Stability and Security of Distribution Networks with High-Penetration Renewables

by

Shammya Shananda Saha

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

Approved June 2021 by the
Graduate Supervisory Committee:

Nathan Johnson, Co-Chair
Anna Scaglione, Co-Chair
Daniel Arnold
Dragan Boscovic

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Rapid increases in the installed amounts of Distributed Energy Resources are forcing a paradigm shift to guarantee stability, security, and economics of power distribution systems. This dissertation explores these challenges and proposes solutions to enable higher penetrations of grid-edge devices. The thesis shows that integrating Graph Signal Processing with State Estimation formulation allows accurate estimation of voltage phasors for radial feeders under low-observability conditions using traditional measurements. Furthermore, the Optimal Power Flow formulation presented in this work can reduce the solution time of a bus injection-based convex relaxation formulation, as shown through numerical results. The enhanced real-time knowledge of the system state is leveraged to develop new approaches to cyber-security of a transactive energy market by introducing a blockchain-based Electron Volt Exchange framework that includes a distributed protocol for pricing and scheduling prosumers' production/consumption while keeping constraints and bids private. The distributed algorithm prevents power theft and false data injection by comparing prosumers' reported power exchanges to models of expected power exchanges using measurements from grid sensors to estimate system state. Necessary hardware security is described and integrated into underlying grid-edge devices to verify the provenance of messages to and from these devices. These preventive measures for securing energy transactions are accompanied by additional mitigation measures to maintain voltage stability in inverter-dominated networks by expressing local control actions through Lyapunov analysis to mitigate cyber-attack and generation intermittency effects. The proposed formulation is applicable as long as the Volt-Var and Volt-Watt curves of the inverters can be represented as Lipschitz constants. Simulation results demonstrate how smart inverters can mitigate voltage oscillations throughout the distribution network. Approaches are rigorously explored and validated using a combination of real distri-

bution networks and synthetic test cases. Finally, to overcome the scarcity of real data to test distribution systems algorithms a framework is introduced to generate synthetic distribution feeders mapped to real geospatial topologies using available OpenStreetMap data. The methods illustrate how to create synthetic feeders across the entire ZIP Code, with minimal input data for any location. These stackable scientific findings conclude with a brief discussion of physical deployment opportunities to accelerate grid modernization efforts.

DEDICATION

To My Family

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Dr. Nathan Johnson, for his insight, direction, and advice regarding the Ph.D. and otherwise. He has been a mentor guiding me during tough times, made me a better writer, allowed me to participate in proposal writing, believed in me, and taught me how to be a mentor for someone else.

I am grateful for the support that I got from Dr. Anna Scaglione. She taught me to look into a problem from a mathematical standpoint, helped me to overcome my fear regarding the theoretical development of a problem formulation. This work benefited a lot from her guidance and advice, and direct involvement.

I want to thank my other committee member. Dr. Daniel Arnold allowed me to do an internship at Lawrence Berkeley National Lab, helped me escape the Tempe Summer, taught me non-linear controls, and helped me expand my knowledge. Finally, the EVE work presented in this dissertation would not have been possible without the generous support I got from Dr. Dragan Boscovic's team and himself. Thanks for allowing me to work with your students, which provided me the knowledge and background regarding Hyperledger Fabric.

Thank you to all the SINE Lab members and my team members at LEAPS. You guys kept me sane, helped me redefine my future goals, and supported my decisions. Thanks for considering me a part of the family. I consider myself lucky to be part of two teams with different dynamics and mind sets.

Thank you to my parents, my in-laws, my brother, my sisters, as well as my friend Amit Kumar Sikder. Thanks, Amit, for listening to my relentless rant, anxiousness, and suffering regarding the Ph.D. life.

A special thanks to the Polytechnic IT Team and business team for their timely and continuous support regarding purchase and travel. They made sure the process is smooth, and I do not break any ASU regulations in the process.

Finally, I would not have made it to the finish line if my wife, Snigdha, had not been so supportive and provided me her 3 PM coffee. I enjoyed being your guinea pig for coffee testing, which kept me sane during the crazy afternoons. Thanks for making sure I do not get "hangry" and can deliver everything in a timely fashion. And most importantly, thanks for believing in me when I was in doubt.

And thanks to Almighty for keeping me healthy during this global pandemic and keeping my family safe. Please bestow your love as I move forward in my life.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF ALGORITHMS

LIST OF ACRONYMS

Acronym	Meaning
2FA	Two Factor Authentication.
ABAC	Attribute-Based Access Control.
ADMM	Alternative Direction Method of Multipliers.
AMI	Advanced Metering Infrastructure.
CA	Certificate Authority.
CPS	Cyber Physical Systems.
CTC	Cryptographic Trust Center.
DER	Distributed Energy Resources.
DSP	Digital Signal Processing.
EMS	Energy Management Systems.
ESP	Eco-Secure Provisioning.
ESS	Energy Storage Systems.
EV	Electric Vehicles.
EVE	Electron Volt Exchange.
FDIA	False Data Injection Attack.
GAN	Generative Adversarial Network.
GFT	Graph Fourier Transform.
GL	Global Ledger.
GSO	Graph Shift Operator.
GSP	Graph Signal Processing.
HIL	Hardware-In-the-Loop.
HLF	Hyperledger Fabric.
IoT	Internet-of-Things.
IT	Information Technology.
IUR	Industry Utility Registrar.

Acronym	Meaning
LL	Local Ledger.
MAC	Message Authentication Code.
MAE	Mean Absolute Error.
MMSE	Minimum Mean Square Error.
MSP	Membership Service Provider.
OPF	Optimal Power Flow.
OT	Operational Technology.
OTP	One Time Pad.
PHIL	Power Hardware-in-the-Loop.
POO	Proof Of Origin.
PV	Photovoltaic.
RNM	Reference Network Model.
RSA	Rivest–Shamir–Adleman.
RTDS	Real-Time Digital Simulator.
SDP	Semi-Definite Programming.
SE	State Estimation.
SOCP	Second-Order Cone Programming.
TCL	Thermostatically Controlled Loads.
TE	Transactive Energy.
TLS	Transport Layer Security.
VV	Volt-Var.
VW	Volt-Watt.

LIST OF SYMBOLS

Symbol	Description
$\mathcal{A}$	Calligraphic letters are sets
$\lvert\mathcal{A}\rvert$	Denotes the cardinality of set $\mathcal{A}$
T	Number of intervals in the decision horizon
$\mathcal{N}$	Set of aggregators
N	Number of aggregators
$\boldsymbol{x}$	Boldfaced lower-case letters denote vectors and x_i denotes the i^{th} element of a vector $\boldsymbol{x}$
$\boldsymbol{X}$	Boldfaced upper-case letters denote matrices and X_{ij} denotes the ij^{th} entry of a matrix $\boldsymbol{X}$
$\boldsymbol{X}^\top$	Transpose is denoted by $^\top$, so $\boldsymbol{X}^\top$ is the transpose of $\boldsymbol{X}$
$\boldsymbol{X}^\mathsf{H}$	Complex conjugate transpose is denoted by $^\mathsf{H}$, so $\boldsymbol{X}^\mathsf{H}$ is the complex conjugate transpose of $\boldsymbol{X}$
$\boldsymbol{X}^\dagger$	Is the pseudo-inverse of $\boldsymbol{X}$, where $\boldsymbol{X}^\dagger = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$
$\mathcal{E}_e$	Set of edges/lines connecting the set of buses $\mathcal{B}$ in $\mathcal{G}_e$
$\mathcal{E}_c$	Set of communication links between aggregators $\mathcal{N}$
$\mathcal{B}$	Set of buses/nodes b in the power grid
$\mathcal{G}_e$	Electric Grid Graph
$\mathcal{G}_c$	Communication network graph
p_i^c, p_i^g	Real power demand, Real power generation at node i
q_i^c, q_i^g	Reactive power demand, Real power generation at node i
r_{ij}, x_{ij}	Line resistance, Line reactance of line between nodes i and j
P_{ij}, Q_{ij}	Real power flow, Reactive power flow of line between nodes i and j
v_i	Voltage of node i

Symbol	Description
c_{ij}	Current of line between nodes i and j
s_i	Rated apparent power of inverter at node i
$\overline{p_i}$	Maximum real power output of inverter at node i
q_i^{lim}	Hardware limit of reactive power generation of inverter at node i
$f_{p,i}$	Volt-Watt control function of inverter at node i
$f_{q,i}$	Volt-Var control function of inverter at node i
$C_{p,i}$	Lipschitz constant for Volt-Watt control function of inverter at node i
$C_{q,i}$	Lipschitz constant for Volt-Var control function of inverter at node i
$\Re\{m\}$	Real part of a complex number m
$\boldsymbol{x} \odot \boldsymbol{y}$	Hadamard or elementwise product of $\boldsymbol{x}$ and $\boldsymbol{y}$
$\text{vec}(\boldsymbol{A})$	Column vector obtained by stacking the columns of $\boldsymbol{A}$ on top of one another
$\boldsymbol{e}_i$	i^{th} column of identity matrix
$\boldsymbol{x} \otimes \boldsymbol{y}$	Kronecker product of $\boldsymbol{x}$ and $\boldsymbol{y}$
$\text{diag}(\boldsymbol{x})$	Diagonal matrix whose entries are the elements of $\boldsymbol{x}$
$\text{diag}(\boldsymbol{X})$	Vector whose entries are the diagonal elements of $\boldsymbol{X}$
$\|\boldsymbol{x}\|$	2-norm of vector $\boldsymbol{x}$
$\|\boldsymbol{X}\|$	2-norm of matrix $\boldsymbol{X}$

CHAPTER 1

INTRODUCTION

1.1 Motivation

Power systems have evolved following a century-old paradigm following the unidirectional flow of electrons from large centralized generation plants through a transmission grid and then distribution lines with radial flows to load centers. This convention is being challenged by the development and diffusion of Distributed Energy Resources (DER) [Carvallo *et al.*(2019)]. Declining technology costs for DER assets such as solar photovoltaics (PV) and advanced Energy Storage Systems (ESS) — as well as financial incentives and innovative business models — are driving accelerated adoption of smaller, smarter, and cleaner energy systems. Existing centralized power plants powered by coal and nuclear that once benefited from traditional economies of scale are being retired by lower-cost distributed assets installed closer to load centers. This trend is unfolding on a global scale with Figure 1.1.1 showing the expected new DER capacity additions and centralized generation capacity additions on an annual basis.

However, the value of DER assets can only be fully realized if they are integrated into customer value systems, utility energy markets, and technical electron flows in a way that creates shared value and stability for all. Moving towards a system that requires seamless integration among participants in different levels compared to a single utility system increases network complexity. It raises a variety of challenges for grid management. Without direct ownership of many of the assets that integrate into their distribution and transmission networks, utilities focus on reinventing themselves best. There is an urgent need from the market for utilities to still play a key role

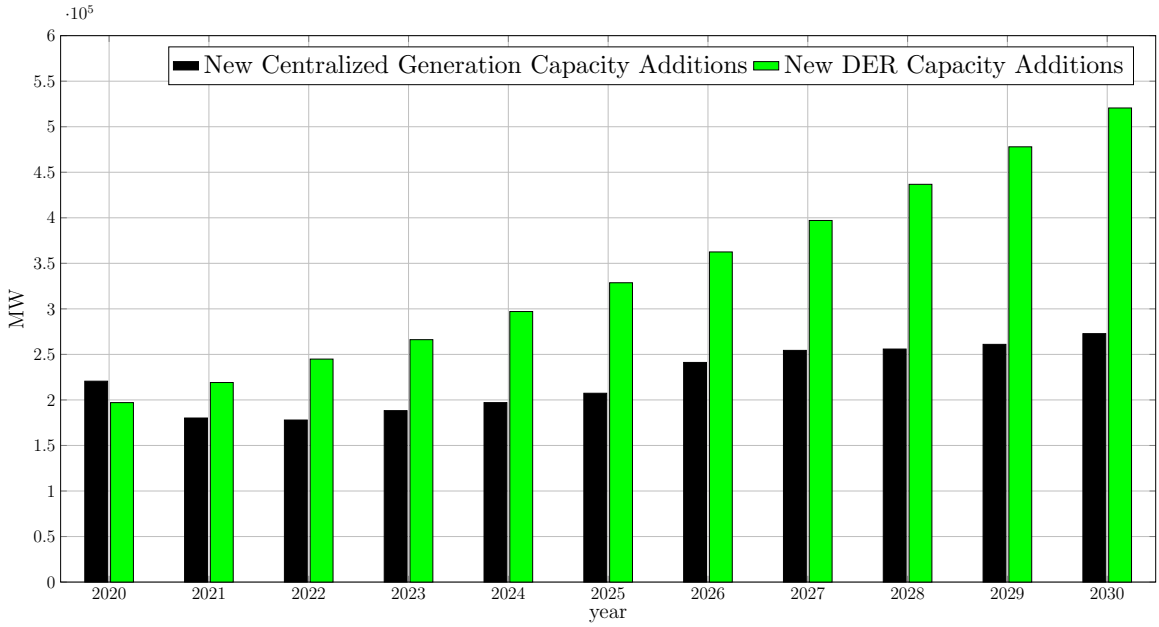


Figure 1.1.1: Global Annual Capacity Additions of Centralized and DER Assets From 2020 to 2030 (Recreated from [Metez *et al.*(2020)])

by using new tools to transform emerging DER concepts into a practical reality and maintain the system's security and reliability. Some of these challenges in this process are:

1. Connecting generation capacity to local infrastructure without expensive network upgrades

2. Ensuring reliable communications between utilities, aggregators and DERs, and more accurately analyzing "technology mix, operation protocols, and consumer behavior"

3. Finding cost-effective pathways to ensure secure and stable integration of DERs with different communication protocols

4. Developing advanced techniques for managing modern distribution networks

These challenges hinder a stable, secured, resilient, and economic power distribution network with open and transparent access to all market participants. The next sections select and describe some of these challenges, prior works completed to address those challenges, and contributions made in this work to address additional gaps in literature and engineering applications.

1.2 Modeling of Distribution Test Cases

Power system test cases are commonly used to evaluate the efficacy and performance of new scientific techniques or translate knowledge through scholarly publications and education. A test case can be an (a) actual feeder model, (b) a generalized, derived, or extended version of an actual feeder, or (c) a completely synthetic feeder with various levels of abstraction to represent actual feeder characteristics. Blended approaches of (b) and (c) can also rapidly generate thousands of new test cases by varying network topology, adjusting placement and parameters of apparatuses (e.g., adding new devices, removing old ones, retrofitting), and changing user energy use patterns based on recorded statistical data from actual utility feeders. Recent years have seen demand rise for realistic distribution feeder test cases due to the wider variety of distribution configurations and system states that now occur from integration of DER, ESS, Electric Vehicles (EV), load control, and microgrids [Liang *et al.*(2021)]. Developing realistic distribution test feeders is imperative to better capture increasing amounts of uncertainty and variability (e.g., changes in topology, variation in model parameters, user patterns). Updated and more representative models of distribution networks can then aid researchers in evaluating approaches to grid modernization research.

However, confidentiality and security policies governing critical infrastructure data can limit researchers from accessing real power systems data needed for scientific

development. For decades, researchers have extensively used a small set of standardized networks such as the IEEE transmission and distribution test cases [Schneider *et al.*(2009), Postigo Marcos *et al.*(2017)]. Recent research has developed additional test cases using algorithms that generate synthetic networks to mimic realistic power grids [Birchfield *et al.*(2017b), Gegner *et al.*(2016), Li *et al.*(2018), Abeysinghe *et al.*(2017),Schweitzer *et al.*(2017),Mateo Domingo *et al.*(2011),Krishnan *et al.*(2017)]. These open source works have provided more example networks for scientific exploration.

Development of synthetic transmission test cases (*ACTIVS*) presented in [Birchfield *et al.*(2017b)] begins with publicly available data including population, measured energy use, and zip codes. Synthetic substations are placed geographically in each zip code with transmission lines connecting substations. Statistical validation of these models using typical topological criteria is provided in [Birchfield *et al.*(2017b),Birchfield *et al.*(2017a)].

Development of synthetic radial distribution feeders in [Schweitzer *et al.*(2017)] views the distribution system as a random graph with nodes and edges that can be imbued with properties of the graph following realistic statistics for the topology and parameters, derived from a large set of real feeders. The approach begins with an analysis step followed by a synthesis step. The analysis step identifies statistical distributions of properties such as load, node, degree or cable length using data from the Netherlands. The statistics are then input to a synthesis algorithm that creates feeders with similar topological and physical characteristics. Important here is that the statistical trends of voltage distributions and phase angles are maintained to make distribution feeders more realistic. Moreover, [Schweitzer *et al.*(2017)] provides a method for validating the results, using metrics of statistical similarity such as the Kulback Leibler distance of the trends of the real and synthetic cases.

Additional work in synthetic distribution feeders is based on the Reference Network Model (RNM) [Mateo Domingo *et al.*(2011)], a large-scale distribution planning tool to help regulators estimate planning and design costs and regulation applied to distribution companies. The synthetic test cases developed using RNM allow for the simultaneous planning of high-, medium-, and low-voltage networks using simultaneity factors and also plan cables in urban areas that take road layouts into consideration. An adaptation of RNM to create RNM-US in [Krishnan *et al.*(2017)] provides full-scale, high-quality synthetic distribution system datasets for testing distribution automation algorithms, distributed control approaches, and other emerging distribution technologies. Broad application of RNM is hindered by the significant amount of data required such as (a) geo-referenced transmission, substation, and consumer data, (b) load profiles, (c) equipment library for all power system components, (d) technical and economic parameters governing operation, and (e) environmental and topography data.

The work in [Liang *et al.*(2021)] also explored the application of machine learning-based techniques for generating synthetic distribution feeders. The method illustrated in [Liang *et al.*(2021)] developed a plug-and-play Generative Adversarial Network (GAN) based framework that can generate synthetic distribution feeders that is scalable and can be customized to the user's choice. However, the methodology relies on learning from training data on actual feeder models, which while generating more active models for a region, will also create synthetic cases that are biased to the input data and not be generally suitable for any geographical area.

1.2.1 Thesis Contributions to Synthetic Feeder Generation

Chapter 2 in this work therefore proposes a framework to generate *radial positive sequence* synthetic distribution feeders mapped to real geo-spatial topologies

using available OpenStreetMap data (a constantly expanding publicly available data source for road network) [Neis and Zipf(2012)]. The methods developed illustrate how to create individual synthetic distribution feeders, and groups of feeders across entire ZIP Code, with minimal input data for any location in the United States. Substation locations are selected from the *ACTIVS* models [Birchfield *et al.*(2017b)], a set of geo-embedded synthetic models of the US transmission system, to enable joint evaluation of synthetic transmission and distribution systems. US census population data is used to assign loads among the nodes in the distribution network. The developed framework also has the capability to provide the distribution feeder model in OpenDSS [Ribeiro *et al.*(2020)] format allowing simulation capabilities like hosting capacity calculation [Bollen and Rönnberg(2017)], complex infrastructure network simulation [Brase and Brown(2009)], and more.

1.3 Operation of Modern Distribution Networks with High Penetration Renewables

Increasing amounts of DER can affect power quality, distribution network stability, overall system balancing, protection strategies, and failure modes and reliability. Essential to understanding and mitigating these challenges are distribution network State Estimation (SE) and Optimal Power Flow (OPF). SE is a data processing algorithm for converting meter readings and other available information into an estimate of the state of an electric power system [Primadianto and Lu(2017)]. OPF problems seek to control the generation/consumption of generators/loads to optimize cost or power loss in the network.

Both problem formulations are increasing in importance for distribution networks due to the advent of distributed generation and controllable loads. While these problems rely on different set of inputs and solve a different objective, they both rely on solving non-convex power flow physical laws as constraints. There are in general three

ways [Gan *et al.*(2015)] to approach this challenge:

1. Linearize the power flow laws: Power flow constraints can be approximated by some linear constraints in transmission networks due to small line losses (high X/R ratio), small voltage angle difference between adjacent buses, and tendency of voltages to stay close to nominal value. Approximation of the constraints allows the OPF problem to reduce to linear program [Stott and Alsac(1974), Stott *et al.*(2009)]. This method is widely used in practice for transmission networks SE and OPF, but does not apply to distribution networks due to its low X/R ratio.

2. Search for local optima : Many nonlinear algorithms that seek a local optimum of have also been developed to avoid these shortcomings. Representative algorithms include successive linear/quadratic programming [Contaxis *et al.*(1986)], trust-region based methods [Sousa *et al.*(2011)], Lagrangian Newton method [Baptista *et al.*(2005)], and interior-point methods [Quintana *et al.*(2000)]. Some of them, especially those based on Newton-Raphson, are quite successful empirically [Gan *et al.*(2015)]. However, when they converge, these algorithms converge to a local minimum without assurance on the suboptimality gap.

3. Convexify power flow laws : Shortcomings of the prior two are resolved using convexification methods proposed in [Lavaei and Low(2012), Low(2014a), Low(2014b)] to transform the nonconvex power flow laws into linear constraints on a positive semi-definite rank-1 matrix, and then remove the rank-one constraint to obtain a Semi Definite Programming (SDP) relaxation or Second-Order Cone Programming (SOCP). When applied to radial systems, such techniques offer an exact solution for AC SE and OPF problems for distribution networks.

One obstacle to applying SDP methods to solve AC OPF in distribution systems is that the number of variables in the relaxed problem increases to $O(n^2)$ for n buses. In the context of SE, the problem is also scalability and the number of measurements required by AC SE using SDP relaxations. For example, authors in [Zhang *et al.*(2018)] developed a conic relaxation method based on a restrictive assumption that voltage magnitude measurements were available and known for all buses, and that knowledge also existed for one active power flow per line of a spanning tree when solving a single-phase balanced electrical network. Authors in [Madani *et al.*(2017)] introduced a penalized conic problem for SE that, while being robust to measurement noise, also requires voltage magnitude and apparent power demand measurements from all buses and the apparent power flow at all branches. An alternative method in [Madani *et al.*(2016)] used a convex relaxation-based technique with measurements that included zero injection buses (equivalent to virtual measurements) and measurements from all PV and PQ buses. However, using virtual and pseudo measurements to compensate for missing data requires appropriate weights to avoid solving an ill-conditioned SE problem [Primadianto and Lu(2017)]. Moreover, distribution networks are rarely equipped with Advanced Metering Infrastructure (AMI) sensors or Phasor Measurement Unit (PMU) at every bus [Zhand *et al.*(2019), Donti *et al.*(2020), Liu *et al.*(2019), Sagan *et al.*(2021)]; therefore, measurements are too sparse to ensure observability (i.e., the number of measurements are equal to or exceed the number of unknowns).

Another line of research proposed a matrix completion formulation to solve SE problems with sparse measurements, relying on the assumption that voltage phasors in space and time are in a low-dimensional space. The formulation can be centralized [Donti *et al.*(2020)] or decentralized [Sagan *et al.*(2021)]. However, matrix completion methods typically interpolate to resolve missing measurements, primarily

because sampling is asynchronous, and therefore the method is limited to estimating gaps of the measured quantities. Moreover, the efficacy of matrix completion relies on the appropriate weighting of error tolerance in the objective function and constraints. Limitations of this approach are addressed using machine learning techniques such as Deep Generative Adversarial Network [Mestav and Tong(2019)], Multi-layer Perceptron Model [Mestav *et al.*(2019)], and others. These methodologies rely on training a neural network to approximate the Bayesian Minimum Mean Square Error (MMSE) estimate of missing or unavailable measurements by simulating several training scenarios for a known topology with demand profiles realistic for the specific network. While it is possible to obtain accurate estimation results, training the network requires a significant amount of historical data that is generally unavailable. Similar machine learning based approaches are discussed in [Hasan *et al.*(2020)] for solving OPF problems, which are again subjected to training scenarios.

1.3.1 Thesis Contributions to AC SE and OPF Formulation

Chapter 3 proposes a Grid-GSP based framework for relaxing bi-linear power flow constraints. Prior works have demonstrated GSP applications for cyber and physical stress detection [Hasnat and Rahnamay-Naeini(2020)], identification of false data injection attack on electric transmission networks [Drayer and Routtenberg(2020)], community detection [Ramakrishna and Scaglione(2019b)], and more. By integrating GSP, this work thus addresses the issues of conventional SDP approaches: (1) limited ability to perform SE due to scarce measurements. (2) the scalability of both SE and OPF problems (3) the inability of matrix completion to go beyond the measured quantities (4) the dependence on training data when using neural networks for interpolating missing measurements.

1.4 Prevention of Cyber-Attack on an Open Energy Market

The proliferation of DER, EV, ESS, and networked grid-edge devices requires extensive data exchange and communication network protocols to develop secure and stable monitoring and control of the underlying assets to ensure resilient Cyber Physical Systems (CPS). These networks' cyber vulnerabilities open the door to cyber attacks that can be highly destructive to power system operations. Power distribution systems are more susceptible to attack because their industrial control systems allow remote access and connection to customer networks. As a result, hackers have several tactics to tap into those systems and potentially disrupt operations. While a cyber-attack on a distribution system may be less significant than one on the bulk power system, the impacts of such an attack could still result in outages of national significance [United States Government Accountability Office(2021)].

Cyber-attacks such as the 2010 Stuxnet attack on Iran's nuclear plant at Natanz [Zetter(2016)], the 2015 BlackEnergy virus attack on Ukraine's three energy companies [SANS Industrial Control Systems(2016)], and the TRITON attack on safety instrumented system in Saudi Arabia [sau(2017)] illustrate that hackers have been creating malicious code that targets Operational Technology (OT). All three being triggered by malware demonstrate the need to understand the differences and connections between Information Technology (IT) and Operation Technology (OT).

IT is common throughout all industries with many moving parts and gateways, making it highly vulnerable and offering a large surface for a wide variety of constantly evolving attacks in the virtual world. Defending against attacks requires safeguarding every layer, and continuously identifying and correcting weaknesses. OT, in contrast, belongs to the physical world to provide correct execution of actions. While IT has to safeguard every layer of the system, OT controls systems that may be on or off, closed

or open. OT focuses on physically moving and controlling devices and processes, with a primary focus on reliability and increased efficiency.

Conventional boundaries between IT and OT are blurring due to the emergence of the Internet of Things (IoT) and the integration of physical machines with networked sensors and software. OT is becoming increasingly accessible, with threat vectors now extending to base-level assets or grid-edge devices such as smart thermostats. As more and more objects connect, communicate and interact with each other, there has been a surge in the number of endpoints and potential ways for cyber-criminals to gain access to networks and infrastructure systems. It is therefore essential to increase focus of security threats during the initial design and development phase of the cyber infrastructure, emphasizing the importance of developing robust preventive measures, fast detection methods, and resilient mitigation frameworks to ensure cyber-secure power distribution systems.

Developing efficient and robust preventive measures can pave the path to develop a trustworthy open energy trading platform for market participants – i.e., a Transactive Energy (TE) framework. TE combines financial signals and dynamic control techniques to shift the timing and quantity of energy usage to achieve greater efficiency, increased use of renewable energy, reduced energy costs, and improved flexibility to manage shifts in net load locally. Such benefits have motivated the increasing body of research whose goal is to manage real-time demand and electricity supply in an open market where prosumers and utilities interact to establish a market clearing price. Examples are auction mechanisms proposed in [Lin *et al.*(2019)], algorithms for co-simulation of transmission and distribution networks [Nguyen *et al.*(2019)], multi-agent models capturing trading behaviors [Janko and Johnson(2018)], and thermostatically controlled loads to participate in TE markets [Behboodi *et al.*(2018)], to name a few examples.

Recently, many researchers have purported blockchain as the ideal enabling platform to implement TE. Blockchain is well-suited to enable a transactive network because grid-edge devices come from a broad matrix of vendors and types, with units that can be added or removed from the network at any time, use rule sets that may be updated frequently, and lack a centralized authority for management and control. It allows non-trusting market participants to trust each through utilizing a common immutable transaction record validated by several peers. It offers a flexible, low cost, and secure means to implement logistics and tracking architecture to manage digital assets and distributed devices [Gorog and Boult(2018)]. This feature addresses one of the most critical issues tied to the security of interconnecting grid-edge devices with critical electricity infrastructure. Blockchain can enhance cyber-security and traceability of Peer-to-Peer (P2P) transactions between mutually non-trusting parties in the TE marketplace [Leeuwen *et al.*(2020)]. There are several benefits to such an implementation: 1) once stored on the ledger, all transactions are transparent to all participants through an identical copy of the ledger, 2) new transactions are *hash-chained* when appended to the ledger, an operation that makes them immutable, mitigating cyber-attacks aimed at reducing the integrity and availability of the data, and 3) all functional aspects of TE enabled by blockchain, from bidding to pricing to billing, can be orchestrated running *Smart Contracts* [Mohanta *et al.*(2018)]. Figure 1.4.1 represents the interactions tied to the TE application layer for a generic blockchain based cyber-physical infrastructure implementation.

Even though blockchain ensures transparency and immutability of bidding and trading records in the ledger. Records in the ledger have security gaps during the submission of bids and the verification of contractual obligations. In fact:

1. Threats exist internal to TE approaches because selfish players have an intrinsic incentive to cheat on reported consumption needs or production capacity during

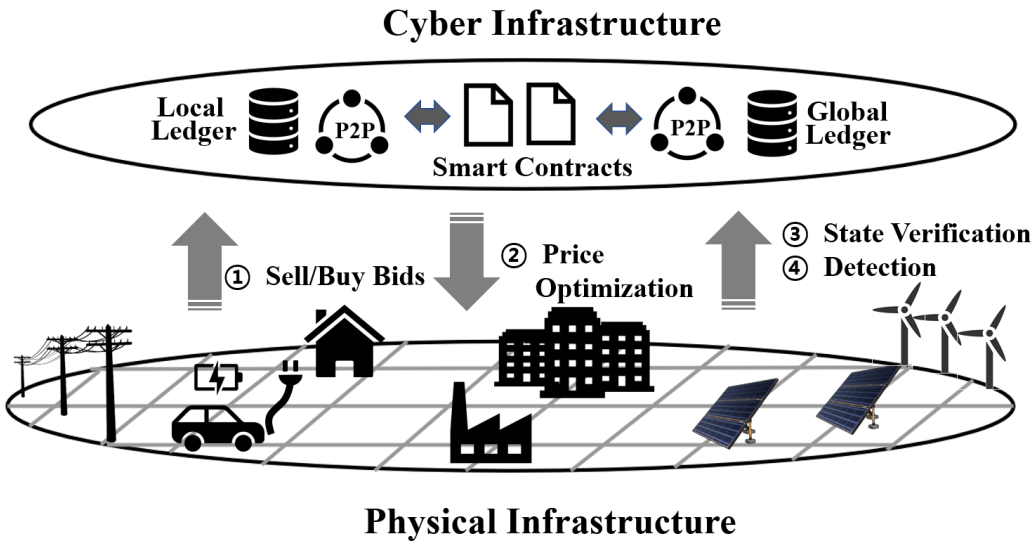


Figure 1.4.1: Conceptual architecture of EVE illustrating cyber and physical layers.

market clearing [Hussain *et al.*(2019)].

2. Ex-post, if the market stakeholders control smart meters, they can inject false data to hide discrepancies that would otherwise reveal cheating.

There is a substantial body of research and industrial efforts in TE; and there is a relatively recent trend that includes blockchain as the backbone for managing P2P communications for market-related operations. These operations involve handling and securing prosumer bids and dispatch values, deciding a market-clearing algorithm, ensuring the balance between demand and supply to meet network constraints, ensuring cyber-security, and more. Readers are referred to [Musleh *et al.*(2019)] for a broad discussion on the applications of blockchain for smart grid cyber-physical infrastructure, to [Mylrea and Gourisetti(2017)] for potential benefits of blockchain for grid resiliency against cyber-attacks, and to [Wu *et al.*(2021)] for details on how smart inverter, advanced metering infrastructure, and energy coordinator can support the digitization and decentralization of TE.

The existing literature has focused on co-simulation of the physical grid in tandem

with blockchain TE implementation [Hayes *et al.*(2020)], quantifying energy losses caused by energy transactions in an energy blockchain [Sanseverino *et al.*(2017)], enforcing proportional fairness among DERs participating in voltage control through smart contracts to ensure voltage stability [Danzi *et al.*(2017)], developing a blockchain based energy trading platform for electric vehicles [Silva *et al.*(2019)], integrating energy and carbon markets through a blockchain based trading framework [Hua *et al.*(2020)], and more. The two-layer blockchain implementation in [Lu *et al.*(2019)] consists of a first layer with smart meters forming a private blockchain and a second layer with aggregators forming a consortium blockchain to coordinate energy transactions within regions. These works assume security is implicit in transactions between aggregators and flexible demand assets. The authors in [Zhao *et al.*(2018)] use a continuous double auction mechanism, but the transaction mechanism exposed the prosumer's unique ID, posing a privacy concern. Similar concerns described in [Agung and Handayani(2020)] indicate a public blockchain exposes the transactions and balances of each prosumer, and further, that the rate of transaction processing limits scalability. Authors in [Kang *et al.*(2017)] present a double auction mechanism for localized energy exchanges between EVs, where local aggregators publicly audit and share transaction records without relying on a trusted third party. The use of local aggregators also appears in [Ferreira and Martins(2018)] to create an energy market that combines blockchain and IoT for two flexible community market players: an EV community and a DER community. In [Münsing *et al.*(2017)], the authors propose a blockchain implementation to clear the market using the Alternative Direction Method of Multipliers (ADMM) in a master-slave distributed architecture, where a central aggregator/master node updates global variables.

The multi-layer smart contract implementation based on Ethereum in [Danzi *et al.*(2018)] addresses the mismatch in the settlement between *System Operators*

and *Balance Responsible Parties*, yet it does not decentralize the Independent System Operator (ISO) nor include verification. Work done in [Jogunola *et al.*(2019)] uses Hyperledger Composer to demonstrate a blockchain based TE market implementation while excluding a market clearing price calculation and assuming energy transactions are automatically verified. A prototype implementation of a TE market using Hyperledger Fabric for metering and billing purposes is proposed in [Gür *et al.*(2019)], yet pricing and verification were not addressed. A related study that uses Hyperledger Composer in [Pipattanasomporn *et al.*(2018)] determines the market clearing price by averaging bid prices offered by all buyers while sorting sellers by first-in, first-out basis. An approach similar to [Pipattanasomporn *et al.*(2018)] using an Ethereum based blockchain architecture is found in [Christidis *et al.*(2021)]. However, such algorithms can be easily exploited by malicious prosumers or attackers to manipulate the clearing price and destabilize the TE market. Malicious prosumers can similarly influence the co-simulation framework presented in [Coignard *et al.*(2018)] where the ratio of total generation and consumption reported by the prosumers is used to determine the price.

The private blockchain solution proposed in [Che *et al.*(2019)] requires a match between the energy producer and consumer regarding the amount of power to be generated and consumed, respectively, which is not practical for many prosumers and can violate physical constraints on the distribution network. The blockchain based energy trading model in [Zhang and Shi(2020)] also lacks sufficient protection from physical constraint violations because the approach allows for an open trading platform across diverse types of power sellers without any optimized market pricing. In most studies, Smart Contracts orchestrate information exchanges among participants and during recording transactions, while still requiring a central entity to be in charge of calculating the market-clearing price in contrast to our fully decentralized solu-

tion. Also, prior methods focusing on security aspects and countermeasures against cyber-attacks to the market-clearing mechanism have ignored physical verification that must be tied to market records to work effectively as a continuous deterrent to theft. Relevant to this study are also prior works integrating smart metering with blockchain, specifically energy trading applications [Pop *et al.*(2018)]. Even if they leverage the immutability of blockchain, these approaches leave data integrity and privacy concerns unresolved [Andoni *et al.*(2019)].

Any TE framework for the market stakeholders based on HLF includes its inherent security features, however there are still following unresolved security issues:

1. HLF uses the JSON Web Token (JWT) authentication method based on claim token [Jones *et al.*(2015)]. However, BASE64 is used as the encoding method without data encryption in claim tokens, allowing malicious users to collect sensitive information by eavesdropping on access tokens of genuine users [Park *et al.*(2018)].

2. HLF provides security credentials for prosumers, but the grid-edge devices owned by prosumers can still be subject to attacks such as physical attacks, communications disruptions, and firmware reprogramming. An excellent example is discussed in [spe(2015)] in which a smart inverter vendor remotely updated control settings for 800,000 inverters in a single day in Hawaii. Were an adversary able to penetrate the back-end system in the same fashion, or hijack the update, that attacker could manipulate power flows or disable or damage equipment, and perhaps cause a localized or grid-wide blackout.

1.4.1 Thesis Contributions on EVE Transactive Energy Framework

Chapter 4 addresses the gaps in the existing literature for TE through developing the **Electron Volt Exchange (EVE)** blockchain architecture that includes a distributed pricing and distributed robust state verification mechanism for TE transactions. Implementation of the EVE blockchain framework onto a distributed ledger [Androulaki *et al.*(2018)] is completed using the open source Hyperledger Fabric (HLF) framework.

Chapter 5 describes how to integrate these concepts and algorithms into hardware security via a Cryptographic Trust Center^{TM} (CTC^{TM}) chip. Design considerations for such a hardware chip are provided in [Saleem *et al.*(2020)] to protect grid-edge devices; however, the solution is expensive and prevents mass scaling. By expanding the ideas of Two Factor Authentication (2FA) [Park *et al.*(2018), Alharbi and Alghazzawi(2019)] and Hardware Root of Trust [Casper and Papa(2011)], the proposed solution integrates hardware security in conjunction with existing security features of the EVE framework.

1.5 Mitigation of Generation Intermittency and Cyber-Attacks on Power Distribution Systems

The fast-acting and complex control mechanisms of DER pose challenges for planning, operations, and reliability of electric distribution systems and microgrids with increasing amounts of DER. Poor DER coordination can create reliability issues even at low penetration levels [Khalili *et al.*(2019)]. At higher penetrations, the net impact of many solar PV generators may accumulate and further affect power quality [Hu *et al.*(2019), Singhal *et al.*(2019)]. Power quality control devices could be added to the distribution network to counteract the problem, but the additional cost can be

avoided if the problem can be resolved directly by solar inverters. One such approach involves using PV inverter capacity to control real and reactive power (generation and consumption) as a means to stabilize voltage swings occurring from changes in load, generation, equipment failure, or cyber-attack. This functionality can be implemented within the allowable scope of interconnection standards [IEEE Standards Coordinating Committee 21(2018), rul(2019)] using only local information at the inverter without the need for additional data monitoring and control equipment to be installed on the network.

Numerous works have studied power quality issues for inverter-dominated distribution networks and recommended mitigation strategies for over-voltage, voltage oscillation, and other issues. Methods to perform voltage control by PV inverters can be categorized by the control approach (centralized or local), control function (Volt-Var, Volt-Watt, combined Volt-Var and Volt-Watt, and none) [Vijayan *et al.*(2019)], and level of detail to which inverter hardware limits are expressed in the control function.

In a centralized approach, a utility or aggregator collects real-time information about the network, processes that information using an optimization formulation (commonly an OPF formulation), and sends updated set-points back to the inverters. Such an approach was taken in [Baker *et al.*(2018)] to determine inverter settings that ensured voltage stability across the entire distribution network, but that approach required the centralized utility to exchange information with inverters and this restricted updates to every 5-15 minutes. Similar approaches could be suitable for smaller networks with fewer inverters and more frequent communication. The authors in [Weckx *et al.*(2014)] provided a centralized optimization formulation with set-points updated at similar timescales to optimize local control curves modeled using a first-order spline model. This approach was well-suited for managing steady-

state voltage levels but posed challenges to resolving fast-acting voltage oscillations occurring from cloud intermittency [Zhu and Liu(2016)]. By incorporating load tap changers and smart inverters, authors in [Long *et al.*(2019)] developed a centralized control method for Volt-Var optimization using sensitivity factors that assumed a constant slope for the Volt-Var controllers of the smart inverters. This approach required recalculation of sensitivity factors by the utility and extensive communication to the smart inverters. The issues created by communication delay in a centralized control system or an equivalent master-slave control system are also described in [Muthukaruppan and Baran(2020)]. A master-slave Volt-Var optimization method was presented in [Shi and Baran(2019)] that minimized real power losses and grouped inverters by spatial distance but did not incorporate the Volt-Var control curves of the smart inverters. Another centralized approach presented in [Ghasemi and Parniani(2016)] showed how to manage voltage by curtailing active power but had the drawback of requiring simultaneous control of all inverters, a challenge given the time delay in communication and inverter control action. The issues associated with communication and synchronization can be avoided using local control which reduces information exchange and is faster to execute.

Among local or decentralized approaches, droop-based voltage control [IEEE Standards Coordinating Committee 21(2018), rul(2019)] is the most common framework from literature and utility practices [Farivar *et al.*(2013), Farivar *et al.*(2015), Zhou *et al.*(2016),Zhou *et al.*(2015),Singhal *et al.*(2019),Jahangiri and Aliprantis(2013),Zhu and Liu(2016),Helou *et al.*(2020)]. However, improper selection of control parameters can lead to control instability and voltage oscillation [Farivar *et al.*(2013), Jahangiri and Aliprantis(2013)]. Work in [Jahangiri and Aliprantis(2013)] highlighted instability concerns and proposed "delayed droop control" to maintain voltages within acceptable bounds by absorbing or supplying reactive power. But as pointed out in [Singhal

et al.(2019)], the method in [Jahangiri and Aliprantis(2013)] cannot adapt to changing operating conditions and external disturbances that may affect control algorithm convergence. The challenge of convergence was addressed in [Singhal *et al.*(2019)] but no generalized method was proposed to choose correction factors (k_i^d; Δv_f) to manage the shifting of the control curves during unstable operating conditions. Results presented in [Pukhrem *et al.*(2017)] showed that a limiter algorithm can mitigate over-voltage issues, but no analytical model was provided to ensure that the algorithm could address voltage oscillations. Authors in [Braslavsky *et al.*(2018)] conducted an analytical stability analysis for a single inverter connected to a strong source (infinite bus/substation) with potential extensions to an inverter dominated distribution network. A related study completed stability analysis for a distribution network [Heidari *et al.*(2018)] using an adaptive decentralized control scheme by assuming an equal ratio of reactance to resistance for all branches and that inverters at load buses could not be controlled. Such assumptions cannot be generally applied to any distribution network. Works in [Sadnan and Dubey(2020)] provided the design of a distributed voltage controller for faster tracking of DER intermittency, however, the communication requirement among agents can still be vulnerable in the event of a cyber-attack.

Inverter control functions are another area of study. Authors in [Pompodakis *et al.*(2016)] introduced reactive power correction methods to reduce over-voltage issues within a low voltage radial grid without incorporating any Volt-Var or Volt-Watt droop control. Research in Volt-Var control has shown ways to achieve voltage stability for smart inverters [Farivar *et al.*(2015), Zhou *et al.*(2016), Zhou *et al.*(2015), Zhu and Liu(2016),Singhal *et al.*(2019),Jahangiri and Aliprantis(2013)]. More recently, interconnection standards [IEEE Standards Coordinating Committee 21(2018),rul(2019)] require smart inverters to provide both Volt-Var and Volt-Watt control with ongoing research focusing on how to employ the combined approach to provide voltage

support and improve reliability [Baker *et al.*(2018), Heidari *et al.*(2018), Ghasemi and Parniani(2016), Olivier *et al.*(2016), Huque(2015)].

Modeling inverter capability presented in [Baker *et al.*(2018), Farivar *et al.*(2013), Farivar *et al.*(2015), Zhou *et al.*(2016), Zhou *et al.*(2015), Jahangiri and Aliprantis(2013), Zhu and Liu(2016), Helou *et al.*(2020), Singhal *et al.*(2019)] often assume the smart inverter can operate anywhere in the power circle . This allows an inverter to set reactive power generation equal to the apparent power rating of the inverter when there is no real power generation. Such behavior violates hardware limits on maximum reactive power generation that can be lower than the apparent power rating of the inverter [Hu *et al.*(2019), Sol(2018), Sarfaraz *et al.*(2016), Yue Yuan *et al.*(2007)].

1.5.1 Thesis Contributions to Mitigate Cyber-Attack Effects on Distribution Systems

Chapter 6 addresses the gaps in the existing works by using piece-wise linear droop models of Volt-Var and Volt-Watt control functions to describe inverter behavior where the functions are expressed as Lipschitz functions. Unlike the modeling techniques presented in existing works, this work includes limitations on reactive power generation that account for physical hardware while deriving the Lipschitz constants. The formulation and derived constants are applicable for all smart inverters with piece-wise linear droop models, and can be used to derive constants for an inverter with reactive power generation capability equal to the apparent power rating, a simplifying assumption common to approaches in prior literature. This work also analytically derives stability policies and provides an approach to reduce voltage oscillations and stabilize the network-wide voltage profile using only local information and control actions by using Lyapunov analysis.

1.6 Dissertation Organization

This dissertation is organized into the following 7 chapters:

- **Chapter 1 - Introduction** : Provides an introduction to distribution networks, integration of distributed energy resources and grid-edge devices, and challenges in creating a robust, stable and secure power distribution networks and prior works to address those issues. A brief description of the work completed in Chapters 2 through 6 are also presented.

- **Chapter 2 - A Framework for Generating Synthetic Distribution Feeders using OpenStreetMap** : This chapter describes the generation of synthetic distribution feeders that facilitate power systems research and development by providing thousands of realistic use cases. Details of the developed algorithm are provided with examples for single-phase positive sequence feeders, with discussion of how the approach can be extended to three-phase unbalanced models.

- **Chapter 3 - Convex Relaxation For State Estimation and Optimal Power Flow using Graph Signal Processing** : This chapter revisits the application of SDP and proposes solutions for AC SE and AC OPF. Inclusion of Grid-GSP allows application of the low-rank representation of the voltage phasor vector in power distribution systems is illustrated that also provides an opportunity for finding optimal sampling patterns that significantly reduce the number of measurements needed for reconstruction and provides guidelines on the optimum sampling pattern (i.e., optimumal placement of sensors).

- **Chapter 4 - A Secure Distributed Ledger for Transactive Energy: The Electron Volt Exchange (EVE) Blockchain** : This chapter describes the

EVE framework and an aggregator-based market model with distributed pricing and a state verification method that is scalable and robust against False Data Injection Attacks (FDIA). Development of the framework using HLF and smart contracts with security features is discussed.

- **Chapter 5 - Integrating Hardware Security into Electron Volt Exchange Platform** : This chapter demonstrates the application of CTC^{TM} to the EVE platform for provisioning cryptographic keys and improving uniformity of the operational network and data management. The process of configuring, installing, and maintaining keys is described with symmetric and asymmetric verification techniques. Laboratory test results illustrate that the approach can resolve several cyber-security gaps in common blockchain frameworks such as Hyperledger Fabric.

- **Chapter 6 - Lyapunov Stability of Smart Inverters Using Linearized DistFlow Approximation** : This chapter provides the underlying mathematical description of the distribution network using a Matrix-based formulation of the linearized DistFlow equations. Proof of the stability criterion and the local control model developed using that stability criterion is described. Simulation results demonstrate how the approach resolves voltage oscillations from generation intermittency and cyber-attacks.

- **Chapter 7 - Discussions** : Results from Chapters 2 through 6 are comprehensively discussed. Benefits of the proposed endeavours to ensure distribution network stability, security and synthetic modeling are described and future research opportunities are discussed.

— PART I —

Generating Geo-Embedded Synthetic Distribution Feeders

CHAPTER 2

A FRAMEWORK FOR GENERATING SYNTHETIC DISTRIBUTION FEEDERS USING OPENSTREETMAP

The work proposed in this chapter uses the location of substations from recent efforts to develop synthetic transmission test cases (*ACTIVS*), with underlying real and reactive power in the distribution network assigned using population information gathered from United States 2010 Census block data. The methods presented illustrate how to create individual multi-phase radial synthetic distribution feeders, and groups of feeders across entire ZIP Code, with minimal input data for any location in the United States. The framework presented also has the capability to output data in OpenDSS format to allow further simulation and analysis.

2.1 Distribution Feeder Generation for a Single Substation

The ACTIVS cases [Birchfield *et al.*(2017b)] are based on a hierarchical clustering algorithm which groups postal codes to the set of substations $\mathcal{N}_s$, where $|\mathcal{N}_s| = N_l + N_g + N_b$, the number of substations containing only loads, only generators, and both respectively. Substations with *positive net real power* are used as the source for distribution feeders in this work.

The set of substations with positive net real power demand is defined as $\mathcal{N}_D$; where $\mathcal{N}_D \in \mathcal{N}_s$. The methods introduced are demonstrated for a substation (S) located at 33.3420° latitude and −111.6739° longitude (Mesa, Arizona) with a load demand of $40.47 + j11.14$ MVA. A simplified flowchart of the framework is shown in Fig. 2.1.1. Results from each step are presented alongside methods to illustrate the incremental process for creating synthetic distribution networks.

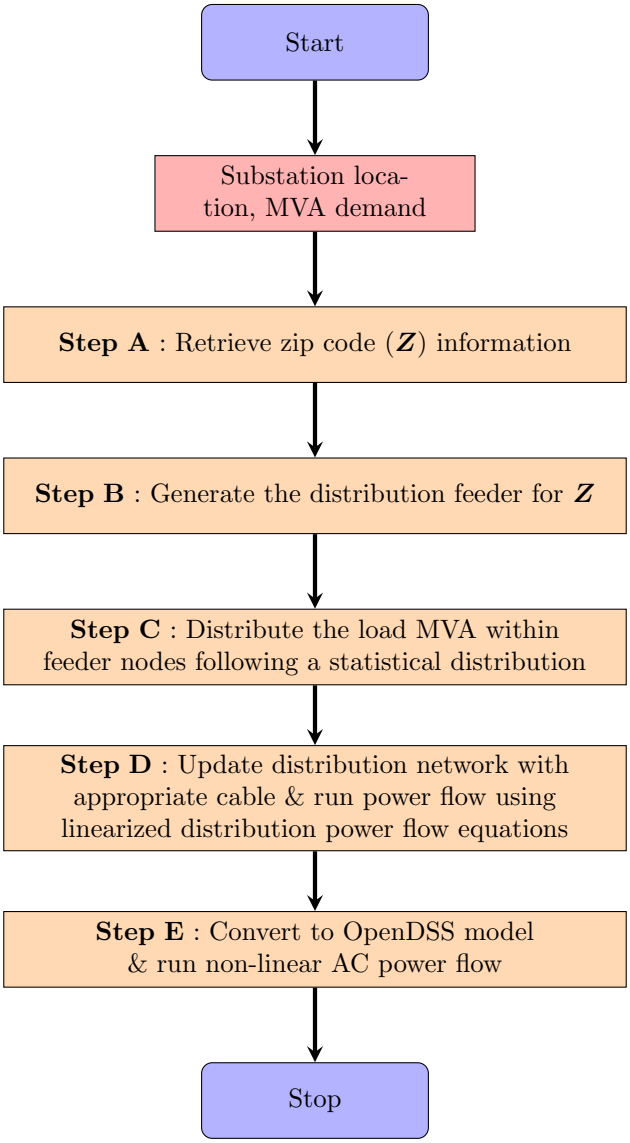
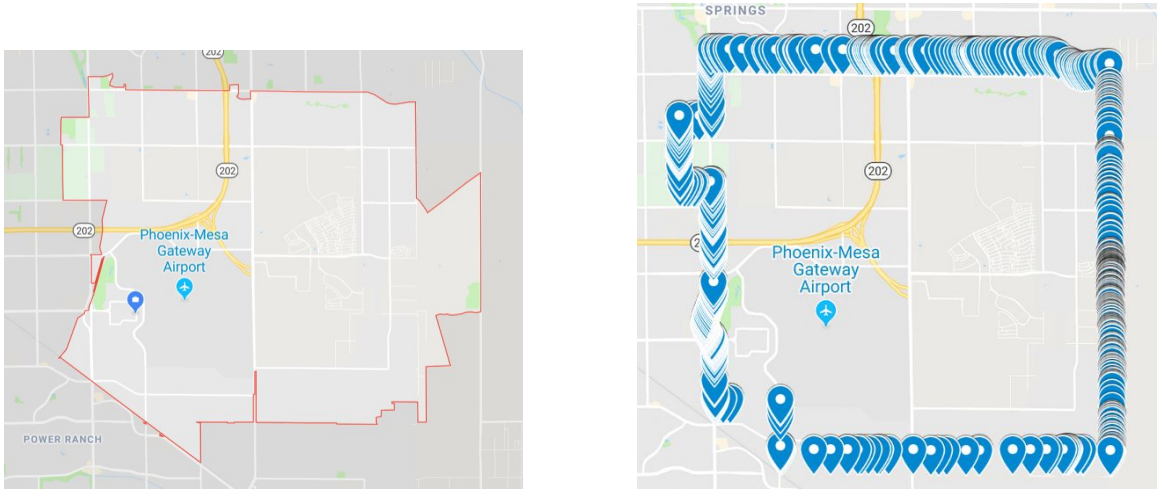


Figure 2.1.1: Flowchart Representing the Synthetic Distribution Feeder Generation Framework

2.1.1 *Finding the ZIP Code (Z) of substation S*

Location information (longitude, latitude) of the substation from *ACTIVS* cases is used to identify the corresponding ZIP Code (Z) from the "US 2010 Census Bureau ZIP Code" shapefile[1] [Sperling(2012)] or the Google Reverse Geocoding API [Kilic and Gülgen(2020)] if unavailable in the shapefile. If both processes fail, then the closest ZIP Code to substation S is assigned. The selected substation for demonstration belongs to the ZIP Code 85212. Fig. 2.1.2a and 2.1.2b illustrate the territory of ZIP Code 85212 from Google Maps and 2010 US Census Data respectively.



(a) (b)

Figure 2.1.2: (a) Overview of ZIP Code 85212 from Google Maps (b) Overview of Zip Code 85212 from 2010 Us Census Data

2.1.2 *Generating the Distribution Feeder Graph for Z*

1. A ZIP Code can be served by multiple substations. Using *ACTIVS*, the set of substation(s) ($\mathcal{S}_Z$) in Z are retrieved, giving $S \in \mathcal{S}_Z \in \mathcal{N}_D$. For ZIP Code 85212,

[1] A shapefile is a simple, non-topological format for storing the geometric location and attribute information of geographic features.

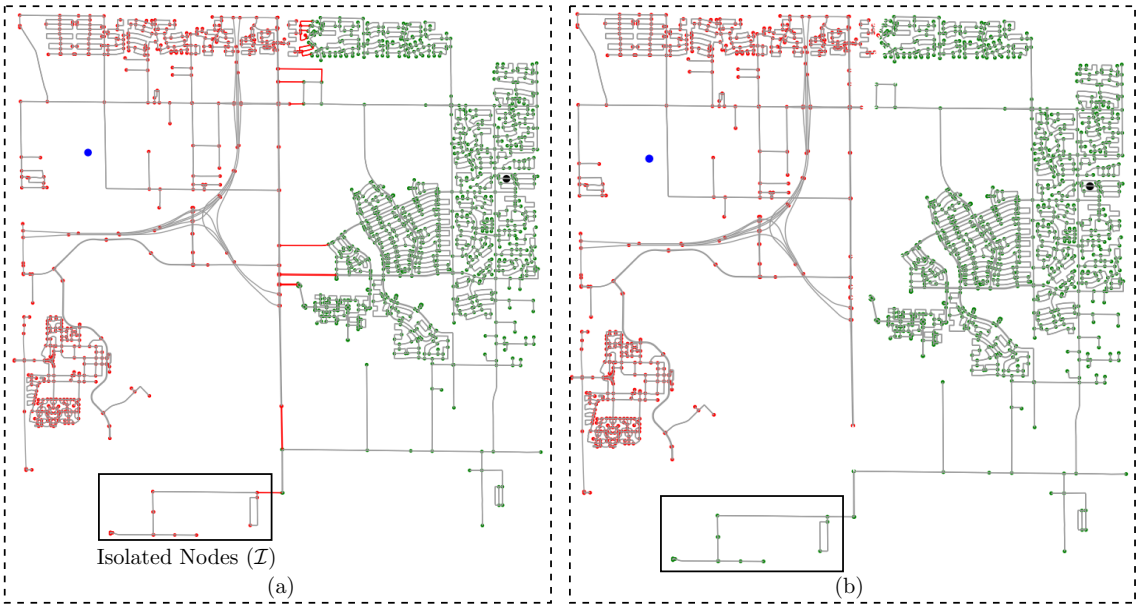


Figure 2.1.3: Creating Sub-graphs and Connecting Isolated Nodes to Appropriate Sub-graph

the *ACTIVS* case returns two substations with positive net loads, the second one to be located at $(33.3376°, -111.5900°)$ with a load demand of $55.07 + j8.51$ MVA.

2. The 'drive' network (drivable public streets excluding service roads) for Z is retrieved using [Boeing(2017)]. The retrieved network is a directional graph (G_Z) with self-loops and parallel edges.

3. G_Z is then split into Voronoi regions [Aurenhammer(1991)] where the number of Voronoi partitions $|N_Z|$ equals to number of substations in ZIP Code Z. Fig. 2.1.3a shows G_Z for the ZIP Code 85212 having two substations with positive net loads. The **blue** and **black** colored circles represent the location of substations in 85212. Fig. 2.1.3a also illustrates the two Voronoi regions corresponding to the two substations. The **red** and **green** nodes are distribution feeder nodes inside the **blue** and **black** substation's Voronoi regions, respectively. The edges marked in red in Fig. 2.1.3a are the edges connecting the two Voronoi regions.

4. $|\mathcal{N}_Z|$ number of sub-graphs (g_z) are then created from G_Z by splitting along the edges connecting the Voronoi regions. A sub-graph only includes the strongly connected nodes [Hong *et al.*(2013)]; hence the sub-graph creation process can create isolated nodes as shown in Fig. 2.1.3a. The isolated nodes either have degree 0 or are part of an isolated graph. Let $\mathcal{G}_Z^n$, $\mathcal{G}_z^n$ be the set of nodes for $\mathcal{G}_Z$, sub-graph $\mathcal{G}_z$ respectively while $\mathcal{I}$ is the set of isolated nodes. It follows that $\mathcal{I} = \mathcal{G}_Z^n \setminus \cup_{z=1}^{|\mathcal{N}_Z|}\mathcal{G}_z^n$, where $\cup_{z=1}^{|\mathcal{N}_Z|}\mathcal{G}_z^n$ represents the union of all the sub-graph nodes. Using Algorithm 1 the isolated nodes are connected to the appropriate sub-graph (g_z) as shown in Fig. 2.1.3b.

Algorithm 1 Algorithm for Connecting Isolated Nodes

for each node $i \in \mathcal{I}$ **do**
  find A_i where A_i is the set of nodes adjacent to node i in G_z
  find the k^{th} sub-graph $g_{z,k}$ for the first node a such that $a \in A_i \cap g_{z,k}^n$
  connect node i to node a and copy the node and edge properties from G_Z to $g_{z,k}^n$
  remove i from $\mathcal{I}$
end for

5. A minimum spanning tree using Kruskal's algorithm [Kruskal(1956)] is calculated for each g_z.

6. The census block data for Z is retrieved using US 2010 Census data. If $\mathcal{B}_Z$ is the set of census blocks for ZIP Code Z, there are $|\mathcal{B}_Z|$ census blocks within Z. There are 662 census blocks for ZIP code 85212, i.e. $|\mathcal{B}_Z| = 662$.

7. For every node in $\cup_z g_z$, the framework finds the census block that a node belongs to using the latitude and longitude information of the node and geometry of census blocks.

8. The framework then assigns a weight (p_w^n) to each node based on the population information of the census block the node belongs to. For instance, if node (n)

belongs to census block $b_z^i \in \mathcal{B}_Z$, $p_w^n = \dfrac{m(b_z^i)}{\sum_{i=1}^{|\mathcal{B}_Z|} m(b_z^i)}$, where $m(b_z^i)$ returns the population of the census block b_z^i. For 85212, total population, or $\sum_{i=1}^{|\mathcal{B}_Z|} m(b_z^i)$, is 25015.

The output of these steps is a set of sub-graphs $\{g_z\}$ with $|\mathcal{N}_Z|$ elements where sub-graph $g_{z,S}$ corresponds to substation S and $g_{z,S} \in \{g_z\}$.

2.1.3 Assigning Load to Sub-graph $g_{z,S}$

1. For the load assignment the node closest to the location of substation S is chosen to be the slack bus/ substation node for distribution feeder for the sub-graph/ feeder $g_{z,S}$. The big **red** circle in Fig. 2.1.4a indicates the substation node.

2. The substation demand is then distributed among the nodes of the sub-graph $g_{z,S}$. Zero load nodes are identified using the methodology in [Schweitzer *et al.*(2017)]. The substation node is designated to have zero load. The load distribution for each node under the sub-graph $g_{z,S}$ is equated using the following equations:

$$P(n) = P_N \left(\frac{1}{n} + \epsilon \right) \times p_w^n \tag{2.1.1a}$$

$$Q(n) = Q_N \left(\frac{1}{n} + \epsilon \right) \times p_w^n \tag{2.1.1b}$$

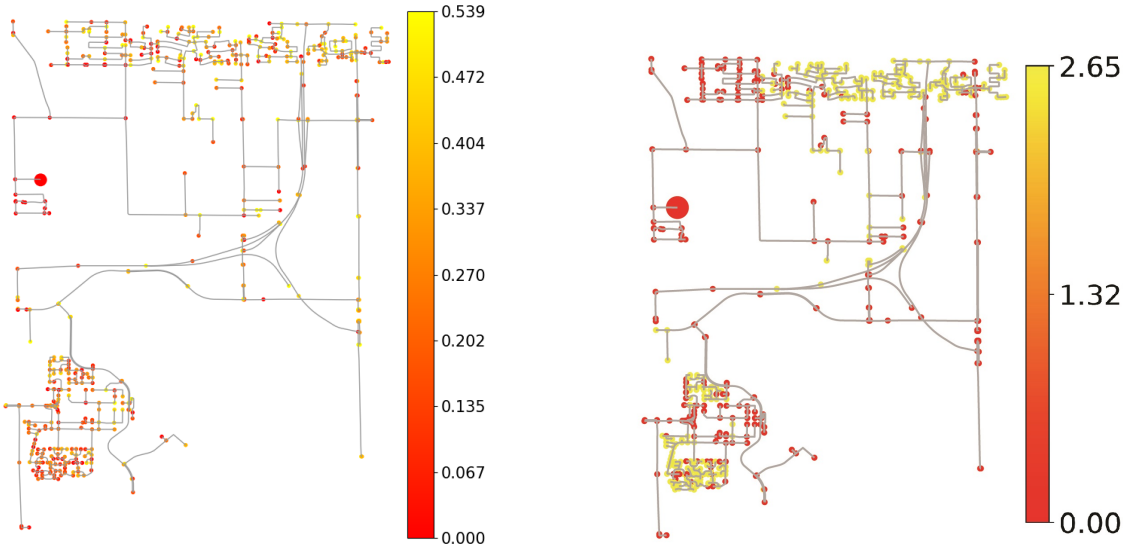
Here, P_S and Q_S are the real and reactive power load demand of substation S, respectively, and ϵ can be chosen from any distribution such as uniform distribution, or t-location scale distribution as in [Schweitzer *et al.*(2017)]. The final $P(n)$ and $Q(n)$ values are then scaled using the following equations to

match the total demand P_S and Q_S.

$$P(n) = P_N \times \frac{P(n)}{\sum P(n)} \tag{2.1.2a}$$

$$Q(n) = Q_N \times \frac{Q(n)}{\sum Q(n)} \tag{2.1.2b}$$

The developed framework does not necessarily require population information for modeling the feeder. Setting the value of p_w^n to 1 for all nodes will ignore the effect of population on the load distribution. Figure 2.1.4a and 2.1.4b shows the heat-map of real power of $g_{z,S}$ done following a t-location scale distribution while excluding and including the population information,respectively.



(a) (b)

Figure 2.1.4: (a) Distribution of Real Power (MW) among the Distribution Feeder Nodes (Population Information Excluded) (b) Distribution of Real Power (MW) among the Distribution Feeder Nodes (Population Information Included)

2.1.4 Steady State Voltage Profile for Sub-graph $g_{z,S}$

Having followed the previous steps, one has a radial graph $g_{z,S}$ of lines with known length, a substation node, and the real and reactive power demand at each node.

Bus 0 denotes the substation node or the point of common coupling (PCC), with a predefined reference voltage (v_0). Additional parameters including line resistance and reactance are required to run a steady state power flow. As the generated networks are radial networks prior works [Baran and Wu(1989a),Baran and Wu(1989b)] can be leveraged to create a matrix based linear formulation allowing a non-iterative power flow solution described Section 2.2. This also helps in developing Algorithm 2 to choose appropriate cables.

2.2 Linearized Branch Flow Equations

Assuming a balanced, radial system; for every line $(i, j) \in \mathcal{E}_e$ the DistFlow equations are :

$$P_{ij} = p_j^c - p_j^g + r_{ij}c_{ij}^2 + \sum_{k:(j,k)\in\mathcal{E}_e} P_{jk} \tag{2.2.1a}$$

$$Q_{ij} = q_j^c - q_j^g + x_{ij}c_{ij}^2 + \sum_{k:(j,k)\in\mathcal{E}_e} Q_{jk} \tag{2.2.1b}$$

$$v_j^2 - v_i^2 = -2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)c_{ij}^2 \tag{2.2.1c}$$

Representing vectors and matrices by boldface letters, the following vectors and matrices are defined:

$$\boldsymbol{v}_o^2 = (v_0^2, \ldots, v_{n-1}^2)^T, \quad \boldsymbol{v}^2 = (v_1^2, \ldots, v_{n-1}^2)^T \tag{2.2.2a}$$

$$\boldsymbol{p} = (p_0, \ldots, p_{n-1})^T, \quad \boldsymbol{q} = (q_0, \ldots, q_{n-1})^T \tag{2.2.2b}$$

$$\boldsymbol{P} = (P_{01}, \ldots)^T, \quad \boldsymbol{Q} = (Q_{01}, \ldots)^T \tag{2.2.2c}$$

$$\boldsymbol{c}^2 = (c_{01}^2, \ldots)^T, \quad \boldsymbol{r} = (r_{01}, \ldots)^T, \quad \boldsymbol{x} = (x_{01}, \ldots)^T \tag{2.2.2d}$$

$$[\boldsymbol{T}]_{e,i} = \begin{cases} 1 & i = t(e) \\ 0 & \text{else} \end{cases} \quad \forall\ e \in \mathcal{E}_e \tag{2.2.2e}$$

$$[\boldsymbol{F}]_{e,i} = \begin{cases} 1 & i = f(e) \\ 0 & \text{else} \end{cases} \quad \forall\ e \in \mathcal{E}_e \tag{2.2.2f}$$

$$\boldsymbol{M}_o = \boldsymbol{F} - \boldsymbol{T} \tag{2.2.2g}$$

Here, $\boldsymbol{T}$ and $\boldsymbol{F}$ are referred as the *to* and *from* matrices, respectively, and $\boldsymbol{M}_o$ represents the incidence matrix. Each are of size $|\mathcal{E}_e| \times |\mathcal{B}| = (n-1) \times n$. Also, $t(e)\text{and}f(e)$ are two functions that return the *to* and *from* node of an edge, respectively.

Note that $\boldsymbol{M}_o\mathbf{1} = 0$, and using the reference voltage v_0 (generally the voltage at the root of the feeder) a unique solution for the DistFlow equations can be found. Let $\boldsymbol{M}$ be the $(n-1) \times (n-1)$ matrix obtained removing the first column from $\boldsymbol{M}_o$, which is also invertible [Zhu and Liu(2016)]. Therefore:

$$\boldsymbol{M}_o(\boldsymbol{v}_o^2 - v_0^2\mathbf{1}_n) \equiv \boldsymbol{M}(\boldsymbol{v}^2 - v_0^2\mathbf{1}_{n-1})$$

Introducing the notation $\text{diag}(\boldsymbol{x})$ to indicate a diagonal matrix with the entries of vector $\boldsymbol{x}$ in its diagonal entries, the following matrices are defined:

$$\boldsymbol{R} = 2\boldsymbol{M}^{-1}\,\text{diag}(\boldsymbol{r})(\boldsymbol{I} - \boldsymbol{T}\boldsymbol{F}^T)^{-1}\boldsymbol{T} \tag{2.2.3a}$$

$$\boldsymbol{X} = 2\boldsymbol{M}^{-1}\,\text{diag}(\boldsymbol{x})(\boldsymbol{I} - \boldsymbol{T}\boldsymbol{F}^T)^{-1}\boldsymbol{T} \tag{2.2.3b}$$

$$\begin{aligned}\boldsymbol{L} = \boldsymbol{M}^{-1}\Big[&2\,\text{diag}(\boldsymbol{r})(\boldsymbol{I}-\boldsymbol{T}\boldsymbol{F}^T)^{-1}\,\text{diag}(\boldsymbol{r}) \\ &+ 2\,\text{diag}(\boldsymbol{x})(\boldsymbol{I}-\boldsymbol{T}\boldsymbol{F}^T)^{-1}\,\text{diag}(\boldsymbol{x}) - \text{diag}(\boldsymbol{r}^2 + \boldsymbol{x}^2)\Big]\end{aligned} \tag{2.2.3c}$$

The absolute diagonal values of $\boldsymbol{R}$ and $\boldsymbol{X}$ matrices represent the electrical distance of the *to* node of a branch from the substation node. Matrix $\boldsymbol{L}$ is associated with the ohmic losses of the network.

Using definitions in (2.2.2) and (2.2.3) and solving for $\boldsymbol{P}$ and $\boldsymbol{Q}$ permits the DistFlow equations from (2.2.1) to be written in terms of $\boldsymbol{v}^2$:

$$\boldsymbol{v}^2 = v_0^2\mathbf{1} + \boldsymbol{R}\boldsymbol{p} + \boldsymbol{X}\boldsymbol{q} + \boldsymbol{L}\boldsymbol{c}^2 \tag{2.2.4}$$

$$\boldsymbol{c}^2 = \text{diag}^{-1}(\boldsymbol{F}\boldsymbol{v}^2)(\boldsymbol{P}^2 + \boldsymbol{Q}^2) \tag{2.2.5}$$

Ignoring the losses associated with the term $\boldsymbol{L}\boldsymbol{c}^2$ as shown in [Baran and Wu(1989b), Farivar *et al.*(2013)], (2.2.4) becomes linear in terms of $\boldsymbol{v}^2$ and can be written as follows:

$$\boldsymbol{v}^2 = v_0^2\mathbf{1} + \boldsymbol{R}(\boldsymbol{p}^c - \boldsymbol{p}^g) + \boldsymbol{X}(\boldsymbol{q}^c - \boldsymbol{q}^g) \tag{2.2.6}$$

Here, the injections in load and generation are partitioned as $\boldsymbol{p} = \boldsymbol{p}^c - \boldsymbol{p}^g$ and $\boldsymbol{q} = \boldsymbol{q}^c - \boldsymbol{q}^g$, respectively. Hence, (2.2.6) can be written as follows:

$$\boldsymbol{v}^2 = \underbrace{v_o^2\mathbf{1} + \boldsymbol{Z}\boldsymbol{s}^c}_{\bar{\boldsymbol{v}}^2} - \boldsymbol{Z}\boldsymbol{s}^g \tag{2.2.7}$$

where:

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{X} \end{bmatrix}, \quad \boldsymbol{s}^c = \begin{bmatrix} \boldsymbol{p}^c \\ \boldsymbol{q}^c \end{bmatrix}, \quad \boldsymbol{s}^g = \begin{bmatrix} \boldsymbol{p}^g \\ \boldsymbol{q}^g \end{bmatrix} \tag{2.2.8}$$

Equation (2.2.1a)-(2.2.1b) can be used to calculate the vector of real and reactive power flows for every line using the nodal demand values. By using a predefined cable database (that includes impedance and apparent power (MVA) carrying capacity information) and the calculated line flows, the framework then follows Algorithm 2 to find the appropriate cable that can meet the power flow through a line with minimum number of parallels. Assigning cables allows the calculation of r_{ij} and x_{ij} for all lines, and equation (2.2.1c) is then used to calculate nodal voltages using $v_0 = 1.00$.

If the minimum of the node voltages is less than a predefined voltage threshold (v_{th}), nodes with degree 1 are removed from $g_{z,S}$ and procedures described in 2.1.3 and 2.1.4 are repeated for the modified version of the graph $g_{z,S}$.

Algorithm 2 Algorithm for Choosing Appropriate Cable with Minimum Number of Parallel Cables

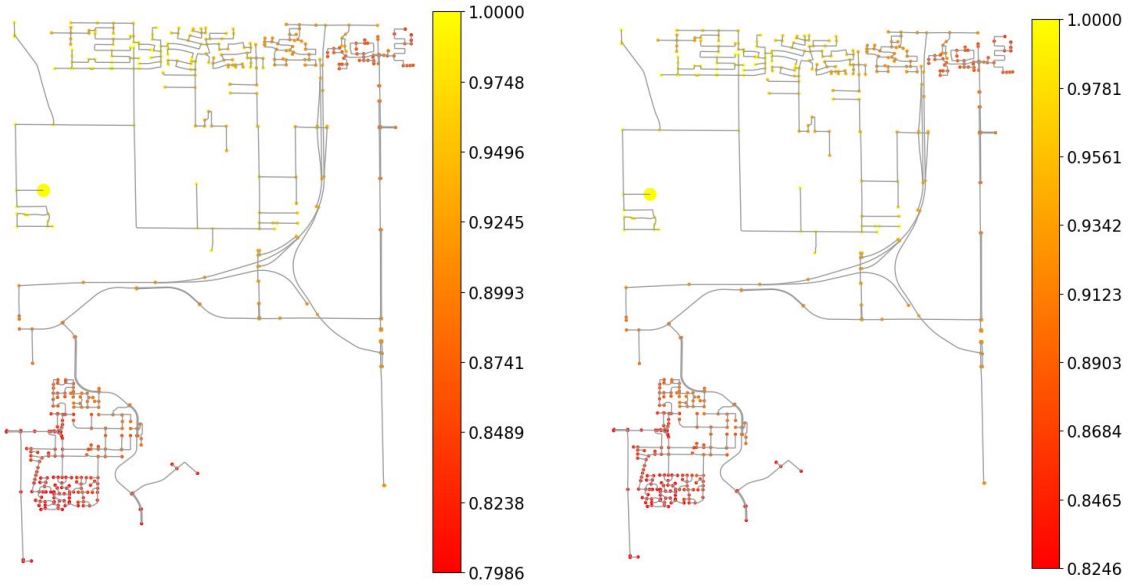
Input : $P_{i,j}$, $Q_{i,j}$ for all lines, cable database ($\mathcal{C}$), maximum number of parallel allowed= $l_{\max}$

Sort the cable database in ascending order of MVA capacity (S_c) of the cables
$\boldsymbol{c} \leftarrow$ vector containing the S_c values for all cables
$\boldsymbol{l} \leftarrow 1 \dots l_{\max}$
$\boldsymbol{C_s} \leftarrow \boldsymbol{c}\, \boldsymbol{l}^T$
for each line segment $(i, j) \in \mathcal{L}$ **do**
    calculate MVA flow ($S_{i,j}$) through line i, j using $P_{i,j}$ and $Q_{i,j}$
    flag $\leftarrow$ False
    for each column in $\boldsymbol{C_s}$ **do**
        for each row in $\boldsymbol{C_s}$ **do**
            if $\boldsymbol{C_s}$[row][column] $> S_{i,j}$ **then**
                calculate $r_{i,j}$ and $x_{i,j}$ using the resistance and reactance of the cable and the length of line i, j
                flag $\leftarrow$ True
                break
            end if
        end for
        if flag **then**
            break
        end if
    end for
end for

Output : Resistance and reactance for each line of the graph ($g_{z,S}$)

2.3 Converting to OpenDSS model

The output of the previous step is converted to an OpenDSS model to calculate a non-linear power flow solution for the positive sequence network model. The framework models all the loads as constant power loads with load and cable information retrieved from the graph $g_{z,S}$. Using the non-linear power flow provides the loss values across the network. The steady state voltage profile for the OpenDSS model is shown in Fig. 2.3.1a.



(a) (b)

Figure 2.3.1: (a) Steady State Voltage Profile Using OpenDSS) (b) Steady State Voltage Profile Using DistFlow

The OpenDSS simulation returns loss values of $3.95832 + j7.63604$ MVA; hence active losses are 8.9% of the load demand of $40.47 + j11.14$ MVA. The non-linear power flow solution from OpenDSS , which accounts for the losses, has larger currents, which are responsible for dragging the voltage profile lower than that calculated by the DistFlow equation in (2.2.1c) as evident in comparing with the voltage heat-map shown in Figure 2.3.1b.

2.4 Extension to Multi-Phase Model

The single-phase synthetic feeder generation framework presented is based on publicly available data. To follow the same path for generating unbalanced models with regulators, transformers, overhead wires, underground cables, and multi-phase lines, this work generated histograms of the above components using IEEE Distribution Test Cases and PNNL taxonomy feeders. Generated histograms for transformers,

overhead wires and underground cables are shown in Figure 2.4.1.

2.4.1 Creating Unbalanced Synthetic Distribution Feeder Model

This section extends the presented framework presented to include multiple phases by assuming the total load per phase is balanced across the entire distribution network. The percentage of the three-phase lines for a network of fixed size is chosen by finding the distributions of the histogram shown in Figure 2.4.1b for underground cable and Figure 2.4.1c for overhead wire. For instance, based on the Bayesian Information Criterion, the fitting of data for overhead wires and underground cables using the Gaussian Mixture Model is presented in Figure 2.4.2.

The following steps then describe the procedure to include single-phase laterals to the developed distribution network.

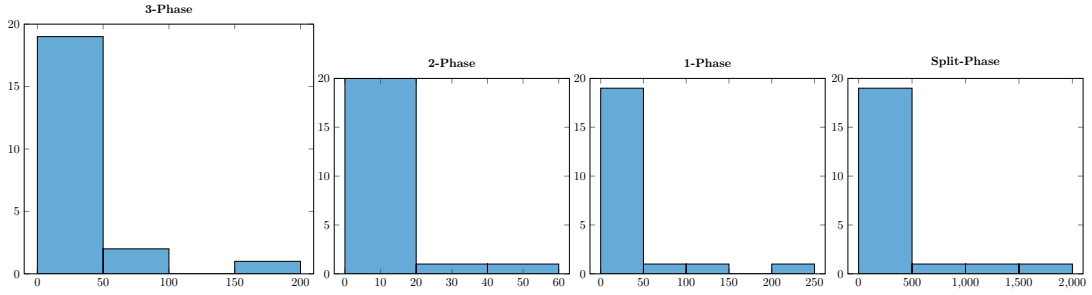
1. Select a total number of three-phase lines as a percentage of the total number of lines (by sampling from the distributions shown in Figure 2.4.2).

2. Incrementally step through the power lines from the source to end-loads. Lines with more power flow are specified as three-phase lines until the total number of three-phase lines is reached. All other lines are specified as single-phase.

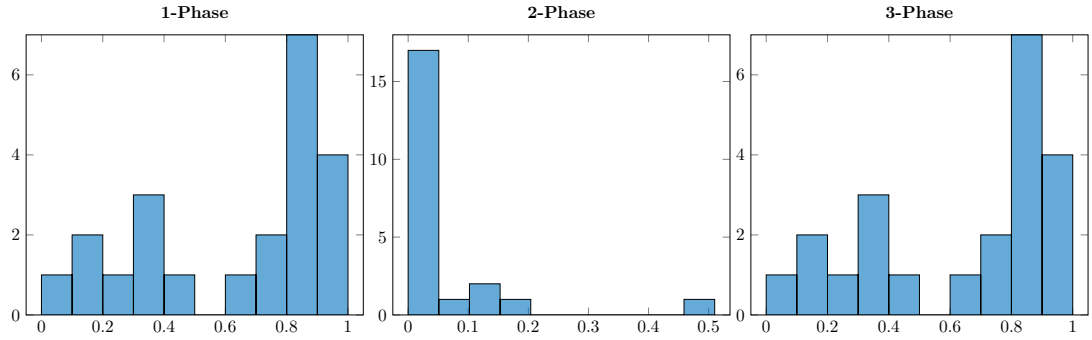
3. Sum the real power demand across individual single-phase lateral to create the set of single phase nodes.

4. Assign phases to the single-phase nodes by solving a minimization problem to achieve balance among total real power load value per phase.

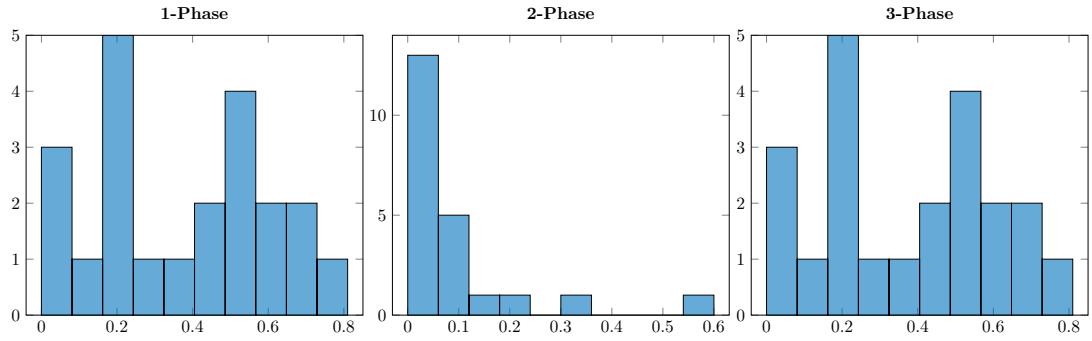
Let $\mathcal{C}_N$ represent the set of unassigned nodes with $\boldsymbol{p}_n$ representing the real power demand vector of nodes in $\mathcal{C}_N$. Also, $\boldsymbol{u}_A, \boldsymbol{u}_B, \boldsymbol{u}_C$ represent the vector of binary variables corresponding to nodes of $\mathcal{C}_N$. The optimization problem presented in (2.4.1)



(a)

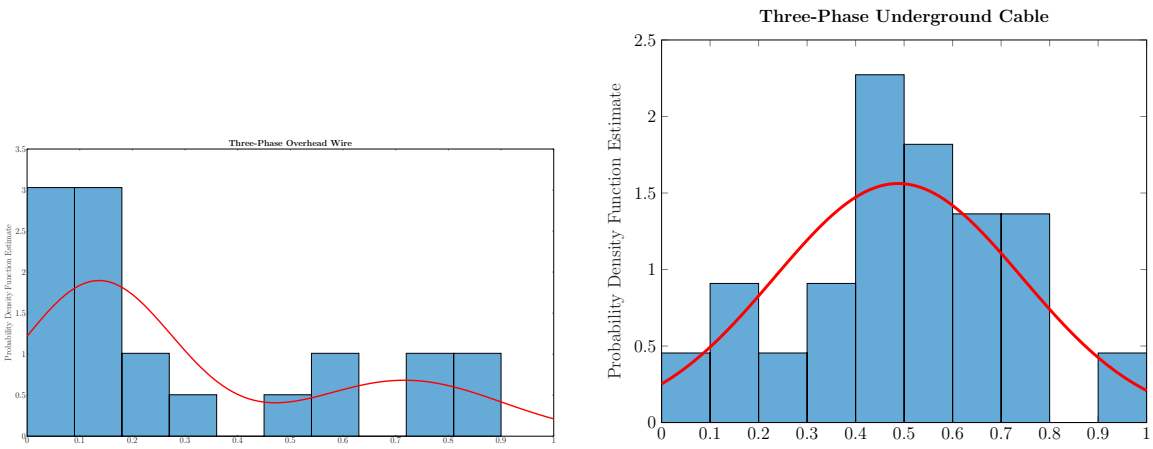


(b)



(c)

Figure 2.4.1: (a) Histogram of Various Types of Transformer (b) Histogram of Multi-phase Overhead Wire Data as a Percentage of Total Number of Lines (c) Histogram of Multi-phase Underground Cable Data as a Percentage of Total Number of Lines



(a) (b)

Figure 2.4.2: (a) Fitting Data into a GMM of Two Components, $\boldsymbol{\mu} = [0.1356\ 0.7172]$, $\boldsymbol{\sigma} = [0.1442\ 0.1846]$, weights, $\boldsymbol{\phi} = [0.684290\ 0.315710]$ (b) Fitting Data into a GMM of One Component, $\mu = 0.4879$, $\sigma = 0.2553$

thus minimizes the sum of differences of total power of phase A and phase B (d_{AB}); phase B and phase C (d_{BC}) ; phase C and phase A (d_{CA}).

$$\min_{d_{AB},d_{BC},d_{CA},\boldsymbol{u}_A,\boldsymbol{u}_B,\boldsymbol{u}_C} \quad d_{AB} + d_{BC} + d_{CA} \tag{2.4.1a}$$

subject to

$$\boldsymbol{u}_A, \boldsymbol{u}_B, \boldsymbol{u}_C \in \{0,1\}^{|\mathcal{C}_N|} \tag{2.4.1b}$$

$$\boldsymbol{p}_n^\top \boldsymbol{u}_A = P_A \ , \ \boldsymbol{p}_n^\top \boldsymbol{u}_B = P_B \ , \ \boldsymbol{p}_n^\top \boldsymbol{u}_C = P_C \tag{2.4.1c}$$

$$d_{AB} \geq 0 \ ; \ d_{BC} \geq 0 \ ; \ d_{CA} \geq 0 \tag{2.4.1d}$$

$$-d_{AB} \leq P_A - P_B \leq d_{AB} \tag{2.4.1e}$$

$$-d_{BC} \leq P_B - P_C \leq d_{BC} \tag{2.4.1f}$$

$$-d_{CA} \leq P_C - P_A \leq d_{CA} \tag{2.4.1g}$$

To model only single-phase literals with three-phase lines, the constraint in (2.4.2) is added to (2.4.1) and synthetic distribution feeder with three-phase and single-phase

edges is shown in Figure 2.4.3a.

$$\boldsymbol{u}_A + \boldsymbol{u}_B + \boldsymbol{u}_C = \mathbf{1} \tag{2.4.2}$$

The developed framework also allows to model both single and two-phase laterals by adding the constraint in (2.4.3) to (2.4.1). Here, $\rho(c)$ returns the parent of node c and (2.4.3b) allows to constrain the phase of a child node with respect to its parent node. The synthetic distribution feeder with multi-phase laterals is shown in Figure 2.4.3b.

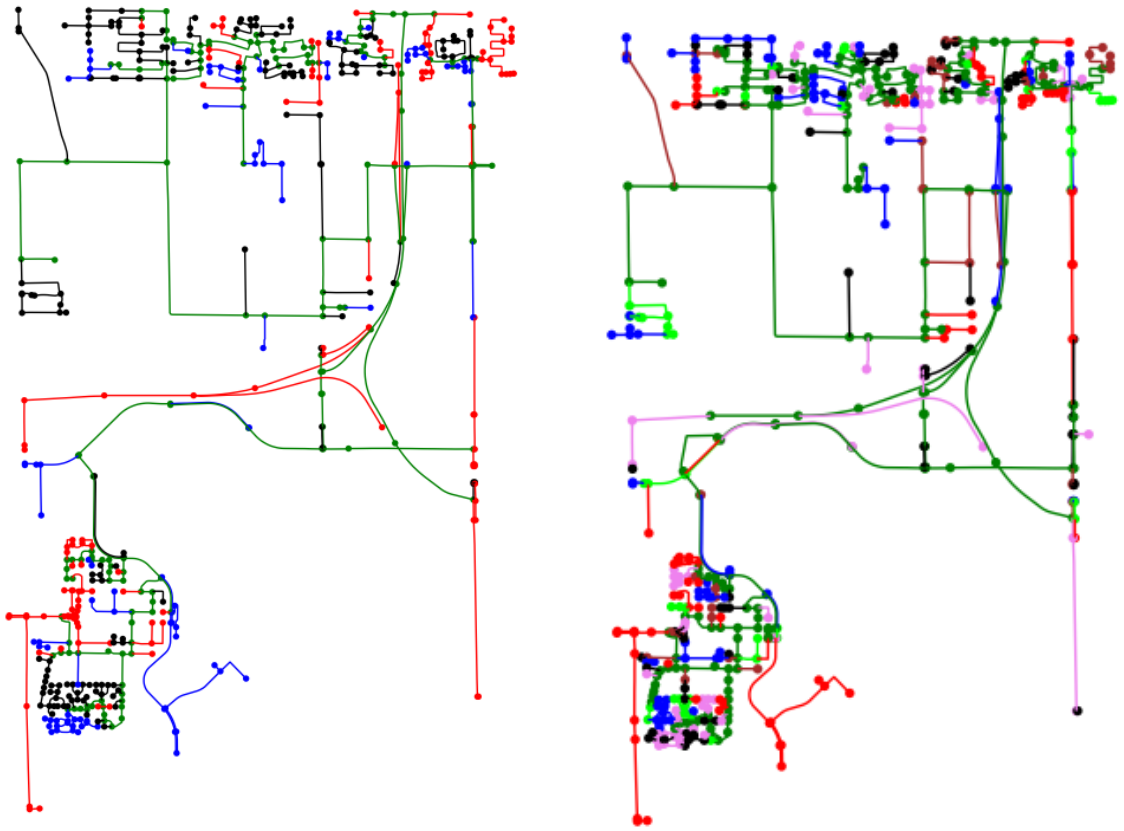
$$\mathbf{1} \leq \boldsymbol{u}_A + \boldsymbol{u}_B + \boldsymbol{u}_C \leq \mathbf{2} \tag{2.4.3a}$$

$$u_A^c \leq u_A^{\rho(c)};\ u_B^c \leq u_B^{\rho(c)};\ u_C^c \leq u_C^{\rho(c)};\ \forall c \in \mathcal{C}_N \tag{2.4.3b}$$

2.5 Modeling City-Wide Synthetic Distribution Feeder

In this section, the process diagram presented in Figure 2.1.1 is used to generate city-wide synthetic distribution feeders. Using 2010 US Census data, the list of ZIP codes for an entire city is retrieved and the algorithm described in Section 2.1 can be repeated for each ZIP code. This procedure is followed to generate the synthetic distribution feeders for "Phoenix" metro as shown in Figure 2.5.1a.

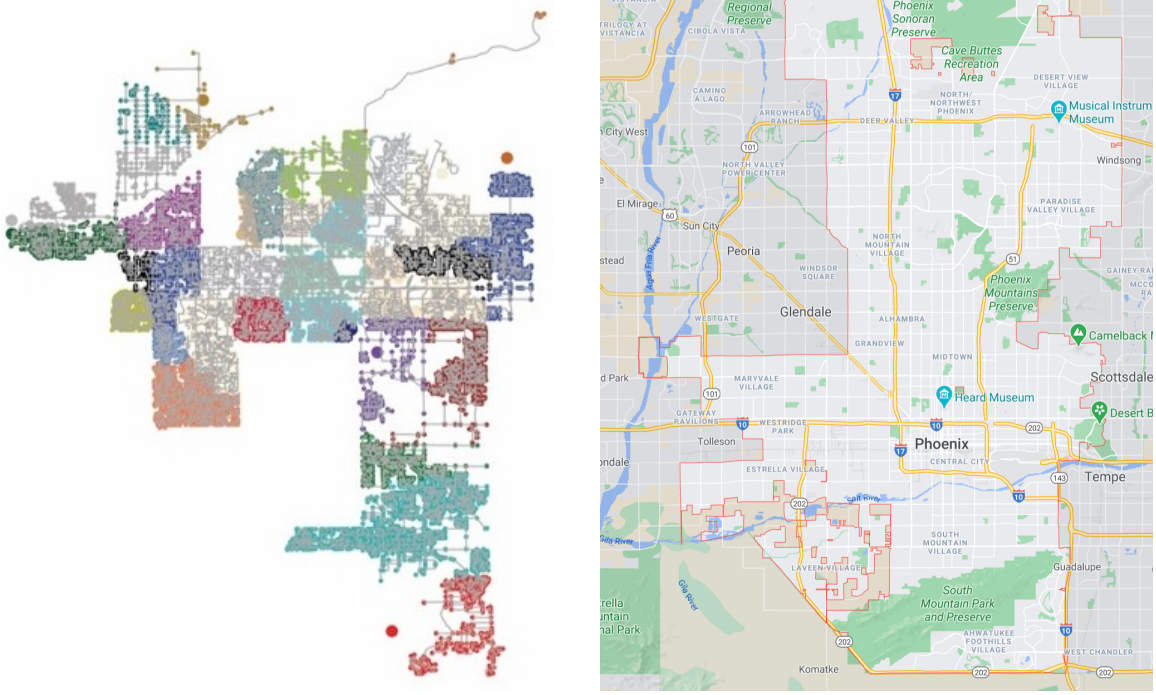
However, comparing Figure 2.5.1a to Figure 2.5.1b shows that, the collection of synthetic feeders span a much greater area than the original Phoenix area. This happens due to the fact that the boundary of ZIP codes do not necessarily follow the boundary of the city itself. To resolve this issue, rather than using ZIP code based approach, the proposed framework is modified to start the algorithm by retrieving the OpenStreetMap of the city, find the substations within the city from ACTIVS and



(a) (b)

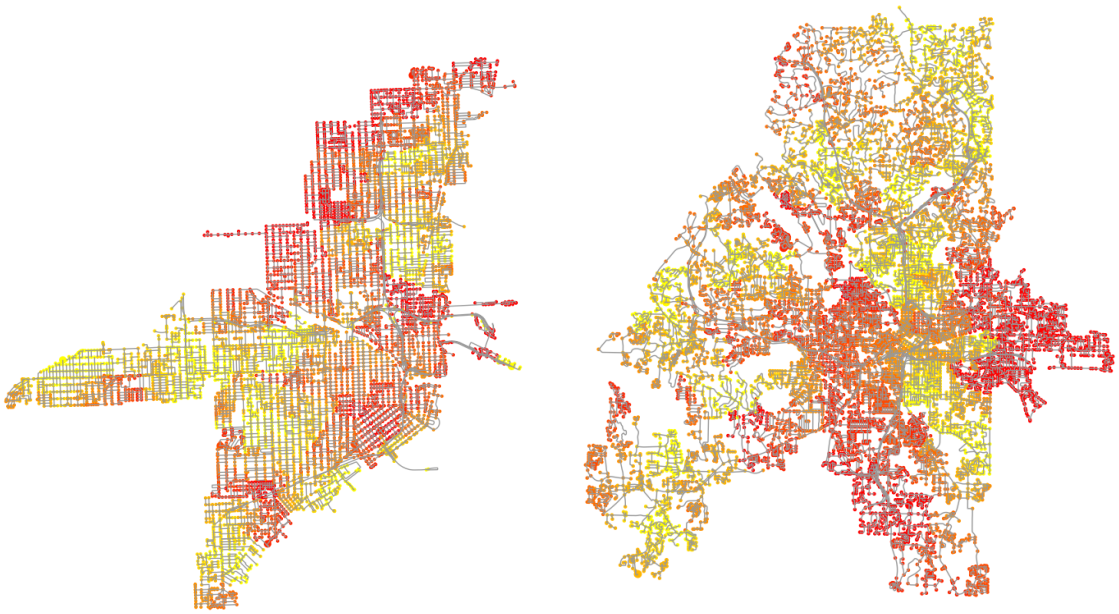
Figure 2.4.3: (a) Single-phase Laterals with Three-phase Lines (Green, Black, Red, and Blue Represent Three-phase, a Phase, B Phase, C Phase Edges Respectively) (b) Three-phase Trunk with Multi-phase Laterals (Green, Violet, Lime, Brown, Black, Red, and Blue Represent Three-phase, AB, BC, CA, A phase, B phase, C Phase Edges Respectively)

then model each substation individually. This framework is then used for generating synthetic distribution feeders for cities in United States as shown in Figure 2.5.2 and Figure 2.5.3.



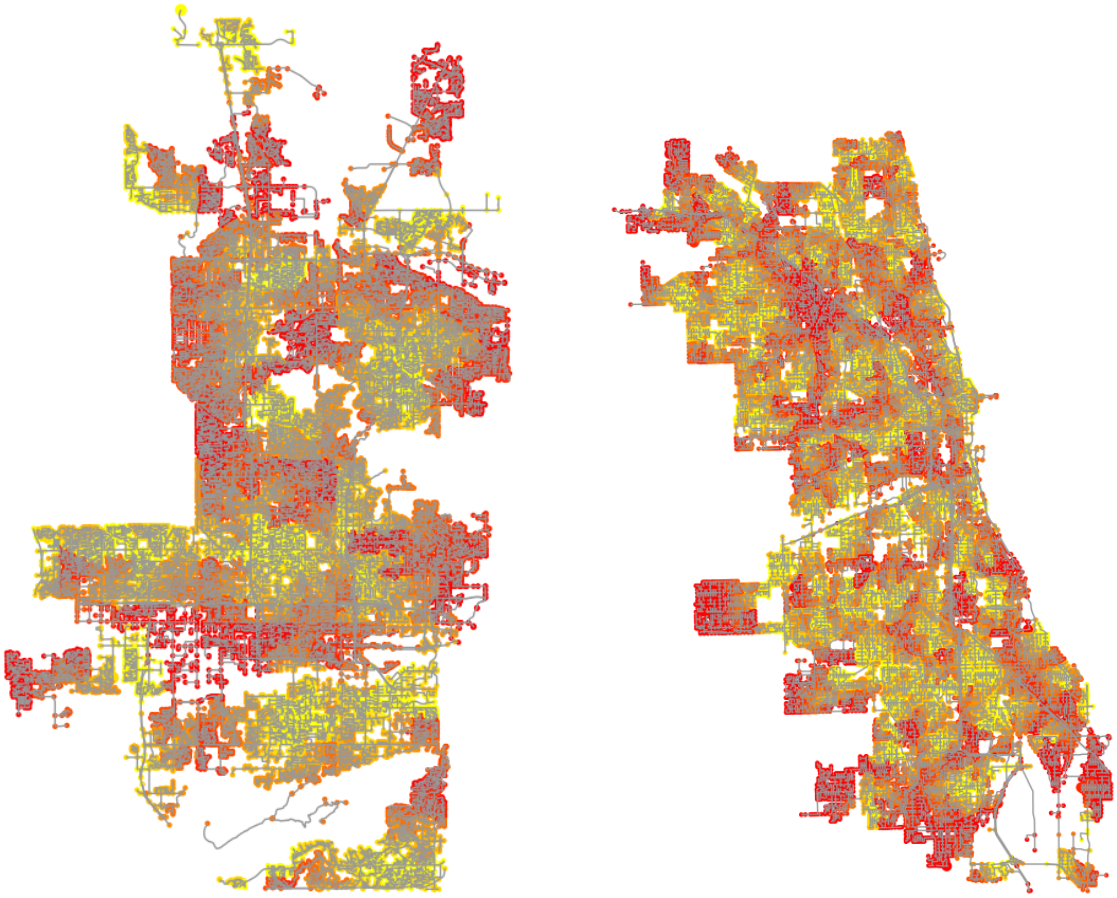
(a) (b)

Figure 2.5.1: (a) Synthetic Distribution Feeders for Phoenix Metro (b) Overview of Phoenix Metro from Google Maps



(a) (b)

Figure 2.5.2: (b) City of **Miami** (11 Substations, 8104 Nodes) (b) City of **Atlanta** (16 Substations, 13382 Nodes)



(a) (b)

Figure 2.5.3: (a) City of **Phoenix** (35 Substations, 25810 Nodes) (b) City of **Chicago** (99 Substations, 28503 Nodes)

— PART II —

Grid Graph Signal Processing For Power Distribution Systems

CHAPTER 3

CONVEX RELAXATION FOR STATE ESTIMATION AND OPTIMAL POWER FLOW USING GRAPH SIGNAL PROCESSING

This chapter first revisits the application of SDP to the solutions of AC SE and AC OPF through the lens of Grid-GSP. Prior works have demonstrated GSP applications for cyber and physical stress detection [Hasnat and Rahnamay-Naeini(2020)], identification of false data injection attack on electric transmission networks [Drayer and Routtenberg(2020)], community detection [Ramakrishna and Scaglione(2019b)], and more. Authors in [Ramakrishna and Scaglione(2019b)] showed that voltage phasors for power transmission systems could be viewed as the output of a complex graph filter through a Graph Shift Operator (GSO) via the system admittance matrix. Importantly, [Ramakrishna and Scaglione(2019b)] also showed that system voltage phasors are a *low-pass* graph signal; this is a key property verified herein and applied to radial distribution systems (as opposed to transmission systems). Also proposed is a formulation to identify the correct Graph Fourier Transform (GFT) basis (c.f. Section 3.1.1) and define a grid graph signal spectrum. Instead of considering the low-pass property as a tool to process PMU measurement as done in prior works, this work proposes a convex relaxation-based framework that uses only AMI measurements (voltage and current magnitude, and apparent power measurement at some nodes) for SE. The proposed approach can outperform matrix completion techniques as matrix completion can only fill gaps in missing measurements and cannot estimate variables that are not measured at all. Therefore, to estimate voltage phasors, the matrix completion technique will require the complex voltage phasor in either Cartesian or polar format. Hence, traditional measurements available through AMIs will be insufficient for estimating voltage phasors through a matrix completion technique,

prompting the use of a GSP based proposed approach to estimate complex voltage phasors using limited data types measured by AMIs [Bernal Heredia *et al.*(2021)].

The technique for solving a centralized convex relaxation problem includes replacing the rank-1 matrix of the outer product of the voltage phasors' vector with a positive semi-definite matrix. Hence, for a b bus system, the variable being solved for is of size b^2, which becomes computationally expensive for large networks. The observation presented is that phasors in distribution systems can be approximated with relatively few components in the graph frequency domain, significantly less than the number of buses. This concept goes beyond generic dimensionality reduction because the principal subspace spanned by the voltage phasors is known. The low-rank representation of the voltage phasor vector opens the door for finding optimal sampling patterns that significantly reduce the number of measurements needed for reconstruction and provides guidelines on the optimum sampling pattern, i.e., the optimum placement for sensors. In addition, the low-pass representation of the voltage phasor reduces the optimization variable resulting in achieving solutions for OPF faster than the traditional SDP approach. The formulations are generic for single-phase and three-phase networks and can include unbalanced networks with multi-phase transformers, loads, underground cables, overhead wires, and more. Simulation results for single-phase and unbalanced three-phase networks under varying operating scenarios are provided in Section 3.4 to illustrate the efficacy of the proposed method.

3.1 Preliminaries

A power distribution network can be represented by a network of lines with an associated weighted graph $\mathcal{G}_e(\mathcal{B}, \mathcal{E}_e, \boldsymbol{y})$ where $\mathcal{B}$ is the set of B number ($|\mathcal{B}| = B$) of buses (nodes), $\mathcal{E}_e$ is the set of edges (overhead wire/underground cable/transformers) and $\boldsymbol{y}$ is a vector of complex edge weights; $y_{ij} \neq 0$ if $(i, j) \in \mathcal{E}_e$ that represent edge

admittances. Irrespective of the single-phase or three-phase representation of the network, the weighted graph can completely express Ohm's law for the underlying distribution network. For the distribution grid, all measurements are signals on the irregular support of $\mathcal{G}_e$, and for such signals, a GSP framework can generalize Digital Signal Processing (DSP) tools. For instance, voltage phasors (system state) are one set of such signals through which all other quantities of interest, such as current and apparent power, can be derived. Moreover, the ability to express the fundamental behaviors of power systems as a graph filter paves the way to predict properties that are useful for dimensionality reduction and can simplify both SE and OPF formulations, as shown in this work.

3.1.1 A Brief Review of GSP

Compared to traditional power systems analysis, GSP provides tools to analyze signals on graphs at a more abstract level. Most importantly, GSP extends Fourier analysis to capture the variability of graph signals across adjacent nodes and sections of the graph in the GFT domain [Ramakrishna *et al.*(2020), Ortega *et al.*(2018)].

Definition 1. *A graph signal is a vector of values indexed by nodes of a graph.*

For instance, for $\mathcal{G}_e$, complex-valued voltage phasor ($\boldsymbol{v} \in \mathbb{C}^{|\mathcal{B}|\times 1}$) defined over $\mathcal{B}$ is considered as a graph signal.

Definition 2. *Analogous to the shift operator in discrete time domain, a GSO, $\boldsymbol{L}$, is a linear neighborhood operator such that each entry of the shifted graph signal $\boldsymbol{Lx}$ is a linear combination of the graph signal $\boldsymbol{x}$ components neighbors' values.*

Defining an appropriate GSO for a class of signals is a pre-requisite in GSP, since almost all operations including filtering, transformation and prediction are directly related to the GSO [Gavili and Zhang(2017)]. Most GSP instances use a Laplacian

operator [Shuman *et al.*(2013)] as GSO, which is similar to the Laplace variable s and behaves much like the Laplacian of a continuous field over the graph topology, in other words, like a differentiator, not a shift. The idea behind extending Fourier analysis to graph signals is that the Fourier basis in DSP corresponds to the eigenvectors of the Laplacian matrix of a directed circular graph where each node represents a particular time index of a periodic signal. For a generic irregular graph, it is still possible to define the Laplacian matrix ($\boldsymbol{L}$) of the graph and use that as the GSO. Let the eigenvalue decomposition be $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ where $\boldsymbol{\Lambda}$ is a diagonal matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ and $\boldsymbol{U}$ be the eigenvector matrix that acts as the basis for *graph* Fourier analysis. The GFT of a graph signal is then $\tilde{\boldsymbol{x}} = \boldsymbol{U}^\top\boldsymbol{x}$ and the eigenvalues $\lambda_\ell, \ell = 1, \ldots, N$ are called as *graph frequencies*. Ordering graph frequencies can be justified by using the total-variation criterion to express the smoothness of a graph signal [Shuman *et al.*(2013)]. The lower the magnitude, the lesser the *variation* (measured by $\|\boldsymbol{L}\boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2$) in signal values among adjacent nodes in a graph.

The concepts of GSO and GFT are helpful in defining a graph filter and its frequency response, respectively.

Definition 3. *A linear shift invariant graph filter is a linear operator $\mathcal{H}(\boldsymbol{L})$ with the property that application of GSO to the input of the filter is equivalent to applying the GSO to the output of the filter i.e.:*

$$\boldsymbol{x} = \mathcal{H}(\boldsymbol{L})\boldsymbol{s} \quad \Longleftrightarrow \quad \boldsymbol{L}\boldsymbol{x} = \mathcal{H}(\boldsymbol{L})\boldsymbol{L}\boldsymbol{s} \tag{3.1.1}$$

The implication of this property is that $\mathcal{H}(\boldsymbol{L})$ is a matrix polynomial in the GSO (possibly of infinite order):

$$\mathcal{H}(\boldsymbol{L}) = \sum_{\ell=-\infty}^{+\infty} h_\ell \boldsymbol{L}^\ell \Longleftrightarrow \mathcal{H}(\boldsymbol{L}) = \boldsymbol{U}\left(\sum_{\ell=-\infty}^{+\infty} h_\ell \boldsymbol{\Lambda}^\ell\right)\boldsymbol{U}^\top \tag{3.1.2}$$

Now, $\sum_{\ell=-\infty}^{+\infty} h_\ell \boldsymbol{\Lambda}^\ell$ in (3.1.2) is a diagonal matrix, with entries $\tilde{h}(\lambda_i) \triangleq \sum_{\ell=-\infty}^{+\infty} h_\ell \lambda_i^\ell$. The conventional notion of transfer function is possible for graph filters through the following:

$$\boldsymbol{x} = \mathcal{H}(\boldsymbol{L})\boldsymbol{s} \quad \Longleftrightarrow \quad \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{h}} \odot \tilde{\boldsymbol{s}} \tag{3.1.3}$$

where $\tilde{\boldsymbol{h}} = [\tilde{h}(\lambda_1), \ldots, \tilde{h}(\lambda_N)]^\top$ is the transfer function or *graph frequency response* of the graph filter.

Definition 4. *For any k, $1 \leq k \leq n-1$, define the low-pass ratio:*

$$\eta_k(\tilde{\boldsymbol{h}}) := \frac{\max\{|h(\lambda_{k+1})|, \ldots, |h(\lambda_n)|\}}{\min\{|h(\lambda_1)|, \ldots, |h(\lambda_k)|\}} \tag{3.1.4}$$

The graph filter $\mathcal{H}(\boldsymbol{L})$ is k-low-pass if and only if the low-pass ratio is less than 1, $\eta_k(\tilde{\boldsymbol{h}}) \in [0, 1)$. Upon passing a graph signal through $\mathcal{H}(\boldsymbol{L})$, the high frequency components (above λ_k) are attenuated by a less than or equal to $\eta_k(\tilde{\boldsymbol{h}})$.

The first k such that $\eta_k(\tilde{\boldsymbol{h}}) \ll 1$ characterizes the bandwidth *(or the cut-off frequency) of the low-pass filter, which is at λ_k.*

A similar definition can be applied to a signal, evaluating for every $k = 1, \ldots, n-1$ the ratio $\eta_k(\tilde{\boldsymbol{s}})$ and it is easy to prove that, for the filter output $\boldsymbol{x}$:

$$\eta_k(\tilde{\boldsymbol{x}}) \leq \eta_k(\tilde{\boldsymbol{h}})\eta_k(\tilde{\boldsymbol{s}}) \tag{3.1.5}$$

This illustrates that, as long as there is a k such that $\eta_k(\tilde{\boldsymbol{s}}) \ll 1/\eta_k(\tilde{\boldsymbol{h}})$, the filtered graph signal $\boldsymbol{x}$ is low-pass.

3.1.2 Grid-GSP for Power Distribution Networks

This section formalizes the notion of Grid-GSP with the goal of finding the appropriate basis $\boldsymbol{U}$ for GFT as well as justifying why the grid state $\boldsymbol{v}$ is a low-pass graph signal.

Let $\boldsymbol{s} = \boldsymbol{p} + j\boldsymbol{q}$ be the vector of net apparent power at buses ($\boldsymbol{s} \in \mathbb{C}^{|\mathcal{B}|}$). Further, let $\boldsymbol{v}$ and $\boldsymbol{V}$ be the bus voltage phasor and magnitude, respectively, with $\boldsymbol{v} \in \mathbb{C}^{|\mathcal{B}|\times 1}$ and $\boldsymbol{V} \in \mathbb{R}_{+}^{|\mathcal{B}|\times 1}$, and let $\boldsymbol{i} \in \mathbb{C}^{|\mathcal{B}|\times 1}$ and $\boldsymbol{I} \in \mathbb{R}_{+}^{|\mathcal{B}|\times 1}$ be the net bus current phasor and magnitude, respectively:

$$v_n = V_n\, e^{j\theta_n^v},\ i_n = I_n\, e^{j\theta_n^i},\ \forall n \in \mathcal{B}\ \Rightarrow\ \boldsymbol{i} = \boldsymbol{Y}\boldsymbol{v} \tag{3.1.6}$$

where $\boldsymbol{Y}$ in (3.1.6) is the system admittance matrix defined as:

$$[\boldsymbol{Y}]_{ij} = \begin{cases} y_{ii}^{sh} + \sum_{k\in\mathcal{B}_i} y_{ik}, i = j \\ -y_{ij}; i \neq j \end{cases} \tag{3.1.7}$$

with y_{ij} as the admittance of the branch between node i and j when $(i,j) \in \mathcal{E}_e$, y_{ii}^{sh} is the shunt element at node i. Moreover, Ohm's law allows us to view voltage as the output *low-pass* filter by $\boldsymbol{v} = \boldsymbol{Y}^{-1}\boldsymbol{i}$. Based on Section 3.1.1, it is clear that $\boldsymbol{Y}^{-1}$ is equivalent to an integrator filter with the GSO as $\boldsymbol{Y}$.

A caveat is that the system matrix $\boldsymbol{Y}$ is complex-valued and symmetric but is non-Hermitian Laplacian due to a low $^{X}/_{R}$ ratio with eigenvalues that generally do not have a partial order. To overcome this issue, authors in [Drayer and Routtenberg(2020)] considered the real and imaginary parts of $\boldsymbol{Y}$ matrix. In [Ramakrishna and Scaglione(2019b), Ramakrishna and Scaglione(2019a)] the authors generalized GSP for complex-valued weights in the GSO using absolute values of the eigenvalues as graph frequencies. Following [Ramakrishna and Scaglione(2019b)], the eigenvectors of the GSO $\boldsymbol{Y}$ are transformed to be complex orthogonal so that:

$$\boldsymbol{Y} = \boldsymbol{U}_{\text{GFT}}\boldsymbol{\Lambda}_{\text{GFT}}\boldsymbol{U}_{\text{GFT}}^{\top}\ ;\ \boldsymbol{U}_{\text{GFT}}^{\top}\boldsymbol{U}_{\text{GFT}} = \boldsymbol{U}_{\text{GFT}}\boldsymbol{U}_{\text{GFT}}^{\top} = \mathbb{I} \tag{3.1.8}$$

For complex valued voltage signals on the graph $\mathcal{G}_e$, the GFT basis is defined using the complex orthogonal basis $\boldsymbol{U}_{\text{GFT}}$; the absolute value of eigenvalues ($|\lambda^{\text{GFT}}|$) as graph

frequencies:

$$\text{The GFT of voltage graph signal: } \tilde{\boldsymbol{v}}^{\text{GFT}} = \boldsymbol{U}_{\text{GFT}}^{\top}\boldsymbol{v} \tag{3.1.9a}$$

$$\text{The inverse GFT: } \boldsymbol{v} = \boldsymbol{U}_{\text{GFT}}\tilde{\boldsymbol{v}}^{\text{GFT}} \tag{3.1.9b}$$

Here, $\boldsymbol{U}_{\text{GFT}}$ forms the GFT basis and $[\tilde{\boldsymbol{v}}_{\text{GFT}}]_m$ is the frequency component corresponding to m^{th} graph frequency $|\lambda_m^{\text{GFT}}|$, where $|\lambda_1^{\text{GFT}}| < |\lambda_2^{\text{GFT}}| < \ldots < |\lambda_N^{\text{GFT}}|$. Having established through Ohm's law that $\boldsymbol{v}$ is the output of the graph filter $\boldsymbol{Y}^{-1}$ [Ramakrishna and Scaglione(2019b),Ramakrishna and Scaglione(2019a)], the transfer function of interest is simply given by $\tilde{h}(\lambda_k^{\text{GFT}}) = 1/|\lambda_k^{\text{GFT}}|$, and thus $\eta_k(\tilde{\boldsymbol{h}}) = {}^{|\lambda_k^{\text{GFT}}|}/_{|\lambda_{k+1}^{\text{GFT}}|} \leq 1$. For cut-off frequencies $k \ll |\mathcal{B}|$, $\eta_k(\tilde{\boldsymbol{h}}) \ll 1$ due to the rapidly decaying nature of the magnitudes of the eigenvalues of $\boldsymbol{Y}$ caused by the weak connectivity of the grid's topology. Similarly, voltage signals GFTs exhibit a relatively small number of significant entries, as shown in Figure 3.1.1. This justifies the approximation that $\boldsymbol{v}$ lies in a low-dimensional space spanned by first k eigenvectors of $\boldsymbol{Y}$ representing the GFT low-frequency basis, that is:

$$\boldsymbol{v} \simeq \sum_{m=1}^{k} \tilde{v}_m^{\text{GFT}} \boldsymbol{u}_m^{\text{GFT}} \equiv \boldsymbol{U}_{\text{GFT},k}\tilde{\boldsymbol{v}}_k^{\text{GFT}}, \quad k < |\mathcal{B}| \tag{3.1.10}$$

where $\tilde{\boldsymbol{v}}_k^{\text{GFT}} \in \mathbb{C}^k$ are the GFT coefficients of the voltage phasors for the first k graph frequencies and $\boldsymbol{U}_{\text{GFT},k} = (\boldsymbol{u}_1^{\text{GFT}}, \boldsymbol{u}_2^{\text{GFT}}, \ldots, \boldsymbol{u}_k^{\text{GFT}})$ is the basis corresponding to the first k frequencies. Using this approximation, it is clear that the number of unknown parameters become k instead of $|\mathcal{B}|$.

This work proposes using a more convenient low dimensional representation for $\boldsymbol{v}$ which relies on an orthonormal basis $\boldsymbol{U}_k$ spanning the same subspace as $\boldsymbol{U}_{\text{GFT},k}$[1],

[1]Transforming the columns in $\boldsymbol{U}_{\text{GFT},k}$ to orthonormal, $\boldsymbol{U}_k$, can be done using the Gram–Schmidt algorithm or Singular Value Decomposition.

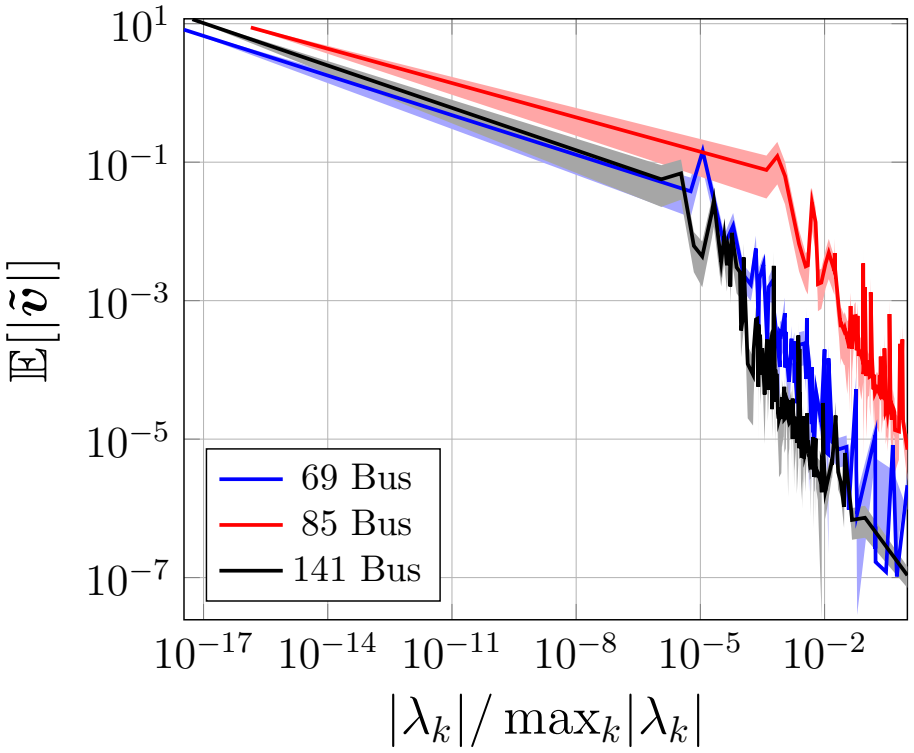


Figure 3.1.1: The Average Plus/Minus the Standard Deviation of the GFT Spectrum $\tilde{\boldsymbol{v}}$ Calculated for MATPOWER $69, 85$ and 141 Bus Radial Test Cases Associated to $n_s = 5000$ of Apparent Power Scenarios and Plotted with Respect to Normalized Graph Frequency. The Low-pass Nature Is Evident in All Three Test Cases.

specifically:

$$(\boldsymbol{U}_k\boldsymbol{U}_k^{\mathsf{H}})\boldsymbol{U}_{\text{GFT},k} = \boldsymbol{U}_{\text{GFT},k} \ , \quad \boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k = \mathbb{I} \tag{3.1.11}$$

With this choice the corresponding complex-valued residual error is orthogonal to the low rank representation:

$$\tilde{\boldsymbol{v}}_k := \boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{v} \quad \Rightarrow \quad \boldsymbol{v} = \boldsymbol{U}_k\tilde{\boldsymbol{v}}_k + \boldsymbol{U}_k^{\perp}\tilde{\boldsymbol{\varepsilon}}_k \tag{3.1.12}$$

where $\boldsymbol{U}_k^{\perp}$ is the projection matrix spanning the subspace orthongonal to $\boldsymbol{U}_k$.

Figure 3.1.2 provides empirical verification that the coefficients of residual error $\tilde{\boldsymbol{\varepsilon}}_k$ are small in magnitude and uncorrelated with $\tilde{\boldsymbol{v}}_k$. This leads to the assumption

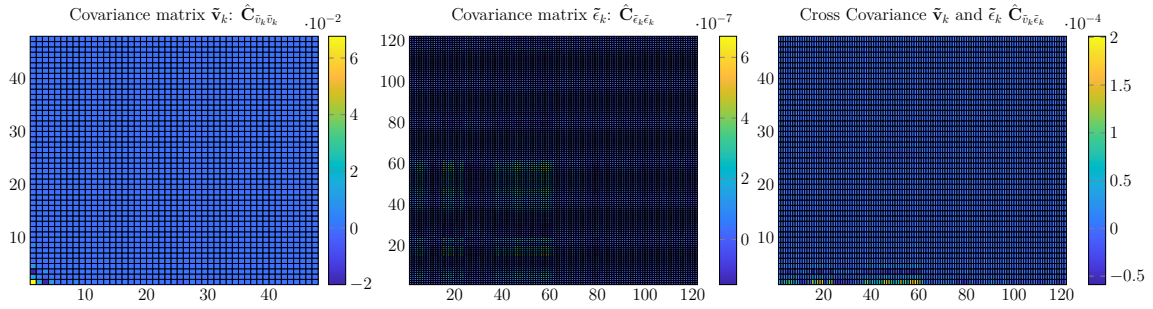


Figure 3.1.2: Sample Covariance Matrix of $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{\varepsilon}}_k$ and Their Cross Covariance for the MATPOWER 85 Bus Radial Test Case with $k = 24$.

that $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{\varepsilon}}_k$ are well approximated as complex Gaussian random variables, whose distribution can be specified in terms of the complex vector mean and complex covariance [Kay(1993)]. The sample covariance matrices for the random variables $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{\varepsilon}}_k$ shown in Figure 3.1.2 illustrate that $\tilde{\boldsymbol{v}}_k$ is a complex Gaussian random variable with low-rank covariance matrix and $\tilde{\boldsymbol{\varepsilon}}_k$ is almost negligible in magnitude. Their cross covariance matrix shown in Figure 3.1.2 indicates that $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{\varepsilon}}_k$ are approximately uncorrelated. Numerically we found that the approximation error $\boldsymbol{U}_k^{\perp}\tilde{\boldsymbol{\varepsilon}}_k$ can be well approximated by additive circularly symmetric Gaussian random vector with a small variance.

Next, the GFT approximation is merged with SDP relaxation used in power systems analysis to replace AC power flow constraints with convex constraints. As mentioned in Section 1.3, even if the SDP provides the exact solution for radial systems, this comes at the price of the number of variables growing to $|\mathcal{B}|^2$. The GFT approach reduces the number of variables to k^2, helping solve an SE problem interpolating the state value when the grid is sparsely sampled, and solves OPF problems with reduced complexity.

3.2 State Estimation Formulation

3.2.1 Measurements Model

Modern AMI sensors can measure apparent power, voltage magnitude and current magnitude, and thereby provide useful information about the system beyond that of power consumption. Values at all buses will use the following notation: $\boldsymbol{V}^2 = \text{diag}(\boldsymbol{v}^\star)\boldsymbol{v}$ as the vector of voltage magnitudes squared, $\boldsymbol{I}^2 = \text{diag}(\boldsymbol{i}^\star)\boldsymbol{i}$ as the vector of current magnitude squared, and $\boldsymbol{s} = \text{diag}(\boldsymbol{i}^\star)\boldsymbol{v}$ as the vector of net apparent power injection. These values can be expressed as linear equations with respect to the vector $(\boldsymbol{v}^\star \otimes \boldsymbol{v})$ by using the following identities:

1. Let $\boldsymbol{a}_i^{(r)}$ and $\boldsymbol{a}_i^{(c)}$ denote the i^{th} row and column of $\boldsymbol{A}$, respectively. For any $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$ if $\boldsymbol{x} = \text{diag}(\boldsymbol{ABC})$:

$$x_i = \boldsymbol{a}_i^{(r)}\boldsymbol{B}\boldsymbol{c}_i^{(c)} = \left((\boldsymbol{c}_i^{(c)})^\top \otimes \boldsymbol{a}_i^{(r)}\right)\text{vec}(\boldsymbol{B}) \tag{3.2.1}$$

2. For vectors $\boldsymbol{a}, \boldsymbol{b}$

$$\text{vec}(\boldsymbol{a}\boldsymbol{b}^{\mathsf{H}}) = \boldsymbol{b}^\star \otimes \boldsymbol{a} \tag{3.2.2}$$

Using (3.2.1) and (3.2.2), the following can be derived:

$$\boldsymbol{x}_i = \begin{bmatrix} V_i^2 \\ s_i^\star \\ I_i^2 \end{bmatrix} = \boldsymbol{H}_i(\boldsymbol{v}^\star \otimes \boldsymbol{v}); \quad \boldsymbol{H}_i := \begin{bmatrix} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)^\top \\ \boldsymbol{e}_i^\top \otimes \boldsymbol{y}_i^{(r)} \\ (\boldsymbol{y}_i^{(r)})^\star \otimes \boldsymbol{y}_i^{(r)} \end{bmatrix} \tag{3.2.3}$$

When considering $\widetilde{\boldsymbol{H}}_i := \boldsymbol{H}_i(\boldsymbol{U}_k \otimes \boldsymbol{U}_k^\star)$, it can be readily inferred that the low-pass

approximation neglects the residual error $\boldsymbol{\epsilon}_k$ in the following equation:

$$(\boldsymbol{v}^\star \otimes \boldsymbol{v}) = (\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)(\tilde{\boldsymbol{v}}_k^\star \otimes \tilde{\boldsymbol{v}}_k) + \boldsymbol{\epsilon}_k \tag{3.2.4a}$$

$$\begin{aligned}\boldsymbol{\epsilon}_k =&(\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k^\perp)(\tilde{\boldsymbol{v}}_k^\star \otimes \tilde{\boldsymbol{\varepsilon}}_k) + ((\boldsymbol{U}_k^\perp)^\star \otimes \boldsymbol{U}_k)(\tilde{\boldsymbol{\varepsilon}}_k^\star \otimes \tilde{\boldsymbol{v}}_k) \\ &+ ((\boldsymbol{U}_k^\perp)^\star \otimes \boldsymbol{U}_k^\perp)(\tilde{\boldsymbol{\varepsilon}}_k^\star \otimes \tilde{\boldsymbol{\varepsilon}}_k)\end{aligned} \tag{3.2.4b}$$

$$\Rightarrow \ \boldsymbol{x}_i = \widetilde{\boldsymbol{H}}_i(\tilde{\boldsymbol{v}}_k^\star \otimes \tilde{\boldsymbol{v}}_k) + \boldsymbol{H}_i\boldsymbol{\epsilon}_k \tag{3.2.4c}$$

The residual error $\boldsymbol{\epsilon}_k$ has the following properties:

Proposition 1. *The vector $\boldsymbol{\epsilon}_k$ lies in the orthogonal subspace with respect to $(\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)$, i.e.*

$$(\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)^{\mathsf{H}}\boldsymbol{\epsilon}_k = \boldsymbol{0} \tag{3.2.5}$$

Furthermore, the $2-$norm of the residual error is:

$$\|\boldsymbol{\epsilon}_k\|^2 = \|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big(\|\tilde{\boldsymbol{\varepsilon}}_k\|^2 + 2\|\tilde{\boldsymbol{v}}_k\|^2\big) \tag{3.2.6}$$

Assuming that the entries of $\tilde{\boldsymbol{\varepsilon}}_k$ are circularly symmetric uncorrelated Gaussian random variables provides:

$$\mathbb{E}\big[\|\boldsymbol{\epsilon}_k\|^2\big] = \mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big]\big(\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big] + 2\mathbb{E}\big[\|\tilde{\boldsymbol{v}}_k\|^2\big]\big) + \big\|\mathrm{diag}\left(\mathbb{E}\big[\tilde{\boldsymbol{\varepsilon}}_k\tilde{\boldsymbol{\varepsilon}}_k^{\mathsf{H}}\big]\right)\big\|^2 \tag{3.2.7}$$

The proof of the proposition is provided in Appendix C.1.

In general, measurements are noisy[2] and only available in a subset of buses $\mathcal{M}$. Therefore, for $i \in \mathcal{M}$,

$$\boldsymbol{x}_i' = \boldsymbol{H}_i(\boldsymbol{v}^\star \otimes \boldsymbol{v}) + \boldsymbol{\xi}_i = \widetilde{\boldsymbol{H}}_i(\tilde{\boldsymbol{v}}_k^\star \otimes \tilde{\boldsymbol{v}}_k) + \boldsymbol{H}_i\boldsymbol{\epsilon}_k + \boldsymbol{\xi}_i \tag{3.2.8}$$

3.2.2 SDP Formulation of the Low-Rank AC SE Problem

A general AC SE problem is naturally non-convex due to the presence of the term $\boldsymbol{v}\boldsymbol{v}^{\mathsf{H}}$ in the formulation. Work in [Low(2014a)] shows that for radial systems,

[2]The measurement noise is assumed to be additive circularly symmetric Gaussian noise.

replacement shown in (3.2.9) leads to the exact recovery of $\boldsymbol{v}$ with an appropriate noiseless measurement set $\mathcal{M}$.

$$\boldsymbol{v}\boldsymbol{v}^{\mathsf{H}} \mapsto \boldsymbol{W}, \quad \text{such that} \;\; \boldsymbol{W} \succeq 0, \; \text{rank}(\boldsymbol{W}) = 1 \tag{3.2.9}$$

Assuming that the indexes of buses with available measurements is $\mathcal{M} = \{1, \ldots, N_m\}$, where $|\mathcal{M}| = N_m$ and denoting

$$\boldsymbol{H}_{\mathcal{M}} = \begin{bmatrix} \boldsymbol{H}_1 & \ldots & \boldsymbol{H}_{N_m} \end{bmatrix}^{\top}; \boldsymbol{x}_{\mathcal{M}} = \begin{bmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_{N_m} \end{bmatrix}^{\top}, \tag{3.2.10}$$

the SDP relaxation that provides the exact solution is the following AC-SDP optimization:

$$\boldsymbol{W}_{\text{opt}} = \underset{\boldsymbol{W}}{\text{argmin}} \, \|\boldsymbol{x}_{\mathcal{M}} - \boldsymbol{H}_{\mathcal{M}} \, \text{vec}(\boldsymbol{W})\|^2 \;\; \text{s.t.} \;\; \boldsymbol{W} \succeq 0 \tag{3.2.11}$$

Upon solving (3.2.11), $\hat{\boldsymbol{v}}$ can be recovered from $\boldsymbol{W}_{\text{opt}}$ for a single-phase network using [Low(2014a)] and for a three-phase network using [Gan and Low(2014)].

Instead, by introducing:

$$\tilde{\boldsymbol{v}}_k \tilde{\boldsymbol{v}}_k^{\mathsf{H}} \mapsto \widetilde{\boldsymbol{W}}, \quad \text{such that} \quad \widetilde{\boldsymbol{W}} \succeq 0, \; \text{rank}(\widetilde{\boldsymbol{W}}) = 1, \tag{3.2.12}$$

and, defining $\widetilde{\boldsymbol{H}}_{\mathcal{M}} = \boldsymbol{H}_{\mathcal{M}}(\boldsymbol{U}_k^{\star} \otimes \boldsymbol{U}_k)$, the low-rank GFT based approximation can be applied to (3.2.11) resulting in a relaxed AC-GSP formulation shown in (3.2.13).

$$\widetilde{\boldsymbol{W}}_{\text{opt}} = \underset{\widetilde{\boldsymbol{W}}}{\text{argmin}} \left\| \boldsymbol{x}_{\mathcal{M}} - \widetilde{\boldsymbol{H}}_{\mathcal{M}} \, \text{vec}(\widetilde{\boldsymbol{W}}) \right\|^2 \text{s.t.} \; , \widetilde{\boldsymbol{W}} \succeq 0 \tag{3.2.13}$$

For this case, $\hat{\boldsymbol{v}}$ can be recovered using $\boldsymbol{U}_k \widetilde{\boldsymbol{W}}_{\text{opt}} \boldsymbol{U}_k^{\mathsf{H}}$ in lieu of $\boldsymbol{W}_{\text{opt}}$.

Remark 3.2.1. *The formulation presented above uses bus measurements for state estimation, however, can be expanded to include branch measurements. For instance, a branch based measurement like real power flow (p_{flow}^{ij}) measured at the* `from` *end of a branch between nodes i and j is the following:*

$$p_{ij} = \Re\{-v_i(v_i - v_j)Y_{ij}^{\star}\}; \; \forall (i,j) \in \mathcal{E}_e \tag{3.2.14}$$

Use of (3.2.1) *allows to introduce the following matrices in* (3.2.15)*:*

$$\widetilde{\boldsymbol{H}}_{F,ij} = \Big(\big(\boldsymbol{U}_k^{\mathsf{H}}\big)_i^{(c)} \Big)^{\top} \otimes \big(\boldsymbol{U}_k\big)_i^{(r)} \tag{3.2.15a}$$

$$\widetilde{\boldsymbol{H}}_{FT,ij} = \Big(\big(\boldsymbol{U}_k^{\mathsf{H}}\big)_j^{(c)} \Big)^{\top} \otimes \big(\boldsymbol{U}_k\big)_i^{(r)} \tag{3.2.15b}$$

$$\widetilde{\boldsymbol{H}}_{T,ij} = \Big(\big(\boldsymbol{U}_k^{\mathsf{H}}\big)_j^{(c)} \Big)^{\top} \otimes \big(\boldsymbol{U}_k\big)_j^{(r)} \tag{3.2.15c}$$

Integrating (3.2.15) *with* (3.2.14) *allows the real power flow through all branches to be expressed as:*

$$\boldsymbol{p}_{\mathit{flow}} = \mathfrak{R}\Big\{ (\widetilde{\boldsymbol{H}}_F - \widetilde{\boldsymbol{H}}_{FT}) \operatorname{vec}(\widetilde{\boldsymbol{W}}) \odot \boldsymbol{y}^{\star} \Big\} \tag{3.2.16}$$

The SE formulation presented above is generic for single-phase or multi-phase networks as long as the formation of $\boldsymbol{Y}$ matrix accounts for the phases of the distribution network properly. For the single-phase case, the formation of $\boldsymbol{Y}$ is straightforward. Compared to a single-phase or balanced three-phase case, the construction of $\boldsymbol{Y}$ for an unbalanced three-phase case is complicated due to multi-phase transformers with different types of winding at primary and secondary sides, unbalanced capacitor banks, multi-phase overhead wires and underground cables, unbalanced loads of various types, and other factors. In this work, the admittance matrix for transformers is modeled according to [Dugan and Santoso(2003)], allowing this work to model any type of transformer. The admittance matrix for the other elements is modeled following [Kersting(2002)]. The individual admittance matrices are then combined to build the system admittance matrix for the three-phase case, which is the GSO for three-phase networks. The size of $\boldsymbol{Y}$ for the three-phase case depends on the number of buses and number of phases per bus. However, the rapidly decaying nature of eigenvalues for single-phase $\boldsymbol{Y}$ is also seen for the $\boldsymbol{Y}$ of three-phase networks.

With the proper form for $\boldsymbol{Y}$, the SE formulation in (3.2.13) can be solved for both single-phase and three-phase cases. Apart from a known topology which results in

extracting the $\boldsymbol{Y}$ for the network, the voltage phasor at the substation is generally known. Adding this information to the SE formulation in (3.2.13) results in (3.2.17):

$$\widetilde{\boldsymbol{W}}_{\text{opt}} = \underset{\widetilde{\boldsymbol{W}}}{\text{argmin}} \left\| \text{diag}(\boldsymbol{\alpha}) \Big(\boldsymbol{x}_{\mathcal{M}} - \widetilde{\boldsymbol{H}}_{\mathcal{M}} \, \text{vec}(\widetilde{\boldsymbol{W}}) \Big) \right\|^2 \tag{3.2.17a}$$

$$\text{subject to } \widetilde{\boldsymbol{W}} \succeq 0, \quad \widetilde{\boldsymbol{H}}_{M,vs} \, \text{vec}(\widetilde{\boldsymbol{W}}) = V_s \tag{3.2.17b}$$

Here, V_s is the substation or slack bus voltage, and $\widetilde{\boldsymbol{H}}_{M,vs}$ denotes the single row and three rows of $\widetilde{\boldsymbol{H}}_{\mathcal{M}}$ corresponding to substation bus voltage for the single-phase and three-phase case, respectively. The introduction of a weighting vector $\boldsymbol{\alpha}$ allows use of actual measurements with different scaling rather than per unit measurements allowing scaling of voltage, current and apparent power values.

3.2.3 AMI Placement

Prior works have applied Grid-GSP for finding optimal locations to place Phasor Measurement Units (PMUs) in transmission networks[3] to minimize reconstruction error. For example, work presented in [Tajer *et al.*(2021)] conducted PMU placement by using a real-valued $\boldsymbol{U}_k$ matrix leveraging the high X/R ratio of transmission lines. As this work focuses on distribution networks, a similar assumption is not appropriate, making that approach inapplicable. The algorithm proposed in [Ramakrishna and Scaglione(2021)] found rows of $\boldsymbol{U}_k$ with minimum correlation, and the corresponding buses in the electrical network were chosen as the PMU locations. Work herein proposes a framework that chooses the rows of $\boldsymbol{U}_k$ and $\boldsymbol{Y}\boldsymbol{U}_k$ simultaneously with the smallest possible coherence, i.e. rows being as close as to being orthogonal. The proposed algorithm thus extends the greedy method proposed in [Tsitsvero *et al.*(2016), Ramakrishna and Scaglione(2021)] and determines optimal AMI place-

[3]The PMUs can measure voltage and current phasors.

ment as outlined in Algorithm 3.

Algorithm 3 Algorithm for Optimal AMI Placement

Input : Total number of AMIs, n_s; $\mathcal{M} = \emptyset$

$m \leftarrow 0$

while $m < n_s$ **do**

  $i^* = \text{argmax}_{i \in (\mathcal{B} \setminus \mathcal{M})} \min[\sigma_{\min}(\boldsymbol{U}_k)\ \sigma_{\min}(\boldsymbol{Y}\boldsymbol{U}_k)]$

  $\mathcal{M} \leftarrow \mathcal{M} + \{i^*\}$

  $m \leftarrow m + 1$

end while

Output : $\mathcal{M}$

3.3 AC-GSP Based OPF Formulation

This section integrates the GSP technique with an AC-SDP based OPF problem that minimizes the cost of generation and includes convexified capacity constraints [Low(2014a), Madani *et al.*(2015)]. Initial assumptions include $\mathcal{B}_G$ as the set of the buses containing generation in which $\mathcal{B}_G \in \mathcal{B}$, $\boldsymbol{p}_g$ and $\boldsymbol{q}_g$ are the vectors of real and reactive power generation for all buses in $\mathcal{B}$, $\boldsymbol{C}_g$ represents the cost function for real power generation, and $\boldsymbol{p}_d$ and $\boldsymbol{q}_d$ are the vectors of real and reactive power demand (non-negative values) for all buses in $\mathcal{B}$. Using the following matrices:

$$\widetilde{\boldsymbol{H}}_V = \left[(\boldsymbol{e}_1 \otimes \boldsymbol{e}_1)^\top, \quad \dots \quad , (\boldsymbol{e}_N \otimes \boldsymbol{e}_N)^\top \right]^\top (\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)$$

$$\widetilde{\boldsymbol{H}}_S = \left[\left(\boldsymbol{e}_1^\top \otimes \boldsymbol{y}_1^{(r)} \right), \quad \dots \quad , \left(\boldsymbol{e}_N^\top \otimes \boldsymbol{y}_N^{(r)} \right) \right]^\top (\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)$$

the AC-GSP OPF formulation using $\widetilde{\boldsymbol{W}}$ is shown in (3.3.2):

$$\min_{\boldsymbol{p}_g,\boldsymbol{q}_g,\widetilde{\boldsymbol{W}},\boldsymbol{\epsilon}_p,\boldsymbol{\epsilon}_q} \mathbf{1}^\top(\boldsymbol{C}_g(\boldsymbol{p}_g)) + w_1\|\boldsymbol{\epsilon}_p\|_2^2 + w_2\|\boldsymbol{\epsilon}_q\|_2^2 \tag{3.3.2a}$$

$$\text{s.t. } \widetilde{\boldsymbol{W}} \succeq 0 \tag{3.3.2b}$$

$$0 \le \boldsymbol{p}_g \le \boldsymbol{p}_g^{\max} \;;\; \boldsymbol{q}_g^{\min} \le \boldsymbol{q}_g \le \boldsymbol{q}_g^{\max} \tag{3.3.2c}$$

$$\boldsymbol{p}_{g,i} = 0, \quad \boldsymbol{q}_{g,i} = 0 \quad \forall i \in \mathcal{B} \setminus \mathcal{B}_G \tag{3.3.2d}$$

$$(\boldsymbol{p}_g - \boldsymbol{p}_d) - j(\boldsymbol{q}_g - \boldsymbol{q}_d) - \widetilde{\boldsymbol{H}}_S \operatorname{vec}(\widetilde{\boldsymbol{W}}) = \boldsymbol{\epsilon}_p + j\boldsymbol{\epsilon}_q \tag{3.3.2e}$$

$$\mathbf{1}^\top(\boldsymbol{p}_g - \boldsymbol{p}_d) - j\ \mathbf{1}^\top(\boldsymbol{q}_g - \boldsymbol{q}_d) = \mathbf{1}^\top(\widetilde{\boldsymbol{H}}_S \operatorname{vec}(\widetilde{\boldsymbol{W}})) \tag{3.3.2f}$$

$$v_{\min}^2 \le \Re\Big\{\widetilde{\boldsymbol{H}}_V \operatorname{vec}(\widetilde{\boldsymbol{W}})\Big\} \le v_{\max}^2 \tag{3.3.2g}$$

$$\boldsymbol{p}_{\min}^{\text{flow}} \le \Re\Big\{\Big(\widetilde{\boldsymbol{H}}_F - \widetilde{\boldsymbol{H}}_{FT}\Big) \operatorname{vec}(\widetilde{\boldsymbol{W}}) \odot \boldsymbol{y}^\star\Big\} \le \boldsymbol{p}_{\max}^{\text{flow}} \tag{3.3.2h}$$

In (3.3.2), nodal balance and line flow constraints are enforced by (3.3.2e) and (3.3.2h), respectively. Nodal voltages are constrained within lower ($\boldsymbol{v}_{\min}$) and upper ($\boldsymbol{v}_{\max}$) limits by (3.3.2g). Constraint (3.3.2c) limits the real and reactive generation capacity. Note that, the left hand side of the node balance constraint in (3.3.2e) should theoretically be equal to zero. However, doing so generally leads to an infeasible problem due to the error caused by the GSP approximation. Therefore, constraint (3.3.2e) is relaxed and additional weights are introduced in the objective function of (3.3.2). The constraint (3.3.2f) is added to model losses to minimize impact of the choice of weight values.

While the formulation presented in (3.3.2) provides a simple representation of the OPF formulation using GSP, the procedure requires calculation of multiple Kronecker products, which can be a memory intensive process. Memory usage can be reduced using a `sparse` matrix representation, such as by calculating $\boldsymbol{H}_V$ using a sparse identity matrix. However, $(\boldsymbol{U}_k^\star \otimes \boldsymbol{U}_k)$ is not a sparse matrix and can cause an *out-of-memory* error. An equivalent memory efficient formulation of the OPF problem is

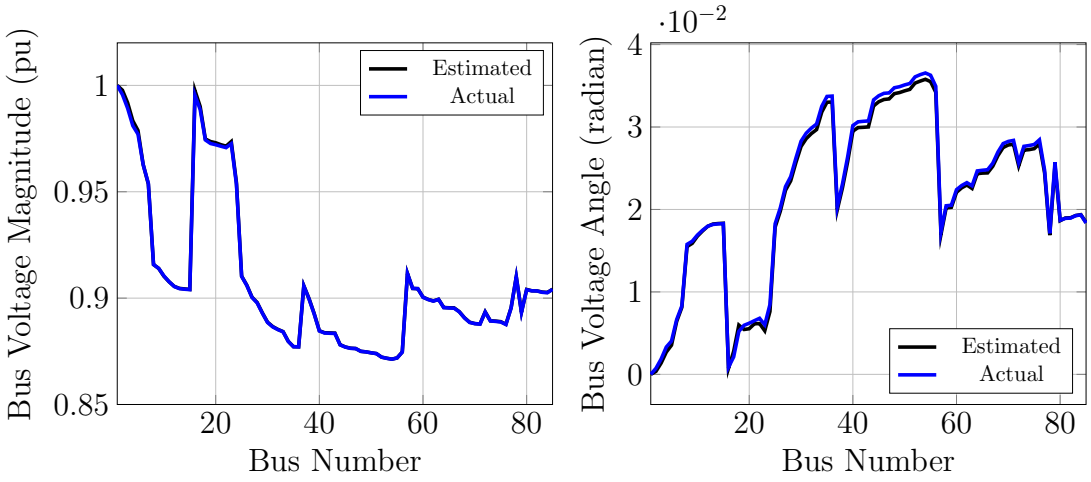


Figure 3.4.1: Voltage phasors averaged for the 85 bus radial test case with 100% measurement (AC-GSP)

provided in Appendix C.2.

3.4 Simulation Results

This section applies GSP for convex relaxation based SE and OPF formulations for single-phase test cases and three-phase unbalanced test cases from MATPOWER and IEEE, respectively. All simulations are completed on an Intel(R) Xeon(R) CPU E5-1630 v3 3.70 GHz computer using CVX as the convex programs solver [Grant and Boyd(2014)]. Voltage magnitude and angle values are reported in pu and radian, respectively.

3.4.1 State Estimation Using AC-GSP with Full Observability

The assumption that voltage phasors lie within a low-dimensional subspace from (3.1.10) is illustrated by applying (3.2.13) for the MATPOWER 85 bus test case [Zimmerman *et al.*(2011)] considering first $k = 24$ frequencies or first 24 eigenvectors to form $\boldsymbol{U}_k$. Access is assumed to bus voltage, current, and apparent power measure-

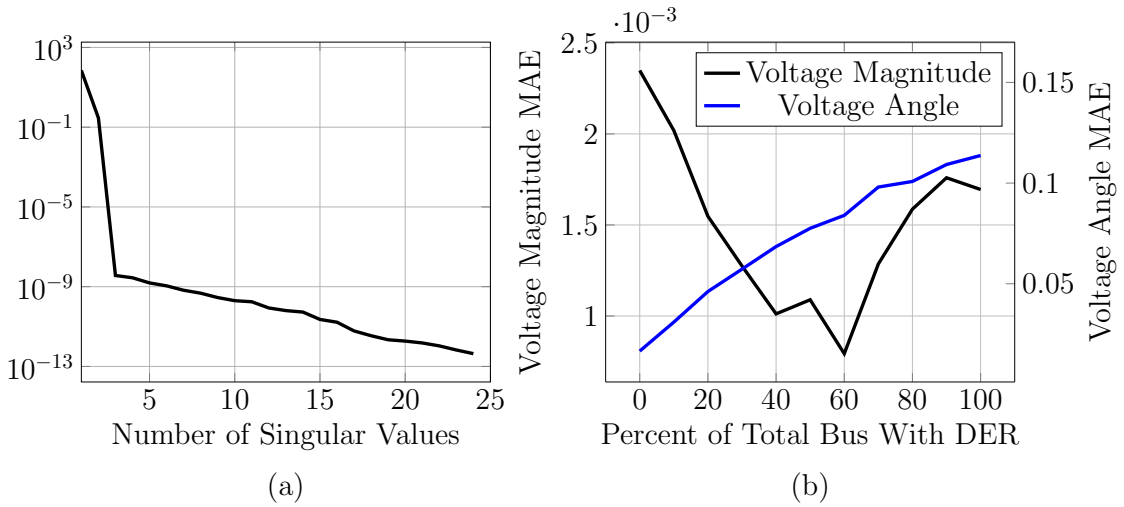


Figure 3.4.2: (a) Singular values of $\widetilde{\boldsymbol{W}}$ (AC-GSP) (b) MAE for voltage magnitude and angle estimation with increasing percent of buses with DERs

ments. Noise in those measurements are assumed as Gaussian random noise.

The first results show that the GSP based approximation approach can estimate voltage phasors with varying noise. For the MATPOWER 85 bus test case, 50 different scenarios with varying noise levels are simulated with AMIs placed at all buses. Figure 3.4.1 shows state estimation results with average voltage magnitude and average voltage angle for $k = 24$ eigenvectors. The MAE is calculated for voltage phasors in each noise scenario with results shown in Table 3.1. The low MAE reported in Table 3.1 in conjunction with Figure 3.4.1 verifies that the GSP approximation permits accurate estimation of voltage phasors with AMI measurements at various noise levels. Results shown in Figure 3.4.1 are calculated using $\widetilde{\boldsymbol{W}}$ with a relaxed rank-1 constraint (3.2.13). Figure 3.4.2a shows that $\widetilde{\boldsymbol{W}}$ holds the rank-1 property, similar to $\boldsymbol{W}$ by demonstrating the presence of dominant singular value. This provides additional verification that AC-GSP approximation can maintain the properties of standard AC-SDP relaxation.

Next, the formulation is applied to the SE problem for the IEEE 34 bus unbalanced

Table 3.1: MAE Statistics for MATPOWER 85 Bus Test Case

	Maximum	**Average**
Voltage Magnitude	0.561×10^{-3}	0.441×10^{-3}
Voltage Angle	0.115×10^{-2}	0.063×10^{-2}

Table 3.2: Voltage Magnitude and Angle Estimation MAE for IEEE 34 Bus Unbalanced Distribution Test Case

	Phase A	Phase B	Phase C
Voltage Magnitude	0.440×10^{-2}	0.530×10^{-2}	0.53×10^{-2}
Voltage Angle	3.3220×10^{-7}	2.1426×10^{-7}	3.3995×10^{-7}

distribution test case. The $\boldsymbol{Y}$ for the IEEE 34 bus test case has dimensions 138×138 including all phases across all buses. Solving this SE using AC-SDP requires solving $\boldsymbol{W}$ with a size of 138×138. Utilization of AC-GSP approximation with $k = 78$ reduces problem complexity from $O(138^2)$ to $O(78^2)$. Solution results for the reduced problem, shown in Table 3.2, verify that integrating GSP into convex relaxation can yield accurate estimation results while solving a smaller optimization problem.

3.4.2 State Estimation Using AC-GSP with Low Observability

An advantage of using the low rank representation of voltage phasors is it allows to spatially sample the signal, which means that fewer measurements are needed for state estimation. However, the placement of sensors should be optimally chosen in the sampling pattern (see algorithm 3) to reconstruct the voltage phasor signal with minimal error. Results are illustrated under the following scenarios:

3.4.2.1 Comparison Between AC-GSP And AC-SDP

Sensors of varying amounts are placed across the distribution network to aid comparing the AC-GSP approach to the conventional AC-SDP approach under low observability. Sensor placement is completed using the algorithm presented in 3. The resulting MAE for voltage and angle estimations for both cases are shown in Table 3.3. It is evident that the GSP based approach provides better estimation results compared to conventional AC-SDP based SE formulation in scenarios where complete observability is not possible.

Table 3.3: Comparison of SE for AC-GSP and AC-SDP After Placing Sensors Following Algorithm 3

Percentage of Buses with AMI	MAE (AC-GSP)		MAE (AC-SDP)	
	Voltage Magnitude	Voltage Angle	Voltage Magnitude	Voltage Angle
100%	0.0010	0.0013	0.0003	0.0005
75%	0.0009	0.0062	0.0333	0.8994
50%	0.0035	0.0302	0.2050	2.5545
25%	0.0075	0.0749	0.1932	3.6761

3.4.2.2 AC-GSP Based SE With DERs

Next experiment shows AC-GSP based SE results obtained when DERs are located throughout a distribution network. Each DER is sized to be three times the demand averaged over all buses. AMIs are placed at 50% of buses chosen by Algorithm 3 and report measurements that include random noise (locations are same as 50% case from Table 3.3). The number of buses with DERs are increased from 0% to 100% in steps of 10%. Results for voltage magnitude and angle MAE shown in Figure 3.4.2b illustrate that the AC-GSP approach achieves minimal estimation error at even high levels of

Table 3.4: MAE Calculated for AC-GSP by Varying Number of AMIs Placed at Non-load Buses

Percentage of Non-load Buses with AMI	Voltage Magnitude	Voltage Angle
0%	0.0010	0.0088
20%	0.0008	0.0067
40%	0.0007	0.0035
60%	0.0007	0.0031
80%	0.0007	0.0032
100%	0.0008	0.0012

DER penetration. MAE plots shown in Figure 3.4.2b illustrate that estimation error increases with increasing penetration of DERs as additional generation increases the high frequency components in $\tilde{\boldsymbol{v}}$ estimation, thus increasing low-pass approximation error.

3.4.2.3 Evaluating Sensor Placement Across Non-Load Buses

This experiment builds onto the last experiment by again assuming all load buses include AMIs and further places AMIs at non-load buses using Algorithm 3 to enhance reconstruction of voltage phasors across the entire network. The MAE for voltage phasor estimation is reported in Table 3.4. With AMIs placed at only load buses, the optimal placement scenario and angle estimation error is highest as the AMIs are not reporting angle measurements. When AMIs are placed across the non-load buses, the results show that estimation error for voltage angle decreases with increasing number of AMIs while solving the SE problem using the low pass approximation.

3.4.2.4 Performance Of AC-GSP Under Different Operating Conditions

The optimal placement Algorithm 3 is applied here for the MATPOWER 85 bus test case under various operating conditions with AMI meters placed at 25% of all the buses. Simulations are then completed by randomly varying apparent power demand in 5000 scenarios and solving the SE problem for each scenario. The average MAE for voltage magnitude and voltage angle are calculated and shown in Figure 3.4.3a and Figure 3.4.3b, respectively. Low estimation errors again show that the AC-GSP approach can successfully estimate voltage phasors without the need for phasor equipment and measurement, allowing us to use the more common and less expensive traditional AMI measurements.

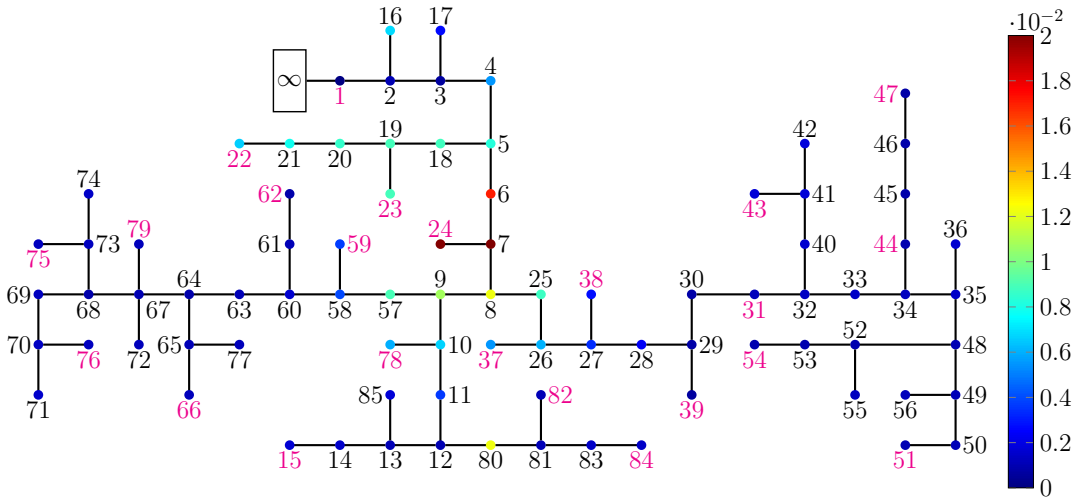
3.4.2.5 AC-GSP For The IEEE 34 Bus Network

Algorithm 3 is now applied for the three-phase case by modifying Step 3 to use a formulation of $\boldsymbol{Y}$ that includes multi-phase transformers, capacitor banks, regulators, and multi-phase distribution lines. Measurements are assumed to be available at 70% of total buses and included Gaussian noise. State estimation errors for voltage magnitude and phase angles are shown in Figure 3.4.4. Results show the AC-GSP approach allows us to calculate voltage phasors using only limited AMI measurements in a low observable region for a three-phase network, expanding applicability of the approach from the single-phase network results shared earlier [4].

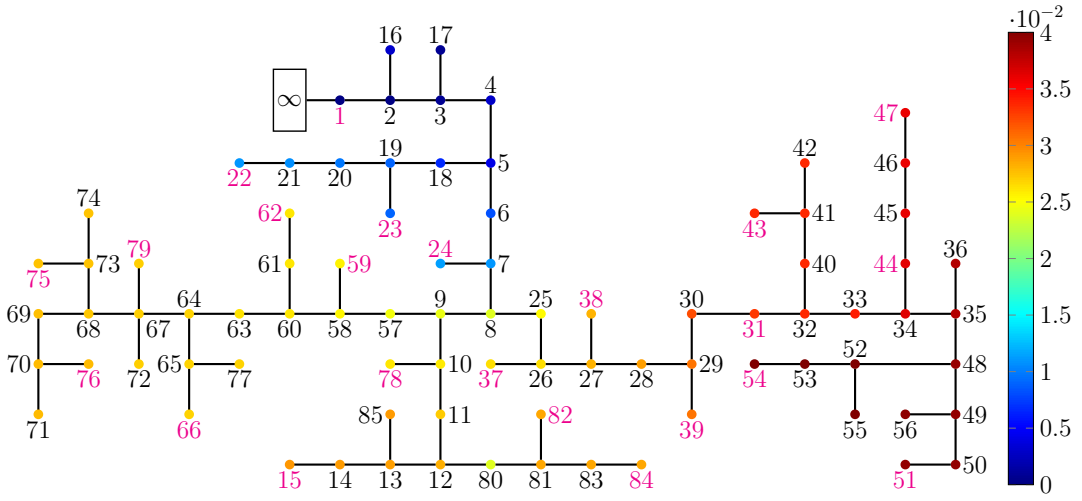
3.4.3 Results for Optimal Power Flow

This work relies on radial synthetic distribution feeders created by [Schweitzer *et al.*(2017)] to illustrate optimal power flow formulations for AC-SDP and AC-GSP.

[4]In the IEEE 34 bus case, distributed loads are modeled by placing the load at an additional bus placed at the middle of the two original buses , thus increasing the number of buses than 34.



(a)



(b)

Figure 3.4.3: (a) Voltage Magnitude MAE and (b) Voltage Angle Mae Averaged for 5000 System State Simulations with Sensor Placed at the Bus Numbers in Magenta.

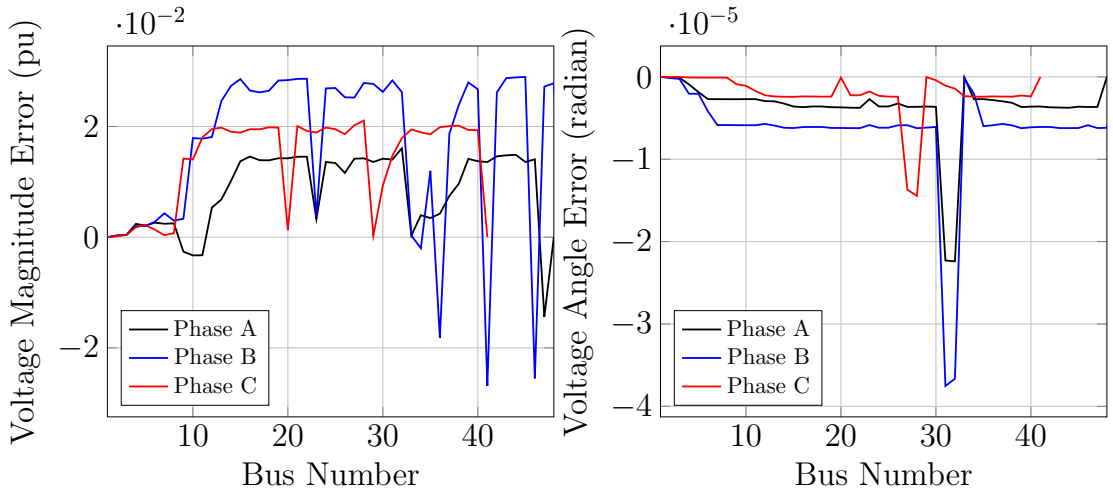


Figure 3.4.4: Three-phase Voltage Phasor Estimation Error Plot for Ieee 34 Bus Unbalanced Distribution Test Case

Table 3.5: Comparison of Optimal Power Flow Results for AC-GSP and AC-SDP

$\lvert\mathcal{B}\rvert$	$\sum(\boldsymbol{p}_d + j\boldsymbol{q}_d)$	AC-GSP					AC-SDP				
		$\sum(\boldsymbol{p}_g + j\boldsymbol{q}_g)$	$\sum \boldsymbol{C}_g(\boldsymbol{p}_g)$	v_μ	v_σ	Solving Time	$\sum(\boldsymbol{p}_g + j\boldsymbol{q}_g)$	$\sum \boldsymbol{C}_g(\boldsymbol{p}_g)$	v_μ	v_σ	Solving Time
200	$9.2263 + j3.8227$	$9.2357 + j1.2728$	1.0752	0.001818	0.001361	3.3726	$9.2427 + j1.4514$	1.0437	9.96×10^{-5}	1.32×10^{-5}	7.3533
400	$18.453 + j7.1233$	$18.462 + j2.8749$	2.8652	0.003896	0.005184	4.5453	$18.502 + j2.92$	7.0888	6.78×10^{-5}	2.06×10^{-5}	53.805
600	$27.679 + j10.616$	$27.688 + j3.68$	3.3781	0.004017	0.003601	6.8481	$27.772 + j3.4596$	7.3943	5.35×10^{-5}	2.24×10^{-5}	188.96
800	$36.905 + j14.957$	$36.953 + j14.973$	3.931	0.00578	0.008612	36.324	$37.185 + j15.048$	19.215	8.78×10^{-5}	9.24×10^{-5}	487.15
2000	$92.263 + j36.599$	$92.269 + j14.298$	9.7595	0.017935	0.007818	385.19	$92.755 + j12.895$	52.141	4.01×10^{-5}	3.27×10^{-5}	5894.9

Synthetic feeder generation allows us to test the optimization formulation on various feeder sizes and compare solution time and generation cost. Distribution feeders of various sizes ($|\mathcal{B}|$) are created using the algorithm in [Schweitzer *et al.*(2017)] with 10% of the total number of buses chosen randomly for placing DERs. The energy cost (\$/kW) is considered to be a second order polynomial with coefficients chosen randomly. Each DER has a maximum real power limit of $\sum \boldsymbol{p}_d/|\mathcal{B}|$ and can inject or absorb reactive power within a minimum power factor limit of 0.8, leading or lagging. Line flow limits are 125% of branch flow values calculated from the base case AC power flow (no DER).

Each size network is solved for the optimization problem using AC-GSP and AC-SDP formulations with corresponding voltage phasors $\boldsymbol{v}_{\text{GSP}}$ and $\boldsymbol{v}_{\text{SDP}}$ calculated. After solving the OPF problem for each feeder size, load demands, DER generation, and substation voltage are used as input to solve the corresponding AC power flow problem that provided $\boldsymbol{v}_{\text{AC-G}}$ and $\boldsymbol{v}_{\text{AC-S}}$ respectively. The difference of voltage magnitude in both approaches are calculated using (3.4.1). Average (v_μ) and standard deviation (v_σ) of $\boldsymbol{v}^e_{\text{GSP}}$ and $\boldsymbol{v}^e_{\text{SDP}}$ are then calculated and reported in Table 3.5.

$$\boldsymbol{v}^e_{\text{GSP}} = |\boldsymbol{v}_{\text{GSP}} - \boldsymbol{v}_{\text{AC-G}}| \qquad \boldsymbol{v}^e_{\text{SDP}} = |\boldsymbol{v}_{\text{SDP}} - \boldsymbol{v}_{\text{AC-S}}| \tag{3.4.1}$$

For both formulations, the average and standard deviation are very small showing that the low-pass approximation can provide an OPF solution. While the most optimal solution is not guaranteed through the AC-GSP approach, there are significant improvements in solution time for larger networks (for instance, almost a 15 times speed gain for a 2000 bus synthetic feeder case). These findings illustrate that GSP can be applied to the OPF problem in situations with non-linear voltage constraints and generate a solution much faster than the conventional AC-SDP approach.

— PART III —

Prevention & Mitigation of Cyber-Threats on Power Distribution Systems

CHAPTER 4

A SECURE DISTRIBUTED LEDGER FOR TRANSACTIVE ENERGY: THE ELECTRON VOLT EXCHANGE (EVE) BLOCKCHAIN

This chapter propose the **Electron Volt Exchange (EVE)** blockchain architecture to address the issues of the a blockchain enabled Transactive Energy framework. Novelty compared to other TE blockchain research lies in the following components:

1. TE blockchain designs commonly consider only bilateral transactions. Instead, the EVE approach utilizes a decentralized solution for the entire market economic dispatch problem whose formulation falls in the class of network utility maximization problems, first proposed for real time pricing in [Li *et al.*(2011b)] (Section 4.2). The closest to the proposed approach is found in [Münsing *et al.*(2017)], where the authors have incorporated controllable loads and generation to develop an iterative pricing algorithm using a smart contract that updates global variables of the distributed optimal power flow problem. The scheme still relies on a central update of variables to achieve convergence. Compared to [Münsing *et al.*(2017)], this work incorporates renewable generation, thermostatically controlled loads (TCLs), storage devices, deferrable appliances (DAs), and electric vehicles (EVs) and relies on a hierarchical, distributed architecture including aggregators [Olivella-Rosell *et al.*(2018)] to "divide and conquer" the communication problem, avoiding congestion and yielding a scalable implementation for optimal price calculation.

2. The EVE architecture includes the first blockchain-based, distributed Robust State Verification (RSV) mechanism for TE transactions, where physical sensor measurements are cross-validated in a decentralized fashion to ensure prosumers

abide by their market commitment (Section 4.3). The proposed algorithm, inspired by the work [Vuković and Dán(2014)] on distributed state estimation in adversarial settings, is shown to be robust against False Data Injection Attacks (FDIAs) aimed at TE market theft.

3. The pricing and verification algorithms are tested via numerical simulations in Section 4.4. Implementation of the EVE blockchain framework onto a distributed ledger [Androulaki *et al.*(2018)] is described in Section 4.5 using the open-source Hyperledger Fabric (HLF) framework. It includes a customized BFT-SMART [Sousa *et al.*(2018)] consensus protocol to provide security and improve performance with Byzantine Fault Tolerance (BFT) in an untrustworthy environment. This improves upon standard security features of HLF, such as Membership Service Provider (MSP), Fabric CA (Certificate Authority), Attribute-Based Access Control (ABAC) [Yuan and Tong(2005)], and others that address common security concerns. Bench-marking for the proposed smart contracts using Hyperledger Caliper [Sukhwani *et al.*(2018)] and a web visualisation using Hyperledger Explorer [Aleksieva *et al.*(2020)] are also described herein.

4.1 EVE as a Cyber-Physical System Architecture

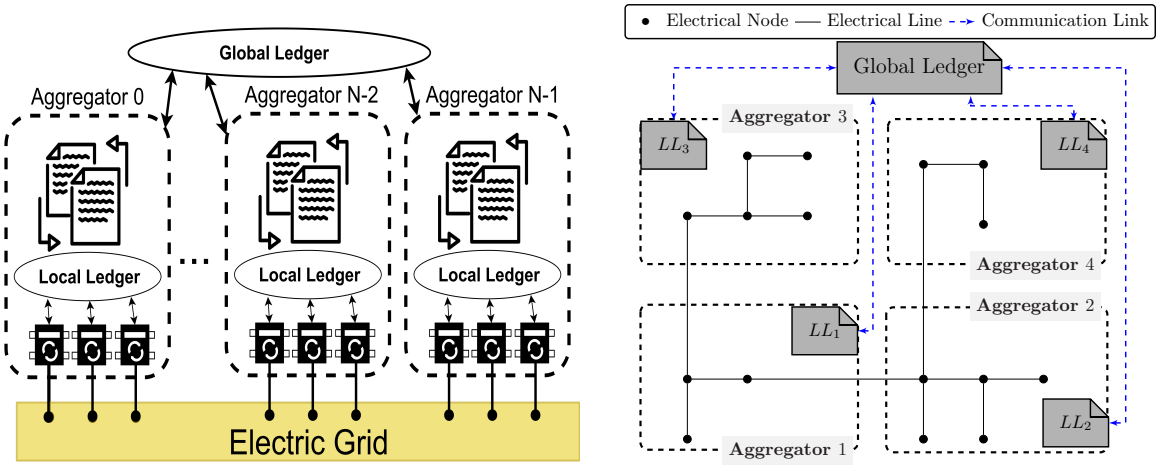
4.1.1 Physical Infrastructure

The physical infrastructure includes:

- The electrical grid modeled as a connected graph $\mathcal{G}_e = (\mathcal{B}, \mathcal{E}_e)$. Lines and transformers are characterized by admittance parameters $y_{ij},\ \forall\ ij \in \mathcal{E}_e$[1].

[1]Voltage control and protective equipment are ignored because these do not have a direct impact on market operation and hence are not required for the description of the EVE architecture.

- Market Participants (prosumers and aggregators).
- Electrical loads, distributed generation, and storage assets that connect to buses on the electrical network. For clarity this formulation models a single prosumer per bus b, making them equivalent[2].
- Electrical sensors and control equipment.



(a) (b)

Figure 4.1.1: (a) The Hierarchical Distributed Ledger Architecture (b) Cyber Infrastructure Overlaid with the Physical Structure for a Sample Distribution Network with 4 Aggregators

4.1.2 Application Layer

In EVE, scalability is achieved by dividing the application layer entities into agents with distinct tasks.

- The bottom layer includes *prosumers* who can buy and/or sell energy and control flexible loads, storage, and generation assets connected to the physical network. This layer can be broken into several physical regions.

[2]In cases where multiple consumers connect to the same physical bus, the modeling considers separate buses connected through zero impedance edges, however omitted in diagrams for simplicity.

- The top layer includes a set of local *aggregators* ($\mathcal{N}$) managing the prosumers connected to a subset of buses $\mathcal{B}^{(n)} \subset \mathcal{B}$ for all $n \in \mathcal{N}$.

The following remark is in order:

Remark 4.1.1. *Individual prosumers have traditionally been unable to participate in energy markets, but aggregators may have access through the lumped capacity bids [Mahmoudi* et al.*(2014a)]*[3]*. Aggregators can act as intermediaries between small consumers/producers and volatile markets, and thereby provide hedging solutions to reduce risk to individual market participants [Koch(2015)]. Aggregators can procure demands from consumers and sell to purchasers through trading frameworks proposed in prior works [Mahmoudi* et al.*(2014b)]. Moreover, resources needed for price optimization and verification processes can be more easily be obtained and justified for aggregators rather than each prosumer. Aggregators do not complicate management because blockchain is suited for a decentralized architecture. A solution without aggregators would otherwise increase communication latency and computation time for market-clearing as new prosumers are added, reducing scalability.*

Policies in EVE can be divided into three main classes:

- **Pricing**: Policies deciding the optimum prosumers schedule and the price.
- **Verification**: Policies processing sensor information to verify that load/generation is correctly reported, contractual obligations have been met and billed accordingly.
- **Billing**: Policies for billing and ensuring compliance.

[3] A practical example can be found in [California's Community Choice Program(2020)].

4.1.3 Cyber Infrastructure

The cyber infrastructure includes security policies for settling transactions, communication/computation resources, and data archival based on blockchain. Building the cyber network requires all market participants to work together as a consortium using a set of policies agreed to during network initialization to determine the participants' permissions. For the shared database or ledger within the cyber architecture, EVE uses CouchDB [Gupta and Rani(2016)] as it supports rich queries when data values are modeled as JSON. The cyber framework is generic to include or exclude Transport Layer Security (TLS); however, including TLS is recommended for additional security.

Application of the policies mentioned in Section 4.1.2 is handled through distributed ledgers. In EVE, a ledger consists of (a) a database that holds current values of a set of ledger states, and (b) a transaction log that records all changes that have resulted in the current system state. The implementation of EVE in this work consists of two types of distributed ledgers, a Local Ledger (LL) and a Global Ledger (GL). The smart contracts in this work handle interactions between the ledgers (GL and LL) and external applications to complete every transaction within EVE. Figure 4.1.1a shows the hierarchical architecture of those ledgers whereas Figure 4.1.1b overlays the distribution of physical nodes into aggregator zones. An aggregator uses the LL to collect bids submitted by prosumers, verify local state information, and update individual prosumer budgets after verification. Aggregators access the GL for distributed pricing and verification algorithms and for sharing global information with other aggregators. All information exchanges and history between participants are handled through smart contracts for reading, writing, and storing in distributed ledgers.

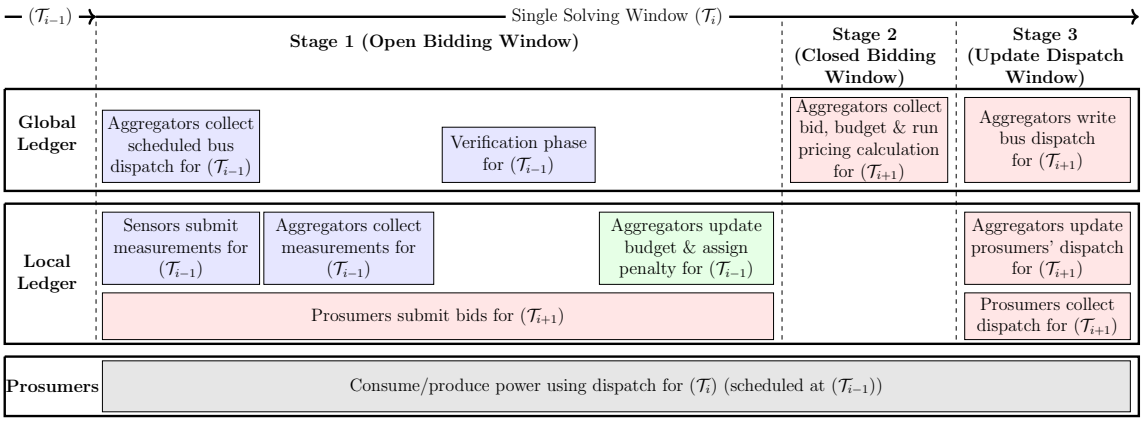


Figure 4.1.2: Overview of Eve Major Tasks and Timeline for a Single Solution Window with Red, Blue, and Green Used to Illustrate Pricing, Verification, and Billing Policy Execution Steps, Respectively

Figure 4.1.2 depicts the implementation of policies under a single solving window $\mathcal{T}_i$ divided into three stages. Stage 1 refers to the Open Bidding Window in which prosumers submit bids for $\mathcal{T}_{i+1}$. Aggregators execute the verification algorithm based on measurements collected for $\mathcal{T}_{i-1}$. Hence stage 1 includes execution of the verification policy and then the billing policy. Stage 2 refers to the Closed Bidding Window in which the aggregators execute the distributed pricing algorithm for $\mathcal{T}_{i+1}$. Stage 3 refers to the Update Dispatch Window in which aggregators update the prosumers' schedules.

4.2 EVE Distributed Pricing Algorithm

Pricing and scheduling decisions are illustrated for a single solving window, as shown in Figure 4.1.2. The period is split into T smaller discrete intervals of unit duration, all in the set $\mathcal{T}_i = \{iT, \ldots, (i+1)T-1\}$. Within each period, $\mathcal{T}_i$, bus b is connected to assets that either supply or demand power. Net real power generation of bus b at time t is denoted by $p_b(t)$. Positive and negative values of $p_b(t)$ indicate

that bus b is supplying power to the grid or consuming from the grid, respectively. Neglecting losses, the total power schedule managed by aggregator $n \in \mathcal{N}$, $p^{(n)}(t)$, is the sum of power from each individual bus $p_b(t)$ associated with it:

$$p^{(n)}(t) = \sum_{b \in \mathcal{B}^{(n)}} p_b(t). \tag{4.2.1}$$

Each component $p_b(t)$ must have a certain cost (disutility) and must satisfy a set of constraints that depend on the generation and storage capacity available at the supply side, as explained in Section 4.2.1, that determines the optimal schedule. Section 4.2.2 describes a decentralized dual decomposition algorithm to solve a power balancing problem between different aggregators with a dual variable reflecting the price of energy. Related works ([Li *et al.*(2011b), Chang *et al.*(2012)]) use a similar dual decomposition algorithm to solve a distributed problem between an aggregator and its customers, with the former reflecting its internal energy procurement cost function through iterative retail pricing and the latter trying to minimize deviation from a pre-determined aggregate power profile. However. this work relies on only modeling a decentralized energy market mechanism (hence no central provider) while leveraging the distributed ledger to ensure the liquidity of purchasers.

4.2.1 Flexible Resource Model

The power injection trajectory $\boldsymbol{p}_b = [p_b(iT), \ldots, p_b((i+1)T-1)]^\top \in \mathbb{R}^{T \times 1}$ is constrained depending on the type of energy service bus b provides. For example, in response to price signals, a participant may accept shifting to a less comfortable thermostat reference temperature, defer use of the dishwasher, dim lights, or other actions. This section describes in general terms[4] the inter-temporal constraints for demand and supply in a single period $\mathcal{T}_i$ (i is omitted for brevity) and discusses

[4]Specific examples are given in Appendix A.

associated cost functions $C_b(\boldsymbol{p}_b)$.

For each aggregator n and bus b, the load profile can be split into a flexible component, that changes based on price, and an inflexible one that prosumers are willing to buy at any price. In the literature, $\boldsymbol{p}_b$ is typically modeled by a linear, affine function of a corresponding control signal $\boldsymbol{u}_b$, i.e., for all $b \in \mathcal{B}^{(n)}, n \in \mathcal{N}$:

$$\boldsymbol{p}_b = \boldsymbol{A}_b \boldsymbol{u}_b + \boldsymbol{\ell}_b, \quad p_b(t) \in \mathcal{S}_b^p, \quad \boldsymbol{u}_b \in \mathcal{U}_b \tag{4.2.2}$$

where $\boldsymbol{\ell}_b$ is the inflexible part of the load, $\boldsymbol{p}$ and $\boldsymbol{u}$ are column vectors, the set $\mathcal{U}_b$ is related to the flexibility that can be offered to adjust the shape of the profile, and $\mathcal{S}_b^p$ expresses operational constraints on how the asset can inject power[5]. Control signal constraints are mapped to $\boldsymbol{p}_b$ using $\boldsymbol{A}_b^\dagger$ as follows:

$$\boldsymbol{p}_b \in \mathcal{P}_b, \quad \mathcal{P}_b = \{\boldsymbol{p} | \boldsymbol{A}_b^\dagger (\boldsymbol{p} - \boldsymbol{\ell_b}) \in \mathcal{U}_b, \quad p(t) \in \mathcal{S}_b^p, \quad \boldsymbol{p}^\top \boldsymbol{\lambda} \leq r_b\} \tag{4.2.3}$$

where the constraints $\mathcal{U}_b$ are linear, meaning that $\boldsymbol{p}_b$ lies within a polygon. In (4.2.3), $\boldsymbol{\lambda}$ denotes the price of energy over the horizon, r_b denotes the budget, and $\boldsymbol{p}_b^\top \boldsymbol{\lambda} \leq r_b$ denotes the affordability constraint.

The cost to prosumer b, $C_b(\boldsymbol{p}_b)$, is the price the customer is willing to pay, or the price the supplier is willing to be paid to generate for a certain amount of power. $C_b(\boldsymbol{p}_b)$ is a convex function of $\boldsymbol{p}_b$. The cost function simultaneously reflects the utility and cost for prosumers that can switch between producing and consuming, respectively, with $p_b(t) \geq 0$ representing the supply utility function and $p_b(t) < 0$ representing the demand cost function.

This generic model can be applied to various resources including renewable generation, TCLs, storage devices, DAs, and EVs. Appendix A shows examples of such

[5]Later integrality constraints will be relaxed in either $\mathcal{U}_b$ or $\mathcal{S}_b^p$, and be replaced with convex constraints to guarantee a convex market clearing problem required for the convergence of the pricing algorithm.

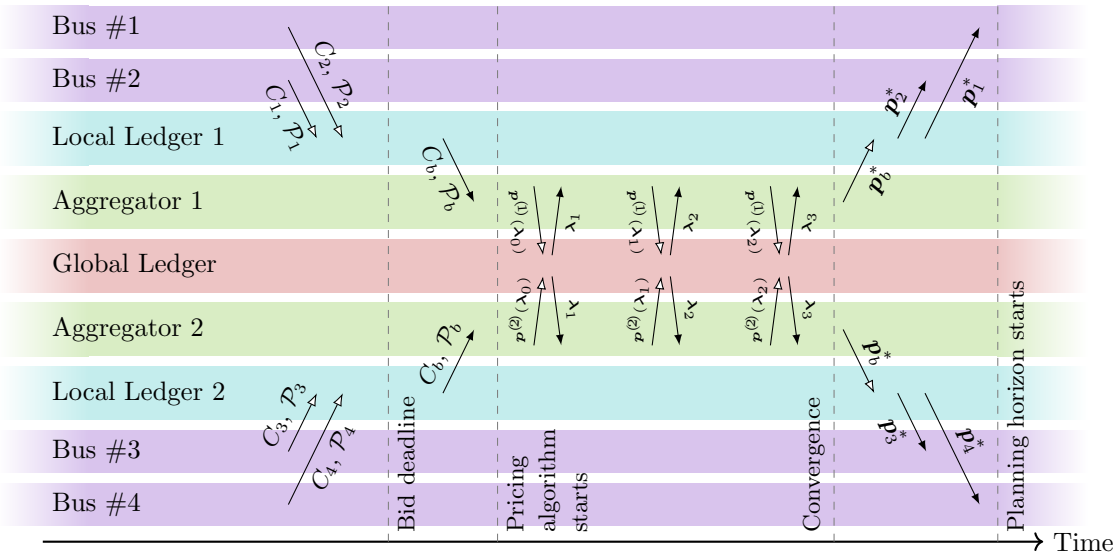


Figure 4.2.1: A Single Solving Window for Distributed Pricing Ordered from (i) Individual Buses Submit Their Costs and Constraints, (ii) Aggregators Collect Bus Costs and Constraints, (iii) Aggregators Iterate Pricing Algorithm, (iv) Aggregators Push Dispatch Solution to Ledger, and (V) Individual Buses Read Dispatch Solution.

models and cost functions for supplemental study.

4.2.2 Distributed Pricing and Scheduling Algorithm

Each aggregator can buy or sell power for the distributed resources connected to $\mathcal{B}^{(n)}$. Let $\mathcal{B}^{(n)} = \{n(1), \ldots, n(|\mathcal{B}^{(n)}|)\}$, where $n(i)$ denotes the i^{th} bus of aggregator n. The power profile of aggregator n is:

$$\boldsymbol{p}^{(n)} = \left(p^{(n)}(iT), \ldots, p^{(n)}((i+1)T-1)\right)^\top = \sum_{b \in \mathcal{B}^{(n)}} \boldsymbol{p}_b \tag{4.2.4}$$

where $\boldsymbol{p}_b$ is the profile of one of the flexible resources at a bus $b \in \mathcal{B}^{(n)}$. Following matrices are also defined:

$$\boldsymbol{P}^{(n)} = \left(\boldsymbol{p}_{n(1)}, \ldots, \boldsymbol{p}_{n(|\mathcal{B}^{(n)}|)}\right)^\top \tag{4.2.5}$$

whose rows are prosumers' profiles contributing to $\boldsymbol{p}^{(n)}$. The aggregated load profile is the sum of individual components, i.e., $\boldsymbol{p}^{(n)} = {\boldsymbol{P}^{(n)}}^\top \mathbf{1}$, where $\mathbf{1} = \{1\}^{\left|\mathcal{B}^{(n)}\right|\times 1}$. The demand cost for aggregator n for a certain set of schedules $\boldsymbol{P}^{(n)}$ is described by:

$$\mathfrak{C}^{(n)}(\boldsymbol{P}^{(n)}) \triangleq \sum_{b\in\mathcal{B}^{(n)}} C_b(\boldsymbol{p}_b) \tag{4.2.6}$$

The feasible set of matrix profiles $\boldsymbol{P}^{(n)}$ that can be chosen by the aggregator must lie in the Cartesian product of the feasible sets for the tuple of profiles $\boldsymbol{p}_b \in \mathcal{P}_b$, $b \in \mathcal{B}^{(n)}$, meaning:

$$\boldsymbol{P}^{(n)} \in \mathcal{P}_{n(1)} \times \cdots \times \mathcal{P}_{n(|\mathcal{B}^{(n)}|)} \triangleq \mathcal{P}^{(n)} \tag{4.2.7}$$

Algorithm 4 Prosumer Pricing Interaction; A Step-by-step Implementation of eqs. (4.2.9) to (4.2.11) from the Perspective of a Prosumer $b \in \mathcal{B}^{(n)}$. LL_n Refers to the Local Ledger to Aggregator $n \in \mathcal{N}$, with $\Leftarrow$ and $\Rightarrow$ Indicating Writing to and Reading from the Ledger, Respectively.

1: Wait until bidding interval starts:
2: $LL_n \Leftarrow \mathcal{P}_b$ (constraints) and C_b (cost function).
3: Wait until dispatch horizon starts.
4: **if** Dispatch instruction ready on ledger **then**
5: $\quad LL_n \Rightarrow$ Read dispatch instruction $\boldsymbol{p}_b$.
6: $\quad$ Execute dispatch instruction $\boldsymbol{p}_b$.
7: **else**
8: $\quad$ Execute most recent $\boldsymbol{p}_b$ instruction.
9: **end if**

If transmission constraints are relaxed and the algorithm only balances instantaneous power, the optimum market clearing requires solving:

$$\min_{\boldsymbol{P}^{(1)},\ldots,\boldsymbol{P}^{(N)}} \quad \sum_{n\in\mathcal{N}} \mathfrak{C}^{(n)}(\boldsymbol{P}^{(n)}) \quad \text{s.t.} \quad \sum_{n\in\mathcal{N}} {\boldsymbol{P}^{(n)}}^\top \mathbf{1} = \mathbf{0},\ \boldsymbol{P}^{(n)} \in \mathcal{P}^{(n)} \tag{4.2.8}$$

The Lagrange multiplier of the balance constraint $\sum_{n\in\mathcal{N}} {\boldsymbol{P}^{(n)}}^\top \mathbf{1} = 0$ in (4.2.8) $\boldsymbol{\lambda}$ expresses the shadow price of energy over the horizon in (4.2.3), $\boldsymbol{\lambda} = (\lambda(iT), \ldots, \lambda((i+$

$1)T-1))^\top$. Also, this is an instance of the network utility maximization problems which can be decomposed as detailed next[6].

For each gradient descent, given the most recent price $\boldsymbol{\lambda}_k$ (the dual variable), each aggregator independently attempts to minimize its cost schedule through the following problem:

$$\min_{\boldsymbol{P}^{(n)}} \mathfrak{C}^{(n)}\boldsymbol{P}^{(n)} + [\boldsymbol{\lambda}_k]^\top {\boldsymbol{P}^{(n)}}^\top \mathbf{1} \;\text{ s.t. }\; \boldsymbol{P}^{(n)} \in \mathcal{P}^{(n)}, \;\; \forall n \in \mathcal{N} \tag{4.2.9}$$

Let $\boldsymbol{P}_\star^{(n)}(\boldsymbol{\lambda}_k) \in \mathcal{P}^{(n)}$ be the solution of eq. (4.2.9) in response to the k^{th} iteration value of the vector $\boldsymbol{\lambda}_k$ and $\boldsymbol{p}_\star^{(n)}(\boldsymbol{\lambda}_k) = [\boldsymbol{\lambda}_k]^\top {\boldsymbol{P}_\star^{(n)}}^\top \mathbf{1}$. Assuming a feasible solution exists, the algorithm updates the price as follows in GL:

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha \sum_{n\in\mathcal{N}} {\boldsymbol{p}_\star^{(n)}}^\top \tag{4.2.10}$$

Note that all the aggregators have to post their total injection based on the current price estimate, which will stop updating as soon as:

$$\boldsymbol{\lambda}_\star = \boldsymbol{\lambda}_{k^\star} : \quad \sum_{n\in\mathcal{N}} \boldsymbol{p}_\star^{(n)} = \mathbf{0} \tag{4.2.11}$$

These equations comprise the distributed and decentralized algorithms explained for prosumers and aggregators in Algorithms 4 and 5, respectively, and are visualized in Figure 4.2.1.

A few interesting observations can be made:

1. Aggregators hide local bids and feasibility constraints from the GL, and instead they only show (expose) how they would react for the particular price scenarios iterated through $\boldsymbol{\lambda}_k$.

2. Because the cost and constraints of an aggregator are decomposable in terms of the prosumer profiles $\boldsymbol{p}_b$, the problem could be decomposed further to al-

[6]See e.g. [Li *et al.*(2011a)] for its application to real time pricing.

Algorithm 5 Aggregator Pricing Algorithm; A Step-by-step Implementation of eqs. (4.2.9) to (4.2.11) from the Perspective of Aggregator $n \in \mathcal{N}$

Define algorithm time limit $\overline{\tau}$ and iteration limit $\overline{k}$.
Wait until prosumer bidding interval ends.
$LL_n \Rightarrow$ read all C_b, r_b, and $\mathcal{P}_b$ values available on ledger, building $\mathcal{B}^{(n)}$ based on submitted prosumers.
Initialize $\boldsymbol{\lambda}_{\triangle}$ and $\boldsymbol{p}_{\triangle}^{(n)}$ to most recent solutions $\boldsymbol{\lambda}_{\star}$ and $\boldsymbol{p}_{\star}^{(n)}$, or $\mathbf{0}$ if no prior solution exists.
Initialize $\boldsymbol{\lambda}_0 \leftarrow \boldsymbol{\lambda}_{\triangle}$, $k \leftarrow 0$ and $\hat{\alpha}$.
Build model eq. (4.2.9) using $\boldsymbol{\lambda}_0$.
while $k < \overline{k}$ **do**
    Solve model eq. (4.2.9) for $\boldsymbol{\lambda}_k$, enforcing the billing constraint for all prosumers.
    Retry writing solution $\boldsymbol{p}_{\star}^{(n)}$ to GL until ACK received.
    Start timer $\tau \leftarrow 0$.
    while Other aggregator solutions are not available from GL **do**
        if $\tau < \overline{\tau}$ **then**
            Wait ε seconds.
        else
            $LL_n \Leftarrow \boldsymbol{p}_b^{\triangle}$ for all prosumers $b \in \mathcal{B}^{(n)}$.
            $LL_n \Leftarrow r_b \leftarrow r_b + {\boldsymbol{p}_b^{\triangle}}^{\top} \boldsymbol{\lambda}_{\triangle}$ for all prosumers $b \in \mathcal{B}^{(n)}$.
            Terminate algorithm, recycling last solutions.
        end if
    end while
    $GL \Rightarrow {\boldsymbol{p}_{\star}^{(m)}}^{\top} \boldsymbol{\lambda}_k$ for all $m \in \mathcal{N} \setminus \{n\}$.
    Update $\boldsymbol{\lambda}_{k+1}$ following eq. (4.2.10) using $\alpha = \hat{\alpha}/(k+1)$.
    if $|\boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k| < \epsilon$ **then**
        $\boldsymbol{\lambda}_{\star} \leftarrow \boldsymbol{\lambda}_k$.
        $LL_n \Leftarrow \boldsymbol{p}_b^{\star}$ for all prosumers $b \in \mathcal{B}^{(n)}$.
        $LL_n \Leftarrow r_b \leftarrow r_b + {\boldsymbol{p}_b^{\star}}^{\top} \boldsymbol{\lambda}_{\star}$ for all prosumers $b \in \mathcal{B}^{(n)}$.
        Terminate.
    end if
    $k \leftarrow k + 1$.
end while
Terminate.

low individual prosumers to keep $C_b(\boldsymbol{p}_b)$ and $\mathcal{P}_b$ private, interacting with the aggregator in a similar manner as shown in [Karakoç *et al.*(2018)].

There are challenges with convergence of the distributed algorithm eq. (4.2.9) not commonly found in comparable centralized formulations. First, the aggregate supply/demand curves must cross for a fixed price to emerge, a standard requirement in market theory. Second, a rogue aggregator may produce malicious values of $p^n_\star$ during the iterative phase of Algorithm 5, potentially preventing the algorithm from converging. There are numerous ways to detect such manipulations such as bounding the gradient step eq. (4.2.10), however, the details of those are beyond the scope of this paper and readers are referred to [Ravi and Scaglione(2019)] for additional information. Third, due to the non-convexity introduced by the budget constraint in eq. (4.2.3), there are no guarantees for convergence to the global optimum. However, as this constraint is only tight for a small number of prosumers, the problem often converges in practice.

4.3 EVE Distributed Robust State Verification Algorithm

While the blockchain TE framework ensures transparency and immutability of transactions placed on the chain, it is still insufficient to ensure that those transactions took place. Physical measurements are needed to confirm power injections at each bus in the electric grid. In a trust-less system, aggregators have to cooperate to verify measurement accuracy, noting that some aggregators (or some of the prosumers they manage) may operate as adversaries by modifying market operations and/or measurements to commit energy theft. The essential tool described in Section 4.3.1 establishes accuracy by checking consistency with the grid's physical laws.

Before delving into algorithmic details, it is worth emphasizing that the notion of *distributed measurement verification* introduced here is new. The approach proposed

for measurement verification consists of solving a regression problem closely related to state estimation in power systems fitting sensor data, but its goal is fundamentally different. Securely estimating the entire state vector for the grid is sufficient but not necessary to cross-validate the self-reported measurements. The state variables themselves are therefore not essential, and rather, this work focuses on the accuracy of reported *real power injections* within an aggregator region and *power flows between* different aggregators' regions. This approach ensures appropriate billing aggregators and prosumers and permits penalties to be levied for discrepancies in reporting. For measurements of the consumption/generation in time window $\mathcal{T}_i$, verification mechanism is run in time window $\mathcal{T}_{i+1}$ as shown in Figure 4.1.2. Considering an FDIA threat [Liu *et al.*(2011)], the proposed mechanism extends the idea in [Vuković and Dán(2014)] and adapts it to serve as a decentralized cross-validation algorithm integrated within the EVE framework. The goal is to extrapolate the actual power injections $\boldsymbol{p}^{(n)}$ of each aggregator $n \in \mathcal{N}$ from sensor measurements. Since $\boldsymbol{p}^{(n)}$ are continuous variables, this is not a binary decision and amounts to solving a regression problem.

4.3.1 Physical Constraints for the Electric Grid

Let $\boldsymbol{x}(t)$ be the vector of system variables at time t, consisting of bus injection variables and branch flow variables. Here, the variables in $\boldsymbol{x}(t)$ include real power injection, reactive power injection, and squared voltage magnitude expressed as $(p_b(t), q_b(t), v_b^2(t))^\top,\ \forall b \in \mathcal{B}$. Branch flow variables include the squared current magnitudes, real power flows, and reactive power flows expressed as $(c_l^2(t), P_l(t), Q_l(t))^\top,\ \forall l \in \mathcal{E}_e$. Let $\boldsymbol{x}_{\mathcal{A}}(t)$ include only the variables in $\boldsymbol{x}(t)$ that are directly-measured through a sensor. Measurements are noisy versions of physical parameters described by:

$$\boldsymbol{z}(t) = \boldsymbol{x}_{\mathcal{A}}(t) + \epsilon(t). \tag{4.3.1}$$

Within a margin of error due to the noise, these physical constraints are a set of non-linear homogeneous equations written in vector form as follows:

$$\boldsymbol{h}(\boldsymbol{x}(t)) = 0 \ ; \quad \forall t \in \mathcal{T} \tag{4.3.2}$$

Appendix B specifies a possible $\boldsymbol{h}(\cdot)$ using the Distflow [Baran and Wu(1989b)] equations .

4.3.2 Malicious Agents Behavior

The following assumptions are made:

1. The adversary (a malicious agent or a group of coordinating agents) is an insider who has legitimate physical and logical access to the network and ledgers through the certification mechanism.

2. The adversary is capable of manipulating sensors measurements, either by compromising sensors or compromising the communication between sensors and aggregators.

3. The insider is motivated to disrupt the verification process and cheat the system for financial gain.

Utilizing sensor measurements reported at the LL level in eq. (4.3.1) and physical constraints in eq. (4.3.2), a decentralized optimization algorithm to complete the verification task using data from all sensors under any aggregator is formulated. The optimization algorithm (detailed later in Section 4.3.3.1) may be generalized into the following form:

$$\min_{\boldsymbol{x}^{(n)}} \quad \sum_{n \in \mathcal{N}} f^{(n)}(\boldsymbol{x}^{(n)}) \tag{4.3.3a}$$

$$\text{s.t.} \quad \text{Consensus Constraints} \tag{4.3.3b}$$

where n is a decentralized agent in set $\mathcal{N}$[7] and $f^{(n)} : \mathbb{R}^l \to \mathbb{R}$ is a cost function that agent n has to minimize while cooperatively minimizing the aggregate of cost functions from all the agents eq. (4.3.3a), subject to some consensus constraints eq. (4.3.3b). Common algorithms used to solve this problem using a certain communication graph $\mathcal{G}_c(\mathcal{N}, \mathcal{E}_c)$ eq. (4.3.3) involve iterative consensus updates to $\boldsymbol{x}^{(n)} \in \mathbb{R}^l$ as follows:

$$\boldsymbol{x}_{k+1}^{(n)} = \sum_{m \in \mathcal{N}} a_{nm} \boldsymbol{r}^{(m)}(\boldsymbol{x}_k^{(m)}) \quad \text{at} \quad k \geq 0 \tag{4.3.4}$$

where k is the iteration index, $a_{nm} \geq 0$ is a mixing weight $(n, m) \in \mathcal{E}$ such that $\sum_{m \in \mathcal{N}} a_{nm} = 1$, and $\boldsymbol{r}^{(m)} : \mathbb{R}^l \to \mathbb{R}^l$ is the information received by n from neighbor m. eq. (4.3.4) illustrates that neighbors on the communication graph exchange information with one another in each iteration.

The communications model is described as:

$$\boldsymbol{r}^{(m)}(\boldsymbol{x}_k^{(m)}) = \begin{cases} \boldsymbol{x}_k^{(m)}, & m \in \mathcal{R} \\ \star, & m \in \mathcal{M} \end{cases} \tag{4.3.5}$$

with $\mathcal{R} \subseteq \mathcal{N}$ and $\mathcal{M} \subseteq \mathcal{N}$ expressing the sets of regular and malicious agents, respectively. That is, regular agents report their true states, whereas malicious agents may inject false data and disrupt convergence of the algorithm to suit their goals.

An attack by an agent results in the following update for a neighbor n:

$$\widetilde{\boldsymbol{x}}_{k+1}^{(n)} = \boldsymbol{x}_{k+1}^{(n)} + \star \tag{4.3.6}$$

where $\widetilde{\boldsymbol{x}}_{k+1}^{(n)}$ is the false update and $\boldsymbol{x}_{k+1}^{(n)}$ is the true update (if all the neighbors communicated truthfully). For instance, a malicious agent will attempt theft by paying less (as a consumer) or receiving a larger compensation (as a producer). This occurs by under reporting energy usage or over reporting generation by altering the appropriate field of $\boldsymbol{z}^{(n)}(t)$ so that: $\widetilde{\boldsymbol{z}}^{(n)}(t) = \boldsymbol{z}^{(n)}(t) + \star$.

[7]The notation for the set of aggregators is abused $\mathcal{N}$ to denote the set of decentralized agents in this subsection.

4.3.3 Robust State Verification in the Presence of FDIAs

Next, details on the decentralized algorithm that allows aggregators to cross verify if measurements of power injections are to be trusted and, if not, which aggregators are likely responsible for the FDIA is provided.

4.3.3.1 Modelling Of The Optimization Problem

The state verification problem is modeled as a decentralized optimization problem in which each aggregator $n \in \mathcal{N}$ has access only to their private cost function $f^{(n)} : \mathbb{R}^{l_n} \to \mathbb{R}$ and act on their own private vector of system variables, $\boldsymbol{x}^{(n)}(t)$. Here, $\boldsymbol{x}^{(n)}(t)$ includes copies of those variables in $\boldsymbol{x}(t)$ that pertain to buses and lines inside and at the periphery of aggregator region n (see Figure 4.3.1). That is,

$$\boldsymbol{x}^{(n)}(t) = \boldsymbol{S}^{(n)}\boldsymbol{x}(t) \tag{4.3.7}$$

where $\boldsymbol{S}^{(n)}$ is a selection matrix that extracts the appropriate entries that make up $\boldsymbol{x}^{(n)}(t)$ from $\boldsymbol{x}(t)$. Since there are tie-lines between aggregator regions, some of the entries in $\boldsymbol{x}^{(n)}(t)$ will have identical counterparts in $\boldsymbol{x}^{(m)}(t)$ of a neighboring aggregator region m. For the extrapolated injections to be valid, the regions have to match values at these tie-lines. Failure to do so indicates an attack. The consensus constraint is, therefore, a consensus on the tie-lines variables, that can be written as:

$$\boldsymbol{S}_{nm}\boldsymbol{x}^{(n)}(t) = \boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t), \qquad \forall n \in \mathcal{N} \text{ and } m : nm \in \mathcal{E}_c \tag{4.3.8}$$

where $\boldsymbol{S}_{nm}$ is the selection matrix that extracts from $\boldsymbol{x}^{(n)}$ common variables between neighboring regions n and m.

Thus, the goal of the set of aggregators is to minimize their individual cost functions $(f^{(n)}, \forall n \in \mathcal{N})$ and the global cost function $(\sum_{n\in\mathcal{N}} f^{(n)})$ simultaneously subject to the consensus constraints eq. (4.3.8) defined over the communication graph,

$\mathcal{G}_c(\mathcal{N}, \mathcal{E}_c)$.

The optimization problem should find values of system variables, $\boldsymbol{x}^{(n)}(t)$, at each aggregator n that:

- Have the least residual error with respect to the measurements, $\boldsymbol{z}^{(n)}(t)$, to reduce measurement deviation from their corresponding system variables ($\boldsymbol{x}_{\mathcal{A}}^{(n)}(t)$) being minimized. Letting $\boldsymbol{S}_{\mathcal{A}}^{(n)}$ be the selection matrix that extracts available measurements from $\boldsymbol{x}^{(n)}(t)$, the following can be written:

$$\boldsymbol{x}_{\mathcal{A}}^{(n)}(t) = \boldsymbol{S}_{\mathcal{A}}^{(n)} \boldsymbol{x}^{(n)}(t) \tag{4.3.9}$$

- Have the least residual error with respect to the scheduled injections, $\mathfrak{p}^{\star(n)}(t)$, where the i^{th} element of the vector is given by $[\mathfrak{p}^{\star(n)}(t)]_i = p_{b_i}(t),\ \forall b_i \in \mathcal{B}^{(n)}$. This seeks to minimize deviation of the scheduled bus dispatch from their corresponding system variables ($\boldsymbol{S}_{\mathcal{P}}^{(n)} \boldsymbol{x}^{(n)}(t)$), where $\boldsymbol{S}_{\mathcal{P}}^{(n)}$ is the selection matrix that extracts active power injections variables from $\boldsymbol{x}^{(n)}(t)$, and

- Fit the physical model equations, $\boldsymbol{h}^{(n)}(\cdot)$[8], with the least residual error.

Power injection components, $\boldsymbol{S}_{\mathcal{P}}^{(n)} \boldsymbol{x}^{(n)}$, from the optimization problem are used in EVE as the closest approximation for ground-truth billing. The severity of deviations from the schedule $\left(\mathfrak{p}^{\star(n)}(t) - \boldsymbol{S}_{\mathcal{P}}^{(n)} \boldsymbol{x}^{(n)}\right)$ that determines the penalties assigned to prosumers under a specific bus.

Finally, the state verification problem can be cast as the following optimization problem, written in a form analogous to eq. (4.3.3a) that is amenable to ADMM

[8]Definitions of the functions $\boldsymbol{h}^{(n)}(\cdot)$ are given in Appendix B.

decomposition [Boyd *et al.*(2011)]:

$$\min_{\{\boldsymbol{x}^{(n)}(t)\}_{n\in\mathcal{N}}^{t\in\mathcal{T}}} \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}} f^{(n)}(\boldsymbol{x}^{(n)}(t)) \tag{4.3.10a}$$

$$\text{s.t. } \boldsymbol{S}_{nm}\boldsymbol{x}^{(n)}(t) = \boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t)\ , \quad \forall nm \in \mathcal{E}_c \tag{4.3.10b}$$

$$\boldsymbol{x}^{(n)}(t) = \boldsymbol{S}^{(n)}\boldsymbol{x}(t)\ , \quad \forall n \in \mathcal{N} \tag{4.3.10c}$$

$$\boldsymbol{x}_{\mathcal{A}}^{(n)}(t) = \boldsymbol{S}_{\mathcal{A}}^{(n)}\boldsymbol{x}(t), \quad \forall n \in \mathcal{N} \tag{4.3.10d}$$

where:

$$f^{(n)}(\boldsymbol{x}^{(n)}(t)) := c_1\left\|\boldsymbol{\Sigma}_{z^{(n)}}^{-1/2}\left(\boldsymbol{z}^{(n)}(t) - \boldsymbol{x}_{\mathcal{A}}^{(n)}(t)\right)\right\|_2^2 + c_2\left\|\boldsymbol{h}^{(n)}(\boldsymbol{x}^{(n)}(t))\right\|_2^2 + c_3\left\|\mathfrak{p}^{\star(n)}(t) - \boldsymbol{S}_{\mathcal{P}}^{(n)}\boldsymbol{x}^{(n)}(t)\right\|_2^2 \tag{4.3.11}$$

Here, $\boldsymbol{\Sigma}_{z^{(n)}}$ is the diagonal matrix such that $[\boldsymbol{\Sigma}_{z^{(n)}}]_{jj} = \text{variance}([\boldsymbol{z}^{(n)}]_j)$. Additionally, eq. (4.3.10b) enforces consensus among the common variables across neighboring regions for the constraints mentioned in eq. (4.3.3b).

Remark 4.3.1. *The terms c_1, c_2, c_3 denote the weight of the penalty levied on the solution if their respective objectives are violated. The third term penalizes a solution that is further from the schedule, and in noting this might hinder the verification step, practical applications can choose c_3 such that $c_3 \lll c_1, c_2$.*

4.3.3.2 Iterative Update Rule

This section shows equations that iteratively update state variables vectors to arrive at the solution to problem eq. (4.3.10). Two matrices are defined that assist in formulating rules for iterative updates. Suppose that l_n is the number of variables pertaining to aggregator region n, then $\boldsymbol{D}^{(n)} \in \mathbb{R}^{l_n \times l_n}$ is a diagonal matrix with $[\boldsymbol{D}^{(n)}]_{jj}$ equal to the number of regions with which region n has the j-th variable of

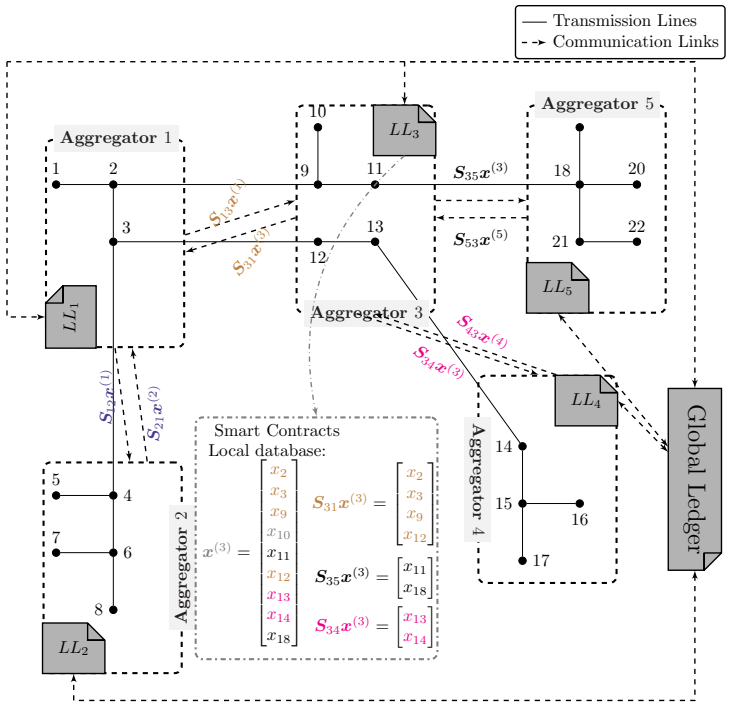


Figure 4.3.1: State Verification Architecture for a Test Case with Five Aggregator Regions

$\boldsymbol{x}^{(n)}$ in common, $\forall j \in [l_n]$. Similarly $\overline{\boldsymbol{D}}^{(n)} \in \mathbb{R}^{l_n \times l_n}$ is also a diagonal matrix such that $[\overline{\boldsymbol{D}}^{(n)}]_{jj} = 1/[\boldsymbol{D}^{(n)}]_{jj}$ if $[\boldsymbol{D}^{(n)}]_{jj} \neq 0$, and $[\overline{\boldsymbol{D}}^{(n)}]_{jj} = 0$ if $[\boldsymbol{D}^{(n)}]_{jj} = 0$.

To arrive at the minimizer of the problem in eq. (4.3.10) for an interval $t \in \mathcal{T}_i$, the updates in eq. (4.3.12) are executed iteratively $\forall k \geq 0$ until a termination condition is satisfied[9]:

$$\boldsymbol{x}_{k+1}^{(n)} = \left(c_1 \boldsymbol{S}_{\mathcal{A}}^{(n)^\top} \boldsymbol{\Sigma}_{z^{(n)}}^{-1} \boldsymbol{S}_{\mathcal{A}}^{(n)} + c_2 \boldsymbol{H}^{(n)^\top} \boldsymbol{H}^{(n)} + c_3 \boldsymbol{S}_{\mathcal{P}}^{(n)^\top} \boldsymbol{S}_{\mathcal{P}}^{(n)} + c_4 \boldsymbol{D}^{(n)}\right)^{-1} \times \left(c_1 \boldsymbol{S}_{\mathcal{A}}^{(n)^\top} \boldsymbol{\Sigma}_{z^{(n)}}^{-1} \boldsymbol{z}^{(n)} + c_3 \boldsymbol{S}_{\mathcal{P}}^{(n)^\top} \mathfrak{p}^{\star(n)} + c_4 \boldsymbol{D}^{(n)} \boldsymbol{v}_k^{(n)}\right) \tag{4.3.12a}$$

$$\boldsymbol{\psi}_{k+1}^{(n)} = \overline{\boldsymbol{D}}^{(n)} \sum_{m:nm \in \mathcal{E}_c} \boldsymbol{S}_{nm}^\top \boldsymbol{S}_{mn} \boldsymbol{x}_{k+1}^{(m)} \tag{4.3.12b}$$

$$\boldsymbol{v}_{k+1}^{(n)} = \boldsymbol{v}_k^{(n)} + \boldsymbol{\psi}_{k+1}^{(n)} - 0.5\left(\boldsymbol{\psi}_k^{(n)} + \boldsymbol{x}_k^{(n)}\right) \tag{4.3.12c}$$

[9]For the sake of readability, $\boldsymbol{x}^{(n)}$ is written instead of $\boldsymbol{x}^{(n)}(t)$ and $\boldsymbol{x}_k^{(n)}$ is used to denote iterates of the algorithm where k is the iterate.

where $\boldsymbol{H}^{(n)}$ is defined in Appendix B with variables initialized as:

$$\boldsymbol{x}_0^{(n)} \in \mathbb{R}^{l_n} \qquad \dots \text{initialize arbitrarily} \tag{4.3.13a}$$

$$\boldsymbol{\psi}_0^{(n)} = \overline{\boldsymbol{D}}^{(n)} \sum_{m:nm\in\mathcal{E}_c} \boldsymbol{S}_{nm}^{\top}\boldsymbol{S}_{mn}\boldsymbol{x}_0^{(m)} \tag{4.3.13b}$$

$$\boldsymbol{v}_0^{(n)} = \frac{1}{2}\left(\boldsymbol{\psi}_0^{(n)} + \boldsymbol{x}_0^{(n)}\right). \tag{4.3.13c}$$

Algorithm 6 describes the iterative process in eq. (4.3.12) with a pictorial representation of the algorithm given in Figure 4.3.1. Buses in each aggregator region n have access to their local ledgers LL_n. The aggregators, in addition to their local ledger, also have access to GL. The aggregator collects measurements from their buses and stores them on LL_n. Results after an ADMM update are also saved in LL_n for each aggregator. Aggregators exchange state elements that correspond to tie-line variables with a neighbor as shown in Figure 4.3.1.

4.3.3.3 Threat Model Specific To RSV

The updates in eqs. (4.3.12b) to (4.3.12c) involve aggregating the shared state variable $\boldsymbol{x}^{(n)}$ from agent n with the same state variable from its neighbors, similar to the updates in eq. (4.3.4). Yet Section 4.3.2 showed that approach to be vulnerable to FDI attacks via eq. (4.3.6), thus indicating eq. (4.3.12b) is also vulnerable to malicious injections by neighbors.

A malicious user in region m may modify values of its input measurements to affect the aggregator's measurement vector as follows:

$$\widetilde{\boldsymbol{z}}^{(m)} = \boldsymbol{z}^{(m)} + \boldsymbol{a}^{(m)}$$

where the perturbation $\boldsymbol{a}^{(m)}$ has non-zero entries in the locations that correspond to false sensor measurements. Similarly, if the aggregator itself acts maliciously, it can inject false data (as seen in eq. (4.3.5)) into the updates of $\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}$ that are passed

Algorithm 6 Robust State Verification; A Step-by-step Implementation From the Perspective of Aggregator n $\forall n \in \mathcal{N}$. Here LL_n $\forall n \in \mathcal{N}$ and GL Represents the Local and Global Ledgers Respectively. The Symbol $a \Leftarrow b$ Corresponds to Upload From b to a.

1: $LL_n \Leftarrow$ Collect local measurements $\boldsymbol{z}^{(n)}(t)$.
2: Initialize ADMM states according to eq. (4.3.13), trust score $\boldsymbol{\pi} = \mathbf{0}$, and disagreements $d_{nm} = 0,\ \forall nm \in \mathcal{E}_c$.
3: **repeat**
4: $\quad \boldsymbol{\pi}^- \leftarrow \boldsymbol{\pi}$ and $[\boldsymbol{x}^{(n)}(t)]^- \leftarrow \boldsymbol{x}^{(n)}(t)$.
5: $\quad$ ADMM update of $\boldsymbol{x}^{(n)}(t)$ according to equation eq. (4.3.12a).
6: $\quad$ Send common variable information to neighbors via the GL with ABAC control: $GL \Leftarrow \{\boldsymbol{S}_{ij}\boldsymbol{x}^{(n)}(t) \mid m : nm \in \mathcal{E}_c\}$.
7: $\quad$ Receive common variable information from neighbors:
$\{\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t) \mid m : nm \in \mathcal{E}_c,\ m \text{ sent information}\} \bigcup$
$\{\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t-1) \mid m : nm \in \mathcal{E}_c, m \text{ didn't sendinformation}\} \Leftarrow GL$.
8: $\quad$ Update intermediate states according to eqs. (4.3.12b) and (4.3.12c).
9: $\quad$ Run Algorithm 7 to update $\boldsymbol{\pi}$ and $d_{nm},\ \forall m : nm \in \mathcal{E}_c$.
10: **until** One of the termination conditions 1 or 2 is satisfied
11: Restart the algorithm with $\{\mathcal{G}_c \setminus (\arg\max_n \pi_n)\}$.

to a neighbor aggregator n in eq. (4.3.12b). Both changes lead to discrepancies in neighboring aggregators updates as follows:

$$\widetilde{\boldsymbol{\psi}}_k^{(n)}(t) = \boldsymbol{\psi}_k^{(n)}(t) + \overline{\boldsymbol{D}}^{(n)}\boldsymbol{S}_{nm}^\top \boldsymbol{S}_{mn}\boldsymbol{a}_k^{(m)}(t) \tag{4.3.14}$$

$$\Rightarrow \quad \widetilde{\boldsymbol{\upsilon}}_k^{(n)}(t) = \boldsymbol{\upsilon}_k^{(n)}(t) + \overline{\boldsymbol{D}}^{(n)}\boldsymbol{S}_{nm}^\top \boldsymbol{S}_{mn}\boldsymbol{a}_k^{(m)}(t) \tag{4.3.15}$$

$$\widetilde{\boldsymbol{x}}_{k+1}^{(n)}(t) = \boldsymbol{x}_{k+1}^{(n)}(t) + c_4\boldsymbol{M}\boldsymbol{D}^{(n)}\overline{\boldsymbol{D}}^{(n)}\boldsymbol{S}_{nm}^\top \boldsymbol{S}_{mn}\boldsymbol{a}_k^{(m)}(t) \tag{4.3.16}$$

where $\boldsymbol{M} = \left(\boldsymbol{H}^{(n)\top}\boldsymbol{H}^{(n)} + \boldsymbol{S}_{\mathcal{A}}^{(n)\top}\boldsymbol{S}_{\mathcal{A}}^{(n)} + \boldsymbol{S}_{\mathcal{P}}^{(n)\top}\boldsymbol{S}_{\mathcal{P}}^{(n)} + c_4\boldsymbol{D}^{(n)}\right)^{-1}$ and eq. (4.3.16) is analogous to the discrepancy in eq. (4.3.6) where the inaccurate update is a modified version ($\widetilde{\boldsymbol{x}}_{k+1}^{(n)}(t)$) of the true update ($\boldsymbol{x}_{k+1}^{(n)}(t)$).

The FDIAs will be successful at creating algorithm divergence, or convergence to a false optimum, if the ADMM updates in eqs. (4.3.12a) to (4.3.12c) are used to solve

eq. (4.3.10a). Algorithm divergence is a special case of a *Denial of Service* attack in which aggregators are unable to complete the verification process.

An additional area of concern is stealth attacks where the attacker injects a sparse vector, $\boldsymbol{a}^{(n)}$. Here, non-zero entries of the attack vector correspond to the sensors being attacked, such that the constraint in eq. (4.3.17) is satisfied even with the perturbed state:

$$\boldsymbol{h}^{(n)}(\boldsymbol{x}^{(n)} + {\boldsymbol{S}_{\mathcal{A}}^{(n)}}^{\top}\boldsymbol{a}^{(n)}) = 0 \tag{4.3.17}$$

where $\boldsymbol{x}^{(n)}$ corresponds to the true variables. Here, without any change in the loss function in the state verification problem eq. (4.3.10a), the attacker is still able to alter the algorithm's output. These types of attacks are only possible when a malicious aggregator can gain complete knowledge about its neighbors' parameters. Such attacks are tough to detect, and even harder to mitigate, in the absence of a specially imposed structure on the actual measurement vectors. In practice, specially designed sparsity patterns for sensors can prevent such attacks.

4.3.3.4 Detection Of The Malicious Agent

Methods proposed in [Vuković and Dán(2014)] are employed to detect an attack as presented in Algorithm 7. Algorithm 7 is a detection subroutine with the robust state verification of Algorithm 6.

In Algorithm 7, each region n calculates a measure of disagreement, d_{nm}, in the shared variables with a neighboring region m as:

$$d_{nm} = (1-\alpha_k)d_{nm} + \frac{\alpha_k/4}{|\boldsymbol{S}_{nm}\boldsymbol{x}^{(n)}(t)||\mathcal{T}_i|}\sum_{t\in\mathcal{T}_i}\left\|\boldsymbol{S}_{nm}\boldsymbol{x}^{(n)}(t) - \boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t)\right\|_F^2 \tag{4.3.18}$$

and the matrix of normalized disagreement scores $\boldsymbol{B}$:

$$[\boldsymbol{B}]_{nm} = \frac{d_{nm}}{\sum_{m':nm'\in\mathcal{E}_c} d_{nm'} + \epsilon}. \tag{4.3.19}$$

Algorithm 7 Detection Loop; $F(d_{nm}, \boldsymbol{S}_{nm}\boldsymbol{x}^{(n)}(t), \boldsymbol{S}_{ji}\boldsymbol{x}^{(m)},\ \forall m : nm \in \mathcal{E}_c)$

1: Set α, $\epsilon = 10^{-16}$
2: Calculate d_{nm} $\forall m : nm \in \mathcal{E}_c$ according to eq. (4.3.18).
3: Calculate $[\boldsymbol{B}]_{nm} \forall m : nm \in \mathcal{E}_c$ according to eq. (4.3.19).
4: Submit $[\boldsymbol{B}]_{n:}$ to GL until ACK received.
5: $[\boldsymbol{B}]_{m:} \Leftarrow GL, \forall m \in \mathcal{N} \setminus \{n\}$
6: Compute $\boldsymbol{\pi}$, the left principle eigenvector of $\boldsymbol{B}$.
7: **return** $\boldsymbol{\pi}$ and $d_{nm},\ \forall m : nm \in \mathcal{E}_c$.

The left principal eigenvector, $\boldsymbol{\pi}$, of $\boldsymbol{B}$ is then calculated. The value of $\|\boldsymbol{\pi}\|_2$ and the location of the highest element of $\boldsymbol{\pi}$ represent the presence of an attack and the index of the most likely attacker, respectively [Vuković and Dán(2014)].

To mitigate the impact of FDIA, line 11 is added to Algorithm 6 to restart the algorithm after isolating the identified attacker. However, the structure of the communication graph can cause misidentification errors, resulting in divergence in the proposed algorithm. Convergence can be guaranteed provided the following condition is met when establishing the structure of the communication graph:

Theorem 1 (Proposition 3 [Vuković and Dán(2014)])**.** *Consider a system with $N > 2$ regions, if (i) there exists a 3-clique in the graph $\mathcal{G}_c$ and if (ii) for finite k the RSV does not converge, then the stationary distribution $\boldsymbol{\pi}_k$ exists and it is unique and can be computed.*

4.3.3.5 Termination Conditions

Algorithm 6 terminates when either of the following two conditions is satisfied.

1. The first condition is met when an aggregator converges, i.e.,

$$\left\| \boldsymbol{x}_k^{(n)}(t) - \boldsymbol{x}_{k-1}^{(n)}(t) \right\|_\infty \leq \epsilon \tag{4.3.20}$$

2. The second condition is met when all aggregators have agreed about the presence and identity of an FDI attacker, i.e.,

$$\|\boldsymbol{\pi}_k - \boldsymbol{\pi}_{k-1}\|_\infty \leq \epsilon_\pi \quad \text{and} \quad \pi_m > \mu_m(\boldsymbol{\pi}) + \beta\sigma_m(\boldsymbol{\pi}) \tag{4.3.21}$$

for some $\beta > 0$, $m \in \mathcal{N}_n$, where each agent n calculates $\mu_m(\boldsymbol{\pi})$ and $\sigma_m(\boldsymbol{\pi})$ as the excluded average and excluded standard deviation, respectively:

$$\begin{aligned} \mu_m(\boldsymbol{\pi}) &= \frac{1}{|\mathcal{N}_n| - 1} \sum_{m' \in \mathcal{N}_n \setminus \{m\}} \pi_{m'} \\ \sigma_m(\boldsymbol{\pi}) &= \sqrt{\frac{1}{|\mathcal{N}_n| - 1} \sum_{m' \in \mathcal{N}_n \setminus \{m\}} |\pi_{m'} - \mu_{m'}(\boldsymbol{\pi})|^2} \end{aligned} \tag{4.3.22}$$

4.3.3.6 Placement Of Measuring Instruments

Partitioning the electric grid into N aggregator regions must be done according to a ruleset. Recovering a unique $\boldsymbol{\pi}_k$ during the convergence failure of the RSV algorithm requires the presence of at least one 3-clique in the communication graph $\mathcal{G}_c$ (Theorem 1). However, the radial structure of the distribution grid and the requirement for each aggregator region to contain a contiguous set of buses limits the number of possible 3-cliques on $\mathcal{G}_c$. A possible solution is to allow aggregators to access a small amount of sensor measurements from the neighboring aggregators to improve accountability at the expense of privacy.

4.4 Numerical Simulation

This section demonstrates performance of the distributed pricing and robust state verification algorithms before describing implementation on HLF in Section 4.5.

4.4.1 Simulation Setup

Figure 4.4.1 shows the radial MATPOWER 141 bus distribution network used as the demonstration case. The distribution network is separated into 7 aggregator

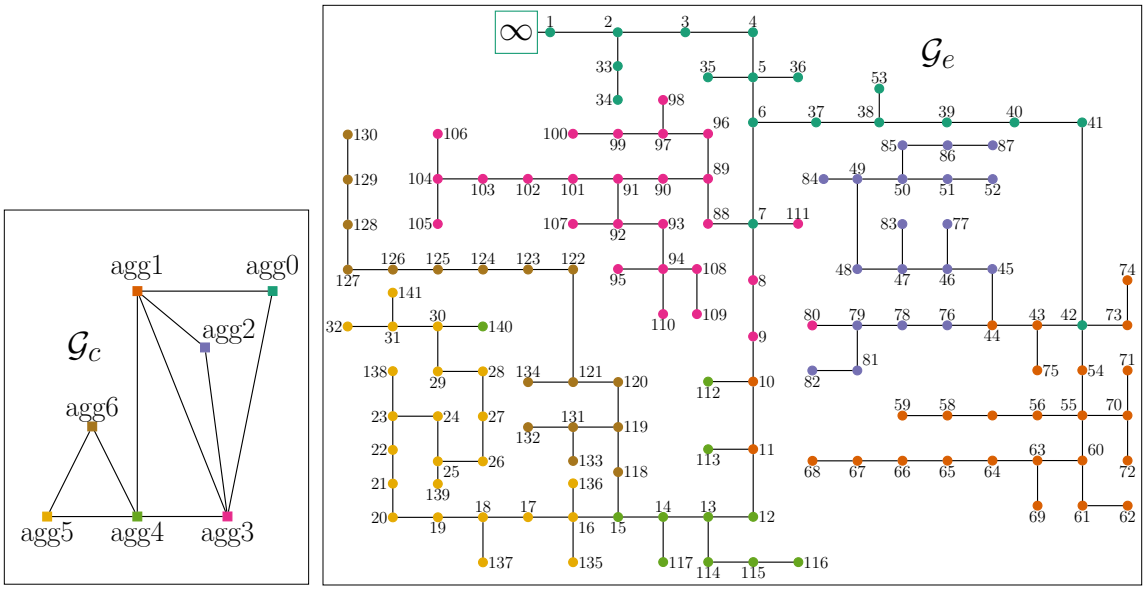


Figure 4.4.1: $\mathcal{G}_e$ Is the Network Graph Corresponding to the 141 Bus Radial Distribution Network. The Network in the Box Shows the Communication Graph $\mathcal{G}_c$ with the Nodes Representing Aggregators

zones ($N = 7$). The choice of N and distribution of buses within N is selected to increase common state variables between aggregators and to increase the number of 3-cliques in the communication graph ($\mathcal{G}_c$ in Figure 4.4.1) to satisfy the required conditions introduced in Section 4.3.3.6. The utility is represented as the substation and belongs to the first aggregator. A total of 3900 prosumers are placed randomly across the distribution network. The total number of prosumers is arbitrary selected and sufficiently high to show algorithm scalability. The distributed pricing algorithm is run for six(6) ten-minute intervals to create a one-hour look-ahead window as a common duration of interest. The choice of interval length and the number of intervals can be selected to match local or regional guidelines on settlement time frames since the generalized formulations are independent of the length and number of intervals.

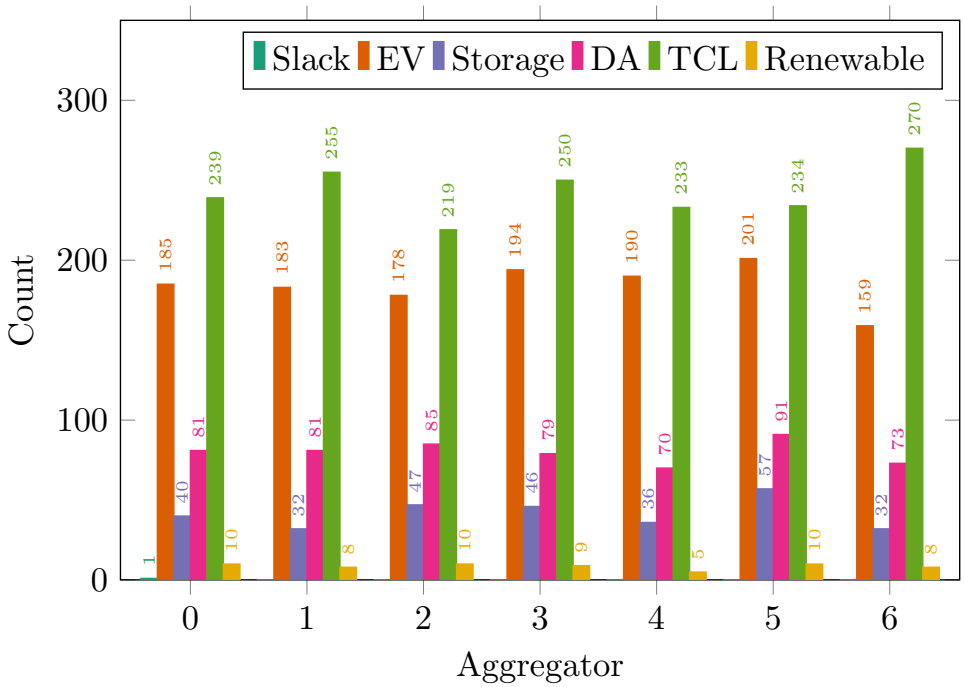


Figure 4.4.2: Asset Types Within Each Aggregator Region

4.4.2 Distributed Pricing

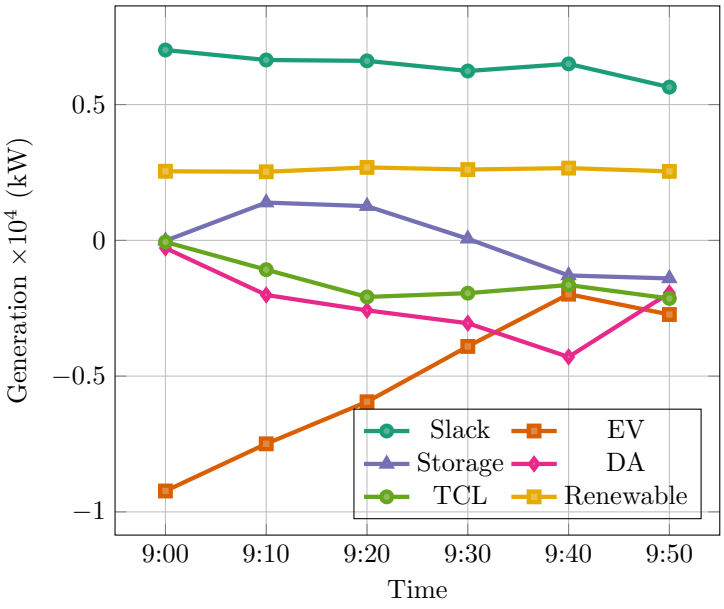
Each aggregator includes EVs, energy storage devices, DAs, thermostatically controlled loads (TCLs), and renewables, as summarized in Figure 4.4.2, with each aggregator including a similar percentage of each prosumer type (generated randomly). Each prosumer has different properties and cost/utility functions uniformly sampled from an identical distribution. Renewables are configured as price-takers (they have no cost/utility), TCLs have a quadratic cost function demanding payments for deviations from their thermostat reference temperature, storage devices require a linear payment proportional to their usage, and DAs and EVs have a linear cost function. The slack bus has a quadratic cost function.

Cost/utility functions and prices are represented by an arbitrary monetary unit in which the price reflects the marginal cost of increasing load by a "single unit".

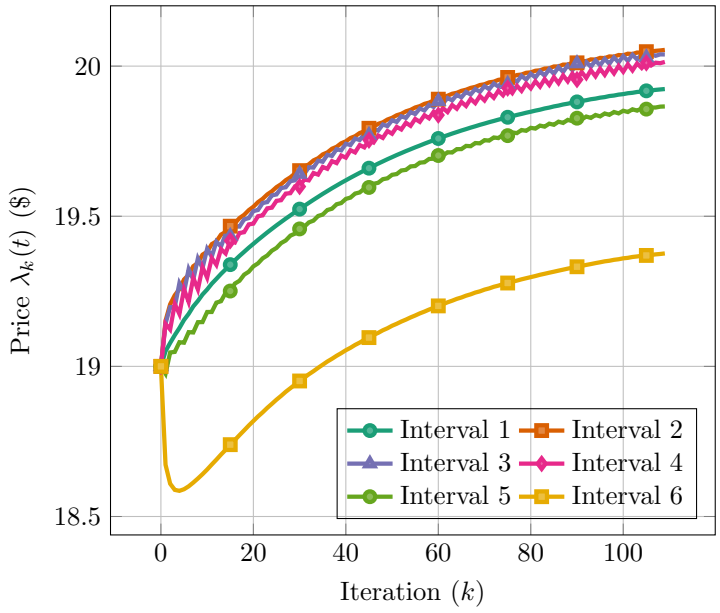
Any changes in asset cost parameters are expected to influence the converged price of eq. (4.2.9). Power transfer between different resources is shown in Figure 4.4.3a. The slack bus is a net generator (as the utility) since all other aggregators are composed mostly of consuming loads. Figure 4.4.3b shows prices for each time interval in the one-hour look-ahead window and how those prices evolve as the distributed algorithm iterates to solve all dispatches and prices for all intervals at once. The price only fluctuates by approximately 2% across the iterations because of the significant load shifting behavior of flexible resources. A total of 28 out of 3900 prosumers are budget-constrained (need to curtail consumption due to insufficient funds); however, the algorithm converges without problems, even though convergence is not guaranteed.

To summarize, given a reasonable initial price for the algorithm, typically obtained from the previous solution, the algorithm solves in a distributed fashion the optimal schedule of all aggregators in a small number of iterations (< 100 in this illustrative example) while managing to honor the individual budget constraint. As mentioned in Section 4.2.2, convergence is not guaranteed due to the non-convexity introduced by the budget constraint in (4.2.3), which may lead to cyclic behavior in the gradient descent. However, for a low number of active constraints, the method often works as was the case in the experiments. This approach chooses an epsilon (Step 22 in Algorithm 5) that expresses the price difference between iterations to determine when prices have stabilized, and hence iterations conclude.

The approach developed here for distributed pricing can easily be adapted to the scenario when each prosumer can trade power directly rather than working through an aggregator. In that case, each prosumer will participate in the iterative pricing algorithm by interacting with the blockchain architecture through smart contracts as shown in Section 6.2. And while that approach is possible with the generalized mathematical framework introduced here, the authors advise caution as the approach will



(a)



(b)

Figure 4.4.3: (a) Power Transfer Between Different Resources (b) Progression of the Shadow Price $\boldsymbol{\lambda}_k$ as Algorithm 5 Iterates, with Each Line Denoting an Element of $\boldsymbol{\lambda}$ Which Reflects Price for a Particular Interval

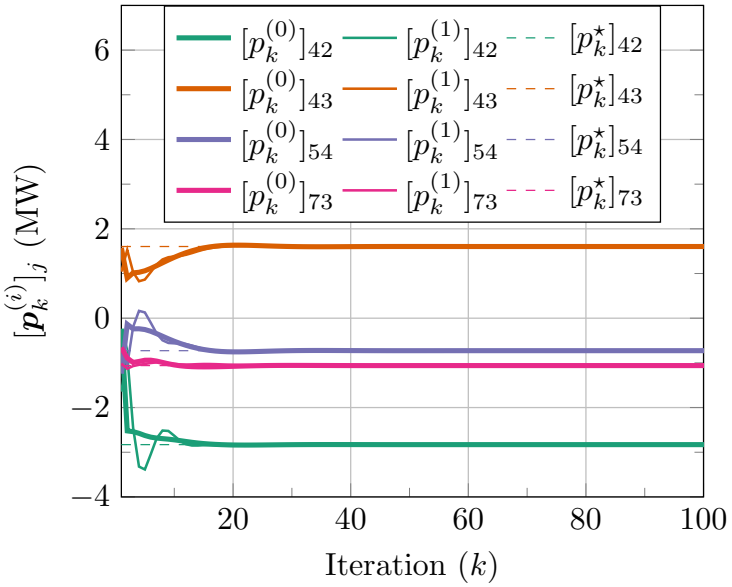
be computationally burdensome because the number of market participants engaging with the GL has an exponential effect on the number of transactions and hence slows the convergence process.

4.4.3 Distributed Verification

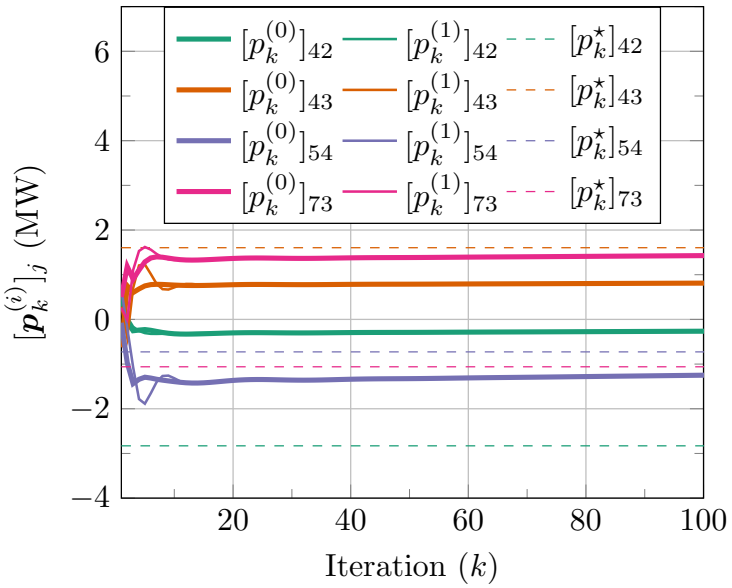
Verification algorithm Algorithms 6 and 7 presented in Section 4.3 are demonstrated here with the following parameters: $\epsilon_\pi = 1e-3$, $\epsilon = 1e-3$, $\alpha_k = 1/k$, $c_3 = c_4 = 0.5$, $\beta_n = 2$, $\forall n \in \mathcal{N}$. Measurements of available variables, $\boldsymbol{x}_{\mathcal{A}}$, were noisy versions of MATPOWER power flow output for the 141 bus radial distribution feeder case. The noise, w_n, was chosen to be Gaussian with zero mean and a variance of 1.

To illustrate the veracity of the verification algorithm, one of the aggregators, m, is set to be a malicious entity capable of injecting false data into its communications with neighboring aggregators. The FDI attack from aggregator m constitutes an injection from attack vector $\boldsymbol{a}^{(mn)}$ to the communications received by a neighboring aggregator n of the attacker in eq. (4.3.12b). In each iteration of the algorithm, the attack vector is chosen randomly subject to $\left\|\boldsymbol{a}^{(mn)}\right\|_2 = 0.5\sqrt{|\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t)|}$, where $\left|\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t)\right|$ is the length of the vector $\boldsymbol{S}_{mn}\boldsymbol{x}^{(m)}(t)$, i.e., the number of common variables shared by region m and its neighbor n.

Consider an example with aggregators 0 and 1 that has buses 42, 43, 54, and 73 as adjacent nodes with common variables between two aggregator regions as shown in the network topology of Figure 4.4.1. Common variables include real power injection, reactive power injection, and voltage magnitude. Figure 4.4.4a shows convergence of the four sets of common variables (one for each bus) shared by aggregators when there is no attack, whereas Figure 4.4.4b shows their divergence when region 0 is an attacker. In Figure 4.4.4a, both aggregators converge to the optimal point $p_b^\star,\ \forall b \in \{42, 43, 54, 73\}$. The attack in Figure 4.4.4b creates a situation in which



(a)



(b)

Figure 4.4.4: Real Power Injections (in MW) by Common Nodes of Aggregators 0 and 1 (a) Convergence to the Optimal Point under No Attack (b) Convergence to a Non-optimal Point When Aggregator 0 Is an Attacker

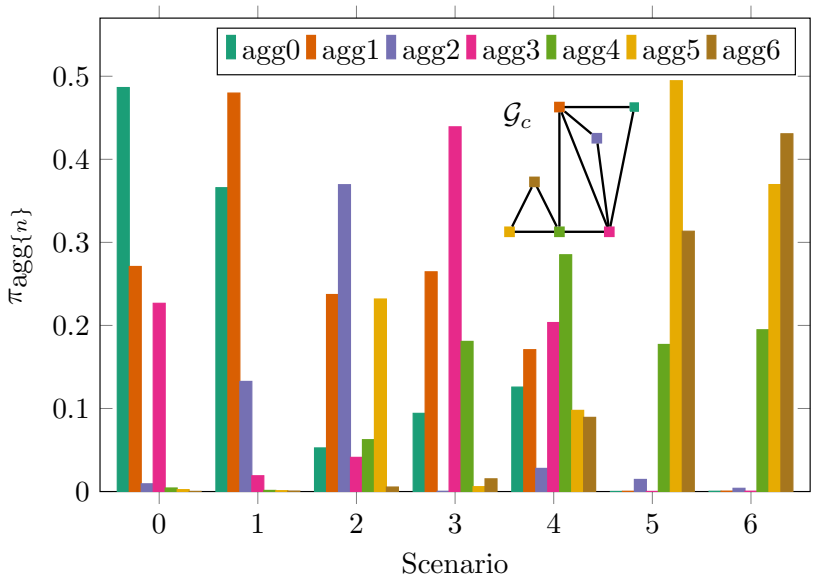


Figure 4.4.5: The Stationary Distribution, $\boldsymbol{\pi}$, of $\boldsymbol{B}$ Under FDI Attack. $\boldsymbol{\pi}$ Represents the Trust Score: Higher the Value of $\pi_{\text{agg}\{n\}}$, Lower is the Trust in Aggregator n. In scenario m, Agg$\{m\}$ is the Attacker.

the parameters reach to the same value yet do not converge at the optimal point. This prevents the verification algorithm from completing. The subroutine described in Algorithm 7 aims to stop such attacks from occurring by tracking disagreements in the common variables between neighbors and identifying the most likely attacker.

The detection of FDI attacks from selfish entities on the TE market is accomplished using the stationary distribution $\boldsymbol{\pi}$ of the disagreement matrix $\boldsymbol{B}$, as discussed in Section 4.3.3.4. Figure 4.4.5 shows results for $N = 7$ in which each aggregator is shown to be the attacker. Each set contains seven bars representing the element of vector $\boldsymbol{\pi}$ corresponding to each of the seven aggregators. As the height of the bar increases, that aggregator is seen as more untrustworthy by the other aggregators. For example, in scenario 0 where agg0 is the attacker, the corresponding bar plot indicates that the network of aggregators trust agg0 the least (i.e., π_{agg0} is highest,

as calculated in line 6 of Algorithm 7). Similarly, in the other scenarios, observation shows that the corresponding attacker amasses the lowest trust level.

It is worthwhile to discuss the distribution of distrust in the network. The distrust is spread among aggregators 0, 1, and 3 considering scenario 0. This can be explained by reflecting on node connectivity of the $\mathcal{G}_c$ graph in Figure 4.4.5, in which agg1 and agg3 are neighbors of agg0. In scenario 4, where agg4 is the dishonest entity, note that agg4 is the most untrustworthy, yet it is not as untrustworthy as the attackers in other scenarios. The distrust is more evenly spread across all aggregators in this scenario. In referring to $\mathcal{G}_c$, the reason for this outcome could occur because aggregator 4 has the highest betweenness centrality and is the only *cut vertex* in the network[10], i.e., agg4 controls information flow between the two clusters (aggregators 0, 1, 2, 3; and aggregators 5, 6). For the communication graph $\mathcal{G}_c$, agg4 dictates the spread of disagreements amongst the aggregators from either cluster. This process to identify attackers will then permit restart of the algorithm to complete the verification process. As billing then occurs the guilty party is penalized, fined, or disconnected from participating in the transactive energy network.

4.5 Design and Implementation on HLF

This section details implementation on HLF of the CPS described in Section 4.1, using the pricing algorithm from Section 4.2, and the verification algorithm in Section 4.3. The choice of a permissioned blockchain architecture, such as HLF, allows consensus protocols that are far less energy-intensive than the proof-of-work consensus protocols employed by permissionless blockchain architectures [Wang *et al.*(2018)].

[10]A vertex in an un-directed connected graph is a cut vertex iff removing it (and edges through it) disconnects the graph or creates more components than the original graph.

4.5.1 Network Setup

The ordering service for any blockchain framework requires a consensus protocol to ensure unambiguous ordering of transactions and guaranteed integrity and consistency of the blockchain across distributed nodes. The developed framework is adaptive to the currently supported Solo, Apache KAFKA, and RAFT algorithms for withstanding crash faults as an ordering service.

Each aggregator is assigned a unique Certificate Authority (CA) [Androulaki *et al.*(2018)] that is responsible for dynamically generating certifications (identities) for authenticating prosumers under each aggregator's purview. When joining the network, an individual prosumer shares information on their type of DER and requests a new set of credentials with a unique prosumer identification (ID) from their associated aggregator. The credential issued by the aggregator includes a resource specific attribute in addition to the unique ID. Unlike prior works [Pipattanasomporn *et al.*(2018)], here the ID is embedded in the prosumer's certificate to add additional security to the issued certificate for that prosumer. Any interaction between the blockchain network and an aggregator requires admin identities [Androulaki *et al.*(2018)]. In contrast, any communication between EVE and a prosumer involves the use of prosumer's unique identity generated by the respective CA.

The use of "channels" here provides the required isolation between individual aggregators and prosumers [Fernandes *et al.*(2020)]. Each transaction in EVE is channel-specific and no data can pass among channels, ensuring privacy and efficient handling of parallel transactions. The EVE blockchain uses $N + 1$ number of separate channels (one channel for each of the N individual aggregator for accessing the corresponding LL, plus one commonchannel among aggregators accessing the GL) to handle access to smart contracts and the ledger. A specific transaction with a specific

Table 4.1: List of Channels and Associated Smart Contracts, Ledgers, and Participant Access For $\mathcal{N} = \{0, \ldots, 6\}$.

Channel	Installed Smart Contract	Ledger Access	Participant Access
commonchannel	Account Contract, Record Contract	GL	aggregator n; $\forall n \in \mathcal{N}$
agg$\{n\}$channel	Bid Contract, Measurement Contract	LL_n	aggregator $n \in \mathcal{N}$ and all prosumers $b \in \mathcal{B}^{(n)}$; $n \in \mathcal{N}$

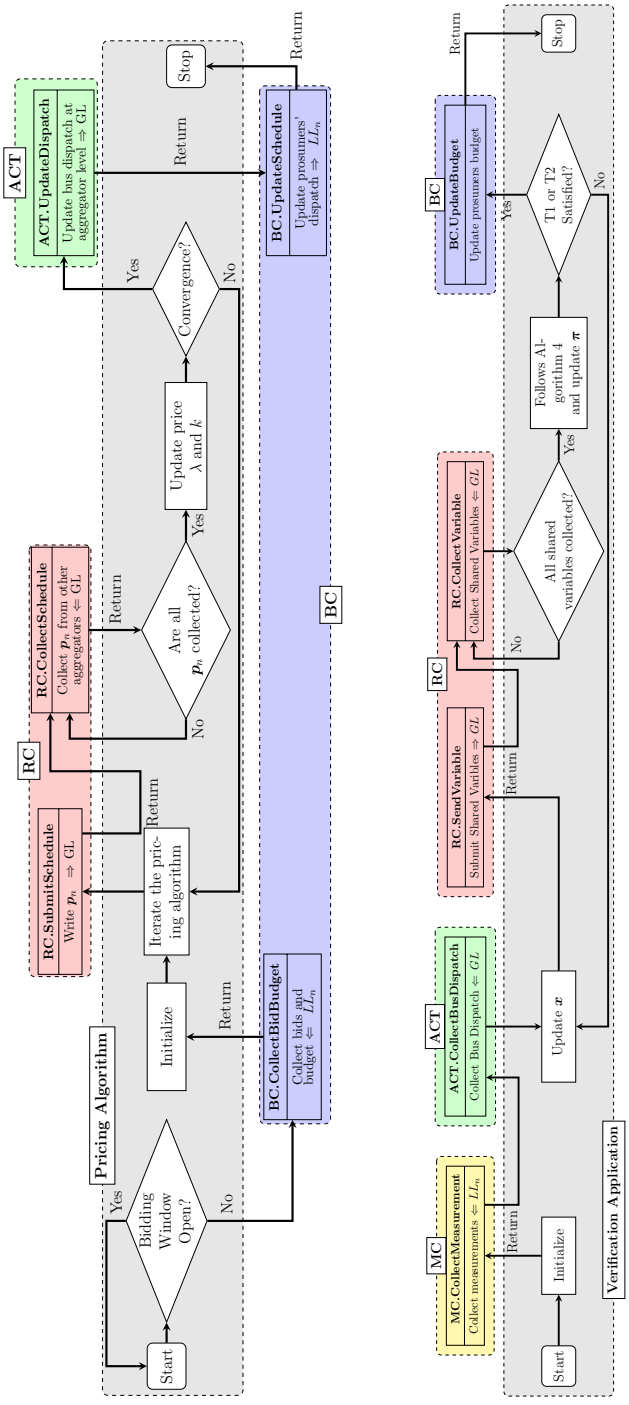
ledger requires invoking the appropriate smart contract.

4.5.2 Implementation of EVE Through Smart Contracts

Table 4.1 lists associations between channels, smart contracts, ledgers, and participant access control. Interactions between smart contracts is summarized in Figure 4.5.1a and Figure 4.5.1b for pricing and verification, respectively. Pricing and verification algorithms are written in Python as external applications. `Node.js` applications are developed to handle communications between external applications and smart contracts allowing read and write operations to appropriate ledgers. Note that each application interacts with different smart contracts to accomplish its required objective.

Account Contract (ACT) manages aggregator-level information such as total production/consumption and associated total cost at the bus-level. At the beginning of the bidding window, each aggregator is required to establish their respective bus information by creating timestamped entries in the GL through **ACT**.

Bid Contract (BC) contains a set of functions to manage individual prosumer bids, dispatch values, and budget information. This contract allows a prosumer to use its certificate to submit any number of bids within the bidding window. However,



(a) (b)

Figure 4.5.1: (a) Graphical Depiction of Pricing Algorithms 4 and 5 Using Smart Contracts and Ledgers for Aggregator $n \in \mathcal{N}$ (b) Graphical Depiction of Verification Algorithm 6 Using Smart Contracts and Ledgers for Aggregator $n \in \mathcal{N}$.

only the last submitted bid within the bidding window is accepted by each aggregator. When submitting a bid, the prosumer is required to provide its asset type (i.e., EV, renewable generation, DA, TCL, storage device, and inflexible load) and corresponding parameters. The smart contract extracts the prosumer's ID from the certificate and ties the submitted bid with the prosumer's unique identity to prevent malicious prosumers from impersonating other prosumers. After the bidding window closes, the aggregator queries submitted bids with their associated budgets from the LL through the aggregator's dedicated channel. The aggregator then executes the distributed pricing algorithm by exchanging information with other aggregators.

Record Contract (RC) is responsible for data exchange for the iterative pricing algorithm through GL. After achieving convergence, each aggregator updates its prosumers' dispatch and cost information in the LL, and bus dispatch values in the GL for future verification. This iteration can be computationally intensive because each interaction (reading and writing) with the GL is a transaction in HLF.

Measurement Contract (MC) handles local measurements from smart meters installed within individual aggregator zones for future verification purposes. Note that the developed algorithm is independent of the sensor location, allowing smart meters to be placed randomly in each aggregator zone for illustration purposes here. The EVE framework can accept measurements at each interval of $\mathcal{T}$ as separate transactions or all measurements for T intervals at the same time as a single transaction. The distributed verification step is always one time step behind the distributed pricing algorithm. For the simulated framework, data from smart meters are generated by solving a non-linear AC power flow problem using the prosumers' dispatch values as input and then adding noise to it. The **RC** handles information exchange for the distributed verification process. Each transaction is assigned a `type` to ensure separation of entries inside **RC** for pricing and verification, thus allowing the use of a

Table 4.2: Benchmark Results for Hyperledger Caliper to Test 200 Iterations of Information Exchange for Algorithm 5.

N	Total Transactions	Sent Rate (tps)	Max Latency (s)	Min Latency (s)	Throughput (tps)
6	1200	6	1.8	0.31	6
7	1400	7	1.55	0.32	7
8	1600	8	1.77	0.36	8
9	1800	9	1.59	0.38	9
10	2000	10	1.49	0.39	10

single smart contract to handle both iterative algorithms.

Algorithms 5 and 6 have different data sharing requirements. Each iteration of Algorithm 5 by one aggregator requires information from all aggregators, whereas each iteration of Algorithm 6 requires information from just neighboring aggregators. Private data sharing between neighboring aggregators can be achieved by creating neighbor-specific channels. Nevertheless, this process is burdensome because 1) more channels are required to handle private data sharing, and 2) changes are required in the existing blockchain network if the communication graph changes (due to changes in sensor deployments among aggregators or distribution network reconfiguration). Therefore, the developed framework leverages the ABAC feature of HLF to handle private communication using the existing commonchannel. When a new edge is created in the communication graph, the new relationship can be added to RC by upgrading the smart contract following the rules of the initial agreement.

4.5.3 Results

Performance results of the proposed framework are generated using predefined use cases in Hyperledger Caliper. For test cases with varying numbers of aggregators (N) between $6 - 10$, and assuming the pricing algorithm requires 200 iterations for convergence, Table 4.2 shows benchmark results in terms of the maximum latency, minimum latency, and throughput[11]. For each scenario shown in Table 4.2, the sending tps rate is equivalent to N because the number of aggregators is the maximum number of writes that can occur to the ledger. Figure 4.5.2 shows a web-based visualization of EVE with transaction numbers from each aggregator during the iterative pricing algorithm. It also demonstrates a graphical representation of the transactions occurring at an hourly basis starting close to 5:30 PM, which includes transactions for the pricing algorithm, updating bus dispatch values for verification in the next hour, and updating prosumers dispatch set points for the next hour.

4.5.4 Security Analysis

Security and privacy of the proposed EVE blockchain introduced here are compared against other leading blockchain approaches. Table 4.3 summaries security considerations of the reviewed works that used blockchain for TE. Most do not fully consider security aspects such as threat model, attack scenario, and verification mechanism at the physical layer. This indicates that the security of these works depends on the built-in security mechanism of the blockchain framework and is not analyzed in-depth for additional threats or weaknesses. In this work, the proposed EVE blockchain framework has been designed and implemented with direct inclusion

[11]Latency = (time when response received - submit time) in second. Throughput = (total valid committed transactions / total time in seconds for all committed nodes in the network) in transactions per second (tps).

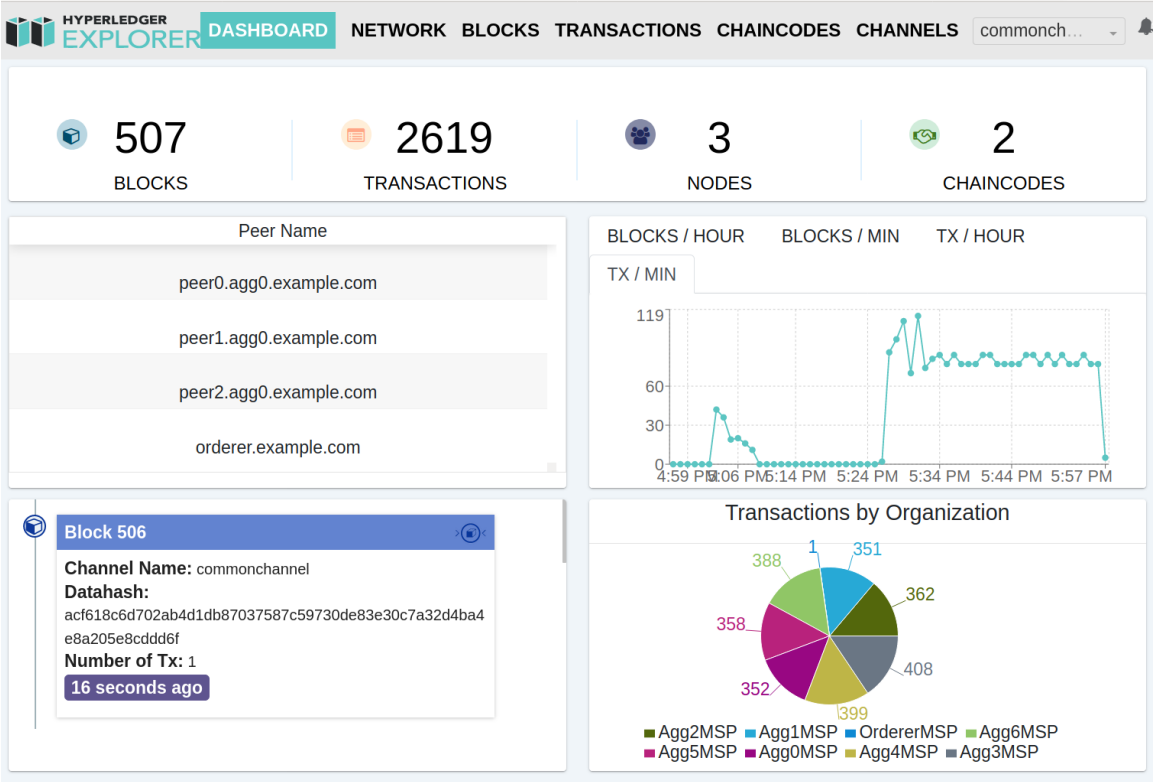


Figure 4.5.2: Visualization of the Commonchannel Through Hyperledger Explorer for $n = 7$

of cyber-security and specific threat models and attack scenarios. Table 4.4 provides a summary of potential threats and countermeasures referenced by the NIST Guide to Industrial Control Systems [Stouffer *et al.*(2015)] and the HLF security mechanism as they relate to the threat model outlined in Section 4.3.2.

As the currently supported ordering mechanisms in HLF only provide crash fault tolerance and do not provide BFT [Hyperledger Architecture Working Group(2017)], a customized BFT-SMART [Sousa *et al.*(2018)] state machine replication and a consensus library has been integrated into this work too. By reinforcing the design with this mechanism, the proposed approach can achieve BFT resilience and durability to avoid a single point of failure. Also, using sensor measurements in the verification stage allows EVE to assure prosumers' compliance with the scheduled transaction.

Table 4.3: Security Analysis of Reviewed Surveys

Reference	Implementation Framework	Threat Model	Attack Scenario	Physical Verification
EVE	HLF	✓	✓	✓
[Wang *et al.*(2019)]	HLF	✗	✗	✗
[Pipattanasomporn *et al.*(2018)]	HLF	✗	✗	✗
[Gai *et al.*(2019)]	HLF	✓	✓	✗
[Laszka *et al.*(2018)]	Ethereum	✓	✗	✗
[Coignard *et al.*(2018)]	Ethereum	✗	✗	✗
[Sabounchi and Wei(2017)]	Ethereum	✗	✗	✗
[Münsing *et al.*(2017)]	Ethereum	✗	✗	✗
[Danzi *et al.*(2017)]	Ethereum	✓	✓	✗

To summarize, this work has incorporated insights from the considerable body of research that has been developed in cyber-security of electric power measurement systems to encompass stealth FDIAs in state estimation [Xie *et al.*(2010)], non-stealth state estimation attacks such as data jamming [Deka *et al.*(2015)], bias injection attack [Luo *et al.*(2019)], and denial of service attacks [Vuković and Dán(2014)]. The only work that discusses possible attack scenarios in blockchain-based energy trading is [Wang *et al.*(2019)]; the scenarios mentioned by the authors include a malicious stakeholder attempting to modify market operations to produce an inaccurate clearing price, a malicious market operator attempting to modify operations of the market algorithm, and a malicious outsider trying to remotely tamper with communications among TE market participants. The RSV presented in this work addresses these security concerns to enhance the blockchain-based TE framework's cyber-security.

Table 4.4: Feasible Threats and Countermeasures in EVE

Layer	**Feasible Threats**	**Countermeasures (HLF)**	**Countermeasures (EVE)**
Application	Stealth FDIA, DoS attack, Smart Contract, Malware	MSP (Fabric CA)	MSP, FDIA Detection (Physical Verification)
Blockchain	Relay attacks, Privilege Elevation, Repudiation, Info disclosure, Byzantine Fault, Civil attack	Read/Write Set Validation, MSP Tracibility with digital signature, Channel isolation	BFT-SMART with features from HLF
Network	DoS attack, Eclipse attack [Heilman *et al.*(2015)]	TLS	ABAC with features from HLF
Client	Identity Theft, Malware	MSP (Fabric CA), Hardware Security Module	ABAC with features from HLF

The market modeling approach in [Wang *et al.*(2019)] uses blockchain only to collect bid information from the prosumers and then utilizes a centralized architecture for determining the market clearing price. On the contrary, this work uses blockchain to manage prosumers, while the proposed decentralized price optimization algorithm uses the decentralized architecture of blockchain to run the iterative price determination algorithm. Moreover, using the inherent security features of Hyperledger Fabric and ABAC allows EVE to avoid the complex attribute based encryption for transaction security introduced in [Guan *et al.*(2021)] while still achieving the same level of privacy.

CHAPTER 5

INTEGRATING HARDWARE SECURITY INTO ELECTRON VOLT EXCHANGE PLATFORM

This chapter introduces a novel framework for integrating hardware security and blockchain functionality with grid-edge devices to establish a distributed cyber-security mechanism that verifies the provenance of messages to and from the devices. Expanding the idea of Two Factor Authentication and Hardware Root of Trust, this work describes the development of a Cryptographic Trust Center^{TM} (CTC^{TM}) chip integrated into grid-edge devices to create uniform cryptographic key management. Product managers, energy system designers, and security architects can utilize this modular framework as a unified approach to manage distributed devices of various vendors, vintages, and sizes. Results demonstrate the application of CTC^{TM} to a blockchain-based TE platform for provisioning of cryptographic keys and improved uniformity of the operational network and data management. This process of configuring, installing, and maintaining keys is described as Eco-Secure Provisioning^{TM} (ESP^{TM}). Laboratory test results show the approach can resolve several cyber-security gaps in common blockchain frameworks such as Hyperledger Fabric.

5.1 Blockchain-Based Transactive Energy Platform

Figure 5.1.1 shows a cyber-physical architecture where the cyber domain is built using EVE, and the physical domain consists of market participants or prosumers who own various types of grid-edge devices. Each prosumer is a "client" to the blockchain network that leverages client applications to interact with the network.

This work assumes that each prosumer is equipped with an EMS responsible for managing the underlying physical assets in the physical domain. The EMS is

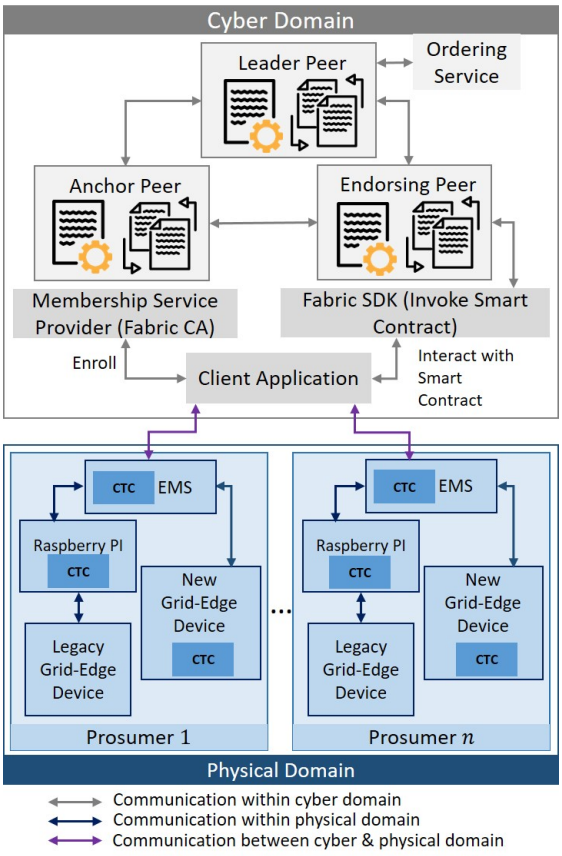


Figure 5.1.1: A Cyber Physical Architecture for Blockchain Based Transactive Energy Network

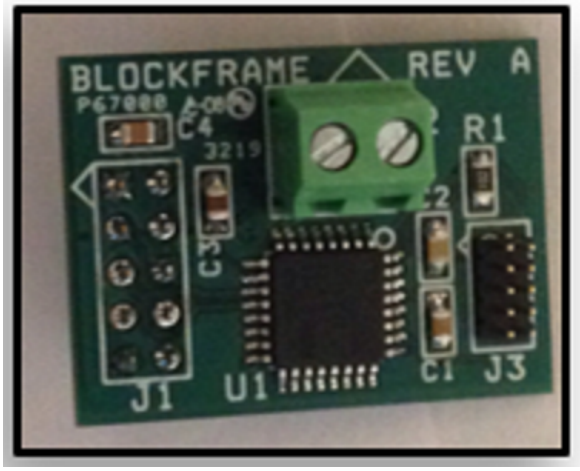


Figure 5.1.2: Cryptographic Trust Center

equipped with the CTC^{TM} (Figure 5.1.2). Physical assets may include legacy assets or new (smart-enabled) infrastructure. Legacy assets integrate with the framework via a Raspberry Pi equipped with the CTC^{TM} chip installed post-manufacture. New grid-edge devices can have the CTC^{TM} installed directly by the manufacturer, thereby skipping the Raspberry Pi for marshaling data from the grid-edge device to the EMS.

The approach provides a secure method to distribute cryptographic keys to grid-edge devices. This approach to configuring, installing, and maintaining keys is called ESP^{TM}). Importantly, EVE records *each provisioning event* and maintain immutable records of such transactions for all grid-edge devices. The logistics and tracking architecture for managing the physical domain through ESP^{TM} is handled through a blockchain that is *separate* from EVE.

The developed EVE framework described in Chapter 4 is capable of performing the following functionalities:

1. Collecting bids from market participants/prosumers;
2. Handling an iterative distributed pricing algorithm;
3. Allowing the local EMS to collect dispatch set points from the TE network through smart contracts and passing those set points onto underlying physical asset(s);
4. Maintaining a record of all transactions for billing.

The first function requires interaction between the cyber and physical domain, and the third function requires interactions within the physical domain. The remaining two functions remain within the cyber domain and hence require no interaction with the physical domain. Any interactions between the cyber and physical domains are considered as a "transaction". This work shows how the CTC^{TM} chip and ED25519

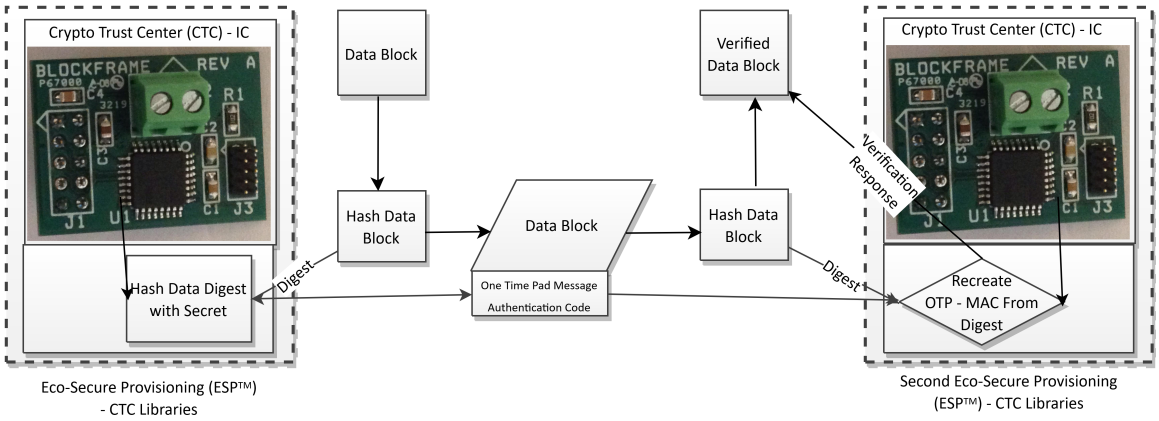


Figure 5.2.1: OTP Message Authentication

verification [Khovratovich and Law(2017)] expands upon 2FA to integrate hardware security into those transactions.

5.2 Security Features Using Eco-Secure Provisioning™ for Cryptographic Key Management

The ESP™ framework includes a small integrated circuit placed within each grid-edge device. A cloud-based Industry Utility Registrar™ (IUR™) service offers `Trust as a Service` to provision trusted components within each device [Hoor and Sheng(2011)]. The provisioning process ensures that no human can access provisioned components, thus preventing hostile actors from circumventing system security without triggering global awareness.

5.2.1 Cyber-Security Layers

The integrity of the cyber-security approach is enabled by the ESP™ framework that uses provisioning protected by logistical tracking of keys contained in each distributed grid-edge device. A cryptographically-secure distributed ledger maintains the record of cryptographic keys providing visibility of provisioned instances while

protecting the actual keys and private values and verifying content without compromising itself. The framework supports any application needs for asynchronous and synchronous communications and supports each case with asymmetric and symmetric crypto-logic operations.

- **Application Layer Secure** – Virtual Private Networks (VPN) sessions are used between any two provisioned devices to give application-level secure communications for one-to-one systems or organizational virtual local area network (VLAN) connections.

- **Data Proof of Origin** – A protected signing capability is used to create Proof of Origin (POO) signatures for any data item. POO forensically proves the device's identity that created it and can do so at any time throughout the life-cycle of the data.

- **Timed Challenge-Response** – Time-restricted challenge-response enforced sequences provide the means to assure that unsanctioned computational analyses on exchanged command components are less likely during the sequence. This is valuable for situations with strict time restrictions when devices perform sensitive or dangerous operations, such as electro-mechanical actuation commands.

- **Uniform Network Segments** – Ensuring operational availability requires network segments such as a local area network, to operate uniformly in the event of the loss of wide-area communications. The framework's distributed uniformity helps sustain the persistence of secure operation without reliance on continuous communication to key management services.

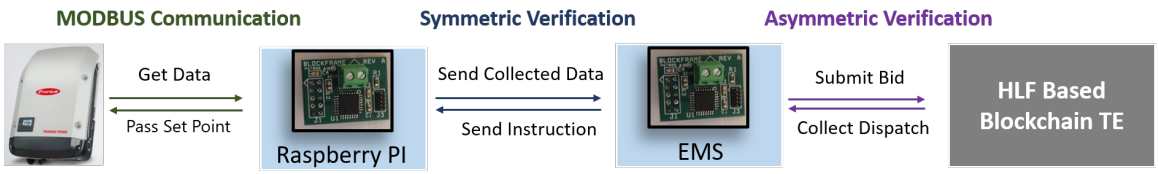


Figure 5.2.2: Architecture of the Hardware Integration Test

5.2.2 Organization-wide Security Support

With a new approach referred to as individualized-uniformity, the framework provisions each device uniquely in a uniform method such that they all work together. Each device can independently perform the sets of security controls. However, other uniform components allow for a synergistic operation of all the devices acting together as a system of systems to increase overall organization-wide security effectiveness. Whether designed for inclusion at the factory or integrated onto legacy applications, the ESP^{TM} framework can support the following functions.

- **Zero-Trust Networking** – The support for individualized-uniformity enables Zero-Trust Networking capabilities because all remote devices can reject unapproved devices, users, or software versions by default until that device can self-validate with multi-factor identification [Flanigan(2018), Kindervag(2010)].

- **Trusted Certificate Provisioning** – Operations that use Transport Layer Security (TLS) connections are strengthened by the distribution and provisioning of trusted certificates [Barker and Barker(2018), NIST(2013)]. ESP^{TM} framework provisions valid certificates and adds features to assure the accuracy of Certificate Revocation Lists (CRL), making legacy TLS security features more robust.

- **Patch and Update Verification** – Individualized confidentiality controls,

enabled by software signatures, can be created for restricted use by individual devices. Single-use codes and restricted-content control can be used to assure that software updates are only installed after verified as trusted.

5.3 Security Features in Grid-edge Devices with Cryptographic Trust Center™ Chips

Integration of ESP™ capabilities with EVE for transmission of data was accomplished using both asymmetric and symmetric capabilities as detailed below:

- **Asymmetric** – These operations use public key cryptography where a mathematically related "public key" and "private key" are used. Only the private key needs to remain secret. This is a common approach of digital signature algorithms in which the private key and data are used to generating a unique signature. The recipient can then independently verify the integrity of the underlying signed data using the sender's public key.

- **Symmetric** – These operations rely on both parties possessing the same cryptographic key, which no outside party can access. For example, most encryption standards use symmetric algorithms that require less system resources, thus enabling faster operations (compared to the asymmetric case) and near run-time performance. Symmetric verification also provides support for checking the authenticity of underlying devices.

5.3.1 Single-Use One Time Pad Verification – Symmetric Message Authentication

In symmetric message authentication, transmitted data is sent with One Time Pad (OTP) Message Authentication Code (MAC) to ensure sensitive command messages remain intact. Upon receiving a message and MAC from a sender, a recipient device

equipped with a CTC^{TM} chip can verify the sender's MAC. MAC messages can be sent alongside any command or response to continuously verify either the receiving or the sending device or in the middle of any transmission sequence. Security in this approach requires both the IUR^{TM} and device to possess the same cryptographic key, which is used to create and verify the MAC. Since the symmetric cryptographic key never leaves the CTC^{TM} chip (cryptographic operations are performed inside), no attacker can clone or impersonate the device, if they gain access to the device's memory or hard drive for a while. Devices which possess the same symmetric key are also able to MAC with each other directly.

The base algorithm used for symmetric authentication is SHA3-512 [Dworkin(2015)], a hash function that, when combined with a secret key, can be used both as a message authenticator and Key Derivation Function. The sequence of operation for symmetric verification is shown in Figure 5.2.1.

5.3.2 Signature Creation and Verification – Digital Signing for Data Integrity Verification

Information from a grid-edge device used for data analytic, controls, user billing transactions, and more functions requires high levels of reliability and integrity. The provisioning framework provides each enabled device a unique elliptic curve asymmetric key pair (private and public key). The asymmetric algorithm used on the chip is ED25519, an elliptic curve with a security level comparable to 3072-bit Rivest–Shamir–Adleman (RSA) crypto-system. Each enabled device signs the data to be exchanged using the private key through an elliptic curve digital signature algorithm. Thus, data from each device can be verified at collection points or after periods of storage using the device's public key. Private keys are used for signing, but do not leave the CTC^{TM}. After signing, data is transmitted to the blockchain network

managing the TE. Smart contracts can verify the integrity of the received information using the public key through ED25519 verification. Hence, the public key belonging to CTC^{TM} is tied into credential management system of EVE without compromising the private keys which remain in the chips. The developed solution is independent of the blockchain framework used to develop EVE because the solution designed does not explicitly depend on HLF. As long as ED25519 verification is possible from the blockchain smart contract, the CTC^{TM} can be used to provide additional verification similar to 2FA verification. A detailed example of a market transaction that uses asymmetrical verification is explained in Section 5.4.3.

5.4 Hardware Integration and Demonstration

The proposed framework was demonstrated using Hardware-In-the-Loop (HIL) simulation described next.

5.4.1 Experimental Setup

The HIL simulation includes an EMS that interacts with TE blockchain to submit consumption and generation bids and to collect dispatch points. The EMS communicates with a legacy inverter through the Raspberry Pi and MODBUS protocol. A 4 KW Fronius Primo inverter was used during the demonstration. Code for managing the EMS is written in `Python` using Django web framework. Though this implementation illustrates secured communication with one physical asset, the Django web framework is suitable for incorporating additional physical assets to the EMS. The Raspberry Pi is equipped with the hardware CTC^{TM} chip and uses I^{2}C protocol to communicate with the chip. Figure 5.2.2 shows the communication links between the associated components with appropriate verification algorithms described in Section 5.3.

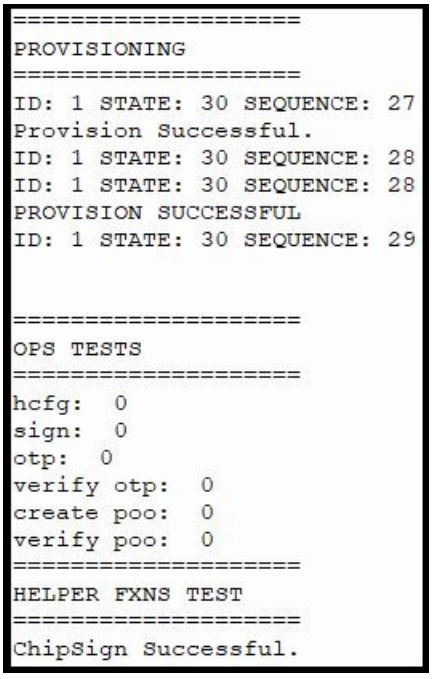


Figure 5.4.1: Provisioning of the CTC Chip

5.4.2 Provisioning of the CTC™ Chip

The provisioning process implements a sequence for updating the CTC™ from the supply chain provisioned status to the end-user application provisioned status. The provisioning IUR™ and CTC™ implement this process using a protocol and sequence consisting of three distinct message-segments, as illustrated in Figure 5.4.1.

In the first step, both the IUR™ and local CTC™ produce random inputs creating symmetric mutual authenticated challenge-response. This process utilizes shared secrets provisioned within CTC™ based on default values, which would be entered into the CTC™ by a manufacturing vendor. These shared secrets are also maintained within the IUR™ server.

The second message-segment contains the packet of cryptographic key information destined for the CTC™, which is encrypted and can only be interpreted by the CTC™. Upon receiving the second message, the CTC™ inserts the new crypto-

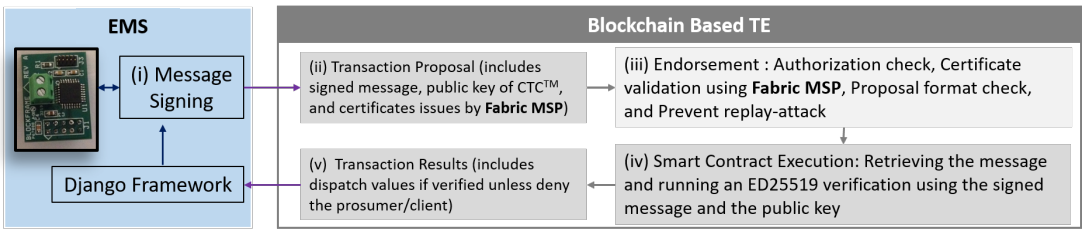


Figure 5.4.2: Two Factor Authentication Using Asymmetric Verification

graphic configuration in temporary status, awaiting the third step.

The third message-segment completes a cryptographic key super-session [NIST(2019)] sequence by verifying both old and new cryptographic configurations are available within the CTC^{TM} by using them together. Upon receiving the super-session verification, the IUR^{TM} creates a final commit stream and returns an encapsulated stream containing this sequence. The CTC^{TM} finalizes super-session and commits the new cryptographic configuration for use only if all signatures for the final sequence content are verified successfully.

5.4.3 Asymmetric Verification in TE Network Communications

An example use case is demonstrated here for asymmetric verification that allows integrating hardware security with the blockchain smart contract. Figure 5.4.2 illustrates the implementation while the EMS is collecting the dispatch schedule from EVE.

The EMS is assigned a unique ID with credentials issued by a Hyperledger Fabric CA. The EMS uses the private key to sign the message where the message includes the prosumers ID. Using the `Node.js` client SDK, a transaction proposal is created that contains the signed message and the chip's public key as the input arguments. After submitting the transaction, the endorsement process in HLF verifies the client's (i.e., EMS's) credential using the MSP. If verified, the smart contract then takes input

arguments from the transactions and uses the public key to perform an ED25519 verification that verifies the prosumer's ID against the credentials issued by the CA. This second verification is analogous to the idea of 2FA, which provides additional security to secure the TE network against malicious prosumers. A similar procedure is followed when the prosumer submits a bid that includes the prosumer's unique ID, type (solar PV/ electric vehicle/ controllable load, and more), and associated parameters.

5.4.4 Symmetric Verification for Two-way Asset Communications

After collecting the dispatch set point, the EMS sends this information to the legacy inverter through the Raspberry Pi and the Django framework. Sending data through the Django framework requires the use of `websocket` protocol, which can transmit a JSON formatted data object. Hence the EMS creates a JSON object that includes the message `type` (to query information from the inverter or to send dispatch instruction) and the message itself. For example, while sending the dispatch information, the message includes the real power dispatch value and the OTP created by the EMS. Upon receiving, the Raspberry Pi parses the message and uses the underlying instruction, and it's own cryptographic key to generate it's OTP. If the generated OTP matches the OTP embedded in the message, the Pi accepts the command and responds accordingly. The process is repeated on the EMS after receiving the response from the Raspberry Pi. This symmetrical verification secures communication between the EMS and any physical asset communicating with the EMS. Figures 5.4.3a-5.4.3c show the use of OTP while sending dispatch information from the EMS to the inverter and receiving output from the inverter[1].

[1]The difference between the command sent and the output of the inverter happens due to rounding issue in the inverter and intermittence in solar irradiance.

```
WS Opened.
Incoming Dispatch Request is Verified.
Received Dispatch Value: 880.0 Watt
Write to Register:Write Resister Response: 40242 =>586
```

(a)

(b)

```
Collecting Current Dispatch from Update.py
A Get Method Posted
Websocket HANDSHAKING /hems/ws/ [192.168.0.152:39962]
Websocket CONNECT /hems/ws/ [192.168.0.152:39962]
Device Check in: {'ID': '1', 'Manufacturer':'Fronious','Model':'Primo','Type':'Inverter',}
Received Data Through Dispatch Update Request in app_controller.py
Peer Received Data: 1
Update Success from app_controller.py: 1
HTTP GET /hems/collectdispatch/?payload=554+53_53_52+39_228_25_130_160_159_197_184_150_157_39_152_125_202_22_0_
163_209_117_89_118_109_52_132_244_107_6_98_152_101_239_1+101_25_145_190_24_239_237_57_70_52
My_function counter 1 from update.py
Sending a Secured Poll Request from update.py as Chip enabled on the remote
Received Data: Device1_V237.6999969482422_P872.0_Q-12.0
```

(c)

Figure 5.4.3: (a) EMS Sends the Dispatch Value to the Raspberry Pi (b) Output from the Fronius Inverter (c) EMS Receives Information after Symmetric Verification

CHAPTER 6

LYAPUNOV STABILITY OF SMART INVERTERS USING LINEARIZED DISTFLOW APPROXIMATION

This part of the dissertation uses linear distribution power flow equations and droop-based Volt-Var and Volt-Watt control curves to analytically derive a stability criterion using Lyapunov analysis that includes the network operating condition. The methodology is generally applicable for control curves that can be represented as Lipschitz functions. The derived Lipschitz constants account for smart inverter hardware limitations for reactive power generation. A local policy is derived from the stability criterion that allows inverters to adapt their control curves by monitoring only local voltage, thus avoiding centralized control or information sharing with other inverters. The criterion is independent of the internal time-delays of smart inverters. Simulation results for inverters with and without the proposed stabilization technique demonstrate how smart inverters can mitigate voltage oscillations locally and mitigate real and reactive power flow disturbances at the substation under multiple scenarios. The study concludes with illustrations of how the control policy can dampen oscillations caused by solar intermittency and cyber-attacks.

6.1 Smart Inverter Models

6.1.1 Overview of the Inverter Logic Design Problem

The operational logic of smart inverters is aimed at selecting real and reactive power generation set-point values and can be expressed through a cost function. The cost function Γ is a linear combination of two objectives [Turitsyn *et al.*(2011)] and

whose minimizer is the optimal dispatch for the inverters:

$$\min_{\boldsymbol{v},\boldsymbol{p},\boldsymbol{q}} \Gamma(\boldsymbol{v},\boldsymbol{p},\boldsymbol{q},\boldsymbol{R},\boldsymbol{X}) \text{ s.t. } (\boldsymbol{v},\boldsymbol{p},\boldsymbol{q}) \in \mathcal{S} \tag{6.1.1}$$

where $\mathcal{S}$ is the feasible set for $\boldsymbol{v},\boldsymbol{p},\boldsymbol{q}$ and:

$$\Gamma(\boldsymbol{v},\boldsymbol{p},\boldsymbol{q},\boldsymbol{R},\boldsymbol{X}) = c_\rho \underbrace{\rho(\boldsymbol{p},\boldsymbol{q},\boldsymbol{R},\boldsymbol{X})}_{M1} + c_\nu \underbrace{\nu(\boldsymbol{v},\boldsymbol{p},\boldsymbol{q})}_{M2} \tag{6.1.2}$$

with $\rho(\boldsymbol{p},\boldsymbol{q},\boldsymbol{R},\boldsymbol{X})$ expressing ohmic losses and $\nu(\boldsymbol{v},\boldsymbol{p},\boldsymbol{q})$ promoting a voltage profile that deviates as little as possible from the nominal voltage. Non-negative weights c_ρ and c_ν can be tuned to adjust the relative importance of each objective.

The feasible set in $\boldsymbol{v},\boldsymbol{p},\boldsymbol{q}$ is defined by the DistFlow equations. When exclusively minimizing losses (term $M1$ in (6.1.2)), the minimization function and constraints can be written as:

$$\min_{\boldsymbol{p}_g,\boldsymbol{q}_g} (\boldsymbol{p}^c - \boldsymbol{p}^g)^\top \boldsymbol{R}(\boldsymbol{p}^c - \boldsymbol{p}^g) + (\boldsymbol{q}^c - \boldsymbol{q}^g)^\top \boldsymbol{X}(\boldsymbol{q}^c - \boldsymbol{q}^g) \tag{6.1.3a}$$

$$\text{s.t. } \boldsymbol{p}^g_{\min} \le \boldsymbol{p}^g \le \boldsymbol{p}^g_{\max}; \quad \boldsymbol{q}^g_{\min} \le \boldsymbol{q}^g \le \boldsymbol{q}^g_{\max} \tag{6.1.3b}$$

Here, $\boldsymbol{p}^g_{\min}, \boldsymbol{p}^g_{\max}, \boldsymbol{q}^q_{\min}, \boldsymbol{q}^g_{\max}$ represent the minimum real power, maximum real power, minimum reactive power, and maximum reactive power that an inverter can generate at any instant of time, respectively. This objective does not include any constraint on the voltage profile; rather, the voltage profile will be the solution of the DistFlow equations and may violate the standard acceptable voltage range.

For the policy focusing only on minimizing voltage deviation from the nominal value (term $M2$ in (6.1.2)), the minimization function and constraints can be written

as:

$$\min_{\boldsymbol{p}^g, \boldsymbol{q}^g} \|\boldsymbol{v} - \boldsymbol{v}_{nom}\|_2^2 \tag{6.1.4a}$$

$$\text{s.t.} \quad \boldsymbol{v}^2 = v_0^2 \mathbf{1} + \boldsymbol{R}(\boldsymbol{p}^c - \boldsymbol{p}^g) + \boldsymbol{X}(\boldsymbol{q}^c - \boldsymbol{q}^g) \tag{6.1.4b}$$

$$\boldsymbol{v}_{\min} \leq \boldsymbol{v} \leq \boldsymbol{v}_{\max} \tag{6.1.4c}$$

$$\boldsymbol{p}^g_{\min} \leq \boldsymbol{p}^g \leq \boldsymbol{p}^g_{\max}; \quad \boldsymbol{q}^g_{\min} \leq \boldsymbol{q}^g \leq \boldsymbol{q}^g_{\max} \tag{6.1.4d}$$

where $\boldsymbol{v}_{nom}$ represents the nominal voltage vector. This includes additional constraints beyond (6.1.3).

The direct minimization of $\Gamma(\boldsymbol{v}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{R}, \boldsymbol{X})$ requires communicating real-time information of all loads and inverters to a centralized solver. The set-points (i.e. the solution of problem (6.1.4a)) can be updated at regular predefined times based on the load components $\boldsymbol{p}_c$ and $\boldsymbol{q}_c$ and the maximum apparent power that can be generated at a current time. Conversely, a purely decentralized method uses local measurements as inputs for deciding inverter set-points. The control policy is instead hardwired at the time of installation, noting, however, that remote firmware updates are possible but excluded from consideration here. Two forms of policies exist for local control: (1) policies that adjust reactive and/or real power to minimize losses, $M1$ [Garg *et al.*(2018)], and (2) policies that respond to voltage measurements and try reduce voltage deviations by adjusting reactive and/or real power, $M2$. The first policy will not cause oscillations but may result in unacceptable voltage levels. The second policy leads to a closed-loop system and may cause oscillations while trying to maintain voltage within operating limits [Farivar *et al.*(2013), Farivar *et al.*(2015), Zhou *et al.*(2016), Zhou *et al.*(2015), Braslavsky *et al.*(2018)] as the DistFlow equations constrain the values $(\boldsymbol{v}, \boldsymbol{p}, \boldsymbol{q})$.

6.1.2 Inverter Modeling Assumptions

This work focuses on the second type of strategy ($M2$) described above while trying to adjust both real and reactive power. This strategy requires defining the Volt-Watt and Volt-Var control functions, respectively, as,

$$\boldsymbol{f_p}(\Delta \boldsymbol{v}) : \mathbb{R}^n \mapsto \mathbb{R}^n \quad \text{and} \quad \boldsymbol{f_q}(\Delta \boldsymbol{v}) : \mathbb{R}^n \mapsto \mathbb{R}^n \tag{6.1.5}$$

For local policies, these functions take each entry of the vector that represents the deviation from the nominal voltage according to (6.1.6) as input, and provide the real and reactive power injection values for the corresponding bus as output.

$$\Delta \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{v}_{nom} \tag{6.1.6}$$

Considering Δv_i as the voltage deviation of the i^{th} bus, and $f_{p,i}(\Delta v_i)$ and $f_{q,i}(\Delta v_i)$ representing the Volt-Watt and Volt-Var control functions of the inverter at i^{th} bus, respectively, the following assumptions can be made regarding the control policy [Farivar *et al.*(2013), Farivar *et al.*(2015), Zhou *et al.*(2016), Zhou *et al.*(2015)]:

A1: Both $f_{p,i}(\Delta v_i)$ and $f_{q,i}(\Delta v_i)$ are monotonically decreasing functions and are continuous and piece-wise differentiable (control functions may include regions where the derivative is zero, such as a dead-band or constant output).

A2: The derivatives of the control functions are bounded, i.e. there exists $C_{p,i} < +\infty$ and $C_{q,i} < +\infty$ such that $\left|f_p'(\Delta v_i)\right| \leq C_{p,i}$ and $\left|f_q'(\Delta v_i)\right| \leq C_{q,i}$ for all Δv_i.

For executing Volt-Watt and Volt-Var control, this work adopts the droop control curves mentioned in [Seal(2017), IEEE Standards Coordinating Committee 21(2018), rul(2019)]. In these works, the modulation of active and reactive power injection indicated in (6.1.5) depends upon (6.1.6). The derivation of the corresponding Lipschitz constants are hereafter shown. This example is used as a case study in the numerical simulations in Section 6.3 that corroborate the analysis.

6.1.3 Definitions

For an inverter at node i of a distribution network, s_i and $\overline{p_i}$ represent the rated apparent power and the maximum real power output at certain irradiance respectively. Following [Braslavsky *et al.*(2018)] and dropping the superscript ^{g} representing generation, $\overline{p_i}$ can be expressed as a fraction of s_i:

$$\overline{p_i} = \mu \; s_i; \quad 0 < \mu \leq 1 \tag{6.1.7}$$

At nominal irradiance, the inverter can generate real power equal to its apparent power rating, resulting in $\mu = 1$.

The maximum reactive power consumption/injection depends on hardware limits (q_i^{lim}) and available reactive power. Hence:

$$\overline{q_i}(\Delta v_i) = \min\left(q_i^{lim}, \sqrt{s_i^2 - f_{p,i}^2(\Delta v_i)}\right) \tag{6.1.8}$$

For constant values of $f_{p,i}(\Delta v_i)$, $\overline{q_i}(\Delta v_i)$ is independent of voltage deviation and can be written as:

$$\overline{q_i}(\Delta v_i) = \min\left(q_i^{lim}, \sqrt{s_i^2 - f_{p,i}^2}\right) = \overline{q_i} \tag{6.1.9}$$

The piece-wise linear Volt-Watt and Volt-Var control curves (referred to as *droop* curves) of the inverter at the i^{th} bus respectively are shown in Figures 6.1.1 and 6.1.2 where, ϵ_p, ϵ_q^+, ϵ_q^-, V_p and V_q^+ are positive, V_q^- is negative, $V_q^+ - \epsilon_q^+/2 > 0$, $V_q^- + \epsilon_q^-/2 < 0$, $V_p - \epsilon_p/2 > 0$. Using (6.1.7)-(6.1.9), (6.1.10) and (6.1.11) provide the mathematical formulation representing Figures 6.1.1 and 6.1.2, respectively.

$$f_{p,i}(\Delta v_i) = \begin{cases} \overline{p_i} & \Delta v_i \leq (V_p - \frac{\epsilon_p}{2}) \\ \dfrac{\overline{p_i}}{-\epsilon_p}\left(\Delta v_i - (V_p + \frac{\epsilon_p}{2})\right) & |\Delta v_i - V_p| < \frac{\epsilon_p}{2} \\ 0 & \Delta v_i > (V_p + \frac{\epsilon_p}{2}) \end{cases} \tag{6.1.10}$$

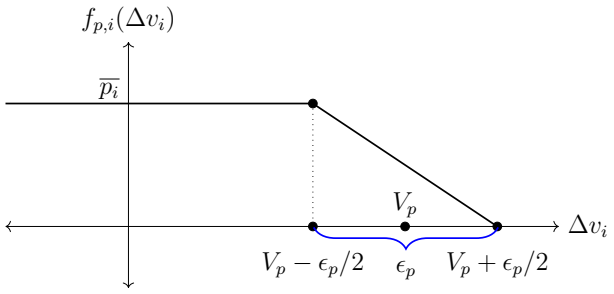


Figure 6.1.1: Inverter Volt-Watt Curve

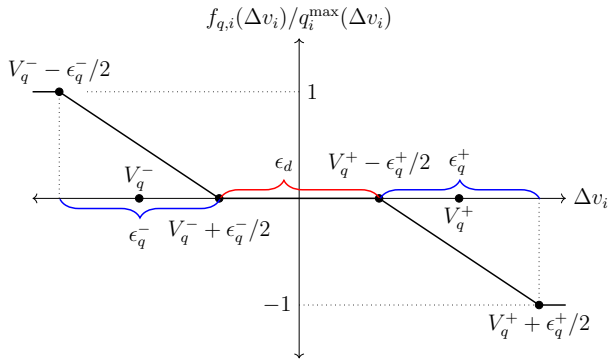


Figure 6.1.2: Inverter Volt-Var Curve

$$f_{q,i}(\Delta v_i) = \begin{cases} \overline{q_i} & \Delta v_i < V_q^- - \frac{\epsilon_q^-}{2} \\ \dfrac{\overline{q_i}}{-\epsilon_q^-}\left(\Delta v_i - (V_q^- + \frac{\epsilon_q^-}{2})\right) & \left|\Delta v_i - V_q^-\right| < \frac{\epsilon_q^-}{2} \\ 0 & |\Delta v_i| < \frac{\epsilon_d}{2} \\ \dfrac{\overline{q_i}(\Delta v_i)}{-\epsilon_q^+}\left(\Delta v_i - (V_q^+ - \frac{\epsilon_q^+}{2})\right) & |\Delta v_i - V_q^+| < \frac{\epsilon_q^+}{2} \\ -\overline{q_i} & \Delta v_i > V_q^+ - \frac{\epsilon_q^+}{2} \end{cases} \tag{6.1.11}$$

Recalling assumptions A1 and A2, the following lemma confirms that the IEEE 1547 policies described above are Lipschitz functions.

Lemma 1. *The functions $f_{p,i}(\Delta v_i)$ and $f_{q,i}(\Delta v_i)$ can be represented as Lipschitz functions where the Lipschitz constants $C_{p,i}$ and $C_{q,i}$ can be written as:*

$$C_{p,i} = \frac{\overline{p_i}}{\epsilon_p}; \qquad C_{q,i} = \frac{\overline{p_i}}{\epsilon_p \sqrt{\frac{1}{\mu^2} - 1}} + \frac{q_i^{lim}}{\epsilon_q^+} \tag{6.1.12}$$

where, μ is defined in (6.1.7).

Proof. Taking the derivative of (6.1.10),

$$\frac{df_{p,i}(\Delta v_i)}{d\Delta v_i} = \begin{cases} -\dfrac{\overline{p_i}}{\epsilon_p} & \text{if: } (\Delta v_i - V_p) \in \ ]-\frac{\epsilon_p}{2}, \frac{\epsilon_p}{2}] \\ 0 & \text{if: otherwise} \end{cases} \tag{6.1.13}$$

Hence, the Lipschitz constant for the Volt-Watt curve of the inverter at i^{th} bus is $C_{p,i} = \dfrac{\overline{p_i}}{\epsilon_p}$, as reported in (6.1.12).
Similarly, taking the derivative of (6.1.11):

$$\frac{df_{q,i}(\Delta v_i)}{d\Delta v_i}(\Delta v_i) = \begin{cases} -\dfrac{\bar{q}_i}{\epsilon_q^-} & \text{if: } (\Delta v_i - V_q^-) \in \ [-\frac{\epsilon_q^-}{2}, \frac{\epsilon_q^-}{2}[\\ \dfrac{d\left(\bar{q}_i(\Delta v_i)\frac{\Delta v_i - (V_q^+ - \frac{\epsilon_q^+}{2})}{-\epsilon_q^+}\right)}{d\Delta v_i} & \text{if: } (\Delta v_i - V_q^+) \in \ ]-\frac{\epsilon_q^+}{2}, \frac{\epsilon_q^+}{2}] \\ 0 & \text{if: otherwise} \end{cases} \tag{6.1.14}$$

For the second case of (6.1.14):

$$\begin{aligned} &\frac{d\left(\bar{q}_i(\Delta v_i)\frac{\Delta v_i - (V_q^+ - \frac{\epsilon_q^+}{2})}{-\epsilon_q^+}\right)}{d\Delta v_i} \\ &= u\Big(q_i^{lim} - \sqrt{s_i^2 - f_{p,i}^2(\Delta v_i)}\Big)\frac{-f_{p,i}(\Delta v_i)\frac{df_{p,i}(\Delta v_i)}{d\Delta v_i}}{\sqrt{s_i^2 - f_{p,i}^2(\Delta v_i)}} \times \Big(\Delta v_i - (V_q^+ - \frac{\epsilon_q^+}{2})\Big) \\ &+ \min\Big(q_i^{lim}, \sqrt{s_i^2 - f_{p,i}^2(\Delta v_i)}\Big) \end{aligned} \tag{6.1.15}$$

where $u(\bullet)$ is the step function. The Lipschitz constant for the Volt-Var curve can be calculated by taking the maximum of (6.1.15). In considering that $f_{p,i}(\Delta v_i)$ has

a maximum value of $\overline{p_i}$, and using (6.1.13) and (6.1.7), a bound for the derivative of the Volt-Var curve of the inverter at i^{th} bus is the value reported in (6.1.12). □

Some additional observations are useful for controlling the Lipschitz constants of smart inverters to ensure stability of the overall network. First, it is desirable to inject as much real power as possible keeping $\epsilon_p > 0$ as small as possible. Also, since the dead-band for the reactive power curve can be at most zero, the parameters ϵ_q^+ and ϵ_q^- are such that:

$$\epsilon_q^+ \leq 2V_q^+ \quad \text{and} \quad \epsilon_q^- \leq -2V_q^- \tag{6.1.16}$$

6.2 Stability Analysis

In this work, the inverter dynamics presented in [Seal(2017)] and (2.2.7) are used to derive the stability criterion of a distribution network with smart inverters. The main result of this paper is a sufficient condition to achieve voltage stability for a network with inverters that use only local information and logic satisfying assumptions A1 and A2 expressed in Section 6.1.

Let $r_i(\boldsymbol{A}) = \sum_j \boldsymbol{A}_{ij}$, which is the row sum of the elements of matrix $\boldsymbol{A}$. The stability condition can be stated as the following Theorem:

Theorem 2. *Let $\boldsymbol{v}^{*2} = \bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^*$ be the voltage vector that is a fixed point of the inverter dynamics. If A1 and A2 (c.f. Section 6.1) hold, a sufficient condition for stability of an inverter dominated network is:*

$$C_{p,i}^2 + C_{q,i}^2 < \frac{2\ v_i^{*2}}{r_i(\boldsymbol{Z}\boldsymbol{Z}^\top)} \tag{6.2.1}$$

Proof. Writing (2.2.7) and dropping the superscript:

$$\boldsymbol{v} = \sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}} \tag{6.2.2}$$

where $\bar{\boldsymbol{v}}^2 = v_o^2\mathbf{1} + \boldsymbol{Z}\boldsymbol{s}^c$ and the square root operates element-wise. Let

$$\boldsymbol{f_S}(\boldsymbol{v} - \boldsymbol{v}_{nom}) = \begin{bmatrix} \boldsymbol{f}_p \\ \boldsymbol{f}_q \end{bmatrix}; \quad \boldsymbol{T} = \begin{bmatrix} \boldsymbol{T}_p \\ \boldsymbol{T}_q \end{bmatrix}; \quad \boldsymbol{C_s} = \begin{bmatrix} \boldsymbol{C}_p \\ \boldsymbol{C}_q \end{bmatrix}$$

where $\boldsymbol{T}_p$, $\boldsymbol{T}_q$, $\boldsymbol{C}_p$, and $\boldsymbol{C}_q$ are diagonal matrices containing the low-pass filter time constants (time delays) and inverter droop values, respectively.

The inverter dynamics can be written as (6.2.3a):

$$\boldsymbol{T}\dot{\boldsymbol{s}} = \boldsymbol{f_S}\Big(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}} - \boldsymbol{v}_{nom}\Big) - \boldsymbol{s} \tag{6.2.3a}$$

and for the fixed point $\boldsymbol{s}^\star$:

$$\mathbf{0} = \boldsymbol{f_S}\Big(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^\star} - \boldsymbol{v}_{nom}\Big) - \boldsymbol{s}^\star \tag{6.2.3b}$$

Defining a shift about the fixed point as $\boldsymbol{\delta s} = \boldsymbol{s} - \boldsymbol{s}^\star$ and using (6.2.3b), the inverter dynamics can be rewritten as:

$$\boldsymbol{T}\boldsymbol{\delta}\dot{\boldsymbol{s}} \;=\; \boldsymbol{f_S}\Big(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^* - \boldsymbol{Z}\boldsymbol{\delta s}} - \boldsymbol{v}_{nom}\Big) \;-\; \boldsymbol{f_S}\Big(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^\star} - \boldsymbol{v}_{nom}\Big) \;-\; \boldsymbol{\delta s} \tag{6.2.4}$$

Now the stability of (6.2.4) can be analyzed with Lyapunov analysis. Considering the candidate Lyapunov function:

$$J(\boldsymbol{\delta s}) = \frac{1}{2}\boldsymbol{\delta s}^\top \, \boldsymbol{T} \, \boldsymbol{\delta s}$$

Taking the derivative of $J(\boldsymbol{\delta s})$ along the trajectory yields:

$$\dot{J}(\boldsymbol{\delta s}) = \boldsymbol{\delta s}^\top \, \boldsymbol{T}\boldsymbol{\delta}\dot{\boldsymbol{s}} \tag{6.2.5}$$

Putting the value of $\boldsymbol{T}\boldsymbol{\delta}\dot{\boldsymbol{s}}$ from (6.2.4) in (6.2.5) provides:

$$\dot{J}(\boldsymbol{\delta s}) = \boldsymbol{\delta s}^\top \left[\boldsymbol{f_S}\left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^\star - \boldsymbol{Z}\boldsymbol{\delta s}} - \boldsymbol{v}_{nom}\right) - \boldsymbol{f_S}\left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z}\boldsymbol{s}^\star} - \boldsymbol{v}_{nom}\right)\right] - \boldsymbol{\delta s}^\top \boldsymbol{\delta s} \tag{6.2.6}$$

Applying the Cauchy-Schwartz inequality on (6.2.6):

$$\dot{J}(\boldsymbol{\delta s}) \leq \|\boldsymbol{\delta s}\| \left\| \boldsymbol{f_S}(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star} - \boldsymbol{Z \delta s}} - \boldsymbol{v}_{nom}) - \boldsymbol{f_S}(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star}} - \boldsymbol{v}_{nom}) \right\| - \|\boldsymbol{\delta s}\|^2 \quad (6.2.7)$$

and then applying the Lipschitz condition:

$$\dot{J}(\boldsymbol{\delta s}) \leq \|\boldsymbol{\delta s}\| \left\| \boldsymbol{C_s} \left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star} - \boldsymbol{Z \delta s}} - \sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star}} \right) \right\| - \|\boldsymbol{\delta s}\|^2 \quad (6.2.8)$$

Now by using $\sqrt{a-x} - \sqrt{a} \leq -\frac{x}{2\sqrt{a}}$, with $a = \bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^*}$ and $x = \boldsymbol{Z \delta s}$, the right-hand side of (6.2.8) can be written as:

$$\left\| \boldsymbol{C_s} \left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star} - \boldsymbol{Z \delta s}} - \sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^*}} \right) \right\| \leq \left\| -\boldsymbol{C_s} \left\{ \frac{1}{2} \operatorname{diag}^{-1} \left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star}} \right) \boldsymbol{Z \delta s} \right\} \right\| \quad (6.2.9)$$

This allows (6.2.8) to be expressed as:

$$\dot{J}(\boldsymbol{\delta s}) \leq \|\boldsymbol{\delta s}\| \left\| \boldsymbol{C_s} \left\{ \frac{1}{2} \operatorname{diag}^{-1} \left(\sqrt{\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^\star}} \right) \boldsymbol{Z \delta s} \right\} \right\| - \|\boldsymbol{\delta s}\|^2 \quad (6.2.10)$$

In recalling that $\bar{\boldsymbol{v}}^2 - \boldsymbol{Z s^*} = \boldsymbol{v}^{*2}$, where $\boldsymbol{v}^*$ is the fixed point voltage vector for the system with active inverters, and using the fact that the product of diagonal matrices is commutative, (6.2.10) can be written as:

$$\dot{J}(\boldsymbol{\delta s}) \leq \frac{\|\boldsymbol{\delta s}\|^2}{\sqrt{2}} \lambda_{max}^{1/2} \left(\left[\boldsymbol{C_p^2} + \boldsymbol{C_q^2} \right] \operatorname{diag}^{-1}(\boldsymbol{v}^{*2}) \boldsymbol{Z Z^\top} \right) - \|\boldsymbol{\delta s}\|^2 \quad (6.2.11)$$

Applying Lyapunov stability condition ($\dot{J} < 0$) on (6.2.11):

$$\lambda_{\max} \left(\left[\boldsymbol{C_p^2} + \boldsymbol{C_q^2} \right] \operatorname{diag}^{-1}(\boldsymbol{v}^{*2}) \boldsymbol{Z Z^\top} \right) < 2 \quad (6.2.12)$$

Applying the theorem $\lambda_{\max}(\boldsymbol{A}) \leq \max(r_i(\boldsymbol{A}))$ [Garren(1968)] to (6.2.13), the stability condition holds if:

$$\max_i \left(r_i \left(\left[\boldsymbol{C_p^2} + \boldsymbol{C_q^2} \right] \operatorname{diag}^{-1}(\boldsymbol{v}^{*2}) \boldsymbol{Z Z^\top} \right) \right) < 2 \quad (6.2.13)$$

Considering that $\left[\boldsymbol{C_p^2} + \boldsymbol{C_q^2} \right] \operatorname{diag}^{-1}(\boldsymbol{v}^{*2})$ are diagonal matrices, (6.2.1) can be readily obtained. □

The criterion derived in (6.2.13) links voltage stability with all inverters' piecewise linear droop control values and network parameters. Unlike other criteria found in the literature such as [Singhal *et al.*(2019), Baker *et al.*(2018)], it also ties the condition to the solution of the DistFlow equations at the fixed point for inverter operations $\boldsymbol{v}^{*2}$. The condition in (6.2.1) also emphasizes that when $\boldsymbol{v}^{*2}$ is relatively small the stability bound becomes harder to satisfy.

Numerical observation shows that the sufficient condition for network stability is quite conservative. However, the benefit of the simple expression in (6.2.1) is that setpoints for inverters can be chosen using a local policy and local voltage information. More specifically, when an inverter measures rapid fluctuations in voltage amplitude, it can react by bringing constants $\boldsymbol{C}_p$ and $\boldsymbol{C}_q$ below the bound established by (6.2.1) to restore stability. To do so, values of the coefficients $r_i(\boldsymbol{Z}\boldsymbol{Z})^\top$ should be known a priori but v_i^* may not be known. However, a conservative value of v_i^* to meet the stability condition at the time step of implementing the policy can be calculated using the voltage of the previous time step or a moving average of voltages calculated according to the time delay value as v_i^*. Considering the smart inverter control curve modeling presented in IEEE 1547 standard, the following analysis illustrates how to change $C_{p,i}$ and $C_{q,i}$ for the i^{th} inverter in response to oscillations.

Using the result in Lemma 1 and relaxing the strict inequality by introducing a stability margin $\epsilon \geq \epsilon_0$, where $\epsilon_0 > 0$ is a desired lower bound on the stability margin, (6.2.1) can be written as:

$$\left(\frac{\overline{p_i}}{\epsilon_{p,i}}\right)^2 + \left(\frac{\overline{p_i}}{\epsilon_{p,i}\sqrt{\frac{1}{\mu^2}-1}} + \frac{q_i^{lim}}{\epsilon_{q,i}^+}\right)^2 \leq \eta_i - \epsilon \tag{6.2.14}$$

where:

$$\eta_i = \frac{2\ v_i^{*2}}{r_i(\boldsymbol{Z}\boldsymbol{Z}^\top)} \tag{6.2.15}$$

By using:

$$x_i = 1/\epsilon_{p,i}^{'}, \quad y_i = 1/\epsilon_{q,i}^{+'} \tag{6.2.16}$$

(6.2.14) can be rearranged as:

$$[x_i, y_i]\boldsymbol{M}_i[x_i, y_i]^\top \leq \eta_i - \epsilon \tag{6.2.17}$$

where the entries of the matrix $\boldsymbol{M}_i$ are:

$$[\boldsymbol{M}_i]_{11} = \frac{\overline{p_i}^2}{1-\mu^2}; \quad [\boldsymbol{M}_i]_{22} = (q_i^{lim})^2; \quad [\boldsymbol{M}_i]_{12} = [\boldsymbol{M}_i]_{21} = \frac{\overline{p_i} q_i^{lim}}{\sqrt{1/\mu^2 - 1}}$$

Now the minimization of ϵ_p allows an inverter to maintain the largest amount of active power injection. Hence, the following optimization problem can be formulated :

$$\max x_i \;\; \text{subj. to (6.2.17)}, \;\; x_i > 0, y_i > 1/(2V_q^+) \tag{6.2.18}$$

Given the various positivity constraints and the bound in $\epsilon_{q,i}^{+'}$, the following is true:

$$\left(\frac{\overline{p_i}}{\epsilon'_{p,i}}\right)^2 + \left(\frac{\overline{p_i}}{\epsilon'_{p,i}\sqrt{\frac{1}{\mu^2}-1}} + \frac{q_i^{lim}}{\epsilon'_{q,i}}\right)^2 < \left(\frac{\overline{p_i}}{\epsilon'_{p,i}\sqrt{1-\mu^2}}\right)^2 \tag{6.2.19}$$

where the right-hand side is simply the left-hand side with the term $\frac{q_i^{lim}}{\epsilon'_{q,i}} = 0$. This forms an upper bound because $0 < \epsilon_{q,i}^{+'} \leq 2V_q^+$ is impossible. Therefore, x_i can be chosen such that the upper-bound matches $(\eta_i - \epsilon)$. This results in choosing $\epsilon'_{p,i}$ following (6.2.20) and setting $\epsilon_{q,i}^{+'}$ to match its upper limit $2V_q^+$. This shrinks the dead-band region in the Volt-Var characteristic and results in choosing $\epsilon'_{p,i}$ through (6.2.20):

$$\epsilon'_{p,i} = \frac{\overline{p_i}}{\sqrt{(1-\mu^2)(\eta_i - \epsilon)}} \tag{6.2.20}$$

The definition of η_i shown in (6.2.15) depicts that η for a node decreases as the electrical distance increases of that node from the substation (due to increase in $r_i(\boldsymbol{Z}\boldsymbol{Z}^\top)$ and less variation in v_i^* compared to $r_i(\boldsymbol{Z}\boldsymbol{Z}^\top)$). This suggests that the

Table 6.1: List of Use Cases

Case	Scenario
Case 1	No inverters (considered as the base case)
Case 2	Inverters without stabilization policy
Case 3	Inverters with stabilization policy

inverter furthest from the substation will have the largest $\epsilon^{'}_{p,i}$. So for multiple equal-sized inverters placed across the entire distribution network, the inverter furthest from the substation can inject more real power than others, thus providing voltage support while satisfying the stability criterion. Moreover, (6.2.20) illustrates that the local control policy for individual inverter relies on its capacity and operational parameters. There is no other independent variable that can affect the calculation of ϵ'_p when an inverter experiences voltage oscillation at its terminal. Integrating the operating condition while deriving Lyapunov analysis provides a conservative estimate to ensure that the control policy can stabilize oscillations under a given operating condition, thus ensuring the robustness of the proposed methodology.

6.3 Simulation Results

Simulations were completed on the 85 bus radial MATPOWER test case [Zimmerman *et al.*(2011)] with maximum active and reactive load being 2.571 MW and 2.622 MVar, respectively. The test case network was used to evaluate three cases as shown in Table 6.1. All simulations were completed on a Intel(R) Xeon(R) CPU E5-1630 v3 3.70 GHz computer.

Smart inverters were randomly placed at 5 of nodes, highlighted in Fig. 6.3.1, with a total installed capacity of 1.05 MW. Each inverter has combined Volt-Var

and Volt-Watt capability (Volt-Watt preference). For the simulation, each load and inverter was assigned load profiles and generation profiles, respectively, using data from Dataport [Helou *et al.*(2020)].

The profiles from Dataport used in this work are at 10 minute resolution for one day. Generation profiles include real power generation data from multiple solar inverters. Load profiles are based on household real power consumption data. For this work, spline interpolation was conducted to create 1 second load and generation profiles from the original data resulting in 86400 data points in seconds. Simulations here use data taken from load and generation profiles between $41000 - 44600$ seconds, one hour over midday, to demonstrate algorithm operation.

The output of the spline interpolation process is based on the size of actual inverters and actual household demand recorded by Dataport. Hence the data is normalized to use with the rated capacities of the inverters and the loads in the 85 bus network. The individual generation profile is used to calculate the rated apparent power s_i for the i^{th} inverter at each time step of the simulation. Figure 6.3.2a shows the average normalized load profile (averaged across 85 nodes) and average normalized generation profile (averaged across 5 nodes) spread across the 3600 second simulation period (between $41000 - 44600$ seconds in the day). The graphs show data that were averaged after the normalization process, and therefore, the peak value is not 1.0 as would be expected in a graph showing normalized profiles of each individual inverter. For cases 2 and 3, the choice of parameters is as follows:

- Time delay T_d^i for the i^{th} inverter considered is the minimum of the low-pass filter constants for real power T_p^i and reactive power T_q^i; $T_d^i = \min(T_p^i, T_q^i)$.

- The i^{th} inverter executes it's control policy when voltage flicker measured by

Table 6.2: Parameter Values for Cases 2 and 3

Parameter	Value	Parameter	Value
V_p	1.035 pu	ϵ_q^-	0.03 pu
ϵ_p	0.03 pu	ϵ	0.000001
V_q^+	1.035 pu	v^{th}	0.01 pu
V_q^-	0.965 pu	ϵ_q^+	0.03 pu
Minimum power factor	0.2		

(6.3.1) is greater than v_i^{th} pu.

$$\text{voltage flicker}(t_0, i) = \frac{\sum_{t=t_0-T_d^i+1}^{t=t_0} |v_{t,i} - v_{t-1,i}|}{T_d^i} \tag{6.3.1}$$

- The i^{th} inverter applies its local control policy at every T_d^i s based on the measured *voltage flicker*(t, i) at time step t following (6.3.1).

Table 6.2 provides the parameters used for initializing the inverters in cases 2 and 3.

The local control policy is now demonstrated under a period of time with solar generation intermittency and then a cyber-attack scenario. Figure 6.3.2b shows the local control algorithm process implemented by each individual inverter.

6.3.1 Damping Oscillation From Generation Intermittency

Simulations were completed with a 1 second time step for each of the three cases and the local policy is employed for every inverter. The time delay (T_d^i) for each inverter was chosen to be $25s$ arbitrarily. The arbitrary choice is justified by considering that the stability criterion in (6.2.1) does not include $\boldsymbol{T}_p$ and $\boldsymbol{T}_q$ and hence the time delay for each inverter dictates the time interval for successive implementation of the policy. The choice of an optimal time delay for inverters will be both network-

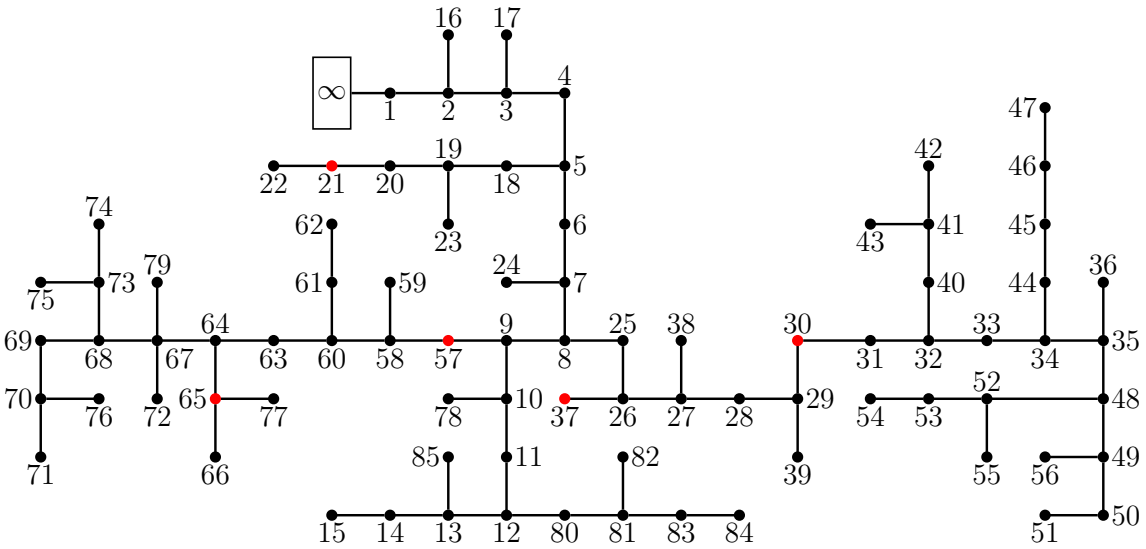
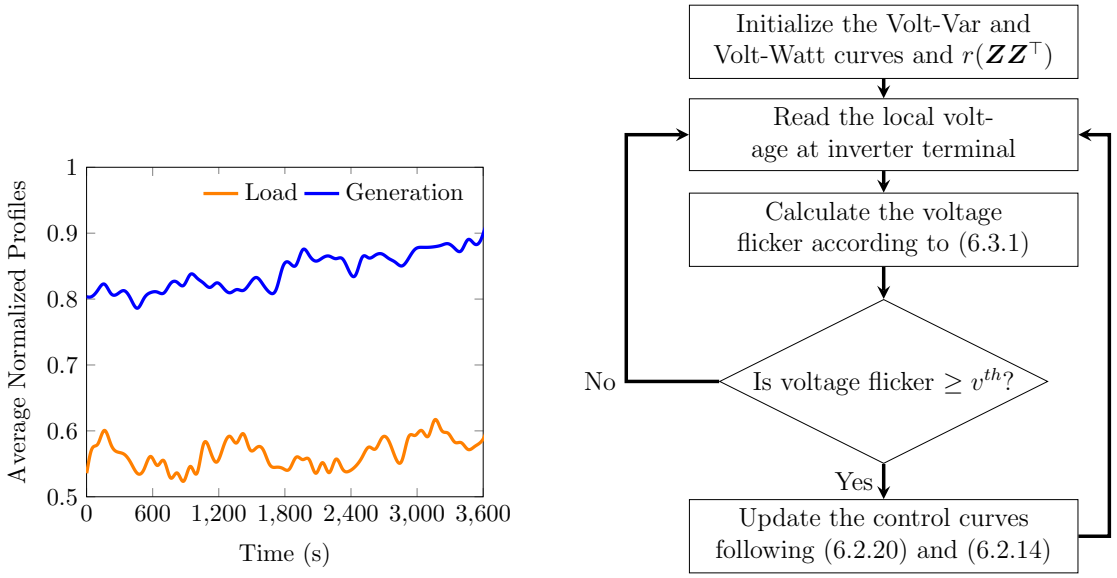


Figure 6.3.1: 85 Bus Radial MATPOWER Test Case With Red Denoting Location of Inverters

specific and inverter limited, and such considerations are out of the scope of this work that demonstrates the fundamental behavior of the policy. The inverter at bus 30 is selected to illustrate the effect of the stabilization policy.

Figure 6.3.3 shows the resulting voltage for the inverter at bus 30 and real power flow at the substation. Oscillations in real power and voltage are significant when inverters are not equipped with the stabilization policy. When the stabilization policy is active and engages according to the criteria shown in (6.3.1), the inverter at bus 30 adjusts its droop curve to mitigate voltage oscillation using only local voltage flicker measurements and without any centralized control or information sharing with others. Note that, though all the inverters are equipped with the policy, the flicker condition is only satisfied for the inverter at bus 30 for this specific load and generation profile. If a tighter limit was chosen (< 0.01 pu), the other inverters may have been triggered to adjust their droop curves to mitigate the oscillation, or it could be that only some of the inverters are called upon to mitigate the oscillation, as shown in this example. The choice of the appropriate voltage flicker threshold can be evaluated through



(a) (b)

Figure 6.3.2: (a) Normalized Average Load and Generation Profile (b) Flowchart Showing the Local Control Algorithm Functions Implemented by Each Individual Inverter

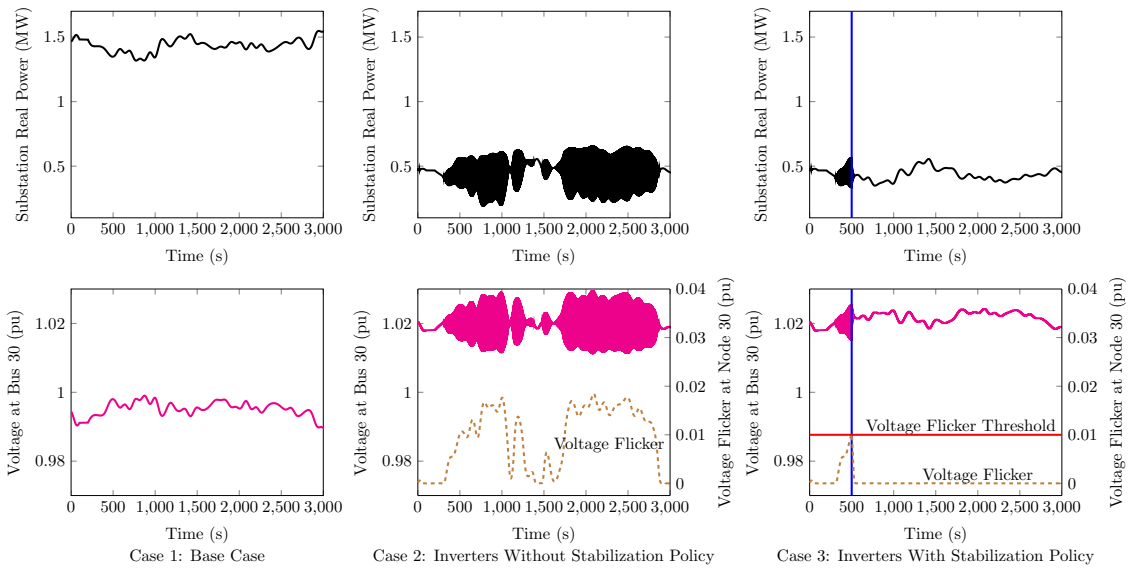


Figure 6.3.3: Substation Real Power (Top) and Bus 30 Voltage (Bottom) under Three Simulated Cases

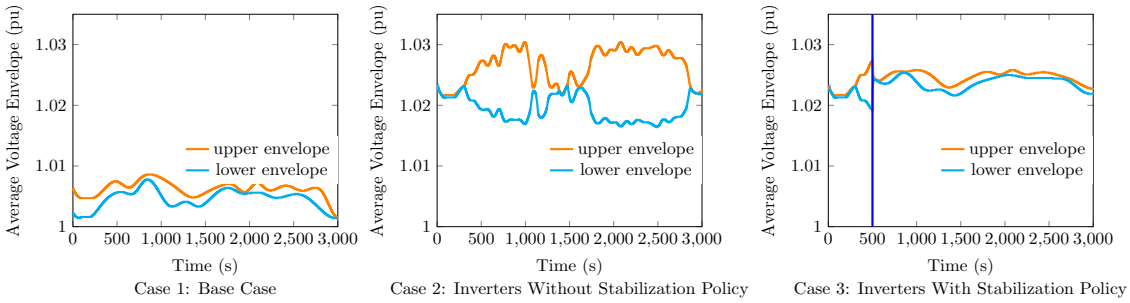


Figure 6.3.4: Average Voltage Envelope for All Nodes under Three Simulated Cases

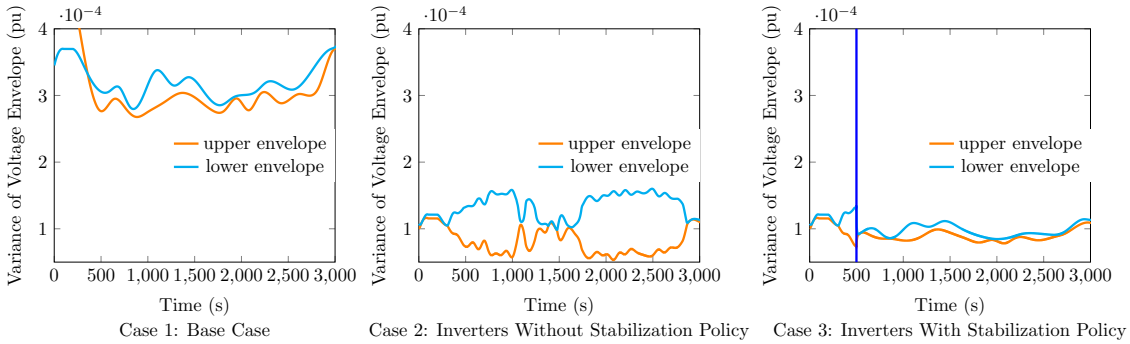


Figure 6.3.5: Variance of Voltage Envelope for All Nodes under Three Simulated Cases

simulations of specific network configurations, inverter types, and network states prior to physical implementation. Case 3 of Figure 6.3.3 shows the voltage at bus 30 and the stabilization policy activating when voltage flicker exceeds the specified threshold at $t = 500s$. Oscillations are rapidly dampened and the bus voltage stabilizes. Results in Figure 6.3.3 also show that oscillations cease in substation real power flow.

To illustrate that the developed policy can stabilize voltage oscillation in the network overall, the upper and lower envelope of voltages calculated across all nodes are shown in Figure 6.3.4 and Figure 6.3.5. This simplifies results for graphical display rather than showing voltages for all 85 nodes. Then the upper envelopes and lower envelopes are averaged for all the 85 nodes for all the simulated cases and shown in Figure 6.3.4. Variance for the upper envelope and lower envelopes for voltages across

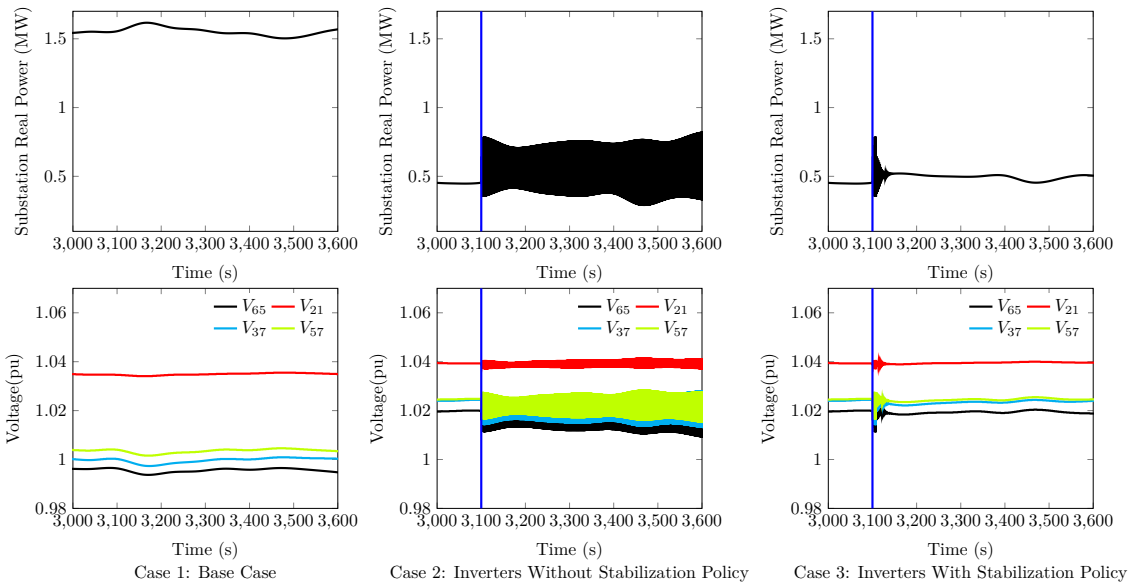


Figure 6.3.6: Substation Real Power (Top) and Unattacked Inverters' Bus Voltages (Bottom) under Three Simulated Cases During a Cyber-attack Scenario

all nodes for all the cases is shown in Figure 6.3.5. It is evident from Figure 6.3.4 that the difference between the upper and lower envelopes of inverter voltages is much lower for case 3 than for case 2. A similar trend is found for the calculated voltage variances as shown in Figure 6.3.5. These two figures provide further insight into network-level behaviors by showing that, in the scenario for inverters with a stabilization policy, the average network voltage profile closely follows the shape and variance of the scenario without inverters, whereas greater deviations exist in the scenario of inverters without a stabilization policy. This demonstrates that the proposed policy was effective in mitigating voltage instability using only local voltage information.

6.3.2 Damping Oscillation From Cyber-Attack

To illustrate that the developed policy can mitigate instability created from a cyber-attack, a scenario was generated with heterogeneous time delays for the inverters. The time delays were sampled from $\mathcal{B}(5s, 10s)$ to acknowledge that inverters from

different manufacturers can have different time delays. The simulation time window was selected during a period of the day with low generation intermittency that was insufficient to trigger the policy by voltage flicker. The severity of a cyber-attack is affected by which inverters are attacked, how inverter parameters are changed, and the operating state of the network. For this demonstration, the inverter at bus 30 was chosen because bus 30 is electrically the furthest from the substation among the buses with an inverter. This results in a higher value of $r_i(\boldsymbol{Z}\boldsymbol{Z}^\top)$, hence making the bound shown in (6.2.1) the easiest to violate from an attacker's perspective. At $t = 3100s$, a cyber-attack modifies V_p and ϵ_p, of the inverter at bus 30 to 1.02 and 0.2, respectively. The inverter also loses its capability to apply the local control policy while under attack. Set-point values for the inverter were selected to increase the slope of the Volt-Watt curve and push C_p into violation of the stability criterion derived in (6.2.1). Case 2 of Figure 6.3.6 shows how a cyber-attack on a single inverter can disrupt bus voltages for the other inverters if they are not equipped with the stability policy. Case 3 shows results for inverters with the stability policy enabled. The four inverters enable the policy according to their local measurements and individual time delays and begin to dampen oscillations and finally achieve stability after approximately 25 seconds. The benefit of the policy is also shown to reduce swings and finally stabilize substation real power flow.

CHAPTER 7
DISCUSSION

The following provides a discussion of findings from Chapters 2-6 and suggests opportunities for future research.

7.1 Scientific Implications for the Research Community

Chapter 2 introduced a framework to generate synthetic distribution feeders with real geospatial topologies. The methods employ a combination of street map data, US population census information, and prior work for synthetic transmission systems to reduce the burden of providing extensive inputs for distribution feeder generation. The software and data are public and freely available, allowing power systems researchers to develop thousands of realistic use cases. The use of publicly available data will enable researchers to voluntarily contribute and expand the framework for anywhere road network data is available. The framework used substations from the literature on synthetic transmission systems [Birchfield *et al.*(2017b)] to permit researchers to develop a combined dataset including transmission and distribution systems. That joint work will allow researchers significant opportunity to run co-simulation of transmission and distribution systems, a growing area of research with few public datasets available to advance knowledge generation. Other domain areas [Ahmad *et al.*(2020)] also using road network data can integrate with the framework and create options for multi-infrastructure simulations. As an example, some of this work has been implemented in interconnected infrastructure simulations to allow planners, engineers, and researchers to explore the effect of stressors with effects that migrate across infrastructure networks [Hamel *et al.*(2020)].

Chapter 3 presented a formulation that integrates Grid-GSP with convex relaxation-

based techniques suitable for solving AC SE and OPF problems. This work builds on prior results that demonstrated voltage phasors for power transmission systems could be viewed as the output of a complex graph filter through a GSO via the system admittance matrix. Compared to prior works that leveraged the low-rank nature of voltage phasors to reconstruct missing measurements, the proposed approach benefits from explicit knowledge of the subspace for a low-rank representation of the grid state to enhance reconstruction capabilities. Simulation results showed that the SE formulation could accurately estimate voltage phasors for single-phase and unbalanced multi-phase radial feeders under low-observability conditions, using measurements of voltage magnitude, current magnitude, and apparent power. Furthermore, the proposed approach outperforms matrix completion techniques as matrix completion can only fill gaps in missing measurements and cannot estimate values for not measured variables. Therefore, to calculate voltage phasors, the matrix completion technique requires the complex voltage phasor in Cartesian or polar formats. Hence, traditional measurements available through AMIs are insufficient for estimating voltage phasors through a matrix completion technique, prompting the use of a GSP-based proposed approach to estimate complex voltage phasors using limited data types measured by AMIs [Bernal Heredia *et al.*(2021)].

The technique for solving a centralized convex relaxation problem includes replacing the rank-1 matrix of the outer product of the voltage phasors' vector with a positive semi-definite matrix. Hence, for an n bus system, the variable being solved for is of size n^2, which becomes computationally expensive for large networks. The observation is that phasors in distribution systems can be approximated with relatively few components in the graph frequency domain, significantly less than the number of buses. This concept goes beyond generic dimensionality reduction because the principal subspace spanned by the voltage phasors is known. The low-rank representation

of the voltage phasor vector opens the door for finding optimal sampling patterns that significantly reduce the number of measurements needed for reconstruction and provides guidelines on the optimum sampling pattern, i.e., the optimum placement for sensors. In addition, the low-pass representation of the voltage phasor reduces the optimization variable resulting in achieving solutions for OPF faster than the traditional SDP approach. The formulations are generic for single-phase and three-phase networks and can include unbalanced networks with multi-phase transformers, loads, underground cables, overhead wires, and more. Simulation results for single-phase and unbalanced three-phase networks under varying operating scenarios are provided to illustrate the efficacy of the proposed method.

Works presented in Chapter 4 and Chapter 5 acknowledge that security of both physical and virtual systems must be embedded in new technologies to improve grid reliability and resiliency and reduce attacks such as unauthorized access, man-in-the-middle attacks, rogue device installation, denial of service attacks, and malicious software patching. Different attack events on cyber-physical systems presented in 1.4 have raised security concerns for both IT and OT infrastructure. These concerns are being magnified by the proliferation of grid-edge devices and IoT installed by consumers that lack appropriate security protocols to prevent intrusion and attacks that expand from distribution infrastructure to more critical systems through undesirable electrical variations or SCADA signals. Bodies such as the Federal Energy Regulatory Commission (FERC), North American Electric Reliability Corporation (NERC), and the National Institute of Standards and Technology (NIST) are increasing requirements to mitigate such vulnerabilities, yet progress is being outpaced by hackers [usc(2012)]. NIST, the Department of the Navy Chief Information Officer, and the Secretary of the Navy realize this threat and have released guidance publications in information assurance [Porche *et al.*(2012)]. These have included specific

evaluations into the value of Blockchain security solutions [Yaga *et al.*(2018)]. These evaluations have led to the adoption of blockchain for TE that has gained significant momentum as it allows mutually non-trusting agents to trade energy services in a trustless energy market. Works presented in Chapter 4 and Chapter 5 address security gaps, including random bilateral transactions that do not guarantee reliable and efficient market operation and market participants having incentives to cheat when reporting actual production/consumption figures.

A blockchain-enabled transactive energy platform entitled Electron Volt Exchange is therefore presented in Chapter 4. The integration of blockchain allowed a secured process for handling individual bids (prosumers) and collective bids (aggregators). Implementing aggregators for the distributed pricing algorithm allowed efficient use of the Hyperledger Fabric distributed architecture, as demonstrated here for a 141 bus radial network. A secure mechanism for pricing and later verifying economic transactions through a distributed consensus process is also presented. The flexibility and robustness of the approach are demonstrated through simulation and implementation using Hyperledger Fabric.

The potential of the Smart Grid is dependent on advanced information and communication technologies to enable and enhance electric grid efficiency, reliability, and resilience. An enabler of this future is having security built-in (rather than bolted onto) the associated hardware and network protocols. Chapter 5 introduced a process for cryptographic key provisioning and demonstrated laboratory evaluation of a hardware chip that embeds cyber-security into grid-edge devices to enable secure peer-to-peer transactive energy trading. The approach and chip were shown to integrate with legacy assets and could be installed by manufacturers on new assets, to allow a rapid vendor-agnostic system to scale cyber-security worldwide while providing governance support on an industry-wide scale. The hardware implementation

also integrated a blockchain to address security gaps using two-factor identification and hardware root of trust. Although transactive energy was the motivating case demonstrated in this dissertation, the approach can be used to secure any data transmitted, such as firmware updates, energy use readings, price signals, and more. This leaves significant flexibility in integrating this solution with other services that need to verify and trust grid-edge devices that can be added or removed from the network.

Chapter 6 developed and simulated an approach to enable voltage stabilization in a distribution network by updating solar PV inverter set points in real-time to counteract voltage oscillation. Piece-wise linear models of both Volt-Var and Volt-Watt functions were represented using Lipschitz functions. Lipschitz constants for both Volt-Var and Volt-Watt control functions were derived while respecting smart inverter hardware limitations for reactive power generation. Lyapunov analysis was used to derive a sufficient condition to ensure network-wide voltage stability. A local control policy was derived using the developed stability condition to adjust droop constants of control curves using only local information. Whereas traditional eigenvalue-based stability analysis can only indicate stability from the sign of eigenvalues, Lyapunov analysis allows the development of a robust control policy using the derived stability criterion. In this work, the network is described using a linear distribution power flow using v^2 instead of v, thus avoiding the assumptions made in [Baker *et al.*(2018), Zhu and Liu(2016), Farivar *et al.*(2013), Farivar *et al.*(2015)]. Performing the linearization using v^2 ties the network operating condition explicitly to the stability condition, in contrast to work found in [Baker *et al.*(2018), Singhal *et al.*(2019), Farivar *et al.*(2015)] that just connected the stability condition to the line parameters of the electrical network. The presented technique applies when droop control curve functions can be represented as Lipschitz functions. Furthermore, by exposing mechanisms that contribute to system instabilities, this work introduced a

completely decentralized policy for restoring a stable voltage profile in a distribution feeder. Simulation results completed for a radial distribution test case network were demonstrated to show how inverters can adapt their control curves to voltage oscillations caused by solar intermittency and cyber-attacks.

Implementation of the proposed algorithm in a real distribution network relies on the validity of the assumptions mentioned in Section 6.1.2. The assumptions hold in any practical application due to: 1) Monotonically increasing functions for an inverter to increase real power output when voltage is beyond the threshold voltage and absorbs or injects reactive power when the voltage is below or above nominal voltage, respectively, and 2) An unbounded derivative for control curves requires a step change rather than a ramp change in smart inverter real or reactive power production, which violates physical constraints of the inverter. In thinking of implementing the local control policy on physical inverters, these controls can be integrated directly by inverter vendors upon manufacture. A legacy inverter can be equipped with a secondary controller (e.g., Raspberry Pi) to implement the control curves and update operating set-points.

7.2 Future Work

The fundamental work presented in Chapter 2 is planned for expansion in future works using unbalanced three-phase power flow, transformers and multiple voltage levels, and voltage supporting devices to more accurately depict a real distribution network. The voltage profile can be improved by selecting the most effective buses for cap bank and voltage regulator placement using machine learning techniques and Grid-GSP. Such a model can also be exercised using time series information for the demand nodes and distributed energy resources to generate many realistic use cases for how localized generation and storage affect power flow and voltages in a distri-

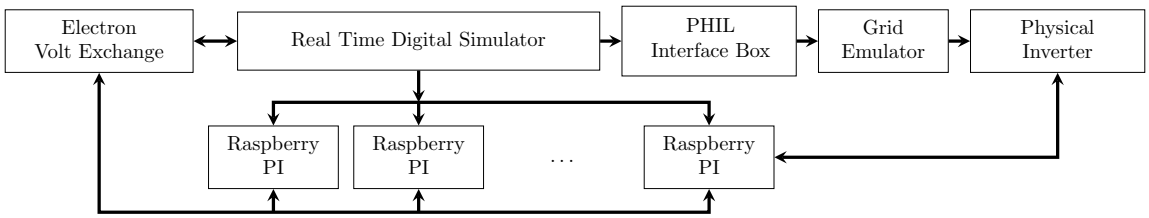


Figure 7.2.1: PHIL Simulation Architecture

bution network. Load profiles generated using historical load profiles and forecasting techniques based on the Long Short-Term Memory technique can be used for this purpose. Future work can also explore the required depth of distribution feeder design depending on grid-edge devices' penetration level by applying the continuation power flow methods. Work can be extended to provide metrics regarding the accuracy of the radial assumption of distribution networks as grid-edge devices' depth increases.

The application of Grid-GSP work presented in Chapter 3 can explore the use of the reduced-order model in a decentralized manner and combination with advanced machine learning methods for security and control applications. OPF simulation results shown in Chapter 3 used a random placement of DER assets. Future work can expand upon integrating the Grid-GSP approach with centrality analysis to find the most optimum locations of placing DER. Treating the voltage phasors as graph signals and leveraging their low-pass property can also be used to find optimum DER placement in case the underlying network information is not available.

Work presented in Chapter 4 and Chapter 5 on prevention and mitigation of cyber threats on power distribution networks can be tested using a Real-Time Digital Simulator (RTDS), a possible illustration is shown in Figure 7.2.1. A distribution network can be modeled in RTDS as the base electrical network. To mimic the behavior of prosumers, Raspberry PI's can be connected to the RTDS. A physical inverter can be connected to the RTDS through a grid emulator with the capability

of Power Hardware-in-the-Loop (PHIL) simulation. Economic dispatch points for each prosumer will reflect the output of an economic optimization problem that is modeled outside the RTDS using physical machines to emulate the EVE network. This approach captures the time delay of the distributed architecture of a blockchain network. The output of the optimization problem then becomes the input as load (positive or negative) values in the RTDS simulator. Adding physical sensors (potential transformers, watt meters, and var meters) at some distribution network nodes will allow measuring voltage, real and reactive power flow. Some sensors will be chosen randomly to report bad data (by causing calibration issues or adding noise to the reported data). Then these data will be collected by EVE from the RTDS and will be used to run an external verification algorithm. Future work will explore implementing additional Hyperledger Fabric features (e.g., idemix, smart contract packaging), other market mechanisms, and verification algorithms through distributed consensus for meshed networks.

Extension of the work presented in Chapter 6 can explore how the proposed approach can counteract voltage instabilities for networks, including the smart inverter capabilities introduced here and other voltage regulating devices such as voltage regulators and capacitor banks. Future work will also evaluate the derived policy in a PHIL simulation, including a RTDS, grid emulator, regenerative solar emulator or solar panels, programmable load bank, and smart inverter(s). An architecture that can be used to create such simulation is illustrated in Figure 7.2.2. A radial distribution network can be modeled in RTDS to represent the underlying distribution network. The inverters connected to the distribution network will be a combination of physical and simulated inverters. Physical inverters will be interfaced with the RTDS through the grid emulator and can take regenerative solar emulators or solar panels as input. Suppose a physical inverter does not include either VV or VW control

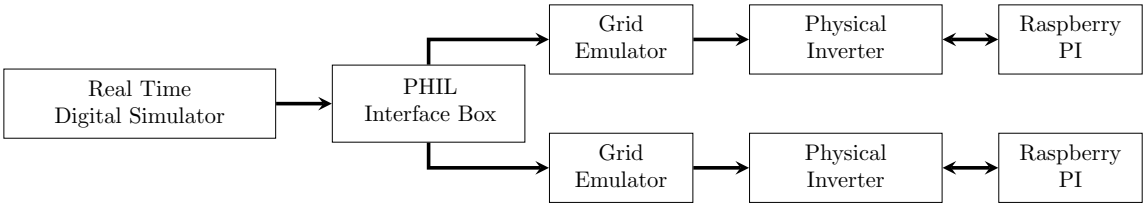


Figure 7.2.2: Real-time Simulation Architecture

capability. In that case, the inverter will be interfaced with a Raspberry PI to adjust the real and reactive power setpoints using voltage feedback. No communication is needed between the physical inverter (i.e., Raspberry PI) and the electrical network (built inside the RTDS). The simulated scenario can illustrate how a cyber-attack on smart inverters attempting to change the VW and VV control curves can create network-wide instability. Thus the effectiveness of the developed control method can be illustrated by real-time simulation.

7.3 Concluding Remarks

Integrating greater DER levels into a utility distribution system is a complex process affecting both distribution and transmission systems. New industry regulations are enabling the participation of DER in power distribution systems converting the old, passive network to an active one. Various levels of DER penetration can occur at different feeders, thus affecting the quality and reliability of service to adjacent customers. Moreover, the proliferation of DER with present-day control techniques will inevitably increase the technical, economic, reliability, and security complexities of the conventional approach to a system with centralized generation only. Advanced modeling, operation, and security features can facilitate integrating and coordinating these assets with higher-level centralized generation and transmission entities to develop the next-generation distribution systems.

[usc(2012)] "COORDINATION OF FEDERAL INFORMATION POLICY SUBCHAPTER III - INFORMATION SECURITY", Tech. rep., United States Congress (2012).

[spe(2015)] "800,000 Microinverters Remotely Retrofitted on Oahu in One Day", https://spectrum.ieee.org/energywise/green-tech/solar/in-one-day-800000-microinverters-remotely-retrofitted-on-oahu, [Online; accessed May-2020] (2015).

[sau(2017)] "Triton Malware Hits Critical Infrastructure in Saudi Arabia", https://resources.infosecinstitute.com/topic/triton-malware-hits-critical-infrastructure-saudi-arabia/, [Online; accessed June-2021] (2017).

[Sol(2018)] "Technical Note: P-Q Diagram of SolarEdge Inverters", Tech. rep., SolarEdge (2018).

[rul(2019)] "Rule 21 Interconnection Standard, CA.", http://www.cpuc.ca.gov/General.aspx?id=3962, Accessed: 2020-03-25 (2019).

[Abeysinghe *et al.*(2017)] Abeysinghe, S., J. Wu and M. Sooriyabandara, "A statistical assessment tool for electricity distribution networks", Energy Procedia **105**, 2595 – 2600, 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China (2017).

[Agung and Handayani(2020)] Agung, A. A. G. and R. Handayani, "Blockchain for smart grid", Journal of King Saud University - Computer and Information Sciences (2020).

[Ahmad *et al.*(2020)] Ahmad, N., M. Chester, E. Bondank, M. Arabi, N. Johnson and B. L. Ruddell, "A synthetic water distribution network model for urban resilience", Sustainable and Resilient Infrastructure **0**, 0, 1–15 (2020).

[Aleksieva *et al.*(2020)] Aleksieva, V., H. Valchanov and A. Huliyan, "Implementation of smart-contract, based on hyperledger fabric blockchain", in "2020 21st International Symposium on Electrical Apparatus Technologies (SIELA)", pp. 1–4 (2020).

[Alharbi and Alghazzawi(2019)] Alharbi, E. T. and D. Alghazzawi, "Two factor authentication framework using otp-sms based on blockchain", Transactions on Machine Learning and Artificial Intelligence **7**, 3 (2019).

[Andoni *et al.*(2019)] Andoni, M., V. Robu, D. Flynn, S. Abram, D. Geach, D. Jenkins, P. McCallum and A. Peacock, "Blockchain technology in the energy sector: A systematic review of challenges and opportunities", Renewable and Sustainable Energy Reviews **100**, 143 – 174 (2019).

[Androulaki *et al.*(2018)] Androulaki, E., A. Barger, V. Bortnikov, C. Cachin, K. Christidis, A. De Caro, D. Enyeart, C. Ferris, G. Laventman, Y. Manevich, S. Muralidharan, C. Murthy, B. Nguyen, M. Sethi, G. Singh, K. Smith, A. Sorniotti, C. Stathakopoulou, M. Vukolić, S. W. Cocco and J. Yellick, "Hyperledger fabric: A distributed operating system for permissioned blockchains", in "Proceedings of the Thirteenth EuroSys Conference", EuroSys '18 (Association for Computing Machinery, New York, NY, USA, 2018).

[Aurenhammer(1991)] Aurenhammer, F., "Voronoi diagrams—a survey of a fundamental geometric data structure", ACM Comput. Surv. **23**, 3, 345–405 (1991).

[Baker *et al.*(2018)] Baker, K., A. Bernstein, E. Dall'Anese and C. Zhao, "Network-cognizant voltage droop control for distribution grids", IEEE Transactions on Power Systems **33**, 2, 2098–2108 (2018).

[Baptista *et al.*(2005)] Baptista, E. C., E. A. Belati and G. R. da Costa, "Logarithmic barrier-augmented lagrangian function to the optimal power flow problem", International Journal of Electrical Power & Energy Systems **27**, 7, 528–532 (2005).

[Baran and Wu(1989a)] Baran, M. and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system", IEEE Trans. Power Del. **4**, 1, 735–743 (1989a).

[Baran and Wu(1989b)] Baran, M. E. and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing", IEEE Transactions on Power Delivery **4**, 2, 1401–1407 (1989b).

[Barker and Barker(2018)] Barker, E. and W. C. Barker, "Recommendations for Key Management", (2018).

[Baudin(2014)] Baudin, P., *Appendix 3: Moments of Normal Random Vectors*, pp. 713–717 (John Wiley & Sons, Ltd, 2014).

[Behboodi *et al.*(2018)] Behboodi, S., D. P. Chassin, N. Djilali and C. Crawford, "Transactive control of fast-acting demand response based on thermostatic loads in real-time retail electricity markets", Applied Energy **210**, 1310 – 1320 (2018).

[Bernal Heredia *et al.*(2021)] Bernal Heredia, W., D. Cutler and J. Dean, "Case study: Field evaluation of a low-cost circuit-level electrical submetering system", Tech. rep. (2021).

[Birchfield *et al.*(2017a)] Birchfield, A. B., E. Schweitzer, M. Athari, T. Xu, T. J. Overbye, A. Scaglione and Z. Wang, "A metric-based validation process to assess the realism of synthetic power grids", Energies **10**, 8, 1233 (2017a).

[Birchfield *et al.*(2017b)] Birchfield, A. B., T. Xu, K. M. Gegner, K. S. Shetye and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks", IEEE Transactions on Power Systems **32**, 4, 3258–3265 (2017b).

[Boeing(2017)] Boeing, G., "OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks", Computers, Environment and Urban Systems **65**, 126 – 139 (2017).

[Bollen and Rönnberg(2017)] Bollen, M. H. J. and S. K. Rönnberg, "Hosting capacity of the power grid for renewable electricity production and new large consumption equipment", Energies (2017).

[Boyd *et al.*(2011)] Boyd, S., N. Parikh, E. Chu, B. Peleato and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers", Foundations and Trends in Machine learning **3**, 1, 1–122 (2011).

[Brase and Brown(2009)] Brase, J. M. and D. L. Brown, "Modeling, simulation and analysis of complex networked systems", Tech. rep., Lawrence Livermore National Laboratory (2009).

[Braslavsky *et al.*(2018)] Braslavsky, J. H., L. D. Collins and J. K. Ward, "Voltage stability in a grid-connected inverter with automatic volt-watt and volt-var functions", IEEE Transactions on Smart Grid pp. 1–1 (2018).

[California's Community Choice Program(2020)] California's Community Choice Program, URL `https://cal-cca.org/about/` (2020).

[Callaway(2009)] Callaway, D. S., "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy", Energy Conversion and Management **50**, 5, 1389–1400 (2009).

[Carvallo *et al.*(2019)] Carvallo, J.-P., J. Taneja, D. Callaway and D. M. Kammen, "Distributed resources shift paradigms on power system design, planning, and operation: An application of the GAP model", Proceedings of the IEEE **107**, 9, 1906–1922 (2019).

[Casper and Papa(2011)] Casper, W. D. and S. M. Papa, *Root of Trust*, pp. 1057–1060 (Springer US, Boston, MA, 2011).

[Chang *et al.*(2012)] Chang, T.-H., M. Alizadeh and A. Scaglione, "Coordinated home energy management for real-time power balancing", in "2012 IEEE Power and Energy Society General Meeting", pp. 1–8 (IEEE, 2012).

[Che *et al.*(2019)] Che, Z., Y. Wang, J. Zhao, Y. Qiang, Y. Ma and J. Liu, "A distributed energy trading authentication mechanism based on a consortium blockchain", Energies **12**, 15 (2019).

[Christidis *et al.*(2021)] Christidis, K., D. Sikeridis, Y. Wang and M. Devetsikiotis, "A framework for designing and evaluating realistic blockchain-based local energy markets", Applied Energy **281**, 115963 (2021).

[Coignard *et al.*(2018)] Coignard, J., E. Munsing, J. MacDonald and J. Mather, "Co-simulation framework for blockchain based market designs and grid simulations", in "2018 IEEE Power Energy Society General Meeting (PESGM)", pp. 1–5 (2018).

[Contaxis *et al.*(1986)] Contaxis, G. C., C. Delkis and G. Korres, "Decoupled optimal load flow using linear or quadratic programming", IEEE Transactions on Power Systems **1**, 2, 1–7 (1986).

[Danzi *et al.*(2017)] Danzi, P., M. Angjelichinoski, Č. Stefanović and P. Popovski, "Distributed proportional-fairness control in microgrids via blockchain smart contracts", in "2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)", pp. 45–51 (IEEE, 2017).

[Danzi *et al.*(2018)] Danzi, P., S. Hambridge, Č. Stefanović and P. Popovski, "Blockchain-based and multi-layered electricity imbalance settlement architecture", in "2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)", pp. 1–7 (2018).

[Deka *et al.*(2015)] Deka, D., R. Baldick and S. Vishwanath, "Optimal data attacks on power grids: Leveraging detection & measurement jamming", in "2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)", pp. 392–397 (IEEE, 2015).

[Donti *et al.*(2020)] Donti, P. L., Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang and Y. Zhang, "Matrix completion for low-observability voltage estimation", IEEE Transactions on Smart Grid **11**, 3, 2520–2530 (2020).

[Drayer and Routtenberg(2020)] Drayer, E. and T. Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing", IEEE Systems Journal **14**, 2, 1886–1896 (2020).

[Dugan and Santoso(2003)] Dugan, R. C. and S. Santoso, "An example of 3-phase transformer modeling for distribution system analysis", in "2003 IEEE PES Transmission and Distribution Conference and Exposition (IEEE Cat. No.03CH37495)", vol. 3, pp. 1028–1032 vol.3 (2003).

[Dworkin(2015)] Dworkin, M. J., "SHA-3 standard: Permutation-based hash and extendable-output functions", (2015).

[Farivar *et al.*(2013)] Farivar, M., L. Chen and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems", in "52nd IEEE Conference on Decision and Control", pp. 4329–4334 (2013).

[Farivar *et al.*(2015)] Farivar, M., X. Zho and L. Che, "Local voltage control in distribution systems: An incremental control algorithm", in "2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)", pp. 732–737 (2015).

[Fernandes *et al.*(2020)] Fernandes, A., V. Rocha, A. F. d. Conceição and F. Horita, "Scalable architecture for sharing ehr using the hyperledger blockchain", in "2020 IEEE International Conference on Software Architecture Companion (ICSA-C)", pp. 130–138 (2020).

[Ferreira and Martins(2018)] Ferreira, J. C. and A. L. Martins, "Building a community of users for open market energy", Energies **11**, 9 (2018).

[Flanigan(2018)] Flanigan, J., "Zero Trust Network Model", (2018).

[Gai *et al.*(2019)] Gai, K., Y. Wu, L. Zhu, M. Qiu and M. Shen, "Privacy-preserving energy trading using consortium blockchain in smart grid", IEEE Transactions on Industrial Informatics (2019).

[Gan *et al.*(2015)] Gan, L., N. Li, U. Topcu and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks", IEEE Transactions on Automatic Control **60**, 1, 72–87 (2015).

[Gan and Low(2014)] Gan, L. and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks", in "2014 Power Systems Computation Conference", pp. 1–9 (2014).

[Garg *et al.*(2018)] Garg, A., M. Jalali, V. Kekatos and N. Gatsis, "Kernel-based learning for smart inverter control", in "2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)", pp. 875–879 (2018).

[Garren(1968)] Garren, K. R., "NASA Technical Note: Bounds For The Eigenvalues of A Matrix ", Tech. rep., National Aeronautics And Space Administration (1968).

[Gavili and Zhang(2017)] Gavili, A. and X. Zhang, "On the shift operator, graph frequency, and optimal filtering in graph signal processing", IEEE Transactions on Signal Processing **65**, 23, 6303–6318 (2017).

[Gegner *et al.*(2016)] Gegner, K. M., A. B. Birchfield, , K. S. Shetye and T. J. Overbye, "A methodology for the creation of geographically realistic synthetic power flow models", in "2016 IEEE Power and Energy Conference at Illinois (PECI)", pp. 1–6 (2016).

[Ghasemi and Parniani(2016)] Ghasemi, M. A. and M. Parniani, "Prevention of distribution network overvoltage by adaptive droop-based active and reactive power control of pv systems", Electric Power Systems Research **133**, 313 – 327 (2016).

[Gorog and Boult(2018)] Gorog, C. and T. E. Boult, "Solving global cybersecurity problems by connecting trust using blockchain", in "2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)", pp. 1425–1432 (2018).

[Grant and Boyd(2014)] Grant, M. and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1", `http://cvxr.com/cvx` (2014).

[Guan *et al.*(2021)] Guan, Z., X. Lu, W. Yang, L. Wu, N. Wang and Z. Zhang, "Achieving efficient and privacy-preserving energy trading based on blockchain and abe in smart grid", Journal of Parallel and Distributed Computing **147**, 34 – 45 (2021).

[Gupta and Rani(2016)] Gupta, S. and R. Rani, "A comparative study of elasticsearch and couchdb document oriented databases", in "2016 International Conference on Inventive Computation Technologies (ICICT)", vol. 1, pp. 1–4 (2016).

[Gür *et al.*(2019)] Gür, A. Ö., Ş. Öksüzer and E. Karaarslan, "Blockchain based metering and billing system proposal with privacy protection for the electric network", in "2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)", pp. 204–208 (2019).

[Hamel *et al.*(2020)] Hamel, D., N. Johnson, A. Scaglione, M. Chester and G. Morris, "Systems and methods for a resilient infrastructure simulation environment", (US Patent:20200287798, 2020).

[Hasan *et al.*(2020)] Hasan, F., A. Kargarian and A. Mohammadi, "A survey on applications of machine learning for optimal power flow", in "2020 IEEE Texas Power and Energy Conference (TPEC)", pp. 1–6 (2020).

[Hasnat and Rahnamay-Naeini(2020)] Hasnat, M. A. and M. Rahnamay-Naeini, "Detection and locating cyber and physical stresses in smart grids using graph signal processing", (2020).

[Hayes *et al.*(2020)] Hayes, B., S. Thakur and J. Breslin, "Co-simulation of electricity distribution networks and peer to peer energy trading platforms", International Journal of Electrical Power and Energy Systems **115**, 105419 (2020).

[Heidari *et al.*(2018)] Heidari, R., M. M. Seron and J. H. Braslavsky, "A decentralised adaptive voltage droop control for inverters in power distribution networks", in "2018 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)", pp. 577–582 (2018).

[Heilman *et al.*(2015)] Heilman, E., A. Kendler, A. Zohar and S. Goldberg, "Eclipse attacks on bitcoin's peer-to-peer network", in "24th USENIX Security Symposium (USENIX Security 15)", pp. 129–144 (2015).

[Helou *et al.*(2020)] Helou, R. E., D. Kalathil and L. Xie, "Communication-free voltage regulation in distribution networks with deep pv penetration", in "Proceedings of the 53rd Hawaii International Conference on System Sciences", (2020).

[Hong *et al.*(2013)] Hong, S., N. C. Rodia and K. Olukotun, "On fast parallel detection of strongly connected components (scc) in small-world graphs", in "Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis", (2013).

[Hoor and Sheng(2011)] Hoor, T. H. and Q. Z. Sheng, "Trust as a Service: A Framework for Trust Management in Cloud Environments", International Conference on Web Information Systems Engineering **Web Information System Engineering**, WISE 2011, 314–321 (2011).

[Hu *et al.*(2019)] Hu, J., Z. Li, J. Zhu and J. M. Guerrero, "Voltage stabilization: A critical step toward high photovoltaic penetration", IEEE Industrial Electronics Magazine **13**, 2, 17–30 (2019).

[Hua *et al.*(2020)] Hua, W., J. Jiang, H. Sun and J. Wu, "A blockchain based peer-to-peer trading framework integrating energy and carbon markets", Applied Energy **279**, 115539 (2020).

[Huque(2015)] Huque, A., "Smart inverter grid support functions and potential impact on reliability", 2015 nrel photovoltaic reliability workshop, Electric Power Research Institute (2015).

[Hussain *et al.*(2019)] Hussain, S. M. S., S. M. Farooq and T. S. Ustun, "Implementation of blockchain technology for energy trading with smart meters", in "2019 Innovations in Power and Advanced Computing Technologies (i-PACT)", vol. 1, pp. 1–5 (2019).

[Hyperledger Architecture Working Group(2017)] Hyperledger Architecture Working Group, "Introduction to hyperledger business blockchain design philosophy and consensus", Tech. rep., Hyperledger Architecture Working Group (2017).

[IEEE Standards Coordinating Committee 21(2018)] IEEE Standards Coordinating Committee 21, "IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces", IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003) pp. 1–138 (2018).

[Jahangiri and Aliprantis(2013)] Jahangiri, P. and D. C. Aliprantis, "Distributed volt/var control by pv inverters", IEEE Transactions on Power Systems **28**, 3, 3429–3439 (2013).

[Janko and Johnson(2018)] Janko, S. A. and N. G. Johnson, "Scalable multi-agent microgrid negotiations for a transactive energy market", Applied Energy **229**, 715 – 727 (2018).

[Jogunola *et al.*(2019)] Jogunola, O., M. Hammoudeh, B. Adebisi and K. Anoh, "Demonstrating blockchain-enabled peer-to-peer energy trading and sharing", in "2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)", pp. 1–4 (2019).

[Jones *et al.*(2015)] Jones, M., J. Bradley and N. Sakimura, "JSON web token (JWT)", Tech. rep., URL `https://doi.org/10.17487/rfc7519` (2015).

[Kang *et al.*(2017)] Kang, J., R. Yu, X. Huang, S. Maharjan, Y. Zhang and E. Hossain, "Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains", IEEE Transactions on Industrial Informatics **13**, 6, 3154–3164 (2017).

[Karakoç *et al.*(2018)] Karakoç, N., A. Scaglione and A. Nedić, "Multi-layer decomposition of optimal resource sharing problems", in "Proc. in IEEE Conf. on Decision and Control (CDC)", pp. 1–6 (Miami Beach, FL, 2018).

[Kay(1993)] Kay, S. M., *Fundamentals of Statistical Signal Processing:Estimation Theory*, vol. 2 of *Prentice-Hall signal processing series* (Prentice Hall PTR, Upper Saddle River, NJ, 1993).

[Kersting(2002)] Kersting, W. H., *Distribution System Modeling and Analysis* (CRC Press, Boca Raton, Florida, 2002).

[Khalili *et al.*(2019)] Khalili, T., A. Jafari, M. Abapour and B. Mohammadi-Ivatloo, "Optimal battery technology selection and incentive-based demand response program utilization for reliability improvement of an insular microgrid", Energy **169**, 92 – 104 (2019).

[Khovratovich and Law(2017)] Khovratovich, D. and J. Law, "BIP32-Ed25519: Hierarchical deterministic keys over a non-linear keyspace", in "2017 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)", pp. 27–31 (2017).

[Kilic and Gülgen(2020)] Kilic, B. and F. Gülgen, "Investigating the quality of reverse geocoding services using text similarity techniques and logistic regression analysis", Cartography and Geographic Information Science **47**, 4, 336–349 (2020).

[Kindervag(2010)] Kindervag, J., "Build Security Into Your Network's DNA: The Zero Trust Network Architecture", (2010).

[Koch(2015)] Koch, S., "Chapter 2 - assessment of revenue potentials of ancillary service provision by flexible unit portfolios", in "Energy Storage for Smart Grids", edited by P. Du and N. Lu, pp. 35 – 66 (Academic Press, Boston, 2015).

[Krishnan *et al.*(2017)] Krishnan, V. K., B. S. Palmintier, B. S. Hodge, E. T. Hale, T. Elgindy, B. Bugbee, M. N. Rossol, A. J. Lopez, D. Krishnamurthy, C. Vergara, C. M. Domingo, F. Postigo, F. de Cuadra, T. Gomez, P. Duenas, M. Luke, V. Li, M. Vinoth and S. Kadankodu, "Smart-ds: Synthetic models for advanced, realistic testing: Distribution systems and scenarios", Tech. rep. (2017).

[Kruskal(1956)] Kruskal, J. B., "On the shortest spanning subtree of a graph and the traveling salesman problem", Proceedings of the American Mathematical Society **7**, 1 (1956).

[Laszka *et al.*(2018)] Laszka, A., S. Eisele, A. Dubey, G. Karsai and K. Kvaternik, "Transax: A blockchain-based decentralized forward-trading energy exchanged for transactive microgrids", in "2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)", pp. 918–927 (2018).

[Lavaei and Low(2012)] Lavaei, J. and S. H. Low, "Zero duality gap in optimal power flow problem", IEEE Transactions on Power Systems **27**, 1, 92–107 (2012).

[Leeuwen *et al.*(2020)] Leeuwen, G. V., T. AlSkaif, M. Gibescu and W. V. Sark, "An integrated blockchain-based energy management platform with bilateral trading for microgrid communities", Applied Energy **263**, 114613 (2020).

[Li *et al.*(2018)] Li, H., A. L. Bornsheuer, T. Xu, A. B. Birchfield and T. J. Overbye, "Load modeling in synthetic electric grids", in "2018 IEEE Texas Power and Energy Conference (TPEC)", pp. 1–6 (2018).

[Li *et al.*(2011a)] Li, L., A. Scaglione and J. H. Manton, "Distributed principal subspace estimation in wireless sensor networks", IEEE Journal of Selected Topics in Signal Processing **5**, 4, 725–738 (2011a).

[Li *et al.*(2011b)] Li, N., L. Chen and S. H. Low, "Optimal demand response based on utility maximization in power networks", in "2011 IEEE power and energy society general meeting", pp. 1–8 (IEEE, 2011b).

[Liang *et al.*(2021)] Liang, M., Y. Meng, J. Wang, D. L. Lubkeman and N. Lu, "Feedergan: Synthetic feeder generation via deep graph adversarial nets", IEEE Transactions on Smart Grid **12**, 2, 1163–1173 (2021).

[Lin *et al.*(2019)] Lin, J., M. Pipattanasomporn and S. Rahman, "Comparative analysis of auction mechanisms and bidding strategies for P2P solar transactive energy markets", Applied Energy **255**, 113687 (2019).

[Liu *et al.*(2019)] Liu, B., H. Wu, Y. Zhang, R. Yang and A. Bernstein, "Robust matrix completion state estimation in distribution systems", in "2019 IEEE Power Energy Society General Meeting (PESGM)", pp. 1–5 (2019).

[Liu *et al.*(2011)] Liu, Y., P. Ning and M. K. Reiter, "False data injection attacks against state estimation in electric power grids", ACM Transactions on Information and System Security (TISSEC) **14**, 1, 13 (2011).

[Long *et al.*(2019)] Long, Q., J. Wang, D. Lubkeman, N. Lu and P. Chen, "Volt-var optimization of distribution systems for coordinating utility voltage control with smart inverters", in "2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)", pp. 1–5 (2019).

[Low(2014a)] Low, S. H., "Convex relaxation of optimal power flow—part i: Formulations and equivalence", IEEE Transactions on Control of Network Systems **1**, 1, 15–27 (2014a).

[Low(2014b)] Low, S. H., "Convex relaxation of optimal power flow—part ii: Exactness", IEEE Transactions on Control of Network Systems **1**, 2, 177–189 (2014b).

[Lu *et al.*(2019)] Lu, X., Z. Guan, X. Zhou, X. Du, L. Wu and M. Guizani, "A secure and efficient renewable energy trading scheme based on blockchain in smart grid", in "2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)", pp. 1839–1844 (2019).

[Luo *et al.*(2019)] Luo, X., X. Wang, M. Zhang and X. Guan, "Distributed detection and isolation of bias injection attack in smart energy grid via interval observer", Applied Energy **256**, 113703 (2019).

[Madani *et al.*(2016)] Madani, R., M. Ashraphijuo, J. Lavaei and R. Baldick, "Power system state estimation with a limited number of measurements", in "2016 IEEE 55th Conference on Decision and Control (CDC)", pp. 672–679 (2016).

[Madani *et al.*(2017)] Madani, R., J. Lavaei, R. Baldick and A. Atamtürk, "Power system state estimation and bad data detection by means of conic relaxation", in "HICSS", (2017).

[Madani *et al.*(2015)] Madani, R., S. Sojoudi and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks", IEEE Transactions on Power Systems **30**, 1, 199–211 (2015).

[Mahmoudi *et al.*(2014a)] Mahmoudi, N., T. K. Saha and M. Eghbal, "A new trading framework for demand response aggregators", in "2014 IEEE PES General Meeting — Conference Exposition", pp. 1–5 (2014a).

[Mahmoudi *et al.*(2014b)] Mahmoudi, N., T. K. Saha and M. Eghbal, "A new trading framework for demand response aggregators", in "2014 IEEE PES General Meeting — Conference Exposition", pp. 1–5 (2014b).

[Mateo Domingo *et al.*(2011)] Mateo Domingo, C., T. Gomez San Roman, A. Sanchez-Miralles, J. P. Peco Gonzalez and A. Candela Martinez, "A reference network model for large-scale distribution planning with automatic street map generation", IEEE Transactions on Power Systems **26**, 1, 190–197 (2011).

[Mestav *et al.*(2019)] Mestav, K. R., J. Luengo-Rozas and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning", IEEE Transactions on Power Systems **34**, 6, 4910–4920 (2019).

[Mestav and Tong(2019)] Mestav, K. R. and L. Tong, "State estimation in smart distribution systems with deep generative adversary networks", in "2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)", pp. 1–6 (2019).

[Metez *et al.*(2020)] Metez, A., P. Gunjan, R. R. Labastida, S. Shepard and E. Woods, "Integrated DER: Orchestrating the Grid's Last Mile", Tech. rep., GridHouse (2020).

[Mohanta *et al.*(2018)] Mohanta, B. K., S. S. Panda and D. Jena, "An overview of smart contract and use cases in blockchain technology", in "2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)", pp. 1–4 (2018).

[Münsing *et al.*(2017)] Münsing, E., J. Mather and S. Moura, "Blockchains for decentralized optimization of energy resources in microgrid networks", in "2017 IEEE Conference on Control Technology and Applications (CCTA)", pp. 2164–2171 (IEEE, 2017).

[Musleh *et al.*(2019)] Musleh, A. S., G. Yao and S. M. Muyeen, "Blockchain applications in smart grid–review and frameworks", IEEE Access **7**, 86746–86757 (2019).

[Muthukaruppan and Baran(2020)] Muthukaruppan, V. and M. E. Baran, "Implementing a decentralized volt/var scheme on a smart distribution system", in "2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)", pp. 1–5 (2020).

[Mylrea and Gourisetti(2017)] Mylrea, M. and S. N. G. Gourisetti, "Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security", in "2017 Resilience Week (RWS)", pp. 18–23 (2017).

[Neis and Zipf(2012)] Neis, P. and A. Zipf, "Analyzing the contributor activity of a volunteered geographic information project — the case of openstreetmap", ISPRS International Journal of Geo-Information **1**, 2, 146–165 (2012).

[Nguyen *et al.*(2019)] Nguyen, H. T., S. Battula, R. R. Takkala, Z. Wang and L. Tesfatsion, "An integrated transmission and distribution test system for evaluation of transactive energy designs", Applied Energy **240**, 666 – 679 (2019).

[NIST(2013)] NIST, "Glossary of Key Information Security Terms", (2013).

[NIST(2019)] NIST, "Recommendation for Key Mnagament: Best practices for key management organizations", (2019).

[Olivella-Rosell *et al.*(2018)] Olivella-Rosell, P., P. Lloret-Gallego, I. Munné-Collado, R. Villafafila-Robles, A. Sumper, S. O. Ottessen, J. Rajasekharan and B. A. Bremdal, "Local flexibility market design for aggregators providing multiple flexibility services at distribution network level", Energies **11**, 4 (2018).

[Olivier *et al.*(2016)] Olivier, F., P. Aristidou, D. Ernst and T. Van Cutsem, "Active management of low-voltage networks for mitigating overvoltages due to photovoltaic units", IEEE Transactions on Smart Grid **7**, 2, 926–936 (2016).

[Ortega *et al.*(2018)] Ortega, A., P. Frossard, J. Kovačević, J. M. F. Moura and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications", Proceedings of the IEEE **106**, 5, 808–828 (2018).

[Park *et al.*(2018)] Park, W., D. Hwang and K. Kim, "A totp-based two factor authentication scheme for hyperledger fabric blockchain", in "2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)", pp. 817–819 (2018).

[Pipattanasomporn *et al.*(2018)] Pipattanasomporn, M., M. Kuzlu and S. Rahman, "A blockchain-based platform for exchange of solar energy: Laboratory-scale implementation", in "2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)", pp. 1–9 (2018).

[Pompodakis *et al.*(2016)] Pompodakis, E. E., I. A. Drougakis, I. S. Lelis and M. C. Alexiadis, "Photovoltaic systems in low-voltage networks and overvoltage correction with reactive power control", IET Renewable Power Generation **10**, 3, 410–417 (2016).

[Pop *et al.*(2018)] Pop, C., T. Cioara, M. Antal, I. Anghel, I. Salomie and M. Bertoncini, "Blockchain based decentralized management of demand response programs in smart energy grids", Sensors **18**, 1, 162 (2018).

[Porche *et al.*(2012)] Porche, I., S. McKay, M. McKErnan, R. W. Button and E. Axelband, *Rapid acquisition and fielding for information assurance and cyber security in the Navy* (RAND, NATIONAL DEFENSE RESEARCH Institute, 2012).

[Postigo Marcos *et al.*(2017)] Postigo Marcos, F. E., C. M. Domingo, T. G. San Roman, B. S. Palmintier, B.-M. Hodge, V. K. Krishnan, F. de Cuadra Garcia and B. A. Mather, "A review of power distribution test feeders in the united states and the need for synthetic representative networks", Energies (Basel) **10**, 11 (2017).

[Primadianto and Lu(2017)] Primadianto, A. and C. Lu, "A review on distribution system state estimation", IEEE Transactions on Power Systems **32**, 5, 3875–3883 (2017).

[Primadianto and Lu(2017)] Primadianto, A. and C.-N. Lu, "A review on distribution system state estimation", IEEE Transactions on Power Systems **32**, 5, 3875–3883 (2017).

[Pukhrem *et al.*(2017)] Pukhrem, S., M. Basu, M. F. Conlon and K. Sunderland, "Enhanced network voltage management techniques under the proliferation of rooftop solar pv installation in low-voltage distribution network", IEEE Journal of Emerging and Selected Topics in Power Electronics **5**, 2, 681–694 (2017).

[Quintana *et al.*(2000)] Quintana, V., G. Torres and J. Medina-Palomo, "Interior-point methods and their applications to power systems: a classification of publications and software codes", IEEE Transactions on Power Systems **15**, 1, 170–176 (2000).

[Ramakrishna and Scaglione(2019a)] Ramakrishna, R. and A. Scaglione, "Detection of false data injection attack using graph signal processing for the power grid", in "2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)", pp. 1–5 (2019a).

[Ramakrishna and Scaglione(2019b)] Ramakrishna, R. and A. Scaglione, "On modeling voltage phasor measurements as graph signals", in "2019 IEEE Data Science Workshop (DSW)", pp. 275–279 (2019b).

[Ramakrishna and Scaglione(2021)] Ramakrishna, R. and A. Scaglione, "Grid-graph signal processing (grid-gsp): A graph signal processing framework for the power grid", (2021).

[Ramakrishna *et al.*(2020)] Ramakrishna, R., H. T. Wai and A. Scaglione, "A user guide to low-pass graph signal processing and its applications: Tools and applications", IEEE Signal Processing Magazine **37**, 6, 74–85 (2020).

[Ravi and Scaglione(2019)] Ravi, N. and A. Scaglione, "Detection and isolation of adversaries in decentralized optimization for non-strongly convex objectives", IFAC-PapersOnLine **52**, 20, 381 – 386, 8th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2019 (2019).

[Ribeiro *et al.*(2020)] Ribeiro, L. C., F. L. Vieira, B. D. Bonatto, A. C. Z. de Souza and P. F. Ribeiro, "Chapter 5 - modeling and simulation of active electrical distribution systems using the opendss", in "Decision Making Applications in Modern Power Systems", pp. 121–152 (Academic Press, 2020).

[Sabounchi and Wei(2017)] Sabounchi, M. and J. Wei, "Towards resilient networked microgrids: Blockchain-enabled peer-to-peer electricity trading mechanism", in "2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)", pp. 1–5 (2017).

[Sadnan and Dubey(2020)] Sadnan, R. and A. Dubey, "Real-time distributed control of smart inverters for network-level optimization", in "2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)", pp. 1–6 (2020).

[Sagan *et al.*(2021)] Sagan, A., Y. Liu and A. Bernstein, "Decentralized low-rank state estimation for power distribution systems", IEEE Transactions on Smart Grid pp. 1–1 (2021).

[Saleem *et al.*(2020)] Saleem, D., A. Hasandka, C. W. Lai, D. Jose and C. M. Howerter, "Design considerations of a cryptographic module for distributed energy resources", in "IEEE International Conference on Consumer Electronics (ICCE)", (2020).

[SANS Industrial Control Systems(2016)] SANS Industrial Control Systems, "Analysis of the Cyber Attack on the Ukrainian Power Grid", Tech. rep., SANS Industrial Control Systems (2016).

[Sanseverino *et al.*(2017)] Sanseverino, E. R., M. L. Di Silvestre, P. Gallo, G. Zizzo and M. Ippolito, "The blockchain in microgrids for transacting energy and attributing losses", in "2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)", pp. 925–930 (2017).

[Sarfaraz *et al.*(2016)] Sarfaraz, A. Bansal and S. Singh, "Optimal allocation and sizing of distributed generation for power loss reduction", in "International Conference Workshop on Electronics Telecommunication Engineering (ICWET 2016)", pp. 15–20 (2016).

[Schneider *et al.*(2009)] Schneider, K. P., Y. Chen, D. Engle and D. Chassin, "A taxonomy of north american radial distribution feeders", in "2009 IEEE Power Energy Society General Meeting", pp. 1–6 (2009).

[Schweitzer *et al.*(2017)] Schweitzer, E., A. Scaglione, A. Monti and G. A. Pagani, "Automated generation algorithm for synthetic medium voltage radial distribution systems", IEEE Journal on Emerging and Selected Topics in Circuits and Systems **7**, 2, 271–284 (2017).

[Seal(2017)] Seal, B., "Common Functions for Smart Inverters, 4th Ed.", Tech. Rep. 3002008217, Electric Power Research Institute (2017).

[Shi and Baran(2019)] Shi, Y. and M. Baran, "A gradient based decentralized volt/-var optimization scheme for distribution systems with high der penetration", in "2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)", pp. 649–654 (2019).

[Shuman *et al.*(2013)] Shuman, D. I., S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains", IEEE Signal Processing Magazine **30**, 3, 83–98 (2013).

[Silva *et al.*(2019)] Silva, F. C., M. A. Ahmed, J. M. Martínez and Y.-C. Kim, "Design and implementation of a blockchain-based energy trading platform for electric vehicles in smart campus parking lots", Energies **12**, 24 (2019).

[Singhal *et al.*(2019)] Singhal, A., V. Ajjarapu, J. Fuller and J. Hansen, "Real-time local volt/var control under external disturbances with high pv penetration", IEEE Transactions on Smart Grid **10**, 4, 3849–3859 (2019).

[Sousa *et al.*(2011)] Sousa, A. A., G. L. Torres and C. A. Cañizares, "Robust optimal power flow solution using trust region and interior-point methods", IEEE Transactions on Power Systems **26**, 2, 487–499 (2011).

[Sousa *et al.*(2018)] Sousa, J., A. Bessani and M. Vukolic, "A byzantine fault-tolerant ordering service for the hyperledger fabric blockchain platform", in "2018 48th annual IEEE/IFIP international conference on dependable systems and networks (DSN)", pp. 51–58 (IEEE, 2018).

[Sperling(2012)] Sperling, J., "The tyranny of census geography: Small-area data and neighborhood statistics", Cityscape **14**, 2, 219–223 (2012).

[Stott and Alsac(1974)] Stott, B. and O. Alsac, "Fast decoupled load flow", IEEE Transactions on Power Apparatus and Systems **PAS-93**, 3, 859–869 (1974).

[Stott *et al.*(2009)] Stott, B., J. Jardim and O. Alsac, "Dc power flow revisited", IEEE Transactions on Power Systems **24**, 3, 1290–1300 (2009).

[Stouffer *et al.*(2015)] Stouffer, K., J. Falco and K. Scarfone, "Guide to industrial control systems (ICS) security", NIST special publication **800**, 82 (2015).

[Sukhwani *et al.*(2018)] Sukhwani, H., N. Wang, K. S. Trivedi and A. Rindos, "Performance modeling of hyperledger fabric (permissioned blockchain network)", in "2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)", pp. 1–8 (2018).

[Tajer *et al.*(2021)] Tajer, A., S. Perlaza and H. Poor, *Advanced Data Analytics for Power Systems* (Cambridge University Press, 2021), URL `https://books.google.com/books?id=0b4OEAAAQBAJ`.

[Tsitsvero *et al.*(2016)] Tsitsvero, M., S. Barbarossa and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling", IEEE Transactions on Signal Processing **64**, 18, 4845–4860 (2016).

[Turitsyn *et al.*(2011)] Turitsyn, K., P. Sulc, S. Backhaus and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators", Proceedings of the IEEE **99**, 6, 1063–1073 (2011).

[United States Government Accountability Office(2021)] United States Government Accountability Office, "ELECTRICITY GRID CYBERSECURITY: DOE Needs to Ensure Its Plans Fully Address Risks to Distribution Systems", Tech. rep., United States Government Accountability Office (2021).

[Vijayan *et al.*(2019)] Vijayan, V., A. Mohapatra and S. N. Singh, "Impact of modes of operation of smart inverters on volt-var optimization", in "2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)", pp. 1–5 (2019).

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

[Wang *et al.*(2018)] Wang, S., A. F. Taha and J. Wang, "Blockchain-assisted crowdsourced energy systems", in "2018 IEEE Power & Energy Society General Meeting (PESGM)", pp. 1–5 (IEEE, 2018).

[Wang *et al.*(2019)] Wang, S., A. F. Taha, J. Wang, K. Kvaternik and A. Hahn, "Energy crowdsourcing and peer-to-peer energy trading in blockchain-enabled smart grids", IEEE Transactions on Systems, Man, and Cybernetics: Systems **49**, 8, 1612–1623 (2019).

[Weckx *et al.*(2014)] Weckx, S., C. Gonzalez and J. Driesen, "Combined central and local active and reactive power control of pv inverters", IEEE Transactions on Sustainable Energy **5**, 3, 776–784 (2014).

[Wu *et al.*(2021)] Wu, Y., Y. Wu, J. M. Guerrero and J. C. Vasquez, "Digitalization and decentralization driving transactive energy internet: Key technologies and infrastructures", International Journal of Electrical Power and Energy Systems **126**, 106593 (2021).

[Xie *et al.*(2010)] Xie, L., Y. Mo and B. Sinopoli, "False data injection attacks in electricity markets", in "2010 First IEEE International Conference on Smart Grid Communications", pp. 226–231 (IEEE, 2010).

[Yaga *et al.*(2018)] Yaga, D., P. Mell, N. Roby and K. Scarfone, "Blockchain Technology Overview", Tech. rep., National Institute of Standards and Technology (2018).

[Yuan and Tong(2005)] Yuan, E. and J. Tong, "Attributed based access control (abac) for web services", in "IEEE International Conference on Web Services (ICWS'05)", p. 569 (2005).

[Yue Yuan *et al.*(2007)] Yue Yuan, Kejun Qian and Chengke Zhou, "The optimal location and penetration level of distributed generation", in "2007 42nd International Universities Power Engineering Conference", pp. 917–923 (2007).

[Zetter(2016)] Zetter, K., *Countdown to zero day: Stuxnet and the launch of the worlds first digital weapon* (Crown Publishers, 2016).

[Zhand *et al.*(2019)] Zhand, Y., A. Bernstein, A. Schmitt and R. Yang, "State estimation in low-observable distribution systems using matrix completion", in "52nd Hawaii International Conference on System Sciences", pp. 1–7 (2019).

[Zhang *et al.*(2018)] Zhang, Y., R. Madani and J. Lavaei, "Conic relaxations for power system state estimation with line measurements", IEEE Transactions on Control of Network Systems **5**, 3, 1193–1205 (2018).

[Zhang and Shi(2020)] Zhang, Y. and Q. Shi, "An intelligent transaction model for energy blockchain based on diversity of subjects", Alexandria Engineering Journal (2020).

[Zhao *et al.*(2018)] Zhao, S., B. Wang, Y. Li and Y. Li, "Integrated energy transaction mechanisms based on blockchain technology", Energies **11**, 9 (2018).

[Zhou *et al.*(2015)] Zhou, X., M. Farivar and L. Chen, "Pseudo-gradient based local voltage control in distribution networks", in "2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)", pp. 173–180 (2015).

[Zhou *et al.*(2016)] Zhou, X., J. Tian, L. Chen and E. Dall'Anese, "Local voltage control in distribution networks: A game-theoretic perspective", in "2016 North American Power Symposium (NAPS)", pp. 1–6 (2016).

[Zhu and Liu(2016)] Zhu, H. and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis", IEEE Transactions on Power Systems **31**, 5, 3794–3803 (2016).

[Zimmerman *et al.*(2011)] Zimmerman, R., C. Murillo-Sánchez and R. Thomas, "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education", IEEE Transactions on Power Systems **26**, 1, 12–19 (2011).

APPENDIX A

DEMAND RESPONSE RESOURCE MODELS

A.1 Electric Vehicles (EV)

An EV requires a certain amount of charge $\overline{u}$ over a period of length τ. Primary constraints are:

1. Rate of charge when grid-connected is $\dot{u}(t) = -p(t)$ for $0 \leq u(t) \leq \overline{u}$ and zero when the EV is full, with $p(t) < 0$ since charging is a load.

2. Charging is permitted only at a constant rate of $-\rho$, i.e., $p(t) \in \{0, -\rho\}$. Discharge to the grid is not permitted.

3. Charging has a deadline, i.e., $u(t_d) = u(t_a + \tau_c + \tau_s) = \overline{u}$.

Here t_a denotes the EV arrival time, t_d the EV departure time, τ_c the time needed for charging, and τ_s the leftover (slack) time. Considering a large number of loads, the constraint on charging rate can be relaxed as $-\rho \leq p(t) \leq 0$. In discrete time and vector form, assuming that the variable $u(t)$ is the energy normalized by the sampling period (i.e., the time that elapses between t and $t+1$), we can write:

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{\ell},\ u(t_d) = \overline{u},\ \ -\rho \leq p(t) \leq 0\ \forall\ t \in \mathcal{T} \tag{A.1.1}$$

In eq. (A.1.1), $\boldsymbol{A}$ computes the finite difference of the state of charge values in $\boldsymbol{u}$. Using $\boldsymbol{J}$ to denote an off-diagonal shift matrix, the following can be written:

$$\boldsymbol{A} = (\boldsymbol{J} - \mathbb{I}) \in \mathbb{R}^{|\mathcal{T}| \times |\mathcal{T}|},\ \boldsymbol{\ell} = [u(t_a), 0, \ldots, 0]^\top \in \mathbb{R}^{|\mathcal{T}| \times 1} \tag{A.1.2}$$

In this case $\boldsymbol{A}^\dagger = \boldsymbol{A}^{-1}$ is a triangular matrix of all negative ones. $\boldsymbol{A}^\dagger$ performs a cumulative sum of the entries of $-\boldsymbol{p}$, which is the inverse operation of taking the finite difference (similar to the integral is the inverse operation of the derivative). In general, we assume a prosumer is willing to pay more for having their EV charged earlier,

and expects a discounted electricity price if not charging at full power. This conceptualization of energy price is a function of time, with price decreasing monotonically with time as $\boldsymbol{c}_1$ and $c_{\mathrm{d}} < [\boldsymbol{c}_1]_t \ \forall t \in \mathcal{T}$:

$$C(\boldsymbol{p}) = \boldsymbol{c}_1^\top \boldsymbol{p} - c_{\mathrm{d}} \min(\overline{u} - u_{|\mathcal{T}|}, \rho \max(0, \tau_c + \tau_s - |\mathcal{T}|)) \tag{A.1.3a}$$

$$= \boldsymbol{c}_1^\top \boldsymbol{p} - c_{\mathrm{d}} \min(\overline{u} + \mathbf{1}^\top \boldsymbol{p}, \rho \max(0, \tau_c + \tau_s - |\mathcal{T}|)) \tag{A.1.3b}$$

A.2 Deferrable Appliances (DA)

These loads include appliances such as washers, dryers, and water pumps that can be programmed to start their cycle at different times of day. The feasible set of power demand is based on a load profile $h(t)$, a minimum activation time $0 \leq t_a \leq |\mathcal{T}| - 1$, and a slack $\tau_s \geq 0$:

$$p(t) = -h(t - t_a - \tau), \quad 0 \leq \tau \leq \tau_s \tag{A.2.1}$$

where the price function depends on the slack τ_s. Let us assume without loss of generality that $t_a = 0$ as the arrival time can be embedded into the the signal $h(t)$. Suppose also that the duration of $h(t)$ is d and that the slack and time are discrete. The signal $u(t)$ can indicate the time at which DA starts its cycle, which naturally means that $u(t) \in \{0, 1\}$ and the ℓ_1 norm $\|\boldsymbol{u}\|_1 = 1$. Considering a large enough population, this constraint can be relaxed to obtain an approximation of the feasible set as follows:

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{u}, \quad \|\boldsymbol{u}\|_1 = 1 \tag{A.2.2}$$

where $\boldsymbol{A} \in \mathbb{R}^{(|\mathcal{T}|+d)\times|\mathcal{T}|}$ equal to:

$$\boldsymbol{A}^\top = -\begin{bmatrix} h(0) & \dots & h(d) & 0 & \dots & 0 \\ 0 & h(0) & \dots & h(d) & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h(0) & \dots & h(d) \end{bmatrix} \tag{A.2.3}$$

The price a prosumer is willing to pay decreases as the delay increases. The price ranges from a maximum price the consumer is willing to pay to a minimum price that is the lowest possible energy cost. Suppose that we want the cost to grow linearly with time, and let c be the constant in the cost expression. Let $\tilde{\boldsymbol{c}}_1 = \tilde{c}_1 \cdot (|\mathcal{T}|-1, |\mathcal{T}|-2, \ldots, 1, 0)$. The cost can be obtained as follows:

$$C(\boldsymbol{p}) = \boldsymbol{c}_1^\top \boldsymbol{p} + c_0 = \tilde{c}_1 \sum_{t=0}^{|\mathcal{T}|-1} (|\mathcal{T}| - 1 - t) \cdot u(t) + \tilde{c}_0, \tag{A.2.4}$$

$$\Rightarrow \quad \boldsymbol{c}_1^\top = \tilde{\boldsymbol{c}}_1 \boldsymbol{A}^\dagger, \quad c_0 = \tilde{c}_0 \tag{A.2.5}$$

A.3 Thermostatically Controlled Loads (TCL)

These loads include space heaters, air conditioners, and water heaters. Similar to prior works [Callaway(2009)], we assume that the temperature dynamics of a heat pump based TCL can be modeled as a first-order differential equation:

$$C\dot{\theta}(t) = (\theta_{\mathrm{o}}(t) - \theta(t))R^{-1} + p(t)\eta + \varepsilon(t)R^{-1}, \quad p(t) \in \{0, -\rho\} \tag{A.3.1}$$

with R being thermal resistance, C thermal capacitance, $\theta(t)$ the inside temperature, $\theta_{\mathrm{o}}(t)$ the outdoor temperature, η the efficiency of the heat pump ($\eta > 0$ for cooling and $\eta < 0$ for heating), ρ continuous electrical power rating [1], and $\varepsilon(t)$ denoting a random perturbation of temperature by external factors such as opening of windows/doors or operation of stoves. We denote the thermostats reference temperature as θ_{r}. Let:

$$\begin{aligned} u(t) &\triangleq \frac{\theta(t) - \theta_r(t)}{R\eta} \\ \tilde{\ell}(t) &\triangleq \frac{\theta_r(t) - \theta_o(t)}{R\eta} + \frac{\varepsilon(t)}{R\eta} + \frac{C}{\eta}\dot{\theta}_r(t), \quad \tau_h \triangleq CR \end{aligned} \tag{A.3.2}$$

[1] Water heaters can be described using the same principles, with an additional energy loss component describing the hot water being replaced by cold water. However, in this paper, we will focus on heat pump based TCL because they are more dependent on external temperatures than water boilers.

To express the inter-temporal constraints as well as the comfort zone of the user, we can rearrange and relax the set of constraints in eq. (A.3.1) as follows:

$$p(t) = \frac{C}{\eta}(\dot{\theta}(t) - \dot{\theta}_r(t)) + \frac{\theta(t) - \theta_r(t)}{R\eta} + \frac{\theta_r(t) - \theta_o(t)}{R\eta} + \frac{\varepsilon(t)}{R\eta} + \frac{C}{\eta}\dot{\theta}_r(t) \tag{A.3.3a}$$

$$= \tau_h \dot{u}(t) + u(t) + \tilde{\ell}(t),\ \underline{u} \leq u(t) \leq \overline{u},\ \ -\rho \leq p(t) \leq 0 \tag{A.3.3b}$$

where we have relaxed the integer constraint $p(t) \in \{0, -\rho\}$.

It is notable that the $\tilde{\ell}(t)$ term in this affine relationship is random, since the outdoor temperature is random and so is the perturbation of the indoor temperature from sources other than the heat pump. In discrete time, we can average the behavior and approximate the relationship as follows:

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{\ell}, \quad \underline{u} \leq u(t) \leq \overline{u}, \quad -\rho \leq p(t) \leq 0 \tag{A.3.4}$$

Using eqs. (A.3.2) and (A.3.3a) the following can be written:

$$\begin{aligned} &\boldsymbol{A} = \tau_h(\mathbb{I} - \boldsymbol{J}) + \mathbb{I},\ \boldsymbol{A}^{\dagger} = \boldsymbol{A}^{-1} = ((\tau_h + 1)\mathbb{I} - \tau_h \boldsymbol{J})^{-1} \\ &\boldsymbol{\ell} = \tilde{\boldsymbol{\ell}} - \tau_h[u(0), \boldsymbol{0}^{\top}]^{\top} \in \mathbb{R}^{|\mathcal{T}| \times 1} \end{aligned} \tag{A.3.5}$$

and $\tilde{\boldsymbol{\ell}} = (\tilde{\ell}(1), \ldots, \tilde{\ell}(|\mathcal{T}|))$ with $\tilde{\ell}(t)$ being defined in eq. (A.3.2). The price demand function should represent the prosumer's willingness to deviate from reference temperature. The prosumer is willing to pay less for larger temperature deviations than expected. We represent this cost as quadratic with the value of $\boldsymbol{u}$, i.e., proportional to $\|\boldsymbol{u}\|^2$. This means that the demand function is $\tilde{c}_0 - \tilde{c}_2\|\boldsymbol{u}\|^2$ and:

$$C(\boldsymbol{p}) = \boldsymbol{p}^{\top}\boldsymbol{C}_2\boldsymbol{p} + \boldsymbol{p}^{\top}\boldsymbol{c}_1 + c_0 = \tilde{c}_0 - \tilde{c}_2\left\|\boldsymbol{A}^{\dagger}(\boldsymbol{p} - \boldsymbol{\ell})\right\|^2 \tag{A.3.6}$$

where:

$$\rightarrow \quad \boldsymbol{C}_2 = -\tilde{c}_2\boldsymbol{A}^{-2},\ \boldsymbol{c}_1 = 2\tilde{c}_2\boldsymbol{\ell}\boldsymbol{A}^{-2},\ \ c_0 = \tilde{c}_0 - \tilde{c}_2\boldsymbol{\ell}^{\top}\boldsymbol{A}^{-2}\boldsymbol{\ell} \tag{A.3.7}$$

A.4 Energy Storage Systems

Typically, battery rate of charge or discharge is constrained to be a constant value, meaning $p(t) \in \{-\rho, 0, \rho\}$. If we relax this non-convex constraint as done before, the load constraints are analogous to that of an EV, except that the unit can discharge:

$$\boldsymbol{p} = (\mathbb{I} - \boldsymbol{J})\boldsymbol{u} - [u(0), \boldsymbol{0}], \quad |p(t)| \leq \rho, \quad 0 \leq u(t) \leq \overline{u} \tag{A.4.1}$$

where $\boldsymbol{J}$ is the shift matrix, shifting to the right each of the entries of $\boldsymbol{u}$. Here it is natural to assume that the price for discharging is higher than the price of charging, but it is also possible to express a cost that depends on the battery state. In addition, if storage is charged by the random injection of, for instance, solar PV, the forecast can be incorporated into the vector $\boldsymbol{\ell}$ in the model. We ignore other complexities here for simplicity. Considering $(a)_+ = \max(0, a)$, we can express the cost as:

$$C(\boldsymbol{p}) = (\boldsymbol{p})_+^\top(\boldsymbol{c}_1^+) + (-\boldsymbol{p})_+^\top(\boldsymbol{c}_1^-) \tag{A.4.2}$$

where $\boldsymbol{c}_1^+$ and $\boldsymbol{c}_1^-$ are non-negative cost vectors.

A.5 Renewables

The power injection from wind or solar PV has no marginal cost, therefore the only meaningful way for renewables to participate is posting a forecast of future production $\boldsymbol{p}$ with zero cost, i.e., $C(\boldsymbol{p}) = 0$.

A.6 Supply from the Transmission Grid

We assume that the transmission grid appears in the system as the slack bus and has a certain cost function for selling and a certain cost function for buying power below and above a schedule $\boldsymbol{p}_s$ that was cleared in previous wholesale market stages.

Therefore, at the substation bus we have a single supplier with supply function:

$$C(\boldsymbol{p}) = (\boldsymbol{p} - \boldsymbol{p}_s)_+^\top(\boldsymbol{c}_1^+) + (-\boldsymbol{p} + \boldsymbol{p}_s)_+^\top(\boldsymbol{c}_1^-) \tag{A.6.1}$$

For simplicity we assume that any deviation is feasible, so that the dispatch is always feasible, and the slack bus compensates for any shortfall or surplus of power subject to physical constraints in Appendix B.

APPENDIX B

ELECTRIC GRID CONSTRAINTS

B.1 Electric Grid Model

A radial electrical distribution system can be represented by a set of buses $\mathcal{B}$, edges $\mathcal{E}_e$, and a root node (commonly the substation node or slack bus) where each edge $\ell = (b, b') \in \mathcal{E}_e$ and $b, b' \in \mathcal{B}$. The *from* and *to* functions are defined as $f(\ell) = b$ and $t(\ell) = b'$ to return the source node and incident node for an edge, respectively. The inverse function $t^{-1}(b)$ gives back the edge pointing to bus b (for a radial graph the *to* bus b for each edge is unique) and $f^{-1}(b)$ returns all the edges originating at bus b. Dropping the time index for brevity, the power flow equations at each of the branches $\ell \in \mathcal{E}_e$ are given by [Baran and Wu(1989b)]:

$$p_{t(\ell)} = P_\ell - \sum_{\ell' \in f^{-1}(t(\ell))} P_{\ell'} - \mathrm{Re}(z_\ell) c_\ell^2 \tag{B.1.1a}$$

$$q_{t(\ell)} = Q_\ell - \sum_{\ell' \in f^{-1}(t(\ell))} Q_{\ell'} - \mathrm{Im}(z_\ell) c_\ell^2 \tag{B.1.1b}$$

$$v_{f(\ell)}^2 = v_{t(\ell)}^2 + 2(\mathrm{Re}(z_\ell) P_\ell + \mathrm{Im}(z_\ell) Q_\ell) - |z_\ell|^2 c_\ell^2 \tag{B.1.1c}$$

$$v_{f(\ell)}^2 = \frac{P_\ell^2 + Q_\ell^2}{c_\ell^2}. \tag{B.1.1d}$$

All these equations are linear in the bus and branch quantities $(p_b(t), q_b(t), v_b^2(t))$ and $(P_\ell(t), Q_\ell(t), c_\ell^2(t))$ except eq. (B.1.1d). However, including an auxiliary variable $x'_\ell = \frac{P_\ell^2 + Q_\ell^2}{c_\ell^2}$ can simplify the description of the physical constraints. By expressing $\boldsymbol{x}(t)$ as the auxiliary variables, then without loss of generality eqs. (B.1.1a) to (B.1.1d) can be written in the following linear form:

$$\boldsymbol{H}\boldsymbol{x} = \boldsymbol{H}_{\mathcal{A}}\boldsymbol{x}_{\mathcal{A}} + \boldsymbol{H}_u\boldsymbol{x}_u = \boldsymbol{0}, \tag{B.1.2}$$

Though constraints in eq. (4.3.2) are non-linear in general, here the constraints are relaxed by ignoring the non-linear relationships among the auxiliary variables and the remaining entries of the vector $\boldsymbol{x}$. The measurements are represented by $\boldsymbol{z} = \boldsymbol{x}_{\mathcal{A}} + \boldsymbol{\epsilon}$ and the physics of the system implies that $\boldsymbol{x}_{\mathcal{A}}$ satisfies eq. (B.1.2).

From the vantage point of each aggregator region n, only a subset of the variables $\boldsymbol{x}_{\mathcal{A}}^{(n)}$ are measured, meaning that $\boldsymbol{z}^{(n)} = \boldsymbol{x}_{\mathcal{A}}^{(n)} + \boldsymbol{\epsilon}^{(n)}$. Also, not all constraints in eq. (B.1.2) include the variables $\boldsymbol{x}^{(n)}$. Hence, the equations that involve buses/lines in region i can be isolated and written as:

$$\boldsymbol{h}^{(n)}(\boldsymbol{x}^{(n)}) = \boldsymbol{H}^{(n)}\boldsymbol{x}^{(n)} = \boldsymbol{0}. \tag{B.1.3}$$

These equations are used to verify that measurements and injections values are consistent with the laws of physics. The available measurements in each zone and the neighboring zones are used to interpolate the unknown variables as discussed in Section 4.3.3.1.

APPENDIX C

ERROR ESTIMATION IN THE GSP APPROXIMATION

C.1 Proof of Proposition 1

By using $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC})\otimes(\boldsymbol{BD})$:

$$(\boldsymbol{U}_k^\star\otimes\boldsymbol{U}_k)^{\mathsf{H}}(\boldsymbol{U}_k^\star\otimes\boldsymbol{U}_k^\perp)=(\boldsymbol{U}_k^\top\boldsymbol{U}_k^\star)\otimes(\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k^\perp)=\boldsymbol{0}$$

$$(\boldsymbol{U}_k^\star\otimes\boldsymbol{U}_k)^{\mathsf{H}}((\boldsymbol{U}_k^\perp)^\star\otimes\boldsymbol{U}_k)=((\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k^\perp)^\star)\otimes(\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k)=\boldsymbol{0}$$

$$(\boldsymbol{U}_k^\star\otimes\boldsymbol{U}_k)^{\mathsf{H}}((\boldsymbol{U}_k^\perp)^\star\otimes\boldsymbol{U}_k^\perp)=((\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k^\perp)^\star)\otimes(\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k^\perp)=\boldsymbol{0}$$

Thus, from (3.2.4b), $(\boldsymbol{U}_k^\star\otimes\boldsymbol{U}_k)^{\mathsf{H}}\boldsymbol{\epsilon}_k=\boldsymbol{0}$.

To prove the second statement (3.2.4a) is substituted in (3.1.12):

$$\boldsymbol{v}\boldsymbol{v}^{\mathsf{H}}=\big(\boldsymbol{U}_k\tilde{\boldsymbol{v}}_k+\boldsymbol{U}_k^\perp\tilde{\boldsymbol{\varepsilon}}_k\big)\big(\boldsymbol{U}_k\tilde{\boldsymbol{v}}_k+\boldsymbol{U}_k^\perp\tilde{\boldsymbol{\varepsilon}}_k\big)^{\mathsf{H}} \tag{C.1.1}$$

Using (3.2.4b) and (3.2.2), and considering, $\boldsymbol{A}=\boldsymbol{U}_k^\perp\tilde{\boldsymbol{\varepsilon}}_k\tilde{\boldsymbol{v}}_k^{\mathsf{H}}\boldsymbol{U}_k^{\mathsf{H}}$, $\boldsymbol{B}=\boldsymbol{U}_k\tilde{\boldsymbol{v}}_k\tilde{\boldsymbol{\varepsilon}}_k^{\mathsf{H}}(\boldsymbol{U}_k^\perp)^{\mathsf{H}}$ and $\boldsymbol{C}=\boldsymbol{U}_k^\perp\tilde{\boldsymbol{\varepsilon}}_k\tilde{\boldsymbol{\varepsilon}}_k^{\mathsf{H}}(\boldsymbol{U}_k^\perp)^{\mathsf{H}}$, the following can be written:

$$\begin{aligned}\|\epsilon_k\|^2=\|\mathrm{vec}(\boldsymbol{A})\|^2+\|\mathrm{vec}(\boldsymbol{B})\|^2+\|\mathrm{vec}(\boldsymbol{C})\|^2+\mathrm{vec}(\boldsymbol{A})^{\mathsf{H}}(\mathrm{vec}\,\boldsymbol{B}+\mathrm{vec}\,\boldsymbol{C})\\+\mathrm{vec}(\boldsymbol{B})^{\mathsf{H}}(\mathrm{vec}\,\boldsymbol{A}+\mathrm{vec}\,\boldsymbol{C})+\mathrm{vec}(\boldsymbol{C})^{\mathsf{H}}(\mathrm{vec}\,\boldsymbol{A}+\mathrm{vec}\,\boldsymbol{B})\end{aligned} \tag{C.1.2}$$

However, for any matrix $\boldsymbol{P}$, $\|\mathrm{vec}(\boldsymbol{P})\|^2=\mathrm{Tr}\big\{\boldsymbol{P}\boldsymbol{P}^{\mathsf{H}}\big\}$. Thus,

$$\|\mathrm{vec}(\boldsymbol{A})\|^2=\mathrm{Tr}\big\{\tilde{\boldsymbol{e}}_k\tilde{\boldsymbol{\varepsilon}}_k{}^{\mathsf{H}}\big\}\Big\|\tilde{\boldsymbol{v}_k^{\mathsf{H}}}\Big\|^2=\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\|\tilde{\boldsymbol{v}}_k\|^2 \tag{C.1.3a}$$

$$\|\mathrm{vec}(\boldsymbol{B})\|^2=\mathrm{Tr}\big\{\tilde{\boldsymbol{v}}_k\tilde{\boldsymbol{v}}_k{}^{\mathsf{H}}\big\}\Big\|\tilde{\boldsymbol{\varepsilon}_k^{\mathsf{H}}}\Big\|^2=\|\tilde{\boldsymbol{v}}_k\|^2\|\tilde{\boldsymbol{\varepsilon}}_k\|^2 \tag{C.1.3b}$$

$$\|\mathrm{vec}(\boldsymbol{C})\|^2=\mathrm{Tr}\big\{\tilde{\boldsymbol{\varepsilon}}_k\tilde{\boldsymbol{\varepsilon}}_k{}^{\mathsf{H}}\big\}\Big\|\tilde{\boldsymbol{\varepsilon}_k^{\mathsf{H}}}\Big\|^2=\|\tilde{\boldsymbol{\varepsilon}}_k\|^4 \tag{C.1.3c}$$

Since $(\boldsymbol{U}_k^\perp)^{\mathsf{H}}\boldsymbol{U}_k=\boldsymbol{0}$, the cross product terms are zero; which proves (3.2.6).

Finally, (3.2.7) follows from applying Isserlis' theorem to evaluate the fourth order moments of complex circularly symmetric Gaussian random variables in (3.2.6). Assuming that random variables $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{\varepsilon}}_k$ are uncorrelated we have:

$$\mathbb{E}\big[\|\boldsymbol{\epsilon}_k\|^2\big]=\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^4\big]+2\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big]\mathbb{E}\big[\|\tilde{\boldsymbol{v}}_k\|^2\big] \tag{C.1.4}$$

$$\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^4\big]=\sum_i\sum_j\mathbb{E}\Big[[\tilde{\boldsymbol{\varepsilon}}_k]_i\,[\tilde{\boldsymbol{\varepsilon}}_k]_i{}^\star[\tilde{\boldsymbol{\varepsilon}}_k]_j[\tilde{\boldsymbol{\varepsilon}}_k]_j{}^\star\Big] \tag{C.1.5}$$

In order to evaluate each term in (C.1.5), recall the Isserlis or Wick's formula to compute fourth order moments for circularly symmetric random variables [Baudin(2014)]

$$\mathbb{E}\Big[[\tilde{\boldsymbol{\varepsilon}}_k]_i\,[\tilde{\boldsymbol{\varepsilon}}_k]_i^{\star}[\tilde{\boldsymbol{\varepsilon}}_k]_j[\tilde{\boldsymbol{\varepsilon}}_k]_j^{\star}\Big] = \mathbb{E}\big[|[\tilde{\boldsymbol{\varepsilon}}_k]_i|^2\big]\mathbb{E}\Big[\Big|[\tilde{\boldsymbol{\varepsilon}}_k]_j\Big|^2\Big] + \mathbb{E}\Big[[\tilde{\boldsymbol{\varepsilon}}_k]_i[\tilde{\boldsymbol{\varepsilon}}_k]_j^{\star}\Big]\mathbb{E}\Big[[\tilde{\boldsymbol{\varepsilon}}_k]_j\,[\tilde{\boldsymbol{\varepsilon}}_k]_i^{\star}\Big] \tag{C.1.6}$$

Assuming that entries of $\tilde{\boldsymbol{\varepsilon}}_k$ are uncorrelated random variables (i.e. diagonal covariance matrix):

$$\mathbb{E}\Big[[\tilde{\boldsymbol{\varepsilon}}_k]_i\,[\tilde{\boldsymbol{\varepsilon}}_k]_i^{\star}[\tilde{\boldsymbol{\varepsilon}}_k]_j[\tilde{\boldsymbol{\varepsilon}}_k]_j^{\star}\Big] = \begin{cases} 2\left(\mathbb{E}\big[|[\tilde{\boldsymbol{\varepsilon}}_k]_i|^2\big]\right)^2, \ i=j \\ \mathbb{E}\big[|[\tilde{\boldsymbol{\varepsilon}}_k]_i|^2\big]\mathbb{E}\Big[\Big|[\tilde{\boldsymbol{\varepsilon}}_k]_j\Big|^2\Big], \ i\neq j \end{cases} \tag{C.1.7}$$

Substitution of (C.1.7) in (C.1.5) results in (3.2.7):

$$\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^4\big] = \sum_i \left(\mathbb{E}\big[|[\tilde{\boldsymbol{\varepsilon}}_k]_i|^2\big]\right)^2 + \left(\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big]\right)^2 = \left\|\mathrm{diag}\left(\mathbb{E}\big[\tilde{\boldsymbol{\varepsilon}}_k\tilde{\boldsymbol{\varepsilon}}_k^H\big]\right)\right\|_2^2 + \left(\mathbb{E}\big[\|\tilde{\boldsymbol{\varepsilon}}_k\|^2\big]\right)^2 \tag{C.1.8a}$$

C.2 Alternate OPF Formulation

Using the following matrices:

$$\widetilde{\boldsymbol{U}}_{k,i}^{(r)} = \left(\mathrm{vec}\left(\left(\boldsymbol{U}_{k,i}^{\mathsf{H}}\right)^{(c)}\boldsymbol{U}_{k,i}^{(r)}\right)\right)^{\top}; \widetilde{\boldsymbol{Y}_{\boldsymbol{U}}}_i^{(r)} \qquad = \left(\mathrm{vec}\left(\left(\boldsymbol{U}_{k,i}^{\mathsf{H}}\right)^{(c)}(\boldsymbol{Y}\boldsymbol{U}_k)_i^{(r)}\right)\right)^{\top},$$

the OPF formulation in (3.3.2) becomes:

$$\min_{\boldsymbol{p}_g,\boldsymbol{q}_g,\widetilde{\boldsymbol{W}},\boldsymbol{\epsilon}_p,\boldsymbol{\epsilon}_q} \sum_{g\in\mathcal{B}_G} \boldsymbol{C}_g(\boldsymbol{p}_g) + w_1\|\boldsymbol{\epsilon}_p\|_2^2 + w_2\|\boldsymbol{\epsilon}_q\|_2^2 \tag{C.2.1a}$$

$$\text{s. t} \quad (3.3.2\text{b}), (3.3.2\text{c}), (3.3.2\text{d}) \tag{C.2.1b}$$

$$(\boldsymbol{p}_g - \boldsymbol{p}_d) - j(\boldsymbol{q}_g - \boldsymbol{q}_d) - \widetilde{\boldsymbol{Y}_U}\,\mathrm{vec}\left(\widetilde{\boldsymbol{W}}^{\top}\right) = \boldsymbol{\epsilon}_p + j\boldsymbol{\epsilon}_q \tag{C.2.1c}$$

$$\sum(\boldsymbol{p}_g - \boldsymbol{p}_d) - \sum j(\boldsymbol{q}_g - \boldsymbol{q}_d) = \sum \widetilde{\boldsymbol{Y}_U}\,\mathrm{vec}\left(\widetilde{\boldsymbol{W}}^{\top}\right) \tag{C.2.1d}$$

$$v_{\min}^2 \leq \Re\Big\{\widetilde{\boldsymbol{U}}_k\,\mathrm{vec}\left(\widetilde{\boldsymbol{W}}^{\top}\right)\Big\} \leq v_{\max}^2 \tag{C.2.1e}$$

$$\boldsymbol{p}_{\min}^{\text{flow}} \leq \Re\Big\{\left(\boldsymbol{U}_{k,i}^{(r)}\widetilde{\boldsymbol{W}}\left(\boldsymbol{U}_{k,j}^{(r)}\right)^{\mathsf{H}} - \boldsymbol{U}_{k,i}^{(r)}\widetilde{\boldsymbol{W}}\left(\boldsymbol{U}_{k,i}^{(r)}\right)^{\mathsf{H}}\right)y_{ij}^{\star}\Big\} \leq \boldsymbol{p}_{\max}^{\text{flow}}, \quad \forall(i,j)\in\mathcal{E}_e \tag{C.2.1f}$$

APPENDIX D

LIST OF PUBLICATIONS SUPPORTED BY THE DISSERTATION

D.1 Accepted

- **Shammya Saha**, N. Ravi, K. Hreinsson, J. Baek, A. Scaglione and N. G. Johnson , "A Secure Distributed Ledger for Transactive Energy: The Electron Volt Exchange (EVE) Blockchain", in *Applied Energy*, Jan. 2021.

- **Shammya Saha**, D. Arnold, A. Scaglione, E. Schweitzer , N. G. Johnson , C. Roberts and S. Peisert, "Lyapunov Stability of Smart Inverters Policies based on Linearized DistFlow Approximation", in *IET Renewable Power Generation*, Jan. 2021.

- E. Schweitzer, **Shammya Saha**, A. Scaglione, N. G. Johnson and D. Arnold, "Lossy DistFlow Formulation for Single and Multiphase Radial Feeders", in *IEEE Transactions on Power Systems*, Nov. 2019.

- C. Roberts, S. Ngo, A. Milesi, S. Peisert, D. Arnold, **Shammya Saha**, A. Scaglione, N. G. Johnson, A. Kocheturov and D. Fradkin , "Deep Reinforcement Learning for DER Cyber-Attack Mitigation", in *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 11-13 Nov. 2020, Tempe, USA.

- **Shammya Saha**, E. Schweitzer, A. Scaglione and N. G. Johnson, "A Framework for Generating Synthetic Distribution Feeders using OpenStreetMap", in *51st North American Power Symposium*, 13-15 Oct 2019, Wichita, Kansas, USA.

- **Shammya Saha**, C. Gorog, A. Moser, A. Scaglione and N. G. Johnson , "Integrating Hardware Security into a Blockchain-Based Transactive Energy Platform", in *North American Power Symposium*, April, 2021, Tempe, USA.

D.2 Under Review

- D. Arnold, C. Roberts, S. Ngo, **Shammya Saha**, A. Scaglione, N. G. Johnson, S. Peisert, D. Penny, "Adaptive Control of Distributed Energy Resources for Distribution Grid Voltage Stability", IEEE Transactions on Power Systems

- **Shammya Saha**, A. Scaglione, R. Ramakrishna and N. G. Johnson, "Convex Relaxation For State Estimation and Optimal Power Flow Using Graph Signal Processing", IEEE Transactions on Power Systems