

© 2017 Hao Jan Liu

DECENTRALIZED OPTIMIZATION APPROACH FOR POWER DISTRIBUTION NETWORK AND
MICROGRID CONTROLS

BY

HAO JAN LIU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Assistant Professor Hao Zhu, Chair
Associate Professor Alejandro Domínguez-García
Professor Peter Sauer
Professor Rayadurgam Srikant

ABSTRACT

The smart grid vision has led to the development of advanced control and management frameworks using distributed generation (DG) and storage resources, commonly referred to together as distributed energy resources (DERs). Albeit environment-friendly, these DERs in distribution networks including microgrids (MGs) could greatly challenge the operational goal of maintaining adequate power system reliability standards because of their high intermittency, uncertainty, and lack of physical inertia. Meanwhile, these networks are inherently unbalanced and lack high-quality communications to a centralized entity as compared to the bulk transmission grid. Both aspects contribute to the challenge of designing voltage and frequency control frameworks therein. To tackle these problems, we propose decentralized control strategies, which account for cyber-physical network interactions automatically and dynamically while being either cognizant of various communication scenarios or resilient to malicious cyber intrusions.

By treating the transmission grid as an infinity bus, voltage stability is the main concern in distribution networks where more DERs are being installed in the near future. Thanks to advances in power electronics, DERs can also be excellent sources of *reactive power* (VAR), a quantity that is known to have a significant impact on the network voltage level. Accordingly, we first formulate the local VAR-based voltage control design by minimizing a weighted quadratic voltage mismatch error objective using gradient-projection (GP) updates. The step-size design under both static and dynamic settings is further analyzed for practical implementation purposes. Nonetheless, such local design suffers degraded performance due to lack of information exchanges, especially under limited VAR resources. To address this issue, we develop the distributed voltage control (DVC) design based on the alternating direction method of multipliers (ADMM) algorithm. The DVC design has simple node-to-node communication architecture while seamlessly adapting to dynamically varying system operating conditions and being robust against random communication link

failures. To further reduce communication complexity and enhance robustness to imperfect communications, especially under the worst-case scenarios of a total communication outage, we integrate both local and distributed control designs to a hybrid voltage control (HVC) scheme that can achieve the dual objectives in terms of flexible adaptivity to variable rate of communications and global optimality of voltage regulation performance. Such an innovative design aims to unify the separated framework of either local or distributed control design. Numerical tests using realistic feeders and real time-series data have been demonstrated for the voltage control designs.

The aforementioned decentralized voltage control designs can improve the power system stability while distribution feeders are interconnecting to the bulk transmission grids. With a high penetration of DERs in the networks, it is possible to build a discrete energy system, namely, a microgrid (MG), that is capable of operating in parallel with, or independently from, the transmission grids. Henceforth, MGs are likely to emerge as a means to advance power and cyber physical resiliency in future grid systems. As MGs may operate independently, these mostly power electronics-interfaced DERs exhibiting low-inertia characteristic have raised significant concern over the frequency stability issues. To tackle this problem, we introduce the concept of virtual inertia of DERs and cast the secondary frequency control design for isolated MGs as a consensus optimization problem. We solve it distributively by adopting the partial primal-dual (PPD) algorithm. Interestingly, parts of our specially designed control algorithm turn out to mimic the dynamics of network power flow and virtual synchronous generator-based inverter. Thus, such dynamics is seamlessly governed by the physical system itself. Given a proper control parameter choice, the convergence of the consensus is guaranteed without assuming the time-scale separation of the hierarchical control design methodologies. By extending this work to a practical industrial MG network that follows the IEC 61850 communication protocol, similar frequency regulation objective is introduced and solved by a decentralized ADMM-based algorithm. The countermeasures for malicious attacks on the communication network for both PPD- and ADMM-based control designs are also investigated. Specifically, we analyze two types of malicious attacks on the communication network, namely, the link and node attacks. Meanwhile, anomaly detection and localization strategies are developed based on the metrics of optimization-related variables. We showcase the microgrid frequency regulation operation to demonstrate the effectiveness of the proposed frequency control designs under a real-time simulation environment.

To my parents, Te-Wu and Feng-Chu.

ACKNOWLEDGMENTS

First, and foremost, I would like to express my sincere gratitude, for their helpful advice and guidance, to my advisors and mentors: Prof. Thomas Overbye, Prof. Hao Zhu, and Alfonso Valdes. I am lucky to have their support through this invaluable experience at UIUC. Thank you, Prof. Overbye, for your advising role during my master's. You are an exemplary role model, which I hope one day to become. Alfonso, I am very grateful for your mentorship regarding industry collaborations and your kind encouragements over the last few years. Especially, I would like to thank my Ph.D. advisor, Prof. Hao Zhu. Over the past four years, Hao has been the driving force behind my academic endeavors. Through many revisions of papers and debates over ideas, I have learned so much from her and know that the impact she has had on my life will continue to influence me well into the future.

Also many thanks to my doctoral committee: Professors Peter Sauer, Alejandro Domínguez-García, and Rayadurgam Srikant. Prof. Sauer's deep understanding of power systems has amazed me and aided my research greatly. He always has a simple explanation to the most complex problem, and his kindness and valuable suggestions have encouraged me throughout my graduate study. I am also very appreciative of the opportunity to learn from Professors Alejandro Domínguez-García and Rayadurgam Srikant. Their knowledge of microgrids and convex optimization has facilitated the development of the thesis.

I am fortunate to have many good friends and colleagues while in graduate school. Special thanks go to Jerry, Candy, Yeo-Jung, and Joseph for their support and putting up with me without much complaint. Our traveling around the world has colored my life profoundly and given me an opportunity to escape the cornfields. These memories are priceless. During my time at UIUC, I also have had the privilege of coaching figure skating in the University Ice Rink. I would like to thank all of my students including Marina, Juliann, Alice, Cadi, Liuqing, Lelani, Kate, Susan,

Sandy, Grace B., Grace G., Mika, and Binyu, for their trust and dedication to the sport I have loved since my childhood. These opportunities have taught me how to be a better teacher and given me excellent insight into some of the life challenges faced outside of academia.

Finally, I would like to thank my parents, Te-Wu and Feng-Chu, for their unconditional love. They have brought me to the USA for a better education and provided me with continuous support. It has been a fantastic journey from being a professional ice skater to having a Ph.D. in electrical engineering. To my family in both Taiwan and the USA, heartfelt thanks for always believing in me. I am incredibly fortunate to have an amazing life story because of you.

TABLE OF CONTENTS

LIST OF FIGURES	ix
CHAPTER 1 INTRODUCTION	1
1.1 Motivation and Context	1
1.2 Organization of the Thesis	2
CHAPTER 2 SYSTEM MODELING FOR VOLTAGE CONTROL IN DISTRIBUTION NETWORKS	6
2.1 Modeling of Distribution Networks	6
2.2 Modeling of Inverters	12
CHAPTER 3 COMMUNICATION-FREE LOCAL VOLTAGE CONTROL	14
3.1 Local Voltage Control Problem	16
3.2 Asynchronous Local Voltage Control	20
3.3 Dynamic Local Voltage Control	22
3.4 Numerical Tests	30
3.5 Summary	40
CHAPTER 4 ONLINE DISTRIBUTED VOLTAGE CONTROL	41
4.1 Distributed Voltage Control Problem	43
4.2 ADMM-based Distributed Solver	45
4.3 Online Voltage Control Design	48
4.4 Robustness to Communication Link Failures	50
4.5 Numerical Tests	52
4.6 Summary	58
CHAPTER 5 COMMUNICATION-COGNIZANT HYBRID VOLTAGE CONTROL	60
5.1 Hybrid Voltage Control Problem	61
5.2 Communication-Cognizant Hybrid Voltage Control Design	63
5.3 Numerical Tests	75
5.4 Summary	81

CHAPTER 6	DISTRIBUTED FREQUENCY CONTROL FOR ISOLATED MICROGRIDS	83
6.1	Modeling of Microgrids	86
6.2	Distributed Secondary Control Design	91
6.3	Attack Models and Countermeasures	98
6.4	Numerical Tests	104
6.5	Summary	114
CHAPTER 7	TESTBED-BASED DISTRIBUTED CONTROL IMPLEMENTATIONS FOR INDUSTRIAL MICROGRIDS	115
7.1	Industrial Microgrid	117
7.2	Problem Statement	120
7.3	ADMM-based Decentralized Solver	121
7.4	Detection and Localization Strategies	123
7.5	Numerical Tests	127
7.6	Summary	131
CHAPTER 8	CONCLUDING REMARKS	133
8.1	Thesis Summary and Contribution	133
8.2	Conclusions	135
APPENDIX A	MULTI-PHASE DVC IMPLEMENTATIONS	137
APPENDIX B	LINE FLOW DYNAMICS	143
REFERENCES	144

LIST OF FIGURES

2.1	A radial multi-phase distribution feeder with bus and line associated variables. . . .	7
2.2	Mode 1 contributes only positive active power output while modes 4 and 5 generate negative and positive VAR representing inductive or capacitive characteristics, respectively. In addition, modes 2 and 3 consist of mixed active and VAR outputs. All are bounded by the physical inverter rating.	12
3.1	Iterative voltage mismatch error performance for the asynchronous decentralized voltage control scheme under various choices of duty cycle η and step-size ϵ	31
3.2	Voltage mismatch error versus the total number of updates across the network for the asynchronous voltage control scheme under various choices of duty cycle η and step-size ϵ	32
3.3	An instance of the nominal voltage series $\{\bar{v}_k\}$ at selected buses under the AR(1) model settings.	33
3.4	Iterative voltage mismatch error performance averaged over 30 random realizations under different ϵ values.	34
3.5	Iterative tracking error averaged over 30 random realizations under different ϵ values.	34
3.6	Iterative voltage mismatch error performance averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α	36
3.7	Iterative tracking error averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α	36
3.8	Iterative voltage mismatch error performance averaged over 30 random realizations under various σ values.	37
3.9	Iterative tracking error averaged over 30 random realizations under various σ values.	37
3.10	Iterative voltage mismatch error performance on the IEEE 123-bus system under both the asynchronous updates and the dynamic network operating conditions with different η values.	39
3.11	Voltage mismatch error versus the total number of updates across the IEEE 123-bus system for the asynchronous voltage control scheme under dynamic operating conditions with various choices of duty cycle η	39

4.1	Online implementation of the proposed DVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of ADMM iterations and control updates, while the two vertical arrows correspond to the communications between bus j and its neighborings buses.	49
4.2	Log-scale voltage mismatch squared error versus iteration index for various c_j values for the static test settings.	53
4.3	Log-scale voltage mismatch squared error comparison between the proposed DVC scheme and the algorithm of [15] with different ρ values.	54
4.4	Voltage mismatch squared error versus iteration index under various rates of link failure.	55
4.5	Sample daily load and solar PV generation profiles of a US home.	57
4.6	Daily voltage mismatch squared error at phase a of the 123-bus feeder.	59
5.1	Online implementation of the proposed HVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of PPD iterations and control updates, while the two vertical arrows correspond to the communications exchange among bus j and its neighboring buses.	73
5.2	Voltage mismatch squared error versus various values of importance factor γ across the network under the static system setting.	77
5.3	Voltage mismatch squared error versus the total number of updates across the network with different step-size choices of α and β under the static system setting.	78
5.4	Voltage mismatch squared error versus the total number of updates across the network with fixed step-size choices of α and β and varying bus activation rate under the static system setting.	79
5.5	Daily voltage mismatch squared error for three different control strategies at phase a. A total communication link failure occurs from hour 16:00 to 24:00.	80
6.1	A cyber-physical MG network with DICs and related controllers	86
6.2	Angle dynamics of VSG governed by a second order swing equation.	88
6.3	Operation of DIC- i and its interaction with the microgrid and the consensus network under the proposed control design.	95
6.4	Proposed control diagrams for individual DIC- i	104
6.5	One-line diagram of the 4-bus/3-DIC microgrid.	104
6.6	DICs' (a) active power outputs; (b) droop frequencies under a severe load change from half to full load at $t = 10$	106
6.7	DICs' droop frequency under varying virtual inertia \mathbf{J} with $\mathbf{J}_b = \mathbf{1} \times 1.061 \text{ kg} \cdot \text{m}^2$	107
6.8	One-line diagram of the 14-bus/6-DIC microgrid.	108
6.9	Case II-A: DICs' (a) active power outputs; (b) droop frequencies.	109
6.10	Case II-A: Detection indices F_{ij}^k of DIC-1.	110
6.11	Case II-B: DICs' (a) active power outputs; (b) droop frequencies.	111
6.12	Case II-B: Detection indices: (a) F_{ij}^k of DIC-1; (b) F_{ij}^k of DIC-2.	111
6.13	Case II-C: DICs' (a) active power outputs; (b) droop frequencies.	112
6.14	Case II-C: Detection indices F_{ij}^k of DIC-1.	112

6.15	Case II-D: DICS' (a) active power outputs; (b) droop frequencies.	113
6.16	Case II-D: Detection indices: (a) F_{ij}^k of DIC-1; (b) F_{ij}^k of DIC-2.	113
7.1	Reference industrial microgrid topology for this work.	117
7.2	Notional representation of a standard IEC 61850 substation architecture.	118
7.3	Reference microgrid communication architecture and data types.	127
7.4	Frequency and active power output response to a load disturbance.	128
7.5	Frequency, active power output, and local x_i update responses to a local controller attack in steady state operation.	129
7.6	Frequency, active power output, and local x_i update responses to a local controller attack during a load disturbance.	130
7.7	Frequency, active power output, and local x_i update responses to an communication link attack on a control command.	130
A.1	Multi-phase online implementation of the proposed DVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of ADMM iterations and control updates, while the two vertical arrows correspond to the communications exchange among node (j, ϕ) and the neighboring nodes.	142

CHAPTER 1

INTRODUCTION

In this chapter, we highlight the need to design the decentralized optimization-based controls proposed in the thesis for power distribution networks and microgrids. Additionally, we summarize the contributions of this work and outline the thesis.

1.1 Motivation and Context

Large system-wide power outages such as the 2003 Northeast blackout can be great catastrophes. Within two days of the 2003 outage, over 50 million people lost power and this event contributed to at least 11 deaths and cost an estimated \$6 billion. Experts assumed that there had been a severe shortage of reactive power in northern Ohio leading to voltage sag and transmission line failures that set the blackout in motion. According to U.S.-Canada Power System Outage Task Force, inadequate reactive power support was one of the primary factors causing the voltage collapse leading to a system-wide blackout [1]. As the power system became one of the most important engineering achievements of the 20th century [2], maintaining its stability, reliability, and sustainability is necessary for modern society. To this end, the principle objectives of the thesis are to design decentralized control frameworks that i) utilize reactive power from power electronics-interfaced distributed energy resources (DERs) to enhance the voltage stability in distribution networks and ii) perform frequency regulation for an isolated microgrid to improve power system reliability, e.g., minimizing the impacts of blackout in the transmission grid.

The smart grid vision has led to an unprecedented level of information flow across the network by capitalizing on state-of-the-art technologies in sensing, control, and communication. Nonetheless, by and large, the communication infrastructure that supports power distribution networks is and will continue to be a challenging factor while designing the aforementioned control objec-

tives. Henceforth, transforming conventional centralized control designs to decentralized ones is becoming increasingly important to the operation of modern distribution networks. To this end, we propose decentralized optimization approach-based control designs to reduce the communication complexity and overhead while guaranteeing satisfactory control performance. Considering practical implementation constraints of this cyber-physical network, we further focus on improvements over our decentralized designs to be either cognizant of the instantaneous availability of communication link or robust to malicious cyber attacks. To sum up, the goal of this thesis is to develop cybersecure decentralized control strategies which are scalable and flexible with respect to the size of the network while enjoying minimal overhead and robustness in sensing, communication, and computation.

1.2 Organization of the Thesis

We give an overview of the ensuing chapters and corresponding contributions.

Chapter 2: A brief overview of the modeling of an unbalanced multi-phase distribution networks and an power electronics-interfaced inverter design is provided. In addition, the linearized counterpart of both single- and multi-phase networks is also introduced for facilitating the development of decentralized control designs and their applications.

Chapter 3: We develop a local dynamic optimization framework for analyzing the performance of a voltage control scheme based on gradient projection methods. Thanks to the physical network couplings, the local voltage measurement serves as the instantaneous system-wide gradient direction by forming a weighted voltage mismatch objective. Compared to earlier results for a static optimization scenario, we significantly extend the analysis on convergence conditions and error performance to account for two dynamic scenarios: i) the nodes perform the local update in an asynchronous fashion; and ii) the network operating point is dynamically changing. Last, we provide the rules of selecting a proper step-size, which shows a trade-off between the convergence speed and the steady-state tracking error for the dynamic control design.

The work presented in this chapter has been published in [3, 4].

Chapter 4: In the context of improving the performance of the local voltage design, we develop

a fully distributed voltage control (DVC) design to manage the VAR-capable DERs deployed to an increasingly dynamic and variable distribution system. Under the limited VAR scenario, the local control in Chap. 3 may attain a suboptimal voltage profile because of lack of information exchanges. To leverage the network-wide VAR support, we cast the voltage control problem as a quadratic programming one that minimizes the unweighted voltage mismatch under VAR resource limits everywhere by formulating the (multi-phase) power flow model to linear equality constraints. The alternating direction method of multipliers (ADMM) distributed optimization algorithm is evoked, which requires only information exchange among neighboring buses and totally local computations. To account for external disturbances and lack of reliable communication links, we further extend the implementation of the proposed ADMM-based control in an online fashion that is robust to random failure of communication links.

The work presented in this chapter has been published in [5].

Chapter 5: In this chapter, we integrate the attractive features from both local and distributed voltage control designs and develop the advanced hybrid control framework. We have cast the specially-designed voltage control problem by combining both weighted and unweighted voltage mismatch objectives under limited VAR resources. The PPD-based algorithm is then evoked and only requires voltage measurement exchanges among neighboring buses with local computations. We have further provided the stable step-size choices to guarantee control stability. To cope with cyber resource constraints and lack of reliable communication links, we have extended the hybrid voltage control (HVC) design to have robustness against random communication link failures and, in particular, communication-cognizant feature to account for the worst-case scenario of a total communication outage. Interestingly, under this worst case, the HVC boils down to a surrogate local control of Chap. 3. Thus, the adequate control performance can still be achieved as it responds to local voltage variations.

The work presented in this chapter has been submitted in [6].

Chapter 6: To date, MGs are foreseen to play an important role in distribution networks to enhance the power system reliability. Under isolated MGs, these mostly power electronics-interfaced distributed energy resources (DERs) exhibiting low-inertia characteristic have raised major concerns over the frequency stability issues. To address these issues, we adopt the concept of the virtual synchronous generator (VSG) and develop a distributed secondary frequency control for

DICs in isolated MGs. Our proposed control architecture consists of a local VSG-based droop control in the primary level and a distributed PPD-based algorithm in the secondary level. Interestingly, our specially designed control algorithm would seamlessly incorporate the dynamics of both power flow and VSG-based DICs which is implemented in the physical system by itself. As most hierarchical control frameworks assume the time-scale separation among different control levels, our optimization-based control design would explicitly account for the original dynamical model of network-coupled DICs in establishing the stability conditions of the distributed control design. This would improve the conventional MGs control framework by incorporating various time-scaled control objectives automatically and dynamically. Additionally, to make the proposed control robust to cyber attacks, we analyze two types of malicious attacks on the communication network, namely, the link and node attacks. Model-based anomaly detection and localization strategies are developed based on the dual variable related metrics.

The work presented in this chapter has been submitted in [7].

Chapter 7: As communication networks continue to advance in electric power systems, the industry standard called IEC 61850 has emerged for metering, protection, and control functions. IEC 61850 provides a standard for configuring various vendors' intelligent electronic devices (IEDs) for MG automation systems to be able to communicate with each other. Based on this protocol, we design the secondary frequency control as a consensus optimization problem. To avoid a single point of failure and enhance DERs' plug-and-play capability, we propose to solve this problem in a decentralized fashion under IEC 61850 by adopting the alternating direction method of multipliers (ADMM) algorithm. Motivated by the work in Ch. 6, we advocate modifying the ADMM updates originally derived for the steady state objective to an online feedback-based scheme, incorporating the instantaneous power measurements. Interestingly, it turns out that we do not need to explicitly model the MG power flow. Meanwhile, the instantaneous power feedback signal would couple DERs with the dynamics of power system networks, steering towards a model-free control approach. Our main contribution lies in the development of a *collaborative* defense strategy against these attacks by leveraging the communication capabilities under the IEC 61850 standard. To enhance the robustness to malicious control command attacks, we employ the round-robin (RR) technique at the central supervisor for generating the consensus variable based on a variant of the ADMM algorithm. This provides a multi-pronged approach to resilient and efficient MG opera-

tion in the face of adversarial conditions. These algorithms are demonstrated through simulation analysis of several use cases of interest.

The work presented in this chapter has been published in [8].

Chapter 8: In the final chapter, we summarize the contributions made by this thesis and include insights and suggestions for future research efforts.

CHAPTER 2

SYSTEM MODELING FOR VOLTAGE CONTROL IN DISTRIBUTION NETWORKS

In this chapter, we provide a brief overview of the modeling of an unbalanced multi-phase distribution network and a power electronics-interfaced inverter design. The linearized counterpart of both single- and multi-phase networks is also introduced for development of control designs and their applications in the ensuing chapters.

2.1 Modeling of Distribution Networks

Distribution systems are inherently unbalanced due to, e.g., untransposed lines, unbalanced loading conditions, and multi-phase feeds (see, e.g., [9, Ch. 2]). Hence, it is necessary to model the coupling effects across multiple phases. To this end, we denote a tree-topology distribution network using $(\mathcal{N}, \mathcal{E})$ with the buses in $\mathcal{N} := \{0, \dots, N\}$ and line segments in $\mathcal{E} := \{(i, j)\}$ (see Fig. 2.1 for a multi-phase radial feeder illustration). Bus 0 corresponds to the secondary side of a voltage regulator at the feeder head, assumed to be of unity reference voltage, which is controlled by an on-load tap changer of transformer (OLTC). Due to a much slower time-scale of OLTC as compared to that of power electronic-interfaced DERs, we consider the slow acting OLTC to have a fixed tap position while performing our ensuing control designs (see [10] for optimal tap settings of an OLTC). Without loss of generality, we assume all buses are connected by three phases. Accordingly, each bus- or line-associated variable is represented by a 3×1 complex vector. For example, vector $\mathbf{V}_j := [V_j^a \ V_j^b \ V_j^c]^T$ has the voltage phasor at phase $\phi \in \mathcal{P} := \{a, b, c\}$ per bus j , and \mathbf{I}_{ij} and \mathbf{S}_{ij} denote the complex line current and power flow from bus i to bus j , respectively. All are per-unit (p.u.) quantities. For each line (i, j) with given phase impedance matrix \mathbf{Z}_{ij} , the multi-phase Ohm's law asserts that

$$\mathbf{V}_j = \mathbf{V}_i - \mathbf{Z}_{ij}\mathbf{I}_{ij}. \quad (2.1)$$

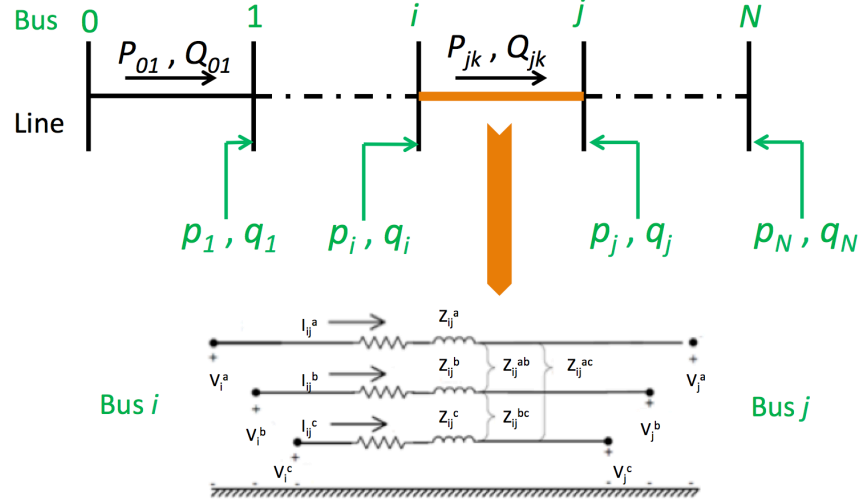


Figure 2.1: A radial multi-phase distribution feeder with bus and line associated variables.

The difference between from-end and to-end line flows equals to $\text{diag}(\mathbf{Z}_{ij}\mathbf{I}_{ij})\mathbf{I}_{ij}^*$ with $*$ denoting conjugate transpose. Accordingly, the flow balance equation for line (i, j) is given by

$$\mathbf{S}_{ij} - \text{diag}(\mathbf{Z}_{ij}\mathbf{I}_{ij})\mathbf{I}_{ij}^* - \sum_{k \in \mathcal{N}_j^+} \mathbf{S}_{jk} = -\mathbf{s}_j, \quad (2.2)$$

where $\mathcal{N}_j^+ \subset \mathcal{N}$ denotes the subset of downstream neighbors, e.g., $\mathcal{N}_i^+ = \{j\}$ in Fig. 2.1, and $\mathbf{s}_j := \mathbf{p}_j + \mathbf{j}\mathbf{q}_j$ the complex power injection per bus j , with the active (reactive) power injection denoted by \mathbf{p}_j (\mathbf{q}_j), respectively. In addition, let $\mathbf{q}_{j,g}$ denote the controllable VAR contributed by local DER inverters and $\mathbf{q}_{j,c}$ represent the VAR level of local loads. We have the total VAR injection $\mathbf{q}_j := \mathbf{q}_{j,g} - \mathbf{q}_{j,c}$. Using (2.1) and (2.2), we can represent the nonlinear relation between voltage and (controllable) VAR injection everywhere in the network. This constitutes the basic nonlinear AC power flow modeling for unbalanced multi-phase feeders; see, e.g., [9, Chap. 3-7] and [11–14]. To facilitate ensuing control designs, we firstly present the linearized model for single-phase networks, and its multi-phase counterpart will be extended later on. The accuracy of these linearized models can be numerically corroborated by several recent works [15–17].

2.1.1 Linear Modeling of Single-Phase Networks

We consider a single-phase distribution network $(\mathcal{N}, \mathcal{E})$. By an abuse of notation, we neglect the phase terms for a single-phase network. Per bus j , let v_j denote its voltage magnitude and p_j and q_j the real and reactive power injection. For each line (i, j) , let r_{ij} and x_{ij} denote its resistance and reactance and P_{ij} and Q_{ij} the power flow from i to j , respectively. Assuming negligible line losses and almost flat voltage, i.e., $v_j \cong 1, \forall j$, the AC single-phase power flow model can be simplified to the so-termed LinDistFlow (linearized) model [18], given for each line (i, j) as

$$P_{ij} - \sum_{k \in \mathcal{N}_j^+} P_{jk} = -p_j, \quad (2.3a)$$

$$Q_{ij} - \sum_{k \in \mathcal{N}_j^+} Q_{jk} = -q_j, \quad (2.3b)$$

$$v_i - v_j = r_{ij}P_{ij} + x_{ij}Q_{ij}. \quad (2.3c)$$

Power-balance equations of bus j correspond to (2.3a)-(2.3b) while (2.3c) relates the voltage drop to line flow. The LinDistFlow approximation leads to a relatively small error of at most 0.25% (1%) if the voltage magnitude deviation is within a 5% (10%) range which is typical under normal operating conditions [17]. In addition, several earlier papers [3, 15, 17] have verified the competitiveness of the LinDistFlow model to the nonlinear AC flow model of (2.1) and (2.2) on realistic feeders.

The matrix form of (2.3) is introduced to further demonstrate its linearity. Let the $(N + 1) \times N$ matrix \mathbf{M}^o denote the graph incidence matrix for $(\mathcal{N}, \mathcal{E})$. Its ℓ -th column corresponds to a line (i, j) , with all zero entries except for the i -th and j -th (see e.g., [19, pg. 6]). We set $M_{il}^o = 1$ and $M_{jl}^o = -1$ if bus j is downstream from i . Let \mathbf{m}_0^T be the first row of \mathbf{M}^o corresponding to bus 0 with the rest of the rows in the $N \times N$ submatrix \mathbf{M} . For tree networks, \mathbf{M} is full-rank and thus invertible; see e.g. [19][Chp. 2]. Upon concatenating all scalar variables into vector form, one can represent (2.3) as follows:

$$-\mathbf{M}\mathbf{P} = -\mathbf{p}, \quad (2.4a)$$

$$-\mathbf{M}\mathbf{Q} = -\mathbf{q}, \quad (2.4b)$$

$$\mathbf{m}_0 + \mathbf{M}^T \mathbf{v} = \mathbf{D}_r \mathbf{P} + \mathbf{D}_x \mathbf{Q}, \quad (2.4c)$$

where \mathbf{D}_r is an $N \times N$ diagonal matrix with diagonals capturing all line r_{ij} 's and similarly to $\mathbf{D}_x := \text{diag}\{x_{ij}\}$.

The VAR injection $\mathbf{q} = \mathbf{q}^g - \mathbf{q}^c$, where \mathbf{q}^g denotes the VAR contributed by DER inverters while \mathbf{q}^c is the uncontrollable VAR consumption of all loads. Solving for \mathbf{P} and \mathbf{Q} and substituting $v_0 = 1$ into (2.4c) give rise to the following linear relation:

$$\mathbf{B}\mathbf{v} = \mathbf{q}^g + \mathbf{w}, \quad (2.5)$$

where the vector $\mathbf{w} := \mathbf{M}\mathbf{D}_x^{-1}\mathbf{D}_r\mathbf{M}^{-1}\mathbf{p} - \mathbf{q}^c - \mathbf{M}\mathbf{D}_x^{-1}\mathbf{m}_0$ captures the system operating conditions, and $\mathbf{B} := \mathbf{M}\mathbf{D}_x^{-1}\mathbf{M}^T$ is the *Bbus matrix* used in the dc power flow model (see e.g., [20, Sec. 6.16]). By definition, \mathbf{B} is a weighted, reduced graph Laplacian matrix and thus has a unique structure according to the network topology, i.e., $B_{ij} = B_{ji} = 0, \forall (i, j) \notin \mathcal{E}$. This sparsity structure will play a crucial role in designing ensuing controls.

2.1.2 Linear Modeling of Multi-Phase Networks

For more realistic representation, linear modeling of single-phase networks can be generalized to unbalanced multi-phase networks which necessitate consideration of coupling effects among multiple phases. Without loss of generality, assume all buses are connected by three phases and, accordingly, the bus- or line-associated variable is represented by a 3×1 complex vector, same as the ones in Sec. 2.1. Assuming negligible power losses as for LinDistFlow derivations, the per-phase flow balance would still hold as in (2.4a) and (2.4b). However, the multi-phase coupling would result in a more complicated voltage drop relation compared to (2.4c). For a given impedance matrix \mathbf{Z}_{ij} per line (i, j) , Ohm's law asserts the voltage drop $\mathbf{V}_i - \mathbf{V}_j = \mathbf{Z}_{ij}\mathbf{I}_{ij}$. Assuming almost flat voltage, one can multiply this relation by the Hermitian of each side and take the diagonals of the resultant matrix, yielding

$$\tilde{\mathbf{v}}_i - \tilde{\mathbf{v}}_j \cong \text{Re}\{\text{diag}(\mathbf{V}_i \mathbf{I}_{ij}^H \mathbf{Z}_{ij}^H)\} \quad (2.6)$$

with the voltage magnitude vector $\tilde{\mathbf{v}}_j := |\mathbf{V}_j|$. As in [12], we assume that the voltage magnitudes between phases are similar, and that angle unbalance is small; i.e.,

$$V_i^b/V_i^a \approx V_i^c/V_i^b \approx V_i^a/V_i^c \approx \tilde{\sigma} := e^{-j2\pi/3}.$$

This assumption leads to the following approximation of (2.6)

$$\tilde{\mathbf{v}}_i - \tilde{\mathbf{v}}_j \cong \text{Re}\{\text{diag}(\tilde{\boldsymbol{\sigma}})\mathbf{Z}_{ij}^H \text{diag}(\tilde{\boldsymbol{\sigma}}^*)\mathbf{S}_{ij}\} \cong \text{Re}\{\tilde{\mathbf{Z}}_{ij}\mathbf{S}_{ij}\}, \quad (2.7)$$

where matrix $\tilde{\mathbf{Z}}_{ij} := \text{diag}(\tilde{\boldsymbol{\sigma}})\mathbf{Z}_{ij}^H \text{diag}(\tilde{\boldsymbol{\sigma}}^*)$ stands for an equivalent impedance by defining $\tilde{\boldsymbol{\sigma}} := [1 \ \tilde{\sigma} \ \tilde{\sigma}^2]^T$.

One can develop the multi-phase counterpart of (2.4c) using (2.7). Define the block diagonal matrix $\mathbb{M} := \text{bdiag}(\mathbf{M}^a, \mathbf{M}^b, \mathbf{M}^c)$ and vector $\mathbf{m}_0 := [\mathbf{m}_0^a; \mathbf{m}_0^b; \mathbf{m}_0^c]$, as the counterparts to \mathbf{M} and \mathbf{m}_0 , respectively. Upon concatenating the full voltage $\mathbf{v} = [\tilde{\mathbf{v}}^a; \tilde{\mathbf{v}}^b; \tilde{\mathbf{v}}^c]$ and similarly for \mathbb{P} and \mathbb{Q} , the approximation (2.7) becomes

$$\mathbb{M}^T \mathbf{v} + \mathbf{m}_0 = \mathbb{D}_r \mathbb{P} + \mathbb{D}_x \mathbb{Q}, \quad (2.8)$$

where resistance and reactance matrices \mathbb{D}_r and \mathbb{D}_x can account for multi-phase coupling; see e.g., [12, 13]. To derive the aforementioned matrices, we let $\tilde{\sigma} = \tilde{\gamma} + j\tilde{\zeta}$ and $\tilde{\sigma}^2 = \tilde{\sigma}^* = \tilde{\alpha} + j\tilde{\beta}$, and $\tilde{\mathbf{Z}}_{ij}$ in (2.7) becomes

$$\tilde{\mathbf{Z}}_{ij} = \begin{bmatrix} Z_{ij}^a & \tilde{\sigma}^* Z_{ij}^{ab} & \tilde{\sigma} Z_{ij}^{ac} \\ \tilde{\sigma} Z_{ij}^{ba} & Z_{ij}^b & \tilde{\sigma}^* Z_{ij}^{bc} \\ \tilde{\sigma}^* Z_{ij}^{ca} & \tilde{\sigma} Z_{ij}^{cb} & Z_{ij}^c \end{bmatrix}^H \quad (2.9)$$

where the superscript of Z_{ij} represents the phase or phase coupling term between line (i, j) as shown in Fig. 2.1. Across phases a and b, we denote diagonal matrix $\mathbf{D}_r^{ab} := \text{diag}(\text{Re}\{\mathbf{Z}^{ab}\})$ collecting all the coupling resistances where complex vector $\mathbf{Z}^{ab} := \{Z_{ij}^{ab}\}_{(i,j) \in \mathcal{E}}$. Similarly, diagonal matrix $\mathbf{D}_x^{ab} := \text{diag}(\text{Im}\{\mathbf{Z}^{ab}\})$ accounts for all the coupling reactances. The same notations are

adopted for couplings among other phases. Thus, we have

$$\mathbb{D}_r = \begin{bmatrix} \mathbf{D}_r^a & \tilde{\alpha}\mathbf{D}_r^{ab} + \tilde{\beta}\mathbf{D}_x^{ab} & \tilde{\gamma}\mathbf{D}_r^{ac} + \tilde{\zeta}\mathbf{D}_x^{ac} \\ \tilde{\gamma}\mathbf{D}_r^{ab} + \tilde{\zeta}\mathbf{D}_x^{ab} & \mathbf{D}_r^b & \tilde{\alpha}\mathbf{D}_r^{bc} + \tilde{\beta}\mathbf{D}_x^{bc} \\ \tilde{\alpha}\mathbf{D}_r^{ac} + \tilde{\beta}\mathbf{D}_x^{ac} & \tilde{\gamma}\mathbf{D}_r^{bc} + \tilde{\zeta}\mathbf{D}_x^{bc} & \mathbf{D}_r^c \end{bmatrix} \quad (2.10)$$

$$\mathbb{D}_x = \begin{bmatrix} \mathbf{D}_x^a & \tilde{\alpha}\mathbf{D}_x^{ab} - \tilde{\beta}\mathbf{D}_r^{ab} & \tilde{\gamma}\mathbf{D}_x^{ac} - \tilde{\zeta}\mathbf{D}_r^{ac} \\ \tilde{\gamma}\mathbf{D}_x^{ab} - \tilde{\zeta}\mathbf{D}_r^{ab} & \mathbf{D}_x^b & \tilde{\alpha}\mathbf{D}_x^{bc} - \tilde{\beta}\mathbf{D}_r^{bc} \\ \tilde{\alpha}\mathbf{D}_x^{ac} - \tilde{\beta}\mathbf{D}_r^{ac} & \tilde{\gamma}\mathbf{D}_x^{bc} - \tilde{\zeta}\mathbf{D}_r^{bc} & \mathbf{D}_x^c \end{bmatrix}. \quad (2.11)$$

Accordingly, the multi-phase counterpart of (2.5) can be similarly derived as

$$\mathbb{B}\mathbf{v} = \mathbf{q}^g + \mathbf{w} \quad (2.12)$$

with $\mathbb{B} := \mathbb{M}\mathbf{D}_x^{-1}\mathbb{M}^T$ and $\mathbf{w} := \mathbb{M}\mathbf{D}_x^{-1}\mathbf{D}_r\mathbb{M}^{-1}\mathbf{p} - \mathbf{q}^c - \mathbb{M}\mathbf{D}_x^{-1}\mathbf{m}_0$ capturing the system operating condition. Again, matrix \mathbb{B} is a weighted graph Laplacian matrix for the multi-phase network, and shares similar sparse structures as \mathbf{B} where entries are zero for buses not connected by a line segment.

Remark 1 (Modeling Considerations). Nonlinearity of the power flow model could be tackled by the formulation of semidefinite programming (SDP) [10]. Generally, a rank relaxation approach is adopted in order to obtain a SDP convex problem formulation. Nonetheless, the SDP-based power flow formulation would face several challenges when applied to power (distribution) networks in practice. First, the resultant power flow solution could be non-exact [11] under this setting when power networks are unbalanced multi-phase. Furthermore, performance guarantees for the SDP approach fail to hold for general network topology such as meshed systems [21]. Last, the SDP solution increases the computational burden significantly as the size of the network grows. Compared to the SDP modeling approach, the LinDistFlow models (2.5) and (2.12) hold for more general scenarios including meshed topology and unbalanced three-phase systems while the resultant algorithms enjoy minimal computational complexity [3, 22]. Thus, we have adopted the LinDisFlow models for developing the ensuing control designs. However, the corresponding numerical tests will be performed using exact nonlinear power flow solvers on practical power

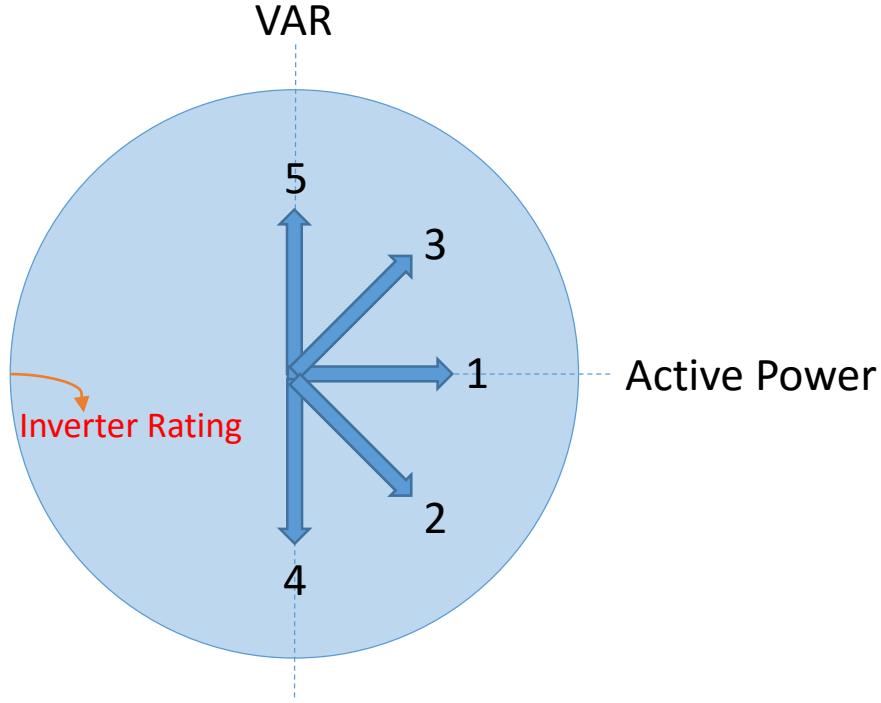


Figure 2.2: Mode 1 contributes only positive active power output while modes 4 and 5 generate negative and positive VAR representing inductive or capacitive characteristics, respectively. In addition, modes 2 and 3 consist of mixed active and VAR outputs. All are bounded by the physical inverter rating.

networks for accurate system responses.

2.2 Modeling of Inverters

There are many existing reactive power resources in the current distribution systems such as capacitor banks and inverters-based DERs, e.g, solar photovoltaic (PV) panels and plug-in hybrid electric vehicles. In the interest of this thesis, we focus on the fast-acting inverters to provide VAR for the voltage regulation objective while fixing other slow-acting devices to be inactive. Without loss in generality, an inverter operates in five different modes as shown in Fig. 2.2. Mode 1 is the base case where only the real power is being generated. Meanwhile, modes 2 and 3 have the apparent power injection at a lagging and a leading power factor, respectively. Last, mode 4 (5) behaves like a purely inductive (capacitive) element. All are bounded by the physical inverter

rating. For example, considering a PV-based inverter, a conventional operating strategy would strictly generate real power during the day in mode 1. Nonetheless, with our ensuing decentralized control designs, we would adjust VAR judiciously among all the modes to take advantage of its VAR supporting capability. Additionally, we do not consider the modeling of inverter dynamics in this thesis except for VAR limit constraints. This approach can be justified under the system-level control setup, as used by [11, 15, 17]. Thanks to a time-scale separation between the inverter internal dynamics and external control [23], it has been shown in [24] that ignoring the former does not significantly affect the stability conditions of the latter.

CHAPTER 3

COMMUNICATION-FREE LOCAL VOLTAGE CONTROL

By treating the transmission grid as an infinity bus, voltage stability is the main concern in distribution networks where more DERs are being installed in the near future. However, the currently limited communication infrastructure supported power distribution network challenges the employment of voltage control approaches by minimizing a centralized voltage mismatch error. Meanwhile, cost concerns for DER products inevitably limit their sensing/computation capabilities. Therefore, centralized voltage control designs are not practically attractive because their success strongly relies on high-quality communications either between a control center and remote devices [25], or among neighboring devices [15, 26, 27]. These optimization-based open-loop control designs may become unstable under communication delays or noises during online implementations. Also, it is unclear whether these approaches will be robust to asynchronous computational speeds among the networked DERs of heterogeneous hardware capabilities.

To cope with the limited cyber infrastructure, a VAR-based local voltage control framework is first proposed for distribution network operations; see e.g., [3, 17, 28, 29]. Under this framework, each node only needs to measure its locally available voltage level as the controller input. Our earlier work [3] has offered an overarching analysis that generalizes a variety of local control designs, along with convergence analysis for static system scenarios. Interestingly, the local voltage measurement naturally provides the instantaneous gradient information for a centralized error objective by weighting the voltage mismatch. Hence, the local voltage control approach using this measurement boils down to the classical gradient-projection (GP) method accounting for the limits on reactive power resources. This voltage-measurement-based local design does not require any real-time communications, and can be implemented with minimal upgrades in sensing hardware.

The goal of this chapter is to analyze the performance of this local GP-based voltage control design under two dynamic scenarios: i) the nodes perform the local update in an asynchronous

fashion, and ii) the network operating condition is dynamically changing. The scenario of asynchronous updates arises from heterogeneous hardware capabilities among different DERs. It is also motivated to allow the “plug-and-play” functionality for flexible DER integration to distribution networks. The classical asynchronous optimization framework always accounts for the case of outdated information from other nodes, and thus the choice of step-size has to be more conservative due to the information delay; see e.g., [30] and [31]. Different from this, the voltage measurement for our local control design always provides the up-to-date gradient information and does not suffer any information delay. Thanks to this physical power network coupling, we can show that the choice of the step-size for the asynchronous local updates is the same for the synchronous case. Hence, its convergence condition is robust to a potential discrepancy in the control update rates among different DERs.

In power networks, voltage control designs under time-varying operating conditions have been implemented as the static optimal power flow solutions to dynamic settings in a heuristic fashion [32, 33]. The scope of these efforts is centered more around dynamic voltage control implementations than providing control performance guarantees. The latter is of high interest when accounting for the variability of networked generations and loads in practice. With a time-varying objective function, this problem becomes one of *stochastic optimization*; see e.g., [34]. Stochastic approximation algorithms such as stochastic (sub-)gradient descent have been developed in, e.g., [35, 36] and have been adopted by [14] for this voltage control problem. Nonetheless, performance analysis for stochastic optimization algorithms has focused on the convergence to the optimal solution that minimizes the expected objective function [34]. Aiming at the error bound in tracking the instantaneous optimal solution, our analysis is more closely related to the body of work on *dynamic convex optimization*; see e.g., [37–39]. This type of problem typically arises from applications in autonomous teams and wireless sensor networks such as target tracking [40] and estimation of the stochastic path [41]. Some of these dynamic optimization algorithms follow a gradient descent update, but none of them has considered the formulation of constrained optimization. This is the key difference from our voltage control problem since the control input has to be feasible under the dynamic reactive power limits. Hence, the main contribution of our work lies in the fact that it explicitly accounts for the time-varying projection operation of *constrained* dynamic optimization. Our tracking error performance bounds will be derived for a quadratic ob-

jective function and an autoregressive dynamic model, motivated by this specific voltage control problem. Nonetheless, our analytical results can be extended to more general constrained dynamic optimization problems with smooth strongly convex objective functions under stochastic processes that have bounded iterative changes.

3.1 Local Voltage Control Problem

For ease of presentation, we present the local control design under a single-phase network $(\mathcal{N}, \mathcal{E})$. Similarly, its multi-phase counterpart could be readily extended by considering (2.12). Thanks to advances in power electronics, DERs' inverters can also provide *reactive power* (VAR), a quantity that is known to have a significant impact on the network voltage level. Accordingly, our goal is to control VAR, such that \mathbf{v} approaches a given desirable voltage profile $\boldsymbol{\mu}$. This is similar to the goal of secondary voltage control design in transmission and microgrid networks to effectively coordinate the network VAR resources and enhance the voltage stability [42, 43]. One can achieve various operational tasks or specifications by designing the preferred voltage profile $\boldsymbol{\mu}$. For high voltage quality or effective conservation voltage reduction [44], a flat voltage profile $\boldsymbol{\mu} = \mathbf{1}$ would be preferred. Moreover, the distribution management system (DMS) can perform network-specific planning studies to determine the best $\boldsymbol{\mu}$, as in earlier work [3, 16, 22]. For example, one can choose a decreasing voltage profile to potentially reduce system losses, compensate voltage unbalance, or correct power factor. Nonetheless, the selection of $\boldsymbol{\mu}$ would be infrequently performed and not affect the implementation of our ensuring algorithms.

With an abuse of notation, we replace \mathbf{q}^g by \mathbf{q} in the rest of the thesis and denote $\mathbf{X} := \mathbf{B}^{-1}$ and $\bar{\mathbf{v}} := \mathbf{X}\mathbf{w}$. Thus, (2.5) becomes

$$\mathbf{v} = \mathbf{X}\mathbf{q} + \bar{\mathbf{v}} \tag{3.1}$$

At every time instance k , let $\bar{\mathbf{v}}_k$ denote the instantaneous nominal voltage profile. To allow for a local control design, it turns out that one can minimize a weighted voltage mismatch error, as

given by

$$\mathbf{q}_k^* = \arg \min_{\mathbf{q} \in \mathcal{Q}_k} f_k(\mathbf{q}) := \frac{1}{2}(\mathbf{X}\mathbf{q} + \bar{\mathbf{v}}_k - \boldsymbol{\mu})^\top \mathbf{B}(\mathbf{X}\mathbf{q} + \bar{\mathbf{v}}_k - \boldsymbol{\mu}) \quad (3.2)$$

where the constraint set $\mathcal{Q}_k := \{\mathbf{q} | \mathbf{q} \in [\underline{\mathbf{q}}_k, \bar{\mathbf{q}}_k]\}$ accounts for the time-varying limits of local reactive power resources at every bus. This constraint on VAR is due to physical inverter ratings and would vary according to their active (real) power outputs as depicted in Fig. 2.2 [29]. In addition, matrix $\mathbf{B} := \mathbf{M}\mathbf{D}_x^{-1}\mathbf{M}^\top$, by definition, is a weighted, reduced graph Laplacian for $(\mathcal{N}, \mathcal{E})$. Since all the reactance values are positive, \mathbf{B} is symmetric and positive definite [3]. Accordingly, the weighted voltage mismatch error objective of (3.2) is convex, and in fact quadratic, in the variable \mathbf{q} . Ideally, the unweighted error norm $\|\mathbf{v} - \boldsymbol{\mu}\|$ is the best objective in order to achieve the flat voltage profile. Compared to the traditional paradigm of maintaining the voltage within limits, this unweighted objective can improve the system-wide voltage profile by coordinating network-wide VAR resources. Albeit the problem (3.2) minimizes a surrogate objective, it has been shown in [3] that \mathbf{q}_k^* can closely approximate the optimal solution to the ideal unweighted error norm, especially if there are abundant reactive power resources. Last, it is possible to use other convex error penalty functions such as the Huber's loss function [45, Ch. 7] instead of the squared error norm objective in (3.2). This approach generalizes the current design to allow for some tolerance in the voltage mismatch error, which may be more attractive under the scenarios of limited VAR resources. Thanks to the separable structure of the box constraint \mathcal{Q}_k , the gradient-projection (GP) method [46, Sec. 2.3] can be invoked to solve (3.2). Upon forming its instantaneous gradient $\nabla f_k(\mathbf{q}_k) := \mathbf{X}\mathbf{q}_k + \bar{\mathbf{v}}_k - \boldsymbol{\mu}$, the GP iteration for a given positive step-size $\epsilon > 0$ becomes

$$\mathbf{q}_{k+1} = \mathbb{P}_k [\mathbf{q}_k - \epsilon \mathbf{D} \nabla f_k(\mathbf{q}_k)], \quad (3.3)$$

where the projection operator \mathbb{P}_k thresholds any input to be within \mathcal{Q}_k , and $\mathbf{D} := \text{diag}(d_1, \dots, d_N)$ is a diagonal scaling matrix that can be designed. As a first-order method, the GP method has a linear convergence rate, while the convergence speed depends on the condition number of the corresponding Hessian matrix [46, Sec. 3.3]. Motivated by this fact, the scaling matrix can be chosen according to the inverse of the diagonals of Hessian matrix by setting $\mathbf{D} := [\text{diag}(\mathbf{X})]^{-1}$

to approximate the Newton gradient. Note that a positive diagonal matrix \mathbf{D} affects neither the separability of operator \mathbb{P}_k , nor the optimality of the update (3.3).

By setting the GP iterate $\mathbf{q}_k \in \mathcal{Q}_k$ to be the control input at any time k , the instantaneous voltage becomes $\mathbf{v}_k = \mathbf{X}\mathbf{q}_k + \bar{\mathbf{v}}_k$ based on (3.1). Thanks to the physical power network coupling, the voltage-based feedback signal \mathbf{v}_k always provides the up-to-date gradient information as $\nabla f_k(\mathbf{q}_k) = \mathbf{v}_k - \boldsymbol{\mu}$. This voltage-based gradient feedback approach is potentially more robust to modeling mismatches [47, Sec. 8.9]. Accordingly, the GP update in (3.3) can be implemented by directly measuring the instantaneous voltage as

$$\mathbf{q}_{k+1} = \mathbb{P}_k [\mathbf{q}_k - \epsilon \mathbf{D}(\mathbf{v}_k - \boldsymbol{\mu})], \quad (3.4)$$

which can be completely decoupled into local updates at each bus because \mathbb{P}_k is separable. This local voltage control is very attractive with minimal hardware requirements as each bus only needs to measure its local voltage and requires no communication. The optimality and convergence conditions for (3.4) are summarized by the following proposition.

Proposition 1. *When $\bar{\mathbf{v}}_k = \bar{\mathbf{v}}$ and $\mathcal{Q}_k = \mathcal{Q}$ (time-invariant case), the local update (3.4) approaches the unique time-invariant optimizer \mathbf{q}^* of problem (3.2) if the step-size $\epsilon \in (0, 2/M)$ where*

$$M := \lambda_{\max}\{\tilde{\mathbf{X}}\} \quad (3.5)$$

is the largest eigenvalue of matrix

$$\tilde{\mathbf{X}} := \mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{\frac{1}{2}}. \quad (3.6)$$

Proof: The update (3.4) approaching the time-invariant optimizer \mathbf{q}^* is equivalent to having the iterative error mismatch $\|\mathbf{q}_k - \mathbf{q}^*\| \rightarrow 0$ as $k \rightarrow \infty$. To this end, we diagonally scale all iterates with $\mathbf{D}^{-\frac{1}{2}}$ and define $\mathbf{y}_k := \mathbf{D}^{-\frac{1}{2}} \mathbf{q}_k$ and $\mathbf{y}^* := \mathbf{D}^{-\frac{1}{2}} \mathbf{q}^*$. Since projection is a nonexpansive

mapping, the scaled error norm

$$\begin{aligned}
\|\mathbf{y}_{k+1} - \mathbf{y}^*\| &= \|\mathbf{D}^{-\frac{1}{2}} \mathbb{P}[\mathbf{q}_k - \epsilon \mathbf{D}(\mathbf{v}_k - \boldsymbol{\mu})] - \mathbf{D}^{-\frac{1}{2}} \mathbb{P}[\mathbf{q}^* - \epsilon \mathbf{D}(\mathbf{v}^* - \boldsymbol{\mu})]\| \\
&\leq \|\mathbf{D}^{-\frac{1}{2}}(\mathbf{q}_k - \mathbf{q}^*) - \epsilon \mathbf{D}^{\frac{1}{2}} \mathbf{X}(\mathbf{q}_k - \mathbf{q}^*)\| \\
&= \|\mathbf{y}_k - \mathbf{y}^* - \epsilon \mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{\frac{1}{2}}(\mathbf{y}_k - \mathbf{y}^*)\| \\
&= \|(\mathbf{I} - \epsilon \mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{\frac{1}{2}})(\mathbf{y}_k - \mathbf{y}^*)\| \\
&\leq \|(\mathbf{I} - \epsilon \mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{\frac{1}{2}})\| \|\mathbf{y}_k - \mathbf{y}^*\|.
\end{aligned} \tag{3.7}$$

Denoting matrix $\tilde{\mathbf{X}} := \mathbf{D}^{\frac{1}{2}} \mathbf{X} \mathbf{D}^{\frac{1}{2}}$, one can establish that its largest eigenvalue

$$M := \lambda_{\max}\{\tilde{\mathbf{X}}\} > 0. \tag{3.8}$$

This inequality holds because \mathbf{D} and \mathbf{X} are both positive definite (PD). A sufficient stability condition is to ensure the non-negative error $\|\mathbf{y}_k - \mathbf{y}^*\|$ is contracting at every iteration. By definition, the matrix Euclidean norm $\|\mathbf{I} - \epsilon \tilde{\mathbf{X}}\|$ equals to the largest singular value of $\mathbf{I} - \epsilon \tilde{\mathbf{X}}$. Because matrix $\tilde{\mathbf{X}}$ is PD [cf. (3.8)], $|1 - \epsilon M|$ is a singular value of $\mathbf{I} - \epsilon \tilde{\mathbf{X}}$. Given $\epsilon > 0$, having $|1 - M| < 1$ for every k ensures that (3.7) is a contraction mapping for every k . Accordingly, the scaled error norm $\|\mathbf{y}_k - \mathbf{y}^*\|$ would go to 0 in the limit and similarly for $\|\mathbf{q}_k - \mathbf{q}^*\|$ since \mathbf{D} is PD. Hence, to ensure the stability of (3.4), we need to have $\epsilon \in (0, 2/M)$. ■

The local voltage control design has to account for a variety of uncertainties in practical system implementations. First, due to heterogeneity of various DERs, it is difficult to perfectly synchronize the local update at different buses. This is especially important to facilitate the “plug-and-play” functionality for flexible DER integration. Second, the volatility and intermittence of electric loads and renewable-based generations challenge the static setting of time-invariant $\bar{\mathbf{v}}_k$. It is of considerable interest to quantify the performance of the local control (3.4) in terms of tracking the time-varying optimizer \mathbf{q}_k^* to the dynamic objective f_k .

3.2 Asynchronous Local Voltage Control

When the DERs are heterogeneous, it is logical that the local voltage update should be performed in an asynchronous fashion. This way, the buses with better computation and sensing capabilities do not need to wait for the slowest one. Accordingly, these buses can execute more updates for the same time interval and hence respond more quickly to localized voltage deviations.

To this end, let the set \mathcal{K}_j collect all the time instances when bus j executes its local update. The asynchronous counterpart of (3.4) can be modeled by

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{s}_k, \forall k \quad (3.9)$$

with the difference at bus j given by

$$s_{j,k} := \begin{cases} \mathbb{P}_{j,k} [q_{j,k} - \epsilon D_j(v_{j,k} - \mu_j)] - q_{j,k}, & k \in \mathcal{K}_j, \\ 0, & k \notin \mathcal{K}_j, \end{cases} \quad (3.10)$$

where D_j is the (j, j) -th entry of \mathbf{D} , while $\mathbb{P}_{j,k}$ projects the input to $[\underline{q}_{j,k}, \bar{q}_{j,k}]$.

To establish the convergence condition, every bus needs to update sufficiently often. Similar to the classical asynchronous algorithm analysis in [30, Ch. 7], we assume the following *bounded update delay* condition.

AS1 (Bounded Update Delay). *For every bus j and time instance $k \geq 0$, there exists a positive integer K such that at least one element in the set $\{k, k+1, \dots, k+K-1\}$ belongs to \mathcal{K}_j . Equivalently, every bus must update at least once every K iterations.*

In addition to Assumption 1, the analysis of classical asynchronous algorithms also assumes the *bounded information delay* condition [30, Ch. 7]. Due to potential communication delays among peer processors, the updates at some buses may not be executed based on the most up-to-date system-wide information. By assuming that the local information used for computing the gradient is potentially obsoleted by at most K iterations, it has been established in [30, Sec. 7.5] that the asynchronous GP algorithm in (3.9)-(3.10) converges to the optimal solution with more

conservative step-size choice given by

$$0 < \epsilon < 1/[M(1 + K + NK)]. \quad (3.11)$$

Due to information delay, the choice of step-size would depend on the slowest processor in the network. Generally, this bound on ϵ can be much smaller than the $2/M$ bound in Proposition 1, resulting in a much slower convergence compared with the synchronous case.

For our local voltage control, the gradient $\nabla f_k(\mathbf{q}_k) = \mathbf{v}_k - \boldsymbol{\mu}$ always holds. Thanks to the physical power network coupling, the local voltage $v_{j,k}$ always provides the up-to-date gradient information at every iteration k . Hence, whenever a node is active, the difference $s_{j,k}$ computed in (3.10) does not suffer from any information delay. This is different from most parallel and distributed algorithms where the updates at every processor require information sent by peer processors. Therefore, convergence of (3.9)-(3.10) no longer requires the more conservative choice of ϵ in [30, Sec. 7.5].

Theorem 1. *Under Assumption 1, when $\bar{\mathbf{v}}_k = \bar{\mathbf{v}}$ and $\mathcal{Q}_k = \mathcal{Q}$ (time-invariant case), the asynchronous update illustrated in (3.9)-(3.10) converges to the time-invariant optimizer \mathbf{q}^* if the step-size $\epsilon \in (0, 2/M)$.*

Proof: First, it is easy to show that the fixed-point of (3.9)-(3.10) is the same as (3.4) using contradiction. As for the convergence, by projecting any scalar q to $[\underline{q}_j, \bar{q}_j]$, it holds that $[\mathbb{P}_j(q) - q_{j,k}][\mathbb{P}_j(q) - q] \leq 0$ for the iterate $q_{j,k} \in [\underline{q}_{j,k}, \bar{q}_{j,k}]$ where $\mathbb{P}_j = \mathbb{P}_{j,k}$; see e.g., [30, Sec. 3.3.1]. This implies that at every iteration $k \in \mathcal{K}_j$ (use $f_k = f$ for time-invariant case)

$$s_{j,k}[s_{j,k} + \epsilon D_j \nabla_j f(\mathbf{q}_k)] = \epsilon D_j s_{j,k} \nabla_j f(\mathbf{q}_k) + (s_{j,k})^2 \leq 0. \quad (3.12)$$

The descent lemma in [30, Sec. 3.3.2] together with the Lipschitz continuity of $f(\cdot)$ entails for

every k

$$\begin{aligned}
f(\mathbf{q}_{k+1}) &= f(\mathbf{q}_k + \mathbf{s}_k) \\
&\leq f(\mathbf{q}_k) + \mathbf{s}_k^T \nabla f(\mathbf{q}_k) + (M/2) \|\mathbf{s}_k\|^2 \\
&\leq f(\mathbf{q}_k) - \left(\frac{1}{\epsilon} - \frac{M}{2} \right) \|\mathbf{s}_k\|^2. \quad [\text{cf. (3.12)}]
\end{aligned}$$

Summing up the inequality over all iterations yields

$$\sum_{k=0}^{\infty} \|\mathbf{s}_k\|^2 \leq \left(\frac{1}{\epsilon} - \frac{M}{2} \right)^{-1} f(\mathbf{q}_0) < \infty,$$

which holds as long as $\left(\frac{1}{\epsilon} - \frac{M}{2} \right)$ is positive. Thus, if $0 < \epsilon < 2/M$, $\|\mathbf{s}_k\|^2$ is summable and the convergence $\lim_{k \rightarrow \infty} s_{j,k} = 0$ holds for every j . And this completes the asymptotic convergence claim for q_k to its fixed point \mathbf{q}^* . ■

Convergence analysis for the asynchronous voltage control updates ensures that the heterogeneity of DERs does not affect the choice of ϵ . As for online implementation, this result can provide guaranteed stability for the proposed control design.

3.3 Dynamic Local Voltage Control

In addition to the asynchronous voltage control updates, the uncertainty in the nominal voltage $\bar{\mathbf{v}}_k$ further challenges the performance of local voltage control. The volatility and intermittence of loads and generations lead to temporal variations in the network operating condition, i.e., a dynamic $\bar{\mathbf{v}}_k$. Thus, it is imperative to analyze the performance of the local voltage control under a dynamic setting.

To this end, the first order autoregressive (AR(1)) process is adapted to model the short-term dynamics.

AS2 (Dynamic Voltage Profile). *For a given constant vector $\bar{\mathbf{c}}$, the nominal voltage $\bar{\mathbf{v}}_k$ follows a*

wide-sense stationary $AR(1)$ process, as given by

$$\bar{\mathbf{v}}_{k+1} = \mathbf{A}\bar{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \bar{\mathbf{c}}, \quad (3.13)$$

where \mathbf{A} is a time-invariant transition matrix with its spectral radius less than 1, while $\boldsymbol{\eta}_{k+1}$ represents a zero-mean white noise process with covariance matrix Σ_η .

The $AR(1)$ model (3.13) can capture both a short-term temporal and spatial correlation of the nominal voltage profile. Its validity has been corroborated by [48] from real data-based tests. Under Assumption 2, for every time k , the nominal voltage $\bar{\mathbf{v}}_k$ has constant mean $\mathbb{E}\bar{\mathbf{v}}_k = (\mathbf{I} - \mathbf{A})^{-1}\bar{\mathbf{c}}$ with a bounded covariance matrix $\Sigma_{\bar{\mathbf{v}}}$ satisfying

$$\Sigma_{\bar{\mathbf{v}}} = \mathbf{A}\Sigma_{\bar{\mathbf{v}}}\mathbf{A}^T + \Sigma_\eta.$$

For ease of exposition, the spatial correlation for power networks is not considered as it is often time negligible [49], corroborated by the structure of inverse of reduced graph Laplacian matrix \mathbf{X} . Since \mathbf{X} is in fact diagonally dominant, the variations in loads and generations tend to have very localized impacts. In addition, we assume an equal variation level in the temporal dynamics across the network; i.e., $\mathbf{A} = \alpha\mathbf{I}$. However, the ensuing analysis holds for the original AR modeling in (3.13) or even higher-order AR modeling as long as it has bounded successive differences. These assumptions simplify the $AR(1)$ model as follows:

$$\bar{\mathbf{v}}_{k+1} = \alpha\bar{\mathbf{v}}_k + \boldsymbol{\eta}_{k+1} + \bar{\mathbf{c}} \quad (3.14)$$

with $\Sigma_\eta = \sigma^2\mathbf{I}$. Accordingly, the stability condition boils down to $|\alpha| < 1$, while $\bar{\mathbf{v}}_k$ the mean $\mathbb{E}\bar{\mathbf{v}}_k = \bar{\mathbf{c}}/(1 - \alpha)$ and the covariance $\Sigma_{\bar{\mathbf{v}}} = \sigma^2/(1 - \alpha^2)\mathbf{I}$. The smaller the value of $|\alpha|$ is, the faster that the nominal voltage $\bar{\mathbf{v}}_k$ evolves.

Proposition 2 (Lemma 1 in [50]) *Under Assumption 2, the expectation of the weighted norm of consecutive difference is bounded, i.e., there exists a bounded constant B_1 such that*

$$\mathbb{E}\|\bar{\mathbf{v}}_{k+1} - \bar{\mathbf{v}}_k\|_{\mathbf{D}}^2 = \frac{2\sigma^2\text{Tr}\mathbf{D}}{1 + \alpha} \leq B_1 \text{ for all } k = 0, 1, \dots \quad (3.15)$$

where the weighted norm $\|\mathbf{v}\|_{\mathbf{D}}^2 := \mathbf{v}^T \mathbf{D} \mathbf{v}$ for any \mathbf{v} .

Under the settings of both dynamic objective and constraint, we formally state the assumption we make for the performance analysis of the gradient projection approach (3.4).

AS3 (Bounded Drift of Optimizer). *The successive difference of the transient optimal solution is bounded, i.e., there exists some bounded constant B_2 such that (see (3.2) for the definition of \mathbf{q}_k^*)*

$$\mathbb{E} \|\mathbf{q}_k^* - \mathbf{q}_{k+1}^*\|_{\mathbf{D}^{-1}}^2 \leq B_2 \text{ for all } k = 0, 1, \dots$$

This assumption is related to the boundedness of voltage drift (see Prop. 2) and the compactness of box constraints. For instance, when the reactive power is unlimited, i.e., $\mathcal{Q}_k = \mathbb{R}^N$, one can easily verify that $\mathbb{E} \|\mathbf{q}_k^* - \mathbf{q}_{k+1}^*\|_{\mathbf{D}^{-1}}^2$ is bounded. When the reactive power is uniformly limited, i.e., \mathcal{Q}_k is always some compact set (double-sided box constraint suffices) for all k , we still have the bounded optimizer drift due to the fact that $\mathbf{q}_k^* \in \mathcal{Q}_k$. Albeit the error bound (stability) we are going to construct will depend on B_2 , intuitively, a smaller voltage drift bound B_1 tends to decrease the drift of the optimizer bound B_2 in power networks.

We first introduce a few quantities to simplify the presentation:

$$\mathbf{y}_k := \mathbf{D}^{-\frac{1}{2}} \mathbf{q}_k, \quad \mathbf{y}_k^* := \mathbf{D}^{-\frac{1}{2}} \mathbf{q}_k^*,$$

$$\mathbf{u}_k := \mathbf{D}^{-\frac{1}{2}} (\bar{\mathbf{v}}_k - \boldsymbol{\mu}),$$

and $\tilde{\mathbb{P}}_k[\cdot]$ is an operator that projects its input onto the set

$$\tilde{\mathcal{Q}}_k := \{\mathbf{q} \mid \mathbf{q} \in [\mathbf{D}^{-\frac{1}{2}} \underline{\mathbf{q}}_k, \mathbf{D}^{-\frac{1}{2}} \bar{\mathbf{q}}_k]\}.$$

This way, the original iterative update in (3.4) is equivalent to

$$\mathbf{y}_{k+1} = \tilde{\mathbb{P}}_k[\mathbf{y}_k - \epsilon(\tilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k)], \quad (3.16)$$

which can be viewed as the standard gradient projection update for the following dynamic con-

strained optimization problem:

$$\min_{\mathbf{y} \in \tilde{\mathcal{Q}}_k} \tilde{f}_k(\mathbf{y}) := \frac{1}{2} \|\mathbf{P}\mathbf{y} + (\mathbf{P}^T)^{-1}\mathbf{u}_k\|^2, \quad (3.17)$$

where \mathbf{P} is obtained by the Cholesky factorization for the symmetric PD matrix $\tilde{\mathbf{X}} = \mathbf{P}^T \mathbf{P}$. Correspondingly, \mathbf{y}_k^* is the optimizer of (3.17). Also, we denote

$$C := \min_{k \geq 0; \mathbf{y} \in \mathbb{R}^N} \left\{ \lambda_{\min} \{ \nabla^2 \tilde{f}_k(\mathbf{y}) \} \right\} = \lambda_{\min} \{ \tilde{\mathbf{X}} \},$$

which is the smallest eigenvalue of $\tilde{\mathbf{X}}$. We say a differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is strongly convex with some positive constant c if for any x and y , we have $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{c}{2} \|x\|^2$ where $\langle \cdot, \cdot \rangle$ is the inner product of two vectors. A differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is gradient Lipschitz continuous with some positive constant m if for any x and y , we have $f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{m}{2} \|x\|^2$. Note that, under this definition, C also serves as the least strong convexity constant of $\tilde{f}_k(\mathbf{y})$, $\forall k$, while M also serves as the greatest gradient Lipschitz continuity constant of $\tilde{f}_k(\mathbf{y})$, $\forall k$. A vector v is called a subgradient of a convex function $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ at point $x \in \mathcal{X}$ if $f(y) \geq f(x) + \langle v, y - x \rangle$ for any $y \in \mathcal{X}$. The set of all subgradients at x is called the subdifferential at x . We use $\tilde{\nabla} \tilde{f}_k(\mathbf{y})$ and $\partial \tilde{f}_k(\mathbf{y})$ to denote a subgradient and the subdifferential of the function \tilde{f}_k at \mathbf{y} , respectively. These notations were also used in [51]. The subgradient used in the algorithm or analysis will be specified in the context, and our analysis will be based on the equivalent update and optimization problem in (3.16) and (3.17), respectively. The following lemma gives the first-order optimality condition of (3.17) and an equivalent recursive relation of (3.16).

Lemma 1 (First-Order Optimality Condition and Recursive Relation). *The instantaneous optimizer \mathbf{y}_k^* to the dynamic optimization problem (3.17) and iterates \mathbf{y}_k satisfy the following conditions, for $k = 0, 1, \dots$,*

$$\tilde{\mathbf{X}}\mathbf{y}_k^* + \mathbf{u}_k + \tilde{\nabla} g_k(\mathbf{y}_k^*) = 0 \quad (3.18)$$

and

$$\mathbf{y}_{k+1} - \mathbf{y}_k = -\epsilon \left[\tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) + \tilde{\nabla} g_k(\mathbf{y}_{k+1}) - \tilde{\nabla} g_k(\mathbf{y}_k^*) \right] \quad (3.19)$$

where

$$\mathcal{I}_k(\mathbf{y}) = \begin{cases} 0, & \text{if } \mathbf{y} \in \tilde{\mathcal{Q}}_k, \\ +\infty, & \text{if } \mathbf{y} \notin \tilde{\mathcal{Q}}_k \end{cases}$$

is the indicator function of the set $\tilde{\mathcal{Q}}_k$ at time k .

Proof: We first replace the projection operation in (3.16) by a subgradient step featured by the indicator function $\mathcal{I}_k(\cdot)$. By definition, the projection of any $\boldsymbol{\omega}$ to $\tilde{\mathcal{Q}}_k$ equals to

$$\tilde{\mathbb{P}}_k(\boldsymbol{\omega}) = \arg \min_{\mathbf{x}} \epsilon \mathcal{I}_k(\mathbf{x}) + \frac{1}{2} \|\mathbf{x} - \boldsymbol{\omega}\|^2. \quad (3.20)$$

The first-order optimality condition leads to $\epsilon \tilde{\nabla} \mathcal{I}_k(\tilde{\mathbb{P}}_k(\boldsymbol{\omega})) + \tilde{\mathbb{P}}_k(\boldsymbol{\omega}) - \boldsymbol{\omega} = 0$. Thus by letting $\boldsymbol{\omega} = \mathbf{y}_k - \epsilon(\tilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k)$ and $\tilde{\mathbb{P}}_k(\boldsymbol{\omega}) = \mathbf{y}_{k+1}$, we obtain

$$\mathbf{y}_{k+1} = \mathbf{y}_k - \epsilon[\tilde{\mathbf{X}}\mathbf{y}_k + \mathbf{u}_k + \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_{k+1})]. \quad (3.21)$$

Furthermore, by using the indicator function, (3.17) is equivalent to

$$\mathbf{y}_k^* = \arg \min_{\mathbf{y}} \frac{1}{2} \|\mathbf{P}\mathbf{y} + (\mathbf{P}^T)^{-1}\mathbf{u}_k\|^2 + \mathcal{I}_k(\mathbf{y}). \quad (3.22)$$

From the first-order optimality condition of (3.22), we have $\tilde{\mathbf{X}}\mathbf{y}_k^* + \mathbf{u}_k + \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_k^*) = 0$. This along with (3.21) proves the recursive relation (3.19). \blacksquare

Note that the subgradient $\tilde{\nabla} \mathcal{I}_k(\mathbf{y}_{k+1})$ used in (3.21) is well-defined because (i) $\mathbf{y}^{k+1} \in \tilde{\mathcal{Q}}_k$, (ii) $g_k(\cdot)$ is continuous over $\tilde{\mathcal{Q}}_k$, and (iii) the minimum in (3.20) is uniquely attained since $\|\mathbf{x} - \boldsymbol{\omega}\|^2$ is real-valued, strictly convex, and coercive. Using the aforementioned notation, our analysis coincides with those earlier results on nonsmooth optimization; e.g., similar notions and analysis schemes have appeared in [51–53] and references therein. Our main result is as follows:

Theorem 2 (BIBO Stability with Geometric Decaying). *Under Assumption 3, for any step-size*

choice

$$\epsilon \in \left(0, \frac{2}{C+M}\right],$$

the expectation of the weighted tracking error between the local control update \mathbf{q}_k of (3.4) and the instantaneous optimal solution \mathbf{q}_k^* can be bounded by

$$\mathbb{E}\|\mathbf{q}_k - \mathbf{q}_k^*\|_{\mathbf{D}^{-1}}^2 \leq \rho^k \mathbb{E}\|\mathbf{q}_0 - \mathbf{q}_0^*\|_{\mathbf{D}^{-1}}^2 + \frac{1-\rho^k}{1-\rho} \Theta, \quad \forall k \quad (3.23)$$

where the geometric rate $\rho \in (0, 1)$ and Θ is a bounded positive constant gap.

Proof: By the smoothness and convexity of \tilde{f}_k , it follows that [54] (see (3.5) for the definition of M)

$$\begin{aligned} & \frac{CM}{C+M} \|\mathbf{y}_k - \mathbf{y}_k^*\|^2 + \frac{1}{C+M} \|\tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\|^2 \\ & \leq \langle \mathbf{y}_k - \mathbf{y}_k^*, \tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) \rangle. \end{aligned} \quad (3.24)$$

By applying the basic inequality

$$2\langle \sqrt{\beta} \mathbf{a}, \frac{1}{\sqrt{\beta}} \mathbf{b} \rangle \leq a \|\mathbf{a}\|^2 + \frac{1}{a} \|\mathbf{b}\|^2,$$

which holds for any $\beta > 0$ and any real vectors \mathbf{a} and \mathbf{b} of the same dimension, the right-hand-side of (3.24) can be upper bounded by

$$\begin{aligned} & \langle \mathbf{y}_k - \mathbf{y}_{k+1} + \mathbf{y}_{k+1} - \mathbf{y}_k^*, \tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) \rangle \\ & \leq \frac{C+M}{4} \|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \frac{1}{C+M} \|\tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*)\|^2 \\ & \quad + \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) \rangle. \end{aligned} \quad (3.25)$$

Substituting (3.25) into (3.24) leads to

$$\begin{aligned} & \frac{CM}{C+M} \|\mathbf{y}_k - \mathbf{y}_k^*\|^2 \\ & \leq \frac{C+M}{4} \|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) \rangle. \end{aligned} \quad (3.26)$$

Since the indicator function $\mathcal{I}_k(\cdot)$ is convex due to the fact that $\tilde{\mathcal{Q}}_k$ is a convex set in our settings, its subgradient $\tilde{\nabla} \mathcal{I}_k(\cdot)$ (subdifferential $\partial \mathcal{I}_k(\cdot)$) is a (set-valued) monotone mapping (this can also

be obtained from the subgradient inequality [53]), i.e.,

$$\langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_{k+1}) - \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_k^*) \rangle \geq 0. \quad (3.27)$$

Combining (3.26) and (3.27) we have

$$\begin{aligned} & \frac{C+M}{4} \|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 \\ & + \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) + \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_{k+1}) - \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_k^*) \rangle \\ & \geq \frac{CM}{C+M} \|\mathbf{y}_k - \mathbf{y}_k^*\|^2. \end{aligned} \quad (3.28)$$

Substituting (3.19) into (3.28) for $\tilde{\mathbf{X}}(\mathbf{y}_k - \mathbf{y}_k^*) + \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_{k+1}) - \tilde{\nabla} \mathcal{I}_k(\mathbf{y}_k^*)$ leads to

$$\begin{aligned} & \frac{C+M}{4} \|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2 + \frac{1}{\epsilon} \langle \mathbf{y}_{k+1} - \mathbf{y}_k^*, \mathbf{y}_k - \mathbf{y}_{k+1} \rangle \\ & \geq \frac{CM}{C+M} \|\mathbf{y}_k - \mathbf{y}_k^*\|^2. \end{aligned}$$

Using the equality $\langle \mathbf{y}_k^* - \mathbf{y}_{k+1}, \mathbf{y}_{k+1} - \mathbf{y}_k \rangle = \|\mathbf{y}_k^* - \mathbf{y}_k\|^2 - \|\mathbf{y}_k^* - \mathbf{y}_{k+1}\|^2 - \|\mathbf{y}_{k+1} - \mathbf{y}_k\|^2$ to expand the inner product, we have

$$\begin{aligned} \|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2 & \leq \left(1 - \frac{2\epsilon CM}{C+M}\right) \|\mathbf{y}_k - \mathbf{y}_k^*\|^2 \\ & \quad + \left(\frac{\epsilon M + \epsilon L}{2} - 1\right) \|\mathbf{y}_k - \mathbf{y}_{k+1}\|^2. \end{aligned} \quad (3.29)$$

By choosing $\epsilon \leq \frac{2}{C+M}$ to ensure the second term on the right-hand-side of (3.29) being nonnegative, the inequality (3.29) can be further relaxed to

$$\|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2 \leq \left(1 - \frac{2\epsilon CM}{C+M}\right) \|\mathbf{y}_k - \mathbf{y}_k^*\|^2. \quad (3.30)$$

By applying another basic inequality

$$\|\mathbf{a}' + \mathbf{b}'\|^2 \leq (1 + \beta') \|\mathbf{a}'\|^2 + \left(1 + \frac{1}{\beta'}\right) \|\mathbf{b}'\|^2,$$

which holds for any $\beta' > 0$ and any real vectors \mathbf{a}' and \mathbf{b}' of the same dimension, we have

$$\begin{aligned}
& \|\mathbf{y}_{k+1} - \mathbf{y}_k^* + \mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2 \\
& \leq (1 + \beta')\|\mathbf{y}_{k+1} - \mathbf{y}_k^*\|^2 + (1 + \frac{1}{\beta'})\|\mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2 \\
& \leq \rho\|\mathbf{y}_k - \mathbf{y}_k^*\|^2 + (1 + \frac{1}{\beta'})\|\mathbf{y}_k^* - \mathbf{y}_{k+1}^*\|^2,
\end{aligned} \tag{3.31}$$

where $\rho := (1 + \beta')(1 - \frac{2\epsilon CM}{C+M})$ while the second inequality comes from (3.30). Let us denote $\Theta := (1 + \frac{1}{\beta'})B_2$, and thus taking expectation on both sides of (3.31) gives us [cf. Assumption 3]

$$\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2 \leq \rho\mathbb{E}\|\mathbf{y}_k - \mathbf{y}_k^*\|^2 + \Theta. \tag{3.32}$$

Applying recursive induction on (3.32), we eventually obtain

$$\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2 \leq \rho^{k+1}\mathbb{E}\|\mathbf{y}_0 - \mathbf{y}_0^*\|^2 + \frac{1 - \rho^{k+1}}{1 - \rho}\Theta,$$

which recovers (3.23) by the definition of \mathbf{y}_{k+1} and \mathbf{y}_{k+1}^* . It shows that as long as $\rho \in (0, 1)$ and $\Theta \in [0, \infty)$, $\mathbb{E}\|\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^*\|^2$ is bounded for all k . Note that the choice of β' can be arbitrarily close to 0. Hence ρ can always achieve a value that is less than 1 as long as $\epsilon > 0$ and $C > 0$ (a simple choice to demonstrate this is $\beta' = \epsilon CM / (C + M - 2\epsilon CM)$). Finally, we conclude that the step-size condition is $0 < \epsilon \leq \frac{2}{C+M}$. ■

Theorem 2 establishes that the tracking error of the local control update (3.4) under dynamical settings exponentially decreases until a constant error bound is reached. Moreover, the AR(1) process assumed to model the $\bar{\mathbf{v}}_k$ series can be potentially extended to a general stochastic process that has bounded iterative changes. This is because the constant Θ in (3.23) is bounded as long as the condition in (3.15) holds. To extract a more specific result, let us choose $\beta' = \epsilon CM / (C + M - 2\epsilon CM)$. In this case, the steady-state ($k \rightarrow \infty$) error is explicitly bounded by

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \frac{1 - \rho^k}{1 - \rho} \Theta \\
& = \frac{1}{1 - (1 + \beta')(1 - \frac{2\epsilon CM}{C+M})} (1 + \frac{1}{\beta'}) B_2 \\
& = \frac{(C+M)(C+M-\epsilon CM)}{(\epsilon CM)^2} B_2.
\end{aligned} \tag{3.33}$$

This error depends on system parameters C , M , step-size ϵ , and the constant B_2 which bounds the successive difference of the instantaneous optimal solutions. It can be seen that the larger step-size would lead to the smaller steady-state error bound. Letting the step-size be $\epsilon = \frac{2}{C+M}$ (best achievable) further yields that the steady-state error does not exceed $\frac{(C+M)^2(C^2+M^2)B_2}{4C^2M^2}$. To sum up, under the settings of both dynamic objective and dynamic constraint, the stable step-size is slightly smaller than the one in the static case (no optimality drift), but under both situations, the achievable step-sizes are on the same order $O\left(\frac{1}{M}\right)$ because $\frac{2}{M} \geq \frac{2}{C+M} \geq \frac{2}{M+M} = \frac{1}{M}$.

Remark 2 (Time-Invariant Box Constraints). For the special case that the box constraints are time-invariant (only objective is time-varying), we can show that the step-size choice to achieve stability is the same to the static case of $\epsilon \in (0, 2/M)$ [50]. This way, the same choice holds for static, dynamic, or asynchronous scenarios. Constant limits on reactive power are the case for photovoltaic inverters if the solar irradiance stays the same during, e.g., nighttime and cloudless scenarios.

3.4 Numerical Tests

We investigate the performance of the decentralized voltage control scheme under the settings of asynchronous update and dynamically time-varying network operating conditions. The desired voltage magnitude μ_j is chosen to be 1 at every bus j . Each bus is assumed to have a certain number of PV panels installed, and thus it is able to control its reactive power via advanced inverter design. All numerical tests are performed in MathWorks® MATLAB 2014a software.

A single-phase radial power distribution network consisting of 21 buses is first used to test the algorithm. This network is equivalent to the system in Fig. 2.1 for $N = 20$. The impedance of each line segment is set to be $(0.233 + j0.366)\Omega$. Hence, the linearized flow equations in (2.4) are only an approximate model. The limit of reactive power resources at every bus is chosen to be $[-100, 100]$ kVA. More realistic test using a 123-bus multi-phase network will be presented later on.

3.4.1 Asynchronous Updates

The impact of asynchronous updates across different buses is first considered under a constant nominal voltage $\bar{\mathbf{v}}_k$. The maximum update delay is set to be $K = 50$. For the 21-bus network, the theoretical upper bound on the step-size is $\epsilon < 2/M = 0.0062$ following Theorem 1. Hence, we set the step-size to be $\epsilon = 1/M = 0.0031$. To model the level of asynchronous updates among multiple buses, we introduce a duty cycle parameter $\eta \in (0, 100\%]$. For a cycle of total $\frac{K}{2}$ time slots, we randomly pick $\lceil \eta \times \frac{K}{2} \rceil$ number of slots for bus j to implement its voltage control update. Hence, the maximum update delay among any two nodes is no more than K . In addition, the larger η is, the more frequently every bus performs an update, and the smaller the effective update delay would be. In particular, the setting of $\eta = 100\%$ provides the benchmark performance of the synchronous scenario where each bus updates at every time slot. Fig. 3.1 plots the iterative voltage

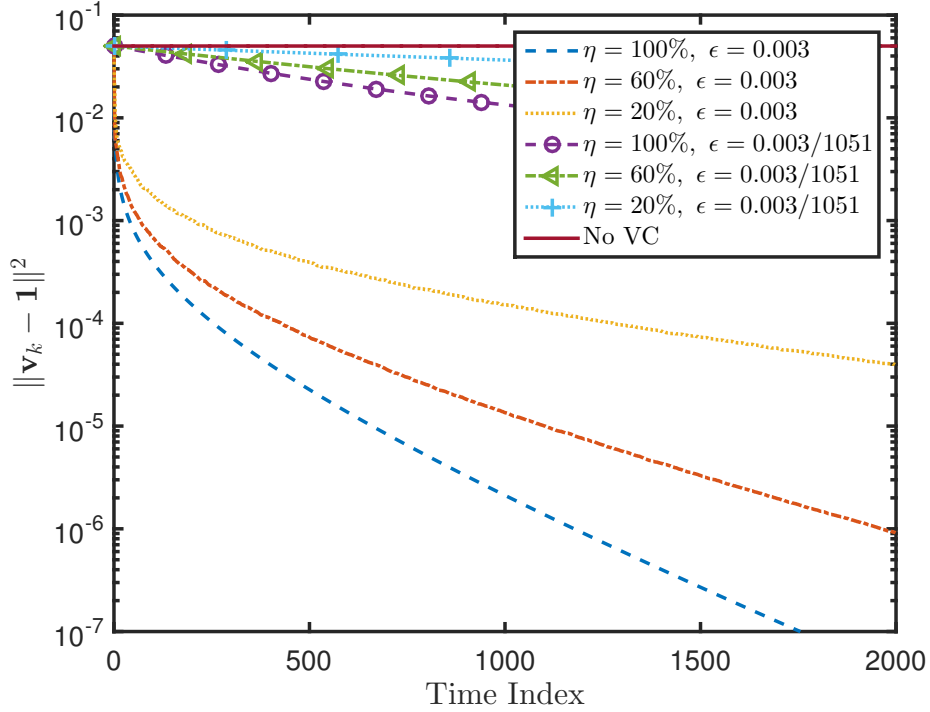


Figure 3.1: Iterative voltage mismatch error performance for the asynchronous decentralized voltage control scheme under various choices of duty cycle η and step-size ϵ .

mismatch error performance for the decentralized voltage control design under different η values and choices of step-size. The case of no voltage control is also plotted with the corresponding error staying constant. Using the classical convergence conditions for asynchronous GP updates

in (3.11), the step-size should be chosen as $\epsilon = 0.0031/1051$ with $N = 20$ and $K = 50$. As shown in Fig. 3.1, this choice of step-size is too conservative. Thus, the resultant convergence speed is much slower than that of the choice $\epsilon = 0.0031$ following Theorem 1. This demonstrates that our theoretical results for asynchronous GP updates are more competitive for the specific decentralized voltage control application here.

Moreover, it is observed that the convergence accuracy would depend on the total number of updates for the whole network. Because of the asynchronous update settings, the expected number of control updates across the network for a cycle of $\frac{K}{2}$ iterations equals to $\lceil \eta \times \frac{K}{2} \rceil \times N$. Fig. 3.2 illustrates the voltage mismatch error performance versus the expected number of total control updates. Interestingly, the convergence speed in this plot is the same for the same step-size ϵ value, regardless of the asynchronous metric η . Hence, the average update rate across all the buses determines the performance of the asynchronous decentralized voltage control scheme.

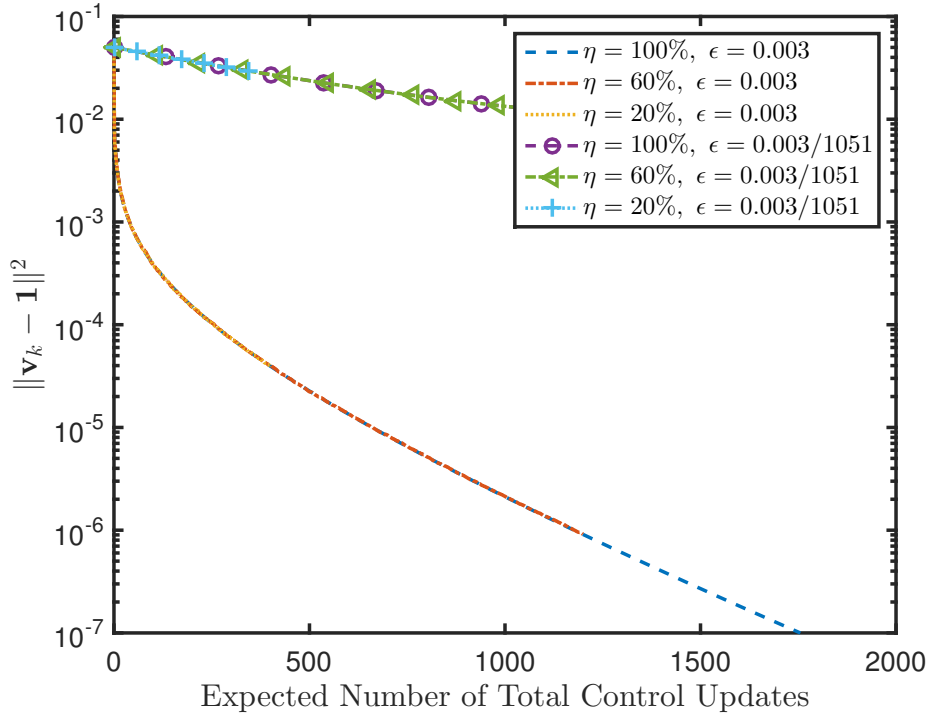


Figure 3.2: Voltage mismatch error versus the total number of updates across the network for the asynchronous voltage control scheme under various choices of duty cycle η and step-size ϵ .

3.4.2 Dynamic Operating Conditions

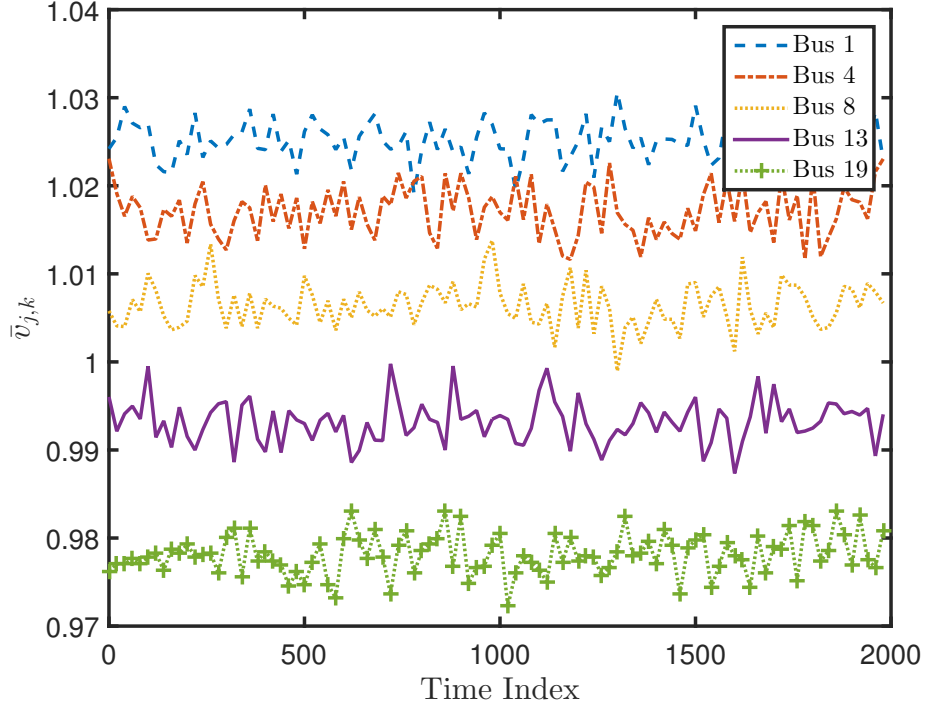


Figure 3.3: An instance of the nominal voltage series $\{\bar{v}_k\}$ at selected buses under the AR(1) model settings.

To verify our results on dynamic voltage control, we generate the nominal voltage series $\{\bar{v}_k\}$ using the AR(1) model in (3.14). Neglecting the effects of voltage regulators, it is well known that the voltage magnitude in power networks tends to decrease monotonically away from the root node, i.e., bus 0. Hence, we set the mean voltage at bus j to be $\bar{c}_j/(1 - \alpha) = 1.025 - \frac{0.05}{19}(j - 1)$ to follow this decreasing voltage rule. Fig. 3.3 plots the nominal voltage sequence at selected network locations for the choice of $\alpha = 0.1$ and noise variance $\sigma^2 = 6 \times 10^{-6}$. This choice of the forgetting factor α value leads to very fast dynamics in the nominal voltage.

The step-size ϵ : Fig. 3.4 plots the iterative voltage mismatch error performance using different choices of ϵ , while Fig. 3.5 plots the weighted tracking error between the iterate \mathbf{q}_k and the corresponding instantaneous optimal \mathbf{q}_k^* . Both plots are averaged over 30 random realizations of the nominal voltage series to approximate the expected values. The maximum value $\epsilon = 0.0061$ is chosen according to the bound $2/(C + M)$ in Theorem 2. As shown more clearly in Fig. 3.5, a larger step-size ϵ leads to slightly faster convergence of the tracking error. However, the steady-

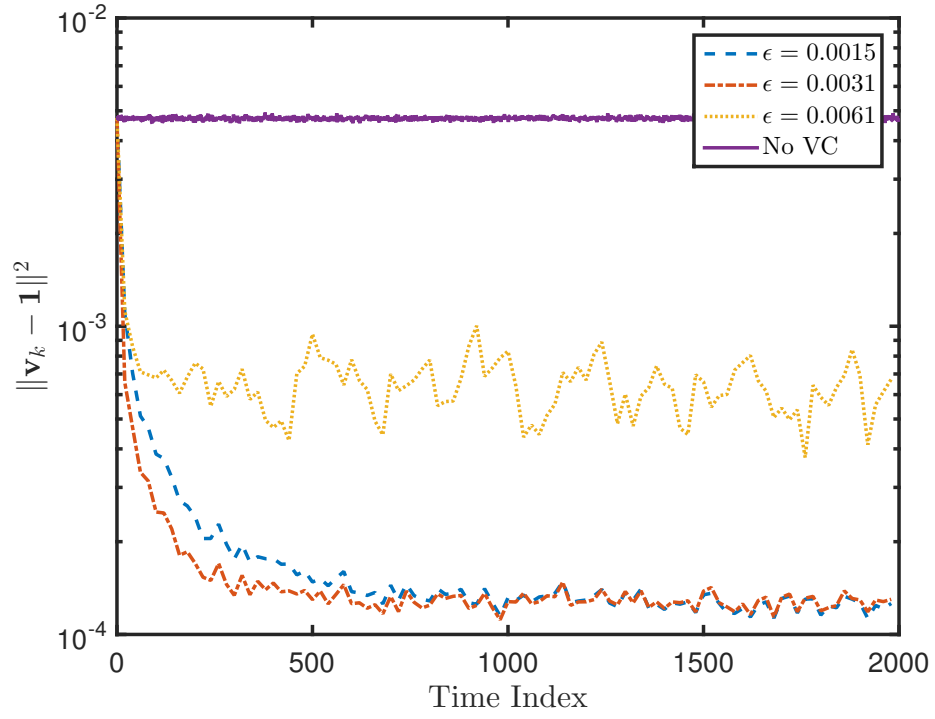


Figure 3.4: Iterative voltage mismatch error performance averaged over 30 random realizations under different ϵ values.

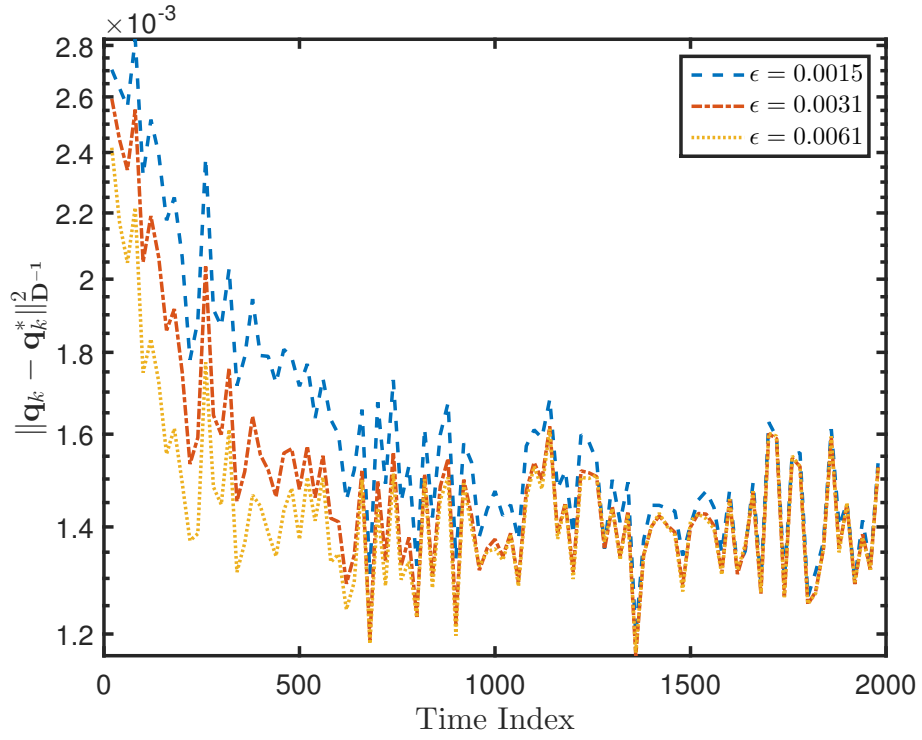


Figure 3.5: Iterative tracking error averaged over 30 random realizations under different ϵ values.

state voltage mismatch error is higher for the largest ϵ as shown in Fig. 3.4. This observation coincides with the analytical results of Theorem 2. The convergence geometric rate ρ depends on an appropriate choice of ϵ , while the steady-state error related constant Θ tends to increase with a larger ϵ choice. Thus, the choice of ϵ would be able to trade the steady-state tracking error off the convergence speed. Under this trade-off, the optimal selection of ϵ would also depend on the dynamics of the AR(1) process. If the nominal voltage evolves very fast, it is preferred to have a large ϵ for a better tracking performance. Otherwise, if the dynamics of the nominal voltage series has a large time constant, we can afford to have a small ϵ in order to achieve a better tracking error performance. By analyzing the bounds in Theorem 2, it is possible to provide a general guideline on selecting a proper ϵ value based on the dynamics of $\bar{\mathbf{v}}_k$.

The forgetting factor α : We have also varied the parameter α for the AR(1) process used to generate the nominal voltage series, with values ranging from 0.1 to 0.999. The step-size ϵ is fixed at 0.0031. For comparison purposes, the variance of the nominal voltage at every bus is aligned to be the same for different α values by setting it to be $\sigma^2/(1 - \alpha^2) = 10^{-5}$. Hence, when $\alpha = 0.999$ very closely approaches 1, the $\{\bar{\mathbf{v}}_k\}$ series would almost stays flat with minimal temporal variations. Accordingly, the consecutive voltage mismatch error bound B_1 in Prop. 2 decreases as α approaches its upper bound 1. Fig. 3.6 plots the iterative voltage mismatch error for various α values, while Fig. 3.7 again plots the corresponding weighted tracking error performance. These curves are also averaged over 30 random realizations. The performance under either mismatch error metrics improves with a larger α value since the constant B_1 and thereby B_2 would decrease. Accordingly, this leads to a smaller Θ value and reduces the steady-state tracking error. This numerical result points out that B_2 , which bounds the optimizer drift, could be related to the consecutive voltage difference B_1 .

The noise variance σ^2 : A similar test has been conducted with varying parameter σ for the AR(1) process ranging from 7.7×10^{-4} to 7.7×10^{-3} . With the same step-size $\epsilon = 0.0031$ and parameter $\alpha = 0.1$, Fig. 3.8 and Fig. 3.9 plot the average voltage mismatch and weighted tracking error over 30 realizations. Similar to the observations under various α values, using a smaller σ^2 would decrease the steady-state error bounds since the constant Θ tends to be positively related to the parameter σ^2 . As the parameter α is fixed, the variance of the nominal voltage at every bus decreases as the noise variance diminishes. Accordingly, the performance of no voltage control

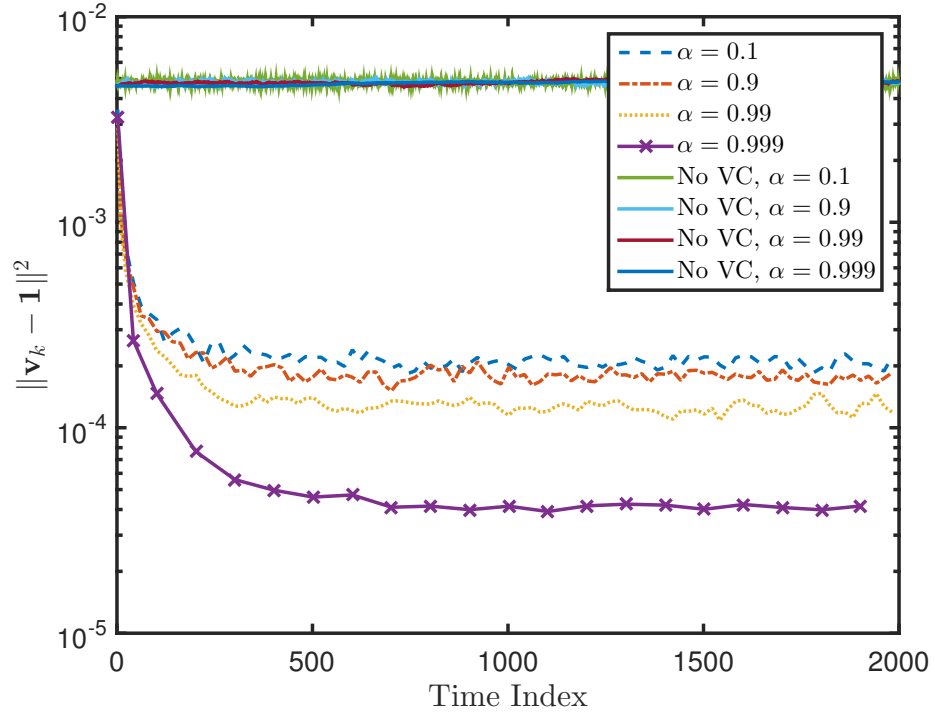


Figure 3.6: Iterative voltage mismatch error performance averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α .

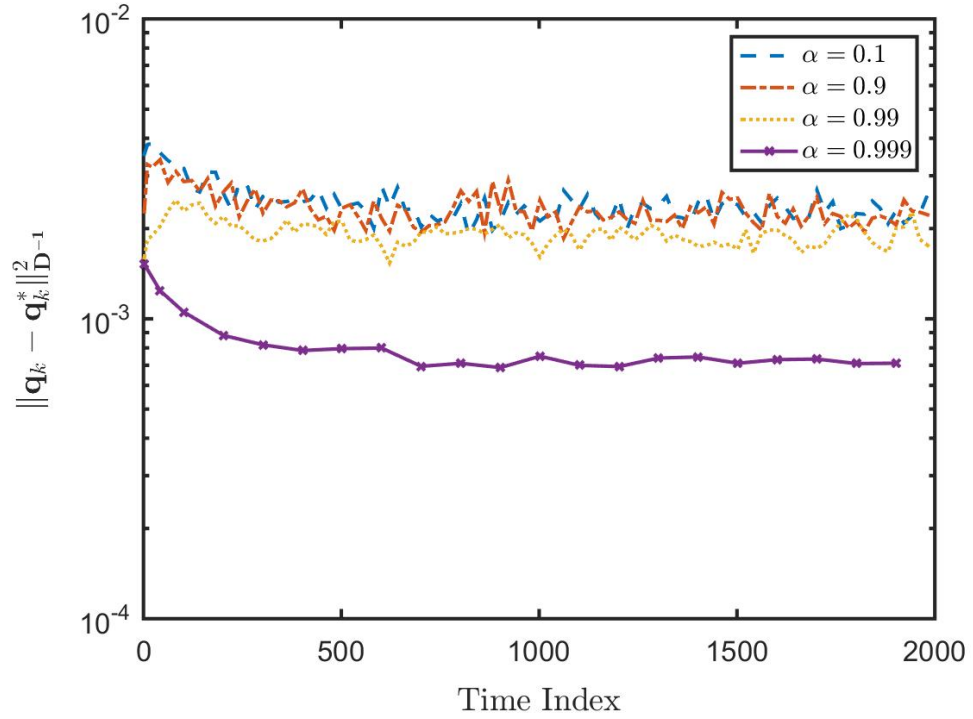


Figure 3.7: Iterative tracking error averaged over 30 random realizations for the voltage control scheme with various values of forgetting factor α .

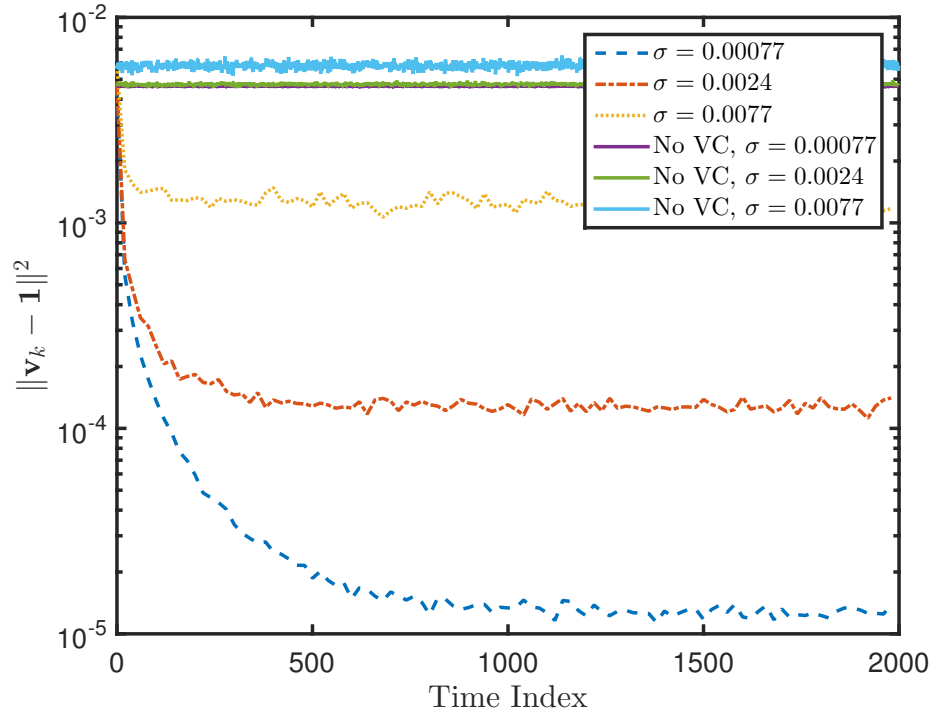


Figure 3.8: Iterative voltage mismatch error performance averaged over 30 random realizations under various σ values.

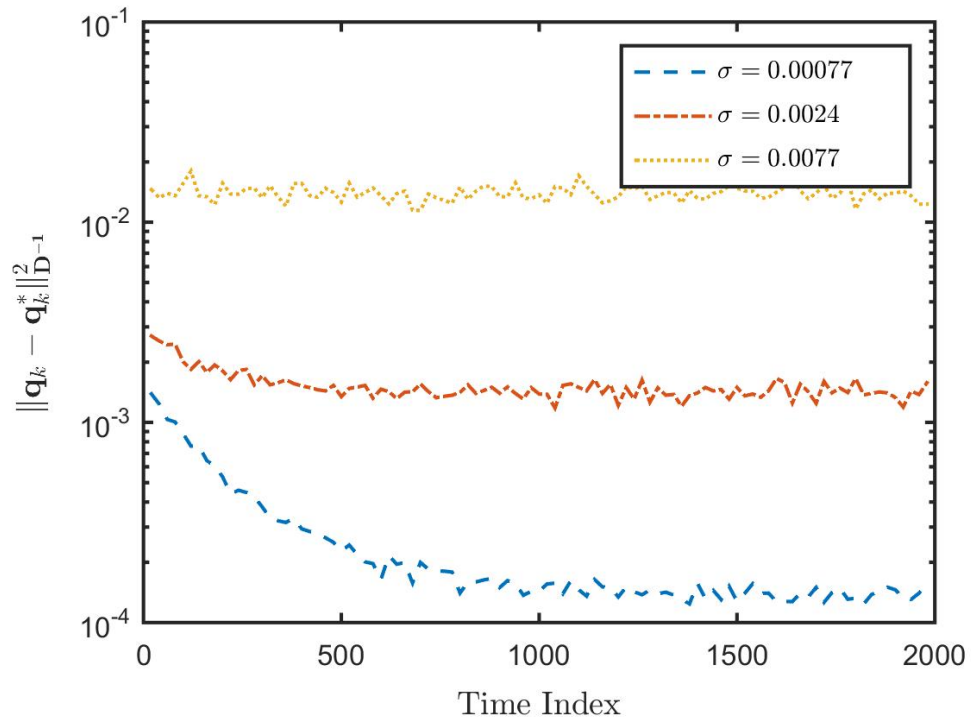


Figure 3.9: Iterative tracking error averaged over 30 random realizations under various σ values.

slightly improves with smaller noise variance.

All numerical results in this test case have verified our analytical bounds on the tracking error performance. To sum up, the convergence speed depends on the choice of step-size ϵ . Depending on the time constant of nominal voltage dynamics, the step-size needs to be properly chosen trading off the convergence speed and the steady-state error performance. The dynamics of nominal voltage series based on the AR(1) process parameters does not affect the convergence speed per se, yet is more significantly related to the steady-state tracking error performance. Note that the analytical bounds of Theorem 2 are not tight, because of the scalar β' used to eliminate the cross-product terms from the squared sum norm. However, they are very effective to characterize the error performance of dynamic decentralized voltage control scheme while facilitating the selection of step-size.

3.4.3 Local Voltage Control under an Unbalanced Multi-Phase Network

We have also tested the decentralized voltage control scheme using a realistic multi-phase distribution network, namely the IEEE 123-bus system model [55]. This test case incorporates both the asynchronous updates among multiple buses and dynamical nominal voltage, with similar settings as in the earlier two test cases. Moreover, the line segments of the 123-bus system involve inter-phase mutual couplings. Hence, this test provides an accurate representation of how the decentralized voltage control would perform in practice with uncertainties in the DER hardware and network operating conditions. Fig. 3.10 plots the iterative network voltage mismatch error under various choices of duty cycle parameter η . The step-size $\epsilon = 0.01$ has been chosen for every scenario to ensure stability. Different from the earlier two test cases, the control implementation has incorporated both the asynchronous updates and the dynamic voltage profile. Although we have not been able to derive the tracking error bounds under both sources of uncertainty, Fig. 3.10 demonstrates that its convergence speed result is similar to the solely asynchronous case as in Fig. 3.1, while the steady-state tracking error may have similar bounds as in the dynamic control analysis. In addition, Fig. 3.11 illustrates the voltage mismatch error performance versus the expected number of total control updates. Similar to its single-phase counterpart, the convergence speed in

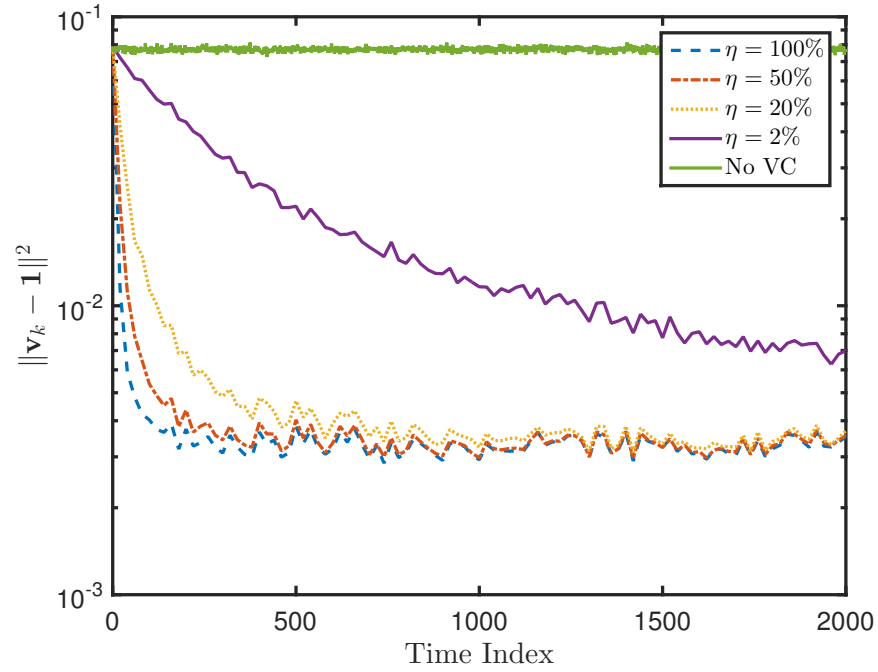


Figure 3.10: Iterative voltage mismatch error performance on the IEEE 123-bus system under both the asynchronous updates and the dynamic network operating conditions with different η values.

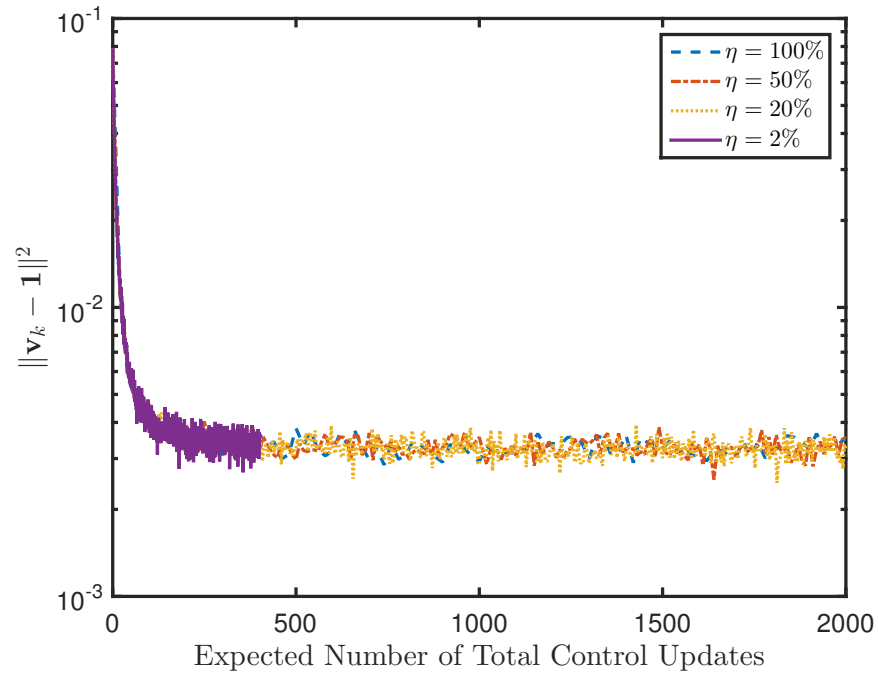


Figure 3.11: Voltage mismatch error versus the total number of updates across the IEEE 123-bus system for the asynchronous voltage control scheme under dynamic operating conditions with various choices of duty cycle η .

this plot is analogous for a fixed step-size ϵ value, regardless of the asynchronous metric η . Thus, we are confident that the analysis of this chapter can be integrated into a joint framework that characterizes the tracking error performance under both uncertain sources.

3.5 Summary

In this chapter, we have developed a local dynamic optimization framework for analyzing the performance of a voltage control scheme based on gradient-projection (GP) methods for online system implementations. By constructing the linearized flow model for power distribution networks, one can design a voltage control scheme by minimizing a surrogate voltage mismatch error using the GP iterations. Thanks to the physical power network coupling, this GP-based scheme boils down to a local voltage control design where every bus can measure its local voltage to obtain the instantaneous gradient direction. Compared to earlier results for a static optimization scenario, we have significantly extended the analysis on convergence conditions and error performance to account for two dynamic scenarios: i) the nodes perform the local update in an asynchronous fashion; and ii) the network operating point is dynamically changing. Assuming the nominal voltage evolves following an AR(1) process, the weighted tracking error can be bounded by an exponentially decreasing term plus a constant term that would depend on the successive difference of the transient optimal solution. Interestingly, the choice of step-size may need to be more conservative depending on the trade-off between the convergence speed and the steady-state tracking error for the dynamic control design. Several numerical tests have been performed to demonstrate and validate our analytical results on the performance of the local voltage control scheme under realistic dynamic scenarios using practical power network models.

CHAPTER 4

ONLINE DISTRIBUTED VOLTAGE CONTROL

To engage heterogeneous DERs for voltage regulation, a key challenge is to design efficient and robust schemes that can coordinate distributed inverters at minimal sensing and communication complexity. Albeit the voltage control problem can be formulated using an optimal power flow (OPF) approach that minimizes certain network-wide operational objective, e.g., the total voltage mismatch error [25] (the unweighted version of (3.2)), such a centralized framework fails to account for the limited and low-quality communication infrastructure currently deployed in distribution networks. Chap. 3 addresses this issue by adopting a voltage control design which only relies on locally available voltage information. This local design, already standardized in [56], may suffer from degraded performance due to lack of information exchange, especially when VAR resources are limited [3, 17]. There also exist several distributed designs by leveraging information exchange among only neighboring nodes (see, e.g., [13, 15]). Nonetheless, they have been developed under a static optimization framework and do not directly apply to an online control implementation. In addition, none of the existing work has considered imperfect communication regimes, which are foreseen to be the reality for distribution networks in the near future.

In this chapter, to address the aforementioned issues, we develop a distributed voltage control (DVC) design that seamlessly adapts to dynamically varying system operating conditions. Accounting for practical communication constraints in bandwidth and network capacity, we further make the proposed DVC design to be robust against random link failures. To this end, we formulate the centralized problem using the linearized approximation modeling of both single- and multi-phase distribution networks (2.5) and (2.12) [3, 12, 15]. To solve this convex quadratic problem of box constraints, a distributed solver adopted from the alternating direction method of multipliers (ADMM) [57] is offered to tackle the coupling of neighboring bus voltage variables. This DVC design is further improved for online dynamic implementations under imperfect communications.

Although the linearized power flow model has been used to develop and analyze the proposed algorithms, we have tested and validated the performance using the non-linear multi-phase systems with line losses and phase coupling.

Compared with existing work on (distributed) voltage regulation, our main contributions are three-fold. First, the proposed DVC enjoys minimal sensing and communication requirements by measuring and sharing voltage magnitude among neighboring nodes only. The open-loop distributed algorithms in [13, 15] require all the variables to be stored and exchanged among neighbors (i.e., real and reactive line power flows), whereas our ADMM-based DVC scheme effectively incorporates only the voltage measurements as the control feedback signal. Thus, our design can cope with fast system dynamics and reduces both sensing and communication overhead. Second, most of the previous work in distributed voltage control has not considered cyber resource constraints, e.g. communication link failures. We further leverage the work by [58, 59] to develop an asynchronous ADMM-based solver using the freezing strategy. Accordingly, our approach is robust against random communication link failures with guaranteed optimality and stability for real-time implementations. Last but not least, numerical tests have corroborated that the feedback-based scheme can effectively mitigate the impacts of modeling mismatches [47, Sec. 8.9] introduced by the linearized power flow model. In practical implementations, this feedback signal also provides the most up-to-date system operating conditions. Thus, our algorithm can be implemented in an online fashion whereas the static problem formulation in [13, 15] needs further modifications to account for system dynamics. Clearly, with its features of feedback control design and asynchronous updates, our proposed DVC approach is very related to a competing solution recently developed in [60]. Nonetheless, the VAR constraint limits in [60] have been relaxed using Lagrangian multipliers, and thus the iterative VAR control signal could become infeasible with respect to the resource limits. Our proposed DVC has successfully tackled this issue by projecting the VAR iterates to achieve feasible instantaneous control signal.

4.1 Distributed Voltage Control Problem

Voltage control aims to manage the output \mathbf{q}^g from local inverters to maintain a desired voltage profile, i.e., $\mathbf{v} \rightarrow \boldsymbol{\mu}$. We formulate the problem under a static setup to develop distributed solvers. Note that \mathbf{q}^g is replaced by \mathbf{q} in the rest of the thesis; hence, the single-phase problem is formulated as

$$\{\mathbf{v}^*, \mathbf{q}^*\} := \arg \min_{\mathbf{v}, \mathbf{q}} f(\mathbf{v}, \mathbf{q}) := \frac{1}{2} \sum_{j=1}^N \left\{ (v_j - \mu_j)^2 + c_j q_j^2 \right\} \quad (4.1a)$$

$$\text{subject to } \sum_{i \in \mathcal{N}_j} B_{ji} v_i = q_j + w_j \quad (4.1b)$$

$$\underline{q}_j \leq q_j \leq \bar{q}_j, \quad \forall j \in \mathcal{N}, \quad (4.1c)$$

where the subset $\mathcal{N}_j \subseteq \mathcal{N}$ contains bus j and all neighboring buses. The coefficient $c_j \geq 0$ is chosen for regularizing q_j . The squared voltage mismatch error norm in (4.1a) is effective for penalizing a high level mismatch of $|v_j - \mu_j|$. This quadratic error objective can also be generalized to more robust error criterion such as the Huber loss function; see e.g., [45, Ch. 7]. Additionally, due to the sparse structure of \mathbf{B} , (4.1b) is equivalent to the power flow constraints in (2.5). Coupling involving only the neighborhood \mathcal{N}_j in (4.1b) is instrumental for distributed control design in Sec. 4.2. Last, the bound constraints in (4.1c) are either due to the inverter apparent power limit or depend on certain inverter power factor limits as depicted in Fig. 2.2 [29].

Clearly, if $c_j = 0$ for every j , the objective in (4.1a) is reduced to minimizing the voltage mismatch with no VAR regularization. A positive c_j would sometimes be preferred. The reason is three-fold. First, as shown in [3, 17], the celebrated droop voltage control naturally enforces certain regularization on \mathbf{q} using a positive c_j , which inversely scales with the droop slope. Second, it is possible that a network with larger VAR bus injections may result in a higher level of line current magnitude and accordingly increase the reactive power flow and network-wide power losses. Thus, a positive c_j to regularize q_j could diminish the total power losses. Finally, from a numerical perspective, a larger c_j would benefit the conditioning of the Hessian of f and thus reduce computation time for solving (4.1) [46, Ch. 2]. This feature may prevent the inverter control from adversely affecting the operation of OLTC and other slower time-scale voltage regulating devices.

Therefore, it is attractive when coping with fast dynamics and disturbances.

Remark 3. (*Voltage Regulation Criterion.*) The classical voltage regulation operations aim to maintain the voltage within $\pm 5\%$ of the rated value. This operational objective can be included in our ensuing approach by adding additional constraints $0.95 \leq v_j^\phi \leq 1.05, \forall j, \phi$ to (4.1), assuming the existence of feasible solutions. Interestingly, it turns out that our proposed solution can accommodate such voltage bounds, as detailed soon in Sec. 4.2. As for the squared Euclidean norm error, it is more effective in mitigating the maximum voltage deviation by coordinating network-wide VAR resources, as compared to the traditional paradigm of maintaining the voltage within limits as in voltage regulators. Additionally, this quadratic voltage mismatch error norm in (4.1a) can also effectively mitigate much larger voltage deviation that is potentially $> 5\%$.

Extending (4.1) to include multi-phase networks, we define vector $\mathbf{q} := [\mathbf{q}^a; \mathbf{q}^b; \mathbf{q}^c]$ to concatenate the controllable VAR injection vector \mathbf{q}^ϕ per phase- ϕ , and similarly vector $\mathbf{v} := [\mathbf{v}^a; \mathbf{v}^b; \mathbf{v}^c]$ for the voltage magnitude. One can formulate the following problem to achieve $\mathbf{v} \rightarrow \mathbf{u}$:

$$\{\mathbf{v}^*, \mathbf{q}^*\} := \arg \min_{\mathbf{v}, \mathbf{q}} \quad f(\mathbf{v}, \mathbf{q}) := \frac{1}{2} \|\mathbf{v} - \mathbf{u}\|_2^2 + \frac{1}{2} \mathbf{q}^T \mathbb{C} \mathbf{q} \quad (4.2a)$$

$$\text{subject to} \quad \mathbb{B} \mathbf{v} = \mathbf{q} + \mathbf{w} \quad (4.2b)$$

$$\underline{\mathbf{q}} \leq \mathbf{q} \leq \overline{\mathbf{q}}, \quad (4.2c)$$

where the diagonal matrix \mathbb{C} contains non-negative coefficients for regularizing \mathbf{q} . The multi-phase problem (4.2) has the same features as (4.1), and both are convex quadratic programming (QP) problems. As matrix $\mathbb{B} := [\mathbf{B}^a; \mathbf{B}^b; \mathbf{B}^c]$ has a block sparse structure that couples only neighboring buses, the control designs developed for (4.1) are also expected to apply for (4.2). Hence, this chapter focuses on presenting voltage control schemes developed for single-phase networks. These can be easily generalized for multi-phase networks; see Remark 4.

Both problems (4.1) and (4.2) are convex quadratic programming (QP) ones, efficiently solvable once all the information on the system model and operating conditions is centrally available. However, this centralized framework will be challenged by the increasing variability of DERs. The ensuing section will develop fast distributed solvers only requiring information exchange among neighboring buses.

4.2 ADMM-based Distributed Solver

To overcome the issues involving centralized control designs, we exploit the decomposable structure of (4.1) to solve it in a distributed fashion. Specifically, we propose to adopt the alternated-direction method-of-multipliers (ADMM), a distributed optimization technique that has been effectively applied to a variety of scientific disciplines including signal processing, statistical learning, and more recently, power systems (see e.g., [11, 15, 57]). Based on ADMM, we design a distributed voltage control (DVC) scheme that needs only to measure and incorporate the dynamic voltage magnitude. This feedback-based design is quite different from existing approaches of distributed operations in power systems, developed as a static optimization problem [11, 15]. In addition to minimal sensing overhead, the voltage-feedback control design is likely to be more robust to mismatch and imperfection in system modeling and implementation (see e.g., [47, Sec. 8.9]). It is also robust to imperfect bus-to-bus communication link as established in Sec. 4.4. Specially, ADMM requires much simpler communication architecture than that of a centralized framework since the optimization problem is divided into sub-area size instead of the full network size. To this end, we first reformulate the problem (4.1) to be suitable for decomposition with consensus version of ADMM. It turns out that the resultant algorithm will introduce simple local update rules at each bus. Consider first the static setup in which bus j has the value w_j available. The dynamic scenario with time-varying w_j will be considered in Sec. 4.3. To solve (4.1) using ADMM, define two groups of variables: $\mathbf{x} := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ and \mathbf{z} as

$$\begin{aligned}\mathbf{x}_j &:= \left\{ \{\nu_i^j\}_{i \in \mathcal{N}_j}, \delta_j \right\}, \quad \forall j \in \mathcal{N}, \\ \mathbf{z} &:= \{v_j, q_j\}_{j \in \mathcal{N}}.\end{aligned}$$

Clearly, the vector \mathbf{z} includes all the optimization variables of (4.1), termed the *decision* variables. Each *auxiliary* variable \mathbf{x}_j contains a local copy of some decision variables in \mathbf{z} relevant to bus j . For example, the auxiliary variable ν_i^j at bus j corresponds to v_i in \mathbf{z} while δ_j corresponds to q_j . By enforcing the auxiliary and decision copies of each variable to consent, one can reformulate

(4.1) as

$$\begin{aligned}
\min_{\{\mathbf{x}_j \in \mathcal{X}_j\}_{j \in \mathcal{N}}, \mathbf{z} \in \mathcal{Z}} \quad & \sum_{j=1}^N f_j(\mathbf{x}_j) := \frac{1}{2} \sum_{j=1}^N \{(\nu_j^j - \mu_j)^2 + c_j \delta_j^2\} \\
\text{subject to} \quad & \lambda_{ji} : \nu_i^j = v_i, \forall j, i \in \mathcal{N}_j \\
& \theta_j : \delta_j = q_j, \forall j \in \mathcal{N},
\end{aligned} \tag{4.3}$$

where the constraint sets are $\mathcal{X}_j := \{\mathbf{x}_j | \sum_{i \in \mathcal{N}_j} B_{ji} \nu_i^j - \delta_j = w_j\}$ and $\mathcal{Z} := \{(v_j, q_j)_{j \in \mathcal{N}} | \underline{q}_j \leq q_j \leq \bar{q}_j\}$. Each equality with the corresponding multiplier denoted in (4.3) represents a consensus constraint for a pair of auxiliary and decision variables, which make (4.1) equivalent to (4.3).

The augmented Lagrangian function for (4.3) with a given parameter $\rho > 0$ is $\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{y}) = \sum_{j=1}^N \mathcal{L}_j(\mathbf{x}_j, \mathbf{z}, \mathbf{y})$ where

$$\begin{aligned}
\mathcal{L}_j(\mathbf{x}_j, \mathbf{z}, \mathbf{y}) = & \frac{1}{2}(\nu_j^j - \mu_j)^2 + c_j \delta_j^2 + \sum_{i \in \mathcal{N}_j} \lambda_{ji}(\nu_i^j - v_i) \\
& + \theta_j(\delta_j - q_j) + \frac{\rho}{2} \sum_{i \in \mathcal{N}_j} (\nu_i^j - v_i)^2 + \frac{\rho}{2}(\delta_j - q_j)^2,
\end{aligned} \tag{4.4}$$

and $\mathbf{y} := \{\{\lambda_{ji}\}_{i \in \mathcal{N}_j}, \theta_j\}_{j \in \mathcal{N}}$ containing all multipliers. Note that \mathcal{L} is different from the typical Lagrangian function; all equality constraints are regularized by quadratic terms. The additional term does not affect the optimality conditions, since it is essentially zero at any feasible solution to (4.3). Based on the augmented Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{y})$, the ADMM works by cyclically minimizing each of the three variable groups, \mathbf{x} , \mathbf{z} , and \mathbf{y} , while fixing others. Its k -th iteration involves the following three steps:

(S1) Update \mathbf{x} : For given $\mathbf{z}[k]$ and $\mathbf{y}[k]$, the augmented Lagrangian \mathcal{L} totally decouples into \mathcal{L}_j for each \mathbf{x}_j . Hence, the local auxiliary variable \mathbf{x}_j can be updated as

$$\mathbf{x}_j[k+1] := \arg \min_{\mathbf{x}_j \in \mathcal{X}_j} \mathcal{L}_j(\mathbf{x}_j, \mathbf{z}[k], \mathbf{y}[k]), \tag{4.5}$$

and this subproblem is a QP with efficient convex solvers. Closed-form solutions can be

obtained under the single-phase setup by re-writing (4.5) as

$$\mathbf{x}_j[k+1] := \arg \min_{\mathbf{h}_j^T \mathbf{x}_j = w_j} \frac{1}{2} \mathbf{x}_j^T \mathbf{A}_j \mathbf{x}_j + \mathbf{b}_j^T[k] \mathbf{x}_j, \quad (4.6)$$

where \mathbf{A}_j is a diagonal and non-singular matrix, $\mathbf{b}_j[k]$ is obtained from $\mathbf{z}[k]$ and $\mathbf{y}[k]$, and \mathbf{h}_j relates to a single linear constraint in \mathcal{X}_j . Upon defining $t_j[k] := (w_j + \mathbf{h}^T \mathbf{A}^{-1} \mathbf{b}[k]) / (\mathbf{h}^T \mathbf{A}^{-1} \mathbf{h})$, the local update (4.5) becomes

$$\begin{aligned} \nu_j^j[k+1] &= \frac{1}{\rho+2} (t_j[k] B_{jj} + 2\mu_j - \lambda_{jj}[k] + \rho v_j[k]), \\ \nu_i^j[k+1] &= \frac{1}{\rho} (t_j[k] B_{ji} - \lambda_{ji}[k] + \rho v_i[k]), \quad \forall i \neq j, \\ \delta_j[k+1] &= \frac{1}{\rho+2w_j} (-t_j[k] - \theta_j[k] + \rho q_j[k]). \end{aligned}$$

(S2) Update \mathbf{z} : Likewise, the decision variables are updated as

$$\mathbf{z}[k+1] := \arg \min_{\mathbf{z} \in \mathcal{Z}} \mathcal{L}(\mathbf{x}[k+1], \mathbf{z}, \mathbf{y}[k]), \quad (4.7)$$

which is a box-constrained QP problem, solved by projecting the unconstrained solutions to \mathcal{Z} [46, Ch. 2]. When all multipliers are initialized to zero by setting $\mathbf{y}[1] = \mathbf{0}$, the summand $\sum_{i \in \mathcal{N}_j} \lambda_{ij}[k]$ remains zero throughout the ADMM iterations. Now, by setting the gradient of (4.7) to zero, (4.7) decouples into the following updates per bus j :

$$v_j[k+1] = \frac{1}{|\mathcal{N}_j|} \sum_{i \in \mathcal{N}_j} \nu_j^i[k+1], \quad (4.8a)$$

$$q_j[k+1] = \mathbb{P}_j [\theta_j[k] / \rho + \delta_j[k+1]], \quad (4.8b)$$

where \mathbb{P}_j projects any input to the interval $[\underline{q}_j, \bar{q}_j]$.

(S3) Update \mathbf{y} : Each multiplier is linearly updated per bus j using the iterative mismatch:

$$\begin{aligned}\lambda_{ji}[k+1] &= \lambda_{ji}[k] + \rho(\nu_i^j[k+1] - v_i[k+1]), \forall i \in \mathcal{N}_j, \\ \theta_j[k+1] &= \theta_j[k] + \rho(\delta_j[k+1] - q_j[k+1]).\end{aligned}\tag{4.9}$$

Thus, the summand $\sum_{i \in \mathcal{N}_j} \lambda_{ij}[k+1]$ is guaranteed to stay zero if properly initialized. This corroborates the derivations for (4.8).

The ADMM iterations in (S1-S3) constitute the basis for our proposed distributed voltage control design. Its optimality can be established using the linear convergence results for general ADMM solvers in [61]. Since each local objective $f_j(\mathbf{x}_j)$ in (4.3) is quadratic, it is closed, proper, and strongly convex, while its gradient is also Lipschitz continuous, leading to Proposition 3.

Proposition 3. *Under fixed \mathbf{w} and for any initialization $\{\mathbf{z}[1], \mathbf{y}[1]\}$, the ADMM iterates given by (S1-S3) converge to the optimal solutions of (4.3) at a rate of at worst $1/k$. Thus, the iterates $\mathbf{q}[k]$ converge to the decision optimal VAR solution \mathbf{q}^* in (4.1) because the two problems are equivalent.*

Remark 4. (*Multi-Phase Networks.*) The ADMM-based voltage control algorithm in (S1-S3) can be readily generalized to solve the multi-phase network problem (4.2). Each scalar variable per bus j in the single-phase problem must be converted to a 3×1 vector containing all three-phase variables. For example, the voltage variable v_j in (4.1) becomes the vector $\tilde{\mathbf{v}}_j$ in (2.6). Hence, the *block* sparse structure of \mathbb{B} allows for a local update of $\mathbf{x}_j[k+1]$ by solving a linearly constrained QP. The update of $\mathbf{z}[k+1]$ is again a box-constrained QP which follows the same computing rules in (4.8b). Thus, it suffices to replace all scalar variables in the single-phase problem by their corresponding vector counterparts to extend (S1-S3) to a multi-phase problem. The optimality and convergence claims in Proposition 3 still hold. See Appendix A for detailed derivations.

4.3 Online Voltage Control Design

The ADMM iterations in (S1-S3) assume the availability of w_j at every bus j , which depends on system operating conditions. If full information on complex power injection becomes available, it is possible to compute \mathbf{w} based on (2.5). However, such a centralized approach would require

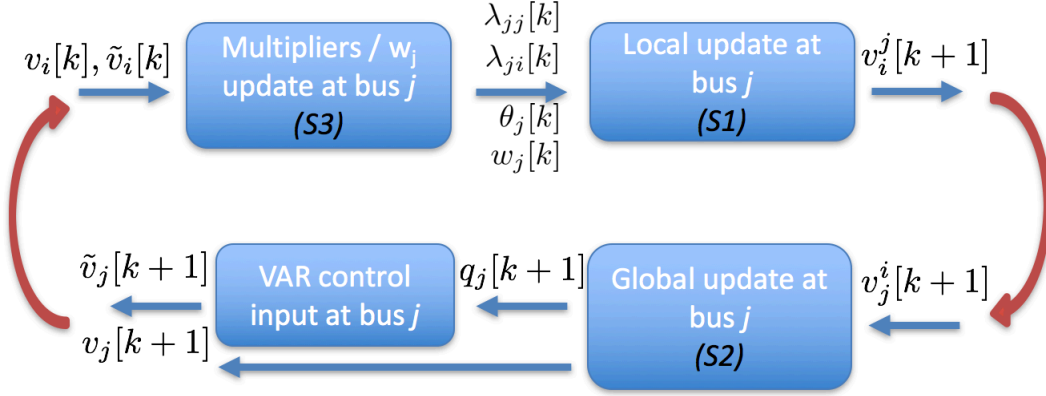


Figure 4.1: Online implementation of the proposed DVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of ADMM iterations and control updates, while the two vertical arrows correspond to the communications between bus j and its neighborings buses.

two-way communications between each local controller and a centralized computer. This is greatly challenged by fast system dynamics. Moreover, a direct approach to calculate \mathbf{w} suffers from the approximation error since (2.5) is obtained by linearizing the actual flow model.

To allow for fully distributed control design, we propose to obtain the possibly time-varying $w_j[k]$ at every iteration k through local information exchange. Specifically, we design an online voltage control scheme that uses the latest ADMM iterate $q_j[k]$ as the instantaneous VAR control input. Under the ADMM update rule (4.8b), it is guaranteed that every $q_j[k] \in [\underline{q}_j, \bar{q}_j]$ thanks to projection operator \mathbb{P}_j . Hence, the local VAR injection output is set to be $q_j[k]$ once its latest value is updated using (S2), without requiring the algorithm to achieve convergence. Under this setup, the instantaneous bus voltage measurement is updated to $\bar{v}_j[k]$ per bus j . Thus, bus j can obtain w_j based on all neighboring bus voltage measurements, as

$$w_j[k] = \sum_{i \in \mathcal{N}_j} B_{ji} \bar{v}_i[k] - q_j[k]. \quad (4.10)$$

This update only needs all incident line reactance values, assumed available locally at each bus. In addition, this control design nicely adapts to dynamically varying operating conditions because it is updated with the latest system information by incorporating the instantaneous voltage measurements. Such an online feedback-based control design is attractive for dealing with the increasing variability of renewable generations and elastic loads in future distribution networks.

The proposed online implementation of DVC design follows the stochastic optimization framework. It resembles the stochastic approximation iterations to determining $\mathbf{q}[k + 1]$ by adapting to the instantaneous $\mathbf{w}[k]$ information (see e.g., [14, 35]). Stochastic approximation is known to lead to the optimum solutions in the averaging sense. This shows that our proposed online DVC design can take random disturbance into account and provide a smooth control action on the fly that closely approximates the optimal control strategy under real-time implementation.

The implementation of our proposed ADMM-based online DVC design is illustrated in Fig. 5.1. Every iteration consists of the local computations of (S1-S3) per bus j (denoted by the blocks), in addition to two steps of information exchange among all neighboring buses (denoted by vertical arrows). The sensing hardware requirement is minimal while the communication architecture is fully local since the local computations would just require voltage measurement from neighboring buses. Furthermore, the local computation subproblem, either QP in (S1) or linear updates in (S2-S3), can be solved efficiently. Upon obtaining $q_j[k + 1]$ in (S2), bus j uses it to control local VAR resources for maintaining the voltage $\bar{v}_j[k + 1] \rightarrow \mu_j$.

Remark 5. (*Cyber Network Topology*) The node-to-node architecture of the DVC design can be generalized to instead coordinate clusters of buses as long as the cyber network is connected; see e.g., [27]. This way, it is not necessary to have DERs to be connected to each other by a line segment. Additionally, even if the distribution network is not a complete entity (i.e., a DER is not necessarily attached to every bus), we may eliminate all the buses with no DERs to create an equivalent network by adopting the Kron reduction method [62]. Thus, this reduced network consists of only buses with DERs installed. By adopting the voltage-based feedback signal \mathbf{w} in (4.10), the corresponding reduced *Bbus matrix* \mathbf{B} would explicitly account for all system characteristics as to the original network. Accordingly, the performance of our DVC design can still be guaranteed under a generalized distribution network where DERs are not attached to each and every bus.

4.4 Robustness to Communication Link Failures

The proposed DVC performance depends on the quality of the communication links. Although it has been assumed that the node-to-node communication is perfect with ideal links throughout the

algorithmic implementation, for contemporary digital communication systems, random link failures are common due to either network congestion, or poor signal-to-noise ratios in some wireless environments. Investigating how the ADMM-based DVC design copes with random link failures is related to the so-termed asynchronous ADMM problem. This has attracted considerable attention from the parallel optimization/computation communities (see e.g., [63,64] and references therein). Most of these approaches consider parallel implementation of ADMM algorithms with potential information exchange delays among peer processors. Under this setup, updating the decision variable $\mathbf{z}[k+1]$ is implemented at a centralized processor and thus would incur no synchronization-related issues. For our DVC design as illustrated in Fig. 5.1, the updates of $v_j[k+1]$ and $q_j[k+1]$ in (S2) are executed in each local bus j and could become asynchronous due to random link failures.

To tackle this, we extend our earlier work [59] on developing ADMM-based consensus averaging algorithms under random link failures. A “freezing” strategy is advocated for every ADMM variable update, which is unchanged until new information becomes available from neighboring nodes’ message packets. The asynchronous consensus averaging algorithm under such a freezing strategy has been proven convergent in [59] if the failure of each link follows a Bernoulli distribution, independent across time. We propose to extend the DVC design shown by Fig. 5.1, by modifying update steps (S1) - (S3) to incorporate only the information from the active links at every iteration k . We denote $\mathcal{B}_j[k] \in \mathcal{N}_j$ as the subset of neighboring buses actively connected to bus j at iteration k . Hence, $j \in \mathcal{B}_j[k]$ always holds. Our proposed asynchronous (A-)DVC design

Algorithm 1 Asynchronous DVC (A-DVC) algorithm

```

1: for every iteration  $k = 1, 2, \dots$  do
2:   for every bus  $j$  with  $|\mathcal{B}_j[k]| \geq 2$  do
3:     (AS1): update  $\mathbf{x}_j[k+1]$  as in (4.5);
4:     (AS2): update  $v_j[k+1] = \frac{1}{|\mathcal{B}_j[k]|} \sum_{i \in \mathcal{B}_j[k]} \nu_j^i[k+1]$ 
5:       update  $q_j[k+1]$  as in (4.8b);
6:     Freeze  $\bar{v}_i[k+1] = \bar{v}_i[k], \forall i \notin \mathcal{B}_j[k]$ ;
7:     Update  $w_j[k+1]$  as in (4.10);
8:     (AS3): update  $\theta_j[k+1], \{\lambda_{ji}[k+1]\}_{i \in \mathcal{B}_j[k]}$  as in (4.9), and  $\{\lambda_{ji}[k+1] = \lambda_{ji}[k]\}_{i \notin \mathcal{B}_j[k]}$ ;
9:   end for
10:  for any other bus  $j$  with  $|\mathcal{B}_j[k]| = 1$  do
11:    Set all associated  $\mathbf{x}_j[k+1], \mathbf{z}[k+1]$ , and  $\mathbf{y}[k+1]$  values as the last iterates.
12:  end for
13: end for

```

is tabulated in Algorithm 1 assuming $\mathcal{B}_j[k]$ is known per bus j .

Proposition 4. *The failure of each link follows an independent Bernoulli distribution. The iterates of the A-DVC method as tabulated in Algorithm 1 asymptotically converge to the optimal solution in (4.3), for both single- and multi-phase voltage control problems.*

Proof of this proposition follows directly from the convergence analysis results of the randomized ADMM algorithm in [58]. By introducing the so-termed Douglas-Rachford (DR) operator, a randomized DR splitting algorithm is convergent based on the non-expansiveness of random Gauss-Seidel iterations. This framework with minor assumptions becomes equivalent to an asynchronous ADMM algorithm. Thus, the convergence proof follows. The proposed A-DVC design adopting the freezing strategy exhibits robustness against potential random communication link failures. This will be further verified through numerical simulations.

4.5 Numerical Tests

Numerical tests are performed to demonstrate the effectiveness of the proposed (A-)DVC scheme for single- and multi-phase networks. Both static and dynamic tests using realistic system loading and generation data are considered. For simplicity, the desired voltage magnitude μ_j is chosen to be 1 in p.u. at every bus j , whereas the substation OLTC-based voltage regulation maintains unit voltage level at the feeder head. Each bus is assumed to have a certain number of PV panels installed and be able to provide voltage support via efficient inverter design. To demonstrate the practical values of the proposed methods, the open-source simulator OpenDSS [65] is used to solve for the non-linear multi-phase power flow and generate the output voltage profile.

4.5.1 Single-Phase 21-bus Radial Feeder

We first test on a simple case for better investigating the performance of ADMM-based voltage control design. Specifically, consider a 12kV single-phase radial distribution feeder of 21 buses; see, e.g., Fig. 2.1 for $N = 20$. Hence, the phase index ϕ can be ignored as well. Each line segment has impedance value $(0.233 + j0.366)\Omega$ and power base $S_{\text{base}}=100\text{MVA}$. Static tests are

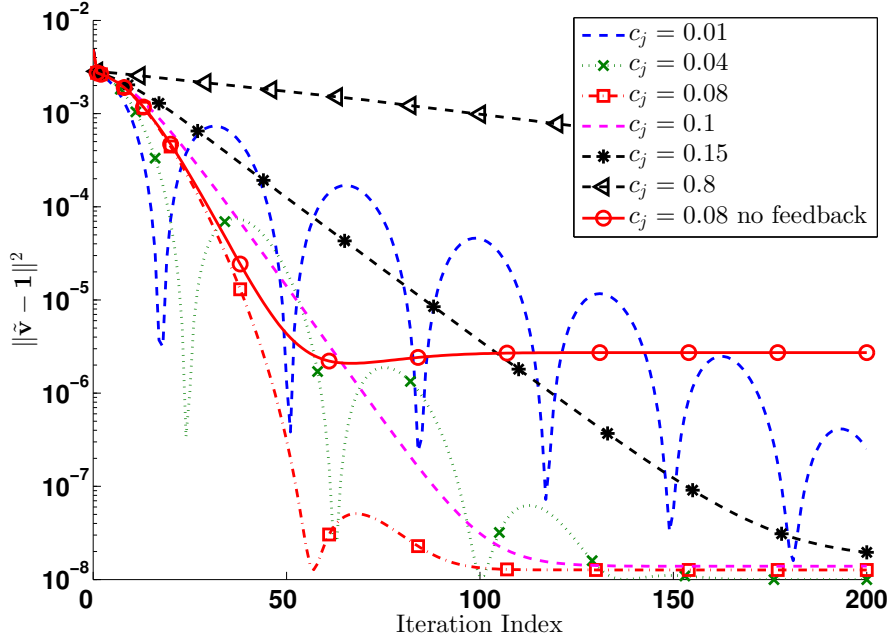


Figure 4.2: Log-scale voltage mismatch squared error versus iteration index for various c_j values for the static test settings.

implemented for comparison purposes where each bus has a constant load at $(70 + j20)$ kVA and abundant VAR resources $q_j \in [-100, 100]$ kVA corresponding to constraint (4.1c) such that a solution for $\mathbf{v} = \mathbf{1}$ always exists under no VAR penalization ($\mathbf{c} = \mathbf{0}$).

We test the DVC algorithm with various choices of regularization coefficient c_j in (4.1a). To verify the aforementioned claim on improving the computational speed and dynamic stability using a larger c_j , we plot in Fig. 4.2 the iterative log-scale voltage mismatch squared error norm $\|\tilde{\mathbf{v}} - \mathbf{1}\|^2$ for each c_j with all other settings the same, e.g., $\rho = 1$. Again, this squared error norm is based on the voltage profile solved by the AC power flow. With an increasing c_j , the oscillation level of the corresponding curve decreases gradually, corroborating the numerical conditioning comment in Sec. 4.1. Moreover, to verify our feedback-based design, we also test the scenario using a constant and known \mathbf{w} with $c_j = 0.08$, as opposed to our DVC design that updates \mathbf{w} recursively using (4.10). Fig. 4.2 illustrates that our feedback-based approach has reduced the steady-state mismatch error of the constant- \mathbf{w} approach, justifying its capability in mitigating modeling mismatch of the linear approximation in (2.5). The value of $c_j = 0.08$ has been chosen for this feeder, as it induces no voltage oscillation with preferred convergence speed (around 50 iterations until convergence).

Last, we let $c_j = 0.8$ to investigate the effect on optimal voltage profile to a highly regulated q_j . Based on (3), a larger regularization coefficient would result in increasing voltage mismatch error. Thanks to our feedback-based design in Sec. 4.3, Fig. 4.2 verifies that the obtained voltage mismatch is not sensitive to the varying level of VAR penalty, as the voltage mismatch $\|\tilde{\mathbf{v}} - \mathbf{1}\|^2 \rightarrow 0$ for a wide range of c_j values. Accordingly, this results in the same system operating conditions when $\tilde{\mathbf{v}} = \mathbf{1}$, e.g., the total line losses and optimal VAR setting are the same for different c_j considered. To sum up, one may tune the regularization coefficient c_j to achieve certain control performance specifications in terms of convergence rate and oscillation level while still being able to closely approach the flat voltage profile.

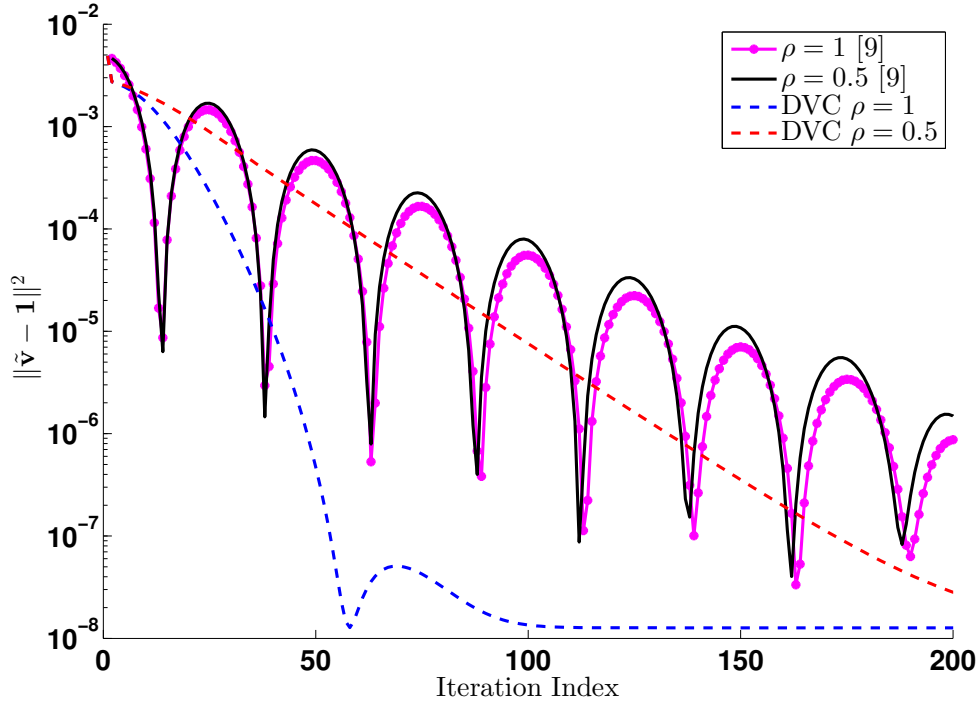


Figure 4.3: Log-scale voltage mismatch squared error comparison between the proposed DVC scheme and the algorithm of [15] with different ρ values.

We further investigate the effects of ρ value with $c_j=0.08$ fixed. Fig. 4.3 plots the log-scale voltage mismatch error for both the proposed DVC method and an earlier open-loop design in [15]. The latter algorithm, developed solely for single-phase networks but also based on ADMM, requires measuring and exchanging information regarding active/reactive flow of incident lines. It incurs much higher complexity in sensing and communication as compared to our DVC approach. Generally for both ADMM-typed methods, the level of regularization in the Lagrangian function

(4.4) would affect the convergence speed. For our DVC design, extensive testing has shown that a reasonably large ρ , such as the choice of $\rho = 1$, would be preferred. Generally speaking, it is challenging to obtain the best choice of ρ , which is beyond the scope of this thesis (see, e.g., [66] for relevant discussions on this topic). For practical implementations, we suggest adaptively adjusting the ρ value at each node depending on the oscillation level of the resultant voltage. More interestingly, with proper choice of ρ our proposed DVC design significantly outperforms the approach in [15], in terms of both convergence speed and the oscillation level. Although the oscillation of the open-loop method is observed to be bounded by an envelope that asymptotically converges, such behavior could still be problematic as it would negatively affect the operations of capacitors and voltage regulators. We expect the same observations would hold when comparing our voltage-feedback design and other open-loop methods in [13, 15], also when applied for multi-phase networks. Henceforth, all static tests have validated the effectiveness of our DVC scheme, which can achieve the optimal VAR control solution at minimal sensing/communication overhead and no centralized coordination.

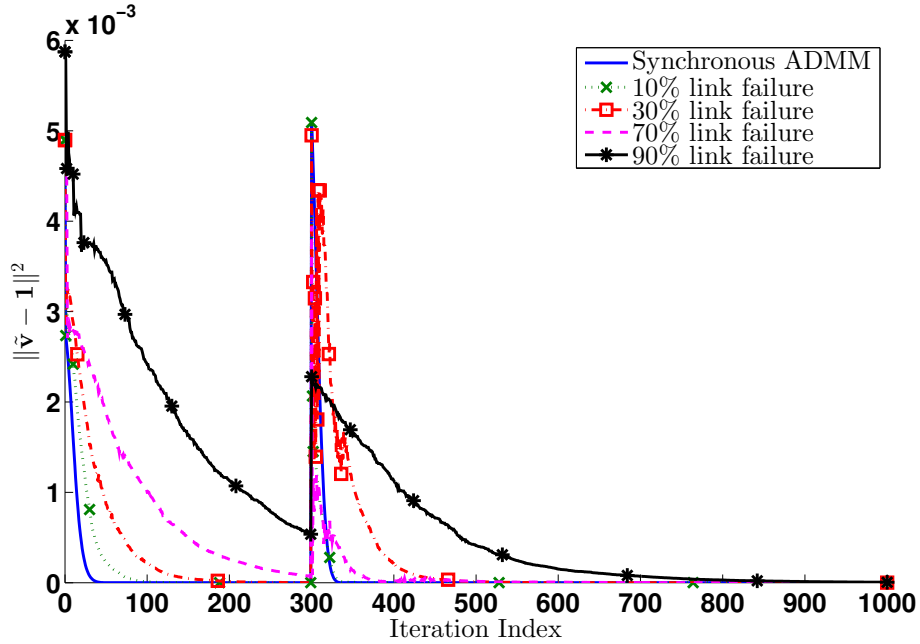


Figure 4.4: Voltage mismatch squared error versus iteration index under various rates of link failure.

Furthermore, we validate the robustness of our DVC scheme against imperfect communication links. The failure of every link is set to follow a Bernoulli distribution with the same probability

of being active; all other settings follow from earlier tests with $\rho = 1$ and $c_j = 0.08$. Fig. 4.4 compares the error performance of the proposed A-DVC design under various rates of link failure, ranging from 0% to 90%. The case of zero failure rate corresponds to that of all ideally perfect links, under which the A-DVC algorithm boils down original DVC one. Fig. 4.4 demonstrates the guaranteed convergence to the optimum for any failure rate, as established by Proposition 4. Intuitively, a lower failure rate would lead to a faster convergence speed, where the synchronous DVC scenario with no link failure at all exhibits the fastest convergence. To test the dynamic performance of the A-DVC algorithm, a decrease in generation output is introduced at iteration $k = 300$ that could represent the scenario of PV variability under sudden cloud coverage. Under this setting, our A-DVC scheme is still convergent to a nearly flat steady-state voltage profile. This test validates the robustness and stability of the proposed A-DVC design under imperfect communication scenarios.

4.5.2 IEEE 123-bus Multi-Phase Feeder

Our proposed DVC scheme is also validated using the unbalanced 123-bus feeder case considering dynamic operating conditions [55]. In particular, we use the real data on residential load and solar generation profiles shown in Fig. 4.5. These minute-sampled profiles including reactive power loading are taken from an online data repository [67], and were collected at a US residential home on Friday, June 20th, 2010. As for the test settings, the tap positions of voltage regulators are held constant as the original test case in order to better capture the performance of different inverter-based voltage control designs. The voltage at the substation of the feeder head is kept to be 1 p.u. A certain number of residential homes with solar generation rated at 3.5kW peak capacity are attached to each node. This number is determined and rounded up by dividing the IEEE 123-bus active spot load profile by the maximum daily active load ($\approx 6\text{kW}$) of a home as shown in Fig. 4.5. Thus, the location of DERs corresponds to nodes with active power loading in the IEEE 123-bus test case. Once the number of homes for each node is computed, both the active and VAR loading can be directly assigned according to the dataset in [67]. Additionally, we diversify the loading by small random additive noises. At every slot, the VAR limits $[\underline{q}, \overline{q}]$ are updated based on the given inverter rating (i.e., 3.5kVA per inverter at a home) and the instantaneous solar

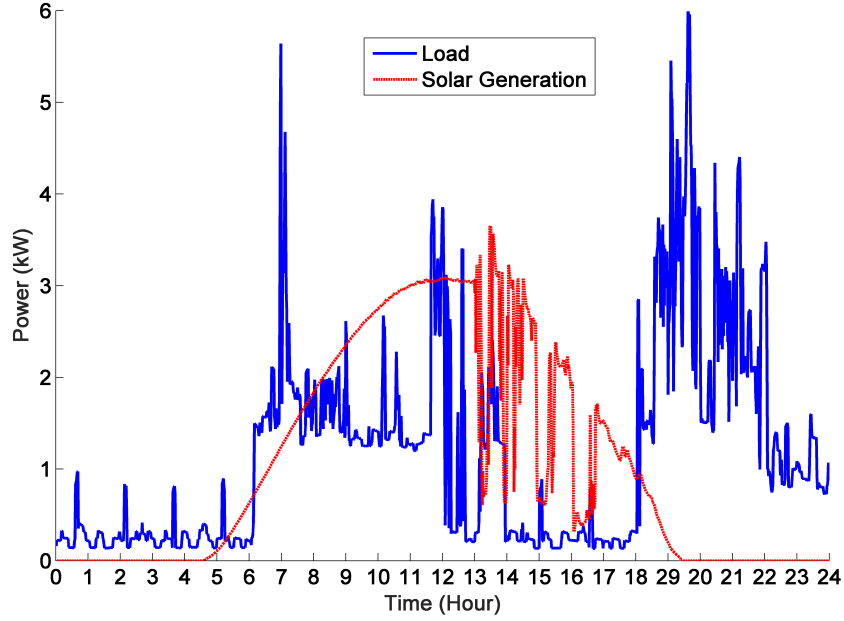


Figure 4.5: Sample daily load and solar PV generation profiles of a US home.

power output. Under these settings, it turns out that VAR resources of inverters are insufficient to achieve perfectly flat voltage at all times. This implies that the control problem (4.2) turns out to be constrained by limited VAR resources. Thus, one can encourage a higher level of VAR support by choosing a reasonably small $c_j^\phi = 0.08$.

The proposed online DVC implementation in Sec. 4.3 is tested to tackle the under- and over-voltage dynamics in this feeder due to load and solar variations. Each ADMM iteration is updated every 2 seconds (a total of 30 updates per minute) assuming a constant minute-to-minute system operation condition \mathbf{w} . This update rate is sufficient for DVC design with an uniform $c_j^\phi = 0.08$ to achieve a satisfactory convergence property within a minute interval. The local control design developed in [3, 17] is also included to demonstrate the improved performance achieved by coordinating network-wide VAR resources through information exchange. It uses locally available voltage measurements and follows either a droop curve [17] or the integral control principle [3], with no communications among nodes. Fig. 4.6 plots the daily voltage mismatch error for the a-phase of the 123-bus, where the other two phases are observed to exhibit similar comparisons. The scenario of no inverter VAR output is included as a benchmark. For all three schemes, the voltage mismatch error is almost the same around noon due to higher solar generation and accordingly

smaller limits of VAR capability. For the rest of the day, especially morning and evening hours when the loading is at the highest, the voltage violation is extremely severe at no VAR support while the inverter-based DVC ideally maintains the nearly flat voltage profile. The local control scheme is also able to reduce the high-level voltage violation based on local voltage information. However, due to lack of network-wide information to support centralized coordination of VAR resources, the local design suffers from visible mismatch error, especially during evening hours in the zoom-in view. Thanks to information exchange among neighboring nodes, the DVC scheme efficiently percolates local voltage violation to the rest of network to solicit VAR support from remote inverters. It attains the globally optimal performance as if all information were available at a centralized controller, leading to around 50% reduction in voltage mismatch error from local control.

In summary, the proposed (A-)DVC design can efficiently improve the performance of local strategies by coordinating network-wide VAR resources. This capability is more valuable when VAR resources are significantly constrained at some locations, a scenario that will represent a majority of distribution feeders into the near future. Meanwhile, its robustness to communication link failures is also attractive, considering the limited deployment of cyber infrastructure in distribution systems. Thus, we envision the proposed DVC designs will be instrumental in engaging inverter-based VAR resources to improve voltage support by accounting for practical constraints in both physical and cyber layers.

4.6 Summary

This chapter has developed a fully distributed voltage control (DVC) design to manage VAR-capable DERs deployed in an increasingly dynamic and variable distribution network. We have cast the voltage control problem as one of quadratic programming that minimizes the voltage mismatch under VAR resource limits. The ADMM distributed optimization algorithm is adopted to develop the decentralized structure for the proposed DVC design. To reduce the communication complexity, the power flow coupling is linearized to involve only neighboring nodes. Thus, the DVC design only relies on information exchange among neighboring nodes with local computa-

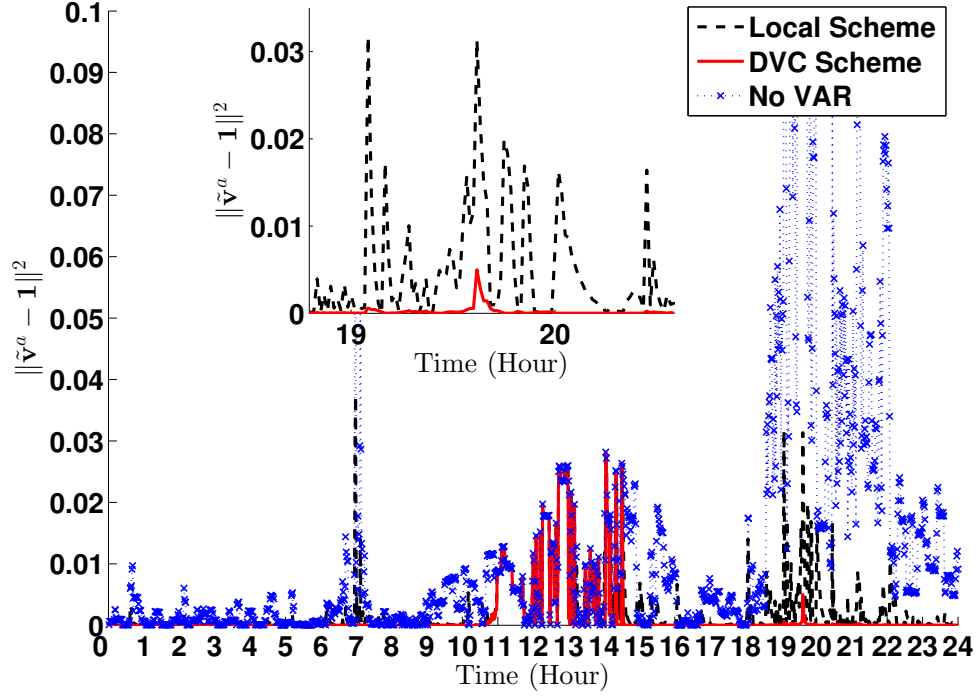


Figure 4.6: Daily voltage mismatch squared error at phase a of the 123-bus feeder.

tions. Furthermore, we have extended the implementation of the proposed ADMM-based control in an online fashion that is robust to random communication link failures. The effectiveness and robustness of our proposed design have been validated through extensive numerical tests using realistic multi-phase feeders under dynamic testing environments.

CHAPTER 5

COMMUNICATION-COGNIZANT HYBRID VOLTAGE CONTROL

The distributed voltage control (DVC) design in Chap. 4 can effectively coordinate network-wide VAR resources to attain a globally optimal voltage profile. Albeit the DVC design is robust against random communication failures, it would fail to work and incorporate any local voltage information under the worst-case scenario of total communication outages. This is because each node needs to freeze all variable updates in order to stay coordinated with neighboring nodes. Hence, to reduce communication complexity and enhance robustness to imperfect communications, it is imperative to develop an integrated design that can achieve the dual objectives of flexible adaptivity to variable rate of communications and global optimality of voltage regulation performance. Such an innovative design has the potential of unifying the currently separated framework of either local or distributed control design of [3–5].

The present chapter aims to design a hybrid voltage control (HVC) strategy that can dynamically adapt to varying system operating conditions while being fully cognizant of the instantaneous rates of communication links. To cope with practical communication limitations, the proposed HVC scheme consists of both distributed and local control architectures and does not require a centralized authority. We formulate the network-wide VAR optimization problem based on the linearized power flow model for analysis purposes only. The resultant quadratic programming problem is solved by a *partial* primal-dual gradient (PPD) algorithm in discrete-time domain, which is a variant of the classical primal-dual (sub)gradient method; see e.g., [68]. We provide the analysis in step-size choices that can guarantee convergence. We further use the PPD-based solver to design an online HVC strategy where each bus can integrate both the local voltage measurement and the communication information shared by neighboring buses. Although a linearized model has been adopted for the algorithmic development and analysis, performance of the proposed HVC design has been verified using the full ac power flow model for unbalanced and lossy distribu-

tion networks. Compared to existing voltage control approaches, the main contributions of our HVC design are three-fold. First, it explicitly accounts for the VAR limits by using the projection operator. Due to the discontinuity of projection mapping, general Krasovskii's methods [69] for analyzing the stability of primal-dual gradient flow method would not hold. To tackle this problem, we have expressed the operation as a subgradient step featured by an indicator function in order to establish the stability of the PPD-based HVC design in Sec. 5.2.1. Second, the HVC design only requires each bus to measure its local voltage magnitude and communicate it to neighboring buses. The sensing requirement and communication overhead are minimal in addition to simple control algorithmics compared to most (de)centralized strategies. Last but not least, our hybrid design can integrate both the neighboring bus voltage information and local voltage measurements regardless of the communication link conditions. This way, the HVC design is cognizant of the instantaneous availability of communication links while effectively tracking the globally optimal VAR setting. Interestingly, although the HVC updates have been developed using the distributed PPD-based solver, it would boil down to a surrogate local voltage control update during a total communication outage. Under this worst-case scenario, satisfactory performance can still be achieved as it responds to local voltage variations.

5.1 Hybrid Voltage Control Problem

For ease of presentation, we design the HVC design based on the single-phase network. Its multi-phase counterpart could be readily extended by considering the power flow model (2.12). To this end, we start with the static problem formulation where the operating condition \mathbf{w} stays constant. The HVC algorithm design will be eventually extended to be fully online and distributed where each bus j only needs the neighboring voltage information to dynamically track w_j . The voltage mismatch error objective $f_1(\mathbf{v}) := \frac{1}{2}\|\mathbf{v} - \boldsymbol{\mu}\|^2$ is introduced to provide the *globally* optimal VAR setting. Minimizing f_1 subject to (2.5) could be solved collaboratively by the buses using distributed optimization techniques; see e.g., [5, 16, 27]. Nonetheless, these distributed designs would fail to adapt to a total communication outage, during which isolated buses have to freeze all local variables in order to remain coordinated with neighboring buses. To make the control design more

robust, we introduce also a *weighted* error objective $f_2(\mathbf{v}) := \frac{1}{2}\|\mathbf{v} - \boldsymbol{\mu}\|_{\mathbf{B}}^2$ where $\|\mathbf{y}\|_{\mathbf{B}}^2 := \mathbf{y}^T \mathbf{B} \mathbf{y}$ for any vector \mathbf{y} . It will turn out that the gradient direction for minimizing f_2 only depends on the local voltage deviation, as in the droop design [3]. Under limited VAR resources, this weighted objective f_2 is not equivalent to the original one f_1 . However, incorporating the former turns out to be extremely useful for achieving the HVC design that can seamlessly perform local updates even with no incoming messages from neighboring buses.

Towards achieving the attractive features of both distributed and local control designs, we cast the HVC problem as

$$(\mathbf{v}^*, \mathbf{q}^*) := \arg \min_{\mathbf{v}, \mathbf{q}} [f_1(\mathbf{v}) + \gamma f_2(s(\mathbf{q}))] \quad (5.1a)$$

$$\text{subject to } \mathbf{B}\mathbf{v} = \mathbf{q} + \mathbf{w} \quad (5.1b)$$

$$\underline{\mathbf{q}} \leq \mathbf{q} \leq \bar{\mathbf{q}}, \quad (5.1c)$$

where $s(\mathbf{q}) := \mathbf{X}(\mathbf{q} + \mathbf{w}) = \mathbf{v}$ according to (2.5), and \mathbf{q}^g is replaced by \mathbf{q} in the rest of the thesis. Both \mathbf{v} and \mathbf{q} are the decision variables in the problem (5.1), with either one uniquely determined by the other. However, expressing f_2 as a function of \mathbf{q} would allow for conveniently forming its instantaneous gradient direction using local voltage mismatch error. Note that the problem (5.1) is strongly convex because \mathbf{B} is positive definite [3], and thus the optimum $(\mathbf{v}^*, \mathbf{q}^*)$ is unique. The objective function (5.1a) uses a positive parameter $\gamma > 0$ to balance between f_1 and f_2 , depending on the communication rates. Under very frequent link failures, γ needs to increase to weigh more on f_2 and thus the local voltage error information. If high-rate communications are available, a small γ value would be preferred to minimize the global voltage mismatch error f_1 . Hence, the choice of γ can be determined upon assessing the quality of communication network in order to balance between the global voltage mismatch error and the local one. Furthermore, the sparsity structure of \mathbf{B} based on the network topology would induce coupling only among the neighborhood \mathcal{N}_j in (5.1b). This is instrumental for the development of a distributed HVC solution. Moreover, the bounds on each local VAR resource in (5.1c) are constrained by inverters' apparent power limit and/or power factor limit, and thus they depend on the instantaneous active power outputs. If there are abundant VAR resources, (5.1c) would become inactive. Under this scenario,

the optimal VAR output \mathbf{q}^* is the same regardless of the choice of γ , as $\mathbf{v}^* = \boldsymbol{\mu}$ can always be achieved. This fact justifies the inclusion of f_2 as its solution could closely approximate the globally optimal one to f_1 . Last, the modeling of inverter controllers is not needed for our system-level problem [15, 17, 27]. Thanks to the time-scale separation between the dynamics of internal inverter control and distribution system level disturbances [23], the former has very minimal effect on the network VAR control problem [24].

5.2 Communication-Cognizant Hybrid Voltage Control Design

This section presents our proposed hybrid voltage control (HVC) framework. Combining the distributed and local control features, we aim to solve (5.1) by adopting the partial primal-dual (PPD) algorithm, a distributed optimization technique based on saddle point flow method that has been studied and applied in different literature (see, e.g., [69–71]). Based on the PPD, we design the HVC scheme that needs only to measure and incorporate the dynamic voltage magnitude. As detailed soon, the proposed feedback approach is very different from existing distributed control schemes in power systems since most of them are developed as a static optimization problem and overlook the communication imperfectness and online implementations. Most importantly, our HVC scheme would boil down to a surrogate local voltage control problem under the worst-case scenario of a total communication failure. Thus, it enjoys a satisfactory performance by having the cognizance of varying communication scenarios.

Introducing the Lagrangian multiplier $\boldsymbol{\lambda}$ for the equality constraints in (5.1b), we obtain the following Lagrangian function for the given static optimization problem (5.1):

$$\begin{aligned} \mathcal{L}(\mathbf{v}, \mathbf{q}, \boldsymbol{\lambda}) \underset{\mathbf{q} \in \mathcal{Q}}{=} & \frac{1}{2} \|\mathbf{v} - \boldsymbol{\mu}\|^2 + \frac{\gamma}{2} \|\mathbf{X}(\mathbf{q} + \mathbf{w}) - \boldsymbol{\mu}\|_{\mathbf{B}}^2 \\ & + \langle \boldsymbol{\lambda}, \mathbf{B}\mathbf{v} - \mathbf{q} - \mathbf{w} \rangle, \end{aligned} \quad (5.2)$$

where we define $\mathcal{Q} := \{\mathbf{q} | \mathbf{q} \in [\underline{\mathbf{q}}, \overline{\mathbf{q}}]\}$ and $\langle \cdot, \cdot \rangle$ represents the inner product. Based on this Lagrangian function, the primal-dual algorithm works by cyclically minimizing the primal variables \mathbf{v} and \mathbf{q} , followed by the gradient ascent-based update of the dual variable $\boldsymbol{\lambda}$. To account for total communication failures, we adopt a modified version, namely, the partial primal-dual (PPD) algo-

rithm, which iteratively updates \mathbf{q} in a gradient-descent fashion instead of finding the minimum at every step [72]. Specifically, the $(k + 1)$ -th iteration at bus j by communicating with neighboring buses consists of the following three steps:

- (S1) Update \mathbf{v} : For given \mathbf{q}^k and $\boldsymbol{\lambda}^k$, \mathbf{v} is updated by solving $\mathbf{v}^{k+1} = \arg \min_{\mathbf{v}} \mathcal{L}(\mathbf{v}, \mathbf{q}^k, \boldsymbol{\lambda}^k)$, which is an unconstrained quadratic program, for which there exists closed-form solution. Note that the Lagrangian \mathcal{L} in (5.2) is separable over v_j 's, and thus the update decouples into each bus j as

$$v_j^{k+1} = - \sum_{i \in \mathcal{N}_j} B_{ji} \lambda_i^k + \mu_j. \quad (5.3)$$

The iterates \mathbf{v}^k can be viewed as an estimate of the network voltage based on the power flow model (2.5). This \mathbf{v} update via direct minimization is different from the classical gradient flow algorithm, which would resort to a gradient descent-type update like $\dot{\mathbf{v}} = -\frac{\partial}{\partial \mathbf{v}} \mathcal{L}(\mathbf{v}, \mathbf{q}^k, \boldsymbol{\lambda}^k)$. Hence, the term *partial* is used here to acknowledge this difference. A similar strategy of “partial gradient update” has been explored in [71] for a continuous-time setting whereas our focus is on the discrete-time update.

- (S2) Update \mathbf{q} : Using the step-size $\alpha > 0$, we perform a gradient-projection based update on \mathbf{q} , as given by

$$\mathbf{q}^{k+1} := \mathbb{P}\{\mathbf{q}^k - \alpha \nabla_{\mathbf{q}} \mathcal{L}(\mathbf{v}^{k+1}, \mathbf{q}^k, \boldsymbol{\lambda}^k)\}, \quad (5.4)$$

where the operator $\mathbb{P}\{\cdot\}$ projects any input into the box \mathcal{Q} . Hence, $\mathbf{q}^k \in [\underline{\mathbf{q}}, \bar{\mathbf{q}}]$ always holds at any time k , and this iterate serves as the feasible control signal to command the VAR resources. Under this setting and using $\mathbf{XB} = \mathbf{I}$, the gradient direction in (5.4) becomes

$$\begin{aligned} \nabla_{\mathbf{q}} \mathcal{L}(\mathbf{v}^{k+1}, \mathbf{q}^k, \boldsymbol{\lambda}^k) &= \gamma \mathbf{XB}[\mathbf{X}(\mathbf{q}^k + \mathbf{w}) - \boldsymbol{\mu}] - \boldsymbol{\lambda}^k \\ &\approx \gamma(\tilde{\mathbf{v}}^k - \boldsymbol{\mu}) - \boldsymbol{\lambda}^k, \end{aligned} \quad (5.5)$$

where the actual system voltage $\tilde{\mathbf{v}}^k \approx \mathbf{X}(\mathbf{q}^k + \mathbf{w})$ if the VAR outputs equals to \mathbf{q}^k , following from the linearized model (2.5). Note that each \tilde{v}_j^k can be easily measured per bus j . Thus,

computing the gradient direction in (5.5) completely decouples per bus j upon measuring its voltage \tilde{v}_j^k and forming its multiplier λ_j^k . As both matrices \mathbf{X} and \mathbf{B} are uniquely invertible of each other, any weighting matrices other than \mathbf{B} will not lead to this decoupling feature. And this is exactly the reason for choosing f_2 in the objective function. Thanks to the separability of box constraints, we further decouple the update (5.4) into each bus j as

$$q_j^{k+1} := \mathbb{P}_j\{q_j^k - \alpha [\gamma(\tilde{v}_j^k - \mu_j) - \lambda_j^k]\}, \quad (5.6)$$

where \mathbb{P}_j denotes the projection at bus j to $[\underline{q}_j, \bar{q}_j]$. Even though these bounds would vary under dynamic setting based on instantaneous active power outputs of inverters, they can be easily updated at each bus. Note that the gradient update in (5.6) using the voltage measurement also nicely adapts to dynamic setting, as the physical power flow coupling always guarantees it has the latest network-wide information. This unique feature will enable the robustness of our proposed HVC design against link failures. Thus, the iterate q_j^{k+1} can be considered as an implementable *transient solution* that is actively tracking the dynamic optimality under time-varying operating conditions.

(S3) Update λ : For a given step-size $\beta > 0$, each multiplier per bus j is linearly updated using the iterative mismatch of the respective equality constraint, as given by

$$\lambda_j^{k+1} := \lambda_j^k + \beta \left[\sum_{i \in \mathcal{N}_j} B_{ji} v_i^{k+1} - q_j^{k+1} - w_j \right]. \quad (5.7)$$

The PPD-based iterations in (S1-S3) constitute the basis for our proposed hybrid voltage control design. Similar to DVC design in Ch. 4, for multi-phase networks, the generalized matrix \mathbb{B} in Sec. 2.1.2 would also have a *block* sparse structure [12]. Hence, it suffices to replace each scalar variable in the single-phase formulation by its three-phase vector counterpart for the multi-phase extensions of (S1)-(S3). In addition, the bus-to-bus architecture of the HVC design can also be generalized to coordinate clusters of buses likes those for DVC designs as long as the cyber network is still connected; see Remark 5.

5.2.1 Convergence Analysis

With no projection operation in the \mathbf{q} -update (5.4), the convergence problem of PPD boils down to a straightforward stability analysis of a discrete-time linear system. However, the presence of projection is equivalent to the so-called saturation effects in linear systems, for which the stability analysis could be much more challenging [73]. To tackle this problem, we instead adopt a convex analysis approach.

To streamline the presentation, consider the following general form of (5.2):

$$\mathcal{L}(\mathbf{v}, \mathbf{q}, \boldsymbol{\lambda}) = f(\mathbf{v}) + g(\mathbf{q}) + \mathcal{I}(\mathbf{q}) + \langle \boldsymbol{\lambda}, \mathbf{B}\mathbf{v} - \mathbf{q} - \mathbf{w} \rangle, \quad (5.8)$$

where both functions f and g are strongly convex and have Lipschitz gradients. To account for \mathcal{Q} , define the following indicator function

$$\mathcal{I}(\mathbf{q}) = \begin{cases} 0, & \text{if } \mathbf{q} \in \mathcal{Q}, \\ +\infty, & \text{if } \mathbf{q} \notin \mathcal{Q}. \end{cases}$$

Thus, (5.2) is a special case of (5.8) with $f(\mathbf{v}) \triangleq f_1(\mathbf{v}) = \frac{1}{2}\|\mathbf{v} - \boldsymbol{\mu}\|^2$ and $g(\mathbf{q}) \triangleq f_2(s(\mathbf{q})) = \frac{\gamma}{2}\|\mathbf{X}(\mathbf{q} + \mathbf{w}) - \boldsymbol{\mu}\|_{\mathbf{B}}^2$. Thus, there exist positive constants η and L such that $g(\cdot)$ is a η -strongly convex function with L -Lipschitz gradient; i.e., it holds that

$$\begin{aligned} \langle \mathbf{x} - \mathbf{y}, \nabla g(\mathbf{x}) - \nabla g(\mathbf{y}) \rangle &\geq \eta \|\mathbf{x} - \mathbf{y}\|_2^2, \quad \forall \mathbf{x}, \mathbf{y} \\ \|\nabla g(\mathbf{x}) - \nabla g(\mathbf{y})\|_2 &\leq L \|\mathbf{x} - \mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y}. \end{aligned}$$

In addition, $f(\cdot)$ is c -strongly convex for any $c = 1$ while $\mathcal{I}(\cdot)$ is a convex function.

The saddle point $(\mathbf{v}^*, \mathbf{q}^*, \boldsymbol{\lambda}^*)$ to (5.8) containing the optimal solution $(\mathbf{v}^*, \mathbf{q}^*)$ to (5.1) satisfies the following KKT conditions:

$$\nabla f(\mathbf{v}^*) + \mathbf{B}^T \boldsymbol{\lambda}^* = \mathbf{0}, \quad (5.9a)$$

$$\nabla g(\mathbf{q}^*) - \boldsymbol{\lambda}^* + \partial \mathcal{I}(\mathbf{q}^*) \ni \mathbf{0}, \quad (5.9b)$$

$$\mathbf{B}\mathbf{v}^* - \mathbf{q}^* - \mathbf{w} = \mathbf{0}, \quad (5.9c)$$

where $\partial\mathcal{I}(\mathbf{q}^*)$ is the subdifferential set of $\mathcal{I}(\cdot)$ at \mathbf{q}^* , which contains any subgradient of $\mathcal{I}(\cdot)$ at $\mathbf{q} \in \mathcal{Q}$, denoted by $\tilde{\nabla}\mathcal{I}(\mathbf{q})$; i.e., $\tilde{\nabla}\mathcal{I}(\mathbf{q}) \in \partial\mathcal{I}(\mathbf{q})$. Thus, the subdifferential inclusion condition in (5.9b) can be replaced by

$$\nabla g(\mathbf{q}^*) - \boldsymbol{\lambda}^* + \tilde{\nabla}\mathcal{I}(\mathbf{q}^*) = \mathbf{0}, \text{ for some } \tilde{\nabla}\mathcal{I}(\mathbf{q}^*). \quad (5.10)$$

The ensuing analysis will rely on (5.10) as a more tractable version of (5.9b). This subgradient based treatment of KKT conditions is popular among many recent references including [74, 75]. The subgradient $\tilde{\nabla}\mathcal{I}(\mathbf{q}^*)$ in (5.10) can be expressed as $\boldsymbol{\lambda}^* - \nabla g(\tilde{\mathbf{q}}^*)$ where

$$\tilde{\mathbf{q}}^* \triangleq \arg \min_{\mathbf{q} \in \mathcal{Q}} g(\mathbf{q}) - \langle \boldsymbol{\lambda}^*, \mathbf{q} \rangle.$$

Note that $\tilde{\mathbf{q}}^*$ is well-defined when $g(\mathbf{q})$ is strongly convex or \mathcal{Q} is compact, exactly the case in our problem. Using these notations, the PPD updates in (S1)-(S3) are equivalent to

$$\begin{aligned} \mathbf{v}\text{-update: } \mathbf{v}^{k+1} &= \arg \min_{\mathbf{v}} f(\mathbf{v}) + \langle \boldsymbol{\lambda}^k, \mathbf{B}\mathbf{v} \rangle; \\ \mathbf{q}\text{-update: } \mathbf{q}^{k+1} &= \mathbb{P}\{\mathbf{q}^k - \alpha \nabla g(\mathbf{q}^k) + \alpha \boldsymbol{\lambda}^k\}; \\ \boldsymbol{\lambda}\text{-update: } \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}^k + \beta(\mathbf{B}\mathbf{v}^{k+1} - \mathbf{q}^{k+1} - \mathbf{w}). \end{aligned} \quad (5.11)$$

The recursive relation of the iterate $\{\mathbf{v}^k, \mathbf{q}^k, \boldsymbol{\lambda}^k\}$ in (5.11) can be further written as¹

$$\nabla f(\mathbf{v}^{k+1}) + \mathbf{B}^T \boldsymbol{\lambda}^k = \mathbf{0}, \quad (5.12a)$$

$$\mathbf{q}^{k+1} = \mathbf{q}^k - \alpha \nabla g(\mathbf{q}^k) + \alpha \boldsymbol{\lambda}^k - \alpha \tilde{\nabla}\mathcal{I}(\mathbf{q}^{k+1}), \quad (5.12b)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta(\mathbf{B}\mathbf{v}^{k+1} - \mathbf{q}^{k+1} - \mathbf{w}). \quad (5.12c)$$

The gist of our analysis is to show that all first-order residuals of (5.12), namely $\|\nabla f(\mathbf{v}^{k+1}) + \mathbf{B}^T \boldsymbol{\lambda}^k\|$, $\|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2$, and $\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2$, will asymptotically converge to zero. Due to the existence and uniqueness of $(\mathbf{v}^*, \mathbf{q}^*)$ under strong convexity, a vanishing property of first-order residuals leads to the asymptotic convergence of the iterates $(\mathbf{v}^k, \mathbf{q}^k)$. The conditions of step-size

¹Using the subgradient and indicator function, one can write a general gradient projection update as a subgradient update.

choice will be given here for achieve the vanishing property.

Theorem 3. *Let $\tilde{\eta}$ and \tilde{L} be the smallest and largest singular values of matrix \mathbf{B} , respectively. If the positive α and β are chosen such that*

$$\begin{cases} \alpha < 2/[\gamma(\tilde{L}^{-1} + \tilde{\eta}^{-1})], \\ \beta < 2/[\tilde{L}^2 + \gamma^{-1}(\tilde{L} + \tilde{\eta})], \end{cases} \quad (5.13)$$

then the sequence $\{\mathbf{q}^k\}$ generated by the PPD updates (S1)-(S3) converges to the optimum \mathbf{q}^ .*

Proof: We first show that the successive difference $\{\|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 + \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2\}$ is an infinitely summable sequence and thus converges to zero. We will form the difference between the iterates and their corresponding optimum solutions. To this end, substituting the update (5.12c) for $\boldsymbol{\lambda}^k$ into (5.12a) and (5.12b), respectively, and then subtracting the two latter equations by the KKT conditions (5.9)-(5.10), we have

$$\begin{aligned} & \nabla f(\mathbf{v}^{k+1}) - \nabla f(\mathbf{v}^*) + \mathbf{B}^T(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*) \\ & - \beta \mathbf{B}^T(\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) - (\mathbf{q}^{k+1} - \mathbf{q}^*)) = \mathbf{0}, \end{aligned} \quad (5.14a)$$

$$\begin{aligned} \mathbf{q}^{k+1} &= \mathbf{q}^k - \alpha(\nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*)) + \alpha(\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*) \\ & - \alpha\beta(\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) - (\mathbf{q}^{k+1} - \mathbf{q}^*)) \\ & - \alpha(\tilde{\nabla} \mathcal{I}(\mathbf{q}^{k+1}) - \tilde{\nabla} \mathcal{I}(\mathbf{q}^*)), \end{aligned} \quad (5.14b)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta(\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) - (\mathbf{q}^{k+1} - \mathbf{q}^*)). \quad (5.14c)$$

Additionally, by the strong convexity and gradient Lipschitz continuity of $g(\cdot)$, we have

$$\begin{aligned} & \frac{2\eta L}{\eta+L} \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 + \frac{2}{\eta+L} \|\nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*)\|_2^2 \\ & \leq 2\langle \mathbf{q}^k - \mathbf{q}^*, \nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*) \rangle \\ & \leq 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*) \rangle \\ & + \frac{\eta+L}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 + \frac{2}{\eta+L} \|\nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*)\|_2^2. \end{aligned} \quad (5.15)$$

The first inequality of (5.15) is a standard result under strong convexity and gradient Lipschitz

continuity; see e.g, [76, Thm 2.1.11]. Reorganizing (5.15) gives rise to

$$\begin{aligned} \frac{2\alpha\eta L}{\eta+L} \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 &\leq 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \alpha(\nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*)) \rangle \\ &\quad + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2. \end{aligned} \quad (5.16)$$

Substituting (5.14b) into (5.16) for $\alpha(\nabla g(\mathbf{q}^k) - \nabla g(\mathbf{q}^*))$ leads to

$$\begin{aligned} &\frac{2\alpha\eta L}{\eta+L} \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 \\ \leq & 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{q}^k - \mathbf{q}^{k+1} \rangle + 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \alpha\boldsymbol{\lambda}^{k+1} - \alpha\boldsymbol{\lambda}^* \rangle \\ & - 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \alpha\beta\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) \rangle \\ & + 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \alpha\beta(\mathbf{q}^{k+1} - \mathbf{q}^*) \rangle \\ & - 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \alpha(\tilde{\nabla}\mathcal{I}(\mathbf{q}^{k+1}) - \tilde{\nabla}\mathcal{I}(\mathbf{q}^*)) \rangle \\ & + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 \\ \leq & 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{q}^k - \mathbf{q}^{k+1} \rangle + 2\alpha\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^* \rangle \\ & - 2\alpha\beta\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) \rangle + 2\alpha\beta\|\mathbf{q}^{k+1} - \mathbf{q}^*\|_2^2 \\ & + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2. \end{aligned} \quad (5.17)$$

As for the function $f(\cdot)$, by the strong convexity we have

$$2c\|\mathbf{v}^{k+1} - \mathbf{v}^*\|_2^2 \leq 2\langle \mathbf{v}^{k+1} - \mathbf{v}^*, \nabla f(\mathbf{v}^{k+1}) - \nabla f(\mathbf{v}^*) \rangle. \quad (5.18)$$

Substituting $\nabla f(\mathbf{v}^{k+1}) - \nabla f(\mathbf{v}^*)$ of (5.14a) into (5.18) leads to

$$\begin{aligned} &2\alpha c\|\mathbf{v}^{k+1} - \mathbf{v}^*\|_2^2 \\ \leq & 2\alpha\langle \mathbf{v}^{k+1} - \mathbf{v}^*, \mathbf{B}^T(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1}) \rangle \\ & + \beta\mathbf{B}^T(\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) - (\mathbf{q}^{k+1} - \mathbf{q}^*)) \rangle \\ = & 2\alpha\langle \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*), \boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1} \rangle + 2\alpha\beta\|\mathbf{v}^{k+1} - \mathbf{v}^*\|_{\mathbf{B}^T\mathbf{B}}^2 \\ & - 2\alpha\beta\langle \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*), \mathbf{q}^{k+1} - \mathbf{q}^* \rangle. \end{aligned} \quad (5.19)$$

Summing up (5.17) and (5.19) results in

$$\begin{aligned}
& \frac{2\alpha\eta L}{\eta+L} \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 + 2\alpha c \|\mathbf{v}^{k+1} - \mathbf{v}^*\|_2^2 \\
& \leq 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{q}^k - \mathbf{q}^{k+1} \rangle + 2\alpha \langle \mathbf{q}^{k+1} - \mathbf{q}^*, \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^* \rangle \\
& \quad - 2\alpha\beta \langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*) \rangle + 2\alpha\beta \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_2^2 \\
& \quad + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 \\
& \quad + 2\alpha \langle \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*), \boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{k+1} \rangle + 2\alpha\beta \|\mathbf{v}^{k+1} - \mathbf{v}^*\|_{\mathbf{B}^T \mathbf{B}}^2 \\
& \quad - 2\alpha\beta \langle \mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*), \mathbf{q}^{k+1} - \mathbf{q}^* \rangle.
\end{aligned} \tag{5.20}$$

Consider a basic inequality rule for Euclidean norm

$$2\langle \sqrt{\rho} \mathbf{a}, \sqrt{\rho^{-1}} \mathbf{b} \rangle \leq \rho \|\mathbf{a}\|^2 + \rho^{-1} \|\mathbf{b}\|^2,$$

which holds for any $\rho > 0$ and vectors \mathbf{a} and \mathbf{b} of the same dimension. Using this inequality and (5.14c), the right-hand-side of (5.20) can be relaxed by

$$\begin{aligned}
& 2\langle \mathbf{q}^{k+1} - \mathbf{q}^*, \mathbf{q}^k - \mathbf{q}^{k+1} \rangle + \frac{2\alpha}{\beta} \langle \boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}, \boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^* \rangle \\
& + \left(\frac{(1-\varepsilon)\alpha}{\beta} + \frac{(1+\varepsilon)\alpha}{\beta} \right) \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2 + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 \\
& \leq \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 - \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_2^2 - \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 \\
& + \frac{\alpha}{\beta} (\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|_2^2 - \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|_2^2) - \varepsilon \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2 \\
& + \frac{\alpha(\eta+L)}{2} \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 + (1+\varepsilon)\alpha\beta(1+\rho) \|\mathbf{B}(\mathbf{v}^{k+1} - \mathbf{v}^*)\|_2^2 \\
& + (1+\varepsilon)\alpha\beta(1+\frac{1}{\rho}) \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_2^2,
\end{aligned} \tag{5.21}$$

where ε is any arbitrary constant within $(0, 1)$. It then follows from (5.20) and (5.21) that

$$\begin{aligned}
& \left(1 - \frac{\alpha(\eta+L)}{2}\right) \|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 + \varepsilon \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2 \\
& \leq \left(1 - \frac{2\alpha\eta L}{\eta+L}\right) \|\mathbf{q}^k - \mathbf{q}^*\|_2^2 + \frac{\alpha}{\beta} (\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|_2^2 - \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|_2^2) \\
& \quad - (1 - (1+\varepsilon)\alpha\beta(1+\frac{1}{\rho})) \|\mathbf{q}^{k+1} - \mathbf{q}^*\|_2^2 \\
& \quad - (2\alpha c - (1+\varepsilon)\alpha\beta(1+\rho)\sigma_{\max}) \|\mathbf{v}^{k+1} - \mathbf{v}^*\|_2^2,
\end{aligned}$$

where σ_{\max} is defined as the largest singular value of $\mathbf{B}^T \mathbf{B}$. This inequality is crucial for showing

the iterative difference is infinitely summable. By comparing the coefficients corresponding to each term, we come up with the following set of necessary conditions:

$$\begin{cases} 1 - \frac{\alpha(\eta+L)}{2} > 0, \\ \varepsilon > 0, \\ 1 - \frac{2\alpha\eta L}{\eta+L} \leq 1 - (1+\varepsilon)\alpha\beta(1+\frac{1}{\rho}), \\ 2\alpha c - (1+\varepsilon)\alpha\beta(1+\rho)\sigma_{\max} \geq 0, \end{cases} \quad (5.22)$$

which is equivalent to having

$$\begin{cases} \alpha < \frac{2}{\eta+L}, \\ \beta < \min \left\{ \frac{2\eta L}{(\eta+L)(1+\frac{1}{\rho})}, \frac{2c}{(1+\rho)\sigma_{\max}} \right\}. \end{cases} \quad (5.23)$$

The best range of β is achieved by setting $\rho = \frac{c(\eta+L)}{\eta L \sigma_{\max}}$. Accordingly, the step-size rule for β becomes

$$\beta < \frac{2c\eta L}{\eta L \sigma_{\max} + c(\eta+L)}. \quad (5.24)$$

The choice of α and β given by (5.23)-(5.24) guarantees the infinite summability of the sequence $\{\|\mathbf{q}^k - \mathbf{q}^{k+1}\|_2^2 + \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2^2\}$. Hence, each one of the summands, namely $\|\mathbf{q}^k - \mathbf{q}^{k+1}\|$ and $\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k+1}\|_2$, converges to zero. Based on (5.12), we thus assert

$$\|\nabla f(\mathbf{v}^{k+1}) + \mathbf{B}^T \boldsymbol{\lambda}^k\|_2 = 0, \quad (5.25a)$$

$$\lim_{k \rightarrow \infty} \|\nabla g(\mathbf{q}^k) - \boldsymbol{\lambda}^k + \tilde{\nabla} \mathcal{I}(\mathbf{q}^{k+1})\|_2 = 0, \quad (5.25b)$$

$$\lim_{k \rightarrow \infty} \|\mathbf{B} \mathbf{v}^{k+1} - \mathbf{q}^{k+1} - \mathbf{w}\|_2 = 0. \quad (5.25c)$$

Comparing (5.25) with (5.9)-(5.10), one can conclude that the KKT conditions hold for the limit and thus the sequence $\{\mathbf{q}^k\}$ converges to \mathbf{q}^* .

Last, for the specific forms of $f = f_1$ and $g = f_2$, their respective Hessian is $\nabla^2 f = \mathbf{I}$ and $\nabla^2 g = \gamma \mathbf{X}$. Since $\tilde{\eta}$ and \tilde{L} denote the smallest and largest singular values of \mathbf{B} , respectively, we have $\eta = \gamma \tilde{L}^{-1}$ and $L = \gamma \tilde{\eta}^{-1}$ for its inverse \mathbf{X} . As $c = 1$ for the function f_1 , the step-size conditions in (5.24) are equivalent to (5.13).

Remark 6. (*General Error Objective Functions*) Our convergence analysis is performed by assuming both functions f and g are strongly convex. This assumption can be relaxed to the so-called restricted strongly convex functions [77], such as the Huber loss function. To this end, the first inequality in (5.15) needs to be modified to have more conservative coefficients on the left-hand side, leading to a narrower region of step-size choice. In addition, it is possible to completely remove the (restricted) strong convexity assumption on f_1 . Specifically, the \mathbf{v} -update in (5.11) can be simply modified as

$$\mathbf{v}^{k+1} = \arg \min_{\mathbf{v}} f_1(\mathbf{v}) + \langle \boldsymbol{\lambda}^k, \mathbf{B}\mathbf{v} \rangle + \vartheta \|\mathbf{v} - \mathbf{v}^k\|^2$$

to tackle the case of lacking (restricted) strong convexity. Compared to the previous update scheme, one more proximal term is augmented to stabilize the system. Informally speaking, the additional term $\vartheta \|\mathbf{v} - \mathbf{v}^k\|^2$ can enhance the stability region of step-sizes for non-strongly convex functions. In this case, the step-size rule would depend on the choice of ϑ . Although the modified update would allow more general error objective functions, such generalization is likely to reduce convergence speed in practice and thus is not employed in our current HVC strategy.

In general, a larger network size increases the value of \tilde{L} , and thus may reduce the range of β . Fortunately, our empirical experience suggests that the choice of β minimally affects the convergence speed, while the step-size α for \mathbf{q} -update plays a more important role. As for the smallest eigenvalue $\tilde{\eta}$ for the reduced graph Laplacian, it is strongly related to the connectivity² of the network. As long as the network is connected, this quantity is lower bounded away from zero. Empirically, a larger α value leads to increased convergence speed. Thus, a more clustered power network would result in a faster convergence, which coincides with the common wisdom. A thorough investigation of convergence rate would involve the spectral analysis on the eigenvalues of matrix \mathbf{B} , which is an interesting future research direction.

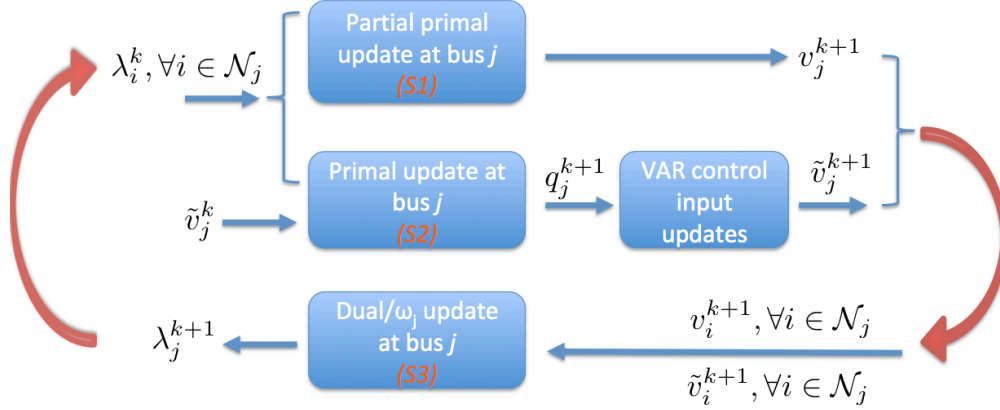


Figure 5.1: Online implementation of the proposed HVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of PPD iterations and control updates, while the two vertical arrows correspond to the communications exchange among bus j and its neighboring buses.

5.2.2 Online Feedback Design

Thus far, we assume the availability of static w_j per bus j . However, it would change in accordance with the system operating condition. To account for system dynamics based on (2.5), we need to compute and update w_j which requires network-wide complex power injections. To this end, two-way communications between each bus and a centralized computer are necessary. Generally, this is not feasible due to limited communication in the distribution network in addition to fast system dynamics. Similar to Sec. 4.3, thanks to the sparsity of \mathbf{B} , its entry is zero for any pair of buses whose corresponding buses are not connected by a line segment. Accordingly, we propose to obtain a time-varying w_j^k through neighboring voltage measurement exchanges. Each bus j measures its voltage magnitude \tilde{v}_j^{k+1} after (S2) and broadcasts to its neighboring buses. Assuming all incident line reactance values are known at each bus, we update w_j^{k+1} as

$$w_j^{k+1} := \sum_{i \in \mathcal{N}_j} B_{ji} \tilde{v}_i^{k+1} - q_j^{k+1}, \quad (5.26)$$

which is computed locally and adopted in (S3). The attractive features of the proposed feedback design are three-fold. First, we can obtain w_j^{k+1} locally at each bus by adopting bus-to-bus communication architecture. This is nicely designed to our HVC scheme as the distributed feature

²Note that the algebraic connectivity of the graph is defined as the second smallest eigenvalue of the standard Laplacian. Thus, the exact relation between the algebraic connectivity and $\tilde{\eta}$ needs further investigation in future.

is maintained. Second, the instantaneous voltage measurements contain the latest system information, and hence \mathbf{w}^{k+1} accurately approximates the dynamically varying operating conditions. Last but not least, the voltage feedback control design improves the robustness to mismatch and imperfection in system modeling (see e.g., [47, Sec. 8.9]) since the voltage measurements could potentially capture the underlying non-linearity in the power networks. Fig. 5.1 offers a schematic of the proposed PPD-based HVC implementation. Every iteration consists of the *local computations* of (S1)-(S3) per bus j (the blocks), in addition to two steps of information exchange among all neighboring buses (the vertical arrows). All the computational tasks for (S1)-(S3) are simple linear updates and thus can be executed efficiently. After updating q_j^{k+1} in (S2), bus j inputs it as the control signal for the local VAR resource and then measures and broadcasts the voltage \tilde{v}_j^{k+1} to neighboring buses for updating ω_j^{k+1} in (5.26).

5.2.3 Imperfect Communication

The performance of the proposed HVC design relies on the quality of bus-to-bus communication links, which we have assumed to be perfect throughout the algorithmic design. However, random link failures and messaging delays are common because of either network congestion, or poor signal-to-noise ratios in some wireless environments for a contemporary digital communication system. It is imperative to examine how the PPD-based HVC scheme works under imperfect communication, which leads to the following two different scenarios. One is often referred to as asynchronous networking and consists of both link failures and messaging delays [64, 78], while the other only considers link failures. The later scenario can be also referred to as a time-varying network [79, 80]. Albeit the second scenario seems to be a special case of the first, by playing a simple trick of embedding a time stamp in the message between each pair of buses, delays can also be treated as link failures.³ Informally speaking, as long as the delay is bounded and the time-varying communication network is B -connected,⁴ one should be able to choose small enough step-sizes to stabilize the proposed algorithm. There have been some analyses of algorithms under these

³We assume that each bus has a clock that is aligned. As the information exchanges and dual updates could be performed at a relatively low speed, a slight mismatch between clocks does not break the viability of this approach.

⁴It is a connectivity description of graphs under time-varying scenarios. Readers are referred to Assumption 2 at page 7 of reference [80] for detailed definition.

conditions in the literature (see, e.g., [58, 81]). Though rigorous proof of convergence properties under these conditions is beyond the scope of this thesis and will be a project for future work, we will test and validate our proposed design with a real distribution feeder in Sec. 5.3.

To tackle the challenge posed by imperfect communication networks, we leverage the work in [58, 59, 81] regarding the “freezing” strategy for distributed optimization problems. Conventionally, this scheme is advocated for every PPD variable, which remains unchanged until new information is available from neighboring buses’ message packets. The asynchronous version of a related distributed primal-dual algorithm in [81] has been proven to be convergent under random activation of agents, i.e., link $(i, j) \in \mathcal{E}$ is available only when i and j are both randomly activated. It is assumed that the activation of each bus follows a Bernoulli distribution, independent across time. Nonetheless, under this strategy, the aforementioned PPD-based HVC updates would completely halt under the case of a total link failure in the communication network. The novelty of our work lies in the extension of the HVC scheme by modifying update steps (S1)-(S3) to have: a) satisfactory performance under partial link failures, b) capability to continue providing VAR regulation under a total link failure scenario. To this end, we freeze the variables v_j and λ_j associated with the inactive bus j while always updating the VAR control signal according to (5.6) by adapting local gradient information from \tilde{v}_j^k . Since the voltage measurements always contain the most updated network information, the local voltage control design objective f_2 is advocated to continue providing the VAR support. As a result, under a total communication failure, our HVC framework boils down to a surrogate local controller design based on the current value of λ_j . This is similar to the microgrid secondary frequency/voltage control design, where λ_j can be treated as an offset signal to a local droop controller [7]. We denote $\mathcal{N}_a^k \subseteq \mathcal{N}$ as the subset activated nodes at iteration k . Our proposed asynchronous (A-)HVC algorithm is tabulated in Algorithm 2.

5.3 Numerical Tests

The numerical tests presented in this section demonstrate the effectiveness of the proposed HVC design which has an attractive communication-cognizant feature for a practical distribution feeder. We investigate the performance of our scheme under the settings of both static and dynamically

Algorithm 2 Asynchronous HVC (A-HVC) algorithm

```
1: for every iteration  $k = 1, 2, \dots$  do
2:   for bus  $j \in \mathcal{N}_a^k$  do
3:     (AS1): update  $v_j^{k+1}$  as in (5.3);
4:     (AS2): update  $q_j^{k+1}$  as in (5.4);
5:     Update  $w_j^{k+1}$  as in (5.26);
6:     (AS3): update  $\lambda_j^{k+1}$  as in (5.7);
7:   end for
8:   for bus  $j \notin \mathcal{N}_a^k$  do
9:      $v_j^{k+1} = v_j^k$ ;
10:    (AS2): update  $q_j^{k+1}$  as in (5.4);
11:     $w_j^{k+1} = w_j^k$ ;
12:     $\lambda_j^{k+1} = \lambda_j^k$ ;
13:   end for
14: end for
```

time-varying network operating conditions. A single-phase radial power distribution feeder that consists of 21 buses with $v_0 = 1$ at the substation is adopted to test the algorithm. This network is equivalent to the system in Fig. 2.1 for $N = 20$ with the impedance of each line segment setting to be $(0.233 + j0.366)\Omega$. The desired voltage magnitude μ_j is chosen to be 1 p.u. at every bus j . Furthermore, each bus is assumed to have a certain number of PV panels installed, and thus it is able to control/provide VAR via advanced inverter design. Albeit the HVC design is based on the linearized model (2.5), we test and validate the performance using the full AC power flow model. All numerical tests are performed in MathWorks® MATLAB 2014a software and OepnDSS for solving the actual power flow. Accordingly, the bus voltage magnitude, instead of the one obtained from (2.5), is used for VAR control outputs in (5.6) and the following numerical tests.

5.3.1 Static System Operating Conditions

Per bus j , we fix the loading $p_j^c = 70\text{kW}$ and $q_j^c = 20\text{kVAR}$ while choosing the inverter rating to be $(70 + \psi)\text{kVA}$ where ψ is zero-mean Gaussian having variance 13.33, thus modeling the variation in inverter sizing by 50%. Accordingly, the VAR constraints in (5.1c) would become active at some locations. We test the HVC algorithm with various choices of importance factor γ . To demonstrate the trade-off between distributed and local control designs, we plot in Fig. 5.2 the optimal voltage

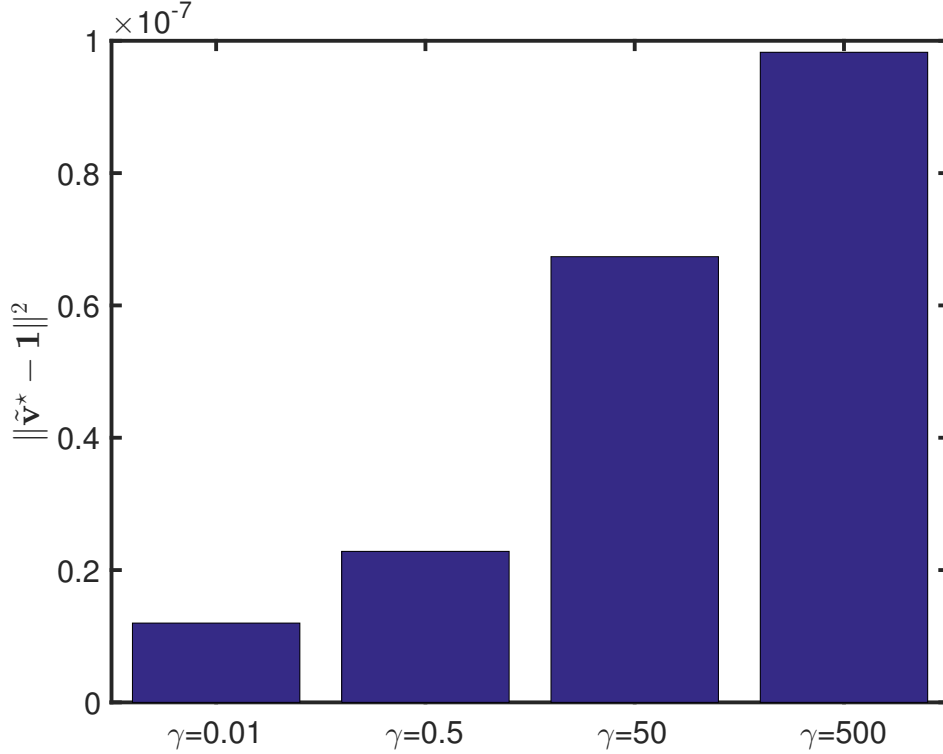


Figure 5.2: Voltage mismatch squared error versus various values of importance factor γ across the network under the static system setting.

mismatch squared error norm $\|\mathbf{v}^* - \mathbf{1}\|^2$ for each γ value with all other settings the same. Notice that increasing the γ value by adding more weight on the local control objective f_2 results in a larger voltage mismatch error. This corroborates our earlier claim that local control schemes attain a sub-optimal VAR setting under limited VAR resources. Given a system model, one may study this trade-off offline to tune the importance factor γ accordingly.

Based on the convergence properties in Theorem 3 with $\gamma = 0.5$, we have $\alpha < 0.092$ and $\beta < 0.0073$. Fig. 5.3 plots the iterative voltage mismatch squared error norm $\|\mathbf{v}^k - \mathbf{1}\|^2$ in log-scale for various step-size choices assuming a perfect communication. To violate the steps-size constraints, we let $\alpha = 0.099$ and $\beta = 0.01$ in two different scenarios, respectively. It clearly shows that the HVC design fails to converge under these cases. To stabilize our design, we bound the step-size values to be within their limits as depicted in Fig. 5.3. Note that the effect of step-size choices shows a trade-off between the stability and convergence rate. Accordingly, the larger α and β are, the faster the updates converge. Nonetheless, this could potentially lead to oscillations in

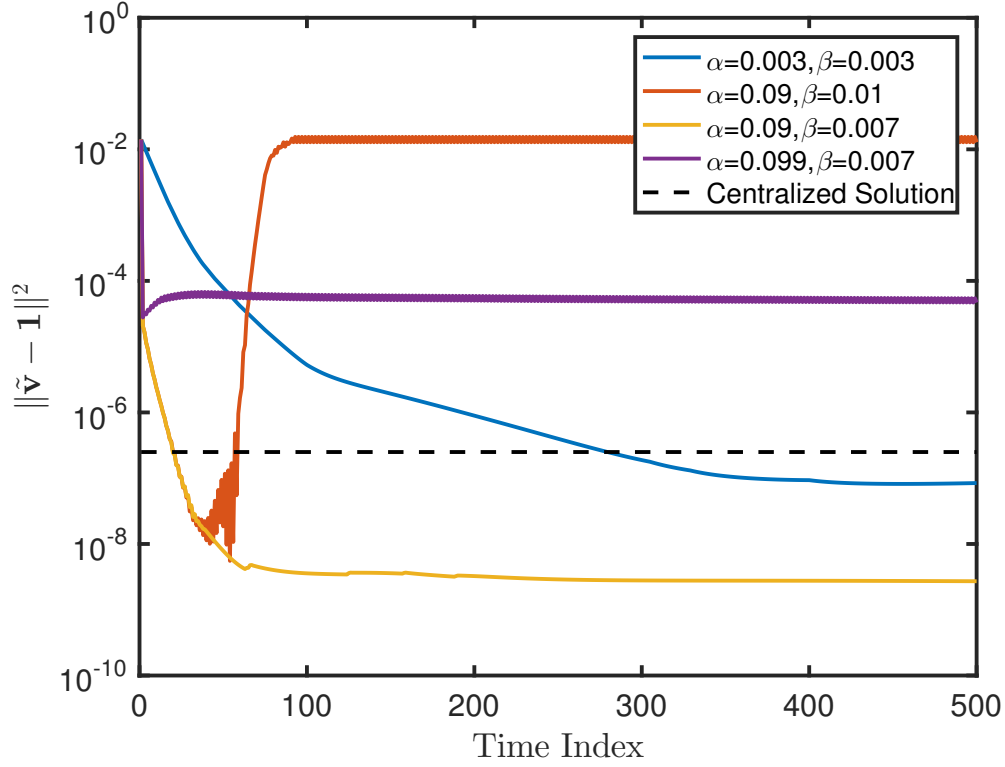


Figure 5.3: Voltage mismatch squared error versus the total number of updates across the network with different step-size choices of α and β under the static system setting.

the error performance, exhibiting instability under fast dynamics. To tackle this problem, once the full feeder information becomes available, our convergence properties in Theorem 3 are very useful in terms of proper step-size choices. Otherwise, it is also possible to adjust the step-size on-the-fly by decreasing the values based on the local bus's voltage oscillation intensity. To sum up, under appropriate step-size choices, Fig. 5.3 validates the effectiveness of our scheme, in terms of achieving the optimal VAR setting while requiring no centralized coordination.

Moreover, to validate the robustness of our A-HVC scheme under imperfect communication links, we model the activation of every bus j as a Bernoulli distribution with the same probability. Since the distributed parts of the HVC are most likely to be affected by random communication link failures, we let $\beta > \alpha$ to investigate the performance of A-HVC design. To this end, we have $\alpha = 0.005$, $\beta = 0.007$, and $\gamma = 0.05$ while fixing other settings to be the same as earlier tests. Fig. 5.4 plots the iterative voltage mismatch squared error in log-scale under various bus activation rate ranging from 10% to 100%, where the case of 100% corresponds to the perfect communication

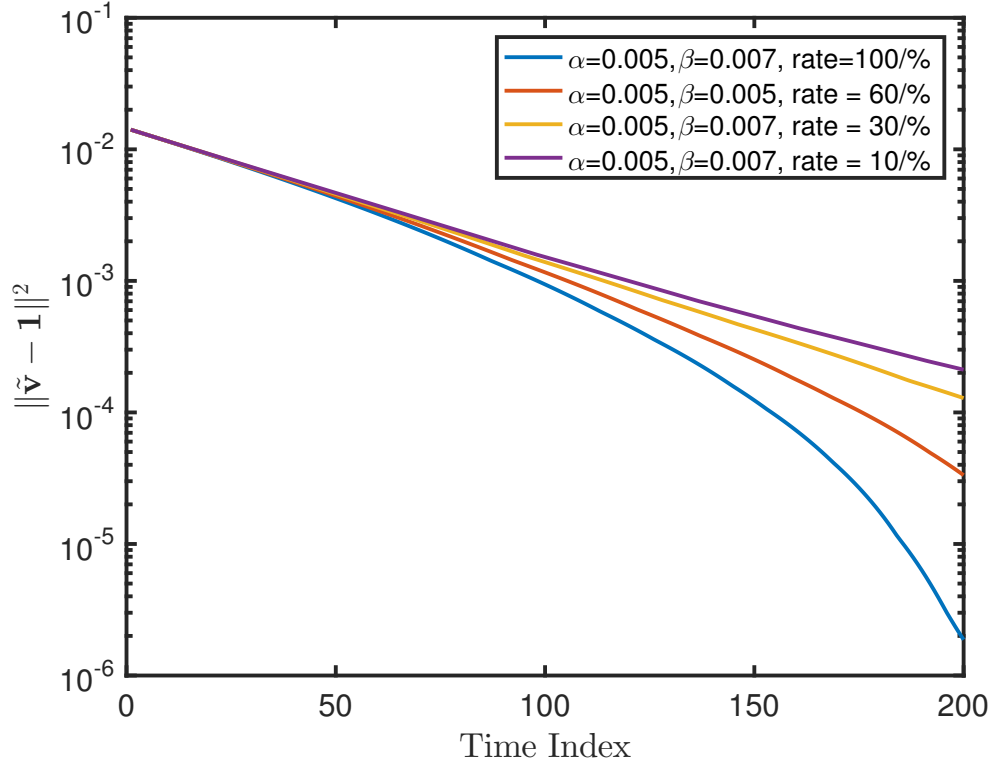


Figure 5.4: Voltage mismatch squared error versus the total number of updates across the network with fixed step-size choices of α and β and varying bus activation rate under the static system setting.

scenario (synchronous case). It clearly depicts that our design enjoys a satisfactory performance guarantee under random link failures for regulating the network voltage. Informally speaking, a lower bus activation rate would lead to a slower convergence speed, with a no link failure scenario exhibiting the fastest convergence. This test validates that our proposed A-HVC design is robust against imperfect communication and thus able to cope with cyber resource constraints.

5.3.2 Dynamic System Operating Conditions

The proposed DVC scheme is also validated using the unbalanced 123-bus feeder case [55]. To corroborate our HVC scheme for online implementation, we consider dynamic system operating conditions caused by actual loading changes gathered from an online data repository [67] as shown in Fig. 4.5. These residential minute-sampled data including reactive power consumption were

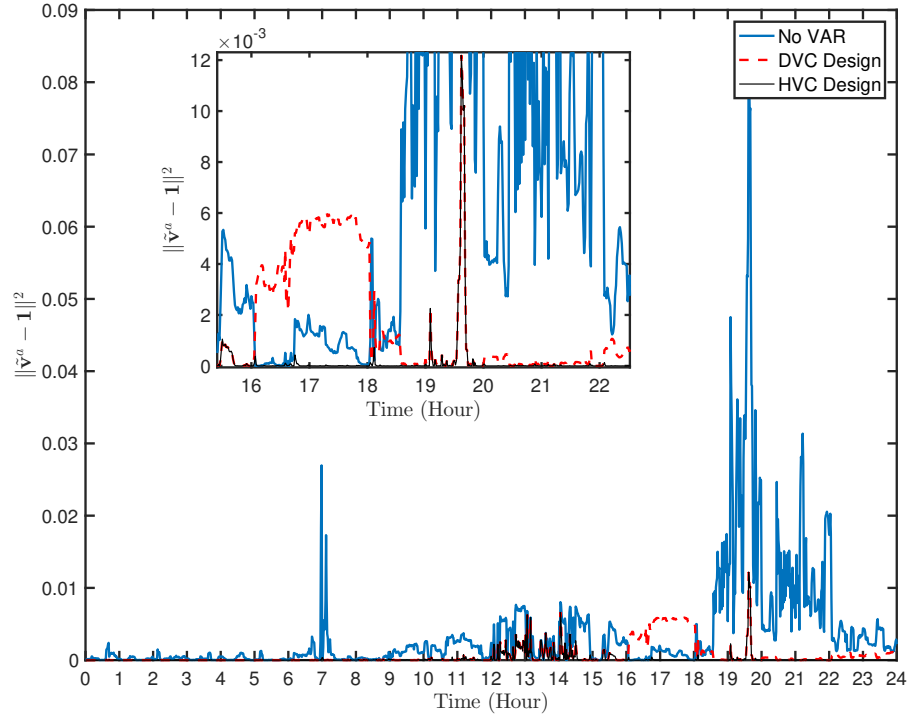


Figure 5.5: Daily voltage mismatch squared error for three different control strategies at phase a. A total communication link failure occurs from hour 16:00 to 24:00.

collected on a particular Friday in Summer 2010. Meanwhile, all settings are the same as the ones in Sec. 4.5.2. For each minute time slot, physical VAR limits $[\underline{q}, \bar{q}]$ are updated according to their inverter ratings and instantaneous active power from solar generations.

Fig. 5.5 plots the daily network-wide voltage mismatch squared error norm for the a-phase of the 123-bus. The plots for the other two phases are of similar trends and neglected here. Three different control strategies including no VAR support, distributed design, and the proposed HVC scheme are plotted. Under a quasi-static operating condition within each minute, the proposed A-HVC algorithm is implemented with each iteration updated every 2 seconds (a total of 30 iterations per minute). The same settings hold for the DVC strategy. This update rate turns out to be sufficient for both the distributed and A-HVC algorithms to achieve satisfactory convergence within a minute interval. For the benchmark case of no VAR support, there are some under- and over-voltage issues due to load and solar variations. Around the noon hours, all three strategies have similarly high voltage mismatch error. This is because the active solar power generation has reached inverter

limits during these hours, and accordingly the VAR limits are nearly zero.

To corroborate A-HVC's ability to adapt to varying communication rates, we have considered the worst-case scenario of total communication outage from hour 16:00 to 24:00. The limits of VAR resources would gradually increase from 16:00 on as the solar power generation reduces and the inverters have more capability to control the VAR outputs. Nonetheless, the total link failures during these hours render the distributed control design not responsive to the dynamic operating point at all, resulting in a highest voltage mismatch error from 16:00-18:00 even compared to the benchmark no VAR scenario. This is due to the fact that distributed design (DVC) would freeze its VAR solution by setting $q_j^{k+1} = q_j^k$ at every node. Meanwhile, the proposed A-HVC design can still effectively minimize the network-wide voltage mismatch error and gracefully maintain a nearly flat voltage profile. Thanks to its flexible adaptivity to communication availability, our A-HVC design enjoys a satisfactory worst-case performance, and significantly outperforms the distributed one under the total communication outage. To sum up, the proposed design can efficiently regulate the voltage level by coordinating network-wide VAR resources. Meanwhile, its cognizance to the instantaneous availability of communication links is also attractive, considering the limited deployment of cyber infrastructure in distribution networks. Therefore, the proposed HVC design would facilitate the future engagements in inverter-based VAR resources to improve voltage support by accounting for practical constraints in both physical and cyber layers.

5.4 Summary

This chapter has developed the communication-cognizant hybrid voltage control (HVC) scheme to coordinate network-wide VAR support in power distribution networks. We have cast the specially-designed voltage control problem combining both attractive features of distributed and local control architectures in order to minimize the voltage mismatch error under limited VAR resources. The PPD-based algorithm is evoked and only requires voltage measurement exchanges among neighboring buses with local computations. In addition, we have provided the convergence properties of the aforementioned algorithm for proper step-size choices. To cope with cyber resource constraints and lack of reliable communication links, we have further extended the HVC design

to have robustness against random communication link failures and, in particular, communication-cognizant feature to account for the worst-case scenario of a total communication outage. We have extensively validated the effectiveness of the HVC design using a realistic distribution feeder under both static and dynamic testing environments.

CHAPTER 6

DISTRIBUTED FREQUENCY CONTROL FOR ISOLATED MICROGRIDS

Thus far, we have proposed decentralized voltage control designs to improve the power system stability while distribution feeders are interconnecting with the transmission grids. Meanwhile, with a high penetration of DERs in the distribution networks, it is possible to build a discrete energy system, namely, a microgrid (MG), that is capable of operating in parallel with, or independently from, the main transmission grids. To further enhance the power system reliability, MGs are foreseen to play an important role in distribution networks. However, these mostly power electronics-interfaced distributed energy resources (DERs) exhibiting low-inertia characteristic have raised major concerns over the frequency stability issues under isolated MGs [82]. Hierarchical control based on the time-scale separation principle for distributed energy resources (DERs) interface converters (DICs) has recently become a standard operational paradigm for isolated MGs [83, 84]. At the primary control level, a conventional power-frequency (P - ω) droop control assisted by the faster inner-loop controls can help reduce the frequency and/or voltage mismatch error while providing power sharing capability [85]. Meanwhile, at the secondary control level, the grid-wide information regarding the status of all DICs can be further used to minimize the mismatch error attained by a local control in a centralized fashion.

Similar to distribution networks, the traditional centralized paradigm of secondary control falls short in achieving scalability and flexibility goals of MG operations. To reduce the communication overhead and enhance DICs' plug-n-play ability, the distributed architecture has been increasingly advocated [7, 86–88]. Under a connected communication network, the proportional power sharing objective is equivalent to having pair-wise consensus between any two neighboring nodes. By considering the DICs as the first-order systems, i.e., modeling of a zero inertia generator, the distributed control problem becomes to minimize the frequency mismatch under linear consensus constraints (see [7] and references therein). Nonetheless, such approach could face two issues re-

garding control performance and stability. First, because DICs have no rotation mass, MGs with a high penetration of DICs would exhibit a low inertia characteristic leading to a poor frequency response under large disturbances [89]. Second, the formulation related to steady-state objectives does not necessarily guarantee the stability of resultant online control updates [7, 86–88]. In [86], the first-order consensus-based updates have been proposed and the stability is established by linearizing around the preferred equilibrium point. A variant of this consensus-based design is developed in [90] while ensuring economic optimality. The convergence of the consensus is guaranteed therein by the time-scale separation assumption in [87]. In addition, a ratio consensus algorithm is proposed in [88] to account for lower and upper limits of DER outputs. To the best of our knowledge, none of the existing work has explicitly included the original dynamical model of network-coupled DICs in establishing the stability conditions of consensus-based distributed control updates.

Meanwhile, a communication-based distributed frequency control framework also exposes the microgrid assets to potential malicious cyber attacks. In general, attack detection for distributed consensus algorithms under false-data injecting attacks has been considered in [91–93]. This type of attack is also related to the so-called Byzantine consensus, a fairly popular research topic in distributed computing. The goal of the Byzantine consensus algorithm is to find a near optimal solution for an optimization problem despite the presence of malicious agents [94–97]. However, in practice, this approach has significant drawbacks: i) online implementations are infeasible as diminishing step-sizes fail to incorporate the most up-to-date system operating conditions; and ii) most system operators are more concerned about determining the *identity* of a malicious agent as settling with a near optimal solution with unidentified attackers presents a notable security threat within a cyber communication network. Thus, it is extremely important to distinguish malicious agents and isolate them within control frameworks. The concern about false-data injection attacks has increasingly challenged the power grid infrastructures along with more smart grid deployments, see e.g. [98–104] and references therein, particularly for distributed power system state estimation in [99, 101]. Earlier work has considered the impacts of cyber attacks on grid monitoring or its control operations, but mostly for wide-area transmission systems. Since there exist fewer resources for cyber defense and less inertia for frequency stability in isolated microgrids, it is of higher interest to investigate the cyber-security problems therein. However, none of the

aforementioned approaches has considered (microgrid) distributed control settings, where physical grid measurements and optimization-related variables can provide additional information for attack detection.

In this chapter, we tackle the low system inertia problem by adopting the concept of a virtual synchronous generator (VSG). To this end, the DICs are designed with virtual inertia by modeling a second-order system which results in effects of damping similar to the machine inertia in [89]. Based on the virtual inertia concept, we consider the full MG network dynamics in developing the distributed control designs. The dynamics of the droop controlled DICs and power flow is first introduced. The problem to minimize the frequency error while ensuring an accurate power sharing operation is then formulated as a consensus optimization one. Motivated by the work in [72] where the power transmission system dynamics can be interpreted as a partial primal-dual (PPD) algorithm, we adopt this algorithm to solve the steady-state problem in closed-form under a connected communication graph among DICs in a MG. Interestingly, two of the update rules boil down to the network dynamics: i) the dynamics associated with power flow, and ii) the dynamics relating to the VSG-base second-order DICs systems; both are seamlessly implemented by the physical system itself. Accordingly, the proposed control design only requires the exchange of a few variables, while its stability follows directly from that for the PPD algorithm. Distinct from most of the previous work where the microgrid stability has only been ensured by the conventional small-signal analysis, the proposed control design can guarantee the microgrid stability simply through the selection of optimization stepsize. In addition to stability guarantees, the PPD-based design with localized dual variable information could improve the capability of attack detection. Earlier attack detection work for general distributed consensus methods typically requires system-wide information collection and accordingly has a very high computational burden [91, 92]. To overcome these limitations, we have developed the metrics for detection and identification only using local physical measurements and neighboring dual variable information. With the centralized energy management system (EMS) making the final decision based on localized metrics, the proposed implementation is very scalable. Therefore, compared with the previous work where the cyber-security of the microgrid secondary control has not been considered [86, 88, 105, 106], we have taken it into consideration and provided possible response strategies for practical microgrids.

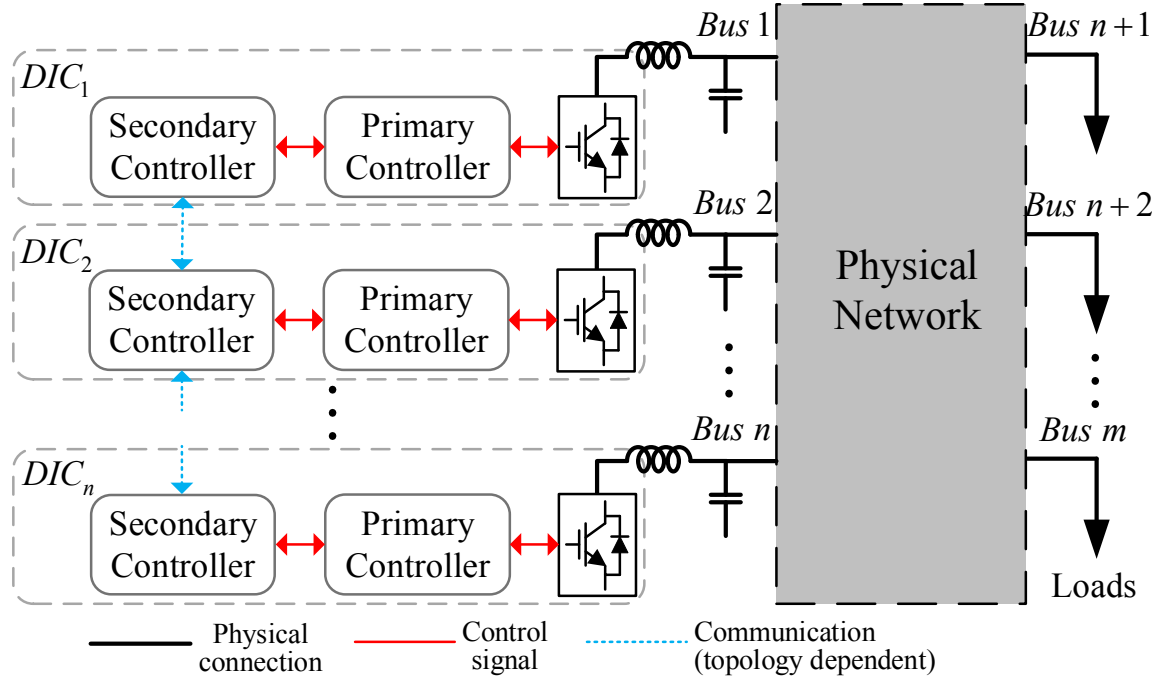


Figure 6.1: A cyber-physical MG network with DICs and related controllers

6.1 Modeling of Microgrids

Fig. 6.1 depicts an isolated MG, including all grid components such as DICs, loads, and lines in addition to controller communication of the cyber layer. This MG is modeled by a connected graph $(\mathcal{N}_M, \mathcal{E}_M)$ where the set \mathcal{N}_M consists of the subsets $\mathcal{N}_D := \{1, \dots, n\}$ and $\mathcal{N}_L := \{n+1, \dots, m\}$ representing the DIC and load buses, respectively. In addition, the set \mathcal{E}_M accounts for the line segments connecting among the buses. We use $j : i \rightarrow j$ and $k : k \rightarrow i$ respectively to represent the set of buses j such that $(i, j) \in \mathcal{E}_M$ and the set of buses k such that $(k, i) \in \mathcal{E}_M$. Per bus- i , the voltage magnitude and phase angle are denoted by v_i and θ_i , respectively. In addition, we let P_i represent the active power injection, P_i^M denote the active power rating of DIC- $i \in \mathcal{N}_D$, and P_i^L to be the active power demand of a load at bus $i \in \mathcal{N}_L$. Meanwhile, $\omega_i := (\dot{\theta}_i - \omega_b)$ is the frequency deviation where $\dot{\theta}_i := d\theta_i/dt$ is the frequency, and ω_b is the nominal frequency set-point. To develop the ensuing control design, we also make the following valid assumptions:

AS4. *The power lines are lossless and relatively short.*

AS5. *Each bus voltage magnitude v_i is constant.*

AS6. *All possible load variations under the isolated MG are supported by DICs without violating active power rating limits of DICs.*

AS7. *The active power demand $P_i^L, \forall i \in \mathcal{N}_L$ is constant while the proposed control design is being executed.*

The short distance property in Assumption 4 commonly holds for power lines in microgrids, and thus line losses are negligible compared to line flows. Constant voltage magnitude in Assumption 5 is also typically ensured through the fast inner-loop voltage control design at DICs. Together with voltage-droop control, the inner voltage control can manager DICs' reactive power output to track a reference voltage level, at much faster time-scale than that of the frequency control. This time-scale separation between frequency and voltage dynamics is well supported by earlier work on microgrid modeling [88]. Consequently, the voltage magnitude at all nodes can be assumed to be fixed, as in [86,88]. As for Assumption 6, it can be guaranteed through careful system planning at the microgrid deployment stage. Last, the constant power demand assumption in Assumption 7 comes from designing the proposed controller to be sufficiently fast to restore the system nominal frequency before another load disturbance occurs. Note that these assumptions have been made to facilitate the development of the proposed secondary frequency control design. Nonetheless, the effectiveness of this design will be demonstrated using a realistic microgrid network with e.g., the voltage control loop.

6.1.1 P - ω Droop Control of DICs

The operational objectives of a secondary active power control in isolated MGs are two-fold:

- (i) Zero frequency deviation from a nominal frequency under steady-state (synchronization).

$$\omega_1 = \omega_2 = \cdots \omega_n = 0. \quad (6.1)$$

- (ii) Autonomous active power sharing among all DICs. Specifically, DICs share the total loads

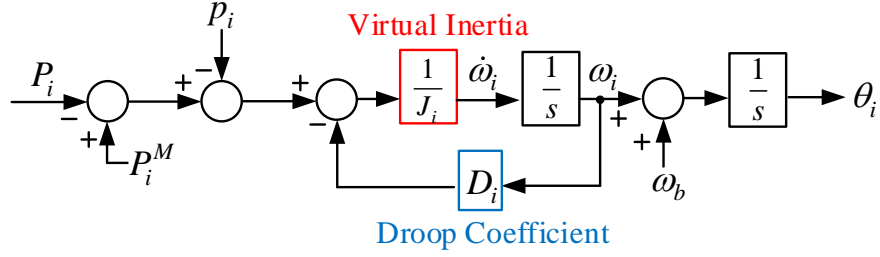


Figure 6.2: Angle dynamics of VSG governed by a second order swing equation.

according to their nominal ratings such that

$$\frac{P_1}{P_1^M} = \frac{P_2}{P_2^M} = \cdots = \frac{P_n}{P_n^M}. \quad (6.2)$$

The proliferation of DICs in MGs raises frequency stability concerns over insufficient system inertia. To address this issue, integrating the concept of a virtual synchronous generator (VSG) into the power-frequency droop control has been recently proposed in [43, 107, 108]. To this end, the control diagram in Fig. 6.2 is motivated by mimicking the dynamical second-order swing equation of a synchronous generator, as given by

$$J_i \dot{\omega}_i + D_i \omega_i = P_i^M - P_i - p_i, \forall i \in \mathcal{N}_D, \quad (6.3)$$

where $J_i > 0$ relates to the virtual moment of inertia for DIC- i , and the positive droop coefficient D_i is designed in accordance with the rating of DIC- i . We set a uniform D_i/P_i^M among DICs. Compared to conventional P - ω droop control, an additional control input p_i is appended in (6.3). Considering J_i , D_i , and P_i^M are fixed parameters based on the size of DIC- i , one can only change the operating set-point by judiciously choosing p_i , which serves as our secondary control input for the ensuing control design. Accordingly, the resultant angle θ_i is used for inner voltage and current controllers of the power electronics.

6.1.2 Microgrid Dynamics

Power loads typically depend on both bus voltage (which is assumed constant under Assumption 5) and frequency. Without loss of generality, the load- $i, \forall i \in \mathcal{N}_L$ is a frequency-sensitive load where the power consumption increases linearly with the frequency deviation ω_i and its velocity $\dot{\omega}_i$, e.g., motor-type [109]. Hence, the dynamics of a frequency dependent load is modeled as

$$J_i \dot{\omega}_i + D_i \omega_i = -P_i^L - P_i, \forall i \in \mathcal{N}_L, \quad (6.4)$$

where $J_i > 0$ and $D_i > 0$ can be thought of as the physical inertia and time constant associated with the dynamics of a load bus at bus $i \in \mathcal{N}_L$, respectively [110]. Note that both (6.3) and (6.4) are again based on the standard time-scale separation between frequency and voltage dynamics. Hence, the frequency is regulated to adjust active power injections.

As for the microgrid network power flow dynamics, based on Assumption 4, we separate the bus injection P_i into branch flows as follows:

$$P_i(\mathbf{f}) = \sum_{j=1}^n \frac{v_i v_j}{x_{ij}} \sin(\theta_i - \theta_j) = \sum_{j:i \rightarrow j} f_{ij} - \sum_{k:k \rightarrow i} f_{ki}, \quad (6.5)$$

where $\mathbf{f} =: \{f_{ij}\}_{\forall (i,j) \in \mathcal{E}_M}$ with f_{ij} denoting the line power flow from bus i to bus j , and x_{ij} is the reactance of line (i, j) . Additionally, the angular difference between any power line is relatively small. Thus, the dynamics of the branch flow between node- i and node- j can be further characterized by

$$\dot{f}_{ij} = b_{ij} (\omega_i - \omega_j), \quad (6.6a)$$

$$b_{ij} = \frac{v_i v_j}{x_{ij}} \cos(\theta_i^0 - \theta_j^0), \quad (6.6b)$$

where b_{ij} is a constant with θ_i^0 being the nominal phase angle at bus i [111]; see Appendix B for detailed derivations. Interestingly, it turns out that by specially designing our ensuing distributed control, the update rules would include the network dynamics of (6.6a) which is seamlessly implemented by the physical system itself.

6.1.3 Equilibrium Points of Droop Controlled DICs

By concatenating all scalar variables into vector form, the desired equilibrium point fulfilling operational objectives of the droop control can be characterized into the following proposition:

Proposition 5. *An equilibrium point $(\omega^*, \mathbf{P}^*(\mathbf{f}^*), \mathbf{p}^*)$ that fulfils objective (6.2) must also satisfy*

$$\frac{p_i^*}{D_i} = \frac{p_j^*}{D_j}, \forall i, j \in \mathcal{N}_D.$$

This proposition is based on the synchronization of all DICs. By considering (6.6a) at the equilibrium point, we have

$$\dot{f}_{ij}^* = b_{ij} (\omega_i^* - \omega_j^*) = 0, \forall i, j \in \mathcal{N}_M.$$

This implies that a system-wise synchronization frequency is attained, i.e., $\omega_i^* = \omega_j^*, \forall i, j \in \mathcal{N}_M$. Furthermore, dividing both side of (6.3) by P_i^M , we have

$$\frac{P_i^*}{P_i^M} = 1 - \frac{D_i}{P_i^M} \frac{p_i^*}{D_i} - \frac{D_i}{P_i^M} \omega_i^*, \forall i \in \mathcal{N}_D.$$

Because of the aforementioned uniform setting of $D_i/P_i^M, \forall i \in \mathcal{N}_D$, achieving an accurate power sharing operation among DICs leads to the result in Prop. 5.

6.1.4 Consensus Optimization

We assume a connected communication network $\mathcal{G}_D := (\mathcal{N}_D, \mathcal{E}_D)$ where the set \mathcal{E}_D accounts for the communication links connecting among the DIC buses. Based on Prop. 5 under quasi-steady-state (i.e., $\dot{\omega} = 0$), the secondary control problem can be cast as a *consensus optimization* one:

$$(\boldsymbol{\omega}^*, \mathbf{f}^*, \mathbf{p}^*) := \arg \min_{\boldsymbol{\omega}, \mathbf{f}, \mathbf{p}} \sum_{i=1}^n \frac{D_i}{2} \omega_i^2 \quad (6.7a)$$

$$\text{subject to } D_i \omega_i = P_i^M - P_i(\mathbf{f}) - p_i, \forall i \in \mathcal{N}_D \quad (6.7b)$$

$$D_i \omega_i = -P_i^L - P_i(\mathbf{f}), \forall i \in \mathcal{N}_L \quad (6.7c)$$

$$\frac{p_i}{D_i} = \frac{p_j}{D_j}, \forall (i, j) \in \mathcal{E}_D. \quad (6.7d)$$

The constraints (6.7b) ensure the DICs to mimic the dynamical second-order swing equation of synchronous generators and thus are coupled with the network frequency. Similarly, the characteristic of a frequency-dependent load is captured by (6.7c). Additionally, the equality constraints (6.7d) are equivalent to satisfying Prop. 5 under a connected communication graph \mathcal{G}_D . This is a quadratic program and thus can be solved using off-the-shelf convex solvers. However, the difficulty in solving (6.7) lies in that active power injection $\mathbf{P}(\mathbf{f})$ is dynamical and dependent on the power network couplings. It is imperative to find a method which respects the MG system dynamics to steer the control to the optimum. In our previous work [7], this issue is tackled by adopting a feedback approach to account for system dynamics. Interestingly, as detailed soon, our proposed distributed partial primal-dual (PPD)-based control design not only effectively attains a zero steady-state frequency deviation but also seamlessly integrates the MG network dynamics into the control algorithm.

6.2 Distributed Secondary Control Design

This section introduces the proposed distributed secondary control framework. Introducing Lagrangian multipliers $\boldsymbol{\lambda} := \{\lambda_i\}_{\forall i \in \mathcal{N}_M}$ and $\boldsymbol{\eta} := \{\eta_{ij}\}_{\forall (i,j) \in \mathcal{E}_D}$ for the equality constraints in (6.7),

we form the following Lagrangian function:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\omega}, \mathbf{f}, \mathbf{p}, \boldsymbol{\eta}, \boldsymbol{\lambda}) = & \sum_{i=1}^n \frac{D_i}{2} \omega_i^2 + \sum_{i \in \mathcal{N}_D} \left\{ \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij} \left(\frac{p_i}{D_i} - \frac{p_j}{D_j} \right) \right. \\
& + \lambda_i \left[P_i^M - \left(\sum_{j:i \rightarrow j} f_{ij} - \sum_{k:k \rightarrow i} f_{ki} \right) - p_i - D_i \omega_i \right] \Big\} \\
& + \sum_{i \in \mathcal{N}_L} \left\{ \lambda_i \left[-P_i^L - \left(\sum_{j:i \rightarrow j} f_{ij} - \sum_{k:k \rightarrow i} f_{ki} \right) - D_i \omega_i \right] \right\}. \quad (6.8)
\end{aligned}$$

Based on the Lagrangian function (6.8), we adopt the partial-primal dual (PPD) algorithm which works by cyclically *partially* and fully minimizing the primal variable $(\boldsymbol{\omega}, \mathbf{f}, \mathbf{p})$ while performing gradient ascent-based update on the dual variable $(\boldsymbol{\eta}, \boldsymbol{\lambda})$ [72]. To reduce communication cost by keeping the dual variables as local ones for each DIC- i , we modify the standard dual-ascent update design by letting the dual variables associated to communication link $(i, j) \in \mathcal{E}_D$ of DIC- i to have the property of $\eta_{ij} = -\eta_{ji}$. This would always hold if $\boldsymbol{\eta}$ is initialed to be zero. The resultant PPD algorithm of (6.7) then becomes

$$\omega_i = \lambda_i, \forall i \in \mathcal{N}_M \quad (6.9a)$$

$$\dot{f}_{ij} = -\epsilon_{f_{ij}} (-\lambda_i + \lambda_j), \forall (i, j) \in \mathcal{E}_M \quad (6.9b)$$

$$\dot{p}_i = -\epsilon_{p_i} \left(2D_i^{-1} \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij} - \lambda_i \right), \forall i \in \mathcal{N}_D \quad (6.9c)$$

$$\dot{\eta}_{ij} = \epsilon_{\eta_{ij}} \left(\frac{p_i}{D_i} - \frac{p_j}{D_j} \right), \forall (i, j) \in \mathcal{E}_D \quad (6.9d)$$

$$\dot{\lambda}_i = \epsilon_{\lambda_i} \left[P_i^M - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - p_i - D_i \omega_i \right], \forall i \in \mathcal{N}_D \quad (6.9e)$$

$$\dot{\lambda}_i = \epsilon_{\lambda_i} \left[-P_i^L - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - D_i \omega_i \right], \forall i \in \mathcal{N}_L, \quad (6.9f)$$

where $\epsilon_{f_{ij}}, \epsilon_{p_i}, \epsilon_{\eta_{ij}}$, and ϵ_{λ_i} are positive parameter choices. Note that we solve $\min_{\omega_i} \frac{D_i}{2} \omega_i^2 - \lambda_i D_i \omega_i, \forall i \in \mathcal{N}_M$ to obtain (6.9a) rather than follow the primal gradient algorithm with respect to ω_i . Thus, the so-termed *partial* primal-dual algorithm is adopted.

Applying linear transform of (6.9a) from λ_i to ω_i , we have

$$\begin{aligned} \dot{f}_{ij} &= \epsilon_{f_{ij}} (\omega_i - \omega_j), \forall i \in \mathcal{N}_M \\ \dot{\omega}_i &= \epsilon_{\lambda_i} \left[P_i^M - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - p_i - D_i \omega_i \right], \forall i \in \mathcal{N}_D \\ \dot{\omega}_i &= \epsilon_{\lambda_i} \left[-P_i^L - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - D_i \omega_i \right], \forall i \in \mathcal{N}_L. \end{aligned}$$

By selecting control parameter values specially to adapt to MG network dynamics, our PPD-based algorithm (6.9) features the dynamics of both microgrid network and VSG-based DICs. Let $\epsilon_{f_{ij}} = b_{ij}, \forall (i, j) \in \mathcal{E}_M$ and $\epsilon_{\lambda_i} = \frac{1}{J_i}, \forall i \in \mathcal{N}_M$. The algorithm (6.9) becomes

$$\dot{f}_{ij} = b_{ij} (\omega_i - \omega_j), \forall (i, j) \in \mathcal{E}_M \quad (6.11a)$$

$$\dot{p}_i = -\epsilon_{p_i} \left(2D_i^{-1} \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij} - \omega_i \right), \forall i \in \mathcal{N}_D \quad (6.11b)$$

$$\dot{\eta}_{ij} = \epsilon_{\eta_{ij}} \left(\frac{p_i}{D_i} - \frac{p_j}{D_j} \right), \forall (i, j) \in \mathcal{E}_D \quad (6.11c)$$

$$\dot{\omega}_i = \frac{1}{J_i} \left[P_i^M - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - p_i - D_i \omega_i \right], \forall i \in \mathcal{N}_D \quad (6.11d)$$

$$\dot{\omega}_i = \frac{1}{J_i} \left[-P_i^L - \sum_{j:i \rightarrow j} f_{ij} + \sum_{k:k \rightarrow i} f_{ki} - D_i \omega_i \right], \forall i \in \mathcal{N}_L. \quad (6.11e)$$

The dynamics of (6.11a), (6.11d), and (6.11e) exactly corresponds to that in MG network and droop controlled DICs. This feature is very attractive as parts of the algorithm seamlessly incorporate MG network dynamics which is implemented by the physical system itself. Therefore, the dynamical model of network-coupled DICs and power flow is inherently accounted for. Meanwhile, it turns out that the dynamics of (6.11b)-(6.11c) resonates with the distributed-averaging proportional-integral (DAPI) controller where the stability analysis is based on the classic *Kuramoto* model and phase cohesiveness principle in [86]. However, the DAPI controller is developed for the inertialess microgrid whereas the algorithm (6.11) incorporates virtual inertia for improving frequency responses. Accordingly, our design has the capability to quickly regulate

network frequency under large and rapid disturbances and preserve the power sharing property while enhancing frequency stability. To sum up, the PPD-based algorithm (6.11) guarantees the convergence to $(\omega^*, \mathbf{f}^*, \mathbf{p}^*)$ under proper parameter values of ϵ_{p_i} and $\epsilon_{\eta_{ij}}$.

Lemma 2. $(\omega^*, \mathbf{f}^*, \mathbf{p}^*)$ is an unique equilibrium point for (6.7) only if $\omega_i^* = 0, \forall i \in \mathcal{N}_M$, $P_i^M = P_i^*(\mathbf{f}^*) + p_i^*, \forall i \in \mathcal{N}_D$, and $P_i^L = -P_i^*(\mathbf{f}^*), \forall i \in \mathcal{N}_L$.

Proof: Based on (6.11a), we have a system-wide synchronization frequency, i.e., $\omega_i = \omega_j, \forall i, j \in \mathcal{N}_M$ at the equilibrium point. This corroborates the results in Prop. (5). Furthermore, if $\boldsymbol{\eta}$ is initialed to be zero, $\sum_{i=1}^n \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij}^* = 0$ always holds because of the Laplacian for graph \mathcal{G}_D . Additionally, from (6.11b), we have $2D_i^{-1} \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij}^* - \omega_i^* = 0, \forall i \in \mathcal{N}_D$. Summing over all the DIC buses in \mathcal{N}_D under a synchronized frequency leads to $\omega^* = 0$. Meanwhile, the dynamics of (6.11d) implies $P_i^M - \sum_{j:i \rightarrow j} f_{ij}^* + \sum_{k:k \rightarrow i} f_{ki}^* - p_i^* - D_i \omega_i^* = 0, \forall i \in \mathcal{N}_D$. Because of $\omega^* = 0$ and the linear mapping in (6.5), we have $P_i^M = P_i^*(\mathbf{f}^*) + p_i^*, \forall i \in \mathcal{N}_D$. Similarly, $P_i^L = -P_i^*(\mathbf{f}^*), \forall i \in \mathcal{N}_L$ holds. Last, as (6.7a) is strongly convex on $\boldsymbol{\omega}$, $(\omega^*, \mathbf{f}^*, \mathbf{p}^*)$ indeed is the unique saddle point of (6.8) by applying the strong duality theorem [57]. ■

For the updates in (6.11b) and (6.11c), a digital control implementation would convert them to the discrete-time counterpart for each DIC- i . Nonetheless, it is difficult to measure ω_i locally due to fast power electronics' switching frequency. Under the notion of a much faster dynamics of the inverter-based DICs than that of secondary control updates (6.11b)-(6.11c), the dynamics of $\dot{\omega}_i$ is insignificant to the secondary updates. Based on (6.3), we can approximate local frequency deviation as

$$\omega_i \approx \frac{P_i^M - P_i - p_i}{D_i}, \forall i \in \mathcal{N}_D. \quad (6.12)$$

Thus, by denoting S_t as the sampling time and letting $\frac{S_t \epsilon_{p_i}}{D_i} = 1, \forall i \in \mathcal{N}_D$, the control updates at $(k+1)$ -st iteration for DIC- $i \in \mathcal{N}_D$ with an uniform step-size $\epsilon := \epsilon_{\eta_{ij}} S_t, \forall (i, j) \in \mathcal{E}_D$ and frequency approximation (6.12) become

$$x_i^{k+1} = -2D_i^{-1} \left(\sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij}^k \right) + c_i^k, \forall i \in \mathcal{N}_D \quad (6.13a)$$

$$\eta_{ij}^{k+1} = \eta_{ij}^k + \epsilon(x_i^{k+1} - x_j^{k+1}), \forall (i, j) \in \mathcal{E}_D, \quad (6.13b)$$

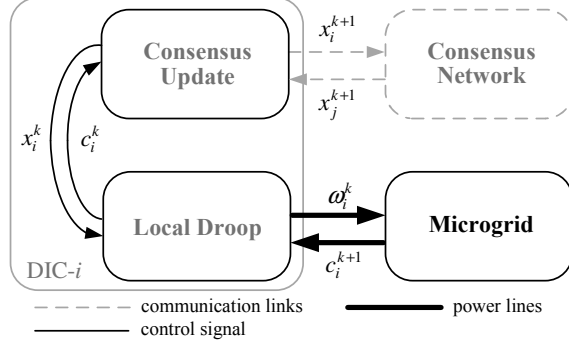


Figure 6.3: Operation of DIC- i and its interaction with the microgrid and the consensus network under the proposed control design.

where $x_i := p_i/D_i$ and $c_i := (P_i^M - P_i)/D_i$. Note that the input variable c_i is locally measurable at DIC- i . Under a connected communication network \mathcal{G}_D , the overall operation of DIC- i at k iteration for the proposed distributed control can be illustrated in Fig. 6.3. Given the local dual variable $\{\eta_{ij}^k\}_{\forall(i,j) \in \mathcal{E}_D}$, the consensus update unit takes in the current local measurement c_i^k from the local droop controller and updates its primal variable x_i^{k+1} as (6.13a) which is broadcast to the neighboring DICs. Next, after receiving $\{x_j^{k+1}\}_{\forall j \in \{j | (i,j) \in \mathcal{E}_D\}}$ via communication links- $(i,j) \in \mathcal{E}_D$ from the neighboring DICs, we update $\{\eta_{ij}\}_{\forall(i,j) \in \mathcal{E}_D}$ according to (6.13b). Meanwhile, the local droop unit adjusts its output frequency ω_i^k based on (6.3). The microgrid reacts to the variations in frequencies of DICs and reaches a new power flow dispatch according to (6.5). This results in the updated measurement c_i^{k+1} which is used for the next iteration. Thanks to the communication network, each DIC is able to obtain neighbors' information, which means the proposed secondary control (6.13) is fully distributed. To sum up, our distributed control implementation is scalable and flexible with respect to the size of the network while enjoying minimal overhead in sensing, communication, and computation.

6.2.1 Choice of Step-size

Albeit the convergence of the proposed control algorithm (6.13) can always be guaranteed, its implementation in microgrid control requires the proper design of update step-size ϵ . To this end, the following remarks of P - ω droop controlled DICs under isolated microgrids are introduced.

Remark 7 (Center-Of-Mass Frequency). We further assume that loads are much less sensitive to

frequency variation compared to DICs, i.e., $D_i \ll D_j, \forall i \in \mathcal{N}_L, j \in \mathcal{N}_D$. Hence, it is possible to directly relate the power balance to the system frequency in isolated microgrids, which is independent of state and can be determined directly from the power injections [112]. This feature is instrumental for ensuing attack scenario analysis. Accordingly, the joint behavior of all DIC frequencies follows the *center-of-mass* frequency [113]

$$\omega_c \approx \frac{\sum_{i=1}^n D_i \omega_i}{\sum_{i=1}^n D_i} = \frac{\sum_{i=1}^n (P_i^M - P_i - p_i)}{\sum_{i=1}^n D_i}. \quad (6.14)$$

By implementing (6.13a) and (6.13b) to the DICs, another remark regarding the characteristics of ω_c is made.

Remark 8 (Characteristics of ω_c). Initializing $\boldsymbol{\eta}^0 = \mathbf{0}$, we sum (6.13a) over all DICs

$$\sum_{i=1}^n \frac{D_i}{2} (c_i^k - x_i^{k+1}) = \sum_{i=1}^n \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij}^k = 0. \quad (6.15)$$

This leads to the following two observations:

- (i) The first updates of all x_i take the mismatch between power outputs of DICs as the initial condition, i.e.,

$$\sum_{i=1}^n \frac{D_i}{2} x_i^1 = \sum_{i=1}^n \frac{D_i}{2} c_i^0. \quad (6.16)$$

- (ii) Substituting x_i^{k+1} in (6.13a) into (6.14), ω_c at $(k+1)$ -st iteration is

$$\omega_c^{k+1} = \frac{\sum_{i=1}^n \left[D_i (c_i^{k+1} - c_i^k) + 2 \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \eta_{ij}^k \right]}{\sum_{i=1}^n D_i}. \quad (6.17)$$

Interestingly, any power imbalance is compensated after one iteration of the proposed update design (6.13). Under Assumption 7 of a constant loading in addition to its less sensitivity to frequency variations, we have $\sum_{i=1}^n P_i^{k+1} \approx \sum_{i=1}^n P_i^k$. This equivalently leads to $(c_i^{k+1} - c_i^k) \approx 0, \forall i \in \mathcal{N}_D$. Thus, $\omega_c^k \approx 0$ is assured for $k \geq 2$, and any changes in \mathbf{p} has minimal effect on steady-state frequency ω_c^k .

To formally analyze the step-size design, we define a vector $\boldsymbol{\zeta}^k := \{\zeta_i^k\}_{\forall i \in \mathcal{N}_D}$ where $\zeta_i^k := \sum_{j \in \{j \mid (i,j) \in \mathcal{E}_D\}} \mu_{ij}^k$ and a weighting matrix $\mathbf{D} := \text{diag}(D_1, \dots, D_n)$, the updates in (6.13a) and

(6.13b) become

$$\mathbf{x}^{k+1} = (-2\epsilon \mathbf{D}^{-1} \mathbf{L} + \mathbf{I}_n) \mathbf{x}^k - 2\mathbf{D}^{-1} \boldsymbol{\zeta}^{k-1} + (\mathbf{c}^k - \mathbf{x}^k), \quad (6.18a)$$

$$\boldsymbol{\zeta}^{k+1} = \boldsymbol{\zeta}^k + \epsilon \mathbf{L} \mathbf{x}^{k+1}, \quad (6.18b)$$

where \mathbf{L} and \mathbf{I}_n are the Laplacian of graph \mathcal{G}_D and an $n \times n$ identity matrix, respectively. Substituting $\mathbf{c}^{k-1} = \mathbf{x}^k + 2\mathbf{D}^{-1} \boldsymbol{\zeta}^{k-1}$ from (6.13) into (6.18a), we have

$$\mathbf{x}^{k+1} = \mathbf{W} \mathbf{x}^k + (\mathbf{c}^k - \mathbf{c}^{k-1}), \quad (6.19)$$

where $\mathbf{W} := (-2\epsilon \mathbf{D}^{-1} \mathbf{L} + \mathbf{I}_n)$. The linear dynamical system in (6.19) can be viewed as a consensus iteration of \mathbf{x}^k with a disturbance depended on $(\mathbf{c}^k - \mathbf{c}^{k-1}) = \mathbf{D}^{-1}(\mathbf{P}^{k-1} - \mathbf{P}^k)$. Based on Lemma 2, the disturbance is bounded by the total load variations and diminishing as $k \rightarrow \infty$. Selecting the step-size ϵ properly ensures matrix \mathbf{W} to have all but one of the eigenvalues strictly within the unit circle. Accordingly, the iterate \mathbf{x}^{k+1} under the update in (6.19) would converge to the average consensus vector $\bar{\mathbf{x}}$. This consensus vector can be determined by further examining the matrix \mathbf{W}^k , which denotes the constant matrix \mathbf{W} to the k -th power. To this end, the work in [114] with slight modifications leads to

$$\lim_{k \rightarrow \infty} \mathbf{W}^k = \frac{\mathbf{1}\mathbf{1}^T}{\mathbf{1}^T \mathbf{D} \mathbf{1}} \mathbf{D}, \quad (6.20)$$

where vector $\mathbf{1}$ denotes the all one-vector while T is the transpose operation. This differs from [114] because of the weighting matrix \mathbf{D}^{-1} in \mathbf{W} . Since the disturbance $(\mathbf{c}^k - \mathbf{c}^{k-1})$ in (6.19) is bounded and diminishing, the consensus iteration of (6.20) becomes

$$\lim_{k \rightarrow \infty} \mathbf{x}^k = \lim_{k \rightarrow \infty} \mathbf{W}^k \mathbf{x}^1 = \frac{\mathbf{1}\mathbf{1}^T}{\mathbf{1}^T \mathbf{D} \mathbf{1}} \mathbf{D} \mathbf{x}^1. \quad (6.21)$$

According to (6.16), it suggests that the initial condition for x_i takes the form of

$$\frac{\sum_{i=1}^n D_i x_i^1}{\sum_{i=1}^n D_i} = \mathbf{1}^T \frac{\mathbf{D} \mathbf{x}^1}{\mathbf{1}^T \mathbf{D} \mathbf{1}}. \quad (6.22)$$

Thus, the average consensus \bar{x} is given by

$$\bar{x} = \mathbf{1}^T \frac{D\mathbf{x}^1}{\mathbf{1}^T D \mathbf{1}} \mathbf{1} = \frac{\mathbf{1}\mathbf{1}^T}{\mathbf{1}^T D \mathbf{1}} D\mathbf{x}^1, \quad (6.23)$$

which is the same as the steady-state value of \mathbf{x}^k shown in (6.21).

6.3 Attack Models and Countermeasures

In this section, the attack models against the proposed distributed control are introduced, and the ensuing attack detection and localization strategies are offered. Such strategies are anomaly tests based on examining the secondary control objectives under steady-state, namely, $\omega_c = 0$ and $p_i/D_i = p_j/D_j, \forall i, j \in \mathcal{N}_D$. In this work, we consider constant malicious communication signal inputs which attempt to alter the microgrid operating points to be the worst-case attack scenario. As detailed soon, such attack can effectively drive the frequency deviation ω_i everywhere away from zero, endangering stability. Regarding the attack strategy other than constant signals, it is possible to improve its design assuming the attacker had the full knowledge of the communication graph as well as the detection mechanism. This scenario is, nonetheless, unlikely from the practical standpoint, as almost all recent cyber attacks in the energy sector are agnostic to the full network information; see e.g., [115]. Studying the interactions between the design of attack strategies and counter-measure algorithms is an important direction for our future work. Based on the complexity of the malicious inputs, the attack models are categorized into link and node attack scenarios which both may further be extended to individual and coordinated cases. As the malicious inputs are implemented to alter the operation of consensus iterations, the formulation in (6.19) is adopted to constitute the basis for developing the ensuing attack models.

6.3.1 Link Attack

We first study the link attack scenario, where the malicious inputs are applied only to the information sent to specific neighbors of the attacked DICs. Given the undirected communication link- $(i, j) \in \mathcal{E}_D$, we define the attack signal ℓ_{ij} as the malicious input sent from DIC- j to DIC- i .

Per DIC- $i \in \mathcal{N}_D$, the update (6.13b) becomes

$$\eta_{ij}^{k+1} = \eta_{ij}^k + \epsilon[x_i^{k+1} - (x_j^{k+1} + \ell_{ij})], \forall j \in \{j | (i, j) \in \mathcal{E}_D\}. \quad (6.24)$$

By reformulating (6.13a) and (6.24) into the compact form, (6.19) is rewritten as

$$\mathbf{x}^{k+1} = \mathbf{W}\mathbf{x}^k + 2\epsilon\mathbf{D}^{-1}\tilde{\mathbf{L}}\boldsymbol{\ell} + (\mathbf{c}^k - \mathbf{c}^{k-1}), \quad (6.25)$$

where vector $\boldsymbol{\ell} := \{\ell_{ij}\}_{\forall(i,j) \in \mathcal{E}_D}$, and $\tilde{\mathbf{L}}$ specifies the malicious link indices of $\boldsymbol{\ell}$. The attack vector $2\epsilon\mathbf{D}^{-1}\tilde{\mathbf{L}}\boldsymbol{\ell}$ would equivalently offset the consensus vector $\bar{\mathbf{x}}$. Accordingly, $\boldsymbol{\zeta}$, the sum of dual variables, is manipulated by the malicious inputs. The change of ω_c^k can be derived from (6.17) as

$$\omega_c^k - \omega_c^{k-1} = \frac{\mathbf{1}^T(\boldsymbol{\zeta}^{k-1} - \boldsymbol{\zeta}^{k-2})}{\mathbf{1}^T\mathbf{D}\mathbf{1}} = \frac{-\epsilon\mathbf{1}^T\tilde{\mathbf{L}}\boldsymbol{\ell}}{\mathbf{1}^T\mathbf{D}\mathbf{1}}, k \rightarrow \infty. \quad (6.26)$$

Interestingly, one may observe the structure of (6.26) and alter the frequency by coordinating the link attack inputs. This would lead to the following two cases for a sufficiently large value of k :

1. Individual link attack: Malicious inputs ℓ_{ij} 's are appended to the information sent from DIC- j individually, resulting in $(\omega_c^k - \omega_c^{k-1}) \neq 0$ and $\omega_c^k \neq 0$.
2. Coordinated link attack: Multiple malicious inputs are deliberately deployed such that $(\omega_c^k - \omega_c^{k-1}) = 0$ and thus $\omega_c^k = 0$.

Note that the link-based attack input is not in the null space of \mathbf{L} in \mathbf{W} . In fact, the cases where $\tilde{\mathbf{L}}\boldsymbol{\ell}$ lies in the null space of \mathbf{L} should be categorized into node attacks, which is analyzed in Sec. 6.3.2. The system (6.25) cannot reach the consensus among \mathbf{x} in both the individual and coordinated cases. Albeit the system frequency behaves differently in these scenarios, the objective of power sharing cannot be achieved. This would potentially lead to the violation of DIC ratings and thus damage the equipment. It is imperative to identify the malicious links and isolate them from the network. To this end, we check the values of the dual variables which provide essential information on cyber-physical interactions. Because of not achieving a consensus, the

dual variables of the proposed algorithm would keep integrating, i.e.

$$(\eta_{ij}^{k+1} - \eta_{ij}^k) \neq 0, \forall (i, j) \in \mathcal{E}_D. \quad (6.27)$$

Given this dynamics, we are able to detect such attack by checking the convergence of the dual variables. Furthermore, to localize the link attack, we adopt the fact based on (6.24) that $|\eta_{ij}^k|$ associated with the malicious link is always greater than those without the malicious inputs. Consequently, each node would be able to localize the attack by comparing such value among the ones related to the neighboring nodes.

6.3.2 Node Attack

Under the node attack scenario, the malicious inputs are applied to the information sent to attacked nodes' neighbors, as well as the attacked nodes themselves. By denoting u_i as the malicious input at DIC- $i \in \mathcal{N}_D$, the update (6.13b) for DIC- i becomes

$$\eta_{ij}^{k+1} = \eta_{ij}^k + \epsilon[(x_i^{k+1} + u_i) - (x_j^{k+1} + u_j)], \forall i, j \in \mathcal{N}_D. \quad (6.28)$$

Reformulating (6.13a) and (6.28) into the compact form and denoting $\mathbf{u} := \{u_i\}_{\forall i \in \mathcal{N}_D}$, we have

$$\mathbf{x}^{k+1} = \mathbf{W}(\mathbf{x}^k + \mathbf{u}) + (\mathbf{c}^k - \mathbf{c}^{k-1}). \quad (6.29)$$

Different from the link attack scenario, $(\mathbf{x} + \mathbf{u})$ lies in the null space of \mathbf{L} , and thus the consensus is achieved. Nonetheless, instead of the consensus vector $\bar{\mathbf{x}}$ of (6.23), a false consensus dictating by \mathbf{u} is attained from (6.29). Recall from Prop. 5 that the proportional power sharing can always be achieved whenever $p_i/D_i, \forall i \in \mathcal{N}_D$ coincides across the network. Hence, the steady-state DIC power output $P_i, \forall i \in \mathcal{N}_D$ is not affected by node-based malicious attacks. Regarding the frequency deviation, similar to the link attack scenario, we examine ω_c^k based on (6.14) and Remark 8 as

$$\omega_c^k = \frac{\mathbf{1}^T \mathbf{D}[\mathbf{c}^k - (\mathbf{x}^k + \mathbf{u})]}{\mathbf{1}^T \mathbf{D} \mathbf{1}} = \frac{-\mathbf{1}^T \mathbf{D} \mathbf{u}}{\mathbf{1}^T \mathbf{D} \mathbf{1}}, k \rightarrow \infty. \quad (6.30)$$

Given a sufficient large value of k , this leads to the following two different cases depending on the malicious signal u :

1. Individual node attack: The attack input is implemented individually and results in a deviated system frequency, i.e., $\omega_c^k \neq 0$.
2. Coordinated node attack: Multiple attack inputs are purposely implanted such that $\omega_c^k = 0$.

The consensus among DICs is achieved in both cases. Meanwhile, there is no dynamic change that can be utilized as a malicious detection/localization index in the system. Fortunately, the information from the inherent dual variable characteristics contains cognizance regarding such attack. Considering a pair of normal DIC- i and attacked DIC- j under steady-state (i.e., $k \rightarrow \infty$), this leads to $x_i^k = x_j^k + u_j$ and $c_i^k = c_j^k$. Thus, (6.14) becomes $\omega_c^k = c_i^k - x_i^k = c_j^k - (x_j^k + u_j)$. In addition, based on (6.13a), we have

$$\zeta_i^k = \frac{D_i}{2} \omega_c^k, \quad (6.31a)$$

$$\zeta_j^k = \frac{D_j}{2} (c_j^k - x_j^k - u_j + u_j) = \frac{D_j}{2} (\omega_c^k + u_j), k \rightarrow \infty. \quad (6.31b)$$

Assuming $\{\zeta_j^k, D_j\}_{\forall j \in \mathcal{N}_i}$ is known at DIC- i , it is plausible to localized the node attack by comparing these values. Nonetheless, this would result in a higher communication cost due to communicating these additional variables. Similar to the aforementioned link attack scenarios, $|\eta_{ij}^k|$ associated with the malicious nodes in (6.28) would be larger compared to the ones without the attacks. This additional knowledge offers the ability of localizing node attacks using only the inherent information of the consensus update design.

6.3.3 Detection/Localization Strategies

Thanks to the information from the dual variables, the following conclusions are made for DIC- i within a given detection time window while experiencing an attack from either its direct link- $(i, j) \in \mathcal{E}_D$ or neighboring DIC- $j \in \{j | (i, j) \in \mathcal{E}_D\}$:

1. $\eta_{ij}^k \neq 0$ under either malicious signals of ℓ_{ij} or u_j according to (6.27) and (6.31). Also, $|\eta_{ij}^k|$ is larger than neighboring ones associated with non-malicious inputs.

Algorithm 3 Attack Detection/Localization for DIC- i

Input: $\{\eta_{ij}^{k+1}, \eta_{ij}^k\}_{\forall (i,j) \in \mathcal{E}_D}$ **Require:** DIC- i has at least 2 neighbors

```
1: for  $j \in \{j | (i, j) \in \mathcal{E}_D\}$  do
2:   Compute  $F_{ij}^k = |\eta_{ij}^k|$  and  $\Delta_{ij}^k = |\eta_{ij}^{k+1} - \eta_{ij}^k|$ ;
3:   if  $\max(F_{ij}^k) > 0$  then
4:     if  $\Delta_{ij}^k > 0$  then
5:       Link- $(i, j)$  is recognized as malicious;
6:     else
7:       DIC- $j$  is recognized as malicious;
8:     end if
9:   end if
10: end for
11: Report events to microgrid energy management system.
```

2. η_{ij}^k diverges under the malicious signal ℓ_{ij} whereas η_{ij}^k converges to a fixed point for some node attack scenario.

Based on these facts, the following detection indices for each DIC- i are introduced:

$$\begin{cases} F_{ij}^k = |\mu_{ij}^k| \\ \Delta_{ij}^k = |\mu_{ij}^{k+1} - \mu_{ij}^k|, \end{cases} \quad (6.32)$$

which both can be obtained locally based on neighboring information. Note that when $\max(F_{ij}^k) = 0$, there is no anomaly in the system. Given these indices, the overall malicious attack detection and localization strategies are tabulated in Algorithm 3. To perform appropriate attack isolation actions, the adverse event report is sent to an energy management system (EMS), which is inherent for typical microgrid structure for decision making.

6.3.4 Decision Making in Energy Management System

In our proposed cyber-security framework, the task of isolating the malicious node or link from the consensus algorithm is handled by a centralized agent such as EMS which is an essential component in typical microgrid frameworks. Communication links between the EMS and DICs only require very low bandwidth and are highly reliable. Thus, based on the report events from all DICs, the proposed decision making strategies would therefore be efficiently implemented in

Algorithm 4 Decision Making in the EMS

Input: Events reported from all DICs**Require:** Reports from all DICs received

```
1: for  $i \in \mathcal{N}_D$  and  $|\{j | (i, j) \in \mathcal{E}_D\}| \geq 2$  do
2:   Link attack events:
3:   if DIC- $i$  reports link- $(i, j)$  as malicious then
4:     Examine the flag raised by DIC- $j$ ;
5:     if DIC- $j$  also indicates link- $(i, j)$  as malicious then
6:       Remove link- $(i, j)$  from consensus network;
7:     end if
8:   end if
9:   Node attack events:
10:  if DIC- $i$  reports DIC- $j$  as malicious then
11:    Examine DIC- $m \in \{m | (j, m) \in \mathcal{E}_D\}$ ;
12:    if DIC- $m, \forall m \in \{m | (j, m) \in \mathcal{E}_D\}$  reports DIC- $j$  as malicious then
13:      Isolate DIC- $j$  from the consensus network;
14:    end if
15:  end if
16: end for
17: Reconfigure the consensus network in case of disconnected graph after the isolation action.
```

existing microgrids.

The detectability of malicious attacks for the system (6.19) has been greatly studied in [91, 92], and thus is beyond the scope of this work. Generalizing our current framework to include detectability will be a focus of our future study. In the present context, we assume that the attack inputs in the microgrid are sparse and do not exceed the theoretical bound under which the detection and localization strategies are no longer feasible. Accordingly, the attack isolation strategies involve the following two stages: (i) identify the source of link or node attack; (ii) isolate the malicious link or node from the consensus network. Motivated by the work of [116], the mechanism behind the former stage adopts the characteristics of information propagation in typical consensus networks. Thus, once the malicious source is pinpointed, the attack isolation intelligence would either command to switch attacked nodes to the local primary control under a node attack scenario or remove the malicious links from the communication graph under a link attack scenario. Under either case, the reconfiguration of consensus network would be carried out if the graph is no longer connected after the attack isolation. To sum up, the decision-making process tabulated in Algorithm 4 is capable of isolating malicious attacks and its effectiveness will be validated extensively

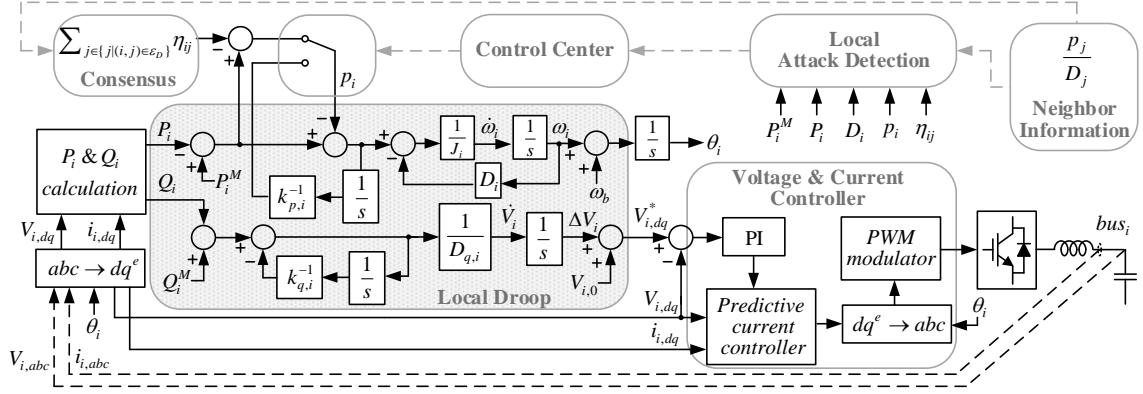


Figure 6.4: Proposed control diagrams for individual DIC- i .

using numerical tests.

6.4 Numerical Tests

The control block diagram of individual DIC- i is depicted in Fig. 6.4. The control architecture consists of the primary droop control, distributed secondary frequency control levels and the attack detection/localization mechanisms. Specifically, the voltage-droop controller and the fast inverter inner voltage control loop have been included for tracking the voltage reference, corresponding to

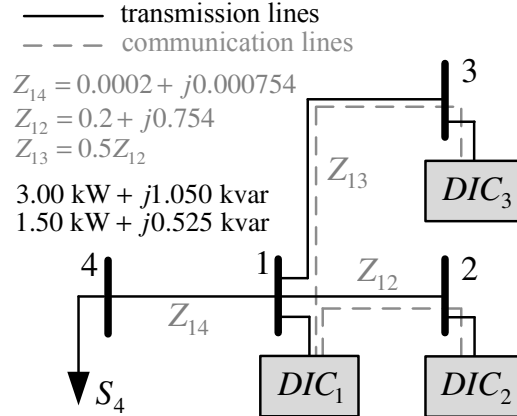


Figure 6.5: One-line diagram of the 4-bus/3-DIC microgrid.

Assumption 5. The local control in the primary level works with a sampling rate of 20kHz. This fast rate is necessitated to maintain a good output power quality, namely, the minimal frequency harmonics. Meanwhile, the distributed control updates at a much slower rate of a 10Hz because of a limited communication infrastructure in practical implementations. Fig. 6.5 depicts the system configuration of the underlying microgrid, where DIC-1 is connected to both DIC-2 and DIC-3, and thus receives information from both neighbors. Physical details of the MG including pulse width modulation emulation are included in the tests and implemented in the real-time simulation environment Opal-RT [117]. For ease of observation, we fix the ratings of these DICs to be the same as 2kW while the droop gain is set uniformly as $5 \times 10^4 \text{W} \cdot \text{s} \cdot \text{rad}^{-1}$. Typical load variations are implemented to show the effectiveness of the proposed control under nominal disturbances. In addition, varying virtual inertia of DICs is further considered. Last, the link and node attacks are further considered to certify the effectiveness of the isolation strategies of EMS.

6.4.1 Case I: Convergence Analysis

Case I-A: Load variations

For the given inertialess MG (i.e. $\mathbf{J} = \mathbf{0}$), we set $\epsilon = 2.1 \times 10^4$ and increase the system loading from half to full at $t = 10$. The resultant DIC active power injections and bus frequencies are shown in Fig. 6.6. Under this severe disturbance, the proposed distributed secondary control is capable of effectively achieving a zero system frequency deviation while maintaining accurate power sharing within a few seconds. This corroborates our earlier claims in Lemma 2 that the equilibrium point $(\omega^*, \mathbf{f}^*, \mathbf{p}^*)$ is attained with $\omega^* = 0$. Based on the amount of time to attain convergence which depends on system configuration, communication rate, and step-size choice, we design the detection time window to be slightly longer than this period. Additionally, we can vary the rate of control update according to load dynamics to achieve a satisfactory frequency performance, i.e., Assumption 7 holds. Therefore, the attack detection and localization mechanisms would not be falsely triggered by load variations.

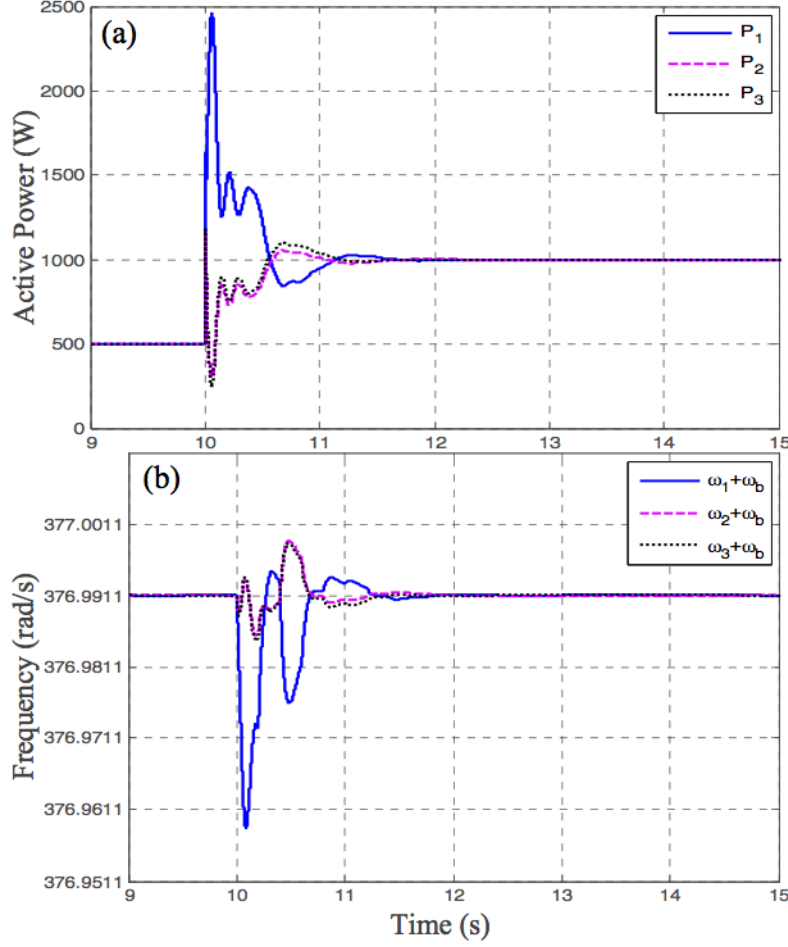


Figure 6.6: DICs' (a) active power outputs; (b) droop frequencies under a severe load change from half to full load at $t = 10$.

Case I-B: Virtual Inertia

With the same MG network under similar settings, the effects of virtual inertia of DICs are also investigated. Comparisons among various uniformly set J_i 's are illustrated in Fig. 6.7. It is clear that an increasing J significantly reduces the initial frequency dip when the load change occurs. In addition, the level of frequency oscillations is slightly improved as well. Since the DICs are modeled as second order systems, the droop gain would also contribute to the damping of the frequency. A larger D_i in terms of smaller droop ratio would slow down the convergence of the power sharing operation and thus degrade the control performance. Meanwhile, varying J does not affect the speed of convergence as corroborated in Fig. 6.7. Therefore, it is preferable in supplying additional damping in the MG supported by droop controlled DICs to enhance frequency

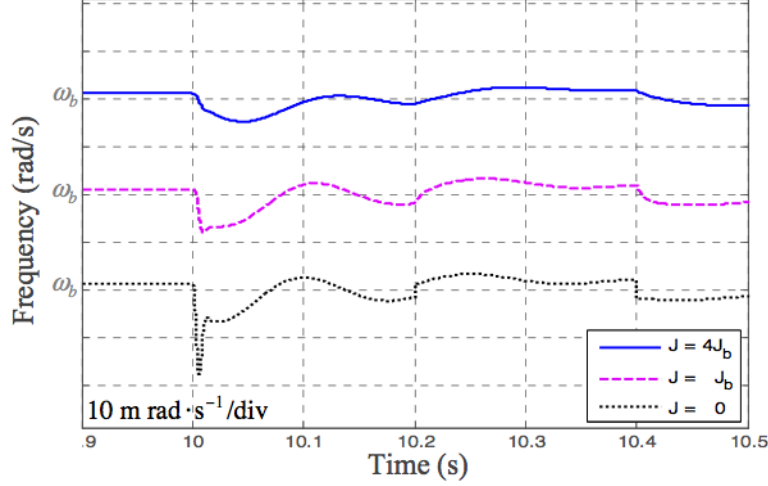


Figure 6.7: DFIGs' droop frequency under varying virtual inertia \mathbf{J} with $\mathbf{J}_b = 1 \times 1.061 \text{ kg} \cdot \text{m}^2$.

responses. To sum up, the proposed virtual inertia-based DFIGs governed by our PPD-based control algorithm (6.11) would effectively achieve a zero steady-state frequency deviation while enhancing the frequency transient stability.

6.4.2 Case II: Malicious Attacks on Communication Networks

To validate the effectiveness of the proposed attack countermeasures in Sec. 6.3, we adopt a more complex MG consisting of 14 buses and 6 DFIGs as shown in Fig. 6.8, which is same as that deployed in [87]. Under the same setting for the previous test, we carry out the following attacks on communication networks.

Case II-A: Individual link attack

An attack signal, 20% of the steady-state x_1 , is introduced at $t = 35$ to link-(1,4) and received by DFIG-1. The resultant plot of all DFIG responses is shown in Fig. 6.9. Clearly, due to this individual link attack, the center-of-mass frequency ω_c^k diverges as derived in (6.26). In addition, the power sharing is deteriorated since the consensus can no longer be attained. To validate our proposed detection indices, we plot $\{F_{1j}^k\}_{\forall j \in \mathcal{N}_1}$ in Fig. 6.10, which attests our claim in Sec. 6.3.1 that a dual variable associated with the malicious link would always be greater than the neighboring ones

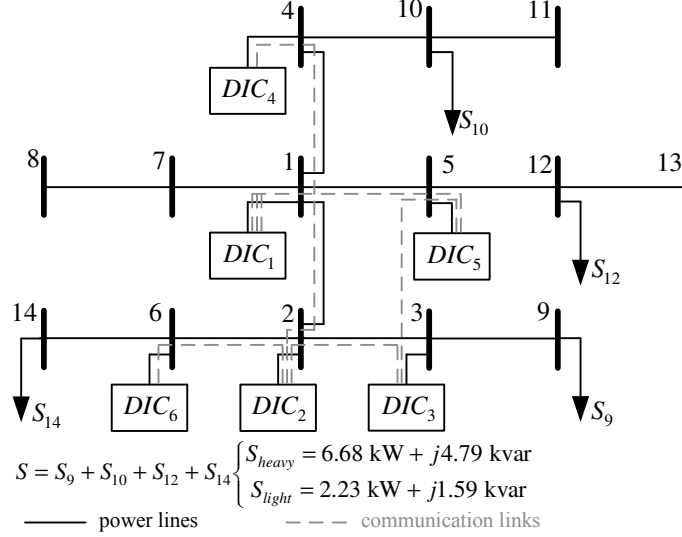


Figure 6.8: One-line diagram of the 14-bus/6-DIC microgrid.

without the malicious attacks; i.e., F_{14}^k exhibits the largest detection index during the detection time window. According to Algorithm 3, one can then localize the malicious link-(1,4) based on the fact that F_{14}^k is the largest. For a 20-second detection time window, the malicious link-(1,4) is flagged and reported to the EMS. Since DIC-4 only has one link, the control center corresponding to Algorithm 4 disables DIC-4's communication and commands it to operate in primary droop mode at $t = 55$, which leads to no participation of power sharing operation from DIC-4 as depicted in Fig. 6.9(a). After eliminating malicious link-(1,4), the active power injections of all DICs except for DIC-4 reach a new consensus, and the steady-state system frequency is zero after $t = 55$.

Case II-B: Coordinated link attack

Under a more complex attack scenario, two malicious inputs received by DIC-1 and DIC-2 are simultaneously appended at $t = 35$ to link-(1,4) and link-(2,6), respectively. These two inputs which have opposite signs are both 20% of the steady-state x_1 in their magnitudes. Fig. 6.11 depicts the responses of all DICs. Under this attack, the center-of-mass frequency ω_c^k in (6.17) would remain at zero regardless of the coordinated attack inputs while the proportional power sharing no longer holds. Conventionally, such attack would be challenging to identify since ω_c^k stays at its optimal value throughout the process as shown in Fig. 6.11(b). Thanks to the information obtained from

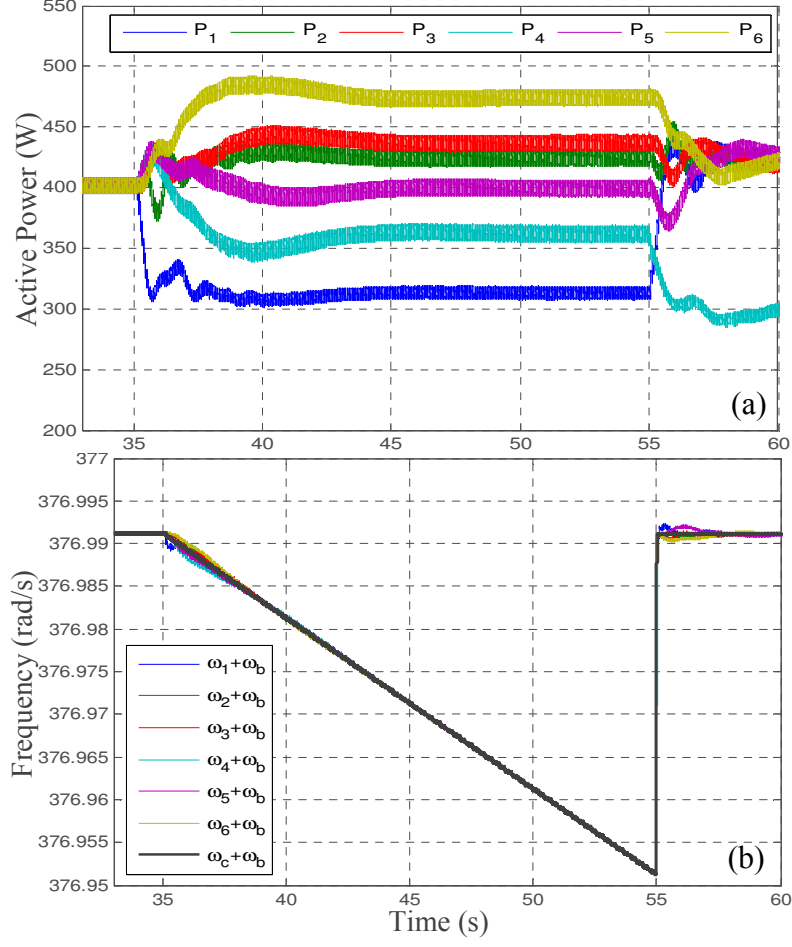


Figure 6.9: Case II-A: DICS' (a) active power outputs; (b) droop frequencies.

the dual variables, the detection indices of DIC-1 and DIC-2 are computed and plotted in Fig. 6.12. Similarly to the aforementioned individual link attack scenario, the malicious links are localized by the largest F_{14}^k and F_{26}^k of DIC-1 and DIC-2, respectively. Under a 20-second detection time window, malicious link-(1,4) and link-(2,6) are reported to the EMS. Since both DIC-4 and DIC-6 have only one neighbor, they are therefore removed from the consensus network after $t = 55$ and thus operate in primary droop mode. Fig. 6.11(a) depicts a recovery of a proportional power sharing operation among all DICS except for the DIC-4 and DIC-6.

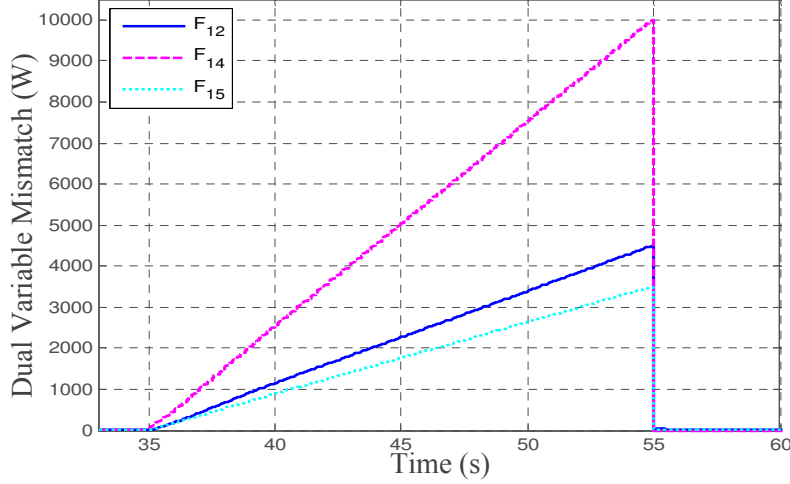


Figure 6.10: Case II-A: Detection indices F_{ij}^k of DIC-1.

Case II-C: Individual node attack

With a similar setting to case II-A, a node attack signal with the same level is inserted to DIC-4 at $t = 35$, resulting a false received information at DIC-1. The consequent plots of DIC output responses and detection indices F_{ij}^k are illustrated in Fig. 6.13 and Fig. 6.14, respectively. Under this attack scenario, as $(\mathbf{x}^k + \mathbf{u})$ lies in the null space of \mathbf{L} , the center-of-mass frequency ω_c^k in (6.30) would depend on \mathbf{u} and become non-zero while maintaining a proportional power sharing operation as shown in Fig. 6.13. This corroborates our attack analysis in Sec. 6.3.2. To isolate the attacked DIC-4, Algorithm 3 and 4 are executed. Accordingly, the EMS receives the flag reported by DIC-1 regarding the largest detection index of F_{14}^k , and therefore commands DIC-4 to switch to primary droop control mode at $t = 55$. After isolating malicious signals, similar results to earlier link attack scenarios can be observed in Fig. 6.13(a) where the nominal frequency is restored with a proper power sharing operation among all DICs except for DIC-4.

Case II-D: Coordinated node attack

To validate our proposed detection and isolation strategies under a coordinated node attack scenarios, we introduce malicious inputs which are simultaneously appended to DIC-4 and DIC-6 at $t = 35$. These inputs are with the same magnitude as 20% of the steady-state value of x_1 but opposite in their signs. Fig. 6.15 plots the resultant output response of DICs. As mentioned in Sec.

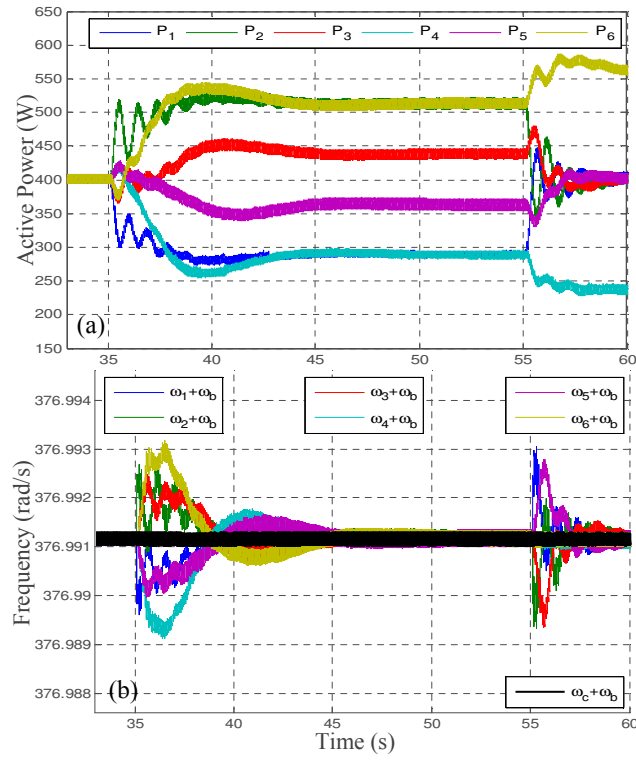


Figure 6.11: Case II-B: DICS' (a) active power outputs; (b) droop frequencies.

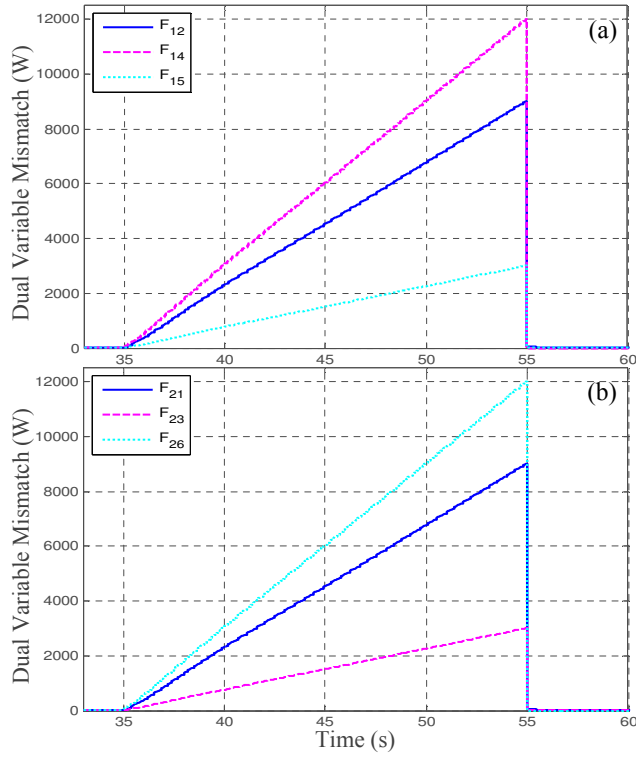


Figure 6.12: Case II-B: Detection indices: (a) F_{ij}^k of DIC-1; (b) F_{ij}^k of DIC-2.

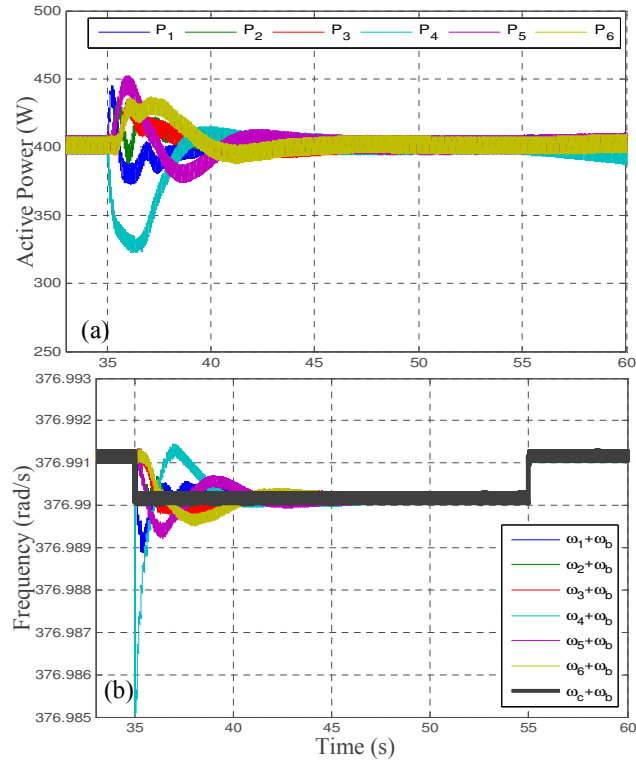


Figure 6.13: Case II-C: DICS' (a) active power outputs; (b) droop frequencies.

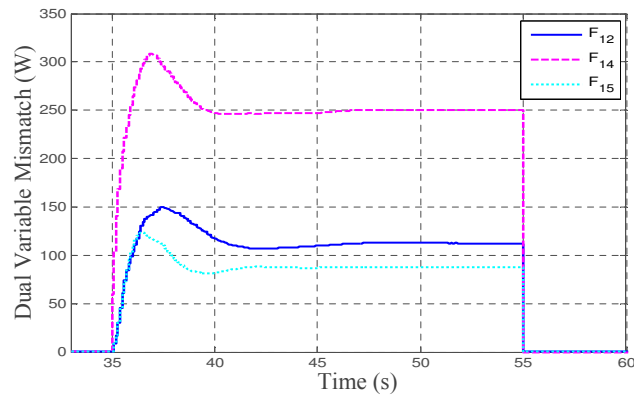


Figure 6.14: Case II-C: Detection indices F_{ij}^k of DIC-1.

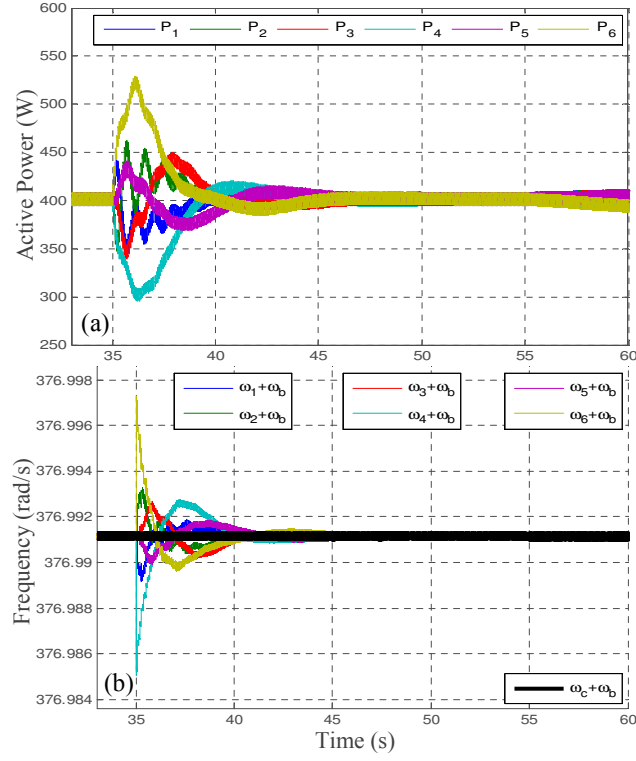


Figure 6.15: Case II-D: DICS' (a) active power outputs; (b) droop frequencies.

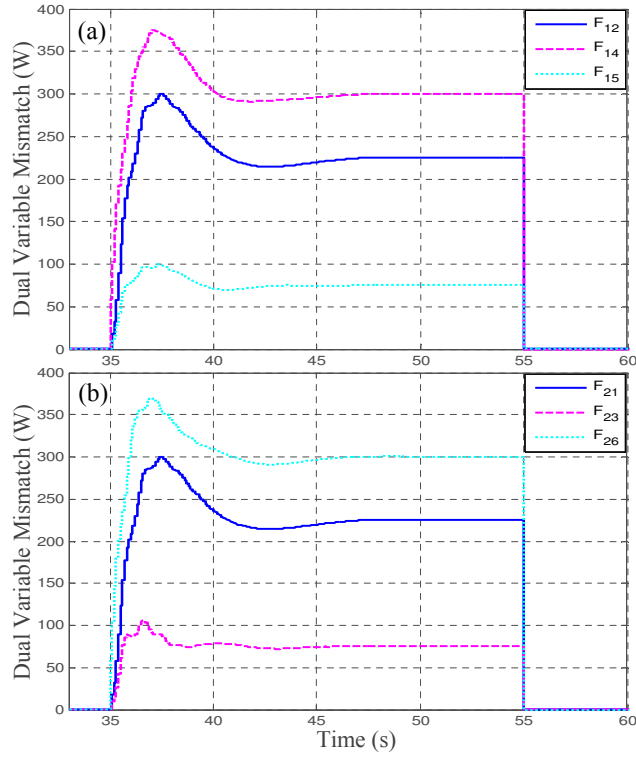


Figure 6.16: Case II-D: Detection indices: (a) F_{ij}^k of DIC-1; (b) F_{ij}^k of DIC-2.

6.3.2, such attack scenario would not deviate the center-of-mass frequency ω_c^k away from its optimal value and deteriorate the power sharing scheme. As shown in Fig. 6.15, albeit the transient of $\omega_i^k, \forall i \in \mathcal{N}$ is disturbed by the coordinated node attack, the overall system would eventually settle back to the pre-attack conditions. Thanks to the information from the dual variables, the detection indices of DIC-1 and DIC-2 shown in Fig. 6.16 are used to localize these two attacks through the largest F_{14}^k and F_{26}^k . For a 20-second detection time window, both DIC-4 and DIC-6 are flagged and reported to EMS which commands the removal of DIC-4 and DIC-6 from the consensus network at $t = 55$. After the isolation of malicious inputs, a slight change of power sharing operation is illustrated in Fig. 6.15(a) due to switching of local droop control mode in DIC-4 and DIC-6. To sum up, the aforementioned test cases II and III manifest the effectiveness of our proposed strategies in terms of detection and isolation malicious link and node attacks.

6.5 Summary

This chapter showcases our proposed control architecture, which consists of a local VSG-based droop control in the primary level and a distributed PPD-based algorithm in the secondary level. Interestingly, our specially designed control algorithm would seamlessly incorporate the dynamics of both power flow and VSG-based DICs which is implemented in the physical system by itself. The convergence is guaranteed by properly selecting stable parameter values based on the analysis of the PPD algorithm. To account for the cyber-security aspects of control design, malicious attack models are investigated along with detection and localization strategies. Numerical tests implemented in the real-time simulator demonstrate the effectiveness of the proposed control design in terms of achieving the objectives. Additionally, the proposed dual variable-based detection indices provide sufficient information to locate and isolate the malicious link or node.

CHAPTER 7

TESTBED-BASED DISTRIBUTED CONTROL IMPLEMENTATIONS FOR INDUSTRIAL MICROGRIDS

The smart grid vision of incorporating “the benefits of distributed computing and communications to deliver real-time information and enable the near-instantaneous balance of supply and demand at the device level”, stated in EPRI’s report [118], implies grid interoperability among multilayer information and control systems. The complexity of this task involves changing to horizontal communication, negotiation and collaborative decision makings. As communication networks continue to advance in electric power systems, the industry standard IEC 61850 has emerged for metering, protection, and control functions. IEC 61850 provides a standard for configuring various vendors’ intelligent electronic devices (IEDs) for electrical substation automation systems to be able to communicate with each other [119]. It has since found applications in new domains, including MGs, see e.g., [120]. Hence, this protocol is a key element in building (de)centralized automation functions in modern MG networks. We adopt the IEC 61850 standard [119, 121] and its architecture for facilitating the ensuing decentralized control framework design in Section 7.2. With a similar objective to Ch. 6, this chapter will focus on frequency regulation for an islanded industrial IEC 61850-based MG; see, e.g., ABB IEC 61850 protocol [122]. Under this protocol, the resultant capability to customize control design and automation logic would greatly enhance both the flexibility and adaptability of a MG network, steering towards the realization of the smart grid vision.

We formulate the secondary frequency control under quasi-steady-state as a consensus optimization problem, as in [7]. To avoid a single point of failure and enhance DERs’ plug-and-play capability, we propose to solve this problem in a decentralized fashion by adopting the alternating direction method of multipliers (ADMM) algorithm [57]. Accordingly, a DER controller uses local sensor measurements of voltage and current to perform a simple algorithmic computation for generating a local estimate. This estimate is then communicated to a utility central supervi-

sor, e.g., ABB MicroSCADA [123], which computes the average consensus of all estimates and broadcasts this consensus variable back to each DER controller. Our implementation differs from that of most decentralized frequency control designs in that we advocate modifying the ADMM updates originally derived for the steady state objective to an online feedback-based scheme, incorporating the instantaneous power measurements. Interestingly, it turns out that we do not need to explicitly model the MG power flow as the instantaneous power feedback signal couples DERs with the dynamics of power system networks.

While the IEC 61850-based cyber infrastructure enables the proposed decentralized control design, there is growing concern that it also exposes an attack surface for cyber adversaries. This is not hypothetical, as evidenced by recent cyber-induced outages in the Ukraine power system [124]. Hence, our control framework requires cyber defenses for controls and DERs against potential malicious cyber attacks. We consider an adversary model whereby an adversary can inject syntactically correct but destabilizing spoofed measurements and control commands, causing the secondary frequency control to fail and possibly resulting in an outage. The utilization of syntactically correct control commands during an attack to cause power outages has recently been reported [125], which motivates this work as addressing realistic attack scenarios.

Our main contribution lies in the development of a *collaborative* defense strategy against these attacks by leveraging the communication capabilities under the IEC 61850 standard. To enhance the robustness to malicious control command attacks, we employ the round-robin (RR) technique at the central supervisor for generating the consensus variable based on a variant of the ADMM algorithm [126]. Interestingly, by tracking the evolution of this RR-related variable, we are able to effectively identify compromised DER controllers. As for the measurement attack, we adopt a complementary defense using an agreement algorithm (AA) to detect and locate false measurements on which the secondary control is based [127, 128]. It should be noted that these two approaches provide insight into where the attack is happening. Thus, this can not only enable appropriate response with the correct mitigation, but also can alert an operator to the *specific* root cause. Together with the RR and AA detection algorithms, the central supervisor would be able to either isolate the malicious communication links from control updates or trip the malicious DERs off-line. This provides a multi-pronged approach to resilient and efficient MG operation in the face of adversarial conditions. These algorithms are demonstrated through simulation analysis of

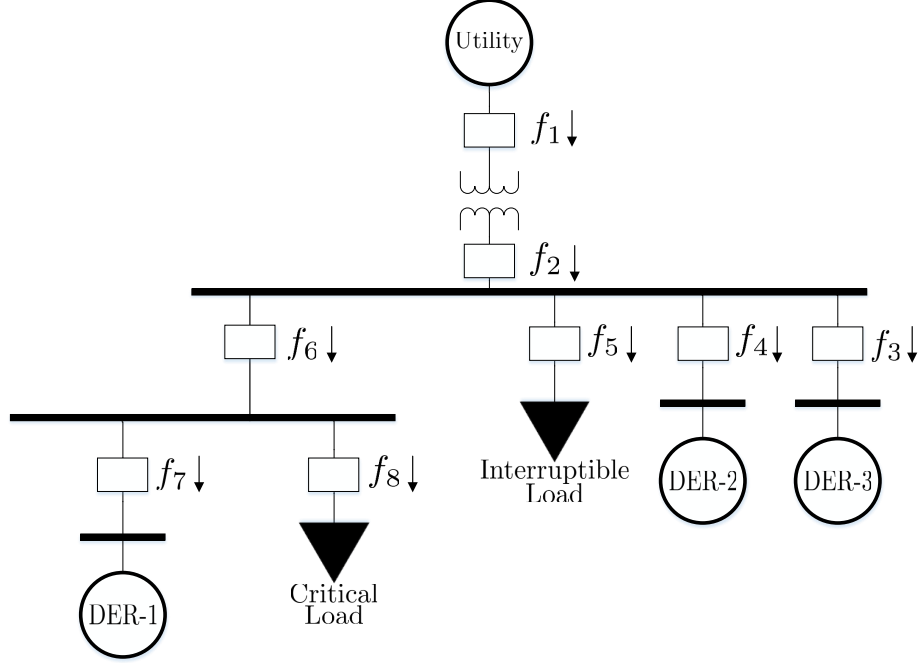


Figure 7.1: Reference industrial microgrid topology for this work.

several use cases of interest.

7.1 Industrial Microgrid

In this section, we define a reference industrial MG topology for this study and build a narrative around the attack scenarios. Additionally, we qualitatively detail the attack scenarios that the proposed mitigation strategies attempt to address.

7.1.1 Reference Industrial Microgrid Topology

Fig. 7.1 depicts the reference topology considered herein. The MG is connected to the area electric power system (AEPS) via a substation, with a corresponding point of interconnection (POI) where islanding decisions and requests can be executed. In normal operations, the MG will be connected to the AEPS, but as a strategy for resilience, the MG has the ability to island from the AEPS in the event of an outage or other degraded operation, including cyber or physical attacks as well as widespread outage due to a major storm.

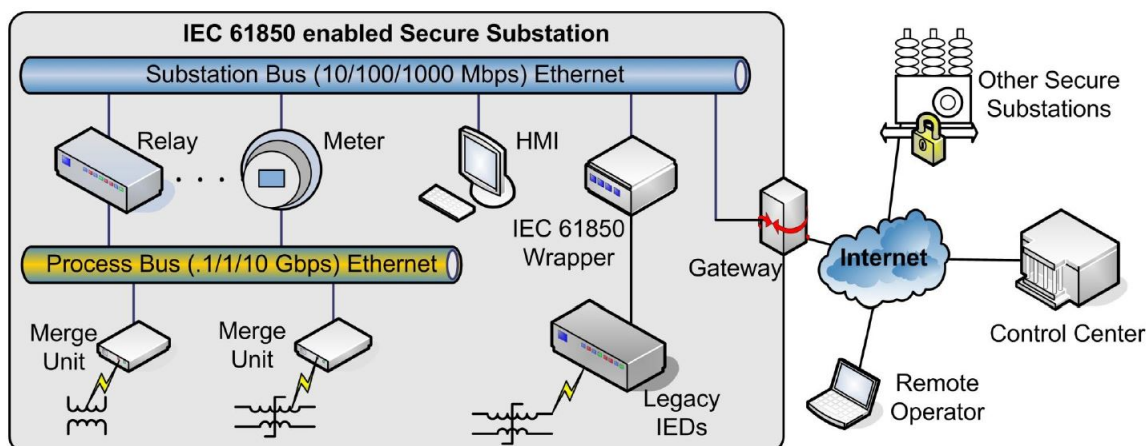


Figure 7.2: Notional representation of a standard IEC 61850 substation architecture.

There are a variety of DERs and loads within the MG. Two DERs and an interruptible load are connected directly to the MG feeder head. A load is considered interruptible if, in the event of a power system disturbance, the load is not the first priority to be served. This would be according to contractual agreements with the MG customer. The MG also contains a critical load and a DER that essentially serves as the backup source dedicated to the critical load. With both critical load and DER buses, they have the ability to island themselves from the rest of the MG as a resilient strategy for the critical load. In effect, DER-1 and the critical load, with the associated bus, would become a nested microgrid.

7.1.2 Communication Architecture: IEC 61850

IEC 61850 defines a number of protocols for various classes of substation messages. Among the protocols relevant for our proposed MG control system are Sampled Values (SV) and Generic Object Oriented Substation Events (GOOSE). Sampled Values transmit digitized measurements of voltage and current from a merging unit to an IED. A merging unit accepts inputs from current transformers (CTs) and potential transformers (PTs), and produces digital, time-synchronized outputs communicated to other nodes via an Ethernet bus, known as the Process Bus in IEC 61850; see Fig. 7.2. GOOSE messages containing status, data, and control commands can be sent from one IED to another. The reason for introducing this architecture is two-fold. First, this standard

is seeing increased applications in MGs, and as such, we find it relevant to design practical algorithms for field implementation. Second, using this standard provides a realistic attack surface that adversaries search out. We find it useful to provide specific solutions for a widely-used standard, especially considering a recent cyber attack impacting IEC 61850 [125]. Since we are dealing with MGs and not with bulk power systems, the number of nodes such as DER and other components is in the tens to at most low hundreds. Thus, modern substation communication architectures based on Ethernet can easily meet the transmission time and bandwidth requirements of the ensuing control architecture. As detailed soon, we consider malicious communication and control signal inputs which attempt to alter the MG operating points. Based on IEC 61850, such attacks can effectively drive the frequency away from the nominal, greatly threatening stability.

7.1.3 IEC 61850-Based Attacks and Countermeasures

Emerging MGs include a central MG controller, denoted in this thesis as Microgrid Controller (MGC), which communicates with individual DER controllers. Measurements and commands travel over communication networks, as given by IEC 61850. This communication structure potentially exposes the system to cyber attack, which can assume the form of invalid commands (which can cause a DER to perform potentially destabilizing power injections) as well as falsified measurements (which can lead even a correctly functioning MGC to issue erroneous control commands).

Attack 1: Communication Link Attack on Control Command

The scenario is a communication link attack on the control command (not measurements) from the MGC which is used to exchange ADMM-related variables via the Ethernet-based IEC 61850 station bus. The attack would result in the MGC calculating the wrong consensus variable, which would thus send the MG to a calculated off-nominal frequency setpoint. The attack detection mechanism examines the consensus variable and monitors for any rapid changes that exceed a threshold. If one is found, the mechanism looks for the errant local variable and sets the corresponding DER to local droop control. The remaining DERs participate in the secondary frequency

control while the spoofed DER operates in local droop mode only.

Attack 2: Local Attack on DER Control Command

The second scenario considers a local control command attack. An attacker compromises the DER controller by some mechanism. The attacker can then cause the system frequency and consensus variable to deviate from the appropriate references. This attack detection again relies on monitoring the consensus variable. Therefore, when it detects which DER is malicious, it again sets the malicious DER to local control mode since it is not yet known if it is a communication link or local controller attack at this point. By setting the malicious DER to local droop mode, if the system frequency and consensus variable are not converging to reference setpoints after a short time period, the MGC then determines such attack must be a local DER controller attack and issues a trip signal to the relay connecting the DER unit to the MG.

Attack 3: Local Measurement Attack

For the last attack scenario, we assume that the attacker either has access locally to the merging unit (the sensor), or can have access to the Ethernet-based process bus and thus is able to inject false measurement data. The DER controller that subscribes to the measurement messages would thus calculate incorrect power injection commands due to the faulty measurements. This could drive the MG to an unstable state. We assume that the attack is large enough to cause a reasonable frequency disturbance within the MG. Accordingly, a local merging unit attack may result in the tripping of the DER while a communication link attack on data measurements leads to reconfiguring the control algorithm and excluding the malicious DERs by setting them to the local droop mode. We next present the secondary frequency control problem statement with the proposed ADMM-based solver and round-robin-based detection mechanism.

7.2 Problem Statement

According to the microgrid model mentioned in Sec. 6.1, the goal of a secondary frequency control as mentioned is to 1) ensure a steady state zero frequency deviation (i.e., $\omega_i = 0, \forall i \in \mathcal{N}_M$) and 2)

guarantee autonomous active power sharing in proportion to active power ratings among all DERs such that

$$\frac{P_1}{P_1^M} = \frac{P_2}{P_2^M} = \cdots = \frac{P_n}{P_n^M}.$$

To this end, the active power-frequency (P - ω) droop control is adopted to achieve these objectives [85]. Upon concatenating all scalar variables into vector form, we formulate the secondary control problem as a consensus optimization problem under quasi-steady-state, as given by

$$\begin{aligned} & \min_{\mathbf{P}} \frac{1}{2} \|\mathbf{P}^M - \mathbf{P} - \mathbf{p}\|_{\mathbf{D}^{-1}}^2 \\ & \text{subject to } \frac{p_i}{D_i} = \frac{p_j}{D_j}, \forall (i, j) \in \mathcal{E}_D, \end{aligned} \tag{7.1}$$

where $\mathbf{D} := \text{diag}(D_1, \dots, D_n)$ is an $n \times n$ diagonal matrix and the weighted norm $\|\mathbf{v}\|_{\mathbf{D}}^2 := \mathbf{v}^T \mathbf{D} \mathbf{v}$ for any vector \mathbf{v} . Based on (6.3), we formulate the objective of (7.1), aiming to achieving a zero system frequency deviation. Similarly to Ch. 6, due to a default uniform D_i/P_i^M setting, the equality constraints in (7.1) equivalently enforce a proportional active power sharing. Note that the quadratic program (7.1) could be solved using off-the-shelf convex solvers. Nonetheless, the challenge lies in that the active power injection \mathbf{P} is dynamical and coupled to the power system network. To tackle this problem, as detailed soon, we adopt the feedback approach from [7] to account for system dynamics.

7.3 ADMM-based Decentralized Solver

This section introduces our proposed ADMM-based decentralized secondary control design. The dynamics coupling \mathbf{P} and \mathbf{p} are neglected initially. As detailed below, the feedback approach will be introduced to account for such interactions. Hence, the objective in (7.1) is fully separable. Using the IEC 61850 communication protocol for measurement and control messages, we can solve the consensus optimization problem (7.1) in a fully decentralized fashion. For notational convenience, we let the optimization variable $x_i := p_i/D_i$ and the input variable $c_i := (P_i^M - P_i)/D_i$, where P_i is the active power injection from DER- i and locally measurable. Accordingly, (7.1) can

be reformulated as

$$\min_{\mathbf{x}, z} \frac{1}{2} \|\mathbf{c} - \mathbf{x}\|_{\mathbf{D}}^2 \quad (7.2a)$$

$$\text{subject to } \mathbf{x} = z\mathbf{1}, \quad (7.2b)$$

where z is a consensus variable among the DERs. Note that the equality constraints in (7.1) are equivalent to (7.2b). Defining the multipliers $\boldsymbol{\lambda}$ and a constant $\rho > 0$, we introduce the augmented Lagrangian function as $\mathcal{L} = \sum_{i \in \mathcal{N}_D} \mathcal{L}_i(x_i, z, \lambda_i)$ where

$$\mathcal{L}_i(x_i, z, \lambda_i) = \frac{D_i}{2}(c_i - x_i)^2 + \lambda_i(x_i - z) + \frac{\rho}{2}(x_i - z)^2. \quad (7.3)$$

Based on the (7.3), the ADMM algorithm is invoked and its $(k + 1)$ -st iteration for DER- i has the following three steps [57]:

(S1) Update \mathbf{x} : As \mathcal{L} totally decouples into \mathcal{L}_i for each DER- i , minimizing x_i involves only the variables z^k and λ_i^k . Thus, upon receiving z^k from the MGC, the update is

$$x_i^{k+1} := \arg \min_{x_i} \mathcal{L}_i(x_i, z^k, \lambda_i^k). \quad (7.4)$$

Taking the gradient of \mathcal{L}_i with respect to x_i and setting it to zero, we have

$$x_i^{k+1} = \frac{\rho z^k + D_i c_i^k - \lambda_i^k}{D_i + \rho}, \quad (7.5)$$

where c_i^k is the feedback measurement signal, corresponding to the active power injection of DER- i .

(S2) Update \mathbf{z} : Likewise, the consensus variable is updated as

$$z^{k+1} := \arg \min_z \mathcal{L}(\mathbf{x}^{k+1}, z, \boldsymbol{\lambda}^k).$$

By initializing $\boldsymbol{\lambda}^0 = \mathbf{0}$, the summation $\sum_{i \in \mathcal{N}_D} \lambda_i^{k+1}$ is guaranteed to stay zero. Thus, we

have

$$z^{k+1} = \frac{\sum_{i=1}^n x_i^{k+1}}{|\mathcal{N}_D|}. \quad (7.6)$$

(S3) Update λ : Each multiplier is linearly updated by the iterative mismatch of the constraint (7.2b), as given by

$$\lambda_i^{k+1} = \lambda_i^k + \rho(x_i^{k+1} - z^{k+1}). \quad (7.7)$$

Because $\lambda^0 = 0$, we have

$$\sum_{i \in \mathcal{N}} \lambda_i^{k+1} = \rho \sum_{i \in \mathcal{N}_D} \sum_{t=1}^{k+1} (x_i^t - z^t) = 0.$$

This fact corroborates the derivation in (7.6).

7.4 Detection and Localization Strategies

Under IEC 61850 communication network, we assume that attackers have compromised the local DER controllers such that the local variable \mathbf{x} is altered, e.g., $\bar{x}_i^{k+1} = x_i^{k+1} + \delta_i^{k+1}$ where δ_i^{k+1} is the bias appended to x_i^{k+1} at the DER- i . Therefore, z^{k+1} in (7.6) at the MGC becomes

$$z^{k+1} = \frac{\sum_{i=1}^n (x_i^{k+1} + \delta_i^{k+1})}{n} = \Delta^{k+1} + \frac{\sum_{i=1}^n x_i^{k+1}}{n} \quad (7.8)$$

with $\Delta^{k+1} := \frac{\sum_{i=1}^n \delta_i^{k+1}}{n}$ being the average attack bias signal with time-varying and arbitrary magnitude. Under the presence of this attack, the consensus variable z^{k+1} would diverge unless Δ^{k+1} is designed specifically so the effect on the consensus variable is trivial. This is, however, unlikely to happen as the attacker does not have the full system information. In any case, such an attack bias signal may drive the MG to unstable conditions and/or damage system equipment, e.g., causing divergence of z^{k+1} .

It is imperative to detect and localize the malicious attack signals promptly since the control

design is based on z^{k+1} . To this end, we monitor the evolution of z^{k+1} and design a flag to trigger the ensuing detection algorithm. Assuming the convergence of z^{k+1} after k^* iterations, we would trigger the detection algorithm once the following condition has been satisfied:

$$|z^{k+1} - z^k| > \epsilon,$$

where $\epsilon > 0$ is a pre-defined threshold.

7.4.1 Round-Robin-Based ADMM Detection Algorithm

The RR-ADMM detection algorithm to discover the malicious DERs is adapted from [126]. The RR is an arrangement of selecting the DER in a fixed rational order, i.e., DER-1, DER-2, \dots , DER- n . For notational convenience, we denote the consensus variable for the RR-ADMM at iteration k as \tilde{z}^k . Given $\alpha > 0$, the steps (S1)-(S3) become

$$\mathbf{x}^{k+1} = (\mathbf{D} + \rho \mathbf{I})^{-1}(\rho \tilde{z}^k \mathbf{1} + \mathbf{D} \mathbf{c}^k - \boldsymbol{\lambda}^k), \quad (7.9a)$$

$$\tilde{z}^{k+1} = \alpha(x_i^{k+1} + \delta_i^{k+1}), \quad (7.9b)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho(\mathbf{x}^{k+1} - \tilde{z}^{k+1} \mathbf{1}), \quad (7.9c)$$

where \mathbf{I} is the identity matrix with $\tilde{i} = 1, \dots, n$ representing the fixed round-robin iteration index. For a non-malicious DER, we set $\delta_i^{k+1} = 0$. Hence, we have the consensus variable \tilde{z}^{k+1} as

$$\tilde{z}^{k+1} = \alpha \delta_i^{k+1} + \alpha \frac{\rho \tilde{z}^k + D_{\tilde{i}} c_{\tilde{i}}^k - \lambda_{\tilde{i}}^k}{D_{\tilde{i}} + \rho}. \quad (7.10)$$

For $k \geq 1$, (7.10) can be expressed as

$$\tilde{z}^{k+1} = \alpha \delta_i^{k+1} + \alpha \frac{\rho \tilde{z}^k + D_{\tilde{i}} c_{\tilde{i}}^k - \rho \sum_{t=1}^k (x_{\tilde{i}}^t - \tilde{z}^t)}{D_{\tilde{i}} + \rho}. \quad (7.11)$$

Let $\tilde{\mathbf{z}}_r := \{\tilde{z}_{r,1}, \dots, \tilde{z}_{r,n}\} \in \mathbb{R}^n$ gather the all the values of the consensus variable at the r -th round of the RR-ADMM algorithm. To determine a threshold to separate malicious DER controllers from the rest of the system, we assume that the bias δ_i^{k+1} is sufficiently large. Based on (7.11),

Algorithm 5 Detection and Localization Strategies

```
1: for every iteration  $k = 0, 1, 2, \dots$  do
2:   for  $i \in \mathcal{N}_D$  do
3:     Compute  $x_i^{k+1}$  as in (7.5) and send it to MGC
4:   end for
5:   MGC computes  $z^{k+1}$  as in (7.6)
6:   if  $|z^{k+1} - z^k| > \epsilon \wedge (k > k^*)$  then
7:     if  $r = 1$  then
8:       MGC computes  $\tilde{z}^{k+1}$  as in (7.9b)
9:       Broadcast the value of  $\tilde{z}^{k+1}$  to all DERs
10:      Determine the index  $\tilde{n}$  for the minimum entry of  $\tilde{z}_1$ 
11:    end if
12:    if  $(r = 2) \wedge (k \leq k^* + n + \tilde{n})$  then
13:      MGC computes  $\tilde{z}^{k+1}$  as in (7.9b)
14:      Broadcast the value of  $\tilde{z}^{k+1}$  to all DERs
15:      Identify malicious DER- $\tilde{i}$  where  $\{\tilde{i} \mid \tilde{z}_{1,\tilde{i}} > \tilde{z}_{2,\tilde{n}}, \forall \tilde{i} \in \mathcal{N}_D \setminus \tilde{n}\}$ 
16:      MGC reconfigures the communication network, resets  $\lambda^k = 0$ , and/or trip malicious DERs off-line
17:    end if
18:  else
19:    Broadcast the value of  $z^{k+1}$  to all DERs
20:  end if
21:  for  $i \in \mathcal{N}_D$  do
22:    Compute  $\lambda_i^{k+1}$  as in (7.7)
23:  end for
24: end for
```

one of the values from the non-malicious DERs during the r -th round must be $\tilde{z}_{r,\tilde{n}}$ with the index \tilde{n} corresponding to the smallest element of $\tilde{\mathbf{z}}_r$. Given this index, the $(r + 1)$ -st round is carried out for obtaining the value of $\tilde{z}_{r+1,\tilde{n}}$ where \tilde{n} is the same index as round 1, and this serves as the detection threshold. Hence, any $\tilde{z}_{r,\tilde{i}} > \tilde{z}_{r+1,\tilde{n}}, \forall \tilde{i} \in \mathcal{N}_D \setminus \tilde{n}$ is identified as the malicious DER in the MG. For a given initialization time index k^* , Algorithm 5 tabulates the detection strategy. As for the localization strategy to isolate the aforementioned malicious attack signals, the MGC first reconfigures the communication network so the malicious DERs no longer participate the ADMM updates in (S1)-(S3) and thus switch to only local droop (primary) control mode. Meanwhile, if zero frequency deviation is achieved, we conclude the isolation process. Otherwise, the malicious DERs are tripped off-line because of either measurement or control signal attack. Last, note that there must be at least one non-malicious DER in the system for the RR-ADMM detection scheme

to work. Such a scheme is only for detection purposes. Thus, once the malicious attacks are localized, the control design is reverted back to follow the ADMM algorithm in (S1)-(S3).

7.4.2 Measurement Attack Detection

We now describe a defense against false measurement injection to complement defenses against control attacks given above. We adopt the agreement algorithm (AA), developed in [127], to determine and locate malicious measurement attacks on substation IEDs and controllers. Accordingly, assuming the loads as constant impedances, the Kirchhoff's voltage and current laws along with Ohm's law are used to facilitate the development of agreement matrix \mathbf{A} for a particular topology. Albeit we assume loads as constant impedance, a general assumption in power flow studies, the method for developing the AA presented herein remains valid for other load models. Elements of \mathbf{A} corresponding to the currents reflect the signed topology of the corresponding merging unit while others corresponding to voltages are reciprocal complex impedances on the corresponding lines. Fig. 7.1 showcases the reference MG topology with corresponding measurement locations. The polarity of the complex current f_i measured at the i -th merging unit is positive when current flows *into* the loads and DERs. By concatenating as $\mathbf{x} = (\mathbf{f}, \mathbf{v})$, the physical equation can be rewritten as

$$\mathbf{Ax} = \mathbf{0}. \quad (7.12)$$

Considering that (7.12) is similar to the error correcting code formulation from [127], if an attacker falsifies one of the measurements, we would have a non-zero corresponding element of the resultant vector, known as the Syndrome vector. By injecting the malicious vectors $\Delta\mathbf{f}$ and $\Delta\mathbf{v}$ to the measurements, we have $\bar{\mathbf{x}} = (\mathbf{f} + \Delta\mathbf{f}, \mathbf{v} + \Delta\mathbf{v})$. Thus, the Syndrome vector is

$$\mathbf{s} = \mathbf{A}\bar{\mathbf{x}}. \quad (7.13)$$

By observing the pattern of vector \mathbf{s} , we can classify multiple subsets of potential malicious merging units. Accordingly, the largest magnitude element of a subset corresponds to the malicious location. This detection mechanism is valid for a limited number of attacks. We refer the reader to [128] for a detailed discussion.

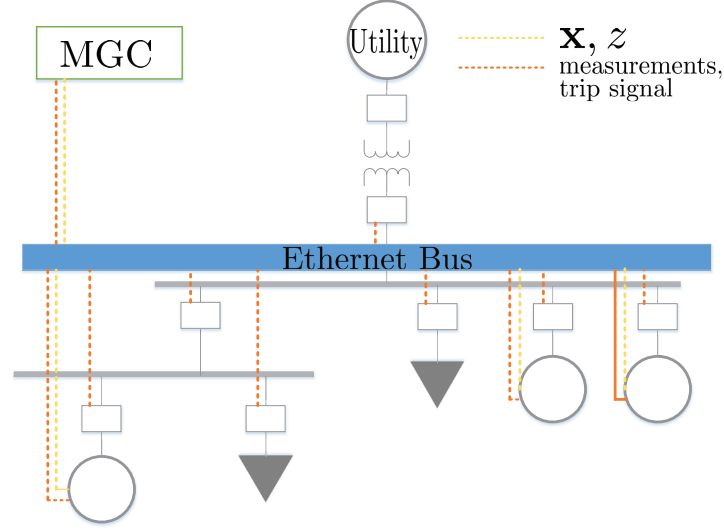


Figure 7.3: Reference microgrid communication architecture and data types.

7.5 Numerical Tests

In this section, we evaluate the proposed mitigation strategies and responses for the communication and measurement link attack scenarios. The three-phase MG topology and power system parameters are given in Fig. 7.1. The load is modeled as a constant impedance load, which is frequency independent. Fig. 7.3 depicts the MG control system communication network topology. To reiterate, the measurements are sent to the local DERs from a merging unit (which we omit from the figure), and the DERs and MGC communicate updates for the ADMM algorithm. This is done over switched Ethernet, denoted by the Ethernet bus in Fig. 7.3. All numerical tests are performed in Mathworks® MATLAB 2013a and Simulink software.

7.5.1 Load Perturbations

In this scenario, we increase the system load by 100% at $t = 4\text{s}$. Each DER is rated at $P_i^* = 1500\text{ W}$, $\forall i$, and we let $D_i = 5 \times 10^4\text{ W} \cdot \text{s} \cdot \text{rad}^{-1}$, $\forall i \in \mathcal{N}_D$ to satisfy the active power sharing. The ADMM algorithm is executed every 100 ms. The resulting bus frequencies and active power output are shown in Fig. 7.4. Within approximately 1.5 seconds, the secondary frequency control is able to obtain zero system frequency deviation from nominal, and the DERs have correctly achieved equal power sharing. Accordingly, each DER archives the steady state frequency of 60 Hz.

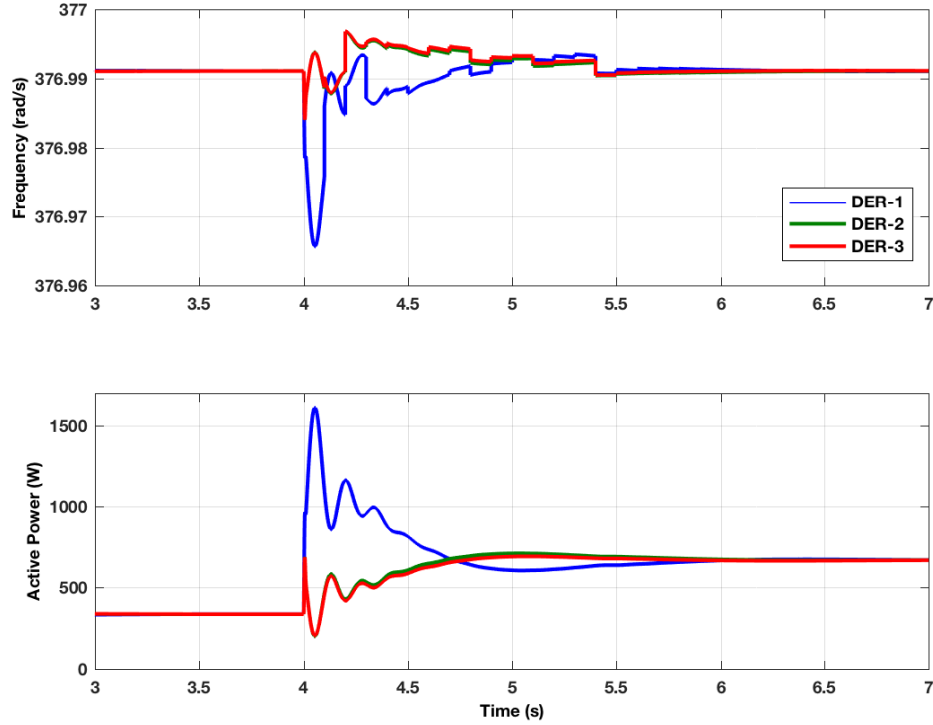


Figure 7.4: Frequency and active power output response to a load disturbance.

7.5.2 Local Attack on DER Controller

We generate an attack signal as a time-varying random number from a uniform (0,3) distribution and draw a new random value at a time step of 100 ms. We multiply this by the steady state x_i value at the attack location, so that the attack is effectively a random re-scaling of this value. Given the steady state conditions, the attack is introduced at $t = 4.1$ s on the local x_i issued to DER-3. The resulting system response and RR-ADMM attack detection and mitigation algorithm results are shown in Fig. 7.5. From the plot of the local x_i update, this particular attack introduces a signal that is approximately 275% of the steady state x_3 signal. Clearly, the system diverges away from its steady state while the attack is present. At $t = 5.4$ s with ϵ setting at 10% of the steady state x_3 signal, the RR-ADMM algorithm successfully detects DER-3 as malicious and trips it off-line, i.e., $P_3 = 0$. For $t > 5.5$ s, the ADMM algorithm changes to only include DER-1 and DER-2, achieving the nominal frequency of 60 Hz.

Next, we investigate the effectiveness of the RR-ADMM detection algorithm for an attack during a load disturbance. While a coincidental simultaneous occurrence of these two events may

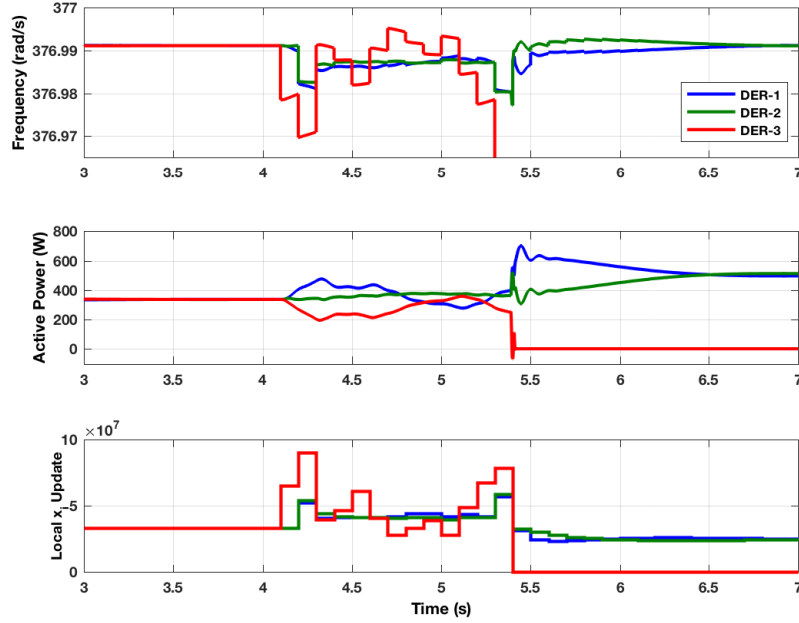


Figure 7.5: Frequency, active power output, and local x_i update responses to a local controller attack in steady state operation.

seem unlikely, we are motivated to seek solutions to coordinated attacks, i.e., the attacker causes a load disruption *and* alters the local controller updates, as depicted in Fig. 7.6. At $t = 4$ s, we introduce a load disturbance of 25% and then subsequently cause an attack at $t = 4.1$ s on the x_i update to DER-3, similar to the attack scenario in the steady state case. We see that the random attack signal is approximately 200% of the transient x_3 signal. The RR-ADMM algorithm is still able to identify the malicious DER even in the presence of a load disturbance. After reconfiguring the ADMM algorithm and tripping DER-3 off-line at $t = 5.5$ s, the system achieves the nominal frequency of 60 Hz.

7.5.3 Communication Link Attack on Control Command

We consider that an attacker has gained access to the station bus (Ethernet bus) that is exchanging control commands between the local DER controllers and the MGC. The attacker is able to spoof the MAC address of a DER controller and thus can alter control commands over the link. This is contrasted with the previous attack since it is not on the local DER controller, and thus the time-varying attack signal does not directly affect the power injection command to the DER. The

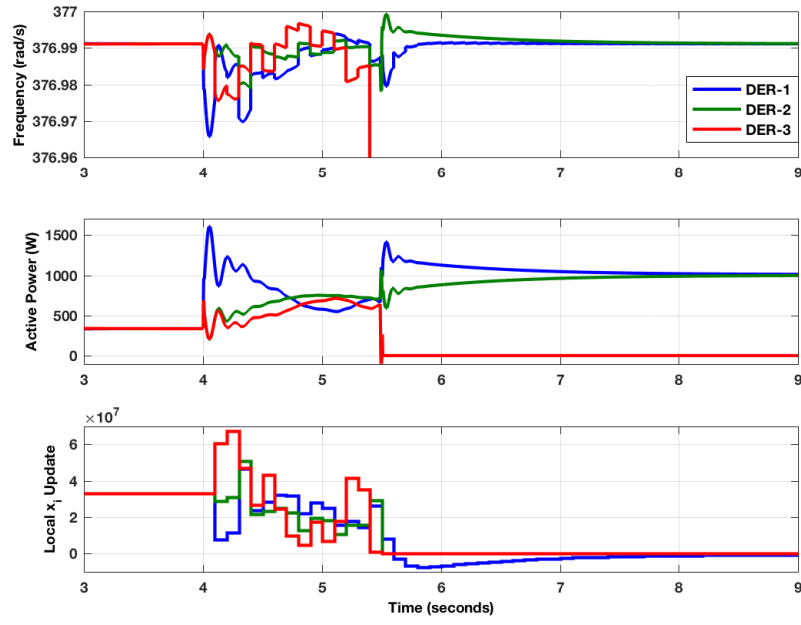


Figure 7.6: Frequency, active power output, and local x_i update responses to a local controller attack during a load disturbance.

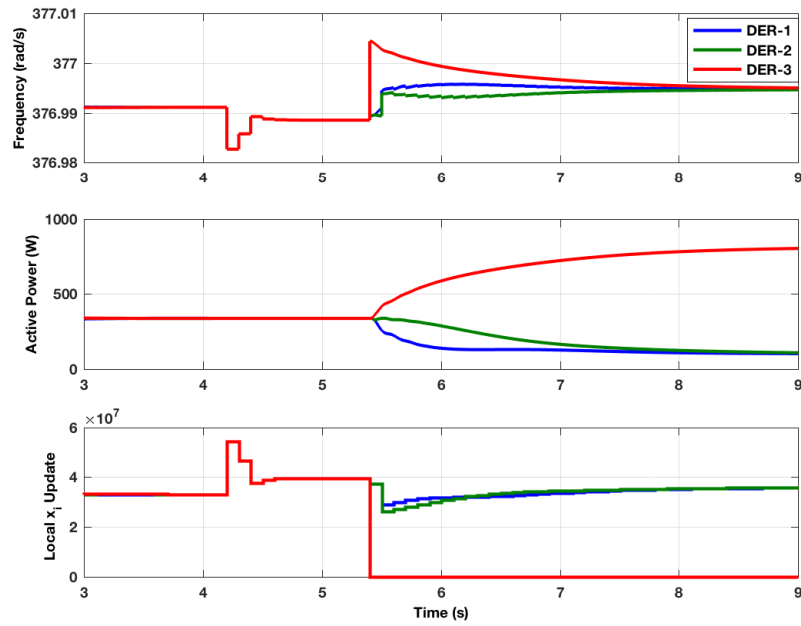


Figure 7.7: Frequency, active power output, and local x_i update responses to an communication link attack on a control command.

attack detection monitors the consensus variable and raises a flag when a deviation occurs that exceeds a threshold. In our simulation, we again use a 10% deviation as the threshold. As the consensus variable is the average across n DERs, an attack bias may not be large, so that is the motivation for setting a relatively sensitive ϵ . After the flag is raised, the RR-ADMM is executed to determine which DER is malicious. The MGC then reconfigures the ADMM update to only include the non-malicious DERs while issuing a configuration command to the spoofed DER to revert to local frequency droop control. The spoofed DER should then eventually return to its initial power setpoint while the non-malicious ones continue to regulate the system frequency for achieving 60 Hz. Fig. 7.7 depicts the results of this attack scenario. At $t = 4.1$ s, an attack signal is introduced on the x_3 update sent from DER-3 controller to the MGC. With the same random attack signal, the MGC then runs the detection mechanism from the RR-ADMM to find the malicious DER. At $t = 5.3$ s, the MGC identifies DER-3 as malicious and removes it from the ADMM update by setting it to local frequency control mode. Note that the x_3 update is a function of p_3 . By setting x_3 to zero, the corresponding DER controller equivalently becomes the local droop control. By reconfiguring the ADMM algorithm, the DER-1 and DER-2 continue to execute the secondary frequency control while maintaining power sharing and achieving the nominal frequency of 60 Hz.

7.6 Summary

In this chapter, we introduce a decentralized secondary frequency control that can successfully achieve frequency regulation in islanded ac microgrids. This approach is based on formulating the DER droop characteristic equations as a consensus optimization problem with a power injection offset command as the control variable. This quadratic program is solved with an ADMM-based decentralized algorithm. To this end, DER controllers locally compute their power injection offsets and communicate these values with the central controller, which then calculates the consensus of all DERs and broadcasts over the network. This decentralized approach allows for cyber attack detection mechanisms on local controllers and communication link attacks. The proposed detection algorithm is based on a round-robin ADMM algorithm which sequentially updates the consensus variable as a function of local controller updates to identify malicious DERs. We pair this

with a so-called agreement algorithm, a complementary false data injection detection mechanism. Mitigation strategies such as isolating attackers from the control algorithm or tripping a compromised DER off-line entirely are discussed. Together with these algorithms, we can implement a cybersecure resilient closed-loop control architecture. Finally, we demonstrate the effectiveness of our decentralized secondary frequency control design and detection algorithms using three case studies.

CHAPTER 8

CONCLUDING REMARKS

In this chapter, we highlight the main contributions of the thesis and conclude with some remarks.

8.1 Thesis Summary and Contribution

Chapter 3: In this chapter, we have developed a local dynamic optimization framework for analyzing the performance of a voltage control scheme based on gradient projection methods. Thanks to the physical network couplings, the local voltage measurement serves as the instantaneous system-wide gradient direction by forming a weighted voltage mismatch objective. Compared to earlier results for a static optimization scenario, we have significantly extended the analysis on convergence conditions and error performance to account for two dynamic scenarios: i) the nodes perform the local update in an asynchronous fashion; and ii) the network operating point is dynamically changing. Last, we provide the rules of selecting a proper step-size, which shows a trade-off between the convergence speed and the steady-state tracking error for the dynamic control design.

Chapter 4: In this chapter, we have developed a fully distributed voltage control (DVC) design to manage the VAR-capable DERs deployed to an increasingly dynamic and variable distribution system. Under the limited VAR scenario, the local control in Chap. 3 may attain a suboptimal voltage profile because of lack of information exchanges. To leverage the network-wide VAR support, we have cast the voltage control problem as one of quadratic programming that minimizes the unweighted voltage mismatch under VAR resource limits everywhere by formulating the (multi-phase) power flow model to linear equality constraints. The ADMM distributed optimization algorithm is evoked, which requires only information exchange among neighboring buses and totally local computations. To account for external disturbances and lack of reliable communication links, we have further extended the implementation of the proposed ADMM-based control

in an online fashion that is robust to random failure of communication links. We showcase the effectiveness of the DVC design regarding mitigation of the variability introduced by DERs and improvement of the voltage stability.

Chapter 5: We integrate the attractive features from both local and distributed voltage control designs and develop the advanced hybrid control framework while being cognizant of various communication scenarios. We have cast the specially-designed voltage control problem by combining both weighted and unweighted voltage mismatch objectives under limited VAR resources. The PPD-based algorithm is then evoked and only requires voltage measurement exchanges among neighboring buses with local computations. We have further provided the rigorous proof for stable step-size choices to guarantee control stability. To cope with cyber resource constraints and lack of reliable communication links, we have extended the hybrid voltage control (HVC) design to have robustness against random communication link failures and, in particular, communication-cognizant feature to account for the worst-case scenario of a total communication outage. Interestingly, under this worst case, the HVC boils down to a surrogate local control of Chap. 3. Henceforth, the adequate control performance can still be achieved as it responds to local voltage variations.

Chapter 6: We have thus far considered distribution networks which are fully connected to the transmission grids. Assuming DERs are capable of supplying the loads in feeders, it is a great interest to develop control designs when feeders are disconnected from the main grid and formed isolated microgrids (MGs). MGs are foreseen to play an important role in distribution networks to enhance the power system reliability. Under isolated MGs, these mostly power electronics-interfaced distributed energy resources (DERs) exhibiting low-inertia characteristic have raised major concerns over the frequency stability issues. To address these issues, we have adopted the concept of the virtual synchronous generator (VSG) and developed a distributed secondary frequency control for DICs in isolated MGs. Our proposed control architecture consists of a local VSG-based droop control in the primary level and a distributed PPD-based algorithm in the secondary level. Interestingly, our specially designed control algorithm would seamlessly incorporate the dynamics of both power flow and VSG-based DICs which is implemented in the physical system by itself. Our future work includes the analysis of the stable control parameter choices in addition to incorporating the economics into our optimization framework. As most hierarchical control frameworks assume

the time-scale separation among different control levels, our optimization-based control design would explicitly account for the original dynamical model of network-coupled DICs in establishing the stability conditions of the distributed control design. This would improve the conventional MG control framework by incorporating various time-scaled control objectives automatically and dynamically. Moreover, we have analyzed two types of malicious attacks on the communication network, namely, the link and node attacks, and proposed model-based anomaly detection and localization strategies based on the dual variable related metrics.

Chapter 7: Extending the work proposed in Ch. 6, we have focused on the development of a secondary frequency control design where the supporting communication network follows the IEC 61850 protocol. Formulating the DER droop characteristic under quasi-steady-state as a consensus optimization problem, we have solved this quadratic program by an ADMM-based decentralized algorithm. Our main contribution lies in the development of a *collaborative* defense strategy against these attacks by leveraging the communication capabilities under the IEC 61850 standard. Specifically, this decentralized approach allows for cyber attack detection mechanisms on local controllers and communication link attacks. Meanwhile, a so-termed agreement algorithm, a complementary false data injection detection mechanism, is also implemented for monitoring the measurement trustfulness. Mitigation strategies such as isolating attackers from the control algorithm or tripping a compromised DER off-line entirely are discussed. This work has facilitated the development of a multi-pronged approach to resilient and efficient MG operation in the face of adversarial conditions.

8.2 Conclusions

In this thesis, we proposed several decentralized control strategies in power distribution networks including microgrids with the objective of either regulating voltage or frequency. In Ch. 3, we formulated the weighted voltage objective under a single-phase network and solved the (local) control problem in a totally communication-free fashion. While this approach has a simple implementation and can be readily extended to multi-phase networks, the stability and optimality could become a significant concern. To address these problems, we developed the distributed control un-

der a globally voltage objective in Ch. 4. Thanks to the information exchanges among neighboring nodes, we can quickly and accurately track a desired voltage set-point while dynamically updating the system operation conditions and thus enabling online implementations. Regarding limited communication resources, we further improved the distributed control design to be robust to random communication failures. However, under the worst-case scenario of total communication outage, this distributed design would fail to work as it cannot even incorporate local information such as voltage magnitude measurements. We envision that a future voltage controller should be cognizant of various communication scenarios. To this end, we develop an integrated design that can achieve the dual objectives in terms of flexible adaptivity to variable rate of communications and global optimality of voltage regulation performance in Ch. 5. Such an innovative hybrid design has the potential of unifying the currently separated framework of either local or distributed control schemes.

With a high penetration of DERs in the networks, MGs are likely to emerge as a means to advance power and cyber-physical resiliency in future distribution networks. In Ch. 6, we developed the distributed secondary frequency control design for isolated MGs for improving the frequency stability. Interestingly, network dynamics is seamlessly governed by the physical system itself and the convergence is guaranteed without assuming the time-scale separation of the hierarchical control design methodologies. Extending this work to an industrial MG network that follows the IEC 61850 communication protocol, similar frequency regulation objective is introduced and solved by a decentralized ADMM-based algorithm in Ch. 7. As cyber threats are significant concern in smart grid, we proposed several countermeasures for malicious attacks on the communication network for both designs of Chs. 6-7. Anomaly detection and localization strategies are developed based on the metrics of optimization-related variables.

This thesis provides a number of control strategies that reduce communication complexity and potentially enhance cybersecure resiliency. Nonetheless, future work is needed to realize these novel schemes. First, the aforementioned fast-acting control designs would require real-time sensing and communications, and thus cyber attack surfaces should be well defined. Second, machine learning tools could be deployed to analyze real-time data and facilitate automatic anomaly detection schemes. Last but not least, the interactions among slow-acting devices should also be considered and carefully examined to further improve overall control performance.

APPENDIX A

MULTI-PHASE DVC IMPLEMENTATIONS

The multi-phase voltage regulation problem is given by

$$\{\mathbf{v}^*, \mathbf{q}^*\} := \arg \min_{\mathbf{v}, \mathbf{q}} \quad f(\mathbf{v}, \mathbf{q}) := \|\mathbf{v} - \mathbf{u}\|_2^2 + \mathbf{q}^T \mathbb{C} \mathbf{q} \quad (\text{A.1a})$$

$$\text{subject to} \quad \mathbf{g}(\mathbf{v}, \mathbf{q}) = \mathbf{0} \quad (\text{A.1b})$$

$$\underline{\mathbf{q}} \leq \mathbf{q} \leq \overline{\mathbf{q}}, \quad (\text{A.1c})$$

where the diagonal matrix $\mathbb{C} := \text{diag}(\{c_j^\phi\}_{j \in \mathcal{N}, \phi \in \mathcal{P}})$ has the non-negative coefficients used for regularizing the network-wide DER output \mathbf{q} . The constraint (A.1b) represents the nonlinear relation between \mathbf{v} and \mathbf{q} based on the aforementioned AC power flow model (2.1)-(2.2) for given system operating conditions (see [9, Chs. 3-7] for explicit expression of $\mathbf{g}(\mathbf{v}, \mathbf{q})$). Last, the bound constraints in (A.1c) are used to account for VAR resource limits due to either the apparent power rating or power factor constraints of DER inverters [29].

We let (i, ψ) represent the node at phase- ψ of bus i , while the subset $\mathcal{N}_j^\phi \subseteq \mathcal{N}_A := \{(i, \psi)\}_{i \in \mathcal{N}, \psi \in \mathcal{P}}$ consists of all the nodes across the network coupled to node (j, ϕ) according to the AC power flow-based equality constraints in (2.1)-(2.2). We reformulate (A.1) to demonstrate the decomposable problem structure using a consensus version of ADMM. To this end, we define two groups of variables: $\mathbf{x} := \{\mathbf{x}_j^\phi\}_{j \in \mathcal{N}, \phi \in \mathcal{P}}$ and $\mathbf{z} := \{\mathbf{z}_j^\phi\}_{j \in \mathcal{N}, \phi \in \mathcal{P}}$, where

$$\mathbf{x}_j^\phi := \left\{ \left\{ \nu_{(i, \psi)j}^\phi \right\}_{(i, \psi) \in \mathcal{N}_j^\phi}, \delta_j^\phi \right\}, \text{ and } \mathbf{z}_j^\phi := \{v_j^\phi, q_j^\phi\}.$$

Clearly, the vector \mathbf{z} includes all the optimization variables of (A.1), the *decision* variables. Each vector \mathbf{x}_j^ϕ contains local *estimates* of some decision variables in \mathbf{z} related to node (j, ϕ) , termed as *auxiliary variables*. For example, the auxiliary variable $\nu_{(i, \psi)j}^\phi$ at node (j, ϕ) estimates the decision

variable v_i^ψ of node (i, ψ) in \mathbf{z} while δ_j^ϕ maps to q_j^ϕ . By enforcing each decision variable to consent with its local auxiliary copies, one can reformulate (A.1) as

$$\min_{\mathbf{x}, \mathbf{z}} \sum_{j=1}^N \sum_{\phi \in \mathcal{P}} f_j^\phi(\mathbf{x}_j^\phi) \quad (\text{A.2a})$$

$$\text{subject to } \nu_{(i,\psi)j}^\phi = v_i^\psi, \forall \phi, \forall j, \forall (i, \psi) \in \mathcal{N}_j^\phi \quad (\text{A.2b})$$

$$\delta_j^\phi = q_j^\phi, \forall \phi, \forall j \in \mathcal{N} \quad (\text{A.2c})$$

$$\mathbf{x}_j^\phi \in \mathcal{X}_j^\phi, \forall \phi, \forall j \in \mathcal{N} \quad (\text{A.2d})$$

$$\mathbf{z}_j^\phi \in \mathcal{Z}_j^\phi, \forall \phi, \forall j \in \mathcal{N}, \quad (\text{A.2e})$$

where the summand $f_j^\phi(\mathbf{x}_j^\phi) := (\nu_{(j,\phi)j}^\phi - \mu_j^\phi)^2 + c_j^\phi(\delta_j^\phi)^2$ while the constraints (A.2d)-(A.2e) represent those in (A.1) by defining

$$\mathcal{X}_j^\phi := \{\mathbf{x}_j^\phi | g_j^\phi(\mathbf{x}_j^\phi) = 0\}, \quad \mathcal{Z}_j^\phi := \{(v_j^\phi, q_j^\phi) | \underline{q}_j^\phi \leq q_j^\phi \leq \bar{q}_j^\phi\}.$$

Since the equality constraints in (A.2b)-(A.2c) enforce the consensus between each decision variable and its local auxiliary copy, the reformulated problem (A.2) is equivalent to (A.1). Thanks to this decomposable problem structure, each node (j, ϕ) in the network has its own local objective function f_j^ϕ with a local set of constraints coupled by the decision variables among neighboring nodes. As detailed soon, this structure is key to the development of our control design. Note that vector \mathbf{z} includes all the optimization variables in (A.1), termed as *decision* variables. Per node (j, ϕ) , each vector \mathbf{x}_j^ϕ contains a local copy of relevant decision variables, termed as *auxiliary* variables. For example, the auxiliary variable $\nu_{(i,\psi)j}^\phi$ locally updated at node (j, ϕ) maps to v_i^ψ in \mathbf{z} , while δ_j^ϕ to q_j^ϕ . By enforcing the local auxiliary and decision variables to consent, one can reformulate (A.1).

To develop the ADMM updates for (A.2), we form the augmented Lagrangian function by denoting $\mathbf{y}_j^\phi := \{\{\lambda_{(i,\psi)j}^\phi\}_{(i,\psi) \in \mathcal{N}_j^\phi}, \theta_j^\phi\}$ as the multipliers for the consensus constraints (A.2d)-(A.2e) per node (j, ϕ) . For a given penalty coefficient $\rho > 0$ and by concatenating all multipliers in

$\mathbf{y} := \{\mathbf{y}_j^\phi\}_{j,\phi}$, the overall Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{y}) = \sum_{j,\phi} \mathcal{L}_j^\phi(\mathbf{x}_j^\phi, \mathbf{z}, \mathbf{y}_j^\phi)$ where

$$\begin{aligned} \mathcal{L}_j^\phi(\mathbf{x}_j^\phi, \mathbf{z}, \mathbf{y}_j^\phi) &= (\nu_{(j,\phi)j}^\phi - \mu_j^\phi)^2 + c_j^\phi (\delta_j^\phi)^2 \\ &+ \sum_{(i,\psi) \in \mathcal{N}_j^\phi} \lambda_{(i,\psi)j}^\phi (\nu_{(i,\psi)j}^\phi - v_i^\psi) + \theta_j^\phi (\delta_j^\phi - q_j^\phi) \\ &+ \sum_{(i,\psi) \in \mathcal{N}_j^\phi} \frac{\rho}{2} (\nu_{(i,\psi)j}^\phi - v_i^\psi)^2 + \frac{\rho}{2} (\delta_j^\phi - q_j^\phi)^2. \end{aligned} \quad (\text{A.3})$$

Note that \mathcal{L} is different from the conventional Lagrangian function; all equality constraints are regularized by an additional quadratic term. Interestingly, this modification does not affect the optimal solution in the primal domain since its corresponding gradient is essentially zero at any feasible point of (A.2b). Using the augmented Lagrangian $\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{y})$, the ADMM works by cyclically minimizing each of the three variable groups, \mathbf{x} , \mathbf{z} , and \mathbf{y} , while fixing others. Thanks to the decomposable reformulation in (A.2), each of these three sub-problems at the k -th iteration can be solved locally per node (j, ϕ) . The following three steps in (S1-S3) constitute the basis of our proposed DVC design.

(S1) Update \mathbf{x} : As \mathcal{L} totally decouples into \mathcal{L}_j^ϕ for each node (j, ϕ) , minimizing \mathbf{x}_j^ϕ involves only the variables in $\mathbf{z}[k]$ and $\mathbf{y}[k]$ that are related to \mathbf{x}_j^ϕ . Thus, upon receiving these variables from all neighboring nodes in \mathcal{N}_j^ϕ , the following update can be implemented locally at node (j, ϕ) :

$$\mathbf{x}_j^\phi[k+1] := \arg \min_{\mathbf{x}_j^\phi \in \mathcal{X}_j^\phi} \mathcal{L}_j^\phi(\mathbf{x}_j^\phi, \mathbf{z}[k], \mathbf{y}[k]). \quad (\text{A.4})$$

(S2) Update \mathbf{z} : Likewise, the decision variables are updated as

$$\mathbf{z}[k+1] := \arg \min_{\mathbf{z} \in \mathcal{Z}} \mathcal{L}(\mathbf{x}[k+1], \mathbf{z}, \mathbf{y}[k]), \quad (\text{A.5})$$

which is a box-constrained quadratic programming (QP) problem. It can be solved by projecting the unconstrained solutions to \mathcal{Z} [46, Ch. 2]. Interestingly, if all multipliers are initialized to zero by setting $\mathbf{y}[1] = \mathbf{0}$, the sum $\sum_{(i,\psi) \in \mathcal{N}_j^\phi} \lambda_{j(i,\psi)}^\phi[k]$ remains zero throughout all ADMM iterations. Thus, we can solve (A.5) by setting the gradient of (A.5) to zero,

which would decouple into the following updates per node (j, ϕ) :

$$v_j^\phi[k+1] = \frac{1}{|\mathcal{N}_j^\phi|} \sum_{(i,\psi) \in \mathcal{N}_j^\phi} \nu_{(j,\phi)i}^\psi[k+1], \quad (\text{A.6a})$$

$$q_j^\phi[k+1] = \mathbb{P}_j^\phi \left[\theta_j^\phi[k]/\rho + \delta_j^\phi[k+1] \right], \quad (\text{A.6b})$$

where \mathbb{P}_j^ϕ projects any input into the interval $[\underline{q}_j^\phi, \bar{q}_j^\phi]$.

(S3) Update \mathbf{y} : Each multiplier is linearly updated by the iterative mismatch of the corresponding constraint per node (j, ϕ) :

$$\begin{aligned} \lambda_{(i,\psi)j}^\phi[k+1] &= \lambda_{(i,\psi)j}^\phi[k] + \rho(\nu_{(i,\psi),j}^\phi[k+1] - v_i^\psi[k+1]), \\ &\quad \forall (i, \psi) \in \mathcal{N}_j^\phi, \\ \theta_j^\phi[k+1] &= \theta_j^\phi[k] + \rho(\delta_j^\phi[k+1] - q_j^\phi[k+1]). \end{aligned} \quad (\text{A.7})$$

Clearly, the sum $\sum_{(i,\psi) \in \mathcal{N}_j^\phi} \lambda_{j(i,\psi)}^\phi[k+1]$ is guaranteed to stay zero if initialized to be zero. This corroborates the derivations in (A.6).

The number of auxiliary variables, and thus the complexity, of the proposed ADMM-based DVC design grows as the size of neighborhood $|\mathcal{N}_j^\phi|$ increases. Due to the non-linear power flow couplings, the subset \mathcal{N}_j^ϕ may contain almost all network nodes under the (lossy) AC power flow models. This network-wide coupling greatly challenges the practical implementation of the updates in (S1)-(S3), as it would require a fully-connected communication network across all nodes. To tackle this problem, one could reduce the size of \mathcal{N}_j^ϕ using the linearized \mathbf{q} - \mathbf{v} sensitivity factors in (2.12). Using (2.12), the (S1) update in (A.4) boils down to the following linearly-constrained QP:

$$\mathbf{x}_j^\phi[k+1] := \arg \min_{(\mathbf{h}_j^\phi)^T \mathbf{x}_j^\phi = w_j^\phi} \frac{1}{2} (\mathbf{x}_j^\phi)^T \mathbf{A}_j^\phi \mathbf{x}_j^\phi + (\mathbf{b}_j^\phi[k])^T \mathbf{x}_j^\phi. \quad (\text{A.8})$$

Matrix \mathbf{A}_j^ϕ is a diagonal and non-singular matrix, and $\mathbf{b}_j^\phi[k]$ contains the information from $\mathbf{z}[k]$ and $\mathbf{y}[k]$. The single linear constraint in (A.8) represents the set $\mathcal{X}_j^\phi := \{\mathbf{x}_j^\phi \mid \sum_{(i,\psi) \in \mathcal{N}_j^\phi} B_{j(i,\psi)}^\phi \nu_{(i,\psi)j}^\phi -$

$q_j^\phi = w_j^\phi\}$ where $B_{j(i,\psi)}^\phi$ is the $(j, (i, \psi))$ -th entry of \mathbf{B}^ϕ . Analysis of the first-order optimality conditions of (A.8) leads to a closed-form solution. Upon defining

$$t_j^\phi[k] := (w_j^\phi + (\mathbf{h}_j^\phi)^T (\mathbf{A}_j^\phi)^{-1} \mathbf{b}_j^\phi[k]) / ((\mathbf{h}_j^\phi)^T (\mathbf{A}_j^\phi)^{-1} \mathbf{h}_j^\phi),$$

(S1) becomes

$$\begin{aligned} \nu_{(j,\phi)j}^\phi[k+1] &= \frac{1}{\rho+2} (t_j^\phi[k] B_{j(j,\phi)}^\phi + 2\mu_j^\phi - \lambda_{(j,\phi)j}^\phi[k] + \rho v_j^\phi[k]), \\ \nu_{(i,\psi)j}^\phi[k+1] &= \frac{1}{\rho} (t_j^\phi[k] B_{j(i,\psi)}^\phi - \lambda_{(i,\psi)j}^\phi[k] + \rho v_i^\psi[k]), \\ \forall (i, \psi) &\in \mathcal{N}_j^\phi \setminus (j, \phi), \\ \delta_j^\phi[k+1] &= \frac{1}{\rho+2w_j^\phi} (-t_j^\phi[k] - \theta_j^\phi[k] + \rho q_j^\phi[k]). \end{aligned}$$

Our multi-phase DVC design not only requires minimal communication network deployment, but also enjoys efficient computations with only linear updates and simple projection in (S1)-(S3).

The ADMM iterations in (S1)-(S3) have been developed assuming the availability of w_j^ϕ at every node (j, ϕ) , which will change accordingly with network operating conditions. With the full knowledge of complex power injection everywhere, one can compute \mathbf{w} using (2.12). However, such a centralized approach would require two-way communications between each local sensor/controller and a centralized computer, infeasible under fast system dynamics. Moreover, this direct approach to calculate \mathbf{w} also suffers from the approximation error introduced by the linearized model (2.12).

To conveniently adapt to time-varying system dynamics, we propose to enhance the proposed DVC design by updating $w_j^\phi[k]$ at every iteration k through local information exchange. Specifically, the proposed online control design would seamlessly utilize the latest ADMM iterate $q_j^\phi[k]$ updated in (S2) as the instantaneous VAR input signal to control local inverters. Note that its update rule in (A.6b) guarantees that every instance of $q_j^\phi[k]$ is feasible to dispatch as it lies within $[q_j^\phi, \bar{q}_j^\phi]$ due to the projection by \mathbb{P}_j^ϕ . This way, the local VAR injection output is set to be $q_j^\phi[k]$ once its latest value is updated in (S2) on-the-fly, without requiring the algorithm to converge. Accordingly, by measuring the instantaneous bus voltage to $\bar{v}_j^\phi[k]$, each node (j, ϕ) can collect all

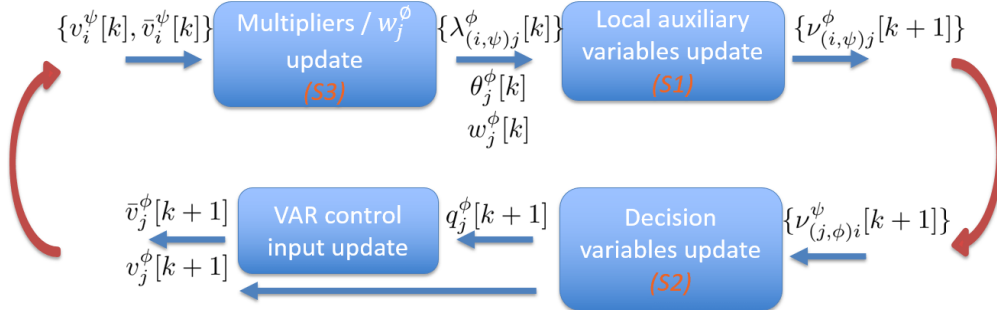


Figure A.1: Multi-phase online implementation of the proposed DVC design that adapts to dynamic system conditions and constantly updates the VAR control inputs. Rectangular blocks denote local computational tasks of ADMM iterations and control updates, while the two vertical arrows correspond to the communications exchange among node (j, ϕ) and the neighboring nodes.

neighboring nodes' voltage data to update

$$w_j^\phi[k] = \sum_{(i,\psi) \in \mathcal{N}_j^\phi} B_{j(i,\psi)}^\phi \bar{v}_i^\psi[k] - q_j^\phi[k]. \quad (\text{A.9})$$

This update simply needs node (j, ϕ) to know the self- and mutual-reactance of all lines incident to bus j . Clearly, this proposed design of updating $w_j^\phi[k]$ nicely adapts to dynamically varying operating conditions because (A.9) seamlessly incorporates the instantaneous system voltage information. Such an online feedback-based DVC design is attractive when dealing with the increasing variability of renewable generations and elastic loads in up-and-coming distribution networks.

Fig. A.1 offers a schematic of the proposed ADMM-based DVC implementation. Every iteration consists of the *local computations* of (S1)-(S3) per node (j, ϕ) (the blocks), in addition to two steps of information exchange among all neighboring nodes (the vertical arrows). The requirement on *sensing hardware* is minimal while the *communication architecture* is fully localized since each node's computations would just need voltage measurement and decision/auxiliary variables from neighboring nodes. Furthermore, the local computational tasks, either the sub-problem in (S1) or the linear updates in (S2-S3), can be executed very efficiently. Upon updating $q_j^\phi[k+1]$ in (S2), node (j, ϕ) inputs it as the control signal for local VAR resources and then measures the voltage $\bar{v}_j^\phi[k+1]$.

APPENDIX B

LINE FLOW DYNAMICS

Given a lossless network under Assumption 4, we define the branch power flow from node i to node j as

$$f_{ij}(\theta_{ij}^0) = \frac{V_i V_j}{X_{ij}} \sin(\theta_{ij}^0),$$

where the nominal phase angle difference $\theta_{ij}^0 := (\theta_i^0 - \theta_j^0)$. By using the Taylor series expansion and defining phase angle difference between nodes i and j as $\theta_{ij} := (\theta_i - \theta_j)$, the approximated branch power flow at θ_{ij} is

$$f_{ij}(\theta_{ij}) = \frac{V_i V_j}{X_{ij}} \sin \theta_{ij}^0 + \frac{V_i V_j}{X_{ij}} \cos \theta_{ij}^0 (\theta_{ij} - \theta_{ij}^0).$$

Taking its time derivative yields

$$\dot{f}_{ij}(\theta_{ij}) = \frac{V_i V_j}{X_{ij}} \cos \theta_{ij}^0 \dot{\theta}_{ij} = \frac{V_i V_j}{X_{ij}} \cos(\theta_i^0 - \theta_j^0) (\dot{\theta}_i - \dot{\theta}_j) = B_{ij}(\omega_i - \omega_j).$$

REFERENCES

- [1] “Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations,” U.S. Department of Energy, Tech. Rep., April 2004.
- [2] W. A. Wulf, “Great achievement and grand challenges,” *The Bridge*, vol. 30, pp. 5–10, 2010.
- [3] H. Zhu and H. J. Liu, “Fast local voltage control under limited reactive power: Optimality and stability analysis,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sept 2016.
- [4] H. J. Liu, W. Shi, and H. Zhu, “Decentralized dynamic optimization for power network voltage control,” *IEEE Transactions on Signal and Information Processing over Networks*, 2016, to be published.
- [5] —, “Distributed voltage control in distribution networks: Online and robust implementations,” *IEEE Trans. Smart Grid*, 2016, to be published.
- [6] —, “Hybrid voltage control in distribution networks under limited communication rates,” *IEEE Trans. Smart Grid*, 2016, submitted.
- [7] L. Y. Lu, H. J. Liu, and H. Zhu, “Distributed secondary control for isolated microgrids under malicious attacks,” in *Proc. IEEE NAPS*, Sept 2016, pp. 1–6.
- [8] H. J. Liu, M. Backes, R. Macwan, and A. Valdes, “Coordination of ders in microgrids with cybersecure resilient decentralized secondary frequency control,” in *2018 51st Hawaii International Conference on System Sciences*, Jan 2018, to be published.
- [9] W. H. Kersting, *Distribution System Modeling and Analysis*. CRC Press LLC, 2002, ch. 1.
- [10] B. A. Robbins, H. Zhu, and A. D. Dominguez-Garcia, “Optimal tap setting of voltage regulation transformers in unbalanced distribution systems,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 256–267, Jan 2016.
- [11] E. Dall’Anese, H. Zhu, and G. Giannakis, “Distributed optimal power flow for smart microgrids,” *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.
- [12] L. Gan and S. Low, “Convex relaxations and linear approximation for optimal power flow in multiphase radial networks,” in *Power Systems Computation Conference (PSCC), 2014*, Aug 2014, pp. 1–9.

- [13] B. A. Robbins and A. D. Domínguez-García, “Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems,” *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, July 2016.
- [14] V. Kekatos, G. Wang, A. Conejo, and G. Giannakis, “Stochastic reactive power management in microgrids with renewables,” *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3386–3395, Nov 2015.
- [15] P. Sulc, S. Backhaus, and M. Chertkov, “Optimal distributed control of reactive power via the alternating direction method of multipliers,” *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 968–977, Dec 2014.
- [16] B. A. Robbins and A. D. Dominguez-Garcia, “Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems,” *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, July 2016.
- [17] M. Farivar, L. Chen, and S. Low, “Equilibrium and dynamics of local voltage control in distribution systems,” in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, Dec 2013, pp. 4329–4334.
- [18] M. Baran and F. Wu, “Optimal capacitor placement on radial distribution systems,” *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan 1989.
- [19] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River: Prentice hall, 2001.
- [20] A. J. Wood, B. F. Wollengery, and G. B. Sheble, *Power Generation, Operation, and Control*, 3rd ed. New York, NY: John Wiley and Sons, 2013, ch. 6.
- [21] S. H. Low, “Convex relaxation of optimal power flow-part ii: Exactness,” *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 2, pp. 177–189, June 2014.
- [22] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, “Voltage regulation algorithms for multiphase power distribution grids,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3913–3923, Sept 2016.
- [23] J. A. Mueller, M. Rasheduzzaman, and J. W. Kimball, “A model modification process for grid-connected inverters used in islanded microgrids,” *IEEE Trans. Energy Conversion*, vol. 31, no. 1, pp. 240–250, March 2016.
- [24] E. Dall’Anese, S. V. Dhople, and G. B. Giannakis, “Photovoltaic inverter controllers seeking ac optimal power flow solutions,” *IEEE Trans. Power Syst.*, 2016, (early access).
- [25] M. Farivar, R. Neal, C. Clarke, and S. Low, “Optimal inverter control in distribution systems with high PV penetration,” in *2012 IEEE PES GM*, July 2012, pp. 1–7.
- [26] B. Robbins, C. Hadjicostis, and A. Dominguez-Garcia, “A two-stage distributed architecture for voltage control in power distribution systems,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1470–1482, May 2013.

- [27] E. Dall’Anese, H. Zhu, and G. Giannakis, “Distributed optimal power flow for smart microgrids,” *IEEE Trans. on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sept 2013.
- [28] P. Carvalho, P. F. Correia, and L. Ferreira, “Distributed reactive power generation control for voltage rise mitigation in distribution networks,” *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 766–772, May 2008.
- [29] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, “Options for control of reactive power by distributed photovoltaic generators,” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063–1073, June 2011.
- [30] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.
- [31] H. R. Feyzmahdavian and M. Johansson, “On the convergence rates of asynchronous iterations,” in *IEEE 53rd Conf. Decision and Control (CDC)*. IEEE, 2014, pp. 153–159.
- [32] P. Scott and S. Thiébaux, “Dynamic optimal power flow in microgrids using the alternating direction method of multipliers,” *ArXiv e-prints*, Oct. 2014.
- [33] S. Gill, I. Kockar, and G. W. Ault, “Dynamic optimal power flow for active distribution networks,” *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 121–131, Jan 2014.
- [34] K. Slavakis, S.-J. Kim, G. Mateos, and G. Giannakis, “Stochastic approximation vis-a-vis online learning for big data analytics [lecture notes],” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 124–129, 2014.
- [35] L. Bottou, “Large-scale machine learning with stochastic gradient descent,” in *Proceedings of the 19th International Conference on Computational Statistics*, 2010.
- [36] R. Johnson and T. Zhang, “Accelerating stochastic gradient descent using predictive variance reduction,” in *Advances in Neural Information Processing Systems 26*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 315–323.
- [37] Q. Ling and A. Ribeiro, “Decentralized dynamic optimization through the alternating direction method of multipliers,” *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1185–1197, March 2014.
- [38] Z. J. Towfic and A. H. Sayed, “Adaptive penalty-based distributed stochastic convex optimization,” *IEEE Trans. Signal Processing*, vol. 62, no. 15, pp. 3924–3938, 2014.
- [39] A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, “A class of prediction-correction methods for time-varying convex optimization,” *ArXiv e-prints*, Sep. 2015. [Online]. Available: <http://arxiv.org/abs/1509.05196>
- [40] K. Zhou and S. Roumeliotis, “Multirobot active target tracking with combinations of relative observations,” *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 678–695, Aug 2011.

- [41] F. Jakubiec and A. Ribeiro, "D-map: Distributed maximum a posteriori probability estimation of dynamic systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 450–466, Jan 2013.
- [42] H. Y. Su, F. M. Kang, and C. W. Liu, "Transmission grid secondary voltage control method using pmu data," *IEEE Trans. Smart Grid*, 2016, to be published.
- [43] L. Y. Lu and C. C. Chu, "Consensus-based secondary frequency and voltage droop control of virtual synchronous generators for isolated ac micro-grids," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 5, no. 3, pp. 443–455, Sept 2015.
- [44] H. J. Liu, R. Macwan, N. Alexander, and H. Zhu, "A methodology to analyze conservation voltage reduction performance using field test data," in *Proc. 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2014, pp. 529–534.
- [45] P. J. Huber, *Robust Statistics*. Wiley New York, 1981.
- [46] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [47] D. G. Luenberger, *Introduction to Dynamic Systems : Theory, Models, and Applications*. J. Wiley & Sons, 1979.
- [48] M. Hassanzadeh, C. Evrenosoglu, and L. Mili, "A short-term nodal voltage phasor forecasting method using temporal and spatial correlation," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–10, 2015.
- [49] E. Cotilla-Sanchez, P. Hines, C. Barrows, and S. Blumsack, "Comparing the topological and electrical structure of the north american electric power infrastructure," *IEEE Systems Journal*, vol. 6, no. 4, pp. 616–626, Dec 2012.
- [50] H. J. Liu, W. Shi, and H. Zhu, "Dynamic decentralized voltage control for power distribution networks," in *Proc. 2016 IEEE Statistical Signal Processing Workshop (SSP)*, June 2016, pp. 1–5.
- [51] D. Bertsekas, "Incremental gradient, subgradient, and proximal methods for convex optimization: A Survey," *Optimization for Machine Learning*, vol. 2010, pp. 1–38, 2011.
- [52] W. Shi, Q. Ling, G. Wu, and W. Yin, "A proximal gradient algorithm for decentralized composite optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6013–6023, 2015.
- [53] D. Davis and W. Yin, "Convergence rate analysis of several splitting schemes," *arXiv preprint arXiv:1406.4834*, 2014.
- [54] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Science & Business Media, 2013, vol. 87.
- [55] "IEEE PES Distribution Test Feeders," Sep. 2010. [Online]. Available: <http://www.ewh.ieee.org/soc/pes/dsacom/testfeeders.html>

- [56] “IEEE application guide for IEEE std 1547(TM), IEEE standard for interconnecting distributed resources with electric power systems,” *IEEE Std 1547.2-2008*, pp. 1–217, April 2009.
- [57] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, 2011.
- [58] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem, “Asynchronous distributed optimization using a randomized alternating direction method of multipliers,” in *IEEE 52nd Conf. Decision & Control (CDC)*, Dec 2013, pp. 3671–3676.
- [59] H. Zhu, G. Giannakis, and A. Cano, “Distributed in-network channel decoding,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3970–3983, Oct 2009.
- [60] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, “Distributed reactive power feedback control for voltage regulation and loss minimization,” *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 966–981, April 2015.
- [61] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, “On the linear convergence of the admm in decentralized consensus optimization,” *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1750–1761, April 2014.
- [62] A. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Pearson/Prentice Hall, 2000.
- [63] R. Zhang and J. Kwok, “Asynchronous distributed admm for consensus optimization,” in *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, 2014.
- [64] T.-H. Chang, M. Hong, W.-C. Liao, and X. Wang, “Asynchronous distributed admm for large-scale optimization- part i: Algorithm and convergence analysis,” *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3118–3130, 2015.
- [65] *OpenDSS Simulation Tool*, EPRI. [Online]. Available: <http://smartgrid.epri.com/SimulationTool>
- [66] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson, “Optimal parameter selection for the alternating direction method of multipliers (ADMM): quadratic problems,” *ArXiv e-prints*, Jun. 2013.
- [67] *UCI Machine Learning Repository, Individual Household Electric Power Consumption Dataset*. [Online]. Available: <http://archive.ics.uci.edu/ml/>
- [68] A. Nedic and A. Ozdaglar, “Subgradient methods for saddle-point problems,” *Journal of Optimization Theory and Applications*, vol. 142, no. 1, pp. 205–228, 07 2009. [Online]. Available: <http://search.proquest.com/docview/196601816?accountid=14553>
- [69] D. Feijer and F. Paganini, “Krasovskii’s method in the stability of network control,” in *2009 American Control Conference*, June 2009, pp. 3292–3297.

- [70] —, “Stability of primal-dual gradient dynamics and applications to network optimization,” *Automatica*, vol. 46, no. 12, pp. 1974–1981, Dec. 2010.
- [71] L. Chen, C. Zhao, and N. Li, “Connecting automatic generation control and economic dispatch from an optimization view,” *IEEE Trans. Control Netw. Syst.*, vol. PP, no. 99, pp. 1–1, 2015.
- [72] N. Li, C. Zhao, and L. Chen, “Connecting automatic generation control and economic dispatch from an optimization view,” *IEEE Trans. Control Netw. Syst.*, vol. 3, pp. 254–264, Sept 2016.
- [73] V. D. Blondel, O. Bournez, P. Koiran, and J. N. Tsitsiklis, “The stability of saturated linear dynamical systems is undecidable,” *Journal of Computer and System Sciences*, vol. 62, no. 3, pp. 442–462, 2001.
- [74] D. P. Bertsekas, “Incremental gradient, subgradient, and proximal methods for convex optimization: A survey,” *Optimization for Machine Learning*, vol. 2010, no. 1-38, p. 3, 2011.
- [75] W. Shi, Q. Ling, G. Wu, and W. Yin, “A proximal gradient algorithm for decentralized composite optimization,” *IEEE Transactions on Signal Processing*, vol. 63, no. 22, pp. 6013–6023, 2015.
- [76] Y. Nesterov, *Introductory Lectures on Convex Optimization: A basic course*. Springer Science & Business Media, 2013, vol. 87.
- [77] W. Shi, Q. Ling, G. Wu, and W. Yin, “EXTRA: An Exact First-Order Algorithm for Decentralized Consensus Optimization,” *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.
- [78] Z. Peng, Y. Xu, M. Yan, and W. Yin, “Arock: an algorithmic framework for asynchronous parallel coordinate updates,” *SIAM Journal on Scientific Computing*, vol. 38, no. 5, 2016.
- [79] M. Hong and T.-H. Chang, “Stochastic proximal gradient consensus over time-varying networks,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 4776–4780.
- [80] A. Nedich, A. Olshevsky, and W. Shi, “Achieving geometric convergence for distributed optimization over time-varying graphs,” *arXiv preprint arXiv:1607.03218*, 2016.
- [81] P. Bianchi, W. Hachem, and I. Franck, “A stochastic primal-dual algorithm for distributed asynchronous composite optimization,” in *Signal and Information Processing, 2014 IEEE Global Conference on (GlobalSip)*, December 2014.
- [82] IEEE-PES Task Force on Microgrid Control, “Trends in microgrid control,” *IEEE Trans. Smart Grid*, vol. 5, pp. 1905–1919, Jul. 2014.
- [83] J. Guerrero, J. Vasquez, J. Matas, L. de Vicuña, and M. Castilla, “Hierarchical control of droop-controlled AC and DC microgrids – general approach toward standardization,” *IEEE Trans. Ind. Electron.*, vol. 58, pp. 158–172, Jan 2011.

- [84] A. Bidram and A. Davoudi, “Hierarchical structure of microgrids control system,” *IEEE Trans. Smart Grid*, vol. 3, pp. 1963–1976, Dec 2012.
- [85] M. C. Chandorkar, D. M. Divan, and R. Adapa, “Control of parallel connected inverters in standalone AC supply systems,” *IEEE Trans. Ind. Appl.*, vol. 29, pp. 136–143, Jan 1993.
- [86] J. W. Simpson-Porco, F. Dörfler, and F. Bullo, “Synchronization and power sharing for droop-controlled inverters in islanded microgrids,” *Automatica*, vol. 49, pp. 2603–2611, Sept. 2013.
- [87] L. Y. Lu and C. C. Chu, “Consensus-based droop control synthesis for multiple DICs in isolated micro-grids,” *IEEE Trans. Power Syst.*, vol. 30, pp. 2243–2256, Nov 2014.
- [88] S. T. Cady, A. D. Domínguez-García, and C. N. Hadjicostis, “A distributed generation control architecture for islanded AC microgrids,” *IEEE Trans. Control Syst. Technol.*, vol. 23, pp. 1717–1735, Sept 2015.
- [89] N. Soni, S. Doolla, and M. C. Chandorkar, “Improvement of transient response in microgrids using virtual inertia,” *IEEE Trans. Power Del.*, vol. 28, no. 3, pp. 1830–1838, July 2013.
- [90] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, “Breaking the hierarchy: Distributed control and economic optimality in microgrids,” *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 3, pp. 241–253, Sept 2016.
- [91] S. Sundaram and C. N. Hadjicostis, “Distributed function calculation via linear iterative strategies in the presence of malicious agents,” *IEEE Trans. Autom. Control*, vol. 56, pp. 1495–1508, July 2011.
- [92] F. Pasqualetti, A. Bicchi, and F. Bullo, “Consensus computation in unreliable networks: A system theoretic approach,” *IEEE Trans. Autom. Control*, vol. 57, pp. 90–104, Jan 2012.
- [93] H. J. LeBlanc, H. Zhang, S. Sundaram, and X. Koutsoukos, “Consensus of multi-agent networks in the presence of adversaries using only local information,” in *Proc. Intl. Conf. on High Confidence Netw. Syst.*, ser. HiCoNS ’12. New York, NY, USA: ACM, 2012, pp. 1–10.
- [94] L. Tseng and N. H. Vaidya, “Fault-tolerant consensus in directed graphs,” in *Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing*, ser. PODC ’15. New York, NY, USA: ACM, 2015, pp. 451–460. [Online]. Available: <http://doi.acm.org/10.1145/2767386.2767399>
- [95] B. Kailkhura, V. S. S. Nadendla, and P. K. Varshney, “Distributed inference in the presence of eavesdroppers: a survey,” *IEEE Communications Magazine*, vol. 53, no. 6, pp. 40–46, June 2015.
- [96] L. Su and N. Vaidya, “Multi-agent optimization in the presence of byzantine adversaries: Fundamental limits,” in *2016 American Control Conference (ACC)*, July 2016, pp. 7183–7188.

- [97] L. Su and N. H. Vaidya, "Fault-tolerant multi-agent optimization: Optimal iterative distributed algorithms," in *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, ser. PODC '16. New York, NY, USA: ACM, 2016, pp. 425–434. [Online]. Available: <http://doi.acm.org/10.1145/2933057.2933105>
- [98] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, pp. 645–658, Dec 2011.
- [99] O. Vuković and G. Dàn, "Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks," *IEEE J. Sel. Areas in Commun.*, vol. 32, pp. 1500–1508, July 2014.
- [100] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Trans. Smart Grid*, vol. 5, pp. 580–591, March 2014.
- [101] S. Nabavi and A. Chakraborty, "An intrusion-resilient distributed optimization algorithm for modal estimation in power systems," in *Proc. IEEE CDC*, 2015, pp. 39–44.
- [102] P. Srikantha and D. Kundur, "A der attack-mitigation differential game for smart grid security analysis," *IEEE Trans. Smart Grid*, vol. 7, pp. 1476–1485, May 2016.
- [103] A. Teixeira, K. Paridari, H. Sandberg, and K. H. Johansson, "Voltage control for interconnected microgrids under adversarial actions," in *Proc. IEEE ETFA*, Sept 2015, pp. 1–8.
- [104] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems," *IEEE Trans. Smart Grid*, to be published.
- [105] M. Zholbaryssov and A. D. Dominguez-Garcia, "Distributed enforcement of phase-cohesiveness for frequency control of islanded inverter-based microgrids," *IEEE Trans. Control Netw. Syst.*, 2017, to be published.
- [106] X. Wu, C. Shen, and R. Iravani, "A distributed, cooperative frequency and voltage control for microgrids," *IEEE Trans. Smart Grid*, 2016, to be published.
- [107] Q. C. Zhong and G. Weiss, "Synchronverters: Inverters that mimic synchronous generators," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 4, pp. 1259–1267, April 2011.
- [108] J. Alipoor, Y. Miura, and T. Ise, "Power system stabilization using virtual synchronous generator with alternating moment of inertia," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 3, no. 2, pp. 451–458, June 2015.
- [109] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1177–1189, May 2014.
- [110] P. Anderson and A. Fouad, *Power System Control and Stability*, 2nd ed. Wiley India Pvt. Limited, 2008. [Online]. Available: <https://books.google.com/books?id=2BXOzA34qBkC>

- [111] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. Autom. Control*, vol. 59, pp. 1177–1189, May 2014.
- [112] N. Ainsworth and S. Grijalva, "A structure-preserving model and sufficient condition for frequency synchronization of lossless droop inverter-based AC networks," *IEEE Trans. Power Syst.*, vol. 28, pp. 4310–4319, Nov. 2013.
- [113] F. Dörfler, M. Chertkov, and F. Bullo, "Synchronization in complex oscillator networks and smart grids," *Proceedings of the National Academy of Sciences*, vol. 110, no. 6, pp. 2005–2010, 2013. [Online]. Available: <http://www.pnas.org/content/110/6/2005.abstract>
- [114] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. & Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.
- [115] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," *IEEE Trans. Power Syst.*, 2016, to be published.
- [116] R. Olfati-Saber, J. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, pp. 215–233, Jan. 2007.
- [117] C. Dufour and J. Belanger, "On the use of real-time simulation technology in smart grid research and development," *IEEE Trans. Ind. Appl.*, vol. 50, pp. 3963–3970, Nov 2014.
- [118] D. V. Dollen, "Report to nist on the smart grid interoperability standards roadmap," EPRI, Tech. Rep., 2009.
- [119] *Communication Networks and Systems for Power Utility Automation - All Parts*, IEC 61850 Std.
- [120] A. Ruiz-Alvarez, A. Colet-Subirachs, F. A.-C. Figuerola, O. Gomis-Bellmunt, and A. Sudria-Andreu, "Operation of a utility connected microgrid using an IEC 61850-based multi-level management system," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 858–865, June 2012.
- [121] *Communication Networks and Systems for Power Utility Automation - Part 7-420: Basic Communication Structure - Distributed Energy Resources Logical Nodes*, IEC 61850-7-420 Std.
- [122] *IEC 61850 Communication Protocol Manual*, ABB, 2 2011, available at https://library.e.abb.com/public/fe4d2c659c6ccad2c125783a00454ae3/1MRK511242-UEN_-_en_Communication_protocol_manual__IEC_61850__650_series__IEC.pdf, version 1.1.
- [123] "Microscada pro sys600," <http://new.abb.com/substation-automation/products/software/microscada-pro/sys600>, accessed: 2017-10-27.
- [124] R. Lee, M. Assante, and T. Conway, "Analysis of the cyber attack on the Ukrainian power grid," E-ISAC Report, March 2016.

- [125] R. Lee, “CRASHOVERRIDE: Analysis of the threat to electric grid operations,” Dragos Inc., March 2017.
- [126] M. Liao and A. Chakraborty, “Optimization algorithms for catching data manipulators in power system estimation loops,” *CoRR*, vol. abs/1608.00299, 2016. [Online]. Available: <http://arxiv.org/abs/1608.00299>
- [127] A. Valdes, C. Hang, P. Panumpabi, N. Vaidya, C. Drew, and D. Ischenko, “Design and simulation of fast substation protection in IEC 61850 environments,” in *2015 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, April 2015, pp. 1–6.
- [128] R. Macwan, C. Drew, P. Panumpabi, A. Valdes, N. Vaidya, P. Sauer, and D. Ishchenko, “Collaborative defense against data injection attack in IEC 61850 based smart substations,” in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.