and the action space $\mathcal{A}_j$ for agent j, $j \in \{1, 2, \ldots, n\}$.
Initialization: the learning rate α, the discount rate γ, the decay factor η, and the size of replay buffer N_b.
for $t = 1$ to N_{ep} **do**
  Initialize state s, and obtain action a by the ε -greedy policy (6.21).
  for $j = 1$ to n **do**
    Get reward r^t by (6.17) and (6.18), and new state s' by power flow calculation, and store them as a transition $(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{r}^t, \boldsymbol{s}')$ in replay buffer.
    Get the current Q vector at state s by agent j
    Sample from the replay buffer to obtain tuple $\left(\boldsymbol{s}(i), \boldsymbol{a}(i), \boldsymbol{r}^{\boldsymbol{t}}(i), \boldsymbol{s}'(i)\right)$, and $i = 1, 2, \ldots, N_d$.
    Set $y = r^t + \gamma \max Q_j^t(\boldsymbol{s'}, \boldsymbol{a'})$
    Train and update agent j by performing gradient descent on (20).
  end for
end for
Output:All agents with parameters $\boldsymbol{\theta}_j^*$.

6.4. Case Study

We test the proposed algorithm on the radial three-phase unbalanced IEEE 13-bus and 123-bus distribution systems [25]. We modify the 13-bus system by adding two single-phase PV units at buses 675 and 684, and a three-phase wind turbine generator at bus

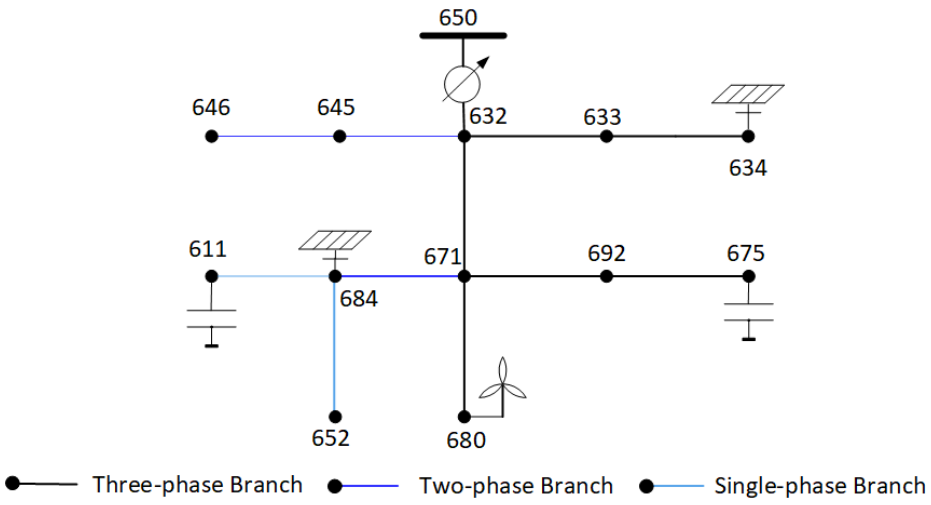

Figure 6.4. Three-phase unbalanced 13-bus distribution system

680, illustrated in Fig. 6.4. Six DG units are added at buses 13, 18, 60, 151, 250, and 300 in the 123-bus systems, and the installation capacity of these DG units is set as 300 or 500 kVA [130]. Here we randomly place the DG units, and similar settings are also used in [130] and [15]. The optimal placement method of DGs can be found in [131] and is not the focus of this chapter. Table 6.2 lists the specifications of control devices, including smart inverters installed at the DG units in both test systems. We can see that taking the action variables for all these devices into account results in high-dimensional joint decision space. We assume that there is no distinction on the per-phase action of regulators located on the same bus, and this assumption also applies on the smart inverters. For example, in the 13-bus system, all six actions produce a $2^4 \times 21 \times 33 = 11,088$ combinations of all these actions; for the 123-bus test system, and the dimension of the combinations of all 13 actions is $2^6 \times 5^3 \times 5^4 = 5,000,000$. The details of the adopted multiple agents can be found in Table 6.3.

In the baseline, the actions are initially set as default to calculate (6.18), where no capacitor bank is connected in these systems, and the steps of AVRs and the power factors of smart inverters are set as 1 as default [132]. The proposed algorithm runs in MATLAB

Table 6.2. Installation settings of control devices

Device Type	13-bus System		123-bus System	
	No.bus/branch	Phase	No.bus/branch	Phase
Regulators	650-632	A,B,C	150-149	A,B,C
			9-14	A
			25-26	A,C
			60-67	A,B,C
Capacitors	611	C	83	A,B,C
			88	A
	675	A,B,C	90	B
			92	C
Smart Inverters	684	C	13	C
			18	A,B,C
			16	A,B,C

by using the Reinforcement Learning Toolbox, and we build the learning environment, *i.e.*, three-phase distribution system operation. The load data are collected at each node from smart meters in the test systems, according to [116] and [115]. For offline training and online test, we randomly generate 9,000 and 13,000 operating conditions (episodes) with 80% to 120% of random fluctuations of base loads/DGs in these two systems, respectively [120]. Moreover, in the test phase, new operating conditions are used as test cases, and Table 6.3 gives the number of training episodes and test cases, N_{ep} and N_{test}.

DQN Specification: The NNs used here have three fully connected layers, and the learning rate is chosen as 0.0001. The NN agents in the proposed method use rectified linear unit (ReLU) activation functions in the hidden layers and sigmoid functions in the output layer. Table 6.4 also summarizes the parameters of the adopted multi-agent DQN in the two test systems.

6.4.1. Learning Performance

We investigate the learning performance of the proposed method. We test the proposed

Table 6.3. Parameter settings of multi-agent DQN

Parameter	13-bus System	123-bus system
Size of hidden layers	{35,20,30}	{120,60,30}
η	0.999	0.9995
γ	0.95	0.95
N_d	500	2000
N_{ep}	8000	9000
N_{test}	1000	4000
n	3	6
N_j for $j \in \{1, 2, \ldots, n\}$	$N_1 = 16$ $N_2 = 21$ $N_3 = 33$	$N_1 = 64$ $N_2 \sim N_4 = 5$ $N_5, N_6 = 25$
M^+	1	10
M	100	100

algorithm on the 123-bus system, and Fig.6.5 shows the reward values in the training process and moving average rewards in successive 200 episodes. We observe that the DQN agents' control policies result in negative rewards ($r^t < 0$) due to limited positive learning experiences and not yet optimized action policies in an earlier learning phase. These negative rewards illustrate that at the beginning, the action policies are incapable of maintaining the system voltages within 0.95∼1.05 p.u. and reducing power loss at the same time, according to (6.17) and (6.18). However, during the training process, the agents gradually evolve and obtain positive rewards ($r^t > 0$) more and more frequently. Moreover, $r^t > 0$ implies that there is no voltage violation and the power loss is further reduced by taking action $\boldsymbol{a}'$, compared with the performance before $\boldsymbol{a}'$ is taken. Also, the average reward in the training process continuously increases, which shows the DQN's ability in realizing the VVO.

Furthermore, for online test, 4,000 new cases are fed to these trained agents. These agents demonstrate an effective control performance for VVO, which is characterized by the positive rewards in these cases, as shown in episodes from 9,000 to 13,000 of Fig. 6.5. We conclude that the proposed DQN enables the power grid to self-learn with the "cognitive"

function of VVO control by mimicking the human mind. Eventually, these trained agents can implement effective control policies when confronted with new operating conditions.

Table 6.4. Performance of voltage regulation

System Scale	N_v	M_m	S_v
13-bus System	998	1,000	99.8%
123-bus System	4,000	4,000	100%

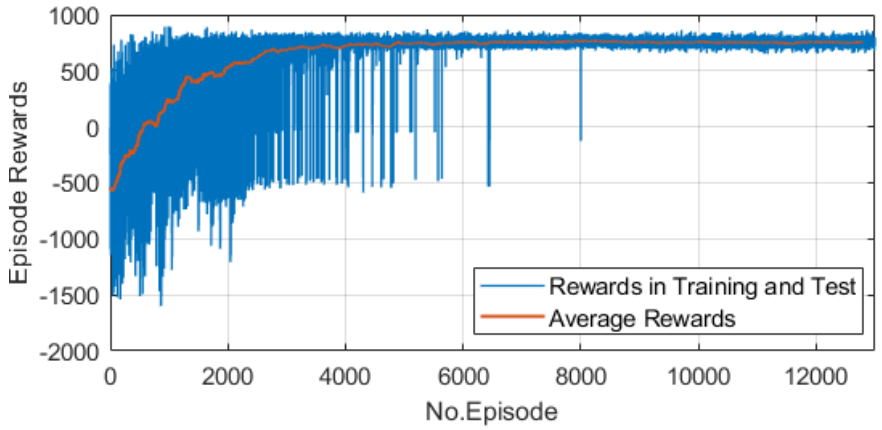

Figure 6.5. Multi-agent DQN training process in the IEEE 123-bus system

6.4.2. VVO Performance

This section demonstrates the VVO performance of the proposed model-free MADRL method when facing random operating conditions, in terms of voltage regulation and power loss reduction, both of which are implemented online by this method simultaneously.

1) Voltage Regulation

To evaluate the voltage control performance of the proposed VVO method in the test cases, we define the success rate in voltage regulation, S_v, as follows.

$$S_v = N_v / M_v \tag{6.23}$$

where M_v denotes the number of those test cases that exist voltage violation before adopting the proposed VVO algorithm, and N_v denotes the number of those cases that avoid the voltage violation issue by using this method. Moreover,a higher S_v illustrates that the proposed algorithm has a better control performance in voltage regulation.

Table 6.4 summarizes the statistical results of the proposed algorithm in test cases. Specifically, before employing the proposed method in the 13-bus system, voltage violation exists in the 1,000 test cases, and our algorithm achieves a control success rate of 99.80% in voltage regulation. For a larger-scale three-phase 123-bus system, the VVO task becomes more challenging due to the so-called "curse of dimensionality." However, the proposed MADRL method obtains a success rate of 100%.

We randomly select a test case in the 13-bus system to compare the voltage magnitudes with the control from the proposed method and the baseline, and Fig. 6.6 depicts these three-phase voltage magnitudes at each bus. It can be seen that in the baseline control, the A-phase and C-phase voltage magnitudes at buses 611, 652, 671, 675, and 684 violate the normal voltage operation limits; after adopting the control actions provided by the agents, the voltages at all these buses fall within the normal operating range. Moreover, we compare the voltage profiles with the control from the baseline and the proposed VVO method, in terms of the minimum, maximum, and average of voltage magnitudes, in Table 6.5. It is shown that the proposed method avoids voltage violation by jointly dispatching these devices, and these DG units with smart inverters are directed to provide reactive powers for voltage lift.

2) Power Loss Reduction

Here we show the performance of power loss reduction in those test cases that are discussed above for voltage regulation. Fig. 6.7 demonstrates the power loss of the baseline and the proposed approach in 50 test cases that are randomly selected in the 13-bus system. We conclude that the proposed method enables effective power loss reduction.

To further quantify the loss reduction performance in these two test systems, we calculate

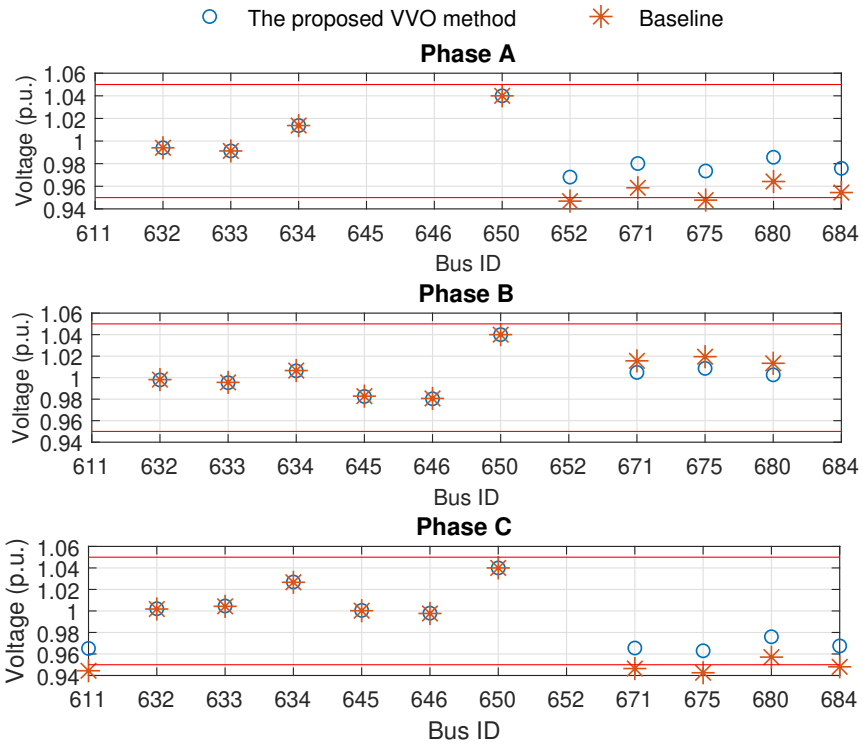

Figure 6.6. Three-phase voltage magnitude comparison before and after VVO control in the 13-bus system.

the reduction of power loss in each test case by

$$\Delta P^i = P^i_{loss,0} - P^i_{loss} \tag{6.24}$$

where ΔP^i denotes the difference in the ith test case between the active power losses with the control strategies from the baseline and the proposed method, *i.e.*, $P^i_{loss,0}$ and P^i_{loss}; $i = 1,\ 2, \ldots, N_{test}$, and these active power losses are calculated by (6.8).

Table 6.6 summarizes the average and maximum of ΔP^i in two test systems and implies that the proposed method enables power loss reduction in all test cases. Specifically, in the 13-bus system, the average power loss reduction obtained by the proposed algorithm is 45.38 kW. In the 123-bus system, these statistic data are 75.23 kW for the mean power loss reduction and 28.38% for the loss reduction percentage. We conclude that the pro-

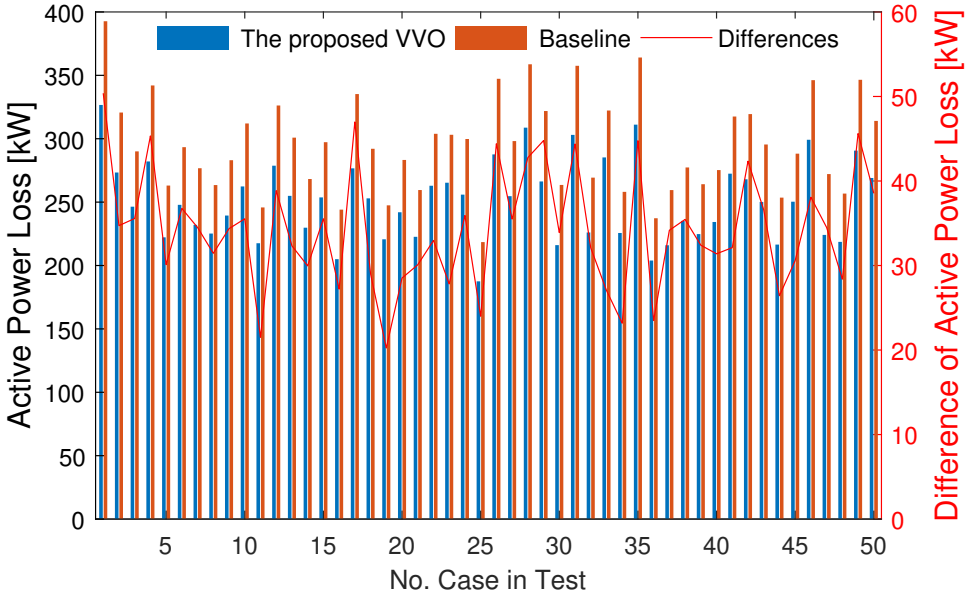

Figure 6.7. Power loss comparison before and after control in the 13-bus system.

posed DQN-based method effectively realizes dual goals on power loss reduction and voltage regulation.

6.4.3. Computation Time

We carry out numerical experiments to investigate the computational efficiency of the proposed method. All the tests are performed using MATLAB on the machine equipped

Table 6.5. Voltage comparison in the IEEE 123-bus system

Voltage Magnitude Profile [p.u.]		Phase A	Phase B	Phase C
Baseline	Min. Voltage	0.9070	0.9115	0.9191
	Max. Voltage	1.0300	1.0300	1.0300
	Aver. Voltage	0.9372	0.9526	0.9552
The proposed method	Min. Voltage	0.9710	0.9958	0.9819
	Max. Voltage	1.0393	1.0418	1.0421
	Aver. Voltage	1.0002	1.0256	1.0213

Table 6.6. Performance of power loss reduction

System Scale	Average ΔP^i [kW]	Maximum ΔP^i[kW]	The prop. of $\Delta P^i > 0$
13-bus System	45.38	67.79	100%
123-bus System	75.23	86.36	100%

with a 2.5 GHz Intel Core i5 CPU and 8 GB of RAM.

In the online test phase, the average executive time of all test cases in the 13-bus and 123-bus systems is 21.7 and 46.2 milliseconds, respectively, which is promising to meet with the requirement of real-time implementation in power systems. The proposed algorithm still shows the high computational efficiency in the unbalanced 123-bus system. Moreover, the proposed MADRL method is competitive when dealing with a high-dimensional action space that exists in three-phase distribution systems.

6.4.4. Comparison with Existing Reward Design

Besides the baseline, where all regulation devices keep their default actions, another benchmark is set up by using the reward design proposed in [116], which does not consider voltage regulation as the primary objective of VVO and directly uses the big penalty. For a fair comparison in the DRL-based VVO methods, we adopt the following reward function in this benchmark and apply the same environment and training conditions as the proposed method.

$$r^t = -P^t_{loss} - M[\text{count}(\ |U^{\varphi}_k| > 1.05 \text{ or } |U^{\varphi}_k| < 0.95)] \tag{6.25}$$

where P^t_{loss} denotes the power loss, and the function count$(\cdot)$ is defined as the number of per-phase voltage violations across all nodes in episode t.

Table 6.7. Comparison of voltage regulation in the 13-bus system

Reward Function	M_v in each experiment	S_v	
		Aver.	Min
(6.17) and (6.18) in the proposed method	1,000	99.40%	99.76%
(6.25)	1,000	90.02%	77.10%

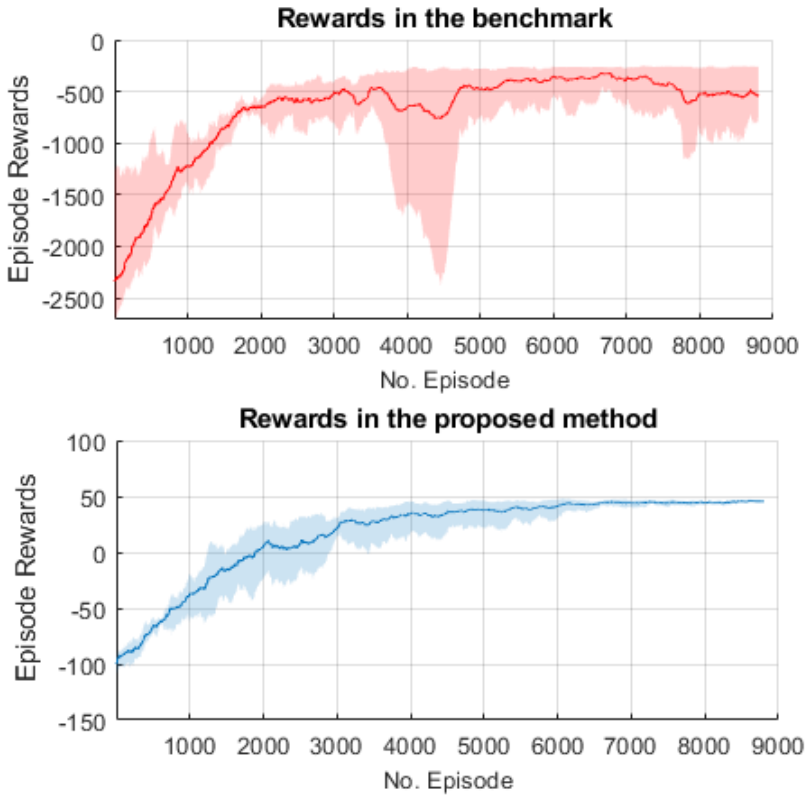

Figure 6.8. The comparison of the learning performance for VVO

Table 6.7 compares the online test performances of both of the reward functions in the multi-agent DRL approach. It is shown that with no consideration of the priority of voltage violation in the VVO problem, the reward function in (6.25) may fail to avoid voltage violation, and the average success rate in voltage regulation is only 90.02%. In contrast, the proposed method has a 99.76% average success rate.

Fig. 6.8 depicts the learning performance of these two reward design schemes in five

random experiments [115]. The dark-colored curves denote the average performances of rewards in all these experiments, and the light-colored regions are the error bounds. We conclude that with no consideration of the priority of voltage violation in the reward design, the reward function used in (6.25) results in unstable learning performance for VVO.

6.5. Conclusion

This chapter proposes a novel and real-time DQN-based VVO algorithm in unbalanced distribution systems. Integrating the voltage-dependent loads, DG penetration, and three types of voltage regulating devices into distribution system operation, we establish the efficient power flow calculation as the environment of the DQN. Via the interaction between the environment and multiple agents, the proposed VVO method adaptively chooses control actions to enable voltage regulation and power loss reduction. This algorithm realizes a promising VVO performance in two unbalanced distribution systems.

Several state-of-the-art policy gradient methods, such as proximal policy optimization (PPO) [133], have the capability of further mitigating the scalability issue. On the other hand, the proposed MADRL method can use the distributed learning by dividing the whole distribution systems into multiple sub-areas to improve the control performance. The distributed control algorithms are already proposed in model-based voltage control methods such as [134], and our future work also combines these approaches with state-of-the-art DRL techniques for VVO.

Chapter 7

Conclusion and Future Work

7.1. Summary

To construct a next-generation smart grid, this dissertation explores novel methodologies for the design, modeling, and application of situational awareness for distribution system monitoring and control. The proposed situational awareness framework consists of a series of promising solutions in distribution system state estimation (DSSE), cyberattack detection, and voltage control. These solutions cover two aspects – methodological innovations and interdisciplinary development using the state-of-the-art of applied mathematics and machine learning, and they are summarized below:

- **Methodology Innovation for Renewable Integration**: We explore efficient mathematical analytic models and methodologies to address the uncertainty issue from high penetration of distributed energy resources (DERs) in practical distribution system operation. Chapter 2 proposes a highly efficient DSSE algorithm using the Taylor series of complex numbers and interval arithmetic techniques. Furthermore, We discuss the impacts of DERs on distribution system operation and the subsequent challenges. Chapter 3 establishes a general interval state estimation (ISE) framework that considers imprecise line parameters, measurements with noises, and uncertain DER outputs simultaneously. A modified Krawczyk-operator algorithm is proposed to solve these ISE models and obtain the upper and lower bounds of state variables for better monitoring systems under the coordinated impacts of these multiple uncertainties. Case studies illustrate the proposed DSSE methods are capable of overcoming the ever-present poor observability issue in the distribution management system (DMS) due to lack of meters and stochastic outputs of renewables, and works conveniently with limited phasor

measurement units (PMUs) in three-phase unbalanced distribution systems. Chapter 4 applies a distributed DSSE algorithm to faulted-line identification by integrating PMU data. This method running in a distributed manner has lower computation cost and enables fast faulted line identification within several tens of milliseconds. Our approach considers the impact of distributed generation (DG) penetration on distribution system operation, and its location performance is independent of fault types and fault impedances. Furthermore, the proposed algorithm is robust against high-level noises in measurements.

- **Knowledge Integration Through Interdisciplinary Studies**: We expect to bridge the gap between modern power system research and state-of-the-art communication, machine learning (ML), and artificial intelligence (AI) techniques by using measurement data from the supervisory control and data acquisition (SCADA) system, PMUs, and advanced metering infrastructure (AMI) systems. The cyclic tendency of power system operation exists, and the strong daily, weekly, and seasonal cycles ought to be used for data mining and algorithmic training. As novel endeavors, we apply DSSE to data-driven cyberattack detection via deep semi-supervised learning and generative models in Chapter 5 and volt-VAR optimization (VVO) via multi-agent deep reinforcement learning, represented by deep Q-networks, in Chapter 6. We show the proposed data-driven situational awareness technologies, combining sophistication and flexibility, are highly applicable to practical systems and have good potentials to improve operation efficiency in real-world power grids. Moreover, in the proposed DRL-based control scheme for voltage regulation, dispatching DG outputs via smart inverters synergizes with power loss reduction.

7.2. Technical Prospect and Outlook

Upon this dissertation, a hybrid model-based and data-driven situational awareness framework is proposed. As far as methodology is concerned, the proposed ISE algorithms can outperform the weighted least square (WLS)-based state estimation that is dominate for a long time, in terms of the computational efficiency and superiority in capturing the

effects of stochasticity and variability of renewable resources on DSSE. This method can be further extended to other DSSE-based operational scenarios. Meanwhile, data-driven and learning-based techniques are expected to upgrade the current monitoring and control paradigm of distribution systems in broader DSSE applications. Taking the DER uncertainty into account, distribution system monitoring and control still demand more plausible and interpretable solutions. We discuss briefly future works in the following aspects, which serve robust and reliable system operation facing more practical and harsh conditions:

- **Topology identification**. Topology identification in distribution networks is to determine the connectivity among nodes in the case of unknown topology changes due to network reconfiguration, repairs, maintenance, and load balancing. In practice, distribution system operators (DSOs) might have partial or no knowledge of basic network topologies, especially on the connectivity of DERs in secondary distribution networks since the utilities may not own them. Furthermore, there are many ad-hoc connections of plug-and-play DER components, which makes the problem even more complicated. Some existing works in the literature, such as [135], formulate this problem as a mixed-integer programming problem, and however, how to leverage abundant historical data from micro-PMUs or smart meters to identify the topology structure by voltage correlation analysis awaits further investigation.

- **Network parameter calibration**. The assumption that network parameters are accurate is a prerequisite in the conventional DSSE procedure. However, it is rarely the case in practice, since the network parameters change in field ambient and aging wirings. Higher penetrations of DERs call for improvement in the accuracy of those varying parameters for accurate system monitoring and other line model-based applications. Therefore, parameter estimation procedures are designed to calibrate these network parameters, also for avoiding manual inspections that are hard to carry out in densely populated urban areas.

- **State recovery with missing data**. Missing measurement data at key buses, which can be caused by various reasons such as sensor failures and unreliable communication, are extremely possible to result in poor observability and even unobservability of system

states. In this case, the existing DSSE methods fail to provide effective states for system monitoring and control. Hence, reliable system operation demands an effective solution to state recovery with missing data.

- **Dynamic DSSE**. With the increasing penetration of DERs, responsive loads, and microgrids, distribution systems are subject to different types of dynamics such as sudden load shedding. The currently used steady-state power system models demand further update to capture the system dynamics in a short time-frame [136]. Computationally efficient and flexible dynamic state estimation approaches still urgently need to be extended to distribution systems with high penetration of DERs.

- **Data synergy and fusion of heterogeneous data**. Distribution systems are undergoing a rapid growth of advanced metering infrastructure, especially the development of PMUs and popularization of smart meters. Nevertheless, this leads to a complicated scenario where a set of diverse sensors feed the DSSE with heterogeneous data at different sampling and update frequencies. Furthermore, only a handful of efforts are working to implement adaptive synchronization of multi-rate data for situational awareness. There is an expectation about how the lagging SCADA data automatically follow the faster updating PMU measurements, and it still remains an open issue.

Chapter 8

Publications

Published Journal Papers

- **Y. Zhang**, J. Wang, and B. Chen, Detecting false data injection attacks in smart grids: a semi-supervised deep learning approach, *IEEE Transactions on Smart Grid*, in press.
- **Y. Zhang**, X. Wang, J. Wang, and Y. Zhang, Deep reinforcement learning based volt-VAR optimization in smart distribution systems, *IEEE Transactions on Smart Grid*, in press.
- **Y. Zhang**, J. Wang, and M. Khodayar, Graph-based faulted line identification using micro-PMU data in distribution systems, *IEEE Transactions on Smart Grid*, in press.
- **Y. Zhang**, and J. Wang, Towards highly efficient state estimation with nonlinear measurements in distribution systems, *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 2471-2474, May 2020.
- **Y. Zhang**, J. Wang, and Z. Li, Interval state estimation with uncertainty of distributed generation and line parameters in unbalanced distribution systems, *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 762 -772, Jan. 2020.
- **Y. Zhang**, J. Wang, and J. Liu, Attack identification and correction for PMU GPS spoofing in unbalanced distribution systems, *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 762-773, Jan. 2020.
- **Y. Zhang**, J. Wang, and Z. Li, Uncertainty modeling of distributed energy resources: techniques and challenges, *Current Sustainable/Renewable Energy Report*, vol. 6, no. 2, pp. 42–51, June 2019.

- M. Cui, M. Khodayar, C. Chen, X. Wang, and **Y. Zhang**, Deep learning based time-varying parameter identification for system-wide load modeling, *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6102-6114, Nov. 2019.

Submitted Journal Papers

- Y. Chen, Y. Y, **Y. Zhang**, and J. Wang, Robust SOCP-based state estimation in integrated electricity-heat systems, *IEEE Transactions on Smart Grid*. (Under the 2nd review)

- T. Zhao, **Y. Zhang**, and J. Wang, Online convex optimization-based real-time distributed control in reconfigurable distribution networks, *IEEE Transactions on Power Systems*. (Under the 1st review)

Published Conference Papers

- **Y. Zhang**, J. Wang, and Z. Li, Interval state estimation with measurement and network parameter uncertainty in unbalanced distribution systems. *2019 IEEE Power Engineering Society General Meeting*, GA, Atlanta, pp. 1-5.

Appendix A

PMU Latency for Faulted Line Identification

The proposed method in Chapter 4 leverages post-fault phasors from PMUs. Fig. A.1 shows the diagram of a waveform for fault currents and the latency of the proposed method for faulted line identification. According to [68], considering the duration of transients, *i.e.*, T_1, there is a short latency between 0 and 20 ms before measuring the steady-state phasors by PMUs. Later, to obtain accurate post-fault synchrophasors, the discrete Fourier transform (DFT) method is used to process a dataset of raw-sampled waveforms [55], [137]. The time window T_2 for PMUs to get post-fault phasors is about two to three periods derived by a fundamental frequency, *e.g.*, 30 ms. Moreover, according to IEEE Standard C37.118, the shortest length of the observation window can reach up to 17 ms [52].

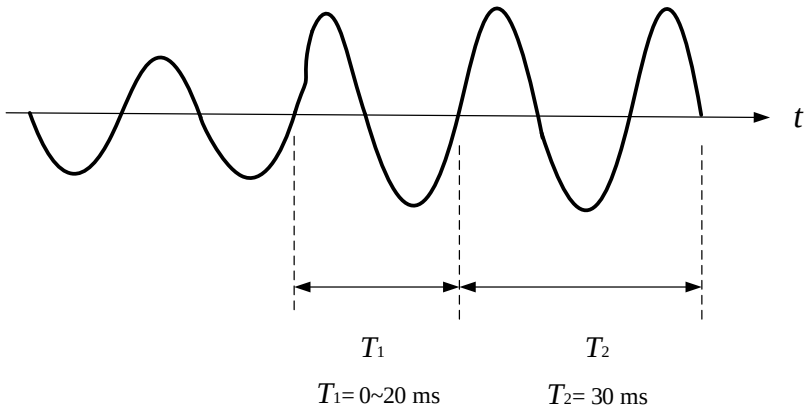

Figure A.1. The diagram of a waveform of fault current and the latency of the proposed method in identifying a fault.

BIBLIOGRAPHY

[1] M. Amini and M. Almassalkhi, "Trading off robustness and performance in receding horizon control with uncertain energy resources," in *2018 Power Systems Computation Conference (PSCC)*, pp. 1–7, 2018.

[2] H. Jiayi, J. Chuanwen, and X. Rong, "A review on distributed energy resources and microgrid," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 9, pp. 2472–2483, 2008.

[3] "Distributed energy resources technical considerations for the bulk power system," 2018.

[4] W. El-Khattam and M. M. Salama, "Distributed generation technologies, definitions and benefits," *Electric power systems research*, vol. 71, no. 2, pp. 119–128, 2004.

[5] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep reinforcement learning based volt-var optimization in smart distribution systems," *arXiv preprint arXiv:2003.03681*, 2020.

[6] Y. Zhang, J. Wang, and M. E. Khodayar, "Graph-based faulted line identification using micro-PMU data in distribution systems," *IEEE Transactions on Smart Grid*, pp. 1–1, 2020.

[7] "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid — WIRED. [Online]. Available: https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/," 18-Jul-2019.

[8] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, pp. 1–33, 2011.

[9] Y. Zhang, J. Wang, and J. Liu, "Attack identification and correction for PMU gps spoofing in unbalanced distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 762–773, 2020.

[10] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2019.

[11] R. Singh, E. Manitsas, B. C. Pal, and G. Strbac, "A recursive bayesian approach for identification of network configuration changes in distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1329–1336, 2010.

[12] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5113–5122, 2018.

[13] C. N. Lu, J. H. Teng, and W. E. Liu, "Distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 10, pp. 229–240, Feb 1995.

[14] A. Primadianto and C. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, pp. 3875–3883, Sep. 2017.

[15] H. Wang, W. Zhang, and Y. Liu, "A robust measurement placement method for active distribution system state estimation considering network reconfiguration," *IEEE Transactions on Smart Grid*, vol. 9, pp. 2108–2117, May 2018.

[16] *Evaluation of measurement data—Supplement 1 to the 'guide to the expression of uncertainty in measurement'—Propagation of distributions using a Monte Carlo method*. JCGM YYY, 2006.

[17] A. K. Al-Othman and M. R. Irving, "A comparative study of two methods for uncertainty analysis in power system state estimation," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1181–1182, 2005.

[18] J. Xu, Z. Wu, X. Yu, Q. Hu, and X. Dou, "An interval arithmetic-based state estimation framework for power distribution networks," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 11, pp. 8509–8520, 2019.

[19] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

[20] Y. Zhang, J. Wang, and Z. Li, "Interval state estimation with uncertainty of distributed generation and line parameters in unbalanced distribution systems," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 762–772, 2020.

[21] W. H. Kersting, *Distribution System Modeling and Analysis*. CRC Press, 2006.

[22] L. Yu, D. Czarkowski, and F. de Leon, "Optimal distributed voltage regulation for secondary networks with dgs," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 959–967, 2012.

[23] F. Flanigan, *Complex Variables*. Dover Books on Mathematics, New York, NY, USA, 2010.

[24] A. Neumaier, *Interval Methods for Systems of Equations*. Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1991.

[25] "IEEE test feeder specifications," 2017. [Online]. available: http://sites.ieee.org/pes-testfeeders/resources.

[26] S. Wang, L. Han, and L. Wu, "Uncertainty tracing of distributed generations via complex affine arithmetic based unbalanced three-phase power flow," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3053–3062, 2015.

[27] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1845–1856, 2010.

[28] J. Liu, F. Ponci, A. Monti, C. Muscas, P. A. Pegoraro, and S. Sulis, "Optimal meter placement for robust measurement systems in active distribution grids," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 5, pp. 1096–1105, 2014.

[29] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, "Effects of measurements and pseudomeasurements correlation in distribution system state estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 12, pp. 2813–2823, 2014.

[30] Y. Zhang, J. Wang, and Z. Li, "Uncertainty modeling of distributed energy resources: techniques and challenges," *Current Sustainable/Renewable Energy Reports*, vol. 6, no. 2, pp. 42–51, 2019.

[31] P. Chen, S. Tao, X. Xiao, and L. Li, "Uncertainty level of voltage in distribution network: An analysis model with elastic net and application in storage configuration," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2563–2573, 2018.

[32] U. Kuhar, M. Pantoš, G. Kosec, and A. Švigelj, "The impact of model and measurement uncertainties on a state estimation in three-phase distribution networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3301–3310, 2019.

[33] B. Brinkmann and M. Negnevitsky, "A probabilistic approach to observability of distribution networks," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1169–1178, 2017.

[34] J. Liu, J. Tang, F. Ponci, A. Monti, C. Muscas, and P. A. Pegoraro, "Trade-offs in PMU deployment for state estimation in active distribution grids," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 915–924, 2012.

[35] B. Wang, G. He, K. Liu, H. Lv, W. Yin, and S. Mei, "Guaranteed state estimation of power system via interval constraints propagation," *IET Generation, Transmission Distribution*, vol. 7, no. 2, pp. 138–144, 2013.

[36] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng, "Deep learning-based interval state estimation of ac smart grids against sparse cyber attacks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4766–4778, 2018.

[37] J. Xu, Z. Wu, X. Dou, and Q. Hu, "An interval arithmetic-based state estimation for unbalanced active distribution networks," in *2017 IEEE Power Energy Society General Meeting*, pp. 1–5, 2017.

[38] C. Zhang, H. Chen, K. Shi, M. Qiu, D. Hua, and H. Ngan, "An interval power flow analysis through optimizing-scenarios method," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5217–5226, 2018.

[39] B. Hu, K. Xie, and H. Tai, "Inverse problem of power system reliability evaluation: Analytical model and solution method," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6569–6578, 2018.

[40] C. Rakpenthai, S. Uatrongjit, and S. Premrudeepreechacharn, "State estimation of power system considering network parameter uncertainty based on parametric interval linear systems," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 305–313, 2012.

[41] Z. Wu, H. Zhan, W. Gu, S. Zheng, and B. Li, "Interval state estimation of distribution network with power flow constraint," *IEEE Access*, vol. 6, pp. 40826–40835, 2018.

[42] Whei-Min Lin, Jen-Hao Teng, and Shi-Jaw Chen, "A highly efficient algorithm in treating current measurements for the branch-current-based distribution state estimation," *IEEE Transactions on Power Delivery*, vol. 16, no. 3, pp. 433–439, 2001.

[43] M. Pau, P. A. Pegoraro, and S. Sulis, "Efficient branch-current-based distribution system state estimation including synchronized measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 9, pp. 2419–2429, 2013.

[44] P. Zhuang, R. Deng, and H. Liang, "False data injection attacks against state estimation in multiphase and unbalanced smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6000–6013, 2019.

[45] A. Von Meier, D. Culler, A. McEachern, and R. Arghandeh, *Micro-synchrophasors for Distribution Systems*. IEEE, 2014.

[46] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 337–346, 2011.

[47] S. M. Rump, "Solving algebraic problems with high accuracy," in *A New Approach to Scientific Computation*, pp. 51–120, Elsevier, 1983.

[48] R. Krawczyk, "Newton-algorithmen zur bestimmung von nullstellen mit fehlerschranken," *Computing*, vol. 4, no. 3, pp. 187–201, 1969.

[49] R. E. Moore, "A test for existence of solutions to nonlinear systems," *SIAM Journal on Numerical Analysis*, vol. 14, no. 4, pp. 611–615, 1977.

[50] J. Rohn, "On overestimations produced by the interval gaussian algorithm," *Reliable Computing*, vol. 3, no. 4, pp. 363–368, 1997.

[51] "National Renewable Energy Laboratory (NREL): Measurement and Instrumentation Data Center," [Online]. Available: https:// midcdmz.nrel.gov/apps.

[52] "IEEE standard for synchrophasor measurements for power systems, IEEE Standard C37.118-2005," Dec. 2011.

[53] V. Murugesan, Y. Chakhchoukh, V. Vittal, G. T. Heydt, N. Logic, and S. Sturgill, "PMU data buffering for power system state estimators," *IEEE Power and Energy Technology Systems Journal*, vol. 2, no. 3, pp. 94–102, 2015.

[54] "Blackout Tracker: United States Annual Report 2018," [Online]. Availabe: https://switchon.eaton.com/plug/blackout-tracker.

[55] P. V. Navalkar and S. A. Soman, "Secure remote backup protection of transmission lines using synchrophasors," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 87–96, 2011.

[56] Y. Zhang, J. Liang, Z. Yun, and X. Dong, "A new fault-location algorithm for series-compensated double-circuit transmission lines based on the distributed parameter model," *IEEE Transactions on Power Delivery*, vol. 32, no. 6, pp. 2398–2407, 2017.

[57] G. Feng and A. Abur, "Fault location using wide-area measurements and sparse estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2938–2945, 2016.

[58] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2926–2936, 2017.

[59] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Anomaly detection using optimally placed μPMU sensors in distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3611–3623, 2018.

[60] J. Ren, S. S. Venkata, and E. Sortomme, "An accurate synchrophasor based fault location method for emerging distribution systems," *IEEE Transactions on Power Delivery*, vol. 29, no. 1, pp. 297–298, 2014.

[61] W. C. Santos, F. V. Lopes, N. S. D. Brito, and B. A. Souza, "High-impedance fault identification on distribution networks," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 23–32, 2017.

[62] R. Razzaghi, G. Lugrin, H. Manesh, C. Romero, M. Paolone, and F. Rachidi, "An efficient method based on the electromagnetic time reversal to locate faults in power networks," *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1663–1673, 2013.

[63] H. Liu, T. Jia, L. Mou, X. Zheng, Y. Zhang, and F. Zhang, "Improved traveling wave based fault location scheme for transmission lines," in *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, pp. 993–998, 2015.

[64] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, "Artificial neural network and support vector machine approach for locating faults in radial distribution systems," *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 710–721, 2005.

[65] J. Mora-Florez, V. Barrera-Nunez, and G. Carrillo-Caicedo, "Fault location in power distribution systems using a learning algorithm for multivariable data analysis," *IEEE Transactions on Power Delivery*, vol. 22, no. 3, pp. 1715–1721, 2007.

[66] K. Jia, T. Bi, Z. Ren, D. W. P. Thomas, and M. Sumner, "High frequency impedance based fault location in distribution system with dgs," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 807–816, 2018.

[67] M. A. Barik, A. Gargoom, M. A. Mahmud, M. E. Haque, H. Al-Khalidi, and A. M. Than Oo, "A decentralized fault detection technique for detecting single phase to ground faults in power distribution systems with resonant grounding," *IEEE Transactions on Power Delivery*, vol. 33, no. 5, pp. 2462–2473, 2018.

[68] 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*, vol. 32, no. 1, pp. 381–392, 2017.

[69] M. Majidi and M. Etezadi-Amoli, "A new fault location technique in smart distribution networks using synchronized/nonsynchronized measurements," *IEEE Transactions on Power Delivery*, vol. 33, no. 3, pp. 1358–1368, 2018.

[70] M. Farajollahi, A. Shahsavari, E. M. Stewart, and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using micro-PMU data," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6343–6354, 2018.

[71] M. Pau, F. Ponci, A. Monti, S. Sulis, C. Muscas, and P. A. Pegoraro, "An efficient and accurate solution for distribution system state estimation with multiarea architecture," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 5, pp. 910–919, 2017.

[72] Haibin Wang and N. N. Schulz, "A revised branch current-based distribution system state estimation algorithm and meter placement impact," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 207–213, 2004.

[73] S. Amini, F. Pasqualetti, M. Abbaszadeh, and H. Mohsenian-Rad, "Hierarchical location identification of destabilizing faults and attacks in power systems: A frequency-domain approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2036–2045, 2019.

[74] A. A. P. Bíscaro, R. A. F. Pereira, and J. R. S. Mantovani, "Optimal phasor measurement units placement for fault location on overhead electric power distribution feeders," in *2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T D-LA)*, pp. 37–43, 2010.

[75] J. Xu, Z. Wu, X. Yu, and C. Zhu, "Robust faulted line identification in power distribution networks via hybrid state estimator," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5365–5377, 2019.

[76] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, 2011.

[77] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting false data injection attacks in ac state estimation," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2476–2483, 2015.

[78] X. Liu and Z. Li, "False data attacks against ac state estimation with incomplete network information," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2239–2248, 2017.

[79] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

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

[81] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 4, pp. 370–379, 2014.

[82] M. Cui, M. Khodayar, C. Chen, X. Wang, Y. Zhang, and M. E. Khodayar, "Deep learning-based time-varying parameter identification for system-wide load modeling," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6102–6114, 2019.

[83] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.

[84] S. A. Foroutan and F. R. Salmasi, "Detection of false data injection attacks against state estimation in smart grids based on a mixture gaussian distribution learning method," *IET Cyber-Physical Systems: Theory Applications*, vol. 2, no. 4, pp. 161–171, 2017.

[85] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773–1786, 2016.

[86] J. J. Q. Yu, Y. Hou, and V. O. K. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, 2018.

[87] L. Yao and Z. Ge, "Scalable semisupervised GMM for big data quality prediction in multimode processes," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3681–3692, 2019.

[88] Q. Dai, L. Shi, and Y. Ni, "Risk assessment for cyberattack in active distribution systems considering the role of feeder automation," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3230–3240, 2019.

[89] R. Deng, P. Zhuang, and H. Liang, "False data injection attacks against state estimation in power distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2871–2881, 2019.

[90] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.

[91] S. Bolognani and S. Zampieri, "On the existence and linear approximation of the power flow solution in power distribution networks," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 163–172, 2016.

[92] J. Deng, X. Xu, Z. Zhang, S. Frühholz, and B. Schuller, "Semisupervised autoencoders for speech emotion recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 1, pp. 31–43, 2018.

[93] A. Radhakrishnan, K. Yang, M. Belkin, and C. Uhler, "Memorization in overparameterized autoencoders," *arXiv preprint arXiv:1810.10333*, 2018.

[94] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.

[95] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[96] Y. Zhang, J. Wang, and J. Liu, "Attack identification and correction for PMU gps spoofing in unbalanced distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 762–773, 2020.

[97] "Smart meters and smart meter systems: A metering industry perspective," March 2011. [Online]. available: https://aeic.org/smartmetersfinal032511/.

[98] "A comprehensive guide to fine-tuning deep learning models in keras (part i)," October 2016. [Online]. available: https://flyyufelix.github.io/2016/10/03/finetuning-in-keras-part1.html.

[99] L. Cai, N. F. Thornhill, S. Kuenzel, and B. C. Pal, "Wide-area monitoring of power systems using principal component analysis and k-nearest neighbor analysis," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4913–4923, 2018.

[100] X. Feng, W. Peterson, F. Yang, G. M. Wickramasekara, and J. Finney, "Smarter grids are more efficient: Voltage and Var optimization reduces energy losses and peak demands," tech. rep., Switzerland, Rev. Tech. Rep. 9AKK104295D7260, Mar 2009.

[101] Y. Ueda, K. Kurokawa, T. Tanabe, K. Kitamura, and H. Sugihara, "Analysis results of output power loss due to the grid voltage rise in grid-connected photovoltaic power generation systems," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 7, pp. 2744–2751, 2008.

[102] J. Li, Z. Xu, J. Zhao, and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 10, pp. 5519–5530, 2019.

[103] R. Anilkumar, G. Devriese, and A. K. Srivastava, "Voltage and reactive power control to maximize the energy savings in power distribution system with wind energy," *IEEE Transactions on Industry Applications*, vol. 54, no. 1, pp. 656–664, 2018.

[104] Z. Wang, J. Wang, B. Chen, M. M. Begovic, and Y. He, "Mpc-based voltage/Var optimization for distribution circuits with distributed generators and exponential load models," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2412–2420, 2014.

[105] B. A. De Souza and A. M. F. De Almeida, "Multiobjective optimization and fuzzy logic applied to planning of the volt/Var problem in distributions systems," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1274–1281, 2010.

[106] W. Sheng, K. Liu, S. Cheng, X. Meng, and W. Dai, "A trust region sqp method for coordinated voltage control in smart distribution grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 381–391, 2016.

[107] F. Capitanescu, I. Bilibin, and E. Romero Ramos, "A comprehensive centralized approach for voltage constraints management in active distribution grid," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 933–942, 2014.

[108] M. Chamana and B. H. Chowdhury, "Optimal voltage regulation of distribution networks with cascaded voltage regulators in the presence of high PV penetration," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1427–1436, 2018.

[109] I. Roytelman, B. K. Wee, and R. L. Lugtu, "Volt/Var control algorithm for modern distribution management system," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1454–1460, 1995.

[110] H. Ahmadi, J. R. Martí, and H. W. Dommel, "A framework for volt-VAR optimization in distribution systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1473–1483, 2015.

[111] C. Zhang, Y. Xu, Z. Dong, and J. Ravishankar, "Three-stage robust inverter-based voltage/Var control for distribution networks with high-level PV," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 782–793, 2019.

[112] M. H. K. Tushar and C. Assi, "Volt-var control through joint optimization of capacitor bank switching, renewable energy, and home appliances," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4077–4086, 2018.

[113] X. Zhang, A. J. Flueck, and C. P. Nguyen, "Agent-based distributed volt/var control with distributed power flow solver in smart grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 600–607, 2016.

[114] R. R. Jha, A. Dubey, C. Liu, and K. P. Schneider, "Bi-level volt-var optimization to coordinate smart inverters with voltage control devices," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1801–1813, 2019.

[115] W. Wang, N. Yu, J. Shi, and Y. Gao, "Volt-var control in power distribution systems with deep reinforcement learning," in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 1–7, 2019.

[116] W. Wang, N. Yu, Y. Gao, and J. Shi, "Safe off-policy deep reinforcement learning algorithm for volt-var control in power distribution systems," *IEEE Transactions on Smart Grid*, pp. 1–1, 2019.

[117] J. Duan, Z. Yi, D. Shi, C. Lin, X. Lu, and Z. Wang, "Reinforcement-learning-based optimal control of hybrid energy storage systems in hybrid ac–dc microgrids," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5355–5364, 2019.

[118] H. Xu, A. D. Domínguez-García, and P. W. Sauer, "Optimal tap setting of voltage regulation transformers using batch reinforcement learning," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1990–2001, 2020.

[119] R. Diao, Z. Wang, D. Shi, Q. Chang, J. Duan, and X. Zhang, "Autonomous voltage control for grid operation using deep reinforcement learning," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, 2019.

[120] J. Duan, D. Shi, R. Diao, H. Li, Z. Wang, B. Zhang, D. Bian, and Z. Yi, "Deep-reinforcement-learning-based autonomous voltage control for power grid operations," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 814–817, 2020.

[121] P. M. Anagnostopoulos and S. A. Papathanassiou, "A power flow method for radial distribution feeders with der penetration," *Journal of Technology Innovations in Renewable Energy*, no. 8, pp. 1–12, 2019.

[122] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multi-agent systems: A review of challenges, solutions and applications," *arXiv preprint arXiv:1812.11794*, 2018.

[123] Z. Tian, W. Wu, and B. Zhang, "A mixed integer quadratic programming model for topology identification in distribution network," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 823–824, 2016.

[124] J. Rocabert, A. Luna, F. Blaabjerg, and P. Rodríguez, "Control of power converters in ac microgrids," *IEEE Transactions on Power Electronics*, vol. 27, no. 11, pp. 4734–4749, 2012.

[125] G. Díaz, J. Gómez-Aleixandre, and J. Coto, "Direct backward/forward sweep algorithm for solving load power flows in ac droop-regulated microgrids," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2208–2217, 2016.

[126] X. Wang, Y. Wang, D. Shi, J. Wang, and Z. Wang, "Two-stage wecc composite load modeling: A double deep q-learning networks approach," *IEEE Transactions on Smart Grid*, pp. 1–1, 2020.

[127] L. Metz, J. Ibarz, N. Jaitly, and J. Davidson, "Discrete sequential prediction of continuous actions for deep rl," *arXiv preprint arXiv:1705.05035*, 2017.

[128] K. Kouno, M. Hirose, W. Hashimoto, T. Nagase, M. Hojo, and T. Ohnishi, "Effectiveness of constant power factor control of mega-solar system for voltage regulation with long distribution line," in *2013 IEEE 10th International Conference on Power Electronics and Drive Systems (PEDS)*, pp. 317–322, 2013.

[129] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[130] Y. Zhang, J. Wang, and Z. Li, "Interval state estimation with measurement and network parameter uncertainty in unbalanced distribution systems," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, 2019.

[131] P. S. Georgilakis and N. D. Hatziargyriou, "Optimal distributed generation placement in power distribution networks: Models, methods, and future research," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3420–3428, 2013.

[132] "Regulating Volatge: Recommendations for Smart Inverters," 2019. [Online]. Availabe: http://gridlab.org/wp-content/uploads/2019/09/GridLab_Regulating-Voltage-report.pdf.

[133] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[134] G. Cavraro and R. Carli, "Local and distributed voltage control algorithms in distribution networks," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1420–1430, 2018.

[135] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, "Topology identification in distribution systems using line current sensors: An milp approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1159–1170, 2020.

[136] J. Zhao and L. Mili, "A decentralized h-infinity unscented kalman filter for dynamic state estimation against uncertainties," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4870–4880, 2019.

[137] P. Romano and M. Paolone, "Enhanced interpolated-DFT for synchrophasor estimation in fpgas: Theory, implementation, and validation of a PMU prototype," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 12, pp. 2824–2836, 2014.