

Secure decentralized state estimation and control for smart distribution
grids

by

Biswajeet Rout

B.Tech., Biju Patnaik University of Technology, Bhubaneswar, 2016

M.Tech., Indian Institute of Technology Gandhinagar, 2018

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wiegiers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Abstract

Modern power distribution networks face significant challenges due to the increasing integration of renewable energy resources and electric vehicles, leading to higher uncertainty and dynamics in grid states. This necessitates accurate state estimation for effective control actions. This dissertation proposes multiple novel approaches for distribution system state estimation (DSSE) that can be scalable to large distribution networks and control of inverter-based resources, offering a robust solution to these challenges.

Current sparsity-based (DSSE) methods face limitations in scalability to large distribution networks or fail to effectively account for the temporal dependency within the measurement data. To address the scalability issue, we propose a hierarchical spectral clustering-based network partitioning algorithm followed by distributed compressive sensing (DCS)-based state estimation. This method utilizes an alternating direction method of multipliers (ADMM) for iterative information exchange among sub-areas for DSSE. We also demonstrate the robustness of ADMM-based distributed CS against various cyber-attacks through simulations on IEEE 37-bus and IEEE 123-bus networks. We address the second challenge by proposing two dynamic sparsity-based state estimation strategies, locally weighted matrix completion (LW-MC) and Bayesian matrix completion with Kalman filter prediction (BMC-KF). Evaluations on IEEE 37 and IEEE 123 bus test systems demonstrate that BMC-KF outperforms LW-MC.

The above-proposed methods work efficiently for large distribution networks. However, they suffer from high computational complexity or slow convergence time. Hence, they become unsuitable for real-time state estimation. Therefore, we propose a model-based neural network called deep unfolding, for fast and accurate estimation of system states in low-observable distribution networks. This approach achieves accurate state estimation with reduced simulation time by incorporating linearized power flow constraints, and the efficacy

is validated through extensive simulations on IEEE-123 bus and 559 bus networks. This deep learning approach does not provide any guarantee with regard to the estimate fidelity. Therefore, we propose a deep unfolding approach augmented with Bayesian neural networks to estimate the mean and covariance of system states, demonstrating its efficacy in estimating system states and providing associated confidence intervals through simulations on the IEEE 123-bus network.

In addition to state estimation, the increasing integration of IBRs necessitates robust and decentralized control strategies to ensure system stability and efficient power sharing. The existing literature uses PID control for voltage and frequency restoration that suffers from asymptotic convergence. Hence, this dissertation introduces a novel approach that leverages game theory principles to design distributed secondary control for inverter-based islanded microgrids. The proposed approach optimizes power sharing among interconnected IBRs while enhancing system frequency stability. Simulation results on a four-bus microgrid demonstrate that the game-theoretic secondary control offers improved power-sharing accuracy, stability, and load balancing compared to the distributed-averaging proportional-integral (DAPI) method. In addition to this, to make a microgrid stable against cyber attacks, we propose an adversarial robust multiagent reinforcement learning secondary control that mitigates adversarial impacts and ensures system resilience.

Together, these contributions address the evolving landscape of power distribution systems, by providing effective decentralized state estimation and control strategies to ensure stability and efficiency amidst large-scale integration of renewable energy resources.

Secure decentralized state estimation and control for smart distribution
grids

by

Biswajeet Rout

B.Tech., Biju Patnaik University of Technology, Bhubaneswar, 2016

M.Tech., Indian Institute of Technology Gandhinagar, 2018

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wieggers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Prof. Balasubhramaniam Natarajan

Copyright

© Biswajeet Rout 2024.

Abstract

Modern power distribution networks face significant challenges due to the increasing integration of renewable energy resources and electric vehicles, leading to higher uncertainty and dynamics in grid states. This necessitates accurate state estimation for effective control actions. This dissertation proposes multiple novel approaches for distribution system state estimation (DSSE) that can be scalable to large distribution networks and control of inverter-based resources, offering a robust solution to these challenges.

Current sparsity-based (DSSE) methods face limitations in scalability to large distribution networks or fail to effectively account for the temporal dependency within the measurement data. To address the scalability issue, we propose a hierarchical spectral clustering-based network partitioning algorithm followed by distributed compressive sensing (DCS)-based state estimation. This method utilizes an alternating direction method of multipliers (ADMM) for iterative information exchange among sub-areas for DSSE. We also demonstrate the robustness of ADMM-based distributed CS against various cyber-attacks through simulations on IEEE 37-bus and IEEE 123-bus networks. We address the second challenge by proposing two dynamic sparsity-based state estimation strategies, locally weighted matrix completion (LW-MC) and Bayesian matrix completion with Kalman filter prediction (BMC-KF). Evaluations on IEEE 37 and IEEE 123 bus test systems demonstrate that BMC-KF outperforms LW-MC.

The above-proposed methods work efficiently for large distribution networks. However, they suffer from high computational complexity or slow convergence time. Hence, they become unsuitable for real-time state estimation. Therefore, we propose a model-based neural network called deep unfolding, for fast and accurate estimation of system states in low-observable distribution networks. This approach achieves accurate state estimation with reduced simulation time by incorporating linearized power flow constraints, and the efficacy

is validated through extensive simulations on IEEE-123 bus and 559 bus networks. This deep learning approach does not provide any guarantee with regard to the estimate fidelity. Therefore, we propose a deep unfolding approach augmented with Bayesian neural networks to estimate the mean and covariance of system states, demonstrating its efficacy in estimating system states and providing associated confidence intervals through simulations on the IEEE 123-bus network.

In addition to state estimation, the increasing integration of IBRs necessitates robust and decentralized control strategies to ensure system stability and efficient power sharing. The existing literature uses PID control for voltage and frequency restoration that suffers from asymptotic convergence. Hence, this dissertation introduces a novel approach that leverages game theory principles to design distributed secondary control for inverter-based islanded microgrids. The proposed approach optimizes power sharing among interconnected IBRs while enhancing system frequency stability. Simulation results on a four-bus microgrid demonstrate that the game-theoretic secondary control offers improved power-sharing accuracy, stability, and load balancing compared to the distributed-averaging proportional-integral (DAPI) method. In addition to this, to make a microgrid stable against cyber attacks, we propose an adversarial robust multiagent reinforcement learning secondary control that mitigates adversarial impacts and ensures system resilience.

Together, these contributions address the evolving landscape of power distribution systems, by providing effective decentralized state estimation and control strategies to ensure stability and efficiency amidst large-scale integration of renewable energy resources.

Contents

List of Figures	xiii
List of Tables	xvi
List of Nomenclature	xvii
Acknowledgements	xix
Dedication	xx
Preface	xxi
1 Introduction	1
1.1 Research Questions	5
1.2 Research Contributions	6
1.3 Dissertation Organization	13
2 Literature Review	15
2.1 State Estimation in Power Distribution Network	15
2.2 Distributed State Estimation	17
2.3 Effect of cyber attack on DSSE	19
2.4 Deep Unfolding for State Estimation	19
2.5 Microgrid Secondary Control	21
2.6 Reinforcement learning for MG Secondary Control	22
3 Background	24

3.1	State Estimation	24
3.1.1	WLS for state Estimation	26
3.1.2	Compressive Sensing	27
3.1.3	Matrix completion	29
3.2	Microgrid Secondary control	30
3.2.1	Primary control of MG	31
3.2.2	Secondary Control of Microgrid	33
3.2.3	Communication network modeling	33
4	Dynamic State Estimation	35
4.1	Introduction	35
4.2	Background	36
4.3	Proposed Approach	38
4.3.1	Locally Weighted matrix completion (LW-MC)	39
4.3.2	Dynamic Bayesian Matrix completion with KF Prediction	40
4.4	Computational Complexity Analysis	45
4.5	Simulation Results and Discussions	46
4.6	Summary	51
5	Distributed State Estimation	52
5.1	Introduction	52
5.2	Problem formulation	54
5.2.1	Decentralised Linear Power Flow Model	54
5.2.2	Fully Distributed Model	57
5.2.3	Computational complexity	61
5.3	Simulation Results	61
5.4	Distributed State Estimation in Large Power Distribution Network	66
5.5	Hierarchical Network Partitioning	66

5.5.1	Distributed Compressive Sensing	69
5.5.2	Computational Complexity	70
5.6	Simulation Results	70
5.7	Summary	72
6	Deep unfolding for Distribution Grid State Estimation	75
6.1	Introduction	75
6.2	Background and preliminaries	76
6.3	Proposed Approach	79
6.3.1	Model based DSSE-Net	81
6.3.2	Model-free DSSE-Net	84
6.4	Computational complexity	87
6.5	Simulation Results	90
6.5.1	Data Preparation and network training	90
6.5.2	Experimental results	91
6.5.3	Robustness to Measurement Noise	94
6.5.4	Robustness to Simulation Model	95
6.6	Uncertainty Estimation	96
6.7	Proposed Approach	97
6.7.1	Deep unfolding from ADMM iterations - DSSE-net	98
6.7.2	Deep unfolding using BNN	100
6.8	Simulation Results	103
6.8.1	Data Preparation and network training	103
6.8.2	Experimental results	104
6.9	Summary	105
7	Cyber Attack Analysis	107
7.1	Cyber Attack modeling	107

7.2	Cyber Attack Models	108
7.3	Simulation Results	109
7.4	Summary	113
8	Microgrid Secondary Control	114
8.1	Introduction	114
8.2	Background and Preliminaries	115
8.2.1	Primary control of MG	115
8.2.2	Secondary Control of Microgrid	117
8.2.3	Communication network modeling	117
8.3	Proposed Approach	118
8.3.1	Game design for distributed frequency control	118
8.3.2	Existence and Uniqueness of NE	120
8.3.3	Stability of NE	123
8.3.4	Convergence to NE	123
8.4	Simulation Results	125
8.4.1	General load variation	126
8.4.2	Effect of payoff parameter β on Nash equilibrium	131
8.4.3	Plug and Play Operation	132
8.5	Summary	132
9	RL for Secondary Control of Microgrids	133
9.1	Introduction	133
9.2	Background and Problem Description	134
9.2.1	Large Signal Model of Inverters	134
9.2.2	Communication network modeling	138
9.2.3	Problem Description	138
9.3	Proposed Approach	139

9.3.1	Error dynamic evaluation	140
9.3.2	RL for Function Approximation	143
9.3.3	Critic Network Architecture	145
9.3.4	Actor Network Architecture	146
9.4	Adversarial Attack on RL agents	148
9.4.1	Strategic Attack design	149
9.4.2	When to Attack	151
9.4.3	How much to attack	151
9.5	Simulation Results	152
10	Conclusion and Future Work	155
10.1	Future Work	157
	Bibliography	160
A	ADMM Convergence	178
A.1	Appendix	178
B	Reuse permissions from publishers	180

List of Figures

1.1	Thesis Organisation	14
3.1	The basic control architecture in an inverter	31
3.2	Microgrid architecture and the communication infrastructure	34
4.1	Block diagram of the proposed BMC-KF approach	40
4.2	Dynamic state estimation due to a change in load at bus 6 of the IEEE 37 bus system	48
4.3	Voltage magnitude recovery with 30 % FAD	49
4.4	Voltage angle recovery with 30 % FAD	49
4.5	IEEE 123 bus system voltage magnitude recovery with 30% FAD	50
4.6	IEEE 123 bus system voltage angle recovery with 30% FAD	51
5.1	Distributed architecture of the IEEE37-bus test system	55
5.2	Schematic of a 3-area architecture of the proposed approach showing the internal and overlapping boundary variables	57
5.3	Voltage recovery in IEEE 37 bus test system	63
5.4	Voltage recovery with noisy data	64
5.5	Convergence characteristics of the proposed ADMM approach	65
5.6	Voltage recovery in IEEE 123 bus test system	65
5.7	Flow diagram of hierarchical network partitioning	68
5.8	Five-area optimal partitioning of the IEEE 123-bus network	71
5.9	Twelve-area optimal partitioning of the IEEE 8,500-node system	71
5.10	Eighteen-area optimal partitioning of the real 6,000+ node system	72

5.11	Voltage recovery of 6,000+ node system	73
6.1	Proposed architecture	76
6.2	Voltage magnitude recovery in different phases of the IEEE-123 bus network for 50% CMR	90
6.3	Voltage magnitude recovery in different phases of the IEEE-123 bus network for 50% CMR	90
6.4	Voltage magnitude and angle estimation in the phase-a of the 559 bus distri- bution network at 30% of CMR	93
6.5	Voltage magnitude and angle estimation in the phase-a of the IEEE 123 bus distribution network at 30% of CMR with different measurement noise . . .	95
6.6	State recovery with inaccurate simulation model data	96
6.7	Voltage estimation with debiased DSSE-net (30% CMR)	102
6.8	Voltage estimation with biased DSSE-net (30% CMR)	104
7.1	Voltage recovery under loss of measurement attack	110
7.2	Voltage recovery under FDI attack	111
7.3	Voltage recovery under neighborhood attack	111
7.4	Voltage recovery under replay attack	112
7.5	Voltage recovery in multi-cyber attack (loss of data + FDI)	113
8.1	Case-1: Microgrid operation with same droop gain and no load change . . .	126
8.2	Case-2: Microgrid operation with same droop gain and load transients . . .	128
8.3	Case-3: Microgrid operation with unequal droop gain and no load transients	128
8.4	Case-4: Microgrid operation with unequal droop gain and load transients . .	130
8.5	Nash equilibrium with different β values	131
8.6	Performance of the Game-theoretic controller during plug-and-play operation	131
9.1	Frequency restoration with adversarial training	153

9.2	Frequency restoration with adversarial training	153
-----	---	-----

List of Tables

4.1	Time requirement for 1 run to estimate 4 time steps	50
5.1	Simulation time for 1 run and 1 CMR	61
5.2	Area partition of IEEE networks	65
5.3	Performance of proposed partitioning algorithm	72
6.1	Simulation time of different algorithms with varying CMR%	89
6.2	Performance comparison of the proposed approaches	94
6.3	DSSE-net performance with 30% CMR	105
8.1	Parameters of the microgrid test system	125
9.1	Parameters of the microgrid test system	152

Nomenclature

DSSE Distribution system state estimation

AMI Advanced metering infrastructure

PMU Phasor measurement unit

CS Compressive Sensing

MC Matrix Completion

DMS Distribution management database

FAD Fraction of Available Data

NLL Negative log-likelihood

KL KL divergence

WLS Weighted least square

SR Sparsity Ratio

RIP Restricted isometry property

KF Kalman filter

DCT Discrete cosine transform

ODE Ordinary differential equation

ML Maximum likelihood

MG Microgrid

- CI Communication infrastructure
- DCS Distributed compressive sensing

Acknowledgments

I extend my deepest gratitude to my advisor, Dr. Balasubramaniam Natarajan, for his unwavering guidance, support, and encouragement throughout my Ph.D. journey. His invaluable patience and feedback have been instrumental in my growth. I also wish to express my sincere appreciation to my doctoral committee members, Dr. Anil Pahwa, Dr. Sanjoy Das, and Dr. Pietro Poggi-Corradini, for their insightful advice and constructive comments, which have significantly enhanced the quality of my research. I am indebted to Dr. Govind Saraswat and Dr. Yashodhan Agalgaonkar for their mentorship and encouragement during my doctoral studies.

I am grateful to my colleagues at CPWin Lab, including Sai, Mohammad, Deepesh, Shweta, Dylan, Aabila, Prudhviraaj, and Priyanka, for their collaboration and constant motivation. I also want to acknowledge my dear friends at K-STATE, Amogh, Sneha, Gunvant, Laxmi, Prakash, Priya, Snehal, and Ganesh, for their unwavering support and care. The memories and joy we shared will always hold a special place in my heart.

My heartfelt thanks go to my parents, Adhir Rout and Abala Rout, and my brother, Dr. Ranjeet Ku. Rout, Sujit Rout, for their unconditional love, support, and encouragement. Their prayers have been my strength throughout this journey. Finally, I owe immense gratitude to my wife, Dr. Shweta Dahale, for her unwavering support and encouragement in every aspect of my life, especially during this challenging experience.

This material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Solar Energy Technologies Office, under Award Number DE-EE0008767 and DE-EE0010416.

Dedication

I would like to dedicate this dissertation to my loving wife Dr. Shweta Dahale and my parents Adhir Ku. Rout and Abala Rout for their constant support and trust in me.

Preface

This dissertation, “Secure Decentralized State Estimation and Control for Smart Distribution Grids,” is submitted for the degree of Doctor of Philosophy in the Mike Weigers Department of Electrical and Computer Engineering at Kansas State University. The research was conducted under the supervision of Professor Balasubramaniam Natarajan. This work is original to the best of my knowledge, except where acknowledgments and references are indicated. Most of the work has been presented in the following published peer-reviewed journals and conferences:

1. B. Rout, S. Dahale and B. Natarajan, ”Dynamic Matrix Completion Based State Estimation in Distribution Grids,” in IEEE Transactions on Industrial Informatics, vol. 18, no. 11, pp. 7504-7511, Nov. 2022, doi: 10.1109/TII.2022.3162210.
2. Biswajeet Rout, Balasubramaniam Natarajan, ”Impact of cyber attacks on distributed compressive sensing based state estimation in power distribution grids,” in International Journal of Electrical Power & Energy Systems, Volume 142, Part B, 2022, 108295, ISSN 0142-0615, <https://doi.org/10.1016/j.ijepes.2022.108295>.
3. B. Rout, G. Saraswat and B. Natarajan, ”Efficient Network Partitioning: Application for Decentralized State Estimation in Power Distribution Grids,” 2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2023, pp. 1-5, doi: 10.1109/ISGT51731.2023.10066393.
4. B. Rout and B. Natarajan, ”Deep Unfolding for State Estimation in Power Distribution Grids,” in IEEE Transactions on Industrial Informatics (in 2nd review)
5. B. Rout, and B. Natarajan, ”Bayesian Deep Unfolding for State Estimation in Power Distribution Grids,” *2024 IEEE Kansas Power and Energy Conference (KPEC-2024)*, (under review).

6. B. Rout, and B. Natarajan, "Distributed Secondary Control of Inverter-based Islanded Microgrid: A Game Theoretic Approach," *IEEE Trans. Power Delivery*, (under review).
7. B. Rout, B. Natarajan, and Y.P. Agalgaonkar, "Reinforcement Learning for Secondary Control of Inverter-based Microgrid under Adversarial Attack," *IEEE Trans. Power Delivery*, (Under review).

Chapter 1

Introduction

The power distribution network is experiencing a major change in its grid characteristics. This characteristic change is driven by increased renewable energy integration, advancements in smart grid technologies, and the growing demand for sustainable energy solutions. Under these circumstances, efficient and reliable delivery of electricity is paramount. Therefore, the need for accurate monitoring and control of distribution networks has become more critical than ever before. State estimation emerges as a fundamental tool in achieving this objective, providing valuable insights into the real-time operational status of the distribution system.

State estimation, also known as network monitoring or observability analysis, is a computational technique used to infer the current state of a power distribution network based on available measurements. State estimation algorithms can estimate the voltage magnitudes, phase angles, and power injections at different nodes along with the line flows by utilizing data from various sensors, like smart meters, phasor measurement units (PMUs), supervisory control and data acquisition (SCADA) and other monitoring devices deployed throughout the network [1]. These estimates enable the network operators to gain a comprehensive understanding of the network's operating conditions, identify potential issues or anomalies, and make informed decisions to optimize grid operating performance and improve stability.

The significance of state estimation in power distribution networks lies in its ability to provide actionable insights in real-time, enabling operators to detect and address opera-

tional challenges promptly. However, achieving accurate distribution system state estimation (DSSE) poses a considerable challenge due to the network’s limited observability, stemming from the scarcity of available real-time measurements. To further aggravate the issue, outliers in measurements and errors in system model information introduce new sources of uncertainty in DSSE. Since optimal control and operation of the distribution grid is contingent upon the ability to estimate the system states accurately, it is critical to investigate methods to estimate the states accurately with a small number of measurements in the presence of measurement uncertainties and inaccurate system model information. This is one of the fundamental problems we seek to study in the context of an unbalanced three-phase power distribution network.

Effective situational awareness relies on our ability to analyze and interpret the data provided by monitoring and control systems. In this regard, there has been an increase in the proliferation of advanced metering devices in the smart distribution grid paradigm. According to Energy Information Administration (EIA) data [2], the advanced meter deployments have increased in the United States by approximately 8.3 million meters, representing a 8.8% annual increase in advanced meters from 2019 to 2020. The use of SCADA, PMUs, Intelligent electronic devices (IEDs), automated feeder switches and voltage regulators, and smart inverters for distributed energy resources (DERs) have provided an opportunity to increase system observability. This increase in sensing infrastructure results in increased information aggregated at a centralized distributed management system (DMS). However, reliable aggregation and management of such a huge volume of data at a central location is challenging, even under optimal conditions. Sometimes, it is difficult for the underlying communication network to keep up with the sensors’ volume and velocity of information. The information aggregation challenge for centralized state estimation grows with the size of the distribution network. As a result, the use of large volumes of data for real-time state estimation and control of the power distribution network becomes difficult. This is another fundamental problem we seek to study. We propose local data aggregation and information processing for control centers to manage a subset of the network nodes and information exchange across control centers for state estimation and control. Moreover, as these systems become more

complex and connected, they become more vulnerable to cyber-attacks. To address this issue, researchers have been exploring using decentralized state estimation and control techniques to improve the security of smart distribution grids. In this dissertation, we discuss the challenges faced by smart distribution grids and how decentralized state estimation and control can be used to overcome them.

The global energy landscape has radically transformed in recent years. Some driving forces for the transformation include (1) pressing concerns posed by climate change, (2) the rise in energy prices, and (3) increasing concerns about energy security and reliability. As traditional centralized power systems grapple with these challenges, microgrids (MGs) with distributed energy resources (DERs) emerge as a novel alternative. However, ensuring frequency and voltage stability is crucial in MGs. In the conventional power grid, voltage and frequency stability is achieved by synchronous generators because of their inbuilt inertia [3]. On the other hand, most MGs are sourced by solid-state power electronics devices that lack inertia. Droop control has emerged as a dominant solution to regulate voltage and frequency in MGs because of its simple implementation and reliance on local measurements without requiring any communication infrastructure. However, due to its inherent characteristics, droop control may not be precise and can cause voltage and frequency deviations from their nominal values [4]. Therefore, a secondary level of hierarchical control becomes necessary and serves as a compensation mechanism. Based on the communication requirements, MG secondary control is broadly classified into three types: (1) Centralized control: This enables easier decision-making and coordination. However, centralized control suffers from large communication bandwidth requirements, which leads to decision delays and low reliability; (2) Decentralized secondary control: This approach does not require exclusive communication infrastructure (CI) for control. Each DER restores its frequency and voltage amplitudes to nominal values individually. However, the CI is required for transferring data from the tertiary control and MG centralized control layers. (3) Distributed secondary control: This uses a low bandwidth sparse communication network in which the inverters are only required to interact with the neighboring inverters.

The research community is continuously working on developing sophisticated distributed

and/or decentralized state estimation and control methodologies for the smart distribution grid. A detailed literature review of the existing methodologies is given in Chapter 2. However, the existing methods suffer from various shortcomings, which are listed below.

1. **Measurement Redundancy and network architecture:** The conventional approaches for state estimation require high measurement redundancy, e.g., for reliable estimation accuracy, the conventional weighted least square (WLS) method of state estimation requires the number of measurements to be greater than the number of states in the power distribution grid. Furthermore, the direct application WLS in the distribution grid is not applicable because of the high $\frac{R}{X}$ ratio of the power distribution network.
2. **Low accuracy:** Pseudo measurements, derived from historical or forecasted data, are frequently combined with current measurements to offset the scarcity of measurements in DSSE. However, the historical data have limited accuracy. Therefore, the estimation performance accuracy deteriorates further.
3. **High computational complexity:** The current methods for DSSE and control face challenges due to their significant computational complexity stemming from centralized data aggregation and processing. The computational burden increases with an increase in system size. Hence, these methods are not suitable for real-time state estimation and control in large distribution networks.
4. **High bandwidth communication infrastructure requirements:** Existing centralized state estimation and control methods depend on a centralized state estimation engine and control center. All the on-field sensors communicate the data to the centralized control center. As the sensor deployment increases, the data volume increases; hence, a high bandwidth communication infrastructure is required for reliable information exchange.
5. **Prone to cyber attack:** As reliance on network-connected smart sensing devices grows, the distribution network becomes more vulnerable to cyber-attacks. However,

current DSSE and control methods lack the capability to assess the impact of such cyber intrusions.

Based on these perspectives, we seek to address a few fundamental research questions in this dissertation. These questions and our approach to address them are the key contributions of this dissertation, as discussed in the following section.

1.1 Research Questions

To tackle the challenges associated with efficient scalable and secure state estimation and control in power distribution networks, this dissertation aims to answer the following questions:

1. **Question 1:** *For distribution network state estimation*

- (a) How can we develop distribution system state estimators that are computationally efficient and achieve enhanced accuracy with limited measurements?
- (b) How can we develop DSSE that can be scalable and resilient against cyber attacks?
- (c) Can DSSE methodologies rooted in learning approaches provide low complexity and scalable to large distribution networks for real-time monitoring? Furthermore, is it possible to derive confidence intervals for state estimators developed through machine learning to aid utilities in making informed decisions?

2. **Question 2:** *Secondary control of Microgrid*

- (a) How can we develop a secondary control architecture for microgrids using a low bandwidth communication infrastructure with improved control performance?
- (b) How can we inject intelligence into microgrid agents to enable them to adapt and respond appropriately to changing microgrid operational conditions? Furthermore, can we enhance the intelligence of these agents to efficiently mitigate cyber intrusions and uphold grid stability during adversarial attacks?

1.2 Research Contributions

We address question 1 (a) in Chapter 4 of the dissertation. The state-of-the-art state estimation approaches, such as WLS and its variants, require a large number of measurements. To overcome this, researchers have proposed sparsity-based techniques like compressive sensing (CS), tensor completion (TC), or matrix completion (MC). These methods leverage the inherent sparsity in raw measurements (MC and TC), or on a sparsifying domain (CS) to recover the system states from limited measurements. However, existing approaches primarily focus on spatial sparsity, neglecting the dynamics of temporal correlation/sparsity within the measurements. Hence, for the first time, we formulate the DSSE as a dynamic matrix completion problem that exploits spatial and dynamics of temporal sparsity in the raw measurements. The major contributions of Chapter 4 are listed below.

- We propose two new methods: locally weighted matrix completion (LW-MC) and Bayesian matrix completion with Kalman filter prediction (BMC-KF) to impute the missing entries of the underlying low-rank matrix dynamically.
- Extensive simulation is conducted on the IEEE 37 and IEEE 123 bus test systems. The performance of both methods (LW-MC and BMC-KF) is compared with the existing static-MC approach. It is found that the accuracy in state estimation is best in BMC-KF followed by LW-MC and static-MC state estimation methods.
- While the computational complexity of the LW-MC approach is shown to be comparable to BMC-KF for smaller networks, the BMC-KF complexity scales favorably as network size increases.

More details on the dynamic matrix completion-based DSSE formulations can be found in Chapters 4 and in the following published article:

- B. Rout, S. Dahale and B. Natarajan, "Dynamic Matrix Completion Based State Estimation in Distribution Grids," in IEEE Transactions on Industrial Informatics, vol. 18, no. 11, pp. 7504-7511, Nov. 2022, doi: 10.1109/TII.2022.3162210.

The dynamic MC completion approach proposed in Chapter 4 demands large computational power, and the computational complexity increases with increased system size. Hence, it lacks scalability for large distribution networks. To overcome this challenge, we develop a distributed compressing-based (CS) DSSE that addresses question 1 (b). Chapter 5 of the dissertation introduces a distributed states estimation approach that uses CS to recover the system states from very limited measurements. The scalability of the proposed approach is evaluated on a real 6000-node distribution network. Finding an optimal network partition for a large distribution network is difficult. Hence, we propose a novel hierarchical network partitioning algorithm to divide the distribution network before applying the distributed CS-based DSSE. Furthermore, we have analyzed the impact of the loss of measurements, FDI, replay, and neighborhood cyber-attacks on the distributed CS-based DSSE in Chapter 7. The major contributions of Chapter 5 and 7 are listed below.

- A novel and efficient hierarchical spectral clustering-based network partitioning algorithm that partitions the network with minimized overlap is proposed. The proposed approach can be applied to all types of distributed/decentralized tasks in a power distribution network.
- The efficacy of the proposed approach is tested on a standard IEEE 123-bus network, an IEEE 8,500-node system, and a real 6000+ node distribution network.
- The network partitioning approach is applied to distributed compressive sensing (DCS)-based state estimation. The distributed multi-area DSSE framework uses the alternating direction method of multiplier (ADMM) to find global solutions.
- The proposed method uses local measurements and limited consensus variables among the boundary busses to reduce the communication burden and computation time.
- Extensive simulation is performed on IEEE 37-bus, IEEE 129-bus test systems, and 6000+ node distribution networks to demonstrate the effectiveness of the multi-area DSSE.

- Furthermore, the dissertation provides a comparative performance analysis of distributed CS-based state estimation and centralized CS-based state estimation approach and highlights the benefits offered in terms of their computational complexity/computation time.

More details on the hierarchical network partition and distributed DSSE formulations can be found in Chapters 5 and Chapter 7 and in the following published articles:

- Biswajeet Rout, Balasubramaniam Natarajan, "Impact of cyber attacks on distributed compressive sensing based state estimation in power distribution grids," in International Journal of Electrical Power & Energy Systems, Volume 142, Part B, 2022, 108295, ISSN 0142-0615, <https://doi.org/10.1016/j.ijepes.2022.108295>.
- B. Rout, G. Saraswat and B. Natarajan, "Efficient Network Partitioning: Application for Decentralized State Estimation in Power Distribution Grids," 2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2023, pp. 1-5, doi: 10.1109/ISGT51731.2023.10066393.

Although the distributed DSSE proposed in Chapter 5 is scalable to large distribution networks, the simulation time depends on the larger part of the partition and the number of shared variables among the neighboring areas. Hence, state estimation for larger networks in real time is infeasible. We address this challenge with a deep learning-based state estimation approach. We propose a model-free deep unfolding architecture for compressive sensing-based DSSE. Firstly, the alternating direction method of multiplier (ADMM) is used to solve the compressive sensing-based DSSE. Next, we unfold the structures of the iterations of the ADMM into L layers of a neural network by introducing a number of trainable parameters, specifically the sparsifying basis matrix Ψ for our case. Ψ is then learned using a deep learning architecture with an appropriate choice of a loss function, back-propagation, and stochastic gradient descent optimization algorithm. The resulting unfolded algorithm with the learned parameters is then used to estimate the system states in a low observable distribution network. By mapping iterations of the optimization algorithm to layers of a

neural network, the deep unfolding approach avoids the need for specifying a sparsifying basis matrix or a detailed distribution network model. Finally, we obtain the closed-form solution of the ADMM algorithm, which solves the DSSE. Furthermore, to completely address question 1 (c), we propose a Bayesian deep unfolding-based state estimation to estimate both the mean and variance of the state estimates. The key contributions of these works are as follows:

- The proposed approach uses CS as the optimization technique to recover the system states in a low observable distribution network. However, the sparsity-aware CS method uses an optimization problem, and the computational complexity grows as the network size increases. To overcome the complexity issue, we propose a learning-based approach. However, unlike traditional learning-based approaches, where a black box approach to function approximation is used, our approach uses only the mathematical model information to estimate the system states from very limited measurements.
- We propose two deep unfolding architectures: (1) MFDSSE-net is a purely data-driven deep unfolding architecture. MFDSSE-net learns the sparsifying basis matrix Ψ by exploiting the inter-dependencies of the data. This approach is helpful when the distribution network model information is outdated due to network aging or poor documentation. (2) MBDSSE-net incorporates linearized power flow constraints into the l_1 -norm optimization problem. As a result, MBDSSE-net uses both the data and the model information to learn the sparsifying basis matrix Ψ .
- We compare the performance of the proposed methods with the ADMM-based l_1 -norm optimization problem, which considers a discrete cosine transform as the sparsifying matrix. Extensive simulation is done on the IEEE 123 bus and a 559 bus network to evaluate the performance of the proposed methods. Simulation results show that both the deep unfolding approaches MFDSSE-net and MBDSSE-net perform equally well in estimating the system states. However, there is a significant reduction in computation time in the deep unfolding approaches relative to the ADMM-based l_1 -norm optimization problem. This is because the computational cost is shifted to the training phase

of the deep unfolding networks.

- DSSE-net provides a point estimate without any information about the prediction confidence. Hence, we leverage the concept of Bayesian neural networks (BNN) in the context of deep unfolding to estimate both the mean and covariance of the estimated states.

More details related to the deep unfolding and Bayesian deep unfolding for DSSE can be found in Chapter 6 and in the following articles:

- B. Rout and B. Natarajan, "Deep Unfolding for State Estimation in Power Distribution Grids," in IEEE Transactions on Industrial Informatics (in 2nd review)
- B. Rout, and B. Natarajan, "Bayesian Deep Unfolding for State Estimation in Power Distribution Grids," *2024 IEEE Kansas Power and Energy Conference (KPEC-2024)*, (under review).

We address question 2(a) in Chapter 8. Secondary control in microgrids is used to compensate for the deviations that occur in voltages and frequencies after the primary control operation. The existing research employs a PID-based controller, which has an asymptotic convergence. However, for microgrid operation asymptotic convergence is undesirable. Hence, we propose a game theoretic framework for multi-agent secondary control of islanded microgrids. We formulate the game theoretic control by treating the microgrid as a multiagent system, where each agent takes appropriate actions to stabilize the frequency and voltage deviations in the network. While game theoretical approaches have been leveraged for distributed control of multiagent systems, a rigorous theoretical analysis of its use for MG frequency control is currently lacking. Hence, we design a game to depict the interaction among inverters in a microgrid and propose a new consensus protocol-based game theoretic approach for the secondary control of microgrids. The key contributions of this work are highlighted below.

- We address the intricacies of inverter-based microgrid (MG) systems by adopting a multi-agent framework. In this novel approach, each IBR within the microgrid is

modeled as an autonomous agent. We delve into the challenge of alleviating frequency deviations within the microgrid, considering it a non-cooperative game encompassing the participating agents. The core objective of this game is to collaboratively rectify the frequency deviation and realign it with the system frequency. Our proposed non-cooperative game involves the agents minimizing a cost function. This cost function is a composite of the agent’s strategy and the strategy taken by its neighbors. This necessitates strategic modifications both in the agent’s personal approach and in its response to the strategies employed by neighboring agents.

- We adopt the potential function approach to rigorously investigate the Nash equilibrium for our game formulations. This equilibrium point represents the state at which no agent can unilaterally deviate from its strategy to attain a better outcome. Through the lens of the potential function, we explore, for the first time, the existence, uniqueness, stability, and convergence of this Nash equilibrium within a microgrid. This step is pivotal in understanding the conditions that allow the system to settle into a balanced state, benefiting individual agents and the overall microgrid stability.
- We offer a comparative examination of the proposed game theoretic secondary control with the distributed-average proportional-integral (DAPI) control. Simulation results demonstrate that our proposed method achieves a speed improvement of 70% faster over the DAPI control in restoring the system frequency following load disturbances.
- Our study examines the robustness of the proposed secondary controller in the face of communication network disruptions. Specifically, we examine how the microgrid responds when communication links fail. This analysis sheds light on the system’s ability to adapt and maintain stability despite unforeseen challenges.

More details related to the game theoretic secondary control of microgrids can be found in Chapter 8 and the following article:

- B. Rout, and B. Natarajan, "Distributed Secondary Control of Inverter-based Islanded

Microgrid: A Game Theoretic Approach,” *IEEE Trans. Power Delivery*, (under review).

With the advancement of network-connected smart sensing devices, the microgrid becomes vulnerable to cyber-attacks. This necessitates the requirement of adversarial robust secondary control. Hence, in this dissertation, we propose, an adversarial robust reinforcement learning-based secondary control of microgrid to address the questions 2(b). We propose an online multi-agent actor-critic-based deep reinforcement learning (DRL) approach for secondary frequency and voltage control of microgrids (MGs). Unlike offline training, which relies on historical data and may lead to confusion when faced with new data during implementation, our method adapts dynamically to changing conditions. The proposed adaptive secondary controller supplements the primary droop control to mitigate deviations after the primary control’s action. Notably, this controller effectively addresses uncertainties stemming from various sources, such as DG dynamics, network parameter change (caused by network aging), load fluctuations, and rectifies steady-state errors inherent in the primary droop control. It has been observed that the RL agents suffer from adversarial attacks because of uncertainty in the measurements or attacks initiated by a cyber actor. Hence, we propose robust adversarial attack learning to cater to the problem of cyber attacks within the RL architecture. Some of the major contributions of this work are given below.

- We propose, for the first time, a multiagent actor-critic-based online reinforcement learning architecture for secondary control of inverter-based MG.
- As the RL agents can suffer from adversarial attacks, we develop a robust adversarial RL framework for the voltage and frequency restoration in an inverter-based MG.
- We design and evaluate a strategically timed attack to attack a deep RL agent.
- We perform extensive simulation on a four-inverter-based islanded MG to assess the performance of the proposed approach, and results show that the adversarial robust RL approach for secondary control surpasses RL-based control in mitigating adversarial impacts and ensuring system resilience.

More details related to the multi-agent RL-based secondary control of microgrids can be found in Chapter 9 and the following article:

- B. Rout, and B. Natarajan, "Reinforcement Learning for Secondary Control of Inverter-based Microgrid under Adversarial Attack" (Manuscript under preparation).

1.3 Dissertation Organization

Figure 1.1 outlines the structure of this dissertation. The remainder of this dissertation is organized as follows. Chapter 2 describes the related work and their limitations in addressing the challenges of a real distribution grid. Chapter 3 discusses the fundamental idea of state estimation in power distribution grids and secondary control of islanded microgrids. In chapter 4, we propose dynamic matrix completion for DSSE. Then, we develop distributed CS-based state estimation and its scalability to large distribution networks in Chapter 5. In chapter 6, we develop a deep unfolding-based state estimation approach for the distribution grid. A Bayesian deep unfolding-based state estimation method for quantifying the uncertainty in estimated states is discussed in Chapter 6. We discuss the robustness of the proposed distributed DSSE to various cyber attack scenarios in Chapter 7. Chapter 8 introduces a novel game theoretic secondary control of island microgrids and an adversarial robust reinforcement learning for secondary control of islanded microgrid is discussed in Chapter 9. The summary of this dissertation and future ideas constitute chapter 10.

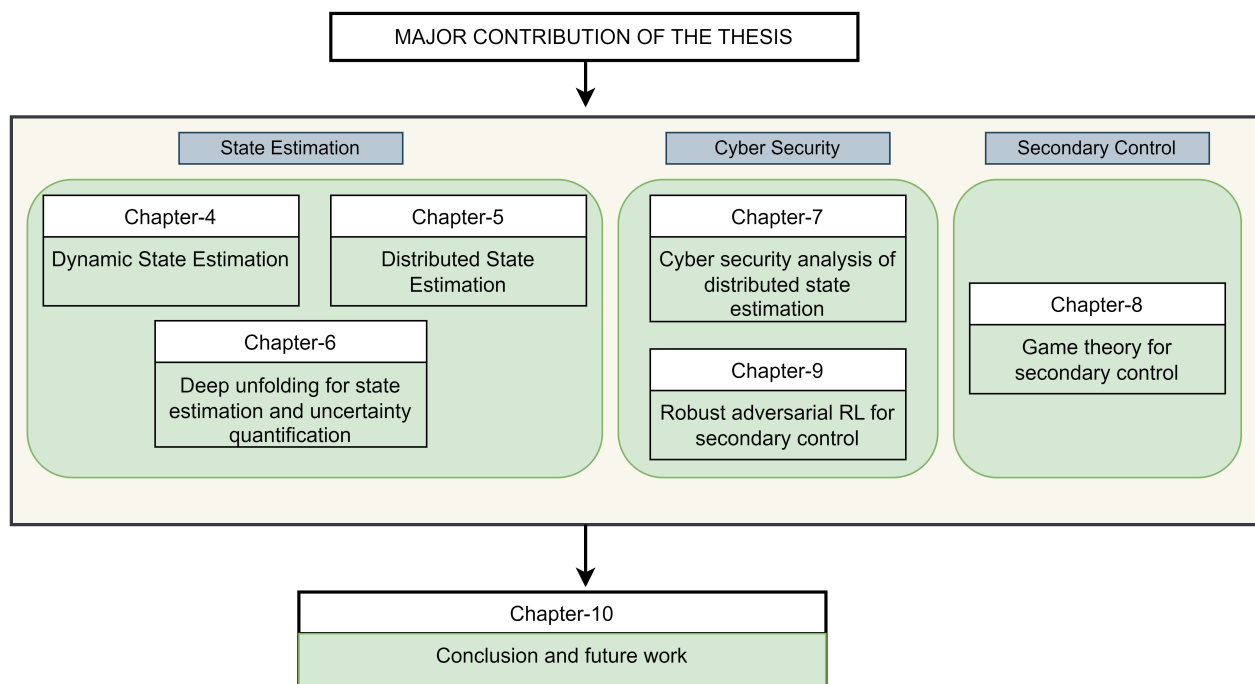


Figure 1.1: Thesis Organisation

Chapter 2

Literature Review

This chapter surveys the existing research on the topic of distributed state estimation and control. It examines articles, books, and studies to understand current knowledge. This analysis identifies gaps and common themes in research. The chapter sets the stage for the current study. It shows how the study fits into existing knowledge. This review helps readers understand the significance of the research. We start the chapter by introducing existing literature on state estimation, followed by distributed state estimation and their robustness to various cyber-attacks. Next, we explore the machine learning-based state estimation, followed by a brief literature on distributed secondary control in islanded microgrids.

2.1 State Estimation in Power Distribution Network

Optimal management of the power distribution grid has become challenging due to increased penetration of renewable energy resources and different dynamical loads like heat pumps and electric vehicles [5]. In this context, real-time control is important in dealing with intermittency and uncertainties. Distribution system state estimation (DSSE) enables the monitoring of the grid operating conditions and serves as the precursor to real-time optimal control. Weighted least square (WLS) based [6] state estimation has been widely used in the power transmission network. However, the applicability of WLS to the distribution network is

limited due to the system being unobservable. Recently, multiple researchers have attempted to address the issue of unobservability while performing state estimation in distribution networks. To make a low observable network fully observable, optimal measurement locations have been identified in [7]. Additionally, the difference between the measured and calculated voltages and complex power is used to obtain the voltage profile of the whole network. The methods adopted in [7] still need a significant number of sensors to observe the system state completely. [8] suggests using pseudo measurements to aid DSSE. However, pseudo-measurements are inaccurate and lead to errors in state estimation. In contrast to WLS, that treats the power system variables as static parameters that need to be estimated, the performance of a pseudo-dynamic system is evaluated using three different algorithms in [9]. In [10], a discrete-time dynamic model for state estimation is modeled by combining the power flow equations and load forecasts. A comparative performance analysis between WLS and Extended Kalman Filter (EKF) is given in [11], which also analyses the relative importance of process and measurement matrices. There are many other efforts related to dynamic state estimation [12, 13, 14, 15]. Iterative Kalman Filter (IKF) based state estimation for a non-linear dynamical system can be found in [16, 17].

However, all these dynamic state estimation strategies rely on the system being observable. Recently, sparsity-based approaches have been developed for static state estimation in low observable distribution systems [18],[19]. These methods estimate the states of the entire grid by exploiting the sparsity of measurements on a linear transformation basis. Matrix completion methods are employed for distribution system state estimation in [20]. Here, the authors employ a standard matrix completion method along with the linearized power flow constraints to capture the physical network constraints. Although the method effectively estimates the voltage phasors, it uses a semi-definite programming problem, which is computationally inefficient for large matrices. A comparative analysis of sparsity-based state estimation techniques, including matrix completion-based approaches, is presented in [21]. An attempt to exploit spatial and temporal correlation among the states using tensor completion is presented in [22]. All the above methods use a snapshot-based state estimation. Recently, authors in [23] attempted to extend MC to multiple snapshots (time series).

However, the method in [23] uses batch processing of data sets. That is, a large data matrix that includes different time steps is first created before implementing the state estimation approach. Therefore, this method requires large storage space and greater computational complexity that grows with size.

2.2 Distributed State Estimation

Sparsity-based state estimation approaches (compressive sensing(CS), matrix completion, and tensor completion) have been recently proposed to address the unobservability issue at the grid edge. These approaches exploit the network and data structure for state estimation at low levels of data availability. Matrix completion-based distribution system state estimation is employed in [20], [24]. Here, the authors employ a standard matrix completion method along with the linearized power flow constraints to capture the physical network constraints. Robust formulation of matrix completion against bad data for DSSE was proposed in [25]. A tensor completion-based SE approach is discussed in [26], which exploits both spatial and temporal correlation of the system states. Matrix completion and tensor completion methods exploit the underlying spatial correlation among the raw data and impute the missing entries by obtaining a suitable low-rank approximation of the incomplete matrix/tensor. In contrast, the CS method of state estimation considers that the original signal of interest is sparse in a specific transform domain. This is a reasonable assumption due to the spatial correlation between the loads and distributed generations as discussed in [27], [21],[28], [29]. A detailed comparative study among CS, matrix completion, and tensor completion based on estimation accuracy and computational complexity is presented in [21] and shows that CS-based state estimation outperforms matrix completion and tensor completion. Since the computational complexity of all sparsity-based state estimation strategies grows with the size of the distribution network, decentralized or distributed implementations are being considered to enable scalability. Additionally, aggregating sensor/meter data over a large network imposes significant stress on the communication infrastructure. This motivates the need to consider distributed compressive sensing-based state estimation. [30]

proposes a unique state estimation approach that employs matrix completion across space and CS across time. Authors in [23] propose an alternating minimization algorithm that reformulates the matrix completion approach for processing time-series measurements.

As discussed earlier, all the above methods treat the entire distribution network as a single entity and perform centralized estimation of the system states. On the contrary, several distributed methods focus on observing the system states by dividing the networks into smaller sub-networks. This reduces the communication overhead and computation time. Microgrid operation using a multi-agent distributed diffusion strategy is proposed in [31]. It uses a centralized operation strategy during normal operating conditions for resource optimization and a decentralized operation strategy during abnormal conditions. A PMU-based decomposition method for distributed state estimation and the effect of measurement error on state estimation accuracy is given in [32]. SE by dividing the whole system into different sub-regions based on their observability is studied in [33]. The improved robustness of state estimation using the matrix splitting strategy is discussed in [34]. [35] proposed a consensus-based multi-area state estimation using the Lagrange multiplayer function for the system modeling.

While all of the above research is devoted to estimating the states for a single time instant, there is a plethora of research that discusses distributed and dynamic state estimation [36], [37], [38]. A consensus filter-based dynamic state estimation is proposed in [36] to obtain the system states using local and neighboring gains. The local gain is obtained by minimizing the least square error between the actual and estimated variables. In contrast, the neighboring gain is obtained by solving a semi-definite program. Both gains are used to calculate the distributed consistent estimator after exchanging the neighboring node information. A fully distributed diffusion strategy for multi-agent microgrid control is proposed in [37] for distributed resources economic dispatch. It employs a two-level cooperative optimization and an additional gradient term for faster information sharing. The lower level control uses an adaptive droop scheme, and the upper level employs diffusion-based distributed communication among the neighbors for optimal microgrid management. Dynamic state estimation using an adaptive-and-then combine method is proposed in [38]. This paper proposes a mean

square error and SDP approach based on dynamic state estimation.

2.3 Effect of cyber attack on DSSE

The effect of cyber attacks on state estimation strategies is limited to the WLS approach [39],[40], [41]. A review of the FDI-type cyber attacks and their physical and economic impacts on power system state estimation is presented in [42]. Denial of service (DoS) attack on distributed state estimation is studied in [43]. A Kalman-filtered CS-based estimation approach in the presence of measurement loss is discussed in [44]. Authors in [45] construct a tolerable FDI attack that can bypass the bad data detection and degrade the performance of the state estimation. A recursive systematic convolutional (RSC) code to add redundancy in the power system states (cyber-attack scenario creation) and Kalman filter to recover the system states and SDP-based optimal feedback controller to regulate the system is proposed in [46]. Understanding the effects of cyber attacks on new state estimation strategies such as matrix completion (MC), tensor completion (TC), and compressive sensing-based state estimation is not explored. As per the author's knowledge, [47] is the only research paper that discusses the impact of cyber attacks on MC-based distribution system state estimation. In [47], a conventional centralized power distribution system is considered for state estimation, and the robustness to different cyber-attacks is analyzed. The effect of cyber-attacks on the distributed state estimation approach is limited.

2.4 Deep Unfolding for State Estimation

The complex voltages at each node and the branch currents of the power distribution network are usually considered to be the states of a power distribution network [48]. A complete survey on DSSE with their associated challenges for the modern power distribution network is presented [49]. Considering the limited available measurements, weighted least square (WLS) based DSSE is challenging for distribution networks. Researchers have proposed alternative mathematical formulations like least median of squares (LMS) and least trimmed squares

(LTS) in [50], least absolute value (LAV) estimator in [51]. However, these methods suffer from large memory requirements, high computational costs, and require measurement redundancy. A plethora of approaches based on probabilistic and statistical techniques have also been proposed. This includes empirical methods [52], expectation maximization [53], Gaussian mixture model [54], multivariate complex Gaussian modeling [55] and generalized maximum likelihood estimator in [56]. All these methods exploit the probability distribution of historical data and estimate the system states by generating pseudo-measurements. Hence, they suffer from low accuracy and high parameter sensitivity.

A recent trend is the application of sparsity-based approaches for DSSE. These methods successfully handle the unobservability issue due to limited measurement at the grid edge. Sparsity-based DSSE methods take advantage of the underlying smoothness of the measurement data, or sparsity of the system states on a linear transformation basis. These methods avoid the requirement of generating pseudo-measurements based on historical data. Sparsity-based DSSE using linearized power flow constraints is formulated as a low-rank matrix completion (MC) problem in [20]. A robust formulation of the MC approach to handling bad data is presented in [25]. To capture temporal dependency for better accuracy, a dynamic MC approach for DSSE is presented in [24]. Authors in [57] propose a tensor completion approach, exploiting both spatial and temporal correlation of the system states. Another sparsity-based technique, compressive sensing (CS), also finds application in DSSE [58]. [59]. DSSE, using compressive sensing, considers the signal to be sparse in a specific transformation basis, i.e., discrete cosine transform (DCT) or wavelet transform. A comparative analysis of CS, MC, and TC-based DSSE is provided in [21]. Results in [21] show that CS-based DSSE outperforms both MC and TC approaches. These estimators are found to have greater accuracy, even with limited measurements. However, the computational complexity of these estimators grows with an increase in system size; hence, these methods lack scalability.

This disadvantage of computational complexity and scalability in sparsity-aware techniques can be overcome by using deep learning (DL) approaches. The computation cost is shifted to the training phase of DL techniques, which need substantially less computational

effort during their real-time use. As a result, DL techniques are frequently utilized to replace computationally intensive algorithms. DSSE using neural network (NN) is formulated in [60], [8]. Authors in [60] develop an NN approximation of a minimum mean squared estimator, whereas in [8], a pseudo-measurement modeling strategy is proposed to generate pseudo-measurements for active and reactive power injections using artificial neural networks. [8] applies a weighted least squares estimator to estimate the system states. DSSE using a feed-forward neural network (FNN) is proposed in [61], which uses autoencoders to estimate the distribution grid state. Unfortunately, the FNN-based DSSE schemes are model-agnostic and necessitate a complicated tuning effort.

2.5 Microgrid Secondary Control

The secondary control in a microgrid is responsible for compensating for the voltage and frequency deviations generated from the primary droop control. To restore frequency and voltage to nominal values, an input-output feedback linearization approach is proposed in [62, 63, 64]. The authors used feedback linearization to transform the nonlinear model of microgrid systems into a linear multi-agent system, assuming that only a few DGs can directly obtain the reference values. A finite-time voltage restoration strategy that uses a simplified dynamic model of MG is discussed in [65]. A finite time frequency regulation strategy with a consensus-based distributed active power regulator is presented in [66]. A distributed averaging algorithm to overcome the frequency and voltage deviations by maintaining active power-sharing is proposed in [67, 68]. Critical bus voltage and frequency regulation by maintaining equal incremental generation cost in droop control and accurate reactive power sharing is discussed in [69]. Linear consensus protocols for frequency synchronization cum proportional active power-sharing problems have been widely used in microgrids [62, 64, 70]. However, these controllers suffer from asymptotic convergence. The asymptotic convergence in power system networks is undesirable because of fast-changing operating conditions due to highly intermittent renewable generation and fast dynamic loads [71]. Therefore, finite-time consensus protocols have gained importance in the secondary control of micro-

grids [72, 73, 74, 75, 76]. The distributed control of voltage and frequency in MG can be treated as a multiagent system, where the goal of each agent is to operate at the system frequency and voltage. Game theory offers solutions to various difficult and interesting problems in multi-agent (player) decision-making and control. Game theory [77] has successfully modeled strategic behavior, where the outcome for each player depends on the actions of himself/herself and all other players. Every player chooses a control action to minimize/maximize cost/performance objectives independently from others. The application of game theory in designing demand-side management in power systems has been studied [78, 79]. A coalition-based game theory approach for load balance is studied in [78]. A distributed consensus energy management system for false information and data detection is studied in [79]. Cooperative game theory was used to solve the optimal control problem of multiagent systems. To ensure consensus-seeking, [80] considers a combination of individual and team cost, and the minimization of this cost function results in a set of Pareto-efficient solutions. A tracking problem using a cooperative multiplayer game is solved in [81].

2.6 Reinforcement learning for MG Secondary Control

The large signal dynamic model of an inverter is non-linear, and the control strategy for a simplified model is not always applicable, especially for large disturbances. Hence, researchers are moving toward model-free, data-driven solutions. Data-driven algorithms like reinforcement learning (RL) have attracted much attention for solving complex control problems [82]. In the RL paradigm, an agent learns to make decisions by interacting with the environment. The agent aims to maximize some notion of cumulative reward through its actions. It learns through trial and error, receiving feedback through rewards or penalties. The combination of deep neural networks (DNN) with RL, known as deep reinforcement learning (DRL), has enabled the approximation of the inputs and outputs of complex systems. This approach is effective in addressing high-dimensional and highly complex control problems. Thus, RL can be very helpful in the control and management of large-scale renewable energy-integrated power grids.

Recently, authors in [83, 84, 85, 86] reported the use of multi-agent reinforcement learning (MARL) for secondary control of inverters. As agents interact with the environment and learn offline by simulating other agents' cooperative strategies, they can find the optimal strategy. After training, agents can adapt well to real-time decision-making in unknown power grids or microgrids. Distribution network management using maximum entropy reinforcement learning framework for voltage control is proposed in [84]. It also improves resilience to controller failures by introducing a communication consensus strategy to reduce network losses. An attention-based MATD3 algorithm for decentralized voltage control in an active distribution network is proposed in [85]. Authors in [86] use the DDPG algorithm for reactive power sharing and voltage control using a static var compensator (SVC). This method helps reduce the number of tap-changing operations in the transformer. All the above-proposed approaches need offline training of DRL agents which can be tedious and prohibitively complex for large systems.

Chapter 3

Background

This chapter provides a concise overview of research in state estimation and microgrid control. We explore different state estimation techniques such as WLS, CS, and MC, along with inverter modeling and primary and secondary control in inverter-based microgrids. These discussions form the basis for subsequent chapters in the dissertation.

3.1 State Estimation

State estimation, also known as network monitoring or observability analysis, is a computational technique used to infer the current state of a power distribution network based on available measurements. State estimation algorithms can estimate the voltage magnitudes, phase angles, and power injections at different nodes along with the line flows by utilizing data from various sensors, like smart meters, phasor measurement units (PMUs), supervisory control and data acquisition (SCADA) and other monitoring devices deployed throughout the network [1]. These estimates enable the network operators to gain a comprehensive understanding of the network's operating conditions, identify potential issues or anomalies, and make informed decisions to optimize grid operating performance and improve stability. The power distribution system is typically unobservable due to insufficient real-time measurements. While deploying more sensors can alleviate this issue. However, it presents new

challenges related to data aggregation and the underlying communication infrastructure. Therefore, developing state estimation methods that enhance situational awareness at the grid edge with limited measurements is critical. In this chapter, we focus on the different state estimation approaches used for state estimation with reliable accuracy using limited measurements in distribution grids. Before we jump for the sparsity-based approaches we would like to discuss the widely used WLS-based state estimation in power transmission networks.

Consider a power distribution system with $|\mathcal{P}|$ three-phase non-slack buses. The measurements in a distribution system include power injections at each bus, power flow, and current over each distribution line, along with bus voltages, which are related to the states (voltage magnitude and voltage angle) by the non-linear equation (3.1),

$$\mathbf{z} = h(\mathbf{x}) + \boldsymbol{\epsilon} \quad (3.1)$$

where, $\mathbf{z} \in \mathbb{C}^m$ represents the vector of measurements, $\mathbf{x} \in \mathbb{C}^n$ denotes the vector of states, and the function $h(\mathbf{x})$ capture the relationship between states and measurements. $\boldsymbol{\epsilon}$ represents the measurement noise vector. The state estimation task is to estimate $\mathbf{x}$ based on $\mathbf{z}$ and knowledge of $h(\cdot)$. This non-linear function $h(\cdot)$ relating the voltage states $\mathbf{x}$ and measurements $\mathbf{z}$ can be linearized around an operating point as in [87] resulting in,

$$\mathbf{x} = \mathbf{B}\mathbf{z} + \mathbf{w} \quad (3.2)$$

$$|\mathbf{x}| = \mathbf{C}\mathbf{z} + |\mathbf{w}|, \quad (3.3)$$

where $\mathbf{z} = ((\mathbf{P})^T, (\mathbf{Q})^T)^T$ represents the measurement vector of active and reactive power injections, $|\mathbf{x}|$ represents the magnitudes at all the buses.

$$\mathbf{B} = \left(\mathbf{Y}_{\mathbf{LL}}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1}, -j \mathbf{Y}_{\mathbf{LL}}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1} \right) \text{ and,}$$

$$\mathbf{C} = \left(|\text{diag}(\bar{\mathbf{w}})|^{-1} \Re(\text{diag}(\bar{\mathbf{w}}) \mathbf{M}) \right),$$

$\mathbf{w} = -\mathbf{Y}_{\text{LL}}^{-1}\mathbf{Y}_{\text{L0}}\mathbf{v}_0$ is the zero-load voltage

where $\mathbf{v}_0$ denotes the complex vectors collecting the three-phase nodal voltage at the slack bus. Here, $\mathbf{Y}_{\text{LL}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{P}|}$ and $\mathbf{Y}_{\text{L0}} \in \mathbb{C}^{|\mathcal{P}|\times 3}$ are the submatrices of the three-phase admittance matrix,

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{00} & \mathbf{Y}_{0\text{L}} \\ \mathbf{Y}_{\text{L0}} & \mathbf{Y}_{\text{LL}} \end{bmatrix} \in \mathbb{C}^{(|\mathcal{P}|+3)\times(|\mathcal{P}|+3)} \quad (3.4)$$

Next, we will discuss the weighted least squares (WLS) approach conventionally used for state estimation in the distribution grid.

3.1.1 WLS for state Estimation

In a traditional DSSE approach, the state vector is estimated from the measurement equation in (3.1) using the weighted least squares method. In particular, the DSSE problem is solved by.

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} [\mathbf{z} - h(\mathbf{x})]^\top \mathbf{W}^{-1} [\mathbf{z} - h(\mathbf{x})] \quad (3.5)$$

where, $\hat{\mathbf{x}}$ is the estimated state vector, $\top$ is the matrix transposition operation, and $\mathbf{W}$ denotes the weight matrix that represents the user's confidence in the measured data. The weighting matrix $\mathbf{W}$ is commonly taken as diagonal elements related to background noise covariance as $\mathbf{W} = \text{diag}\{\sigma_1^{-2}, \dots, \sigma_m^{-2}\}$, where σ_j^2 represents the variance of the measurement error corresponding to the j^{th} element of $\mathbf{z}$. This choice of the weight matrix is based on two assumptions: 1) the error vector (ϵ) has a Gaussian distribution with zero mean, and 2) the measurement errors of different elements of the measurement vector are statistically independent. Under these assumptions, the WLS problem is transformed to the maximum likelihood problem.

The solution for $\hat{\mathbf{x}}$ is obtained in an iterative fashion by linearizing (3.1) around the available estimate (at iteration k) and applying the Gauss-Newton algorithm to improve the

estimate, using the following equations:

$$\mathbf{G}(j)\Delta\mathbf{x}(j) = \mathbf{H}^\top(j)\mathbf{W}^{-1}[\mathbf{z} - h(\mathbf{x}(j))] \quad (3.6)$$

$$\hat{\mathbf{x}}(j+1) = \hat{\mathbf{x}}(j) + \Delta\mathbf{x}(j) \quad (3.7)$$

where $\mathbf{G}(j) = \mathbf{H}^\top(j)\mathbf{W}^{-1}\mathbf{H}(j)$ is the gain matrix at iteration j . The Jacobian matrix $\mathbf{H}(j)$ needed at each iteration is the first-order derivative of $h(\mathbf{x})$, with respect to $\mathbf{x}$, evaluated at $\hat{\mathbf{x}}(j)$, i.e., $\mathbf{H}(j) = \{\frac{\partial h(\mathbf{x})}{\partial \mathbf{x}}\}_{\mathbf{x}=\hat{\mathbf{x}}}$. The iterative procedure is terminated when the norm of the residual falls below a predefined value, i.e., for some $\delta > 0$, $\|\mathbf{z} - h(\hat{\mathbf{x}}(j))\|^2 \leq \delta$.

According to the update equation in (3.6), an invertible Jacobian matrix is required to implement WLS. However, this condition cannot be satisfied due to insufficient measurements at the grid edge. In other words, the low observability in distribution grids hinders the applicability of the conventional WLS method for DSSE. Correlations between states and measurements are exploited to address the low observability issue in DSSE. Specifically, a spatial correlation exists among DERs based on their geographical proximity [27]. Moreover, power injection/consumption patterns of nodes may exhibit temporal correlation. The presence of spatial and/or temporal correlation results in the sparsity of states/measurements on a linear transformation basis. Thus, very few measurements from network buses are required for state estimation. The following DSSE methods can exploit the sparsity characteristics.

3.1.2 Compressive Sensing

This technique exploits the sparsity of the signal of interest in a linear transformation basis. It is used to efficiently reconstruct a signal by finding sparse solutions to an under-determined linear system. Let $\mathbf{x} \in \mathbb{R}^n$ be the original state, compressible in a linear transformation basis such that,

$$\mathbf{x} = \Psi\mathbf{a} \quad (3.8)$$

where $\mathbf{a}$ has at most $k \ll n$ significant coefficients i.e., $\mathbf{x}$ is k -sparse in sparsifying basis Ψ . If the sensing mechanism is such that,

$$\mathbf{y} = \Phi \mathbf{x}; \mathbf{y} \in \mathbb{R}^m, \Phi \in \mathbb{R}^{m \times n}, \quad (3.9)$$

where Φ is a random measurement/projection matrix (e.g., matrix elements distributed as i.i.d. Gaussian random variable with mean 0 and variance $1/m$ or Bernoulli random variables), then the original state $\mathbf{x}$ can be reconstructed by solving the following l_1 minimization problem

$$\begin{aligned} \hat{\mathbf{a}} &= \min_{\mathbf{s}} \|\mathbf{s}\|_1 \\ \text{subject to } \mathbf{y} &= \Phi \Psi \mathbf{s} \end{aligned} \quad (3.10)$$

(3.2), (3.3)

Here, (3.2) represents the linearized power-flow constraints. This formulation can include noisy measurements as shown in [88]. The reconstructed state is

$$\hat{\mathbf{x}} = \Psi \hat{\mathbf{a}} \quad (3.11)$$

The result of the optimization problem in (3.10) provides an exact reconstruction with overwhelming probability [88] if there exists a $\delta \in (0, 1)$ such that,

$$(1 - \delta) \|\mathbf{s}\|_2^2 \leq \|\Phi \Psi \mathbf{s}\|_2^2 \leq (1 + \delta) \|\mathbf{s}\|_2^2, \quad (3.12)$$

holds for all K -sparse signal $\mathbf{s}$. This is called the Restricted Isometry property (RIP) of order K . The ratio M/N is termed the Compressed Measurement Ratio (CMR). More details regarding 1-D CS-based DSSE can be found in [27].

3.1.3 Matrix completion

To utilize the matrix completion approach for DSSE, a structured matrix $\mathbf{X}$ whose rows correspond to measurement locations and columns correspond to measurement types (e.g. power or voltage) is constructed for a given time. For example, each row represents a bus and each column represents a measurement associated with the bus [20]. Therefore, for every bus $b \in \mathcal{P}$, the corresponding row in the matrix $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ contains:

$$[\Re(v_b), \Im(v_b), |v_b|, \Re(s_b), \Im(s_b)] \quad (3.13)$$

where $n_1 = |\mathcal{P}|$ and $n_2 = 5$ quantities per row. Since the distribution grid is unobservable, $\mathbf{X}$ will be an incomplete matrix. The smoothness in the spatial variation of the physical power system quantities (voltage, power, etc) translates into the low-rank property for this matrix $\mathbf{X}$. This low-rank property can be exploited to impute the unknown entries of $\mathbf{X}$. Specifically, matrix completion aims to determine the unknown elements in the matrix by minimizing the rank of the matrix. The convex relaxation for the rank of a matrix is the nuclear norm. The sampling operator $P_\Omega(\mathbf{X})$ represents the observation matrix as,

$$[P_\Omega(\mathbf{X})]_{ij} = \begin{cases} \mathbf{X}_{ij}, & (ij) \in \Omega \\ 0, & \text{otherwise} \end{cases} \quad (3.14)$$

Therefore, the optimization formulation (3.15) recovers the complete low rank matrix, as

$$\begin{aligned} \hat{\mathbf{L}} = \min_{\mathbf{L} \in \mathbb{R}^{n_1 \times n_2}} \quad & \|\mathbf{L}\|_* \\ \text{subject to} \quad & \|P_\Omega(\mathbf{X}) - P_\Omega(\mathbf{L})\|_2^2 \leq \epsilon \end{aligned} \quad (3.15)$$

(3.2), (3.3)

where, (3.2) and (3.3) captures the linearized power-flow constraint relating voltage states $\mathbf{x}$ to power measurements $\mathbf{z}$. More information on matrix completion-based DSSE can be found in [20].

3.2 Microgrid Secondary control

The global energy landscape has radically transformed in recent years. Some driving forces for the transformation include: (1) pressing concerns posed by climate change; (2) the rise in energy prices, and (3) increasing concerns about energy security and reliability. As traditional centralized power systems grapple with these challenges, microgrids (MGs) with distributed energy resources (DERs) are emerging as a novel alternative. MGs are localized, autonomous energy systems integrating various DERs, energy storage devices, loads, and other power electronic equipment [89, 90]. MGs can operate with the main grid or independently (island mode of operation) to enhance the overall grid resilience and efficiency [90, 91].

In the conventional power grid, synchronous generators are responsible for maintaining voltage and frequency stability because of their inbuilt inertia [3]. At the same time, ensuring frequency and voltage stability is crucial in MGs. Generally, the microgrids are flooded with IBRs that lack inertia. Droop control has emerged as a dominant solution to regulate voltage and frequency in MGs because of its simple implementation [92]. Moreover, droop control only relies on local measurements and requires no communication [93]. However, due to its inherent characteristics, droop control may not be precise and can cause voltage and frequency deviations from their nominal values [4]. Therefore, a secondary level of hierarchical control becomes necessary to compensate for the voltage and frequency deviations from its normal values [94]. The secondary control acts as a compensation mechanism. The approaches to implement secondary control in MG architecture are centralized, decentralized, or distributed. Details of the different secondary control architecture in MG setup can be found in [95]. Centralized control has an easier decision-making process and coordination [4]. However, it suffers from large communication bandwidth requirements, which leads to decision delays and low reliability. On the other hand, the decentralized controller does not require exclusive communication infrastructure (CI) to restore the frequency and voltage. Each DER restores its frequency and voltage amplitudes to nominal values individually. However, the CI is required for transferring data from the tertiary control and MG centralized control layers [96]. On the contrary, distributed secondary controllers use

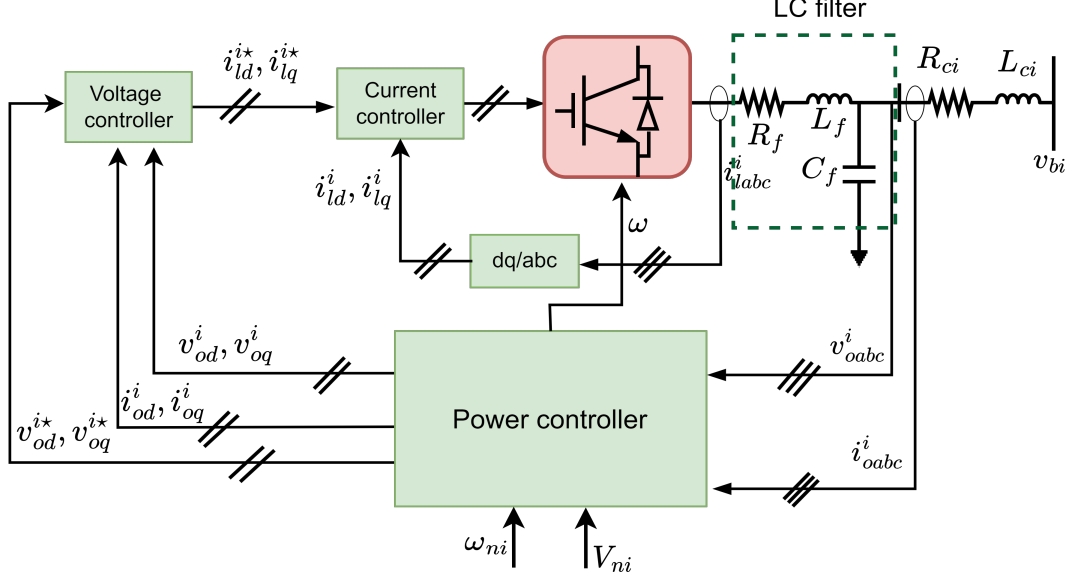


Figure 3.1: The basic control architecture in an inverter

a low bandwidth sparse communication network in which the inverters are only required to interact with the neighboring inverters. In the next subsection, we will briefly explain the primary and secondary control of inverter-based microgrids.

3.2.1 Primary control of MG

AC microgrid consists of DC sources behind a DC/AC inverter and an LC filter connected to the point of common coupling (PCC) through an RL output connector [97]. The inverters normally operate on voltage control mode during the islanded operation of MG [98]. A typical DG model with the inverter control loops is shown in Fig. 3.1. The major control loops of the inverter control are the current control loop, the voltage control loop, and the power control loop, which employs the droop controller. A detailed mathematical model of the control loops is explained in [99]. The power control loop generates the voltage references V_{od}^*, V_{oq}^* based on the set points provided by the secondary control, where, V_{od}^* and V_{oq}^* are direct and quadrature axis voltages, respectively. The voltage controller generates the direct and quadrature axis current references i_{Ld}^{i*} and i_{Lq}^{i*} , respectively. Then, the current control loop is activated to regulate the outputs of the inverter by generating sinusoidal pulse

width modulation (SPWM). The power control loop employs droop control. Droop-based control is an autonomous control that requires no communication links among the DERs. In an islanded microgrid, the objective of droop control is to maintain voltage and frequency stability keeping the active and reactive power sharing. The two governing equations for the primary droop control of i^h DG are as follows [94].

$$\begin{cases} \omega_i = \omega_n - m_{pi}P_i \\ V_i^{od} = V_n - n_{qi}Q_i \\ V_i^{oq} = 0 \end{cases} \quad (3.16)$$

where, ω_n and V_n are the nominal system frequency and nominal voltage amplitude, respectively. ω_i is the output frequency. V_i^{od} , and V_i^{oq} are the direct and quadrature axis voltages, respectively. P_i and Q_i are the measured active and reactive power, respectively. m_{pi} and n_{qi} are the droop gain for frequency and voltage, respectively. The value of m_{pi} and n_{qi} are calculated as,

$$m_{pi} = \frac{\omega_{\max} - \omega_{\min}}{P_{\max}}, \quad n_{qi} = \frac{V_{i\max} - V_{i\min}}{Q_{\max}}.$$

The measured power P_i , and Q_i are obtained by filtering the converter output through a low pass filter with a cutoff frequency of ω_{ci} . Mathematically,

$$P_i = \frac{1}{s + \omega_{ci}} p_i, \quad (3.17)$$

$$Q_i = \frac{1}{s + \omega_{ci}} q_i, \quad (3.18)$$

where, p_i and q_i are the DG active and reactive power output, respectively. Generally, for simple control purposes V_{oq} is set to zero and hence, the voltage magnitude is expressed as $V_i = \sqrt{(V_i^{od})^2 + (V_i^{oq})^2} = V_i^{od}$. Hence, the primary droop control in (9.2b) is expressed as,

$$\omega_i = \omega_n - m_{pi}P_i \quad (3.19)$$

$$V_i = V_n - n_{qi}Q_i \quad (3.20)$$

3.2.2 Secondary Control of Microgrid

The conventional way of compensating for the deviation in frequency and voltage amplitude after the primary droop control is through a centralized controller. The centralized controller employs a proportional-integral controller given as [94, 100],

$$\delta\omega_i = K_{P\omega}(\omega_n - \omega_i) + K_{I\omega} \int (\omega_n - \omega_i) dt. \quad (3.21)$$

$$\delta V_i = K_{PV}(V_n - V_i) + K_{IV} \int (V_n - V_i) dt, \quad (3.22)$$

Where, K_{PV} , K_{IV} , $K_{P\omega}$ and $K_{I\omega}$, are the proportional and integral (PI) gain for the voltage and frequency controller, respectively. V_n and ω_n is the nominal voltage and reference frequency respectively. δV_i and $\delta\omega_i$ are the offsets in voltage and frequency, respectively (generated by the secondary control). These offset parameters are sent to the primary droop controller to compensate for the voltage and frequency. Although the controller works well, designing and tuning the PI parameters is difficult. Furthermore, the centralized controller needs a high bandwidth communication network and suffers from the single-point failure issue. Therefore, we propose a game-theoretic approach for the secondary control of frequency in islanded microgrid. This approach requires localized communications limited among neighboring IBRs. Next, we discuss a general model for the communication network in MG.

3.2.3 Communication network modeling

Consider a communication network that enables information exchange among inverters. This network is represented by a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a nonempty finite set of N vertices (an example is shown in Fig. 3.2). The node-set, the edges, and the graph adjacency matrix are defined as $\mathcal{V} = \{v_1, v_2 \dots v_N\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$,

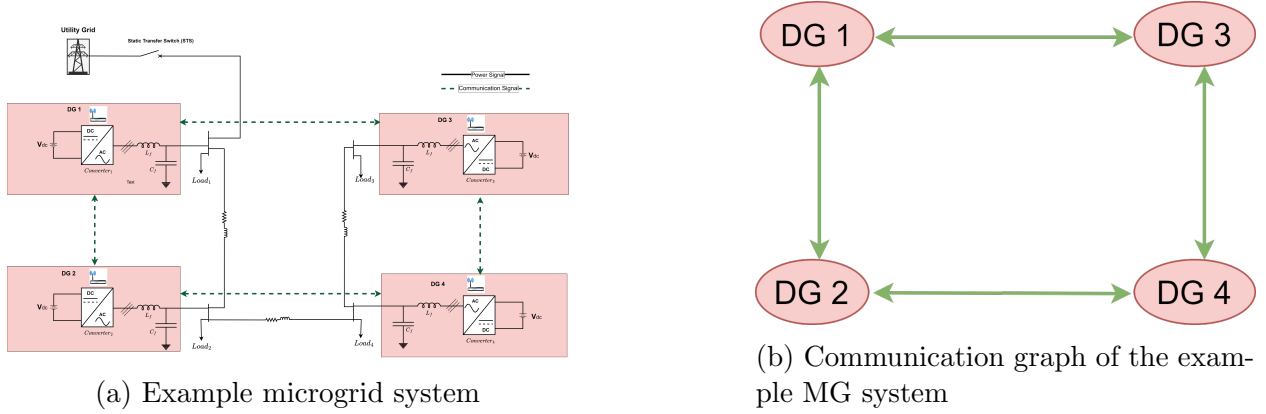


Figure 3.2: Microgrid architecture and the communication infrastructure

respectively. The IBRs are considered the vertices of the communication graph, and the edges of the communication network form the links. The edges from node v_j to v_i is denoted by (v_j, v_i) , i.e., node v_j receives information from v_i . The weight of the communication link is given by $a_{ij} > 0$ if $(v_j, v_i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The in-degree matrix is defined by $\mathbf{D} = \text{diag}(d_i) \in \mathbb{R}^{N \times N}$ with $d_i = \sum_{j=1}^N a_{ij}$. Finally, the graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The Laplacian matrix $\mathbf{L}$ has all row sums equal to zero, i.e., $\mathbf{L}\mathbf{1}_N = 0$, where $\mathbf{1}_N$ is a column of all ones. The graph $\mathcal{G}$ is called balanced if its Laplacian matrix $\mathbf{L}$ meets $\mathbf{1}_N^T \mathbf{L} = 0$. A graph is said to have a spanning tree if there exists a node (called the root) that has a directed path from this node to other nodes in the graph [101]. The graph is said to be connected if the algebraic connectivity i.e., the second smallest eigenvalue, is non-zero.

Chapter 4

Dynamic State Estimation

In this chapter, we discuss the proposed novel dynamic matrix completion-based state estimations, LW-MC and BMC-KF, for dynamic state estimation in low-observable distribution grids.

4.1 Introduction

Matrix completion methods are employed for distribution system state estimation in [20]. Here, the authors employ a standard matrix completion method along with the linearized power flow constraints to capture the physical network constraints. Although the method effectively estimates the voltage phasors, it uses a semi-definite programming problem, which is computationally inefficient for large matrices. A comparative analysis of sparsity-based state estimation techniques, including matrix completion-based approaches, is presented in [21]. An attempt to exploit both spatial and temporal correlation among the states using tensor completion is presented in [22]. All the above methods use a snapshot-based state estimation. Recently, authors in [23] attempted to extend MC to multiple snapshots (time series). However, the method in [23] uses batch processing of data sets. That is, a large data matrix that includes different time steps is first created before implementing the state estimation approach. Therefore, this method requires large storage space and greater

computational. Hence, we propose a dynamic matrix completion approach for DSSE that exploits both spatial and temporal sparsity in the raw measurements.

4.2 Background

Matrix completion-based state estimation for a three-phase distribution network based on a single snapshot of information is developed in [20]. In this context, consider a multi-phase power distribution network with bus 1 as the slack bus; i.e., the bus which connects the distribution grid with the rest of the power system and the set of phases at non-slack buses represented by $\mathcal{B}$. We construct the data matrix $\mathbf{Z}_t \in \mathbb{R}^{n_1 \times n_2}$ at a particular time instant t as,

$$\mathbf{Z}_t = [\Re(\mathbf{v}_t^b), \Im(\mathbf{v}_t^b), |\mathbf{v}_t^b|, \Re(\mathbf{s}_t^b), \Im(\mathbf{s}_t^b)], \quad (4.1)$$

where, $\Re(\mathbf{v}_t^b)$ and $\Im(\mathbf{v}_t^b)$ are the real and imaginary parts of the voltage phasor at $b \in \mathcal{B}$; $|\mathbf{v}_t^b|$ is the voltage magnitude at bus $b \in \mathcal{B}$; $\Re(\mathbf{s}_t^b)$ and $\Im(\mathbf{s}_t^b)$ are the active and reactive power injections at bus $b \in \mathcal{B}$; $n_1 = |\mathcal{B}|$ and $n_2 = 5$ (representing the total number of power system variables of interest). Each row $b \in \mathcal{B}$ in the matrix $\mathbf{Z}_t$ represents a bus in the network and the column represents the five measurement variables. In a typical distribution system, not all the buses have sensors/meters. Therefore, $\mathbf{Z}_t$ has many missing elements. The goal of state estimation is to estimate the missing elements of $\mathbf{Z}_t$ based on the limited number of observed entries. In classic matrix completion, the unknown entries of a matrix $\mathbf{Z}_t$ are determined by a suitable low rank approximation[102]. Let $\Psi \subseteq \{1, \dots, n_1\} \times \{1, \dots, n_2\}$ describe the known elements in $\mathbf{Z}_t$. The sampling operator $\mathbf{S}_\Psi(\mathbf{Z}_t)$ represents the observation matrix as,

$$[\mathbf{S}_\Psi(\mathbf{Z}_t)]_{ij} = \begin{cases} \mathbf{Z}_{t_{ij}}, & (i, j) \in \Psi \\ 0, & \text{otherwise} \end{cases}$$

The goal of MC is to determine the low-rank matrix $\mathbf{X}_t$ based on $\mathbf{S}_\Psi(\mathbf{Z}_t)$. This is accomplished by exploiting the smoothness in the data that composes $\mathbf{Z}_t$. That is, the unknown entries of matrix $\mathbf{Z}_t$ are imputed by minimizing its rank. As the rank function is non-convex,

matrix completion is NP-hard. To ensure convexity, the nuclear norm is minimized resulting in a constrained formulation corresponding to relax the convexity [102]. The MC problem formulation is given as,

$$\begin{aligned} \hat{\mathbf{X}}_t = \arg \min_{\mathbf{X}_t \in \mathbb{R}^{n_1 \times n_2}} \quad & ||\mathbf{X}_t||_* \\ \text{subject to} \quad & ||\mathbf{S}_\Psi(\mathbf{X}_t) - \mathbf{S}_\Psi(\mathbf{Z}_t)||_F \leq \delta \end{aligned} \quad (4.2)$$

where, $\mathbf{X}_t \in \mathbb{R}^{n_1 \times n_2}$ is the decision variable, $||\cdot||_F$ is the Frobenius norm, $||\cdot||_*$ is the nuclear norm, $\delta \geq 0$ is the parameter tuned based on the measurement noise [103]. The formulation (4.2) can be improved by including power flow constraints. Specifically, linearized power flow constraints (4.3), (4.4) can be included in (4.2) and maintain the convexity. These constraints correspond to [104],

$$\mathbf{v}_t = \mathbf{B} \begin{bmatrix} \Re(\mathbf{s}_t) \\ \Im(\mathbf{s}_t) \end{bmatrix} + \mathbf{w}, \quad (4.3)$$

$$|\mathbf{v}_t| = \mathbf{C} \begin{bmatrix} \Re(\mathbf{s}_t) \\ \Im(\mathbf{s}_t) \end{bmatrix} + |\mathbf{w}|. \quad (4.4)$$

here, $\mathbf{v}_t$ and $|\mathbf{v}_t|$ are the nodal complex voltage vector and voltage magnitude vector at time t respectively, $\mathbf{s}_t$ is the complex power injection at PQ buses; the matrices $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{w}$ are given as [104],

$$\mathbf{B} = \left(\mathbf{Y}_{\mathbf{LL}}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1}, -j \mathbf{Y}_{\mathbf{LL}}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1} \right) \text{ and,}$$

$$\mathbf{C} = \left(\text{diag}(\bar{\mathbf{w}})^{-1} \Re(\text{diag}(\bar{\mathbf{w}})) \mathbf{B} \right) \text{ and,}$$

$$\mathbf{w} = -\mathbf{Y}_{\mathbf{LL}}^{-1} \mathbf{Y}_{\mathbf{L1}} \mathbf{v}_1 \text{ is the zero-load voltage,}$$

where, $\mathbf{v}_1$ denote the complex three phase nodal voltage vector at the slack bus; $\mathbf{Y}_{\mathbf{LL}}$ and $\mathbf{Y}_{\mathbf{L1}}$ are the sub-matrices of the three-phase admittance matrix given in (6.2) as,

$$\mathbf{Y} := \begin{bmatrix} \mathbf{Y}_{11} \in \mathbb{C}^3 & \mathbf{Y}_{1L} \in \mathbb{C}^{3 \times n_1} \\ \mathbf{Y}_{L1} \in \mathbb{C}^{n_1 \times 3} & \mathbf{Y}_{LL} \in \mathbb{C}^{n_1 \times n_1} \end{bmatrix} \quad (4.5)$$

Although state estimation based on the above method (i.e., solving (2) subject to constraints (4.3) and (4.4)) can accurately estimate the voltage phasors under low-observability conditions, it only uses the measurement information of the current time step t . It does not account for temporal dependencies among the states. In the next section, we introduce two dynamic state estimation methods to obtain the missing entries of the matrix $\mathbf{Z}_t$ by considering both current measurements and the previously estimated states.

4.3 Proposed Approach

As discussed earlier, each row of the data matrix $\mathbf{Z}_t$ represents a location and each column represents a measurement variable (voltage, power, etc.). Thanks to the physics of power flow, there are two dependencies that exist among the rows and columns of $\mathbf{Z}_t$. Among the rows, there exists a system dependent spatial correlation and among the columns there exist a co-relation among different type of measurements [105]. Therefore, the data matrix is low rank. Furthermore, the power system is a quasi-static system with slowly varying states, we can assume the current state value will be related to the previous state value with the level of dependence being a function of the data sampling rate. Hence, we could use the previously available measurements or estimated state information to estimate the current state. The power system dynamics corresponding to each system variable can be approximated by a linear dynamical model given as [6], [16], [106],

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{u}_t \quad (4.6)$$

where, $\mathbf{x}_t \in \mathbb{R}^{5n_1}$, $\mathbf{A}_t \in \mathbb{R}^{5n_1 \times 5n_1}$ are the state variable and state transition matrix respectively, and $\mathbf{u}_t \in \mathbb{R}^{5n_1}$ is the white Gaussian noise with zero mean and known variance.

Leveraging this dependency across time, we propose two methods to dynamically recover a low-rank matrix for distribution system dynamic state estimation in the following subsections.

4.3.1 Locally Weighted matrix completion (LW-MC)

As discussed earlier the observed low rank matrix $\mathbf{X}_t$ can be modelled as a discrete time dynamic process. The temporal variations of the state matrix are captured in a general form as,

$$\mathbf{X}_t = f(\mathbf{X}_1, \dots, \mathbf{X}_{t-1}) + \boldsymbol{\epsilon}_t, \quad (4.7)$$

where, $\boldsymbol{\epsilon}_t$ is i.i.d $\mathcal{N}(0, \sigma_\epsilon^2 \mathbf{I})$ and represents the model uncertainty.

The low rank matrix $\mathbf{X}_t$ at each time step t is observed through a linear operator $\mathbf{M}_t : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^m$

$$\mathbf{z}_t = \mathbf{M}_t \mathbf{X}_t + \boldsymbol{\zeta}_t, \quad (4.8)$$

where, $\mathbf{z}_t \in \mathbb{R}^m$ and $\boldsymbol{\zeta}_t \in \mathbb{R}^m$ are the measurement vector and measurement noise at time t , respectively. It is assumed that the measurement noise is i.i.d. $\mathcal{N}(0, \sigma_\zeta^2 \mathbf{I})$. The sensing matrices $\mathbf{M}_t$ are i.i.d. uniformly distributed on the set given as,

$$\mathcal{X} = \{e_i(n_1)e_j^\top(n_2), 1 \leq i \leq n_1, 1 \leq j \leq n_2\},$$

where, $e_i(n)$ are the canonical basis vectors in $\mathbb{R}^n$. We have considered that states up to d time steps are observed and our goal is to recover $\{\mathbf{X}_t\}_{t=1}^d$ jointly from $\{\mathbf{z}_t\}_{t=1}^d$. The LW-MC approach assigns non-negative weights to the previously estimated matrices similar to the approach in [107]. However, in contrast to [107], we introduce the linearized power flow constraints to guide the estimation process. The complete formulation of the LW-MC

correspond to,

$$\hat{\mathbf{X}}_t = \arg \min_{\mathbf{X}_t \in \mathbb{R}^{n_1 \times n_2}} \|\mathbf{X}_t\|_* + \frac{\lambda}{2} \sum_{t=1}^d w_t \|\mathbf{M}_t \mathbf{X}_t - \mathbf{z}_t\|_2^2 \quad (4.9)$$

subject to: (4.3) and (4.4)

where, λ is the penalty parameter, $\{w_t\}_1^d$ are the non negative weights and $\sum_{t=1}^d w_t = 1$. The optimal weights are calculated by solving the quadratic program as,

$$\{w_t^*\}_{t=1}^d = \arg \min_{\sum_t w_t = 1; w_t \geq 0} \sum_{t=1}^d w_t^2 + \sum_{t=1}^{d-1} p_t k w_t^2, \quad (4.10)$$

where, $k := \sigma_\epsilon^2 / \sigma_\zeta^2$ and $p_t = (d - t), 1 \leq t \leq d$. The corresponding analytical solution of (4.10) is [107].

$$w_t^* = \frac{1}{\sum_{t=1}^d \frac{1}{p_t k}} \frac{1}{1 + p_t k}, \quad 1 \leq t \leq d. \quad (4.11)$$

4.3.2 Dynamic Bayesian Matrix completion with KF Prediction

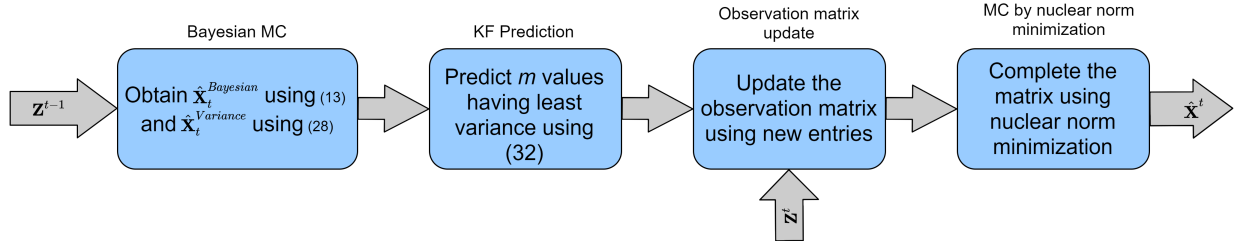


Figure 4.1: Block diagram of the proposed BMC-KF approach

The semidefinite program presented in (4.2) along with the constraints (4.3) and (4.4) can be computationally inefficient for large networks. In addition to this, prior knowledge on the rank of the matrix is generally required for this formulation [108]. Furthermore, this formulation requires extensive parameter tuning as well as data and application-dependent supervision. In order to overcome these shortcomings, a hierarchical sparse Bayesian matrix completion framework with KF prediction (as shown in Fig.4.1) is proposed. The dynamic

Bayesian MC completion based state estimation involves two steps; matrix completion and state prediction. In the first step, we impute the missing entries by applying the Bayesian MC on the incomplete observation matrix. In step 2, we apply the Kalman filter prediction to estimate the states from the current observation and previously estimated estimates. A pseudo code of the dynamic BMC-KF approach is also presented in Algorithm 1. In this section, we will first introduce the Bayesian matrix completion process prior to integration with KF prediction.

Bayesian Matrix Completion

The objective of Bayesian matrix completion is to get the Bayesian estimate $\hat{\mathbf{X}}_t^{Bayesian}$ along with a measure of the uncertainty associated with the estimate. In essence, the goal is to get a posterior distribution of $\mathbf{X}_t$ given the observation matrix $\mathbf{Z}_t$. The unknown low-rank matrix $\mathbf{X}_t$ at any time instant t can be factorized into two matrices as,

$$\mathbf{X}_t = \mathbf{U}_t \mathbf{V}_t^\top, \quad (4.12)$$

where, $\mathbf{U}_t$ is an $n_1 \times r$ matrix, and $\mathbf{V}_t$ is an $n_2 \times r$ matrix such that $\text{rank}(\mathbf{X}_t) = r$. Although $\mathbf{U}_t$ and $\mathbf{V}_t$ are time dependent, it is fixed for a particular time step; hence, we omit the time subscript for notational simplicity in the rest of the article. Let $\mathbf{u}_{\cdot i}$ and $\mathbf{v}_{\cdot i}$ as the i^{th} column of matrix $\mathbf{U}$ and $\mathbf{V}$, respectively. The estimated matrix $\hat{\mathbf{X}}_t^{Bayesian}$ is expressed as the sum of the outer-products of the columns of $\mathbf{U}$ and $\mathbf{V}$ as,

$$\hat{\mathbf{X}}_t^{Bayesian} = \sum_{i=1}^k \mathbf{u}_{\cdot i} \mathbf{v}_{\cdot i}^\top, \quad (4.13)$$

where, $k \geq r$, The i^{th} row of matrix $\mathbf{U}$ and $\mathbf{V}$ is represented by $\mathbf{u}_{i\cdot}$ and $\mathbf{v}_{i\cdot}$, respectively. The columns of the matrix $\mathbf{U}$ and $\mathbf{V}$ are set to zero to obtain the low rank matrix $\hat{\mathbf{X}}_t^{Bayesian}$. This condition is achieved by associating Gaussian priors of precisions γ_i to the columns of $\mathbf{U}$ and $\mathbf{V}$ as given in [108]. The conditional distribution of the observation matrix $\mathbf{Z}_t$ is given

as,

$$\mathbf{p}(\mathbf{Z}_t|\mathbf{U}, \mathbf{V}) = \prod_{(i,j) \in \Psi} \mathcal{N}(Z_{t_{ij}}|X_{t_{ij}}, \beta^{-1}), \quad (4.14)$$

where, β is the noise precision. The joint distribution is therefore given as,

$$p(\mathbf{Z}_t, \mathbf{U}, \mathbf{V}, \boldsymbol{\gamma}) = p(\mathbf{Z}_t|\mathbf{U}, \mathbf{V})\mathbf{p}(\mathbf{U}|\boldsymbol{\gamma})\mathbf{p}(\mathbf{V}|\boldsymbol{\gamma})\mathbf{p}(\boldsymbol{\gamma}). \quad (4.15)$$

The complete evaluation of posterior distributions of the latent variables $(\mathbf{U}, \mathbf{V}, \boldsymbol{\gamma})$ as given in [108] are,

$$\mathbf{q}(\mathbf{u}_{i\cdot}) = \mathcal{N}(\mathbf{u}_{i\cdot} | \langle \mathbf{u}_{i\cdot} \rangle, \boldsymbol{\Sigma}_i^u), \quad (4.16)$$

where, the mean and covariance is defined as,

$$\langle \mathbf{u}_{i\cdot} \rangle^\top = \langle \beta \rangle \boldsymbol{\Sigma}_i^u \langle \mathbf{V}_i \rangle^\top \mathbf{z}_{i\cdot}^\top, \quad (4.17)$$

$$\boldsymbol{\Sigma}_i^u = (\langle \beta \rangle \langle \mathbf{V}_i^\top \mathbf{V}_i \rangle + \boldsymbol{\Gamma})^{-1}, \quad (4.18)$$

here,

$$\langle \mathbf{V}_i^\top \mathbf{V}_i \rangle = \sum_{j:(i,j) \in \Psi} (\langle \mathbf{v}_{j\cdot}^\top \rangle \langle \mathbf{v}_{j\cdot} \rangle + \boldsymbol{\Sigma}_j^v), \quad (4.19)$$

and

$$\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma}).$$

Similarly, the posterior density of j^{th} row of $\mathbf{V}$ is found as,

$$\mathbf{q}(\mathbf{v}_{j\cdot}) = \mathcal{N}(\mathbf{v}_{j\cdot} | \langle \mathbf{v}_{j\cdot} \rangle, \boldsymbol{\Sigma}_j^v), \quad (4.20)$$

where, the mean and covariance are defined as,

$$\langle \mathbf{v}_{j\cdot} \rangle^\top = \langle \beta \rangle \boldsymbol{\Sigma}_j^v \langle \mathbf{U}_j \rangle^\top \mathbf{z}_{j\cdot}^\top, \quad (4.21)$$

$$\boldsymbol{\Sigma}_j^v = (\langle \beta \rangle \langle \mathbf{U}_j^\top \mathbf{U}_j \rangle + \boldsymbol{\Gamma})^{-1}. \quad (4.22)$$

The posterior density of γ_i becomes a gamma distribution

$$\mathbf{q}(\gamma_i) \propto \gamma_i^{(a-1+\frac{n_1+n_2}{2})} \exp\left(-\gamma_i \frac{2b + \langle \mathbf{u}_i^\top \mathbf{u}_i \rangle + \langle \mathbf{v}_i^\top \mathbf{v}_i \rangle}{2}\right) \quad (4.23)$$

with mean,

$$\langle \gamma_i \rangle = \frac{2a + n_1 + n_2}{2b + \langle \mathbf{u}_i^\top \mathbf{u}_i \rangle + \langle \mathbf{v}_i^\top \mathbf{v}_i \rangle} \quad (4.24)$$

The required expectations are given by

$$\langle \mathbf{u}_i^\top \mathbf{u}_i \rangle = \langle \mathbf{u}_i \rangle^\top \langle \mathbf{u}_i \rangle + \sum_j \left(\Sigma_i^u \right)_{ii}, \quad (4.25)$$

$$\langle \mathbf{v}_i^\top \mathbf{v}_i \rangle = \langle \mathbf{v}_i \rangle^\top \langle \mathbf{v}_i \rangle + \sum_j \left(\Sigma_i^v \right)_{ii}, \quad (4.26)$$

The approximate posterior distribution of β is given as,

$$\langle \beta \rangle = \frac{pn_1n_2}{\langle \|\mathbf{Z}_t - S_\Psi(\mathbf{U}\mathbf{V}^\top)\|_F^2 \rangle}, \quad (4.27)$$

where, p is the fraction of the observed entries of $\mathbf{X}_t^{Bayesian} \in \mathbb{R}^{n_1 \times n_2}$.

The estimated variance $\hat{\mathbf{X}}_t^{Variance} = Var(\mathbf{U}\mathbf{V}^\top)$ is given as,

$$Var(\mathbf{U}\mathbf{V}^\top)_{i,j} = tr(\mathbf{v}_j \cdot \Sigma_i^u \mathbf{v}_j^\top) + tr(\mathbf{u}_i \cdot \Sigma_j^v \mathbf{u}_i^\top) + tr(\Sigma_j^v \Sigma_i^u), \quad (4.28)$$

where, $Var(\cdot)$ is the variance and $tr(\cdot)$ is the trace.

Kalman Filter Prediction

As discussed earlier, the Bayesian matrix completion framework estimates the low rank matrix (4.13) and the associated uncertainty (4.28). Next, we aim to predict a subset of state values for the next time step. For this purpose we choose m number of values having the least variance from the matrix $\hat{\mathbf{X}}_t^{Bayesian}$. This choice ensures that the prediction error variance is limited. We predict these values using Kalman filter prediction as discussed next.

The power distribution network is a quasi-static model, where the states vary slowly. The dynamic model of a power system state variables in discrete-time can be represented as [109]

$$\mathbf{x}_{t+1} = \mathbf{F}_t \mathbf{x}_t + \mathbf{g}_t + \boldsymbol{\zeta}_t, \quad (4.29)$$

where, $\mathbf{x}_t \in \mathbb{R}^{5n_1}$ is obtained by vectorizing the matrix $\mathbf{X}_t$, $\mathbf{F}_t \in \mathbb{R}^{5n_1 \times 5n_1}$ is the matrix representing the state transition between two time instants, $\mathbf{g}_t \in \mathbb{R}^{5n_1}$ is the trend of the state trajectory, and $\boldsymbol{\zeta}_t \in \mathbb{R}^{5n_1}$ is the white Gaussian noise with zero mean and known variance. There have been various methods proposed in the literature to identify the model parameters $(\mathbf{F}_t, \mathbf{g}_t)$. Linear exponential smoothing method is proposed in [16]. In [110], authors have captured the dynamics of the states through the process noise $(\boldsymbol{\zeta}_t)$ only. This is achieved by replacing $\mathbf{F}_t$ by an identity matrix and $\mathbf{g}_t$ by zero. In this paper, we have used Holt's 2 parameter linear exponential smoothing method [109]. This model gives decreasing weights to past estimations. The parameters, $\mathbf{F}_t$ and $\mathbf{g}_t$ in every time step are calculated as,

$$\mathbf{F}_t = \alpha(1 + \beta)\mathbf{I}, \quad (4.30)$$

$$\mathbf{g}_t = (1 + \beta)(1 - \alpha)\tilde{\mathbf{x}}_{t-1} - \beta\mathbf{s}_{t-2} + (1 - \beta)\mathbf{b}_{t-1}, \quad (4.31)$$

where,

$$\mathbf{s}_{t-2} = \alpha\hat{\mathbf{x}}_{t-1} + (1 - \alpha)\tilde{\mathbf{x}}_{t-2},$$

$$\mathbf{b}_{t-1} = \beta[\mathbf{s}_{t-2} - \mathbf{s}_{t-3}] + (1 - \beta)\mathbf{b}_{t-3},$$

where, $\mathbf{I}$ is the identity matrix and α, β are scalar tuning parameters that take values in the range of 0 to 1. The parameter $\boldsymbol{\zeta}_t$ represents the random fluctuations in the system states due to the load generation schedule, and network configuration associated with the corresponding time step. The mean value of $\boldsymbol{\zeta}_t$ is assumed to be zero and its variance is determined using historical data. (4.29) suggests that, the system states at time instant t can be predicted based on the information obtained at $t - 1$. This is achieved by applying

the conditional expectation operator on (4.29) leading to,

$$\tilde{\mathbf{x}}_{t+1} = \mathbf{F}_t \hat{\mathbf{x}}_t + \mathbf{g}_t, \quad (4.32)$$

where, $\tilde{\mathbf{x}}_t \in \mathbb{R}^{5n_1}$ and $\hat{\mathbf{x}}_t \in \mathbb{R}^{5n_1}$ represents the predicted and the estimated states respectively.

The BMC-KF approach is summarized in Algorithm 1.

Algorithm 1 Bayesian MC with KF

Input: Observation matrix $\mathbf{Z}_t \in \mathbb{R}^{n_1 \times n_2}$ at $t = 1$, Set Ψ of observed indices

Output: $\hat{\mathbf{X}}_t$

Initialization: Find SVD of $\mathbf{S}_\Psi(\mathbf{Z}_t)$ and initialize matrices $\mathbf{U}$ and $\mathbf{V}$

- 1: **for** $t = 2$: T **do**
 - 2: Compute posterior distribution of $\mathbf{U}, \mathbf{V}, \gamma$ and β as given in (4.16), (4.20), (4.23) and (4.27) respectively.
 - 3: Estimate $\hat{\mathbf{X}}_{t-1}^{Bayesian}$ augmented with linearized powerflow constraints (4.3), (4.4) and the corresponding covariance $\hat{\mathbf{X}}_{t-1}^{Variance}$ using (4.13) and (4.28) respectively.
 - 4: Select m elements with least variance from $\hat{\mathbf{X}}_{t-1}^{Variance}$.
 - 5: Predict the least variance values for the next time step t using (4.32)
 - 6: Update the observation matrix $\mathbf{Z}_t$ with the new entries obtained from step 5
 - 7: Complete the unknown matrix using nuclear norm minimisation and obtain $\hat{\mathbf{X}}_t$
 - 8: return $\hat{\mathbf{X}}_t$
 - 9: **end for**
-

4.4 Computational Complexity Analysis

In matrix completion, the missing entries of the low rank matrix $\mathbf{X}_t \in \mathbb{R}^{n_1 \times n_2}$ are recovered by minimizing the nuclear norm. This involves a semi-definite optimization problem solved by the interior point method [111]. The solution to the optimization problem has a computational complexity $\mathcal{O}(n_1 n_2 \min(n_1, n_2))$. The linearized power flow constraints introduce a complexity of $\mathcal{O}(n_1^3)$. Hence, the worst case computational complexity for classic matrix completion based state estimation is $\mathcal{O}(n_1^3)$.

In locally weighted matrix completion, the states at the current time instant is obtained by appropriate weight assignment to the previously obtained states. Hence, the computational complexity of the algorithm depends on the number of previously obtained states

taken into account while calculating the current states in the network. Hence, the complexity involved in recovering a matrix $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$, of $\text{rank}(\mathbf{X}) \leq r$ using locally weighted matrix completion (LW-MC) is given by

$$\Theta(n_{\min} \log^2(n_1 + n_2) \Phi'(w))$$

where, $n_{\min} = \min(n_1, n_2)$, and

$$\Phi'(w) = \frac{\max_t w_t^2 ((d-t)r\sigma_\zeta^2/n_1 + \sigma_\epsilon^2)}{\sum_{t=1}^d w_t^2 ((d-t)\sigma_\zeta^2 + \sigma_\epsilon^2)}.$$

The BMC-KF approach involves three calculation steps: Bayesian matrix completion, KF-prediction, and finally the nuclear norm minimization. As mentioned earlier, the deterministic matrix completion has a computational complexity $\mathcal{O}(n_1 n_2 \min(n_1, n_2))$ [111]. The state estimation at each time also consists of KF prediction of m states with computational complexity $\mathcal{O}(m^2)$. Thus, the overall complexity of BMC-KF approach is $\mathcal{O}(n_1 \cdot \min(p^3 n_1^3, r^3) + n_2 \cdot \min(p^3 n_2^3, r^3)) + 2m^2 + n_1^3$ [108]. Here, p is the fraction of observed entries and r is the estimated rank at each iteration. The computational complexity of BMC-KF also depends on p and r . In a sparse Bayesian matrix completion, the effective rank of the matrix is reduced in the first few iterations leading to faster convergence.

4.5 Simulation Results and Discussions

In this section, the efficacy of LW-MC and BMC-KF based dynamic DSSE approaches is illustrated. The results are compared with static/classic-MC based DSSE. All three approaches of matrix completion are analyzed on IEEE standard test systems: IEEE 37 bus test system and IEEE 123 bus test system. Two test systems are implemented to evaluate the scalability of the methods. The IEEE 37 bus test system is a 4.8 kV standard three-phase unbalanced delta connected distribution system whereas IEEE 123 bus is an unbalanced network operating at 4.16 kV [112]. The reason for considering the IEEE 123 bus test sys-

tem for simulation is two-fold. Firstly, the IEEE 123 bus system is an unbalanced network with single-phase, two-phase and three-phase overhead and underground lines. Hence, the network characteristics resemble that of a real network. Secondly, it offers insights into the scalability of the proposed method. We have considered data streaming at 15 minute time interval, i.e., at every 15-minute interval the DSSE receives an incomplete and noisy observation matrix. Based on the incomplete observation matrix at each time step, dynamic DSSE is performed using LW-MC and BMC-KF strategies. It is important to note that the proposed method and the associated theoretical foundations are applicable to any time scale. That is, the algorithm can accurately estimate the system states even with a time resolution less than 15 minutes. However, as the current field measurement data typically has a 15-minute resolution, we have performed the simulations for a 15-minute time resolution.

In the classic matrix completion approach, we impute the missing entries of the low-rank observation matrix at a single snapshot of time. The temporal dependencies among the states are not taken into account while estimating the states. In LW-MC approach we estimate the missing entries of the observed matrix at the current time step by exponentially weighing three previously estimated states i.e., $d = 4$. The weights are obtained using equation (4.11). In the BMC-KF filter approach, we follow the method presented in Algorithm 1. The source code of the implemented algorithms can be found in the link <https://github.com/broutksu/Dynamic-Matrix-Completion>.

The CVX solver [113] is used to solve the optimization problem for all the approaches discussed above. The performance of DSSE is quantified using MAPE (Mean absolute percentage error = $\frac{1}{n_1} \sum_{b=1}^{n_1} \left| \frac{x_b - \hat{x}_b}{x_b} \right|$) for voltage magnitude and MIAE (Mean Integrated Absolute Error = $\sum_{b=1}^{n_1} \frac{|\theta_b - \hat{\theta}_b|}{n_1}$) for voltage phase angle. Here, x_b and $\hat{x}_b$ are the true values and estimated values of the voltage magnitude, and θ_b and $\hat{\theta}_b$ are the true and estimated values of the phase angle of the voltage phasor at bus $b \in \mathcal{B}$ respectively.

For demonstrating the results we have considered ten time instants. The MAPE and MIAE results for different time instants are plotted. Fig.4.3 and Fig.4.4 shows the voltage magnitude and phase angle recovery corresponding to 30% of the fraction of available data (FAD). For static MC and LW-MC, we have considered 30% observed entries relative to the

size of the observation matrix $\mathbf{Z}_t$. At each time instant t , the locations of the observed entries in $\mathbf{Z}_t$ are randomly sampled. However, BMC-KF, at each time t only 15% observations are chosen corresponding to $\mathbf{Z}_t$. The rest 15% are obtained from BMC-KF prediction done at $t - 1$ time step. We have considered 2 cases for selecting the measurement at time step t : (a) BMC-KF (Case-1) - In this approach, measurements are sampled randomly and (b) BMC-KF (Case-2) - In this approach, we have fixed the measurement locations for the entire duration of the simulation. Case-2 resembles a practical distribution network, where the sensor locations are fixed once the sensors are installed in the distribution network. The state transition parameters of (4.32) used for simulation are as follows: the value of the $\mathbf{s}_0$ is initialized to the state vector at $t = 1$ and $\mathbf{b}_0 = 0$. The tuning parameters α and β are assumed to be 0.8 and 0.5 respectively.

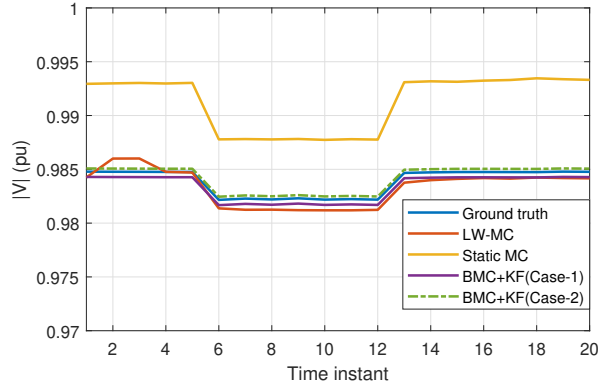


Figure 4.2: Dynamic state estimation due to a change in load at bus 6 of the IEEE 37 bus system

Fig. 4.2 shows the dynamic state estimation performance using the proposed algorithms. To illustrate the ability of our algorithms to track voltage dynamics, we consider an increase in the load demand from 6th time instant to 12th time instant at bus 6 of the IEEE 37 bus test system; as a result, the voltage at bus 6 reduces. This is indicated by ground truth results using load-flow. The proposed approach can track the changes in the voltage profile due to load change. It can be seen that the two proposed approaches - BMC-KF and LW-MC can accurately track the dynamics in the system states. It can be further seen that the estimation error is more in static-MC compared to the proposed dynamic approaches.

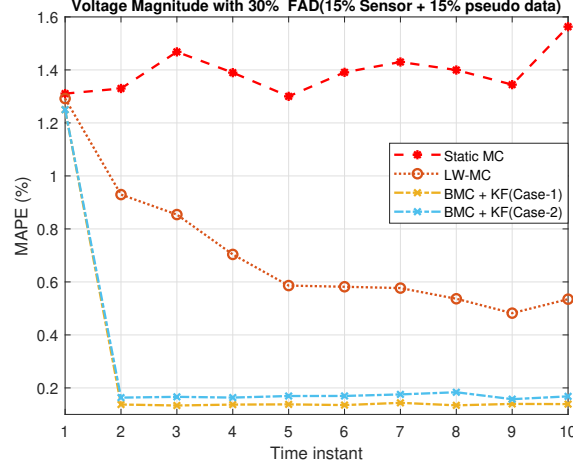


Figure 4.3: Voltage magnitude recovery with 30 % FAD

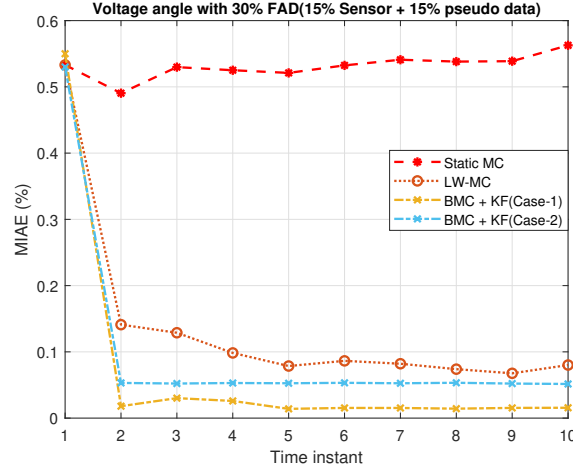


Figure 4.4: Voltage angle recovery with 30 % FAD

It can be observed that Bayesian matrix completion based DSSE using Kalman filter prediction gives the best performance in estimating voltage magnitude and angle as shown in Fig.4.3 and Fig.4.4 respectively. This is because BMC-KF prediction employs a two step matrix completion approach. Firstly, it employs the Bayesian MC approach to estimate the missing entries at $t - 1$ time-step and selects m most certain (least variance) values for KF prediction. We perform KF predictions for m values to use at time t . Secondly, it augments the observation matrix available at time t with BMC-KF estimates. Therefore, accurate estimates are obtained in the BMC-KF approach as compared to LW-MC and static MC. The LW-MC method takes three previously obtained time-step information to estimate the

states in the current time step. Therefore, the error in estimating the states is more in this approach than the BMC-KF approach and less than the classic MC approach. The static-MC does not account for the temporal correlation among the data and hence gives the worst estimate for the voltage magnitude and angle. To validate the applicability of the proposed methods for the large distribution network, we have employed the proposed methods on IEEE 123 bus system. The recovery of the voltage magnitude and phase angle at different time instants are shown in Fig.4.5 and Fig.4.6. It is found that the observations are consistent with IEEE 37 bus test network. It is observed that in the first time instant, all three approaches provide similar performance as temporal correlation among data is not exploited. However, in the subsequent time instants, the estimation error reduces in BMC-KF and LW-MC as these methods leverage the temporal correlations among the data sets.

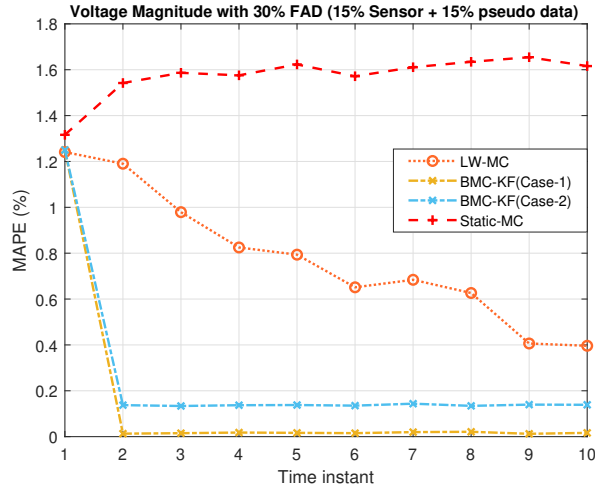


Figure 4.5: IEEE 123 bus system voltage magnitude recovery with 30% FAD

Table 4.1: Time requirement for 1 run to estimate 4 time steps

Matrix Completion Method	Time (Second)	
	IEEE 37 bus	IEEE 123 bus
Classic MC	15.8895	127.4091
LW-MC	16.4725	162.0538
BMC-KF	16.7325	128.3922

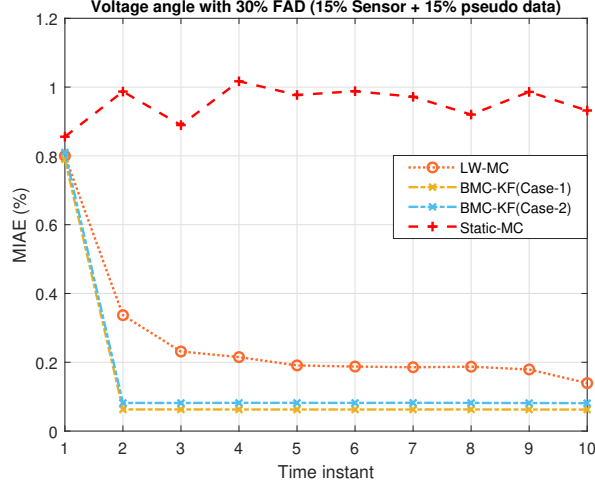


Figure 4.6: IEEE 123 bus system voltage angle recovery with 30% FAD

The time required to run each of the MC techniques (for one Monte-Carlo simulation) for four-time instants is given in Table 5.3. Four-time instants are considered because LW-MC requires four-time steps to capture the dynamic nature of the states. The simulations are performed on Intel core i7, a 16GB RAM processor. It can be observed from Table 5.3 that for smaller network sizes, the time complexity of LW-MC and BMC-KF are comparable. However, the time complexity of BMC-KF scales much more favorably relative to LW-MC for larger networks.

4.6 Summary

We present two algorithms, LW-MC and BMC-KF, for dynamic state estimation in a low-observable distribution grid. Numerical studies demonstrate that the BMC-KF-based DSSE approach outperforms LW-MC and static-MC.

The MC approaches proposed in this chapter efficiently estimate the system states in a nominal size of the power distribution network. However, the estimation process becomes computationally inefficient as network size increases. Hence, we propose a distributed DSSE approach in the next chapter.

Chapter 5

Distributed State Estimation

The MC approaches proposed in Chapter 4 efficiently estimate the system states in a nominal size of the power distribution network. However, the estimation process becomes computationally inefficient as network size increases. Hence, we propose a distributed DSSE approach in this chapter.

5.1 Introduction

The power distribution grid has limited measurements. Hence, the system is unobservable. However, there is a continuous effort to make the distribution network observable by installing more sensors. Under this scenario, it becomes challenging to aggregate all the data reliably, even under optimal conditions. Sometimes, it is difficult for the underlying communication network to keep up with the sensors' volume and velocity of information. The information aggregation challenge for centralized state estimation grows with the size of the distribution network. Hence, distributed SE is gaining attention. In this framework, local information processing and control centers manage a subset of the network nodes and exchange information across control centers to recover system states. This reduces the communication overhead and computation time.

We propose a distributed compressive sensing-based DSSE. In the proposed approach,

the distribution network is segregated into multiple sub-networks. Each sub-network has a controller that estimates its states based on available local measurements. Limited information on consensus and boundary bus state variables is shared coherently among adjacent sub-networks to obtain a globally optimal solution for the system states. The optimal global solution is obtained within several iterations.

Here, we briefly describe the basic concept of CS; more details are available in [114]. Consider the original signal $\mathbf{x} \in \mathbb{R}^N$ that is compressible in a linear transformation basis Ψ such that

$$\mathbf{x} = \Psi \mathbf{b} \quad (5.1)$$

where, $\mathbf{b}$ has at most $K \ll N$ significant coefficients i.e., $\mathbf{x}$ is K sparse in sparsifying basis Ψ . If the sensing mechanism is such that:

$$\mathbf{z} = \Phi \mathbf{x}; \quad \mathbf{z} \in \mathbb{R}^M, \quad \Phi \in \mathbb{R}^{M \times N} \quad (5.2)$$

where, Φ is a random measurement/projection matrix. For example, the elements of Φ could be distributed as *i.i.d.* Gaussian random variables with mean 0 and variance $1/M$ or Bernoulli random variables. The original state vector $\mathbf{x}$ can be reconstructed by solving the l_1 -norm minimization problem.

$$\begin{aligned} \hat{\mathbf{a}} &= \underset{\mathbf{q}}{\text{minimize}} && ||\mathbf{q}||_1 \\ &\text{subject to:} && \end{aligned} \quad (5.3)$$

$$\mathbf{z} = \Phi \Psi \mathbf{q}$$

$$\Psi \mathbf{q} = \mathbf{M} \mathbf{y} + \mathbf{w} \quad (5.4)$$

where, (5.4) captures the linearized power flow constraints. The reconstructed states correspond to

$$\hat{\mathbf{x}} = \Psi \hat{\mathbf{a}}, \quad (5.5)$$

where, $\hat{\mathbf{x}}$ can be an exact reconstruction with overwhelming probability if there exists a $\delta \in (0, 1)$ such that,

$$(1 - \delta)\|\mathbf{q}\|_2^2 \leq \|\Phi\Psi\mathbf{q}\|_2^2 \leq (1 + \delta)\|\mathbf{q}\|_2^2, \quad (5.6)$$

holds for all K -sparse signal $\mathbf{q}$ [115]. This is called the Restricted Isometry Property (RIP) of order K [116].

5.2 Problem formulation

In this paper, we consider an interconnected power distribution network consisting of n nodes in a node-set $\mathcal{N} = \{1, 2, \dots, n\}$. Node 1 is the slack bus, connecting the distribution network to the rest of the power system network. The rest of the $n - 1$ nodes are considered PQ as nodes. The network is partitioned on a geographical basis into $\mathcal{D}$ non-overlapping areas (control centers). It is assumed that each control center can collect the electrical measurements available in their respective control areas, communicate to neighboring control centers reliably, and has minimum computation capabilities. A control area is not confined to a DSO. Rather it can be an area under a regional balancing authority's footprint, or a micro-grid setup, or it can be as small as a single node of the distribution network. As an example, a three area partition of the IEEE 37-bus test system is shown in Figure 5.1. Next, we will discuss the multi-area partition of the distributed network.

5.2.1 Decentralised Linear Power Flow Model

This section introduces the linear power flow model and then the procedure to obtain the decentralized linear power flow model for decentralized state estimation. This paper uses the linearized power flow model developed as,

$$v \approx \mathbf{M}\mathbf{y} + \mathbf{w}, \quad (5.7)$$

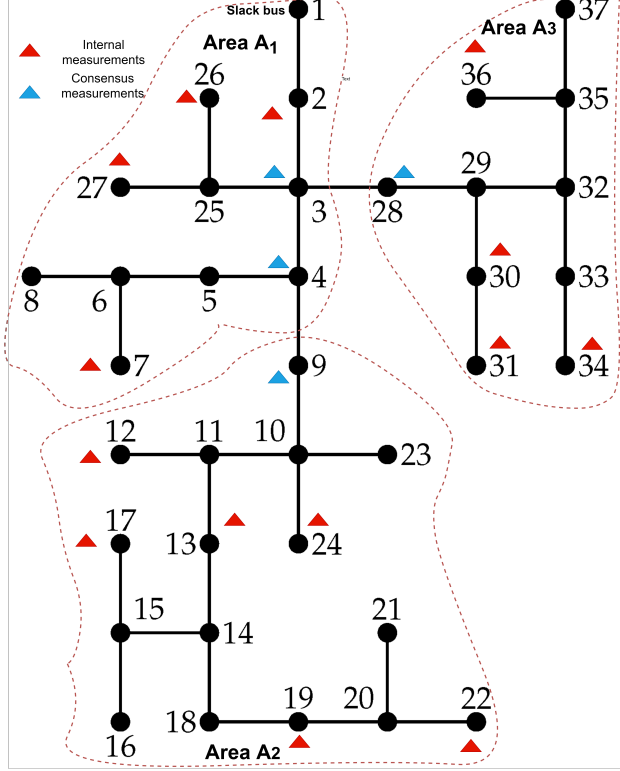


Figure 5.1: Distributed architecture of the IEEE37-bus test system

where, $\mathbf{y} = [\Re(\mathbf{s})^T, \Im(\mathbf{s})^T]^T$. Details related to the computation of the coefficients $\mathbf{M} \in \mathbb{C}^{n \times 2n}$ and $w \in \mathbb{R}^{|n|}$ can be found in [104].

To find a decentralized approximation to the linear power flow model of (5.7), the distribution network is partitioned into $\mathcal{D}$ sub networks. From the linear model (5.7), for a phase i in sub-network ($1 \leq a \leq \mathcal{D}$), we have

$$v_i = \sum_{j=1}^{2n} M_{ij} y_j + w_i$$

in which

$$\sum_{j=1}^{2n} M_{ij} y_j = \sum_{j \in \mathcal{J}^a} M_{ij} y_j + \sum_{b \in \mathcal{J}(a)} \sum_{j \in \mathcal{J}^b} M_{ij} y_j + \sum_{b \notin \mathcal{J}(a)} \sum_{j \in \mathcal{J}^b} M_{ij} y_j$$

where, M_{ij} is the $(i, j)^{th}$ element of the matrix $\mathbf{M}$ defined in (5.7). $\mathcal{J}^a$ denote the set of phases in sub-network a and $\mathcal{J}(a)$ denote sub-network neighbor to a . It is observed that the third term in the above model is the contribution in calculating the voltage at node i

because of power injections at the nodes non-adjacent to the sub-network. Intuitively, the effect is very less compared to the power injection from the neighboring areas [117]. Hence, we can have the approximated decentralized linear power flow model as,

$$v_i \approx \sum_{j \in \mathcal{J}^a} M_{ij} y_j + \sum_{b \in \mathcal{J}(a)} \sum_{j \in \mathcal{J}^b} M_{ij} y_j + w_i \quad (5.8)$$

In matrix form, (5.8) can be written as

$$\mathbf{v}_a = \mathbf{M}_a \mathbf{s}_{a,b} + \mathbf{w}_a, \quad (5.9)$$

where, $\mathbf{M}_a = \left[\sum_{j \in \mathcal{J}^a} M_{ij}, \sum_{b \in \mathcal{J}(a)} \sum_{j \in \mathcal{J}^b} M_{ij} \right]$, $\forall i \in a, j \in \{a, b\}$, and $b \in \mathcal{J}(a)$; $\mathbf{s}_{a,b} = [1, 1]^T y_j, \forall j \in \{a, b\}$.

With this approximation every sub-network b communicates the term $\sum_{j \in \mathcal{J}^b} M_{ij} y_j$ to all buses a in the neighbouring areas. This paper uses the above decentralized system matrix $\mathbf{M}_a$ in (5.10) for the multi area distributed state estimation.

We define $\mathbf{x} = [\Re(\mathbf{s})^T, \Im(\mathbf{s})^T, \Re(\mathbf{v})^T, \Im(\mathbf{v})^T, |\mathbf{v}|^T]^T$ as the concatenation of variables of real power, reactive power, real part of voltage phasor, imaginary part of voltage phasor and the magnitude of the voltage phasor, respectively. Since the distribution network is divided into $\mathcal{D}$ sub-areas. Let $\mathbf{x}_a$ denotes the variables of area a , which include the internal variables of its own $\mathbf{x}_{a,int}$ and boundary variables $\mathbf{x}_{a,adj}$. $\mathbf{x}_{a,adj}$ collects copies of boundary variables both in local and neighbouring areas. Figure 5.2 shows a schematic diagram of a 3-area system. It can be seen that, *area a* shares its boundary nodes with neighboring *areas b* and the same is true for *area c* which shares its boundary nodes with *area b*. *Area b* has two neighbouring areas, *area a* and *area c*. *Area b* shares its boundary nodes i with *area a* and f with *area c*, respectively. *Area b* holds its own internal variables $\mathbf{x}_{b,int}$ and the boundary node variables $\mathbf{x}_{b,adj}^{ba}$ and $\mathbf{x}_{b,adj}^{bc}$ which are subset of $\mathbf{x}_{a,adj}$ and $\mathbf{x}_{c,adj}$ respectively. where, $\mathbf{x}_{a,adj} := \cup_{b \in \mathcal{J}^a} \mathbf{x}_{a,adj}^{ab}$ and $\mathbf{x}_{c,adj} := \cup_{b \in \mathcal{J}^c} \mathbf{x}_{c,adj}^{cb}$. The aforementioned partition is fair and equal to all areas where distributed control centers co-exist.

Consider a global variable $g_{a,b}$ equivalent to the boundary variable value of the *area a*

and b , such that,

$$g_{a,b} = \mathbf{x}_{a,adj}^{ab} = \mathbf{x}_{b,adj}^{ab}.$$

It is worth noting that for any two adjacent areas a and b , the global variable is essentially unique, i.e., $g_{a,b} = g_{b,a}$. The original centralised model of DSSE model described in (7) can be rewritten as a distributed model of CS based DSSE as,

$$\begin{aligned} \hat{x}_a = \underset{x_a}{\operatorname{argmin}} \quad & \sum_{a=1}^D f_a(\mathbf{x}_a) \\ \text{subject to: } \quad & \mathbf{x}_a \in \mathcal{X}_a \quad a = 1 \cdots D \\ & \mathbf{x}_{a,adj}^{ab} = g_{a,b} \quad \forall (a,b) \in \mathcal{J} \end{aligned} \quad (5.10)$$

where, $f_a(\mathbf{x}_a) = \|\mathbf{x}_a\|_1 + \frac{\lambda_1}{2} \|\mathbf{y}_a - \Phi \Psi \mathbf{x}_a\|_2^2 + \frac{\lambda_2}{2} \|\Psi_a \mathbf{x}_a - (\mathbf{M}_a \mathbf{s}_a + \mathbf{w}_a)\|_2^2$. However, the solution is still not fully distributed due to the global variable $g_{a,b}$. The fully distributed solution is presented in the next sub-section.

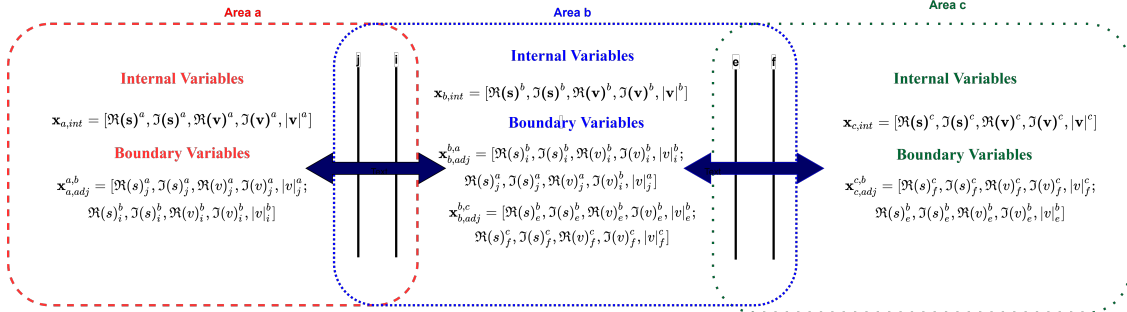


Figure 5.2: Schematic of a 3-area architecture of the proposed approach showing the internal and overlapping boundary variables

5.2.2 Fully Distributed Model

ADMM is a powerful solver for distributed optimization problems. ADMM takes the form of a *decomposition-coordination* procedure, in which solutions to small local problems are coordinated to find a solution to a large global problem[118]. In the context of the formulation in (5.10), the dual variables $\lambda_{a,b}$ correspond to consensus boundary constraints of area a and

b ; the cardinality of neighbor sets $|\mathcal{J}^a|$ is the number of neighbors of area a . Define local variables $\boldsymbol{\lambda}_a = \{\lambda_{a,b} | b \in \mathcal{J}^a\}$ and $\mathbf{g}_a = \{\mathbf{g}_{a,b} | b \in \mathcal{J}^a\}$ as the concatenation of dual and global variables, respectively; and $\boldsymbol{\lambda}$ and $\mathbf{g}$ as the concatenation of corresponding local and global variables. The augmented Lagrangian formulation for (5.10) is given as,

$$\mathcal{L}(\mathbf{x}, \mathbf{g}, \boldsymbol{\lambda}) = \sum_{a=1}^D \mathcal{L}_a(\{\mathbf{x}_a\}, \{\mathbf{g}_a\}, \{\boldsymbol{\lambda}_a\}) \quad (5.11)$$

where,

$$\mathcal{L}_a(\{\mathbf{x}_a\}, \{\mathbf{g}_a\}, \{\boldsymbol{\lambda}_a\}) = f_a(\mathbf{x}_a) + \sum_{b \in \mathcal{J}^a} \left[\boldsymbol{\lambda}_a^T (\mathbf{x}_{a,adj} - \mathbf{g}_a) + \frac{\rho}{2} \|\mathbf{x}_{a,adj} - \mathbf{g}_a\|_2^2 \right] \quad (5.12)$$

where, ρ is a predefined constant. The ADMM iterates through

$$\mathbf{x}_a^{t+1} = \underset{\mathbf{x}_a}{\operatorname{argmin}} \mathcal{L}_a(\mathbf{x}_a, \mathbf{g}_a^t, \boldsymbol{\lambda}_a^t) \quad (5.13)$$

$$\mathbf{g}_a^{t+1} = \underset{\mathbf{g}_a}{\operatorname{argmin}} \mathcal{L}_a(\mathbf{x}_a^{t+1}, \mathbf{g}_a, \boldsymbol{\lambda}_a^t) \quad (5.14)$$

$$\boldsymbol{\lambda}_a^{t+1} = \boldsymbol{\lambda}_a^t + \rho(\mathbf{x}_{a,adj}^{t+1} - \mathbf{g}_a^{t+1}) \quad \forall b \in \mathcal{J}^a \quad (5.15)$$

It can be observed that (5.13) and (5.15) can be executed with available local information. However, an update of (5.14) needs a central coordinator to collect the global data and make a global update.

Lemma 1 : The Lagrangian update using (5.15) satisfies $\boldsymbol{\lambda}_a^t + \boldsymbol{\lambda}_b^t = 0$ for every iteration t .

Proof: For any two neighboring areas a and b , (5.15) gives,

$$\boldsymbol{\lambda}_a^{t+1} = \boldsymbol{\lambda}_a^t + \rho(\mathbf{x}_{a,b}^{t+1} - \mathbf{g}_{a,b}^{t+1}), \quad (5.16)$$

$$\lambda_b^{t+1} = \lambda_b^t + \rho(\mathbf{x}_{b,a}^{t+1} - \mathbf{g}_{b,a}^{t+1}) \quad (5.17)$$

Adding (5.16) and (5.17), we get

$$\lambda_a^{t+1} + \lambda_b^{t+1} = \lambda_a^t + \lambda_b^t + \rho(\mathbf{x}_{a,b}^{t+1} + \mathbf{x}_{b,a}^{t+1}) - (\mathbf{g}_{a,b}^{t+1} + \mathbf{g}_{b,a}^{t+1}) \quad (5.18)$$

As the global variables $\mathbf{g}_{a,b}^{t+1} = \mathbf{g}_{b,a}^{t+1}$ for neighboring areas, from (5.18) we get

$$\mathbf{g}_{a,b}^{t+1} = \mathbf{g}_{b,a}^{t+1} = \frac{\lambda_a^t + \lambda_b^t}{2\rho} + \frac{\mathbf{x}_{a,b}^{t+1} + \mathbf{x}_{b,a}^{t+1}}{2} - \frac{\lambda_a^{t+1} + \lambda_b^{t+1}}{2\rho} \quad (5.19)$$

Further observation on (5.14) gives the closed form solution as,

$$\mathbf{g}_{a,b}^{t+1} = \mathbf{x}_{a,b}^{t+1} + \frac{\lambda_a^t}{\rho}, \mathbf{g}_{b,a}^{t+1} = \mathbf{x}_{b,a}^{t+1} + \frac{\lambda_b^t}{\rho} \quad (5.20)$$

comparing (5.19) and (5.20), we get $\lambda_a^{t+1} + \lambda_b^{t+1} = 0$.

Using the results of *Lemma 1* in (5.19) and replacing $\mathbf{x}_{a,b}^t = \mathbf{x}_{a,adj}^t$ and $\mathbf{x}_{b,a}^t = \mathbf{x}_{b,adj}^t$, we obtain

$$\mathbf{g}_{a,b} = \frac{\mathbf{x}_{a,adj}^t + \mathbf{x}_{b,adj}^t}{2} \quad \forall b \in \mathcal{J}^a. \quad (5.21)$$

Hence, a fully distributed ADMM results to,

$$\mathbf{x}_a^{t+1} = \underset{\mathbf{x}_a}{\operatorname{argmin}} \mathcal{L}_a \left(\{\mathbf{x}_a\}, \left\{ \frac{\mathbf{x}_{a,adj}^t + \mathbf{x}_{b,adj}^t}{2} \right\}, \{\lambda_{a,b}^t\} \right) \quad \forall b \in \mathcal{J}^a, \quad (5.22)$$

$$\lambda_a^{t+1} = \lambda_a^t + \rho \left(\frac{\mathbf{x}_{a,adj}^t - \mathbf{x}_{b,adj}^t}{2} \right) \quad \forall b \in \mathcal{J}^a. \quad (5.23)$$

The termination criteria comprises of the primal and the dual residual convergence. Set the primal and dual residual as $\epsilon_{primal}^t = \{|\mathbf{x}_{a,adj}^t - \mathbf{g}_a^t|, \forall a\}$ and $\epsilon_{dual}^t = \{|\mathbf{g}_a^t - \mathbf{g}_a^{t-1}|, \forall a\}$, respectively. Considering σ a feasible tolerance, ADMM convergences with the condition $\epsilon = \|\epsilon_{primal}^t, \epsilon_{dual}^t\|_\infty \leq \sigma$.

Proposition 1 (Convergence criteria): Given a bounded penalty parameter $\rho \leq \infty$, the

iteration described by (5.22) and (5.23) will attain both feasibility and optimality i.e,

$$\lim_{t \rightarrow \infty} \epsilon^t = 0 \quad (5.24)$$

$$\lim_{t \rightarrow \infty} \mathbf{x}_a^t = \mathbf{x}_a^* \quad (5.25)$$

and

$$\lim_{t \rightarrow \infty} \sum_a^D f_a(\mathbf{x}_a^t) = r^*. \quad (5.26)$$

where, $\mathbf{x}_a^*$ is the optimal solution and r^* optimal value of (5.10).

Proof: See Appendix for proof.

The fully distributed algorithm described in (5.22) and (5.23) can be solved in a parallel manner. Each sub-network solves their own state estimation based on their available data using (5.22), exchange the boundary data with the neighbor and then dual variable is updated via the current boundary information as in (5.23). A pseudo-code is presented in Algorithm 2:

Algorithm 2 Distributed CS based State estimation using ADMM

Data: Network parameter: $\mathbf{M}$, Measurement vector: $\mathbf{y}$

Result: $\hat{\mathbf{x}}$

initialization $\lambda^0, \rho > 0, \epsilon^{feasible} > 0, t = 0$

while $\epsilon > \epsilon^{feasible}$ **do**

 Obtain $\mathbf{x}_a^{t+1}$ using (17)

 Obtain $\mathbf{g}_a^{t+1}$ using (18)

 Obtain λ_a using (19)

 Calculate $\epsilon = \|\mathbf{g}_a^{t+1} - \mathbf{g}_a^t\|$

 Update $t = t+1$ **if** $\epsilon < \epsilon^{feasible}$ **then**

 break

end

end

5.2.3 Computational complexity

Compressive sensing based state estimation method involves an l_1 minimization which is solved by linear programming. Let M be the number of variables and N the number of constraints. The computational complexity of solving one Newton step is $O(MN.min(M, N))$. The same argument is valid for the decentralized CS. However, the number of variables in a particular area reduces, which reduces the execution time. The execution time in the decentralized state estimation method is always less than the centralized state estimation method. The total time for the decentralized state estimation method depends on the time taken for the estimation process on the area having the highest number of variables and the intercommunication among the neighboring areas. Table 1 shows the simulation time for three area partitions of the standard IEEE37 and IEEE123 bus system.

Table 5.1: Simulation time for 1 run and 1 CMR

Test System	Simulation Time (Sec)	
	Centralized SE	Decentralized SE
IEEE 37 bus	3.85	1.03
IEEE 123 bus	20.48	7.28

5.3 Simulation Results

The efficacy of the proposed distributed CS-based state estimation is verified on the three-phase unbalanced IEEE 37 bus test system [119] and the scalability is verified via the IEEE 123 bus test system [120]. The IEEE 37-bus test system is a highly unbalanced delta-connected system with an operating voltage of 4.8 kV and spot loads. The IEEE 123-bus test system is unbalanced, with several three-phase and single phase laterals operating at 4.16 kV and spot loads. The optimisation problem (5.11) is solved using CVX solver [113]. The results obtained using distributed CS-based state estimation are compared with the states of the whole distribution network using centralized CS state estimation with varying compressed measurement ratio (CMR). The performance measures used in this paper to quantify the estimation accuracy are mean absolute percentage error (MAPE) for voltage

magnitude and mean integrated absolute error (MIAE) for voltage angle, respectively and are defined as follows;

$$MAPE = \frac{1}{N} \sum_{b=1}^N \left| \frac{x_b - \hat{x}_b}{x_b} \right|, \quad (5.27)$$

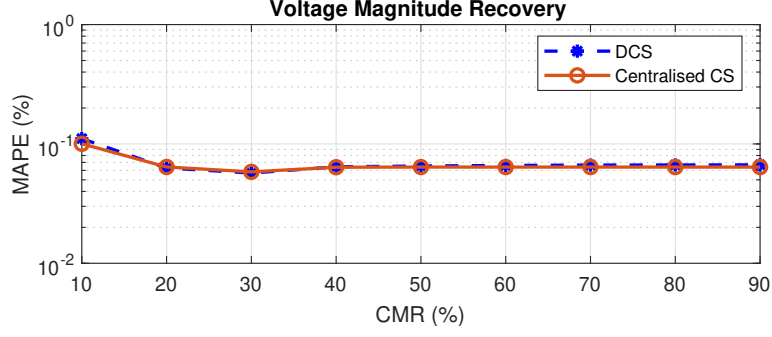
$$MIAE = \sum_{b=1}^N \frac{|\theta_b - \hat{\theta}_b|}{N}, \quad (5.28)$$

where, x_b and θ_b are the true values of the magnitude and phase angle at bus b , respectively; similarly, the estimated values of the magnitude and phase angle at bus b are represented by $\hat{x}_b$ and $\hat{\theta}_b$ respectively. In the results section, we have presented the MAPE and MIAE results for voltage magnitude and angle estimations, respectively.

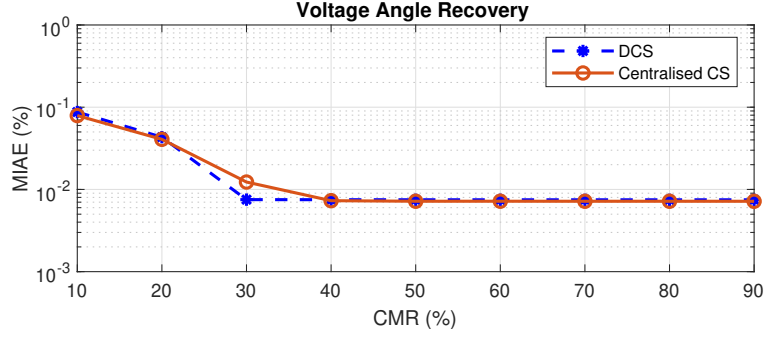
We first evaluate the performance of the distributed CS approach DSSE in the IEEE 37-bus test system. The IEEE 37-bus test system is divided into three sub-areas as shown in Figure 5.1. Each area has a certain number of measuring devices local to its control center (Marked as red triangles) and some boundary measurement nodes that share the information with the neighboring control center (Marked as green). Figure 5.3 shows the voltage magnitude and angle recovery in IEEE 37 bus test system. The MAPE and MIAE for different CMR are plotted. It is observed that the plots of both MAPE (%) in voltage magnitude and MIAE (%) in voltage angle estimation using DCS-based state estimation and the centralized state estimation overlap with each other. This implies that the distribution grid state estimation using DCS approach is equivalent to the centralized CS state estimation approach.

Next, we discuss the effectiveness of the proposed ADMM based distributed CS state estimation on the IEEE 123 bus test system. The sub-area partition of the network is given in Table 9.1. Figure 5.6 shows the voltage magnitude and angle recovery. Similar observations like the IEEE 37-bus system can be observed.

In practical applications, measurement data is impacted by noise and it is important to evaluate the estimator performance with varying level of measurement noise. The proposed distributed CS based estimator performance is evaluated by considering measurement data with varying noise. We have considered Gaussian measurement noise $\mathcal{N}(0, \sigma^2)$, where, σ^2



(a) Voltage magnitude recovery with varying CMR



(b) Voltage angle recovery with varying CMR

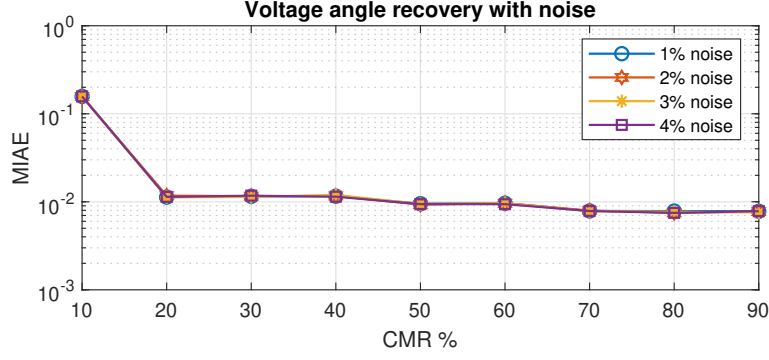
Figure 5.3: Voltage recovery in IEEE 37 bus test system

varies from 1% to 4% of measurement data value. Figure 5.4(a) and Figure 5.4(b) show the estimator performance for different noise levels in the measurement data. The estimator performance is evaluated on the IEEE 37 bus test system. It is observed that the estimator performance remains the same for all the noise levels, which shows the robustness of the proposed algorithm to measurement noise.

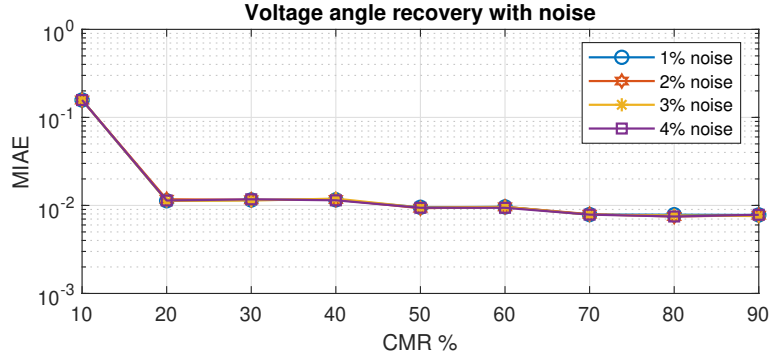
The proposed DSSE method is also sensitive to the choice of of penalty parameter (ρ). In the proposed approach, we employ both adaptive and fixed ρ values for better convergence as shown in (5.29).

$$\rho^{t+1} = \begin{cases} \tau^{incr} \rho^t & \text{if } \epsilon_{primal}^t > \mu \epsilon_{dual}^t \\ \rho^t / \tau^{decr} & \text{if } \epsilon_{dual}^t > \mu \epsilon_{primal}^t \\ \rho^t & \text{otherwise,} \end{cases} \quad (5.29)$$

where, $\mu > 1$, $\tau^{incr} > 1$, and $\tau^{decr} > 1$ are parameters. Initially ρ is updated based on (5.29) and after a fixed number of iterations it holds the final value till the algorithm converges.



(a) Voltage magnitude recovery with noisy data



(b) Voltage angle recovery with noisy data

Figure 5.4: Voltage recovery with noisy data

The idea of varying penalty parameter is to limit the primal and dual residual norms within μ . More details about the parameter update can be found in [118].

The convergence characteristics of the proposed approach is shown in Figure 5.5. For the distributed method, the number of communication exchanges among neighboring variables is decided by the number of iterations. Thus, smaller the number of iterations, lower is the computational complexity. As discussed earlier, the proposed approach is tested on a 3-area IEEE 37 bus system and 3-area IEEE 123 bus system. The partition lists are given in Table 1. Figure 5.5 shows that the CS-based distributed state estimation approach converges faster in the IEEE 37 bus network than the IEEE 123 network as the number of neighboring states in the IEEE 37 bus is less than the IEEE 123 bus system. However, both the networks reaches an acceptable tolerance of less than 10^{-2} in less than 30 iterations.

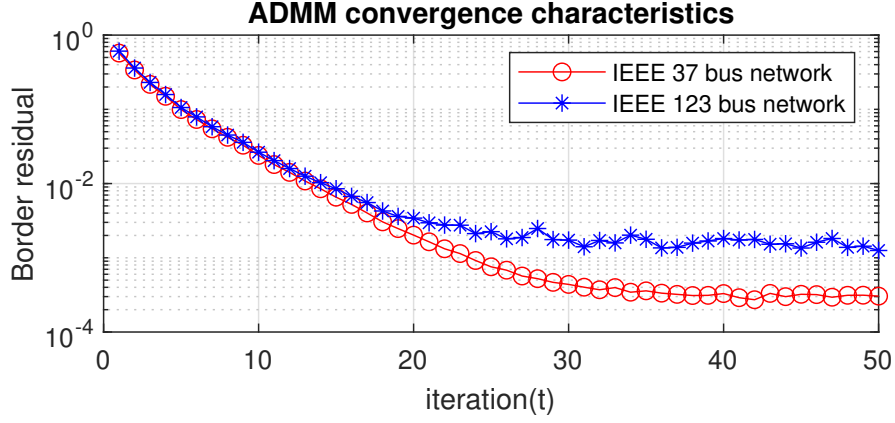
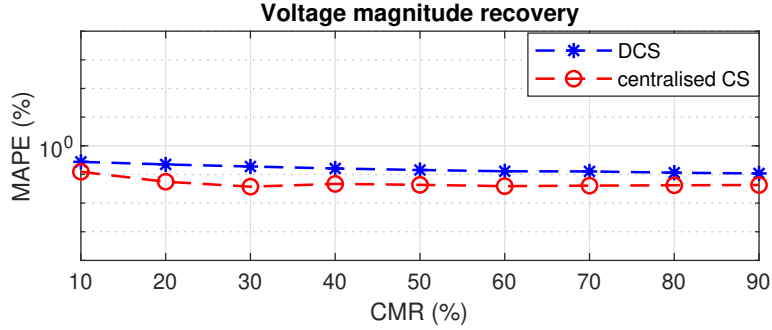


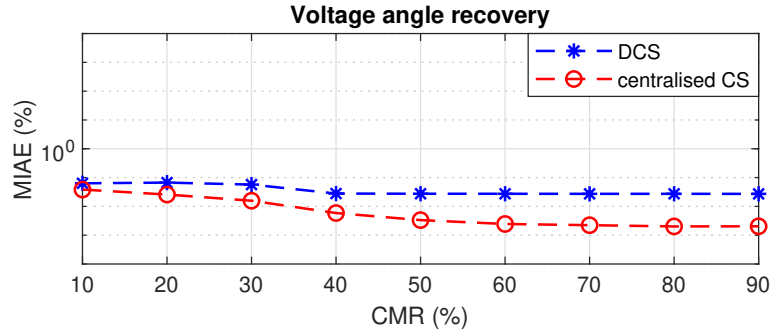
Figure 5.5: Convergence characteristics of the proposed ADMM approach

Table 5.2: Area partition of IEEE networks

Network	Bus sets		
	Area A1	Area A2	Area A3
IEEE 37	$A1 = \{1-8, 25-27\}$	$A2 = \{9-24\}$	$A3 = \{28-37\}$
IEEE 123	$A1 = \{1-33\}$	$A2 = \{34-66\}$	$A3 = \{67-123\}$



(a) Voltage magnitude recovery with varying CMR



(b) Voltage angle recovery with varying CMR

Figure 5.6: Voltage recovery in IEEE 123 bus test system

5.4 Distributed State Estimation in Large Power Distribution Network

A direct application of the distributed DSSE proposed in 5.2.2 is difficult. This is because, manual partition of a large distribution grid is extremely time and resource-consuming. Hence, it is very important to partition the large network into multiple sub-networks with minimum shared edges among neighboring sub-networks. Addressing this concern, [121] proposed algorithms for identifying network communities by iteratively removing edges. This approach measures the most "between" other vertices, i.e., the edge is responsible for connecting many other pair of other vertices. Additionally, an efficient heuristic method for graph partitioning is presented in [122], which aims to minimize a cost function dependent on shared edges among partitions. Nonetheless, these approaches suffer from high computational complexity, constraining their scalability. Hence, we propose a hierarchical network partition approach to optimally partition the distribution network before applying distributed state estimation.

5.5 Hierarchical Network Partitioning

We model the power distribution network as a connected and undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and then apply hierarchical spectral clustering to the graph to divide it into multiple sub-graphs.

Consider a power distribution network that consists of n nodes and e edges. The network can be represented as a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = 1, 2, \dots, n$ is the set of nodes, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. The graph adjacency matrix, $\mathbf{A}$, is used to represent the connection among the nodes and is defined as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad (5.30)$$

where $a_{i,j}$ represents the connection between nodes i and j . The entries of $\mathbf{A}$ are defined as:

$$a_{i,j} = \begin{cases} 1, & \text{if there is a connection from node } i \text{ to node } j \\ 0, & \text{otherwise} \end{cases} \quad (5.31)$$

For efficient optimal network partitioning, we need a similarity matrix of the network graph. For our study, we used the graph Laplacian matrix, $\mathbf{L}$, as the similarity matrix. The entries of matrix $\mathbf{L}$ are defined as:

$$\mathbf{L}_{i,j} = \begin{cases} d_i, & \text{if } i = j \\ -1, & \text{if } i \text{ and } j \text{ are adjacent} \\ 0, & \text{Otherwise} \end{cases} \quad (5.32)$$

where d_i is the degree of node i and is calculated as $d_i = \sum_{j=1}^n A_{i,j}$. Because the Laplacian matrix of a planner graph is positive semidefinite, the vector of all ones is always an eigenvector corresponding to the 0 eigenvalues, and all other eigenvalues are nonnegative. We focus on the second smallest eigenvalue, λ_2 , or the *algebraic connectivity of a graph*, and the corresponding eigenvector, v , called the *Fiedler vector*. For any vector, $\mathbf{u}$, the Fiedler value, λ_2 , of a graph is given by:

$$\lambda_2 = \min_{\mathbf{u} \perp \mathbf{1}} \frac{\mathbf{u}^T \mathbf{L} \mathbf{u}}{\mathbf{u}^T \mathbf{u}}, \quad (5.33)$$

The minimum value of (6.7) is obtained when $\mathbf{u} = v$, i.e, the unknown vector is the Fiedler vector. In this work, we exploit a key theorem validated in [123].

Theorem 1. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph on n nodes of maximum degree, d_{max} . Let L be the Laplacian matrix and ϕ be the isoperimetric number. For any vector $\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{u}^T \mathbf{1} = 0$:

$$\frac{\mathbf{u}^T \mathbf{L} \mathbf{u}}{\mathbf{u}^T \mathbf{u}} \geq \frac{\phi^2}{2d_{max}}. \quad (5.34)$$

Moreover, there exists a splitting value, s , for which $(\{i : u_i \leq s\}, \{i : u_i > s\})$ has a ratio at most $\frac{\phi^2}{2d_{max}}$. More details on this theorem can be found in [123].

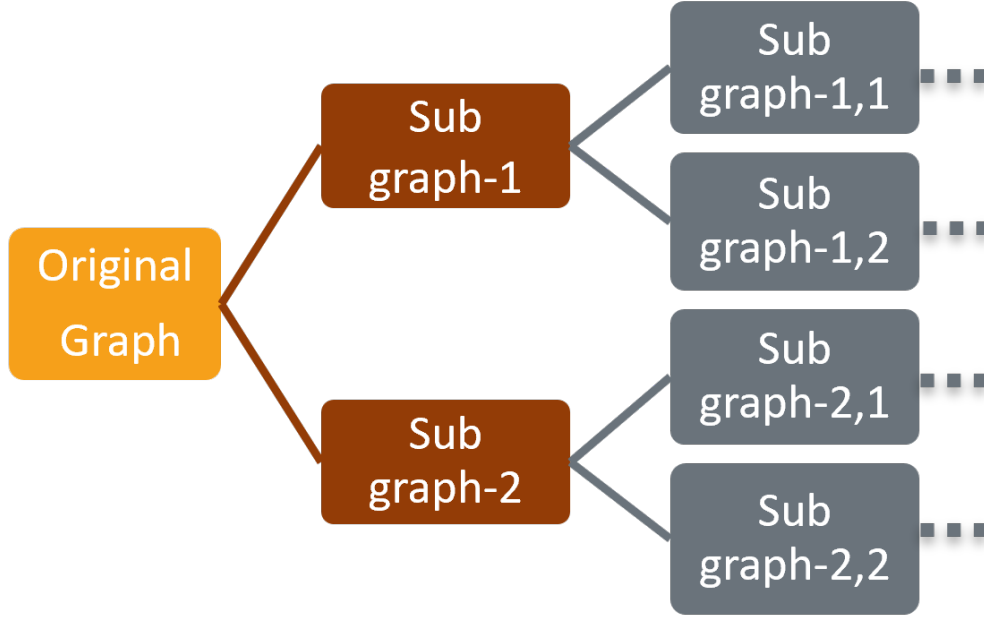


Figure 5.7: Flow diagram of hierarchical network partitioning

Using Theorem 1, with v the Fiedler vector of the graph Laplacian, the spectral partitioning is to find a splitting value, s , that partitions the vertices of $\mathcal{G}$ into sets such that $v_i \leq s$ and $v_i > s$. Although there are several values to choose for s , we choose the *sign cut* for which $s = 0$. We successively apply the discussed method to obtain the desired number of subgraphs for an equivalent large graph. We limit the subgraph creation based on the computational capability of the particular region. Fig. 5.7 shows a pictorial representation of the hierarchical spectral graph partitioning, and the step-by-step procedure is given in Algorithm 3. Next, we discuss the DSE proposed in [124] using the optimally partitioning algorithm discussed.

Algorithm 3 Hierarchical spectral clustering-based network partitioning

Input Admittance matrix: $\mathbf{Y}$, Minimum number of nodes in a single partition k

Initialize: $K = 0$

while $k > K$ **do**

 Obtain: Adjacency matrix $\mathbf{A}$

 Obtain: Degree matrix of graph $\mathbf{D}$

 Obtain: graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$

 Obtain: The Fiedler vector $\mathbf{v}$ of $\mathbf{L}$

 Obtain: $Subgraph_1 \leftarrow \mathbf{v} > 0$ and $Subgraph_2 \leftarrow \mathbf{v} \leq 0$

 Update $K = \max(\text{Volume}(Subgraph_1), \text{Volume}(Subgraph_2))$

end

Return: Number of partitions with their node index

5.5.1 Distributed Compressive Sensing

To reduce the computational complexity of compressive sensing-based state estimation in large distribution networks, we propose a DCS-based state estimation approach in power distribution grids. First, we use the hierarchical network partitioning algorithm, as discussed in Section 5.5, to divide the power distribution network into $\mathcal{D}$ subareas. Next, we use DCS to recover the states in each subarea.

Let us define the states of the power distribution network for a particular area, a , as $\mathbf{x}_a = [\Re(\mathbf{s}_a)^T, \Im(\mathbf{s}_a)^T, \Re(\mathbf{v}_a)^T, \Im(\mathbf{v}_a)^T, |\mathbf{v}_a|^T]^T$, i.e., the concatenation of active power, reactive power, real part of the voltage, imaginary part of the voltage, and voltage magnitude. The state variables, $\mathbf{x}_a$, include internal variables, $\mathbf{x}_{a,int}$, and boundary variables, $\mathbf{x}_{a,adj}$. Boundary variables are the variables that overlap with other subareas. Although the optimization problem (6.6) is used to recover the system states in each subarea, we must ensure that the boundary variables in each area are equal. To make the boundary variables equal, we introduce a global variable, $g_{a,b}$, for two neighboring areas, a and b , as:

$$g_{a,b} = \mathbf{x}_{a,adj}^{ab} = \mathbf{x}_{b,adj}^{ba}. \quad (5.35)$$

Note that the global variable for any two adjacent areas, a and b , should be equal, i.e., $g_{a,b} = g_{b,a}$. The decentralized model of the compressive sensing-based state estimation

problem corresponds to:

$$\begin{aligned}
& \underset{\mathbf{x}_a}{\operatorname{argmin}} && \sum_{a=1}^D f_a(\mathbf{x}_a) \\
& \text{subject to: } && \mathbf{x}_a \in \mathcal{X}_a \quad a = 1 \cdots D \\
& && \mathbf{x}_{a,adj}^{ab} = g_{a,b} \quad \forall (a,b) \in \mathcal{J}
\end{aligned} \tag{5.36}$$

where, $f_a(\mathbf{x}_a) = \|\Psi_a \mathbf{x}_a\|_1 + \frac{\lambda_a}{2} \|\mathbf{y}_a - \Phi_a \Psi_a \mathbf{x}_a\|_2^2$, and $\mathbf{y}_a$ is the available measurement in area a . The most popular way to find the solutions to (5.36) is to employ the alternating direction method of multipliers (ADMM). ADMM takes the form of a *decomposition-coordination* procedure, in which solutions to small local problems are coordinated to find a solution to a large global problem [118]. We apply the

5.5.2 Computational Complexity

The compressive sensing-based state estimation method involves an l_1 minimization that is solved by linear programming. Let M be the number of variables and N the number of constraints. The computational complexity of solving one Newton step is $O(MN \cdot \min(M, N))$ [19]. The same argument is valid for DCS; however, the number of variables in a particular area reduces, reducing execution time. The execution time in the DSE method is always less than the centralized state estimation method. The total time for the DSE method depends on the time taken for the estimation process in the area having the highest number of variables and the intercommunication among the neighboring areas. The proposed hierarchical spectral clustering approach minimizes the overlap among partitions, reducing the communication burden across subareas/partitions.

5.6 Simulation Results

In this section, we present the simulation results of the efficient hierarchical graph partitioning algorithms followed by the DSE. We test the proposed hierarchical spectral clustering-

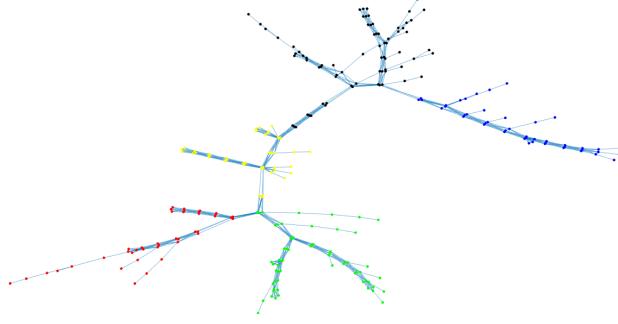


Figure 5.8: Five-area optimal partitioning of the IEEE 123-bus network

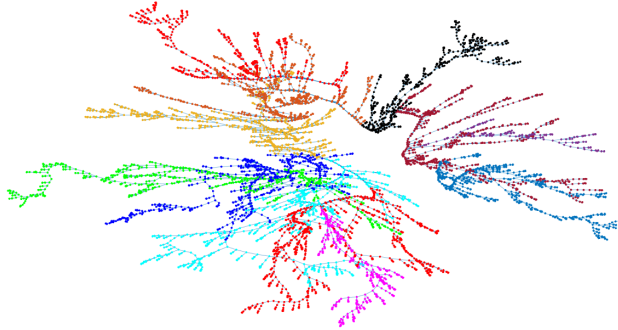


Figure 5.9: Twelve-area optimal partitioning of the IEEE 8,500-node system

based graph partitioning algorithm on the standard IEEE 123- bus system and the IEEE 8,500-node system. We also test the algorithm on a real power distribution network having 6,000+ single-phase nodes. The distribution network data is taken from a distribution network operator in Kansas, USA. The 5-area partitioning of the IEEE 123-bus system and 12-area partitioning of the IEEE 8,500-node system are shown in Fig. 5.8 and 5.9, respectively. As shown, the hierarchical spectral clustering divides the network such that the neighboring area shares minimum edges. Similar performance is also observed for the 6,000+ node distribution network. An 18-area partition of the network is shown in Fig. 5.10. A comparative analysis of the proposed hierarchical spectral network partitioning with the k-means partitioning is given in 5.3. It can be seen, that the total number of shared links among the neighbors using the proposed algorithm is much less than k-means network clustering.

Next, we evaluate the performance of the DCS-based state estimation on the 18-area partitioned distribution network. We evaluate the performance of the DSE methods using mean absolute percentage error ($MAPE$) for the voltage magnitude and the mean integrated

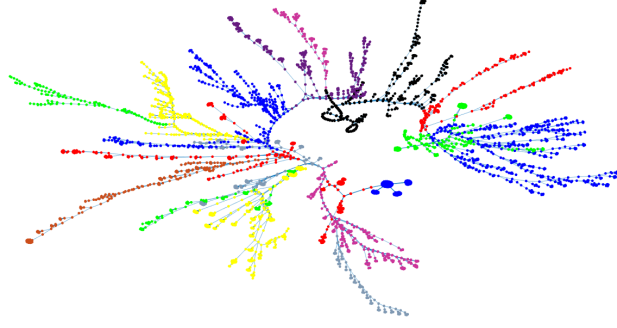


Figure 5.10: Eighteen-area optimal partitioning of the real 6,000+ node system

Table 5.3: Performance of proposed partitioning algorithm

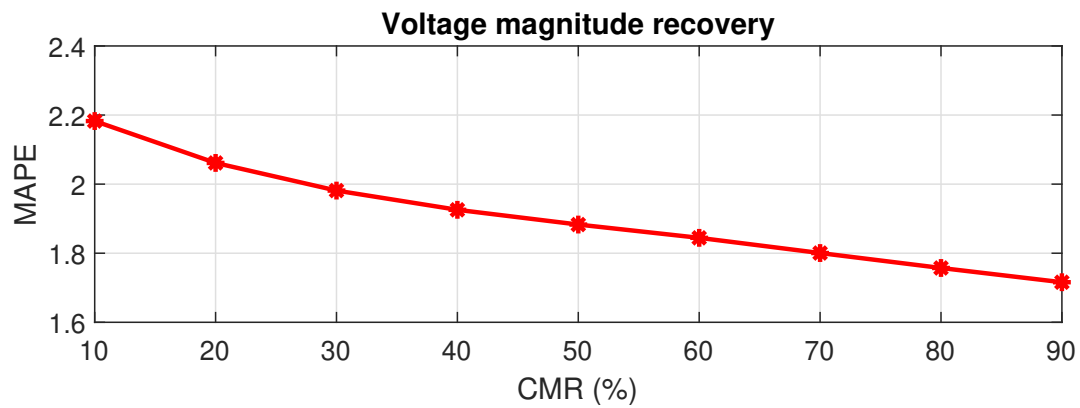
Test system	No of partitions	No of shared links	
		k-means partition	Hierarchical network partition
IEEE 8500	12	123	86
6000+	18	500+	140

absolute error ($MIAE$) for the voltage angle, respectively [30]. In the results, we present the MAPE and MIAE for the voltage magnitude and angle estimations, respectively.

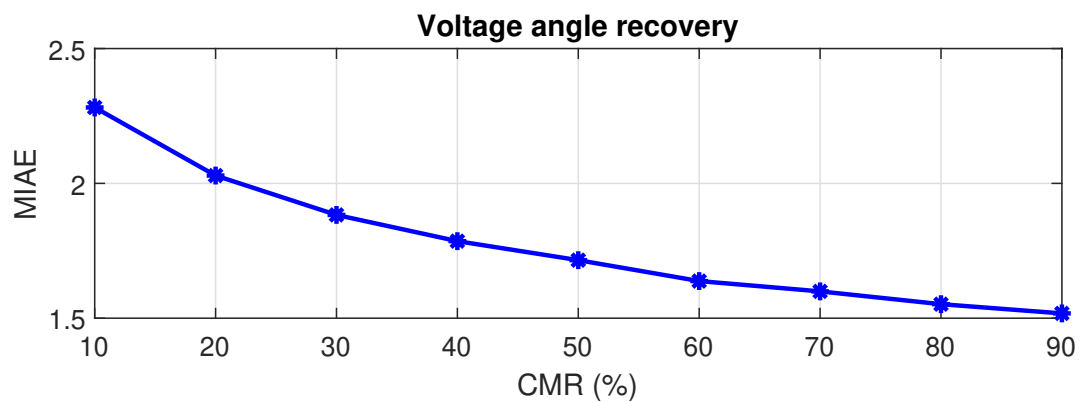
Fig. 5.11a and 5.11b show the voltage magnitude and angle recovery of the 6,000+ node distribution network. It is observed that the estimation error both in voltage magnitude and angle is maximum when the compressed measurement ratio (CMR) is 10%. The MAPE and MIAE in the voltage magnitude and the angle decrease as the CMR (%) increases. This is because as the CMR value increases, the available measurement increases, which increases the estimation accuracy and reduces the estimation error. TABLE 5.3 shows a comparative analysis of simulation time for centralized and decentralized CS-based state estimation. The simulation is performed for one run at 50% CMR.

5.7 Summary

This chapter proposes an efficient hierarchical spectral clustering-based network partitioning algorithm. We tested the algorithm performance on standard IEEE systems and a real 6,000+ distribution network. Next, we performed decentralized compressive sensing-based



(a) Voltage magnitude recovery of 6,000+ node system



(b) Voltage angle recovery of 6,000+ node system

Figure 5.11: Voltage recovery of 6,000+ node system

state estimation on the IEEE 37 bus, IEEE 123 bus, and 6,000+ real distribution network by optimally dividing the network into 18 smaller sub-networks and reporting the simulation results.

Although the distributed DSSE proposed in this chapter is scalable to large distribution networks, the simulation time depends on the larger part of the partition and the number of shared variables among the neighboring areas. Hence, state estimation for larger networks in real time is infeasible. Therefore, we propose a deep learning-based state estimation in the next chapter.

Chapter 6

Deep unfolding for Distribution Grid State Estimation

Although the distributed DSSE proposed in Chapter 5 is scalable to large distribution networks, the simulation time depends on the larger part of the partition and the number of shared variables among the neighboring areas. Hence, state estimation for larger networks in real time is infeasible. Hence, we propose a deep learning-based state estimation in this chapter.

6.1 Introduction

In recent years, the integration of distributed energy resources (DERs), electric vehicles, and demand-side management has brought significant changes to the power distribution networks. These changes have resulted in frequent voltage limit violations, reverse power flow, and mal-operation of protective devices [125]. As a result, distribution system operators are yearning for better situational awareness to better monitor and control the network. Distribution system state estimation (DSSE) is important for gaining better situational awareness. State estimation methods in power transmission networks assume full observability due to the availability of an adequate number of measurements. We pursue model-based

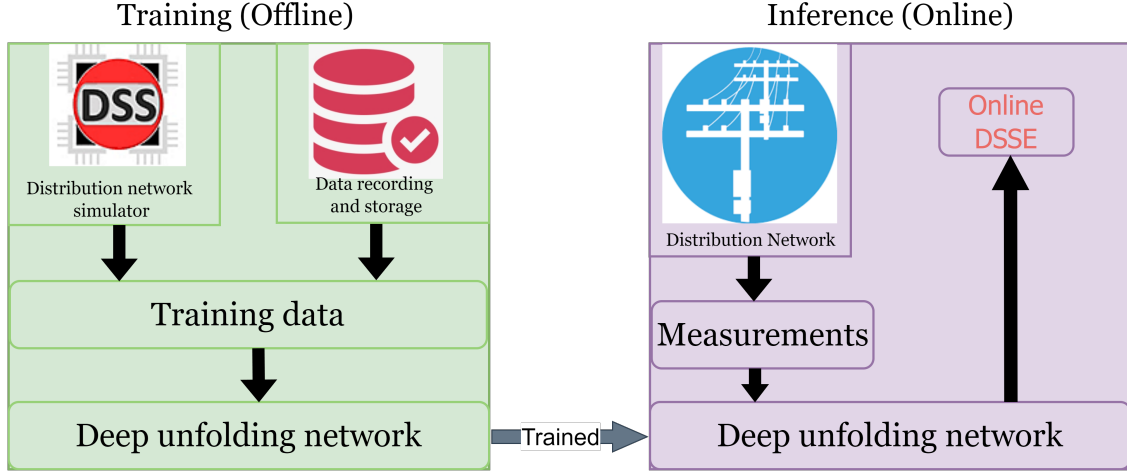


Figure 6.1: Proposed architecture

deep neural networks for DSSE by unrolling an existing iterative optimization-based DSSE. Specifically, we consider a CS-based DSSE solver, which uses l_1 minimization. In CS, l_1 norm minimization is a technique used to recover sparse signals from undersampled data. It involves minimizing the sum of the absolute values of the signal's coefficients. Mathematically, $\|\mathbf{x}\|_1 = \sum |x_i|$, where $\mathbf{x}$ is the signal being reconstructed, and $\|\mathbf{x}\|_1$ is its l_1 norm. We first derive the least absolute shrinkage activation operator (LASSO) equivalent of the l_1 norm minimization problem and solve the LASSO problem using the ADMM algorithm. ADMM involves several iterations to find the optimal solution to the l_1 norm minimization problem. The idea behind deep unfolding is to formulate a deep neural network based on the iterations of the ADMM algorithm, wherein the iterations are considered layers of the neural network. Figure 6.1 shows a block diagram of the proposed approach for real-time DSSE in power distribution grids.

6.2 Background and preliminaries

The nodal voltage and branch current phasors are the state variables in a power distribution network. State variables are responsible for the complete characterization of the power distribution network. DSSE refers to estimating the correct voltage phasors at each node of

a power distribution network based on available measurements. The state variables share a non-linear relationship with the measurement $\mathbf{y}$ as,

$$\mathbf{y} = \mathcal{G}(\mathbf{x}) + \boldsymbol{\eta}, \quad (6.1)$$

where, $\mathbf{y} = [|V|_n^T, \theta_n^T, P_n^T, Q_n^T]^T \in \mathbb{R}^M$ is the measurement vector containing voltage magnitude ($|V|_n$), phase angle (θ_n), real power (P_n), and reactive power (Q_n), respectively, $\mathbf{x} \in \mathbb{R}^N$ is the state vector, and $\mathcal{G} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ represents the nonlinear relationship between the state and measurement vector; $\boldsymbol{\eta} \in \mathbb{R}^M$ represents the measurement noise. The goal of DSSE is to estimate $\mathbf{x}$ with the knowledge of $\mathbf{y}$ and the non-linear function $\mathcal{G}(\mathbf{x})$. As discussed in [87], this can be achieved by linearizing $\mathcal{G}(\mathbf{x})$ around an operating point as,

$$\mathbf{x} = \mathbf{M}\mathbf{y}_s + \mathbf{w} \quad (6.2)$$

where, $\mathbf{y}_s = [\Re(\mathbf{s})^T, \Im(\mathbf{s})^T]^T$ is the measurement vector of real and reactive power injections;

$$\mathbf{M} = \left(\mathbf{Y}_{LL}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1}, -j\mathbf{Y}_{LL}^{-1} \text{diag}(\bar{\mathbf{w}})^{-1} \right) \text{ and,}$$

$$\mathbf{w} = -\mathbf{Y}_{LL}^{-1} \mathbf{Y}_{L1} \mathbf{v}_1 \text{ is the zero-load voltage,}$$

where, $\mathbf{v}_1$ denotes the complex vectors containing the three-phase nodal voltage at the slack bus. $\mathbf{Y}_{LL} \in \mathbb{C}^{N \times N}$ and $\mathbf{Y}_{L1} \in \mathbb{C}^{(N-3) \times 3}$ are the block matrices of the network admittance matrix $\mathbf{Y}$ defined as,

$$\mathbf{Y} := \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{1L} \\ \mathbf{Y}_{L1} & \mathbf{Y}_{LL} \end{bmatrix} \in \mathbb{C}^{N \times N}$$

where, $\mathbf{Y}_{11} \in \mathbb{C}^{3 \times 3}$ and $\mathbf{Y}_{1L} \in \mathbb{C}^{3 \times (N-3)}$ are the other two block matrices of the network admittance matrix $\mathbf{Y}$.

The exact solution of (6.1) can be obtained by using the Newton-Raphson (NR) method. However, in a power distribution network, (6.1) is typically an underdetermined system of equations which makes the NR method unsuitable. WLS and its variants have found appli-

cation in solving (1) [6]. However, WLS requires an invertible Jacobian matrix calculation. This inversion is computationally expensive, and sometimes it is impossible to calculate the Jacobian matrix because of fewer available real-time measurements. The low observability in DSSE is addressed by exploiting the correlation among the measurements and system states [58]. The states exhibit spatial and temporal correlation because of geographical proximity and correlation in power injections or consumption patterns as discussed in [58]. The spatial and/or temporal correlation results in a sparsity of states in a linear transformation basis. This sparsity trait is exploited by compressive sensing-based techniques, enabling DSSE even under low observable conditions. Next, we will briefly describe the basics of compressive sensing.

Compressive sensing is a popular technique that enables signal reconstruction from limited measurements (i.e., undetermined system of equations), provided the signal admits a sparse representation in a linear transformation basis [126]. Let's consider $\mathbf{x} \in \mathbb{R}^N$ be the original signal of interest compressible in a linear transformation basis $\Psi \in \mathbb{R}^{N \times N}$. If $\mathbf{s} \in \mathbb{R}^N$ is the sparse representation of $\mathbf{x}$ on the sparsifying basis matrix Ψ , then $\mathbf{x}$ is given as,

$$\mathbf{x} = \Psi \mathbf{s}, \quad (6.3)$$

where, $\mathbf{s}$ has at most $K \ll N$ significant coefficients. Let $\mathbf{y} \in \mathbb{R}^M$ be the measurement vector and $\Phi \in \mathbb{R}^{M \times N}$ be the projection/measurement matrix. That is,

$$\mathbf{y} = \Phi \mathbf{x}, \quad (6.4)$$

For example, the entries of Φ could be i.i.d random Gaussian variables with zero mean and $1/M$ variance. Replacing (6.3) in (6.4), the sensing model can be represented as,

$$\mathbf{y} = \Phi \Psi \mathbf{s} = \mathbf{A} \mathbf{s}, \quad (6.5)$$

where, $\mathbf{A} = \Phi \Psi \in \mathbb{R}^{M \times N}$ is the sensing matrix. The most common approach to recover $\mathbf{x}$

based on observing $\mathbf{y}$ is to solve the l_1 -norm minimization problem as:

$$\begin{aligned} & \underset{\mathbf{s} \in \mathbb{R}^N}{\text{minimize}} && \|\mathbf{s}\|_1 \\ & \text{subject to} && \mathbf{y} = \mathbf{A}\mathbf{s} \end{aligned} \quad (6.6)$$

Where, $\|\mathbf{s}\|_1$ is the l_1 -norm of $\mathbf{s}$ and is defined as $\|\mathbf{s}\|_1 = \sum |s_i|$. The optimization problem in (6.6) seeks to recover the sparse signal $\mathbf{s}$ from the under-sampled data $\mathbf{y}$. It reconstructs $\hat{\mathbf{x}} = \mathbf{\Psi}\mathbf{s}$ with overwhelming probability if $\mathbf{s}$ satisfies the restricted isometry property for any $\delta \in (0, 1)$ [127] i.e.,

$$(1 - \delta)\|\mathbf{s}\|_2^2 \leq \|\mathbf{\Phi}\mathbf{\Psi}\mathbf{s}\|_2^2 \leq (1 + \delta)\|\mathbf{s}\|_2^2 \quad (6.7)$$

In the next section, we explain the proposed novel approach for solving (6.6) for distribution network state estimation.

6.3 Proposed Approach

We model the power distribution network as a connected and undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$; where, $\mathcal{V} = 1, 2, \dots, N$ is the set of nodes, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. Let $|V_n|e^{j\theta_n}$ and $S_n = P_n + jQ_n$ represents the complex voltage and complex power injections at bus $n \in \mathcal{V}$, respectively. The measurement vector $\mathbf{y} = [|V_n^T, \theta_n^T, P_n^T, Q_n^T]^T$ includes all the aggregated SCADA sensor and/or μ PMU data from the network. Using the linearized power flow equations as defined in (6.2), the CS-based DSSE in (6.6) boils down to solving the following l_1 -norm optimization problem:

$$\begin{aligned} & \underset{\mathbf{s} \in \mathbb{R}^N}{\text{minimize}} && \|\mathbf{s}\|_1 \\ & \text{subject to} && \mathbf{y} = \mathbf{A}\mathbf{s}, \\ & && \mathbf{\Psi}\mathbf{s} = \mathbf{M}\mathbf{y}_s + \mathbf{w}. \end{aligned} \quad (6.8)$$

Remark 1. *This is to be noted that, in power distribution network, the measurements can be obtained using various measuring devices, such as, remote terminal units (RTUs), smart*

meter concentrators or advance meter infrastructure (AMI), and phasor measurement units (PMUs). However, these sensors record data at different time intervals, ranging from high-frequency readings in PMUs (several samples/sec) to less frequent data collection, such as 15-minute intervals in AMI. The amalgamation of these multi-time scale measurements poses a considerable challenge in unifying them into a coherent time frame. Several research efforts have emerged to address this issue before feeding the data into the state estimation engine. Notably, our previous work in [128] introduces a multitask Gaussian process tailored to handle unevenly sampled time series data by leveraging the inherent spatio-temporal correlations. The application of a similar method can be instrumental in processing the disparate multi-time scale data obtained from various sources within the power distribution network, harmonizing them into a shared time frame. Subsequently, these processed data can be fed as measurements to the deep unfolding-based state estimation approaches proposed in this work.

A well-known algorithm that solves (6.8) is the alternating direction method of multipliers (ADMM) [118], which considers an equivalent generalized least absolute shrinkage and selection operator (LASSO) form of (6.8) as,

$$\begin{aligned} & \underset{\mathbf{s}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda \|\mathbf{s}\|_1 \\ & \text{subject to:} \quad \mathbf{\Psi}\mathbf{s} = \mathbf{M}\mathbf{y}_s + \mathbf{w}, \end{aligned} \tag{6.9}$$

where, $\lambda > 0$ is scalar regularization parameter. Introducing an auxiliary variables $\mathbf{z} \in \mathbb{R}^N$, the (6.9) can be rewritten as,

$$\begin{aligned} & \underset{\mathbf{s}, \mathbf{z}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda \|\mathbf{z}\|_1 \\ & \text{subject to:} \quad \mathbf{z} - \mathbf{s} = \mathbf{0}, \\ & \quad \mathbf{\Psi}\mathbf{s} = \mathbf{M}\mathbf{y}_s + \mathbf{w} \end{aligned} \tag{6.10}$$

Now, introducing $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^N$ and initializing $(\mathbf{x}^0, \mathbf{z}^0, \mathbf{u}_1^0, \mathbf{u}_2^0) = (0, 0, 0, 0)$, the optimization

problem (6.10) iterates as,

$$\begin{aligned} \mathbf{s}^{k+1} = & (\mathbf{A}^T \mathbf{A} + \rho \mathbf{\Psi}^T \mathbf{\Psi} + \rho \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{y} + \rho(\mathbf{z}^k - \mathbf{u}_1^k) \\ & + \rho \mathbf{\Psi}^T (\mathbf{M} \mathbf{y}_s + \mathbf{w} - \mathbf{u}_2^k)) \end{aligned} \quad (6.11)$$

$$\mathbf{z}^{k+1} = S_{\lambda/\rho}(\mathbf{s}^{k+1} + \mathbf{u}_1^k) \quad (6.12)$$

$$\mathbf{u}_1^{k+1} = \mathbf{u}_1^k + \mathbf{s}^{k+1} - \mathbf{z}^{k+1} \quad (6.13)$$

$$\mathbf{u}_2^{k+1} = \mathbf{u}_2^k + \mathbf{\Psi} \mathbf{s}^{k+1} - \mathbf{M} \mathbf{y}_s - \mathbf{w} \quad (6.14)$$

where, $\rho > 0$ is the penalty parameter and $S_{\lambda/\rho}$ is a soft threshold operator. The iterates (6.11) - (6.14) converge to p^* of (6.10), i.e., minimum $(\|\mathbf{y} - \mathbf{A} \mathbf{s}^k\|_2^2 + \|\mathbf{s}\|_1) \rightarrow p^*$ and $\mathbf{z}^k - \mathbf{s}^k = 0$ as $k \rightarrow \infty$ [118]. However, this process may take a large time to converge, and the convergence time increases with increase in system size. Hence, we propose a new distribution system state estimation method using deep unfolding and refer to this technique as DSSE-net. This method takes advantage of both the iterative algorithm and the function approximation techniques of neural networks to reduce the convergence time. As the problem formulation in (6.10) takes the linearized power flow constraints into account for the estimation approach, we name it model-based DSSSE-net (MBDSSE-net). Later in this section, we propose a model-free DSSSE-net, which excludes model-related constraints in the problem formulation (referred to as MFDSSE-net). In the following subsection, we discuss the MBDSSE-net formulation, followed by the MFDSSE-net.

6.3.1 Model based DSSE-Net

Deep unfolding is a model-based neural network (NN) that fuses principled algorithms with performance guarantees with NN tools. The authors of [129] summarize deep unfolding as *"given a model-based approach that requires an iterative inference method, we unfold the iterations into a layer-wise structure analogous to a neural network."* In other words, deep unfolding takes an iterative algorithm with a fixed number of iterations L , unfolds its struc-

ture, and introduces several trainable parameters. These parameters are then learned using techniques from deep learning (with suitable loss functions, stochastic gradient descent, and backpropagation). The resulting unfolded algorithm with learned parameters can then solve various tasks.

The ADMM iteration (6.13) derived in the previous section can be rewritten as,

$$\begin{aligned} \mathbf{u}_1^{k+1} &= (\mathbf{I} - \rho\mathbf{H})\mathbf{u}_1^k - \rho\mathbf{H}\Psi^T\mathbf{u}_2^k + \rho\mathbf{H}\mathbf{z}^k + \mathbf{q}_1 - \\ &\quad S_{\lambda/\rho}[(\mathbf{I} - \rho\mathbf{H})\mathbf{u}_1^k - \rho\mathbf{H}\Psi^T\mathbf{u}_2^k + \rho\mathbf{H}\mathbf{z}^k + \mathbf{q}_1], \end{aligned} \quad (6.15)$$

where,

$$\mathbf{H} = (\mathbf{A}^T\mathbf{A} + \rho\Psi^T\Psi + \rho\mathbf{I})^{-1} \in \mathbb{R}^{N \times N},$$

and

$$\mathbf{q}_1 = \mathbf{q}_1(\mathbf{y}) = (\mathbf{H}\mathbf{A}^T + \rho\mathbf{H}\Psi^T\mathbf{M})\mathbf{y} + \rho\mathbf{H}\Psi^T\mathbf{w} \in \mathbb{R}^N.$$

Similarly, substituting the value of (6.11) and (6.13) into (6.14), we get the update equation of $\mathbf{u}_2^{k+1}$ as,

$$\mathbf{u}_2^{k+1} = \rho\Psi\mathbf{H}\Psi^T\mathbf{u}_1^k + (\mathbf{I} - \rho\Psi\mathbf{H}\Psi^T)\mathbf{u}_2^k + \rho\Psi\mathbf{H}\mathbf{z}^k + \mathbf{q}_2, \quad (6.16)$$

where,

$$\mathbf{q}_2 = \mathbf{q}_2(\mathbf{y}) = \Psi\mathbf{H}\mathbf{A}^T + [(\rho\Psi\mathbf{H}\Psi^T - \mathbf{I})\mathbf{M}]\mathbf{y} + (\rho\Psi\mathbf{H}\Psi^T - \mathbf{I})\mathbf{w}.$$

Finally substituting the value of (6.11) into (6.12), we get the update for $\mathbf{z}^{k+1}$ as,

$$\mathbf{z}^{k+1} = S_{\lambda/\rho}[(\mathbf{I} - \rho\mathbf{H})\mathbf{u}_1^k - \rho\mathbf{H}\Psi^T\mathbf{u}_2^k + \rho\mathbf{H}\mathbf{z}^k + \mathbf{q}_1] \quad (6.17)$$

Concatenating the variables $\mathbf{u}_1^{k+1}$, $\mathbf{u}_2^{k+1}$, $\mathbf{z}^{k+1}$ into a single vector $\boldsymbol{\vartheta}^{k+1}$; (6.15), (6.16) and (6.17) can be written in the matrix form as,

$$\boldsymbol{\vartheta}^{k+1} = \boldsymbol{\chi}\boldsymbol{\vartheta}^k + \mathbf{I}_1\mathbf{b}_1 + \mathbf{I}_2S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}}\boldsymbol{\vartheta}^k + \mathbf{b}_2), \quad (6.18)$$

where, $\boldsymbol{\vartheta}^k = \begin{bmatrix} \mathbf{u}_1^k; & \mathbf{u}_2^k; & \mathbf{z}^k \end{bmatrix}^T \in \mathbb{R}^{3N}$,

$$\boldsymbol{\chi} = \begin{bmatrix} (\mathbf{I} - \rho\mathbf{H}) & -\rho\mathbf{H}\boldsymbol{\Psi}^T & \rho\mathbf{H} \\ \rho\boldsymbol{\Psi}\mathbf{H}\boldsymbol{\Psi}^T & (\mathbf{I} - \rho\boldsymbol{\Psi}\mathbf{H}\boldsymbol{\Psi}^T) & \rho\boldsymbol{\Psi}\mathbf{H} \\ \mathbf{O}_{N \times N} & \mathbf{O}_{N \times N} & \mathbf{O}_{N \times N} \end{bmatrix} \in \mathbb{R}^{3N \times 3N};$$

$$\mathbf{I}_1 = \begin{bmatrix} \mathbf{I}_{N \times N}; \mathbf{I}_{N \times N}; \mathbf{O}_{N \times N} \end{bmatrix} \in \mathbb{R}^{3N \times N};$$

$$\mathbf{I}_2 = \begin{bmatrix} -\mathbf{I}_{N \times N}; \mathbf{O}_{N \times N}; \mathbf{I}_{N \times N} \end{bmatrix} \in \mathbb{R}^{3N \times N};$$

$$\mathbf{b}_1 = \begin{bmatrix} \mathbf{q}_1; \mathbf{q}_2; \mathbf{0} \end{bmatrix} \in \mathbb{R}^{3N};$$

$$\mathbf{b}_2 = \begin{bmatrix} \mathbf{q}_1; \mathbf{0}; \mathbf{q}_2 \end{bmatrix} \in \mathbb{R}^{3N};$$

$$\tilde{\boldsymbol{\chi}} = \begin{bmatrix} (\mathbf{I} - \rho\mathbf{H}); & -\rho\mathbf{H}\boldsymbol{\Psi}^T; & \rho\mathbf{H} \end{bmatrix} \in \mathbb{R}^{N \times 3N}$$

We map the update (6.18) to L layers/iterations in an equivalent neural network as,

$$\begin{aligned} \eta_1(y) &= \mathbf{I}_1 \mathbf{b}_1 + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta}^k + \mathbf{b}_2), \\ \eta_k(\vartheta) &= \boldsymbol{\chi} \boldsymbol{\vartheta} + \mathbf{I}_1 \mathbf{b}_1 + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta} + \mathbf{b}_2), \quad k = 2, \dots, L. \end{aligned} \tag{6.19}$$

Here, $\boldsymbol{\Psi}$ represents the trainable parameters for the neural network, and $\eta_k(\vartheta)$ represents the k^{th} output layer of the neural network. From a CS standpoint, this is equivalent to training/learning the best transformation basis as part of the estimation process. We denote the concatenation of L such layers/iterations (all having the same $\boldsymbol{\Psi}$) as,

$$\eta_{\boldsymbol{\Psi}}^L(y) = \eta_L \circ \eta_{L-1} \circ \dots \circ \eta_1(y), \tag{6.20}$$

where, ‘ $\circ$ ’ represents the function composition operator. Finally, the output $\mathbf{s}$ is obtained by

applying an affine map Γ , which uses (6.11) to the final layer L as,

$$\begin{aligned} \mathbf{s} &= \Gamma(\eta_{\Psi}^L(y)) \\ &= (\mathbf{A}^T \mathbf{A} + \rho \Psi^T \Psi + \rho \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{y} + \rho(\mathbf{z} - \mathbf{u}_1^L) \\ &\quad + \rho \Psi^T (\mathbf{M} \mathbf{y} + \mathbf{w} - \mathbf{u}_2^L)), \end{aligned} \tag{6.21}$$

where, $\mathbf{u}_1^L$, $\mathbf{u}_2^L$, and $\mathbf{z}^L$ are obtained from $\boldsymbol{\vartheta}^L$. The final states are recovered as $\hat{\mathbf{x}} = \Psi \mathbf{s}$. While training the DSSE-net, there is a chance that the output norm falls out of a reasonable range. To avoid unreasonable out-of-bound norm, $\hat{\mathbf{x}}$ is passed through a limiting function $\zeta : \mathbb{R}^N \rightarrow \mathbb{R}^N$ defined as,

$$\zeta(x) = \begin{cases} x, & \text{if } \|x\|_2 \leq K_{out}, \\ K_{out}x/\|x\|_2, & \text{otherwise,} \end{cases}$$

for some constant value of $K_{out} > 0$.

The MBDSSE-net formulation requires distribution network model information (i.e., equation (6.2)) to estimate the system states accurately. However, in real-world applications, accurate system model information may not be available. This is mainly because of aging, missing, or inappropriate/inaccurate documentation of the distribution network information. Hence, we develop a purely data-driven, model-free DSSE-net in the next subsection.

6.3.2 Model-free DSSE-Net

A purely data-driven, model-free DSSE-net (MFDSSE-net) estimates the system states based on the available measurements. The network is trained offline and can estimate the system states in real-time. The optimization formulation for the MFDSSE-net can be derived from (6.10) by removing the linearized power flow constraints. The resulting formulation corre-

sponds to,

$$\begin{aligned} & \underset{\mathbf{s}, \mathbf{z}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda \|\mathbf{z}\|_1 \\ & \text{subject to:} \quad \mathbf{z} - \mathbf{s} = 0, \end{aligned} \tag{6.22}$$

The ADMM iterations to solve (6.28) are given by,

$$\mathbf{s}^{k+1} = (\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{y} + \rho(\mathbf{z}^k - \mathbf{u}^k)) \tag{6.23}$$

$$\mathbf{z}^{k+1} = S_{\lambda/\rho}(\mathbf{s}^{k+1} + \mathbf{u}^k) \tag{6.24}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{s}^{k+1} - \mathbf{z}^{k+1} \tag{6.25}$$

The ADMM iterations (6.30) and (6.31) can be further simplified as

$$\begin{aligned} \mathbf{u}^{k+1} &= (\mathbf{I} - \mathbf{H})\mathbf{u}^k + \mathbf{H}\mathbf{z}^k + \mathbf{q} - \\ & \quad S_{\lambda/\rho}((\mathbf{I} + \mathbf{H})\mathbf{u}^k + \mathbf{H}\mathbf{z}^k + \mathbf{q}), \\ \mathbf{z}^{k+1} &= S_{\lambda/\rho}((\mathbf{I} + \mathbf{H})\mathbf{u}^k + \mathbf{H}\mathbf{z}^k + \mathbf{q}), \end{aligned} \tag{6.26}$$

where,

$$\begin{aligned} \mathbf{H} &= \rho(\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} \in \mathbb{R}^{N \times N} \\ \mathbf{q} &= \mathbf{q}(\mathbf{y}) = (\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} \mathbf{A}^T \mathbf{y} \in \mathbb{R}^N \end{aligned}$$

Similar to the derivation for MBDSSE-net, we use the update rule (6.32) for MFDSSE-net derivation, and the resulting equation is given by,

$$\boldsymbol{\vartheta}^{k+1} = \boldsymbol{\chi} \boldsymbol{\vartheta}^k + \mathbf{I}_1 \mathbf{q} + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta}^k + \mathbf{q}) \tag{6.27}$$

$$\text{where, } \boldsymbol{\vartheta}^k = \begin{bmatrix} \mathbf{u}^k \\ \mathbf{z}^k \end{bmatrix} \in \mathbb{R}^{2N}; \boldsymbol{\chi} = \begin{bmatrix} \mathbf{I} - \mathbf{W} & \mathbf{W} \\ \mathbf{O}_{N \times N} & \mathbf{O}_{N \times N} \end{bmatrix} \in \mathbb{R}^{2N \times 2N}; \tilde{\boldsymbol{\chi}} = \begin{bmatrix} -(\mathbf{I} + \mathbf{W}) & \mathbf{W} \end{bmatrix} \in \mathbb{R}^{N \times 2N}; \mathbf{I}_1 = \begin{bmatrix} \mathbf{q} \\ \mathbf{O}_N \end{bmatrix} \in \mathbb{R}^{2N \times N}; \mathbf{I}_2 = \begin{bmatrix} -\mathbf{I} & \mathbf{O}_{N \times N} \\ \mathbf{O}_{N \times N} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{2N \times 2N}.$$

Once (6.33) is obtained, a L layered MFDSSE-net can be formulated by following the similar procedure discussed for MBDSSE-net. We give a pseudo-code of the proposed MBDSSE-net in Algorithm 4. A similar approach can be used for the MFDSSE-net.

Remark 2. *The model-based distribution system state estimation network (MBDSSE-net) is trained with the assumption of a known and unchanging underlying topology of the power distribution network. Nevertheless, practical power distribution networks often undergo topology alterations due to switch/circuit breaker statuses or contingencies within the network. Addressing these changes necessitates modifying the current formulation, aligning with the methodologies outlined in our prior work [130]. The approach in [130] involves a consolidated framework for topology identification and state estimation in low-observable power distribution networks. The core of the state estimation method aligns with a compressive sensing framework akin to the one employed in our present work. However, the authors present a mixed integer nonlinear programming (MINLP) to accommodate the topology change in the power distribution network. Furthermore, they explore mixed integer linear programming (MILP) employing convex relaxation and integrating auxiliary variables. Our future research direction can encompass adopting a similar formulation within our approach to accommodate topology variations within distribution networks. This potential modification will facilitate handling dynamic changes within the network topology, aligning more closely with real-world scenarios. By integrating elements akin to the MINLP and MILP strategies explored in the referenced work, our method can potentially enhance accuracy in state estimation while robustly addressing topology changes in power distribution networks. Alternatively, we have proposed a model-free approach for distribution system state estimation (MFDSSE-net). This method accurately estimates system states by leveraging the inherent sparsity of measurement data, eliminating the need for knowledge of the underlying topology of the distribution net-*

work.

Algorithm 4 Deep unfolding for power distribution network state estimation

Data: Network parameter: $\mathbf{M}, \mathbf{w}$, Actual state: $\mathbf{x}$, Measurement: $\mathbf{y}$

Result: Estimated state: $\hat{\mathbf{x}}$

Initialization $\lambda, \rho > 0, \mathbf{s}, \mathbf{z}, \mathbf{u}_1, \mathbf{u}_2, \Psi$, epoch =1, epochs, k=1, L **while** $epoch \neq epochs$ **do**

while $k \neq L$ **do**

 Obtain the first dual variable $\mathbf{u}_1^{k+1}$ using (15)

 Obtain the second dual variable $\mathbf{u}_2^{k+1}$ using (16)

 Obtain the auxiliary variable $\mathbf{z}^{k+1}$ using (17)

 Obtain the concatenation of $\mathbf{u}_1^{k+1}, \mathbf{u}_2^{k+1}, \mathbf{z}^{k+1}$ as $\boldsymbol{\vartheta}^{k+1}$ using (18)

 Obtain the value at last layer $\eta(\boldsymbol{\vartheta})$ using (19)

end

 Obtain the sparse vector $\mathbf{s}$ using (21)

 Obtain $\hat{\mathbf{x}}^{epoch} = \Psi \mathbf{s}$

 Calculate loss = $\text{MSE}(\hat{\mathbf{x}}^{epoch} - \mathbf{x})$

 Calculate gradient and update Ψ

end

Out: $\hat{\mathbf{x}} = \Psi \mathbf{s}$

6.4 Computational complexity

The compressive sensing-based state estimator uses ADMM for the solution of the l_1 norm minimization problem. The computational complexity of the ADMM method is decided by the size of the sensing matrix $\mathbf{A}$. For DSSE, the size of $\mathbf{A}$ is $M \times N$, where M is the length of the measurement vector and N is the total number of states (with $M < N$). Hence, the computational complexity of one iteration of the ADMM is $\mathcal{O}(N^2M)$. The worst case computational complexity for K iterations is $\mathcal{O}(KN^2M)$.

In the deep unfolding approach, all the computation burden is shifted to the training phase. Once the network is trained, the estimation process involves a matrix and vector multiplication. Hence the computational complexity reduces drastically to $\mathcal{O}(N^2)$. Here, $\mathcal{O}$ gives the worst-case computational complexity. However, the simulation time varies based on the number of constraints. For instance, the model-based approach includes the power flow constraints, whereas the model-free approach does not have it. Hence the simulation time is different for different methods proposed in this paper.

A comparative analysis of the computation time for DSSE using different methods evaluated on IEEE 123 bus network and a 559 bus network is given in Table 6.1. The table shows the simulation time required to estimate the system states for both distribution networks with varying CMR%. It can be seen from Table 6.1 that the simulation time increases with the increase in CMR% for CS-ADMM (model-free) and CS-ADMM(model-based) algorithms. This is because an increase in CMR % increases the size of sensing matrix $\mathbf{A}$ and hence the inverse calculation of the first part of (6.11) of CS-ADMM(model-based) and (6.29) of CS-ADMM(model-free) become complex and hence the simulation time increases. Furthermore, for the optimization approaches, CS-ADMM model-free and CS-ADMM model-based simulation times are different. This is because of the inclusion of the power flow constraints in the optimization framework. On the other hand, in the proposed MFDSSE-net and MBDSSE-net, the simulation time is almost constant. This is because, the computation burden is shifted to the training phase, and once the training is done, the evaluation phase only considers algebraic addition and multiplication. Hence, the addition of the power flow constraints does not have any effect in the inference phase. It can also be observed from Table 6.1 that as the network size increases from IEEE 123 bus to 559 bus system, the simulation time almost increases by three times for both CS-ADMM (model-free) and CS-ADMM (model-based) algorithms. However, in the proposed MFDSSE-net and MBDSSE-net, the increase in simulation time is less than two times.

Algorithm	CMR (%)	Simulation time (ms)	
		IEEE 123 bus network	559 node network
MFDSSE-net	10	14.20	25.8
MBDSSE-net		14.50	25.93
CS-ADMM (Model free)		60.20	181.22
CS-ADMM (Model based)		80.28	218.32
MFDSSE-net	20	14.19	26.23
MBDSSE-net		14.52	26.47
CS-ADMM (Model free)		60.28	185.27
CS-ADMM (Model based)		80.59	232.47
MFDSSE-net	30	14.57	26.29
MBDSSE-net		14.59	26.52
CS-ADMM (Model free)		68.32	210.27
CS-ADMM (Model based)		85.28	247.41
MFDSSE-net	50	15.57	28.20
MBDSSE-net		15.65	28.32
CS-ADMM (Model free)		83.32	240.29
CS-ADMM (Model based)		98.28	280.53
MFDSSE-net	70	15.68	28.82
MBDSSE-net		15.93	28.98
CS-ADMM (Model free)		89.37	270.37
CS-ADMM (Model based)		120.28	290.23
MFDSSE-net	80	15.93	28.85
MBDSSE-net		15.96	29.21
CS-ADMM (Model free)		93.35	290.53
CS-ADMM (Model based)		128.76	311.56

Table 6.1: Simulation time of different algorithms with varying CMR%

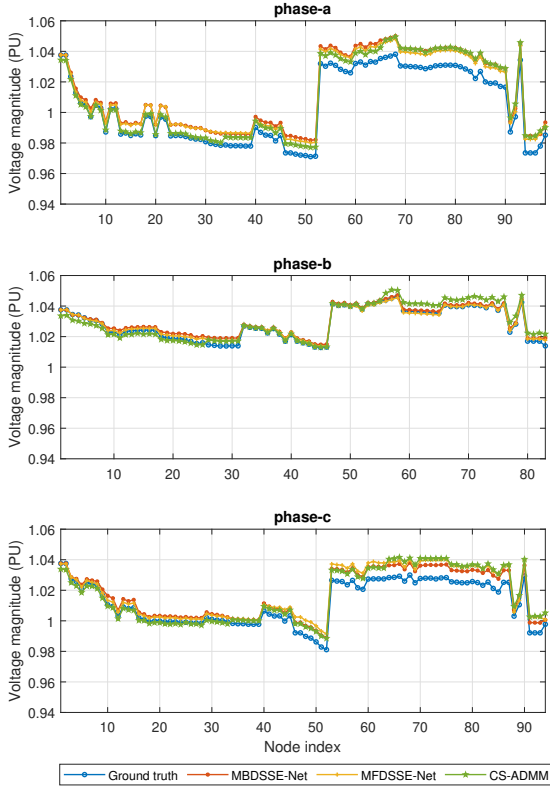


Figure 6.2: Voltage magnitude recovery in different phases of the IEEE-123 bus network for 50% CMR

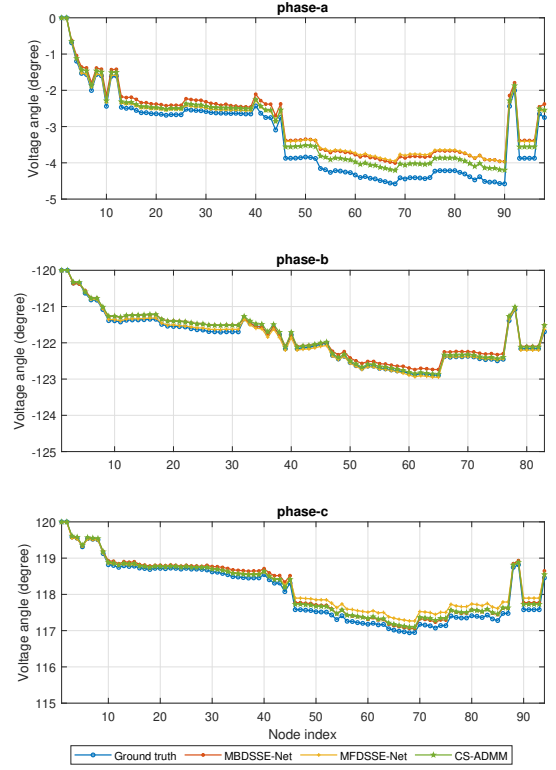


Figure 6.3: Voltage angle recovery in different phases of the IEEE-123 bus network for 50% CMR

6.5 Simulation Results

This section discusses the performance of the proposed deep unfolding-based networks, MFDSSE-net and MBDSSE-net. We compare the results with the ADMM-based optimization framework. First, we present the data used for the training and testing of the proposed approaches, followed by experimental results.

6.5.1 Data Preparation and network training

We evaluate the performance of the proposed deep unfolding-based state estimation methods MBDSSE-net and MFDSSE-net using the IEEE-123 bus benchmark system and a 559 bus distribution network introduced in [119]. We use real smart meter data from [131] for

generating training and testing data sets. We use the OpenDSSDirect python package [132] to solve the AC power flow and to generate the ground truth data for both the power distribution networks. Fifteen simulation days with a 15-minute resolution are generated to generate a total data set for 1440 instances. Out of which 1008 (70%) time instants data is used for training, 216 (15%) for testing, and 216 (15%) for validating the deep unfolding network. The training and testing measurements are obtained from the ground truth data generated, corrupting the data with an additive Gaussian noise of zero mean and a 10% variation of the measurement value. The deep unfolding NNs are trained using PyTorch in an NVIDIA GeForce RTX GPU with 32GB RAM machine. The weights (Ψ in our case) of both the networks MBDSSE-net and MFDSSE-net are learned by the backpropagation algorithm with ‘Adam’ optimizer and learning rate of 10^{-5} . Both MBDSSE-net and MFDSSE-net are trained for 100 epochs. The penalty parameter $\rho = 8$ is used for training and testing, and the number of layers/iterations chosen is 20 for both networks. These parameters are obtained using the validation data set of 216 (15%) time instants of the total data set. Given the network parameters and a training and testing data set $D_{Train} = (y, x)$, $D_{Test} = (\tilde{y}, \tilde{x})$, respectively. The network MBDSSE-net maps a function h_{D_T} to reconstruct $\mathbf{x}$ from $\mathbf{y} = \mathbf{Ax}$. To measure the difference between x_i and the $\hat{x}_i = h_{D_T}(y_i)$, the mean square error is chosen as the loss function for both the training and testing phases of the neural network.

6.5.2 Experimental results

In this subsection, we first evaluate the performance of the proposed method on the IEEE-123 bus network [120] followed by the 559 bus network.

The IEEE-123 bus network is unbalanced with single-phase, two-phase, and three-phase loads and overhead and underground lines. The training, testing and validation data are obtained by following the procedure explained in subsection 6.8.1. The estimation performance of both the methods MBDSSE-net and MFDSSE-net are evaluated in terms of mean absolute percentage error (MAPE) and mean integrated absolute error (MIAE) for voltage magnitude and angle, respectively, and are defined as, $MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right|$,

$MIAE = \sum_{i=1}^N \frac{|\theta_i - \hat{\theta}_i|}{N}$ [21]. Where, x_i and $\hat{x}_i$ are the true and estimated voltage magnitudes at bus i . Similarly, θ_i and $\hat{\theta}_i$ are the true and estimated voltage angles at bus i .

We compare the performance of our proposed deep unfolding-based neural networks MBDSSE-net and MFDSSE-net with the standard ADMM solver; hereafter, we mention it as CS-ADMM. We simulate both model-based CS-ADMM and model-free CS-ADMM. We use the penalty parameter $\rho = 8$ for all the methods, including the CS-ADMM. The simulation results for voltage magnitude and angle estimation in the different phases of the IEEE-123 bus network with 50% compressed measurement ratio (CMR) are shown in Figure 6.2 and Figure 6.3, respectively. In CS-ADMM, we have performed the simulation for both model-free and model-based approaches. However, we present the results in Figure 6.2 and Figure 6.3 only for the model-based CS-ADMM approach. The reason for not showing the results is that in the model-free CS-ADMM approach, the estimation accuracy is very low. Hence, the estimated values are far from the nominal value making the plot difficult to visualize. However, we present the results for all four methods in a tabular form in Table 9.1.

It is observed from Figure 6.2 and Figure 6.3 that the proposed MFDSSE-net and MBDSSE-net can estimate the system states accurately. The accuracy is comparable to the CS-ADMM optimization technique. One important point is that the MFDSSE-net does not require the network model information to learn the sparsifying basis matrix Ψ and still estimates the system states with very good accuracy comparable to the counterpart MBDSSE-net and model-based CS-ADMM optimization technique. Hence, the MFDSSE-net can be very useful in real-world applications where the network information is unavailable or poorly documented.

We evaluated the scalability of the proposed approaches on a larger distribution network introduced in [119]. This is a three-phase unbalanced distribution network consisting of both the primary and secondary level distribution systems that account for a total of 559 buses. The recovery of the system states on the 559 bus system for 30% CMR is shown in Figure 6.4. The top and bottom of Figure 6.4 show voltage magnitude and angle estimation for phase-a of the 559 bus network. Similar results are also observed for phase-b and phase-c. Figure 6.4

demonstrates the effective recovery of system states by the proposed deep unfolding-based state estimation methods, affirming their accuracy in large-scale distribution systems. This shows the scalability of the proposed methods for application in larger distribution networks.

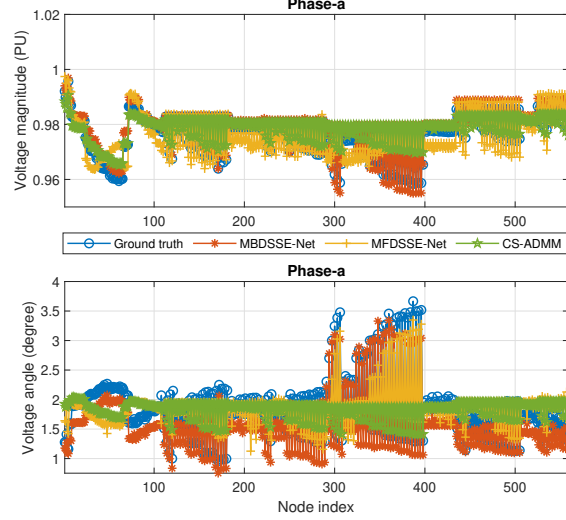


Figure 6.4: Voltage magnitude and angle estimation in the phase-a of the 559 bus distribution network at 30% of CMR

A comparative performance of all four methods, MFDSSE-net, MBDSSE-net, CS-ADMM (model-free), and CS-ADMM (model-based), is given in Table 9.1. The performance is compared based on errors in estimating the system states using MAPE and MIAE for voltage magnitude and angles, respectively, for the IEEE 123 bus and the 559-node distribution network. We also present the estimator performance for different levels of available measurements by varying the CMR from 10% to 80%. It can be observed from Table 9.1 that error in voltage magnitude estimation at 10% CMR is least for MFDSSE-net, followed by MBDSSE-net, CS-ADMM (model-based) and CS-ADMM (model-free) for both the distribution networks. It is observed that the model-free CS-ADMM fails to estimate the system states accurately. However, the MFDSSE-net can learn the sparsifying basis matrix from the data and accurately estimate the system states at CMR as low as 10%. The errors in voltage angle estimation also follow a similar observation to the error in voltage magnitude estimation, i.e., the angle error estimation is less in MFDSSE-net, followed by MBDSSE-net,

Test network	Algorithm	CMR (%)	Evaluation metric		Algorithm	CMR (%)	Evaluation metric	
			MAPE (%)	MIAE (%)			MAPE (%)	MIAE (%)
IEEE 123 bus network	MFDSSE-net	10	1.3403	0.5396	MFDSSE-net	50	1.5256	0.5741
	MBDSSE-net		1.6185	0.6431	MBDSSE-net		1.6189	0.6434
	CS-ADMM (Model free)		12.8627	14.64	CS-ADMM (Model free)		5.0820	10.7514
	CS-ADMM (Model based)		2.2921	2.35	CS-ADMM (Model based)		2.2176	2.355
	MFDSSE-net	20	1.6134	0.6199	MFDSSE-net	70	1.6101	0.5741
	MBDSSE-net		1.6195	0.6433	MBDSSE-net		1.6178	0.6434
	CS-ADMM (Model free)		9.7246	13.2192	CS-ADMM (Model free)		3.6877	10.7514
	CS-ADMM (Model based)		2.264	2.35	CS-ADMM (Model based)		2.2114	2.35
	MFDSSE-net	30	1.2126	0.4220	MFDSSE-net	80	0.5585	0.3860
	MBDSSE-net		1.6192	0.6432	MBDSSE-net		0.4835	0.3321
	CS-ADMM (Model free)		7.4971	12.9353	CS-ADMM (Model free)		3.3373	6.6379
	CS-ADMM (Model based)		2.2384	2.35	CS-ADMM (Model based)		2.2105	2.35
559 bus network	MFDSSE-net	10	1.1397	0.8669	MFDSSE-net	50	0.6992	0.4989
	MBDSSE-net		1.1587	0.9280	MBDSSE-net		0.2895	0.3793
	CS-ADMM (Model free)		14.3421	15.7290	CS-ADMM (Model free)		3.1419	0.2083
	CS-ADMM (Model based)		2.7467	2.9201	CS-ADMM (Model based)		2.7193	0.1833
	MFDSSE-net	20	0.8514	0.6282	MFDSSE-net	70	0.6984	0.4283
	MBDSSE-net		0.8728	0.7929	MBDSSE-net		0.7235	0.4528
	CS-ADMM (Model free)		11.2354	13.2323	CS-ADMM (Model free)		2.8461	0.2075
	CS-ADMM (Model based)		2.7306	2.1852	CS-ADMM (Model based)		2.7173	0.1825
	MFDSSE-net	30	0.7046	0.5289	MFDSSE-net	80	0.4259	0.4280
	MBDSSE-net		0.8569	0.6289	MBDSSE-net		0.2753	0.3585
	CS-ADMM (Model free)		4.0242	0.2233	CS-ADMM (Model free)		2.7920	0.2052
	CS-ADMM (Model based)		2.7244	0.1843	CS-ADMM (Model based)		2.7165	0.1623

Table 6.2: Performance comparison of the proposed approaches

CS-ADMM (model-based) and CS-ADMM (model-free) for 10% CMR. Similar observations are observed for the rest of the CMR values. However, for 80% CMR value, the MAPE and MIAE for voltage magnitude and angle are least in MBDSSE-net followed by MFDSSE-net, CS-ADMM (model-based), and, CS-ADMM (model-free). This may be because the network becomes observable at such a high value of CMR, for which the MBDSSE-net approach takes advantage and learns the sparsifying matrix using the data and model information. As a result, the estimation accuracy is higher. However, exploiting the sparsity for the MFDSSE-net approach is difficult as the basis matrix is only learned using the data. Hence, it recovers the system states with an accuracy slightly less than the MBDSSE-net.

6.5.3 Robustness to Measurement Noise

In this subsection, we evaluate the performance of the proposed approaches to non-Gaussian measurement noise. The state estimation method proposed in our study does not operate under the constraint that the measurement noise confirms a Gaussian distribution rather uses this assumption for simulation studies. This assumption aligns with common practices adopted in numerous prior studies, as indicated in (6.1) of the revised manuscript and further supported by the works of references [53, 59, 128, 130]. The current form of our

proposed approach is not confined to exclusively Gaussian measurement noise. To assess its adaptability to non-Gaussian noise distributions, we have evaluated its performance under Laplacian measurement noise. The performance assessment revealed that the MBDSSE-net exhibits commendable accuracy in reconstructing system states, irrespective of the nature of the measurement noise. Figure 6.5 shows the recovery of system states on the IEEE 123 bus network with a 30% (CMR) having a 10% variation of the measurement value following a Laplacian noise distribution. We have shown the results for phase-a only in Figure 6.5. However, similar results are obtained for the other two phases. The simulation result shows that the proposed method is robust against various types of measurement noise.

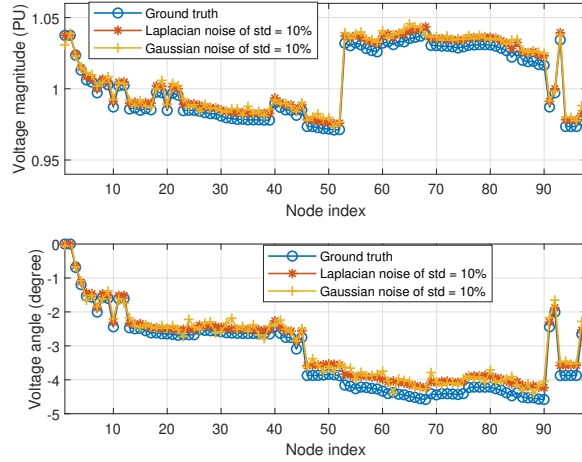


Figure 6.5: Voltage magnitude and angle estimation in the phase-a of the IEEE 123 bus distribution network at 30% of CMR with different measurement noise

6.5.4 Robustness to Simulation Model

The proposed MFDSSE-net and MBDSSE-net are trained on the data generated from a simulation environment. Hence, the fidelity of the simulation environment may directly affect the performance of the proposed MFDSSE-net model and, hence, the accuracy of the estimation. We have used the OpenDSS distribution network simulation tool to produce the training data. The use of OpenDSS for power distribution network modeling and analysis is accepted by the research community and the industry, which gives the first level of validation

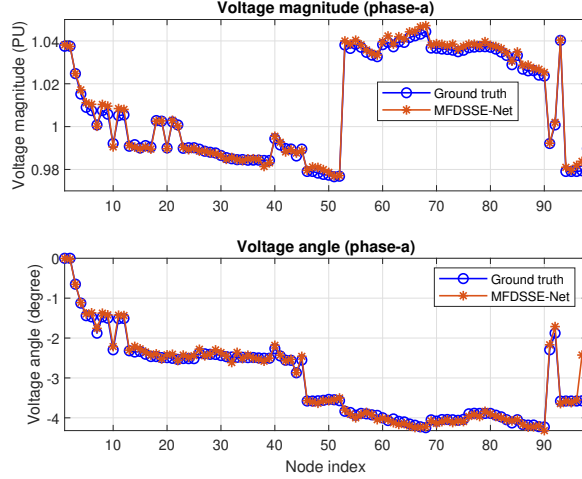


Figure 6.6: State recovery with inaccurate simulation model data

of the training data generation. To check the fidelity of the OpenDSS simulation tool further, we have intentionally generated a test data set by opening one of the normally closed switches of the IEEE 123 bus network. This process alters the network topology and produces a wrong test data set. However, the training data remain unchanged. We estimate the system states using this test data, considering 30% CMR. The estimation of the system states under the changed topology of the IEEE-123 bus network is shown in Figure 6.6. We have shown the results for the phase only. However, similar results are obtained for the other two phases. It is observed from Figure 6.6 that MFDSSSE-net can recover the system states accurately, illustrating the robustness of the approach to training model inaccuracies. The voltage magnitude estimation error MAPE and the voltage angle estimation MIAE are 1.41% and 0.38%, respectively.

6.6 Uncertainty Estimation

we propose a Bayesian deep unfolding to estimate both the mean and variance of the estimated states.

6.7 Proposed Approach

We model the power distribution network as a connected and undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$; where, $\mathcal{V} = 1, 2, \dots, N$ is the set of nodes, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. Let $|V_n|e^{j\theta_n}$ and $S_n = P_n + jQ_n$ represents the complex voltage and complex power injections at bus $n \in \mathcal{V}$, respectively. The measurement vector $\mathbf{y} = [|V_n|^T, \theta_n^T, P_n^T, Q_n^T]^T$ includes all the aggregated SCADA sensor and/or μ PMU data from the network. We solve the compressive sensing problem (6.6) to estimate the system states using the measurement vector $\mathbf{y}$. In order to solve (6.6), we first formulate the generalized least absolute shrinkage and selection operator (LASSO) of (6.6) and then solve using alternating direction method of multipliers (ADMM) [118] as follows,

$$\begin{aligned} & \underset{\mathbf{s}, \mathbf{z}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda \|\mathbf{z}\|_1 \\ & \text{subject to:} \quad \mathbf{z} - \mathbf{s} = 0, \end{aligned} \tag{6.28}$$

Here, $\mathbf{z} \in \mathbb{R}^N$ is an intermediate variable and $\lambda > 0$ is the regularization parameter. The ADMM iterations to solve (6.28) are given by,

$$\mathbf{s}^{k+1} = (\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{y} + \rho(\mathbf{z}^k - \mathbf{u}^k)) \tag{6.29}$$

$$\mathbf{z}^{k+1} = S_{\lambda/\rho}(\mathbf{s}^{k+1} + \mathbf{u}^k) \tag{6.30}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{s}^{k+1} - \mathbf{z}^{k+1} \tag{6.31}$$

Where, $\mathbf{u} \in \mathbb{R}^N$ is the dual variable and $S_{\lambda/\rho}$ is the soft threshold operator.

6.7.1 Deep unfolding from ADMM iterations - DSSE-net

The ADMM iterations (6.30) and (6.31) can be further simplified as

$$\begin{aligned}\mathbf{u}^{k+1} &= (\mathbf{I} - \mathbf{W})\mathbf{u}^k + \mathbf{W}\mathbf{z}^k + \mathbf{q} - \\ &\quad S_{\lambda/\rho}((\mathbf{I} + \mathbf{W})\mathbf{u}^k + \mathbf{W}\mathbf{z}^k + \mathbf{q}), \\ \mathbf{z}^{k+1} &= S_{\lambda/\rho}((\mathbf{I} + \mathbf{W})\mathbf{u}^k + \mathbf{W}\mathbf{z}^k + \mathbf{q}),\end{aligned}\tag{6.32}$$

where,

$$\begin{aligned}\mathbf{W} &= \rho(\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} \in \mathbb{R}^{N \times N} \\ \mathbf{q} &= \mathbf{q}(\mathbf{y}) = (\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} \mathbf{A}^T \mathbf{y} \in \mathbb{R}^N\end{aligned}$$

We map the update rule (6.32) to form L layers/iterations in an equivalent neural network as,

$$\boldsymbol{\vartheta}^{k+1} = \boldsymbol{\chi} \boldsymbol{\vartheta}^k + \mathbf{I}_1 \mathbf{q} + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta}^k + \mathbf{q})\tag{6.33}$$

$$\begin{aligned}\text{where, } \boldsymbol{\vartheta}^k &= \begin{bmatrix} \mathbf{u}^k \\ \mathbf{z}^k \end{bmatrix} \in \mathbb{R}^{2N}; \boldsymbol{\chi} = \begin{bmatrix} \mathbf{I} - \mathbf{W} & \mathbf{W} \\ \mathbf{O}_{N \times N} & \mathbf{O}_{N \times N} \end{bmatrix} \in \mathbb{R}^{2N \times 2N}; \tilde{\boldsymbol{\chi}} = \begin{bmatrix} -(\mathbf{I} + \mathbf{W}) & \mathbf{W} \end{bmatrix} \in \\ \mathbb{R}^{N \times 2N}; \mathbf{I}_1 &= \begin{bmatrix} \mathbf{q} \\ \mathbf{O}_N \end{bmatrix} \in \mathbb{R}^{2N \times N}; \mathbf{I}_2 = \begin{bmatrix} -\mathbf{I} & \mathbf{O}_{N \times N} \\ \mathbf{O}_{N \times N} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{2N \times 2N}.\end{aligned}$$

We map the update (6.33) to L layers/iterations in an equivalent neural network, and named it “DSSE-net”. The output of each layers of DSSE-net is given as,

$$\eta_1(\mathbf{y}) = \mathbf{I}_1 \mathbf{q} + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta}^k + \mathbf{q}),\tag{6.34}$$

$$\eta_k(\boldsymbol{\vartheta}) = \boldsymbol{\chi} \boldsymbol{\vartheta} + \mathbf{I}_1 \mathbf{q} + \mathbf{I}_2 S_{\lambda/\rho}(\tilde{\boldsymbol{\chi}} \boldsymbol{\vartheta} + \mathbf{q}), \quad k = 2, \dots, L.\tag{6.35}$$

$$\mathbf{s} = (\mathbf{A}^T \mathbf{A} + \rho \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{y} + \rho(\mathbf{z}^L - \mathbf{u}^L)),\tag{6.36}$$

$$\hat{\mathbf{x}} = \boldsymbol{\Psi} \mathbf{s}.\tag{6.37}$$

Here, $\mathbf{u}^L$, and $\mathbf{z}^L$ are obtained from $\boldsymbol{\vartheta}^L$. $\eta_k(\boldsymbol{\vartheta})$ represents the k^{th} output layer of the

neural network. Ψ represents the trainable parameters for the neural network. From a CS standpoint, this is equivalent to training/learning the best transformation basis as part of the estimation process. While training the DSSE-net, there is a chance that the output norm falls out of a reasonable range. To avoid unreasonable out-of-bound norm, $\hat{\mathbf{x}}$ is passed through a limiting function $\zeta : \mathbb{R}^N \rightarrow \mathbb{R}^N$ defined as,

$$\zeta(x) = \begin{cases} x, & \text{if } \|x\|_2 \leq K_{out}, \\ K_{out}x/\|x\|_2, & \text{otherwise,} \end{cases}$$

for some constant value of $K_{out} > 0$.

The solution (6.37) obtained using the DSSE-net is not centered around the ground truth value $\mathbf{x}^*$. Hence, we need to debias it. The debiased LASSO construction proposed in [?] is based on the KKT conditions and is given by,

$$\mathbf{x}^d = \hat{\mathbf{x}} + \frac{1}{M}\mathbf{\Omega}\mathbf{A}^T(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}), \quad (6.38)$$

where, $\mathbf{x}^d$ is the debiased recovered state vector and $\mathbf{\Omega}$ is a matrix selected to approximately invert the sample covariance matrix $\mathbf{\Sigma} = \frac{1}{M}\mathbf{A}^T\mathbf{A}$, i.e., $\mathbf{\Omega}\mathbf{\Sigma} \approx \mathbf{I}_{N \times N}$. The difference between the debiased LASSO and the ground truth is given as,

$$\sqrt{M}(\mathbf{x}^d - \mathbf{x}^*) = \frac{\mathbf{\Omega}\mathbf{A}^T\boldsymbol{\eta}}{\sqrt{M}} - \sqrt{M}(\mathbf{\Omega}\mathbf{\Sigma} - \mathbf{I}_{N \times N})(\hat{\mathbf{x}} - \mathbf{x}^*) \quad (6.39)$$

Where, the first term in (6.39) has a normal distribution i.e., $\frac{\mathbf{\Omega}\mathbf{A}^T\boldsymbol{\eta}}{\sqrt{M}} \sim \mathcal{N}(0, \sigma^2\mathbf{\Omega}\mathbf{\Sigma}\mathbf{\Omega}^T)$ and the second term is the remainder term which vanishes asymptotically [133],[134]. Hence, it is possible to construct the confidence intervals for the ground truth $\mathbf{x}^*$ based on $\mathbf{x}^d$.

In the next sub section, we discuss the Bayesian NN (BNN) construction for estimating the mean and covariance of the estimates obtained using DSSE-net.

6.7.2 Deep unfolding using BNN

The DSSE-net discussed in subsection 6.7.1 provides point estimate and does not say anything about the estimation confidence. For safety-critical applications, it is imperative to give a confidence interval of the predictions. In this subsection, we explore the use of BNN in deep unfolding to estimate both the mean and the variances of the estimated states.

The likelihood function $p(\mathbf{x}|\mathbf{y}, \psi)$ of the problem in (6.6) has the form.

$$p(\mathbf{x}|\mathbf{y}, \psi) = \mathcal{N}(\mathbf{x}|f_\psi(\mathbf{y}), \sigma^2\mathbf{I}), \quad (6.40)$$

where, $f_\psi : \mathbb{R}^M \rightarrow \mathbb{R}^N$ is the deep unfolding network parameterized by ψ , and σ is a fixed constant. Considering the training dataset consisting of measurement vectors and the true states, we can compute the MAP estimate of the set of parameters by solving the following optimization:

$$\hat{\psi}_{\text{MAP}} = \underset{\psi}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} \sum_{i=1}^{N_D} \|\mathbf{x}^{[i]} - f_\psi(\mathbf{y}^{[i]})\|_2^2 - \log p(\psi) \right\}$$

where, $(\mathbf{y}^{[i]}, \mathbf{x}^{[i]})$ is the i^{th} training data pair. N_D is the number of training datasets, $p(\psi)$ is the prior distribution of the parameter set. In the inference stage, for a given $\mathbf{y}_*$, we can calculate the prediction distribution $p(\mathbf{x}_*|\mathbf{y}_*, \hat{\psi}_{\text{MAP}})$. The probabilistic formulation implicitly appears in the training and inference phase of the deep unfolding network discussed in subsection 6.7.1. However, the discussed approach has two major problems in characterizing the uncertainties of prediction; (1) the formulation fails to model the uncertainty in prediction for a given measurement vector as it assumes a fixed parameter model of the covariance matrix, and (2) the formulation does not model any parameter uncertainties as it only uses a point estimate by following MAP estimation principle.

A Bayesian Neural network can solve the above two problems[135]. The first problem is solved by defining a likelihood function to model the inherent uncertainty on the reconstruction task. In the case of a Gaussian likelihood, this can be accomplished by defining a neural

network for the covariance matrix. BNN solves the second problem by defining the prior distribution on the set of likelihood function parameters and then computing the posterior distribution of the parameters given a training dataset. In the inference stage, BNN finds the predictive distribution for a given measurement vector $\mathbf{y}_*$ using

$$p(\mathbf{x}_*|\mathbf{y}_*, \mathcal{D}) = \int p(\mathbf{x}_*|\mathbf{y}_*, \boldsymbol{\psi})p(\boldsymbol{\psi}|\mathcal{D})d\boldsymbol{\psi}, \quad (6.41)$$

where, $p(\mathbf{x}_*|\mathbf{y}_*, \mathcal{D})$ is the prediction distribution. The first term in (6.41) is the likelihood function, which incorporates the aleatoric uncertainty. The second term in (6.41) is the posterior distribution of the network parameters and incorporates the epistemic uncertainty. Thus a BNN can quantify both aleatoric and epistemic uncertainty for a deep unfolding based NN for distribution system state estimation.

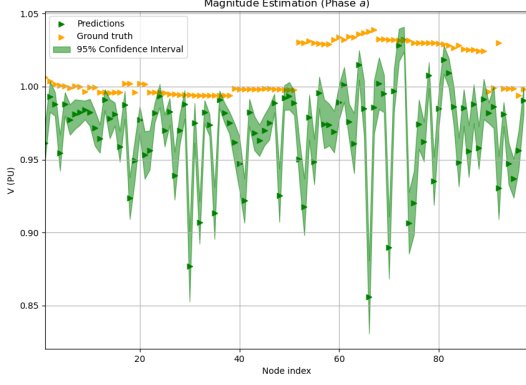
We define the likelihood function as

$$p(\mathbf{x}|\mathbf{y}, \boldsymbol{\psi}) = \mathcal{N}(\mathbf{x}|f_\mu(\mathbf{y}), \text{diag}(\sigma_\rho^2(\mathbf{y}))),$$

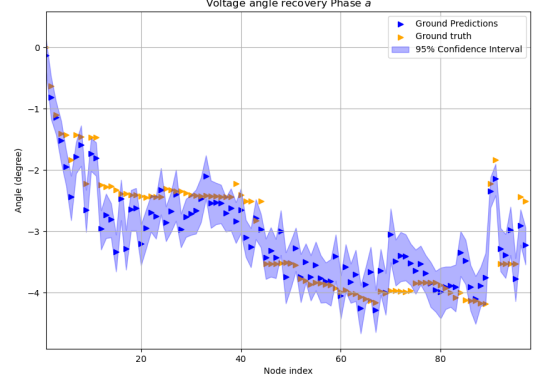
where, $\boldsymbol{\psi} = \mu \cup \rho$, and $f_\mu : \mathbb{R}^M \rightarrow \mathbb{R}^N$ is the neural network parameterized by sets of parameters μ maps a given measurement to a point estimate. Similarly, $\sigma_\rho^2 : \mathbb{R}^M \rightarrow \mathbb{R}^N$ is the neural network parameterized by ρ , aims to capture the uncertainty of the estimates for a given measurement. The likelihood implicitly assumes that $\mathbf{x} = f_\mu(y) + \boldsymbol{\epsilon}_l$, where, $\boldsymbol{\epsilon}_l \sim \mathcal{N}(0, \text{diag}(\sigma_\rho^2(\mathbf{y})))$, i.e., the aleatoric uncertainty modeled as additive Gaussian noise.

The true posterior distribution of the network parameters is intractable because of the complex functional form and large number of network parameters. Hence, it is important to approximate the posterior distribution. One of the methods for posterior approximation is to use variational inference called MC dropout [136]. For a detailed discussion on variational inference using MC dropout, the readers can refer to [137].

Suppose the neural networks f_μ and σ_ρ^2 to estimate the mean and variance contain L_1 and L_2 number of layers/iterations, respectively. Then, we can write the two sets μ and ρ



(a) Voltage magnitude recovery with 95% CI



(b) Voltage angle recovery with 95% CI

Figure 6.7: Voltage estimation with debiased DSSE-net (30% CMR)

as,

$$\mu = \bigcup_{i=1}^{L_1} \{\mathbf{W}_i^f\} \text{ and } \rho = \bigcup_{j=1}^{L_2} \{\mathbf{W}_j^\sigma\} \quad (6.42)$$

where, $\mathbf{W}_i^f$ and $\mathbf{W}_j^\sigma$ are the vectors consisting of the parameters of the NNs f_μ and σ_ρ^2 , respectively. We use variational inference MC dropout parametric distribution $q(\mathbf{w}|\theta)$ for approximation of the true posterior distribution $p(\boldsymbol{\psi}|D)$. We assume that the layers of the two neural networks f_μ and σ_ρ^2 are independent and the layers within the NNs are mutually independent i.e., posterior parameterized by θ , $q(\boldsymbol{\psi}|\theta)$. In this approach, the BNN learns the parameter θ^* , which minimizes Kullback-Leibler (KL) divergence between the variational and the true posterior. Thus the variational parameters θ of the approximate variational distribution are computed by minimizing the the loss function given by (6.43),

$$\begin{aligned} \theta^* &= \underset{\theta}{\operatorname{argmin}} KL[q(\boldsymbol{\psi}|\theta)||p(\boldsymbol{\psi}|D)] \\ &= \underset{\theta}{\operatorname{argmin}} KL[q(\boldsymbol{\psi}|\theta)||p(\boldsymbol{\psi})] - \mathbb{E}_{q(\boldsymbol{\psi}|\theta)}[\ln(p(D|\boldsymbol{\psi}))] \end{aligned} \quad (6.43)$$

Training the NN using the loss function given in (6.43) is called BayesByBackprop [138]. For a Gaussian variational prior the first term of (6.43) is calculated analytically where as the expectation in the second term is calculated by Monte Carlo sampling, i.e., $\mathbb{E}_{q(\boldsymbol{\psi}|\theta)}[\ln(p(D|\boldsymbol{\psi}))] = \frac{1}{N_D} \sum_{i=1}^{N_D} \ln(p(D/\boldsymbol{\psi}^{(i)}))$, with $\boldsymbol{\psi}^{(i)} \sim q(\theta|\boldsymbol{\psi})$.

After finding the approximate posterior distribution, in the inference stage we compute the inference probability of the output $\mathbf{x}^*$ for a measurement $\mathbf{y}^*$ by using the two NNs f_μ and σ_ρ^2 for the mean and variance is given by,

$$\mathbb{E}[\mathbf{x}^*|\mathbf{y}^*, \boldsymbol{\psi}] = \mathbb{E}_{p(\boldsymbol{\psi}|D)}[f_\mu(\mathbf{y}^*)] \quad (6.44)$$

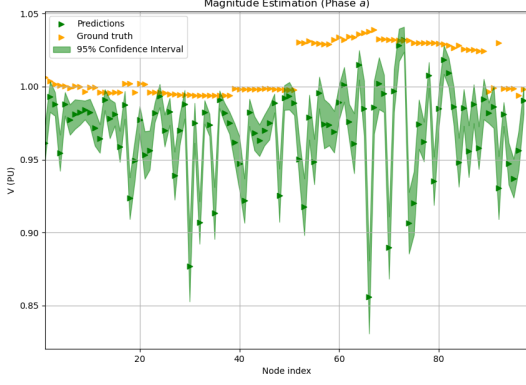
$$\begin{aligned} \text{Var}(\mathbf{x}^*|\mathbf{y}^*, \mathcal{D}) &= \mathbb{E}_{p(\boldsymbol{\psi}|D)}[\text{Var}(\mathbf{x}^*|\mathbf{y}^*, \boldsymbol{\psi})] + \\ &\quad \text{Var}_{p(\boldsymbol{\psi}|D)}(\mathbb{E}[\mathbf{x}^*|\mathbf{y}^*, \boldsymbol{\psi}]) \\ &= \underbrace{\sigma(\mathbf{x}^*)^2}_{\text{Aleatoric variance}} + \underbrace{\text{Var}_{p(\boldsymbol{\psi}|D)}(f_\mu(\mathbf{y}^*))}_{\text{epistemic uncertainty}} \end{aligned} \quad (6.45)$$

6.8 Simulation Results

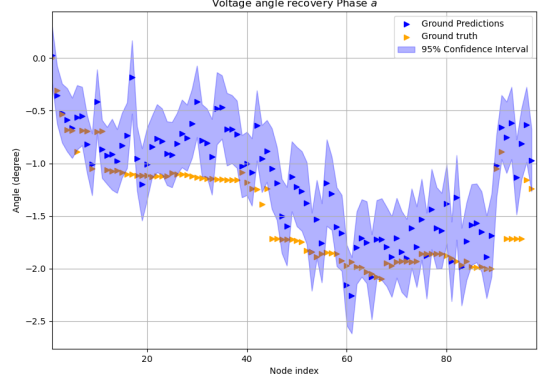
This section discusses the performance of the proposed Bayesian deep unfolding-based DSSE. Before diving into the simulation results, we present the data used for training the network, followed by experimental results.

6.8.1 Data Preparation and network training

We evaluate the performance of the proposed Bayesian deep unfolding network using the IEEE-123 bus benchmark system. We use real smart meter data from [131] for generating training and testing data sets. We use the OpenDSSDirect python package [132] to solve the AC power flow and to generate the ground truth data. Fifteen simulation days with a 15-minute resolution are generated to form a data set corresponding to 1440 instances. 1008 (70%) time instants data is used for training, 216 (15%) for testing, and 216 (15%) for validating the deep unfolding network. The training and testing measurements are obtained from the ground truth data generated, corrupting the data with an additive Gaussian noise of zero mean and a 10% variation of the measurement value. The proposed network is trained using PyTorch in an NVIDIA GeForce RTX GPU with 32GB RAM machine. The mean and variance of the sparsifying matrix $\boldsymbol{\Psi}$ of both the networks f_μ for mean prediction



(a) Voltage magnitude recovery with 95% CI



(b) Voltage angle recovery with 95% CI

Figure 6.8: Voltage estimation with biased DSSE-net (30% CMR)

and σ_p^2 variance prediction, respectively are learned by the back propagation algorithm with ‘Adam’ optimizer and learning rate of 10^{-5} . The network is trained for 250 iterations. The penalty parameter $\rho = 8$ is used for training and testing, and the number of layers/iterations chosen is 15 for both mean prediction network f_μ and variance prediction network σ_p^2 . These parameters are obtained using the validation data set of 216 (15%) time instants of the total data set.

6.8.2 Experimental results

In this subsection, we evaluate the performance of the proposed method on the IEEE-123 bus network [120].

The IEEE-123 bus network is unbalanced with single-phase, two-phase, and three-phase loads and overhead and underground lines. The training, testing, and validation data are obtained by following the procedure explained in subsection 6.8.1. The estimation performance of the proposed Bayesian deep unfolding network is evaluated in terms of mean absolute percentage error (MAPE) and mean integrated absolute error (MIAE) for voltage magnitude and angle, respectively, and are defined as, $MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right|$, $MIAE = \sum_{i=1}^N \frac{|\theta_i - \hat{\theta}_i|}{N}$ [21]. Where x_i and $\hat{x}_i$ are the true and estimated voltage magnitudes at bus i , similarly, θ_i and $\hat{\theta}_i$ are the true and estimated voltage angles at bus i .

Network	MAPE (%)	MIAE
biased DSSE-net	5.01	0.38
debiased DSSE-net	4.48	0.32

Table 6.3: DSSE-net performance with 30% CMR

We present the simulation results for both the biased and debiased Bayesian DSSE-net. In case of biased DSSE-net, the neural network construction excludes the use of (6.38) in its architecture. However, all other parameters are the same as in the debiased DSSE-net. The simulation results for voltage magnitude and angle estimation for a 30% compressed measurement ratio (CMR) using the debiased and biased DSSE-net is shown in Fig. 6.7 and Fig. 6.8, respectively. Here, we include the results of the only phase-a. However, similar observations can be obtained for the other two phases. A careful observation of Fig. 6.7a and Fig. 6.8a shows that, in the case of debiased DSSE-net, the use of (6.38) removes the bias and ensures that the estimates are centered around the ground truth value. As the estimates are near the ground truth value, the variances in the estimation are low. On the other hand, the estimates obtained using biased DSSE-net have a skewed structure, i.e. all the estimated values lie below the ground truth values (see Fig. 6.8a). As the estimates are away from their true value, the variance is more. Similar observations can be drawn from angle estimates as shown in Fig. 6.7b and Fig. 6.8b for debiased and biased DSSE-net, respectively. The performance of both the networks in estimating the mean voltage magnitude and angle is given in Table 6.3. It can be seen from Table 6.3 that both networks perform equally well in estimating the mean values of the system states. However, the Bayesian debiased DSSE-net estimation has less variance than the biased DSSE-net.

6.9 Summary

This chapter proposes a novel model-informed, learning-based distribution system state estimation technique. We formulate two deep unfolding NNs, namely MFDSSE-net and MBDSSE-net. The MFDSSE-net is a purely data-driven neural network, whereas the MBDSSE-net is a physics-informed deep unfolding neural network.

The performance of the proposed networks is evaluated on the IEEE-123 bus distribution network by considering real-world load data. Simulation results show that the MFDSSE-net, MBDSSE-net, and model-based CS-ADMM estimate the system states with reasonable accuracy even for 10% CMR. The estimation process proposed here uses deep learning techniques. Hence, it is imperative to give the associated confidence bound of the estimations. Hence, we propose a Bayesian deep unfolding network that can give point estimates and the associated variance. By offering insights into the uncertainties of the state estimates, this approach improves situational awareness and empowers distribution network operators to make more informed control and management decisions.

Chapter 7

Cyber Attack Analysis

This chapter discusses the robustness of the proposed CS- based DSSE to various cyber-attacks.

7.1 Cyber Attack modeling

The modern distribution grid is configured with smart sensing devices like smart meters, remote terminal units (RTUs), phasor measurement units (PMUs), and communication infrastructure. All these smart devices may communicate with each other and the control center wirelessly or via Ethernet connections. This exposes the distribution grid measurements to cyber intrusions [139]. Cyber-attacks on measurements in power grids can adversely impact state estimation, which in turn can cause catastrophic system failures. For example, a successful synchronized cyber-attack on Ukrainian power distribution grids in 2015 led to a system-wide power outage for around 6 hours switching off 30 substations, causing considerable economic loss and seriously affecting citizens' lives [140]. Similarly, deception attacks on memory-based event-triggering schemes used for load frequency control have been shown to have a disastrous effect on the power system [141]. Therefore, it is essential to analyze the robustness of the distributed state estimation strategies to cyber intrusions. This chapter analyzes the impacts of some of the critical cyber-attacks (e.g., loss of measurement

data, false data injection (FDI), replay, and neighborhood attack) on distributed grid state estimation.

7.2 Cyber Attack Models

1. *Loss of measurement*: As discussed earlier, sparsity-based state estimations like CS exploit the spatial correlation of the measurement data. In the distribution grid, the number of measurement data is scarce; any further loss in measurement data can have a detrimental effect on estimation accuracy. Further details can be obtained from sections AMI.19, AMI.28, WAMPAC.6, etc. of [142].
2. *False data injection (FDI) attack*: It is a process of compromising the sensor reading such that the estimation process fails to detect and produces wrong values for the estimation variables. Here, we have considered two ways to corrupt the data, namely additive corruption and multiplicative corruption. The detailed procedure to realize the FDI attack is given below;

Multiplicative attack model:

$$\hat{z}_i = (1 \pm 0.1r)z_i, \quad (7.1)$$

where, z_i and $\hat{z}_i$ is the actual and corrupted measurement at node i respectively. The degree of corruption is decided by r , and its value is chosen from a uniform probability distribution between 0 and 1.

Additive attack model:

$$\hat{z}_i = z_i \pm 0.1r\bar{z}, \quad (7.2)$$

where, $\bar{z}$ is the mean of all the similar measurements, i.e., if z_i is the voltage measurement, then $\bar{z}$ is the mean of all the available voltage measurements. The $\pm$ ve sign in (7.1) and (7.2) shows the corruption direction. The +ve sign leads to over-attack, and the -ve sign leads to under-attack. Related scenarios from [142] include AMI.27, WAMPAC.3, etc.

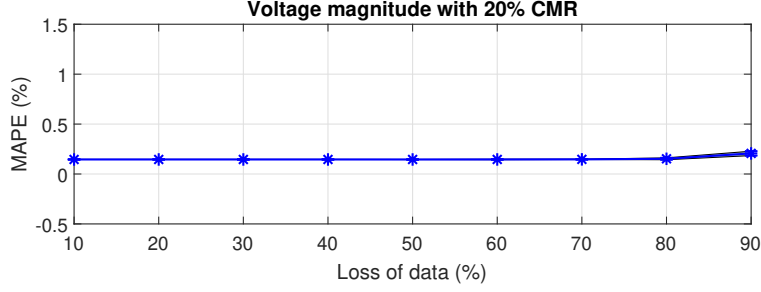
3. *Neighborhood attack*: Analysis of this attack plays an important role in identifying the areas of attack that can create maximum disruption to distribution network state estimation. This attack is realized by compromising all the measurements from a physical neighborhood. Related scenarios from [142] include AMI.27, WAMPAC.3, etc.
4. *Replay attack*: This attack is considered the easiest attack to implement. In this framework, measurements from the previous time step are passed on as a current measurement to the estimator. This attack is realized by considering maximum variability in voltage and power to simulate maximum corruption.

7.3 Simulation Results

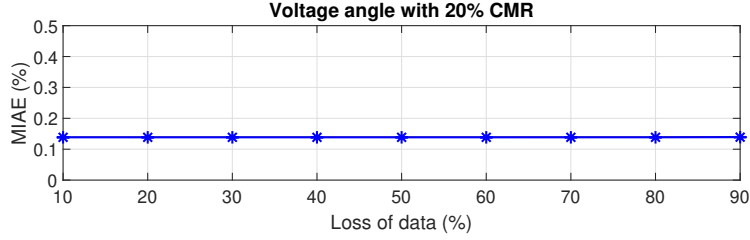
In order to study the effect of cyber attacks, we first design our sensing architecture and attack setup on the IEEE 123 bus test system, as discussed earlier. The IEEE 123-bus test system is a standard network, and it has 260 single-phase nodes; this excludes the slack bus. We have considered measuring devices that measure active power P , reactive power Q injections, and the voltage phasor. The total number of measurements possible is $(260 \times 3) = 780$. For the cyber intrusion scenario creation, the total number of measurements available is considered to be 20% of the maximum possible measurements, i.e., $(780 \times 0.2) = 156$. These measurements are randomly selected for each scenario.

We first present the simulation results for the loss of measurements attack. Out of the 20% of the available measurements data, we consider that 10% to 90% of the data is removed. We have performed ten simulations for each random removal of data. Figure 7.1 shows the voltage magnitude recovery and voltage angle recovery to loss of measurement attack. It can be observed that the performance of the voltage magnitude estimation deteriorates only marginally as the % of measurement loss increases beyond 80%. This shows the robustness of the distributed CS-based state estimation to the loss of measurement data.

Now we will discuss the robustness of the algorithm to FDI attacks. In this case out of



(a) Voltage magnitude recovery with varying CMR



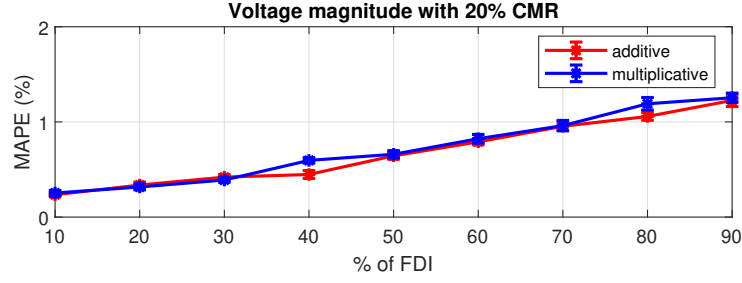
(b) Voltage angle recovery with varying CMR

Figure 7.1: Voltage recovery under loss of measurement attack

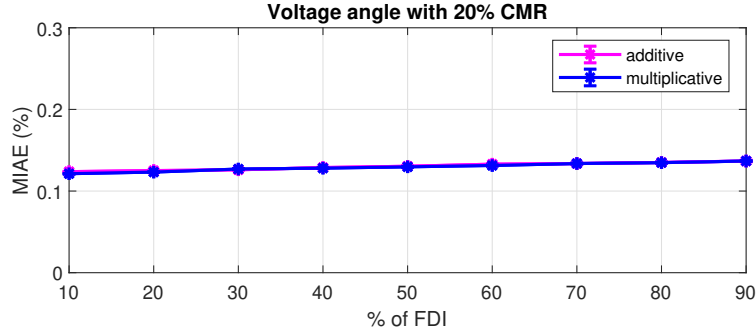
the 20% available data, 10% to 90% of data is randomly corrupted using (7.1) for additive and using (7.2) for multiplicative corruption, respectively. Here, we present only the over attack scenario. However, similar observations can be obtained for under attack scenarios as well. Figure 7.2 shows the plot of MAPE and MIAE against the increase in % of data corruption. It is observed that the error in voltage magnitude and the angle increases with an increase in % of data corruption. However, the maximum value of MAPE and MIAE is well below the permissible limit, which shows the algorithm's robustness to FDI attacks.

Next, we present the neighborhood attack as shown in Figure 7.3. For our simulation, neighborhoods are selected based on geographical proximity. As discussed earlier, the IEEE123-bus network is divided into three sub-networks of an approximately equal number of single-phase nodes. Figure 7.3(a) and 7.3(b) shows the MAPE in voltage magnitude and MIAE in voltage angle, respectively. Each plot shows the error in estimation when all the measurements of that area are corrupted. It can be observed that MAPE is more when the measurements of area 2 are corrupted. Hence it can be inferred that area 2 has the highest impact on the state estimation accuracy and should be given extra preference in security.

Next, we discuss the impact of replay attack in Figure 7.4. The attack is realized by

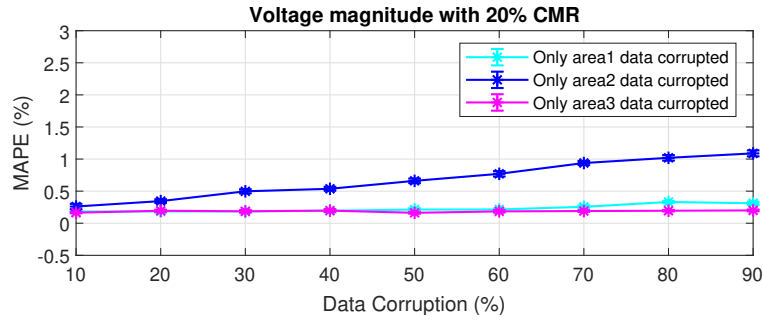


(a) Voltage magnitude recovery with varying CMR

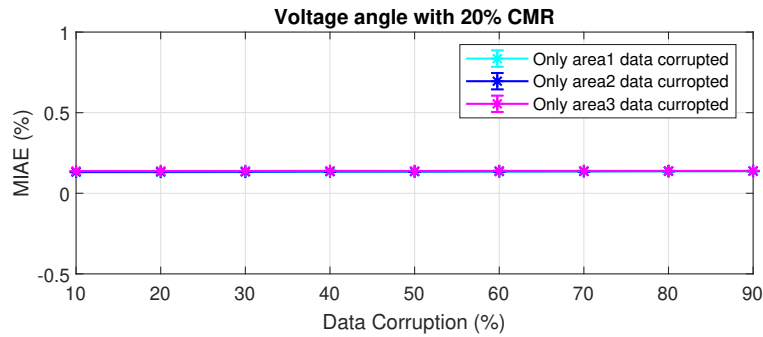


(b) Voltage angle recovery with varying CMR

Figure 7.2: Voltage recovery under FDI attack

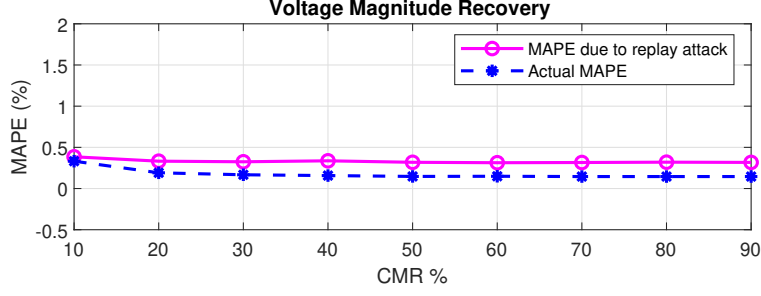


(a) Voltage magnitude recovery with varying CMR

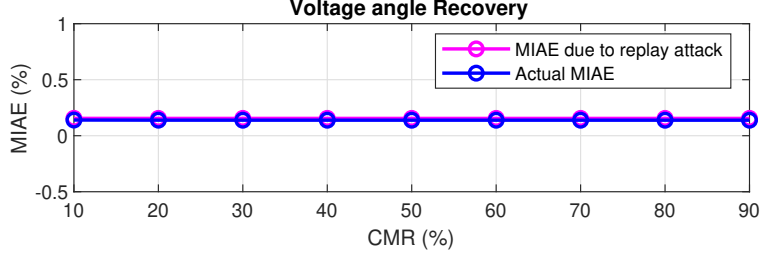


(b) Voltage angle recovery with varying CMR

Figure 7.3: Voltage recovery under neighborhood attack



(a) Voltage magnitude recovery with varying CMR

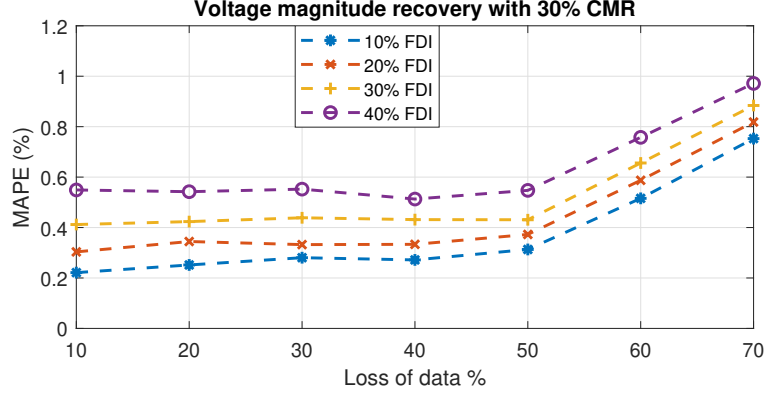


(b) Voltage angle recovery with varying CMR

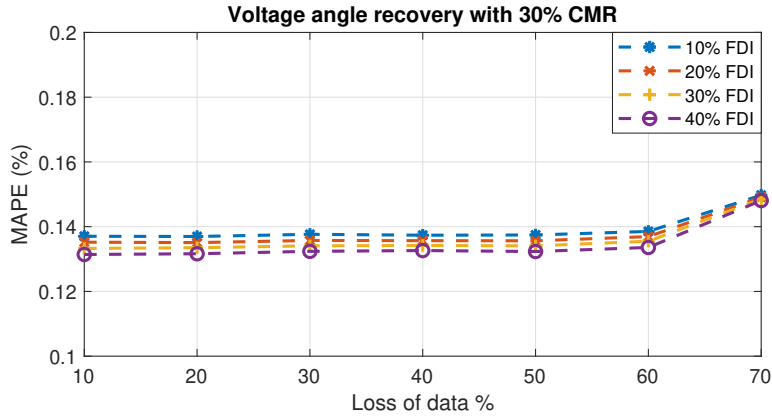
Figure 7.4: Voltage recovery under replay attack

considering a previous time step measured value as the current input to the estimator. The attack is checked under two highly heterogeneous load profiles, i.e., the measurements of the peak load duration are fed as the input to the estimator when the load curve attends its valley during the day. Figure 7.4(a) shows the MAPE in voltage magnitude, and Figure 7.4(b) shows the MIAE in voltage angle. It can be observed from the figure that MAPE due to replay attack is more at each CMR compared to the current time step estimation. This is because of the difference in load variation. However, both MAPE and MIAE are well within the permissible limits. The effect of a replay attack on the state estimator output can be different for different networks.

Finally, we discuss the multi-attack scenario and its effect on the distributed CS-based state estimation. We have considered a multi-cyber attack case that includes loss of measurement (e.g., due to DOS attack) and FDI (multiplicative) attack simultaneously. The attack scenario is evaluated on the IEEE 37 bus test system at 30% CMR. The voltage magnitude and angle recovery in the multi-attack scenario are shown in Figure 7.5 (a) and Figure 7.5 (b), respectively. As expected, it is observed that the estimation accuracy decreases with an increase in data loss and FDI attacks.



(a) Voltage magnitude recovery



(b) Voltage angle recovery

Figure 7.5: Voltage recovery in multi-cyber attack (loss of data + FDI)

7.4 Summary

In this chapter, we have analyzed the robustness of the proposed distributed algorithm to four commonly occurring cyber-attacks, namely loss of measurements, FDI, replay, and neighborhood attacks. Simulation results show that the CS-based distributed state estimation is highly robust to all four attacks, with the replay attack being the least effective in creating a significant error in state estimation results.

Chapter 8

Microgrid Secondary Control

In this chapter, we discuss our proposed game theoretic-based microgrid secondary control.

8.1 Introduction

The global energy landscape has radically transformed in recent years. Some driving forces for the transformation include (1) pressing concerns posed by climate change, (2) the rise in energy prices, and (3) increasing concerns about energy security and reliability. As traditional centralized power systems grapple with these challenges, microgrids (MGs) with distributed energy resources (DERs) emerge as a novel alternative. MGs are localized, autonomous energy systems integrating various DERs, energy storage devices, loads, and other power electronic equipment [89, 90].

In the conventional power grid, synchronous generators maintain voltage and frequency stability because of their inbuilt inertia [3]. At the same time, ensuring frequency and voltage stability is crucial in MGs. Generally, the microgrids are flooded with IBRs that lack inertia. Droop control has emerged as a dominant solution to regulate voltage and frequency in MGs because of its simple implementation [92]. Moreover, droop control only relies on local measurements and requires no communication [93]. However, due to its inherent characteristics, droop control may not be precise and can cause voltage and frequency deviations

from their nominal values [4]. Therefore, a secondary level of hierarchical control becomes necessary to compensate for the voltage and frequency deviations from its normal values [94]. The secondary control acts as a compensation mechanism. The approaches to implement secondary control in MG architecture are centralized, decentralized, or distributed. Details of the different secondary control architecture in MG setup can be found in [95]. Centralized control has an easier decision-making process and coordination [4]. However, it suffers from large communication bandwidth requirements, which leads to decision delays and low reliability. On the other hand, the decentralized controller does not require exclusive communication infrastructure (CI) to restore the frequency and voltage. Each DER restores its frequency and voltage amplitudes to nominal values individually. However, the CI is required for transferring data from the tertiary control and MG centralized control layers [96]. On the contrary, distributed secondary controllers use a low bandwidth sparse communication network in which the inverters are only required to interact with the neighboring inverters.

8.2 Background and Preliminaries

In this section, we introduce some useful background related to microgrids that will drive our follow-on secondary control design discussions.

8.2.1 Primary control of MG

AC microgrid consists of DC sources behind a DC/AC inverter and an LC filter connected to the point of common coupling (PCC) through an RL output connector [97]. The inverters normally operate on voltage control mode during the islanded operation of MG [98]. A typical DG model with the inverter control loops is shown in Fig. 3.1. The major control loops of the inverter control are the current control loop, the voltage control loop, and the power control loop, which employs the droop controller. A detailed mathematical model of the control loops is explained in [99]. The power control loop generates the voltage references V_{od}^* , V_{oq}^* based on the set points provided by the secondary control, where, V_{od}^* and V_{oq}^* are direct and quadrature axis voltages, respectively. The voltage controller generates

the direct and quadrature axis current references i_{Ld}^* and i_{Lq}^* , respectively. Then, the current control loop is activated to regulate the outputs of the inverter by generating sinusoidal pulse width modulation (SPWM). The power control loop employs droop control. Droop-based control is an autonomous control that requires no communication links among the DERs. In an islanded microgrid, the objective of droop control is to maintain voltage and frequency stability keeping the active and reactive power sharing. The two governing equations for the primary droop control of i^h DG are as follows [94].

$$\begin{cases} \omega_i = \omega_n - m_{pi}P_i \\ V_i^{od} = V_n - n_{qi}Q_i \\ V_i^{oq} = 0 \end{cases} \quad (8.1)$$

where, ω_n and V_n are the nominal system frequency and nominal voltage amplitude, respectively. ω_i is the output frequency. V_i^{od} , and V_i^{oq} are the direct and quadrature axis voltages, respectively. P_i and Q_i are the measured active and reactive power, respectively. m_{pi} and n_{qi} are the droop gain for frequency and voltage, respectively. The value of m_{pi} and n_{qi} are calculated as, $m_{pi} = \frac{\omega_{\max} - \omega_{\min}}{P_{\max}}$, $n_{qi} = \frac{V_{i\max} - V_{i\min}}{Q_{\max}}$. The measured power P_i , and Q_i are obtained by filtering the converter output through a low pass filter with a cutoff frequency of ω_{ci} . Mathematically, $P_i = \frac{1}{s + \omega_{ci}}p_i$, $Q_i = \frac{1}{s + \omega_{ci}}q_i$. Where, p_i and q_i are the DG active and reactive power output, respectively. Generally, for simple control purposes, V_{oq} is set to zero, and hence, the voltage magnitude is expressed as $V_i = \sqrt{(V_i^{od})^2 + (V_i^{oq})^2} = V_i^{od}$. Hence, the primary droop control in (9.2b) is expressed as,

$$\omega_i = \omega_n - m_{pi}P_i \quad (8.2)$$

$$V_i = V_n - n_{qi}Q_i \quad (8.3)$$

8.2.2 Secondary Control of Microgrid

The conventional way of compensating for the deviation in frequency and voltage amplitude after the primary droop control is through a centralized controller. The centralized controller employs a proportional-integral controller given as [94, 100],

$$\delta\omega_i = K_{P\omega} (\omega_n - \omega_i) + K_{I\omega} \int (\omega_n - \omega_i) dt. \quad (8.4)$$

$$\delta V_i = K_{PV} (V_n - V_i) + K_{IV} \int (V_n - V_i) dt, \quad (8.5)$$

Where, K_{PV} , K_{IV} , $K_{P\omega}$ and $K_{I\omega}$, are the proportional and integral (PI) gain for the voltage and frequency controller, respectively. V_n and ω_n is the nominal voltage and reference frequency respectively. δV_i and $\delta\omega_i$ are the offsets in voltage and frequency, respectively (generated by the secondary control). These offset parameters are sent to the primary droop controller to compensate for the voltage and frequency. Although the controller works well, designing and tuning the PI parameters is difficult. Furthermore, the centralized controller needs a high bandwidth communication network and suffers from the single-point failure issue. Therefore, we propose a game-theoretic approach for the secondary control of frequency in an islanded microgrid. This approach requires localized communications limited among neighboring IBRs. Next, we discuss a general model for the communication network in MG.

8.2.3 Communication network modeling

Consider a communication network that enables information exchange among inverters. This network is represented by a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a nonempty finite set of N vertices (an example is shown in Fig. 3.2). The node-set, the edges, and the graph adjacency matrix are defined as $\mathcal{V} = \{v_1, v_2 \dots v_N\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$, respectively. The IBRs are considered the vertices of the communication graph, and the edges of the communication network form the links. The edges from node v_j to v_i is denoted by (v_j, v_i) , i.e., node v_j receives information from v_i . The weight of the communication link

is given by $a_{ij} > 0$ if $(v_j, v_i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The in-degree matrix is defined by $\mathbf{D} = \text{diag}(d_i) \in \mathbb{R}^{N \times N}$ with $d_i = \sum_{j=1}^N a_{ij}$. Finally, the graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The Laplacian matrix $\mathbf{L}$ has all row sums equal to zero, i.e., $\mathbf{L}\mathbf{1}_N = 0$, where $\mathbf{1}_N$ is a column of all ones. The graph $\mathcal{G}$ is called balanced if its Laplacian matrix $\mathbf{L}$ meets $\mathbf{1}_N^T \mathbf{L} = 0$. A graph is said to have a spanning tree if there exists a node (called the root) that has a directed path from this node to other nodes in the graph [101]. The graph is said to be connected if the algebraic connectivity i.e., the second smallest eigenvalue, is non-zero.

8.3 Proposed Approach

In this paper, we propose a game-theoretic distributed secondary control of frequency for islanded microgrids. As the dynamics of the inner control loops are much faster than the droop control, the secondary control is designed by considering only the power dynamics [64]. We first introduce the game design for distributed frequency control.

8.3.1 Game design for distributed frequency control

The frequency-active power droop control action for i^{th} DG is given in (8.2). Differentiating (8.2) with respect to time yields,

$$\dot{\omega}_{ni} = \dot{\omega}_i + m_{pi} \dot{P}_i \quad (8.6)$$

By denoting $\dot{\omega}_i = u_i$, (8.6) can be written as,

$$\dot{\omega}_{ni} = u_i + m_{pi} \dot{P}_i. \quad (8.7)$$

Hence, a careful design of the secondary control of the i^{th} DG, u_i can yield the nominal frequency as,

$$\omega_{ni} = \int_0^t (u_i + m_{pi} \dot{P}_i) dt. \quad (8.8)$$

Based on the above discussion, we try to address the following problem.

Problem 1. Consider the AC microgrid discussed in 9.2.2 with frequency dynamics in (8.6). The goal is to design the secondary controller u_i in (8.8) such that,

$$\lim_{t \rightarrow \infty} (\omega_i(t) - \omega_j(t)) = 0 \text{ and } \omega_i(t) = \omega^*, \forall i, j \in \mathcal{V}, \quad (8.9)$$

where, ω^* is the desired frequency of the microgrid. To solve *Problem 1*, we formulate the controller design problem as a game played among the IBRs. The cost function for each IBR corresponds to,

$$J_i(\omega_i, \boldsymbol{\omega}_{i-}) = -\omega_n \omega_i + \frac{1}{2} \omega_i^2 + \beta \omega_i \sum_{j \in N_i} a_{i,j} \omega_j \quad \forall \omega_i \in \mathcal{V} \quad (8.10)$$

where, $\omega_i \in \Omega$ is the action taken by i^{th} IBR, and $\boldsymbol{\omega}_{i-}$ is the action taken by all other IBRs in the MG. ω_n is the nominal frequency. N_i is the neighbor set of IBR i and it is defined as $N_i = \{j = 1, \dots, N | (i, j) \in \mathcal{E}\}$. $\beta \in [0, 1]$ is the payoff parameter that measures the neighbors' effect on the cost function of i^{th} IBR.

Next, we define the frequency error of i^{th} IBR from the microgrid desired frequency ω^* as,

$$\delta_i = \omega_i - \omega^*, \quad \forall i \in \mathcal{V}. \quad (8.11)$$

The time derivative of (8.11) gives error dynamics of i^{th} IBR as,

$$\dot{\delta}_i = \dot{\omega}_i = u_i, \quad \forall i \in \mathcal{V}. \quad (8.12)$$

From (8.11) and (8.12), it is clear that, we have to design u_i such that $\delta_i = \dot{\delta}_i = 0$. To achieve that, we design the controller as,

$$u_i = \min_{\omega_i} J_i(\omega_i, \boldsymbol{\omega}_{i-}), \quad (8.13)$$

where, $J_i(\omega_i, \boldsymbol{\omega}_{i-})$ is the cost function of each player defined in (8.10).

Consider a N player game (i.e., an MG with N IBRs), where the cost function of each player of the game is defined as in (8.10). Let, $\boldsymbol{\omega} = [\omega_1, \omega_2, \dots, \omega_N]^T \in \mathbb{R}^N$ be the vector of all player's actions; $\omega_i \in \mathbb{R}_+$ is the action of player i and $\boldsymbol{\omega}_{i-} = [\omega_i, \omega_2, \dots, \omega_{i-1}, \omega_{i+1}, \dots, \omega_N] \in \mathbb{R}^{N-1}$ is the actions of all players but i .

Assumption 1: The players can communicate with their neighbors via the un-directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, discussed in 9.2.2.

The Nash equilibrium of the game is $\boldsymbol{\omega}^* = (\omega_i^*, \boldsymbol{\omega}_{i-}^*)$ which minimizes the cost function is defined as follows.

Note: The Nash equilibrium achieved in the described game is a pure Nash equilibrium, in contrast to scenarios with discrete strategy spaces where mixed Nash equilibria are possible.

Definition 1. *Nash equilibrium* (NE) of an N -player game is defined as, “the action profile on which no player can reduce its own cost by unilaterally changing its own action i.e., $\boldsymbol{\omega}^* = (\omega_i^*, \boldsymbol{\omega}_{i-}^*)$ is the Nash equilibrium of the N -player game if for $\omega_i \in \mathbb{R}$ ”,

$$J_i(\omega_i^*, \boldsymbol{\omega}_{i-}^*) \leq J_i(\omega_i, \boldsymbol{\omega}_{i-}^*), \forall i. \quad (8.14)$$

8.3.2 Existence and Uniqueness of NE

It is clear from the game formulation that each player chooses their own action ω_i to minimize its cost function $J_i(\omega_i, \boldsymbol{\omega}_{i-})$, given others' actions $\boldsymbol{\omega}_{i-}$, the communication graph $\mathcal{G}$, represented by the adjacency matrix $\mathbf{A}$, and the parameter β . If the cost function of each player J_i is continuous and twice differentiable in ω_i , minimizing J_i with respect to ω_i gives the best response,

$$\omega_i = \omega_n - \beta \sum_{j=N_i} a_{i,j} \omega_j \quad (8.15)$$

where, ω_n is the nominal frequency of the microgrid. Each player tries to reach ω_n by combining its own and neighbor's actions. If the weighted sum of the neighbor's actions is less than ω_n , player i makes up the difference; otherwise, the player takes no action. Let $\boldsymbol{\omega}$ be the best reply of all the players in the communication graph represented by the adjacency

matrix $\mathbf{A}$. Using matrix notation in (8.15) yields,

Proposition 1. *Action $\boldsymbol{\omega}$ is a Nash equilibrium of the N -player game if and only if,*

$$(\mathbf{I} + \beta \mathbf{A})\boldsymbol{\omega} = \mathbf{1}\omega_n, \quad (8.16)$$

where, $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix and $\mathbf{1} \in \mathbb{R}^N$ is a column vector of all 1. *Proof:* We write the cost function defined in (8.10) for all players in matrix form as,

$$\mathbf{J} = -\mathbf{1}\omega_n \circ \boldsymbol{\omega} + \frac{1}{2}\boldsymbol{\omega} \circ \boldsymbol{\omega} + \beta \boldsymbol{\omega} \circ (A\boldsymbol{\omega}) \quad (8.17)$$

Given other players actions $\boldsymbol{\omega}_{i-}$, finding the best reply of i^{th} IBR is to find the minimum value of (8.17), which can be obtained by the first order necessary condition. Applying the first-order necessary condition on (8.17) yields the NE condition defined by (8.16).

Proposition 1 can be used to find all the equilibria of any graph $\mathbf{A}$ and β . However, the equilibrium point changes with change in β or the change in graph structure. Hence, next, we define the criteria for a unique and stable equilibrium. To achieve this, we use the theory of potential games [143].

Consider a game with cost functions $J_i(\omega_i, \boldsymbol{\omega}_{i-})$. A function $\Phi_i(\omega_i, \boldsymbol{\omega}_{i-})$ is the potential function for this game if and only if $\forall \omega_i, \omega_i'$ and $\boldsymbol{\omega}_{i-}$ satisfies,

$$\Phi_i(\omega_i, \boldsymbol{\omega}_{i-}) - \Phi_i(\omega_i', \boldsymbol{\omega}_{i-}) = J_i(\omega_i, \boldsymbol{\omega}_{i-}) - J_i(\omega_i', \boldsymbol{\omega}_{i-}), \forall i. \quad (8.18)$$

A potential function mirrors the cost function of each player. If there is a change in the action of i^{th} player from ω_i to ω_i' , the potential function changes by the same amount as its cost function would change. A potential function exists for a game with cost function J_i , if and only if (8.19) holds.

$$\frac{\partial^2 J_i(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} = \frac{\partial^2 J_j(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j}, \forall i \neq j \quad (8.19)$$

For the game played among the IBRs, with each having a cost function J_i , there exists a potential function as $\frac{\partial^2 J_i(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} = \frac{\partial^2 J_j(\boldsymbol{\omega})}{\partial \omega_i \partial \omega_j} = \beta a_{i,j}$, and it is given as,

$$\phi(\omega_i, \boldsymbol{\omega}_{i-}) = \boldsymbol{\omega}^T \mathbf{1} \omega_n + \frac{1}{2} \boldsymbol{\omega}^T (\mathbf{I} + \beta \mathbf{A}) \boldsymbol{\omega}. \quad (8.20)$$

Note: The second term in (8.20), $(\mathbf{I} + \beta \mathbf{A})$ relates to the Nash equilibrium obtained in proposition 1. This term will be the central term of our discussion.

Lemma 1. *For given communication infrastructure $\mathbf{A}$ and parameter β , the set of Nash equilibria is either a saddle point or minima for the potential function $\Phi(\boldsymbol{\omega})$.*

Proof: Consider minimizing the potential function,

$$\min_{\boldsymbol{\omega}} \Phi(\omega_i, \boldsymbol{\omega}_{i-}), \forall i \quad (8.21)$$

Actions $\boldsymbol{\omega}$ is a Nash equilibrium if and only if $\boldsymbol{\omega}$ satisfies the Kuhn-Tucker conditions of (8.21). The first-order conditions for (8.21) mimic the agent's individual best responses. Given other agents' actions, each agent chooses his/her action in the game as if he/she wants to minimize the potential function. Hence, the first-order solution can either be a saddle point or a minimum for the potential function.

Lemma 2. *The N -player game has a unique Nash equilibrium if $|\lambda_{\min}(\mathbf{A})| < 1/\beta$.*

Proof: Each player is trying to minimize the potential function defined in (8.20), given others actions. There will be a unique Nash equilibrium if and only if the first-order conditions of (8.20) have only one solution. A sufficient condition for the unique solution is the strict convexity of the potential function. The strict convexity of $\Phi(\boldsymbol{\omega})$ can be obtained by taking the second derivative, which yields $(\mathbf{I} + \beta \mathbf{A})$. Hence, the potential function is strictly convex if $\mathbf{I} + \beta \mathbf{A}$ is negative definite. The matrix $(\mathbf{I} + \beta \mathbf{A})$ is negative definite if $|\lambda_{\min}(\mathbf{A})| < 1/\beta$, where, $\lambda_{\min}(A)$ is the smallest eigenvalue of $\mathbf{A}$.

8.3.3 Stability of NE

The equilibrium point's stability refers to the equilibrium point's deviations due to a slight change in the player's actions around the equilibrium point. In other words, an equilibrium point is asymptotically stable if it converges back to the equilibrium point after a slight perturbation around the equilibrium point. Mathematically, let $\mathcal{B}(\mathbf{x}, \epsilon) = \{\mathbf{y} \in \mathbb{R}_+^n : \|\mathbf{y} - \mathbf{x}\| < \epsilon\}$ and $\vartheta(t, \mathbf{y})$, be the value of the unique solution to the system of differential equations at time t . An equilibrium $\mathbf{x}$ is asymptotically stable if $\forall \epsilon > 0, \exists \eta > 0 : \forall \mathbf{y} \in \mathcal{B}(\mathbf{x}, \eta), \forall t \geq 0, \vartheta(t, \mathbf{y}) \in \mathcal{B}(\mathbf{x}, \epsilon)$ and $\exists \epsilon > 0 : \mathbf{y} \in \mathcal{B}(\mathbf{x}, \eta), \lim_{t \rightarrow \infty} \vartheta(t, \mathbf{y}) = \mathbf{x}$.

Proposition 2. *The Nash equilibrium ω will be a stable Nash Equilibrium if and only if $\lambda_{\min}(\mathbf{A}) < 1/\beta$.*

Proof: For the stability of the NE, we will again analyze the second-order condition for minimization of the potential function defined in (8.20) as follows,

ω is a minimum of Φ on $\mathbb{R}_+^n$ iff, $\epsilon^T(\mathbf{I} + \beta\mathbf{A})\epsilon \leq 0, \forall \epsilon \in \mathbb{R}^n$;

ω is a strict minimum of Φ $\mathbb{R}_+^n$ $\epsilon^T(\mathbf{I} + \beta\mathbf{A})\epsilon < 0, \forall \epsilon \in \mathbb{R}^n$.

If ω is the minima, then $(\mathbf{I} + \beta\mathbf{A})$ is negative semi-definite matrix. Similarly, if ω is the strict minima, then $(\mathbf{I} + \beta\mathbf{A})$ is negative definite matrix. On the contrary ω is maxima of Φ , if $(\mathbf{I} + \beta\mathbf{A})$ is a positive definite matrix, which is impossible. Therefore, any NE which is not a minimum of the potential function is a saddle point.

8.3.4 Convergence to NE

Convergence to Nash equilibrium refers to the process by which players in a game adjust their strategies over time in such a way that they eventually reach a Nash equilibrium.

Theorem 1. *The game defined by the cost function of each player in (8.10) converges to the η - pure Nash equilibrium (η - PNE) in $\left(\frac{N}{\eta} \log \frac{\Phi(\omega^0)}{\Phi_{\min}}\right)$ iterations.*

The proof of *Theorem 1* needs some background of the game dynamics and some definitions; hence, we will present the backgrounds here before we move to the final proof.

Definition 2. *Best response dynamic* is a straightforward procedure by which the players adjust their strategies to reach the Nash equilibrium of the game.

Proposition 3. *In a finite potential game, the best response dynamic converges to a pure Nash equilibrium [143].*

Proof: In every iteration of best response dynamics, the deviator's cost strictly decreases, and hence by (8.20) the potential function Φ decreases strictly. Since the game is finite, best response dynamics eventually halt at Nash equilibrium (PNE).

Definition 3. η -*Nash equilibrium* is a state in which no player has more than a η incentive to move. Mathematically, a strategy set $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_N) \in \prod_{i=1}^N \Omega_i$ is an η -Nash equilibrium if for every player i , (8.22) holds,

$$J_i(\omega'_i, \boldsymbol{\omega}_{i-}) \leq (1 - \eta)J_i(\omega_1, \dots, \omega_i, \dots, \omega_N) \quad \forall \omega'_i \in \Omega_i. \quad (8.22)$$

Assumption-2 In a potential game, where the players follows a η -best dynamic strategies response, there is always a player whose utility function has a value at least greater than the potential function i.e., $\max_{i=1}^N J_i(\boldsymbol{\omega}) \geq \frac{\Phi(\boldsymbol{\omega})}{N}$.

Proof of Theorem 1: We consider the players of the game, where each player has a utility function defined in (8.10) follows a η -best response dynamic to reach the Nash equilibrium. We use the potential function relation with the utility function defined in (8.18) to prove the convergence. We know from (8.18),

$$\Phi_i(\omega_i, \boldsymbol{\omega}_{i-}) - \Phi_i(\omega'_i, \boldsymbol{\omega}_{i-}) = J_i(\omega_i, \boldsymbol{\omega}_{i-}) - J_i(\omega'_i, \boldsymbol{\omega}_{i-}), \forall i.$$

Using (8.22) in the above equation,

$$\begin{aligned} \Phi_i(\omega_i, \boldsymbol{\omega}_{i-}) - \Phi_i(\omega'_i, \boldsymbol{\omega}_{i-}) &\geq \eta J_i(\omega'_i, \boldsymbol{\omega}_{i-}) \\ &\geq \frac{\eta}{N} \Phi(\boldsymbol{\omega}) \quad \text{using Assumption ??} \end{aligned}$$

Parameters	Symbol	Quantity	Value
Electrical	V _{dc}	DC Voltage	700 V
	V _{MG}	MG system voltage	380 (phase-phase rms) V
	f	MG nominal frequency	60 Hz
	L _f	filter inductance	3.5 mH
	R _f	filter resistance	0.1 Ω
	C _f	filter capacitance	50 μ F
Lines	R	Line resistance	0.15 Ω
	L	Line inductance	3500 μ H
PI controller	K _{pc}	Proportional current gain	54.97
	K _{ic}	Integral current gain	1.57×10^3
	K _{pv}	Proportional voltage gain	0.2859
	K _{iv}	Integral voltage gain	594.85
Droop	m _p	Frequency droop	0.9×10^{-6}
	n _q	Voltage droop	3×10^{-3}
Game	β	Payoff parameter	0.001

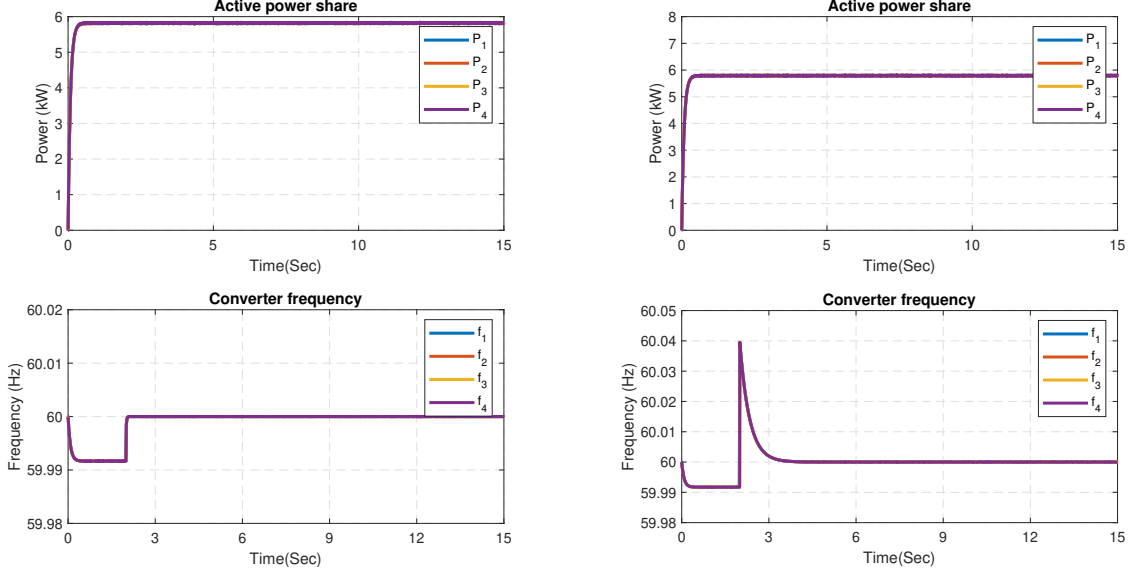
Table 8.1: Parameters of the microgrid test system

The crust here, in every iteration of η -best response dynamics, decreases the potential function by at least by a factor of $(1 - \frac{\eta}{N})$. Thus, every $\frac{N}{\eta}$ iteration decreases the potential function by a constant factor. Considering the initial values of the potential function $\Phi(\omega^0)$ and its value can not drop less than Φ_{min} , η -best response dynamic converges in $\left(\mathcal{O}(\frac{N}{\eta} \log \frac{\Phi(\omega^0)}{\Phi_{min}})\right)$ iterations.

8.4 Simulation Results

The effectiveness of the proposed game theoretic secondary control (GaTSeC) on an islanded four IBR-based microgrid test system is implemented in MATLAB/Simulink. The test system and the corresponding communication graph are shown in Fig. 3.2 (a) and Fig. 3.2 (b), respectively. The DGs are connected through a series of RL branches. The electrical specification of the microgrid is shown in Table 9.1. The results of the proposed distributed game-theoretic controller are compared with the distributed averaging proportional integral (DAPI) controller proposed in [144].

We organize the simulation results into four studies. First, we evaluate the controller



(a) Proposed (GaTSeC) approach

(b) Secondary control using DAPI approach

Figure 8.1: Case-1: Microgrid operation with same droop gain and no load change

performance for general load variations, followed by the performance of the controller to communication link failure. Next, we evaluate the controller performance under different values of the payoff parameter β , and finally, we evaluate the controller adaptability for the plug-and-play operation of the microgrid. It is assumed that DGs can only communicate with their neighboring DGs.

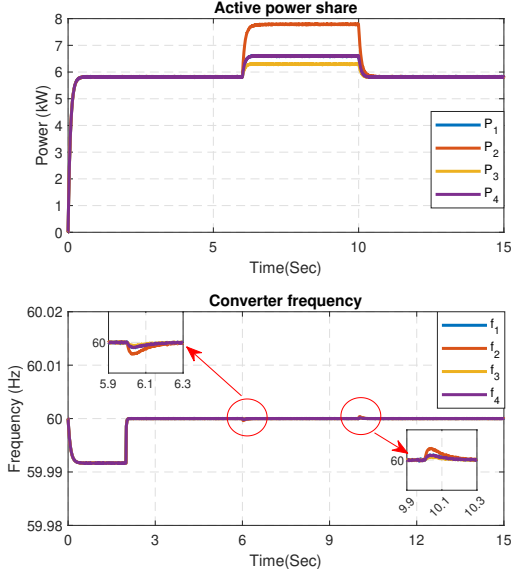
8.4.1 General load variation

In this subsection, we assess the effectiveness of the proposed secondary control across different loading conditions of the DGs while varying the droop gain values. We consider four distinct study scenarios as detailed below. Throughout these cases, the microgrid initially operates solely with the primary droop controller, adhering to (8.2) for frequency control. The secondary controller is activated at the two-second mark, influencing the system's response.

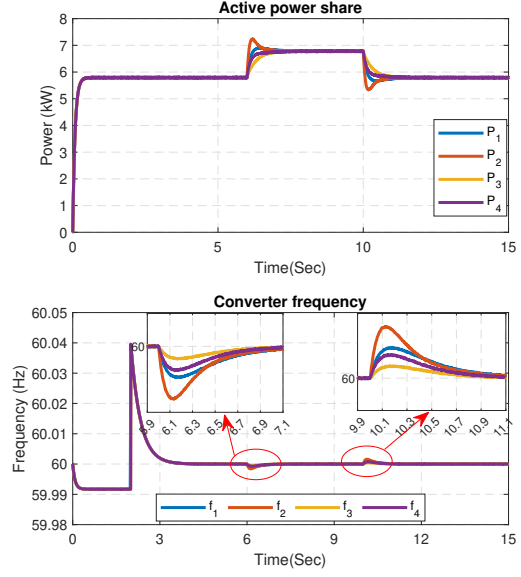
Case-1 (Equal droop gain and no load transients): This case considers an equal amount of load connected at each DG terminal, and all the power converters have equal droop gain as given in Table 9.1. The active and reactive power load connected at each generator terminal is $P_1 = P_2 = P_3 = P_4 = 5.8$ kW and $Q_1 = Q_2 = Q_3 = Q_4 = 3.48$ kVAR, respectively. The

active power share among the DGs and the frequency of each DG unit using the proposed game theoretic secondary controller and the DAPI controller are shown in Fig. 8.1 (a) and Fig. 8.1 (b), respectively. As mentioned earlier, the microgrid initially operates using the local primary droop controller until two seconds, causing a frequency deviation from the nominal value of 60 Hz. At two seconds, the secondary controller switches on, as a result of which the proposed game-theoretic approach swiftly restores converter frequencies to the nominal value, as depicted in Fig.8.1(a). However, there is a minor frequency overshoot that occurs before stabilizing the deviations at 3 seconds by the (DAPI) controller, illustrated in Fig.8.1(b). Examination of Fig.8.1(a) (top) and Fig.8.1(b) (top) reveals that both controllers ensure equal power distribution among DGs, transitioning from zero loads to the full load of 5.8 kW in almost equal time.

Case-2 ((Equal droop gain with load transients): In this case, all the DGs have equal droop gain and are connected with equal loads as in Case-1. However, unlike Case-1, there is a sudden increase in load $P'_2 = 5.8$ kW and $Q'_2 = 4.8$ KVAR at DG 2 terminal at 6 seconds, and the same load is removed at 10 seconds. The response of the proposed GaTSeC and the DAPI controller for active power share among the DGs and the frequency restoration is shown in Fig. 8.2 (a) and Fig. 8.2 (b), respectively. In this case, the microgrid also operates with only the primary controller for two seconds, which is why we see the frequency deviation from the nominal value for two seconds. At two seconds, the secondary control turns on and restores the DG frequencies to the nominal value. As discussed earlier, there is a sudden increase of load at six seconds at the DG2 terminal, as a result of which the frequency of the converters drops below 60 Hz. However, the secondary control immediately acts and restores the frequencies to its nominal values. It can be observed from Fig. 8.2 (a) (bottom) that the frequency restoration only takes 0.3 seconds using the proposed GaTSeC approach. Because the proposed approach has a polynomial degree of convergence as discussed in Theorem 1. On the other hand, the DAPI controller takes more than 1.1 seconds to restore the system frequency, as depicted in Fig. 8.2 (b). This shows that the proposed GTC is almost 70% faster than the DAPI controller in achieving frequency restoration. Similar observations are also noticed when there is a sudden removal of load at the DG2 terminal.

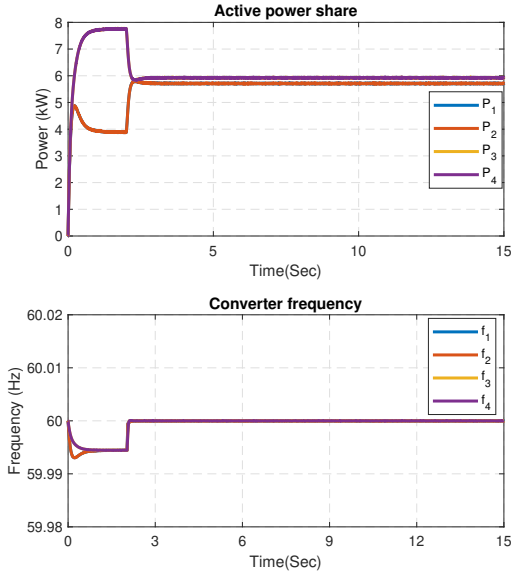


(a) Proposed GaTSeC approach

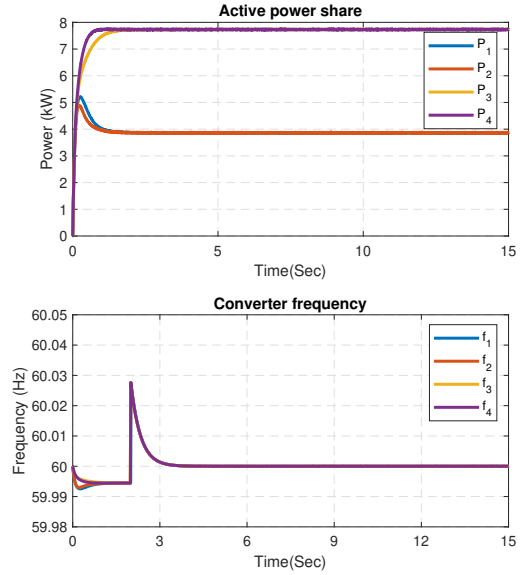


(b) Secondary control using DAPI approach

Figure 8.2: Case-2: Microgrid operation with same droop gain and load transients



(a) Proposed (GaTSeC) approach



(b) Secondary control using DAPI approach

Figure 8.3: Case-3: Microgrid operation with unequal droop gain and no load transients

Case-3 (Unequal droop gain and no load transients): In this case, we observe the proposed game theoretic secondary controller response when the converters have unequal droop gain. This case is also built upon Case-1. However, the converters have different droop gains. Specifically, converter one and converter 2 have droop gain as given in Table 9.1, but converter

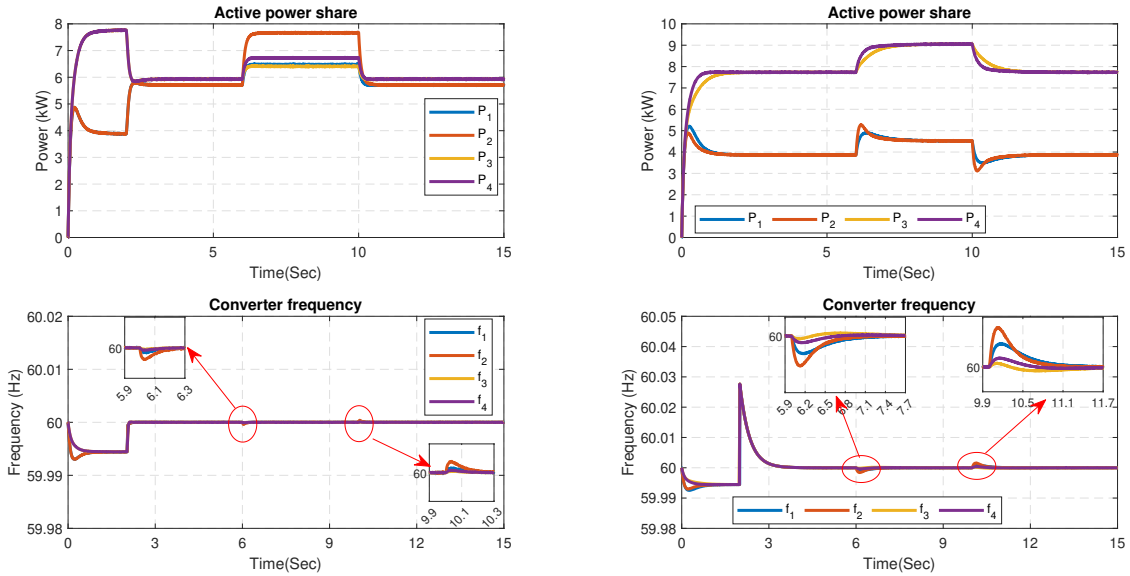
three and converter 4 has a droop gain half of the converter one and converter 2, i.e., $m_{p3} = m_{p4} = m_{p1}/2 = m_{p2}/2$. The response of the proposed secondary GaTSeC and the DAPI controller for power sharing among the DGs and frequency restoration is shown in Fig. 8.3 (a) and Fig. 8.3 (b), respectively. It is clear from Fig. 8.3 (a) and Fig. 8.3 (b) that till two seconds the DGs follow the local primary droop controllers. As a result, there is frequency deviation from the nominal value, and the power share by each DG shares an inverse relationship with their respective droop gains i.e., DG3 and DG4 share twice the power DG1 and DG2 as $m_{p3} = m_{p4} = m_{p1}/2 = m_{p2}/2$. At two seconds, the secondary controller switches on, and it successfully restores the system frequency in both controllers. In the DAPI controller, the power share among the DGs remains the same as the primary controller, as depicted in Fig. 8.3 (b) (top). However, with the proposed secondary GTC, the power share among the DGs follows the same pattern as the primary droop controller, but the difference in power share is almost zero, making the power-sharing among the converters almost equal. This ensures that the converters share almost equal powers irrespective of their droop gain.

Case-4 (Unequal droop gain with load transients): In this study, we investigate the response of the proposed game-theoretic secondary controller to sudden load changes in the presence of converters with unequal droop gain, thereby combining elements from Case-2 and Case-3 scenarios. Similar to the preceding cases, the microgrid operates solely under the primary controller for the initial two seconds, exhibiting a response akin to Case-2. To assess the controller's performance during abrupt load variations, we introduce an additional load of $P'_2 = 5.8$ kW and $Q'_2 = 4.8$ at DG2 terminal, followed by its removal at ten seconds. The response of the proposed GaTSeC and the (DAPI) controller in restoring the system frequency post a sudden load increase is illustrated in Fig. 8.4(a) and Fig. 8.4(b), respectively.

Upon a sudden load increase at DG2 terminal, the power supplied by DGs escalates to meet the increase load demand. Notably, in the proposed game-theoretic control, the DG2 terminal shares the maximum load compared to other DGs, as depicted in Fig 8.4(a) (top). Conversely, the DAPI controller evenly distributes power among DGs based on the droop gain of the converter units, as shown in Fig 8.4(b) (top). Despite the load disturbance, the

proposed game-theoretic control swiftly restores the system frequency to its nominal value within 0.3 seconds, as evidenced in Fig 8.4(a) (bottom). In contrast, the DAPI controller requires approximately 1.7 seconds for frequency restoration, as illustrated in Fig 8.4(b) (bottom). This implies that our proposed controller exhibits a restoration speed that is approximately 80% faster than the DAPI controller.

Furthermore, it is noteworthy that the frequency overshoot is more pronounced in the DAPI controller compared to our proposed controller in response to sudden load increases. Similar trends are observed during load decreases. A crucial finding between Case-2 and Case-4 is that the proposed game-theoretic control consistently restores the system frequency in 0.3 seconds, regardless of the droop gain of the converters. In contrast, the frequency restoration time increases from 1.1 seconds in the case of equal droop gain to 1.7 seconds when converters have unequal droop gain. This underscores the robustness and efficiency of the proposed game-theoretic control in managing system dynamics under varying conditions.



(a) Proposed (GaTSeC) approach

(b) Secondary control using DAPI approach

Figure 8.4: Case-4: Microgrid operation with unequal droop gain and load transients

8.4.2 Effect of payoff parameter β on Nash equilibrium

In this subsection, we will discuss the effect of β on the Nash equilibrium of the proposed game theoretic distributed secondary control of MG. As discussed in 8.3.2, the communication infrastructure and the value of β decide the uniqueness of the NE. As long as $|\lambda_{\min}(\mathbf{A})| < 1/\beta$, the game has a unique and stable NE. The Nash equilibrium of the game with two different values of β for power sharing and frequency of the converters is shown in Fig. 8.5 (a) and Fig. 8.5 (b), respectively. To demonstrate the effect of β on the NE, we consider the loading on the converters as discussed in 8.4.1, Case-1. The MG was operating steadily in NE with $\beta = 0.001$. At 0.8 sec, the value β increases from 0.001 to 0.01. As a result, *Proposition 2* fails. This leads to two Nash equilibria for the game. Hence, converters operate at two different frequencies in two groups. The power share among the converters became imbalanced, leading to the failure of the NE.

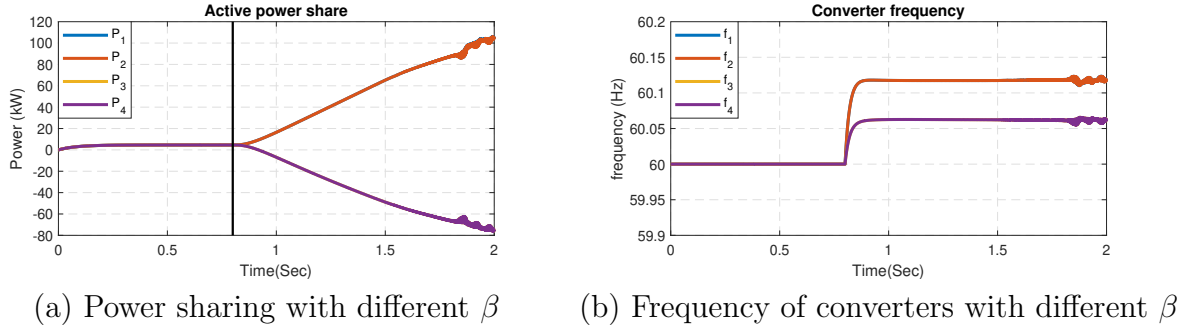


Figure 8.5: Nash equilibrium with different β values

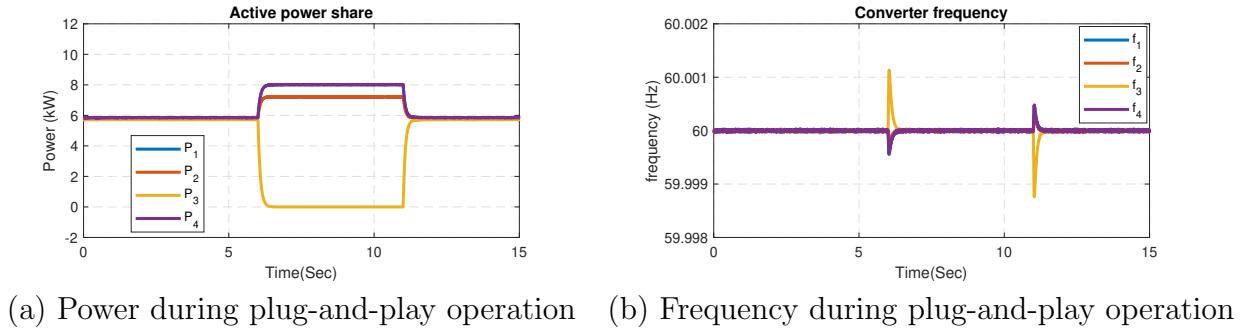


Figure 8.6: Performance of the Game-theoretic controller during plug-and-play operation

8.4.3 Plug and Play Operation

In this specific section, we explore the adept plug-and-play performance of the proposed game theoretic controller. The controller's efficacy during the removal and subsequent re-connection of DG3 from the Microgrid (MG) at the sixth and eleventh seconds, respectively, is visually represented in Fig. 8.6. Initially, the MG operates in a stable state under the conditions outlined in 8.4.1, Case-1. The uniform droop gain across all converters ensures an equitable power distribution among them. At the sixth second, DG3 is detached from the MG, introducing transients within the system. Consequently, the power contribution of DG3 reduces to zero, causing an augmented load distribution among the remaining DGs, as depicted in Fig. 8.6(a). A parallel response is observed when DG3 is reintegrated into the MG at the eleventh second, restoring an equal power distribution among all DGs. Throughout these dynamic adjustments, the removal and re-connection of DG3 induce frequency deviations from the nominal value. Nonetheless, the game theoretic controller adeptly rectifies these deviations, swiftly restoring the system frequency. This underscores the demonstrated proficiency of the proposed controller in facilitating seamless plug-and-play operations within the MG.

8.5 Summary

This chapter paper proposes an innovative solution to address distributed secondary control in microgrids (MG) for frequency restoration. The approach involves formulating the secondary frequency control problem as a non-cooperative game within a multiagent system framework. The paper derives both necessary and sufficient conditions for achieving the NE. By conducting extensive simulations on a representative MG model under various loading conditions, the proposed approach is validated and demonstrates its superiority over the DAPI method.

Chapter 9

RL for Secondary Control of Microgrids

This chapter introduces the RL-based microgrid secondary control followed by an adversarial robust formulation of the RL agents.

9.1 Introduction

Over the recent years, microgrids (MGs) have emerged as versatile and adaptive solutions for addressing various challenges the energy sector is facing, starting from enhancing grid resiliency and sustainability to promoting energy access and innovation. MGs are localized, autonomous energy systems integrating various distributed energy resources (DERs), energy storage devices, loads, and other power electronic equipment [89, 90]. MGs can operate with the main grid or independently (island mode of operation) to enhance the overall grid resilience and efficiency [90, 91]. However, ensuring frequency and voltage stability is crucial in MGs. In the conventional power grid, voltage and frequency stability is achieved by synchronous generators because of their inbuilt inertia [3]. On the other hand, most MGs are sourced by solid-state power electronics devices that lack inertia. Droop control has emerged as a dominant solution to regulate voltage and frequency in MGs because of its

simple implementation [92] and reliance on local measurements without requiring any communication infrastructure [93]. However, due to its inherent characteristics, droop control may not be precise and can cause voltage and frequency deviations from their nominal values [4]. Therefore, a secondary level of hierarchical control becomes necessary [94] and serves as a compensation mechanism. Based on the communication requirements, MG secondary control is broadly classified into three types [95]: (1) Centralized control: This enables easier decision-making and coordination [4]. However, centralized control suffers from large communication bandwidth requirements, which leads to decision delays and low reliability; (2) Decentralized secondary control: This approach does not require exclusive communication infrastructure (CI) for control. Each DER restores its frequency and voltage amplitudes to nominal values individually. However, the CI is required for transferring data from the tertiary control and MG centralized control layers [96]. (3) Distributed secondary control: This uses a low bandwidth sparse communication network in which the inverters are only required to interact with the neighboring inverters.

9.2 Background and Problem Description

In this section, we briefly discuss some useful background related to microgrids that will drive our follow-up secondary control design discussions.

9.2.1 Large Signal Model of Inverters

AC microgrids consist of a number of inverter-based DERs. Every inverter-based resource can be represented by DC sources behind a DC/AC inverter and an LC filter connected to the point of common coupling (PCC) through an RL output connector [97]. The inverters normally operate on voltage control mode during the islanded operation of MG [98]. A typical DG model with the inverter control loops is shown in Fig. 3.1. The major control loops of the inverter control are the current control loop, the voltage control loop, and the power control loop, which employs the droop controller. In this subsection, we briefly

summarize three control loops.

The DG dynamics are expressed with respect to their own d-q (direct-quadrature) reference frame. Hence, the reference frame of each DG is derived from the common reference frame is given as,

$$\dot{\delta}_i = \omega_i - \omega_{com} \quad (9.1)$$

where, δ_i is the angle of inverter i , ω_i is the angular frequency of inverter i and ω_{com} is the angular frequency of common reference frame. The power control loop is responsible for generating the reference voltage signals V_i for the voltage controller loop and the operating frequency ω_i for the inverter by employing the primary droop equation given as,

$$\omega_i = \omega_{ni} - m_{pi}P_i, \quad (9.2a)$$

$$V_i = V_{ni} - n_{qi}Q_i. \quad (9.2b)$$

The power controller block has two low pass filters with cutoff frequency ω_c to extract the fundamental components of active and reactive powers P_i and Q_i , respectively. The first-order differential equation of the filter is given as,

$$\dot{P}_i = -\omega_{ci}P_i + \omega_{ci}(v_{odi}i_{odi} + v_{oqi}i_{oqi}), \quad (9.3a)$$

$$\dot{Q}_i = -\omega_{ci}Q_i + \omega_{ci}(v_{oqi}i_{odi} - v_{odi}i_{oqi}) \quad (9.3b)$$

where, v_{odi} and v_{oqi} are the direct and quadrature axis inverter output voltage, respectively. Similarly, i_{odi} and i_{oqi} represent the direct and quadrature axis inverter output currents, respectively.

The voltage controller is responsible for generating the required reference current values for the current controller [145]. The differential-algebraic equations of the voltage control

loop are given by,

$$\dot{\phi}_{di} = v_{odi}^* - v_{odi}, \quad (9.4a)$$

$$\dot{\phi}_{qi} = v_{oqi}^* - v_{oqi}, \quad (9.4b)$$

$$i_{ldi}^* = F_i i_{odi} - \omega_b C_{fi} v_{oqi} + K_{PVi} (v_{odi}^* - v_{odi}) + K_{IVi} \phi_{di}, \quad (9.4c)$$

$$i_{lqi}^* = F_i i_{oqi} + \omega_b C_{fi} v_{odi} + K_{PVi} (v_{oqi}^* - v_{oqi}) + K_{IVi} \phi_{qi} \quad (9.4d)$$

where, K_{PVi} and K_{IVi} are the PI-controller gains, ω_b is the nominal frequency, $F_i > 0$ is a positive constant, and C_{fi} is the filter capacitance. v_{odi}^*, v_{oqi}^* are the direct and quadrature components of reference voltages, respectively. Similarly, i_{ldi}^*, i_{lqi}^* are the direct and quadrature components of the input reference currents, respectively. ϕ_{di} and ϕ_{qi} are the auxiliary state variables introduced for the PI controller.

Similarly, the current controller produces the required reference voltage signal for the inverter input voltages [145]. The differential-algebraic equations of the current controller correspond to,

$$\dot{\psi}_{di} = i_{ldi}^* - i_{ldi} \quad (9.5a)$$

$$\dot{\psi}_{qi} = i_{lqi}^* - i_{lqi} \quad (9.5b)$$

$$v_{idi}^* = -\omega_b L_{fi} i_{lqi} + K_{PCi} (i_{ldi}^* - i_{ldi}) + K_{ICi} \psi_{di} \quad (9.5c)$$

$$v_{iqi}^* = \omega_b L_{fi} i_{ldi} + K_{PCi} (i_{lqi}^* - i_{lqi}) + K_{ICi} \psi_{qi} \quad (9.5d)$$

Here, K_{PCi} and K_{ICi} are the PI-controller gains, ω_b is the nominal frequency, and L_{fi} is the filter inductance. v_{idi}^*, v_{iqi}^* are the direct and quadrature components for input reference voltages, respectively. i_{ldi}, i_{lqi} are the direct and quadrature components of the actual inverter output currents, respectively. ψ_{di} and ψ_{qi} are the auxiliary state variables introduced for the PI controller.

The differential equations of the output LC filter correspond to,

$$\dot{i}_{ldi} = -\frac{R_{fi}}{L_{fi}}i_{ldi} + \omega_i i_{lqi} + \frac{1}{L_{fi}}v_{idi} - \frac{1}{L_{fi}}v_{odi}, \quad (9.6a)$$

$$\dot{i}_{lqi} = -\frac{R_{fi}}{L_{fi}}i_{lqi} - \omega_i i_{ldi} + \frac{1}{L_{fi}}v_{iqi} - \frac{1}{L_{fi}}v_{oqi}, \quad (9.6b)$$

$$\dot{v}_{odi} = \omega_i v_{oqi} + \frac{1}{C_{fi}}i_{ldi} - \frac{1}{C_{fi}}i_{odi}, \quad (9.6c)$$

$$\dot{v}_{oqi} = -\omega_i v_{odi} + \frac{1}{C_{fi}}i_{lqi} - \frac{1}{C_{fi}}i_{oqi}. \quad (9.6d)$$

Here, R_{fi} is the filter resistance and ω_i is the inverter angular frequency. v_{od}, v_{oq} are direct and quadrature components of the output voltages, respectively.

Finally, the differential equations of the output RC connector are given as,

$$\dot{i}_{odi} = -\frac{R_{ci}}{L_{ci}}i_{odi} + \omega_i i_{oqi} + \frac{1}{L_{ci}}v_{odi} - \frac{1}{L_{ci}}v_{bdi}, \quad (9.7a)$$

$$\dot{i}_{oqi} = -\frac{R_{ci}}{L_{ci}}i_{oqi} - \omega_i i_{odi} + \frac{1}{L_{ci}}v_{oqi} - \frac{1}{L_{ci}}v_{bqi}. \quad (9.7b)$$

Here, R_{ci}, L_{ci} are the output RC connector resistance and inductance, respectively. v_{bdi} and v_{bqi} are the direct and quadrature components of the bus voltage, respectively.

Let the state vector of an inverter i be represented via χ , where

$$\chi = [\delta_i, P_i, Q_i, \phi_{di}, \phi_{qi}, \psi_{di}, \psi_{qi}, i_{ldi}, i_{lqi}, v_{odi}, v_{oqi}, i_{odi}, i_{oqi}]^T$$

The large signal non-linear dynamical model of an inverter i is given by,

$$\begin{cases} \dot{\chi}_i = \mathbf{f}_i(\chi_i) + \mathbf{g}_i(\chi_i)\boldsymbol{\eta}_i + \mathbf{h}_i(\chi_i)\mathbf{u}_i \\ \mathbf{y}_i = \mathbf{o}(\chi_i). \end{cases} \quad (9.8)$$

where, $\mathbf{f}_i(\mathbf{x}_i)$, $\mathbf{g}_i(\mathbf{x}_i)$ and $\mathbf{h}_i(\mathbf{x}_i)$ are derived from (9.3) to (9.7), and $\boldsymbol{\eta}_i = [\omega_{comm}, v_{bdi}, v_{bqi}]$ is the known disturbance.

9.2.2 Communication network modeling

We consider the microgrid to have a communication network over which the inverters exchange information. This network is represented by a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where, the node-set and the edges are represented by $\mathcal{V} = \{v_1, v_2 \dots v_N\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, respectively. Assume that the network has finite set vertices i.e., $|\mathcal{V}| = N$ and $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is the graph adjacent matrix. The IBRs are considered the vertices of the communication graph, and the edges of the communication network form the links. The edges from node v_j to v_i is denoted by (v_j, v_i) , i.e., node v_j receives information from v_i . The weight of the communication link is given by $a_{ij} > 0$ if $(v_j, v_i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The in-degree matrix is defined by $\mathbf{D} = \text{diag}(d_i) \in \mathbb{R}^{N \times N}$ with $d_i = \sum_{j=1}^N a_{ij}$. Finally, the graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The Laplacian matrix $\mathbf{L}$ has all row sums equal to zero, i.e., $\mathbf{L}\mathbf{1}_N = 0$, where $\mathbf{1}_N$ is a column of all ones. The graph $\mathcal{G}$ is called balanced if its Laplacian matrix $\mathbf{L}$ meets $\mathbf{1}_N^T \mathbf{L} = 0$. A graph is said to have a spanning tree if there exists a node (called the root) that has a directed path from this node to other nodes in the graph [101]. The graph is said to be connected if the algebraic connectivity, i.e., the second smallest eigenvalue, is non-zero.

9.2.3 Problem Description

The droop controller described in (9.2a) and (9.2b) leads to steady-state error because of load imbalance or unequal power changes. Hence, it needs secondary compensation to restore the system voltages and frequencies to their nominal steady-state values by adding correction terms. These correction terms are generated by the secondary controller and correspond to,

$$\omega_i = \omega_{ni} - m_{pi}P_i + \delta\omega_i, \quad (9.9a)$$

$$V_i = V_{ni} - n_{qi}Q_i + \delta V_i \quad (9.9b)$$

Here, $\delta\omega_i$ and δV_i are the correction terms generated by the secondary controller. The goal of the secondary controller is to design $\delta\omega_i$ and δV_i such that the following holds true [146]

$$\lim_{t \rightarrow \infty} (\omega_i(t) - \omega_j(t)) = 0 \text{ and } \omega_i(t) = \omega^*, \forall i, j \in \mathcal{V}, \quad (9.10a)$$

$$\lim_{t \rightarrow \infty} (V_i(t) - V_j(t)) = 0 \text{ and } V_i(t) = V^*, \forall i, j \in \mathcal{V}, \quad (9.10b)$$

9.3 Proposed Approach

For designing the secondary control, we consider the microgrid to be represented by a graph described in Section 9.2.2. A representative microgrid and the associated communication network are given in Fig. 3.2a and Fig. 3.2b, respectively. For designing the secondary control, we consider the controller only having information about the inverter angles δ , angular frequencies ω , and the terminal voltages v . Considering only three variables, we find the discrete-time dynamics of the MG as,

$$\mathbf{x}_1(n+1) = f_1(\mathbf{x}_1(n)), \quad (9.11a)$$

$$\mathbf{x}_2(n+1) = f_2(\mathbf{x}_2(n)) + \mathbf{u}_\omega(n), \quad (9.11b)$$

$$\mathbf{x}_3(n+1) = f_3(\mathbf{x}_3(n)) + \mathbf{u}_v(n), \quad (9.11c)$$

where,

$$\mathbf{x}_1(n) = [\delta_1(n), \delta_2(n), \dots, \delta_N(n)]^T \in \mathbb{T}^N, \quad (9.12a)$$

$$\mathbf{x}_2(n) = [\omega_1(n), \omega_2(n), \dots, \omega_N(n)]^T \in \mathbb{R}^N, \quad (9.12b)$$

$$\mathbf{x}_3(n) = [V_1(n), V_2(n), \dots, V_N(n)]^T \in \mathbb{R}^N, \quad (9.12c)$$

$$\mathbf{u}_\omega(n) = [u_{\omega 1}(n), u_{\omega 2}(n), \dots, u_{\omega N}(n)]^T \in \mathbb{R}^N, \quad (9.12d)$$

$$\mathbf{u}_v(n) = [u_{v 1}(n), u_{v 2}(n), \dots, u_{v N}(n)]^T \in \mathbb{R}^N, \quad (9.12e)$$

$$\mathbf{x} = [\mathbf{x}_1^T(n), \mathbf{x}_2^T(n), \mathbf{x}_3^T(n)]^T \in \mathbb{R}^{3N}. \quad (9.12f)$$

Here, $\mathbf{u}_\omega$ and $\mathbf{u}_v$ are the secondary controllers for frequency and voltage restoration to be designed in the next subsection. Before we design the controller $\mathbf{u}_\omega$ and $\mathbf{u}_v$, we would like to briefly describe the conditions of Lyapunov stability criteria for the discrete system in (9.11).

Definition-1: Consider a nonlinear dynamical system described as, $\mathbf{x}(n+1) = f(\mathbf{x}(n), n) + \mathbf{b}(n)$ with $\mathbf{b}(n)$ being an unknown bounded disturbance. The system is said to be ultimately uniformly bounded [147], and the norm of the state is bounded within a neighborhood of $\epsilon > 0$, if there exists a function $\vartheta(x(n), n)$ with partial differences $\dot{\vartheta}(x(n), n)$ and $\mathbf{x}$ is in a compact set $\mathbf{S} \in \mathbb{R}^N$ such that,

1. $\vartheta(x(n), n)$ is positive definite i.e., $\vartheta(x(n), n) > 0$ and
2. $\dot{\vartheta}(x(n), n) < 0$ for $\|x\| > \epsilon$

9.3.1 Error dynamic evaluation

Let us consider a microgrid defined by the non-linear system dynamics in (9.11). We assume $f_2(\mathbf{x})$ and $f_3(\mathbf{x})$ are unknown functions and need to design $\mathbf{u}_\omega$ and $\mathbf{u}_v$ for frequency and voltage deviations resulting after the primary droop control. Consider the vector of nominal system frequency $\mathbf{x}_2^* = \mathbf{1}_N \omega^* \in \mathbb{R}^N$ and vector of reference value of the voltages at PCC is $\mathbf{x}_3^{ref} = \mathbf{1}_N v_{ref} \in \mathbb{R}^N$. The error in frequency deviation from the nominal frequency value and error in voltage deviation from the reference value is E_ω and E_v , respectively, and are defined as,

$$E_\omega(n) = \mathbf{x}_2^* - \mathbf{x}_2(n) \tag{9.13a}$$

$$E_v(n) = \mathbf{x}_3^{ref} - \mathbf{x}_3(n) \tag{9.13b}$$

This results in the error dynamics,

$$E_\omega(n+1) = \mathbf{x}_2^* - \mathbf{x}_2(n+1) \quad (9.14a)$$

$$E_v(n+1) = \mathbf{x}_3^{ref} - \mathbf{x}_3(n+1) \quad (9.14b)$$

The goal is to design $\mathbf{u}_\omega$ and $\mathbf{u}_v$, such that the error dynamics defined in (9.14) stabilize for both frequency and voltage. From now on, we focus on the derivation of stabilizing the error dynamic in voltage. A similar argument can be obtained for the frequency error stability.

Using (9.11c) in (9.14b), the error dynamics in voltage can be written as,

$$E_v(n+1) = \mathbf{x}_3^{ref} - [f_3(\mathbf{x}(\mathbf{n})) + \mathbf{u}_v(n)] \quad (9.15)$$

We define the candidate Lyapunov function as,

$$\mathcal{V}_{E_v} = E_v^T(n)E_v(n) \quad (9.16)$$

The difference of (9.16) gives,

$$\Delta \mathcal{V}_{E_v} = E_v^T(n+1)E_v(n+1) - E_v^T(n)E_v(n) \quad (9.17)$$

Using (9.15) in (9.17) gives,

$$\begin{aligned} \Delta \mathcal{V}_{E_v} &= [\mathbf{x}_3^{ref} - [f_3(\mathbf{x}(\mathbf{n})) + \mathbf{u}_v(n)]]^T \\ &\quad [\mathbf{x}_3^{ref} - [f_3(\mathbf{x}(\mathbf{n})) + \mathbf{u}_v(n)]] - E_v^T(n)E_v(n) \end{aligned} \quad (9.18)$$

From *Definition-1*, in order to stabilize E_v , $\Delta \mathcal{V}_{E_v} < 0$. Hence, we select the control input as,

$$\mathbf{u}_v(n) = \mathbf{x}_3^{ref} - f_3(\mathbf{x}(\mathbf{n})) + \mathbf{\Lambda}E_v(n), \quad (9.19)$$

where, $\mathbf{\Lambda}$ is a diagonal positive definite matrix. Substituting (9.19) in (9.18) gives

$$\Delta\mathcal{V}_{E_v} = \sum_{i=1}^N (\lambda_i^2 - 1) E_{vi}^2, \quad (9.20)$$

where, λ_i is the i^{th} eigenvalue of the matrix $\mathbf{\Lambda}$ and E_{vi} is the i^{th} element of the error vector E_v . Hence, the error E_v is asymptotically stable if $\Delta\mathcal{V}_{E_v} < 0$ or in other words $0 \leq \lambda_{max} < 1$. Here λ_{max} is the maximum eigenvalue of the positive definite matrix $\mathbf{\Lambda}$.

The above discussion holds if the $f_3(\mathbf{x})$ is known. However, in practice, the $f_3(\mathbf{x})$ is unknown because of wrong documentation of the network parameters, network aging, or network topology changes. Hence, we estimate the approximate value of $f_3(\mathbf{x})$ using the actor-network of reinforcement learning in the next section of the article. Let's assume the approximate value of $f_3(\mathbf{x}) \approx h(\mathbf{x})$. With the approximation, we can design the control input $\mathbf{u}_v$ as,

$$\mathbf{u}_v(n) = \mathbf{x}_3^{ref} - h(\mathbf{x}(n)) + \mathbf{\Lambda}E_v(n), \quad (9.21)$$

Hence, (9.18) can be written as,

$$\begin{aligned} \Delta\mathcal{V}_{E_v} &= [\hat{h}(\mathbf{x}(x)) + \mathbf{\Lambda}E_v(n)]^T \\ &[\hat{h}(\mathbf{x}(x)) + \mathbf{\Lambda}E_v(n)] - E_v^T(n)E_v(n), \end{aligned} \quad (9.22)$$

where, $\hat{h}(\mathbf{x}) = f_3(\mathbf{x}) - h(\mathbf{x})$ is the function estimation error. Therefore, from (9.22), $\Delta\mathcal{V}_{E_v}$ has a negative value i.e., $\Delta\mathcal{V}_{E_v} < 0$, if

$$\|\hat{h}(\mathbf{x}(x)) + \mathbf{\Lambda}E_v(n)\| < \|E_v(n)\|. \quad (9.23)$$

Let's assume the maximum estimation error is h_{max} , i.e., $\|\hat{h}(\mathbf{x})\| < h_{max}$, then (9.23) can be

written as,

$$\begin{aligned}
\|\hat{h}(\mathbf{x}(x)) + \Lambda E_v(n)\| &\leq \|\hat{h}(\mathbf{x}(x))\| + \|\Lambda E_v(n)\| \\
&\leq h_{max} + \lambda_{max} \|E_v(n)\| \\
&< E_v(n)
\end{aligned} \tag{9.24}$$

solving further (9.24) results

$$E_v^{max} < \|E_v(n)\|, \tag{9.25}$$

where, $E_v^{max} = \frac{h_{max}}{1-\lambda_{max}}$. This follows that $\Delta \mathcal{V}_{E_v} < 0$ if $E_v^{max} < \|E_v(n)\|$. In other words, the error dynamic stabilizes asymptotically for a compact set $\mathcal{S} = \{E_v^{max} < \|E_v(n)\|\}$.

9.3.2 RL for Function Approximation

Actor-critic reinforcement Learning (RL) represents a pivotal advancement in function approximation within decision-making systems. This paradigmatic approach amalgamates the best of both worlds, leveraging the strengths of both policy-based (Actor) and value-based (Critic) methods to enhance learning efficiency and stability. At its core, Actor-Critic RL revolves around two key components: the actor, responsible for selecting actions based on a policy, and the critic, tasked with evaluating the chosen actions. This duality enables the system to simultaneously learn a policy and approximate the value function, offering a more robust framework for navigating complex environments. Through function approximation, Actor-Critic RL transcends the limitations of tabular methods by generalizing across states, enabling scalability to high-dimensional state spaces. By utilizing function approximation techniques such as neural networks, the Actor-Critic architecture can effectively capture the underlying dynamics of the environment, facilitating adaptive and flexible decision-making. Moreover, in actor-critic method of RL, the Actor uses estimates of the discounted sum of the reward signals, subtracting the Critic output from the actor output lowers the variance and speed up the learning process. Hence, we adopt the Actor-critic RL for the function approximation discussed in 9.3.

We propose to design the secondary controller problem in the MG as a multi-agent

reinforcement learning (MARL) problem expressed as a Markov decision process (MDP). The MDP is described by the tuple $(N, \mathcal{S}, \mathcal{A}_i, \mathcal{P}, \mathcal{R}_i, \gamma)$, where,

1. N is the number of agents (i.e., the number of inverters in the microgrid)
2. $\mathcal{S}$ is the state space of the system, $\mathbf{x}(n) \in \mathcal{S}$ is the state of the system at time instant n . For designing the secondary controller, we only consider the states described in (9.12f) rather than all states of the MG as defined in (9.8).
3. $\mathcal{A}_i$ is the action set of each agent. $a_i(n) \in \mathcal{A}_i$ is the action of agent a_i at time instant n . We define the joint action of all the agents at time n as $\mathbf{a}(n) = [a_1(n), a_2(n), \dots, a_N(n)] \in \mathcal{A}$, where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_N$. We consider the actions taken by each agent are the correction terms $\delta\omega_i$ and δV_i used for frequency and voltage restoration, respectively (as defined in (9.9)).
4. $\mathcal{P}(\mathbf{x}(\mathbf{n} + \mathbf{1})|\mathbf{x}(\mathbf{n}), a(n))$ is the transition probability of states.
5. $\mathcal{R}_i$ is the reward function of each agent as defined in (9.27).
6. $\gamma \in [0, 1)$ is the future discount factor.

We propose the use of neural networks for both actor and critic networks. We measure the long-term performance of the cost function $\mathcal{J}(n) \in \mathbb{R}^N$ using the instantaneous costs as [148],

$$\begin{aligned} \mathcal{J}(n) &= \sum_{m=n}^{\infty} \gamma^{m-n} r(m+1) \\ &= r(n+1) + \gamma r(n+2) + \gamma^2 r(n+3) + \dots \end{aligned} \tag{9.26}$$

Here, $0 < \gamma < 1$ is the discount factor and $r(n) = [r_1(n), r_2(n), \dots, r_N(n)] \in \mathbb{R}^N$ is the instantaneous cost vector defined as,

$$r_i(n) = \begin{cases} 0 & \text{if } |E_v(n)| \leq \xi \\ 1 & \text{if } |E_v(n)| > \xi \end{cases} \tag{9.27}$$

Here, $i \in \mathcal{V}$ and $\xi \in \mathbb{R}^+$ is a fixed threshold. $r_i(n)$ measures the current performance of the i^{th} DG. $r_i(n) = 0$ signifies the successful restoration of the voltage at the i^{th} DG and $r_i(n) = 1$ signifies unsuccessful voltage restoration. In a microgrid, the functions $f_1(\mathbf{x}(n)), f_2(\mathbf{x}(n)), f_3(\mathbf{x}(n))$ are unknown; hence, we define a critic network to learn the cost function $\mathcal{J}(n)$ in real-time as discussed in the next subsection.

9.3.3 Critic Network Architecture

The output of the critic network $\hat{\mathcal{J}}(n) \in \mathbb{R}^N$ learns to approximate the cost function $\mathcal{J}(n) \in \mathbb{R}^N$. With $\mathbf{x}_3(n)$ as the input of the critic network, the output $\hat{\mathcal{J}}(n) \in \mathbb{R}^N$ is defined as,

$$\hat{\mathcal{J}}(n) = \mathbf{W}_{ho}^{cT}(n) \eta^c(\mathbf{W}_{ih}^{cT}(\mathbf{x}_3)) \quad (9.28)$$

Here, $\mathbf{W}_{ih}^{cT} \in \mathbb{R}^{n_1 \times N}$ is the weight matrix between the input and hidden layers and the hidden layer has n_1 number of layers. Similarly, the weight matrix between the hidden and output layer is represented by $\mathbf{W}_{ho}^{cT}(n) \in \mathbb{R}^{N \times n_1}$. The function $\eta^c(\mathbf{W}_{ih}^{cT}(\mathbf{x}_3)) \in \mathbb{R}^{n_1}$ is the basis function. We will use the short hand notation η^c for $\eta^c(\mathbf{W}_{th}^{cT}(\mathbf{x}_3))$ for further discussion.

We define the temporal difference error of the critic network as,

$$\begin{aligned} E^c(n) &= r(n) + \gamma \hat{\mathcal{J}}(n) - \hat{\mathcal{J}}(n-1) \\ &= r(n) + \gamma \mathbf{W}_{ho}^{cT}(n) \eta^c(n) - \mathbf{W}_{ho}^{cT}(n-1) \eta^c(n-1) \end{aligned} \quad (9.29)$$

Now, we want minimize the objective function of the critic network $\mathcal{J}^c$ defined as,

$$\mathcal{J}^c(n) = \frac{1}{2} E^{cT}(n) E^c(n) \quad (9.30)$$

For low computational complexity, we fix the weights of the input to hidden layer matrix $\mathbf{W}_{ih}^{cT}$ to unity. Hence, only update the weights of the matrix connecting the hidden layers to the output layers $\mathbf{W}_{ho}^{cT}(n)$. This assumption can be relaxed if needed. Taking the gradient

of (9.30) with respect to $\mathbf{W}_{ho}^c{}^T(n)$ gives,

$$\frac{\partial \mathcal{J}^c(n)}{\partial \mathbf{W}_{ho}^c(n)} = \frac{\partial \mathcal{J}^c(n)}{\partial E^c(n)} \frac{\partial E^c(n)}{\partial \hat{\mathcal{J}}(n)} \frac{\partial \hat{\mathcal{J}}(n)}{\partial \mathbf{W}_{ho}^c(n)} \quad (9.31)$$

Using, (9.28), (9.29) and (9.30) in (9.31) we get,

$$\frac{\partial \mathcal{J}^c(n)}{\partial \mathbf{W}_{ho}^c(n)} = \gamma \eta^c(n) E^{cT}(n) \quad (9.32)$$

Now, we update the weight matrix with a learning rate $\alpha^c \in \mathbb{R}^+$ as,

$$\begin{aligned} \mathbf{W}_{ho}^c(n+1) &= \mathbf{W}_{ho}^c(n) - \alpha^c \gamma E^{cT}(n) \eta^c(n) \\ &= \mathbf{W}_{ho}^c(n) - \alpha^c \gamma \eta^c(n) \times \\ &\quad (r(n) + \gamma \mathbf{W}_{ho}^c{}^T(n) \eta^c(n) - \mathbf{W}_{ho}^c{}^T(n-1) \eta^c(n-1))^T \end{aligned} \quad (9.33)$$

In the next subsection, we will discuss the architecture of the actor network to minimize both the function approximation error $\hat{h}(\mathbf{x})$ and the cost function $\mathcal{J}(n)$.

9.3.4 Actor Network Architecture

The goal of the actor-network is to find the function approximation $h(\mathbf{x})$ of the true unknown function $f_3(\mathbf{x})$. Once $h(\mathbf{x})$ is calculated, the next task is to plug it into the control policy defined in (9.21). We define the function estimation as a single-layer neural network similar to the critic network parameterized by,

$$h(\mathbf{x}) = \mathbf{W}_{ho}^a{}^T(n) \eta^a(\mathbf{W}_{ih}^a{}^T(n) \mathbf{x}_3(n)) \quad (9.34)$$

where, $\mathbf{W}_{ho}^a{}^T \in \mathbb{R}^{N \times n_2}$ is the matrix connecting the hidden layers to the output layer, $\mathbf{W}_{ih}^a{}^T \in \mathbb{R}^{n_2 \times N}$ is the weight matrix connecting the input layer to the hidden layer, and $\eta^a(\mathbf{W}_{ih}^a{}^T \mathbf{x}_3(n))$ is the basis function. Similar to the critic network, we fix the weights of $\mathbf{W}_{ih}^a$ to an identity matrix and we only update the weights of $\mathbf{W}_{ho}^a$ using gradient descent

algorithm. Again, this assumption can be relaxed.

As per the discussion in 9.3.1, the function estimation error $\hat{h}(\mathbf{x}) \in \mathbb{R}^N$ is given as,

$$\hat{h}(n) = f_3(n) - h(n). \quad (9.35)$$

Let the desired cost function is $\mathcal{J}^*(n) \in \mathbb{R}^N$ and its approximation is the output of the critic network $\hat{\mathcal{J}}(n)$ defined in (9.28). The error between these two functions is defined as,

$$\tilde{\mathcal{J}}(n) = \hat{\mathcal{J}}(n) - \mathcal{J}^*(n) \quad (9.36)$$

We train the actor network using $\hat{h}(\mathbf{x})$ and $\tilde{\mathcal{J}}(n)$ by defining a prediction error of the actor network $E^a(n) \in \mathbb{R}^N$ as

$$E^a(n) = \hat{h}(n) + \tilde{\mathcal{J}}(n) \quad (9.37)$$

As per (9.26) and (9.27) at the steady state $\mathcal{J}^*(n) = 0$. Hence, we write (9.37) as,

$$E^a(n) = \hat{h}(n) + \hat{\mathcal{J}}(n) \quad (9.38)$$

Now, we consider the objective function for the actor network to minimize as,

$$\mathcal{J}^a(n) = \frac{1}{2} E^{aT}(n) E^a(n) \quad (9.39)$$

We calculate the gradient of (9.39) with respect to $\mathbf{W}_{ho}^a$ as,

$$\frac{\partial \mathcal{J}^a(n)}{\partial \mathbf{W}_{ho}^a(n)} = \frac{\partial \mathcal{J}^a(n)}{\partial E^a(n)} \frac{\partial E^a(n)}{\partial \hat{h}(n)} \frac{\partial \hat{h}(n)}{\partial \mathbf{W}_{ho}^a(n)} \quad (9.40)$$

Using (9.34), (9.37) and (9.39) in (9.40) we get,

$$\frac{\partial \mathcal{J}^a(n)}{\partial \mathbf{W}_{ho}^a(n)} = \eta^a(n) E^{aT}(n) \quad (9.41)$$

Now considering the learning rate of actor-network as $\alpha^a \in \mathbb{R}^+$, we calculate the update rule

of the actor-network as,

$$\begin{aligned}\mathbf{W}_{ho}^a(n+1) &= \mathbf{W}_{ho}^a(n) - \alpha^a \eta^a(n) E^{aT}(n) \\ &= \mathbf{W}_{ho}^a(n) - \alpha^a \eta^a(n) (\hat{h}(n) + \hat{\mathcal{J}}(n))^T\end{aligned}\tag{9.42}$$

However, we can not use (9.42) for the weight update rule of actor-network as it has the unknown function $f_3(\mathbf{x}_3(n))$. In order to eliminate the dependency on $f_3(\mathbf{x}_3(n))$, we use (9.21) in (9.15) to find the value of $\hat{h}(n)$ as,

$$\hat{h}(n) = E_v(n+1) + \mathbf{\Lambda} E_v(n)\tag{9.43}$$

Finally, we obtain the weight update rule of the actor-network using (9.43) in (9.42) as,

$$\begin{aligned}\mathbf{W}_{ho}^a(n+1) &= \mathbf{W}_{ho}^a(n) - \alpha^a \eta^a(n) (\hat{h}(n) + \hat{\mathcal{J}}(n))^T \\ &= \mathbf{W}_{ho}^a(n) - \alpha^a \eta^a(n) \\ &\quad (E_v(n+1) + \mathbf{\Lambda} E_v(n) + \hat{\mathcal{J}}(n))^T\end{aligned}\tag{9.44}$$

The actor is responsible for estimating the nonlinear dynamics of the system and generating the control input (9.21) such that it minimizes the cost function $\hat{\mathcal{J}}(n)$ (estimated by the critic network). The critic adapts the estimation of the cost function, given $\mathbf{x}(n)$, and the voltage regulation error signal. This process is repeated until we reach our control goal. In the next section, we focus our discussion on the design of adversarial attacks on the RL agents.

9.4 Adversarial Attack on RL agents

In Section 9.3, we delved into the structural components of RL agents, which consist of both an actor-network and a critic network. These agents engage with the environment through a reward signal, a pivotal element in reinforcement learning. At every discrete time step n , the

RL agent decides upon an action based on its current observation of the environment, aiming to minimize the cumulative future rewards as stipulated in equation (9.26). A control policy governs this decision-making process denoted as $\mathbf{u}_v$, hereafter referred to as Π to align with standard RL terminology. Given that this policy is typically modeled using neural networks, it renders the RL agent vulnerable to adversarial attacks. Such attacks involve perturbing observations that coerce the RL agent into undertaking sub-optimal actions, thereby increasing its cumulative future rewards. The success of an adversarial attack can be quantified by the extent to which it diminishes the rewards accrued by the RL agent. In contrast to conventional adversarial strategies aimed at degrading the RL agent’s performance, we propose an innovative approach: designing adversarial attacks that reduce the total accumulated reward. This paradigm shift introduces novel challenges, particularly considering the intricate relationship between the actions taken by the RL agent and its subsequent observations. A strategically designed adversary can exploit this relationship to guide the agent towards states that maximize cumulative rewards, potentially leading to significant degradation in overall performance. However, it’s essential to acknowledge the potential risks associated with such strategic attacks. Given that actions taken by the RL agent influence its future observations, an optimal design of an adversary could inadvertently steer the agent towards states with catastrophic outcomes. Consequently, the design of strategic attacks must be approached with caution, balancing the quest for improved rewards with the mitigation of potential risks. In the subsequent subsection, we will delve deeper into the intricacies of strategic attack design, exploring methodologies to systematically perturb observations and guide the RL agent towards states conducive to enhanced reward accumulation.

9.4.1 Strategic Attack design

In each episode, an RL agent observes a sequence of observations or states $\mathbf{x}_1, \mathbf{x}, \dots, \mathbf{x}_L$. Instead of attacking at every time step in an episode, we design a strategically-timed attack that selects a subset of time steps to attack the agent. Let $\delta_1, \delta_2, \dots, \delta_L$ be the sequence of perturbations. Let R_1 be the expected return at the first time step. We formulate the above

intuition as an optimization problem as,

$$\begin{aligned}
& \min_{b_1, b_2, \dots, b_L, \delta_1, \delta_2, \dots, \delta_L} R_1(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_L) \\
& \bar{\mathbf{x}}_n = \mathbf{x}_n + b_n \delta_n \quad \forall \quad n = 1, \dots, L \\
& b_n \in \{0, 1\} \quad \forall \quad n = 1, \dots, L \\
& \sum_{n=1}^L b_n \leq \Gamma
\end{aligned} \tag{9.45}$$

Here $b_1, b_2, \dots, b_L$ are binary variables signifying adversarial attack time in an episode and δ_n is the magnitude of adversarial. If $b_n = 1$, the perturbation is applied. Otherwise, we do not alter the state. The constant Γ limits the total number of attacks. In other words, the adversary maximizes the expected accumulated rewards by strategically attacking less than $\Gamma \leq L$ time instants.

The optimization problem presented in (9.45) poses a formidable challenge due to its nature as a mixed-integer programming problem. Additionally, in reinforcement learning (RL) scenarios, the observation at any given time step n is contingent upon all preceding observations, compounding the complexity of devising a solution to (9.45). This complexity escalates exponentially with the length of the observation sequence denoted by L . Given these daunting challenges, a direct solution to (9.45) becomes impractical. In investigating adversarial attacks on deep RL agents, we circumvent these computational limitations by proposing a heuristic algorithm. This algorithm aims to compute the sequence of attack timing decisions $b_1, b_2, \dots, b_L$ (addressing the "when-to-attack" problem) and the corresponding attack magnitudes $\delta_1, \delta_2, \dots, \delta_L$ (resolving the "how-to-attack" problem). By adopting a heuristic approach, we trade off optimality for computational tractability, thereby enabling practical exploration of adversarial strategies. In the subsequent discussion, we first outline our methodology for addressing the "when-to-attack" problem, followed by an elucidation of our approach to tackling the "how-to-attack" problem.

9.4.2 When to Attack

We introduce a relative action preference function c for solving the when-to-attack problem. The function c computes the preference of the agent in taking the most preferred action over the least preferred action at the current state $\mathbf{x}(n)$ [149]. A large c implies that the agent strongly prefers one action. For policy gradient-based methods such as the actor-critic algorithm, if the action distribution of a well-trained agent is uniform at state $\mathbf{x}(n)$, it means that taking any action is equally good. But, when an agent strongly prefers a specific action (The action has a relatively high probability.), it means that it is critical to perform the action; otherwise, the accumulated reward will be reduced. Based on this intuition, we define the c function as,

$$c(\mathbf{x}(n)) = \max_{\mathbf{a}(n)} \Pi(\mathbf{x}(n), \mathbf{a}(n)) - \min_{\mathbf{a}(n)} \Pi(\mathbf{x}(n), \mathbf{a}(n)) \quad (9.46)$$

where Π is the policy network that maps a state–action pair $(\mathbf{x}(n), \mathbf{a}(n))$ to a probability, representing the likelihood that the action at is chosen. In our strategically-timed attack, the adversary attacks the deep RL agent at time step n when the relative action preference function c has a value greater a threshold parameter β . In other words, $b_n = 1$ if and only if $c(\mathbf{x}(n)) \geq \beta$. We note the β parameter controls how often it attacks the RL agent and is related to Γ .

9.4.3 How much to attack

To craft an adversarial magnitude at time step δ_n , we search for a perturbation to be added to the observation that can change the preferred action of the agent from the originally (before applying the perturbation) most preferred one to the originally least preferred one. We use the attack method similar to a false data injection attack method to corrupt the state at time instant n as,

$$\delta_n = \mu \mathbf{x} \quad (9.47)$$

Parameters	Symbol	Quantity	Value
Electrical	V _{dc}	DC Voltage	700 V
	V _{MG}	MG system voltage	380 (phase-phase rms) V
	f	MG nominal frequency	60 Hz
	L _f	filter inductance	3.5 mH
	R _f	filter resistance	0.1 Ω
	C _f	filter capacitance	50 μ F
Lines	R	Line resistance	0.15 Ω
	L	Line inductance	3500 μ H
PI controller	K _{pc}	Proportional current gain	54.97
	K _{ic}	Integral current gain	1.57×10^3
	K _{pv}	Proportional voltage gain	0.2859
	K _{iv}	Integral voltage gain	594.85
Droop	m _p	Frequency droop	0.9×10^{-6}
	n _q	Voltage droop	3×10^{-3}

Table 9.1: Parameters of the microgrid test system

Where, μ decides the degree of corruption, and its value is chosen from a uniform probability distribution between 0 and 1. This approach allows us to leverage the output of a trained deep RL agent as a cue to craft effective adversarial attack for increasing the cumulative rewards.

9.5 Simulation Results

In this section, we discuss the effectiveness of the proposed adversarial robust MARL for secondary control of an islanded four IBR-based microgrid test system. The test system and the corresponding communication graph are shown in Fig. 3.2a and Fig. 3.2b, respectively. The DGs are connected through a series of RL branches. The electrical specification of the microgrid is shown in Table 9.1. It is assumed that DGs can only communicate with their

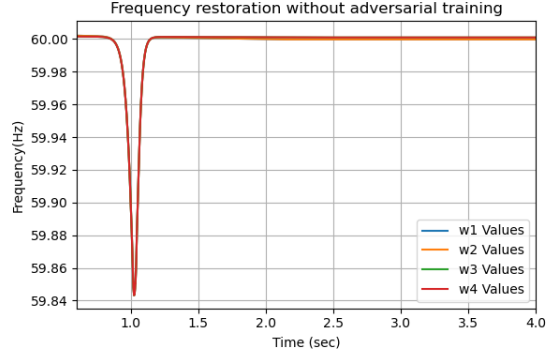


Figure 9.1: Frequency restoration with adversarial training

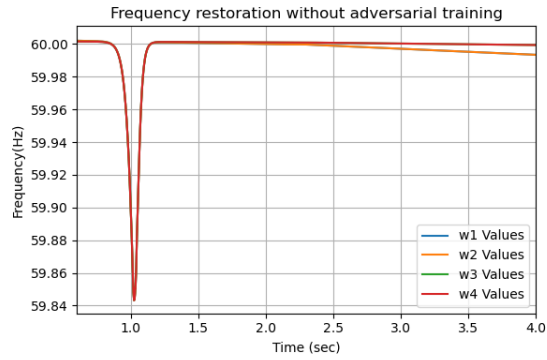


Figure 9.2: Frequency restoration with adversarial training

neighboring DGs. The associated adjacency matrix of the communication graph is,

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

We train two MARL algorithms, one without adversarial robustness and the other with the adding adversaries in the policy function. The time of adversaries' introduction to the policy function is decided by following (9.46). The magnitude of adversaries is decided by (9.47). We have selected the adversaries' magnitude to 0.1 for this study. The simulation results for secondary frequency restoration with and without adversarial training of RL agents are shown in Fig. (9.1) and Fig. 9.2, respectively. It can be seen that, initially there is a droop

in the frequency values till one second. This is because the microgrid is operating with only the primary droop control on. At one second, MARL agents come into action and restore the system frequency to the nominal frequency values. At two seconds, there is an attack introduced to the microgrid agents. The effect of cyber attacks on the agents trained with adversarial training is negligible, and the microgrid operates without any disturbance in the system frequency, see Fig. 9.1. On the contrary, it can be seen from Fig 9.2 that the attack makes the microgrid unstable, with two agents operating at one frequency and the other two at another. This makes the system unstable.

Chapter 10

Conclusion and Future Work

This chapter provides a summary of the major contributions of the dissertation and outlines the potential avenues for future work. The dissertation focuses on three major areas of distribution system operation and control; (1) efficient state estimation for better situational awareness; (2) improved secondary control for voltage and frequency stability, and (3) Robustness of the proposed state estimation and control architecture to potential cyber threats. To achieve the stated objectives, various optimization and data-driven formulations are developed in the context of the power distribution grid. These approaches are useful for utility companies to enhance grid visibility and improved control actions for better distribution grid operation. The key accomplishments of the dissertation can be summarized as follows:

The state-of-the-art state estimation approaches, such as WLS and its variants, require a large number of measurements. To overcome this, researchers have proposed sparsity-based techniques like compressive sensing (CS), tensor completion (TC), or matrix completion (MC). These methods leverage the inherent sparsity in raw measurements (MC and TC), or on a sparsifying domain (CS) to recover the system states from limited measurements. However, existing approaches primarily focus on spatial sparsity, neglecting the dynamics of temporal correlation/sparsity within the measurements. Hence, the dissertation proposes matrix completion approaches for dynamic state estimation in power distribution grids in

Chapter 4. More specifically, we develop two approaches: locally weighted matrix completion (LW-MC) and Bayesian matrix completion with Kalman filter prediction (BMC-KF). The efficacy of the proposed methods is evaluated against static MC of state estimation and found to have better estimation accuracy with improved computational complexity compared to the existing static MC approach. However, the matrix completion approach lacks scalability because of the Bayesian kernel matrix inversion. Hence, we propose distributed state estimation.

Chapter 5 of the dissertation proposes a distributed compressive sensing (CS) state estimation approach in a distribution grid by dividing the large distribution network into smaller partitions. This method utilizes an alternating direction of multipliers (ADMM) for iterative information exchange among sub-areas. Although the proposed distributed DSSE approach was successful in reducing the computational burden from a centralized state estimation engine to multiple distributed engines. Still, the computation time depends on the speed of the larger partition of the network and the number of information exchanges among the neighboring areas. Hence, we propose a hierarchical network partitioning algorithm to partition the large distribution network optimally with a minimum number of sharing edges. The performance of the proposed approaches evaluated on standard IEEE test systems and large practical distribution networks confirm the effectiveness of this approach in recovering system states accurately with improved complexity.

Furthermore, to reduce the computation time further and achieve real-time state estimation in large power distribution networks, we propose a novel data-driven and model-based neural network approach called deep unfolding in Chapter 6. This method achieves better computational performance by shifting the computation burden to the training phase and uses only algebraic equations in the implementation phase. In this method, we formulate a neural network by unfolding the optimization iterations to layers of a neural network. Simulations on the standard IEEE 123 bus network and 549 node networks confirm the scalability of the proposed approach to large distribution networks. Applying deep learning for safety-critical applications like DSSE is challenging and requires confidence in the estimates. Hence, the dissertation proposes a Bayesian framework in deep unfolding to provide both

mean and variance of the estimated estimates.

The effectiveness of the proposed distributed CS-based DSSE to commonly occurring cyber threats is presented in Chapter 7. More specifically, the impact of loss of measurement data, false data injection (FDI), replay, and neighborhood cyber-attacks is analyzed. Simulation results on standard IEEE test systems show that the CS-based distributed state estimation is highly robust to all four attacks, with the replay attack being the least effective in creating a significant error in state estimation results.

Chapter 8 proposes an innovative solution to address distributed secondary control in microgrids (MG) for frequency restoration. This method eliminates the necessity for a PI control-based secondary control, which traditionally exhibits asymptotic convergence with microgrid agents. The approach involves formulating the secondary frequency control problem as a non-cooperative game within a multiagent system framework. The paper derives both necessary and sufficient conditions for achieving the NE. By conducting extensive simulations on a representative MG model under various loading conditions, the proposed approach is validated and demonstrates its superiority over the distributed averaging proportional integral control method.

The dissertation delves further to develop an intelligent and adversarial robust reinforcement learning-based secondary control of islanded microgrids in Chapter 9. This method uses a multiagent actor-critic RL approach for the frequency and voltage restoration in an islanded microgrid. Simulation results on a four-inverter microgrid show the efficacy of the proposed approach.

10.1 Future Work

This section presents possible future directions to solve the challenges faced by distribution systems and effective methodologies for addressing them. The following are the potential extensions to the work presented in this dissertation:

- In Chapter 4, we propose two algorithms, LW-MC and BMC-KF, for dynamic state estimation in a low-observable distribution grid. This work only considers the dynamic

estimation of system states in distribution grids under limited measurements with an assumption of accurately defined network topology. However, the distribution network topology can change based on the switch and protective device status. Hence, developing a framework for topology identification with dynamic state estimation can be pursued as a part of future research efforts.

- Cyber attacks in power distribution networks have increased. Hence, a robust cyber-secured dynamic state estimation with topology identification can also be a part of future research efforts.
- In Chapter 6, we propose a centralized deep unfolding approach for state estimation in the power distribution network. However, this method exposes the entire data set to a centralized estimation engine, posing severe challenges to a single-point failure. A privacy-preserving federated learning in the context of DSSE can be an interesting topic for future research.
- Chapter 8 provides a game theoretic formulation for distributed secondary frequency control in an islanded microgrid. However, we have not evaluated its efficacy in grid-connected mode. Furthermore, we have only considered inverter-based resources. However, a real microgrid will have heterogeneous DERs, and the system's dynamics will change with other DERs like small-scale wind power, diesel generators, etc. It will be interesting to evaluate the controller's performance in a grid-connected microgrid. Hence, this can be pursued as a part of future research.
- Finding the Nash equilibrium is difficult to apply game theory in control theory applications. The time taken to reach a Nash equilibrium can be asymptotic, which is not desired in microgrid control applications. Finding a finite-time Nash equilibrium convergence can be pursued as the future direction of work
- Game theory finds its vast application in detecting and managing cyber attacks. It will be interesting to extend the proposed game theoretic control approach for cyber threat detection in microgrid control.

- Chapter 9 provides adversarial robust actor-critic RL for microgrid secondary control. For this study, we assumed that the microgrid is lossless. However, a practical microgrid has losses, and hence, managing the terminal voltage and frequency is challenging. Hence, developing an RL framework for lossy microgrids is one of the works we would like to pursue in the future.

Bibliography

- [1] Yih-Fang Huang, Stefan Werner, Jing Huang, Neelabh Kashyap, and Vijay Gupta. State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid. *IEEE Signal Processing Magazine*, 29(5):33–43, 2012.
- [2] Federal Energy Regulatory Commission et al. Assessment of demand response and advanced metering staff report, 2022.
- [3] Kwang Woo Joung, Taewan Kim, and Jung-Wook Park. Decoupled frequency and voltage control for stand-alone microgrid with high renewable penetration. *IEEE Transactions on Industry Applications*, 55(1):122–133, 2018.
- [4] Ali Bidram and Ali Davoudi. Hierarchical structure of microgrids control system. *IEEE Transactions on Smart Grid*, 3(4):1963–1976, 2012. doi: 10.1109/TSG.2012.2197425.
- [5] Qipeng Chen, Dritan Kaleshi, Zhong Fan, and Simon Armour. Impact of smart metering data aggregation on distribution system state estimation. *IEEE Transactions on Industrial Informatics*, 12(4):1426–1437, 2016.
- [6] Ali Abur and Antonio Gomez Exposito. *Power system state estimation: theory and implementation*. CRC press, 2004.
- [7] Marinko Barukčić, Marko Vukobratović, Dino Masle, Dalibor Buljić, and Željko Herderić. The evolutionary optimization approach for voltage profile estimation in a radial distribution network with a decreased number of measurements. In *2017 15th International Conference on Electrical Machines, Drives and Power Systems (ELMA)*, pages 26–31. IEEE, 2017.
- [8] Efthymios Manitsas, Ravindra Singh, Bikash C Pal, and Goran Strbac. Distribution

- system state estimation using an artificial neural network approach for pseudo measurement modeling. *IEEE Transactions on Power Systems*, 27(4):1888–1896, 2012.
- [9] E. A. Blood, M. D. Ilic, J. Ilic, and B. H. Krogh. A kalman filter approach to quasi-static state estimation in electric power systems. In *2006 38th North American Power Symposium*, pages 417–422, 2006. doi: 10.1109/NAPS.2006.359606.
- [10] EA Blood, BH Krogh, and MD Ilic. Electric power system static state estimation through kalman filtering and load forecasting. In *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–6. IEEE, 2008.
- [11] R. Gelagaev, P. Vermeyen, and J. Driesen. State estimation in distribution grids. In *2008 13th International Conference on Harmonics and Quality of Power*, pages 1–6, 2008. doi: 10.1109/ICHQP.2008.4668803.
- [12] G Durgaprasad and SS Thakur. Robust dynamic state estimation of power systems based on m-estimation and realistic modeling of system dynamics. *IEEE Transactions on Power Systems*, 13(4):1331–1336, 1998.
- [13] Kuang-Rong Shih and Shyh-Jier Huang. Application of a robust algorithm for dynamic state estimation of a power system. *IEEE Transactions on Power Systems*, 17(1):141–147, 2002.
- [14] Esmail Ghahremani and Innocent Kamwa. Dynamic state estimation in power system by applying the extended kalman filter with unknown inputs to phasor measurements. *IEEE Transactions on Power Systems*, 26(4):2556–2566, 2011.
- [15] Mohammadali Rostami and Saeed Lotfifard. Distributed dynamic state estimation of power systems. *IEEE Transactions on Industrial Informatics*, 14(8):3395–3404, 2018. doi: 10.1109/TII.2017.2777495.
- [16] A. S. Debs and R. E. Larson. A dynamic estimator for tracking the state of a power

- system. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(7):1670–1678, 1970. doi: 10.1109/TPAS.1970.292822.
- [17] Bradley M Bell and Frederick W Cathey. The iterated kalman filter update as a gauss-newton method. *IEEE Transactions on Automatic Control*, 38(2):294–297, 1993.
 - [18] S. M. S. Alam, B. Natarajan, and A. Pahwa. Distribution grid state estimation from compressed measurements. *IEEE Transactions on Smart Grid*, 5(4):1631–1642, 2014. doi: 10.1109/TSG.2013.2296534.
 - [19] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Joint topology identification and state estimation in unobservable distribution grids. *IEEE Transactions on Smart Grid*, 12(6):5299–5309, 2021.
 - [20] Priya L Donti, Yajing Liu, Andreas J Schmitt, Andrey Bernstein, Rui Yang, and Yingchen Zhang. Matrix completion for low-observability voltage estimation. *IEEE Transactions on Smart Grid*, 11(3):2520–2530, 2019.
 - [21] S. Dahale, H. S. Karimi, K. Lai, and B. Natarajan. Sparsity based approaches for distribution grid state estimation - a comparative study. *IEEE Access*, 8:198317–198327, 2020. doi: 10.1109/ACCESS.2020.3035378.
 - [22] Rahul Madbhavi, Balasubramaniam Natarajan, and Babji Srinivasan. Enhanced tensor completion based approaches for state estimation in distribution systems. *IEEE Transactions on Industrial Informatics*, pages 1–1, 2020. doi: 10.1109/TII.2020.3035449.
 - [23] Yajing Liu, April Sagan, Andrey Bernstein, Rui Yang, Xinyang Zhou, and Yingchen Zhang. Matrix completion using alternating minimization for distribution system state estimation. *arXiv preprint arXiv:1909.12459*, 2019.
 - [24] Biswajeet Rout, Shweta Dahale, and Balasubramaniam Natarajan. Dynamic matrix completion based state estimation in distribution grids. *IEEE Transactions on Industrial Informatics*, pages 1–1, 2022. doi: 10.1109/TII.2022.3162210.

- [25] Bo Liu, Hongyu Wu, Yingchen Zhang, Rui Yang, and Andrey Bernstein. Robust matrix completion state estimation in distribution systems. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2019.
- [26] Rahul Madbhavi, Balasubramaniam Natarajan, and Babji Srinivasan. Enhanced tensor completion based approaches for state estimation in distribution systems. *IEEE Transactions on Industrial Informatics*, 2020.
- [27] SM Shafiul Alam, Balasubramaniam Natarajan, and Anil Pahwa. Distribution grid state estimation from compressed measurements. *IEEE Transactions on Smart Grid*, 5(4):1631–1642, 2014.
- [28] Amit Joshi, Laya Das, Balasubramaniam Natarajan, and Babji Srinivasan. A framework for efficient information aggregation in smart grid. *IEEE Transactions on Industrial Informatics*, 15(4):2233–2243, 2018.
- [29] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Recursive dynamic compressive sensing in smart distribution systems. In *2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, 2020. doi: 10.1109/ISGT45199.2020.9087784.
- [30] Shweta Dahale and Balasubramaniam Natarajan. Joint matrix completion and compressed sensing for state estimation in low-observable distribution system. In *2021 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, pages 1–5, 2021. doi: 10.1109/ISGTLatinAmerica52371.2021.9543006.
- [31] Van-Hai Bui, Akhtar Hussain, and Hak-Man Kim. Diffusion strategy-based distributed operation of microgrids using multiagent system. *Energies*, 10(7):903, 2017.
- [32] Yafei Yang, Xiaohong Guan, Yuxun Zhou, Jiang Wu, and Ting Liu. Impact of information security on pmu-based distributed state estimation. In *IEEE PES Innovative Smart Grid Technologies*, pages 1–4. IEEE, 2012.

- [33] Ye Guo, Wenchuan Wu, Boming Zhang, and Hongbin Sun. A distributed state estimation method for power systems incorporating linear and nonlinear models. *International Journal of Electrical Power and Energy Systems*, 64:608–616, 2015.
- [34] Ariana Minot and Na Li. A fully distributed state estimation using matrix splitting methods. In *2015 American Control Conference (ACC)*, pages 2488–2493. IEEE, 2015.
- [35] Shiwei Xia, Qian Zhang, Jiangping Jing, Zhaohao Ding, Jing Yu, Bing Chen, and Haiwei Wu. Distributed state estimation of multi-region power system based on consensus theory. *Energies*, 12(5):900, 2019.
- [36] Md Masud Rana, Li Li, Steven W Su, and Wei Xiang. Consensus-based smart grid state estimation algorithm. *IEEE Transactions on Industrial Informatics*, 14(8):3368–3375, 2017.
- [37] Ricardo de Azevedo, Mehmet Hazar Cintuglu, Tan Ma, and Osama A Mohammed. Multiagent-based optimal microgrid control using fully distributed diffusion strategy. *IEEE Transactions on Smart Grid*, 8(4):1997–2008, 2017.
- [38] Md Masud Rana, Li Li, and Steven W Su. An adaptive-then-combine dynamic state estimation considering renewable generations in smart grids. *IEEE Journal on Selected Areas in Communications*, 34(12):3954–3961, 2016.
- [39] Oliver Kosut, Liyan Jia, Robert J Thomas, and Lang Tong. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In *2010 first IEEE international conference on smart grid communications*, pages 220–225. IEEE, 2010.
- [40] André Teixeira, Saurabh Amin, Henrik Sandberg, Karl H Johansson, and Shankar S Sastry. Cyber security analysis of state estimators in electric power systems. In *49th IEEE conference on decision and control (CDC)*, pages 5991–5998. IEEE, 2010.
- [41] Tung T Kim and H Vincent Poor. Strategic protection against data injection attacks on power grids. *IEEE Transactions on Smart Grid*, 2(2):326–333, 2011.

- [42] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven R Weller, and Zhao Yang Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 2016.
- [43] Ognjen Vuković and György Dán. Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks. *IEEE Journal on Selected Areas in Communications*, 32(7):1500–1508, 2014.
- [44] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Kalman filtered compressive sensing with intermittent observations. *Signal Processing*, 163:49–58, 2019. ISSN 0165-1684. doi: <https://doi.org/10.1016/j.sigpro.2019.05.004>. URL <https://www.sciencedirect.com/science/article/pii/S0165168419301628>.
- [45] Dai Wang, Xiaohong Guan, Ting Liu, Yun Gu, Chao Shen, and Zhanbo Xu. Extended distributed state estimation: A detection method against tolerable false data injection attacks in smart grids. *Energies*, 7(3):1517–1538, 2014.
- [46] Md Masud Rana, Li Li, and Steven W Su. Cyber attack protection and control of microgrids. *IEEE/CAA Journal of Automatica Sinica*, 5(2):602–609, 2017.
- [47] Govind Saraswat, Rui Yang, Yajing Liu, and Yingchen Zhang. Analyzing the effects of cyberattacks on distribution system state estimation. In *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 01–05, 2021. doi: 10.1109/ISGT49243.2021.9372262.
- [48] Mesut E Baran and Arthur W Kelley. State estimation for real-time monitoring of distribution systems. *IEEE Transactions on Power systems*, 9(3):1601–1609, 1994.
- [49] Kaveh Dehghanpour, Zhaoyu Wang, Jianhui Wang, Yuxuan Yuan, and Fankun Bu. A survey on state estimation techniques and challenges in smart distribution systems. *IEEE Transactions on Smart Grid*, 10(2):2312–2322, 2019. doi: 10.1109/TSG.2018.2870600.

- [50] Lamine Mili, Michael G Cheniae, and Peter J Rousseeuw. Robust state estimation of electric power systems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 41(5):349–358, 1994.
- [51] Murat Göl and Ali Abur. Lav based robust state estimation for systems measured by pmus. *IEEE Transactions on Smart Grid*, 5(4):1808–1814, 2014.
- [52] Atish K Ghosh, David L Lubkeman, Matthew J Downey, and Robert H Jones. Distribution circuit state estimation using a probabilistic approach. *IEEE Transactions on Power Systems*, 12(1):45–51, 1997.
- [53] R Singh, BC Pal, and RA Jabr. Distribution system state estimation through gaussian mixture model of the load as pseudo-measurement. *IET generation, transmission & distribution*, 4(1):50–59, 2010.
- [54] Ravindra Singh, Bikash C Pal, and Rabih A Jabr. Statistical representation of distribution system loads using gaussian mixture model. *IEEE Transactions on Power Systems*, 25(1):29–37, 2009.
- [55] Ali Arefi, Gerard Ledwich, and Behnaz Behi. An efficient dse using conditional multivariate complex gaussian distribution. *IEEE Transactions on Smart Grid*, 6(4):2147–2156, 2015.
- [56] Junbo Zhao, Gexiang Zhang, Massimo La Scala, and Zhaoyu Wang. Enhanced robustness of state estimator to bad data processing through multi-innovation analysis. *IEEE Transactions on Industrial Informatics*, 13(4):1610–1619, 2016.
- [57] Rahul Madbhavi, Hazhar Sufi Karimi, Balasubramaniam Natarajan, and Babji Srinivasan. Tensor completion based state estimation in distribution systems. *11th Conference on Innovative Smart Grid Technologies, Washington DC, Feb*, 2020.
- [58] SM Shafiul Alam, Balasubramaniam Natarajan, and Anil Pahwa. Distribution grid state estimation from compressed measurements. *IEEE Transactions on Smart Grid*, 5(4):1631–1642, 2014.

- [59] Biswajeet Rout and Balasubramaniam Natarajan. Impact of cyber attacks on distributed compressive sensing based state estimation in power distribution grids. *International Journal of Electrical Power and Energy Systems*, 142:108295, 2022.
- [60] Kursat Rasim Mestav, Jaime Luengo-Rozas, and Lang Tong. Bayesian state estimation for unobservable distribution systems via deep learning. *IEEE Transactions on Power Systems*, 34(6):4910–4920, 2019.
- [61] PN Pereira Barbeiro, J Krstulovic, H Teixeira, J Pereira, Filipe Joel Soares, and José Pedro Iria. State estimation in distribution smart grids using autoencoders. In *2014 IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014)*, pages 358–363. IEEE, 2014.
- [62] Ali Bidram, Ali Davoudi, Frank L Lewis, and Josep M Guerrero. Distributed cooperative secondary control of microgrids using feedback linearization. *IEEE transactions on power systems*, 28(3):3462–3470, 2013.
- [63] Ali Bidram, Ali Davoudi, and Frank L Lewis. A multiobjective distributed control framework for islanded ac microgrids. *IEEE Transactions on industrial informatics*, 10(3):1785–1798, 2014.
- [64] Ali Bidram, Ali Davoudi, Frank L Lewis, and Zhihua Qu. Secondary control of microgrids based on distributed cooperative control of multi-agent systems. *IET Generation, Transmission & Distribution*, 7(8):822–831, 2013.
- [65] Fanghong Guo, Changyun Wen, Jianfeng Mao, and Yong-Duan Song. Distributed secondary voltage and frequency restoration control of droop-controlled inverter-based microgrids. *IEEE Transactions on industrial Electronics*, 62(7):4355–4364, 2014.
- [66] Xiaoqing Lu, Xinghuo Yu, Jingang Lai, Yaonan Wang, and Josep M Guerrero. A novel distributed secondary coordination control approach for islanded microgrids. *IEEE Transactions on Smart Grid*, 9(4):2726–2740, 2016.

- [67] John W Simpson-Porco, Qobad Shafiee, Florian Dörfler, Juan C Vasquez, Josep M Guerrero, and Francesco Bullo. Secondary frequency and voltage control of islanded microgrids via distributed averaging. *IEEE Transactions on Industrial Electronics*, 62(11):7025–7038, 2015.
- [68] John W Simpson-Porco, Florian Dörfler, and Francesco Bullo. Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica*, 49(9):2603–2611, 2013.
- [69] Xiangyu Wu, Chen Shen, and Reza Iravani. A distributed, cooperative frequency and voltage control for microgrids. *IEEE Transactions on Smart Grid*, 9(4):2764–2776, 2016.
- [70] Nima Mahdian Dehkordi, Nasser Sadati, and Mohsen Hamzeh. Fully distributed cooperative secondary frequency and voltage control of islanded microgrids. *IEEE Transactions on Energy Conversion*, 32(2):675–685, 2017. doi: 10.1109/TEC.2016.2638858.
- [71] Yinliang Xu. Robust finite-time control for autonomous operation of an inverter-based microgrid. *IEEE Transactions on Industrial Informatics*, 13(5):2717–2725, 2017.
- [72] Shan Zuo, Ali Davoudi, Yongduan Song, and Frank L Lewis. Distributed finite-time voltage and frequency restoration in islanded ac microgrids. *IEEE Transactions on Industrial Electronics*, 63(10):5988–5997, 2016.
- [73] Zicong Deng, Yinliang Xu, Hongbin Sun, and Xinwei Shen. Distributed, bounded and finite-time convergence secondary frequency control in an autonomous microgrid. *IEEE Transactions on Smart Grid*, 10(3):2776–2788, 2018.
- [74] Nima Mahdian Dehkordi, Nasser Sadati, and Mohsen Hamzeh. Distributed robust finite-time secondary voltage and frequency control of islanded microgrids. *IEEE Transactions on Power systems*, 32(5):3648–3659, 2016.

- [75] Yinliang Xu and Hongbin Sun. Distributed finite-time convergence control of an islanded low-voltage ac microgrid. *IEEE Transactions on Power Systems*, 33(3):2339–2348, 2017.
- [76] Boda Ning, Qing-Long Han, and Lei Ding. Distributed finite-time secondary frequency and voltage control for islanded microgrids with communication delays and switching topologies. *IEEE Transactions on Cybernetics*, 51(8):3988–3999, 2020.
- [77] Stef Tijs. *Introduction to game theory*. Springer, 2003.
- [78] Adia Khalid and Nadeem Javaid. Coalition based game theoretic energy management system of a building as-service-over fog. *Sustainable Cities and Society*, 48:101509, 2019.
- [79] Jie Duan and Mo-Yuen Chow. A resilient consensus-based distributed energy management algorithm against data integrity attacks. *IEEE Transactions on smart grid*, 10(5):4729–4740, 2018.
- [80] Elham Semsar-Kazerooni and Khashayar Khorasani. Multi-agent team cooperation: A game theory approach. *Automatica*, 45(10):2205–2213, 2009.
- [81] Kyriakos G Vamvoudakis, Frank L Lewis, and Greg R Hudas. Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality. *Automatica*, 48(8):1598–1611, 2012.
- [82] Thanh Thi Nguyen, Ngoc Duy Nguyen, and Saeid Nahavandi. Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. *IEEE Transactions on Cybernetics*, 50(9):3826–3839, 2020. doi: 10.1109/TCYB.2020.2977374.
- [83] Jiajun Duan, Di Shi, Ruisheng Diao, Haifeng Li, Zhiwei Wang, Bei Zhang, Desong Bian, and Zhehan Yi. Deep-reinforcement-learning-based autonomous voltage control for power grid operations. *IEEE Transactions on Power Systems*, 35(1):814–817, 2019.

- [84] Yuanqi Gao, Wei Wang, and Nanpeng Yu. Consensus multi-agent reinforcement learning for volt-var control in power distribution networks. *IEEE Transactions on Smart Grid*, 12(4):3594–3604, 2021.
- [85] Di Cao, Junbo Zhao, Weihao Hu, Fei Ding, Qi Huang, and Zhe Chen. Attention enabled multi-agent drl for decentralized volt-var control of active distribution system using pv inverters and svcs. *IEEE transactions on sustainable energy*, 12(3):1582–1592, 2021.
- [86] Xi Zhang, Youbo Liu, Jiajun Duan, Gao Qiu, Tingjian Liu, and Junyong Liu. Ddpq-based multi-agent framework for svc tuning in urban power grid with renewable energy resources. *IEEE Transactions on Power Systems*, 36(6):5465–5475, 2021.
- [87] Andrey Bernstein and Emiliano Dall’Anese. Linear power-flow models in multiphase distribution networks. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2017.
- [88] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008.
- [89] Robert H Lasseter. Microgrids. In *2002 IEEE power engineering society winter meeting. Conference proceedings (Cat. No. 02CH37309)*, volume 1, pages 305–308. IEEE, 2002.
- [90] Daniel E Olivares, Ali Mehrizi-Sani, Amir H Etemadi, Claudio A Cañizares, Reza Iravani, Mehrdad Kazerani, Amir H Hajimiragha, Oriol Gomis-Bellmunt, Maryam Saeedifard, Rodrigo Palma-Behnke, et al. Trends in microgrid control. *IEEE Transactions on smart grid*, 5(4):1905–1919, 2014.
- [91] Mehrdad Yazdanian and Ali Mehrizi-Sani. Distributed control techniques in microgrids. *IEEE transactions on smart grid*, 5(6):2901–2909, 2014.
- [92] Mukul C Chandorkar, Deepakraj M Divan, and Rambabu Adapa. Control of parallel

- connected inverters in standalone ac supply systems. *IEEE transactions on industry applications*, 29(1):136–143, 1993.
- [93] Daniel E. Olivares, Ali Mehrizi-Sani, Amir H. Etemadi, Claudio A. Cañizares, Reza Iravani, Mehrdad Kazerani, Amir H. Hajimiragha, Oriol Gomis-Bellmunt, Maryam Saeedifard, Rodrigo Palma-Behnke, Guillermo A. Jiménez-Estévez, and Nikos D. Hatziargyriou. Trends in microgrid control. *IEEE Transactions on Smart Grid*, 5(4):1905–1919, 2014. doi: 10.1109/TSG.2013.2295514.
- [94] Josep M Guerrero, Juan C Vasquez, José Matas, Luis García De Vicuña, and Miguel Castilla. Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. *IEEE Transactions on industrial electronics*, 58(1):158–172, 2010.
- [95] Yousef Khayat, Qobad Shafiee, Rasool Heydari, Mobin Naderi, Tomislav Dragičević, John W. Simpson-Porco, Florian Dörfler, Mohammad Fathi, Frede Blaabjerg, Josep M. Guerrero, and Hassan Bevrani. On the secondary control architectures of ac microgrids: An overview. *IEEE Transactions on Power Electronics*, 35(6):6482–6500, 2020. doi: 10.1109/TPEL.2019.2951694.
- [96] Rasool Heydari, Yousef Khayat, Mobin Naderi, Amjad Anvari-Moghaddam, Tomislav Dragicevic, and Frede Blaabjerg. A decentralized adaptive control method for frequency regulation and power sharing in autonomous microgrids. In *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, pages 2427–2432. IEEE, 2019.
- [97] Josep M. Guerrero, Juan C. Vasquez, José Matas, Luis García de Vicuña, and Miguel Castilla. Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. *IEEE Transactions on Industrial Electronics*, 58(1):158–172, 2011. doi: 10.1109/TIE.2010.2066534.
- [98] J.A.P. Lopes, C.L. Moreira, and A.G. Madureira. Defining control strategies for micro-

- grids islanded operation. *IEEE Transactions on Power Systems*, 21(2):916–924, 2006. doi: 10.1109/TPWRS.2006.873018.
- [99] Nagaraju Pogaku, Milan Prodanovic, and Timothy C. Green. Modeling, analysis and testing of autonomous operation of an inverter-based microgrid. *IEEE Transactions on Power Electronics*, 22(2):613–625, 2007. doi: 10.1109/TPEL.2006.890003.
- [100] Ali Bidram and Ali Davoudi. Hierarchical structure of microgrids control system. *IEEE Transactions on Smart Grid*, 3(4):1963–1976, 2012.
- [101] Zhihua Qu, Jing Wang, and Richard A Hull. Cooperative control of dynamical systems with application to autonomous vehicles. *IEEE Transactions on Automatic Control*, 53(4):894–911, 2008.
- [102] Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.
- [103] David Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Transactions on Information Theory*, 57(3):1548–1566, 2011.
- [104] A. Bernstein and E. Dall’Anese. Linear power-flow models in multiphase distribution networks. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6, 2017. doi: 10.1109/ISGTEurope.2017.8260205.
- [105] M. Baran and F. F. Wu. Optimal sizing of capacitors placed on a radial distribution system. *IEEE Transactions on Power Delivery*, 4(1):735–743, 1989. doi: 10.1109/61.19266.
- [106] K Nishiya, J Hasegawa, and T Koike. Dynamic state estimation including anomaly detection and identification for power systems. In *IEE proceedings C (generation, transmission and distribution)*, volume 129, pages 192–198. IET, 1982.

- [107] Liangbei Xu and Mark Davenport. Dynamic matrix recovery from incomplete observations under an exact low-rank constraint. In *Advances in Neural Information Processing Systems*, pages 3585–3593, 2016.
- [108] S. Derin Babacan, Martin Luessi, Rafael Molina, and Aggelos K. Katsaggelos. Sparse bayesian methods for low-rank matrix estimation. *IEEE Transactions on Signal Processing*, 60(8):3964–3977, 2012. doi: 10.1109/TSP.2012.2197748.
- [109] AM Leite Da Silva, MB Do Coutto Filho, and JF De Queiroz. State forecasting in electric power systems. In *IEE Proceedings C (Generation, Transmission and Distribution)*, volume 130, pages 237–244. IET, 1983.
- [110] A. M. Leite da Silva, M. B. Do Coutto Filho, and J. M. C. Cantera. An efficient dynamic state estimation algorithm including bad data processing. *IEEE Transactions on Power Systems*, 2(4):1050–1058, 1987. doi: 10.1109/TPWRS.1987.4335300.
- [111] Thierry Bouwmans and El Hadi Zahzah. Robust pca via principal component pursuit: A review for a comparative evaluation in video surveillance. *Computer Vision and Image Understanding*, 122:22–34, 2014.
- [112] K. P. Schneider, B. A. Mather, B. C. Pal, C. . W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W. Kersting. Analytic considerations and design basis for the iee distribution test feeders. *IEEE Transactions on Power Systems*, 33(3):3181–3188, 2018. doi: 10.1109/TPWRS.2017.2760011.
- [113] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>, March 2014.
- [114] Emmanuel J Candes and Terence Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.
- [115] Emmanuel J. Candes and Michael B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008. doi: 10.1109/MSP.2007.914731.

- [116] Marco F. Duarte and Yonina C. Eldar. Structured compressed sensing: From theory to applications. *IEEE Transactions on Signal Processing*, 59(9):4053–4085, 2011. doi: 10.1109/TSP.2011.2161982.
- [117] April Sagan, Yajing Liu, and Andrey Bernstein. Decentralized low-rank state estimation for power distribution systems. *IEEE Transactions on Smart Grid*, 2021.
- [118] Stephen Boyd, Neal Parikh, and Eric Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.
- [119] Ahmad Reza Malekpour and Anil Pahwa. Radial test feeder including primary and secondary distribution network. In *2015 North American Power Symposium (NAPS)*, pages 1–9. IEEE, 2015.
- [120] Kevin P Schneider, BA Mather, BC Pal, C-W Ten, Greg J Shirek, Hao Zhu, Jason C Fuller, José Luiz Rezende Pereira, Luis Fernando Ochoa, Leandro Ramos de Araujo, et al. Analytic considerations and design basis for the iee distribution test feeders. *IEEE Transactions on power systems*, 33(3):3181–3188, 2017.
- [121] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.
- [122] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, 49(2):291–307, 1970. doi: 10.1002/j.1538-7305.1970.tb01770.x.
- [123] Daniel A Spielman and Shang-Hua Teng. Spectral partitioning works: Planar graphs and finite element meshes. In *Proceedings of 37th conference on foundations of computer science*, pages 96–105. IEEE, 1996.
- [124] Biswajeet Rout and Balasubramaniam Natarajan. Impact of cyber attacks on distributed compressive sensing based state estimation in power distribution grids. *International Journal of Electrical Power and Energy Systems*, 142:108295, 2022. ISSN 0142-0615.

- [125] Biswajeet Rout and Naran M. Pindoriya. Active distribution network analysis-a case study. In *2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, pages 828–833, 2018. doi: 10.1109/ISGT-Asia.2018.8467974.
- [126] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.
- [127] Marco F Duarte and Yonina C Eldar. Structured compressed sensing: From theory to applications. *IEEE Transactions on signal processing*, 59(9):4053–4085, 2011.
- [128] Shweta Dahale and Bala Natarajan. Bayesian framework for multi-timescale state estimation in low-observable distribution systems. *IEEE Transactions on Power Systems*, pages 1–1, 2022. doi: 10.1109/TPWRS.2022.3155151.
- [129] John R. Hershey, Jonathan Le Roux, and Felix Weninger. Deep unfolding: Model-based inspiration of novel deep architectures, 2014.
- [130] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Joint topology identification and state estimation in unobservable distribution grids. *IEEE Transactions on Smart Grid*, 12(6):5299–5309, 2021. doi: 10.1109/TSG.2021.3102179.
- [131] Zhaoyu Wang. URL <https://wzy.ece.iastate.edu/Testsystem.html>.
- [132] Opendssdirect.py, version 00, 7 . URL <https://www.osti.gov//servlets/purl/1374532>.
- [133] Adel Javanmard and Andrea Montanari. Debiasing the lasso: Optimal sample size for gaussian designs. 2018.
- [134] Adel Javanmard and Andrea Montanari. Confidence intervals and hypothesis testing for high-dimensional regression. *The Journal of Machine Learning Research*, 15(1): 2869–2909, 2014.

- [135] Laurent Valentin Jospin, Hamid Laga, Farid Boussaid, Wray Buntine, and Mohammed Bennamoun. Hands-on bayesian neural networks—a tutorial for deep learning users. *IEEE Computational Intelligence Magazine*, 17(2):29–48, 2022.
- [136] Alex Kendall and Roberto Cipolla. Modelling uncertainty in deep learning for camera relocalization. In *2016 IEEE international conference on Robotics and Automation (ICRA)*, pages 4762–4769. IEEE, 2016.
- [137] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.
- [138] Alex Graves. Practical variational inference for neural networks. *Advances in neural information processing systems*, 24, 2011.
- [139] Jingxuan Wang, Lucas CK Hui, Siu-Ming Yiu, Eric Ke Wang, and Junbin Fang. A survey on cyber attacks against nonlinear state estimation in power systems of ubiquitous cities. *Pervasive and Mobile Computing*, 39:52–64, 2017.
- [140] Julia E Sullivan and Dmitriy Kamensky. How cyber-attacks in ukraine show the vulnerability of the us power grid. *The Electricity Journal*, 30(3):30–35, 2017.
- [141] Engang Tian and Chen Peng. Memory-based event-triggering h_∞ load frequency control for power systems under deception attacks. *IEEE transactions on cybernetics*, 50(11):4610–4618, 2020.
- [142] A Lee. Electric sector failure scenarios and impact analyses-version 3.0. *National Electric Sector Cybersecurity Organization Resource (NESCOR) Technical Working Group: Palo Alto, CA, USA*, 2015.
- [143] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

- [144] John W. Simpson-Porco, Qobad Shafiee, Florian Dörfler, Juan C. Vasquez, Josep M. Guerrero, and Francesco Bullo. Secondary frequency and voltage control of islanded microgrids via distributed averaging. *IEEE Transactions on Industrial Electronics*, 62(11):7025–7038, 2015. doi: 10.1109/TIE.2015.2436879.
- [145] Mohammad N Marwali and Ali Keyhani. Control of distributed generation systems-part i: Voltages and currents control. *IEEE Transactions on power electronics*, 19(6):1541–1550, 2004.
- [146] Vaibhav Vaishnav, Dushyant Sharma, and Anoop Jain. Quadratic-droop-based distributed secondary control of microgrid with detail-balanced communication topology. *IEEE Systems Journal*, 17(3):3401–3412, 2023. doi: 10.1109/JSYST.2023.3240171.
- [147] Jagannathan Sarangapani. *Neural network control of nonlinear discrete-time systems*. CRC press, 2018.
- [148] FW Lewis, Suresh Jagannathan, and Aydin Yesildirak. *Neural network control of robot manipulators and non-linear systems*. CRC press, 1998.
- [149] Amir-massoud Farahmand. Action-gap phenomenon in reinforcement learning. *Advances in Neural Information Processing Systems*, 24, 2011.

Appendix A

ADMM Convergence

A.1 Appendix

Problem discussed in (5.10) can be viewed as variable substitution for a two area system as follows;

$$\begin{aligned} \hat{x} = \underset{x}{\operatorname{argmin}} \quad & f_a(\mathbf{x}_a) + f_b(\mathbf{x}_b) \\ \text{subject to: } & \mathbf{A}_a \mathbf{x}_{ab} + \mathbf{A}_b \mathbf{x}_{ba} = \mathbf{q} \end{aligned} \quad (\text{A.1})$$

where, $\mathbf{A}_a$ and $\mathbf{A}_b$ are identity matrices of appropriate size, $\mathbf{x}_{ab}$ and $\mathbf{x}_{ba}$ are obtained by dropping the superscript of $\mathbf{x}_{a,adj}^{ab}$ and $\mathbf{x}_{a,adj}^{ab}$ for convenience. We will limit our discussion to two area system only and the same argument can be extended for multi area power system network.

The optimal value of (A.1) can be denoted by

$$\mathbf{x}^* = \inf \{f_a(\mathbf{x}_a) + f_b(\mathbf{x}_b) | \mathbf{A}_a \mathbf{x}_{ab} + \mathbf{A}_b \mathbf{x}_{ba} = \mathbf{q}\} \quad (\text{A.2})$$

The augmented Lagrangian of (A.1) can be formed as,

$$\mathcal{L}_\rho(\mathbf{x}_a, \mathbf{x}_b, \lambda) = \inf \{f_a(\mathbf{x}_a) + f_b(\mathbf{x}_b) + \lambda^T (\mathbf{A}_a \mathbf{x}_{ab} + \mathbf{A}_b \mathbf{x}_{ba} - \mathbf{q}) + \frac{\rho}{2} \|\mathbf{A}_a \mathbf{x}_{ab} + \mathbf{A}_b \mathbf{x}_{ba} - \mathbf{q}\|_2^2 \} \quad (\text{A.3})$$

Assuming both f_a and f_b are convex sets, the augmented Lagrangian has a saddle point i.e., there exists a $(\mathbf{x}_a^*, \mathbf{x}_b^*, \lambda^*)$, for which

$$\mathcal{L}_0(\mathbf{x}_a^*, \mathbf{x}_b^*, \lambda) \leq \mathcal{L}_0(\mathbf{x}_a^*, \mathbf{x}_b^*, \lambda^*) \leq \mathcal{L}_0(\mathbf{x}_a, \mathbf{x}_b, \lambda^*) \quad (\text{A.4})$$

holds for all $\mathbf{x}_a, \mathbf{x}_b, \lambda$.

The necessary and sufficient Optimality conditions for the ADMM are primal feasibility,

$$\mathbf{A}_a \mathbf{x}_{ab} + \mathbf{A}_b \mathbf{x}_{ba} - \mathbf{q} = 0, \quad (\text{A.5})$$

and dual feasibility.

$$0 \in \partial f_a(\mathbf{x}_a^*) + \mathbf{A}_a^T \lambda^* \quad (\text{A.6})$$

$$0 \in \partial f_b(\mathbf{x}_b^*) + \mathbf{A}_b^T \lambda^*. \quad (\text{A.7})$$

Where, ∂ denotes the sub-differential operator. The optimality condition always holds for (A.7). The primal residual (A.5) and the dual residual (A.6) converges to 0 as the ADMM proceeds [118].

Appendix B

Reuse permissions from publishers



Requesting permission to reuse content from an IEEE publication

Dynamic Matrix Completion Based State Estimation in Distribution Grids

Author: Biswajeet Rout
Publication: IEEE Transactions on Industrial Informatics
Publisher: IEEE
Date: November 2022
Copyright © 2022, IEEE

Thesis / Dissertation Reuse

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

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

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

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

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

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

BACK

CLOSE WINDOW



ELECTRICAL POWER & ENERGY SYSTEMS

Impact of cyber attacks on distributed compressive sensing based state estimation in power distribution grids

Author: Biswajeet Rout, Balasubramaniam Natarajan
Publication: International Journal of Electrical Power & Energy Systems
Publisher: Elsevier
Date: November 2022
© 2022 Elsevier Ltd. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW



Efficient Network Partitioning: Application for Decentralized State Estimation in Power Distribution Grids

Conference Proceedings: 2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)

Author: Biswajeet Rout

Publisher: IEEE

Date: 16 January 2023

Copyright © 2023, IEEE

Thesis / Dissertation Reuse

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

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

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

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

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

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

BACK

CLOSE WINDOW