DATA-DRIVEN COORDINATION OF ASSETS IN POWER DISTRIBUTION SYSTEMS FOR ANCILLARY SERVICE PROVISION

BY

HANCHEN XU

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, 2019

Urbana, Illinois

Doctoral Committee:

Professor Alejandro D. Domínguez-García, Co-chair
Professor Peter W. Sauer, Co-chair
Professor Venugopal V. Veeravalli
Professor Tamer Başar

ABSTRACT

In this dissertation, we develop several data-driven frameworks for coordinating distributed energy resources (DERs) in power distribution systems to provide ancillary services including active power provision and reactive power regulation. The proposed frameworks generally consist of three components, namely (i) an input-output (IO) model of the system describing the relation between the variables of interest to the problem, (ii) an estimator that provides estimates of the parameters that populate the model in (i), and (iii) a controller that uses the model in (i) with the parameters estimated via (ii) to determine the active and/or reactive power injection set-points of the DERs by solving the optimal DER coordination problem (ODCP), which is cast as a static optimization problem. We develop efficient estimation algorithms that utilize measurements to estimate the parameters in the IO model. Special emphasis is devoted to algorithms that address the potential collinearity issue in the measurements, and formulations that significantly reduce the number of parameters to be estimated.

The idea of data-driven coordination is also applied to address the problem of coordinating load tap changers (LTCs)—an important class of assets used for voltage control in distribution networks—using only measurements of voltage magnitudes. Different from the ODCP that is cast as a static optimization problem, the optimal LTC coordination problem is cast as a multi-stage decision-making problem and formulated as a Markov decision process (MDP), in which the unknown power injections are modeled as uncertainty sources. The MDP is solved via a reinforcement learning algorithm to obtain a control policy that maps the voltage magnitude measurements to the optimal tap positions.

The data-driven nature makes the proposed frameworks intrinsically adaptive and robust to changes in operating conditions and power distribution system models, which are illustrated via extensive case studies.

To my family.

ACKNOWLEDGMENTS

First and foremost, I am deeply indebted to my advisors, Prof. Alejandro Domínguez-García and Prof. Peter Sauer. Prof. Domínguez-García is an incredible advisor who has invested a tremendous amount of effort in helping me grow academically as well as personally. He not only shared with me his knowledge and experience, but also showed me the immense fun of research. His diligence, persistence, and passion for teaching and research are highly infectious. I wish I could find as much enthusiasm and love in my future endeavors. Prof. Sauer, who is always kind and patient, has been the light that led me through my toughest time, for which I could never thank him enough. His love for students, power systems, as well as life, has made him a charismatic person, from whom I can learn so much. In addition, I have benefited from his insightful feedback on my research and enjoyed his humor all the time. I could not be blessed with better advisors and role models.

I would also like to thank my other doctoral committee members, Prof. Venugopal Veeravalli and Prof. Tamer Başar. I have enjoyed discussion with Prof. Veeravalli, who can always provide unique perspectives on problems discussed. The collaboration with Prof. Veeravalli has led to the work in Chapter 4. I am also very grateful to Prof. Başar, who has been with me at every milestone of my PhD program, including the qualifying exam, the preliminary exam, and the final exam. He has made my journey in the graduate school more memorable.

My advisor at Tsinghua University, Prof. Chao Lu, has been very supportive throughout my graduate study—from the stage of application till the end of my PhD program. I am very grateful for his trust in me. I would like to extend my thanks to Prof. Fei Wang, who has shared unreservedly with me his experiences and understanding of life. I cannot appreciate more to have such a caring friend and mentor. Special thanks to my physics teacher during high school, Xu Wang, who revealed the true beauty of science to me and

shaped the way I understand the world, and to my chemistry teacher during middle school, Liansheng Tang, whose trust and encouragement have given me the confidence to take all the challenges.

During the past five years, I also had the privilege to work with many extraordinary people in the industry. I would like to express my gratitude to all my mentors and colleagues who have enriched my experiences and broadened my understanding of power systems: Dr. Bri-Mathias Hodge and Dr. Carlo Brancucci from NREL gave me an opportunity to analyze the impacts of renewable energy resources on multiple aspects of power systems; Richard Brown, Gary Helm, and Rami Dirani from PJM provided me with a precious platform where I learned the frontier of electricity markets; Dr. Hongbo Sun from MERL made it possible for me to explore more in the area of reinforcement learning. All these experiences have directly or indirectly contributed to this dissertation, and moreover, my professional development.

My life in graduate school would be incomplete without my friends and colleagues here. I would like to express my deep gratitude to Ti Xu, who offered me tremendous help on every aspect from my life to my research. Ti was a trustworthy and caring friend, with whom I shared a lot of joy and pain. He was a man who loved the life so much but left us unexpectedly. May he rest in peace. I am also grateful to have Kaiqing Zhang as a great collaborator, and more importantly, a trustworthy friend whom I can always rely on. Heartfelt thanks to Hanzhong Ke and Chongda Liu for their care and company. Sincere thanks go to my colleagues Siddhartha Nigam, Archana Manjunath, Dariush Fooladivanda, Madi Zholbaryssov, and Olaolu Ajala, for all the discussions and jokes.

Many thanks to Yue Zheng, Yipeng Dong, and Yuzhang Lin, who have accompanied me since undergrad, and to Xiangyu Zhang, who always has my back since we started our graduate school application together. Special thanks to Prof. Junbo Zhang, who has truly been an inspiring friend, a mentor, and a role model to me.

Finally, my deepest gratitude goes to my grandparents, my parents, my brother, and my girlfriend Xiao Li, whose unconditional love and support have always been the source of strength for me to move forward.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

a.s.	almost surely
DER	Distributed Energy Resource
IO	Input-Output
LF	Loss Factor
LTC	Load Tap Changer
LTV	Linear Time-Varying
LSPI	Least Squares Policy Iteration
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDP	Markov Decision Process
ODCP	Optimal DER Coordination Problem
OPF	Optimal Power Flow
PF	Participation Factor
PGD	Projected Gradient Descent
RL	Reinforcement Learning
RWLS	Recursive Weighted Least-Squares
SD	Standard Deviation
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TD	Temporal Difference

LIST OF SYMBOLS

$\mathbb{E}$	Expectation
$\mathbb{P}$	Probability
$\mathbb{N}$	Set of natural numbers
$\mathbb{R}$	Set of real numbers
$\exp(\cdot)$	Exponential function
$\text{diag}(\cdot)$	Diagonal matrix
$\text{rank}(\cdot)$	Rank of a matrix
$\mathbf{0}$	All-zeros vector
$\mathbf{1}$	All-ones vector
$\|\| \cdot \|\|$	L_2-norm

CHAPTER 1

INTRODUCTION

In this chapter, we motivate the need to develop data-driven coordination frameworks for assets in power distribution systems to provide ancillary services. We then describe the power distribution model adopted throughout this dissertation, and review some existing works that are relevant to the problems addressed in the dissertation. Finally, we summarize the contents of each chapter.

1.1 Motivation and Background

Modern power distribution systems are experiencing fundamental transformations in structure and functionality facilitated by rapid technological developments [1]. In particular, an increasing number of distributed energy resources (DERs), e.g., rooftop solar PV installations, energy storage resources, and demand response resources, are being integrated in power distribution systems; this significantly increases the complexity of the system behavior. These transformations pose serious challenges to system operations but also bring numerous possibilities for innovations.

The uncertainty and variability in renewable-based generations may lead to imbalances in both active and reactive power, imposing additional requirements on the amount of regulation power to be provided by the bulk system [2], and causing voltage regulation problems in the power distribution system to which they are connected [3]. Consequently, this requires efficient active and reactive power controls in power distribution systems to provide ancillary services so as to mitigate the impacts of the aforementioned problems. However, active power control is typically limited in power distribution systems; meanwhile, reactive power control is provided by devices such as load tap changers (LTCs) and fixed/switched capacitors, which are not designed to

manage high variability in voltage fluctuations induced by renewable-based generation. As such, active and reactive power controls that can efficiently handle highly variable power changes are needed.

Fortunately, controllable DERs, which typically have fast-responding characteristics, could potentially be utilized to provide necessary ancillary services under appropriate coordination. Many research efforts addressed the development of DER coordination schemes for active and reactive power control. While these approaches could effectively regulate frequency or voltage given an accurate model of the power distribution system, they may fail in the absence of an accurate power distribution system model. As such, DER coordination schemes that are robust against model errors are yet to be investigated. The extensive deployment of advanced meters and communications nowadays makes it possible to access more detailed operating states of power distribution systems. Using system state measurements, DER coordination frameworks that do not explicitly rely on system models can be developed as an alternative approach for power control in power distribution networks; thus, the ultimate objective of this dissertation is

> **Objective:** To develop data-driven coordination frameworks for fully utilizing the capability of assets, particularly DERs, in power distribution systems to provide ancillary services. Specifically, the following two problems will be investigated:
>
> - **active power provision problem**, in which the total active power exchanged between the power distribution system and the bulk power system needs to equal to some amount requested by the bulk system operator; and
> - **reactive power regulation problem**, in which the voltage magnitude at each bus needs be maintained within some prespecified range or stay close to some reference value.

The problem of coordinating the response of a collection of DERs to provide ancillary services, which we refer to as the optimal DER coordination problem (ODCP), can be cast as a static optimization problem, the solution of which provides the optimal active and/or reactive power injection set-points for the DERs. Thus, we will develop algorithms that utilize measurements acquired from meters to estimate an input-output (IO) model

relevant to the ODCP formulation associated with the specific problem. Indeed, effective and efficient algorithms that can estimate an IO model are key in the data-driven frameworks. Therefore, special emphasis is devoted to algorithms that address the potential collinearity issue in the measurements, as well as formulations that significantly reduce the number of parameters to be estimated.

The data-driven coordination framework is also extended to address the problem of coordinating load tap changers (LTCs)—an important class of assets used for voltage control—using only measurements of voltage magnitudes. Different from the ODCP that is cast as a static optimization problem, the optimal LTC coordination problem is cast as a multi-stage decision-making problem and is formulated as a Markov decision process (MDP), in which the unknown power injections are modeled as uncertainty sources. Efficient reinforcement learning algorithms need to be developed to solve the MDP so as to obtain a control policy that maps the voltage magnitude measurements to the optimal tap positions.

1.2 Preliminaries

In this section, we introduce the power flow model for balanced power distribution systems adopted throughout this work. Specifically, a general power flow model and a power flow model for radial power distribution systems are reviewed.

1.2.1 Power Distribution System Model

Consider a balanced power distribution system represented by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ that consists of a set of buses indexed by the elements in the set $\mathcal{N} = \{0, 1, \cdots, N\}$, and a set of distribution lines indexed by the elements in the set $\mathcal{L} = \{1, 2, \cdots, L\}$. Each line $\ell \in \mathcal{L}$ is associated with a tuple $(i, j) \in \mathcal{N} \times \mathcal{N}$, where i is the sending end and j is the receiving end of a line, with the direction from i to j defined to be positive. Let $\mathcal{E} = \{(i, j)\} \subseteq \mathcal{N} \times \mathcal{N}$ denote the set of tuples that represent distribution lines. Let r_ℓ and x_ℓ denote the resistance and reactance of line $\ell \in \mathcal{L}$. Define $\boldsymbol{r} = [r_1, \cdots, r_L]^\top$ and $\boldsymbol{x} = [x_1, \cdots, x_L]^\top$. Let $\boldsymbol{Y} = \boldsymbol{G} + \mathrm{i}\boldsymbol{B}$ denote the

bus admittance matrix of the power distribution system, where $\mathrm{i} = \sqrt{-1}$, and $\boldsymbol{G}, \boldsymbol{B} \in \mathbb{R}^{(N+1)\times(N+1)}$. Let V_i and θ_i denote the magnitude and angle of the voltage at bus $i \in \mathcal{N}$, respectively. Define $\boldsymbol{V} = [V_1, \cdots, V_N]^\top$ and $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_N]^\top$. Assume bus 0 corresponds to a substation bus and $\theta_0 = 0$, which is the only connection of the power distribution system to the bulk power system. Furthermore, assume that bus 0 is an ideal voltage source that maintains a constant voltage magnitude, i.e., V_0 is a constant.

Without loss of generality, assume there is at most one DER and/or load at each bus, except bus 0, which does not have any DER or load. Let $\mathcal{N}^g = \{1, \cdots, n\}$ denote the DER index set. Throughout this work, DERs are assumed to be controllable. Uncontrollable DERs are modeled as negative loads. Let p_i^d and q_i^d respectively denote the active and reactive power loads at bus $i \in \mathcal{N}$, and define $\boldsymbol{p}^d = [p_1^d, \cdots, p_N^d]^\top$, and $\boldsymbol{q}^d = [q_1^d, \cdots, q_N]^\top$. Let p_i^g and q_i^g respectively denote the active and reactive power injections from DER $i \in \mathcal{N}^g$, and define $\boldsymbol{p}^g = [p_1^g, \cdots, p_n^g]^\top$, and $\boldsymbol{q}^g = [q_1^g, \cdots, q_n^g]^\top$. Let $\underline{p}_i^g$ and $\overline{p}_i^g$ respectively denote the minimum and maximum active power that can be provided by DER $i \in \mathcal{N}^g$, and define $\underline{\boldsymbol{p}}^g = [\underline{p}_1^g, \cdots, \underline{p}_n^g]^\top$, and $\overline{\boldsymbol{p}}^g = [\overline{p}_1^g, \cdots, \overline{p}_n^g]^\top$. Similarly, let $\underline{q}_i^g$ and $\overline{q}_i^g$ respectively denote the minimum and maximum reactive power that can be provided by DER $i \in \mathcal{N}^g$, and define $\underline{\boldsymbol{q}}^g = [\underline{q}_1^g, \cdots, \underline{q}_n^g]^\top$, and $\overline{\boldsymbol{q}}^g = [\overline{q}_1^g, \cdots, \overline{q}_n^g]^\top$.

Let p_i and q_i denote the net active and reactive power injections, respectively, at bus $i \in \mathcal{N}$, and define $\boldsymbol{p} = [p_1, \cdots, p_N]^\top$, $\boldsymbol{q} = [q_1, \cdots, q_N]^\top$; then

$$\boldsymbol{p} = \boldsymbol{C}\boldsymbol{p}^g - \boldsymbol{p}^d, \tag{1.1}$$

$$\boldsymbol{q} = \boldsymbol{C}\boldsymbol{q}^g - \boldsymbol{q}^d, \tag{1.2}$$

where $\boldsymbol{C} \in \mathbb{R}^{N\times n}$ is a matrix that maps the DER indices to the buses as follows: the entry at the i^{th} row, j^{th} column of $\boldsymbol{C}$ is 1 if DER j is at bus i and 0 otherwise.

General Power Flow Model

The quasi-steady-state behavior of the power distribution system at some operating point, $(\boldsymbol{\theta}, \boldsymbol{V}, \boldsymbol{p}, \boldsymbol{q})$, can be characterized by the power balance equa-

tions as follows:

$$\boldsymbol{g}^p(\boldsymbol{\theta}, \boldsymbol{V}) - \boldsymbol{p} = \mathbf{0}_N, \tag{1.3}$$

$$\boldsymbol{g}^q(\boldsymbol{\theta}, \boldsymbol{V}) - \boldsymbol{q} = \mathbf{0}_N, \tag{1.4}$$

where $\mathbf{0}_N$ denotes an N-dimensional all-zeros vector, and

$$\boldsymbol{g}^p(\boldsymbol{\theta}, \boldsymbol{V}) = [g_1^p(\boldsymbol{\theta}, \boldsymbol{V}), \cdots, g_N^p(\boldsymbol{\theta}, \boldsymbol{V})]^\top,$$
$$\boldsymbol{g}^q(\boldsymbol{\theta}, \boldsymbol{V}) = [g_1^q(\boldsymbol{\theta}, \boldsymbol{V}), \cdots, g_N^q(\boldsymbol{\theta}, \boldsymbol{V})]^\top,$$

with

$$g_i^p(\boldsymbol{\theta}, \boldsymbol{V}) = \sum_{j \in \mathcal{N}} V_i V_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)), \tag{1.5}$$

$$g_i^q(\boldsymbol{\theta}, \boldsymbol{V}) = \sum_{j \in \mathcal{N}} V_i V_j (G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j)), \tag{1.6}$$

where G_{ij} and B_{ij} denote the $(i,j)^{\text{th}}$ entries of the matrices $\boldsymbol{G}$ and $\boldsymbol{B}$, respectively. Note that given $\boldsymbol{p}$ and $\boldsymbol{q}$, $\boldsymbol{\theta}$ and $\boldsymbol{V}$ can be computed by solving (1.3) and (1.4). Once $\boldsymbol{\theta}$ and $\boldsymbol{V}$ are known, p_0 and q_0 can be determined as

$$p_0 = g_0^p(\boldsymbol{\theta}, \boldsymbol{V}) = \sum_{j \in \mathcal{N}} V_0 V_j (G_{0j} \cos(\theta_j) - B_{0j} \sin(\theta_j)), \tag{1.7}$$

$$q_0 = g_0^q(\boldsymbol{\theta}, \boldsymbol{V}) = \sum_{j \in \mathcal{N}} V_0 V_j (-G_{0j} \sin(\theta_j) + B_{0j} \cos(\theta_j)). \tag{1.8}$$

Note that the active and reactive power that flows into the bulk power system equals to $-p_0$ and $-q_0$, respectively.

Distribution Power Flow Model

We next introduce a power flow model for radial power distribution systems. Let f_ℓ^p and f_ℓ^q denote the active and reactive power that flows on line $\ell \in \mathcal{L}$, respectively, and define $\boldsymbol{f}^p = [f_1^p, \cdots, f_L^p]^\top$ and $\boldsymbol{f}^q = [f_1^q, \cdots, f_L^q]^\top$. The value of f_ℓ^p is positive when the active power flow is from the sending end to the receiving end of line ℓ, and negative otherwise. The value of f_ℓ^p is defined similarly. Let $\overline{f}_\ell$ denote maximum power flows on line $\ell \in \mathcal{L}$, and define $\overline{\boldsymbol{f}} = [\overline{f}_1, \cdots, \overline{f}_L]^\top$. With a little abuse of notation, we allow the substitution of the subscript ℓ with ij, if line ℓ is associated with (i,j), e.g.,

$f_{ij}^p = f_\ell^p$, $r_{ij} = r_\ell$, and $x_{ij} = x_\ell$.

Let v_i denote the squared voltage magnitude at bus $i \in \mathcal{N}$, i.e., $v_i = V_i^2$. For a radial power distribution system, the relation between squared voltage magnitudes, power injections, and line power flows, can be captured by the following distribution flow model [4]:

$$f_{ij}^p = -p_j + \sum_{k:(j,k)\in\mathcal{E}} f_{jk}^p + r_{ij}\frac{(f_{ij}^p)^2 + (f_{ij}^q)^2}{v_i}, \tag{1.9}$$

$$f_{ij}^q = -q_j + \sum_{k:(j,k)\in\mathcal{E}} f_{jk}^q + x_{ij}\frac{(f_{ij}^p)^2 + (f_{ij}^q)^2}{v_i}, \tag{1.10}$$

$$v_i - v_j = 2(r_{ij}f_{ij}^p + x_{ij}f_{ij}^q) - (r_{ij}^2 + x_{ij}^2)\frac{(f_{ij}^p)^2 + (f_{ij}^q)^2}{v_i}. \tag{1.11}$$

After dropping the nonlinear terms (nonlinear with respect to the squared voltage magnitude), (1.9)–(1.11) can be further simplified to a linearized distribution flow model, which is referred to as the LinDistFlow model, as follows:

$$f_{ij}^p = -p_j + \sum_{k:(j,k)\in\mathcal{E}} f_{jk}^p, \tag{1.12}$$

$$f_{ij}^q = -q_j + \sum_{k:(j,k)\in\mathcal{E}} f_{jk}^q, \tag{1.13}$$

$$v_i - v_j = 2(r_{ij}f_{ij}^p + x_{ij}f_{ij}^q). \tag{1.14}$$

Next, we write (1.12)–(1.14) in a matrix form. Let $\tilde{\boldsymbol{M}} = [M_{i\ell}] \in^{(N+1)\times L}$ denote the node-to-edge incidence matrix of the graph $\mathcal{G}$ defined as follows: $M_{i\ell} = 1$ and $M_{j\ell} = -1$ if line ℓ starts from bus i and ends at bus j, and all other entries equal to zero. Let $\boldsymbol{m}^\top$ denote the first row and $\boldsymbol{M}$ the remaining $(N \times L)$-dimensional matrix in $\tilde{\boldsymbol{M}}$. If the power distribution system is radial, $L = N$, and $\boldsymbol{M}$ is invertible. Then, (1.12) and (1.13) can be written as follows:

$$\boldsymbol{f}^p = \boldsymbol{M}^{-1}\boldsymbol{p}, \tag{1.15}$$

$$\boldsymbol{f}^q = \boldsymbol{M}^{-1}\boldsymbol{q}. \tag{1.16}$$

In practice, we are mostly concerned with active power flows on the distribution lines. To ease the notation, we drop the superscript p and let f_ℓ denote

the active power flow on line $\ell \in \mathcal{L}$, and define $\boldsymbol{f} = [f_1, \cdots, f_L]^\top$.

Define $\boldsymbol{v} = [v_1, \cdots, v_N]^\top$ and $\tilde{\boldsymbol{v}} = \boldsymbol{v} - v_0 \mathbf{1}_N$, where $\mathbf{1}_N$ is an N-dimensional all-ones vector. Then, (1.14) can be written in the following matrix from:

$$\tilde{\boldsymbol{v}} = \boldsymbol{R}\boldsymbol{p} + \boldsymbol{X}\boldsymbol{q}, \tag{1.17}$$

with

$$\boldsymbol{R} = 2(\boldsymbol{M}^{-1})^\top \mathrm{diag}(\boldsymbol{r})\boldsymbol{M}^{-1}, \tag{1.18}$$

$$\boldsymbol{X} = 2(\boldsymbol{M}^{-1})^\top \mathrm{diag}(\boldsymbol{x})\boldsymbol{M}^{-1}, \tag{1.19}$$

where $\mathrm{diag}(\cdot)$ returns a diagonal matrix with the entries of the argument on its diagonals. We refer to the matrices $\boldsymbol{R}$ and $\boldsymbol{X}$ collectively as voltage sensitivity matrices.

1.3 Related Work

Previous works that are relevant to the problems addressed in this dissertation fall into three categories, which are to be detailed in the following sections. Notwithstanding the abundance of work on model-based asset coordination and data-driven approaches, there is very little prior work on developing data-driven coordination frameworks that are adaptive to system condition changes and robust to model errors for ancillary services.

1.3.1 Asset Coordination for Ancillary Service Provision

Model-based DER coordination for provision of ancillary services, including frequency regulation [5–7], and voltage regulation [8–13], is an active area of research. In [5], the authors proposed a distributed coordination scheme for DERs to provide ancillary services in lossless systems. However, neglecting system losses will result in a solution that fails to provide the exact amount active power. In [6], the authors presented a framework to coordinate the DERs to provide the primary frequency control using a droop-control scheme, based on a power flow model with explicit consideration of losses. They also proposed in [7] a real-time coordination algorithm for a virtual power plant,

which is essentially a collection of DERs, to provide regulation leveraging online primal-dual-type methods. With respect to the voltage regulation problem, in [8], the authors proposed a two-stage distributed architecture for voltage regulation in a power distribution system, where the required reactive power injections are determined by each local controller in the first stage and any deficiency is compensated in the second stage. Optimal power flow (OPF) based DER coordination schemes for voltage regulation have also been proposed in works such as [10–12]. For example, in [12], the authors investigated the voltage regulation problem for an unbalanced power distribution system, and proposed distributed algorithms for optimally setting the tap ratios of LTCs and DERs based on a linearized power flow model.

In practice, LTCs are the most widely utilized assets in power distribution systems to regulate the voltage magnitudes along the feeder. Conventionally, the tap position of each LTC is controlled through an automatic voltage regulator based on local voltage measurements [14]. This approach, albeit simple and effective, is not optimal in any sense. Particularly, the voltage deviation may not be minimized. In the context of transmission systems, transformer tap positions are optimized jointly with active and reactive power generation by solving an OPF problem, which is typically cast as a mixed-integer programming problem (see, e.g., [15–17] and references therein). Similar OPF-based approaches are also adopted in power distribution systems. For example, in [18], the authors cast the optimal tap setting problem as a rank-constrained semidefinite program, which avoids the non-convexity and integer variables, and thus can be solved efficiently.

These model-based approaches require complete system knowledge, including active and reactive power injections, and transmission/distribution line parameters. While it may be reasonable to assume that such information is available for transmission systems, the situation in distribution systems is quite different. Accurate line parameters may not be known and power injections at each bus may not be available in real time, which prevents the application of model-based approaches.

1.3.2 Data-Driven Methods for Estimation and Control

The second line of works relevant to this dissertation focuses on data-driven approaches to estimation and control problems. As an alternative to the model-based approaches, data-driven methods have been demonstrated to be very effective in such situations where models are not readily available [19–26]. The fundamental idea behind data-driven approaches is to describe the system behavior by a linear time-varying (LTV) IO model, and estimate the parameters of this model—the so-called sensitivities—via regression using measurements of pertinent variables [19, 20].

Many previous works have applied data-driven approaches to power system problems, both in a steady-state setting [21–24], and a dynamical setting [25,26]. For example, in [21], the authors developed a data-driven framework to estimate linear sensitivity distribution factors such as injection shifting factors [22]; they further proposed a data-efficient sparse representation to estimate these sensitivities [22]. This framework was later tailored to the problem of estimating the power flow Jacobian [23]. In [24], the authors used the estimation framework proposed in [21] to solve the security constrained economic dispatch problem. The sensitivities have also been utilized in voltage regulation problems [27–29]. For example in [29], the authors proposed ambient signal based estimation methods for volt-to-var sensitivities in transmission systems. They further developed data-driven sequential voltage control methods based on estimated volt-to-var sensitivities and have proven the effectiveness via simulations using realistic data. Data-driven approaches have also been used to develop power systems stabilizers [25], and damping controls [26]. We refer interested readers to [30] for a review of data-driven approaches and their applications in a variety of other areas.

The sensitivities estimated from measurements enjoy several advantageous properties, including adaptivity to changes in system conditions such as topology reconfigurations or parameter changes. Yet, due to collinearity in the measurements [31], the regression problem may be ill-conditioned, thus resulting in large estimation errors [32]. Though numerical approaches such as locally weighted ridge regression [31] and noise-assisted ensemble regression [32] can be used to mitigate the impacts of collinearity, there is no theoretical guarantee on the estimation errors. In addition, existing approaches require a significant number of measurements in order to obtain

accurate estimates of the sensitivities. This may be feasible in transmission systems equipped with phasor measurement units (PMUs), but it may be impractical for power distribution systems.

1.3.3 Power System Multi-Stage Decision-Making under Uncertainty

The last area of research relevant to this dissertation is multi-stage decision making for power systems under uncertainty. Several methodologies can be applied to solve such problems, including stochastic programming [33–35], robust programming [36, 37], and reinforcement learning (RL) [38–41]. Uncertainty in the uncontrollable variables is characterized using scenarios in stochastic programming, uncertainty sets in robust programming, and Markov processes in RL. Yet, a complete model of the system is required in the stochastic or robust programming approaches. Models may not be necessary in RL algorithms; however, a large number of samples that potentially cover all possible control actions are needed to achieve satisfactory performance. In the absence of a model, such samples are costly or even impossible to obtain due to operational reliability concerns. As such, approaches that can effectively solve multi-stage decision-making problems for the purpose of coordinating assets in power distribution systems to provide ancillary services are still to be investigated.

1.4 Dissertation Organization

In this section, we give an overview of the organization of this dissertation.

Chapter 2: This chapter is concerned with the problem of DER coordination for active power provision, particularly in frequency regulation. We formulate the ODCP of interest as a static optimization problem, the objective of which is to minimize some cost function constrained by the power balance equation and DER capacity limits. This formulation requires an IO model where the inputs are power injections and the output is the incremental total system loss. The coefficients in the IO model, or the so-called loss factors (LFs), are estimated recursively by an estimator using power injection

measurements so as to adapt to various phenomena that impact the operation of the power system such as changes in the system operating conditions. Using the estimated IO model, the controller can determine the DER power injection set-points by solving the ODCP. In the proposed framework, the estimator and the controller work alternately on the same timescale; therefore, we refer to this framework as a one-timescale framework.

Chapter 3: In this chapter, we continue studying the same problem studied in Chapter 2 and extend the data-driven DER coordination framework to a more general setting. A major shortcoming of the framework proposed in Chapter 2 is that the performance of estimator may be poor in the presence of measurement collinearity. To resolve this issue and mitigate its impact on estimation accuracy, we introduce random perturbations in the DER power injections during the estimation process. The random perturbations are injected on a timescale that is faster than the one at which the controller updates the DER power injection set-points. Therefore, we refer to the framework proposed in this chapter as a two-timescale framework, so as to distinguish it from the one proposed in the Chapter 2.

Chapter 4: In this chapter, we adapt the data-driven DER coordination framework to address the problem of DER coordination for reactive power regulation. The goal here is to determine optimal DER power injections that minimize the voltage excursions outside a desirable voltage range, while relying only on an incomplete model of the power distribution system. To this end, we adopt the LinDistFlow model as the IO model where the inputs are power injections and the outputs are the deviations of squared voltage magnitudes. The coefficients in the IO model are voltage sensitivities that can be computed using information on the network topology and the line parameters. Assuming the knowledge of feasible network topology configurations and distribution line resistance-to-reactance (r-to-x) ratios, we propose an estimator that can determine the true network topology configuration and estimate the corresponding line parameters from a few measurements of voltage magnitudes and power injections. Using the estimated voltage sensitivities, the optimal DER power injections can be readily determined by solving a convex ODCP.

Chapter 5: In this chapter, we address the problem of setting the tap positions of LTCs for voltage regulation under uncertain load dynamics. The objective is to find a policy to determine the tap positions that only uses measurements of voltage magnitudes and topology information so as to minimize the voltage deviation across the system. We formulate this problem as an MDP, and propose a batch RL algorithm to solve it. By taking advantage of the LinDistFlow model, we propose an effective algorithm to estimate the voltage magnitudes under different tap settings, which allows the RL algorithm to explore the state and action spaces freely offline without impacting the system operation. To circumvent the "curse of dimensionality" resulting from the large state and action spaces, we propose a sequential learning algorithm to learn an action-value function for each LTC, based on which the optimal tap positions can be directly determined.

Chapter 6: In this chapter, we summarize the works presented in this dissertation, draw some conclusions, and provide some directions for future work.

CHAPTER 2

DER COORDINATION FOR ACTIVE POWER PROVISION: A ONE-TIMESCALE FRAMEWORK

2.1 Introduction

Frequency regulation services are by-and-large provided by conventional synchronous generators. However, with the deepening penetration of renewable-based generation resources, the frequency regulation requirements will be significantly increased [42]. Consequently, synchronous generators alone may be insufficient to meet the regulation requirements. Moreover, such conventional synchronous generators usually suffer from poor performance when providing regulation services [43]. To overcome the aforementioned challenges, DERs such as energy storage resources and plug-in vehicles, are also allowed to provide such services [44–46]. In the meantime in the US, driven by Order 755 from the Federal Energy Regulatory Commission [47], performance-based regulation markets have emerged as an effective means to incentivize the provision of high-quality frequency regulation services from resources including conventional generators and DERs [43, 48, 49]. In a performance-based regulation market, resources are incentivized to track the instructed regulation signal accurately since otherwise they will incur loss of performance payments [43]. Compared to conventional generators, DERs typically enjoy much faster-responding capabilities, and can potentially track the regulation signal better.

In this chapter, we focus on the problem of DER coordination for active power provision, specifically, the provision of frequency regulation services. The ODCP of interest here is to coordinate the response of the DERs in a power distribution system to collectively provide a requested amount of active power, the so-called regulation power, to the bulk power system to which the power distribution system is connected. We address the ODCP by taking a data-driven approach that explicitly takes into account system losses

yet without reliance on power system models. To this end, we formulate the ODCP as a static optimization problem that minimizes some cost function and is constrained by the power balance equation and DER capacity limits.

In order to provide the exact amount of regulation power, it is necessary to explicitly consider the impacts of total system active power loss (simply referred to as the total system loss), which is a nonlinear function of the voltage magnitudes and angles, as well as the network parameters, when determining the regulation power provided by each DER. As such, in the ODCP formulation, we adopt an IO model where the inputs are power injections and the output is the incremental total system loss. The coefficients in the IO model are LFs—linear sensitivities of the total system loss with respect to changes in power injections—that explicitly capture the impacts of both active and reactive power injections on system losses. The LFs are conventionally computed from power flow models and have been applied in locational marginal price computation in electricity markets [50, 51] and sizing and allocation of DERs in power distribution systems [52]. In the proposed framework, the LF will be estimated in an online fashion using measurements acquired from the system, and updated in real time so as to adapt to various phenomena that impact the operation of the power system.

An important characteristic of the proposed framework is that the timescale at which the estimator updates its sensitivity estimates and that at which the controller updates the DER active power injection set-points are the same; thus, we refer to this framework as a one-timescale framework.

The remainder of this chapter is organized as follows. The LFs and the ODCP of interest are introduced in Section 2.2. A recursive estimator for the LFs is proposed in Section 2.4. Then, in the same section, a data-driven coordination framework is developed based on the estimated LFs. We illustrate the application of the proposed framework in Section 2.5 and summarize this chapter in Section 2.6.

2.2 Preliminaries

In this section, we introduce the LFs that are used to approximate the total system loss in a power system. In addition, we describe the ODCP for the provision of frequency regulation services.

2.2.1 Loss Factors

Consider the balanced power distribution system introduced in Section 1.2.1. Let y denote the active power that flows from the power distribution system into the bulk power system. We adopt the convention that y is negative when the power distribution system is importing power and positive when it is exporting power. We can conceptually express y as a function of $\boldsymbol{p}$ and $\boldsymbol{q}$ as follows:

$$y = h(\boldsymbol{p}, \boldsymbol{q}). \tag{2.1}$$

For given $\boldsymbol{p}$ and $\boldsymbol{q}$, the total system loss, denoted by $l(\boldsymbol{p}, \boldsymbol{q})$, is given by

$$l(\boldsymbol{p}, \boldsymbol{q}) = \mathbf{1}_N^\top \boldsymbol{p} - h(\boldsymbol{p}, \boldsymbol{q}), \tag{2.2}$$

where $\mathbf{1}_N$ is an N-dimensional all-ones vector. The partial derivatives of the total system loss with respect to active and reactive power injections at each bus are referred to as the LFs. Let Λ_i^p and Λ_i^q respectively denote the partial derivatives of the total system loss with respect to the active and reactive net power injections at bus $i \in \mathcal{N}$; we refer to Λ_i^p as the active LF at bus i, and Λ_i^q as the reactive LF at bus i. Define $\boldsymbol{\Lambda}^p = [\Lambda_1^p, \cdots, \Lambda_N^p]^\top$ and $\boldsymbol{\Lambda}^q = [\Lambda_1^q, \cdots, \Lambda_N^q]^\top$. Then, it follows from (2.2) that

$$(\boldsymbol{\Lambda}^p)^\top = \left.\frac{\partial l}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p},\boldsymbol{q}} = \mathbf{1}_N^\top - \left.\frac{\partial h}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p},\boldsymbol{q}}, \tag{2.3}$$

$$(\boldsymbol{\Lambda}^q)^\top = \left.\frac{\partial l}{\partial \boldsymbol{q}}\right|_{\boldsymbol{p},\boldsymbol{q}} = -\left.\frac{\partial h}{\partial \boldsymbol{q}}\right|_{\boldsymbol{p},\boldsymbol{q}}, \tag{2.4}$$

where the partial derivatives of h with respect to $\boldsymbol{p}$ and $\boldsymbol{q}$ are row vectors.

Therefore, for given small $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{q}$ (with respect to some $\boldsymbol{p}$ and $\boldsymbol{q}$), the incremental total system loss associated with $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{q}$ can be approximated by

$$\Delta l(\boldsymbol{p}, \boldsymbol{q}) \approx (\boldsymbol{\Lambda}^p)^\top \Delta\boldsymbol{p} + (\boldsymbol{\Lambda}^q)^\top \Delta\boldsymbol{q}. \tag{2.5}$$

Then, by using (2.1), (2.3), and (2.4), we can obtain the change in y, denoted

by Δy, as follows

$$\Delta y \approx (\mathbf{1}_N - \mathbf{\Lambda}^p)^\top \Delta \boldsymbol{p} - (\mathbf{\Lambda}^q)^\top \Delta \boldsymbol{q}. \tag{2.6}$$

It is clear from (2.6) that the impacts from reactive power injections on the total system loss cannot be neglected. When active power injections change, the reactive power injections may change accordingly based on some specific rules (possibly as a result of a feedback control action). Both $\Delta \boldsymbol{p}$ and $\Delta \boldsymbol{q}$ will lead to changes in the total system loss. Thus, in order to determine the incremental total system loss after active power injections change, it is necessary to know the reactive power control policies implemented throughout the system. If $\Delta \boldsymbol{p}$ is small, we can approximately represent $\Delta \boldsymbol{q}$ as $\Delta \boldsymbol{q} \approx \mathbf{\Phi} \Delta \boldsymbol{p}$, where $\mathbf{\Phi} \in \mathbb{R}^{N \times N}$ is the reactive power response sensitivity matrix at the operating point defined by $(\boldsymbol{\theta}, \boldsymbol{V}, \boldsymbol{p}, \boldsymbol{q})$. Using this approximation, together with (2.5) and (2.6), $\Delta l(\boldsymbol{p})$ and Δy can be represented as a linear function of $\Delta \boldsymbol{p}$ as follows:

$$\Delta l(\boldsymbol{p}, \boldsymbol{q}) \approx \mathbf{\Lambda}^\top \Delta \boldsymbol{p}, \tag{2.7}$$

$$\Delta y \approx (\mathbf{1}_N - \mathbf{\Lambda})^\top \Delta \boldsymbol{p}, \tag{2.8}$$

where $\mathbf{\Lambda} = \mathbf{\Lambda}^p + \mathbf{\Phi}^\top \mathbf{\Lambda}^q$; we refer to the entries of $\mathbf{\Lambda}$ as the total LFs.

It is important to emphasize that the total LFs depend strongly on the reactive power control policies. In the simple case where no reactive power control is employed, i.e., $\mathbf{\Phi} = \mathbf{0}$, the total LFs are identical to the active LFs. However, in general, $\mathbf{\Phi}$ can be hardly derived from models.

2.2.2 Optimal DER Coordination Problem

Let $\boldsymbol{p}_i^{g0}$ denote the nominal power injection from DER i, and define $\boldsymbol{p}^{g0} = [p_1^{g0}, \cdots, p_n^{g0}]^\top$. Similarly, let $\boldsymbol{p}_i^{d0}$ denote the nominal power demanded by the load at bus i, and define $\boldsymbol{p}^{d0} = [p_1^{d0}, \cdots, p_N^{d0}]^\top$. In addition, let y^0 denote the nominal active power that flows from the power distribution system to the bulk power system. Let ρ denote the regulation signal sent by the bulk power system operator to the power distribution system and let $\rho[k]$ denote its value at time instant k. We adopt the convention that if ρ is positive, then the power distribution system must regulate upwards, i.e., the total

amount of power it imports must be decreased or the total amount of power it exports must be increased.

Usually, the operator sends a single value every 2 to 4 seconds. Therefore, the duration between two time instants is 2 to 4 seconds. Let $p_i^{g0}[k]$ and $p_i^g[k]$ denote the nominal power and actual power injected by DER i at time instant k. The nominal power injection may vary with time; hence its dependence of k. For example, $p_i^{g0}[k]$ may increase/decrease during a period so as to follow forecasted load changes. Then, we can define the regulation power provided by DER i in time interval k as $\rho_i^g[k] = p_i^g[k] - p_i^{g0}[k]$. In addition, define $\boldsymbol{\rho}^g = [\rho_1^g, \cdots, \rho_n^g]^\top$. Let $\overline{\rho}_i^g$ and $\underline{\rho}_i^g$ respectively denote the maximum up and down regulation capacities of DER i, and define $\overline{\boldsymbol{\rho}}^g = [\overline{\rho}_1^g, \cdots, \overline{\rho}_n^g]^\top$ and $\underline{\boldsymbol{\rho}}^g = [\underline{\rho}_1^g, \cdots, \underline{\rho}_n^g]^\top$. Then, at time instant k, the problem is to determine $\boldsymbol{\rho}^g[k]$ satisfying the following two constraints:

[C1.] $\underline{\boldsymbol{\rho}}^g \leq \boldsymbol{\rho}^g[k] \leq \overline{\boldsymbol{\rho}}^g$, i.e., DER up and down regulation capacity constraints are satisfied; and

[C2.] $y[k] = y^0[k] + \rho[k]$, i.e., the total amount of regulation power provided to the bulk power system is $\rho[k]$,

while all loads within the power distribution system are balanced, i.e., equation (2.8), which is used to approximately model the power balance in the system, is satisfied for some $\boldsymbol{q}$.[1] Note that $\rho[k]$ and all the measurements are obtained at time instant k, and the loads are assumed to be constant between two time instants. Also, the computation of $\boldsymbol{\rho}^g[k]$ is assumed to be instantaneous. The DERs will be instructed to provide $\boldsymbol{p}^g[k]$ immediately after the computation is completed, and their power outputs will remain fixed until the beginning of the next time instant.

In a lossless power distribution system, the sum of $\rho_i^g[k]$'s must equal to $\rho[k]$; however, this is not the case in the actual power distribution system that has losses. As such, it is important to explicitly take into account the system losses in order for the DERs to provide the exact amount of regulation power to the bulk power system. Also, while there may exist multiple ways to choose a $\boldsymbol{\rho}^g[k]$ that solves the DER coordination problem, we can select one that is optimal with respect to some objective function. For example,

[1] Some q_i's are purely determined by the reactive power consumed by the loads, but some may be determined as the result of reactive power control schemes as explained later.

we may select a $\boldsymbol{\rho}^g[k]$ that minimizes the incremental total system loss or costs, or one that minimizes the norm of the regulation power vector, or a combination of these various objectives.

2.3 Model-based Loss Factors

In this section, we first derive an expression for the LFs based on the nonlinear power flow model. Consider the same power distribution system model described in Section 2.2.1. Assume g_i^p and g_i^q, $\forall i \in \tilde{\mathcal{N}}$, are continuously differentiable. Changes in the net active and reactive power injections with some nominal values $\boldsymbol{p}$ and $\boldsymbol{q}$, denoted respectively by $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{q}$, will result in changes in $\Delta\boldsymbol{\theta}$ and $\Delta\boldsymbol{V}$. Assume $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{q}$ are sufficiently small, then the following relation holds:

$$\begin{bmatrix} \boldsymbol{J}_{p\theta} & \boldsymbol{J}_{q\theta} \\ \boldsymbol{J}_{pV} & \boldsymbol{J}_{qV} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{\theta} \\ \Delta\boldsymbol{V} \end{bmatrix} \approx \begin{bmatrix} \Delta\boldsymbol{p} \\ \Delta\boldsymbol{q} \end{bmatrix}, \tag{2.9}$$

where

$$\boldsymbol{J}_{p\theta} = \left.\frac{\partial \boldsymbol{g}^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \in \mathbb{R}^{N\times N}, \quad \boldsymbol{J}_{pV} = \left.\frac{\partial \boldsymbol{g}^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \in \mathbb{R}^{N\times N},$$

$$\boldsymbol{J}_{q\theta} = \left.\frac{\partial \boldsymbol{g}^q}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \in \mathbb{R}^{N\times N}, \quad \boldsymbol{J}_{qV} = \left.\frac{\partial \boldsymbol{g}^q}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \in \mathbb{R}^{N\times N}.$$

Assuming $\begin{bmatrix} \boldsymbol{J}_{p\theta} & \boldsymbol{J}_{q\theta} \\ \boldsymbol{J}_{pV} & \boldsymbol{J}_{qV} \end{bmatrix}$ is invertible around the operating point $(\boldsymbol{\theta}, \boldsymbol{V}, \boldsymbol{p}, \boldsymbol{q})$, we further define

$$\begin{bmatrix} \boldsymbol{\Gamma}_{\theta p} & \boldsymbol{\Gamma}_{\theta q} \\ \boldsymbol{\Gamma}_{Vp} & \boldsymbol{\Gamma}_{Vq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}_{p\theta} & \boldsymbol{J}_{q\theta} \\ \boldsymbol{J}_{pV} & \boldsymbol{J}_{qV} \end{bmatrix}^{-1}, \tag{2.10}$$

where $\boldsymbol{\Gamma}_{\theta p}, \boldsymbol{\Gamma}_{\theta q}, \boldsymbol{\Gamma}_{Vp}, \boldsymbol{\Gamma}_{Vq} \in \mathbb{R}^{N\times N}$. It follows from (2.9) that

$$\begin{bmatrix} \Delta\boldsymbol{\theta} \\ \Delta\boldsymbol{V} \end{bmatrix} \approx \begin{bmatrix} \boldsymbol{\Gamma}_{\theta p} & \boldsymbol{\Gamma}_{\theta q} \\ \boldsymbol{\Gamma}_{Vp} & \boldsymbol{\Gamma}_{Vq} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{p} \\ \Delta\boldsymbol{q} \end{bmatrix}. \tag{2.11}$$

Changes in the net active power injections at bus 1, which equals to $-\Delta y$,

can be approximated by

$$-\Delta y \approx \left.\frac{\partial g_0^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \Delta\boldsymbol{\theta} + \left.\frac{\partial g_0^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \Delta\boldsymbol{V}. \tag{2.12}$$

Plugging (2.11) into (2.12) leads to

$$\begin{aligned} -\Delta y \approx & \left(\left.\frac{\partial g_0^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{\theta p} + \left.\frac{\partial g_0^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{Vp} \right) \Delta\boldsymbol{p} \\ & + \left(\left.\frac{\partial g_0^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{\theta q} + \left.\frac{\partial g_0^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{Vq} \right) \Delta\boldsymbol{q}. \end{aligned} \tag{2.13}$$

As such, it follows from (2.13) that

$$-\left(\left.\frac{\partial h}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p},\boldsymbol{q}} \right)^\top \approx \left.\frac{\partial g_0^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{\theta p} + \left.\frac{\partial g_0^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{Vp}, \tag{2.14}$$

$$-\left(\left.\frac{\partial h}{\partial \boldsymbol{q}}\right|_{\boldsymbol{p},\boldsymbol{q}} \right)^\top \approx \left.\frac{\partial g_0^p}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{\theta q} + \left.\frac{\partial g_0^p}{\partial \boldsymbol{V}}\right|_{\boldsymbol{\theta},\boldsymbol{V}} \boldsymbol{\Gamma}_{Vq}. \tag{2.15}$$

Substituting (2.14) and (2.15) into (2.3) and (2.4) gives the model-based active and reactive LFs.

2.4 Coordination Framework

In this section, we propose a data-driven framework for coordinating the DERs so as to provide frequency regulation services. We first propose a recursive algorithm to estimate the LFs. The LFs obtained via this estimation method capture the most up-to-date operating conditions. We will then use the estimated LFs to develop a formulation of the ODCP. Thus, when choosing their set-points, as determined by the solution of the ODCP problem, the DERs are adapting to changes in operating conditions.

2.4.1 Recursive Loss Factor Estimator

The computation of the LFs using the model-based approach requires the knowledge of system operating conditions, as well as accurate models of any

reactive power control schemes implemented; however, in practice, it may be difficult to obtain all this information. As an alternative approach, we propose a data-driven approach for estimating LFs, which is built upon the recursive weighted least-squares estimation (RWLS) method.

Suppose at time instant k, we have k consecutive measurements of y and $\boldsymbol{p}$, denoted by $y[0], \cdots, y[k-1]$, $\boldsymbol{p}[0], \cdots, \boldsymbol{p}[k-1]$, respectively. Define the change in the power injection vector at time instant t as $\Delta\boldsymbol{p}[t] = \boldsymbol{p}[t] - \boldsymbol{p}[t-1]$, where

$$\boldsymbol{p}[t] = \boldsymbol{C}(\boldsymbol{p}^{g0}[t] + \boldsymbol{\rho}^{g}[t]) - \boldsymbol{p}^{d}[t], \tag{2.16}$$

$$\boldsymbol{p}[t-1] = \boldsymbol{C}(\boldsymbol{p}^{g0}[t-1] + \boldsymbol{\rho}^{g}[t-1]) - \boldsymbol{p}^{d}[t-1], \tag{2.17}$$

with $\boldsymbol{C}$ as in (1.1). Let $\Delta y[t]$ denote the change in the active power injection into the network from the bulk power system at time instant t that results from $\Delta\boldsymbol{p}[t]$, i.e., $\Delta y[t] = y[t] - y[t-1]$. Then, for sufficiently small $\Delta\boldsymbol{p}[t]$, and assuming $\boldsymbol{\Lambda}$ remains relatively constant over time, it follows from (2.8) that

$$\Delta\boldsymbol{p}[t]^{\top}(\mathbf{1}_N - \boldsymbol{\Lambda}) \approx \Delta y[t]. \tag{2.18}$$

Then, by stacking the equation in (2.18) for $t = 0, 1, \cdots, k-1$, we obtain the following system of equations:

$$\begin{bmatrix} \Delta\boldsymbol{p}[1]^{\top} \\ \vdots \\ \Delta\boldsymbol{p}[k-1]^{\top} \end{bmatrix} (\mathbf{1}_N - \boldsymbol{\Lambda}) \approx \begin{bmatrix} \Delta y[1] \\ \vdots \\ \Delta y[k-1] \end{bmatrix}. \tag{2.19}$$

When k is sufficiently large, (2.19) becomes overdetermined, and thus, we can use weighted least-squares to obtain an estimate of $\boldsymbol{\Lambda}$ at instant k, denoted by $\hat{\boldsymbol{\Lambda}}[k]$, as follows:

$$\hat{\boldsymbol{\Lambda}}[k] = \mathbf{1}_N - \boldsymbol{F}[k]\boldsymbol{A}[k]^{\top}\boldsymbol{W}[k]\boldsymbol{b}[k], \tag{2.20}$$

where $\boldsymbol{W}[k] \in \mathbb{R}^{(k-1)\times(k-1)}$ is a positive definite symmetric weight matrix,

and

$$\boldsymbol{A}[k] = \begin{bmatrix} \Delta \boldsymbol{p}[1]^\top \\ \vdots \\ \Delta \boldsymbol{p}[k-1]^\top \end{bmatrix}, \tag{2.21}$$

$$\boldsymbol{b}[k] = \begin{bmatrix} \Delta y[1] \\ \vdots \\ \Delta y[k-1] \end{bmatrix}, \tag{2.22}$$

$$\boldsymbol{F}[k] = \left(\boldsymbol{A}[k]^\top \boldsymbol{W}[k] \boldsymbol{A}[k]\right)^{-1}. \tag{2.23}$$

When a new set of measurements, $y[k]$ and $\boldsymbol{p}[k]$, becomes available at k, we can update the LF estimate from $\hat{\boldsymbol{\Lambda}}[k]$ to $\hat{\boldsymbol{\Lambda}}[k+1]$ as follows [53]:

$$\hat{\boldsymbol{\Lambda}}[k+1] = \hat{\boldsymbol{\Lambda}}[k] - \boldsymbol{F}[k+1]\Delta \boldsymbol{p}[k](\Delta y[k] - \Delta \boldsymbol{p}[k]^\top (\boldsymbol{1}_N - \hat{\boldsymbol{\Lambda}}[k])), \tag{2.24}$$

where $\gamma \in (0, 1]$ is the forgetting factor for the measurements, and

$$\boldsymbol{F}[k+1] = \frac{1}{\gamma}\left(\boldsymbol{F}[k] - \frac{1}{\gamma + \Delta \boldsymbol{p}[k]^\top \boldsymbol{F}[k] \Delta \boldsymbol{p}[k]} \boldsymbol{F}[k] \Delta \boldsymbol{p}[k] \Delta \boldsymbol{p}[k]^\top \boldsymbol{F}[k]\right). \tag{2.25}$$

Thus, using (2.24), the LFs can be updated using new measurements with little computational effort.

2.4.2 LF-based ODCP Formulation

Using the LFs estimated using the procedure above, we can now develop a formulation for the ODCP. At time instant k, the estimated total LFs, $\hat{\boldsymbol{\Lambda}}[k]$, is available. Then, for sufficiently small $\Delta \boldsymbol{p}[k]$, it follows from (2.18) that

$$\Delta y[k] \approx (\boldsymbol{1}_N - \hat{\boldsymbol{\Lambda}}[k])^\top \Delta \boldsymbol{p}[k]. \tag{2.26}$$

Now, given $\rho[k]$, we would like to choose $\boldsymbol{\rho}^g[k]$ in some optimal fashion so that $y[k]$ tracks the regulation signal, i.e., $y[k] = y^0[k] + \rho[k]$. Thus, by using

the incremental model in (2.26), we have that

$$\begin{aligned}\Delta y[k] &= y[k] - y[k-1] \\ &= (y^0[k] + \rho[k]) - y[k-1] \\ &= (\mathbf{1}_N - \hat{\mathbf{\Lambda}}[k])^\top \Delta \boldsymbol{p}[k], \end{aligned} \tag{2.27}$$

with

$$\begin{aligned}\Delta \boldsymbol{p}[k] &= \boldsymbol{p}[k] - \boldsymbol{p}[k-1] \\ &= \boldsymbol{C}(\boldsymbol{p}^{g0}[k] + \boldsymbol{\rho}^g[k]) - \boldsymbol{p}^d[k] - \boldsymbol{C}(\boldsymbol{p}^{g0}[k-1] + \boldsymbol{\rho}^g[k-1]) + \boldsymbol{p}^d[k-1]. \end{aligned} \tag{2.28}$$

Then, by choosing the minimization of a weighted sum of the incremental losses and the L_2-norm (denoted by $\|\cdot\|$) of the regulation power vector as the optimality criterion, the ODCP to be solved at time instant k can be formulated as follows:

$$\boldsymbol{\rho}^g[k] = \underset{\boldsymbol{z} \in [\underline{\boldsymbol{\rho}}^g, \overline{\boldsymbol{p}}^g]}{\arg\min} \ \hat{\mathbf{\Lambda}}[k]^\top \Delta \boldsymbol{p}[k] + \frac{\eta}{2} \|\boldsymbol{z}\|^2 \tag{2.29}$$

subject to

$$y^0[k] + \rho[k] - y[k-1] = (\mathbf{1}_N - \hat{\mathbf{\Lambda}}[k])^\top \Delta \boldsymbol{p}[k], \tag{2.30}$$

$$\boldsymbol{C}(\boldsymbol{p}^{g0}[k-1] + \boldsymbol{\rho}^g[k-1]) - \boldsymbol{p}^d[k-1] = \boldsymbol{C}(\boldsymbol{p}^{g0}[k] + \boldsymbol{z}) - \boldsymbol{p}^d[k] - \Delta \boldsymbol{p}[k], \tag{2.31}$$

where $\eta \geq 0$, $\boldsymbol{z} = [z_1, \cdots, z_n]^\top$, with z_i being the regulation power provided by DER i, which is to be determined. Note that at time instant k, $y[k-1]$, $\boldsymbol{p}^d[k-1]$, $\boldsymbol{p}^{g0}[k-1]$, $\boldsymbol{\rho}^g[k-1]$, $y^0[k]$, $\boldsymbol{p}^d[k]$, $\boldsymbol{p}^{g0}[k]$, and $\rho[k]$ are known. The ODCP is a quadratic program with one equality constraint, which can be solved efficiently by the lambda iteration method [54].

Alternatively, we can also formulate the ODCP using $\mathbf{\Lambda}^p$, $\mathbf{\Lambda}^q$, $\mathbf{\Phi}$, which can be estimated in a similar manner as that described in Section 2.4.1. Yet, the estimation process would require at least twice the number of measurements (both active and reactive power injections), and more computational effort. As such, the total LF-based formulation is more desirable.

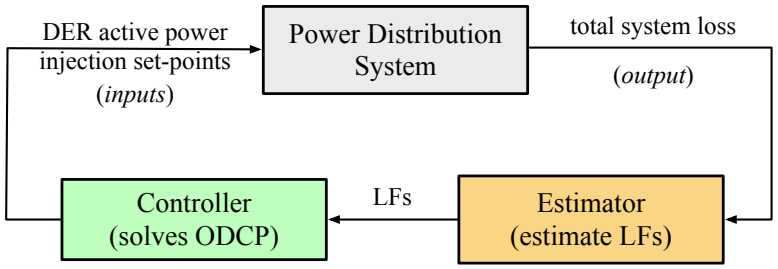

Figure 2.1: The closed-loop DER coordination framework.

2.4.3 Interaction between LF Estimator and ODCP Solver

The closed-loop interaction between the LF estimator, the controller that solves the ODCP, and the power distribution system, is illustrated via the block diagram in Fig. 2.1. At time instant k, the LF estimator provides $\hat{\mathbf{\Lambda}}[k]$, i.e., the estimate of the total LFs based on some measurements taken up to time instant $k-1$. Then, the vector of estimated LFs, $\hat{\mathbf{\Lambda}}[k]$, is sent to the controller. The controller also receives loads $\boldsymbol{p}^d[k]$, the requested regulation power $\rho[k]$, and solves the ODCP to determine $\boldsymbol{\rho}^g[k]$. Then, DERs are instructed to change their set-points so as to provide $\boldsymbol{\rho}^g[k]$. A new set of measurements will be available after the DER set-points are modified. The new measurements will be used in the LF estimator to dynamically update the estimated values of the total LFs and obtain $\hat{\mathbf{\Lambda}}[k+1]$, which will be used in ODCP to compute the optimal regulation power from DERs for the next time interval.

Note that in the proposed framework, the estimator and the controller work on the same timescale, as illustrated in Fig. 2.2. In this one-timescale framework, the estimator updates its sensitivity estimates at each time instant, after which the controller updates the DER active power injection set-points at the same time instant.

2.5 Numerical Simulation

In this section, we illustrate the application of the proposed framework through numerical simulations. The differences between active, reactive, and total LFs are also illustrated. The accuracy of the estimated LFs, as well as the performance of the framework, is studied for a case where the nominal loads are kept constant, and a case where there are load changes.

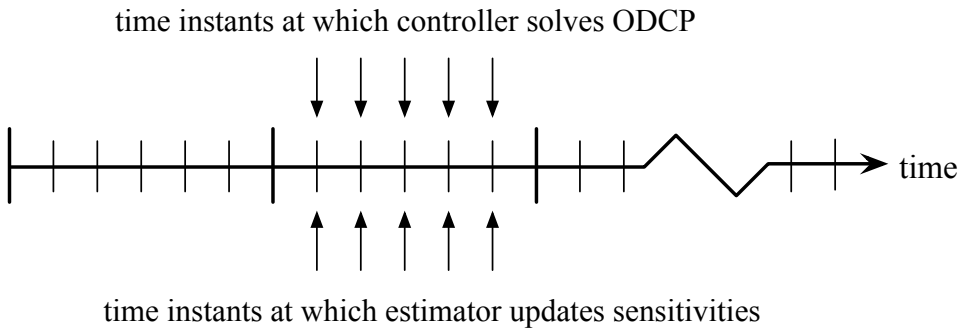

Figure 2.2: Timeline of the DER coordination framework.

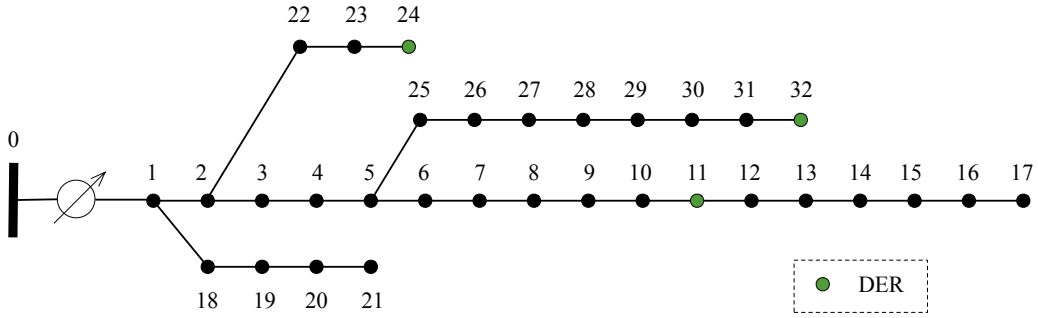

Figure 2.3: The IEEE 33-bus distribution test feeder.

2.5.1 Simulation Setup

A modified 33-bus distribution test feeder from [4], the topology of which is shown in Fig. 2.3, is used for all numerical simulations. The total nominal active load is 3720 kW, and the total reactive load is 2300 kVar. The nominal loads are shown in Fig. 2.4. There are three DERs, DER 1, DER 2, and DER 3, located at buses 11, 24, and 32, respectively, with their respective capacities being 2300 kW, 1500 kW, and 1200 kW. We assume the regulation capacity of each DER equals to 10% of its total capacity. The nominal active powers of DER 1, DER 2, and DER 3 are 2000 kW, 1000 kW, and 800 kW, respectively, while their respective nominal reactive powers are 0 kVar, 0 kVar, and 162.6 kVar. We assume the reactive power control at buses 0 and 11 aims to maintain a constant voltage magnitude of 1 p.u. while no reactive power control is employed at any other buses. Throughout the simulation, we set $\eta = 1$ and $\gamma = 0.97$.

In the simulation, the power demanded by load i is generated using $P_i^d[k] = P_i^{d0}(1 + \nu_i)$, where P_i^{d0} denotes the demand nominal value, and ν_i is a zero-mean Gaussian random variable with standard deviation being $\sigma_i = 0.01$

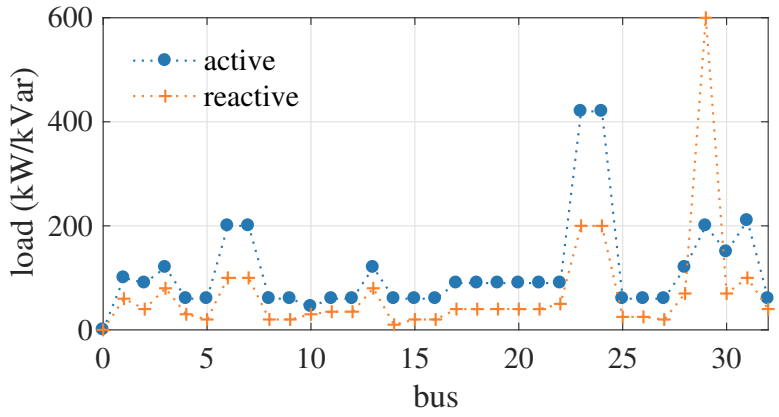

Figure 2.4: The nominal loads in the IEEE 33-bus distribution test feeder.

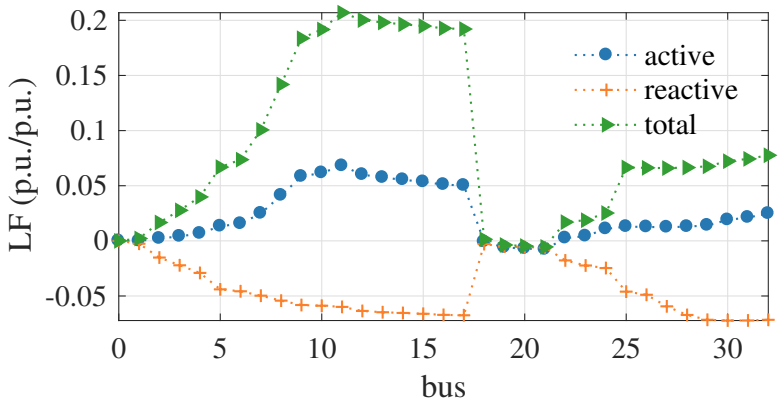

Figure 2.5: Actual LFs at nominal loads.

[23]. The term $P_i^{d0}\nu_i$ models random load fluctuations, which are assumed to be independent of each other. The regulation signal is taken from PJM [55]. This signal is updated every two seconds; correspondingly, we update the DER set-pints every two seconds.

The actual LFs are obtained through the following procedures. First, by solving the ac power flow using Matpower [56], we compute the change in the total system loss caused by a sufficiently small amount of change in the net injection at each bus. Then, the actual LFs can be obtained as the quotient of the change in the total system loss that results from the aforementioned computation and the change in the net injection at each bus. The actual LFs at the nominal operating point are shown in Fig. 2.5. As can be seen from Fig. 2.5, the active LFs, reactive LFs, and the total LFs have distinct values. Therefore, in the absence of the knowledge of $\mathbf{\Phi}$, we can hardly obtain the accurate total LFs from the model.

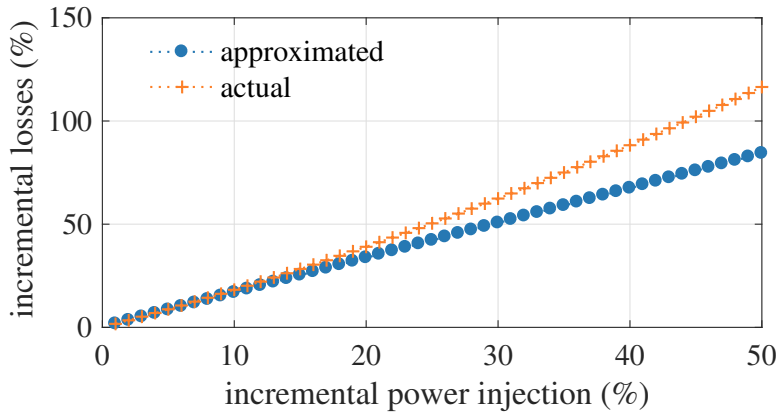

Figure 2.6: Error of LF approximation.

2.5.2 LF Approximation Error

The LFs can be used to linearly approximate the nonlinear loss function; as such, computing system losses using LFs may incur errors. We compare the incremental total system loss approximated using the total LFs with their actual values obtained by directly solving the power flow equations. The results are presented in Fig. 2.6, where the incremental losses are normalized using the total system loss at the nominal operating point, and the incremental power injection is normalized using the nominal total loads. The results in Fig. 2.6 indicate that the LFs are effective in computing the approximate losses when the incremental power injections are relatively small, which is suitable for the application considered in this chapter, i.e., coordinating DERs to provide frequency regulation services.

2.5.3 LF Estimation Accuracy

Next, we show the accuracy of the total LFs estimated using the data-driven approach as measured by the mean absolute error (MAE). Given the actual total LFs, $\mathbf{\Lambda}[k]$, and the estimated total LFs, $\hat{\mathbf{\Lambda}}[k]$, the MAE of the LF estimation, denoted by $\varepsilon[k]$, is calculated as follows:

$$\varepsilon[k] = \frac{\mathbf{1}_N^\top |\hat{\mathbf{\Lambda}}[k] - \mathbf{\Lambda}[k]|}{N}. \tag{2.32}$$

Assume we have an initial estimation of $\hat{\mathbf{\Lambda}}$ and $\boldsymbol{F}$ at 0 s, which corresponds to time instant 0, obtained by solving (2.19). The LFs estimated using the

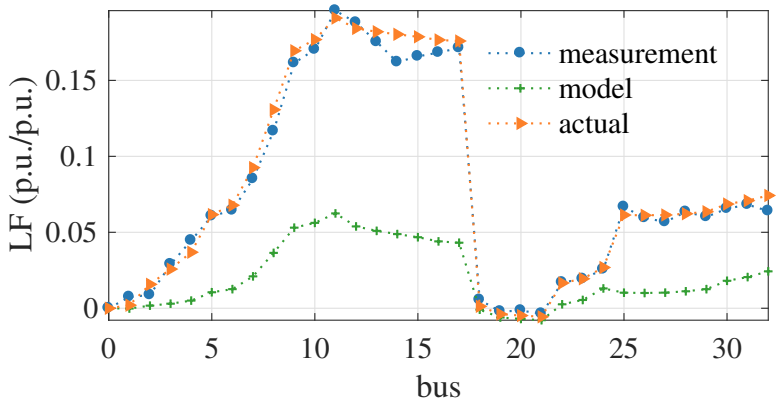

Figure 2.7: LF estimation results at 0 s.

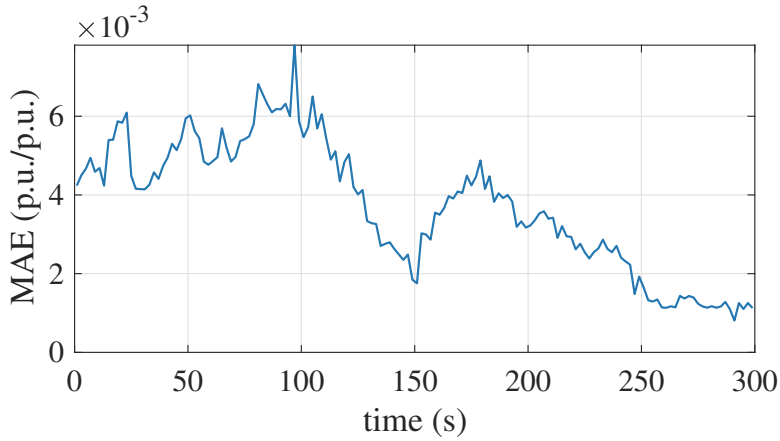

Figure 2.8: LF estimation errors.

data-driven approach (indicated by "measurement") and the model-based approach (indicated by "model") are shown in Fig. 2.7, together with the actual LFs (indicated by "actual"). Note that in the model-based approach, the active LFs are used as estimates for the total LFs. At time instant 0, $\varepsilon[0] = 0.0051$ in the data-driven approach, which is relatively small. Yet, we have $\varepsilon[0] = 0.0593$ for the model-based approach, which is one order of magnitude greater than the values obtained using the data-driven approach. This is expected since the impacts of reactive LFs on the total LFs are ignored in the model-based approach due to the lack of knowledge of $\mathbf{\Phi}$. As such, in later simulations, results from the model-based approach are no longer presented. After the first estimation, the RWLS algorithm is used to dynamically update $\hat{\mathbf{\Lambda}}$ and $\boldsymbol{F}$. The LF estimation errors between 0 s and 300 s are presented in Fig. 2.8. The average estimation error in the data-driven approach is 0.0037.

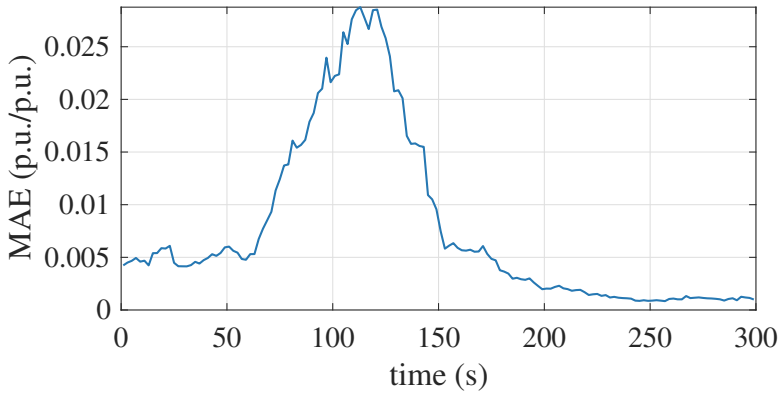

Figure 2.9: LF estimation errors under operating point change.

To illustrate the adaptivity of the data-driven approach to operating point changes, we simulate a case where the nominal loads increase linearly by 20% from 60 s to 120 s. In the meantime, the nominal active power injections from DERs also increase commensurably. Under this setup, the LF estimation errors during the time interval $(0, 300]$ s are presented in Fig. 2.9. The average estimation error in the data-driven approach is 0.0077, while the maximum estimation error is 0.0288. The estimation errors start to increase when the operating point begins to change at 60 s and starts to decrease when the changes end at 120 s. This is intuitively reasonable since the during the time when the operating point changes, the majority of existing measurements provide information on old operating points, based on which we can hardly get accurate estimates at the current operating point. After the operating point change ends, the newly obtained measurements provide more information on the new operating point and the impacts from the measurements on the old one decay due to the weights we assigned.

2.5.4 Quantification of Frequency Regulation Performance

The proposed framework is compared with the participation factor (PF) based coordination approach, where the PFs are nonnegative real numbers that are proportional to the regulation capacity of each DER and sum up to 1. The total incremental changes in the active loads and the requested regulation power are allocated to each DER based on the PFs. The effectiveness of the proposed framework for frequency regulation is measured by a performance score, the value of which at time instant k is computed as

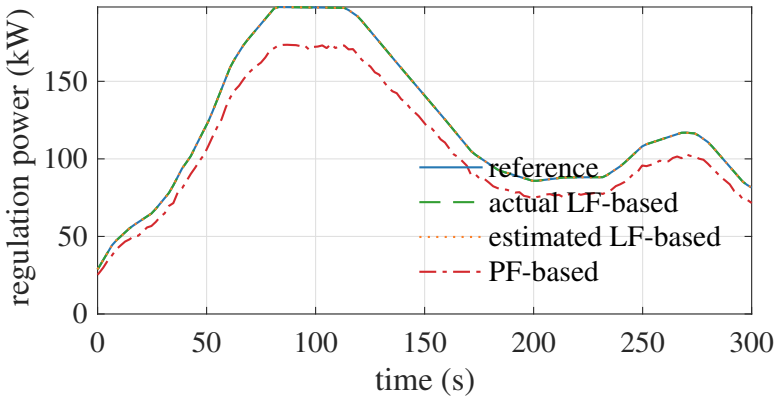

Figure 2.10: Trajectory of regulation power.

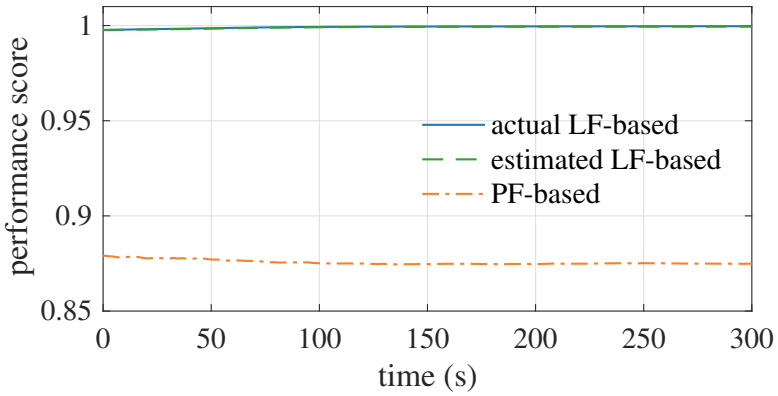

Figure 2.11: Performance scores of DERs providing frequency regulation services.

follows:

$$1 - \frac{\sum_{t=0}^{k} |y[t] - y^0[t] - \rho[t]|}{\sum_{t=0}^{k} |\rho[t]|}. \tag{2.33}$$

The frequency regulation performances using the PF-based approach, the LF-based approach with estimated total LFs, as well as the LF-based approach with actual total LFs, are compared in Figs. 2.10 and 2.11. The average performance score obtained using the PF-based approach is 0.8756; that obtained using the LF-based approach with estimated total LFs is 0.9991; and that obtained using the LF-based approach with actual total LFs is 0.9992. Obviously, the LF-based approaches perform much better than the PF-based approach where system losses are ignored. Moreover, the estimated-LF based approach performs almost equally as well as the actual-LF based approach.

Next, we evaluate the performance scores under the same operating point

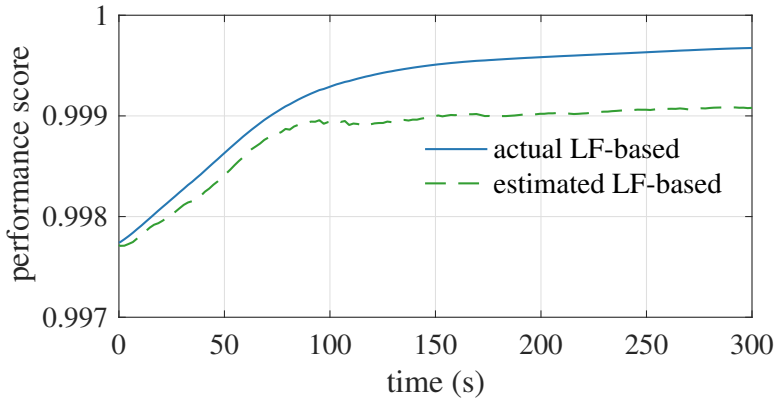

Figure 2.12: Performance scores of DERs under operating point change.

changes described in Section 2.5.3. Figure 2.12 shows the frequency regulation performances obtained using the LF-based approach with actual LFs and estimated LFs. As is shown in Fig. 2.12, the regulation power still tracks regulation signal accurately under operating point change. The average performance score using LF-based approach with estimated total LFs is 0.9988, and that obtained using the LF-based approach with actual total LFs is 0.9992. The slight decrease in the performance score obtained using the estimated LFs is caused by the decrease in the estimation accuracy of the total LFs. As explained in Section 2.5.3, the operating point changes have impacts on the estimation accuracy of the total LFs, which directly impacts the determination of regulation power provided by each DER. Yet, the performance score decrease is negligible, showing the adaptivity of the propose framework.

2.6 Summary

In this chapter, we have proposed a data-driven coordination framework for DERs connected to a lossy power distribution system to collectively provide frequency regulation services to a bulk power system. The framework consists of a recursive LF estimator and a controller that solves the ODCP formulated using the estimated LFs. The estimator updates its sensitivity estimates at the same timescale at which the controller updates the DER active power injection set-points. The proposed LF estimator is capable of obtaining an accurate estimate of the LFs that can capture the impacts of both active and reactive power injections on system losses with little computational effort. The inherent nature of the estimator makes it adaptive to system condition changes. Numerical simulation demonstrated that the proposed framework is more effective in coordinating the DERs to provide frequency regulation services, compared to the approach in which system losses are neglected.

The work presented in this chapter was published in [57].

CHAPTER 3

DER COORDINATION FOR ACTIVE POWER PROVISION: A TWO-TIMESCALE FRAMEWORK

3.1 Introduction

In this chapter, we continue studying the problem of coordinating the response of a set of DERs in a power distribution system so that they collectively provide some amount of active power to the bulk power system. A major shortcoming of the framework proposed in Chapter 2 is that the performance of the estimator may be poor in the presence of measurement collinearity. In this chapter, we propose an improved data-driven DER coordination framework that can resolve the measurement collinearity issue and mitigate its impact on estimation accuracy.

The proposed framework consists of three components, namely (i) a model of the system describing the relation between the variables of interest to the problem, i.e., DER active power injections and power exchanged between the distribution and bulk power systems, (ii) an estimator, which provides estimates of the parameters that populate the model in (i), and (iii) a controller that uses the model in (i) with the parameters estimated via (ii) to determine the active power injection set-points of the DERs by solving the ODCP. Specifically, an LTV IO model is adopted as the system model to capture the relation between the DER active power injections (inputs) and the total active power exchange (output). The parameters in this model are estimated by the estimator via the solution of a box-constrained quadratic program, obtained by using the projected gradient descent (PGD) algorithm.

The focus in this chapter is on the development of an estimation algorithm that can efficiently and effectively estimate the parameters in the IO model. Inspired by ideas in power system identification [58, 59], we introduce random perturbations in the DER active power injections during the estimation process to resolve the potential collinearity issue in the measurements used

by the estimator. The random perturbations are injected on a timescale that is faster than the one at which the controller updates the DER active power injection set-points. Therefore, we refer to the data-driven framework proposed in this chapter as a two-timescale framework, so as to distinguish the one proposed in Chapter 2.

We show that the estimation algorithm converges almost surely (a.s.) under some mild conditions, i.e., the estimated parameters converge to the true parameters a.s., and the total provided active power also converges to the required amount, as long as the DERs have collectively enough incremental capacity. Using the estimated IO model, the ODCP to be solved by the controller can be formulated as a convex optimization problem, which can be solved easily.

The remainder of this chapter is organized as follows. The IO system model and the ODCP of interest are described in Section 3.2. The components of the data-driven DER coordination framework are presented in Section 3.3. A description of the algorithm used in the framework, as well as its convergence analysis, is provided in Section 3.4. The proposed framework is demonstrated via numerical simulations on a IEEE 123-bus distribution test feeder in Section 3.5. We summarize this chapter in Section 3.6.

3.2 Preliminaries

In this section, we introduce the IO system model adopted in this chapter and then review the ODCP for the active power provision problem.

3.2.1 Input-Output System Model

Let y denote the active power exchanged between the distribution and bulk power systems via bus 0, defined to be positive if the flow is from the substation to the bulk power system. Conceptually, y can be represented as a function of $\boldsymbol{p}^g$, $\boldsymbol{q}^g$, $\boldsymbol{p}^d$, $\boldsymbol{q}^d$. Note also $\boldsymbol{q}^g$ is typically set according to some specific reactive power control rules to achieve certain objectives such as constant voltage magnitude or constant power factor, and thus is a function of

$\boldsymbol{p}^g, \boldsymbol{p}^d, \boldsymbol{q}^d$. Then, y can be written as a function of $\boldsymbol{p}^g, \boldsymbol{p}^d, \boldsymbol{q}^d$ as follows:

$$y = h(\boldsymbol{p}^g, \boldsymbol{p}^d, \boldsymbol{q}^d), \tag{3.1}$$

where $h(\cdot)$ captures the impacts from both the physical laws as well as reactive power control. We emphasize that although the voltage control problem is not explicitly modeled here, we assume certain voltage control schemes exist in the power distribution system such that the voltage profile will be maintained within an acceptable range. Indeed, voltage control schemes may have a significant impact on h.

Note that the explicit form of h is difficult to obtain; however, we can make the following assumption on h:

Assumption 1. The function h is differentiable and its first order partial derivatives with respect to $\boldsymbol{p}^g$ belong to $[\underline{b}_1, \overline{b}_1]$, where $\underline{b}_1$, $\overline{b}_1$ are some known constants. In addition, $\frac{\partial h}{\partial \boldsymbol{p}^g}$ is a Lipschitz function, i.e., there exists $b_2 > 0$ such that

$$\left\| \left. \frac{\partial h}{\partial \boldsymbol{p}^g} \right|_{\boldsymbol{a}} - \left. \frac{\partial h}{\partial \boldsymbol{p}^g} \right|_{\boldsymbol{b}} \right\| \leq b_2 \|\boldsymbol{a} - \boldsymbol{b}\|,$$

where $\boldsymbol{a}, \boldsymbol{b} \in [\underline{\boldsymbol{p}}^g, \overline{\boldsymbol{p}}^g]$.

Assumption 1 implies that, for fixed loads, the rate of change in y is bounded for bounded changes in the DER active power injections. In addition, the total active power provided to the bulk power system will increase when more active power is injected in the power distribution system. This assumption holds when the system is at a normal operating condition without line congestions.

3.2.2 Optimal DER Coordination Problem

The DERs in the distribution system can collectively provide active power to the bulk power system as quantified by the power exchange between both systems at the substation bus. For example, the DERs can provide demand response services or frequency regulation services to the bulk power system; in both cases, the DERs need to be coordinated in such a way that the total active power provided to the bulk power system, y, tracks some pre-specified value, denoted by $y^\star$. The objective of the ODCP is to determine the DER

active power injections, $\boldsymbol{p}^g$, that minimize some cost function, e.g., one that reflects the cost of active power provision, while respecting to the following constraints:

[C1.] the active power exchanged between the distribution and bulk power systems via bus 0, y, tracks some pre-specified value $y^\star$;

[C2.] the active power injection from each DER $i \in \mathcal{N}^g$, does not exceed its corresponding capacity limits, i.e., $\underline{\boldsymbol{p}}^g \leq \boldsymbol{p}^g \leq \overline{\boldsymbol{p}}^g$;

[C3.] the power flow on each line $\ell \in \mathcal{L}$, does not exceed its maximum capacity, i.e., $-\overline{\boldsymbol{f}} \leq \boldsymbol{f} \leq \overline{\boldsymbol{f}}$.

While constraint **C2** is a hard constraint that cannot be violated, constraint **C3** may be allowed to be violated slightly for a short period. The ODCP can be formulated as the following optimization problem:

$$\underset{\boldsymbol{p}^g \in [\underline{\boldsymbol{p}}^g, \overline{\boldsymbol{p}}^g]}{\text{minimize}} \; c(\boldsymbol{p}^g), \tag{3.2}$$

subject to

$$h(\boldsymbol{p}^g, \boldsymbol{p}^d, \boldsymbol{q}^d) = y^\star, \tag{3.3}$$

$$-\overline{\boldsymbol{f}} \leq \boldsymbol{M}^{-1}(\boldsymbol{C}\boldsymbol{p}^g - \boldsymbol{p}^d) \leq \overline{\boldsymbol{f}}, \tag{3.4}$$

where $c(\cdot)$ denotes the cost function of the active power injections, $\boldsymbol{M}$ is the reduced node-to-edge incidence matrix as in (1.15), $\boldsymbol{C}$ is the mapping matrix between DER indices and buses as in (1.1). Note that active power losses are taken into account when computing the active power exchange in (3.3), but are ignored when computing the active power line flows in (3.4), which is based on (1.15). The ODCP in (3.2)–(3.4) extends the ODCP in (2.29)–(2.31) with a more general cost function and the consideration of line flow constraints.

This problem is difficult, however, when the model describing the power exchange with the bulk power system, as captured by h, is unknown.

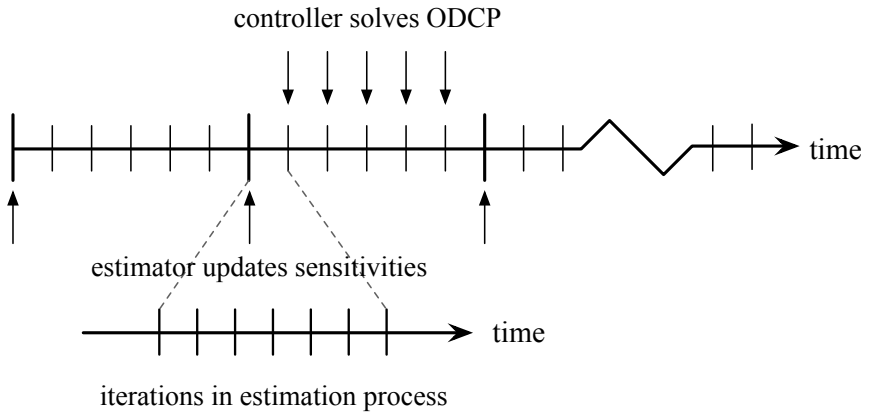

Figure 3.1: Timescale separation of actions taken.

3.3 Coordination Framework

In this section, we describe the building blocks of the proposed framework, namely an LTV IO model, an estimator, and a controller.

3.3.1 Overview

The DER coordination framework consists of three components, namely (i) an LTC model of the system describing the relation between y and $\boldsymbol{u}$, (ii) an estimator that provides estimates of the parameters—the so-called sensitivity vector—that populate the model in (i), and (iii) a controller that uses the model in (i) with the parameters estimated via (ii) to solve the ODCP in (3.2)–(3.4). This framework works on two timescales—a slow one and a fast one, as illustrated in Fig. 3.1. On the slow timescale, the controller determines the DER active power injection set-points by solving the ODCP periodically. The sensitivity vector is also updated by the estimator periodically. However, during each update of the sensitivity vector, the estimator needs to take actions in several iterations on a fast timescale. Since the sensitivity vector may not change significantly in a short time, it is used in the ODCP for several time instants before it is updated again. Before we proceed to presenting the detailed components in the proposed framework, we make the following assumption:

Assumption 2. $\boldsymbol{p}^d$ and $\boldsymbol{q}^d$ are constant during the estimation process; therefore, changes in y that occur across iterations in the estimation process de-

pend only on changes in $\boldsymbol{p}^g$.

Remark 1. Assumption 2 allows us to determine the impacts of the DER active power injections on the output. When the load variability is significant enough so that it cannot be neglected during the estimation process, it becomes necessary to measure the loads and determine their impacts on the output as well. This is beyond the scope of the dissertation, and therefore we will leave it as future work.

3.3.2 Input-Output System Model

For notational simplicity in the later development, define $\boldsymbol{u} = \boldsymbol{p}^g$, $\underline{\boldsymbol{u}} = \underline{\boldsymbol{p}}^g$, $\overline{\boldsymbol{u}} = \overline{\boldsymbol{p}}^g$, and $\boldsymbol{\pi} = [(\boldsymbol{p}^d)^\top, (\boldsymbol{q}^d)^\top]^\top$; then, (3.1) can be written as:

$$y = h(\boldsymbol{u}, \boldsymbol{\pi}). \tag{3.5}$$

Unless otherwise noted, throughout this chapter, $x[k]$ denotes the value that some variable x takes at iteration k. It follows from (3.5) and Assumption 2 that $y[k-1] = h(\boldsymbol{u}[k-1], \boldsymbol{\pi})$ and $y[k] = h(\boldsymbol{u}[k], \boldsymbol{\pi})$. Then, by the Mean Value Theorem, there exists $a_k \in [0, 1]$ and $\tilde{\boldsymbol{u}}[k] = a_k \boldsymbol{u}[k] + (1 - a_k)\boldsymbol{u}[k-1]$ such that

$$\begin{aligned} y[k] - y[k-1] &= h(\boldsymbol{u}[k], \boldsymbol{\pi}) - h(\boldsymbol{u}[k-1], \boldsymbol{\pi}) \\ &= \boldsymbol{\phi}[k]^\top (\boldsymbol{u}[k] - \boldsymbol{u}[k-1]), \end{aligned}$$

where $\boldsymbol{\phi}[k]^\top = [\phi_1[k], \cdots, \phi_n[k]] = \left.\dfrac{\partial h}{\partial \boldsymbol{u}}\right|_{\tilde{\boldsymbol{u}}[k]}$,[1] is referred to as the sensitivity vector at iteration k. Note that $\boldsymbol{\phi} = \mathbf{1}_n - \boldsymbol{C}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is the total LF vector as in (2.7). It follows from Assumption 2 that $\phi_i[k] \in [\underline{b}_1, \overline{b}_1]$, $i = 1, \cdots, n$. Therefore, at any iteration k, (3.5) can be transformed into the following equivalent LTV IO model:

$$y[k] = y[k-1] + \boldsymbol{\phi}[k]^\top (\boldsymbol{u}[k] - \boldsymbol{u}[k-1]). \tag{3.6}$$

[1] We adopt the convention that the partial derivative of a scalar function with respect to a vector is a row vector.

3.3.3 Estimator on Fast Timescale

As illustrated in Fig. 3.1, the estimator updates the sensitivity vector across several iterations on the fast timescale. At iteration k, the objective of the estimator is to obtain an estimate of $\boldsymbol{\phi}[k]$, denoted by $\hat{\boldsymbol{\phi}}[k]$, using measurements collected up to iteration k, i.e., $y[k-1], \boldsymbol{u}[k-1], y[k-2], \boldsymbol{u}[k-2], \cdots$; we formulate this estimation problem as follows:

$$\hat{\boldsymbol{\phi}}[k] = \underset{\hat{\boldsymbol{\phi}} \in \mathcal{B}}{\arg\min} \; J^e(\hat{\boldsymbol{\phi}}) = \frac{1}{2}(y[k-1] - \hat{y}[k-1])^2, \tag{3.7}$$

subject to

$$\hat{y}[k-1] = y[k-2] + \hat{\boldsymbol{\phi}}^\top (\boldsymbol{u}[k-1] - \boldsymbol{u}[k-2]), \tag{3.8}$$

where $\mathcal{B} = [\underline{b}_1, \overline{b}_1]^n$, $J^e(\cdot)$ is the cost function of the estimator, and $\hat{y}[k-1]$ is the value of $y[k-1]$ estimated by the IO model at iteration k. Essentially, the estimator aims to find $\hat{\boldsymbol{\phi}}$ that minimizes the squared error between the estimated value and the true value of y. Then, $\hat{\boldsymbol{\phi}}[k]$ is used in the controller to determine the control for the next time instant.

During the estimation process, it is still necessary to track the output target. Therefore, at each iteration, the control is set based on the solution to the following problem:

$$\boldsymbol{u}[k] = \underset{\boldsymbol{u} \in \mathcal{U}}{\arg\min} \; J^c(\boldsymbol{u}) = \frac{1}{2}(y^\star - \hat{y}[k])^2, \tag{3.9}$$

subject to

$$\hat{y}[k] = y[k-1] + \hat{\boldsymbol{\phi}}[k]^\top (\boldsymbol{u} - \boldsymbol{u}[k-1]), \tag{3.10}$$

where $\mathcal{U} = [\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}}]$, and $J^c(\cdot)$ is the cost function. Note that $\hat{\boldsymbol{\phi}}[k]$ is used in (3.10) to predict the value of $y[k]$ for a given $\boldsymbol{u}$. Different from the ODCP in (3.2)–(3.4), the objective of the controller during the estimation process is to ensure that the output tracks the target and there may exist multiple solutions to this problem. This objective is chosen such that the DER active power injections behave in a way that can improve the estimation accuracy, as will be shown later in Section 3.4.

3.3.4 Controller on Slow Timescale

As illustrated in Fig. 3.1, the controller solves the ODCP to determine the least-cost active power set-points for DERs on the slow timescale. Meanwhile, it also forces the DERs to inject random active power perturbations at each iteration on the fast timescale. The ODCP to be solved by the controller is as follows:

$$\underset{\boldsymbol{p}^g \in [\underline{\boldsymbol{p}}^g, \overline{\boldsymbol{p}}^g]}{\text{minimize}} \; c(\boldsymbol{p}^g), \tag{3.11}$$

subject to

$$y + \hat{\boldsymbol{\phi}}^\top (\boldsymbol{p}^g - \tilde{\boldsymbol{p}}^g) = y^\star, \tag{3.12}$$

$$-\overline{\boldsymbol{f}} \leq \boldsymbol{M}^{-1}(\boldsymbol{C}\boldsymbol{p}^g - \boldsymbol{p}^d) \leq \overline{\boldsymbol{f}}, \tag{3.13}$$

where y is the output and $\tilde{\boldsymbol{p}}^g$ is the DER active power injection vector at the beginning of the current time instant, and $\hat{\boldsymbol{\phi}}$ is the up-to-date sensitivity vector. When c is a convex function, (3.11)–(3.13) is a convex problem and therefore can be solved easily with convergence guarantees. This formulation is obtained by replacing (3.3) with the estimated IO model in (3.12).

3.4 Estimation Algorithm and Its Convergence

The ODCP in (3.11)–(3.13) can be solved using existing algorithms for convex optimization and thus we do not discuss it in more detail here. In this section, we focus on the problem faced by the estimator and propose a PGD based algorithm to solve it. We then provide convergence results for the proposed algorithm.

3.4.1 Estimation Algorithm

We first describe the basic workflow of the proposed algorithm. Each iteration consists of an estimation step and a control step. At the beginning of iteration k, $y[k-1]$ is available to the estimator, which uses it to update the estimate of the sensitivity vector. The updated estimate of the sensitivity vector, $\hat{\boldsymbol{\phi}}[k]$, is then used in the controller to determine the control,

$\boldsymbol{u}[k]$. Then, the DERs are instructed to change their active power injection set-points based on $\boldsymbol{u}[k]$. At time instant $k+1$, the estimation and control iterations are repeated once $y[k]$ becomes available. The sequential process described above, which happens on the fast timescale, is illustrated as follows:

$$\cdots \boldsymbol{u}[k-1] \to \underbrace{\boldsymbol{y}[k-1] \to \hat{\boldsymbol{\phi}}}_{\text{estimation step}} \overbrace{[k] \to \boldsymbol{u}[k]}^{\text{control step}} \to \boldsymbol{y}[k] \to \hat{\boldsymbol{\phi}}[k+1] \cdots$$

Problems (3.7)–(3.8) and (3.9)–(3.10) can be solved using the PGD method (see, e.g., [60]). Let $\mathsf{P}_{\mathbb{V}_1 \to \mathbb{V}_2}$ denote the projection operator from a vector space $\mathbb{V}_1$ to its (arbitrary) subspace $\mathbb{V}_2$, i.e.,

$$\mathsf{P}_{\mathbb{V}_1 \to \mathbb{V}_2}(\boldsymbol{v}_1) = \underset{\boldsymbol{v}_2 \in \mathbb{V}_2}{\arg\min} \|\boldsymbol{v}_2 - \boldsymbol{v}_1\|,$$

where $\boldsymbol{v}_1 \in \mathbb{V}_1$. For ease of notation, when the vector space to which $\boldsymbol{v}_1$ belongs is unambiguous, we simply write $\mathsf{P}_{\mathbb{V}_2}(\boldsymbol{v}_1)$ instead of $\mathsf{P}_{\mathbb{V}_1 \to \mathbb{V}_2}(\boldsymbol{v}_1)$.

Define the tracking error at iteration k as $e[k] = y[k] - y^\star$. In addition, define $\Delta y[k] = y[k] - y[k-1]$ and $\Delta \boldsymbol{u}[k] = \boldsymbol{u}[k] - \boldsymbol{u}[k-1]$. The partial derivative vector of $J^e(\hat{\boldsymbol{\phi}})$ with respect to $\hat{\boldsymbol{\phi}}$ is

$$\frac{\partial J^e(\hat{\boldsymbol{\phi}})}{\partial \hat{\boldsymbol{\phi}}} = \Delta \boldsymbol{u}[k-1](\Delta \boldsymbol{u}[k-1]^\top \hat{\boldsymbol{\phi}} - \Delta y[k-1]), \tag{3.14}$$

and that of $J^c(\boldsymbol{u})$ with respect to $\boldsymbol{u}$ is

$$\frac{\partial J^c(\boldsymbol{u})}{\partial \boldsymbol{u}} = \hat{\boldsymbol{\phi}}[k](\hat{\boldsymbol{\phi}}[k]^\top (\boldsymbol{u} - \boldsymbol{u}[k-1]) + e[k-1]). \tag{3.15}$$

Instead of solving these two problems to completion, we iterate the PGD algorithm that would solve them for one step at each iteration. Specifically, at iteration k, we evaluate the new gradient at $\hat{\boldsymbol{\phi}}[k-1]$ and $\boldsymbol{u}[k-1]$ and iterate once. Thus, by using (3.14) and (3.15), the update rules for $\hat{\boldsymbol{\phi}}$ and $\boldsymbol{u}$, respectively, are

$$\hat{\boldsymbol{\phi}}[k] = \mathsf{P}_{\mathcal{B}} \left(\hat{\boldsymbol{\phi}}[k-1] - \alpha_k \Delta \boldsymbol{u}[k-1](\Delta \boldsymbol{u}[k-1]^\top \hat{\boldsymbol{\phi}}[k-1] - \Delta y[k-1]) \right), \tag{3.16}$$

$$\boldsymbol{u}[k] = \mathsf{P}_{\mathcal{U}} \left(\boldsymbol{u}[k-1] - \beta_k e[k-1] \hat{\boldsymbol{\phi}}[k] \right), \tag{3.17}$$

Algorithm 3.1: Estimation Algorithm

Input:

y: output
$y^\star$: output tracking target
δ: maximum allowed tracking error
$\hat{\boldsymbol{\phi}}^0$: initial estimate of sensitivity vector
$\boldsymbol{u}^0$: initial DER active power injection set-point

Output:

$\boldsymbol{u}$: DER active power injection set-point
$\hat{\boldsymbol{\phi}}$: estimate of sensitivity vector

Initialization: set $\hat{\boldsymbol{\phi}}[0] = \hat{\boldsymbol{\phi}}^0$, $\boldsymbol{u}[-1] = \boldsymbol{u}[0] = \boldsymbol{u}^0$, obtain measurement of $y[-1]$, set $k = 1$

while $|e[k]| > \delta$ **do**

obtain new measurement of $y[k-1]$
compute $e[k-1] = y[k-1] - y^\star$
compute $\Delta y[k-1] = y[k-1] - y[k-2]$
compute $\Delta \boldsymbol{u}[k-1] = \boldsymbol{u}[k-1] - \boldsymbol{u}[k-2]$
update the sensitivity vector estimate according to

$$\hat{\boldsymbol{\phi}}[k] = \mathsf{P}_{\mathcal{B}} \Big(\hat{\boldsymbol{\phi}}[k-1] - \alpha_k \Delta \boldsymbol{u}[k-1] \big(\Delta \boldsymbol{u}[k-1]^\top \hat{\boldsymbol{\phi}}[k-1] - \Delta y[k-1] \big) \Big)$$

update the control vector, $\boldsymbol{u}$, according to

$$\boldsymbol{u}[k] = \mathsf{P}_{\mathcal{U}} \Big(\boldsymbol{u}[k-1] - \beta_k e[k-1] \boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k] \Big)$$

change DER active power injections to $\boldsymbol{u}[k]$
set $k = k + 1$

end

where $\alpha_k > 0$ and $\beta_k > 0$ are estimation and control step sizes at iteration k.

In order to resolve the potential issue of collinearity in the measurements used by the estimator, we introduce random perturbations during the estimation process. Define $\boldsymbol{W}[k] = \text{diag}(w_1[k], \ldots, w_n[k])$, where $w_i[k]$'s are independent random variables that follow a Bernoulli distribution with a probability parameter of 0.5. Then, the control update rule in (3.17) is modified, resulting in:

$$\boldsymbol{u}[k] = \mathsf{P}_{\mathcal{U}} \Big(\boldsymbol{u}[k-1] - \beta_k e[k-1] \boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k] \Big) . \tag{3.18}$$

Intuitively, this means that, at each iteration, the control of each DER is updated with probability 0.5. The random perturbation in the control is key to establish convergence of the parameter estimation process. The estimation algorithm, along with its initialization, is summarized in Algorithm 3.1, where $\boldsymbol{u}^0$ is the vector of DER active power injections at the beginning of the time instant at which the estimation starts and $\hat{\boldsymbol{\phi}}^0$ is the up-to-date estimate of the sensitivity vector at the beginning of the same time instant.

3.4.2 Convergence Analysis

The convergence analysis of the control step during the estimation process relies on the following two lemmas.

Lemma 1. There exists $\bar{\boldsymbol{\phi}}[k]$ satisfying $\mathbf{0}_n \leq \bar{\boldsymbol{\phi}}[k] \leq \hat{\boldsymbol{\phi}}[k]$, such that (3.18) is equivalent to

$$\boldsymbol{u}[k] = \boldsymbol{u}[k-1] - \beta_k e[k-1] \boldsymbol{W}[k] \bar{\boldsymbol{\phi}}[k].$$

Also, $\bar{\boldsymbol{\phi}}[k] = \mathbf{0}_n$ if and only if $\boldsymbol{u}[k] = \underline{\boldsymbol{u}}$ or $\boldsymbol{u}[k] = \overline{\boldsymbol{u}}$. Furthermore, if $\boldsymbol{u}[k] \neq \underline{\boldsymbol{u}}$ and $\boldsymbol{u}[k] \neq \overline{\boldsymbol{u}}$, there exists $i \in \mathcal{N}^g$ such that $\bar{\phi}_i[k] = \hat{\phi}_i[k] \in [\underline{b}_1, \overline{b}_1]$.

Proof. If $\boldsymbol{u}[k-1] - \beta_k e[k-1] \boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k] \in \mathcal{U}$, then we simply set $\bar{\boldsymbol{\phi}}[k] = \hat{\boldsymbol{\phi}}[k]$. Without loss of generality, first consider the case where the following holds for some $i \in \mathcal{N}^g$:

$$u_i[k-1] - \beta_k e[k-1] w_i[k] \hat{\phi}_i[k] > \overline{u}_i. \tag{3.19}$$

Then, $e[k-1] < 0$ and $w_i[k] > 0$ since otherwise (3.19) cannot hold. Therefore,

$$u_i[k] = \mathsf{P}_{\mathcal{U}}(u_i[k-1] - \beta_k e[k-1] w_i[k] \hat{\phi}_i[k]) = \overline{u}_i. \tag{3.20}$$

Let $\bar{\phi}_i[k] = \dfrac{u_i[k-1] - \overline{u}_i}{\beta_k e[k-1] w_i[k]}$; by definition, $\bar{\phi}_i[k] = 0$ if and only if $u_i[k-1] = \overline{u}_i$. Then, we have that:

$$u_i[k] = u_i[k-1] - \beta_k e[k-1] w_i[k] \bar{\phi}_i[k]. \tag{3.21}$$

If follows from (3.19), (3.20), and (3.21) that

$$\beta_k e[k-1]\hat{\phi}_i[k]w_i[k] < \beta_k e[k-1]\bar{\phi}_i[k]w_i[k], \tag{3.22}$$

which leads to $0 \leq \bar{\phi}_i[k] < \hat{\phi}_i[k]$. A similar argument can be used to for the case where $u_i[k-1] - \beta_k e[k-1]w_i[k]\hat{\phi}_i[k] < \underline{u}_i$ and for some $i \in \mathcal{N}^g$.

If $\boldsymbol{u}[k] \neq \underline{\boldsymbol{u}}$ and $\boldsymbol{u}[k] \neq \overline{\boldsymbol{u}}$, then there exists $i \in \mathcal{N}^g$ such that $\underline{u}_i < u_i[k] < \overline{u}_i$, which implies

$$u_i[k] = u_i[k-1] - \beta_k e[k-1]w_i[k]\hat{\phi}_i[k]. \tag{3.23}$$

Therefore, $\bar{\phi}_i[k] = \hat{\phi}_i[k]$. Consequently, $\bar{\phi}_i[k] = \hat{\phi}_i[k] \in [\underline{b}_1, \overline{b}_1]$. It can be easily seen that if $\mathcal{U}$ is sufficiently large and no DER hits its capacity limits, then $\bar{\boldsymbol{\phi}}[k] = \hat{\boldsymbol{\phi}}[k]$. □

Lemma 2. Let X_k, $k = 1, 2, \cdots$, be independently identically distributed (i.i.d.) random variables. Assume $X_k > 0$ and $\mathbb{E}[X_k] \in (0, 1)$, where $\mathbb{E}$ denotes expectation. Let $Y_k = \prod_{i=1}^{k} X_i$. Then, $\lim_{k\to\infty} Y_k = 0$ a.s.

Proof. Note that $Y_k = \exp\left(\sum_{i=1}^{k} \log X_i\right)$, where $\exp(\cdot)$ denotes the exponential function. By the Strong Law of Large Numbers (see Proposition 2.15 in [61]), we have that

$$\lim_{k\to\infty} \sum_{i=1}^{k} \frac{1}{k} \log X_i = \mathbb{E}[\log X_1], \text{ a.s.} \tag{3.24}$$

By Jensen's inequality (see Theorem 2.18 in [61]), we have that

$$\mathbb{E}[\log X_1] \leq \log \mathbb{E}[X_1] < 0. \tag{3.25}$$

Therefore,

$$\lim_{k\to\infty} \sum_{i=1}^{k} k\frac{1}{k} \log X_i = -\infty, \text{ a.s.}, \tag{3.26}$$

which leads to

$$\begin{aligned}\lim_{k\to\infty} Y_k &= \lim_{k\to\infty} \exp\left(\sum_{i=1}^{k} k\frac{1}{k}\log X_i\right) \\ &= \exp\left(\lim_{k\to\infty}\sum_{i=1}^{k} k\frac{1}{k}\log X_i\right) \\ &= 0, \text{ a.s.;} \end{aligned} \tag{3.27}$$

this completes the proof. □

Using Lemma 1 and Lemma 2, we can prove the following convergence result for the control step:

Theorem 1. Using the estimation update rule in (3.16) and the control update rule in (3.18) with $\beta_k \in (\frac{\epsilon}{\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, where $0 < \epsilon < \frac{\underline{b}_1^2}{n\bar{b}_1^2}$ is a given parameter, the system attains one of the following equilibria: 1) $e[k]$ converges to 0 a.s.; 2) $e[k]$ converges to some positive constant and $\boldsymbol{u}[k]$ stays at $\underline{\boldsymbol{u}}$; 3) $e[k]$ converges to some negative constant and $\boldsymbol{u}[k]$ stays at $\overline{\boldsymbol{u}}$. In all cases, $\lim_{k\to\infty} \Delta\boldsymbol{u}[k] = \mathbf{0}_n$, where $\mathbf{0}_n \in \mathbb{R}^n$ is an all-zeros vector.

Proof. By (3.6), we have that

$$e[k] - e[k-1] = \boldsymbol{\phi}[k]^\top \Delta\boldsymbol{u}[k]. \tag{3.28}$$

By Lemma 1, we have that

$$\Delta\boldsymbol{u}[k] = -\beta_k e[k-1]\boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k], \tag{3.29}$$

where $\mathbf{0}_n \leq \bar{\boldsymbol{\phi}}[k] \leq \hat{\boldsymbol{\phi}}[k]$. Substituting (3.29) into (3.28) leads to

$$e[k] = (1 - \beta_k \boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k])e[k-1]. \tag{3.30}$$

Define $\rho_k = 1 - \beta_k \boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k]$, then

$$e[k] = e[0]\prod_{i=1}^{k} \rho_i. \tag{3.31}$$

By Assumption 2, $0 < \underline{b}_1 \leq \phi_i[k] \leq \overline{b}_1$. In addition, it follows from Lemma 1 that $0 \leq \bar{\phi}_i[k] \leq \hat{\phi}_i[k] \leq \overline{b}_1$. Therefore, $\boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k]$ can be bounded as

follows:

$$0 \leq \boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k] = \sum_{i=1}^{n} w_i[k]\phi_i[k]\bar{\phi}_i[k] \leq n\overline{b}_1^2. \tag{3.32}$$

Since $\beta_k < \frac{1}{n\overline{b}_1^2}$, then all $e[k]$ has the same sign for all k (positive if $e[0] > 0$, and negative otherwise). As a result, the entries of $\Delta\boldsymbol{u}[k]$ always have the same sign by (3.29).

(a) If $e[k] = 0$ for some $k \in \mathbb{N}$, then $\Delta\boldsymbol{u}[k+1] = \mathbf{0}_n$. The control and estimation algorithms will stop updating according to (3.16) and (3.18). In this case, $\boldsymbol{u}[k]$ may equal to $\underline{\boldsymbol{u}}$ or $\overline{\boldsymbol{u}}$ or neither, and the system attains an equilibrium.

(b) Now suppose $e[k] \neq 0, \forall k \in \mathbb{N}$. Since the increments of $\boldsymbol{u}$ always have the same sign, the entries of $\boldsymbol{u}$ cannot hit their bounds in different directions, i.e., some hit their lower bounds while others hit their upper bounds.

$(b.1)$ If $\boldsymbol{u}[k] = \underline{\boldsymbol{u}}$ for some iteration k, then $e[k] > 0, \forall k \in \mathbb{N}$. By (3.18), we have that

$$\boldsymbol{u}[k+1] = \mathsf{P}_{\mathcal{U}}(\underline{\boldsymbol{u}} - \beta_k e[k]\boldsymbol{W}[k+1]\hat{\boldsymbol{\phi}}[k+1]) = \underline{\boldsymbol{u}}. \tag{3.33}$$

Thus, $\Delta\boldsymbol{u}[k+1] = \mathbf{0}_n$, which leads to $e[k+1] = e[k]$ by (3.28). Therefore, $\boldsymbol{u}$ will equal to $\underline{\boldsymbol{u}}$ and $e[k'] = e[k] > 0$ for all $k' > k$.

Similarly, when $\boldsymbol{u}[k] = \overline{\boldsymbol{u}}$, $\boldsymbol{u}$ will be equal to $\overline{\boldsymbol{u}}$, and e will be equal to $e[k] < 0$ in all future time intervals. The system attains an equilibrium in both cases.

$(b.2)$ If $\boldsymbol{u}[k] \neq \underline{\boldsymbol{u}}$ and $\boldsymbol{u}[k] \neq \overline{\boldsymbol{u}}$, $\forall k \in \mathbb{N}$, by Lemma 1, there exists $i \in \mathcal{N}^g$ such that $\bar{\phi}_i[k] \in [\underline{b}_1, \overline{b}_1]$. Then,

$$\boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k] = \sum_{i=1}^{n} \phi_i[k]\bar{\phi}_i[k]w_i[k] \geq \underline{b}_1^2 w_i[k]. \tag{3.34}$$

Thus, by using (3.32) and (3.34), it follows that $\rho_k \in [1-\beta_k n\overline{b}_1^2, 1-\beta_k \underline{b}_1^2 w_i[k]]$. Define $\overline{\rho}_k = 1 - \epsilon w_i[k]$, then $\overline{\rho}_k$ equals to $1-\epsilon$ or 1, each with probability 0.5, and $\mathbb{E}[\overline{\rho}_k] = 1 - \frac{\epsilon}{2} \in (0,1)$. Note that $0 < \epsilon < \frac{\underline{b}_1^2}{n\overline{b}_1^2}$ implies $\overline{\rho}_k > 0$. By

Lemma 2,

$$\lim_{k\to\infty}\prod_{i=1}^{k}\overline{\rho}_i = 0, \text{ a.s.} \tag{3.35}$$

When $\beta_k \in (\frac{\epsilon}{\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, $0 \le \rho_k \le \overline{\rho}_k$. Then, in an a.s. sense,

$$\lim_{k\to\infty}|e[k]| = |e[0]|\lim_{k\to\infty}\prod_{i=1}^{k}\rho_i \le |e[0]|\lim_{k\to\infty}\prod_{i=1}^{k}\overline{\rho}_i = 0. \tag{3.36}$$

Since $|e[k]| \ge 0$, $\lim_{k\to\infty}|e[k]| = 0$ a.s. In addition, by (3.29), $\lim_{k\to\infty}\Delta\boldsymbol{u}[k] = \mathbf{0}_n$ a.s. □

Remark 2. If $\mathcal{U}$ is sufficiently large and no DER hits the capacity limits, then $\bar{\boldsymbol{\phi}}[k] = \hat{\boldsymbol{\phi}}[k]$ and $\boldsymbol{\phi}[k]^\top \boldsymbol{W}[k]\bar{\boldsymbol{\phi}}[k] \ge \underline{b}_1^2\sum_{i=1}^{n} w_i[k]$. Following a similar argument as in part $(b.2)$ in the proof of Theorem 1, we can show $e[k]$ converges to 0 a.s. when $\beta_k \in (\frac{\epsilon}{n\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, where $0 < \epsilon < \frac{\underline{b}_1^2}{\bar{b}_1^2}$.

Theorem 1 shows something intuitive, i.e., the tracking error will be positive (negative) if the requested active power is less (more) than the minimum (maximum) amount of active power the DERs can provide; otherwise, the tracking error goes to zero a.s.

We note that ϵ has a direct impact on the convergence rate of the control algorithm. This is more obvious in a deterministic setting, when the control update rule in (3.17) is used instead of the one in (3.18). A result on the convergence rate is given in the following corollary.

Corollary 1. Assume $\boldsymbol{u}[k] \ne \underline{\boldsymbol{u}}$ and $\boldsymbol{u}[k] \ne \overline{\boldsymbol{u}}$, $\forall k \in \mathbb{N}$. Using the estimation update rule in (3.16) and the control update rule in (3.17) with $\beta_k \in (\frac{\epsilon}{\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, where $\epsilon > 0$ is a given parameter, $e[k]$ converges to 0 at a rate smaller that $1-\epsilon$, i.e., $\left|\frac{e[k]}{e[k-1]}\right| < 1-\epsilon$.

Proof. When the control update rule in (3.17) is used instead of the one in (3.18),

$$e[k] = (1 - \beta_k\boldsymbol{\phi}[k]^\top\bar{\boldsymbol{\phi}}[k])e[k-1]. \tag{3.37}$$

If $\boldsymbol{u}[k] \ne \underline{\boldsymbol{u}}$ and $\boldsymbol{u}[k] \ne \overline{\boldsymbol{u}}$, $\boldsymbol{\phi}[k]^\top\bar{\boldsymbol{\phi}}[k] = \phi_i[k]\bar{\phi}_i[k] \ge \underline{b}_1^2$. Define $\rho_k = 1 - \beta_k\boldsymbol{\phi}[k]^\top\bar{\boldsymbol{\phi}}[k]$. When $\beta_k \in (\frac{\epsilon}{\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, $\rho_k < 1-\epsilon$. Therefore, $\left|\frac{e[k]}{e[k-1]}\right| = \rho_k < 1-\epsilon$. □

Next, we establish the convergence of the estimation step. Define the estimation error vector at iteration k as $\boldsymbol{\varepsilon}[k] = \hat{\boldsymbol{\phi}}[k] - \boldsymbol{\phi}[k]$. Since both $\hat{\boldsymbol{\phi}}[k]$ and $\boldsymbol{\phi}[k]$ are bounded, $\boldsymbol{\varepsilon}[k]$ is also bounded. Define $\Delta\boldsymbol{\phi}[k] = \boldsymbol{\phi}[k] - \boldsymbol{\phi}[k-1]$.

The convergence analysis of the estimation step uses some convergence results for $\Delta\boldsymbol{\phi}[k]$, which are presented next.

Lemma 3. Let X_k, $k = 1, 2, \cdots$, be i.i.d. random variables that take value 1 with probability 0.5, or some constant $x \in (0, 1)$, also with probability 0.5. Let $Y_k = \prod_{i=1}^{k} X_i$ and $Z = \sum_{i=1}^{\infty} Y_i$. Then, Z is bounded a.s.

Proof. Let K denote the maximum number of 1's that appear continuously in the sequence $\{X_k\}$; then, the sequence $\{Y_k\}$ will have a new (smaller) value at most after $K + 1$ steps. We claim Z is unbounded only if K is infinite. Suppose $X_j = x$, and $X_k = 1$ for $k = j + 1, \cdots, j + m$, then $Y_j = Y_{j+1} = \cdots = Y_{j+m}$ and $\sum_{i=j}^{j+m} Y_j = (m + 1)Y_j \leq (K + 1)Y_j$. Therefore,

$$Z = \sum_{i=1}^{\infty} Y_i \leq (K + 1) \sum_{i=0}^{\infty} x^i = \frac{K + 1}{1 - x}. \tag{3.38}$$

It follows from (3.38) that Z is unbounded only if K is infinite. However, $\mathbb{P}\{M = \infty\} \leq \mathbb{P}\{X_{i+1} = \cdots = X_{i+K} = 1, \text{ for some } i\} = \frac{1}{2^\infty} = 0$, where $\mathbb{P}$ denotes probability. Thus, Z is bounded a.s. □

Lemma 4. Using estimation update rule (3.16) and control update rule (3.18), with $\beta_k \in (\frac{\epsilon}{n\underline{b}_1^2}, \frac{1}{n\bar{b}_1^2})$, where $0 < \epsilon < \frac{\underline{b}_1^2}{\bar{b}_1^2}$ is a given parameter, then

$$\lim_{k\to\infty} \|\Delta\boldsymbol{\phi}[k]\| = 0, \text{ a.s.} \tag{3.39}$$

and

$$\sum_{k=1}^{\infty} \|\Delta\boldsymbol{\phi}[k]\| < \infty, \text{a.s.} \tag{3.40}$$

Proof. If follows from the proof of Theorem 1 that the entries of $\Delta\boldsymbol{u}[k]$ always have the same sign. First consider the case where $\Delta\boldsymbol{u}[k] \geq \mathbf{0}_n$ for all $k \in \mathbb{N}$. Note that $\boldsymbol{\phi}[k]^\top = \left.\dfrac{\partial h}{\partial \boldsymbol{u}}\right|_{\tilde{\boldsymbol{u}}[k]}$, where $\tilde{\boldsymbol{u}}[k] = a_k\boldsymbol{u}[k] + (1 - a_k)\boldsymbol{u}[k-1]$ with $a_k \in [0, 1]$, i.e., $\boldsymbol{u}[k-1] \leq \tilde{\boldsymbol{u}}[k] \leq \boldsymbol{u}[k]$. Similarly, $\boldsymbol{\phi}[k-1]^\top = \left.\dfrac{\partial h}{\partial \boldsymbol{u}}\right|_{\tilde{\boldsymbol{u}}[k-1]}$,

where $\boldsymbol{u}[k-2] \leq \tilde{\boldsymbol{u}}[k-1] \leq \boldsymbol{u}[k-1]$. Thus, by Assumption 2, we have that

$$\begin{aligned}\|\Delta\boldsymbol{\phi}[k]\| &\leq b_2\|\tilde{\boldsymbol{u}}[k] - \tilde{\boldsymbol{u}}[k-1]\| \\ &\leq b_2\|\boldsymbol{u}[k] - \boldsymbol{u}[k-2]\| \\ &= b_2\|\Delta\boldsymbol{u}[k] + \Delta\boldsymbol{u}[k-1]\| \\ &\leq b_2(\|\Delta\boldsymbol{u}[k]\| + \|\Delta\boldsymbol{u}[k-1]\|). \end{aligned} \tag{3.41}$$

Since $\lim_{k\to\infty}\|\Delta\boldsymbol{u}[k]\| = 0$ a.s. by Theorem 1, as a result, $\lim_{k\to\infty}(\|\Delta\boldsymbol{u}[k]\| + \|\Delta\boldsymbol{u}[k-1]\|) = 0$ a.s., which gives

$$\lim_{k\to\infty}\|\Delta\boldsymbol{\phi}[k]\| = 0, \text{ a.s.} \tag{3.42}$$

Assume $\boldsymbol{u}[k] = \mathbf{0}_n$ for all $k < 0$, then we have that

$$\begin{aligned}\sum_{k=1}^{\infty}\|\Delta\boldsymbol{\phi}[k]\| &\leq \sum_{k=1}^{\infty} b_2(\|\Delta\boldsymbol{u}[k]\| + \|\Delta\boldsymbol{u}[k-1]\|) \\ &\leq 2b_2\sum_{k=0}^{\infty}\|\Delta\boldsymbol{u}[k]\| \\ &\leq 2b_2\sum_{k=0}^{\infty}\|\beta_k\boldsymbol{W}[k]\hat{\boldsymbol{\phi}}[k]e[k-1]\| \\ &\leq \frac{2b_2}{n\underline{b}_1^2}\sqrt{n}\bar{b}_1\sum_{k=0}^{\infty}|e[k-1]| \\ &= \frac{2b_2}{\sqrt{n}\bar{b}_1}\sum_{k=-1}^{\infty}|e[k]|. \end{aligned} \tag{3.43}$$

Recall that $\overline{\rho}_k$ equals to $1-\epsilon$ or 1, each with probability 0.5, where $\overline{\rho}_k$ is defined in the proof of Theorem 1. Therefore, by Lemma 3, $\sum_{k=1}^{\infty}\prod_{i=1}^{k}\overline{\rho}_i$ is bounded a.s. When $\beta_k \in (\frac{\epsilon}{\underline{b}_1^2}, \frac{1}{n\underline{b}_1^2})$, $0 \leq \rho_k \leq \overline{\rho}_k$, and

$$\sum_{k=0}^{\infty}|e[k]| = |e[0]|(1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\rho_i) \leq |e[0]|(1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\overline{\rho}_i). \tag{3.44}$$

As a result, $\sum_{k=-1}^{\infty}|e[k]|$ is bounded a.s. since $|e[-1]|$ is also bounded. The case where $\Delta\boldsymbol{u}[k] \leq \mathbf{0}_n$ for all $k \in \mathbb{N}$ can be proved similarly. □

The convergence analysis of the estimation step also relies on the following lemma (see Theorem 1 in [62]).

Lemma 5. Let X_k, Y_k, Z_k, $k = 1, 2, \cdots$, be non-negative variables in $\mathbb{R}$ such that $\sum_{k=0}^{\infty} Y_k < \infty$, and $X_{k+1} \leq X_k + Y_k - Z_k$, then X_k converges and $\sum_{k=0}^{\infty} Z_k < \infty$.

Using Lemma 4 and Lemma 5, we can prove the following convergence result for the estimation step:

Theorem 2. Using the estimation update rule in (3.16) and the control update rule (3.18), with $\alpha_{k+1} = \frac{2}{\|\Delta \boldsymbol{u}[k]\|^2}$, $\beta_k \in (\frac{\epsilon}{n\underline{b}_1^2}, \frac{1}{n\overline{b}_1^2})$, where $0 < \epsilon < \frac{\underline{b}_1^2}{\overline{b}_1^2}$ is a given parameter, if $\boldsymbol{u}[k] \in (\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}})$ and $e[k] \neq 0$, $\forall k \in \mathbb{N}$, then $\|\boldsymbol{\varepsilon}[k]\|$ converges to 0 a.s.

Proof. Consider an arbitrary sample path. Without loss of generality, assume $e[k] < 0$, it follows from Theorem 1 that $e[k] < 0, \forall k \in \mathbb{N}$. Since $\boldsymbol{u}[k] \in (\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}}), \forall k \in \mathbb{N}$, (3.18) becomes

$$\Delta \boldsymbol{u}[k] = -\beta_k e[k-1] \boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k]. \tag{3.45}$$

It follows from (3.6) and (3.16) that

$$\hat{\boldsymbol{\phi}}[k+1] = \mathsf{P}_{\mathcal{B}}(\hat{\boldsymbol{\phi}}[k] - \alpha_{k+1} \Delta \boldsymbol{u}[k] \Delta \boldsymbol{u}[k]^\top \boldsymbol{\varepsilon}[k]). \tag{3.46}$$

By definition, the estimation error at iteration k is

$$\boldsymbol{\varepsilon}[k+1] = \mathsf{P}_{\mathcal{B}}(\hat{\boldsymbol{\phi}}[k] - \alpha_{k+1} \Delta \boldsymbol{u}[k] \Delta \boldsymbol{u}[k]^\top \boldsymbol{\varepsilon}[k]) - \boldsymbol{\phi}[k+1]. \tag{3.47}$$

Since $\boldsymbol{\phi}[k+1] = \mathsf{P}_{\mathcal{B}}(\boldsymbol{\phi}[k+1])$, by the non-expansiveness of the projection operation (see Proposition 1.1.9 in [63]), then

$$\begin{aligned} \|\boldsymbol{\varepsilon}[k+1]\| &\leq \|\boldsymbol{\varepsilon}[k] - \alpha_{k+1} \Delta \boldsymbol{u}[k] \Delta \boldsymbol{u}[k]^\top \boldsymbol{\varepsilon}[k] - \Delta \boldsymbol{\phi}[k+1]\| \\ &\leq \|\boldsymbol{\varepsilon}[k] - \alpha_{k+1} \Delta \boldsymbol{u}[k] \Delta \boldsymbol{u}[k]^\top \boldsymbol{\varepsilon}[k]\| + \|\Delta \boldsymbol{\phi}[k+1]\|. \end{aligned} \tag{3.48}$$

Let $g(\alpha_{k+1}) = \|\boldsymbol{\varepsilon}[k] - \alpha_{k+1} \Delta \boldsymbol{u}[k] \Delta \boldsymbol{u}[k]^\top \boldsymbol{\varepsilon}[k]\|^2$; then, g attains its minimum at $\alpha_{k+1} = \frac{1}{\|\Delta \boldsymbol{u}[k]\|^2}$, which is

$$\|\boldsymbol{\varepsilon}[k]\|^2 - (\boldsymbol{\varepsilon}[k]^\top \frac{\Delta \boldsymbol{u}[k]}{\|\Delta \boldsymbol{u}[k]\|})^2 = \|\boldsymbol{\varepsilon}[k]\|^2 - (\boldsymbol{\varepsilon}[k]^\top \frac{\boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k]}{\|\boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k]\|})^2. \tag{3.49}$$

Define $\cos \theta_k = \frac{\boldsymbol{\varepsilon}[k]^\top}{\|\boldsymbol{\varepsilon}[k]\|} \frac{\boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k]}{\|\boldsymbol{W}[k] \hat{\boldsymbol{\phi}}[k]\|}$. Consequently, $g(\alpha_{k+1}) = (1 - \sin^2 \theta_k) \|\boldsymbol{\varepsilon}[k]\|^2$,

and

$$\|\boldsymbol{\varepsilon}[k+1]\| \leq |\sin\theta_k|\|\boldsymbol{\varepsilon}[k]\| + \|\Delta\boldsymbol{\phi}[k+1]\|. \tag{3.50}$$

Let $X_k = \|\boldsymbol{\varepsilon}[k]\|$, $Y_k = \|\Delta\boldsymbol{\phi}[k+1]\|$, and $Z_k = (1-|\sin\theta_k|)\|\boldsymbol{\varepsilon}[k]\|$. Then, $X_{k+1} \leq X_k + Y_k - Z_k$. Also, $\sum_{k=0}^{\infty} Y_k = \sum_{k=1}^{\infty}\|\Delta\boldsymbol{\phi}[k]\| < \infty$ by Lemma 4. Therefore, by Lemma 5, $\|\boldsymbol{\varepsilon}[k]\|$ converges, and $\sum_{k=1}^{\infty}(1-|\sin\theta_k|)\|\boldsymbol{\varepsilon}[k]\| < \infty$, which further implies $\lim_{k\to\infty}(1-|\sin\theta_k|)\|\boldsymbol{\varepsilon}[k]\| = 0$. Let $\varepsilon^\star$ denote the limit of $\|\boldsymbol{\varepsilon}[k]\|$; then,

$$\begin{aligned}\lim_{k\to\infty}|\sin\theta_k|\|\boldsymbol{\varepsilon}[k]\| &= \lim_{k\to\infty}(|\sin\theta_k|-1)\|\boldsymbol{\varepsilon}[k]\| + \lim_{k\to\infty}\|\boldsymbol{\varepsilon}[k]\| \\ &= \varepsilon^\star. \end{aligned}\tag{3.51}$$

Next, we show $\varepsilon^\star = 0$ by contradiction. Assume $\varepsilon^\star > 0$. Since both $\|\boldsymbol{\varepsilon}[k]\|$ and $|\sin\theta_k|\|\boldsymbol{\varepsilon}[k]\|$ converge to $\varepsilon^\star$,

$$\lim_{k\to\infty}|\sin\theta_k| = \frac{\lim_{k\to\infty}|\sin\theta_k|\|\boldsymbol{\varepsilon}[k]\|}{\lim_{k\to\infty}\|\boldsymbol{\varepsilon}[k]\|} = 1, \tag{3.52}$$

which implies $|\cos\theta_k|$ converges to 0. Since $\|\boldsymbol{\varepsilon}[k]\|$ and $\|\boldsymbol{W}[k]\hat{\boldsymbol{\phi}}[k]\|$ are bounded, then $|\boldsymbol{\varepsilon}[k]^\top\boldsymbol{W}[k]\hat{\boldsymbol{\phi}}[k]|$ converges to 0. Define $\mathsf{E}_i[k] = \{w_j[k] = 1 \text{ if } j = i, w_j[k] = 0 \text{ otherwise}\}$; then $\mathbb{P}\{\mathsf{E}_i[k]\} = \frac{1}{2^n}$. Thus, $\sum_{k=1}^{\infty}\mathbb{P}\{\mathsf{E}_i[k]\}$ equals to ∞. Also note that $\mathsf{E}_i[k]$, $k \in \mathbb{N}$, are independent. By the Borel-Cantelli Lemma (see Lemma 1.3 in [61]), $\mathbb{P}\{\mathsf{E}_i[k] \text{ infinitely often}\} = 1$; therefore, there are infinitely many time instants that $w_i[k] = 1$ and $w_j[k] = 0$ for all $j \neq i$. Let $\mathcal{K}_i$ denote the set of such time instants. Then $|\boldsymbol{\varepsilon}[k]^\top\boldsymbol{W}[k]\hat{\boldsymbol{\phi}}[k]| = |\varepsilon_i[k]\hat{\phi}_i[k]|$ for $k \in \mathcal{K}_i$. The sequence $\{|\varepsilon_i[k]\hat{\phi}_i[k]|, k \in \mathcal{K}_i\}$ is a subsequence of $\{|\boldsymbol{\varepsilon}[k]^\top\boldsymbol{W}[k]\hat{\boldsymbol{\phi}}[k]|\}$; therefore, it also converges to 0. Note that $\hat{\boldsymbol{\phi}}[k] > 0$; thus, $\varepsilon_i[k]$ converges to 0. Since i is arbitrary, we conclude that $\|\boldsymbol{\varepsilon}\|[k]$ converges to 0, which implies $\varepsilon^\star = 0$, contradiction. Since this result holds for all sample paths, then we conclude that $\|\boldsymbol{\varepsilon}[k]\|$ converges to 0 a.s. □

The intuition is that the estimation error goes to zero if the system can be continuously excited (guaranteed by the condition $\boldsymbol{u}[k] \in (\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}})$ and $e[k] \neq 0$, $\forall k \in \mathbb{N}$).

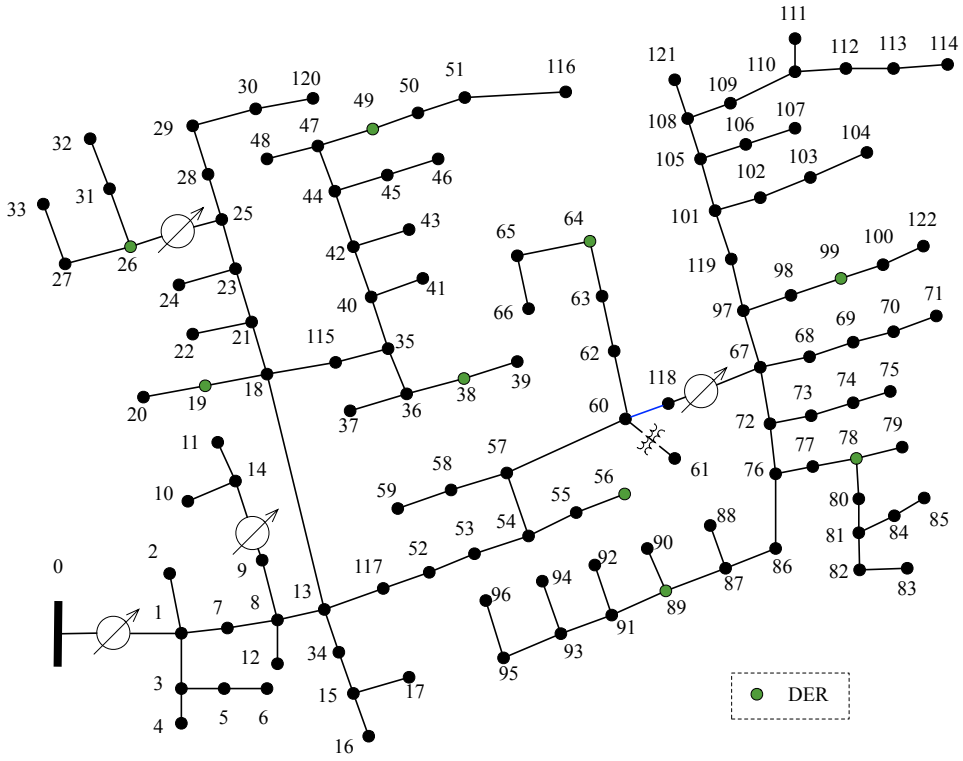

Figure 3.2: IEEE 123-bus distribution test feeder.

3.5 Numerical Simulation

In this section, we illustrate the application of the proposed DER coordination framework and validate the theoretical results presented earlier. From a practical point of view, the timescale separation illustrated in Fig. 3.1 is critical for the applicability of the proposed framework. Specifically, the estimation process needs to be much faster than the timescale governing the load changes. The DERs, which are typically power electronics interfaced, can respond very quickly, on a timescale of millisecond to second [64]. In this simulation, we set the duration between two iterations to be 100 ms. We will show later that under this setup, the requirements on the time separation can be reasonably met.

A modified three-phase balanced IEEE 123-bus distribution test feeder from [65] (see Fig. 3.2 for the one-line diagram) is used for all numerical simulations. This balanced test feeder has a total active power load of 3000 kW, and a total reactive power load of 1575 kVAr. DERs are added at buses 19, 26, 38, 49, 56, 64, 78, 89, 99. We assume each DER can output active

power from 0 kW to 100 kW. Therefore, the maximum DER active power injections account for 30% of the nominal loads. To illustrate the impacts of reactive power control, we assume all DERs operate at unity power factor except DERs at buses 78 and 89, which are assumed to have enough reactive power capacity and maintain a constant voltage magnitude of 0.95 p.u. Yet, we would like to emphasize that the proposed algorithm is agnostic to the underlying reactive power control scheme and also works under other reactive power control schemes. In addition, to validate the effectiveness of the proposed algorithm under different operating conditions of the power distribution system, we assume there are some uncontrollable renewable energy resources in the power distribution system, which are modeled as negative loads. The underlying nonlinear power flow problem is solved using Matpower [56].

In all subsequent simulations, we set $\underline{b}_1 = 0.8$, $\overline{b}_1 = 1.2$, which are reasonable values for real power systems. Intuitively, these values indicate that the percentage of active power losses will be no greater than 20% of the total active power injections. Note that the exact value of b_2 is not necessary. Under this simulation setting, as given in Theorem 1, the upper bound of the control step size is $\frac{1}{n\overline{b}_1^2} \approx 0.0694$.

We note that comprehensive simulations including the two timescales can be done using data such as ones adopted in [18]. However, since the ODCP to be solved on the slow timescale is a standard problem, we will mainly focus on simulations for the fast timescale, where our major contributions lie.

3.5.1 Case I

In this case, we assume the power distribution system is importing energy from the bulk power system with $y = -3110$ kW. This corresponds to the situation where the uncontrollable renewable energy resources are not generating more active power than that needed by the loads. In addition, we set $\hat{\boldsymbol{\phi}}^0 = \mathbf{1}_n$ and $\boldsymbol{u}^0 = \mathbf{0}_n$.

Tracking Performance During Estimation

For $y^\star = -3000$ kW and a constant step size $\beta_k = 0.02$, the DER active

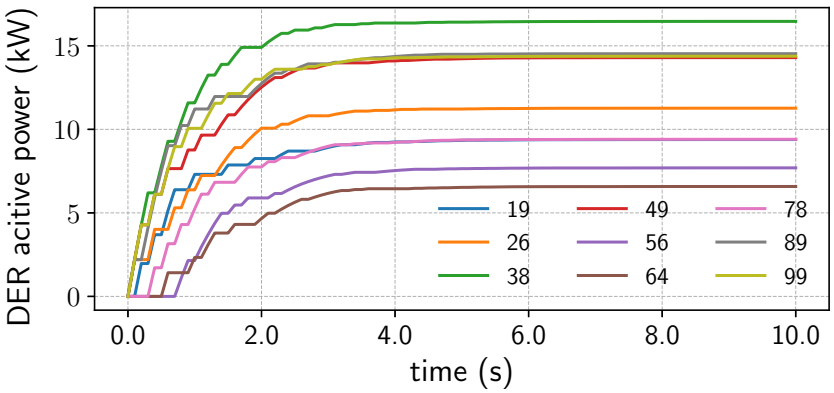

Figure 3.3: DER active power injections for $\beta_k = 0.02$ and $y^\star = -3000$ kW in Case I. (Legends indicate DER buses.)

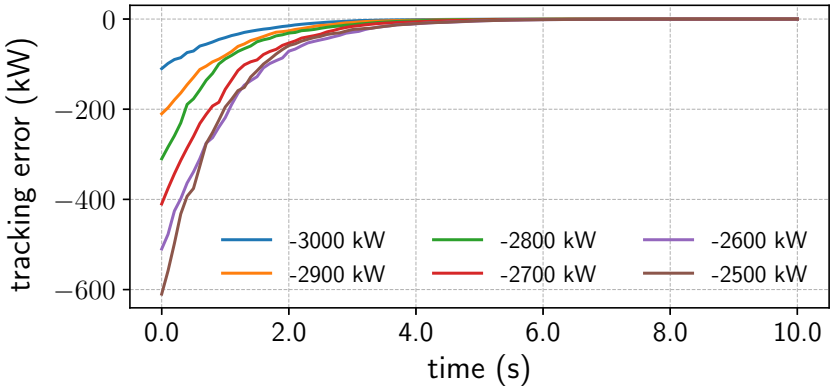

Figure 3.4: Tracking error for $\beta_k = 0.02$ under various tracking targets in Case I. (Legends indicate values of $y^\star$.)

power injections are shown in Fig. 3.3. The non-smoothness in the active power profiles is caused by the random perturbation imposed on the control step. Also as shown in Fig. 3.4, the convergence rate of the tracking error is not affected by the tracking target, i.e., the total active power required from the bulk power system. The tracking error $e[k]$ under various constant control step sizes is shown in Fig. 3.5. As expected, a larger step size will reduce the tracking error faster than a small step size.

Estimation Accuracy

With $\beta_k = 0.02$ and $y^\star = -3000$ kW, true and estimated sensitivities are compared in Table 3.1 and the MAE of estimation errors, i.e., $\sum_{i=1}^{n} |\varepsilon_i[k]|/n$,

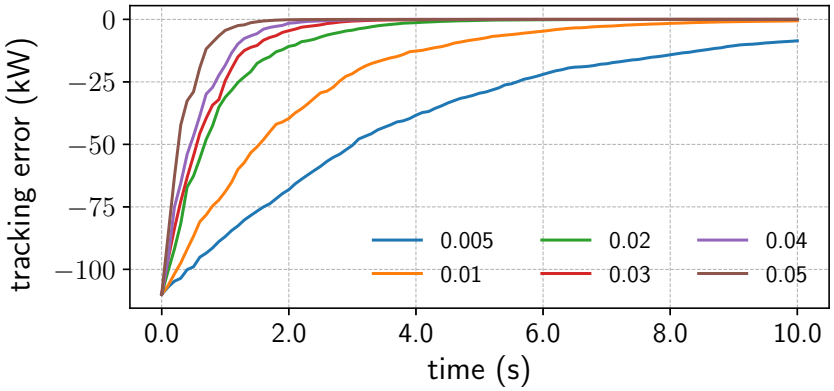

Figure 3.5: Tracking error for $y^\star = -3000$ kW and various constant control step sizes in Case I. (Legends indicate values of β_k.)

Table 3.1: Estimated Sensitivities in Case I After 60 Iterations

bus	19	26	38	49	56
true	1.0394	1.0413	1.0426	1.0454	1.0467
estimate	1.0342	1.0390	1.0440	1.0468	1.0421
bus	64	78	89	99	
true	1.0702	1.0703	1.0749	1.0711	
estimate	1.0696	1.0697	1.0817	1.0702	

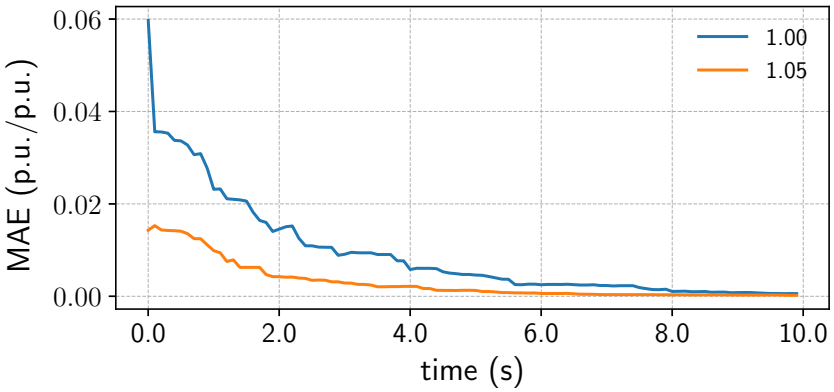

Figure 3.6: MAE of estimation errors with $\beta_k = 0.02$ in Case I. (Legends indicate the values of initial estimates.)

is plotted in Fig. 3.6. The estimated sensitivities are very close to their true values after 60 steps, which corresponds to 6 s. Note that $\hat{\phi}^0$ has an important impact on the convergence of the sensitivity estimation algorithm. As can

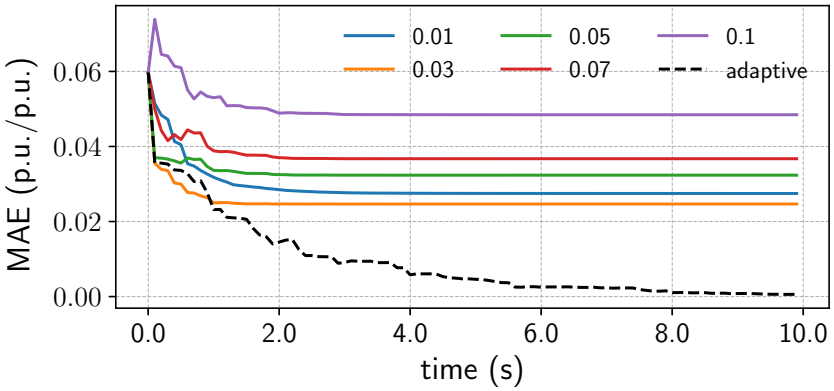

Figure 3.7: MAE of estimation errors with $\beta_k = 0.02$ under various estimation step sizes in Case I. (Legends indicate values of α_k.)

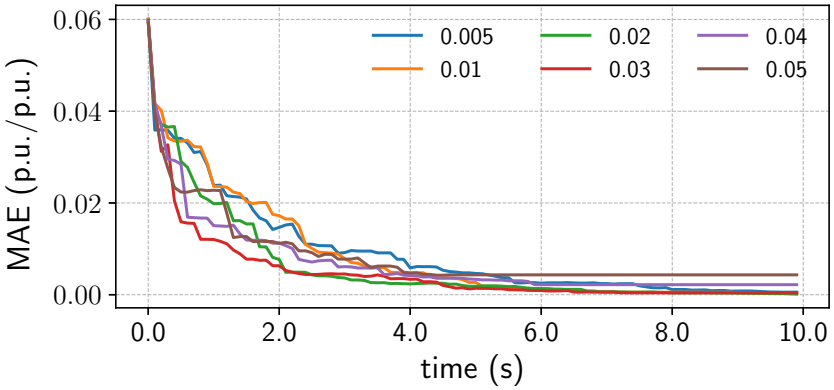

Figure 3.8: MAE of estimation errors under various control step sizes in Case I. (Legends indicate values of β_k.)

be seen from Fig. 3.6, when the initial values of the estimated sensitivities are set to 1.05, which is closer to their true values, it takes much less time to obtain a small estimation error.

While the estimation step size, α_k, in the proposed algorithm is adaptive, we also investigate the case when α_k is chosen to be constant. Figure 3.7 shows the MAE of estimation error under various constant estimation step sizes. As can be seen from Fig. 3.7, the MAE of estimation will converge to some non-zero constant under constant estimation step sizes.

The impact of the control step sizes on the estimation accuracy is also investigated. Figure 3.8 shows the MAE of estimation errors under various control step sizes. With a large control step size, the tracking error converges

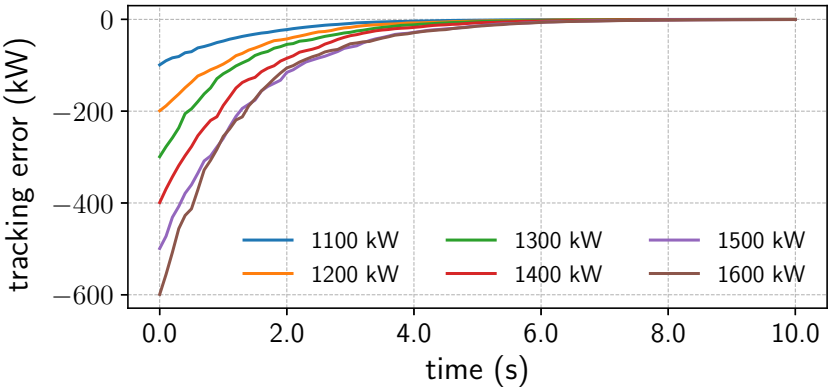

Figure 3.9: Tracking error for $\beta_k = 0.02$ under various tracking targets in Case II. (Legends indicate values of $y^\star$.)

to 0 quickly, leading to a situation where the system cannot get sufficient excitation and consequently, the estimation errors cannot be further reduced.

3.5.2 Case II

In this case, we assume the power distribution system is exporting energy to the bulk power system with $y = 1000$ kW. This corresponds to the situation where the uncontrollable renewable energy resources are generating much more active power than that needed by the loads. We set $\hat{\boldsymbol{\phi}}^0 = \mathbf{1}_n$ and $\boldsymbol{u}^0 = \mathbf{0}_n$.

Tracking Performance During Estimation

Using a constant step size $\beta_k = 0.02$, the convergence rate of the tracking error under various tracking target is shown in Fig. 3.9. Similar to results in Case I, the convergence rate is not affected by the tracking target.

Estimation Accuracy

With $\beta_k = 0.02$ and $y^\star = 1100$ kW, true and estimated sensitivities are compared in Table 3.2 and the MAE of estimation errors is plotted in Fig. 3.10, respectively. Similar to the results in Case I, the estimated sensitivities are very close to their true values after 60 steps. This verifies that the proposed estimation algorithm can effectively estimate the sensitivities under different operating conditions of the power distribution system.

Table 3.2: Estimated Sensitivities in Case II After 60 Iterations

bus	19	26	38	49	56
true	0.9533	0.9526	0.9518	0.9509	0.9254
estimate	0.9588	0.9512	0.9497	0.9475	0.9285
bus	64	78	89	99	
true	0.8872	0.8477	0.8488	0.8700	
estimate	0.8854	0.8536	0.8396	0.8707	

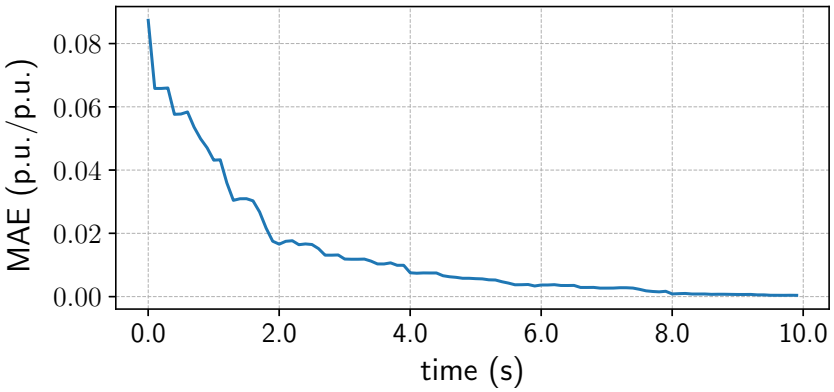

Figure 3.10: MAE of estimation errors with $\beta_k = 0.02$ in Case II.

We note that the performance of the proposed estimation algorithm is independent of the number of DERs. To see this, we simulate a case where the DER at bus 99 gets disconnected and consequently there are 8 DERs left in the power distribution system. The sensitivities at these 8 DER buses can still be estimated effectively, as is shown in Fig. 3.11.

3.5.3 Case III

In this case, we illustrate how the proposed framework handles line congestions. The setup is the same as Case II except that the tracking target is $y^\star = 1500$ kW. We set the capacity limit of line $(55, 56)$ to 40 kW to create congestion. For simplicity, all other lines are assumed to have an infinite capacity; yet, we would like to emphasize that the proposed framework can handle multiple congestion of multiple lines. The objective function in the ODCP is assumed to be $c(\boldsymbol{p}^g) = \|\boldsymbol{p}^g - \tilde{\boldsymbol{p}}^g\|^2$, where $\tilde{\boldsymbol{p}}^g$ is the current DER active power injection vector as used in (3.11). Intuitively, this objective

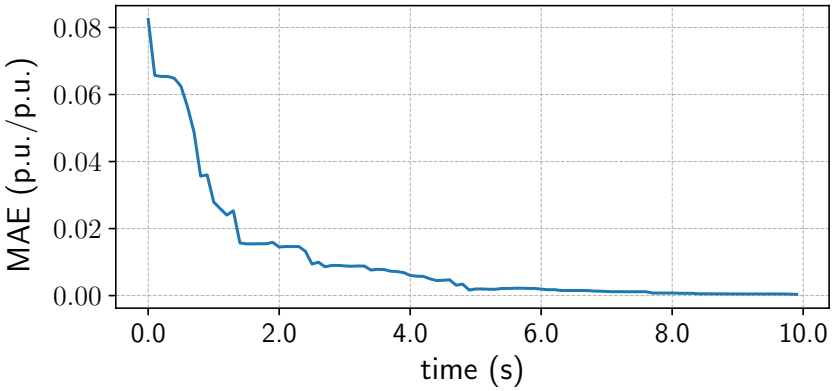

Figure 3.11: MAE of estimation errors with $\beta_k = 0.02$ in Case II with 8 DERs.

function will favor the solution with the least change in the DER active power injections.

The estimation algorithm is first run to obtain an estimate of the sensitivity vector. After the estimation algorithm ends, the DER at bus 56 is generating 51.3 kW, which exceeds the capacity limit of line $(55, 56)$. The ODCP is run afterwards to adjust the active power set-points of the DERs. Figure 3.12 shows the DER active power set-point before and after solving the ODCP. The DER at bus 56 is dispatched down to 40 kW, which conforms with the capacity limit of line $(55, 56)$. All other DERs are dispatched up such that the active power exchanged between the distribution and bulk power systems still equals to 1500 kW. We note that line $(55, 56)$ is overloaded for a short period during the estimation process but it is quickly restored to a normal loading level after the DER active power set-points are adjusted via the ODCP.

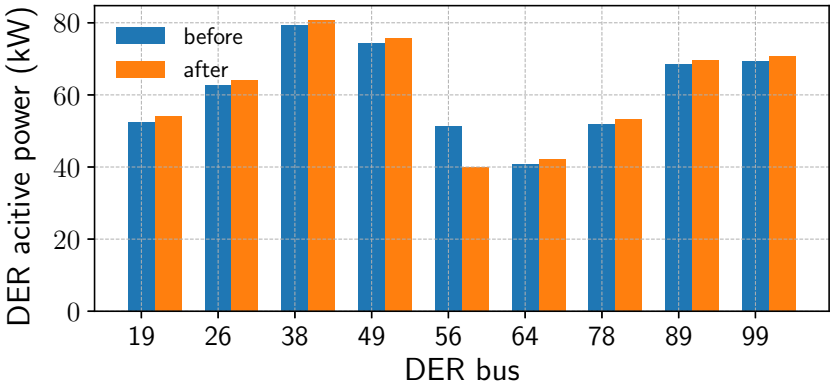

Figure 3.12: DER active power set-point before and after solving the ODCP in Case III.

3.6 Summary

In this chapter, we have proposed a data-driven coordination framework for DERs in a lossy power distribution system, the model of which is not completely known, to collectively provide some pre-specified amount of active power to a bulk power system at a minimum generating cost, while respecting distribution line capacity limits. The proposed framework consists of an LTV IO model, an estimator, and a controller.

We have shown that using the estimation algorithm proposed in the framework, the estimated parameters converge to the true parameters a.s., and the total provided active power converges to the required amount during the estimation process. The data-driven nature of this framework makes it adaptive to various system operating conditions. We validated the effectiveness of the proposed framework through numerical simulations on a modified version of the IEEE 123-bus test feeder.

The work presented in this chapter was published in [66].

CHAPTER 4

DER COORDINATION FOR REACTIVE POWER REGULATION

4.1 Introduction

In this chapter, we address the problem of DER coordination for reactive power regulation, or equivalently, voltage control, so as to maintain the voltage magnitude at each bus in a radial power distribution system within a desirable range. Following the same spirit of the data-driven approaches pursued in previous chapters, we propose a data-driven DER coordination framework for voltage control that consists of three fundamental building blocks: (i) an IO model that captures the relation between the voltage magnitudes and power injections, (ii) an estimator of the IO model based on measurements, and (iii) a controller that solves an ODCP—the objective of which is to minimize the voltage excursions outside a desirable voltage range—that uses the IO model estimated in (ii).

Specifically, we will take advantage of the LinDistFlow model to simplify the nonlinear relationships between voltage magnitudes and power injections. The coefficients of the LinDistFlow model are essentially the sensitivities of the squared voltage magnitudes with respect to active and reactive power injections, referred to as the voltage sensitivities, and can be computed directly using system topology and line parameter information. Assuming the knowledge of feasible topology configurations and distribution line "r-to-x" ratios, which are typically available and do not change during a relatively short time period, the true topology configuration and corresponding line parameters can be estimated effectively using a few voltage magnitude and power injection measurements. Using the estimated voltage sensitivities, the optimal DER power injections can be readily determined by solving a convex ODCP. Theoretical analysis shows that the voltage sensitivities of interest are easily identifiable.

The remainder of the chapter is organized as follows. In Section 4.2, the IO model adopted in this chapter and the voltage regulation problem are introduced. The details of the proposed data-driven DER coordination framework for reactive power regulation are presented in Section 4.3. The identifiability of the voltage sensitivities is analyzed in Section 4.4. The effectiveness of the proposed framework is validated in Section 4.5 through numerical simulations. We summarize this chapter in Section 4.6.

4.2 Preliminaries

In this section, we briefly review the IO model adopted in this chapter and describe the ODCP of interest.

4.2.1 Input-Output System Model

Consider the power distribution system described in Section 1.2.1. The relation between the vector of squared voltage magnitude deviation $\boldsymbol{v}$, and the vectors of active and reactive power injections, $\boldsymbol{p}$ and $\boldsymbol{q}$, is recapped here as follows:

$$\tilde{\boldsymbol{v}} = \boldsymbol{R}\boldsymbol{p} + \boldsymbol{X}\boldsymbol{q}, \tag{4.1}$$

where $\tilde{\boldsymbol{v}}$ is a vector of difference between squared voltage magnitudes for buses 1 to N and that for bus 0; $\boldsymbol{p}$ and $\boldsymbol{q}$ are vectors of active and reactive power injections, respectively; $\boldsymbol{R} = 2(\boldsymbol{M}^{-1})^{\top}\mathrm{diag}(\boldsymbol{r})\boldsymbol{M}^{-1}$, and $\boldsymbol{X} = 2(\boldsymbol{M}^{-1})^{\top}\mathrm{diag}(\boldsymbol{x})\boldsymbol{M}^{-1}$ in which $\boldsymbol{r}$ and $\boldsymbol{x}$ are vectors of line resistances and reactances, respectively; and $\boldsymbol{M}$ is the reduced node-to-edge incidence matrix. The inputs in (4.1) are the active and reactive power injections, while the outputs are the deviations of squared voltage magnitudes.

Note that the topology of the power distribution system is uniquely determined by $\boldsymbol{M}$; therefore, we also refer to $\boldsymbol{M}$ as the system topology configuration. A power distribution system may be operated under various feasible topology configurations. Let $\mathcal{M} = \{\boldsymbol{M}_1, \cdots, \boldsymbol{M}_\tau\}$ denote the set of τ feasible topology configurations of the power distribution system. Note that each topology configuration is associated with a vector of "r-to-x" ra-

tios. Let ζ_ℓ denote the "r-to-x" ratio of line ℓ, i.e., $r_\ell/x_\ell = \zeta_\ell$, and define $\boldsymbol{\zeta} = [\zeta_1, \cdots, \zeta_L]^\top$. Let $\mathcal{Z} = \{\boldsymbol{\zeta}_1, \cdots, \boldsymbol{\zeta}_\tau\}$ denote the set of τ "r-to-x" ratio vectors associated with $\mathcal{M}$.

4.2.2 Optimal DER Coordination Problem

The objective here is to maintain the voltage magnitude at each bus $i \in \mathcal{N}$, of the power distribution system within a pre-specified interval $[\underline{V}_i, \overline{V}_i], 0 \leq \underline{V}_i \leq \overline{V}_i$. Voltage regulation in power distribution systems can be effectively accomplished by a two-timescale architecture (see, e.g., [10]), where on the slow timescale, slower actuation devices including load tap changers and capacitor banks are adjusted to minimize voltage deviations from the desired range, and on the fast timescale, fast actuators such as DERs are dispatched. In this chapter, we focus solely in the later mechanism for achieving voltage regulation on the fast timescale. Then, the problem is to determine DER active and reactive power injections so that

[C1.] the active and reactive power injections from each DER $i \in \mathcal{N}^g$ do not exceed its corresponding capacity limits, i.e., $\underline{\boldsymbol{p}}^g \leq \boldsymbol{p}^g \leq \overline{\boldsymbol{p}}^g$, $\underline{\boldsymbol{q}}^g \leq \boldsymbol{q}^g \leq \overline{\boldsymbol{q}}^g$; and

[C2.] the voltage magnitude at each bus $i \in \mathcal{N}$ is within the pre-specified interval, i.e., $\underline{V}_i \leq V_i \leq \overline{V}_i$.

In addition, among all feasible values of $\boldsymbol{p}^g$ and $\boldsymbol{q}^g$, we would like to select the ones that minimize some cost function, which reflects the cost of voltage deviations as well as the cost of active and reactive power provision.

Except for $\mathcal{M}$ and $\mathcal{Z}$, which are known, we assume no prior information on the voltage sensitivity matrices. The voltage regulation problem cannot be solved without knowing the voltage sensitivity matrices. Therefore, we will resort to a data-driven approach to estimate voltage sensitivity matrices from measurements of voltage magnitudes and power injections.

4.3 Coordination Framework

In this section, we propose a data-driven DER coordination framework for voltage regulation. We first give an overview of the framework and then

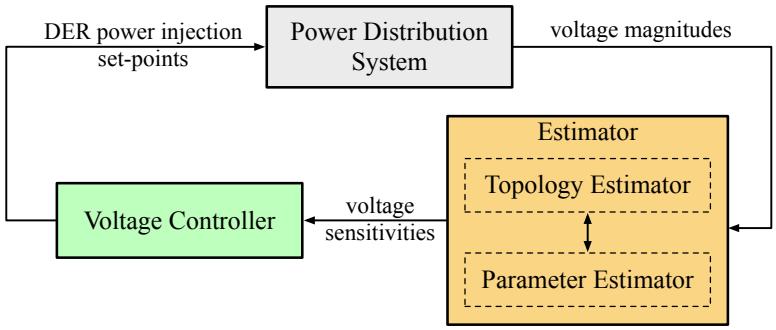

Figure 4.1: Data-driven voltage regulation framework.

present the details of its fundamental building blocks.

4.3.1 Framework Overview

The proposed DER coordination framework consists of three components: an IO model that takes DER power injections as inputs and squared voltage magnitudes deviations as outputs, a voltage sensitivity estimator and a voltage controller. The interaction between the different components and the power distribution system is illustrated via the block diagram in Fig. 4.1. The estimator component contains a topology estimator that determines the topology configuration of the power distribution system, i.e., it determines $\boldsymbol{M}$, and a parameter estimator that estimates the line parameters $\boldsymbol{r}$ and $\boldsymbol{x}$ using measurements of power injections and voltage magnitudes. The estimates of the voltage sensitivity matrices $\boldsymbol{R}$ and $\boldsymbol{X}$, denoted respectively by $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{X}}$, are computed from estimates of $\boldsymbol{M}$, $\boldsymbol{r}$, and $\boldsymbol{x}$. After that, $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{X}}$ are sent to the voltage controller. The voltage controller then computes the set-points for the DER active and reactive power injections that minimize some cost function subject to constraints **C1** and **C2**. The DERs will be instructed to inject the amount of active and reactive power determined by the voltage controller. A new set of measurements will be available once the DERs have modified their power injections. These measurements will be used by the estimator to update $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{X}}$ so as to reflect any changes in them. The detailed formulations for the voltage sensitivity problem and the voltage regulation problem are presented next.

4.3.2 Voltage Sensitivity Estimator

At time instant $k+1$, assume we have measurements $V_0[t]$, $\boldsymbol{V}[t]$, $\boldsymbol{p}[t]$, $\boldsymbol{q}[t]$, $t = 0, 1, \cdots, k$, where the index t indicates that the corresponding measurement is obtained at time instant t. To reduce the computational burden, we select a subset of measurements, denoted by $\mathcal{K} = \{k - m, \cdots, k\}$. The voltage sensitivities can be estimated based on the LinDistFlow model in (4.1). The goal of the voltage sensitivity estimator at time instant k is to estimate the values of $\boldsymbol{R}$ and $\boldsymbol{X}$ from the measurements obtained at the time instants in $\mathcal{K}$. We propose a voltage sensitivity estimator that consists of two components: a parameter estimator and a topology estimator. The former aims to estimate the line parameters given the topology configuration, i.e., $\boldsymbol{M}$, while the later aims to determine $\boldsymbol{M}$ from $\mathcal{M}$ based on the results from the parameter estimator.

Parameter estimator

For a given $\boldsymbol{M} \in \mathcal{M}$, estimating $\boldsymbol{R}$ and $\boldsymbol{X}$ boils down to estimating $\boldsymbol{r}$ and $\boldsymbol{x}$. Let $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{x}}$ denote the estimates of $\boldsymbol{r}$ and $\boldsymbol{x}$, respectively. We can then formulate the parameter estimation problem by using the relation in (4.1) as

$$(\hat{\boldsymbol{r}}, \hat{\boldsymbol{x}}) = \underset{(\boldsymbol{r}, \boldsymbol{x})}{\arg\min} \sum_{t \in \mathcal{K}} \eta^{k-t} \|\boldsymbol{R}\boldsymbol{p}[t] + \boldsymbol{X}\boldsymbol{q}[t] - \tilde{\boldsymbol{v}}[t]\|^2, \tag{4.2}$$

subject to

$$\boldsymbol{R} = 2(\boldsymbol{M}^{-1})^\top \mathrm{diag}(\boldsymbol{r}) \boldsymbol{M}^{-1}, \tag{4.3}$$

$$\boldsymbol{X} = 2(\boldsymbol{M}^{-1})^\top \mathrm{diag}(\boldsymbol{x}) \boldsymbol{M}^{-1}, \tag{4.4}$$

where $\eta \in (0, 1]$ is a forgetting factor. Essentially, the objective of the parameter estimator is to find the line parameters that fit the LinDistFlow model best for a given topology configuration.

We next show that (4.2)–(4.4) has a closed-form solution. First note that the matrix $\mathrm{diag}(\boldsymbol{x})$ can be decomposed as follows:

$$\mathrm{diag}(\boldsymbol{x}) = \sum_{\ell=1}^{L} x_\ell \boldsymbol{e}_\ell \boldsymbol{e}_\ell^\top, \tag{4.5}$$

where $\boldsymbol{e}_\ell$ is the ℓ^{th} basis vector in $\mathbb{R}^L$, i.e., all entries in $\boldsymbol{e}_\ell$ are 0 except the

ℓ^{th} entry, which is equal to 1. Using (4.5), we obtain that

$$\begin{aligned}\boldsymbol{X}\boldsymbol{q}[t] &= 2(\boldsymbol{M}^{-1})^\top \text{diag}(\boldsymbol{x})\boldsymbol{M}^{-1}\boldsymbol{q}[t] \\ &= 2(\boldsymbol{M}^{-1})^\top \sum_{\ell=1}^{L} x_\ell \boldsymbol{e}_\ell \boldsymbol{e}_\ell^\top \boldsymbol{M}^{-1}\boldsymbol{q}[t] \\ &= \sum_{\ell=1}^{L} \boldsymbol{\Gamma}_\ell \boldsymbol{q}[t] x_\ell, \end{aligned} \tag{4.6}$$

where $\boldsymbol{\Gamma}_\ell = 2(\boldsymbol{M}^{-1})^\top \boldsymbol{e}_\ell \boldsymbol{e}_\ell^\top \boldsymbol{M}^{-1}$. Similarly,

$$\begin{aligned}\boldsymbol{R}\boldsymbol{p}[t] &= \sum_{\ell=1}^{L} \boldsymbol{\Gamma}_\ell \boldsymbol{p}[t] r_\ell \\ &= \sum_{\ell=1}^{L} \boldsymbol{\Gamma}_\ell \zeta_\ell \boldsymbol{p}[t] x_\ell. \end{aligned} \tag{4.7}$$

Let $\boldsymbol{\rho}_\ell[t] = \eta^{\frac{k-t}{2}}(\zeta_\ell \boldsymbol{p}[t] + \boldsymbol{q}[t])$, $\ell \in \mathcal{L}$, and define

$$\boldsymbol{\Psi}[k] = \begin{bmatrix} \boldsymbol{\Gamma}_1\boldsymbol{\rho}_1[k-m] & \cdots & \boldsymbol{\Gamma}_L\boldsymbol{\rho}_L[k-m] \\ \vdots & \vdots & \vdots \\ \boldsymbol{\Gamma}_1\boldsymbol{\rho}_1[k] & \cdots & \boldsymbol{\Gamma}_L\boldsymbol{\rho}_L[k] \end{bmatrix}, \tag{4.8}$$

and

$$\boldsymbol{\psi}[k] = [\eta^{\frac{m}{2}}\tilde{\boldsymbol{v}}[k-m]^\top, \cdots, \eta^{\frac{0}{2}}\tilde{\boldsymbol{v}}[k]^\top]^\top. \tag{4.9}$$

Note that $\boldsymbol{\Psi}[k] \in \mathbb{R}^{(m+1)N \times L}$ and $\boldsymbol{\psi}[k] \in \mathbb{R}^{(m+1)N}$ are dependent on $\mathcal{K}$. Then (4.2) can be equivalently formulated in the classical form of a linear regression problem as follows:

$$\underset{\boldsymbol{x}}{\text{minimize}} \|\boldsymbol{\Psi}[k]\boldsymbol{x} - \boldsymbol{\psi}[k]\|^2, \tag{4.10}$$

the solution of which is given by

$$\hat{\boldsymbol{x}} = \boldsymbol{\Psi}[k]^\dagger \boldsymbol{\psi}[k], \tag{4.11}$$

where $\boldsymbol{\Psi}[k]^\dagger = (\boldsymbol{\Psi}[k]^\top \boldsymbol{\Psi}[k])^{-1}\boldsymbol{\Psi}[k]^\top$ denotes the pseudo-inverse of $\boldsymbol{\Psi}[k]$. When $\boldsymbol{\Psi}[k]^\top \boldsymbol{\Psi}[k]$ is singular, $\boldsymbol{\Psi}[k]^\dagger$ can be obtained via singular value decom-

position (SVD). Note that $\boldsymbol{\Psi}[k]$ needs to have full rank, i.e., $\text{rank}(\boldsymbol{\Psi}[k]) = L$, in order to be able to estimate $\boldsymbol{x}$. The resistance vector is estimated as

$$\hat{\boldsymbol{r}} = \text{diag}(\boldsymbol{\zeta})\hat{\boldsymbol{x}}. \tag{4.12}$$

Topology estimator

Define a residual vector, denoted by $\boldsymbol{\varepsilon}$, as follows:

$$\boldsymbol{\varepsilon} = \hat{\boldsymbol{R}}\boldsymbol{p} + \hat{\boldsymbol{X}}\boldsymbol{q} - \tilde{\boldsymbol{v}}, \tag{4.13}$$

where

$$\hat{\boldsymbol{R}} = 2(\boldsymbol{M}^{-1})^{\top}\text{diag}(\hat{\boldsymbol{r}})\boldsymbol{M}^{-1}, \tag{4.14}$$

$$\hat{\boldsymbol{X}} = 2(\boldsymbol{M}^{-1})^{\top}\text{diag}(\hat{\boldsymbol{x}})\boldsymbol{M}^{-1}. \tag{4.15}$$

Given a set of measurements, we can compute a residual vector for each $\boldsymbol{M} \in \mathcal{M}$.

The objective of the topology estimator is to find $\boldsymbol{M} \in \mathcal{M}$ such that a weighted sum of $\|\boldsymbol{\varepsilon}\|$ over several time instants is minimized. At time instant $k+1$, the topology estimation problem can be formulated as:

$$\hat{\boldsymbol{M}} = \underset{\boldsymbol{M}\in\mathcal{M}}{\arg\min}\, \epsilon_{\boldsymbol{M}}, \tag{4.16}$$

with

$$\epsilon_{\boldsymbol{M}} = \sum_{t\in\mathcal{K}} \eta^{k-t}\|\boldsymbol{\varepsilon}[t]\|, \tag{4.17}$$

where $\boldsymbol{\varepsilon}[t]$ is computed through (4.13) to (4.15). We refer to $\epsilon_{\boldsymbol{M}}$ as the *residual error* associated with topology configuration $\boldsymbol{M}$. Essentially, the topology estimator selects the topology under which the residual error $\epsilon_{\boldsymbol{M}}$ is minimized, where the line parameters are estimated by the parameter estimator. The intuition here is that different topology configurations will impose different structural constraints on voltage sensitivity matrices, which consequently impacts the residual error. The true topology configuration is expected to result in the least residual error. The voltage sensitivity estima-

Algorithm 4.1: Voltage Sensitivity Estimation

Input:

$\mathcal{M}$: set of feasible topology configurations

$\mathcal{Z}$: set of "r-to-x" ratio vectors

$\boldsymbol{p}[t], \boldsymbol{q}[t], \boldsymbol{v}[t]$: active power, reactive power, voltage magnitude measurements, $t \in \mathcal{K}$

Output:

$\hat{\boldsymbol{M}}$: estimated topology configuration

$\hat{\boldsymbol{r}}, \hat{\boldsymbol{x}}$: estimated line parameters

for $\boldsymbol{M} \in \mathcal{M}$, $\boldsymbol{\zeta} \in \mathcal{Z}$ **do**

Construct $\boldsymbol{\Psi}$ and $\boldsymbol{\psi}$ according to (4.8) and (4.9)

Compute pseudo-inverse of $\boldsymbol{\Psi}$, i.e., $\boldsymbol{\Psi}^{\dagger}$

Compute line parameters using

$$\hat{\boldsymbol{x}} = \boldsymbol{\Psi}^{\dagger}\boldsymbol{\psi},$$
$$\hat{\boldsymbol{r}} = \mathrm{diag}(\boldsymbol{\zeta})\hat{\boldsymbol{x}}$$

Compute voltage sensitivities using

$$\hat{\boldsymbol{R}} = 2(\boldsymbol{M}^{-1})^{\top}\mathrm{diag}(\hat{\boldsymbol{r}})\boldsymbol{M}^{-1}$$
$$\hat{\boldsymbol{X}} = 2(\boldsymbol{M}^{-1})^{\top}\mathrm{diag}(\hat{\boldsymbol{x}})\boldsymbol{M}^{-1}$$

Compute the residual error via

$$\epsilon_{\boldsymbol{M}} = \sum_{t \in \mathcal{K}} \eta^{k-t} \|\hat{\boldsymbol{R}}\boldsymbol{p}[t] + \hat{\boldsymbol{X}}\boldsymbol{q}[t] - \tilde{\boldsymbol{v}}[t]\|^2$$

end

Select topology configuration $\hat{\boldsymbol{M}}$ according to

$$\hat{\boldsymbol{M}} = \arg\min_{\boldsymbol{M} \in \mathcal{M}} \epsilon_{\boldsymbol{M}}$$

Select line parameters $\hat{\boldsymbol{r}}, \hat{\boldsymbol{x}}$ to be the ones associated with $\hat{\boldsymbol{M}}$

tion algorithm is summarized in Algorithm 4.1.

4.3.3 Voltage Controller

The voltage controller aims to determine the set-points for the DER active and reactive power injections while meeting all requirements discussed in Section 4.2.2. Note that for a given set of power injections, the resulting

voltage magnitude at each bus can be estimated using (4.1), where $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{X}}$ are used instead of $\boldsymbol{R}$ and $\boldsymbol{X}$. Define $\underline{\boldsymbol{v}} = [\underline{V}_1^2, \cdots, \underline{V}_N^2]^\top$ and $\overline{\boldsymbol{v}} = [\overline{V}_1^2, \cdots, \overline{V}_N^2]^\top$. Then, the voltage control problem can be formulated as the following convex program:

$$\underset{\boldsymbol{p}^g, \boldsymbol{q}^g}{\text{minimize}} \;\; c(\boldsymbol{p}^g, \boldsymbol{q}^g) \tag{4.18}$$

subject to

$$\boldsymbol{v} = \hat{\boldsymbol{R}}(\boldsymbol{C}\boldsymbol{p}^g - \boldsymbol{p}^d) + \hat{\boldsymbol{X}}(\boldsymbol{C}\boldsymbol{q}^g - \boldsymbol{q}^d) + v_0 \mathbf{1}_N, \tag{4.19}$$

$$\underline{\boldsymbol{p}}^g \leq \boldsymbol{p}^g \leq \overline{\boldsymbol{p}}^g, \tag{4.20}$$

$$\underline{\boldsymbol{q}}^g \leq \boldsymbol{q}^g \leq \overline{\boldsymbol{q}}^g, \tag{4.21}$$

with

$$\begin{aligned} c(\boldsymbol{q}^g) =& (\boldsymbol{p}^g)^\top \boldsymbol{W}^p \boldsymbol{p}^g + (\boldsymbol{q}^g)^\top \boldsymbol{W}^q \boldsymbol{q}^g \\ &+ \beta_1 \|[\underline{\boldsymbol{v}} - \boldsymbol{v}]_+\|^2 + \beta_2 \|[\boldsymbol{v} - \overline{\boldsymbol{v}}]_+\|^2, \end{aligned} \tag{4.22}$$

where $\boldsymbol{W}^p = \text{diag}(w_1^p, \cdots, w_n^p)$, $\boldsymbol{W}^q = \text{diag}(w_1^q, \cdots, w_n^q)$ are non-negative diagonal matrices, and the operator $[\cdot]_+$ returns its argument if the argument is positive and zero otherwise, and β_1 and β_2 are non-negative weights. The first two terms of $c(\cdot)$ are the costs associated with the active and reactive power injections, respectively, while the last two terms penalize the violation of constraint **C2**.

Constraint (4.19) is the LinDistFlow model, which is used to predict the voltage magnitudes for given power injections. Note that $\boldsymbol{p}^d$ and $\boldsymbol{q}^d$ are measured before solving the voltage control problem. Solving (4.18)–(4.21) gives the optimal set-points for the DER active and reactive power injections.

4.4 Voltage Sensitivity Identifiability

In this section, we first introduce the path matrix and then analyze the conditions under which the line parameters and correspondingly voltage sensitivities can be identified.

4.4.1 Path Matrix

Let $\mathcal{P}_i \subseteq \mathcal{L}$ denote the set of lines that form a path from bus 0—referred to as the root—to bus i. Since the power distribution system is radial, then $\mathcal{P}_i$ is unique (see Theorem 2.1.4 in [67]). Bus i is a leaf if for all $j \in \mathcal{N} \setminus \{i\}$, $(i,j) \notin \mathcal{E}$, i.e., there are no distribution lines starting at bus i. We say bus i is closer to the root than bus j if $|\mathcal{P}_i| < |\mathcal{P}_j|$, where $|\cdot|$ denotes the cardinality of a set. In defining $\boldsymbol{M}$, we choose the sending end of a line to be the bus that is closer to the root. Let $\boldsymbol{P} = [P_{\ell i}] \in \mathbb{R}^{L \times N}$ denote the path matrix of $\mathcal{G}$, with $P_{\ell i} = 1$ if line ℓ is on $\mathcal{P}_i$, and all other entries equal to zero. Under this setup, the relation between $\boldsymbol{P}$ and $\boldsymbol{M}$ is given by the following lemma.

Lemma 6. $\boldsymbol{P}$ and $\boldsymbol{M}$ are related as follows: $\boldsymbol{M}^{-1} = -\boldsymbol{P}$. (see also Theorem 2.10 in [68].)

Proof. Consider the entry at the i^{th} row and j^{th} column in $\boldsymbol{MP}$, which is $\sum_{\ell=1}^{L} M_{i\ell} P_{\ell j}$.

1. Consider first the case when $i = j$. If line ℓ is not connected to bus i, then $M_{i\ell} = 0$. If line ℓ starts at bus i, then $M_{i\ell} = 1$ and $P_{\ell i} = 0$ (because line ℓ is not on $\mathcal{P}_i$—the path from the root to bus i). If line ℓ ends at bus i, then $M_{i\ell} = -1$ and $P_{\ell i} = 1$. Obviously, there is one line that ends at bus i. Moreover, such a line is unique since otherwise there would be two paths from the root to bus i. Therefore, $\sum_{\ell=1}^{L} M_{i\ell} P_{\ell i} = M_{i\ell_i} P_{\ell_i i} = -1$, where line $\ell_i \in \mathcal{L}$ is the line that ends at bus i.

2. Next consider the case where $i \neq j$. Similar to the previous case, we only need to consider the lines that start from or end at bus i.

   (a) If line ℓ ends at bus i, then $M_{i\ell} = -1$. If $\ell \notin \mathcal{P}_j$, then $P_{\ell j} = 0$ and $M_{i\ell} P_{\ell j} = 0$. If $\ell \in \mathcal{P}_j$, then $P_{\ell j} = 1$. In the latter case, there must exist a unique line $\ell' \in \mathcal{P}_j$ that starts at bus i. Then $M_{i\ell} P_{\ell j} + M_{i\ell'} P_{\ell' j} = -1 + 1 = 0$. Therefore, $\sum_{\ell=1}^{L} M_{i\ell} P_{\ell j} = 0$.

   (b) If line ℓ starts at bus i, then $M_{i\ell} = 1$. If $\ell \notin \mathcal{P}_j$, then $P_{\ell j} = 0$ and $M_{i\ell} P_{\ell j} = 0$. If $\ell \in \mathcal{P}_j$, then $P_{\ell j} = 1$. In the latter case, there must exist a unique line $\ell' \in \mathcal{P}_j$ that ends at bus i. Similar to the previous argument, $\sum_{\ell=1}^{L} M_{i\ell} P_{\ell j} = 0$.

To summarize, $\sum_{\ell=1}^{L} M_{i\ell}P_{\ell j}$ equals to 1 if $i = j$ and 0 otherwise; therefore, $\boldsymbol{M}^{-1} = -\boldsymbol{P}$. □

The path matrix plays an important role in the identifiability analysis of the voltage sensitivities to be detailed next.

4.4.2 Identifiability Analysis

Before presenting the result on the identifiability of voltage sensitivities, we introduce the concept of downstream buses.

Definition 1. If line $\ell \in \mathcal{P}_i$, $\ell \in \mathcal{L}$, i.e., line ℓ is on the path from the root to bus i, then bus i is a downstream bus of line ℓ. The set of downstream buses of line ℓ is denoted by $\mathcal{N}_\ell$.

As discussed in Section 4.3.2, $\boldsymbol{\Psi}[k]$ needs to have a full rank, i.e., rank($\boldsymbol{\Psi}[k]$) equals to L, in order to estimate $\boldsymbol{x}$ according to (4.11). When $\boldsymbol{\Psi}$ does not have a full rank, some of the line parameters cannot be estimated from the measurements. The main results for the voltage sensitivity identifiability is stated as follows:

Theorem 3. The parameters of line $\ell \in \mathcal{L}$ are identifiable if and only if the following condition is satisfied for some $t \in \mathcal{K}$:

$$\sum_{i \in \mathcal{N}_\ell} \zeta_\ell p_i[t] + q_i[t] \neq 0. \tag{4.23}$$

Proof. Using the path matrix, $\boldsymbol{\Gamma}_\ell$ can be written as $\boldsymbol{\Gamma}_\ell = 2\boldsymbol{P}^\top \boldsymbol{e}_\ell \boldsymbol{e}_\ell^\top \boldsymbol{P}$. Note that $\boldsymbol{P}^\top \boldsymbol{e}_\ell$ is the ℓ^{th} column of $\boldsymbol{P}^\top$ and $\boldsymbol{\Gamma}_\ell$ is a rank-one matrix. Let $\boldsymbol{\pi}_\ell = \boldsymbol{P}^\top \boldsymbol{e}_\ell$, then $\boldsymbol{P}^\top = [\boldsymbol{\pi}_1, \cdots, \boldsymbol{\pi}_L]$. Thus,

$$\begin{aligned} \boldsymbol{\Gamma}_\ell &= 2\boldsymbol{P}^\top \boldsymbol{e}_\ell \boldsymbol{e}_\ell^\top \boldsymbol{P} \\ &= 2\boldsymbol{\pi}_\ell \boldsymbol{\pi}_\ell^\top, \end{aligned} \tag{4.24}$$

and $\boldsymbol{\Psi}[k]$ can be written as

$$\boldsymbol{\Psi}[k] = 2 \begin{bmatrix} \boldsymbol{\pi}_1 \boldsymbol{\pi}_1^\top \boldsymbol{\rho}_1[k-m] & \cdots & \boldsymbol{\pi}_L \boldsymbol{\pi}_L^\top \boldsymbol{\rho}_L[k-m] \\ \vdots & \vdots & \vdots \\ \boldsymbol{\pi}_1 \boldsymbol{\pi}_1^\top \boldsymbol{\rho}_1[k] & \cdots & \boldsymbol{\pi}_L \boldsymbol{\pi}_L^\top \boldsymbol{\rho}_L[k] \end{bmatrix}. \tag{4.25}$$

Let $\mathcal{L} = \mathcal{L}_0 \cup \mathcal{L}_1$, where $\mathcal{L}_1$ and $\mathcal{L}_0$ are the index sets of lines that meet and do not meet the conditions in (4.23), respectively. Without loss of generality, the lines can be re-labeled so that $\mathcal{L}_0$ corresponds to the left columns of $\mathbf{\Psi}$ and $\mathcal{L}_1$ to the right columns of $\mathbf{\Psi}$. If line $\ell \in \mathcal{L}_0$, then $\forall t \in \mathcal{K}$,

$$\sum_{i \in \mathcal{N}_\ell} \zeta_\ell p_i[t] + q_i[t] = 0. \tag{4.26}$$

Note that the i^{th} entry in $\boldsymbol{\pi}_\ell$ is 1 if and only if bus i is a downstream bus of line ℓ. Essentially, the non-zero entries in $\boldsymbol{\pi}_\ell$, which are ones, indicate the downstream buses of line ℓ. Therefore, it follows from (4.26) that, $\forall t \in \mathcal{K}$:

$$\boldsymbol{\pi}_\ell^\top \boldsymbol{\rho}_\ell[t] = 0. \tag{4.27}$$

Consequently, all entries in the ℓ^{th} column of $\mathbf{\Psi}[k]$ are zero, and the value of x_ℓ does not affect the objective function in (4.10). Under such condition, x_ℓ cannot be identified. For line $\ell \in \mathcal{L}_0$, we can remove the ℓ^{th} column of $\mathbf{\Psi}[k]$, the ℓ^{th} entry of $\boldsymbol{x}$ and $\boldsymbol{\psi}[k]$ to obtain a reduced-size estimation problem.

Next we show that the line parameter can be identified as long as condition (4.23) is satisfied. Without loss of generality, we assume $\mathcal{L}_1 = \mathcal{L}$ since otherwise we can remove the left columns in $\mathbf{\Psi}$ that correspond to $\mathcal{L}_0$ to obtain a reduced problem. Then, (4.27) is satisfied for all $\ell \in \mathcal{L}$ and for some $t \in \mathcal{K}$. Assume $\text{rank}(\mathbf{\Psi}[k]) < L$, then there exist $a_1, \cdots, a_L \in \mathbb{R}$, which are not all zero, such that

$$\mathbf{\Psi}[k][a_1, \cdots, a_L]^\top = \mathbf{0}_{mL}, \tag{4.28}$$

where $\mathbf{0}_{mL}$ is an mL-dimensional all-zeros vector. Without loss of generality, assume $a_1, \cdots, a_{L'}$ are not zero, while $a_{L'+1}, \cdots, a_L$ are all zero, where $1 < L' \leq L$. Then, it follows from (4.25) and (4.28) that

$$\sum_{l=1}^{L'} a_l \boldsymbol{\pi}_l^\top \boldsymbol{\rho}_l[t] \boldsymbol{\pi}_l = \mathbf{0}_L. \tag{4.29}$$

Since $\boldsymbol{\pi}_1, \cdots, \boldsymbol{\pi}_{L'}$ are linear independent, then $a_l \boldsymbol{\pi}_l^\top \boldsymbol{\rho}_l[t] = 0$ for $\ell = 1, \cdots, L'$. However, since for any $\ell \in \mathcal{L}$ there exists some $t \in \mathcal{K}$ such that $\boldsymbol{\pi}_\ell^\top \boldsymbol{\rho}_\ell[t] = 0$, then $a_\ell = 0$ for $\ell = 1, \cdots, L'$, which leads to a contradiction. Therefore,

$\text{rank}(\boldsymbol{\Psi}[k]) = L$ and the line parameters can be identified. $\square$

The voltage sensitivity matrices can be readily computed if all line parameters can be identified. If some line parameter cannot be identified, the resulting voltage sensitivity matrices may not be accurate. This, however, will not have any impact on the approximate relation between voltage magnitudes and the active and reactive power injections in (4.19) since in such cases the line parameter does not affect the voltage magnitudes anyway. Specifically, if follows from (4.1), (4.6), and (4.7) that

$$\begin{aligned}\tilde{\boldsymbol{v}}[t] &= \boldsymbol{R}\boldsymbol{p}[t] + \boldsymbol{X}\boldsymbol{q}[t] \\ &= 2\sum_{\ell=1}^{L} \boldsymbol{\pi}_\ell \boldsymbol{\pi}_\ell^\top \boldsymbol{\rho}[t] x_\ell, \end{aligned} \tag{4.30}$$

in which $\boldsymbol{\pi}_\ell^\top \boldsymbol{\rho}[t] = 0$ if x_ℓ cannot be identified. Therefore, for the purpose of solving the optimization problem in (4.18), the proposed voltage sensitivity estimation algorithm is still effective.

If we think of $\zeta_\ell p_i[t] + q_i[t]$ as some "combined power" (in the sense that it is a combination of active and reactive power), then (4.23) essentially indicates that the sum of combined power injection at all downstream buses of line ℓ is nonzero, or equivalently, there exists some combined power flow on line ℓ. For any line for which the receiving end is a leaf, its parameter can be identified as long as the combined power injection at the receiving end is nonzero.

4.5 Numerical Simulation

In this section, we validate the effectiveness of the proposed framework using a modified three-phase balanced IEEE 123-bus distribution test feeder from [65], the topology of which is shown in Fig. 4.2. There are six switches, four of which are normally closed while the other two are open so as to ensure the system maintains a radial topology at all times. Under Assumption **A2**, this feeder has nine possible topology configurations as listed in Table 4.1, among which configuration 0 is the nominal one.

The loads are constructed based on historical hourly active power load data from a residential building in San Diego [69]. Specifically, the historical

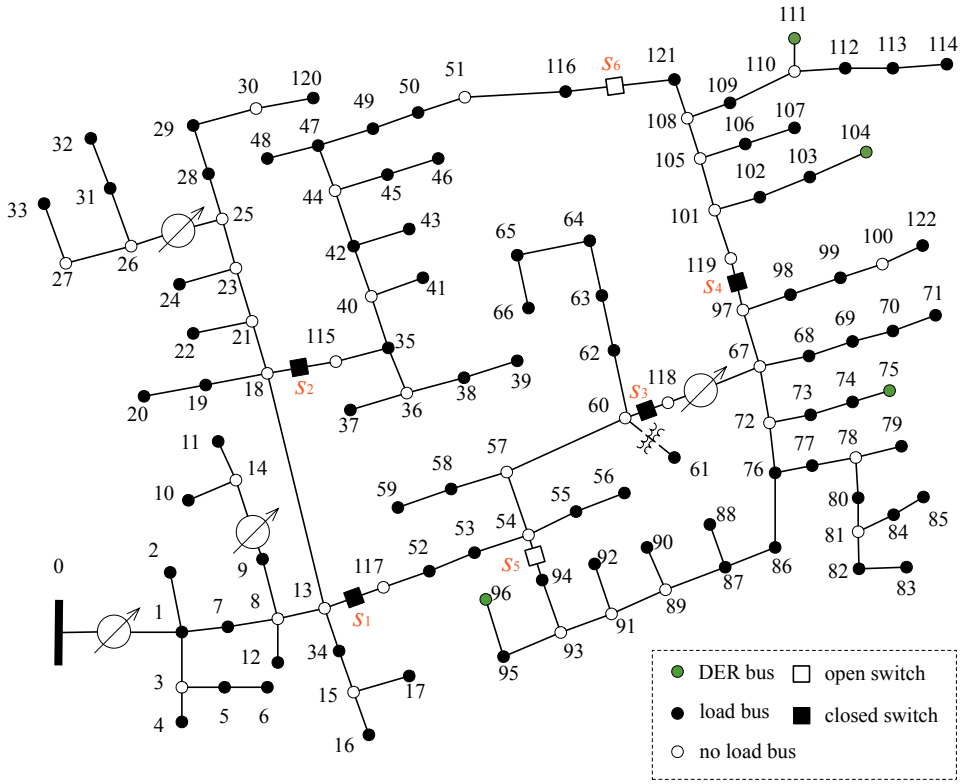

Figure 4.2: IEEE 123-bus distribution test feeder.

Table 4.1: Switch Status Under Feasible Topology Configurations

config.	s_1	s_2	s_3	s_4	s_5	s_6
0	closed	closed	closed	closed	open	open
1	closed	closed	open	closed	closed	open
2	closed	closed	closed	open	open	closed
3	closed	closed	open	closed	open	closed
4	closed	open	closed	closed	open	closed
5	open	closed	closed	closed	open	closed
6	closed	closed	open	open	closed	closed
7	closed	open	open	closed	closed	closed
8	open	closed	open	closed	closed	closed

hourly active power load data are first normalized such that the maximum active load becomes 1. Then, the time granularity of the normalized active power load is increased to 1 second through a linear interpolation. Let $d[k]$ denote the k^{th} value in the normalized 1-second active power load time series. Each value in the resulting normalized 1-second system total active power load data time series is further multiplied by a normally distributed variable,

the mean and standard deviation of which is 1 and 0.01, respectively. Then, the active and reactive power demanded by load i is constructed as follows:

$$p_i^d[k] = p_i^{d0} d[k](1 + 0.01\mu^p[k]),$$
$$q_i^d[k] = q_i^{d0} d[k](1 + 0.01\mu^q[k]),$$

where p_i^{d0} and q_i^{d0} are the nominal active and reactive power demanded by load i, μ^p and μ^q are two standard Gaussian random variables.

Four DERs are added at buses 76, 97, 105, 112, respectively, with reactive power outputs taking values in the interval $[-200, 200]$ kVAr. We set the weights in (4.22) to $w_i^p = 1 + 0.1i$ and $w_i^q = 1 + 0.1i$, for $i \in \mathcal{N}^g$. For simplicity, the active power outputs from DERs are fixed at zero; yet, the proposed methodology can be directly applied to cases in which the active power outputs from DERs are nonzero. The minimum and maximum voltage magnitudes are 0.95 p.u. and 1.05 p.u., respectively. In addition, in (4.22), we set $\beta_1 = \beta_2 = 1 \times 10^5$. Unless otherwise specified, the forgetting factor η is set to 1, and the underlying topology configuration is configuration 0, i.e., the nominal one. While the LinDistFlow model was adopted for the analysis, in the simulations, we use a full nonlinear power flow model and solve it using Matpower [56].

4.5.1 Estimation Accuracy

Throughout this part, the DERs do not inject any reactive power into the power distribution system.

Noise-free case

We first evaluate the accuracy of the proposed estimation algorithm in the case where the measurements are noise-free. The algorithm is evaluated in 100 Monte Carlo simulation runs under various loading conditions. In each simulation run, 10 sets of measurements are used to compute the residual error. Residual errors are computed for each feasible topology configuration in $\mathcal{M}$, while the underlying true topology configuration is one of them. Residual errors associated with topology configurations $0-3$ and 6, when the underlying topology configuration is configuration 6, are shown in Fig. 4.3.

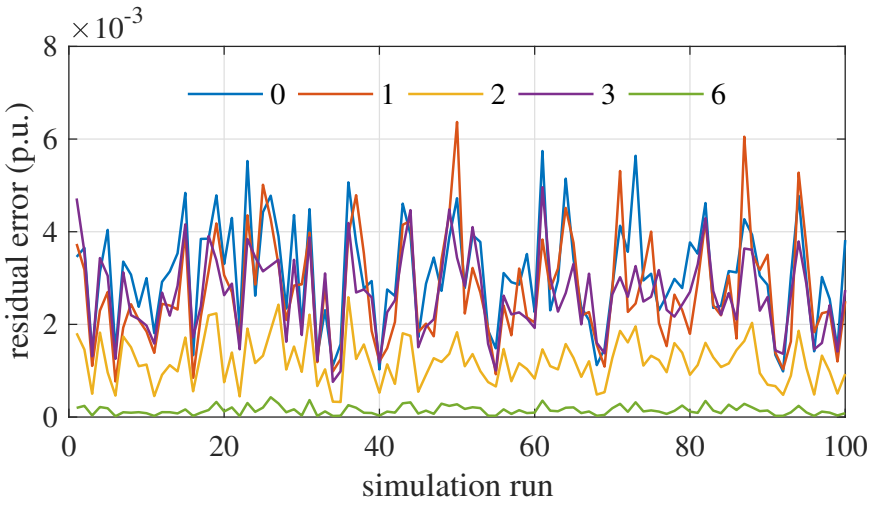

Figure 4.3: Residual errors under different topology configurations with 10 sets of noise-free measurements. (Legends indicate topology configurations.)

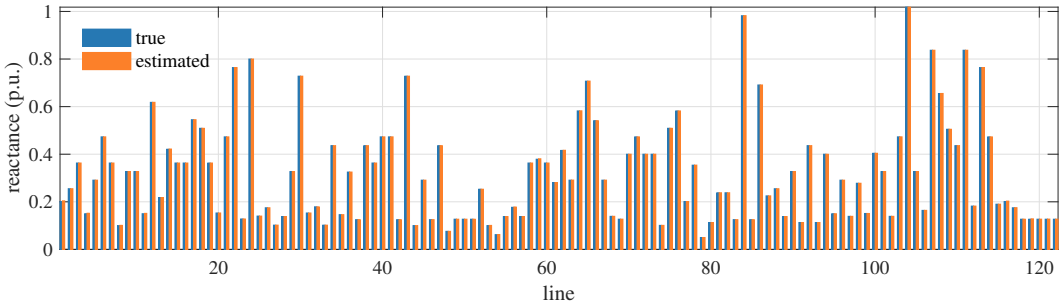

Figure 4.4: Estimated and true values of line reactances using 10 sets of noise-free measurements.

Residual errors associated with topology configurations $4, 5, 7, 8$ are at least one order of magnitude larger than those of the other configurations, and are hence not plotted. Note that the case with the underlying topology configuration being configuration 6 is the one where the residual error differences between topology configurations are the smallest. Yet, it is still obvious that the true topology configuration results in the minimum residual error, which is one order of magnitude smaller than those of other configurations.

The parameter estimation accuracy is evaluated using the mean absolute percentage error (MAPE), given by $\frac{1}{L}\sum_{\ell=1}^{L}|\hat{x}_\ell/x_\ell - 1|$ for $\hat{\boldsymbol{x}}$, and given by $\frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L}|\hat{X}_{ij}/X_{ij} - 1|$ for $\hat{\boldsymbol{X}}$. Figure 4.4 shows the estimated line reactances in one simulation run and one set of measurements is utilized. When one set of measurements is utilized, a typical MAPE is 0.11% for $\hat{\boldsymbol{x}}$ and 1.16% for $\hat{\boldsymbol{X}}$, both of which are really small. We note that the loading conditions

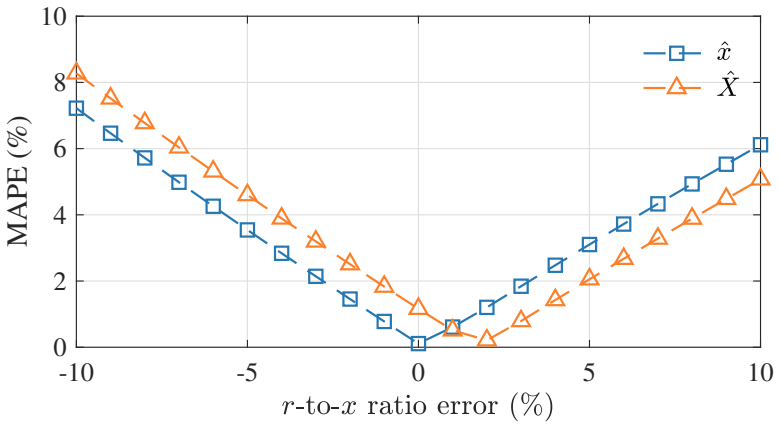

Figure 4.5: Impacts of errors in r-to-x ratios on parameter estimation accuracy in the noise-free case.

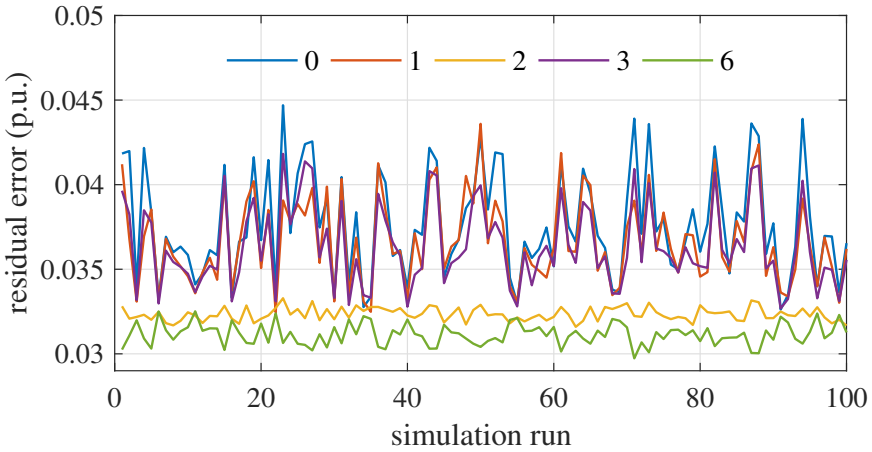

Figure 4.6: Residual errors under different topology configurations with 60 sets of noisy measurements. (Legends indicate topology configurations.)

of the power distribution system do not affect the accuracy of the proposed algorithm. The r-to-x ratios of all lines are assumed to be known. Figure 4.5 shows that the MAPE is almost linear with respect to the r-to-x ratio errors. Therefore, relatively small error in the r-to-x ratios will not result in a significant increase in the parameter estimator error.

Noisy case

To see the impacts of measurement noise, we add a white Gaussian noise to all measurements such that the signal-to-noise ratio (SNR) is 92 dB, as adopted in [70]. More measurements across time are required to obtain a good estimation accuracy in the presence of measurement noise. The algo-

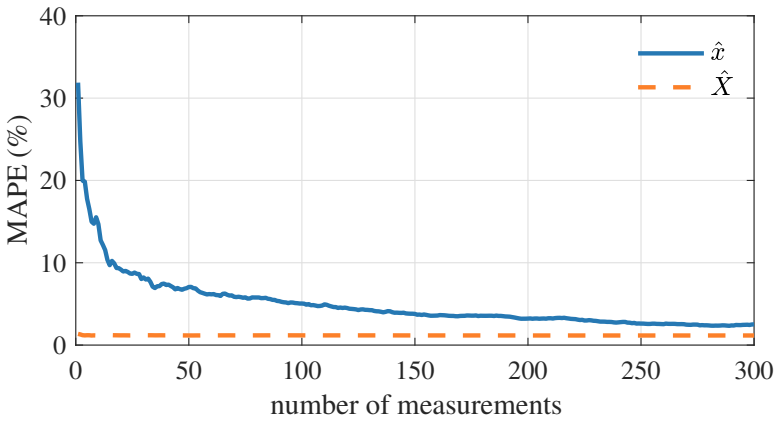

Figure 4.7: Impacts of measurement numbers on parameter estimation accuracy in the noisy case.

rithm is again evaluated in 100 Monte Carlo runs under the same setup as the noise-free case, except that 60 sets of measurements—corresponding to measurements collected in 1 minute—are used to compute the residual error. Residual errors associated with configurations $0-3$ and 6 are shown in Fig. 4.6. Note that residual errors associated with topology configurations $4, 5, 7, 8$ are one order magnitude larger than those of the other configurations, and are hence not plotted. The true topology configuration, i.e., configuration 6, still results in the minimum residual error. We note that increasing the number of measurements across time generally leads to a higher accuracy in identifying the topology configuration.

The number of measurements has a direct impact on the estimation accuracy. As shown in Fig. 4.7, the MAPE of $\hat{\boldsymbol{x}}$ drops quickly when increasing the number of measurements, approximately from 31.9% with one set of measurements to 2.51% when 300 sets of measurements are used. The MAPE of $\hat{\boldsymbol{X}}$—which is what really matters—is relatively insensitive to the number of measurements, being around 1.17%. Indeed, this result illustrates the effectiveness of the proposed estimation algorithm.

Figure 4.8 shows the impacts of SNR on parameter estimation accuracy when 300 sets of measurements are used. When the SNR is beyond 50 dB, the MAPE of the voltage sensitivity matrix is within 3.6%, which is relatively small. In the rest of the simulation, we assume a SNR of 92 dB for all measurements.

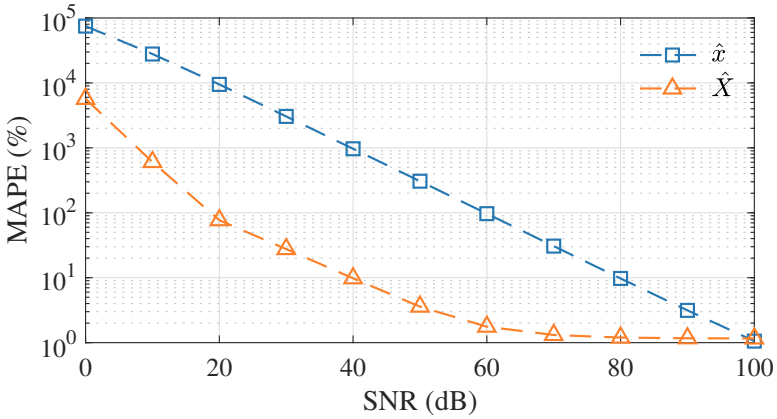

Figure 4.8: Impacts of SNR on parameter estimation accuracy when 300 sets of measurements are used.

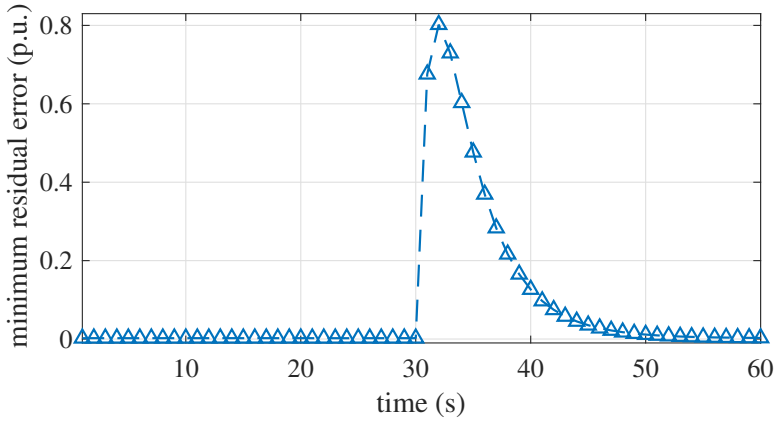

Figure 4.9: Minimum residual error under topology reconfiguration.

Accuracy under topology reconfiguration

The proposed algorithm works well not only under a fixed topology configuration but also when topology reconfiguration occurs. To illustrate this, we simulate a case where the underlying topology configuration is changed from configuration 0 to configuration 3 at 31 s. A total of 60 sets of measurements are used to compute the voltage sensitivities, i.e., $|\mathcal{K}| = 60$. The forgetting factor η is set to 0.6. The minimum residual error and the corresponding estimated topology configuration are shown in Figs. 4.9 and 4.10, respectively. A jump in the minimum residual error is observed when the topology is reconfigured. The new topology is successfully identified after 6 s. Correspondingly, the MAPE of $\hat{\boldsymbol{X}}$ is also reduced to less than 2% after 6 s, as shown in Fig. 4.11.

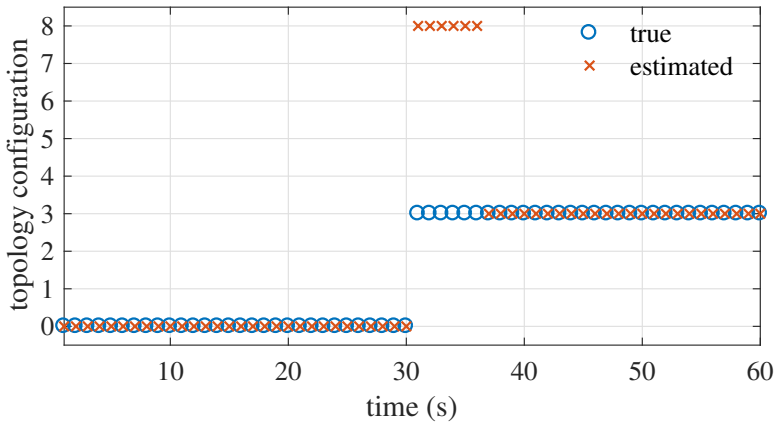

Figure 4.10: Estimated topology configuration under topology reconfiguration.

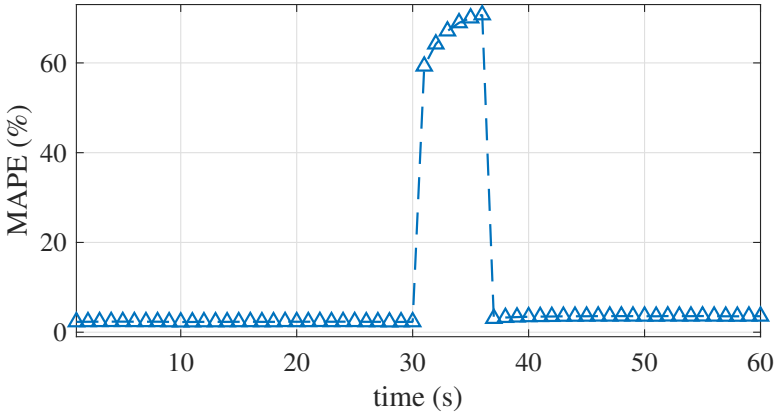

Figure 4.11: MAPE of $\hat{\boldsymbol{X}}$ under topology reconfiguration.

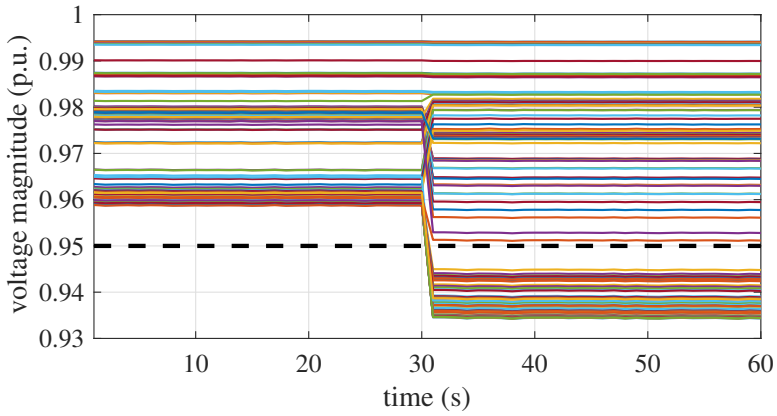

Figure 4.12: Voltage profiles with model-based voltage regulation scheme under topology reconfiguration.

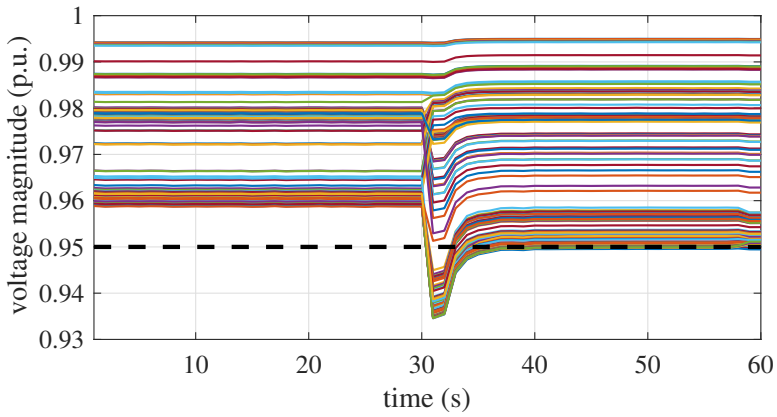

Figure 4.13: Voltage profiles with proposed voltage regulation scheme under topology reconfiguration.

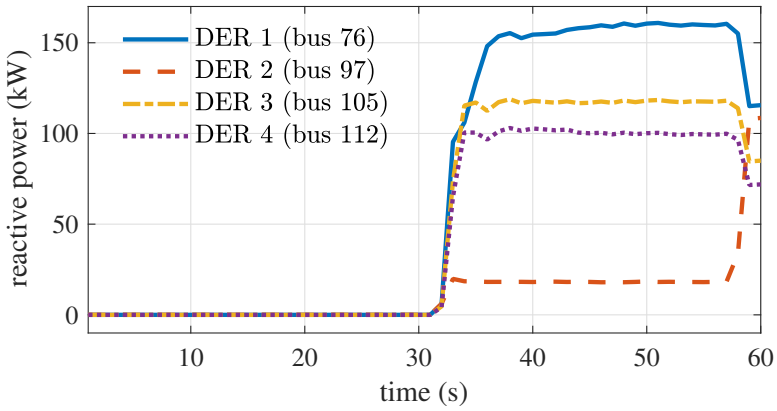

Figure 4.14: DER reactive power injections with proposed voltage regulation under topology reconfiguration.

4.5.2 Voltage Control Performance

Next, we show the performance of the voltage regulation framework proposed in Section 4.3 in the same case as the one in the previous section with topology reconfiguration, where the underlying topology configuration is changed from configuration 0 to configuration 3 at 31 s. A model-based voltage regulation scheme, which uses the voltage sensitivity matrices in (4.19) computed from the LinDistFlow model but is not aware of the topology reconfiguration, is used for the purpose of comparison. The voltage profiles with the model-based and the proposed voltage regulation schemes are presented in Figs. 4.12 and 4.13, respectively, and the DER reactive power injections are shown in Fig. 4.14. It is clear from the results that the proposed data-driven voltage regulation framework is very effective and efficient in restoring the voltage

magnitudes to the desirable range. This illustrates the adaptivity of our voltage regulation framework to system condition changes such as topology reconfiguration.

4.6 Summary

In this chapter, we have proposed a data-driven voltage regulation framework for DERs in a balanced radial power distribution system. This framework utilizes the LinDistFlow model that approximates the nonlinear relation between the voltage magnitudes and power injections, and estimates its parameters indirectly by estimating the topology configuration and the corresponding line parameters. In particular, by exploiting the structural characteristics of the power distribution system, the proposed estimation algorithm for the voltage sensitivities requires much less data than existing algorithms. Using the estimated voltage sensitivities, the optimal DER power injections can be readily determined by solving a convex optimization problem.

Our theoretical analysis shows that the voltage sensitivities of interest are easily identifiable. The inherent data-driven nature of the framework makes it adaptive to changes in system operational conditions, such as topology reconfigurations. Numerical simulations demonstrated that the voltage sensitivities can be estimated accurately using a few sets of measurements even under topology reconfiguration, consequently guaranteeing good voltage regulation performance.

The work presented in this chapter was published in [71] and submitted for publication in [72].

CHAPTER 5

OPTIMAL TAP SETTING OF LOAD TAP CHANGERS

5.1 Introduction

In this chapter, we develop an algorithm that can find a policy for determining the optimal tap positions of the LTCs in a power distribution system under uncertain load dynamics without any information on power injections or line parameters; the algorithm requires only voltage magnitude measurements and system topology information. Specifically, the optimal tap setting problem is cast as an MDP, which can be solved using RL algorithms. Yet, adequate state and action samples that sufficiently explore the MDP state and action spaces are needed. However, it is hard to obtain such samples in real power systems since this requires changing tap settings and other controls to excite the system and record voltage responses, which may jeopardize system operational reliability and incur economic costs. To circumvent this issue, we take advantage of a linearized power flow model and develop an effective algorithm to estimate voltage magnitudes under different tap settings so that the state and action spaces can be explored freely offline without impacting the real system.

The dimension of the state and action spaces increases exponentially as the number of LTCs grows, which causes the issue known as the "curse of dimensionality" and makes the computation of the optimal policy intractable [73]. To circumvent the "curse of dimensionality," we propose an efficient batch RL algorithm—the least squares policy iteration (LSPI) based sequential learning algorithm—to learn an action-value function sequentially for each LTC. Once the learning of the action-value function is completed, we can determine the policy for optimally setting the LTC taps. We emphasize that the optimal policy can be computed offline, where most computational burden takes place. However, when executed online, the required computation

to find the optimal tap positions is minimal. The effectiveness of the proposed algorithm is validated through simulations on two IEEE distribution test feeders.

The remainder of the chapter is organized as follows. A modified version of the LinDistFlow model that includes the effect of LTCs, and the optimal tap setting problem, are introduced in Section 5.2. A primer on MDPs and the LSPI algorithm is provided in Section 5.3. An MDP-based formulation for the optimal tap setting problem is presented in Section 5.4 and an algorithm to solve this problem is proposed in Section 5.5. Numerical simulation results on two IEEE test feeders are presented in Section 5.6. We summarize this chapter in Section 5.7.

5.2 Preliminaries

In this section, we modify the LinDistFlow (see Section 1.2.1) to include the effect of LTCs.

5.2.1 Power Flow Model

The LinDistFlow model in (1.17) can be rewritten as follows:

$$\boldsymbol{M}^\top \boldsymbol{v} + \boldsymbol{m} v_0 = 2\mathrm{diag}(\boldsymbol{r})\boldsymbol{M}^{-1}\boldsymbol{p} + 2\mathrm{diag}(\boldsymbol{x})\boldsymbol{M}^{-1}\boldsymbol{q}, \tag{5.1}$$

where $\boldsymbol{M}$ is the reduced node-to-edge incidence matrix, $\boldsymbol{p}$ and $\boldsymbol{q}$ are vectors of active and reactive power injections, respectively, $\boldsymbol{r}$ and $\boldsymbol{x}$ are vectors of line resistances and reactances, respectively, $\boldsymbol{v}$ is a vector of squared voltage magnitudes for buses 1 to N, and v_0 is the squared voltage magnitude for bus 0.

The standard model for an LTC in the literature is shown in Fig. 5.1 (see, e.g., [14]), where $\mathrm{i} = \sqrt{-1}$, line ℓ is associated with (i, j), and t_ℓ is the tap ratio of the LTC on line ℓ. Typically, the tap ratio can possibly take on 33 discrete values ranging from 0.9 to 1.1, by an increment of 5/8% p.u., i.e., $t_\ell \in \mathcal{T} = \{0.9, 0.90625, \cdots, 1.09375, 1.1\}$ [14]. Let $\Delta t_\ell \in \Delta\mathcal{T} = \{0, \pm 0.00625, \cdots, \pm 0.19375, \pm 0.2\}$ denote a feasible LTC tap ratio change. We index the 33 tap positions by $-16, \cdots, -1, 0, 1, \cdots, 16$ for convenience.

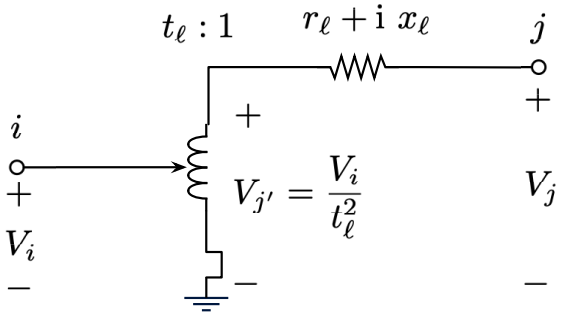

Figure 5.1: Load tap changer model.

Let $\mathcal{L}^t$ denote the set of lines with LTCs and let $|\mathcal{L}^t| = L^t$, where $|\cdot|$ denotes the cardinality of a set. For line ℓ that is associated with (i, j), if $\ell \in \mathcal{L}^t$, the voltage relation in the LinDistFlow model needs to be modified as follows:

$$\frac{1}{t_\ell^2} v_i - v_j = 2(r_\ell p_{ij} + x_\ell q_{ij}). \tag{5.2}$$

Define $\boldsymbol{t} = [t_\ell]^\top$ and $\Delta \boldsymbol{t} = [\Delta t_\ell]^\top$, $\ell \in \mathcal{L}^t$. Let $\tilde{\boldsymbol{M}}(\boldsymbol{t}) = [\tilde{M}_{i\ell}(\boldsymbol{t})] \in \mathbb{R}^{(N+1)\times L}$, with $\tilde{M}_{i\ell}(\boldsymbol{t}) = 1$ and $\tilde{M}_{j\ell}(\boldsymbol{t}) = -1$ if line $\ell \in \mathcal{L} \setminus \mathcal{L}^t$, $\tilde{M}_{i\ell}(\boldsymbol{t}) = \frac{1}{t_\ell^2}$ and $\tilde{M}_{j\ell}(\boldsymbol{t}) = -1$ if line $\ell \in \mathcal{L}^t$, and all other entries equal to zero. Let $\boldsymbol{m}(\boldsymbol{t})^\top$ denote the first row of $\tilde{\boldsymbol{M}}(\boldsymbol{t})$ and $\boldsymbol{M}(\boldsymbol{t})$ the matrix that results by removing $\boldsymbol{m}(\boldsymbol{t})^\top$ from $\tilde{\boldsymbol{M}}(\boldsymbol{t})$. The matrix $\boldsymbol{M}(\boldsymbol{t})$ is non-singular when the power distribution system is connected. Then, the modified matrix-form LinDistFlow model that takes into account the LTCs is given by:

$$\boldsymbol{M}(\boldsymbol{t})^\top \boldsymbol{v} + \boldsymbol{m}(\boldsymbol{t}) v_0 = 2\mathrm{diag}(\boldsymbol{r}) \boldsymbol{M}^{-1} \boldsymbol{p} + 2\mathrm{diag}(\boldsymbol{x}) \boldsymbol{M}^{-1} \boldsymbol{q}. \tag{5.3}$$

5.2.2 Optimal Tap Setting Problem

To effectively regulate the voltages in a power distribution system, the tap positions of LTCs need to be set appropriately. The objective of the optimal tap setting problem is to find a policy $\boldsymbol{\pi}$ that determines the LTC tap ratio so as to minimize the voltage deviation from some reference value, denoted by $\boldsymbol{v}^\star$, based on current tap ratios and measurements of the voltage magnitudes, i.e., $\boldsymbol{\pi} : (\boldsymbol{t}, \boldsymbol{v}) \to \Delta \boldsymbol{t},\ \boldsymbol{t} \in \mathcal{T}^{L^t},\ \boldsymbol{v} \in \mathbb{R}^N, \Delta \boldsymbol{t} \in \Delta \mathcal{T}^{L^t}$. Throughout this chapter, we make the following two assumptions:

A1. The system topology is known but line parameters are unknown.

A2. The active and reactive power injections are not measured and their probability distributions are unknown.

5.3 Markov Decision Process and Batch Reinforcement Learning

In this section, we provide some background on MDPs and the batch RL algorithm, a type of data-efficient and stable algorithm for solving MDPs with unknown models.

5.3.1 Markov Decision Process

An MDP is defined as a 5-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where $\mathcal{S}$ is a finite set of states; $\mathcal{A}$ is a finite set of actions; $\mathcal{P}$ is a Markovian transition model that denotes the probability of transitioning from one state into another after taking an action; $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ is a reward function such that, for $\boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}$ and $\boldsymbol{a} \in \mathcal{A}$, $r = \mathcal{R}(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{s}')$ is the reward obtained when the system transitions from state $\boldsymbol{s}$ into state $\boldsymbol{s}'$ after taking action $\boldsymbol{a}$; and $\gamma \in [0, 1)$ is a discount factor (see, e.g., [74]).[1] We refer to the 4-tuple $(\boldsymbol{s}, \boldsymbol{a}, r, \boldsymbol{s}')$, where $\boldsymbol{s}'$ is the state following $\boldsymbol{s}$ after taking action $\boldsymbol{a}$ and $r = \mathcal{R}(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{s}')$, as a transition.

Let $\boldsymbol{S}[k]$ and $\boldsymbol{A}[k]$ denote the state and action at time instant k, respectively, and $R[k]$ the reward received after taking action $\boldsymbol{A}[k]$ in state $\boldsymbol{S}[k]$. Then, $\mathcal{P}_k(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a}) := \mathbb{P}\{\boldsymbol{S}_{k+1} = \boldsymbol{s}'|\boldsymbol{S}[k] = \boldsymbol{s}, \boldsymbol{A}[k] = \boldsymbol{a}\}$ is the probability of transitioning from state $\boldsymbol{s}$ into state $\boldsymbol{s}'$ after taking action $\boldsymbol{a}$ at time instant k. Throughout this chapter, we assume time-homogeneous transition probabilities, hence we drop the subindex k and just write $\mathcal{P}(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a})$.

Let $\bar{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ denote the expected reward for a state-action pair $(\boldsymbol{s}, \boldsymbol{a})$; then, we have

$$\bar{R}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}[R] = \sum_{\boldsymbol{s}' \in \mathcal{S}} \mathcal{R}(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{s}')\mathcal{P}(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a}), \tag{5.4}$$

[1] These definitions can be directly extended to the case where the the set of states is infinite. Due to space limitation, this case is not discussed in detail here.

where $\mathbb{E}[\cdot]$ denotes the expectation operation. The total discounted reward from time instant k and onwards, denoted by $G[k]$, also referred to as the return, is given by

$$G[k] = \sum_{i=k}^{\infty} \gamma^{i-k} R[i]. \tag{5.5}$$

A deterministic policy $\boldsymbol{\pi}$ is a mapping from $\mathcal{S}$ to $\mathcal{A}$, i.e., $\boldsymbol{a} = \boldsymbol{\pi}(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}, \boldsymbol{a} \in \mathcal{A}$. The action-value function under policy $\boldsymbol{\pi}$ is defined as follows:

$$Q^{\boldsymbol{\pi}}(\boldsymbol{s}, \boldsymbol{a}) = \mathbb{E}\left[G[k] | \boldsymbol{S}[k] = \boldsymbol{s}, \boldsymbol{A}[k] = \boldsymbol{a}; \boldsymbol{\pi}\right], \tag{5.6}$$

which is the expected return when taking action $\boldsymbol{a}$ in state $\boldsymbol{s}$, and following policy $\boldsymbol{\pi}$ afterwards. Intuitively, the action-value function quantifies, for a given policy π, how "good" the state-action pair $(\boldsymbol{s}, \boldsymbol{a})$ is in the long run.

Let $Q^*(\cdot, \cdot)$ denote the optimal action-value function—the maximum action-value function over all policies, i.e., $Q^*(\boldsymbol{s}, \boldsymbol{a}) = \max_{\boldsymbol{\pi}} Q^{\boldsymbol{\pi}}(\boldsymbol{s}, \boldsymbol{a})$. All optimal policies share the same optimal action-value function. Also, the greedy policy with respect to $Q^*(\boldsymbol{s}, \boldsymbol{a})$, i.e., $\boldsymbol{\pi}^*(\boldsymbol{s}) = \arg\max_{\boldsymbol{a}} Q^*(\boldsymbol{s}, \boldsymbol{a})$ is an optimal policy. Then, it follows from (5.5) and (5.6) that $Q^*(\boldsymbol{s}, \boldsymbol{a})$ satisfies the following Bellman optimality equation (see, e.g., [73]):

$$Q^*(\boldsymbol{s}, \boldsymbol{a}) = \bar{R}(\boldsymbol{s}, \boldsymbol{a}) + \gamma \sum_{\boldsymbol{s}' \in \mathcal{S}} \mathcal{P}(\boldsymbol{s}' | \boldsymbol{s}, \boldsymbol{a}) \max_{\boldsymbol{a}' \in \mathcal{A}} Q^*(\boldsymbol{s}', \boldsymbol{a}'). \tag{5.7}$$

The MDP is solved if we find $Q^*(\boldsymbol{s}, \boldsymbol{a})$, and correspondingly, the optimal policy $\boldsymbol{\pi}^*$. It is important to emphasize that (5.7) is key in solving the MDP. For ease of notation, in the rest of this chapter, we simply write the $Q^*(\boldsymbol{s}, \boldsymbol{a})$ as $Q(\boldsymbol{s}, \boldsymbol{a})$.

When both the state and the action sets are finite, the action-value function can be exactly represented in a tabular form that covers all possible pairs $(\boldsymbol{s}, \boldsymbol{a}) \in \mathcal{S} \times \mathcal{A}$. In this case, if $\mathcal{P}$ is also known, then the MDP can be solved using, e.g., the so-called policy iteration and value iteration algorithms (see, e.g., [73]). If $\mathcal{P}$ is unknown but samples of transitions are available, the MDP can be solved by using RL algorithms such as the Q-learning algorithm (see, e.g., [75]).

5.3.2 Batch Reinforcement Learning

When $\mathcal{S}$ is not finite, conventional Q-learning based approaches require discretization of $\mathcal{S}$ (see, e.g., [39] and [76]). The discretized state space will better approximate the original state space if a small step size is used in the discretization process, yet the resulting MDP will face the "curse of dimensionality." A large step size can alleviate the computational burden caused by the high dimensionality of the state space, but at the cost of potentially degrading performance significantly.

More practically, when the number of elements in $\mathcal{S}$ is large or $\mathcal{S}$ is not finite, the action-value function can be approximated by some parametric functions such as linear functions [74] and neural networks [77], or non-parametric functions such as decision trees [78]. Let $\hat{Q}(\cdot, \cdot)$ denote the approximate optimal action-value function. Using a linear function approximation, $\hat{Q}(\boldsymbol{s}, \boldsymbol{a})$ can be represented as follows:

$$\hat{Q}(\boldsymbol{s}, \boldsymbol{a}) = \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{s}, \boldsymbol{a}), \tag{5.8}$$

where $\boldsymbol{\phi} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^f$ is a feature mapping for $(\boldsymbol{s}, \boldsymbol{a})$, which is also referred to as the basis function, and $\boldsymbol{w} \in \mathbb{R}^f$ is the parameter vector.

A class of stable and data-efficient RL algorithms that can solve an MDP with function approximations are the batch RL algorithms—"batch" in the sense that a set of transition samples are utilized each time—such as the LSPI algorithm [74], which is considered to be the most efficient one in this class. We next explain the fundamental idea behind the LSPI algorithm. Let $\mathcal{D} = \{(\boldsymbol{s}, \boldsymbol{a}, r, \boldsymbol{s}') : \boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}, \boldsymbol{a} \in \mathcal{A}\}$ denote a set (batch) of transition samples obtained via observation or simulation. The LSPI algorithm finds the best $\boldsymbol{w}$ that fits the transition samples in $\mathcal{D}$ in an iterative manner. One way to explain the intuition behind the LSPI algorithm is as follows (the readers are referred to [74] for a more rigorous development). Define

$$g(\boldsymbol{w}) = \sum_{(\boldsymbol{s}, \boldsymbol{a}, r, \boldsymbol{s}') \in \mathcal{D}} (Q(\boldsymbol{s}, \boldsymbol{a}) - \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{s}, \boldsymbol{a}))^2. \tag{5.9}$$

Let $\boldsymbol{w}_i$ denote the value of $\boldsymbol{w}$ that is available at the beginning of iteration

i. At iteration i, the algorithm finds $\boldsymbol{w}_{i+1}$ by solving the following problem:

$$\boldsymbol{w}_{i+1} = \arg\min_{\boldsymbol{w}} g(\boldsymbol{w}), \tag{5.10}$$

which is an unconstrained optimization problem. The solution of (5.10) can be computed by setting the gradient of $g(\cdot)$ to zero as follows:

$$\frac{\partial g}{\partial \boldsymbol{w}} = -2 \sum_{(\boldsymbol{s},\boldsymbol{a},r,\boldsymbol{s}')\in\mathcal{D}} (Q(\boldsymbol{s},\boldsymbol{a}) - \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{s},\boldsymbol{a}))\boldsymbol{\phi}(\boldsymbol{s},\boldsymbol{a}) = \mathbf{0}_f. \tag{5.11}$$

Note that the true value of $Q(\boldsymbol{s},\boldsymbol{a})$ is not known and is substituted by the so-called temporal-difference (TD) target, $r + \gamma \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{s}',\boldsymbol{a}')$, where $\boldsymbol{a}' = \arg\max_{\boldsymbol{a}\in\mathcal{A}} \boldsymbol{w}_i^\top \boldsymbol{\phi}(\boldsymbol{s}',\boldsymbol{a})$ is the optimal action in state $\boldsymbol{s}'$ determined based on $\boldsymbol{w}_i$. Note that the TD target is a sample of the right-hand-side (RHS) of (5.7), which serves as an estimate for the RHS of (5.7). We emphasize that despite $Q(\boldsymbol{s},\boldsymbol{a})$ being substituted by $r + \gamma \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{s}',\boldsymbol{a}')$, the true optimal action-value function is not a function of $\boldsymbol{w}$; therefore, the gradient of g with respect to $\boldsymbol{w}$ is taken before the $Q(\boldsymbol{s},\boldsymbol{a})$ is approximated by the TD target, which does depends on $\boldsymbol{w}$. Then, after replacing $Q(\boldsymbol{s},\boldsymbol{a})$ with the TD target, (5.11) has the following closed-form solution:

$$\boldsymbol{w}_{i+1} = \left(\sum_{(\boldsymbol{s},\boldsymbol{a},r,\boldsymbol{s}')\in\mathcal{D}} \boldsymbol{\phi}(\boldsymbol{s},\boldsymbol{a})(\boldsymbol{\phi}(\boldsymbol{s},\boldsymbol{a}) - \gamma\boldsymbol{\phi}(\boldsymbol{s}',\boldsymbol{a}'))^\top \right)^{-1} \sum_{(\boldsymbol{s},\boldsymbol{a},r,\boldsymbol{s}')\in\mathcal{D}} \boldsymbol{\phi}(\boldsymbol{s},\boldsymbol{a})r. \tag{5.12}$$

Intuitively, at each iteration, the LSPI algorithm finds the $\boldsymbol{w}$ that minimizes the mean squared error between the TD target and $\hat{Q}(\boldsymbol{s},\boldsymbol{a})$ over all transition samples in $\mathcal{D}$. This process is repeated until change of $\boldsymbol{w}$, defined as $\|\boldsymbol{w}_{i+1} - \boldsymbol{w}_i\|$, where $\|\cdot\|$ denotes the L_2-norm, becomes smaller than a threshold ε, upon which the algorithm is considered to have converged.

The LSPI algorithm has the following three advantageous properties. First, linear functions are used to approximate the optimal action-value function, which allows the algorithm to handle MDPs with high-dimensional or continuous state spaces. Second, at each iteration, a batch of transition samples is used to update the vector $\boldsymbol{w}$ parameterizing $\hat{Q}(\cdot,\cdot)$, and these samples are reused at each iteration, thus increasing data efficiency. Third, the optimal

parameter vector is found by solving a least-squares problem, resulting in a stable algorithm. We refer interested readers to [74] for more details on the convergence and performance guarantee of the LSPI algorithm.

5.4 Optimal Tap Setting Problem Formulation

In this section, we formulate the optimal tap setting problem as an MDP as follows:

State space

Define the squared voltage magnitudes at all buses but bus 0 and the tap ratios as the state, i.e., $\boldsymbol{s} = (\boldsymbol{t}, \boldsymbol{v})$, which has both continuous and discrete variables. Then, the state space is $\mathcal{S} \subseteq \mathcal{T}^{L^t} \times \mathbb{R}^N$.

Action space

The actions are the LTC tap ratio changes, i.e., $\boldsymbol{a} = \Delta\boldsymbol{t}$, and the action space is the set of all feasible values of LTC tap ratios, i.e., $\mathcal{A} = \Delta\mathcal{T}^{L^t}$. In the optimal tap setting problem, the action is discrete. The size of the action space increases exponentially with the number of LTCs.

Reward function

The objective of voltage regulation is to minimize the voltage deviation as measured by the L_2 norm. As such, when the system transitions from state $\boldsymbol{s} = (\boldsymbol{t}, \boldsymbol{v})$ into state $\boldsymbol{s}' = (\boldsymbol{t}', \boldsymbol{v}')$ after taking action $\boldsymbol{a} = \Delta\boldsymbol{t} := \boldsymbol{t}' - \boldsymbol{t}$, the reward is computed by the following function:

$$\mathcal{R}(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{s}') = -\frac{1}{N}\|\boldsymbol{v}' - \boldsymbol{v}^\star\|. \tag{5.13}$$

Transition model

To derive the transition model $\mathcal{P}$, note that it follows from (5.3) that

$$\boldsymbol{v}' = (\boldsymbol{M}(\boldsymbol{t}')^\top)^{-1}(\boldsymbol{\xi} + \boldsymbol{M}(\boldsymbol{t})^\top \boldsymbol{v} + \boldsymbol{m}(\boldsymbol{t})v_0 - \boldsymbol{m}(\boldsymbol{t}')v_0), \tag{5.14}$$

where $\boldsymbol{\xi} = 2\text{diag}(\boldsymbol{r})\boldsymbol{M}^{-1}(\boldsymbol{p}' - \boldsymbol{p}) + 2\text{diag}(\boldsymbol{x})\boldsymbol{M}^{-1}(\boldsymbol{q}' - \boldsymbol{q})$, and $\boldsymbol{p}'$ and $\boldsymbol{q}'$ are

active and reactive power injections that result into $\boldsymbol{v}'$, respectively. Then, the transition model $\mathcal{P}(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a})$ can be derived from the probability density function (pdf) of $(\boldsymbol{v}'|\boldsymbol{v}, \boldsymbol{t}, \Delta\boldsymbol{t})$, which can be further computed from the pdf of $(\boldsymbol{\xi}|\boldsymbol{v}, \boldsymbol{t}, \Delta\boldsymbol{t})$. However, under Assumptions **A1** and **A2**, the line parameters as well as the probability distributions of active and reactive power injections are unknown; thus, the transition model is not known a priori. Therefore, we need to resort to RL algorithms that do not require an explicit transition model to solve the MDP.

5.5 Optimal Tap Setting Algorithm

In this section, we propose an optimal tap setting algorithm, which consists of a transition generating algorithm that can generate samples of transitions in $\mathcal{D}$, and an LSPI-based sequential learning algorithm to solve the MDP. Implementation details such as the feature selection are also discussed.

5.5.1 Overview

The overall structure of the optimal tap setting framework is illustrated in Fig. 5.2. The framework consists of an environment that is the power distribution system, a learning agent that learns the action-value function from a set of transition samples, and an acting agent that determines the optimal action from the action-value function. Define the history to be the sequence of states, actions, and rewards, and denote it by $\mathcal{H}$, i.e., $\mathcal{H} = \{\boldsymbol{s}[0], \boldsymbol{a}[0], r[0], \boldsymbol{s}[1], \boldsymbol{a}[1], r[1], \cdots\}$. Specifically, the learning agent will use the elements in the set $\mathcal{H}$ together with a virtual transition generator to generate a set of transition samples $\mathcal{D}$ according to some exploratory behavior defined in the exploratory actor. The set of transition samples in $\mathcal{D}$ is then used by the action-value function estimator—also referred to as the critic—to fit an approximate action-value function using the LSPI algorithm described earlier. The learning agent, which has a copy of the up-to-date approximate action-value function from the learning agent, finds a greedy action for the current state and instructs the LTCs to follow it.

Note that the learning of the action-value function can be done offline by the learning agent, which is capable of exploring various system conditions

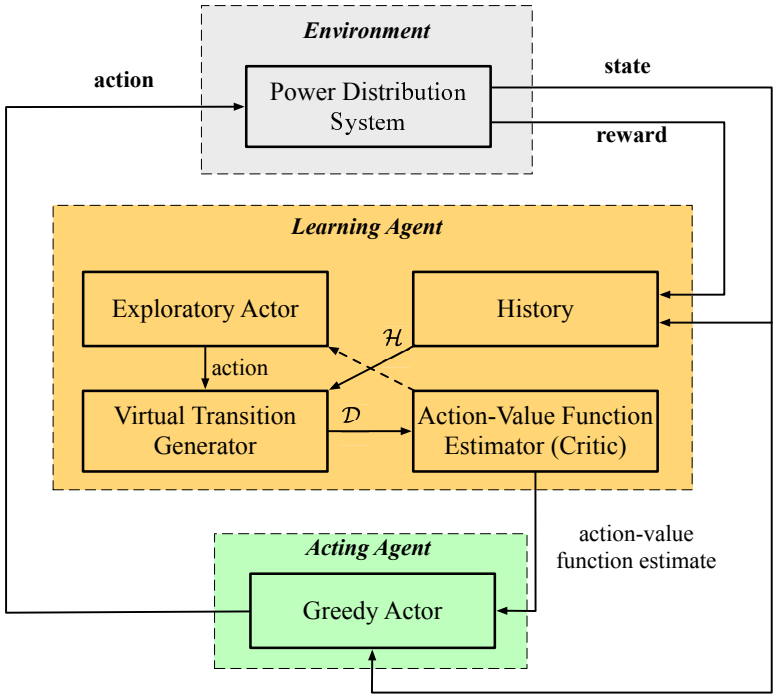

Figure 5.2: The batch RL based framework for optimal tap setting. (Dotted line indicates the critic is optional for the exploratory actor.)

through the virtual transition generator based on the history $\mathcal{H}$, yet without directly interacting with the power distribution system. This avoids jeopardizing system operational reliability, which is a major concern when applying RL algorithms to power system applications [40].

5.5.2 Virtual Transition Generator

The LSPI algorithm (as well as all other RL algorithms) requires adequate transition samples that spread over the state and action spaces $\mathcal{S} \times \mathcal{A}$. However, this is challenging in power systems since the system operational reliability might be jeopardized when exploring randomly. One way to work around this issue is to use simulation models, rather than the physical system, to generate virtual transitions. To this end, we develop a data-driven virtual transition generator that simulates transitions without any knowledge of the active and reactive power injections (neither measurements nor probability distributions) or the line parameters.

The fundamental idea is the following. For an actual transition sample $(\boldsymbol{s}, \boldsymbol{a}^\dagger, r^\dagger, \boldsymbol{s}^\dagger = (\boldsymbol{t}^\dagger, \boldsymbol{v}^\dagger))$ that is obtained from $\mathcal{H}$, the virtual transition gener-

ator generates a new transition sample $(\boldsymbol{s}, \boldsymbol{a}^\ddagger, r^\ddagger, \boldsymbol{s}^\ddagger = (\boldsymbol{t}^\ddagger, \boldsymbol{v}^\ddagger))$, where $\boldsymbol{a}^\ddagger$ is determined from $\boldsymbol{s}$ according to some exploration policy (to be defined later) that aims to explore the state and action spaces. Replacing $\boldsymbol{a}^\dagger$ in the first transition sample with $\boldsymbol{a}^\ddagger$, the voltage magnitudes will change accordingly. Assume the same transition of the power injections in these two samples, then the RHS of (5.3) does not change. Thus, $\boldsymbol{v}^\ddagger$ can be readily computed from $\boldsymbol{v}^\dagger$ by solving the following set of linear equations:

$$\boldsymbol{M}(\boldsymbol{t}^\ddagger)^\top \boldsymbol{v}^\ddagger + \boldsymbol{m}(\boldsymbol{t}^\ddagger) v_0 = \boldsymbol{M}(\boldsymbol{t}^\dagger)^\top \boldsymbol{v}^\dagger + \boldsymbol{m}(\boldsymbol{t}^\dagger) v_0. \tag{5.15}$$

Since the only unknown in (5.15) is $\boldsymbol{v}^\ddagger \in \mathbb{R}$ and $M(\boldsymbol{t}^\ddagger) \in \mathbb{R}^{N \times N}$ is invertible, we can solve for $\boldsymbol{v}^\ddagger$ as follows:

$$\boldsymbol{v}^\ddagger = (\boldsymbol{M}(\boldsymbol{t}^\ddagger)^\top)^{-1} (\boldsymbol{M}(\boldsymbol{t}^\dagger)^\top \boldsymbol{v}^\dagger + \boldsymbol{m}(\boldsymbol{t}^\dagger) v_0 - \boldsymbol{m}(\boldsymbol{t}^\ddagger) v_0). \tag{5.16}$$

For ease of notation, we simply write (5.16) as

$$\boldsymbol{v}^\ddagger = \varphi(\boldsymbol{v}^\dagger, \boldsymbol{t}^\dagger, \boldsymbol{t}^\ddagger). \tag{5.17}$$

This property allows us to estimate the new values of voltage magnitudes when the tap positions change without knowing the exact values of power injections and line parameters. The virtual transition generating procedure is summarized in Algorithm 5.1.

5.5.3 LSPI-based Sequential Action-Value Function Learning

Given the transition sample set $\mathcal{D}$, we can now develop a learning algorithm for $\hat{Q}(\boldsymbol{s}, \boldsymbol{a})$ based on the LSPI algorithm. While the LSPI is very efficient when the action space is relatively small, it becomes computationally intractable when the action space is large, since the number of unknown parameters in the approximate action-value function is typically proportional to $|\mathcal{A}|$, which increases exponentially with the number of LTCs. To overcome the "curse of dimensionality" that results from the size of the action space, we propose an LSPI-based sequential learning algorithm to learn the action-value function.

The key idea is the following. Instead of learning an approximate optimal

Algorithm 5.1: Virtual transition Generating

Input:

$\mathcal{H}$: history

D: number of transition samples

$\boldsymbol{v}^\star$: reference value of squared voltage magnitudes

exploration policy

Output:

$\mathcal{D}$: transition sample set

Initialize $\mathcal{D} \leftarrow \varnothing$

for $d = 1, \cdots, D$ **do**

Choose a transition sample $(\boldsymbol{s}, \boldsymbol{a}^\dagger, r^\dagger, \boldsymbol{s}^\dagger = (\boldsymbol{t}^\dagger, \boldsymbol{v}^\dagger))$ from $\mathcal{H}$

Select $\boldsymbol{a}^\ddagger$ according to exploration policy and set $\boldsymbol{t}^\ddagger = \boldsymbol{t}^\dagger + \boldsymbol{a}^\ddagger$ ($\boldsymbol{t}^\ddagger$ is restricted to elements in $\mathcal{T}^{L^t}$)

Estimate $\boldsymbol{v}^\ddagger$ following $\boldsymbol{a}^\ddagger$ as $\boldsymbol{v}^\ddagger = \varphi(\boldsymbol{v}^\dagger, \boldsymbol{t}^\dagger, \boldsymbol{t}^\ddagger)$

Compute the reward by $r^\ddagger = -\frac{1}{N}\|\boldsymbol{v}^\ddagger - \boldsymbol{v}^\star\|$

Add $(\boldsymbol{s}, \boldsymbol{a}^\ddagger, r^\ddagger, \boldsymbol{s}^\ddagger = (\boldsymbol{t}^\ddagger, \boldsymbol{v}^\ddagger))$ to $\mathcal{D}$

end

action-value function for the action vector $\boldsymbol{a}$, we learn a separate approximate action-value function for each component of $\boldsymbol{a}$. To be more specific, for each LTC l, $l = 1, \cdots, L^t$, we learn an approximate optimal action-value function $\hat{Q}^{(l)}(\boldsymbol{s}, a^{(l)}) = \boldsymbol{\phi}^{(l)}(\boldsymbol{s}, a^{(l)})^\top \boldsymbol{w}^{(l)}$, where $a^{(l)}$ is the l^{th} component of $\boldsymbol{a}$, and $\boldsymbol{\phi}^{(l)}(\cdot,\cdot)$ is a feature mapping from $\mathcal{S} \times \Delta\mathcal{T}$ to $\mathbb{R}^f$. During the learning process of $\boldsymbol{w}^{(l)}$, the rest of the LTCs are assumed to behave greedily according to their own approximate optimal action-value function. To achieve this, we design the following exploration policy to generate the virtual transition samples $\mathcal{D}$ used when learning $\boldsymbol{w}^{(l)}$ for LTC l. In the exploration step in Algorithm 5.1, the tap ratio change of LTC l is selected uniformly in $\Delta\mathcal{T}$ (uniform exploration), while those of others are selected greedily with respect to the up-to-date $\hat{Q}^{(l)}(\cdot,\cdot)$ (greedy exploration). Then, the LSPI algorithm detailed in Algorithm 5.2, where c is a small positive pre-condition number and $\boldsymbol{w}_0^{(l)}$ is the initial value for the parameter vector, is applied to learn $\boldsymbol{w}^{(l)}$. This procedure is repeated in a round-robin fashion for all LTCs for J iterations, in each of which $\boldsymbol{w}_0^{(l)}$ is set to the up-to-date $\boldsymbol{w}^{(l)}$ learned in the previous iteration or chosen if it is in the first iteration. The value of J is set to 1 if there is only one LTC and is increased slightly when there are more LTCs. Note that a new set of transitions $\mathcal{D}$ is generated when learning $\boldsymbol{w}^{(l)}$ for different LTCs at each iteration. Using this sequential learning algorithm,

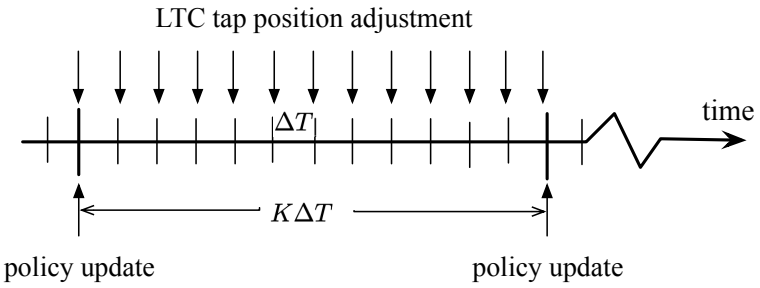

Figure 5.3: Timeline for LTC tap setting.

the total number of unknowns is then proportional to $L^t|\Delta\mathcal{T}|$, which is far fewer compared to $|\Delta\mathcal{T}^{L^t}|$ as in the case where the approximate optimal action-value function for the entire action vector, $\boldsymbol{a}$, is learned.

A critical step in implementing the LSPI algorithm is constructing features from the state-action pair $(\boldsymbol{s}, a^{(l)})$ for LTC l; we use radial basis function (RBFs) to this end. The feature vector for a state-action pair $(\boldsymbol{s}, a^{(l)})$, i.e., $\boldsymbol{\phi}^{(l)}(\boldsymbol{s}, a^{(l)})$, is a vector in $\mathbb{R}^f$, where $f = (\kappa + 1) \times |\Delta\mathcal{T}|$ and κ is a positive integer. The feature vector $\boldsymbol{\phi}^{(l)}(\boldsymbol{s}, a^{(l)})$ has $|\Delta\mathcal{T}|$ segments, each one of length $\kappa + 1$ corresponding to a tap change in $\Delta\mathcal{T}$, i.e, $\boldsymbol{\phi}^{(l)}(\boldsymbol{s}, a^{(l)}) = [\boldsymbol{\psi}_1^\top, \cdots, \boldsymbol{\psi}_{|\Delta\mathcal{T}|}^\top]^\top$, where $\boldsymbol{\psi}_i \in \mathbb{R}^{\kappa+1}, i = 1, \cdots, |\Delta\mathcal{T}|$. Specifically, for $\boldsymbol{s} = (\boldsymbol{t}, \boldsymbol{v})$ and $a^{(l)}$ being the i^{th} tap change in $\Delta\mathcal{T}$, $\boldsymbol{\psi}_j = \mathbf{0}_{\kappa+1}$ for $j \neq i$, and $\boldsymbol{\psi}_i = [1, e^{-\frac{\|\tilde{\boldsymbol{v}} - \bar{\boldsymbol{v}}_1\|}{\sigma^2}}, \cdots, e^{-\frac{\|\tilde{\boldsymbol{v}} - \bar{\boldsymbol{v}}_\kappa\|}{\sigma^2}}]^\top$, where $\sigma > 0$, $\tilde{\boldsymbol{v}} = \varphi(\boldsymbol{v}, \boldsymbol{t}, \tilde{\boldsymbol{t}})$ with $\tilde{\boldsymbol{t}}$ being obtained by replacing the l^{th} entry in $\boldsymbol{t}$ with 1, and $\bar{\boldsymbol{v}}_i$, $i = 1, \cdots, \kappa$ are pre-specified constant vectors in $\mathbb{R}^N$ referred to as the RBF centers. The action $a^{(l)}$ only determines which segment will be non-zero. Thus, $\tilde{\boldsymbol{v}}$ is indeed the squared voltage magnitudes under the same power injections if the tap of LTC l is at position 0. Each RBF computes the distance between $\boldsymbol{v}'$ and some pre-specified squared voltage magnitudes.

5.5.4 Tap Setting Algorithm

The tap setting algorithm, the timeline of which is illustrated in Fig. 5.3, works as follows. At time instant k, a new state $\boldsymbol{s}[k]$ as well as the reward following the action $\boldsymbol{a}_{k-1}$, r_{k-1}, is observed. Let ΔT denote the time elapsed between two time instants. Every K time instants, i.e., every $K\Delta T$ units of time, $\boldsymbol{w}^{(l)}$ is updated by the learning agent by executing the LSPI-based sequential learning algorithm described in Section 5.5.3. The acting agent

Algorithm 5.2: LSPI for Single LTC

Input:

l: index of LTC
$\mathcal{D}$: transition sample set
$\boldsymbol{\phi}(\cdot,\cdot)$: basis function
γ: discount factor
ε: convergence threshold
δ: pre-condition number
$\boldsymbol{w}_0^{(l)}$: initial parameter vector of approximate optimal action-value function for LTC l

Output:

$\boldsymbol{w}^{(l)}$: updated parameter vector of approximate optimal action-value function for LTC l

Initialize $\boldsymbol{w}_{-1}^{(l)} = \mathbf{0}_f$ and $i = 0$
while $\|\boldsymbol{w}_i^{(l)} - \boldsymbol{w}_{i-1}^{(l)}\| > \varepsilon$ *or* $i = 0$ **do**
  Initialize $\boldsymbol{B}_0 = c\boldsymbol{I}_{f\times f}$ and $\boldsymbol{b}_0 = \mathbf{0}_f$, set $j = 1$
  for $(\boldsymbol{s}, \boldsymbol{a}, r, \boldsymbol{s}') \in \mathcal{D}$ **do**
    $a^{(l)'} = \arg\max_{a\in\Delta\mathcal{T}} \boldsymbol{\phi}(\boldsymbol{s}', a)^\top \boldsymbol{w}_i^{(l)}$
    $\boldsymbol{B}_j = \boldsymbol{B}_{j-1} + \boldsymbol{\phi}(\boldsymbol{s}, a^{(l)})(\boldsymbol{\phi}(\boldsymbol{s}, a^{(l)}) - \gamma\boldsymbol{\phi}(\boldsymbol{s}', a^{(l)'}))^\top$
    $\boldsymbol{b}_j = \boldsymbol{b}_{j-1} + \boldsymbol{\phi}(\boldsymbol{s}, a^{(l)})r$
    Increase j by 1
  end
  $\boldsymbol{w}_{i+1}^{(l)} = \boldsymbol{B}_{|\mathcal{D}|}^{-1}\boldsymbol{b}_{|\mathcal{D}|}$, increase i by 1
end
Set $\boldsymbol{w}^{(l)} = \boldsymbol{w}_i^{(l)}$

then finds a greedy action for the current state $\boldsymbol{s}[k]$ and sends it to the LTCs. In order to reduce the wear and tear on the LTCs, the greedy action for the current state $\boldsymbol{s}[k]$ is chosen only if the difference between the action-value resulting from the greedy action, i.e., $\max\limits_{a\in\Delta\mathcal{T}} \boldsymbol{\phi}(\boldsymbol{s}[k], a)^\top \boldsymbol{w}^{(l)}$, and that resulting from the previous action, i.e., $\boldsymbol{\phi}(\boldsymbol{s}[k], a_{k-1}^{(l)})^\top \boldsymbol{w}^{(l)}$, is larger than a threshold ϵ. Otherwise, the tap positions do not change. The above procedure is summarized in Algorithm 5.3.

5.6 Numerical Simulation

In this section, we apply the proposed methodology to the IEEE 13-bus and 123-bus test feeders from [65].

Algorithm 5.3: Optimal Tap Setting

Input:

K: update period

J: number of learning iterations

ϵ: minimum action-value change

for $k = 1, 2, \cdots$ **do**

  Obtain $r[k-1]$ and $\boldsymbol{s}[k]$, and add them into $\mathcal{H}$

  if $k \mod K = 0$ **then**

    for $j = 1, \cdots, J$ **do**

      for $l = 1, \cdots, L^t$ **do**

        Run Algo. 5.1 to generate $\mathcal{D}$ using uniform exploration for LTC l and greedy exploration for other LTCs

        Run Algo. 5.2 with $\boldsymbol{w}_0^{(l)}$ set to the current $\boldsymbol{w}^{(l)}$

      end

    end

  end

  for $l = 1, \cdots, L^t$ **do**

    Set $a^{(l)}[k] = \arg\max_{a \in \Delta\mathcal{T}} \boldsymbol{\phi}(\boldsymbol{s}[k], a)^\top \boldsymbol{w}^{(l)}$ if $\max_{a \in \Delta\mathcal{T}} \boldsymbol{\phi}(\boldsymbol{s}[k], a)^\top \boldsymbol{w}^{(l)} - \boldsymbol{\phi}(\boldsymbol{s}[k], a^{(l)}[k-1])^\top \boldsymbol{w}^{(l)} > \epsilon$

    Set $a^{(l)}[k] = a^{(l)}[k-1]$ otherwise

  end

  Add $\boldsymbol{a}[k]$ to $\mathcal{H}$ and adjust tap ratios based on $\boldsymbol{a}[k]$

end

5.6.1 Simulation Setup

The power injections for these two test feeders are constructed based on historical hourly active power load data from a residential building in San Diego over one year [69]. Specifically, the historical hourly active power load data are first scaled up so that the maximum system total active power load over that year for the IEEE 13-bus and 123-bus distribution test feeders are 6.15 MW and 12.3 MW, respectively. These numbers are chosen so that the resulting voltage magnitudes fall outside of the desired range at some time instants. Then, the time granularity of the scaled system total active power load is increased to 5 minutes through a linear interpolation. Each value in the resulting five-minute load data is further multiplied by a normally distributed variable, the mean and standard deviation (SD) of which are 1 and 0.02, respectively. The active power load profile at each bus is constructed by pseudo-randomly redistributing the system total active power

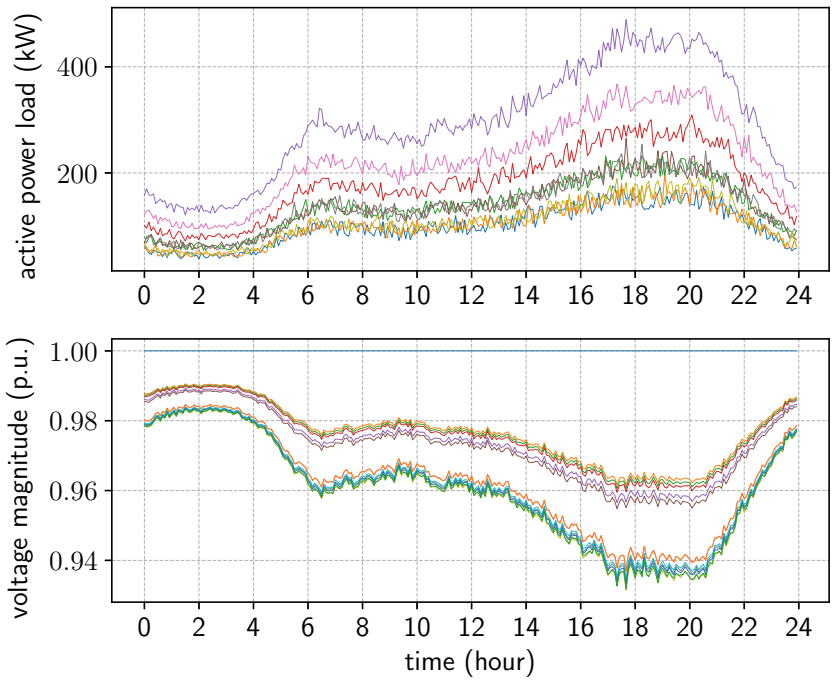

Figure 5.4: Typical load and voltage profiles in the IEEE 13-bus test feeder in July.

load among all load buses. Each load bus is assumed to have a constant power factor of 0.95. While only load variation is considered in the simulation, the proposed methodology can be directly applied to the case with renewable-based resources, which can be modeled as negative loads. Figure 5.4 shows the typical active power load profiles in July at each bus in the IEEE 13-bus distribution test feeder, as well as the corresponding voltage magnitudes under the nominal LTC tap ratio, i.e., the LTC tap ratio equals to 1.

We first verify the accuracy of the virtual transition generating algorithm. Specifically, assume the voltage magnitudes are known for some unknown power injections under a known tap ratio of 1. Then, when the tap ratio changes, we compute the true voltage magnitudes under the new tap ratio, denoted by $\boldsymbol{v}$, by solving the full ac power flow problem, and the estimated voltage magnitudes under new tap ratio, denoted by $\hat{\boldsymbol{v}}$, via (5.17). Simulation results indicate that the voltage approximation error, which is defined to be the maximum absolute difference between the true and the estimated voltage magnitude, i.e., $\|\boldsymbol{v} - \hat{\boldsymbol{v}}\|_\infty$, is smaller than 0.001 p.u., which is accurate enough for the application of voltage regulation addressed here. The voltage

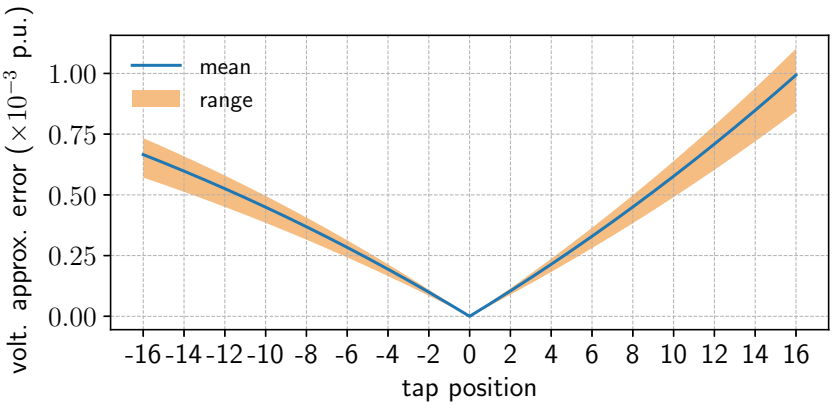

Figure 5.5: Voltage approximation error. (Position 0 corresponds to tap ratio 1.)

approximation error for the IEEE 13-bus test feeder is presented in Fig. 5.5. The result for the IEEE 123-bus test feeder is similar.

5.6.2 Case Study on the IEEE 13-bus Test Feeder

Assume $\boldsymbol{v}^\star = \mathbf{1}_N$, where $\mathbf{1}_N$ is an all-ones vector in $\mathbb{R}^N$ In the simulation, 21 RBF centers are used, i.e., $\kappa = 21$. Specifically, $\bar{\boldsymbol{v}}_i = (0.895 + 0.005i)^2 \times \mathbf{1}_N$, $i = 1, \cdots, 21$. The duration between two time instants is $\Delta T = 5$ min. The policy is updated every 2 hours, i.e., $K = 24$. In each update, actual transition samples are chosen from the history over the same time interval in the previous 5 days, which are part of $\mathcal{H}$, and new actions are chosen according to the exploration policy described in Section 5.5.3. A total number of $D = 6000$ virtual transitions are generated using Algorithm 5.1. Since this test feeder only has one LTC, there is no need to sequentially update the approximate action-value function, so we set $J = 1$. Other parameters are chosen as follows: $\gamma = 0.9$, $\varepsilon = 1 \times 10^{-5}$, $\epsilon = 1 \times 10^{-4}$, $c = 0.1$, and $\sigma = 1$.

Assuming complete and perfect knowledge of the system parameters as well as active and reactive power injections for all time instants, we can find the optimal tap position that results in the highest reward by exhaustively searching the action space, i.e., all feasible tap ratios, at each time instant. It is important to point out that, in practice, the exhaustive search approach is infeasible since we do not have the necessary information, and not practical due to the high computational burden. Results obtained by the exhaustive

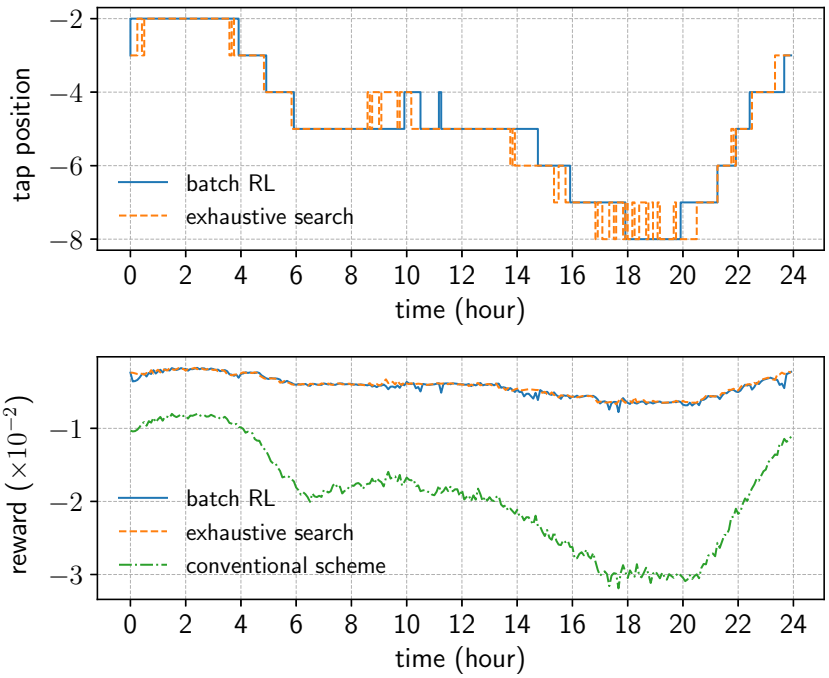

Figure 5.6: LTC tap positions and rewards in IEEE 13-bus test feeder.

search approach and the conventional tap setting scheme (see, e.g., [14]), in which the taps are adjusted only when the voltage magnitudes exceed a desired range, e.g., $[0.9, 1.1]$ p.u., are used to benchmark the proposed methodology.

Figure 5.6 shows the tap positions (top panel) and the rewards (bottom panel) under different approaches. The rewards resulted from these two approaches are very close. The daily mean rewards, i.e., $\rho = \frac{1}{288}\sum_{k=1}^{288} r[k]$, where $r[k]$ is the reward at time instant k as defined in (5.13), obtained by the batch RL approach and the exhaustive search approach are $\rho = -4.279\times10^{-3}$ and $\rho = -4.156 \times 10^{-3}$, respectively, while that under the conventional scheme is $\rho = -19.26 \times 10^{-3}$. The tap positions under the batch RL approach and the exhaustive search approach are aligned during most of the time during the day. Note that the tap position under the conventional scheme remains at 0 since the voltage magnitudes are within $[0.9, 1.1]$ p.u., and is not plotted. Figure 5.7 shows the voltage magnitude profiles under the different tap setting algorithms. The voltage magnitude profiles under the proposed batch RL approach (see Fig. 5.7, center panel) are quite similar to those obtained via the exhaustive search approach (see Fig. 5.7, bottom

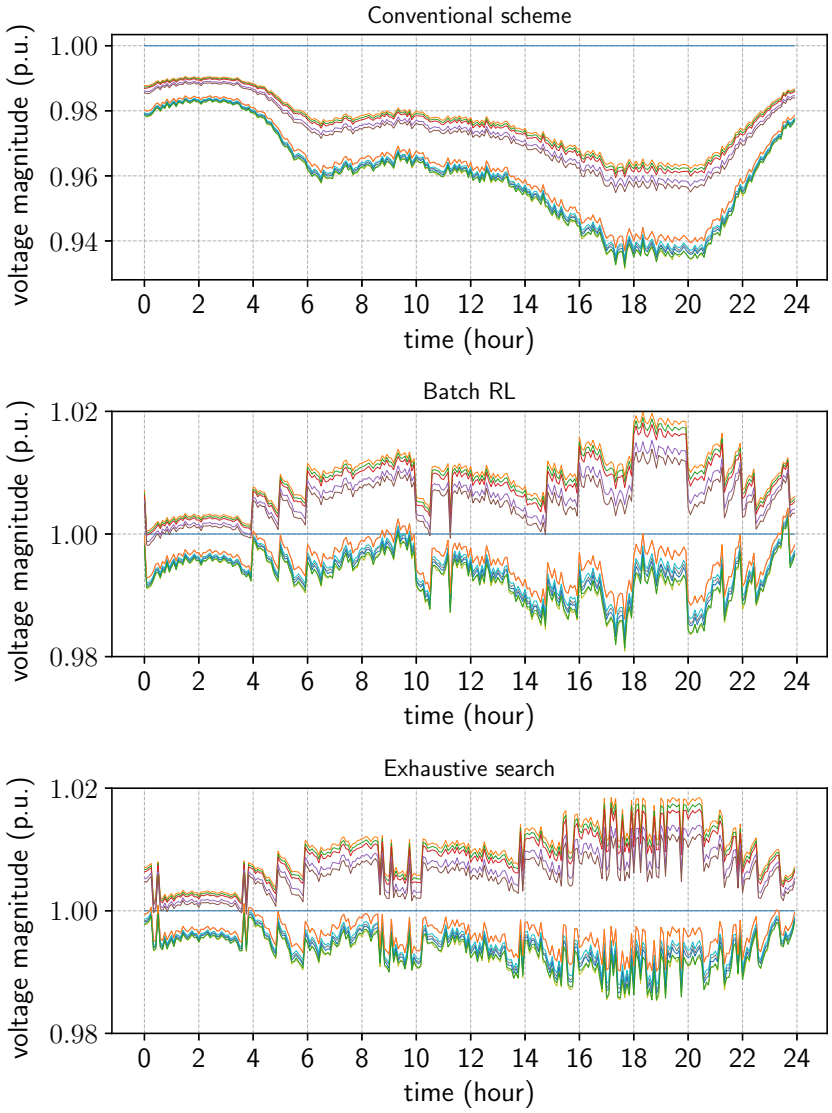

Figure 5.7: Voltage magnitude profiles in IEEE 13-bus test feeder.

panel); both result in a higher daily mean reward than that resulted from the conventional scheme (see Fig. 5.7, top panel). Algorithm 5.2 typically converges in about 30 seconds, and the batch RL approach is faster than the exhaustive search approach by several orders of magnitude.

While the policy in the above simulation is updated using data over 2 hours, we also study cases when the policy is updated using data over a longer time period, i.e., when K is larger than 24. Table 5.1 shows the mean and SD of rewards over one day when the policy is updated using data over different periods, when $|\mathcal{D}|$ is 50 times the number of actual transitions during

Table 5.1: Rewards Under Various Policy Update Periods

update period (hour)	2	4	6	8	24
reward mean ($\times 10^{-3}$)	-4.279	-4.290	-4.390	-4.347	-4.501
reward SD ($\times 10^{-3}$)	1.436	1.423	1.480	1.529	1.601

the same time period over the previous 5 days. As can be seen, learning a policy using data over a shorter period reduces the required amount of data, while achieving an equally good result.

5.6.3 Case Study on the IEEE 123-bus Test Feeder

We next test the proposed methodology on the IEEE 123-bus test feeder. In the results for the IEEE 13-bus test feeder reported earlier, while the LTC has 33 tap positions, only a small portion of them is actually used. This motivates us to further reduce the action space by narrowing the action space to a smaller range. Specifically, we can estimate the voltage magnitudes under various power injections and LTC tap positions using (5.17). After ruling out tap positions under which the voltage magnitudes will exceed the desired range, we eventually allow 9 positions, from -8 to 0, for two LTCs, and 5 positions, from 0 to 5, for the other two LTCs. Here, $\kappa = 11$ RBF centers are used. Specifically, $\bar{\boldsymbol{v}}_i = (0.94 + 0.01i)^2 \times \mathbf{1}_N$ for all LTCs except for the one near the substation, for which $\bar{\boldsymbol{v}}_i = (0.89 + 0.01i)^2 \times \mathbf{1}_N$, $i = 1, 2, \cdots, 11$. A total number of $D = 3600$ virtual transitions are generated in a similar manner as in the IEEE 13-bus test feeder case. The number of iterations in the LSPI-based sequential learning algorithm is set to $J = 3$. Other parameters are the same as in the IEEE 13-bus test feeder case.

Figures 5.8 and 5.9 show the tap positions and rewards, respectively, under the batch RL approach and the exhaustive search. The daily mean rewards obtained by the batch RL approach and the exhaustive search approach are $\rho = -1.646 \times 10^{-3}$ and $\rho = -1.402 \times 10^{-3}$, respectively, while that under the conventional scheme is $\rho = -7.513 \times 10^{-3}$. While the tap positions differ, the rewards resulting from these two approaches are very close. Note that the tap changes are smoother in the proposed approach since we have enforced a minimum action-value function change requirement to trigger a tap change action. The voltage magnitude profiles are shown in Fig. 5.10.

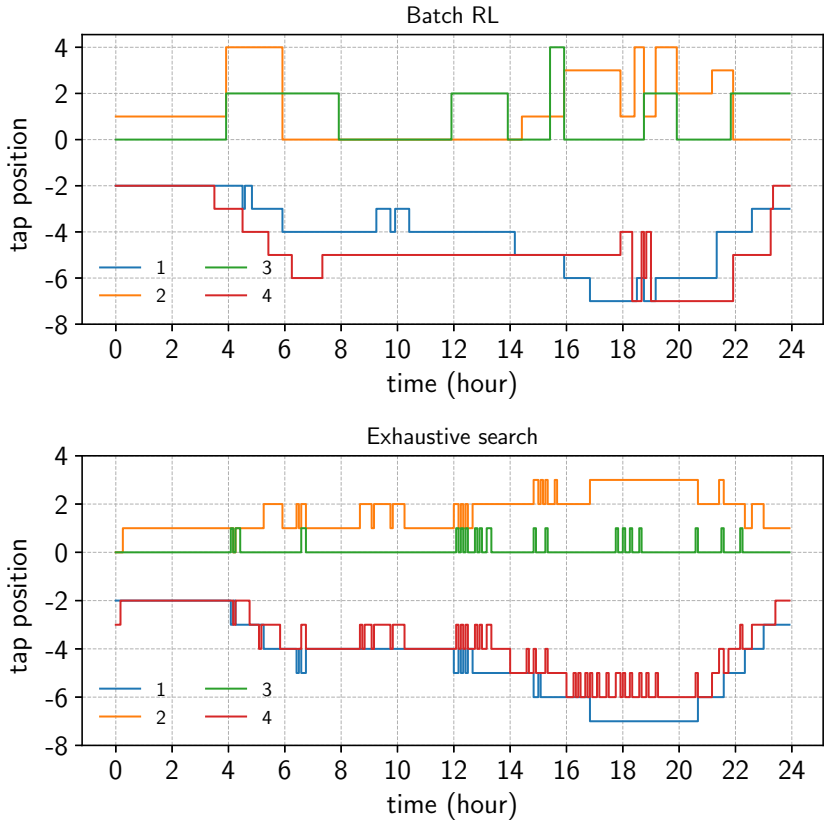

Figure 5.8: LTC tap positions in IEEE 123-bus test feeder.

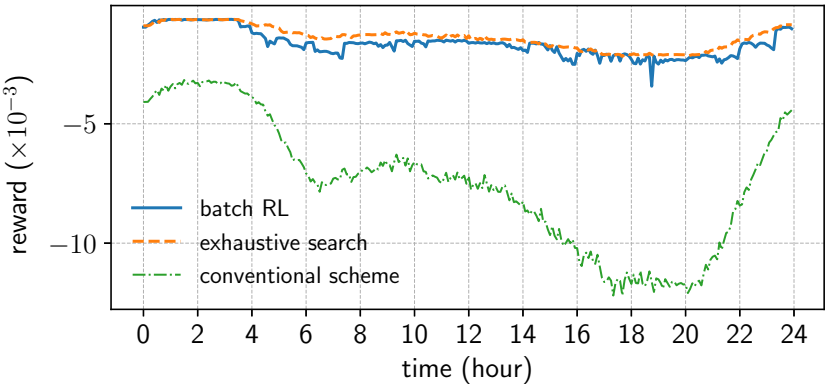

Figure 5.9: Rewards in IEEE 123-bus test feeder.

Figure 5.11 shows the rewards over one week. As can be seen from Fig. 5.11, the voltage regulation performance achieved by the proposed approach is in general close to that achieved by the exhaustive search approach.

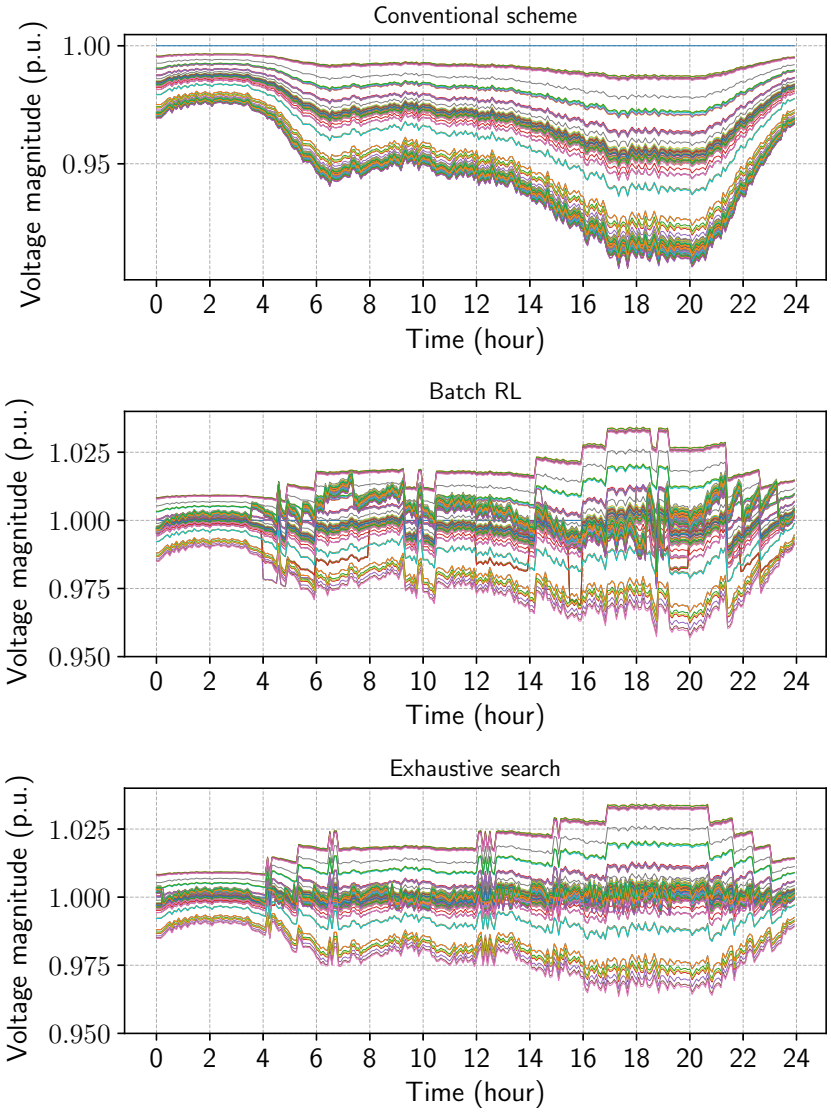

Figure 5.10: Voltage magnitude profiles in IEEE 123-bus distribution test feeder.

5.7 Summary

In this chapter, we formulated the optimal tap setting problem of LTCs in power distribution systems as an MDP and proposed a batch RL algorithm to solve it. To obtain adequate state-action samples, we developed a virtual transition generator that estimates the voltage magnitudes under different tap settings. To circumvent the "curse of dimensionality", we proposed an LSPI-based sequential learning algorithm to learn an action-value function for each LTC, based on which the optimal tap positions can be determined

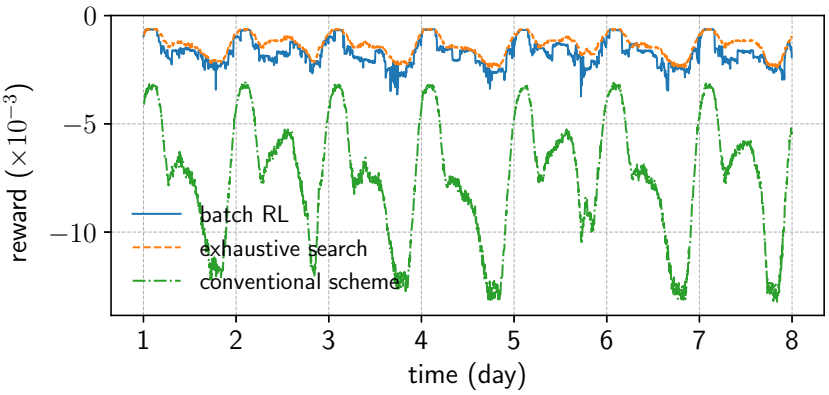

Figure 5.11: Rewards in IEEE 123-bus test feeder over 7 days.

directly. The proposed algorithm can find the policy that determines the optimal tap positions that minimize the voltage deviation across the system, based only on voltage magnitude measurements and network topology information, which makes it more desirable for implementation in practice. Numerical simulation on two IEEE distribution test feeders validated the effectiveness of the proposed methodology.

The work presented in this chapter is submitted for publication in [79].

CHAPTER 6

CONCLUDING REMARKS

In this chapter, we summarize this dissertation and highlight the main contributions to the development of data-driven frameworks for coordinating assets in power distribution systems for the ancillary service provision. We conclude this chapter with a discussion of future work that needs to be done.

6.1 Dissertation Summary and Contributions

In Chapter 1, we first presented the motivation to develop data-driven frameworks for coordinating assets in power distribution systems in order to provide ancillary services. We then described the power distribution system model adopted throughout this dissertation, particularly, the LinDistFlow model, which enables the development of efficient data-driven estimation or coordination algorithms. We also reviewed some of the relevant works and discussed their shortcomings. We ended this chapter by summarizing the contributions and contents of this dissertation.

In Chapter 2, we proposed a data-driven framework for DER coordination for the purpose of active power provision, particularly, provision of frequency regulation services. A key element in this chapter is the concept of loss factors (LFs), which are essentially sensitivities that allow the approximation of the incremental total system losses. We developed a recursive estimator for estimating the LFs and a controller that solves the ODCP formulated using the estimated LFs. The estimator updates its sensitivity estimates at the same timescale at which the controller updates the DER active power injection set-points. The proposed LF estimator is capable of obtaining an accurate estimate of the LFs that can capture the impacts of both active and reactive power injections on system losses with little computational effort. We showed through numerical simulation that the proposed framework

is more effective in coordinating the DERs to provide frequency regulation services, compared to the approach in which system losses are neglected. The main contribution in this chapter is the one-timescale data-driven DER coordination framework.

In Chapter 3, we continued focusing on the problem of coordinating DERs for active power provision. A key issue addressed in this chapter is the potential collinearity in the measurements, which may deteriorate the estimation accuracy. We resolved this issue by introducing random perturbations in the DER active power injections during the estimation process. The estimator works on a fast timescale at which random perturbations are injected, while the controller works on a slow timescale. Under some mild assumptions, we showed that, in an a.s. sense, the estimated parameters converge to their true values, and the total active power exchanged between both systems also converges to the required amount during the estimation process. We validated the effectiveness of the proposed framework through numerical simulations on a modified IEEE 123-bus distribution test feeder. The main contributions in this chapter include the two-timescale coordination framework, the estimation algorithm with random perturbation, as well as its convergence analysis.

In Chapter 4, we developed a data-driven DER coordination framework for reactive power regulation, i.e., voltage regulation, in power distribution systems. A fundamental difficulty in this problem is that the number of parameters to be estimated by the estimator, i.e., the voltage sensitivities, is much larger than that in the active power provision problem; thus, more measurements are required in order to get an accurate estimate. By exploiting the structural characteristics of radial power distribution systems, we reformulated the estimation problem into a much simpler one and reduced the number of parameters to be estimated significantly. Thus, the estimator requires much less data to obtain an accurate estimate. We show through numerical simulations that the framework is effective and efficient in coordinating the DERs to provide voltage regulation. The main contribution in this chapter is the efficient estimation algorithm for the voltage sensitivities.

In Chapter 5, we formulated the optimal tap setting problem of LTCs as an MDP and proposed a batch RL algorithm to solve it. We developed a virtual experience generating algorithm that can estimate the voltage magnitudes under different tap settings, which allows us to obtain adequate state

action samples required by the learning algorithm. To circumvent the "curse of dimensionality", we proposed an LSPI based sequential learning algorithm to learn an action value function for each LTC, based on which the optimal tap positions can be determined directly. The proposed algorithm can find the policy that determines the optimal tap positions that minimize the voltage deviation across the system, based only on the voltage magnitude measurements and the topology information, which makes it more desirable for implementation in practice. We validated the effectiveness of the proposed batch RL based algorithm via numerical simulation on the two IEEE distribution test feeders. The main contributions in this chapter include the formulation of the LTC tap setting problem as an MDP and the development of efficient batch RL algorithm to solve this MDP.

6.2 Conclusions and Future Work

In this dissertation, we have developed several data-driven frameworks for coordinating assets such as DERs and LTCs in power distribution systems to provide ancillary services. The proposed frameworks assume no prior information on the power distribution system model, except knowledge of the feasible network topology configurations and distribution line 'r-to-x' ratios in some instances, and mainly rely on measurements; this makes the framework adaptive and robust to changes in operating conditions and power distribution system models.

While the primary focus in this dissertation is on three-phase balanced power distribution systems that have a radial topology, in future work it will be necessary to extend the data-driven asset coordination frameworks to three-phase unbalanced power distribution systems with possibly a mesh topology. In addition, it would be worthwhile to integrate the several proposed data-driven DER coordination frameworks, which focus on different problems in power distribution systems, into one that can deal with multiple problems simultaneously.

REFERENCES

[1] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, 2010.

[2] H. Bevrani, A. Ghosh, and G. Ledwich, "Renewable energy sources and frequency regulation: survey and new perspectives," *IET Renewable Power Generation*, vol. 4, no. 5, pp. 438–457, 2010.

[3] F. Katiraei and J. R. Aguero, "Solar PV integration challenges," *IEEE Power and Energy Magazine*, vol. 9, no. 3, pp. 62–71, 2011.

[4] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.

[5] A. D. Domínguez-García and C. N. Hadjicostis, "Coordination and control of distributed energy resources for provision of ancillary services," in *Proc. of First IEEE International Conference on Smart Grid Communications*, Oct 2010, pp. 537–542.

[6] S. S. Guggilam, C. Zhao, E. Dall'Anese, Y. C. Chen, and S. V. Dhople, "Primary frequency response with aggregated DERs," in *Proc. of American Control Conference*, May 2017, pp. 3386–3393.

[7] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1868–1881, 2018.

[8] B. A. Robbins, C. N. Hadjicostis, and A. D. Domínguez-García, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1470–1482, 2013.

[9] A. Kulmala, S. Repo, and P. Järventausta, "Coordinated voltage control in distribution networks including several distributed energy resources," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2010–2020, 2014.

[10] B. Zhang, A. Y. Lam, A. D. Domínguez-García, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714–1726, 2015.

[11] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3794–3803, 2016.

[12] B. A. Robbins and A. D. Domínguez-García, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2903–2913, 2016.

[13] H. Li, F. Li, Y. Xu, D. T. Rizy, and J. D. Kueck, "Adaptive voltage control with distributed energy resources: Algorithm, theoretical analysis, simulation, and field test verification," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1638–1647, 2010.

[14] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*. McGraw-hill New York, 1994, vol. 7.

[15] W.-H. Liu, A. D. Papalexopoulos, and W. F. Tinney, "Discrete shunt controls in a Newton optimal power flow," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1509–1518, 1992.

[16] M. Salem, L. Talat, and H. Soliman, "Voltage control by tap-changing transformers for a radial distribution network," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 144, no. 6, pp. 517–520, 1997.

[17] Z. Yang, A. Bose, H. Zhong, N. Zhang, Q. Xia, and C. Kang, "Optimal reactive power dispatch with accurately modeled discrete control devices: A successive linear approximation approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2435–2444, 2017.

[18] B. A. Robbins, H. Zhu, and A. D. Domínguez-García, "Optimal tap setting of voltage regulation transformers in unbalanced distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 256–267, 2016.

[19] Z. Hou and S. Jin, "Data-driven model-free adaptive control for a class of MIMO nonlinear discrete-time systems," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 2173–2188, 2011.

[20] J. Zhang, X. Zheng, Z. Wang, L. Guan, and C. Chung, "Power system sensitivity identification–inherent system properties and data quality," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2756–2766, 2017.

[21] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, "Measurement-based estimation of linear sensitivity distribution factors and applications," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1372–1382, 2014.

[22] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, "A sparse representation approach to online estimation of power system distribution factors," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1727–1738, 2015.

[23] Y. C. Chen, J. Wang, A. D. Domínguez-García, and P. W. Sauer, "Measurement-based estimation of the power flow Jacobian matrix," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2507–2515, 2016.

[24] K. E. Van Horn, A. D. Domínguez-García, and P. W. Sauer, "Measurement-based real-time security-constrained economic dispatch," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3548–3560, 2016.

[25] C. Lu, Y. Zhao, K. Men, L. Tu, and Y. Han, "Wide-area power system stabiliser based on model-free adaptive control," *IET Control Theory & Applications*, vol. 9, no. 13, pp. 1996–2007, 2015.

[26] J. Zhang, C. Chung, and Y. Han, "Online damping ratio prediction using locally weighted linear regression," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1954–1962, 2016.

[27] C. Mugnier, K. Christakou, J. Jaton, M. De Vivo, M. Carpita, and M. Paolone, "Model-less/measurement-based computation of voltage sensitivities in unbalanced electrical distribution networks," in *Proc. of Power Systems Computation Conference*, June 2016, pp. 1–7.

[28] J. Zhang, Z. Chen, C. He, Z. Jiang, and L. Guan, "Data-driven based optimization for power system var-voltage sequential control," *IEEE Transactions on Industrial Informatics*, 2018.

[29] J. Zhang, Z. Chen, and C. He, "Identification of var-voltage characteristics based on ambient signals," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3202–3203, 2018.

[30] Z. Hou, R. Chi, and H. Gao, "An overview of dynamic-linearization-based data-driven control and applications," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4076–4090, 2017.

[31] J. Zhang, Z. Wang, X. Zheng, L. Guan, and C. Chung, "Locally weighted ridge regression for power system online sensitivity identification considering data collinearity," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1624–1634, 2018.

[32] J. Zhang, C. Chung, and L. Guan, "Noise effect and noise-assisted ensemble regression in power system online sensitivity identification," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2302–2310, 2017.

[33] M. V. Pereira and L. M. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, no. 1-3, pp. 359–375, 1991.

[34] C. C. Carøe and R. Schultz, *A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system.* Konrad-Zuse-Zentrum für Informationstechnik, Berlin, 1998.

[35] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2197–2206, 2011.

[36] C. Zhao, J. Wang, J.-P. Watson, and Y. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2708–2717, 2013.

[37] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.

[38] D. Ernst, M. Glavic, and L. Wehenkel, "Power systems stability control: reinforcement learning framework," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 427–435, 2004.

[39] J. G. Vlachogiannis and N. D. Hatziargyriou, "Reinforcement learning for reactive power control," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1317–1325, 2004.

[40] M. Glavic, R. Fonteneau, and D. Ernst, "Reinforcement learning for electric power system decision and control: Past considerations and perspectives," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6918–6927, 2017.

[41] H. Xu, H. Sun, D. Nikovski, S. Kitamura, K. Mori, and H. Hashimoto, "Deep reinforcement learning for joint bidding and pricing of load serving entity," *IEEE Transactions on Smart Grid*, 2019.

[42] Y. V. Makarov, C. Loutan, J. Ma, and P. De Mello, "Operational impacts of wind generation on California power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1039–1050, 2009.

[43] Y. Chen, R. Leonard, M. Keyser, and J. Gardner, "Development of performance-based two-part regulating reserve compensation on MISO energy and ancillary service market," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 142–155, 2015.

[44] F. Zhang, Z. Hu, X. Xie, J. Zhang, and Y. Song, "Assessment of the effectiveness of energy storage resources in the frequency regulation of a single-area power system," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3373–3380, 2017.

[45] J. Song, X. Pan, C. Lu, and H. Xu, "A simulation-based optimization method for hybrid frequency regulation system configuration," *Energies*, vol. 10, no. 9, p. 1302, 2017.

[46] H. Liu, J. Qi, J. Wang, P. Li, C. Li, and H. Wei, "EV dispatch control for supplementary frequency regulation considering the expectation of EV owners," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3763–3772, 2018.

[47] FERC, "Final rule order no. 755: frequency regulation compensation in the organized wholesale power markets," 2011.

[48] A. D. Papalexopoulos and P. E. Andrianesis, "Performance-based pricing of frequency regulation in electricity markets," *IEEE Transactions on Power systems*, vol. 29, no. 1, pp. 441–449, 2014.

[49] D. Fooladivanda, H. Xu, A. D. Domínguez-García, and S. Bose, "Offer strategies for wholesale energy and regulation markets," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7305–7308, 2018.

[50] E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi, "Marginal loss modeling in LMP calculation," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 880–888, 2004.

[51] D. Q. Hung and N. Mithulananthan, "Multiple distributed generator placement in primary distribution networks for loss reduction," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 4, pp. 1700–1708, 2013.

[52] N. Acharya, P. Mahat, and N. Mithulananthan, "An analytical approach for DG allocation in primary distribution network," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 10, pp. 669–678, 2006.

[53] S. O. Haykin, *Adaptive Filter Theory*. Pearson Higher Ed, 2013.

[54] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. John Wiley & Sons, 2012.

[55] PJM, "RTO regulation signal data," 2017. [Online]. Available: http://www.pjm.com/markets-and-operations/ancillary-services.aspx

[56] R. D. Zimmerman, C. E. Murillo-Sánchez, R. J. Thomas et al., "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[57] H. Xu, A. D. Domínguez-García, and P. W. Sauer, "Adaptive coordination of distributed energy resources in lossy power distribution systems," in *Proc. of IEEE Power & Energy Society General Meeting*, Aug 2018, pp. 1–5.

[58] J. Zhang and H. Xu, "Microperturbation method for power system online model identification," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1055–1063, 2016.

[59] J. Zhang and H. Xu, "Online identification of power system equivalent inertia constant," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8098–8107, 2017.

[60] A. Iusem, "On the convergence properties of the projected gradient method for convex optimization," *Computational & Applied Mathematics*, vol. 22, no. 1, pp. 37–52, 2003.

[61] B. Hajek, *Random Processes for Engineers*. Cambridge University Press, 2015.

[62] H. Robbins and D. Siegmund, "A convergence theorem for non negative almost supermartingales and some applications," in *Herbert Robbins Selected Papers*. Springer, 1985, pp. 111–135.

[63] D. P. Bertsekas, *Convex Optimization Theory*. Athena Scientific Belmont, 2009.

[64] O. Ajala, M. Almeida, P. W. I. Celanovic, Sauer, and A. D. Domínguez-García, "A hierarchy of models for microgrids with grid-feeding inverters," in *Proc. of X Bulk Power Systems Dynamics and Control Symposium*, Aug 2017, pp. 1–10.

[65] "IEEE distribution test feeders." [Online]. Available: http://sites.ieee.org/pes-testfeeders/resources/

[66] H. Xu, A. Domínguez-García, and P. W. Sauer, "Data-driven coordination of distributed energy resources for active power provision," *IEEE Transactions on Power Systems*, 2019.

[67] D. B. West et al., *Introduction to Graph Theory*. Prentice hall Upper Saddle River, 2001, vol. 2.

[68] R. B. Bapat, *Graphs and Matrices*. Springer, 2010, vol. 27.

[69] "Commercial and residential hourly load profiles for all TMY3 locations in the United States." [Online]. Available: https://openei.org/doe-opendata/dataset

[70] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784–2794, 2014.

[71] H. Xu, A. D. Domínguez-García, and P. W. Sauer, "A data-driven voltage control framework for power distribution systems," in *Proc. of IEEE Power & Energy Society General Meeting*, Aug 2018, pp. 1–5.

[72] H. Xu, A. D. Domínguez-García, V. V. Veeravalli, and P. W. Sauer, "Data-driven voltage regulation in radial power distribution systems," arXiv preprint arXiv:1901.03173, 2019.

[73] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT press, 2018.

[74] M. G. Lagoudakis and R. Parr, "Least-squares policy iteration," *Journal of Machine Learning Research*, vol. 4, pp. 1107–1149, Dec 2003.

[75] C. J. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[76] Y. Xu, W. Zhang, W. Liu, and F. Ferrese, "Multiagent-based reinforcement learning for optimal reactive power dispatch," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1742–1751, 2012.

[77] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[78] D. Ernst, P. Geurts, and L. Wehenkel, "Tree-based batch mode reinforcement learning," *Journal of Machine Learning Research*, vol. 6, pp. 503–556, Apr 2005.

[79] H. Xu, A. D. Domínguez-García, and P. W. Sauer, "Optimal tap setting of voltage regulation transformers using batch reinforcement learning," arXiv preprint arXiv:1807.10997, 2018.