STOCHASTIC OPTIMIZATION AND OPTIMAL CONTROL FOR COMPLEX NETWORKED INFRASTRUCTURES

by

Yi Guo

APPROVED BY SUPERVISORY COMMITTEE:

Tyler Summers, Chair

Yaoyu Li

Justin Koeln

Justin Ruths

Xinyang Zhou

To my parents.

STOCHASTIC OPTIMIZATION AND OPTIMAL CONTROL FOR COMPLEX NETWORKED INFRASTRUCTURES

by

YI GUO, BS, MS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MECHANICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

December 2020

ACKNOWLEDGMENTS

I am forever grateful for my experience and memories in the Control, Optimization, & Networks Laboratory (CONL) and the Department of Mechanical Engineering at The University of Texas at Dallas. I would first like to thank my advisor, Professor Tyler H. Summers, for offering me a great and valuable research opportunity, for his unwavering support, advice and encouragement in my research pursuits over the past four years, for giving me the perfect balance of guidance and autonomy that taught me to become a control and power researcher, for always holding me up to a high standard and for always acting as a role model in research and life in general. I am very fortunate to have had the opportunity to work with him. His passion, profession, and belief in control and power engineering always provided me inspiration and motivation. I could not have imagined having a better advisor for my graduate study.

I would like to thank the other members of my committee: Prof. Yaoyu Li, Prof. Justin Koeln, Prof. Justin Ruths and Dr. Xinyang Zhou. Their comments and suggestions have greatly helped improve my work. Special thanks to my mentors, Dr. Xinyang Zhou and Prof. Changhong Zhao, for offering a great summer internship position at National Renewable Energy Laboratory (NREL) in 2019, for always being encouraging and taking care of me, and for their patience, trust and endless support, which definitely made a major impact on my academic career path. I always learned from any research discussions we had. Also, I would like to thank other colleagues at NREL for a memorable and joyous summer in 2019. The beauty of the Rocky Mountains, and the time spent with friends in Denver will be never forgotten.

I want to thank Prof. Mario Rotea for a great collaboration on the wind farm power maximization. I am grateful to have this opportunity to work with him, which provided a deep impression on me. His research philosophy sets a role model to a PhD student in Mechanical Engineering, like me, demonstrating how a professor, as well as an experienced control

engineer, pose and solve a problem with a great balance in between the industrial practice and theoretical complexity.

I am grateful to my colleagues and co-authors around the world including, Kyri Baker, Emiliano Dall'Anese, Zechun Hu, Shen Wang, Ahmad Taha, Zhiyuan Liu, Lijun Chen, Chengda Ji, Dennice Gayme, Yue Chen, Chin-Yao Chang, for their fruitful discussions and collaborations. Some of them indeed worked as my mentors at different stages of my research.

I feel very fortunate to work with a group of intelligent colleagues at CONL and UT-Dallas, including Venkatraman Renganathan, Karthik Ganapathy, Benjamin Gravell, Sleiman Safaoui, Vignesh Raghuraman, Chang Liu, Devesh Kumar, Kaveh Fathian, I would like to thank them for intriguing discussions and brain storming. Their passion, profession and enthusiasm are always what I look up to.

Furthermore, I would like to thank all my friends for their unwavering support, for bringing such happiness and taking care of me, and for their company when I was suffering on my way to pursue my PhD degree; without them this great journey to be where I am now would not be possible. Special thanks go to: Lu Zhan, Zhongyou Wu, Yongpeng Huang, Changyuan Li, Xiaomeng Wang, Chao Pan, Lin Jiang, Fenglin Zhou, Zhongfan Zhao, Wenyi Wang, Peiyuan Kang, Liujia Dong, Yujie Zhang, Jianda Tang, Jonathan Horn, Matthew Patak, Wenqing Wang, Ge Lv, Xiaowei Zhu, Jianping Lin, Liangyu Jia, Chen Fang, Bo Chen, Kun-long Chen, Xuerui Lin, Xiaoyang Li, Kebin Ni, Minghui Zhu, Xiangqi Zhu and Xun Ouyang. Some of them are always acting a role model of my life and career in general.

Most importantly, I would like to give my deepest gratitude to my parents, my mother Yuping Xiong, my father Weijun Guo, and other family members for their endless love and support on my graduate studies, which never decreased even over thousands of miles distance. I could not appreciate them enough for everything they have done for me.

In closing, the past four years of research would not have been possible without the financial support from the National Science Foundation under grants CNS-1566127 and CMMI-1728605, the funding from US Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office under contract No. DE-EE-0007998, and the support from The University of Texas at Dallas Office of Research through the SCI program. I would like to thanks the UT-Dallas Center for Wind Energy for its academic atmosphere and potential opportunities.

October 2020

STOCHASTIC OPTIMIZATION AND OPTIMAL CONTROL FOR COMPLEX NETWORKED INFRASTRUCTURES

Yi Guo, PhD
The University of Texas at Dallas, 2020

Supervising Professor: Tyler Summers, Chair

Networked infrastructures (e.g., power grids and water networks) are becoming increasingly complicated systems due to accelerating integration of distributed devices, increasing scale of network systems, heterogeneous distributed control strategies, nonlinear nature of physical models and unpredictable netloads (e.g., renewable energy and water demands). These tendencies promise to deliver unprecedented flexibility and smart features in these systems, but require more sophisticated modelling and operation schemes. In particular, the large and unpredictable uncertainties across networks are challenging the current management and threaten the security of these vital infrastructures.

The main goal of this dissertation is to establish a methodology for designing stochastic operation strategies in complex networks with large variation uncertainties, for optimal performance, stability and robustness, which is applicable to a wide range of networked infrastructures. This dissertation addresses the main goal in various time-scales, ranging from the operation phase in minutes to the dynamic phase in seconds.

In the operation phase, we present a data-based distributionally robust stochastic optimal control framework to attain the real-time optimal adjustment for controllable devices based on a finite training dataset of uncertainties. We consider *ambiguity sets* of probability distributions centered around the finite sampling dataset, and leverage Wasserstein metric to

quantify the distance between the empirical data-based distribution and the real unknown data-generating distribution. This allows the optimal decisions to be robust to the *worst-case* probability distribution within the *ambiguity sets*, which efficiently trades off efficiency, risks of constraint violations and out-of-sample performance. This proposed framework is adapted to solve the stochastic optimal power flow (OPF) problems for power systems (i.e., transmission systems and distribution networks) and the stochastic optimal water pump control for water distribution networks. In addition, we propose a gradient-based optimal control algorithm with state estimation in the loop to facilitate the practical implementation of OPF problems in large-scale distribution networks, where has inadequate monitoring and unreliable measurement infrastructures.

In the dynamic phase, we consider the linear system with multiplicative noise and additive noise, which captures the modelling errors and netload uncertainties, respectively. We adapt this model to study the performance and stability of low-inertia power grids with stochastic system inertia. An analytical expression of system $\mathcal{H}_2$ norm is provided and a mean-square stability criteria is developed. In addition, we present a multi-stage stochastic optimal control problem for wind farm power maximization. We generalize the original actuator disk model (ADM) by incorporating state- and input-dependent multiplicative noises as a stochastic actuator disk model (S-ADM) to capture the stochastic wind fluctuations. The optimal control policies for each wind turbine explicitly incorporate the moment information of multiplicative noise, which establishes a connection between uncertainties sampling dataset and optimal feedback control.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The rapidly growing city modernization and industrialization 4.0 have put advanced requirement on the current vital municipal networked infrastructures (e.g., power grids and water distribution networks). These complex networked infrastructures become more sophisticated physical systems by accelerating integration of distributed self-organized multi-agents, increasing scale of network systems, heterogeneous and distributed control strategies, nonlinearity nature of physical coupling and unpredictable stochastic properties. These tendencies will definitely deliver many smart features and flexibility for system operators, but obviously aggravate the stochasticity of these complex systems, which are challenging the current management and control strategies and may threaten the security of these vital infrastructures. One of the main challenges is how to explicitly balance the operation efficiency, robustness, stability and complexity if the networked systems are under large variation uncertainties.

The main goal of this research dissertation is to address the aforementioned challenge and establish a methodology for designing stochastic operation scheduling and optimal feedback controllers in complex networked infrastructures. Our approaches explicitly incorporate the probability distribution information of uncertainties based on real-time sampling and training dataset. The dissertation utilizes distributionally robust optimization and optimal feedback control techniques to address the main goal in various time-scales, ranged from the operation phase in minutes to the dynamical behavior in seconds.

In the operation phase. We use distributionally robust optimization (DRO) techniques to achieve the controllable conservativeness of optimal decisions. Numerous recent studies have explored aspects of stochastic control in complex networks, but there is a lack of the controllable robustness to inherent sampling errors, which has inadequate understanding and quantification of overestimation/underestimation to risk of operational constraint violations. The works in this dissertation provide a fundamental data-driven discussion on the inherent

tradeoffs between operational risk and efficiency for effective overfitting regularization, and guarantee the superior out-of-sample performance for complex networked infrastructures under large variation uncertainties.

In the dynamic phase. we consider a generic linear system with both additive and multiplicative noise to appropriately capture the system stochasticity of the netload uncertainties (i.e., additive noise) and system modelling errors (i.e., multiplicative noise). The open-loop performance and the system mean-square stability are investigated under various network settings (e.g., network topologies, distribution parameters of disturbances and system parameters) in power grids. A closed-loop optimal controller is designed for wind turbine control in the objective of wind farm power maximization, where the multiplicative noise captures the stochastic wind fluctuations.

1.1 Distributionally Robust Optimization

Distributionally robust optimization (DRO) is a class of state-of-art optimization techniques to solve a wide variety of real-life decision making problems, which naturally come with random parameters. The optimization problems under uncertainties have been discussed over years since 1950's, which are traditionally solved via stochastic and dynamic programming. In practice, the uncertain parameters are inherently subjected to measurement errors or partial unobservable at the planning stage. These factors are challenging our traditional methods for decisions making to have disappointing robustness and poor out-of-sample performance, where the situation is even worse if the size of decision problems have been substantially grown.

Nowadays, a new field of distributionally robust optimization is tackling these challenges in a different way. The key idea of such methodologies is to consider the uncertainty through a *ambiguity set* consisting a group of probability distributions (possibly infinite candidates), supported by finite sampling data or prior structural information. This underlying setting of

DRO comes with a min-max problem to prompt the optimal decision robust to the *worst-case* distribution within the ambiguity set for various optimization objectives (e.g., operational efficiency, risk of constraint violations and etc.).

There are several striking benefits of having DRO formulation for stochastic decision making. First of all, DRO formulations explicitly incorporate the inherent measurement errors into stochastic optimization problems based on data sampling or structural information. This exploits a more realistic modelling of uncertainties than the traditional stochastic and dynamic programming. In addition, the optimal decisions are robust to a family of distributions, instead of making prescribed distributions, which provides a good robust performance. Finally, many DRO problems can be solved completely in low complexity and always have tractability guarantee. This provides a great potential for practical adaption to many large-scale industrial applications and real-life decision making problems.

1.1.1 Related Works in Distributionally Robust Optimization

Distributionally robust optimization has been widely discussed in both operation research and machine learning communities, for their modelling, solutions and applications. Most recent research works focused on the several directions. In the literature, two well-known groups of ambiguity sets, *moment-based* ambiguity sets and *metric-based* ambiguity sets, characterize the specific family of distributions. In general, the moment-based ambiguity sets contain distributions sharing same moments' characters. The metric-based ambiguity sets collect all the distributions in the distance to the nominal distribution in the sense of various discrepancy measure.

In the literature, the moment-based distributionally robust techniques model the ambiguity sets whose satisfy certain moments properties [14, 101, 108, 129, 164, 173, 111, 163, 81, 34, 311, 57, 158]. As for the metric-based distributionally robust techniques, the options of probability metric to model the distributional ambiguity includes: ϕ-divergence [220],

Kolmogorov-Smirnov [272], Wasserstein [96, 70], Anderson-Darling [25], Cramer-von Mises [272], Waston [272] and Kuiper [272]. The literature also consider a variety of risk measures such as 1) Conditional Value-at-Risk (CVAR) [96], 2) Value-at-Risk (VaR) [272], 3) variance [25], 4) absolute deviation of mean from the median [279], 5) negative mean return [45] and etc. For the applications, the DRO problems have been studied in statistical learning [186, 197, 133, 156, 103, 97], energy systems [269, 247, 310, 282, 239, 300, 89, 157], water resource management [205], intelligent transportation [127], blood supply network [263] and robotics [230, 246, 128, 223, 221, 222]. In this dissertation, we are mostly focusing on the applications of distributionally robust optimization for particular stochastic optimization and optimal control tasks in power systems and water distribution networks.

1.1.2 Contributions

The contributions of this dissertation along this line was inspired by [96, 49] and summarized below:

- We extend the recent tractable reformulations of distributionally robust optimization, moment-based DRO [49] and Wasserstein-based DRO [96], to a multi-stage setting as the multi-stage distributionally robust stochastic optimal control framework, whereas mostly of the recent works focus on the single-stage phase; This allows the model predictive controllers to repeatedly update the ambiguous distributional information based on the latest realization of uncertainties, and then explicitly make real-time optimal decisions.

- We explore the applications of the proposed multi-stage DRO in power systems to formulate the stochastic optimal power flow problems with high penetration of renewables.

- We also adapt the proposed multi-stage DRO in water distribution networks to have optimal pump control with large variations of water demands.

- The proposed optimal control framework and applications formulations are with tractability and convexity for low-complexity computation, which can be solved efficiently through the existing commercial solvers.

1.2 Linear Systems with Multiplicative Noise

The systems with multiplicative noise are a special case of stochastic systems, which are more general and practical than the classic linear system with only additive noises. This opens a door to incorporate the statistical description of the noise in both state and input channels for analysis and optimal design. The study of linear systems with multiplicative noise has a long history since 1960's [149, 154, 275, 13]. Such dynamical models can be adapted to many physical systems with applications to signal process systems [107, 264], biological movement systems [130, 251], fluid dynamics [123], power and energy systems [188, 125, 201, 141], robotics, aerospace engineering systems [168, 170] and statistical learning [286, 69, 115, 114].

One important benefit of having multiplicative noise modelling in a stochastic system is that it connects the performance and robustness of dynamic systems to the statistical information of multiplicative noise in a closed-form expression via generalized Lyapunov equations and Riccati equation, for stability and optimal control [43]. In the past, many researches have explored various analysis and discussions on stability and stabilization of stochastic multiplicative uncertainties in the sense of mean-square stability [283, 216, 273, 131, 177, 93]. There are also many interesting design of control and filtering of systems with multiplicative noise, including but not limited to LQR problems [276, 69, 296, 161, 170, 145, 40, 56, 134, 219, 298], learning-based control [115, 114], $\mathcal{H}_\infty$ control [107, 92], estimation and filtering [168, 107, 288, 242], model predictive control (MPC) framework [162, 52, 104] and game theory [116]. In this dissertation, we mostly focus on the applications of stochastic multiplicative modelling to power grids and wind farms, which also come with performance, mean-square stability and optimality analysis.

1.2.1 Contributions

The contributions of this dissertation along this line are summarized as below

- We study the performance and stability of linear systems with multiplicative noise in power grids with stochastic inertia setting. The mean-square stability of frequency dynamics is studied and a closed-form stability criteria is provided.

- We design the optimal feedback controllers for linear systems with multiplicative noise, where the state- and input-dependence noise capture the wind fluctuation for wind farm power maximization. The optimal feedback control policies explicitly incorporate the probability information of multiplicative noise for optimal closed-loop performance.

1.3 Structure of this Dissertation

The structure of the rest of this dissertation are listed below. All chapters can be read individually based on the readers' interests and background.

Chapter 2 proposes a data-based distributionally robust stochastic optimal control framework. Such formulation is applied to solve a stochastic optimal power flow (OPF) in power grids, which explicitly adjust the real-time economic dispatch decisions for controllable devices (e.g., distributed generators and inverters), considering the stochasticity of solar PVs in distribution networks and wind outputs in transmission systems.

Chapter 3 adapts the data-based distributionally robust stochastic optimal control to solve a stochastic optimal water flow (OWF) problem for water distribution networks. The objective is to determine the nominal pump schedule and tank levels with reaction to forecast errors for accommodation of fluctuating water demand, which explicitly tradeoffs pump operational costs, risks of tank level constraint violation and out-of-sample performance.

Chapter 4 proposes an optimal control framework with tightly integrating state estimation for power system distribution networks. Our approach solves an OPF problem via a

primal-dual algorithm with a SE feedback loop based on a limited sensors measurement. This approach can scale to extremely large distribution networks (i.e., 11.000-nodes), which facilitates a practical implementation scenario under limited communication infrastructures and inadequate node-wise measurement. The numerical analysis of having SE feedback loop promotes its excellent robustness to measurement noises on the proposed OPF solvers.

Chapter 5 investigates the performance and stability of low-inertia power grids with stochastic system inertia. We consider system frequency dynamics modelled by a linearized stochastic swing equation, where stochastic system inertia is regarded as multiplicative noises. We develop a closed-form expression of $\mathcal{H}_2$ norms for the stochastic swing equation, which is closely connected to the grid topology, system parameters and probability distribution of disturbances. This discussion has fundamentally different characteristics from systems with only additive noises.

Chapter 6 formulates a multi-stage stochastic optimal control problem for wind farm power maximization, which explicitly incorporates information about the probability distributions of wind fluctuations into control decisions. In particular, our model incorporates state- and input-dependent multiplicative noise whose distributions capture stochastic wind fluctuations. The optimal feedback control policies for each turbine can be attained by analytically solving the proposed problem via stochastic dynamic programming. This provides a direct connection between statistical properties of the unsteady wind flow physics and the optimal feedback control of wind farms.

Chapter 7 concludes this dissertation.

CHAPTER 2

DATA-BASED DISTRIBUTIONALLY ROBUST STOCHASTIC OPTIMAL POWER FLOW[1]

In this chapter, we propose a data-based method to solve a multi-stage stochastic optimal power flow (OPF) problem based on limited information about forecast error distributions. The framework explicitly combines multi-stage feedback policies with any forecasting method and historical forecast error data. The objective is to determine power scheduling policies for controllable devices in a power network to balance operational cost and conditional value-at-risk (CVaR) of device and network constraint violations. These decisions include both nominal power schedules and reserve policies, which specify planned reactions to forecast errors in order to accommodate fluctuating renewable energy sources. Instead of assuming the uncertainties across the networks follow prescribed probability distributions, we consider *ambiguity sets* of distributions centered around a finite training dataset. By utilizing the Wasserstein metric to quantify differences between the empirical data-based distribution and the real unknown data-generating distribution, we formulate a multi-stage distributionally robust OPF problem to compute control policies that are robust to both forecast errors and sampling errors inherent in the dataset. Two specific data-based distributionally robust stochastic OPF problems are proposed for distribution networks and transmission systems.

---

[1]This chapter is based on work supported by the National Science Foundation (NSF) under grant CNS-1566127.

2.1 Introduction

The continued integration of renewable energy sources (RESs) in power systems is making it more complicated for system operators to balance economic efficiency and system reliability and security. As penetration levels of RESs reach substantial fractions of total supplied power, networks will face high operational risks under current operational paradigms. As it becomes more difficult to predict the net load, large forecast errors can lead to power security and reliability issues causing significant damage and costly outages. Future power networks will require more sophisticated methods for managing these risks, at both transmission and distribution levels.

The flexibility of controllable devices, including power-electronics-interfaced RESs, can be utilized to balance efficiency and risk with optimal power flow methods [54, 85, 9, 18, 175, 176, 180, 167, 166], which aim to determine power schedules for controllable devices in a power network to optimize an objective function. However, most OPF methods in the research literature and those widely used in practice are deterministic, assuming point forecasts of exogenous power injections and ignoring forecast errors. Increasing forecast errors push the underlying distributed feedback controllers that must handle the transients caused by these errors closer to stability limits [153].

More recently, research focus has turned to stochastic and robust optimal power flow methods that explicitly incorporate forecast errors, in order to more systematically trade off economic efficiency and risk and to ease the burden on feedback controllers [291, 53, 68, 39, 260, 262, 270, 301, 208, 225, 248, 136, 224, 226, 157, 297, 178, 16, 73, 261, 160, 174, 256, 285]. Many formulations assume that uncertain forecast errors follow a prescribed probability distribution (commonly, Gaussian [39, 178, 225, 256]) and utilize analytically tractable reformulations of probabilistic constraints. However, such assumptions are unjustifiable due to increasingly complex, nonlinear phenomena in emerging power networks, and can significantly underestimate risk.

In practice, forecast error probability distributions are never known; they are only observed indirectly through finite datasets. Sampling-based methods have been applied with a focus on quantifying the probability of constraint violation [260, 262] and for constraining or optimizing conditional value at risk (CVaR) [301, 248, 73]. The prediction-realization approach [190, 189] solves an online stochastic optimal power flow problem by a reconciliation algorithm, which ensures feasibility for any forecast error. Distributionally robust approaches use data to estimate distribution parameters (e.g., mean and variance) and aim to be robust to any data-generating distribution consistent with these parameters [248, 226, 157, 16, 73, 285]. Others take a robust approach, assuming only knowledge of bounds on forecast errors and enforcing constraints for any possible realization, e.g., [270, 136]. Overall, this line of recent research has explored tractable approximations and reformulations of difficult stochastic optimal power flow problems. However, none of the existing work explicitly accounts for sampling errors arising from limited data, which in operation can cause poor out-of-sample performance[2]. Even with sophisticated recent stochastic programming techniques, decisions can be overly dependent on small amounts of relevant data from a high-dimensional space, a phenomenon akin to overfitting in statistical models.

We propose a multi-period data-based method to solve a stochastic optimal power flow problem based on limited information about forecast error distributions available through finite historical training datasets. A preliminary version of this work appeared in [121], and here we significantly expand the work in several directions into this chapter. The main contributions are as follows

1. We formulate a multi-stage distributionally robust optimal control problem for optimal power flow. A distributionally robust model predictive control algorithm is then proposed, which utilizes computationally tractable data-driven distributionally robust

[2]Out-of-sample performance is an evaluation of the optimal decisions using a dataset that is different from the one used to obtain the decision, which can be tested with Monte Carlo simulation

optimization techniques [96] to solve the subproblems at each stage. Whereas distributionally robust optimization approaches focus on single-stage problems, here we extend these approaches to multi-stage settings to obtain closed-loop feedback control policies. This allow us to update forecast error datasets, and in turn re-compute decisions with the latest knowledge. In principle, the framework allows any forecasting methodology and a variety of ambiguity set parameterizations. We focus on Wasserstein balls [109] around an empirical data-based distribution [109, 96], which allows controllable conservativeness by adjusting the Wasserstein radius. In contrast to previous work, we obtain policies that are explicitly robust to sampling errors inherent in the dataset. This approach achieves superior out-of-sample performance guarantees in comparison to other stochastic optimization approaches, effectively regularizing against overfitting the decisions to limited available data.

2. We leverage pertinent linear approximations of the AC power-flow equations (see, e.g., [20, 42, 120, 61, 29]) to facilitate the development of computationally-affordable chance-constrained AC OPF solutions that are robust to distribution mismatches, and provide a unified framework that is applicable to both transmission and distribution systems. Formulations for distribution networks incorporate inverted-based RESs and energy storage systems, and focus on addressing the voltage regulation problem under uncertainty. The transmission system formulation incorporates synchronous generators and power injections from RESs, and it focuses on probabilistic $N-1$ security constraints on active transmission line flows. The framework yields set points and feedback control policies for controllable devices that are robust to variations in solar and wind injections and sampling errors inherent to the finite training datasets.

3. The effectiveness and flexibility of the proposed methodologies are demonstrated with extensive numerical experiments in a 37-node distribution network and in a 118-bus

transmission system. These extensive case studies are presented in Part II [122]. We demonstrate inherent tradeoffs between economic efficiency and robustness to constraint violations and sampling errors due to forecasting. By explicitly incorporating forecast error and sampling uncertainties, the methodology can help network operators to better understand these risks and inherent tradeoffs, and to design effective optimization and control strategies for appropriately balancing efficiency objectives with security requirements.

The rest of the chapter is organized as follows: Section 2.2 presents a general formulation of the proposed data-based distributionally robust stochastic OPF problem. Section 2.3 describes the modeling of network, grid-connected components, and network constraints. Section 2.4 specializes the data-based stochastic OPF for distribution networks with an approximate AC power flow. Section 2.5 adapts the proposed methods for transmission systems with DC power flow. Section 2.6 illustrates the data-based distributionally robust stochastic AC OPF for mitigating overvoltages in a modified IEEE 37-node distribution feeder with high PV penetration using local energy storage devices. Section 2.7 illustrates the data-based distributionally robust stochastic DC OPF for reducing N-1 security line flow constraint risks due to high wind penetration using reserve policies for controllable generators. Section 2.8 concludes the chapter.

Notation: The inner product of two vectors $a, b \in \mathbf{R}^m$ is denoted by $\langle a, b \rangle := a^\intercal b$. The N_s-fold product of distribution $\mathbb{P}$ on a set Ξ is denoted by $\mathbb{P}^{N_s}$, which represents a distribution on the Cartesian product space $\Xi^{N_s} = \Xi \times \ldots \times \Xi$. We use N_s to represent the number of samples inside the training dataset $\hat{\Xi}$. Superscript "$\hat{\cdot}$" is reserved for the objects that depend on a training dataset $\hat{\Xi}^{N_s}$. We use $(\cdot)^\intercal$ to denote vector or matrix transpose. The operators $\Re\{\cdot\}$ and $\Im\{\cdot\}$ return the real and imaginary part of a complex number, respectively. The operator $[\,\cdot\,]_{[a,b]}$ selects the a-th to b-th elements of a vector or rows of a matrix.

2.2 Stochastic OPF as Stochastic Optimal Control

In this section, we formulate a stochastic OPF problem as a distributionally robust stochastic optimal control problem. We first pose the problem generically to highlight the overall approach, and in subsequent sections we detail the model and objective and constraint functions more explicitly for both distribution networks and transmission systems. This framework is more general than most stochastic OPF and distributionally robust optimization approaches in the literature, which typically focus only on individual or single-stage optimization problems.

Let $x_t \in \mathbf{R}^n$ denote a state vector at time t that includes the internal states of all devices in the network. Let $u_t \in \mathbf{R}^m$ denote a control input vector that includes inputs for all controllable devices in the network. Let $\xi_t \in \mathbf{R}^{N_\xi}$ denote a random vector in a probability space $(\Omega, \mathcal{F}, \mathbb{P}_t)$ that includes forecast errors of all uncertainties in the network. Forecast error distributions are never known in practice, so $\mathbb{P}_t$ is assumed to be *unknown* but belonging to an *ambiguity set* $\mathcal{P}_t$ of distributions with a known parameterization, which will be discussed in detail shortly. We define the concatenated forecast error over an operating horizon T as $\boldsymbol{\xi}_{0:T} := [\xi_0^\intercal, \ldots, \xi_T^\intercal]^\intercal \in \mathbf{R}^{N_\xi T}$, which has joint distribution $\mathbb{P}$ and corresponding ambiguity set $\mathcal{P}$.

Since forecast errors are explicitly included, we seek closed-loop feedback policies of the form $u_t = \pi(x_0, \ldots, x_t, \boldsymbol{\xi}_{0:t}, \mathcal{D}_t)$, where the term $\mathcal{D}_t$ collects all network and device model information and the parameterization of the ambiguity set of the forecast error distribution. The control decisions at time t are allowed to be functions of the entire state and forecast error history up to time t; this is called a history-dependent state and disturbance feedback information pattern. This general formulation allows for the design of not only nominal plans for controllable devices inputs, but also for planned reactions to forecast errors as they

are realized.[3] The policy function π maps all available information to control actions and is an element of a set Π of measurable functions.

This leads us to the following multi-stage distributionally robust stochastic optimal control problem

$$\inf_{\pi\in\Pi}\sup_{\mathbb{P}\in\mathcal{P}} \quad \mathbb{E}^{\mathbb{P}}\sum_{t=0}^{T} h_t(x_t,u_t,\xi_t), \tag{2.1a}$$

$$\text{subject to} \quad x_{t+1} = f_t(x_t,u_t,\xi_t), \tag{2.1b}$$

$$u_t = \ \pi(x_0,\ldots,x_t,\boldsymbol{\xi}_t,\mathcal{D}_t), \tag{2.1c}$$

$$(x_t,u_t)\in\mathcal{X}_t. \tag{2.1d}$$

The goal is to compute a feedback policy that minimizes the expected value of an objective function $h_t : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{N_\xi} \to \mathbf{R}$ under the worst-case distribution in the forecast error ambiguity set $\mathcal{P}$. The objective function h_t will include both operating costs and risks of violating various network and device constraints and is assumed to be continuous and convex for every fixed ξ_t. The system dynamics function $f_t : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{N_\xi t} \to \mathbf{R}^n$ models internal dynamics and other temporal interdependencies of devices, such as state of charge for batteries and ramp limits of generators. The constraint set $\mathcal{X}_t$ includes network and device constraints, such as power balance and generator and storage device bounds (some constraints may be modeled deterministically and others may be included as risk terms in the objective function).

The main challenges to solving (2.1) are the multi-stage feedback nature of the problem, the infinite dimensionality of the control policies, the possible nonlinearity of device dynamics, and how to appropriately parameterize and utilize our available knowledge about forecast error distributions. We will tackle these using a distributionally robust model predictive

[3]The reactions can be interpreted as pre-planned secondary frequency control allocations [270] or contingency reactions in response to forecast errors based on device dynamics and parameters describing forecast error distributions.

control scheme with affine feedback policies and linear models for device dynamics, where stage-wise distributionally robust stochastic optimization problems are repeatedly solved over a planning horizon. To incorporate forecast error knowledge, we will use tractable reformulations of ambiguity sets based explicitly on empirical training datasets of forecast errors.

2.2.1 Ambiguity Sets and Wasserstein Balls

There is a variety of ways to reformulate the general stochastic OPF problem (2.1) to obtain tractable subproblems that can be solved by standard convex optimization solvers. These include assuming specific functional forms for the forecast error distribution (e.g., Gaussian) and using specific constraint risk functions, such as those encoding value at risk (i.e., chance constraints), conditional value at risk (CVaR), distributional robustness, and support robustness. In all cases, the out-of-sample performance of the resulting decisions in operational practice ultimately relies on 1) the quality of data describing the forecast errors and 2) the validity of assumptions made about probability distributions. Many existing approaches make either too strong or too weak assumptions that possible lead to underestimation or overestimation of risk. In this chapter, we extend a recently proposed tractable method [96] to a multi-period data-based stochastic OPF, in which the ambiguity set is based on a finite forecast error training dataset $\hat{\Xi}^{N_s}$.

Within the area of distributionally-robust optimization, moment-based ambiguity sets are utilized to model distributions featuring specified moment constraints such as unimodality [157, 158], directional derivatives [272], symmetry [226], and log-concavity [159]. The ambiguity sets can be defined as confidence regions based on goodness-of-fit tests [35]. Another line of works considers ambiguity sets as balls in the probability space, with radii computed based on the Wasserstein metric [278], the Kullback-Leilber divergence [137], and the Prohorov metric [95]. This chapter formulates ambiguity sets by leveraging Wasserstein

balls. Relative to other approaches, Wasserstein balls provide an upper confidence bound, quantified by Wasserstein radius ε [96], to achieve the superior out-of-sample performance; they also enable power system operators to "control" the conservativeness of the solution, thus ensuring the flexibility in the power system operation. Additionally, the approach in this chapter seeks the worst-case expectation subjected to all distributions contained in the ambiguity set. The worst-case expectation of the stochastic OPF problem over Wasserstein ambiguity set can be reformulated as a finite-dimensional convex problem, and can be solved using existing convex optimization solvers.

The Wasserstein metric defines a distance in the space $\mathcal{M}(\Xi)$ of all probability distributions $\mathbb{Q}$ supported on a set Ξ with $\mathbb{E}^{\mathbb{Q}}[\|\xi\|] = \int_{\Xi} \|\xi\| \mathbb{Q}(d\xi) < \infty$. In this chapter, we assume the support set is polytopic of the form the uncertainty set is a polytope $\Xi := \{\xi \in \mathbf{R}^{N_\xi} : H\xi \leq \mathbf{d}\}$.

Definition [Wasserstein Metric]. Let $\mathcal{L}$ be the space of all Lipschitz continuous functions $f : \Xi \to \mathbf{R}$ with Lipschitz constant less than or equal to 1. The Wasserstein metric $d_W : \mathcal{M}(\Xi) \times \mathcal{M}(\Xi) \to \mathbf{R}$ is defined $\forall \mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}(\Xi)$ as

$$d_W(\mathbb{Q}_1, \mathbb{Q}_2) = \sup_{f \in \mathcal{L}} \left(\int_{\Xi} f(\xi) \mathbb{Q}_1(d\xi) - \int_{\Xi} f(\xi) \mathbb{Q}_2(d\xi) \right).$$

Intuitively, the Wasserstein metric quantifies the minimum "transportation" cost to move mass from one distribution to another. We can now use the Wasserstein metric to define an ambiguity set

$$\hat{\mathcal{P}}^{N_s} := \left\{ \mathbb{Q} \in \mathcal{M}(\Xi) : d_w(\hat{\mathbb{P}}^{N_s}, \mathbb{Q}) \leq \varepsilon \right\}, \tag{2.2}$$

which contains all distributions within a Wasserstein ball of radius ε centered at a uniform empirical distribution $\hat{\mathbb{P}}^{N_s}$ on the training dataset $\hat{\Xi}^{N_s}$. The radius ε can be chosen so that the ball contains the true distribution $\mathbb{P}$ with a prescribed confidence level and leads to performance guarantees [96]. The radius ε also explicitly controls the conservativeness of the resulting decision. Large ε will produce decisions that rely less on the specific features

of the dataset $\hat{\Xi}^{N_s}$ and give better robustness to sampling errors. This parameterization will be used in the next subsection to formulate a distributionally robust optimization.

2.2.2 Data-based Distributionally Robust Model Predictive Control

The goal of our data-based distributionally robust stochastic OPF framework is to understand and to illustrate inherent tradeoffs between efficiency and risk of constraint violations. Accordingly, the objective function comprises a weighted sum of an operational cost function and a constraint violation risk function: $h_t = J^t_{\text{Cost}} + \rho J^t_{\text{Risk}}$, where $\rho \in \mathbf{R}_+$ is a weight that quantifies the network operator's risk aversion. The operational cost function is assumed to be linear or convex quadratic

$$J^t_{\text{Cost}}(x_t, u_t) := f_x^\intercal x_t + \frac{1}{2} x_t{}^\intercal H_x x_t + f_u^\intercal u_t + \frac{1}{2} u_t{}^\intercal H_u u_t,$$

where H_x and H_u are positive semidefinite matrices. This function can capture several objectives including thermal generation costs, ramping costs, and active power losses.

The risk function J_{Risk} associated with the constraint violation comprises a sum of the conditional value-at-risk (CVaR) [227] of a set of N_ℓ network and device constraint functions; specifically, it is defined as

$$J^t_{\text{Risk}} := \sum_{i=1}^{N_\ell} \text{CVaR}^\beta_{\mathbb{P}}[\ell_i(x_t, u_t, \xi_t)],$$

where $\beta \in (0, 1]$ refers to the confidence level of the CVaR under the distribution $\mathbb{P}$ of the random variable ξ_t. Intuitively, the constraint violation risk function J_{Risk} could be understood as the sum of networks and devices constraint violation magnitude at a "risk level" β, which penalizes both frequency and expected severity of constraint violations [227]. Further details will be provided in subsequent sections.

Data-based distributionally robust model predictive control for stochastic OPF. The general problem (2.1) will be approached with a distributionally robust model

predictive control (MPC) algorithm. MPC is a feedback control technique that solves a sequence of open-loop optimization problems over a planning horizon $\mathcal{H}_t$ (which in general may be smaller than the overall horizon T). At each time t, we solve the distributionally robust optimization problem over a set Π_{affine} of affine feedback policies using the Wasserstein ambiguity set (3.14)

$$\inf_{\pi \in \Pi_{\text{affine}}} \sup_{\mathbb{P} \in \hat{\mathcal{P}}^{N_s}} \quad \mathbb{E}^{\mathbb{P}} \sum_{\tau=t}^{t+\mathcal{H}_t} J^{\tau}_{\text{Cost}}(x_\tau, u_\tau, \xi_\tau) + \rho J^{\tau}_{\text{Risk}}(x_\tau, u_\tau, \xi_\tau), \tag{2.3a}$$

$$\text{subject to} \quad x_{\tau+1} = f_\tau(x_\tau, u_\tau, \xi_\tau), \tag{2.3b}$$

$$u_\tau = \pi(x_0, \ldots, x_\tau, \boldsymbol{\xi}_\tau, \mathcal{D}_\tau), \tag{2.3c}$$

$$(x_\tau, u_\tau) \in \mathcal{X}_\tau. \tag{2.3d}$$

Only the immediate control decisions for time t are implemented on the controllable device inputs. Then time shifts forward one step, new forecast errors and states are realized, the optimization problem (2.3) is re-solved at time $t+1$, and the process repeats. This approach allows any forecasting methodology to be utilized to predict uncertainties over the planning horizon. Furthermore, the forecast error dataset $\hat{\mathbb{P}}^{N_s}$, which defines the center of the ambiguity set $\hat{\mathcal{P}}^{N_s}$, can be updated online as more forecast error data is obtained. It is also possible to remove outdated data online to account for time-varying distributions.

In the rest of the chapter, we will derive specific models for both distribution and transmission networks and grid devices where the subproblems (2.3) have exact tractable convex reformulations as quadratic programs [96] and can be solved to global optimality with standard solvers.

2.3 System Model

We now consider a symmetric and balanced electric power network model in steady state, where all currents and voltages are assumed to be sinusoidal signals at the same frequency.

2.3.1 Network Model

Consider a power network (either transmission or distribution), denoted by a graph $(\mathcal{N}, \mathcal{E})$, with a set $\mathcal{N} = \{1, 2, ..., N\}$ of buses, and a set $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ of the power lines connecting buses. Let $V_i^t \in \mathbf{C}$ and $I_i^t \in \mathbf{C}$ denote the phasors for the line-to-ground voltage and the current injection at node $i \in \mathcal{N}$. Define the complex vectors $\mathbf{v}^t := [V_1^t, V_2^t, ..., V_N^t]^\intercal \in \mathbf{C}^N$ and $\mathbf{i}^t := [I_1^t, ..., I_N^t] \in \mathbf{C}^N$. Let z_{ij} denote the complex impedance of the line between bus i and bus j, then the line admittance is $y_{ij} = 1/z_{ij} = g_{ij} + \mathrm{j}b_{ij}$. We model the lines using a standard Pi Model. The admittance matrix $\mathbf{Y} \in \mathbf{C}^{N \times N}$ has elements

$$Y_{ij} = \begin{cases} \sum_{l \sim i} y_{il} + y_{ii} & \text{if } i = j \\ -y_{ij} & (i,j) \in \mathcal{E} \\ 0 & (i,j) \notin \mathcal{E} \end{cases}, \tag{2.4}$$

where $l \sim i$ indicates that bus i and bus l are connected. Via Kirchoff's and Ohm's laws, we have $\mathbf{i}^t = \mathbf{Y}\mathbf{v}^t$. Net complex power bus injections are given by

$$\mathbf{s}^t = \mathbf{v}^t \left(\mathbf{i}^t\right)^* = \mathrm{diag}\left(\mathbf{v}^t\right)\left(\mathbf{Y}\mathbf{v}^t\right)^*. \tag{2.5}$$

The components of $\mathbf{s}^t = [S_1^t, S_2^t, \ldots, S_N^t]^\intercal \in \mathbf{C}^N$ can be expressed in rectangular coordinates as $S_i^t = P_i^t + \mathrm{j}Q_i^t$, where P_i^t is active power and Q_i^t is reactive power. Positive P_i and Q_i means that bus i generates active/reactive power, and negative P_i and Q_i mean that bus i absorbs the active/reactive power. Vectors of active and reactive power $\mathbf{p}^t = [P_1^t, P_2^t, \ldots, P_N^t]^\intercal$ and $\mathbf{q}^t = [Q_1^t, Q_2^t, \ldots, Q_N^t]^\intercal$ are further divided into nominal and error terms: $\mathbf{p}^t = \bar{\mathbf{p}}^t(u_t) + \tilde{\mathbf{p}}^t(\xi_t)$ and $\mathbf{q}^t = \bar{\mathbf{q}}^t(u_t) + \tilde{\mathbf{q}}^t(\xi_t)$. The nominal active and reactive power injection vectors $\bar{\mathbf{p}}^t(u_t) \in \mathbf{R}^N$ and $\bar{\mathbf{q}}^t(u_t) \in \mathbf{R}^N$ depend on control decisions, and the forecast errors $\tilde{\mathbf{p}}^t(\xi_t) \in \mathbf{R}^N$ and $\tilde{\mathbf{q}}^t(\xi_t) \in \mathbf{R}^N$ depend on the random vector ξ_t.

To handle nonconvexity of the power flow equations (2.5), we utilize two different linearization methods that are effective in both distribution and transmission networks.

2.3.2 Dynamic of Grid-connected Devices

We consider N_d grid-connected devices, which may include 1) traditional generators and inverter-based RESs; 2) fixed, deferrable, and curtailable loads; 3) storage devices like batteries and plug-in electric vehicles, which are able to act as both generators and loads. There are two types of devices: devices with controllable power flow affected by decision variables (e.g., conventional thermal and RESs generators, deferrable/curtailable loads and storage devices); and devices with fixed or uncertain power flow which will not be affected by decision variables (e.g., fixed loads). The power flow of each controllable device is modeled by a discrete-time linear dynamical system

$$x_{t+1}^d = \bar{A}^d x_t^d + \bar{B}^d u_t^d,$$

where device d at time t has internal state $x_t^d \in \mathbf{R}^{n_d}$, dynamics matrix $\bar{A}^d \in \mathbf{R}^{n_d \times n_d}$, input matrix $\bar{B}^d \in \mathbf{R}^{n_d \times m_d}$, and control input $u_t^d \in \mathbf{R}^{m_d}$. The first element of x_t^d corresponds to the power injection of device d at time t into the network at its bus, and other elements describe internal dynamics, such as state-of-charge (SOC) of energy storage devices. At time t, state and input histories are expressed by $\mathbf{x}_t^d := [(x_1^d)^\intercal, \ldots, (x_t^d)^\intercal]^\intercal \in \mathbf{R}^{n_d t}$ and $\mathbf{u}_t^d := [(u_0^d)^\intercal, \ldots, (u_{t-1}^d)^\intercal]^\intercal \in \mathbf{R}^{m_d t}$ with

$$\mathbf{x}_t^d = A_t^d x_0^d + B_t^d \mathbf{u}_t^d,$$

where

$$A_t^d := \begin{bmatrix} \bar{A}^d \\ (\bar{A}^d)^2 \\ \vdots \\ (\bar{A}^d)^t \end{bmatrix}, B_t^d := \begin{bmatrix} \bar{B}^d & 0 & \ldots & 0 \\ \bar{A}^d \bar{B}^d & \bar{B}^d & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ (\bar{A}^d)^{t-1} \bar{B}^d & \ldots & \bar{A}^d \bar{B}^d & \bar{B}^d \end{bmatrix}.$$

2.3.3 Network Constraints

The AC power flow equations render prototypical AC OPF formulation nonconvex and NP-hard; what is more, in the present context, they hinder the development of computationally-affordable chance-constrained AC OPF formulations where CVaR arguments are leveraged as risk measures.

For distribution systems, we refer the reader to the linear approximation methods proposed in e.g., [20, 42, 120, 61, 29, 82], with the latter suitable for multi-phase systems with both wye and delta connections; these approximate models have been shown to provide high levels of accuracy in many existing test systems. For transmission systems, one can consider the tradition DC power-flow model to approximate the voltage angles and active power flows in the system; see, e.g., [62, 270, 248, 277]. Alternatively, one can consider alternative linearizations; see e.g.,[82]. As long as an accurate linear model exists, the proposed technical approach can be utilized to formulate and solve a distributionally-robust chance-constrained AC OPF problem. In particular, linear models can be utilized to formulate convex (in fact, linear) constraints on *line flows* and *voltage magnitudes*, e.g.

$$V^{\min} \leq |V_i| \leq V^{\max}, \ \forall i \in \mathcal{N}. \tag{2.6}$$

Grid-connected devices have various local constraints including, e.g., state of charge limitations for energy storage devices, allowable power injection ranges, generator ramping limits, and other device limits. These can be modeled (or approximated) as linear inequalities of the form

$$\mathbf{T}_d^t \mathbf{x}_t^d + \mathbf{U}_d^t \mathbf{u}_t^d + \mathbf{Z}_d^t \boldsymbol{\xi}_t \leq w_d, \ \ d = 1, \ldots, N_d, \tag{2.7}$$

where $\mathbf{T}_d^t \in \mathbf{R}^{l_d \times n_d t}$, $\mathbf{U}_d^t \in \mathbf{R}^{l_d \times m_d t}$, and $\mathbf{Z}_d^t \in \mathbf{R}^{l_d \times N_\xi t}$, and $w_d \in \mathbf{R}^{l_d}$ is a local constraint parameter vector.

2.4 Data-based Distributionally Robust Stochastic Optimal Power Flow for Distribution Networks

2.4.1 Distribution Network Model

In this section, we specialize the model to symmetric and balanced power distribution networks, connected to the grid at a point of common coupling (PCC). Loads and distributed generators (e.g., thermal generators, inverter-based RESs, and energy storage devices) may be connected to each bus. We augment the bus set with node 0 as the PCC.

The voltage and injected current at each bus are defined as $V_n^t = |V_n^t|e^{j\angle V_n^t}$, and $I_n^t = |I_n^t|e^{j\angle I_n^t}$. The absolute values $|V_n^t|$ and $|I_n^t|$ correspond to the signal root-mean-square values, and the phase $\angle V_n^t$ and phase $\angle I_n^t$ correspond to the phase of the signal with respect to a global reference.

Node 0 is modeled as a slack bus and the others are PQ buses, in which the injected complex power are specified. The admittance matrix can be partitioned as

$$\begin{bmatrix} I_0^t \\ \mathbf{i}^t \end{bmatrix} = \begin{bmatrix} y_{00} & \bar{y}^\intercal \\ \bar{y} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} V_0 \\ \mathbf{v}^t \end{bmatrix}.$$

The net complex power injection is then

$$\mathbf{s}^t = \mathrm{diag}(\mathbf{v}^t)\Big(\mathbf{Y}^*(\mathbf{v}^t)^* + \bar{y}^*(v_0^t)^*\Big). \tag{2.8}$$

The nonconvexity of this equation in the space of power injections and bus voltages is a source of significant computational difficulty in optimal power flow problems. In the rest of this section, we formulate a convex and computationally efficient data-based stochastic OPF problem based on a particular linear approximation of (4.3) that is appropriate for distributions networks. This approximation occurs on a specific point of a power flow manifold that preserves network structure for both real and reactive power flows and allows direct application of stochastic optimization techniques for incorporating forecast errors.

2.4.2 Leveraging Approximate Power Flow

Collect the voltage magnitudes $\{|V_n^t|\}_{n\in\mathcal{N}}$ into the vector $|\mathbf{v}^t| := [|V_1^t|, \ldots, |V_N^t|]^\intercal \in \mathbf{R}^N$. To develop computationally-feasible approaches, the technical approach in this chapter leverages linear approximations of the AC power-flow equations; in particular, linear approximations for voltages, as a function of the injected power $\mathbf{s}^t$, are given by

$$\mathbf{v}^t \approx \Big(\mathbf{H}\mathbf{p}^t + \mathbf{J}\mathbf{q}^t + \mathbf{c}\Big), \quad |\mathbf{v}^t| \approx \Big(\mathbf{M}\mathbf{p}^t + \mathbf{N}\mathbf{q}^t + \mathbf{a}\Big). \tag{2.9}$$

Using these approximations, the voltage constraints $V^{\min} \leq |V_n^t| \leq V^{\max}$ can be approximated as $V^{\min}\mathbf{1}_N \preceq \mathbf{M}\mathbf{p}^t + \mathbf{N}\mathbf{q}^t + \mathbf{a} \preceq V^{\max}\mathbf{1}_N$. The coefficient matrices of the linearized voltages, and the normalized vectors $\mathbf{a}$ and $\mathbf{c}$ can be obtained as shown in [82, 120, 29, 61]. For completeness, in the remainder of this subsection we briefly outline the approach taken in [82, 120] to derive a linear model for the voltages.

Suppose that $\mathbf{v}^t = \bar{\mathbf{v}} + \Delta v^t$, where $\bar{\mathbf{v}} = |\bar{\mathbf{v}}|\angle\boldsymbol{\theta} \in \mathbf{C}^N$ is a pre-determined nominal voltage vector and $\Delta v^t \in \mathbf{C}^N$ denotes a deviation from nominal. Then we have

$$\mathbf{s}^t = \mathrm{diag}(\bar{\mathbf{v}} + \Delta v^t)\Big(\mathbf{Y}^*(\bar{\mathbf{v}} + \Delta v^t)^* + \bar{y}^* V_0^*\Big). \tag{2.10}$$

Neglecting second-order terms $\mathrm{diag}(\Delta v^t)\mathbf{Y}^*(\Delta v^t)^*$, the power balance (4.3) becomes $\Lambda\Delta v^t + \Phi(\Delta v^t)^* = \mathbf{s}^t + \Psi$, where $\Lambda := \mathrm{diag}(\mathbf{Y}^*\bar{\mathbf{v}}^* + \bar{y}^* V_0^*)$, $\Phi := \mathrm{diag}(\bar{\mathbf{v}})\mathbf{Y}^*$, $\Psi := -\mathrm{diag}(\bar{\mathbf{v}})(\mathbf{Y}^*\bar{\mathbf{v}}^* + \bar{y}^* V_0^*)$. We consider a choice of the nominal voltage $\bar{\mathbf{v}} = \mathbf{Y}^{-1}\bar{y}V_0$, which gives $\Lambda = \mathbf{0}_{N\times N}$ and $\Psi = \mathbf{0}_N$. Therefore the linearized power flow expression is $\mathbf{s}^t = \mathrm{diag}(\bar{\mathbf{v}})\mathbf{Y}^*(\Delta v^t)^*$, the deviation Δv^t becomes $\Delta v^t = \mathbf{Y}^{-1}\mathrm{diag}^{-1}(\bar{\mathbf{v}}^*)(s^t)^*$.

Let us denote $\mathbf{Y}^{-1} = (\mathbf{G} + \mathrm{j}\mathbf{B})^{-1} = \mathbf{Z}_R + \mathrm{j}\mathbf{Z}_I$. Then expanding Δv^t in rectangular form gives

$$\bar{\mathbf{M}} = \left(\mathbf{Z}_R\ \mathrm{diag}\left(\frac{\cos(\theta)}{|\bar{\mathbf{v}}|}\right) - \mathbf{Z}_I\ \mathrm{diag}\left(\frac{\sin(\theta)}{|\bar{\mathbf{v}}|}\right)\right),$$
$$\bar{\mathbf{N}} = \left(\mathbf{Z}_I\ \mathrm{diag}\left(\frac{\cos(\theta)}{|\bar{\mathbf{v}}|}\right) + \mathbf{Z}_R\ \mathrm{diag}\left(\frac{\sin(\theta)}{|\bar{\mathbf{v}}|}\right)\right),$$

which define the rectangular matrices $\mathbf{H} := \bar{\mathbf{M}} + j\bar{\mathbf{N}}$, $\mathbf{J} := \bar{\mathbf{N}} - j\bar{\mathbf{M}}$, and the coefficient $\mathbf{c}$ is $\bar{\mathbf{v}}$. If $\bar{\mathbf{v}}$ dominates Δv^t, then the voltage magnitudes are approximated by $|\bar{\mathbf{v}}| + \Re\{\Delta v^t\}$, and linearized coefficients of voltage magnitudes become $\mathbf{M} := \bar{\mathbf{M}}$, $\mathbf{N} := \bar{\mathbf{N}}$, and $\mathbf{a} := |\bar{\mathbf{v}}|$. It is worth noting that the approach proposed in [29] accounts for multiphase systems with both wye and delta connections. Accordingly, the proposed framework is applicable to generic multiphase feeders with both wye and delta connections.

2.4.3 Data-based Stochastic OPF Formulation for Distribution Networks

Using the introduced linearized relationship between voltage and power injection vectors $\mathbf{p}^t$ and $\mathbf{q}^t$, we express the voltage magnitude in the following form

$$g^t\big[\mathbf{p}^t(u_t,\xi_t),\mathbf{q}^t(u_t,\xi_t)\big] := \mathbf{M}(\mathbf{I} - \mathrm{diag}\{\boldsymbol{\alpha}^t\})\mathbf{p}^t_{\mathrm{av}} + \mathbf{N}\mathbf{q}^t + |\bar{\mathbf{v}}|,$$

where $\boldsymbol{\alpha}^t \in \mathbf{R}^N$ is a control policy defined as the fraction of the active power curtailment by the renewable energy power injection. A system state vector $\mathbf{p}^t_{\mathrm{av}} \in \mathbf{R}^N$ collects the active power injection including loads and the available RES power. We aim to optimize the set points $\{\boldsymbol{\alpha}^t, \mathbf{q}^t\}$ of nodal power injections, which can be achieved by adjusting controllable loads and generators. More details of system modeling and component dynamics will be introduced in Part II.

Broadly speaking, we quantify a violation risk of voltage magnitude constraints (6) and local device constraints (7) for each node and each time as follows

$$\mathbb{E}\ \mathcal{R}\Big\{ g^t_n\big[\mathbf{p}^t(u_t,\xi_t),\mathbf{q}^t(u_t,\xi_t)\big] - V^{\max} \Big\} \le 0, \tag{2.11a}$$

$$\mathbb{E}\ \mathcal{R}\Big\{ V^{\min} - g^t_n\big[\mathbf{p}^t(u_t,\xi_t),\mathbf{q}^t(u_t,\xi_t)\big] \Big\} \le 0, \tag{2.11b}$$

$$\mathbb{E}\ \mathcal{R}\Big\{ \mathbf{T}^t_d\mathbf{x}^d_t + \mathbf{U}^t_d\mathbf{u}^d_t + \mathbf{Z}^t_d\boldsymbol{\xi}_t - w_d \Big\} \le 0, d = 1,\dots,N_d, \tag{2.11c}$$

where $g^t_n(\,\cdot\,)$ is the n-th element of the function value $g^t(\,\cdot\,)$, and $\mathcal{R}$ denotes a generic transformation of the inequality constraints into stochastic versions. Using a prior on the uncertainty

and possibly introducing auxiliary variables, the general risk functions (2.11a)-(2.11c) can be reformulated using e.g., scenario-based approaches, [261, 160] or moment-based distributionally robust optimization [248, 157]. This chapter seeks a reformulation by leveraging the CVaR [227]. A set of constraints will be approximated using the proposed distributionally robust approach, while other constraints will be evaluated using sample average methods.

We define a set $\mathcal{V}_t$ that contains a subset of N_ℓ affine constraints (2.11a)-(2.11c) that will be treated with distributionally robust optimization techniques. This allows some or all of the constraints to be included. We express them in terms of a decision variable vector $\mathbf{y}_t$ and uncertain parameters ξ_t, where $\mathbf{y}_t$ consists of all the decision variables including the RES curtailment ratio vector $\boldsymbol{\alpha}^\tau$ and other controllable device set-points, and ξ_t contains the uncertain parameters across the network including the active and reactive power injection forecast errors. For simplicity, we consider the risk of each constraint individually; it is possible to consider risk of joint constraint violations, but this is more difficult and we leave it for future work. Each individual affine constraint in the set $\mathcal{V}_t$ can be written in a compact form as follows

$$\mathcal{C}_o^t(\mathbf{y}_t, \xi_t) = [\bar{\mathcal{A}}(\mathbf{y}_t)]_o \xi_t + [\bar{\mathcal{B}}(\mathbf{y}_t)]_o, \quad o = 1, ..., N_\ell,$$

where $\mathcal{C}_o^t(\cdot)$ is the o-th affine constraint in the set $\mathcal{V}_t$. We use $[\,\cdot\,]_o$ to denote the o-th element of a vector or o-th row of a matrix. The CVaR with risk level β of the each individual constraint in the set $\mathcal{V}_t$ is

$$\inf_{\kappa_o^t} \mathbb{E}_{\xi_t} \left\{ [\mathcal{C}_o^t(\mathbf{y}_t, \xi_t) + \kappa_o^t]_+ - \kappa_o^t \beta \right\} \leq 0, \tag{2.12}$$

where $\kappa_o^t \in \mathbf{R}$ is an auxiliary variable [227]. The expression inside the expectation in (3.25) can be expressed in the form

$$\bar{\mathcal{Q}}_o^t = \max_{k=1,2} \left[\langle \bar{\mathbf{a}}_{ok}(\mathbf{y}_t), \xi_t \rangle + \bar{\mathbf{b}}_{ok}(\kappa_o^t) \right]. \tag{2.13}$$

This expression is convex in $\mathbf{y}_t$ for each fixed ξ_t since it is the maximum of two affine functions. Our risk objective function is expressed by the distributionally robust optimization of CVaR

$$\hat{J}^t_{\text{Risk}} = \sum_{\tau=t}^{t+\mathcal{H}_t} \sum_{o=1}^{N_\ell} \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}^{N_s}_\tau} \mathbb{E}^{\mathbb{Q}_\tau} \max_{k=1,2} \left[\langle \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}_\tau \rangle + \bar{\mathbf{b}}_{ok}(\kappa^\tau_o) \right].$$

The above multi-period distributionally robust optimization can be equivalently reformulated the following quadratic program, the details of which are described in [96]. The objective is to minimize a weighted sum of an operational cost function and the total worst-case CVaR of the affine constraints in set $\mathcal{V}^t$ (e.g., voltage magnitude and local device constraints).

Data-based distributionally robust stochastic OPF

$$\begin{aligned}
&\inf_{\mathbf{y}_\tau, \kappa^\tau_o} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^\tau_{\text{Cost}}] + \rho \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}^{N_s}_\tau} \sum_{o=1}^{N_\ell} \mathbb{E}^{\mathbb{Q}_\tau}[\bar{\mathcal{Q}}^\tau_o] \right\}, \\
&= \inf_{\substack{\mathbf{y}_\tau, \kappa^\tau_o, \\ \lambda^\tau_o, s^\tau_{io}, \varsigma^\tau_{iko}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^t_{\text{Cost}}] + \sum_{o=1}^{N_\ell} \left(\lambda_o \varepsilon_\tau + \frac{1}{N_s} \sum_{i=1}^{N_s} s^\tau_{io} \right) \right\}, \qquad (2.14a)
\end{aligned}$$

subject to

$$\rho(\bar{\mathbf{b}}_{ok}(\kappa^\tau_o) + \langle \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}^i_\tau \rangle + \langle \varsigma_{iko}, \mathbf{d} - H\hat{\xi}^i_\tau \rangle) \leq s^\tau_{io}, \qquad (2.14b)$$

$$\| H^\intercal \varsigma_{iko} - \rho \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau) \|_\infty \leq \lambda^\tau_o, \qquad (2.14c)$$

$$\varsigma_{iko} \geq 0, \qquad (2.14d)$$

$$\forall i \leq N_s, \forall o \leq N_\ell, k = 1, 2, \tau = t, ..., t + \mathcal{H}_t,$$

where $\rho \in \mathbf{R}_+$ quantifies the power system operators' risk aversion. This is a quadratic program that explicitly uses the training dataset $\hat{\Xi}^{N_s}_\tau = \{\hat{\xi}^i_\tau\}_{i \leq N_s}$. The risk aversion parameter ρ and the Wasserstein radius ε_τ allow us to explicitly balance tradeoffs between efficiency, risk and sampling errors inherent in $\hat{\Xi}^{N_s}_\tau$. The uncertainty set is modeled as a polytope $\Xi := \{\xi \in \mathbf{R}^{N_\xi} : H\xi \leq \mathbf{d}\}$. The constraint $\varsigma_{iko} > 0$ holds since the uncertainty set is not-empty; on the other hand, in a case with no uncertainty (i.e, $\varsigma_{iko} = 0$), the variable λ does not play any role and $s^\tau_{io} = \rho(\bar{\mathbf{b}}_{ok}(\kappa^\tau_o) + \langle \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}^i_\tau \rangle)$.

2.5 Data-based Distributionally Robust Stochastic Optimal Power Flow for Transmission Systems

2.5.1 Illustrative Explanation for the DC Approximation

In transmission networks, we employ a widely used "DC" linearization of the nonlinear AC power flow equations that makes the following assumptions [270, 248]:

- The lines $(i, j) \in \mathcal{E}$ are lossless and characterized by their reactance $\Im\{1/y_{ik}\}$;
- The voltage magnitudes $\{|V_n|\}_{\{n\in\mathcal{N}\}}$ are constant and close to one per unit;
- Reactive power is ignored.

Under a DC model, line flows are linearly related to the nodal power injections; therefore, flow constraints in the transmission lines can be expressed as

$$\sum_{d=1}^{N_d} \Gamma_d^t (r_d^t + G_d^t \boldsymbol{\xi}_t + C_d^t \mathbf{x}_t^d) \leq \bar{p}_t, \tag{2.15}$$

where $\Gamma_d^t \in \mathbf{R}^{2Lt\times t}$ maps the power injection (or consumptions) at each node to line flows and can be constructed from the network line impedances [62, 308]; on the other hand, $\bar{p}_t \in \mathbf{R}^{2Lt}$ denotes a limit on the line flows. The power injections in (2.15) features controllable and non-controllable components; the non-controllable power for a device/node d is modeled as $r_d^t + G_d^t \boldsymbol{\xi}_t$ (with positive values denoting the net power injection in the network), where the vector $r_d^t \in \mathbf{R}^t$ denotes the nominal power injection and $G_d^t \boldsymbol{\xi}_t$ models uncertain injections. Finally, $\mathbf{x}_t^d$ is the vector of controllable powers, and the matrix $C_d^t \in \mathbf{R}^{t\times n_d t}$ selects the first element of $\mathbf{x}_t^d$ at each time, i.e., $C_d^t := I_t \otimes \tilde{C}_d^t$, where $\tilde{C}_d^t = [1, \mathbf{0}_{1\times(n_d-1)}]$, I_t is a t-dimensional identity matrix and $\otimes$ denotes the Kronecker product operator. The power balance constraint is

$$\sum_{d=1}^{N_d} (r_d^t + G_d^t \boldsymbol{\xi}_t + C_d^t \mathbf{x}_t^d) = 0. \tag{2.16}$$

2.5.2 Reserve Policies

Deterministic OPF formulations ignore the uncertainties $\boldsymbol{\xi}_t$ and compute an open-loop input sequence for each device. In a stochastic setting, one must optimize over causal *policies*, which specifies how each device should respond to the current system states and forecast errors as they are discovered. We can now formulate a finite horizon stochastic optimization problem

$$\inf_{\pi \in \Pi_{\text{affine}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \mathbb{E}\Big[J^{\tau}_{\text{Cost}}(x_\tau, \pi_\tau, \xi_\tau) \Big], \tag{2.17a}$$

subject to

$$\sum_{d=1}^{N_d} \Big[r^{\bar{t}}_d + G^{\bar{t}}_d \boldsymbol{\xi}_{\bar{t}} + C^{\bar{t}}_d \mathbf{x}^d_{\bar{t}} \Big]_{[\underline{t},\bar{t}]} = 0, \tag{2.17b}$$

$$\mathbb{E}\ \mathcal{R}\Big\{ \sum_{d=1}^{N_d} \Big[\Gamma^{\bar{t}}_d (r^{\bar{t}}_d + G^{\bar{t}}_d \boldsymbol{\xi}_{\bar{t}} + C^{\bar{t}}_d \mathbf{x}^d_{\bar{t}}) - \bar{p}_{\bar{t}} \Big]_{[2L\underline{t},2L\bar{t}]} \Big\} \leq 0, \tag{2.17c}$$

$$\mathbb{E}\ \mathcal{R}\Big[\mathbf{T}^{\tau}_d \mathbf{x}^d_\tau + \mathbf{U}^{\tau}_d \mathbf{u}^d_\tau + \mathbf{Z}^{\tau}_d \boldsymbol{\xi}_\tau - w_d \Big] \leq 0, \tag{2.17d}$$

$$d = 1, \ldots, N_d, \tau = t, \ldots, t + \mathcal{H}_t,$$

where $\mathcal{R}$ denotes a general constraint risk function. The definition and discussion of the general stochastic transformation of (2.17c)-(2.17d) is same as (2.11a)-(2.11c), and can be found in Section IV.C. Here, $[\underline{t}, \bar{t}]$ denotes the finite time interval $[t, t + \mathcal{H}_t]$ for brevity. The cost function is proportional to the first and second moments of the uncertainties ξ_t, since we assume the operational cost function is convex quadratic. Optimizing over general policies makes the problem (2.17) infinite dimensional, so we restrict attention to affine policies

$$\mathbf{u}^d_t = D^d_t \boldsymbol{\xi}_t + e^d_t,$$

where each participant device d (e.g., traditional generators, flexible loads or energy storage devices) power schedule $\mathbf{u}^d_t \in \mathbf{R}^t$ is parameterized by a nominal schedule $e^d_t \in \mathbf{R}^t$ plus a linear

function $D_t^d \in \mathbf{R}^{t \times N_\xi t}$ of prediction error realizations. To obtain causal policies, D_t^d must be lower-triangular. The D_t^d matrices can be interpreted as planned reserve mechanisms involving secondary frequency feedback controllers, which adjust conventional generator power outputs using frequency deviations to follow power mismatches [270]. Under affine policies, the power balance constraints are linear functions of the distribution of $\boldsymbol{\xi}_t$, which are equivalent to

$$\sum_{d=1}^{N_d}(r_d^t + C_d^t(A_t^d x_0^d + B_t^d e_t^d)) = 0, \sum_{d=1}^{N_d}(G_d^t + C_d^t B_t^d D_t^d) = 0.$$

2.5.3 Data-based Stochastic OPF Formulation for Transmission Systems

We now use the above developments to formulate a data-based distributionally robust OPF problem to balance an operating efficiency metric with CVaR values of line flow and local device constraint violations. We collect the line flow and local device constraints into a set $\mathcal{V}_t$, which will be evaluated with distributionally robust optimization techniques in a total amount of N_ℓ. This allows some or all of the constraints to be included. For simplicity, we again consider the risk of each constraint individually. We express the constraints in terms of decision variables $\{D_t^d, e_t^d\}$ and uncertain parameters ξ_t. Each individual affine constraint in the set $\mathcal{V}_t$ can be written in a compact form as follows

$$\mathcal{C}_o^t(D_t, e_t, \xi_t) = [\underline{\mathcal{A}}(D_t)]_o \xi_t + [\underline{\mathcal{B}}(e_t)]_o, \ o = 1, ..., N_\ell,$$

where $\mathcal{C}_o^t(\cdot)$ is the o-th affine constraint in the set $\mathcal{V}_t$. The decision variables D_t and e_t both appear linearly in $\mathcal{C}_o^t$. The CVaR with risk level β of the each individual constraint in the set $\mathcal{V}_t$ is

$$\inf_{\sigma_o^t} \mathbb{E}_{\xi_t}\left\{[\mathcal{C}_o^t(D_t, e_t, \xi_t) + \sigma_o^t]_+ - \sigma_o^t \beta\right\} \leq 0, \tag{2.18}$$

where $\sigma_o^t \in \mathbf{R}$ is an auxiliary variable [227]. The expression inside the expectation (2.18) is expressed in the form

$$\underline{\mathcal{Q}}_o^t = \max_{k=1,2} \Big[\langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_t), \xi_t \rangle + \underline{\mathbf{b}}_{ok}(\sigma_o^t) \Big], \tag{2.19}$$

where the decision variable vector $\mathbf{y}$ consists of optimization variables $\{D, e, \sigma\}$. The convex expression (2.19) is maximum of two affine functions, and we consider the following distributionally robust total CVaR objective

$$\hat{J}_{\text{Risk}}^t = \sum_{\tau=t}^{t+\mathcal{H}_t} \sum_{o=1}^{N_\ell} \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}_\tau^{N_s}} \mathbb{E}^{\mathbb{Q}_\tau} \max_{k=1,2} \Big[\langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}_\tau \rangle + \underline{\mathbf{b}}_{ok}(\sigma_o^\tau) \Big].$$

The above multi-period distributionally robust optimization can be equivalently reformulated as a quadratic program. The details of the linear reformulation are shown in [96]. The objective is to minimize a weighted sum of an operational cost function and the total worst-case CVaR of the affine constraints in the set $\mathcal{V}^t$ (e.g., line flow constraints and local device constraints).

Data-based Distributionally Robust Stochastic DC OPF

$$\begin{aligned}
&\inf_{\mathbf{y}_\tau, \sigma_o^\tau} \sum_{\tau=t}^{t+\mathcal{H}_t} \Big\{ \mathbb{E}[\hat{J}_{\text{Cost}}^\tau] + \rho \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}_\tau^{N_s}} \sum_{o=1}^{N_\ell} \mathbb{E}^{\mathbb{Q}_\tau}[\underline{\mathcal{Q}}_o^\tau] \Big\}, \\
&= \inf_{\substack{\mathbf{y}_\tau, \sigma_o^\tau, \\ \lambda_o^\tau, s_{io}^\tau, \varsigma_{iko}^\tau}} \sum_{\tau=t}^{t+\mathcal{H}_t} \Big\{ \mathbb{E}[\hat{J}_{\text{Cost}}^\tau] + \sum_{o=1}^{N_\ell} \Big(\lambda_o^\tau \varepsilon_\tau + \frac{1}{N_s} \sum_{i=1}^{N_s} s_{io}^\tau \Big) \Big\}, \qquad (2.20a)
\end{aligned}$$

subject to

$$\sum_{d=1}^{N_d} \Big[(r_d^{\bar{t}} + C_d^{\bar{t}} (A_{\bar{t}}^d x_0^d + B_{\bar{t}}^d e_{\bar{t}}^d)) \Big]_{[\underline{t}, \bar{t}]} = 0, \tag{2.20b}$$

$$\sum_{d=1}^{N_d} \Big[(G_d^{\bar{t}} + C_d^{\bar{t}} B_{\bar{t}}^d D_{\bar{t}}^d) \Big]_{[\underline{t}, \bar{t}]} = 0, \tag{2.20c}$$

$$\rho(\underline{\mathbf{b}}_{ok}(\sigma_o^\tau) + \langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}_\tau^i \rangle + \langle \varsigma_{iko}^\tau, \mathbf{d} - H\hat{\xi}_\tau^i \rangle) \leq s_{io}^\tau, \tag{2.20d}$$

$$\| H^\intercal \varsigma_{iko}^\tau - \rho \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau) \|_\infty \leq \lambda_o^\tau, \tag{2.20e}$$

$$\varsigma_{iko}^\tau \geq 0, \tag{2.20f}$$

$$\forall i \leq N_s, \forall o \leq N_\ell, k = 1, 2, \tau = t, \ldots, \mathcal{H}_t,$$

where $\rho \in \mathbf{R}_+$ quantifies the power system operators' risk aversion. Once again, this is a quadratic program that explicitly uses the training dataset $\hat{\Xi}_\tau^{N_s} := \{\hat{\xi}_\tau^i\}_{i \leq N_s}$, and the risk aversion parameter ρ and the Wasserstein radius ε_τ allow us to explicitly balance tradeoffs between efficiency, risk, and sampling errors inherent in $\hat{\Xi}_\tau^{N_s}$. The uncertainty set is modeled as a polytope $\Xi := \{\xi \in \mathbf{R}^{N_\xi} : H\xi \leq \mathbf{d}\}$. The constraint $\varsigma_{iko} > 0$ holds since the uncertainty set is not-empty; on the other hand, in a case with no uncertainty (i.e, $\varsigma_{iko} = 0$), the variable λ does not play any role and $s_{io}^\tau = \rho(\underline{\mathbf{b}}_{ok}(\sigma_o^\tau) + \langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}_\tau^i \rangle)$.

2.6 Overvoltage Mitigation in Distribution Networks

In this section, we apply the data-based distributionally robust stochastic OPF methodology to mitigate overvoltages in distribution networks by controlling set points in RESs and energy storage devices. We provide further modeling details of the loads, inverter-based RESs, and energy storage devices. The set points of controllable devices are repeatedly optimized over a finite planning horizon within a MPC feedback scheme. The risk conservativeness of the voltage magnitude constraints and the out-of-sample performance robustness to sampling errors are explicitly adjustable by two scalar parameters.

2.6.1 System Model

1) Loads. We use $P_{l,n}^t$ and $Q_{l,n}^t$ to denote the active and reactive power demands at bus $n \in \mathcal{N}$. We also define two vectors $\mathbf{p}_l^t := [P_{l,1}^t, \ldots, P_{l,N}^t]^\intercal$ and $\mathbf{q}_l^t := [Q_{l,1}^t, \ldots, Q_{l,N}^t]^\intercal$. If no load is connected to bus $n \in \mathcal{N}$, then $P_{l,n}^t = 0$ and $Q_{l,n}^t = 0$. Load uncertainties are modeled based on historical data of forecast errors. The active and reactive loads are given by $\mathbf{p}_l^t = \bar{\mathbf{p}}_l^t(u_t) + \tilde{\mathbf{p}}_l^t(\xi_t)$, $\mathbf{q}_l^t = \bar{\mathbf{q}}_l^t(u_t) + \tilde{\mathbf{q}}_l^t(\xi_t)$, where $\bar{\mathbf{p}}_l^t(u_t) \in \mathbf{R}^N$ and $\bar{\mathbf{q}}_l^t(u_t) \in \mathbf{R}^N$ are forecasted nominal loads, which can depend on control decisions (e.g., load curtailment control). The nodal injection errors $\tilde{\mathbf{p}}_l^t(\xi_t) \in \mathbf{R}^N$ and $\tilde{\mathbf{q}}_l^t(\xi_t) \in \mathbf{R}^N$ depend on the aggregate forecast error vector ξ_t.

2) Renewable energy model. Let $P^t_{\text{av},n}$ be the maximum availability renewable energy generation at bus $n \in \mathcal{N}_R \subseteq \mathcal{N}$, where the set $\mathcal{N}_R$ denotes all buses with RESs. With high RES penetration, overvoltages can cause power quality and reliability issues. By intelligently operating set points of RES and energy storage, operators can optimally trade off risk of constraint violation and economic efficiency (e.g., purchase of electricity from the main grid, active power curtailment costs, and reactive compensation costs). The active power injections of RESs are controlled by adjusting an active power curtailment factor $\alpha^t_n \in [0, 1]$. Reactive power set points of RESs can also be adjusted within a limit $\bar{S}_n$ on apparent power as follows

$$\sqrt{((1-\alpha^t_n)P^t_{\text{av},n})^2 + (Q^t_n)^2} \leq \bar{S}_n, n \in \mathcal{N}_R.$$

We define aggregate vectors: $\boldsymbol{\alpha}^t := [\alpha^t_1, \ldots, \alpha^t_N]^\intercal$ and $\mathbf{p}^t_{\text{av}} := [P^t_{\text{av},1}, \ldots, P^t_{\text{av},N}]^\intercal$ and $\mathbf{q}^t_c := [Q^t_1, \ldots, Q^t_N]^\intercal$.

If bus $n \in \mathcal{N} \backslash \mathcal{N}_R$ has no RES, by convention we set $\alpha^t_n = 0, P^t_{\text{av},n} = 0$ and $Q^t_n = 0$. The curtailment factor and reactive power compensation $\{\alpha^t_n, Q^t_n\}$ together set the inverter operating point and are also subject to a power factor constraint

$$|Q^t_n| \leq \tan(\theta_n)[(1-\alpha^t_n)P^t_{\text{av},n}], n \in \mathcal{N}_R,$$

where $\cos(\theta_n) \in (0, 1]$ is the power factor limit for RESs. The power factor constraint is convex, and can be discarded in settings where the inverters are not required to operate at a minimum power factor level. The premise here is that RESs can assist in the regulation of voltages by promptly adjusting the reactive power and curtailing active power as needed; RESs can provide faster voltage regulation capabilities compared to traditional power factor correction devices (i.e., capacitor banks). The proposed data-based distributionally robust OPF will consider adjustments of both active and reactive powers to aid voltage regulation, which in principle can be done in both transmission and distribution networks.

3) Energy storage model. The state-of-charge (SOC) of the energy storage device located at bus $n \in \mathcal{N}_B \subseteq \mathcal{N}$ in kWh is represented as B_n^t. The dynamics of these devices are

$$B_n^{t+1} = B_n^t + \eta_{B,n} P_{B,n}^t \Delta, n \in \mathcal{N}_B, \tag{2.21}$$

where Δ is the duration of the time interval $(t, t+1]$, and $P_{B,n}^t$ is the charging/discharging power of the storage device in kW. We assume the battery state is either charging ($P_{B,n}^t \geq 0$) or discharging ($P_{B,n}^t \leq 0$) during each time interval $(t, t+1]$. For simplicity, we suppose the round-trip efficiency of the storage device $\eta_{B,n} = 1$ to avoid the nonconvexity when introducing binary variables. Additionally, two common operational constraints of energy storage devices are

$$B_n^{\min} \leq B_n^t \leq B_n^{\max}, \quad P_{B,n}^{\min} \leq P_{B,n}^t \leq P_{B,n}^{\max},$$

where $B_n^{\min}$, $B_n^{\max}$ are the rated lower and upper SOC levels, and $P_{B,n}^{\min}$, $P_{B.n}^{\max}$ are the minimum and maximum charging/discharging limits. Other constraints can be added for electric vehicles (EVs), for example, a prescribed SOC $B_n^t = B_n^{\max}$ at a particular time. If no energy device is connected to a certain bus, the charging/discharging power and SOC are fixed to zero: $P_{B,n}^t = 0$, $B_n^t = 0$, for all $n \in \mathcal{N} \backslash \mathcal{N}_B$. We define the aggregate vectors $\mathbf{p}_B^t := [P_{B,1}^t, \ldots, P_{B,N}^t]^\intercal$, and $\mathbf{b}^t := [B_1^t, \ldots, B_N^t]^\intercal$.

2.6.2 Data-based Stochastic OPF Implementation

We now use the methodologies presented in Part I Section IV, and the models of loads, RESs and energy storage devices to develop a data-based stochastic AC OPF for solving a voltage regulation problem. This stochastic OPF aims to balance the operational costs the total CVaR values of the voltage magnitude constraints. We consider an operational cost that captures electricity purchased by customers, excessive solar energy fed back to the utility,

reactive power compensation costs and penalties for active power curtailment

$$
\begin{aligned}
&J^t_{\text{Cost}}(\boldsymbol{\alpha}^t, \mathbf{q}^t_c, \mathbf{p}^t_B, \xi_t) = \\
&= \sum_{n\in\mathcal{N}} a^t_{1,n} \left[P^t_{l,n} + P^t_{B,n} - (1-\alpha^t_n)P^t_{\text{av},n}\right]_+ \\
&+ \sum_{n\in\mathcal{N}} a^t_{2,n} \left[(1-\alpha^t_n)P^t_{\text{av},n} - P^t_{l,n} - P^t_{B,n}\right]_+ \\
&+ \sum_{n\in\mathcal{N}} a^t_{3,n}|Q^t_n| + \sum_{n\in\mathcal{N}} a^t_{4,n}\alpha^t_n P^t_{\text{av},n}.
\end{aligned}
$$

We collect all decision variables into $\mathbf{y}_t = \{\alpha^t, \mathbf{q}^t_c, \mathbf{p}^t_B, \mathbf{b}^t\}$, and all RES and load forecast errors into the random vector ξ_t. Now the MPC subproblems take the following form

Data-based Distributionally Robust Stochastic OPF

$$
\begin{aligned}
&\inf_{\substack{\mathbf{y}_\tau, \kappa^\tau_o, \\ \varpi_{1,n}, \varpi_{2,n}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^\tau_{\text{Cost}}] + \rho \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}^{N_s}_\tau} \sum_{o=1}^{N_\ell} \mathbb{E}^{\mathbb{Q}_\tau}[\bar{\mathcal{Q}}^\tau_o] \right\}, \\
&= \inf_{\substack{\mathbf{y}_\tau, \kappa^\tau_o, \\ \varpi_{1,n}, \varpi_{2,n} \\ \lambda^\tau_o, s^\tau_{io}, \varsigma^\tau_{iko}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^\tau_{\text{Cost}}] + \sum_{o=1}^{N_\ell} \left(\lambda_o \varepsilon_\tau + \frac{1}{N_s} \sum_{i=1}^{N_s} s^\tau_{io} \right) \right\}, \qquad (2.22a)
\end{aligned}
$$

subject to

$$
\rho(\bar{\mathbf{b}}_{ok}(\kappa^\tau_o) + \langle \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}^i_\tau \rangle + \langle \varsigma_{iko}, \mathbf{d} - H\hat{\xi}^i_\tau \rangle) \leq s^\tau_{io}, \qquad (2.22b)
$$

$$
\|H^\intercal \varsigma_{iko} - \rho \bar{\mathbf{a}}_{ok}(\mathbf{y}_\tau)\|_\infty \leq \lambda^\tau_o, \qquad (2.22c)
$$

$$
\varsigma_{iko} \geq 0, \qquad (2.22d)
$$

$$
\frac{1}{N_s} \sum_{i=1}^{N_s} \left[[(1-\alpha^\tau_n)\hat{P}^{\tau,i}_{\text{av},n}]^2 + (Q^\tau_n)^2 - \bar{S}^2_n + \varpi^\tau_{1,n} \right]_+ \leq \varpi^\tau_{1,n}\beta, \qquad (2.22e)
$$

$$
\frac{1}{N_s} \sum_{i=1}^{N_s} \left[\tan(\theta_n)[(1-\alpha^t_n)\hat{P}^{\tau,i}_{\text{av},n}] - |Q^\tau_n| + \varpi^\tau_{2,n} \right]_+ \leq \varpi^\tau_{2,n}\beta, \qquad (2.22f)
$$

$$
B^{\min}_n \leq B^\tau_n \leq B^{\max}_n, \qquad (2.22g)
$$

$$
P^{\min}_{B,n} \leq P^\tau_{B,n} \leq P^{\max}_{B,n}, \qquad (2.22h)
$$

$$
B^{\tau+1}_n = B^\tau_n + \eta_{B,n} P^\tau_{B,n} \Delta, \qquad (2.22i)
$$

$$0 \leq \alpha_n^\tau \leq 1, \tag{2.22j}$$

$$\forall i \leq N_s,\ \forall o \leq N_\ell,\ n \in \mathcal{N}_R,\ k = 1, 2,\ \tau = t, \ldots, t + \mathcal{H}_t,$$

where $\varpi_{1,n}^\tau$, $\varpi_{2,n}^\tau$, and κ_o^τ are CVaR auxiliary valuables, and λ_o^τ, s_{io}^τ, ς_{iko}^τ are auxiliary variables associated with the distributionally robust Wasserstein ball reformulation. For simplicity, the power factor constraints and apparent power limitation constraints are not treated as distributionally robust constraints, and instead are handled using direct sample average approximation.

Remark 2.1 (battery efficiency). To maintain convexity of the underlying problem formulation and therefore facilitate the development of computationally affordable solution methods, we utilized an approximate model for the battery dynamics with no charging and discharging efficiency losses (2.22). At the expense of significantly increasing the problem complexity, charging and discharging efficiencies can be accommodated as [217]

$$B_n^{\tau+1} = B_n^\tau + \eta_c P_{B_c,n}^\tau \Delta - \frac{1}{\eta_d} P_{B_d,n}^\tau \Delta,$$

where $\eta_c, \eta_d \in (0, 1]$ denote the charging and the discharging efficiencies, respectively; $P_{B_c,n}^\tau \geq 0$ represents the charging rate and $P_{B_d,n}^\tau \geq 0$ the discharging rate at time τ. Additional constraints, however, are needed to ensure that the solution avoids meaningless solutions where a battery is required to charge and discharge simultaneously; in particular, one can: a) add a constraint $P_{B_c,n}^\tau P_{B_d,n}^\tau = 0$ [217]; or. b) introduce binary variables to indicate the charging status (e.g., charging/discharging) of the batteries [136]. Either way, given the non-convexity of the resultant problem, possibly sub-optimal solutions can be achieved (2.22). In addition, exact relaxation methods under appropriate assumptions offer an alternative way to maintain convexity of the charging problem; see [169, 83]. Extending the proposed technical approach to a setting with binary variables or exact relaxation methods will be pursued as a future research effort.

Remark 2.2 (battery life). The degradation of energy storage systems may depend on the depth of discharge and the number of charging/discharging cycles [84]. The battery aging process is usually described by partial differential equations [218]; this is a practical model for industrial applications, but it introduces significant computational challenges in optimization tasks [287]. Additional optimization variables as well as penalty functions could be included to limit the number of cycles per day and ensure a minimum state of charge [17, 99, 200, 290, 147, 237, 295]. Pertinent reformulations to account for battery degradation will be pursued in future research activities.

Remark 2.3 (voltage at slack bus). Similar to the majority of the works in the literature, the voltage at the slack bus (i.e., substation) is considered as an input of the problem (and, therefore, it is not controllable). However, it is worth noting that discrete variables modeling changing the tap position of the transformer can be incorporated in (2.22); see e.g., [189, 190]. Branch and bound techniques can then be utilized to solve the problem.

2.6.3 Numerical results

We use a modified IEEE-37 node test feeder to demonstrate our proposed data-based stochastic AC OPF method. As shown in Fig. 2.1, the modified network is a single-phase equivalent and the load data is derived from real measurements from feeders in Anatolia, CA during the week of August 2012 [19]. We place 21 photovoltaic (PV) systems in the network. Their locations are marked by yellow boxes in Fig. 2.1, and their capacities are summarized in Table 2.1. Based on irradiation data from [19, 2], we utilized a greedy gradient boosting method [100] to make multi-step ahead predictions of solar injections, and then computed a set of forecast errors from the dataset. In general forecast errors increase with the prediction horizon. Other parameters of the network, such as line impedances and shunt admittances, are taken from [309]. The total nominal available solar power $\sum_n P^t_{\mathrm{av},n}$ and aggregate load demand over 24 hours is also shown in Fig. 2.2.

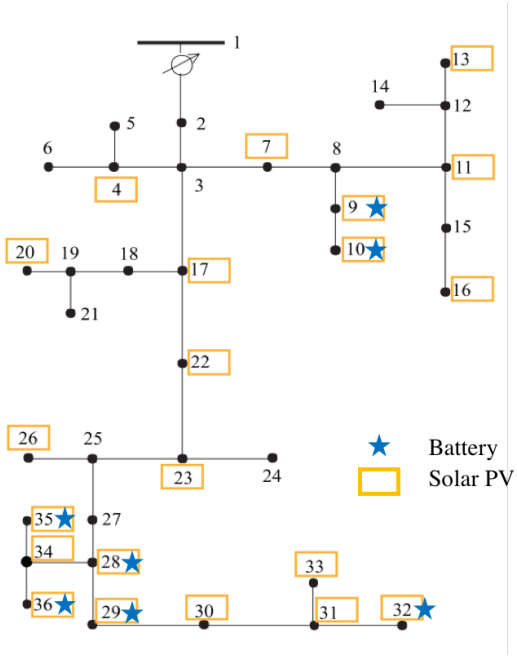


Figure 2.1. IEEE 37-node test feeder with renewable energy resources and storage devices.

Table 2.1. Capacities of inverted-based solar energy generations and energy storage devices

Node	S_n [kVA]	Node	S_n [kVA]	Node	$B_n^{\max}$
4	150	7	300	9	100
9	300	10	600	10	100
11	660	13	360	28	50
16	600	17	360	29	250
20	450	22	150	32	250
23	750	26	300	35	120
28	750	29	300	36	200
30	360	31	600		
32	330	33	750		
34	450	35	450		
36	450				

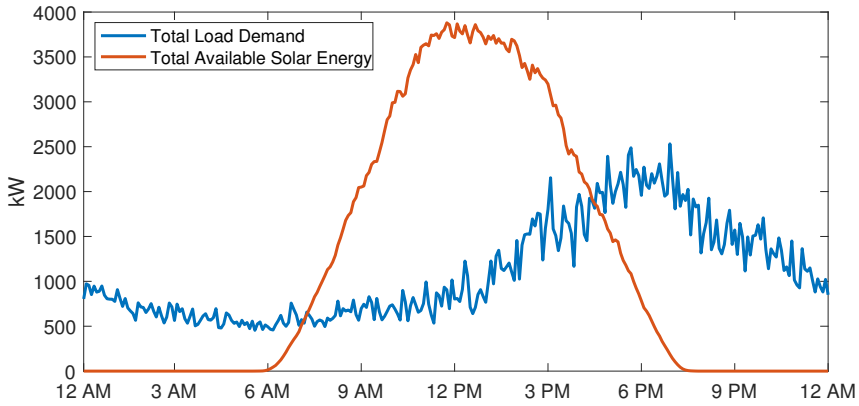


Figure 2.2. Total available solar energy and load demand.

The energy storage systems are placed with PV systems at certain nodes, as shown in Fig. 2.1. Their locations and capacities are listed in Table 2.1. We select the capacities in the range of typical commercial storage systems, or aggregate of 10-12 residential-usage batteries (e.g., electric vehicles), which are connected to the same step-down transformer. The lower limit of SOC, $B_n^{\min}$, is set to be zero for all batteries. The charging/discharging rate $P_{B,n}^t$ is also limited by 10% of their respective energy capacity $B_n^{\max}$. Voltage limits $V^{\max}$ and $V^{\min}$ are 1.05 p.u. and 0.95 p.u., respectively. The cost function parameters are $a_{1,n}^t = 10$, $a_{2,n}^t = 3$, and $a_{3,n}^t = 3$ and $a_{4,n}^t = 6$. The decision making time period is 5 minutes.

Due to high PV penetration, overvoltage conditions can emerge during solar peak irradiation. Given the real data available for the numerical tests, the numerical tests are focused on alleviating over-voltage conditions via the proposed distributionally-robust tools. Other constraints are approximated via sample average methods [171, 248, 72]; however, in general it is straightforward to formulate other constraints as distributionally robust. The power factor PF limit is 0.9 in (2.22f). The risk level parameter η is set to 0.01 for quantifying 1% violation probability of constraints (2.22e)-(2.22f). To emphasize the effect of sampling errors, the number of forecast error samples N_s included in the training dataset $\hat{\Xi}_t^{N_s}$ is limited to 30. The forecast errors are not assumed bounded, so the parameters of the support

polytope $\Xi_t := \{\xi_t \in \mathbf{R}^{N_\xi} : H\xi_t \leq \mathbf{d}\}$ are set to zero in (2.22b)-(2.22c). We solved (2.22) using the MOSEK solver [11] via the MATLAB interface CVX [135] on a laptop with 16 GB of memory and 2.8 GHz Intel Core i7. Solving each time step during solar peak hours with distibutionally robust constraints takes 4.84 seconds. Note that our implementation is not optimized for speed and in principle could easily be sped up and scaled to larger problems since the problem is ultimately convex quadratic.

In our framework, there are two key parameters, ρ and ε, that explicitly adjust trade offs between performance and constraint violation risk, and robustness to sampling errors. Fig. 2.3 illustrates the basic tradeoffs between operational cost and CVaR values of voltage constraint violations during a 24-hour operation for various values of ρ and ε. It can be readily seen that as ρ increases, operational cost increases, but CVaR decreases since the risk term is emphasized. Notice that with the increasing of ε, the *estimated* risk is higher so that the solution is more conservative and leads to a lower risk of constraint violation; larger Wasserstein balls lead to higher robustness to sampling errors. These parameters offer system operators explicit data-based tuning knobs to systematically set the conservativeness of operating conditions.

Fig. 2.4(a)-2.4(c) shows the aggregated solar energy curtailment and substation power purchases for varying risk aversion ρ and Wasserstein radius ε. In order to prevent voltages over 1.05 p.u., the available solar energy must be increasingly curtailed as the risk aversion parameter ρ increases. As a result, the network must import more power from the substation. The increasing curtailment of solar energy and purchase of power drawn from the substation lead to significantly higher operational cost.

However, these decisions will also lead to more stable voltage profiles, as shown in Fig. 2.5. When ρ is small, there is almost no curtailment, causing overvoltages at several buses. As ρ increases, more active power is curtailed, and all voltages move below their upper limit. Similar comments apply for varying the Wasserstein radius ε. For example, fixing ρ and

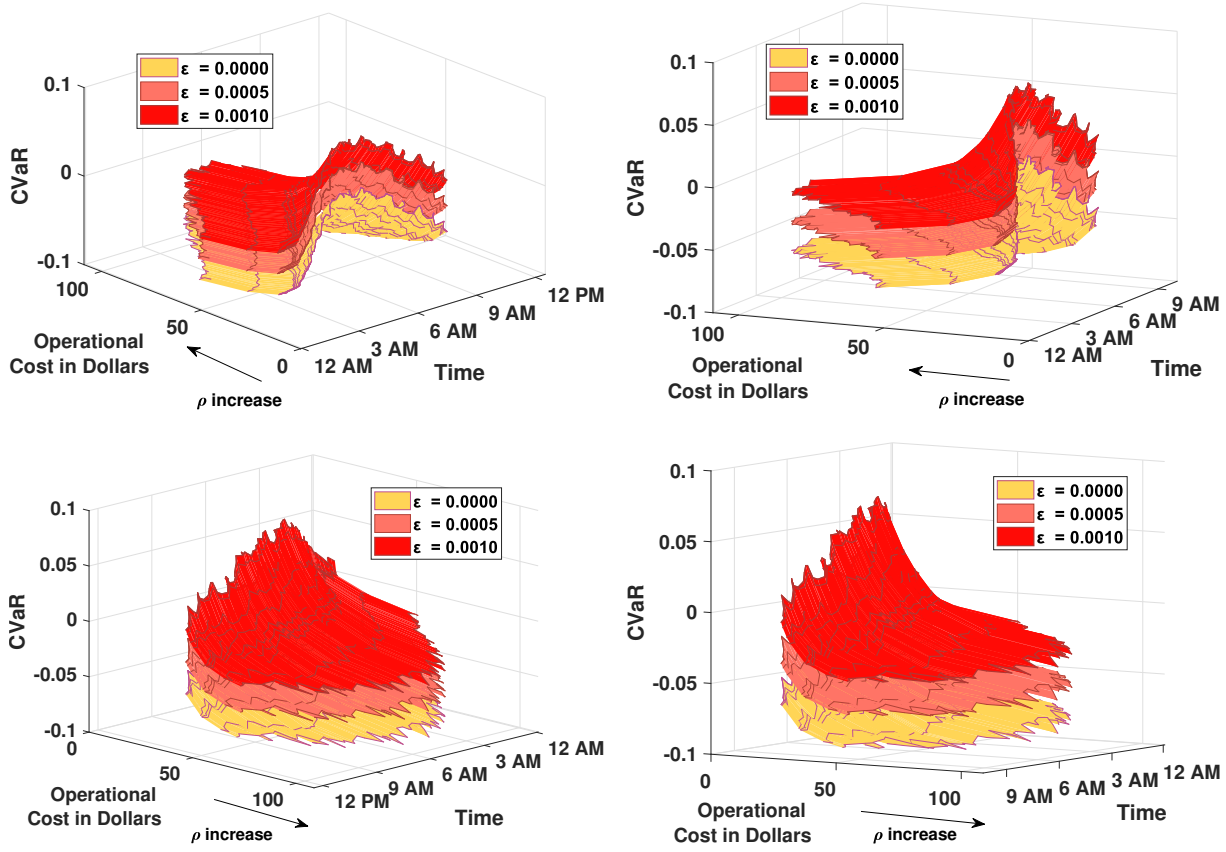
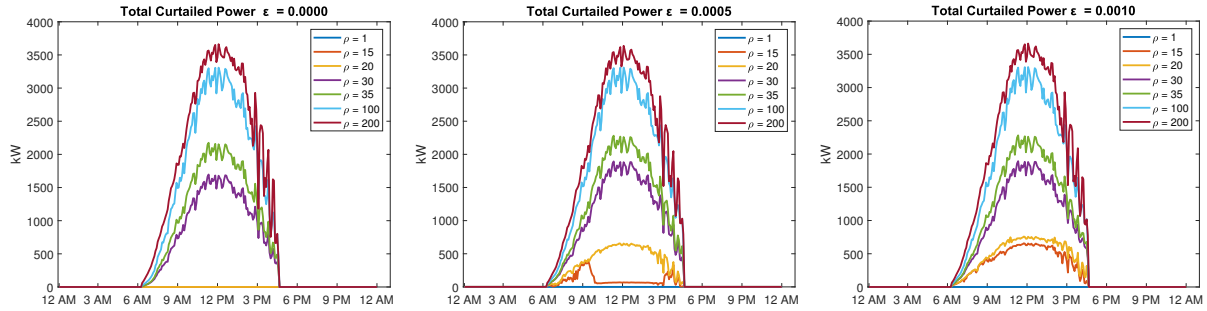


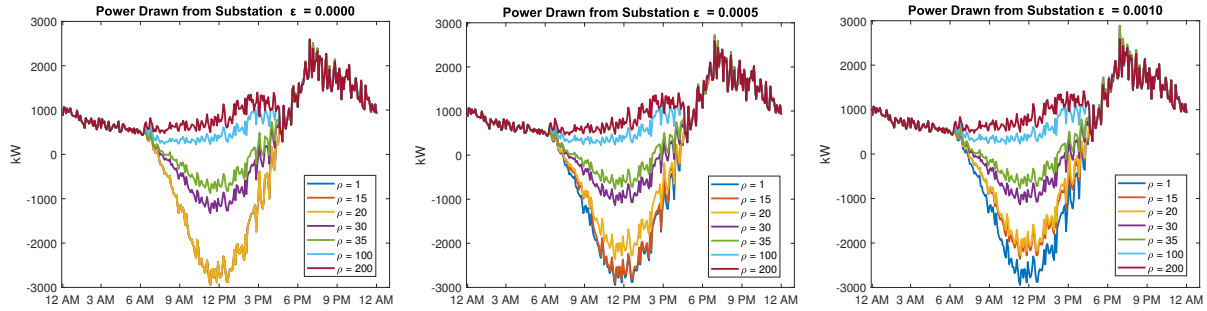
Figure 2.3. Tradeoffs between operational cost, conditional value at risk (CVaR) of voltage constraints, and robustness to sampling errors for $\varepsilon = 0.0000, 0.0005, 0.0010$); parameters ρ and ε are varied to tests different weighting settings and radii of the Wasserstein ball, respectively. We present four views from different directions to avoid occlusion.

increasing ε also results in more curtailment and lower voltage magnitude profiles, which leads to better robustness to solar energy forecast errors.

Finally, we evaluate out-of-sample performance by implementing the full closed-loop distributionally robust MPC scheme over the 24 hour period with a 15 minute planning horizon. Monte Carlo simulations with 100 realizations of forecast errors over the entire horizon are shown in Fig. 2.6. We subsampled new solar energy forecast errors from the training dataset. The closed-loop voltage profiles based on MPC decisions for all scenarios at node 28 are shown (other nodes with overvoltages show qualitatively similar results). Again, it is clearly seen that larger values of ε and ρ yield more conservative voltage profiles.



(a) Total curtailed active power $\varepsilon = 0.0000$ (b) Total curtailed active power $\varepsilon = 0.0005$ (c) Total curtailed active power $\varepsilon = 0.0010$



(d) Power drawn from substation $\varepsilon = 0.0000$ (e) Power drawn from substation $\varepsilon = 0.0005$ (f) Power drawn from substation $\varepsilon = 0.0010$

Figure 2.4. Comparison of active power curtailment and power purchased from substation for various values of risk aversion ρ and Wasserstein radius ε. As these parameters increase, more active power from PV is curtailed and more power is drawn from the substation, leading to a lower risk of constraint violation and a higher operating cost.

In summary, we conclude that the proposed data-based distributionally robust stochastic OPF is able to systematically assess and control tradeoffs between the operational costs, risks, and sampling robustness in distribution networks. The benefits of the open-loop stochastic optimization problems are also observed in the closed-loop multi-period distributionally robust model predictive control scheme.

2.7 *N*-1 Security Problem in Transmission Systems

In this section, we apply the proposed methodology from Part I in a transmission system to handle N-1 line flow security constraints. The basic DC power flow approximation and

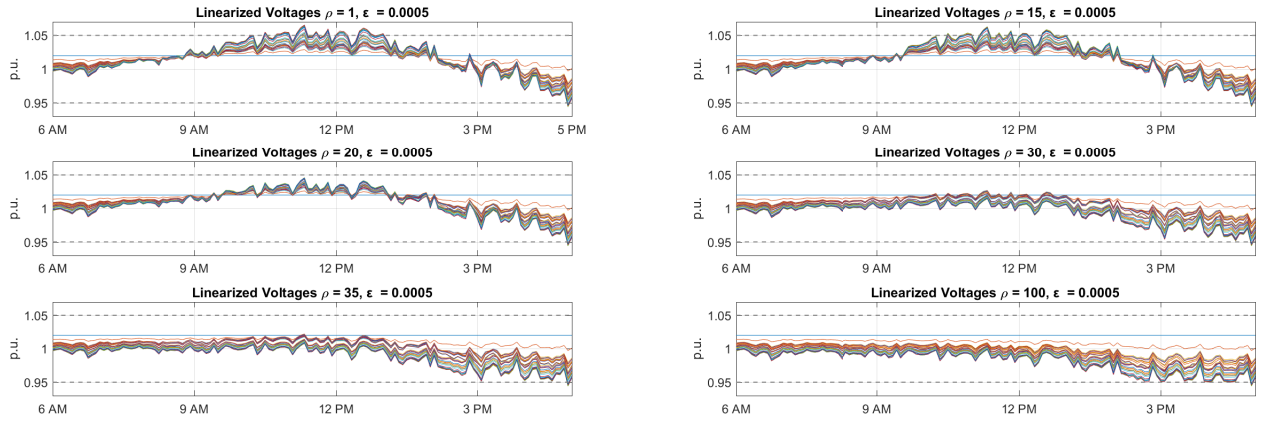


Figure 2.5. Optimal network voltage profiles for varying ρ and ε. Overvoltages are reduced as ρ increases.

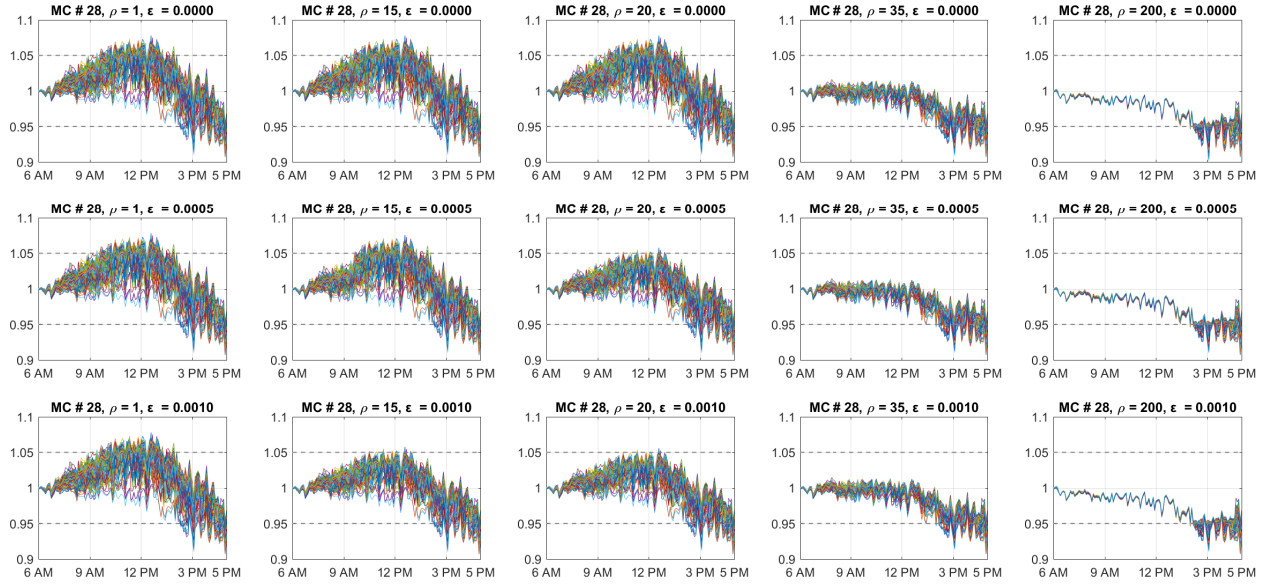


Figure 2.6. Monte Carlo simulation results of the voltage profiles at node #28 resulting from the full distributionally robust closed-loop model predictive control scheme. We validate that in closed-loop larger values of ε and ρ yield more conservative voltage profiles.

device modeling is discussed in Part I. Here, we also incorporate N-1 security constraints and associated contingency reactions due to uncertain wind power injections.

2.7.1 System Model

We consider a transmission system with N_G generators (e.g., conventional thermal and wind) connected to bus subset $\mathcal{N}_G \subseteq \mathcal{N}$. There are N_L loads, N_l lines, and N_b buses. The outages included for N-1 security consist of tripping of any single lines, generators or loads, yielding $N_{\text{out}} = N_G + N_L + N_l$ possible outages. We collect the outages corresponding to a generator, a line and a load in different sets $\mathcal{I}_G$, $\mathcal{I}_l$, and $\mathcal{I}_L$. The outages in total are $\mathcal{I} = \{0\} \cup \mathcal{I}_G \cup \mathcal{I}_L \cup \mathcal{I}_l$, where $\{0\}$ indicates no outage.

The formulation of a N-1 security problem is based on the following assumptions: 1) the power flow equations are approximated with DC power flow, as described in Part I (Section V); 2) each wind farm is connected to a single bus of the network; 3) load forecasting is perfect; 4) a single line outage can cause multiple generator/load failures.

The objective of the data-based distributionally robust stochastic DC OPF is to determine an optimal reserve schedule for responding to the wind energy forecast errors while taking the network security constraints into account. We define $P^j_{\text{mis}} \in \mathbf{R}$ for all $j \in \mathcal{I}$ as the generation-load mismatch given by

$$P^j_{\text{mis}} = \begin{cases} P^j_L - P^j_G, & \text{if } \ j \in \{0\} \cup \mathcal{I}_l \\ +P^j_L, & \text{if } \ j \in \mathcal{I}_L \\ -P^j_G, & \text{if } \ j \in \mathcal{I}_G \end{cases},$$

where $P^j_L, P^j_G \in \mathbf{R}$ denote the power disconnection corresponding to the outage $j \in \mathcal{I}$. Define $P_G \in \mathbf{R}^{N_G}$, and $P_L \in \mathbf{R}^{N_L}$ as nodal generation and load injection vectors. For the generator or load failures, the power disconnection P^j_G or P^j_L is corresponding to the components in the vector P_G or P_L. In the case of line failures $j \in \mathcal{I}_l$, the power disconnection P^j_G or P^j_L

is the sum of the power loss caused by multiple failures. If there is no power disconnection caused by a line outage, or if $j = 0$ (no outage happens), the power mismatch is set to zero: $P^j_{\text{mis}} = 0$.

To respond to contingencies, we can also define another reserve policy response matrix $\mathbf{R}^{j,d}_{\text{mis},t} := [R^{j,d}_1, \ldots, R^{j,d}_t]^\intercal \in \mathbf{R}^t$, so that the affine reserve policy becomes

$$\mathbf{u}^d_t = D^d_t \boldsymbol{\xi}_t + \mathbf{R}^{j,d}_{\text{mis},t} P^j_{\text{mis}} + e^d_t, \forall j \in \mathcal{I}, d = 1, \ldots, N_d. \tag{2.23}$$

The general constraint risk function of the line flow in Part I (Section V, Equation (17c)) is then given by

$$\begin{aligned} f\Bigg(\sum_{d=1}^{N_d} \widetilde{\Gamma}^{t,j}_d \Big\{ r^t_d + G^t_d \boldsymbol{\xi}_t + C^t_d \big[A^d_t x^d_0 + B^d_t (D^d_t \boldsymbol{\xi}_t \\ + \mathbf{R}^{j,d}_{\text{mis},t} P^j_{\text{mis}} + e^d_t) \big] \Big\} - \bar{p}_t \leq 0 \Bigg), \ \forall j \in \mathcal{I}, \end{aligned} \tag{2.24}$$

where $\widetilde{\Gamma}^{t,j}_d \in \mathbf{R}^{2Lt \times t}$ maps the power injection of each device in the case of j-th outage.

2.7.2 Data-based Stochastic OPF Implementation

We now use the modeling here and in Part I and the affine control strategy (2.23) to formulate a data-based distributionally robust stochastic DC OPF for transmission systems that also incorporate N-1 security constraints. The goal is to balance tradeoffs between cost of thermal generation, CVaR values of the line flow constraints, and sampling error robustness. The generation with reserve policy in the cost function is given by $P^{d,t}_G = \big[D^d_t \xi_t + e^d_t \big]_t$. The operational cost of generators is

$$J^t_{\text{Cost}} = \sum_{d \in \mathcal{N}_G} c_{1,d} [P^{d,t}_G]^2 + c_{2,d} [P^{d,t}_G] + c_{3,d},$$

which captures nominal and reserve costs of responding to wind energy forecast errors. The N-1 security reserve cost is not included to simplify presentation, but this can also easily be included in our framework as an additional linear cost [262].

With the proposed modeling in Part I (Section V), the updated data-based stochastic DC OPF is shown as follows. The decision variables are collected into $\mathbf{y}_t = \{\mathbf{D}_t, e_t, \mathbf{R}_{\text{mis},t}\}$. The random vector ξ_t comprises all wind energy forecast errors.

Data-based distributionally robust stochastic DC OPF

$$\begin{aligned}
&\inf_{\mathbf{y}_\tau, \sigma_o^\tau} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}_{\text{Cost}}^\tau] + \rho \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}_\tau^{N_s}} \sum_{o=1}^{N_\ell} \mathbb{E}^{\mathbb{Q}_\tau}[\underline{\mathcal{Q}}_o^\tau] \right\}, \\
&= \inf_{\substack{\mathbf{y}_\tau, \sigma_o^\tau, \\ \lambda_o^\tau, s_{io}^\tau, \varsigma_{iko}^\tau}} \sum_{\tau=t}^{\tau+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}_{\text{Cost}}^\tau] + \sum_{o=1}^{N_\ell} \left(\lambda_o^\tau \varepsilon_\tau + \frac{1}{N_s} \sum_{i=1}^{N_s} s_{io}^\tau \right) \right\}, \qquad (2.25a)
\end{aligned}$$

subject to

$$\rho(\underline{\mathbf{b}}_{ok}(\sigma_o^\tau) + \langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau), \hat{\xi}_\tau^i \rangle + \langle \varsigma_{iko}^\tau, \mathbf{d} - H\hat{\xi}_\tau^i \rangle) \leq s_{io}^\tau, \qquad (2.25b)$$

$$\|H^\intercal \varsigma_{iko}^\tau - \rho \underline{\mathbf{a}}_{ok}(\mathbf{y}_\tau)\|_\infty \leq \lambda_o^\tau, \qquad (2.25c)$$

$$\varsigma_{iko}^\tau \geq 0, \qquad (2.25d)$$

$$\frac{1}{N_s} \sum_{i=1}^{N_s} \sum_{d=1}^{N_d} \left[\widetilde{\Gamma}_d^{\bar{t},j} \Big\{ r_d^{\bar{t}} + G_d^{\bar{t}} \hat{\boldsymbol{\xi}}_{\bar{t}}^i + C_d^{\bar{t}} \big[A_{\bar{t}}^d x_0^d + B_{\bar{t}}^d (D_{\bar{t}}^d \hat{\boldsymbol{\xi}}_{\bar{t}}^i + \mathbf{R}_{\text{mis},\bar{t}}^{j,d} P_{\text{mis}}^j + e_{\bar{t}}^d) \big] \Big\} - \bar{p}_{\bar{t}} \right]_{[\underline{t},\bar{t}]} \leq 0, \qquad (2.25e)$$

$$\sum_{d=1}^{N_d} \left[(r_d^{\bar{t}} + C_d^{\bar{t}} (A_{\bar{t}}^d x_0^d + B_{\bar{t}}^d e_{\bar{t}}^d)) \right]_{[\underline{t},\bar{t}]} = 0, \qquad (2.25f)$$

$$\sum_{d=1}^{N_d} \left[(G_d^{\bar{t}} + C_d^{\bar{t}} B_{\bar{t}}^d D_{\bar{t}}^d) \right]_{[\underline{t},\bar{t}]} = 0, \qquad (2.25g)$$

$$\forall i \leq N_s,\ \forall j \in \mathcal{I},\ \forall o \leq N_\ell,\ k = 1, 2,\ \tau = t, \ldots, t + \mathcal{H}_t.$$

2.7.3 Numerical Results

We consider a modified IEEE 118-bus test system [309] to demonstrate our proposed data-based distributionally robust stochastic DC OPF shown in Fig. 2.7. For simplicity, we only show results of a single-period stochastic optimization problem. As with the distribution

network, it is straightforward to extend to multi-period closed-loop stochastic control using MPC.

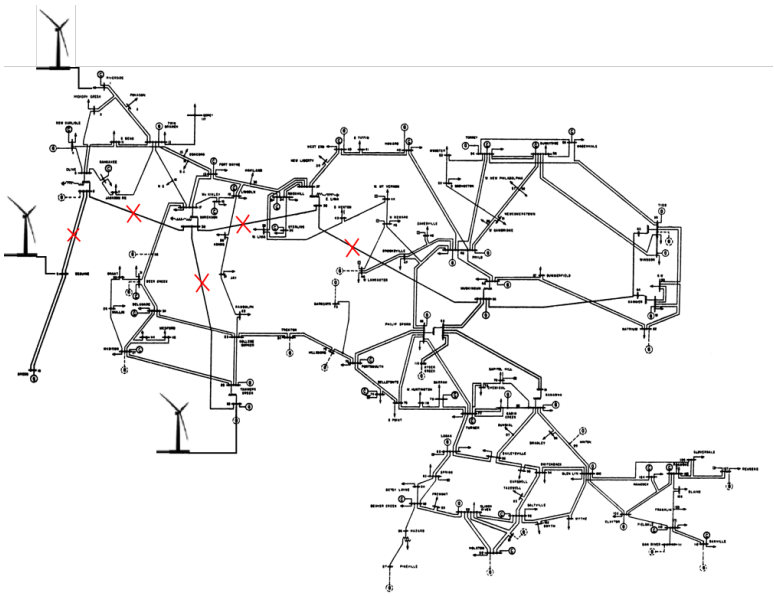
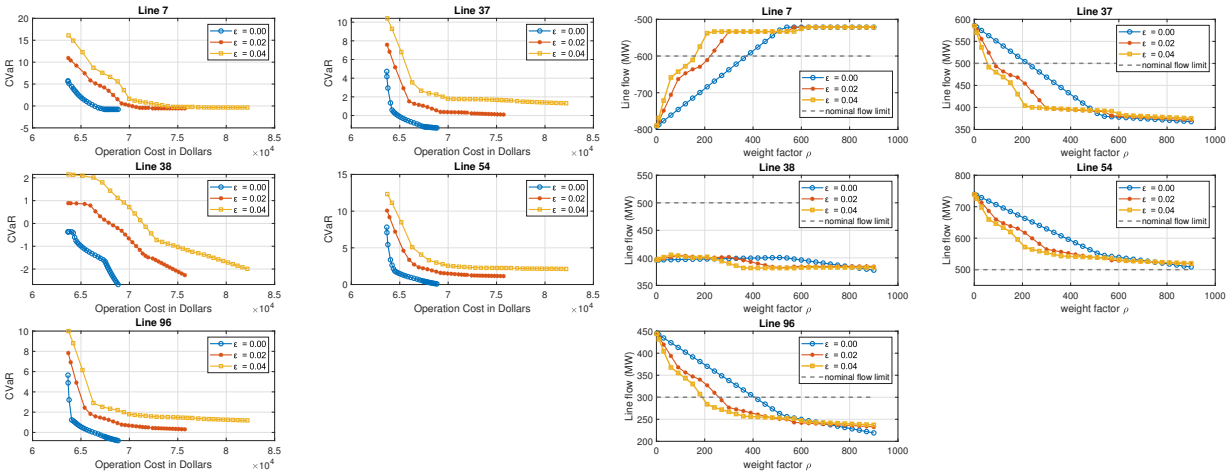


Figure 2.7. IEEE 118-bus test network with multiple wind energy connections.



(a) Predicted tradeoffs between operational cost and CVaR of line constraint violation.

(b) Out-of-sample performance is demonstrated by Monte Carlo simulation, with a controllable level of conservativeness.

Figure 2.8. Five line flows are reported to illustrate the performance of the proposed data-based distributionally robust stochastic DC OPF in a transmission system.

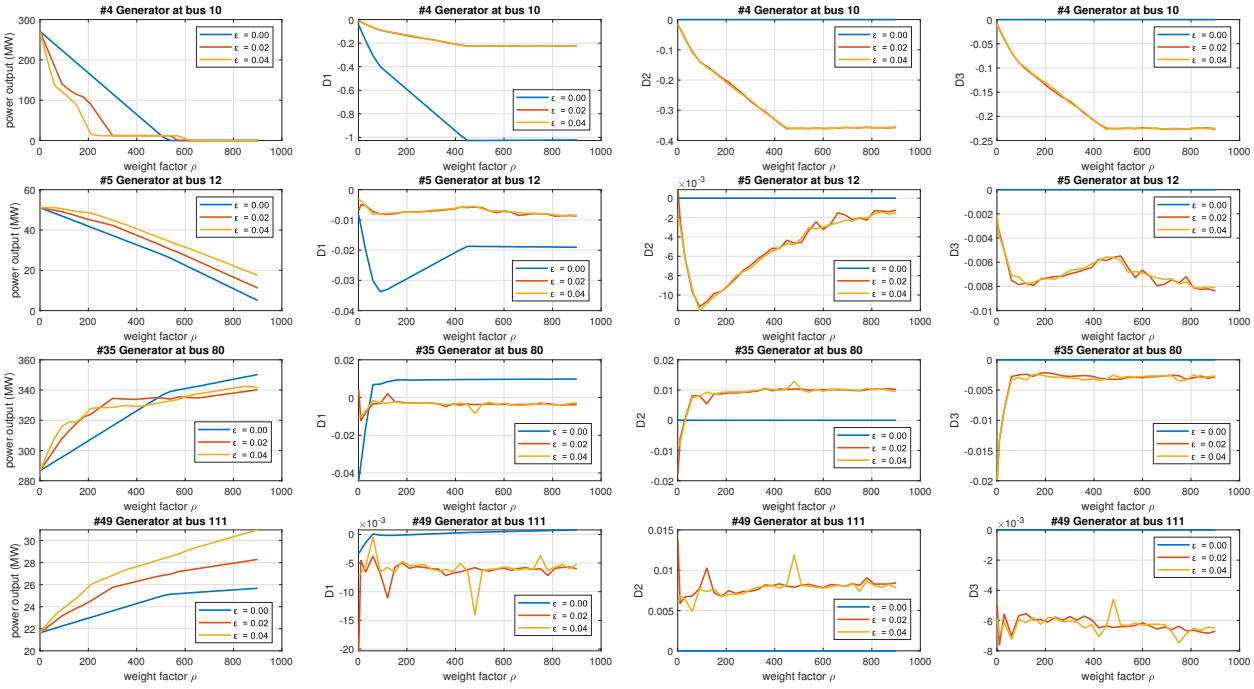


Figure 2.9. Comparison of the coefficients of the policies and output powers of selected generators for various values of risk aversion ρ and Wasserstein metric ε. As these parameters increase, the risk of line flow constraint decreases at the expense of higher operational costs.

Three wind farms are connected to bus #1, bus #9, and bus #26, with the normal feed-in power 500 MW, 500 MW, and 800 MW, respectively. The corresponding conventional generators at bus #1 and bus #26 are removed. The wind power forecast errors are derived from the real wind sampling data from hourly wind power measurements provided in 2012 Global Energy Forecasting competition (GEFCCom2012) [1]. The wind energy forecast errors are evaluated based on a simple persistence forecast, which predicts that the wind injection in the following period remains constant. It can be seen the forecast errors are highly leptokurtic, i.e, that the errors have significant outliers that make the distribution tails much heavier than Gaussian tails. We scale the forecasting errors to have zero mean and the standard deviations of 200 MW, 200 MW and 300 MW for the wind farms at bus #1, #9 and #26, respectively. We consider five key lines, which deliver the wind power from

the left side of the system to the right, as marked by the red crosses in Fig. 2.7. The line flow limits are shown in Table 5.1 and marked by gray dash lines in Fig.2.8(b).

Table 2.2. Five main channel lines data

# of line	From bus	To bus	Line flow limitation [MW]
7	8	9	600
37	8	30	500
38	26	30	500
54	30	38	500
96	38	65	300

To simplify our presentation, only these five lines flows are handled with distributionally robust optimization (2.25b)-(2.25d) in both directions; the remaining line flows are modeled by N-1 security constraints (2.25e) with nominal CVaR and sample average approximation (essentially equivalent to taking $\varepsilon = 0$), and no other local device constraint is included. Additionally, we assume no bounds on the wind power forecast errors ξ_τ, hence H and $\mathbf{d}$ in (2.25b)-(2.25c) are set to zero. It takes 10 seconds to solve (2.25) for each time step using the MOSEK solver [11] via the MATLAB interface with CVX [135] on a laptop with 16 GB of memory and a 2.8 GHz Intel Core i7.

Fig. 2.8(a) illustrates the solutions of the proposed data-based distributionally robust stochastic DC OPF for varying risk aversion and Wasserstein radius. Once again, the numerical results demonstrate fundamental tradeoffs between operational cost, CVaR values of line constraints violations, and robustness to sampling errors. The conservativeness of the generator policies are controlled by adjusting ρ and ε. By explicitly using the forecast error training dataset and accounting for sampling errors, risks are systematically assessed and controlled. Since the forecast errors are non-Gaussian, existing methods may significantly underestimate risk [121]. Increasing ε provides better robustness to sampling errors and guarantees out-of-sample performance. Note that CVaR values can be negative when the worst-case expected line flow is below the constraint boundary.

Fig. 2.8(b) illustrates the out-of-sample performance of the decisions via Monte Carlo simulation. For every value of ρ utilized to obtain the results of Fig. 8(a), we i) sampled (new) values from the training dataset, ii) implemented the decisions based on the solution of the problem, and iii) calculated the empirical line flows. The dashed gray line indicates the line flow limit. The number of scenarios for the Monte Carlo simulations is 1000. From the simulation results, it is seen that larger ε ensures smaller line constraint violation for all lines except line 38. This happens because the risk objective is the sum of all five CVaRs, and a lower overall risk is achieved for certain values of ρ and ε by allowing higher risk of violating the flow limit of line 38. In general, it is possible to prioritize certain important constraints by weighting their associated risk higher compared to lower priority constraints. Again, Monte Carlo simulations demonstrate that conservativeness can be controlled explicitly by changing the Wasserstein radius ε and the risk aversion parameter ρ.

Fig. 2.9 illustrates the output powers and the coefficients of the reserve polices for selected generators for different values of the risk aversion ρ and the Wasserstein radius ε. In order to satisfy the limit on the line flows, the scheduled power output of some generators (mostly located on the left side of the feeder or with cheaper cost profiles) are reduced as the risk aversion parameter ρ increase. As a result, some of the generators (mostly located on the right side of the feeders or with high cost profiles) increase the power injection to supply the demand. With these settings, the risk of line flow constraint decreases as shown in Fig. 2.8(b), at the expense of higher operational costs.

2.8 Conclusions

In this chapter, we have proposed a framework for the data-based distributionally robust stochastic OPF based on finite dataset descriptions of forecast error distributions across the power network. The method allows efficient computation of multi-stage feedback control policies that react to forecast errors and provide robustness to inherent sampling errors

in the finite datasets. Tractability is obtained by exploiting convex reformulations of ambiguity sets based on Wasserstein balls centered on empirical distributions. The general framework is adapted to both distribution networks and transmission systems by allowing general device models and utilizing different linearizations of the AC power flow equations. We performed numerical experiments to balance overvoltages in distribution networks and $N-1$ security line flow risks in transmission networks. The flexibility of controllable devices was exploited to balance efficiency and risk due to high penetration renewable energy sources. In contrast to existing work, our method directly incorporates forecast error training datasets rather than making strong assumptions on the forecast error distribution, which allows us to leverage distributionally robust optimization techniques to achieve superior out-of-sample performance. Parameters in the optimization problems allow system operators to systematically select operating strategies that optimally trade off performance and risk.

Supplementary Materials

Implementation codes for 1) data-based distributionally robust stochastic OPF and 2) data-based distributionally robust stochastic DC OPF can be download from `https://github.com/TSummersLab/Distributionally-robust-stochastic-OPF`.

Appendix

Pertinent reformulations of the probabilistic line flow constraints in transmission systems are described in this section. Consider, for simplicity, the time period between $[0, t]$ and rewrite the line flow constraint as

$$\mathbb{E}\ \mathcal{R}\Big\{ \sum_{d=1}^{N_d} \Gamma_d^t (r_d^t + G_d^t \boldsymbol{\xi}_t + C_d^t \mathbf{x}_t^d) - \bar{p}_t \Big\} \leq 0,$$

where $\mathbf{x}_t^d = A_t^d x_0^d + B_t^d \mathbf{u}_t^d$ and the affine reserve policies are $\mathbf{u}_t^d = D_t^d \boldsymbol{\xi}_t + e_t^d$. The CVaR counterpart of each individual constraints is given by [cf. (2.18)]

$$\begin{aligned}\inf_{\sigma_o^t} \mathbb{E}_{\xi_t} \max\Bigg\{ \Big[\sum_{d=1}^{N_d} \Gamma_d^t (r_d^t + G_d^t \boldsymbol{\xi}_t + C_d^t (A_t^d x_0^d + B_t^d D_t^d \boldsymbol{\xi}_t \\ + B_t^d e_t^d)) - \bar{p}_t \Big]_o + \sigma_o^t \Bigg\} - \sigma_o^t \beta \leq 0.\end{aligned} \tag{2.26}$$

Since the decision variables $\{D_t, e_t, \sigma_o^t\}$ enter as linear terms in (2.26), a compact expression similar to (2.19) can be obtained; i.e.,

$$\underline{\mathcal{Q}}_o^t = \max_{k=1,2} \Big[\langle \underline{\mathbf{a}}_{ok}(\mathbf{y}_t), \xi_t \rangle + \underline{\mathbf{b}}_{ok}(\sigma_o^t) \Big],$$

where

$k = 1$,

$$\underline{\mathbf{a}}_{o1}(\mathbf{y}_t) = \left[\big[\textstyle\sum_{d=1}^{N_d} \Gamma_d^t C_d^t B_t^d D_t^d \big]_o, \big[\textstyle\sum_{d=1}^{N_d} \Gamma_d^t C_d^t B_d^t e_t^d \big]_o \right],$$

$$\underline{\mathbf{b}}_{o1}(\sigma_o^t) = \big[-\bar{p} + \sum_{d=1}^{N_d} \Gamma_d^t (r_d^t + G_d^t \xi + C_d^t A_t^d x_0^d) \big]_o + \sigma_o^t - \sigma_o^t \beta,$$

$k = 2$,

$$\underline{\mathbf{a}}_{o2}(\mathbf{y}_t) = \Big[\mathbf{0}_{N_\xi t}^\intercal, 0 \Big], \underline{\mathbf{b}}_{o2}(\sigma_o^t) = -\sigma_o^t \beta,$$

for all o. Similar steps can be followed to obtain the constraints for each device and the voltage magnitude.

CHAPTER 3

OPTIMAL PUMP CONTROL FOR WATER DISTRIBUTION NETWORKS VIA DATA-BASED DISTRIBUTIONALLY ROBUSTNESS[1]

In this chapter, we propose a data-based methodology to solve a multi-period stochastic optimal water flow (OWF) problem for water distribution networks (WDNs). The framework explicitly considers the pump schedule and water network head level with limited information of demand forecast errors for an extended period simulation. The objective is to determine the optimal feedback decisions of network-connected components, such as nominal pump schedules and tank head levels and reserve policies, which specify device reactions to forecast errors for accommodation of fluctuating water demand. Instead of assuming the uncertainties across the water network are generated by a prescribed certain distribution, we consider *ambiguity sets* of distributions centered at an empirical distribution, which is based directly on a finite training data set. We use a distance-based ambiguity set with the Wasserstein metric to quantify the distance between the real unknown data-generating distribution and the empirical distribution. This allows our multi-period OWF framework to trade off system performance and inherent sampling errors in the training dataset. Case studies on a three-tank water distribution network systematically illustrate the tradeoff between pump operational cost, risks of constraint violation, and out-of-sample performance.

---

[1]This Chapter is based on work supported by the National Science Foundation (NSF) under grant CMMI-1728605.

Chapter 3 in part is a reprint of material published or under review in:

3.1 Introduction

Due to a broad range of future energy and environmental issues [155], water distribution network operators are seeking improved strategies to deliver energy-efficient, reliable, and high quality service to consumers [198]. However, the increasing complexity (e.g., due to high dimensionality, nonlinearities, operation constraints and uncertainties) in municipal water supply network operation is challenging the current management and control strategies and may threaten the security of this vital infrastructure. Future urban water supply systems will require more sophisticated methods to function robustly and efficiently in the presence of this increasing complexity.

The flexibility of water flow manipulators (pumps and valves) in water networks has been utilized to optimize various objectives, including production and transportation costs, water quality, safe storage, smoothness of control actions, etc. [98, 75, 142, 292, 258, 65, 193, 294, 266, 265]. However, most optimal water flow control methods use deterministic point forecasts of exogenous water demands, which neglects their inherent stochasticity. In practice, the variation of water demands in real water distribution networks is high and difficult to predict [113]. Further, as complexity of network topology increases [12], small perturbations can cause significant performance decrease and even infeasibility of optimal water flow problems [113].

Recent research on optimal water network operation has been shifting from deterministic to stochastic models, since uncertainties (e.g., human usage, unexpected component failures, climate change) are increasingly key factors in many sectors of water resource management [12, 44, 113, 271, 231, 118, 198, 117, 268, 55, 232, 267]. Most stochastic formulations assume that the uncertain water demands follow a prescribed distribution (e.g., Gaussian [268, 117]), or enforce constraint for all possible water demand realization by assuming only knowledge of bounds on uncertainties [113, 55], and then utilize robust optimization. In addition, sampling-based stochastic optimization has also been applied to water flow manipulation

problem [231] to quantify the probability of constraint violation based on an assumed data generating mechanism. However, the underlying assumptions in these approaches can be too strong or overly conservative, which can lead to underestimation or overestimation of the actual risks, and therefore to ineffective operation. The methods based on chance-constraints effectively only measure the frequency of constraint violations not the severity, which can underestimate risk. The robust methods can enforce constraints for extreme and highly unlikely uncertainty realizations, effectively overestimating risk. Furthermore, some sampling-based methods are computational intensive due to their requirement of a large numbers of samples. In practice, forecasts of water demand are never perfect, and their distributions must be estimated from finite data.

In this chapter, we investigate a multi-period data-driven optimization problem to tackle a stochastic optimal water flow (OWF) problem for optimal pump schedule and head management in water distribution networks (WDNs). The proposed framework uses the limited information of water demand forecasting errors from a finite training dataset to explicitly balance the tradeoffs between performance and the risk of constraint violations in the presence of large water demand variations. A preliminary version of this work present in [124]; here we significantly extend the work in several directions. The main contributions are:

1) We formulate a general multi-period distributionally robust optimal water flow problem for optimal pump schedule and head management. The distributionally robust OWF model predictive controller uses data-driven distributionally robust optimization [96] to tractably obtain control decisions for network components at each stage. This allows the data-driven distributionally robust MPC OWF controller to update the water demand forecast with a finite time horizon and then re-compute the real-time optimal decision based on the latest and future forecasting information. In general, this OWF controller accepts the training data set from all forecasting frameworks and the decisions can be robust to various *ambiguity sets (i.e., moment-based or metric-based).*

In this chapter, we assume the unknown real data generating distribution is located in a metric-based ambiguity set, which is constructed by a Wasserstein ball with constant radius centered at an empirical distribution supported by the finite training dataset. In contrast to other stochastic OWF formulations, this approach makes the resulting control policies explicitly robust to the inherent sampling errors in the training dataset, which leads to superior out-of-sample performance. We can appropriately parameterize the ambiguity set to avoid the overly conservative decisions due to overfitting and finite sampling.

2) To handle computationally difficulties with the nonlinear/nonconvex water network hydraulics, we leverage a pertinent linear approximation of water network hydraulic coupling (i.e., flow-head coupling) to promote a computationally-efficient stochastic optimal water flow formulation for optimal pump control and nodal pressure management. In contrast with the literature, we further establish a generic matrix linearization in compact format between water flow and nodal head by re-defining a network Laplacian matrix based on linearization coefficients. This provides a unified framework that is applicable for approximation algorithms after linearization (i.e., successive linearization algorithms or piece-wise linearization algorithms). We empirically observe that the convergence of successive linearization algorithm provides an excellent approximation to the nonlinear water flow.

3) The effectiveness and flexibility of our proposed stochastic water flow formulation are demonstrated on a model of the Barcelona water distribution network. We illustrate the inherent tradeoff between the system conservativeness and forecasting errors. The results can help the operators to explicitly prioritize the tradeoff between the pump operational efficiency and the risk of tank head constraint violation, and then develop the appropriate control strategies to balance their our objectives and risk aversion.

The rest of chapter is organized as follows: Section 3.2 describes a generic model of water distribution networks, and the successive linearization approach; Section 3.3 presents the general formulation of proposed data-based multi-period distributionally robust stochastic optimal water flow problems. Section 3.4 specifies the proposed stochastic OWF to a stochastic optimal pump schedule and head management. Section 3.5 demonstrates the flexibility and effectiveness of the proposed methodologies via numerical experiments. Section 3.6 concludes.

Notation: The inner product of two vectors $a, b \in \mathbf{R}^m$ is denoted by $\langle a, b\rangle := a^\intercal b$, and $(\cdot)^\intercal$ denotes the transpose of a vector or matrix. The N_s-fold product of distribution $\mathbb{P}$ on a set Ξ is denoted by $\mathbb{P}^{N_s}$, which represents a distribution on the Cartesian product space $\Xi^{N_s} = \Xi \times \ldots \times \Xi$. We use N_s to represent the number of samples inside the training dataset $\hat{\Xi}$. Superscript " $\hat{\cdot}$ " is reserved for the objects that depend on a training dataset $\hat{\Xi}^{N_s}$. The cardinality of set $\mathcal{J}$ is denoted by $|\mathcal{J}|$. The Kronecker product operator is defined as $\otimes$.

3.2 Hydraulic Model and Leveraging Linear Approximation

In this section, we consider a water distribution network model associated with active and passive networked components, and then we leverage a pertinent linear approximation, which leads to a novel network Laplacian-based matrix expression. This allows us to use successive linearization to approximate the original nonlinear hydraulic relationship for several topologies. WDNs control actions include speeds of pumps and settings of valves. In the rest of this section, we introduce the network and hydraulic modelling of networked components.

3.2.1 Network Modelling

We consider a water distribution network as a directed graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$ with a set $\mathcal{N} := \{1, 2, \ldots, N\}$ of vertices. These vertices include junctions, reservoirs, and tanks that are collected in sets $\mathcal{J}$, $\mathcal{S}$, and $\mathcal{T}$, and $\mathcal{N} = \mathcal{J} \cup \mathcal{S} \cup \mathcal{T}$. Similarly, the set $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ of all

links including the sets of pipes, pumps, and valves represented by $\mathcal{I}$, $\mathcal{M}$, and $\mathcal{V}$ so that $\mathcal{E} = \mathcal{I} \cup \mathcal{M} \cup \mathcal{V}$. Let $\mathcal{N}_i^{\text{in}}$ and $\mathcal{N}_i^{\text{out}}$ collect the supplying and carrying neighboring vertices of i^{th} node , respectively. We use $q_{ij} \in \mathbf{R}$ to denote the water flow through the link $(i, j) \in \mathcal{E}$, and $h_i \in \mathbf{R}_+$ denotes the head of node $i \in \mathcal{N}$. We assume each pipe has the prescribed flow direction and the actual flow direction on pipes are either following the assumption ($q_{ij} \geq 0$) or appearing in the opposite direction ($q_{ij} < 0$).

Junctions

The water demand is assumed to be a constant $d_i(t)$ in gallons per hour (GPM), which is applied for time interval t at junction $i \in \mathcal{J}$. Mass conservation must be hold any time at i^{th} node

$$\sum_{j \in \mathcal{N}_i^{\text{in}}} q_{ji}(t) - \sum_{j \in \mathcal{N}_i^{\text{out}}} q_{ij}(t) = d_i(t), \quad \forall i \in \mathcal{J}, \tag{3.1}$$

where $\mathcal{N}_i^{\text{in}}$ and $\mathcal{N}_i^{\text{out}}$ are the sets of nodes supplying and carrying flow at i^{th} junction, respectively. If there is no water demand consumption for nodes $i \in \mathcal{N} \backslash \mathcal{N}_d$, we have $d_i(t) = 0$ for all time slots. Here we define an aggregated vector $d(t) := [d_1(t), \ldots, d_N(t)]^\intercal \in \mathbf{R}^N$.

Reservoirs

The set $\mathcal{S}$ collects all reservoirs in a water distribution network. We assume that all reservoirs have infinite water resource supply, and that the head of each reservoir is a constant, which can be treated as an operational constraint

$$h_i^{\text{R}} = h_i^{\text{elv}}, \quad \forall i \in \mathcal{S},$$

where h_i^{elv} represents the elevation for i^{th} reservoir.

Tanks

The head of tank at node $i \in \mathcal{T}$ at time t is represented by $h_i^{\mathrm{TK}}(t)$. The dynamics of these elements are given by the discrete-time difference equations

$$h_i^{\mathrm{TK}}(t+1) = h_i^{\mathrm{TK}}(t) + \frac{\Delta t}{A_i^{\mathrm{TK}}} \left(\sum_{j \in \mathcal{N}_i^{\mathrm{in}}} q_{ji}(t) - \sum_{j \in \mathcal{N}_i^{\mathrm{out}}} q_{ij}(t) \right), \tag{3.2}$$

where Δt is the duration of the time interval $(t, t+1]$. The cross-section area of tank is defined by A_i^{TK}.

Pumps

The pumps provide head gain in the water distribution networks on the links $(i,j) \in \mathcal{M}$ connecting the suction j^{th} node and the delivery i^{th} node. The head gain explicitly depends on the pump flow and pump property. Now we consider the variable speed pump (VSP) in the network, and the head gain is given by

$$h_{ij}^{\mathrm{M}}(t) = h_i(t) - h_j(t) = \alpha_{ij} q_{ij}^{\mathrm{M}}(t)^2 + \beta_{ij} q_{ij}^{\mathrm{M}}(t) + \gamma_{ij}, \tag{3.3}$$

where coefficients α_{ij}, β_{ij}, and γ_{ij} are determined by the pump operation curve.

Pipes

The head loss of of pipe $(i,j) \in \mathcal{I}$ described via the empirical Chezy-Manning (C-M) is given as follows

$$h_{ij}^{\mathrm{P}} = h_i(t) - h_j(t) = R_{\mathrm{CM},ij} q_{ij}^{\mathrm{P}}(t)^2, \tag{3.4}$$

where the resistance coefficient is denoted by $R_{\mathrm{CM},ij} \in \mathbf{R}_{++}$ and defined by [228]

$$R_{\mathrm{CM},ij} = 4.66 \frac{L_{\mathrm{CM},ij} C_{\mathrm{CM}}^2}{D_{\mathrm{CM},ij}^{5.33}}.$$

Note that $C_{\mathrm{CM}} \in \mathbf{R}_{++}$ is the Manning roughness coefficient; $D_{\mathrm{CM},ij} \in \mathbf{R}_{++}$ is the diameter of pipeline in feet; and $L_{\mathrm{CM},ij} \in \mathbf{R}_{++}$ is the length of pipeline in feet.

Pressure Reduce Valves

There are several types of controllable valves in a water distribution network, such as pressure reduce valves (PRVs), general purpose valves (GPVs) and flow control valves (FCVs), associated with different control variables: valve openness, pressure reduction, and flow regulation. Here, we utilize PRVs to restrict the pressure to a certain difference $\phi_{ij} \in \mathbf{R}_+, (i,j) \in \mathcal{V}$ along a pipeline when the upstream pressure at i^{th} node is higher than the downstream j^{th} node

$$\phi_{ij}(t) = h_i(t) - h_j(t), \tag{3.5}$$

where $h_i(t)$ the head of upstream junction and $h_j(t)$ is the head of downstream junction, and the variable ϕ_{ij} determines the energy conservation on pipeline (i,j). Note that no reverse flow on PRVs is allowed, and the water flow through PRVs, $q_{ij}, (i,j) \in \mathcal{V}$, is not determined by (3.5), thereby depends on other network coupling constraints (3.1). The implementation of valve control actions depends on valve construction. We refer interested readers to [98, 240, 213, 8, 253] for more details. The deployment of PRVs in water distribution networks can promote the potential control availability. Here, we utilize a "smart" PRV, whose pressure reduce setting can be optimized in the real time.

Network Operational Constraints

We specify several constraints on network states and inputs in our proposed stochastic OWF problem to satisfy the physical operation limitation of water distribution networks (i.e., limits of nodal heads, pipe flows and tank levels)

$$h_i^{\min} \leq h_i(t) \leq h_i^{\max}, \quad \forall i \in \mathcal{N},$$

$$q_{ij}^{\min} \leq q_{ij}(t) \leq q_{ij}^{\max}, \quad \forall (i,j) \in \mathcal{E},$$

where $h_i^{\min}$ and $h_i^{\max}$ are the lower and upper heads on i^{th} node and $q_{ij}^{\min}$ and $q_{ij}^{\max}$ are minimum and maximum flows on link (i,j). We introduce a binary parameter $z_{ij}(t)$ for link

(i, j) to indicate the ON/OFF status of the controllable devices (i.e., pumps, valves). Then the head coupling between two neighboring nodes can be modelled as follows

$$\begin{aligned} -M\left(1-z_{ij}(t)\right) \\ \leq h_{ij}(t) - g(q_{ij}(t), \phi_{ij}(t)) \leq M\left(1-z_{ij}(t)\right), \end{aligned} \tag{3.7}$$

where $h_{ij}(t) := h_i(t) - h_j(t)$, $g(\cdot)$ is a general expression of (3.3), (3.4), and (3.5) as functions of q_{ij} and ϕ_{ij}, and M is a large positive constant. Note that when $z_{ij}(t) = 1$, the device on link (i, j) is ON (i.e., $q_{ij}(t) \neq 0$), then the energy conservation constraints (3.3), (3.4), and (3.5) hold on this link; otherwise $z_{ij}(t) = 0$, and the head at i^{th} node and j^{th} node are decoupled. For the links without a controllable device (e.g., pipes), we let $z_{ij}(t) = 1$ for all time intervals.

Table 3.1. Variable notations

Notation	Description
h_i	Head at node i
$h_i^{\text{TK}}/h_i^{\text{R}}$	Head at tank/reservoir i
$h_{ij}^{\text{P}}/\phi_{ij}$	Head loss on the pipe/valve from node i to node j
h_{ij}^{M}	Head gain on the pump from node i to node j
q_{ij}	Flow through on the link from node i to node j
$q_{ij}^M/q_{ij}^P/q_{ij}^V$	Flow through on pump/pipe/valve from node i to node j

3.2.2 Leveraging Linear Approximation of Hydraulic Coupling

The nonlinear energy conservation (3.7) renders the water flow formulation nonconvex. This hinders the development of a computationally efficient stochastic optimal water flow problem where distributionally robust optimization and risk measures can be utilized to balance system performance and robustness. To that end, we provide a Laplacian-based linearization

that utilizes the successive linearization algorithm to enable a highly accurate approximation of the original nonlinear energy conservation (3.7).

The energy conservation (3.3), (3.4) and (3.5) can be concluded in a compact matrix form

$$\mathbf{B}_f\mathbf{h}(t) = \mathbf{q}^\intercal(t)\mathbf{N}\mathbf{q}(t) + \mathbf{P}\mathbf{q}(t) + \mathbf{q}_0, \tag{3.8}$$

where $\mathbf{h}(t) := [h_1(t), \ldots, h_N(t)]^\intercal \in \mathbf{R}^N$ and $\mathbf{q}(t) := \{q_{ij}(t)|(i,j) \in \mathcal{I} \cup \mathcal{M}\} \cup \{\phi_{ij}(t)|(i,j) \in \mathcal{V}\} \in \mathbf{R}^{|\mathcal{E}|}$ collect network states, e.g., head, flow, and valve settings. The constant matrices/vector $\mathbf{N} \in \mathbf{R}^{|\mathcal{E}|\times|\mathcal{E}|}$, $\mathbf{P} \in \mathbf{R}^{|\mathcal{E}|\times|\mathcal{E}|}$, and $\mathbf{q}_0 \in \mathbf{R}^{|\mathcal{E}|}$ explicitly depend on the property of pipelines, pump, and valves. The incident matrix of graph $\mathcal{G}$ is denoted by $B_f \in \mathbf{R}^{|\mathcal{E}|\times|N|}$, having entries

$$\mathbf{B}_f(n,i) = \begin{cases} 1 & \text{if flow in } n^{\text{th}} \text{ link is away from } i^{\text{th}} \text{ node} \\ -1 & \text{if flow in } n^{\text{th}} \text{ link is towards } i^{\text{th}} \text{ node} \\ 0 & \text{if flow in } n^{\text{th}} \text{ link is not incident on } i^{\text{th}} \text{ node} \end{cases}. \tag{3.9}$$

The nonlinearities in (3.8) make the OWF problem non-convex and computationally challenging. Therefore, in the rest of this subsection, we will seek to linearize (3.8) instead. We express the flow as $\mathbf{q} = \bar{\mathbf{q}} + \Delta\mathbf{q}$, where $\bar{\mathbf{q}} \in \mathbf{R}^{|\mathcal{E}|}$ is the nominal water flow vector, and $\Delta\mathbf{q} \in \mathbf{R}^{|\mathcal{E}|}$ captures disturbances around the nominal values. To lighten notation we omit the time index in the discussion of linearization in this section. Substituting $\mathbf{q} = \bar{\mathbf{q}} + \Delta\mathbf{q}$ into (3.8), we have

$$\begin{aligned}\mathbf{B}_f\mathbf{h} = &\left(\bar{\mathbf{q}} + \Delta\mathbf{q}\right)^\intercal \mathbf{N}\left(\bar{\mathbf{q}} + \Delta\mathbf{q}\right) + \mathbf{P}\left(\bar{\mathbf{q}} + \Delta\mathbf{q}\right) + \mathbf{q}_0 \\ &+ \bar{\mathbf{q}}^\intercal\mathbf{N}\bar{\mathbf{q}} + \mathbf{P}\bar{\mathbf{q}} + \left(2\bar{\mathbf{q}}^\intercal\mathbf{N} + \mathbf{P}\right)\Delta\mathbf{q} + \Delta\mathbf{q}^\intercal\mathbf{N}\Delta\mathbf{q} + \mathbf{q}_0.\end{aligned}$$

Neglecting second-order terms in $\Delta\mathbf{q}$, (3.8) becomes approximately

$$\mathbf{B}_f\mathbf{h} \approx \underbrace{\bar{\mathbf{q}}^\intercal\mathbf{N}\bar{\mathbf{q}} + \mathbf{P}\bar{\mathbf{q}} + \mathbf{q}_0}_{\mathbf{A}} + \underbrace{\left(2\bar{\mathbf{q}}^\intercal\mathbf{N} + \mathbf{P}\right)}_{\mathbf{B}}\Delta\mathbf{q},$$

where matrices $\mathbf{A}, \mathbf{B} \in \mathbf{R}^{|\mathcal{E}|\times|\mathcal{E}|}$. Now, we turn our attention to solving for the water flow perturbation vector $\Delta\mathbf{q}$. Decomposing all energy conservation on each pipeline, we can write the above linearization in the scalar form

$$h_i - h_j = a_{ij} + b_{ij}\Delta q_{ij}, \quad \forall (i,j) \in \mathcal{E},$$

where a_{ij} and b_{ij} denote the elements of $\mathbf{A}$ and $\mathbf{B}$, respectively. The water flow perturbation on each link is

$$\Delta q_{ij} = \frac{1}{b_{ij}}(h_i - h_j) - \frac{1}{b_{ij}}a_{ij}, \quad \forall (i,j) \in \mathcal{E}.$$

The sum of water perturbation carrying away from i^{th} node is defined as $\Delta Q_i \in \mathbf{R}$ around the nominal operation point $\bar{\mathbf{q}}$ given by

$$\Delta Q_i = \sum_j \left[\frac{1}{b_{ij}}(h_i - h_j)\right] - \underbrace{\sum_j \left[\frac{a_{ij}}{b_{ij}}\right]}_{\bar{Q}_i}, \quad \forall i \in \mathcal{N}. \tag{3.10}$$

Note that the first term in (3.10) can be expressed using the network Laplacian matrix $\mathbf{L} \in \mathbf{R}^{N\times N}$ defined by the edge weights $\frac{1}{b_{ij}}$. The second term in (3.10) is the nominal carrying flow $\bar{Q}_i \in \mathbf{R}$ at i^{th} node. Then the linear energy conservation of water distribution network in compact form is

$$\mathbf{L}\mathbf{h} = \Delta\mathbf{Q} + \bar{\mathbf{Q}},$$

where the network's Laplacian matrix $\mathbf{L}$ has elements

$$\mathbf{L}_{ij} = \begin{cases} \sum_{l\sim i} \frac{1}{b_{il}} & \text{if } i = j \\ -\frac{1}{b_{ij}} & (i,j) \in \mathcal{E} \\ 0 & (i,j) \notin \mathcal{E} \end{cases}.$$

We define two vectors as $\Delta\mathbf{Q} := [\Delta Q_1, \ldots, \Delta Q_N]^\intercal$ and $\bar{\mathbf{Q}} := [\{\bar{Q}_i \in \mathbf{R} | \bar{Q}_i = \sum_j \frac{a_{ij}}{b_{ij}}\}$. Given the incidence matrix defined in (3.9), the following energy conservation constraint holds as a function of water flow perturbations

$$\mathbf{L}\mathbf{h} = \mathbf{B}_f^\intercal \Delta\mathbf{q} + \bar{\mathbf{Q}}, \tag{3.11}$$

where the Laplacian-based compact form maps the water flow disturbance $\Delta\mathbf{q}$ to the network head $\mathbf{h}$ around the linearized point of nodal water carrying $\bar{\mathbf{Q}}$.

3.2.3 Verifying Feasibility of Laplacian Approximation

To validate the effectiveness and feasibility of the proposed Laplacian approximation, we solve a water flow feasibility problem

$$\textbf{WFP-0:} \quad \min_{\mathbf{h},\Delta\mathbf{q}} \quad 0 \qquad \text{subject to} \quad (3.1) \text{ and } (3.11). \tag{3.12}$$

by utilizing successive linearization algorithm, and compare the solutions to the water flow results from EPANET [228] modelled via the nonlinear energy conservation constraints (3.8). The overall successive linearization process is presented in Algorithm 1.

We empirically observe that the successive linearization algorithm for WDN in Algorithm 1 provides a nearly exact approximation of the nonlinear hydraulic dynamics $(\mathbf{q}, \mathbf{h})$ and flow directions if the following assumptions holds

Assumption 1. *The water distribution network $\mathcal{G}$ is a pure tree.*

Assumption 2. *Pumps and valves are all active, i.e., $z_{ij} = 1, \forall (i,j) \in \mathcal{M} \cup \mathcal{V}$.*

The main purpose of this chapter is to seek the optimal pump schedules satisfied the network constraints (e.g., flows and head limits) under network uncertainties. The proposed stochastic OWF formulation is designed for the real-time optimal control, which is based on the pre-determined operation status of all active devices, i.e., z_{ij} is not an optimization variable in (3.7). Additionally, we assume the network topology is a pure tree. This ensures that the successive linearization algorithm converges (3.12) to a nonlinear feasible point and gives the correct directions of water flows.

[2]The EPANET network information source file contains the topology of water networks and the properties of the hydraulic components (e.g., pipes, pump, tank and etc..).

Algorithm 1 Successive Linearization Algorithm for Water Distribution Networks

Input:.EPANET Network Information `.inp` source file[2], demand $\mathbf{d}$, incidence matrix B_f

Output: System operating points $\mathbf{h}^*$ and $\mathbf{q}^*$

Initialize: $n = 0$, initial nominal water flow $\bar{\mathbf{q}}_0$ and Linearization error Err $> \delta$
while Err $> \delta$ **do**
    Calculate Laplacian L_n and flow carrying vector $\bar{Q}_n$ for n^{th} iteration
    Solve **WFP-0** for $\Delta\mathbf{q}_n^*$ and $\mathbf{h}_n^*$ from (3.12)
    Compute the linearization error, Err $= \|\Delta\mathbf{q}_n^*\|_2^2$
    Update nominal water flow points $\bar{\mathbf{q}}_{n+1} = \bar{\mathbf{q}}_n + \Delta\mathbf{q}_n^*$
    $n = n + 1$
end while
return solutions $\mathbf{q}^* = \bar{\mathbf{q}}_{n+1}$, $\mathbf{h}^* = \mathbf{h}_n$

Remark 1. *(Hydraulic-Network Simplification). Our proposed framework focuses on optimal operation for water distribution networks instead of hydraulic design and analysis. Therefore, we assume all networks are pure trees, where the proposed linearization approach is highly effective. Many networks can be simplified or approximated as a pure tree network using various techniques, which facilitates a higher-level interpretation of the main network structure [207, 10].*

Remark 2. *(Successive Linearization Initialization). The initial nominal water flow* $\mathbf{q}_0$ *is an input of successive linearization algorithm for WDNs. We suggest here a possible* $\mathbf{q}_0$ *for various components (e.g., pumps and pipes) to initialize Algorithm 1 for improved convergence. For all pipes* $\mathcal{I}$ *in WDNs, the initial water flow is corresponding to the flow speed* 1 *CFS [228]. The actual input initial water flow is adjusted based on the properties of individual pipeline (i.e., length and diameter). The initial linearized water flow point of pumps will come from the pump efficiency curve [258]*

$$E_{ij} = e_{ij}^1 q_{ij}^2 + e_{ij}^2 q_{ij} + e_{ij}^3, \quad \forall (i,j) \in \mathcal{M}.$$

The successive linearization of pumps starts from the most efficient point as

$$\frac{\partial E_{ij}}{\partial q_{ij}} = 2e^1_{ij} q_{ij} + e^2_{ij} = 0, \quad \bar{q}_{ij} = \frac{e^2_{ij}}{2e^1_{ij}}, \quad \forall (i,j) \in \mathcal{M}.$$

Empirical speaking, starting points satisfying the physical constraints often lead to a feasible solution.

3.3 Data-based Multi-period Stochastic Optimal Water Flow

In this section, we formulate a stochastic OWF problem as a distributionally robust stochastic optimal control problem. We first pose the problem generically to highlight the overall approach, and in subsequent sections we incorporate the linearization of hydraulic modelling in Section 3.2 for a tractable and computationally-efficient stochastic OWF. This framework is more general than most stochastic OWF in the literature, which typically focus only on individual or single-stage optimization problems, or has a less sophisticated approach for explicitly incorporating uncertainties. Consider a multi-period data-driven distributionally robust optimization problem

$$\inf_{\pi \in \Pi} \sup_{\mathbb{P} \in \mathcal{P}} \quad \mathbb{E}^{\mathbb{P}} \sum_{t=0}^{T} h_t(x_t, u_t, \xi_t) \tag{3.13a}$$

$$\text{subject to} \quad x_{t+1} = f_t(x_t, u_t, \xi_t) \tag{3.13b}$$

$$u_t = \ \pi(x_0, \ldots, x_t, \boldsymbol{\xi}_t, \mathcal{D}_t) \tag{3.13c}$$

$$(x_t, u_t) \in \mathcal{X}_t \tag{3.13d}$$

where $x_t \in \mathbf{R}^n$ represent the state vector at time t that includes the internal states of all elements (i.e., valves, tanks and pipes). Let $u_t \in \mathbf{R}^m$ denote a control input vector that includes inputs for all controllable components (e.g., pump output and valve settings). The $\xi_t \in \mathbf{R}^{N_\xi}$ denote random vectors in a probability space $(\Omega, \mathcal{F}, \mathbb{P}_t)$ which includes forecast errors of all uncertainties in the network.

The goal of (3.13) is to find a optimal feedback policy that minimizes the expected value of the system objective function $h_t : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{N_\xi} \to \mathbf{R}$ robust to the worst-case distribution in the forecast error ambiguity set $\mathcal{P}$. We consider a setting where the objective function h_t includes both operating costs and risks of violating various network and device constraints and is assumed to be continuous and convex as functions of (x_t, u_t) for any fixed ξ_t. The system dynamics function $f_t : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{N_\xi} \to \mathbf{R}^n$ models internal dynamics of all network-connected components, such as water storage tanks. The general feasible set $\mathcal{X}_t$ includes other network and device constraints, such as mass balance, energy conservation, operational bounds on nodal heads and pipe flows (some constraints may be modeled deterministically with respect to mean values and others may be included as risk terms in the objective function).

Since the real distributions of forecast errors are never known in practice, we explicitly account for uncertainty in their distributions themselves by assuming that the real but unknown distribution $\mathbb{P}_t$ belongs to an *ambiguity set* $\mathcal{P}_t$ of distributions which will be constructed from a forecast sampling dataset.We collect the forecast error over an operating horizon t as $\boldsymbol{\xi}_t := [\xi_1^\intercal, \ldots, \xi_t^\intercal]^\intercal \in \mathbf{R}^{N_\xi t}$, which has joint distribution $\mathbb{P}$ and corresponding ambiguity set $\mathcal{P}$.

In this multi-period stochastic OWF, we are seeking a series of closed-loop feedback policies in the form $u_t = \pi(x_0, \ldots, x_t, \boldsymbol{\xi}_{0:t}, \mathcal{D}_t)$ explicitly considering forecast errors describing historical patterns, where the term $\mathcal{D}_t$ indicates all network component model information and the parameterization of the ambiguity set of the forecast error distribution. This framework allows for design of not only for current nominal reaction, but also reactions to future uncertainty realizations. The policy function π maps all available information to control actions and is an element of a set Π of measurable functions.

3.3.1 Ambiguity Sets based on Wasserstein Metric

One of the main challenges for solving (3.13) is how to utilize our available information of uncertainties to appropriately realize the distributions for a tractable problem reformulation. There is a variety of ways to reformulate the general stochastic OWF problem (3.13) to obtain tractable subproblems that can be solved by standard convex optimization solvers. These include assuming specific functional forms for the forecast error distribution (e.g., Gaussian)[231] and using specific constraint risk functions, such as those encoding value at risk (i.e., chance constraints)[117, 119], conditional value at risk (CVaR) [124], distributional robustness[124], and support robustness[113]. In all cases, the out-of-sample performance of the resulting decisions in operational practice ultimately relies on 1) the quality of data describing the forecast errors and 2) the validity of assumptions made about probability distributions. Many existing approaches make either too strong or too weak assumptions that possible lead to underestimation or overestimation of risk.

In this chapter, we utilize a recently proposed tractable method [96] in a multi-period data-based stochastic OWF, in which the ambiguity set is based on a finite forecast error training dataset $\hat{\Xi}^{N_s}$ via Wasserstein balls. Comparing with others existing ambiguity sets [158, 272, 35, 278, 137, 95], Wasserstein balls offer the powerful out-of-sample performance and provide the water distribution network operators to control the conservativeness of the decisions, which promote the flexibility of water distribution network from a practical perspective. We optimize an expected objective over the worst-case distribution in the ambiguity set $\mathcal{P}$, which can be formulated as a finite-dimensional convex program. The decisions from this stochastic OWF provide an upper confidence bound under forecast errors realization, quantified by the size of the *ambiguity set* (i.e., Wasserstein radius [96]).

The Wasserstein metric defines a distance in the space $\mathcal{M}(\Xi)$ of all probability distributions $\mathbb{Q}$ supported on a set Ξ with $\mathbb{E}^{\mathbb{Q}}[\|\xi\|] = \int_{\Xi} \|\xi\| \mathbb{Q}(d\xi) < \infty$.

Definition 1 (Wasserstein Metric [143]). *Given all distributions* $\mathbb{Q}_1$,$\mathbb{Q}_2$ *supported on* Ξ*, the Wasserstein metric* $d_W : \mathcal{M}(\Xi) \times \mathcal{M}(\Xi) \to \mathbf{R}_+$ *is defined as*

$$d_W\left(\mathbb{Q}_1, \mathbb{Q}_2\right) := \int_{\Xi} \|\xi_1 - \xi_2\| \Theta(d\xi_1, d\xi_2),$$

where Θ *represent a joint distribution of* ξ_1 *and* ξ_2 *with marginals* $\mathbb{Q}_1$ *and* $\mathbb{Q}_2$*, respectively, and* $\|\cdot\|$ *indicates an arbitrary norm on* $\mathbf{R}^{N_\xi}$.

The Wasserstein metric quantifies the "transportation costs" to move mass from one distribution to another. The Wasserstein ambiguity set is defined by

$$\hat{\mathcal{P}}^{N_s} := \left\{ \mathbb{Q} \in \mathcal{M}(\Xi) : d_W(\hat{\mathbb{P}}^{N_s}, \mathbb{Q}) \leq \epsilon \right\}. \tag{3.14}$$

This ambiguity set $\hat{\mathcal{P}}^{N_s}$ constructs a ball with radius ϵ in Wasserstein distance around the empirical distribution $\hat{\mathbb{P}}^{N_s}$ on the training dataset. The radius ϵ can be chosen so that the ball contains the true distribution $\mathbb{P}$ with a prescribed confidence level and leads to performance guarantees [96]. The radius ϵ also explicitly controls the conservativeness of the resulting decision. Large ϵ would produce decisions that rely less on the specific features of the uniform empirical distribution supported by the training dataset $\hat{\Xi}^{N_s}$ and improve robustness to inherent sampling errors. We will discuss the use of this conservativeness index for our stochastic OWF problem.

3.3.2 Data-based Distributionally Robust Model Predictive Control of Optimal Water Flow

The goal of our data-based distributionally robust stochastic OWF framework is to interpret and demonstrate inherent tradeoffs between efficiency and risk of constraint violations. Accordingly, the objective function comprises a weighted sum of an operational cost function and a constraint violation risk function: $h_t = J^t_{\text{Cost}} + \rho J^t_{\text{Risk}}$, where $\rho \in \mathbf{R}_+$ is a weight that

quantifies the network operator's risk aversion. The operational cost function is assumed to be linear or convex quadratic. The cost functions will be discussed in detail in Section 3.4.

The constraint violation risk function J_{Risk} comprises a sum of the conditional value-at-risk (CVaR) [227] of a set of N_ℓ network and device constraint functions. The conditional value-at-risk is a well known and coherent risk measurement in finance [227]. Here we introduce the CVaR risk metric to solve a MPC-based OWF engineering problem, due to the large variation of water demand uncertainties. Minimizing the CVaR of constraint violation limits both the frequency and expected severity of constraints. Specifically, we have

$$J^t_{\text{Risk}} := \sum_{i=1}^{N_\ell} \text{CVaR}^{\beta}_{\mathbb{P}}[\ell_i(x_t, u_t, \xi_t)],$$

where $\beta \in (0, 1]$ refers to the confidence level of CVaR under the distribution $\mathbb{P}$ of random variable ξ_t. Intuitively, the constraint violation risk function J_{Risk} could be understood as the sum of networks and devices constraint violation magnitude at risk level β. The details of CVaR constraint convex reformulation are shown in the next Section.

The general problem (3.13) will be approached with a distributionally robust model predictive control (MPC) algorithm. MPC is a feedback control technique that solves a sequence of open-loop optimization problems over a planning horizon $\mathcal{H}_t$ (which in general may be smaller than the overall horizon T). At each time t, we solve the distributionally robust optimization problem over a set Π_{affine} of affine feedback policies using the Wasserstein ambiguity set (3.14)

Distributionally Robust MPC for Stochastic OWF:

$$\inf_{\pi \in \Pi_{\text{affine}}} \sup_{\mathbb{P} \in \hat{\mathcal{P}}^{N_s}} \quad \mathbb{E}^{\mathbb{P}} \sum_{\tau=t}^{t+\mathcal{H}_t} J^\tau_{\text{Cost}} + \rho J^\tau_{\text{Risk}} \tag{3.15a}$$

$$\text{subject to} \quad x_{\tau+1} = f_\tau(x_\tau, u_\tau, \xi_\tau) \tag{3.15b}$$

$$u_\tau = \pi(x_0, \ldots, x_\tau, \boldsymbol{\xi}_\tau, \mathcal{D}_\tau) \tag{3.15c}$$

$$(x_\tau, u_\tau) \in \mathcal{X}_\tau. \tag{3.15d}$$

Only the immediate control decisions for time t are implemented on the controllable device inputs. Then time shifts forward one step, new forecast errors and states are realized, the optimization problem (3.15) is re-solved at time $t+1$, and the process repeats. This approach allows any forecasting methodology to be utilized to predict uncertainties over the planning horizon. Furthermore, the forecast error dataset $\hat{\mathbb{P}}^{N_s}$, which defines the center of the ambiguity set $\hat{\mathcal{P}}^{N_s}$, can be updated online as more forecast error data is obtained. It is also possible to remove outdated data online to account for time-varying distributions.

In the rest of the chapter, we will use the specific model of water distribution networks discussed in Section 3.2, where the subproblems (3.15) have exact tractable convex reformulations as quadratic programs [96] and can be solved to global optimality with standard solvers.

3.4 Chance-constraints and Distributionally Robustness Formulation

Following our proposed formulation above, we begin this section by introducing the state space expression of WDN hydraulic dynamics, briefly discuss chance constraints, and describe a convex reformulations of the stochastic optimal water flow problem based on conditional value-at-risk and distributionally robust optimization.

3.4.1 Network Dynamics in State-Space Format

The WDN model discussed in the previous section can be summarized in a difference algebraic equation (DAE) model

$$x(t+1) = \bar{A}x(t) + \bar{B}_u u(t) + \bar{B}_v v(t), \tag{3.16a}$$

$$d(t) = \bar{E}_u u(t) + \bar{E}_v v(t), \tag{3.16b}$$

$$\bar{F}_x x(t) + \bar{F}_l l(t) = \bar{F}_u u(t) + \bar{F}_v v^{\mathrm{P}}(t) + \bar{F}_\phi \phi(t) + \bar{F}_0, \tag{3.16c}$$

where the decision variables $\{x, u, l, v, v^{\mathrm{P}}, \phi\}$ are defined in Table 3.2 and the constant matrices $\{A, B, E, F\}$ are derived from the hydraulic dynamics in Section 3.2. We detail these constants in term of the Laplacian-based hydraulic model (3.11) in the Appendix. The dynamics of tank head (3.2) is given in (3.16a), and the mass balance (3.1) and linearized energy conservation (3.11) are summarized in (3.16b) and (3.16c), respectively. For compact notation, we concatenate the states, inputs and demands over the planning horizon as $\mathbf{x}_t = [x(1)^\intercal, \ldots, x(t)^\intercal]^\intercal \in \mathbf{R}^{n_{\mathrm{TK}}t}$, $\mathbf{u}_t = [u(0)^\intercal, \ldots, u(t-1)^\intercal]^\intercal \in \mathbf{R}^{n_u t}$, $\mathbf{v}_t = [v(1)^\intercal, \ldots, v(t)^\intercal]^\intercal \in \mathbf{R}^{n_v t}$, $\mathbf{v}_t^{\mathrm{P}} = [v^{\mathrm{P}}(1)^\intercal, \ldots, v^{\mathrm{P}}(t)^\intercal]^\intercal \in \mathbf{R}^{n_p t}$, $\mathbf{l}_t = [l(1)^\intercal, \ldots, l(t)^\intercal]^\intercal \in \mathbf{R}^{n_l t}$, $\boldsymbol{\phi}_t = [\phi(1)^\intercal, \ldots, \phi(t)^\intercal]^\intercal \in \mathbf{R}^{n_\phi t}$ and $\mathbf{d}_t = [d(0)^\intercal, \ldots, d(t-1)^\intercal]^\intercal \in \mathbf{R}^{Nt}$, yielding

$$\mathbf{x}_t = Ax_0 + B_u\mathbf{u}_t + B_v\mathbf{v}_t,$$

$$\mathbf{d}_t = E_u\mathbf{u}_t + E_v\mathbf{v}_t,$$

$$F_x\mathbf{x}_t + F_l\mathbf{l}_t = F_u\mathbf{u}_t + F_v\mathbf{v}_t^{\mathrm{P}} + F_\phi\boldsymbol{\phi}_t + F_0,$$

where I_t indicates a t-dimensional identity matrix

$$E_u = I_t \otimes \bar{E}_u, \quad E_v = I_t \otimes \bar{E}_v, \quad F_x = I_t \otimes \bar{F}_x,$$

$$F_l = I_t \otimes \bar{F}_l, \quad F_u = I_t \otimes \bar{E}_u, \quad F_v = I_t \otimes \bar{F}_v,$$

$$F_\phi = I_t \otimes \bar{F}_\phi, \quad F_0 = I_t \otimes \bar{F}_0,$$

$$A = \begin{bmatrix} \bar{A} \\ \bar{A}^2 \\ \vdots \\ \bar{A}^t \end{bmatrix}, B_d = \begin{bmatrix} \bar{B}_u & 0 & \cdots & 0 \\ \bar{A}\bar{B}_u & \bar{B}_u & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \bar{A}^{t-1}\bar{B}_u & \cdots & \bar{A}\bar{B}_u & \bar{B}_u \end{bmatrix},$$

$$B_v = \begin{bmatrix} \bar{B} & 0 & \cdots & 0 \\ \bar{A}\bar{B}_v & \bar{B}_v & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \bar{A}^{t-1}\bar{B}_v & \cdots & \bar{A}\bar{B}_v & \bar{B}_v \end{bmatrix}.$$

Table 3.2. Variable description in DAE model

Notation	Description	Dimension
x	a vector collecting heads at tanks	$n_{\text{TK}} = \lvert\mathcal{T}\rvert$
l	a vector collecting heads at junctions & reservoirs	$n_l = \lvert\mathcal{J}\rvert + \lvert\mathcal{S}\rvert$
u	a vector collecting flow at pumps	$n_u = \lvert\mathcal{M}\rvert$
v	a vector collecting flows through pipes & valves	$n_v = \lvert\mathcal{I}\rvert + \lvert\mathcal{V}\rvert$
v^{P}	a vector collecting flows through pipe	$n_p = \lvert\mathcal{I}\rvert$
ϕ	a vector collecting head loss on PRVs	$n_\phi = \lvert\mathcal{V}\rvert$

3.4.2 Cost Functions and Constraints

Multiple objective functions can be included in the stochastic optimal water flow problem

$$J_1^t = u(t)^\intercal H_u(t)u(t) + f_u^\intercal(t)u(t) + f_0, \tag{3.18a}$$

$$J_2^t = \Delta u(t)^\intercal \Delta u(t), \tag{3.18b}$$

$$J_3^t = \left(x(t) - V^{\text{safe}}\right)^\intercal \left(x(t) - V^{\text{safe}}\right), \tag{3.18c}$$

where (3.18a) captures the pump operational cost based on time-varying electricity tariffs. The matrix H_u is positive semi-definite. The control input variation between consecutive time slots (e.g., $\Delta u(t) := u(t) - u(t-1)$) can be also penalized in (3.18b) to avoid large transient in pipes, and to satisfy treatment requirements. Additionally, tank management requires a safety head level V^{safe} to account for unexpected demand given in (3.18c).

The system constraints are introduced due to the physical nature of the decision variables (i.e, $\mathbf{x}$ and $\mathbf{u}$). We seek to enforce state and input constraints

$$\mathbf{u}^{\min} \leq \mathbf{u}_t \leq \mathbf{u}^{\max}, \tag{3.19a}$$

$$\mathbf{x}^{\min} \leq \mathbf{x}_t \leq \mathbf{x}^{\max}. \tag{3.19b}$$

where (3.19a) corresponds to actuator limits (e.g., pumps and valves) and (3.19b) captures bounds on pipe flows, nodal heads and tank levels. Here, $\mathbf{x}_{\min}$ and $\mathbf{x}_{\max}$ denote the minimum

and maximum admissible bounds of states. The lower and upper physical limits of actuators are $\mathbf{u}_{\min}$ and $\mathbf{u}_{\max}$, respectively. In general, these constraints can not be violated strictly due to the mass conservation principles and physical restriction of components. For the rest of this chapter, we assume that these hard bounds can be "softened" to non-physical upper and lower bounds from a pre-specified safe operation zone, which can be violated probabilistically but results in safety or operational risk [119].

3.4.3 Multi-Period Stochastic Optimal Water Flow

In a deterministic optimal water flow control problem, water demand uncertainty is not explicitly considered. Since actual water demands can exhibit large variations and unpredictability [44], we model demand stochastically as $\mathbf{d}_t = \bar{\mathbf{d}}_t + \boldsymbol{\xi}_t$, with a nominal predicted value $\bar{\mathbf{d}} \in \mathbf{R}^{Nt}$ and a zero-mean forecast error $\boldsymbol{\xi}_t = [\xi_1^\intercal, \ldots, \xi_t^\intercal]^\intercal \in \mathbf{R}^{Nt}$ from a probability space $(\Omega, \mathcal{F}, \mathbb{P}_\xi)$. The distribution captures spatiotemporal variations and dependencies among the demands.

To explicitly account for this stochasticity of water demands, we formulate the following general stochastic optimal water flow problem to find an optimal strategy for responding to forecast errors via an optimal control policy for the flow actuators $\mathbf{u}_t = \pi_t(\boldsymbol{\xi}_t)$, where $\pi_t : \mathbf{R}^{Nt} \to \mathbf{R}^{n_u t}$ is a function from a set Π_c of causal policies. Specifically, we consider a multi-period optimal water flow problem with finite time horizon T

$$\inf_{\pi_T \in \Pi_c} \quad \sum_{\tau=1}^{T} \mathbb{E}^{\mathbb{P}_\xi}\big[J^\tau(\mathbf{x}_\tau, \pi_\tau(\boldsymbol{\xi}_\tau), \boldsymbol{\xi}_\tau\big], \tag{3.20a}$$

$$\text{subject to} \quad \mathbf{d}_T = E_u \pi_T(\boldsymbol{\xi}_T) + E_v \mathbf{v}_T, \tag{3.20b}$$

$$\mathbf{x}_T = A x_0 + B_u \pi_T(\boldsymbol{\xi}_T) + B_v \mathbf{v}_T, \tag{3.20c}$$

$$F_x \mathbf{x}_T + F_l \mathbf{l}_T = F_u \mathbf{u}_T + F_v \mathbf{v}_T^{\mathrm{P}} + F_\phi \boldsymbol{\phi}_T + F_0, \tag{3.20d}$$

$$\mathcal{R}\big(\mathbf{u}_{\min} - \pi_T(\boldsymbol{\xi}_T)\big) \leq 0, \tag{3.20e}$$

$$\mathcal{R}\big(\pi_T(\boldsymbol{\xi}_T) - \mathbf{u}_{\max}\big) \leq 0, \tag{3.20f}$$

$$\mathcal{R}\big(\mathbf{x}_{\min} - \mathbf{x}_T\big) \leq 0, \tag{3.20g}$$

$$\mathcal{R}\big(\mathbf{x}_T - \mathbf{x}_{\max}\big) \leq 0, \tag{3.20h}$$

where $\mathcal{R}$ donates *risk measure*, which maps a random variable to a real number and will be described in more detail shortly. Note that this transformation can be different in general for each constraint. For constraints that represent physical limits, we consider tightened non-physical upper and lower bounds on states and inputs from a pre-specified safe operation zone, which can be violated probabilistically but results in safety or operational risks [119]. Since optimizing over general policies makes problem (3.20) infinite dimensional, we optimize instead over a set of affine control policies

$$\mathbf{u}_\tau = D_\tau \boldsymbol{\xi}_\tau + e_\tau, \tag{3.21}$$

where $e_\tau \in \mathbf{R}^{n_u \tau}$ represents a nominal plan for pumps, and the block lower-triangular matrix $D_\tau \in \mathbf{R}^{n_u \tau \times N\tau}$ ensures that the controller is causal. In this case, the input design variables turn to an uncertainty feedback matrix D_τ and nominal input vector e_τ. Unlike traditional state-driven feedback control, the optimal feedback matrix D_τ acts as reserve policies of pumps to respond to realized water demand variations $\boldsymbol{\xi}_\tau$.

Substituting the affine control policies into (3.20), the objective function (3.20a) becomes convex quadratic in D_τ and e_τ and depends on the distributional information of $\boldsymbol{\xi}_\tau$. Since the policy is affine, the robust equality constraint (3.20b) is equivalent to

$$E_u D_T = \mathbf{1}, \quad \bar{\mathbf{d}}_T = E_u e_T + E_v \mathbf{v}_T. \tag{3.22}$$

With affine policies, (3.20e)-(3.20h) become

$$\mathcal{R}\left(D_T \boldsymbol{\xi}_T + e_T - \mathbf{u}_{\max}\right) \leq 0, \tag{3.23a}$$

$$\mathcal{R}\left(\mathbf{u}_{\min} - D_T \boldsymbol{\xi}_T - e_T\right) \leq 0, \tag{3.23b}$$

$$\mathcal{R}\left(Ax_0 + B_u(D_T \boldsymbol{\xi}_T + e_T) + B_v \mathbf{v}_T - \mathbf{x}_{\max}\right) \leq 0, \tag{3.23c}$$

$$\mathcal{R}\left(\mathbf{x}_{\min} - Ax_0 - B_u(D_T\boldsymbol{\xi}_T + e_T) - B_v\mathbf{v}_T\right) \leq 0, \tag{3.23d}$$

We collect all above affine constraints inside the risk measures (3.23a)–(3.23d) into a set $\mathbb{V}_{\{1:T\}}$ of $N_\ell = 2T(n_{\text{TK}} + n_u)$ constraints, and the expressions inside the brackets can be written in a general linear form $a_i(D_T)^\intercal\boldsymbol{\xi}_T + b_i(e_T)$, where index i refers to each individual constraint in $\mathbb{V}_{\{1:T\}}$.

3.4.4 Chance-Constraints

Using a Value-ar Risk measure, the OWF problem can be posed as a chance-constrained optimization problem

$$\begin{aligned}
\inf_{D,e} \quad & \sum_{\tau=1}^{T} \mathbb{E}^{\mathbb{P}_\xi}\left[J^\tau(\mathbf{x}_\tau, \mathbf{u}_\tau, \boldsymbol{\xi}_\tau)\right], \\
\text{subject to} \quad & \mathbb{P}_\xi\left(a_i(D_T)^\intercal\boldsymbol{\xi}_T + b_i(e_T) \leq 0\right) \geq 1 - \beta, \\
& E_u D_T = \mathbf{1}, \bar{\mathbf{d}}_T = E_u e_T + E_v\mathbf{v}_T, \\
& F_x\mathbf{x}_T + F_l\mathbf{l}_T = F_u e_T + F_v\mathbf{v}_T^{\text{P}} + F_\phi\boldsymbol{\phi}_T + F_0, \\
& \forall i \in \mathbb{V}_{\{1:T\}},
\end{aligned}$$

where $\beta \in \mathbf{R}$ is the prescribed safety parameter or "risk budget" for the linear constraint in set $\mathbb{V}_{\{1:T\}}$. The subscript $\{1:T\}$ of set $\mathbb{V}_{\{1:T\}}$ indicates the set exclusively includes the state and input constraints between time interval $[1, T]$. If $\boldsymbol{\xi}_T$ is normally distributed, then it is known that the chance constraint can be written as a second-order cone constraint [81, 49]. However, in general chance constraints only restrict the frequency of constraint violations, not the severity. Since the real distribution is never known in practice, this approach can lead to underestimation of actual risks and poor out-of-sample performance. In this chapter, we leverage a data-driven distributionally robust optimization methodology to account for both frequency and severity of constraint violation via conditional value-at-risk (CVaR) metric without assuming a particular distribution.

3.4.5 Stochastic OWF based on Distributionally Robust Optimization and Conditional Value-at-Risk (CVaR)

We treat the constraints (3.23a)–(3.23d) with a risk measure derived from distributionally robust optimization techniques. It is possible to allow some constraints to be reformulated by other risk measures and optimization techniques, such as sample average approximation, moment-based distributionally robust optimization, robust optimization and Gaussian-based chance constraints. We restrict the model here only for Wasserstein metric distributionally robust techniques, and leave potential combinations for the future work.

For simplicity, we consider the risk of each constraint individually; it is possible to consider risk of joint constraint violations, but this is more difficult and we leave it for future work. Recall each individual affine constraint between the finite time horizon $\mathcal{H}_t$ in the set $\mathbb{V}_{\{\underline{t},\bar{t}\}}$ can be written in a compact form as follows. The $[\underline{t},\bar{t}]$ here refers to the finite time horizon $[t, t+\mathcal{H}_t]$.

$$\mathcal{C}_i^t(D_t, e_t, \boldsymbol{\xi}_t) = a_i(D_t)^\intercal \boldsymbol{\xi}_t + b_i(e_t), t \in [\underline{t},\bar{t}],$$

where $\mathcal{C}_i^t(\cdot)$ is the i^{th} affine constraint in the set $\mathbb{V}_{\{\underline{t},\bar{t}\}}$. The CVaR with risk level β of the each individual constraint in the set $\mathbb{V}_{\{\underline{t},\bar{t}\}}$ is

$$\inf_{\kappa_i^t} \mathbb{E}_{\xi_t}\left\{[\mathcal{C}_i^t(D_t, e_t, \boldsymbol{\xi}_t) + \kappa_i^t]_+ - \kappa_i^t\beta\right\} \leq 0, t \in [\underline{t},\bar{t}], \tag{3.25}$$

where $\kappa_i^t \in \mathbf{R}$ is an auxiliary variable [227]. The expression inside the expectation in (3.25) can be expressed in the form with risk level β

$$\mathcal{Q}_i^t = \max_{k=1,2}\left[\langle \mathbf{a}_{ik}^\beta(D_t, e_t), \xi_t\rangle + \mathbf{b}_{ik}^\beta(\kappa_i^t)\right], t \in [\underline{t},\bar{t}].$$

This expression is convex in (D_t, e_t) for each fixed $\boldsymbol{\xi}_t$ since it is the maximum of two affine functions. Our risk objective function is expressed by the distributionally robust optimization

of CVaR

$$\hat{J}^t_{\text{Risk}} = \sum_{\tau=t}^{t+\mathcal{H}_t} \sum_{i=1}^{N_\ell} \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}_\tau^{N_s}} \mathbb{E}^{\mathbb{Q}_\tau} \max_{k=1,2} \left[\langle \mathbf{a}^\beta_{ik}(D_\tau, e_\tau), \hat{\boldsymbol{\xi}}_\tau \rangle + \mathbf{b}^\beta_{ik}(\kappa_i^\tau) \right].$$

The above multi-period distributionally robust optimization can be equivalently reformulated the following quadratic program, the details of which are described in [96]. The objective is to minimize a weighted sum of an operational cost function and the total worst-case CVaR of the affine constraints in set $\mathbb{V}_{\{\underline{t},\bar{t}\}}$ (e.g., nodal head and tank level).

Data-based Distributionally Robust MPC Stochastic OWF:

$$\inf_{\substack{D_\tau, e_\tau \kappa_i^\tau \\ \mathbf{v},\mathbf{v}^{\text{P}},\mathbf{x},\mathbf{l},\boldsymbol{\phi}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^\tau_{\text{Cost}}] + \rho \sup_{\mathbb{Q}_\tau \in \hat{\mathcal{P}}_\tau^{N_s}} \sum_{i=1}^{N_\ell} \mathbb{E}^{\mathbb{Q}_\tau}[\mathcal{Q}_i^\tau] \right\},$$

$$= \inf_{\substack{D_\tau, e_\tau \kappa_i^\tau, \\ \lambda_i^\tau, s_{io}^\tau, \varsigma_{iko}^\tau \\ \mathbf{v},\mathbf{v}^{\text{P}},\mathbf{x},\mathbf{l},\boldsymbol{\phi}}} \sum_{\tau=t}^{t+\mathcal{H}_t} \left\{ \mathbb{E}[\hat{J}^t_{\text{Cost}}] + \sum_{i=1}^{N_\ell} \left(\lambda_i \epsilon_\tau + \frac{1}{N_s} \sum_{o=1}^{N_s} s_{io}^\tau \right) \right\}, \tag{3.26a}$$

subject to

$$\left[E_u D_{\bar{t}} - \mathbf{1} \right]_{[\underline{t},\bar{t}]} = \mathbf{0}_{N\mathcal{H}_t}, \tag{3.26b}$$

$$\left[E_u e_{\bar{t}} + E_v \mathbf{v}_{\bar{t}} - \bar{\mathbf{d}}_{\bar{t}} \right]_{[\underline{t},\bar{t}]} = \mathbf{0}_{N\mathcal{H}_t}, \tag{3.26c}$$

$$\left[F_x \mathbf{x}_{\bar{t}} + F_l \mathbf{l}_{\bar{t}} - F_u e_{\bar{t}} - F_v \mathbf{v}^{\text{P}}_{\bar{t}} - F_\phi \boldsymbol{\phi}_{\bar{t}} - F_0 \right]_{[\underline{t},\bar{t}]} = \mathbf{0}_{N\mathcal{H}_t}, \tag{3.26d}$$

$$\rho(\langle \mathbf{a}^\beta_{ik}(D_\tau, e_\tau), \hat{\boldsymbol{\xi}}^o_\tau \rangle + \mathbf{b}^\beta_{ik}(\kappa_i^\tau) + \langle \varsigma_{iko}, \mathbf{z}_\tau - \mathbf{F}_\tau \hat{\boldsymbol{\xi}}^o_\tau \rangle) \leq s_{io}^\tau, \tag{3.26e}$$

$$\| \mathbf{F}_\tau^{\intercal} \varsigma_{iko} - \rho \mathbf{a}^\beta_{ik}(D_\tau, e_\tau) \|_\infty \leq \lambda_i^\tau, \tag{3.26f}$$

$$\varsigma_{iko} \geq 0, \tag{3.26g}$$

$$\forall o \leq N_s, \forall i \leq N_\ell, k = 1, 2, \tau = t, ..., t + \mathcal{H}_t$$

where $\rho \in \mathbf{R}_+$ quantifies the water network operators' risk aversion. This is a quadratic program that explicitly uses the training dataset $\hat{\Xi}_\tau^{N_s} = \{\hat{\boldsymbol{\xi}}^o_\tau\}_{o \leq N_s}$. The risk aversion parameter ρ and the Wasserstein radius ϵ_τ allow us to explicitly balance tradeoffs between

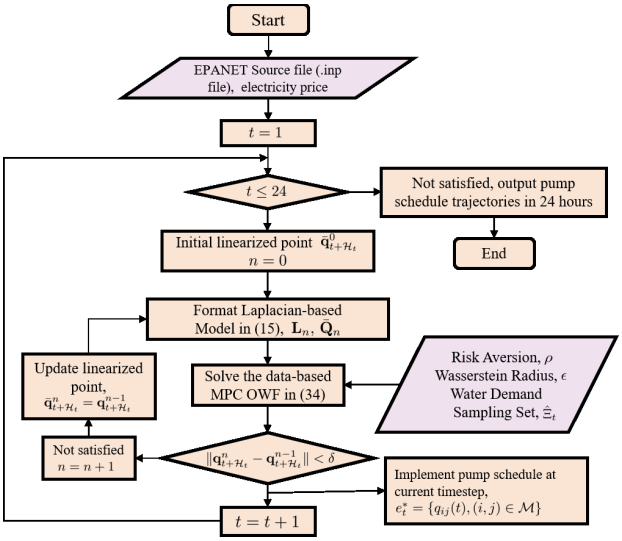


Figure 3.1. Flowchart of data-based distributionally robust stochastic OWF.

efficiency, risk and sampling errors inherent in $\hat{\Xi}^{N_s}_\tau$. The support is modeled as a polytope $\Xi_\tau := \{\boldsymbol{\xi}_\tau \in \mathbf{R}^{N_\xi \tau} : \mathbf{F}_\tau \boldsymbol{\xi}_\tau \leq \mathbf{z}_\tau\}$. The constraint $\varsigma_{iko} > 0$ holds since the uncertainty set is not-empty; on the other hand, in a case with no uncertainty (i.e, $\varsigma_{iko} = 0$), the variable λ does not play any role and $s^\tau_{io} = \rho(\langle \mathbf{a}^\beta_{ik}(D_\tau, e_\tau), \hat{\boldsymbol{\xi}}^o_\tau \rangle + \mathbf{b}^\beta_{ik}(\kappa^\tau_i))$.

Remark 3. *There are three important tuning parameters in our proposed multi-period data-based stochastic OWF* (3.26) *corresponding to different performance-risk tradeoffs, which all function in the ways with their unique interpretation*

- *The* ***Wasserstein radius*** ϵ *improves out-of-sample performance and mitigates the effects of inherent sampling errors, which here is our main focus. The decisions optimize performance under the worst-case distributions within Wasserstein distance* ϵ *of the empirical distribution in probability distribution space. A larger* ϵ *indicates less reliance on the specific training dataset* $\hat{\Xi}$ *that describes the real unknown data-generating distribution, which results in more conservative decisions. The superior out-of-sample performance is achieved by this adjustable Wasserstein metric, as demonstrated in Section 3.5.*

- ***Risk aversion*** *ρ trades off the operational risk and the nominal operational efficiency. The proposed stochastic OWF offers the system operators alternative strategies to run the water distribution networks under different risk levels. The decisions under various ρ achieve various risk levels. Meanwhile, the out-of-sample performance under fixed risk aversion is controlled by the adjustable Wasserstein radius.*

- ***CVaR risk level*** *β indicates the risk level of constraints* (3.23a)–(3.23d)*, which trades off constraint violation magnitudes with nominal operational efficiency.*

We fix the risk level parameter β of CVaR to highlight the effects of ϵ and ρ in the next section. It would also be interesting to explore the additional effects of changing β; however, we leave this discussion for future work.

Remark 4. *(Successive Linearization for Stochastic OWF* (3.26)*). The data-based distributionally robust stochastic optimal water flow* (3.26) *at t^{th} time interval is solved via successive linearization algorithm discussed in Section 3.2. Since all coefficients and affine constraints $\{A, B, E, F, \mathbf{a}_{ik}^{\beta}.\mathbf{b}_{ik}^{\beta}\}$ are derived from Laplacian-based network format* (3.11)*, at each successive linearization iteration for certain time interval, we repeatedly obtain $\{A, B, E, F, \mathbf{a}_{ik}^{\beta}.\mathbf{b}_{ik}^{\beta}\}$ in problem* (3.26) *based on the linearized updated $\{\mathbf{L}, \bar{\mathbf{Q}}\}$ until the linearized errors converge, and move to the next time period. A flowchart of proposed data-based distributionally robust stochastic OWF is demonstrated in Fig. 3.1.*

3.5 Case Studies

We now demonstrate the effectiveness of the proposed framework with numerical experiments. We use a networks model derived from a portion of the Barcelona drinking water network [117]. There are 2 reservoirs, 4 water demands, 3 tanks, 2 pumps, 4 valves, and 20 junctions, the physical properties of nodes and links are given in Tables 3.3 and 3.4, respectively. The nominal water demand pattern over 24 hours shown in Fig. 3.3(a) is derived from

EPANET (a standard software package for analysis of drinking water distribution systems) [228]. Four demands are located at nodes 8, 15, 16, and 17. Realization of demand forecasting errors are generated by evaluating the so-called persistence forecast on the EPANET demand data, which predicts the water demand at the next time step to be equal to that at the previous time step. The time-of-use (TOU) electricity price is given in Fig. 3.3(b).

We placed three tanks at Node 23, Node 24 and Node 25 to accommodate the water demand uncertainties associated with the downstream nodes. The lower and upper tank level in feet are restricted to $h_i^{\min} = 525$ and $h_i^{\max} = 530, \forall i \in \mathcal{T}$. Due to the inherent variability of water demands, tank level constraint violations may occur. Given the forecasting error data of water demand, the numerical tests are focused on reducing potential constraint violation via proposed distributionally robust framework (3.26), and minimizing the operational cost under certain risk aversion as well. To have a clear and straightforward presentation, only the lower level constraints of three tanks are modelled in distributionally robust fashion (3.26e)–(3.26g). Other constraints are handled via sample average approximation (SAA) [171, 36] or deterministic approach, though it is easy reformulate other constraints with distributional robustness.

No bound is enforced on water demand forecast errors, which implies the parameters (i.e., $\mathbf{z}$ and $\mathbf{F}$) of polytope supported set in (3.26e)–(3.26f) are set to zero. The variation of forecast errors increases with the prediction horizon. The number of forecast error samples in the training data set $\hat{\Xi}_t^N$ is $N_s = 100$. The simulation takes 60 seconds or less to solve DRO OWF with finite horizon $\mathcal{H}_t = 4$ (hours) using MOSEK Solver [11] via the MATLAB interface with CVX [135] on a laptop with 16GM of memory and a 2.8GHz Intel Core i7.

Fig. 3.4 visualizes the fundamental tradeoff between the conservativeness of constraint violation and the water network operational costs during 24 hour operation under various risk aversion ρ and Wasserstein radius ϵ. As we increase the Wasserstein radius ϵ, the pump cost will increase as well, but leads to more conservative pumps schedule and lower risk of

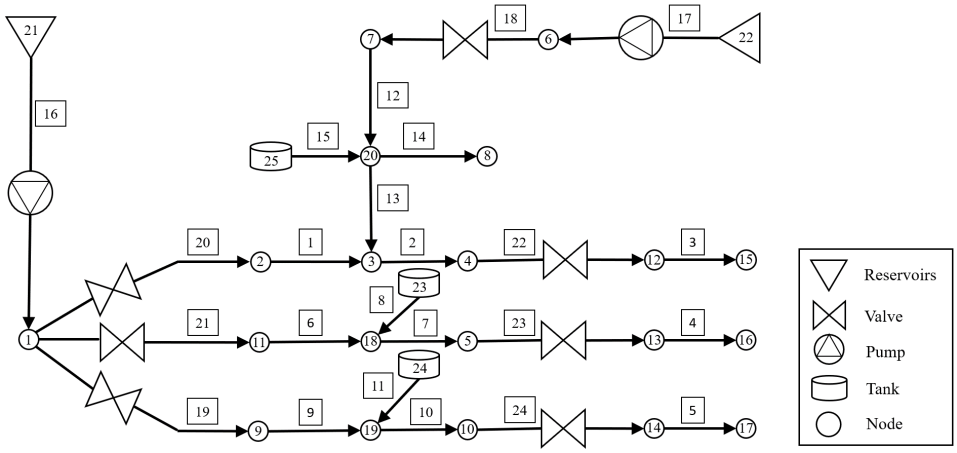


Figure 3.2. Barcelona drinking water network includes 25-node, 3 tanks and 2 reservoirs.

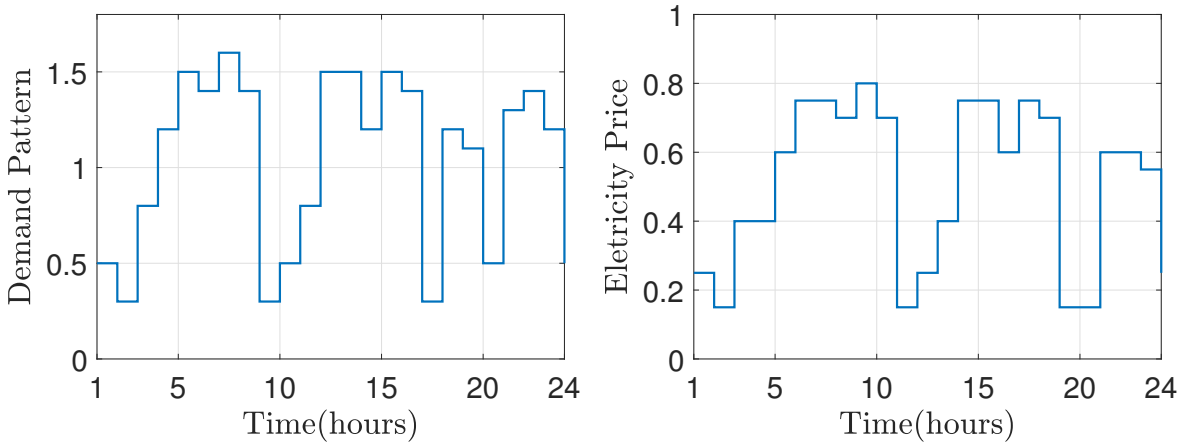
Table 3.3. Node setting of the barcelona water distribution network

Node	Type	Base Demand (GPM)	Node	Type	Base Demand (GPM)
1	Junction	0	2	Junction	0
3	Junction	0	4	Junction	0
5	Junction	0	6	Junction	0
7	Junction	0	8	Junction	100
9	Junction	0	10	Junction	0
11	Junction	0	12	Junction	0
13	Junction	0	14	Junction	0
15	Junction	100	16	Junction	100
17	Junction	100	18	Junction	0
19	Junction	0	20	Junction	0
21	Reservoir	0	22	Reservoir	0
23	Tank	0	24	Tank	0
25	Tank	0			

tank constraint violation. A larger ϵ results in less constraint violation based on the inherent sampling errors, and in turn guarantee the a stronger robustness performance, which will ensure good out-of-sample performance. In addition, with increasing risk aversion ρ, the CVaR of constraint violation is emphasized, which comes to a higher operational costs and lower constraint violation.

Table 3.4. Link setting of the Barcelona water distribution network (Chezy-Manning)

Link	Type	From Node	To Node	Pipe Length (feet)	Pipe Diameter (feet)	Pipe Rough -ness	Link	Type	From Node	To Node	Pipe Length (feet)	Pipe Diamter (feet)	Pipe Rough -ness
1	Pipe	2	3	2000	12	0.03	2	Pipe	3	4	1000	12	0.03
3	Pipe	12	15	3000	12	0.03	4	Pipe	13	16	4000	12	0.03
5	Pipe	14	17	5000	12	0.03	6	Pipe	11	18	1000	12	0.03
7	Pipe	18	5	1000	12	0.03	8	Pipe	23	18	1000	12	0.03
9	Pipe	9	19	1000	12	0.03	10	Pipe	19	10	1000	12	0.03
11	Pipe	24	19	1000	12	0.03	12	Pipe	7	20	1000	12	0.03
13	Pipe	20	3	1000	12	0.03	14	Pipe	20	8	3000	12	0.03
15	Pipe	25	20	1000	6	0.03	16	Pump	21	1	-	-	-
17	Pump	22	6	-	-	-	18	PRV	6	7	-	-	-
19	PRV	1	9	-	-	-	20	PRV	1	2	-	-	-
21	PRV	1	11	-	-	-	22	Pipe	4	12	3000	12	0.03
23	Pipe	5	13	3000	12	0.03	24	Pipe	10	14	3000	12	0.03



(a) Water Demand Pattern (b) Time-of-Use Price

Figure 3.3. Time-varying input profiles including water demand pattern and TOU electricity price. The actual water consumption at each node depends on the based demand setting shown in Table 3.3. The pump cost function is parameterized in proportional to TOU electricity price.

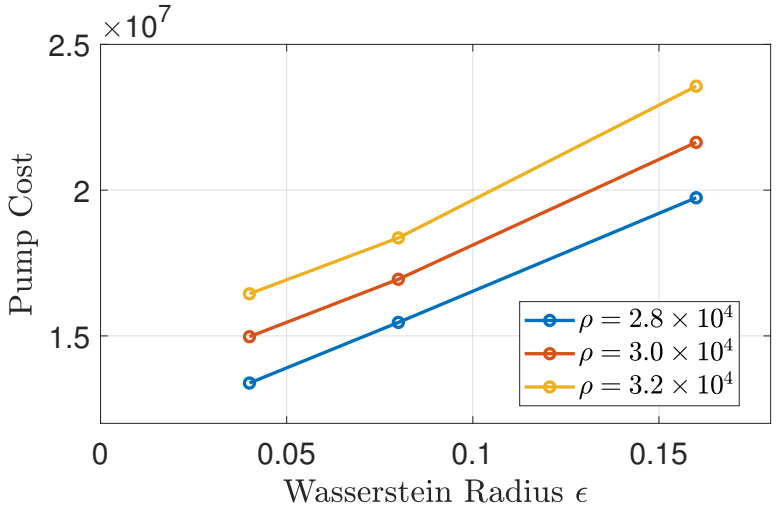


Figure 3.4. Tradeoffs between conservativeness of optimal decisions and pump operational costs under various Wasserstein radius ϵ and risk aversion ρ.

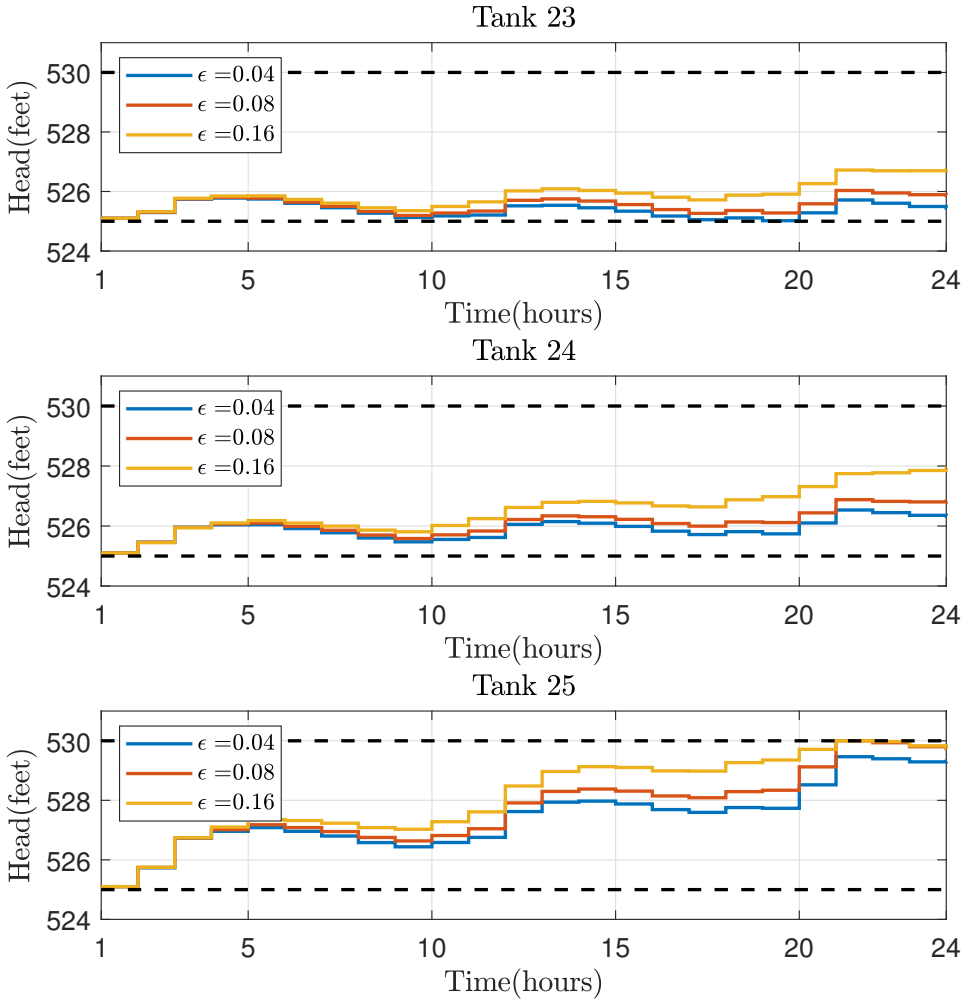


Figure 3.5. Optimal state trajectories of three tanks (i.e., Node 23, Node 24, and Node 25) for varying Wasserstein radii ϵ. The dash lines indicate the upper and lower bounds on tank head. The initial tank level for all three tanks is 525.1 feet. The risk aversion is set to $\rho = 2.8 \times 10^4$.

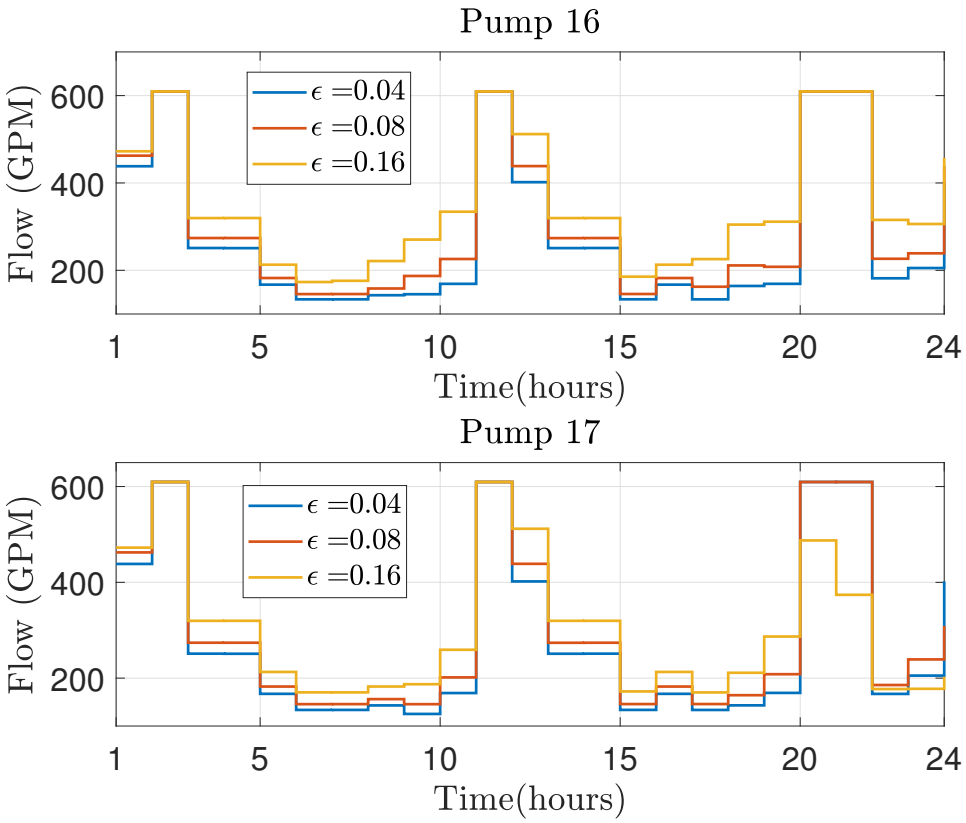


Figure 3.6. Comparison of optimal pump schedule for various value of Wasserstein radius $\epsilon = 0.04, 0.08, 0.16$ under certain risk aversion $\rho = 2.8 \times 10^4$.

Fig. 3.5 and Fig. 3.6 show the water level of tanks $\mathbf{h}_T^{\mathrm{TK}}$ and the optimal nominal pump schedule e_T over $T-24$ hours under varying Wasserstein radius ϵ. The tank head trajectories and pump schedule are re-optimized at each timestep via the closed-loop MPC controller based on the data-based uncertainty representation (i.e., a Wasserstein ball of distributions of water demand forecast errors). To prevent the tank level decrease lower then 525 feet, the pumps need to transport more water to tanks for accommodating the water demand uncertainties. As the results, the pump are more active during the time-slots with higher electricity costs, which leads a significant increase of operational costs. This leads to a safer tank level profiles, as shown in Fig. 3.5. When ϵ is small, the water consumption mostly come from tanks to maintain an economic operation, which cause the constraint violation (e.g., Tank 23 when $\epsilon = 0.04$). As we increase ϵ leading to a more conservative decision, all

pumps sacrifice the operational efficiency and provide more water to increase the tank level and support the water demands. The tank lower level constraints are satisfied due to the better robustness to water demand forecast errors.

To demonstrate the effectiveness of the proposed framework (3.26), we also introduce the EPANET built-in traditional Rule-based Control (RBC) scheme, which has been widely employed for various water engineering problems. The RBC scheme shares the same control constraints in (3.26), limits the water heads of three tanks (i.e., Tanks 23, 24 and 25) within a prescribed safe range (i.e., [525, 530] ft) via binary ON or OFF status of pumps (i.e., Pumps 16 and 17). The time step to control pumps is set to one hour, which implies the pumps check the water levels of tanks every hour and then perform control actions. Note that Tanks 23 and 24 can only be controlled by Pump 16 while Tank 25 can be managed by Pumps 16 and 17 simultaneously. Note that the water levels of Tank 23 and Tank 25 possibly direct Pump 16 to the completely conflict control actions (i.e., ON or OFF) if we do not explicitly prioritize these two tanks. Therefore, we assign the level signal from Tank 23 is the priority for Pump 16 to take control actions if a conflict happens.

Fig. 3.7 illustrates the water levels of three tanks based on the RBC scheme via Monte Carlo simulations. We randomly generate 100 scenarios of water demand forecast errors, which follow the Gaussian distribution with zero mean and 20% standard deviation of nominal water demand shown in Fig. 3.3(a). It readily seen that the RBC mechanism fails to realize the water demand forecast errors, and can not successfully manage the tank heads located at a prescribed safe bound. In general, it is very hard to parameterize the RBC control scheme for low risk constraint violation guarantee under the large demand variation, which is due to its decentralized control structure. The benefit of closed-loop multi-period distributionally robust optimal water flow based on model predictive control scheme can be clearly seen via the comparison to the RBC control framework.

In summary, we conclude that our proposed data-based distributionally robust OWF framework can explicitly incorporate water demand uncertainties, and successfully control

the tradeoff between operational efficiency, risk of constraint violation and out-of-sample performance.

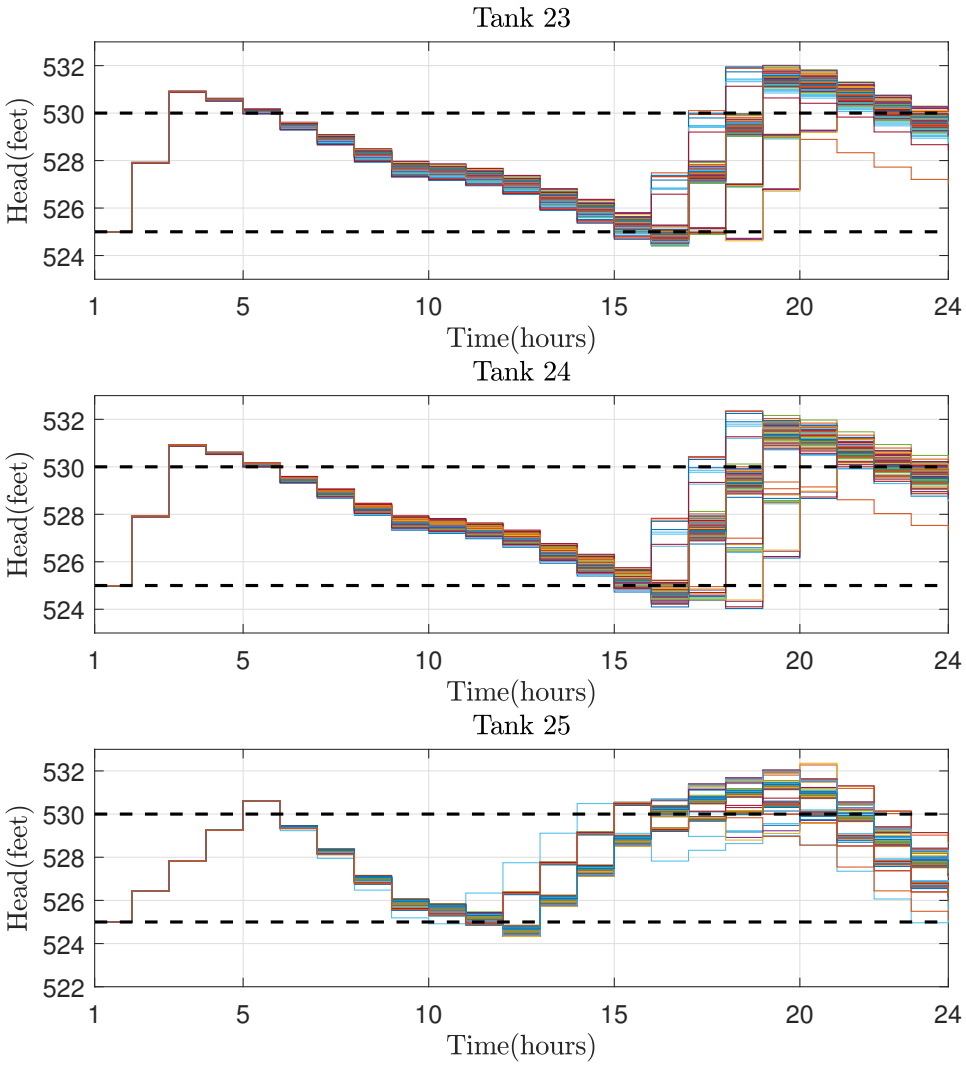


Figure 3.7. State trajectories of tank head (i.e., Tank 23, Tank 24, and Tank 25) after performing RBC via Monte Carlo simulation including 100 demand scenarios. The dash lines indicate the upper and lower bounds on tank head.

3.6 Conclusion and Outlook

We propose a data-based distributionally stochastic robust optimal water flow based on limited information from water demand forecasts. The framework creates and then leverages

a successive linearization of hydraulic coupling for an efficient computation of multi-period feedback control policies, which are robust to inherent sampling errors in the training dataset. We explore the tractability of proposed multi-period OWF problem via the Wasserstein-based distributional information of ambiguity set centered at the empirical distribution. The effectiveness and flexibility has been demonstrated on a 25-node water distribution network for the optimal water pump schedule and tank head management under water demand uncertainties. The numerical results indicate that our proposed framework has superior out-of-sample performance then existing control frameworks, and allows flexible parameterization to systematically exploit the operating strategies of water pumps to explicitly tradeoff the operational efficiency and constraint violations due to large water demand variations.

Future work includes

- exploring the flexibility and feasibility of the proposed data-based multi-period optimal water flow under more general network topologies;
- developing a data-based distributionally robust control framework for optimal water contamination control;
- including operational status of actuators as controllable variables for a distributionally robust stochastic hybrid MPC OWF framework.

Appendix

Here we provide the constant matrices $\{A, B, E, F\}$ shown in (3.16) in terms of the network coefficients in the Laplacian-based linearization (3.11). Recall the tank level dynamics (3.16a)

$$x(t+1) = \bar{A}x(t) + \bar{B}_u u(t) + \bar{B}_v v(t),$$

where $\bar{A}$ is the n_{TK}-dimension identity matrix and the last two terms on the right side are re-organized in terms of the network incident matrix $\mathbf{B}_f^\intercal$, the coefficient of tanks A_i^{TK} and

time interval Δt

$$\begin{bmatrix} \bar{B}_u & \bar{B}_v \end{bmatrix} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\Delta t}{A_1^{\mathrm{TK}}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\Delta t}{A_{n_{\mathrm{TK}}}^{\mathrm{TK}}} \end{bmatrix} [\mathbf{B}_f^{\intercal}]_{\{\mathcal{T}\}}}_{\begin{bmatrix} \bar{B}_u & \bar{B}_v \end{bmatrix}} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}.$$

The operator $[\cdot]_{\{\mathcal{T}\}}$ selects the i^{th}-row of matrix $\mathbf{B}_f^{\intercal}$, where all node i are collected in the set $\mathcal{T}$. Similarly, for mass balance $d(t) = \bar{E}_u u(t) + \bar{E}_v v(t)$ we have

$$\begin{bmatrix} \bar{E}_u & \bar{E}_v \end{bmatrix} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \mathbf{B}_f^{\intercal} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}.$$

Finally, recall the mass conservation in DAE model (3.16c)

$$\bar{F}_x x(t) + \bar{F}_l l(t) = \bar{F}_u u(t) + \bar{F}_v v^{\mathrm{P}}(t) + \bar{F}_\phi \phi(t) + F_0,$$

and re-write the above equation as follows

$$\underbrace{\begin{bmatrix} \bar{F}_x & \bar{F}_l \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} x(t) \\ l(t) \end{bmatrix}}_{\mathbf{h}(t)} = \underbrace{\begin{bmatrix} \bar{F}_u & \bar{F}_v & \bar{F}_\phi \end{bmatrix}}_{\mathbf{B}_f^{\intercal}} \underbrace{\begin{bmatrix} u(t) \\ v^{\mathrm{P}}(t) \\ \phi(t) \end{bmatrix}}_{\Delta\mathbf{q}(t)} + \underbrace{F_0}_{\bar{\mathbf{Q}}(t)}.$$

CHAPTER 4

OPTIMAL POWER FLOW WITH STATE ESTIMATION IN-THE-LOOP FOR DISTRIBUTION NETWORKS[1]

In this chapter, we propose a framework for integrating optimal power flow (OPF) with state estimation (SE) in the loop for distribution networks. Our approach combines a primal-dual gradient-based OPF solver with a SE feedback loop based on a limited set of sensors for system monitoring, instead of assuming exact knowledge of all states. The estimation algorithm reduces uncertainty on unmeasured grid states based on a few appropriate online state measurements and noisy "pseudo-measurements". We analyze the convergence of the proposed algorithm and quantify the statistical estimation errors based on a weighted least squares (WLS) estimator. The numerical results on a 4521-node network demonstrate that this approach can scale to extremely large networks and provide robustness to both large pseudo-measurement variability and inherent sensor measurement noise.

4.1 Introduction

The increasing penetration of distributed energy resources (DERs) has provided more flexibility to better explore the benefits of advanced smart grid technologies in distribution networks. As the heterogeneous control strategies of grid-connected elements dominate distribution networks, many of the customers will become active and motivated end-users to

[1]This chapter is based on work supported by funding from US Department (DOE) of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office under contract No. DE-EE-0007998 and the National Science Foundation (NSF) under grant CMMI-1728605.

Chapter 4 in part is a reprint of material published or under review in:

optimize their own power usage via optimal power flow (OPF) methods [54, 85, 9, 18, 175]. This requires the power system control scheme to have real-time knowledge about the structure and state of the distribution network (e.g., operation states, netload variation, device dynamics, network topology, etc.), and to provide the corresponding real-time responses (e.g., optimal control inputs. set-points of DERs, etc.) for safe and efficient operation. However, the current distribution network control paradigm cannot satisfy the above requirement due to an under-developed information feedback mechanism, and high expense of real-time system states measurement. Future distribution systems will require more sophisticated and tightly integrated control, optimization, and estimation methods for these issues.

Most OPF methods for distribution networks in the literature assume complete availability of network states to implement various optimal control strategies [191, 50, 51, 182, 32, 247, 59, 209, 202, 192, 280, 3]. However, in practice network states must be estimated with a monitoring system from noisy measurements, which itself is a challenging problem due to the increasingly complex, extremely large-scale, and nonlinear time-varying nature of emerging networks. To solve these issues, the recently proposed OPF frameworks [66, 304, 74, 166] leverage measurement feedback-based online optimization method to loop the physical measurement information back to OPF controllers, which adapt the OPF decisions to real-time data to mitigate the effects of inherent disturbances and modelling errors. It is unrealistic to have real-time physical measurements of system states at every corner of distribution networks due to heavy communication burdens, end-user privacy concerns, and high costs.

In this chapter, we propose a more general framework than the existing OPF approaches, which tightly integrates state estimation (SE) techniques [172, 80, 4, 79, 236] into online OPF control algorithms for distribution networks. This OPF with SE in the loop framework allows us to utilize a limited set of sensor measurements together with a power system state estimator instead of exact knowledge of network states. The power system state estimator, which may include data from the Supervisory Control and Data Acquisition (SCADA)

system, phasor measurement units (PMUs), topology processor and pseudo-measurements[2], provides the best available information about network states [214, 5, 80, 79, 172, 293] and in-turn enables implementation and enhances the performance of OPF controllers. Our approach allows OPF decisions to adapt in real-time time-varying stochastic DERs and loads, and compensates for disturbances and modelling errors, since SE results utilize measurement data from the actual nonlinear system response. The recent paper [210] studied the interaction between the dynamic state estimation (i.e., Kalman Filter) and the feedback-based optimization scheme for voltage regulation in distribution networks. In our work, we mostly focus on the gradient-based feedback optimization with static SE in the loop for large-scale distribution networks

A preliminary version of this work appeared in [126], and here we significantly expand the work in several directions. Firstly, we present a more comprehensive discussion on the concept of OPF with SE in the loop. Then, we provide a statistical bound to show the convergence, which explicitly includes the variance of SE in the loop for analysis. This work also builds all numerical studies on an extremely large-scale 11,000-node distribution network to demonstrate the flexibility. Finally, we present the numerical analysis of SE in the loop to further promote its excellent robustness to measurement noises on OPF solvers. The main contributions are as follows:

1) We formulate a general convex OPF problem subject to power flow equations and network-wise coupling constraints. To integrate OPF with SE in the loop, we propose a primal-dual gradient-based OPF algorithm with state estimation feedback. Instead of requiring full knowledge of all system states, the controller utilizes at every gradient step real-time monitoring information from state estimation results to inform control

[2]Due to the lack of real time measurement and stochasticity nature of power netloads in distribution system state estimation, the nodal power injections are measured by their nominal load-pattern (i.e., the real value plus zero-mean random deviations), so-called pseudo-measurement, whose information is derived from the past records of load behaviors [236].

decisions. Whereas OPF and SE problems for distribution networks have been widely studied individually, none of the existing literature explores the connection and bridge the gap between them. Here we are closing the loop between OPF and state estimation in large-scale distribution networks, which guarantees full availability of state estimates [112, 280]. This allows us to react to real-time information of system states with limited number of deployed sensors. In principle, the framework allows a variety of state estimation methodologies and control strategies in distribution networks. Here, we illustrate the approach through a voltage regulation problem with voltage magnitude estimation in the loop.

2) We leverage linear approximations of the AC power flow equations, to facilitate scalable and computationally efficient OPF problems for SE feedback integration [74, 66, 304]. The voltage profile estimation uses a weighted least squares (WLS) estimator. Convergence of the proposed gradient-based algorithm with state estimation feedback is analytically established. Additionally, we quantify the statistical estimation errors of the WLS estimator. This provides a measure of quality of the SE feedback associated with a particular allocation of sensors across the network.

3) The effectiveness, scalability, flexibility and robustness of the proposed algorithm are demonstrated on a 4521-node multi-phase unbalanced distribution network with 1043 (aggregated) netloads. With only 3.6% voltage measurement deployment, the integrated OPF controller with SE feedback effectively regulates network voltage. The distributed algorithm in [306] using linearized distribution flow (LinDistFlow) enables scaling to extremely large networks. The numerical results also indicate that the proposed OPF controller with SE feedback has excellent performance and robustness to the inherent measurement noise and estimation errors.

The rest of this chapter is organized as follows. Section 4.2 discusses the general concept of OPF with SE in the loop for distribution networks. Section 4.3 formulates an OPF problem and introduces gradient algorithm with state estimation feedback. Section 4.4 demonstrates numerical results on voltage regulations and Section 4.5 concludes.

4.2 Optimal Power Flow with State Estimation in the loop

In this section, we propose an OPF solver with state estimation feedback. We first pose a general problem to highlight the overall approach, and in subsequent sections we detail the model, objectives, constraints and state estimator for a certain control and monitoring purpose.

Consider the OPF problem for distribution networks

$$\min_{\mathbf{p},\mathbf{q}} \quad \sum_{i\in\mathcal{N}} C_i(p_i, q_i) + C_0(\mathbf{p}, \mathbf{q}), \tag{4.1a}$$

$$\text{s.t.} \quad \mathbf{g}(\mathbf{r}(\mathbf{p}, \mathbf{q})) \leq 0, \tag{4.1b}$$

$$(p_i, q_i) \in \mathcal{Z}_i, \forall i \in \mathcal{N}, \tag{4.1c}$$

where $C_0(\mathbf{p}, \mathbf{q})$ is a cost function capturing system objectives (e.g., cost of deviation of total power injections into the substation from nominal values), the local objective function $C_i(p_i, q_i)$ captures the generation costs, ramping costs, active power losses, renewable curtailment penalty, auxiliary service expenses and reactive compensation (comprising a weighted sum thereof) at node $i \in \mathcal{N}$. We then define a state vector $\mathbf{r}(\mathbf{p}, \mathbf{q}) \in \mathbf{R}^M$ denoting (combined) electrical quantities of interests (e.g., voltage magnitude, current injections, power injection at the substation, etc.), which depends on nodal power injections $\mathbf{p} := [p_1, \ldots, p_N]^\intercal$ and $\mathbf{q} := [q_1, \ldots, q_N]^\intercal$ through power flow equations[3]. The constraint function $g : \mathbf{R}^M \to \mathbf{R}^{N_g}$

[3]The power flow relations may represent either full nonlinear power flow, SOCP relaxations, SDP relaxations, or various linearization. The general approach of OPF with SE in the loop can be adapted to various (approximate) power flow mappings. In the rest of this chapter, we use a linearized power flow model to illustrate the effectiveness of the proposed framework.

models network constraints, including voltage magnitude, voltage angle, current injection and line flows. The nodal power injections are constrained to convex and compact feasible sets $\mathcal{Z}_i$.

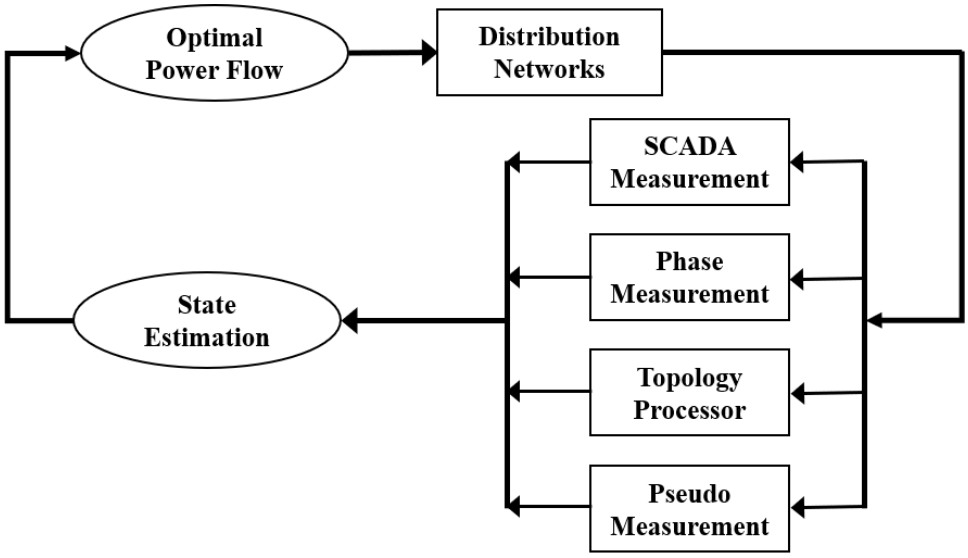


Figure 4.1. The concept of solving optimal power flow with state estimation in the loop.

Problem (4.1) is typically solved assuming that all system states are available. However, in practice, there is generally a lack of reliable measurement devices and communication infrastructure in distribution networks, rendering these conventional OPF approaches impractical. Therefore, the main challenges for solving (4.1) in practice involve how to best integrate *estimates* of current system states $\mathbf{r}$ and understanding tradeoffs between SE performance and sensor deployment quality and OPF controller performance. We will tackle these by integrating a state estimation feedback loop with a limited number of sensor measurements. This allows the OPF controller to respond to real-time information and update control decisions despite nodes without measurement in the grid. The overall approach is illustrated in Fig. 4.1.

In our general framework, the SE techniques may determine the system states using any or all of SCADA measurements, phasor measurement units (PMUs) measurements,

pseudo-measurements and topology information to reduce estimation uncertainty. How to fuse different sources of information into OPF formulations remains largely an open question under exploration. We aim to indicate that there are many possibilities and research direction to potentially improve control and optimization in distribution networks through tight integration with SE. The increasing penetration of renewable energy resources and distributed generators enable the distribution networks with smart features, such as demand response and distributed automation. This allows the networks turn to a more active and complex system with fast system response. An efficient, real-time monitoring of distribution networks should be looped into OPF controllers. In the rest of this chapter, we take the voltage regulation problem with voltage estimation as an illustrative example, which is based on a few PMU voltage measurements and netload pseudo-measurements.

4.3 Gradient-Based OPF Solver with State Estimation Feedback

4.3.1 System Modelling

Consider a distribution network denoted by a directed and connected graph $\mathcal{G}(\mathcal{N}_0, \mathcal{E})$, where $\mathcal{N}_0 := \mathcal{N} \cup \{0\}$ is a set of all "buses" or "nodes" with substation node 0 and $\mathcal{N} := \{1, \ldots, N\}$, and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is a set of "links" or "lines" for all $(i, j) \in \mathcal{E}$. Let $V_i := |V_i| e^{j \angle V_i} \in \mathbf{C}$ and $I_i := |I_i| e^{j \angle I_i} \in \mathbf{C}$ denote the phasor for the line-to-ground voltage and the current injection at node $i \in \mathcal{N}$. The absolute values $|V_i|$ and $|I_i|$ denote the signal root-mean-square values and $\angle V_i$ and $\angle I_i$ corresponding to the phase angles with respect to the global reference. We collect these variables into complex vectors $\mathbf{v} := [V_1, V_2, \ldots, V_N]^\intercal \in \mathbf{C}^N$ and $\mathbf{i} := [I_1, I_2, \ldots, I_N]^\intercal \in \mathbf{C}^N$. We denote the complex admittance of line $(i, j) \in \mathcal{E}$ by $y_{ij} \in \mathbf{C}$. The admittance matrix $\mathbf{Y} \in \mathbf{C}^{N \times N}$ is given by

$$Y_{ij} = \begin{cases} \sum_{l \sim i} y_{il} + y_{ii} & \text{if } i = j \\ -y_{ij} & (i, j) \in \mathcal{E} \\ 0 & (i, j) \notin \mathcal{E} \end{cases}, \tag{4.2}$$

where $l \sim i$ indicates connection between node l and node i, and y_{ii} is the self admittance of node i to the ground.

Node 0 is modelled as a slack bus. The other nodes are modelled as PQ buses for which the injected complex power is specified. The admittance matrix can be partitioned as

$$\begin{bmatrix} I_0 \\ \mathbf{i} \end{bmatrix} = \begin{bmatrix} y_{00} & \bar{y}^\intercal \\ \bar{y} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} V_0 \\ \mathbf{v} \end{bmatrix}.$$

The net complex power injection then reads:

$$\mathbf{s} = \mathrm{diag}(\mathbf{v})\Big(\mathbf{Y}^*(\mathbf{v})^* + \bar{y}^*(V_0)^*\Big), \tag{4.3}$$

where superscript $(\cdot)^*$ indicates the element-wise conjugate of complex vector $\mathbf{v}$.

To facilitate computational efficiency using convex optimization, here we leverage a linearization of (4.3) as follows:

$$\mathbf{r} = \mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{q} + \mathbf{r}_0, \tag{4.4}$$

where the parameters $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{r}_0$ can be attained from various linearization methods, e.g., [30, 42, 102, 21]. In the rest of this chapter, the linearized coefficient matrices $\mathbf{A}$ and $\mathbf{B}$ are fixed over times for simplification. From now on, we limit $\mathbf{r}(\mathbf{p}, \mathbf{q})$ to the above linearized power power (4.4). Recall that $\mathbf{r} \in \mathbf{R}^M$ represents certain electrical quantities of interests (e.g., voltage magnitude, current injections, power injection at the substation, etc.).

4.3.2 OPF Formulation and Primal-Dual Gradient Algorithm

In this section, we introduce a general OPF problem and the pertinent gradient algorithm with idealized measurement feedback[4] from nonlinear power flow to reduce modelling errors. The feasible operating regions $\mathcal{Z}_i$ depend on the terminal properties of various dispatchable devices, e.g., inverter-based distributed generators, energy storage systems or small-scale diesel generators. We make the following assumptions.

[4]The "idealized" refers to the full measurement of vector $\mathbf{r}$ without noise.

Assumption 3 (Slater's condition)**.** *There exists a strictly feasible point within the operation region* $(\mathbf{p}, \mathbf{q}) \in \mathcal{Z}$*, where* $\mathcal{Z} := \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_N$*, so that*

$$\mathbf{g}(\mathbf{r}(\mathbf{p}, \mathbf{q})) < 0.$$

Assumption 3 suffices strong duality for problem (4.1).

Assumption 4. *A set of local objective functions* $C_i(p_i, q_i), \forall i \in \mathcal{N}$ *are continuously differentiable and strongly convex as functions of* (p_i, q_i)*, and their first order derivative are bounded within their operation regions indicated as* $(p_i, q_i) \in \mathcal{Z}_i, \forall i \in \mathcal{N}$*; The system-wise objective function* $C_0(\mathbf{p}, \mathbf{q})$ *is continuously differentiable and convex with its first-order derivative bounded. Furthermore, the constraint function* g *is continuously differentiable and convex with bounded derivatives on its domain.*

The regularized Lagrangian $\mathcal{L}$ for (4.1) is

$$\mathcal{L} = \sum_{i \in \mathcal{N}} C_i(p_i, q_i) + C_0(\mathbf{p}, \mathbf{q}) + \boldsymbol{\mu}^\intercal \mathbf{g}(\mathbf{r}(\mathbf{p}, \mathbf{q})) - \frac{\eta}{2} \|\boldsymbol{\mu}\|_2^2, \tag{4.5}$$

where $\boldsymbol{\mu} \in \mathbf{R}_+^{N_\mu}$ is the dual variable vector associated with the general inequality constraints and we keep the feasible regions $\boldsymbol{\mu} \in \mathbf{R}_+^{N_\mu}$ and $(\mathbf{p}, \mathbf{q}) \in \mathcal{Z}$ implicit. To facilitate proof of convergence, the Lagrangian (4.5) includes a Tikhonov regularization term $-\frac{\eta}{2}\|\boldsymbol{\mu}\|_2^2$ with a prescribed small parameter η that introduces bounded discrepancy. The upshot of having a regularized term in Lagrangian (4.5) is that the gradient-based approaches can be applied to (4.5) to find an approximate solution of the original Lagrangian with improved convergence properties. The discrepancy due to the regularized term discussed and quantified in [148].

Then to solve (4.1) we come to the saddle-point problem

$$\max_{\boldsymbol{\mu} \in \mathbf{R}_+^{N_\mu}} \min_{(\mathbf{p}, \mathbf{q}) \in \mathcal{Z}} \mathcal{L}\left(\mathbf{p}, \mathbf{q}, \boldsymbol{\mu}\right), \tag{4.6}$$

which leads to an iterative primal-dual gradient algorithm to reach the unique saddle-point of (4.6)

$$\mathbf{r}^k = \mathbf{A}\mathbf{p}^k + \mathbf{B}\mathbf{q}^k + \mathbf{r}_0, \tag{4.7a}$$

$$\mathbf{p}^{k+1} = \left[\mathbf{p}^k - \epsilon\nabla_{\mathbf{p}}\mathcal{L}|_{\mathbf{p}^k,\mathbf{q}^k,\boldsymbol{\mu}^k}\right]_{\mathcal{Z}}, \tag{4.7b}$$

$$\mathbf{q}^{k+1} = \left[\mathbf{q}^k - \epsilon\nabla_{\mathbf{q}}\mathcal{L}|_{\mathbf{p}^k,\mathbf{q}^k,\boldsymbol{\mu}^k}\right]_{\mathcal{Z}}, \tag{4.7c}$$

$$\boldsymbol{\mu}^{k+1} = \left[\boldsymbol{\mu}^k + \epsilon\nabla_{\boldsymbol{\mu}}\mathcal{L}|_{\mathbf{r}^k,\boldsymbol{\mu}^k}\right]_{\mathbf{R}_+^{N_\mu}}, \tag{4.7d}$$

where $\epsilon > 0$ is a constant stepsize to be determined and the operator $[\cdot]_{\mathcal{Z}}$ project onto the feasible set $\mathcal{Z} = \times_{i\in\mathcal{N}}\mathcal{Z}_i$ [5]. The operator $[\cdot]_{\mathbf{R}_+^{N_\mu}}$ projects onto nonnegative orthant.

The updates (4.7) are represented compactly by the mapping

$$\Phi : \{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k\} \mapsto \begin{bmatrix} \nabla_{\mathbf{p}}\mathcal{L}|_{\mathbf{p}^k,\mathbf{q}^k,\boldsymbol{\mu}^k} \\ \nabla_{\mathbf{q}}\mathcal{L}|_{\mathbf{p}^k,\mathbf{q}^k,\boldsymbol{\mu}^k} \\ -\nabla_{\boldsymbol{\mu}}\mathcal{L}|_{\mathbf{r}^k(\mathbf{p}^k,\mathbf{q}^k),\boldsymbol{\mu}^k} \end{bmatrix},$$

so that (4.7) can be written as

$$\mathbf{x}^{k+1} = \left[\mathbf{x}^k - \epsilon\Phi(\mathbf{x}^k)\right]_{\mathbf{R}_+^{N_\mu}\times\mathcal{Z}}, \tag{4.8}$$

where $\mathbf{x}^k := [(\mathbf{p}^k)^\intercal, (\mathbf{q}^k)^\intercal, (\boldsymbol{\mu}^k)^\intercal]^\intercal$. Under Assumption 4, it can be shown [306] that Φ is strongly monotone and Lipschitz continuous, i.e., it satisfies for all feasible points $\mathbf{x}_1$ and $\mathbf{x}_2$ and for some constants $M > 0$ and $L > 0$

$$(\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2))^\intercal (\mathbf{x}_1 - \mathbf{x}_2) \geq M\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2, \tag{4.9}$$

$$\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)\|_2^2 \leq L^2\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2. \tag{4.10}$$

We now have the following result

[5] We use the projection operator $[\cdot]_{\mathcal{Z}}$ in primal updates instead of $[\cdot]_{\mathcal{Z}_p}$ and $[\cdot]_{\mathcal{Z}_q}$, since the feasible sets of active and reactive power are not independent, but correlated based on the terminal apparent power limits.

Theorem 1. *Consider the primal-dual gradient algorithm* (4.7) *for the optimization problem* (4.1) *based on the regularized Lagrangian* (4.5). *If the step size* ϵ *satisfies*

$$0 < \epsilon \leq \bar{\epsilon} < 2M/L^2, \tag{4.11}$$

for some $\bar{\epsilon}$*, algorithm* (4.7) *converges to the unique saddle point of* (4.5).

The optimization problem (4.1) and gradient algorithm (4.7) are based on a linearized power flow to guarantee its convexity and prove convergence to the saddle point. However, linearization errors cause the solution of (4.7) to be suboptimal or even infeasible for the system with nonlinear power flow. To address this issue, feedback-based online optimization methods [66, 31, 304] have been leveraged to reduce the effects of modelling error. In particular, by replacing (4.7a) with following nonlinear power flow

$$\mathbf{r}^k \;\; = \;\; f(\mathbf{p}^k, \mathbf{q}^k), \tag{4.12}$$

that obtained from the physical system, these measured values $\mathbf{r}^k$ can be used, instead of an approximate model, to update dual variables in (4.7d). Convergence to a bounded range of the optimum can be analytically shown for such implementations, and this also facilitates a real-time implementation that can track the time-varying grid conditions [74, 304].

However, one crucial issue of such feedback-based algorithm has been largely overlooked: in practice, there are too few monitoring devices in distribution systems to measure all components of $\mathbf{r}$, and therefore it is not possible to directly implement feedback-based algorithms to solve the problem (4.1). Our preliminary results [126] demonstrated that limited knowledge of system states can lead the OPF controller to cause constraint violations.

To enable an implementation of feedback-based OPF algorithms in distribution networks, and also to improve performance of algorithms that make use of "pseudo-measurements", we integrate a state estimation algorithm based on a sparse set of available measurements, before performing the dual variable update (4.7d). This allows us to utilize improved information

on the network state to make decisions, specifically improving information at nodes without measurement of the grid where there are no direct measurements. Fig. 4.2 illustrates the proposed OPF framework with state estimation in the loop.

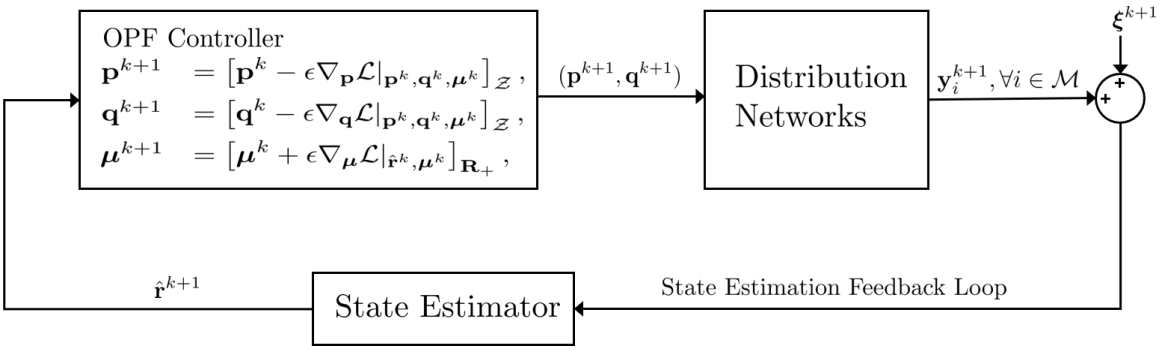


Figure 4.2. The diagram of the proposed optimal power flow problem with SE in the loop.

4.3.3 State Estimation in the loop

We consider $\mathbf{z}^k = [(\mathbf{p}^k)^\intercal, (\mathbf{q}^k)^\intercal]^\intercal$ the system states at iteration k, and the grid measurement model

$$\mathbf{y}^k = \mathbf{h}(\mathbf{z}^k) + \boldsymbol{\xi}^k, \tag{4.13}$$

where $\mathbf{y}^k \in \mathbf{R}^L$ is a measurement vector received at time k comprising raw noisy measurements from sensors and pseudo-measurements (which include real and reactive power injections, real and reactive power flow, and voltage magnitude and angles), and $\mathbf{h} : \mathbf{R}^M \to \mathbf{R}^L$ is a measurement function. The vector $\boldsymbol{\xi}^k \in \mathbf{R}^L$ models measurement errors, which are assumed to be independent and identically distributed and subject to Gaussian distribution $N(0, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} \in \mathbf{R}_+^{L\times L}$ is a diagonal matrix. The pseudo-measurements are modeled as sensor measurements corrupted by high-variance Gaussian noise based on historical data (e.g., customer billing data and typical load profile) that provide rough information about variations in the state of the grid [112].

To estimate grid states from the available measurements, we consider the WLS estimator [112, 280, 305]:

$$\hat{\mathbf{z}}^k = \underset{\mathbf{z}^k}{\operatorname{argmin}} \frac{1}{2} \left(\mathbf{y}^k - \mathbf{h}(\mathbf{z}^k)\right)^\intercal \mathbf{W} \left(\mathbf{y}^k - \mathbf{h}(\mathbf{z}^k)\right), \tag{4.14}$$

where the weight matrix is defined as $\mathbf{W} = \mathbf{\Sigma}^{-1}$. The estimated quantities of interest $\hat{\mathbf{r}}^k(\hat{\mathbf{z}}^k)$ is uniquely determined by the states $\hat{\mathbf{z}}^k$ through power flow equations. Existence and uniqueness of a solution to (4.14) require certain properties of the measurement function.

Definition 2 (Full Observability[6][4, 281])**.** *The system state is called fully observable if* $\mathbf{z}^k = 0$ *is the only solution for* $h(\mathbf{z}^k) = 0$*, which allows a unique solution to* (4.14)*.*

Assumption 5. *We assume that the distribution system state with measurements* (4.13) *is fully observable.*

Since distribution networks typically have only a sparse set of real-time measurements from deployed sensors, we require enough pseudo-measurements to ensure full observability. It is always possible to include enough pseudo-measurements to ensure full observability and satisfy Assumption 5. The effectiveness combining real measurements and pseudo-measurements has been observed in [305, 91].

Fig. 4.2 and Algorithm 1 illustrate and describe the proposed OPF controller with the state-estimation feedback loop. Note that the step 2 in Algorithm 2 is not implemented in the proposed OPF controller, but instead is a realization of the power system state's physical response $\mathbf{r}^k$ resulting from gradient updates on nodal power injections $(\mathbf{p}^k, \mathbf{q}^k)$. We utilize SE in the loop to compute a state estimate $\hat{\mathbf{r}}^k$, which then contributes to the update of dual variables $\boldsymbol{\mu}^{k+1}$ in step 7. Our numerical experiments in Section 4.4 compare this approach with the direct use of noisy measurements and pseudo-measurements without an estimation scheme.

[6]This definition should be distinguished from observability of linear dynamical systems. Here, we limit the definition of observability to power system static state estimation problems [236] throughout this manuscript.

Algorithm 2 (OPF with SE in the loop)

Require: netload initialization $(\mathbf{p}^0, \mathbf{q}^0)$ and $\boldsymbol{\mu}^k$

1: **for** $k = 0 : K$ **do**
2: $\quad \mathbf{r}^k \leftarrow$ nonlinear power flow $(\mathbf{p}^k, \mathbf{q}^k)$
3: $\quad$ receive system measurement $\mathbf{y}^k$
4: $\quad$ estimate the system states $\hat{\mathbf{z}}^k$ based on $\mathbf{y}^k$ and calculate the electrical quantities of interest $\hat{\mathbf{r}}^k(\hat{\mathbf{z}}^k)$
5: $\quad \mathbf{p}^{k+1} = \left[\mathbf{p}^k - \epsilon \nabla_{\mathbf{p}} \mathcal{L}(\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k)\right]_{\mathcal{Z}}$
6: $\quad \mathbf{q}^{k+1} = \left[\mathbf{q}^k - \epsilon \nabla_{\mathbf{q}} \mathcal{L}(\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k)\right]_{\mathcal{Z}}$
7: $\quad \boldsymbol{\mu}^{k+1} = \left[\boldsymbol{\mu}^k + \epsilon \nabla_{\boldsymbol{\mu}} \mathcal{L}(\hat{\mathbf{r}}^k, \boldsymbol{\mu}^k)\right]_{\mathbf{R}_+^{N_\mu}}$
8: **end for**

4.3.4 Convergence Analysis

The computations and updates in steps 5-7 of Algorithm 1 are written more explicitly as

$$\hat{\mathbf{z}}^k = \underset{\mathbf{z}^k}{\operatorname{argmin}} \frac{1}{2} \left(\mathbf{y}^k - h(\mathbf{z}^k)\right)^\intercal \mathbf{W} \left(\mathbf{y}^k - h(\mathbf{z}^k)\right), \tag{4.15a}$$

$$\mathbf{p}^{k+1} = \Big[\mathbf{p}^k - \epsilon\big(\nabla_{\mathbf{p}} C(\mathbf{p}^k, \mathbf{q}^k) + \nabla_{\mathbf{p}} C_0(\mathbf{p}^k, \mathbf{q}^k) + \mathbf{A}^\intercal \nabla_{\mathbf{r}} \mathbf{g}(\mathbf{r})^\intercal \boldsymbol{\mu}^k\big)\Big]_{\mathcal{Z}}, \tag{4.15b}$$

$$\mathbf{q}^{k+1} = \Big[\mathbf{q}^k - \epsilon\big(\nabla_{\mathbf{q}} C(\mathbf{p}^k, \mathbf{q}^k) + \nabla_{\mathbf{q}} C_0(\mathbf{p}^k, \mathbf{q}^k) + \mathbf{B}^\intercal \nabla_{\mathbf{r}} \mathbf{g}(\mathbf{r})^\intercal \boldsymbol{\mu}^k\big)\Big]_{\mathcal{Z}}, \tag{4.15c}$$

$$\boldsymbol{\mu}^{k+1} = \left[\boldsymbol{\mu}^k + \epsilon \left(\mathbf{g}(\hat{\mathbf{r}}^k(\hat{\mathbf{z}}^k)) - \eta \boldsymbol{\mu}^k\right)\right]_{\mathbf{R}_+^{N_\mu}}, \tag{4.15d}$$

This iteration (4.15), associated with step 1 to step 7 in algorithm 1, is performed until convergence.

As the state estimation in distribution networks has been widely discussed for different applications [214], the existing literature shows that these type of methods lead to an accurate and computationally efficient approximation under nominal operating condition.

For the purpose of performance analysis, we next define $\tilde{\Phi}(\mathbf{x}^k)$ as the gradient mapping with nonlinear power flow feedback:

$$\tilde{\Phi} : \{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k\} \mapsto \begin{bmatrix} \nabla_{\mathbf{p}} \mathcal{L}|_{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k} \\ \nabla_{\mathbf{q}} \mathcal{L}|_{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k} \\ -\nabla_{\boldsymbol{\mu}} \mathcal{L}|_{\tilde{\mathbf{r}}^k(\mathbf{p}^k, \mathbf{q}^k), \boldsymbol{\mu}^k} \end{bmatrix},$$

where the linearized power flow $\mathbf{r}^k(\mathbf{p}^k, \mathbf{q}^k)$ is replaced with its nonlinear counterpart $\tilde{\mathbf{r}}^k(\mathbf{p}^k, \mathbf{q}^k) = f(\mathbf{p}^k, \mathbf{q}^k)$.

Finally, we define the gradient mapping with SE in the loop as $\overline{\Phi}(\mathbf{x}^k)$, where the dual variables are updated using the SE results $\hat{\mathbf{r}}^k$ from (4.14) to update the dual variables

$$\overline{\Phi} : \{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k\} \mapsto \begin{bmatrix} \nabla_{\mathbf{p}} \mathcal{L}|_{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k} \\ \nabla_{\mathbf{q}} \mathcal{L}|_{\mathbf{p}^k, \mathbf{q}^k, \boldsymbol{\mu}^k} \\ -\nabla_{\boldsymbol{\mu}} \mathcal{L}|_{\hat{\mathbf{r}}^k(\mathbf{p}^k, \mathbf{q}^k), \boldsymbol{\mu}^k} \end{bmatrix}.$$

Assumption 6. *There is a uniform bound on the squared error of the gradient update due to state estimation, i.e., there exists $\alpha > 0$ such that*

$$\mathbf{E}\Big[\|\Phi(\mathbf{x}^k) - \overline{\Phi}(\mathbf{x}^k)\|_2^2\Big] = \sigma_k^2 \leq \alpha, \quad \forall \mathbf{x}^k. \tag{4.16}$$

This bound always exists and is realistic due to the physical limits of network states. We pose a general bound on the SE in the loop in Assumption 4 to understand how the noisy measurements iteratively interact with the feedback optimization schemes. We define the estimation error variance from SE in the loop of the saddle point of (4.6) as:

$$\sigma_*^2 := \mathbf{E}\Big[\|\Phi(\mathbf{x}^*) - \overline{\Phi}(\mathbf{x}^*)\|_2^2\Big].$$

Assumption 7. *There is a uniform bound on the norm of the squared distance between update with SE in the loop and the update that uses the actual nonlinear power flow in* (4.12)*, i.e., there exists $\rho > 0$ such that*

$$\|\overline{\Phi}(\mathbf{x}^k) - \tilde{\Phi}(\mathbf{x}^k)\|_2^2 \leq \rho, \quad \forall \mathbf{x}^k. \tag{4.17}$$

This discrepancy also includes the difference raised by the linear estimator and real values based on the nonlinear power flow.

Theorem 2. *Suppose the step size ϵ satisfies the condition* (4.11) *from Theorem 1. Under Assumptions 6-7, the sequence $\{\mathbf{x}^k\}$ generated by Algorithm 1 satisfies*

$$\limsup_{k\to\infty} \mathbf{E}\Big[\|\mathbf{x}^{k+1} - \mathbf{x}^*\|_2^2\Big] = \frac{\rho + 3\alpha}{2M/\epsilon - L^2}, \tag{4.18}$$

where $\mathbf{x}^* = [(\mathbf{p}^*)^\intercal, (\mathbf{q}^*)^\intercal, (\boldsymbol{\mu}^*)^\intercal]^\intercal$ is saddle point of $\mathcal{L}$ in (4.6).

We refer the proof to the Appendix.

The result (4.18) from Theorem 2 provides an upper bound on the expected squared distance between the sequence $\{\mathbf{x}^k|\ \mathbf{x}^k := [(\mathbf{p}^k)^\intercal, (\mathbf{q}^k)^\intercal, (\boldsymbol{\mu}^k)^\intercal]^\intercal, k \leq K, K \to \infty\}$ generated by our proposed OPF with SE feedback algorithm (4.15) and the saddle point $\mathbf{x}^*$ of (4.6). This analytical bound indicates that our proposed approach has robust performance to estimation errors and measurement noise.

1. *Inherent Measurement Noise:* The online measurements by PMUs are typically within $1\% \sim 2\%$ of the actual values. The pseudo-measurements of active and reactive power can be regarded as a rough initialization (with up to 50% variations in comparison to actual values). These errors can be reduced through the estimation phase in (4.14), which improves decisions from the OPF controller with SE feedback (4.15a), improving robustness to measurement noise and power variability;

2. *Linearization Approximation Errors:* The OPF-phase (and some of state estimators) in the proposed algorithm utilize the linearized power flow to promote computational efficiency in gradient calculation. The discrepancy between linearized and nonlinear power flow is quantified in (4.17) by ρ, when set-points $(\mathbf{p}^k, \mathbf{q}^k)$ are realized by the nonlinear network power flow.

Intuitively, this discrepancy depends on 1) the step size, the monotonicity and Lipschitz coefficients; 2) the difference between nonlinear power flow and linearized power flow and

3) a general bound on the covariance of the in the loop SE. Overall, the estimation errors of the in-the-loop SE will not be iteratively accumulated with increasing number of OPF iterations. Moreover, the bound characterized in (4.18) can be arbitrarily small if we choose small enough ϵ.

Remark 5. *(Feedback-based OPF). For solving an optimal power flow problem in a large-scale distribution network, our methodology can effectively reduce the computational complexity by adopting the linearized relationship in the problem formulation and compensating the modeling error with the closed-loop feedback from the nonlinear model. State estimation serves to bridge the gap between theoretic design where all quantities of interest are measured and the reality where only a few nodes are equipped with measuring devices. Analytical characterization of the model-based feedback algorithm for solving OPF can be found in a few recent works [66, 74]. Extending the convergence analysis based on the nonlinear power flow will be pursued as a future research effort.*

4.3.5 Estimation Error Analysis

In this subsection, we analytically quantify the errors of the SE algorithm under a linearized measurement model. A linear measurement model is given by

$$\mathbf{y}^k = \mathbf{H}\mathbf{z}^k + \boldsymbol{\xi}^k,$$

where $\mathbf{H} \in \mathbf{R}^{L\times M}$ is a measurement matrix that could be obtained by linearizing the nonlinear measurement function around a particular nominal operating condition.

For the linear WLS problem

$$\min_{\mathbf{z}^k} \frac{1}{2} \left(\mathbf{y}^k - \mathbf{H}\mathbf{z}^k\right)^\intercal \mathbf{W} \left(\mathbf{y}^k - \mathbf{H}\mathbf{z}^k\right),$$

the closed-form analytical solution is given by $\hat{\mathbf{z}}^k = \left(\mathbf{H}^\intercal\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^\intercal\mathbf{W}\mathbf{y}^k$. When the matrix $\mathbf{H}^\intercal\mathbf{W}\mathbf{H}$ is non-singular (which occurs when $\mathbf{W}$ is positive definite and $\mathbf{H}$ is full column

rank), the estimate can be expressed as

$$\begin{aligned}\hat{\mathbf{z}}^k &= \left(\mathbf{H}^\intercal \mathbf{W}\mathbf{H}\right)^{-1} \mathbf{H}^\intercal \mathbf{W}\mathbf{H}\mathbf{z}^k + \left(\mathbf{H}^\intercal \mathbf{W}\mathbf{H}\right)^{-1} \mathbf{H}^\intercal \mathbf{W}\boldsymbol{\xi}^k \\ &= \mathbf{z}^k + \left(\mathbf{H}^\intercal \mathbf{W}\mathbf{H}\right)^{-1} \mathbf{H}^\intercal \mathbf{W}\boldsymbol{\xi}^k.\end{aligned}$$

The WLS estimator is unbiased (since $\mathbf{E}\left[\hat{\mathbf{z}}^k\right] = \mathbf{z}^k$ due to the noise being zero mean), and the variance is given by $\text{Var}\left[\hat{\mathbf{z}}_j^k\right] = \sum_{i=1}^n \Gamma_{ji}\sigma_i^2$, where Γ_{ji} denotes the ji-th element of the $\boldsymbol{\Gamma} = \left(\mathbf{H}^\intercal \mathbf{W}\mathbf{H}\right)^{-1} \mathbf{H}^\intercal \mathbf{W}$, and σ_i^2 is the ith diagonal element of the measurement covariance matrix $\boldsymbol{\Sigma}$. The confidence intervals for components of the state estimate $\hat{\mathbf{z}}^k$ can be constructed as

$$\hat{\mathbf{z}}_j^k \pm c\sqrt{\text{Var}\left(\hat{\mathbf{z}}_j^k\right)} = \hat{\mathbf{z}}_j^k \pm c\sqrt{\sum_{i=1}^n \Gamma_{ji}\sigma_i^2},$$

where c can be chosen based on the prescribed confidence level. In addition to the bound in Theorem 2, the confidence intervals provide a numerical performance metric on the severity of estimation errors within the OPF control loop. In the next section, we will use this analysis to quantify estimation errors from voltage measurements and netoad pseudo-measurements.

4.4 Numerical Results

In this section, we use a modified three-phase unbalanced 11,000-node distribution network shown in Fig. 4.3 to demonstrate the effectiveness and scalability of the proposed OPF solver with SE in the loop. We model the primary loads of this system in detail and merge the secondary loads into distribution transformers, which lumps the system into a 4521-node distribution network. This extremely large system is divided into 5 clusters and then we utilized a spatially distributed optimization algorithm for computational affordability. The details of multi-phase power flow modelling and the feasibility of distributed algorithm were discussed in our companion paper [306]. Here, we focus on closing a loop between OPF and SE to solve a general OPF problem. We explore a tradeoff between sensing

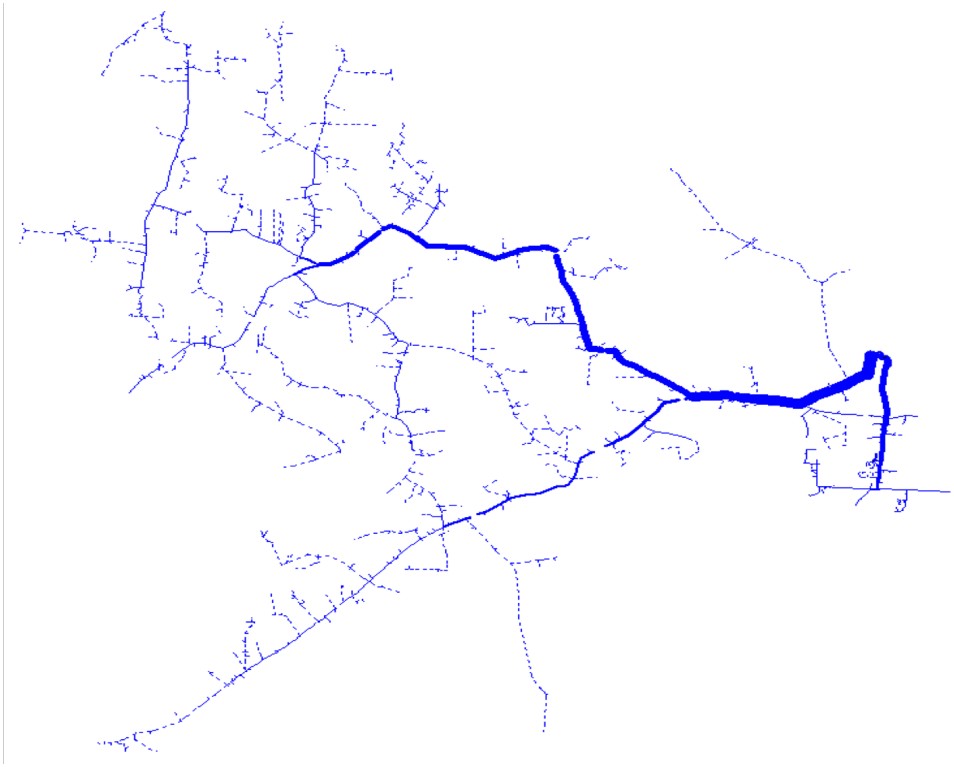
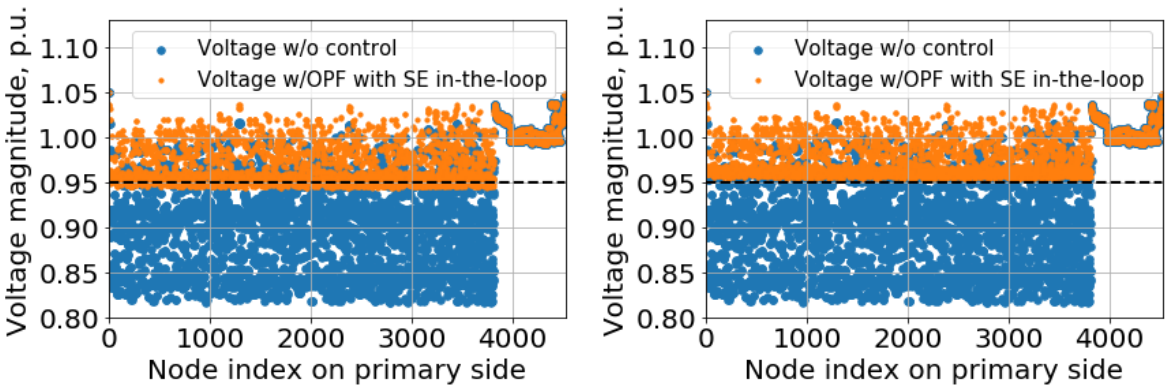


Figure 4.3. A 11,000-node distribution network. This testbed is constructed by connecting an IEEE 8,500-node distribution network and an EPRI Ckt7 test feeder at PCC. The primary side of this modified feeder is modelled in detail, while the loads on secondary side are merged into distribution transformers. This lumps the 11,000-node testbed into a 4521-node distribution network.

and communication effectiveness and performance of OPF controllers in an extremely large network.

We consider a voltage regulation problem where the electrical quantities of interests is voltage magnitude. In particular, we specify the vector $\mathbf{r}$ to voltage magnitude vector $\mathbf{r} := |\mathbf{v}| := [|v_1|, \ldots, |v_N|]^\intercal \in \mathbf{R}^N_{++}$ and consider

$$\begin{aligned} \textbf{OPF-V:} \quad \min_{\mathbf{p},\mathbf{q},|\mathbf{v}|} \quad & \sum_{i\in\mathcal{N}} C_i(p_i, q_i) + C_0(\mathbf{p}, \mathbf{q}), \\ \text{s.t.} \quad & |\mathbf{v}| = \mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{q} + |\mathbf{v}_0|, \\ & \underline{\mathbf{v}} \le |\mathbf{v}| \le \bar{\mathbf{v}}, \\ & (p_i, q_i) \in \mathcal{Z}_i, \forall i \in \mathcal{N}. \end{aligned}$$



(a) normal bound

(b) tighter bound

Figure 4.4. Voltage profile of OPF controller with SE in the loop. The black dash line indicates the lower voltage bound, i.e., 0.95 p.u.. After we utilize a tighter bound $[0.96, 1.05]$ to compensates the inherent errors of SE in the loop, the voltage profile on the right then meets the constraint.

The inequality constraints capture the lower and upper bounds $(\underline{\mathbf{v}}, \bar{\mathbf{v}})$ of voltage magnitudes. In particular, we take linear approximation for AC power flow to express $|\mathbf{v}|$ as a linear function of power injection $(\mathbf{p}, \mathbf{q})$. The coefficient matrices $(\mathbf{A}, \mathbf{B})$ of the linearized voltages and normalized vector $|\mathbf{v}_0|$ can be attained from numerous linearization methods, e.g., [22, 30, 102]. The gradient-based OPF controller (4.15) utilizes the online voltage magnitude measurement and voltage estimation to make the system converge. We consider a cost function $C_i(p_i, q_i) = (p_i - p_i^0)^2 + (q_i - q_i^0)^2$ that minimizes the deviation of the power setpoints (p_i, q_i) from their nominal/preferred level (p_i^0, q_i^0) for node i. The system objective function $C_0(\mathbf{p})$ penalizes the deviation of the total active power injection $P_0(\mathbf{p})$ at substation from its preferred values $\tilde{P}_0$ with a small weighted factor as $C_0(\mathbf{p}) = \alpha(P_0(\mathbf{p}) - \tilde{P}_0)^2$. Here, we choose a very small $\alpha = 0.0005$ so that we focus on the voltage regulation.

The default voltage profile of system in Fig. 4.3 without any control[7] is given by OpenDSS [90] shown in blue dots in Fig. 4.4. The voltage limits $\bar{\mathbf{v}}$ and $\underline{\mathbf{v}}$ are set to 1.05 and 0.95 p.u.

[7]We disable all rule-based control of voltage regulators, local capacitors and low-voltage transformers, etc., and then only solve the nonlinear power flow.

This particular network has a significant under-voltage situation. We implement Algorithm 1 to solve the above voltage regulation problem, while minimizing the objective. We consider distributed energy resources (DERs) with box constraints. The default netload settings can be found in [90, 94]. No storage device, PV and distributed diesel generator is included here for simplicity. Note that the proposed framework has the generality and flexibility to additionally include these controllable devices for various control objectives.. No storage device, PV and distributed diesel generator is included here for simple demonstration. Note that the proposed framework has the generality and flexibility to additionally include these controllable devices for various control objectives.

To have a clear picture of voltage level for OPF controller, most of the literature assumes that we have full knowledge of the real-time voltage information, which requires an unrealistic sensor deployment, extreme communication, and huge investment. To tradeoff these two issues and facilitate a practical OPF controller for acceptable performance, we randomly deploy voltage magnitude measurements at 3.6% of the nodes with measurement noise subject to Gaussian distribution with zero mean and 1% standard deviation. We also have the load pseudo-measurements for all nodal injections (i.e., active and reactive power) with significant noise (e.g., zero mean and 50% standard deviation of real values), which will guarantee the full observability of SE in the loop. The voltage information of the whole network will be fed back to OPF controller based on the voltage estimation results. The estimated active and reactive power in distribution networks are the primary estimation variables from the SE in the loop. The estimated voltage magnitudes are determined by the primary estimation via nonlinear power flow. The detailed problem formulation of voltage magnitude estimation for distribution networks can be found in our companion paper [305]. Note that the simulation is conducted on a desktop with AMD Ryzen 7 2700X Eight-Core Processor CPU@3.7GHz, 64GM RAM, Python 3.7 and Windows 10. We implement Algorithm 1 with the step-size 7×10^{-4} for primal updates and 1×10^{-3} for dual.

Fig. 4.4 visualizes the voltage profile regulated by the OPF controller with SE in the loop. In order to prevent voltage from failing below 0.95 p.u., the netloads must be curtailed based on the SE feedback information. The voltage of most nodes (i.e., orange dots) have been bounded within $[0.95, 1.05]$. There are few voltages just located across the lower bound with slight variations. This is due to the feedback signal containing the voltage estimation errors. To understand the severity of these estimation errors due to sensor noise and large variation of pseudo-measurement, we use the analysis in Section 4.3.5 to quantify the statistical estimation error numerically. Fig. 4.5 and Fig. 4.6 visualize the average and maximum errors of voltage estimation, and the comparison with analytically calculated confidence intervals over each OPF gradient step. Most of average errors over 1000 OPF iterations are bounded by the 99% confidence interval. For comparison purpose, we launch an ideal situation where the OPF solvers explicitly implement decisions based on the noisy measurement of all voltage magnitudes. The noises of the raw measurements are subjected to the Gaussian distribution with zero mean and 1% standard deviation of their true values. Based on the numerical analysis in Fig. 4.5, we conclude that having SE in the loop will significantly reduce errors due to inherent measurement noise compared to direct use of raw measurements. Also, the proposed SE in the loop can mitigates the effects of measurement errors on OPF controller performance.

To resolve the feedback estimation errors and further improve the performance of OPF controllers, we give a tighter lower bound (i.e., $[0.96, 1.05]$) based on the statistical analysis of SE errors. As a result shown in Fig. 4.7, the network curtails more netloads to achieve a more conservative voltage profile, which leads to a higher operational cost. We emphasize that there is always a tradeoff between: 1) cost of the measurement system (e.g., number of deployed sensors, communication infrastructure) and 2) OPF controller performance (e.g., robustness, feasibility and optimality). In general, our proposed approach provides utilities and system operators a framework to systematically design OPF controllers under a limited set of sensor measurements.

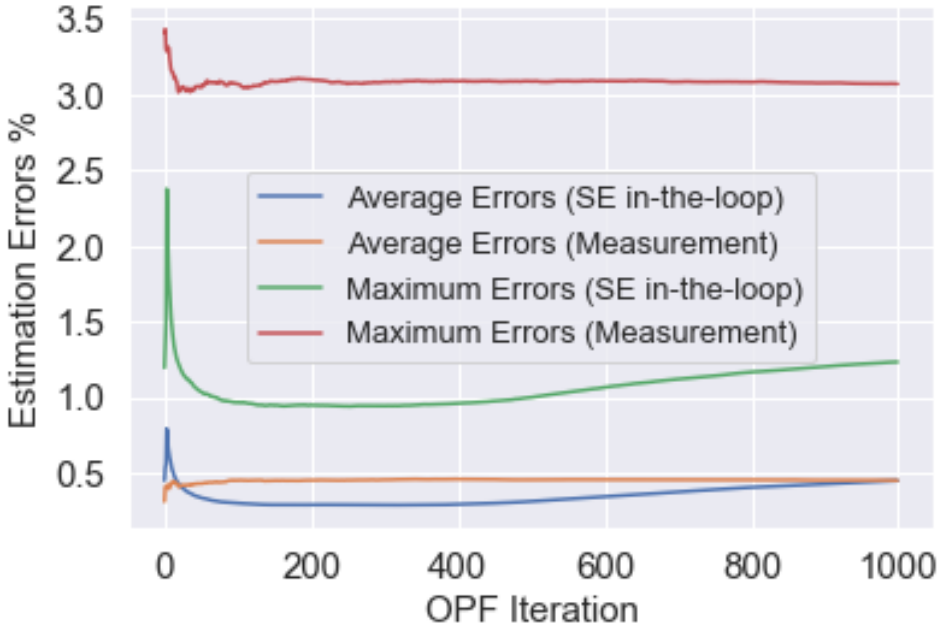


Figure 4.5. Comparison of estimation errors between SE in the loop and the use of raw voltage measurements. The running average of average/maximum errors show that the SE in the loop yields less error than using raw measurements.

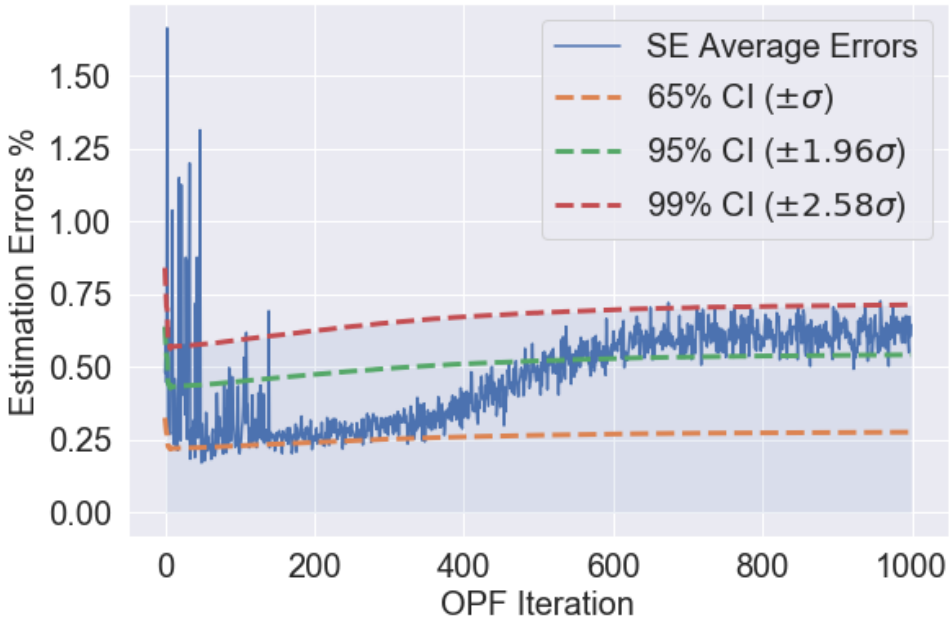


Figure 4.6. Comparison of the average estimation errors with different confidence intervals over 1000 OPF iterations.

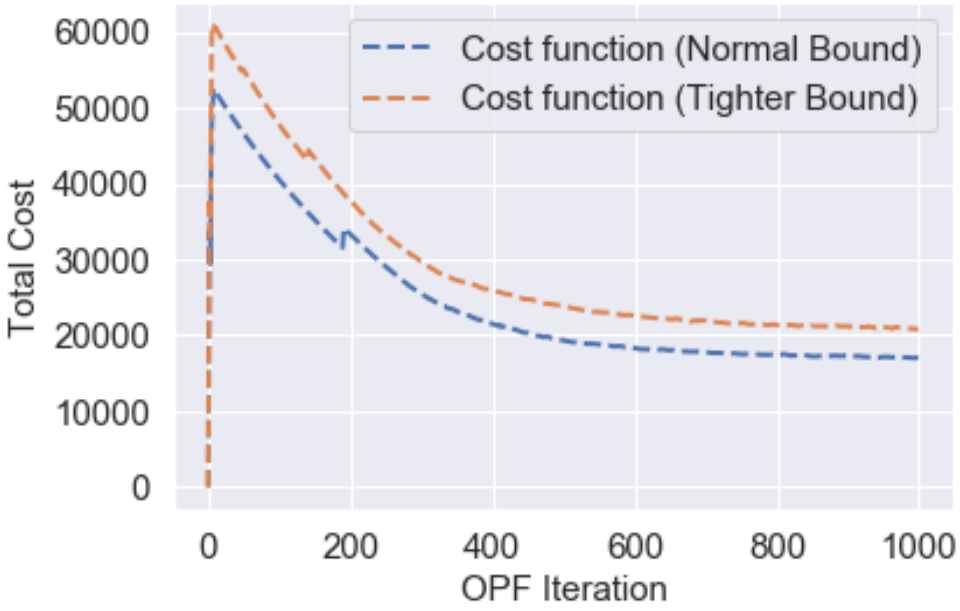


Figure 4.7. Total cost with SE in the loop over 1000 OPF iterations.

Overall, we conclude that the proposed OPF controller with SE feedback is able to systemically reduce estimation error of voltages at unmeasured nodes, successfully achieve voltage regulation, and improve robustness to measurement and estimation errors. The benefits of closing the loop between OPF controllers and state estimators can be clearly observed from the perspectives of effectiveness, robustness and efficiency.

4.5 Conclusions and Outlooks

In this chapter, we proposed a general optimal power flow controller with state estimation feedback to facilitate the operation of modern distribution networks. The controller depends explicitly on the state estimation results derived from system measurements. In contrast to existing works, our method utilizes a feedback loop to the OPF controller to estimate the system voltages from a limited number of sensors rather than making strong assumptions on full observability or requiring full state measurements. The performance of our design is an-

alyzed and numerically demonstrated. The numerical results demonstrate the effectiveness, scalability, and robustness of the proposed OPF controller with SE in the loop.

Our results on OPF problem launched an initial step towards closing a loop between control and state estimation in power systems. There are several lines of future works that can extend the present results in various ways to more fully explore the benefits, and discover the limitations of having SE in the loop for OPF controllers. Future work includes

- performance evaluation of various OPF formulations with different SE techniques in the loop;
- optimal sensor placement with SE in the loop for better OPF performance;
- OPF & SE in the loop co-design considering the estimation errors for a more efficient communication structure in a real network;
- convergence analysis based on the nonlinear power flow setting;
- performance discussions with different distributed algorithms, e.g., ADMM;
- extension with the time-varying power flow linearization.

Appendix

Proof of Theorem 2

Now we are ready to show the convergence of Algorithm 2. The expected distance between the generated sequence $\{\mathbf{x}^{k+1}\}$ and the saddle point of $\mathcal{L}$ can be characterized as:

$$
\begin{aligned}
& \mathbf{E}\Big[\|\mathbf{x}^{k+1}-\mathbf{x}^*\|_2^2\Big] \\
\leq\ & \mathbf{E}\Big[\|\mathbf{x}^k-\epsilon\tilde{\Phi}(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\Phi(\mathbf{x}^*)\|_2^2\Big] \\
=\ & \mathbf{E}\Big[\|\mathbf{x}^k-\epsilon\tilde{\Phi}(\mathbf{x}^k)+\epsilon\overline{\Phi}(\mathbf{x}^k)-\epsilon\overline{\Phi}(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\Phi(\mathbf{x}^*) \\
& +\epsilon\overline{\Phi}(\mathbf{x}^*)-\epsilon\overline{\Phi}(\mathbf{x}^*)\|_2^2\Big] \\
\leq\ & \mathbf{E}\Big[\|\mathbf{x}^k-\epsilon\overline{\Phi}(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\overline{\Phi}(\mathbf{x}^*)\|_2^2+\epsilon^2\|\tilde{\Phi}(\mathbf{x}^k)-\overline{\Phi}(\mathbf{x}^k)\|_2^2 \\
& +\epsilon^2\|\Phi(\mathbf{x}^*)-\overline{\Phi}(\mathbf{x}^*)\|_2^2\Big] \\
=\ & \mathbf{E}\Big[\|\mathbf{x}^k-\epsilon\overline{\Phi}(\mathbf{x}^k)+\epsilon\Phi(\mathbf{x}^k)-\epsilon\Phi(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\overline{\Phi}(\mathbf{x}^*) \\
& +\epsilon\Phi(\mathbf{x}^*)-\epsilon\Phi(\mathbf{x}^*)\|_2^2+\epsilon^2\|\tilde{\Phi}(\mathbf{x}^k)-\overline{\Phi}(\mathbf{x}^k)\|_2^2 \\
& +\epsilon^2\|\Phi(\mathbf{x}^*)-\overline{\Phi}(\mathbf{x}^*)\|_2^2\Big] \\
\leq\ & \mathbf{E}\left[\|\mathbf{x}^k-\epsilon\Phi(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\Phi(\mathbf{x}^*)\|_2^2\right] \\
& +\mathbf{E}\Big[\epsilon^2\|\overline{\Phi}(\mathbf{x}^k)-\Phi(\mathbf{x}^k)\|_2^2+2\epsilon^2\|\Phi(\mathbf{x}^*)-\overline{\Phi}(\mathbf{x}^*)\|_2^2\Big] \\
& +\epsilon^2\|\tilde{\Phi}(\mathbf{x}^k)-\bar{\Phi}(\mathbf{x}^k)\|_2^2 \\
\leq\ & \mathbf{E}\left[\|\mathbf{x}^k-\epsilon\Phi(\mathbf{x}^k)-\mathbf{x}^*+\epsilon\Phi(\mathbf{x}^*)\|_2^2\right]+\epsilon^2\left(\rho+\sigma_k^2+2\sigma_*^2\right) \\
\leq\ & \mathbf{E}\left[\|\mathbf{x}^k-\mathbf{x}^*\|_2^2\right]+\mathbf{E}\left[\|\epsilon\Phi(\mathbf{x}^k)-\epsilon\Phi(\mathbf{x}^*)\|_2^2\right] \\
& -2\epsilon\left(\Phi(\mathbf{x}^k)-\Phi(\mathbf{x}^*)\right)^{\intercal}\left(\mathbf{x}^k-\mathbf{x}^*\right)+\epsilon^2\left(\rho+\sigma_k^2+2\sigma_*^2\right) \\
\leq\ & \left(\epsilon^2L^2-2\epsilon M+1\right)\mathbf{E}\left[\|\mathbf{x}^k-\mathbf{x}^*\|_2^2\right]+\epsilon^2\left(\rho+\sigma_k^2+2\sigma_*^2\right),
\end{aligned}
\tag{4.20}
$$

where the first inequality is due to the non-expansiveness of the projection operator, the next two inequalities depend on the triangle inequality, the fourth inequality comes from

(4.16)–(4.17), and the last two inequalities are because of the strong monotonicity (4.9) and Lipschitz continuity (4.10) of the operator Φ. Next, we let $\Delta = \epsilon^2 L^2 - 2\epsilon M + 1$ and recursively implement the steps above until $k = 0$ to have

$$\begin{aligned}
& \mathbf{E}\Big[\|\mathbf{x}^{k+1} - \mathbf{x}^*\|_2^2\Big] \\
\leq\ & \Delta^{k+1}\|\mathbf{x}^0 - \mathbf{x}^*\|_2^2 + \epsilon^2\left(\rho + 2\sigma_*^2\right)\left(\frac{1-\Delta^{k+1}}{1-\Delta}\right) \\
& + \epsilon^2 \sum_{i=0}^{k} \Delta^{k-i}\sigma_k^2 \qquad (4.21a) \\
<\ & \Delta^{k+1}\|\mathbf{x}^0 - \mathbf{x}^*\|_2^2 + \epsilon^2\left(\rho + 2\alpha\right)\left(\frac{1-\Delta^{k+1}}{1-\Delta}\right) \\
& + \epsilon^2 \frac{1-\Delta^{k+1}}{1-\Delta}\alpha. \qquad (4.21b)
\end{aligned}$$

By applying the SE estimation variance bound in Assumption 6, (4.21a) is relaxed to (4.21b). Then we have the step size chosen as $0 < \epsilon \leq \bar{\epsilon} < 2M/L^2$ from Theorem 1, which leads to $0 < \Delta \leq \bar{\epsilon}^2 L^2 - 2\bar{\epsilon}M + 1 < 1$. As $k \to \infty$, Δ^{k+1} on the right-hand-side in (4.21b) will vanish. Given such Δ and any feasible initial point $\mathbf{x}^0$, we let k approaches the infinite and take supreme on the right side of (4.21b) to get the expectation of this discrepancy as

$$\lim_{k\to\infty} \sup \mathbf{E}\Big[\|\mathbf{x}^{k+1} - \mathbf{x}^*\|_2^2\Big] = \frac{\rho + 3\alpha}{2M/\epsilon - L^2}.$$

This concludes the proof.

CHAPTER 5

A PERFORMANCE AND STABILITY ANALYSIS OF LOW-INERTIA POWER GRIDS WITH STOCHASTIC SYSTEM INERTIA[1]

Traditional synchronous generators with rotational inertia are being replaced by low-inertia renewable energy resources (RESs) in many power grids and operational scenarios. Due to emerging market mechanisms, inherent variability of RESs, and existing control schemes, the resulting system inertia levels can not only be low but also markedly time-varying. In this chapter, we investigate performance and stability of low-inertia power systems with stochastic system inertia. In particular, we consider system dynamics modeled by a linearized stochastic swing equation, where stochastic system inertia is regarded as multiplicative noise. The $\mathcal{H}_2$ norm is used to quantify the performance of the system in the presence of persistent disturbances or transient faults. The performance metric can be computed by solving a generalized Lyapunov equation, which has fundamentally different characteristics from systems with only additive noise. For grids with uniform inertia and damping parameters, we derive closed-form expressions for the $\mathcal{H}_2$ norm of the proposed stochastic swing equation. The analysis gives insights into how the $\mathcal{H}_2$ norm of the stochastic swing equation depends on 1) network topology; 2) system parameters; and 3) distribution parameters of disturbances. A mean-square stability condition is also derived. Numerical results provide additional insights for performance and stability of the stochastic swing equation.

---

[1]This chapter is based on work supported by the National Science Foundation (NSF) under grant CMMI-1728605.

Chapter 5 in part is a reprint of material published in:

5.1 Introduction

Environmental and sustainability concerns are forcing unprecedented charges in the modern electric power system. The continued replacement of traditional synchronous generators by renewable energy sources (RESs) in power systems is raising concerns about their stability. As the penetration levels of RESs reach substantial fractions of total generation, power systems will require more low-inertia RESs to participate in frequency and voltage control. The inherent variability of RESs can produce high amplitude and persistent disturbances, which may adversely affect stability. Due to emerging market mechanisms and deregulated existing control schemes of RESs, the resulting system inertia levels can not only be low but also markedly time-varying. This system-level inertia variation together with unpredictable RESs and net loads make frequency control and power system stabilization more challenging. Future power systems need more sophisticated stochastic dynamic models and stochastic control methods for stability and performance analysis.

Many issues of power system stability have been well studied in recent decades through mathematical analysis and computational techniques [60, 58, 152, 28, 249, 211, 195, 86, 179, 6]. However, many of the underlying assumptions and models must be called into question in the context of low- and variable-inertia power systems. The decreasing system inertia results in higher rate of change of frequency, which requires controllers to respond faster to the system dynamics. In addition, the intermittent disturbances from RESs are spatially distributed over power networks, which requires that performance and stability analysis consider various grid topologies and the diverse dynamics of various grid-connected components.

Many challenges and related solutions for low-inertia power grid stability have been highlighted and discussed in [254, 255, 151, 274, 250]. Recently, stability analysis and control techniques have been proposed and have demonstrated their effectiveness for system stabilization [23, 78, 303, 88, 37, 249, 211, 86, 307, 243, 67, 252, 238, 302, 183, 71, 165, 6, 187, 7]. Vir-

tual inertia emulation is an approach to control the terminal behavior of inverter-interfaced RESs to mimic inertial response of conventional synchronous machines [23, 78, 303, 88, 37]. Recent works on virtual oscillators have shown that oscillator-based control strategies have advantages in faster response and global convergence [67, 252, 238]. Also, distributed control methodologies have been implemented on generation and load sides to provide effective ways for frequency stabilization [88, 302, 183, 71, 165]. Recent stability analyses have demonstrated that the efforts to maintain synchronous stability in low-inertia power grids depend on grid structure, node dynamics and coupling strength [249, 211, 86, 307, 243].

Overall, this line of research has explored useful solutions for the well recognized issues in low-inertia power systems. However, a large portion of unpredictable RESs will possible cause the system inertia to be not just low but also significantly time-varying. A few recent analysis consider the effects of system inertia variability raised by RESs [6, 7, 187, 212], but none of works explicitly consider a rigorous performance and stability analysis with respect to a stochastic system inertia model. As the inverter-based RESs dominated the generation, some RESs are required to participate in frequency and voltage control in a low/no-inertia power grid. The time-varying system inertia profiles and the heterogeneous allocation of inertia can lead to destabilizing effects, which complicates and challenges stability analysis and stabilization control in power systems.

In this chapter, we investigate performance and stability of low-inertia power systems with stochastic system inertia. In particular, we consider system dynamics modeled by a linearized stochastic swing equation, where stochastic system inertia is regarded as multiplicative noise. The main contributions are as follows:

- We consider the frequency dynamics of a low-inertia power grid as a stochastic linear system with both multiplicative (due to inertia variations) and additive noise (due to power injection disturbances), which connects stability analysis of a linearized swing equation with a generalized Lyapunov equation.

- We quantify the system $\mathcal{H}_2$ norm in terms of various outputs, which measures system performance in the presence of multiplicative and additive disturbances. The closed-form $\mathcal{H}_2$ norm of a homogeneous power grid is derived and discussed. In contrast to existing work with only additive disturbances, we observe that the system $\mathcal{H}_2$ norm depends in fundamentally different ways on 1) network topology; 2) system parameters; 3) distribution parameters of disturbances, and is sensitive to system inertia variation for certain outputs.

- A mean-square stability condition is also derived for the stochastic linear system, which indicates that a low-inertia grid can be destabilized in a second-order sense by inertia variability. Numerical results also indicate that a lower inertia grid with larger system inertia variance is less robust to disturbances.

The rest of the chapter is organized as follows: Section 5.2 introduces a stochastic linearized swing equation to model power systems with stochastic system inertia. Section 5.3 develops the system $\mathcal{H}_2$ norm of the stochastic swing equation for three particular outputs and derives a second-order stability condition. Section 5.4 presents some numerical results to illustrate the theory. Section 5.5 concludes the chapter and points out several possible future research directions.

5.2 Problem Formulation

5.2.1 System Modelling

Consider a power network $\mathcal{G}$ with $N+1$ nodes (buses) $\mathcal{V} = \{0\} \cup \{1, \ldots, N\}$ connected by a set of edges (transmission lines) $\mathcal{E} \subset \{\mathcal{V} \times \mathcal{V}\}$. We assume the power network is Kron-reduced [87], where each node represents an equivalent generator[2] with state variables (e.g., voltage

[2]The equivalent generator can be interpreted as the aggregate terminal interaction dynamics of a power system sub-area. Note that the equivalent generator here characterizes the dynamics of a group of grid-connected components (e.g., traditional generators and inverter-based generators) as a synchronous machine.

magnitude $|V_i|$ and voltage angle θ_i) and parameters (e.g., rotational inertia M_i, damping coefficient β_i). The index 0 is reserved for the grounded node/bus in a Kron-reduced power network. We consider the following swing equation to model the generator dynamics of each bus i

$$M_i\ddot{\theta}_i + \beta_i\dot{\theta}_i = P_{m,i} - P_{e,i}, \forall i = 0, 1, \ldots, N, \tag{5.1}$$

where $P_{m,i}$ refers to the mechanical power from the synchronous generator, and $P_{e,i}$ represents the electrical power injection of the generator. The nominal power injection at bus i is given by the power flow equations

$$\begin{aligned} P_{e,i} = \overline{g}_{ii}|V_i|^2 + \sum_{(i,j)\in\mathcal{E}} & g_{ij}|V_i||V_j|\cos(\theta_i - \theta_j) \\ & + \sum_{(i,j)\in\mathcal{E}} b_{ij}|V_i||V_j|\sin(\theta_i - \theta_j). \end{aligned} \tag{5.2}$$

where g_{ij} and b_{ij} denote the line conductance and the line susceptance, respectively. The shunt capacity at bus i is $\overline{g}_{ii}$. The admittance matrix $\mathbf{Y} \in \mathbf{C}^{(N+1)\times(N+1)}$ has elements

$$Y_{ij} = \begin{cases} \sum_{l\sim i}(g_{il} - \mathbf{j}b_{il}) + \overline{g}_{ii} & \text{if } i = j \\ -(g_{ij} - \mathbf{j}b_{ij}) & (i,j) \in \mathcal{E} \\ 0 & (i,j) \notin \mathcal{E}. \end{cases}$$

The admittance matrix can be written in compact matrix form

$$\mathbf{Y} = (L_G + \overline{g}) - \mathbf{j}L_B,$$

where L_G and L_B are the conductance matrix and the susceptance matrix, and $\overline{g} := \text{diag}\{\overline{g}_{ii}\}$ is a diagonal matrix of bus shunt capacitors. The Laplacian matrices L_B and L_G comprise the weights of the line susceptance b_{ij} and the line conductance g_{ij} in the Kron-reduced network, respectively.

The dynamic model (5.1)-(5.2) is often linearized around an operating point, which allows the study of system response in the presence of small faults or persistence disturbances around the linearization point. We use the linearized "DC power flow model", neglecting the line reactance, to approximate (5.2), which assumes $|V_i| \approx 1$ and $|\theta_i - \theta_j| \ll 1$. Detailed analysis and applications of the DC power flow approximation are discussed in [215]. Then the system dynamics (5.1) becomes

$$M_i\ddot{\theta}_i + \beta_i\dot{\theta}_i \approx -\sum_{(i,j)\in\mathcal{E}} b_{ij}(\theta_i - \theta_j) + P_{m,i}. \tag{5.3}$$

We then shift the equilibrium point of (5.3) to the origin and write it in state-space form

$$\begin{aligned} \dot{\theta} &= \omega \\ M\dot{\omega} &= -L_B\theta - D\omega + W, \end{aligned} \tag{5.4}$$

where $\theta = [\theta_1, \ldots, \theta_N]^\intercal$ and $\omega = [\omega_1, \ldots, \omega_N]^\intercal$. The inertia matrix and damping matrix are defined as $M := \text{diag}\{M_i\}$, $D := \text{diag}\{\beta_i\}$. The standard approach to analyze (5.3) considers various disturbances as additive noise W driving the system away from its current equilibrium point. The remainder of this chapter considers both *multiplicative* and *additive* disturbances in (5.4).

5.2.2 Frequency Dynamics with Multiplicative and Additive Noise

Here we consider inertia variations caused by RESs, which are modeled by treating the system inerta matrix as multiplicative noise rather than simply a constant. The inertia parameter at each node M_i can be modeled as a independent Wiener processes on a probability space $(\Omega, \mathcal{F}, P)$ with mean $\bar{M}_i$ and variance σ_i^2 [199]. The nominal inertia matrix is defined as $\mathcal{M} := \text{diag}\{\bar{M}_i\}$, and we rewrite (5.4) as a stochastic linear system with both additive and multiplicative noise

$$\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -(\mathcal{M}^{-1} + \boldsymbol{\delta}\mathcal{M}^{-1})L_B & -(\mathcal{M}^{-1} + \boldsymbol{\delta}\mathcal{M}^{-1})D \end{bmatrix} \begin{bmatrix} \theta \\ \omega \end{bmatrix} + \begin{bmatrix} 0 \\ \eta I \end{bmatrix} W. \tag{5.5}$$

The matrix $\mathcal{M}^{-1} := \text{diag}\{\hat{M}_i^{-1}\}$ collects the mean values of the inverse distribution of $\bar{M}_i$ on the diagonal, and the additive noise W represents independent white-noise with zero mean and unit variance, scaled by η. Each diagonal element of the matrix $\boldsymbol{\delta}\mathcal{M}^{-1} := \text{diag}\{\delta_i\}$ is modeled as an independent Wiener process with zero mean and variance $\hat{\sigma}_i^2$. For simplicity, we write (5.5) in a generalized form with outputs as a multi-input multi-output stochastic linear system with multiplicative noise

$$\begin{aligned} \dot{x} &= A_0 x + \sum_{i=1}^{N+1} A_i \delta_i x + BW, \\ y &= Cx, \end{aligned} \tag{5.6}$$

where,

$$A_0 = \begin{bmatrix} 0 & I \\ -\mathcal{M}^{-1}L_B & -\mathcal{M}^{-1}D \end{bmatrix}, A_i = \begin{bmatrix} 0 & 0 \\ -R_i L_B & -R_i D \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ \eta I \end{bmatrix}, x = \begin{bmatrix} \theta^\intercal, \omega^\intercal \end{bmatrix}^\intercal.$$

The matrix $A_0 \in \mathbf{R}^{2(N+1)}$ characterizes the nominal system with average inertia and damping ratio. We define an *inertia disturbance allocation* matrix $R_i \in \mathbf{R}^{N+1}$ in $A_i \in \mathbf{R}^{2(N+1)}$ associated with each bus i. The elements in R_i are all zeros except for one diagonal element $r_{ii} = 1$, which maps the corresponding inertia disturbance δ_i onto bus i. If the inertia variation at bus i is insignificant, we set $R_i = 0$ to remove the inertia disturbance at ith bus. In the following, we refer to the stochastic system input/output mapping (5.6) as $\Sigma = (A_0, A_i, B, C)$. This stochastic linear dynamical model allows us to investigate the effects of phase angle deviations and frequency changes in the presence of both additive and multiplicative disturbances around the original operating point. To assess the system stability and evaluate the performance of (5.6), we consider the following three outputs [211]:

Phase cohesiveness. This output quantifies real power losses due to phase differences caused by the fluctuations from the nominal operating points [249]. The resistive losses on

transmission lines during transients or due to persistent disturbances can be expressed in terms of Laplacian matrix L_G

$$P_{\text{loss}} = \sum_{(i,j)\in\mathcal{E}} g_{ij}(\theta_i - \theta_j)^2 = \theta^\intercal L_G \theta.$$

Expressed in terms of the output, we have $P_{\text{loss}} = y^\intercal y$ with

$$y = \begin{bmatrix} L_G^{\frac{1}{2}} & 0 \end{bmatrix} x.$$

It is worth to note that the stochastic linear system (5.6) driven from the linearized swing equation neglects the line resistances. The output matrix with conductance matrix L_G can capture resistive losses arising from (5.6).

Frequency. To quantify frequency deviations due to faults or disturbances, we have the output $y = [0, I]x$.

Phase cohesiveness & frequency. This output quantifies both phase and frequency performance with the output matrix

$$y = \text{diag}\{L_G^{\frac{1}{2}}, \kappa I\}x,$$

where $\kappa \in \mathbf{R}_+$ trades off phase angle and frequency deviations.

In this chapter, we use the system $\mathcal{H}_2$ norm to quantify the system performance under above three outputs. The system $\mathcal{H}_2$ norm is the root-mean-square value of the output when the system is driven by multiplicative and additive noise inputs. It has been widely studied in power system models with additive noise [211, 212, 249, 138].

5.3 Coherency Performance Metric

5.3.1 System Reduction

The stochastic swing equation model (5.6) and associated performance outputs consist of two key Laplacian matrices: the conductance matrix L_G and the susceptance matrix L_B,

which both have a zero eigenvalue associated with the eigenvector $\mathbf{1}$. The Laplacian structure implies (5.6) is not asymptotically stable, but the subspace corresponding to the zero eigenvalue does not appear in the output [249]. The physical interpretation of this zero eigenvalue is that we lack a grounded reference bus in the power network. Therefore, we consider the grounded Laplacians by deleting the kth row and column of L_G and L_B, respectively, yielding $\tilde{L}_G$ and $\tilde{L}_B$. The states in the reduced system $\tilde{\theta}$ and $\tilde{\omega}$ are obtained by simply removing the kth element of the original vectors, which can be interpreted as grounding bus k, with θ_k and ω_k forced to zero for the voltage references. Then the grounded system $\tilde{\Sigma}$ can be expressed as

$$\begin{aligned} \dot{\tilde{x}} &= \tilde{A}_0\tilde{x} + \sum_{i=1}^{N} \tilde{A}_i\delta_i\tilde{x} + \tilde{B}\tilde{W}, \\ \tilde{y} &= \tilde{C}\tilde{x}, \end{aligned} \tag{5.7}$$

where,

$$\tilde{A}_0 = \begin{bmatrix} 0 & I \\ -\tilde{\mathcal{M}}^{-1}\tilde{L}_B & -\tilde{\mathcal{M}}^{-1}\tilde{D} \end{bmatrix}, \tilde{A}_i = \begin{bmatrix} 0 & 0 \\ -\tilde{R}_i\tilde{L}_B & -\tilde{R}_i\tilde{D} \end{bmatrix},$$

$$\tilde{B} = \begin{bmatrix} 0 \\ \eta I \end{bmatrix}, \tilde{x} = \begin{bmatrix} \tilde{\theta}^\intercal, \tilde{\omega}^\intercal \end{bmatrix}^\intercal.$$

The kth row and column of matrices $\mathcal{M}$, R_i and D are also discarded in the reduced power system (5.7), which are then written as $\tilde{\mathcal{M}}$, $\tilde{R}_i$ and $\tilde{D}$, respectively. We assume the power network $\mathcal{G}$ in our problem is connected so that the grounded Laplacians $\tilde{L}_G$ and $\tilde{L}_B$ are symmetric positive definite. Thus, all eigenvalues of matrix $\tilde{A}_0$ are located in the open left half of the complex plane. We will detail the system stability with finite $\mathcal{H}_2$ norm in the present of multiplicative noise in the next subsection.

5.3.2 Performance Metric and Stability Conditions

The system $\mathcal{H}_2$ norm is the root-mean-square value of the output when the system is driven by multiplicative and additive noise inputs. For the system $\tilde{\Sigma} = (\tilde{A}_0, \tilde{A}_i, \tilde{B}, \tilde{C})$, the squared

$\mathcal{H}_2$ norm is given by

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \mathrm{Tr}(\tilde{B}^\intercal \tilde{Q} \tilde{B}), \tag{5.8}$$

where $\tilde{Q}$ is the observability Gramian [76]. The observability Gramian $\tilde{Q}$ can be interpreted as the steady-state output covariance [259], which has a stochastic integral expression [299]. When a finite positive semidefinite observability Gramian $\tilde{Q}$ exists, it can be attained by solving the following generalized Lyapunov equation (5.9) [144, 76, 299]

$$\tilde{A}_0^\intercal \tilde{Q} + \tilde{Q}\tilde{A}_0 + \sum_{i=1}^{N} \hat{\sigma}_i^2 \tilde{A}_i^\intercal \tilde{Q} \tilde{A}_i = -\tilde{C}^\intercal \tilde{C}, \tag{5.9}$$

which also implies the system is second-moment bounded (i.e., mean-square stable), resulting in the finite $\mathcal{H}_2$ norm.

The Gramian obtained from this generalized Lyapunov equation is used to compute the $\mathcal{H}_2$ norm of system (5.7), which explicitly incorporates both multiplicative and additive noise. The multiplicative noise can be removed by letting $\hat{\sigma}_i^2 = 0$ in (5.9), yielding a standard Lyapunov equation. In contrast to systems with only additive noise, there are differing notions of stability when multiplicative noise is present. In particular, even when the mean value of the state is stable, (i.e., $\tilde{A}_0$ is stable), the covariance of the state may be unstable due to the multiplicative noise, in which case the $\mathcal{H}_2$ norm becomes infinite. When the multiplicative noise variances are sufficiently small, the system will have second-moment bounded (i.e., be mean-square stable). We have the following second-order stability definition, which is equivalent to existence of a finite positive semidefinite solution to the generalized Lyapunov equation and finiteness of the corresponding $\mathcal{H}_2$ norm [76].

Definition 3. *(Second-moment boundedness). The system* (5.7) *is called second moment bounded, or mean square stable, if there exists a positive constant* α *such that*

$$\lim_{t\to\infty} \mathbb{E}[\tilde{x}(t)^\intercal \tilde{x}(t)] \leq \alpha, \quad \forall \tilde{x}(0) \in \mathbf{R}^N.$$

Remark 6. *(Computation of $\mathcal{H}_2$ norm for non-homogeneous power grids). The generalized Lyapunov equation* (5.9) *is linear in $\tilde{Q}$ and can be solved directly using vectorization and Kronecker products, yielding*

$$vec(\tilde{Q}) = -\left(\tilde{A}_0^\intercal \otimes I + I \otimes \tilde{A}_0^\intercal + \sum_{i=1}^{N} \tilde{A}_i^\intercal \otimes \tilde{A}_i^\intercal\right)^{-1} vec(\tilde{C}^\intercal \tilde{C}).$$

For the standard Lyapunov equation, factorization methods can be used to exploit the structure of the equation and achieve superior computational complexity. However, these methods cannot be easily applied to the generalized Lyapunov equation, and alternative methods have been studied [27, 77, 15, 43], e.g., using Krylov subspaces, semidefinite programming, or differential equations for the state covariance matrix.

To gain additional insights into the effects of multiplicative noise on power networks, we now consider computation of the $\mathcal{H}_2$ norm for power networks with homogeneous nominal inertia and damping. The inverse inertia perturbation at each bus is an independent stochastic process, which has identical mean and variance $\hat{M}_i = \hat{M}$, $\hat{\sigma}_i^2 = \hat{\sigma}^2, \forall i$. The damping ratio is also assumed to be identical $D = \beta I$. We will derive a closed-form expression for the $\mathcal{H}_2$ norm of the stochastic system (5.7), which allows us to highlight several insights regarding system performance with inertia disturbances.

Theorem 3. *($\mathcal{H}_2$ norm for homogeneous power grids). Consider an N-generator power system with both multiplicative and additive noise specified by parameters $\tilde{\Sigma} = (\tilde{A}_0, \tilde{A}_i, \tilde{B}, \tilde{C})$ that define the input-output mapping shown in* (5.7)*. Suppose the inertia and damping are homogeneous, i.e., $\hat{M}_i = \hat{M}$, $\hat{\sigma}_i^2 = \hat{\sigma}^2, \forall i$ and $D = \beta I$. Consider also a general output matrix $\tilde{C} = \begin{bmatrix} \tilde{J}^{\frac{1}{2}} & 0 \\ 0 & \tilde{K}^{\frac{1}{2}} \end{bmatrix}$, where matrices $\tilde{J}$ and $\tilde{K}$ are positive definite. Then the squared $\mathcal{H}_2$ norm is given by*

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}^2} \operatorname{Tr}\left[\tilde{P}^{-1}\left(\hat{M}\tilde{J}\tilde{L}_B^{-1} + \tilde{K}\right)\right], \tag{5.10}$$

where $\tilde{P} = \left[\left(\frac{2\beta}{\hat{M}} - \hat{\sigma}^2\beta^2\right) I - \hat{\sigma}^2 \hat{M} \tilde{L}_B\right]$.

Proof. See the Appendix. □

Remark 7. *($\mathcal{H}_2$ norm for specified outputs). We have the following expressions for the $\mathcal{H}_2$ norm of system* (5.7) *for the three specific outputs mentioned in Section II:*

Phase cohesiveness *($\tilde{C} = [\tilde{L}_G^{\frac{1}{2}}, 0]$)*

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}} \operatorname{Tr}\left(\tilde{P}^{-1}\tilde{L}_G\tilde{L}_B^{-1}\right).$$

Frequency *($\tilde{C} = [0, I]$)*

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}^2} \operatorname{Tr}\left(\tilde{P}^{-1}\right).$$

Phase cohesiveness & frequency *($\tilde{C} = diag\{\tilde{L}_G^{\frac{1}{2}}, \kappa I\}$)*

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}^2} \operatorname{Tr}\left[\tilde{P}^{-1}\left(\hat{M}\tilde{L}_G\tilde{L}_B^{-1} + \kappa^2 I\right)\right].$$

Corollary 1. *(Second-moment bounded (Mean-square stability) condition). The power system in* (5.7) *with both multiplicative and additive noise is second-moment bounded (mean-square stable) and has finite $\mathcal{H}_2$ norm if and only if*

$$\hat{\sigma}^2 < \frac{2\beta}{\hat{M}\left[\beta^2 + \lambda_{max}(\tilde{L}_B)\hat{M}\right]}, \tag{5.11}$$

where $\lambda_{max}(\tilde{L}_B)$ denotes the largest eigenvalue of Laplacian matrix $\tilde{L}_B$, and $\hat{\sigma}^2$ is the variance of the inverse distribution of M.

Proof. Since $\tilde{A}_0$ is stable, the (standard) Lyapunov equation obtained when the inertia variance $\hat{\sigma}^2 = 0$ has a finite positive definite solution, and the corresponding $\mathcal{H}_2$ norm is finite. As $\hat{\sigma}^2$ increases, the smallest eigenvalue of the $\tilde{P}$ matrix defined in the $\mathcal{H}_2$ norm expression decreases (and thus largest eigenvalue of $\tilde{P}^{-1}$ increases), causing the $\mathcal{H}_2$ norm to increase. When $\hat{\sigma}^2$ approaches a critical value where $\tilde{P}$ goes from being positive definite to being singular, the $\mathcal{H}_2$ norm approaches infinity, and the system has unbounded second moment when the smallest eigenvalue of $\tilde{P}$ is zero. Examining the condition where the smallest eigenvalue of $\tilde{P}$ is zero yields the condition that guarantees bounded second moment. □

We will provide discussions of this Corollary in the context of our numerical results in the following section.

5.4 Numerical Results

The results derived in the previous section indicate that the $\mathcal{H}_2$ norm of a power system with both multiplicative and additive noise depends on 1) the system topology (via the Laplacians L_B and L_G); 2) the nominal system parameters (via the nominal inertia and damping coefficients); and 3) the distribution parameters of the multiplicative and additive disturbances. In this section, we present numerical simulations to analyze stability and to evaluate performance for three different outputs. Consider an interconnected power network with four areas (e.g., with homogeneous inertia M and damping ratio β) shown in Fig.5.1. We assume this power network is Kron-reduced and single-phase equivalent, with line data given in Table 5.1.

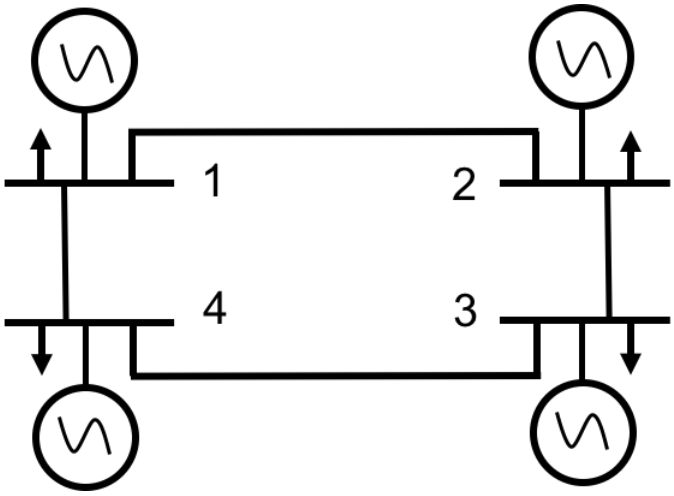
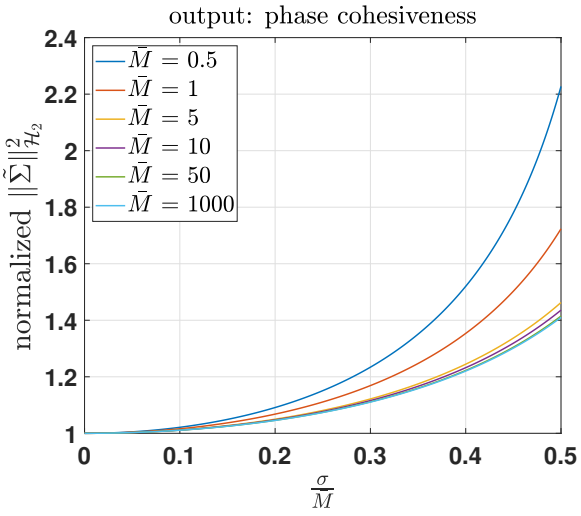


Figure 5.1. A four-area interconnected power system.

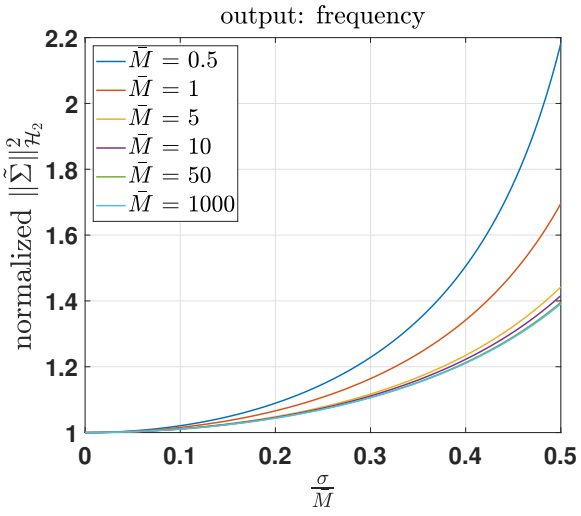
Remark 8. *(Approximation of the inverse distribution of M). To facilitate interpretation of insights derived from our previous analysis in terms of distribution parameters of M instead of the inverse distribution M^{-1}, we seek to express the parameters (e.g., $\bar{M}$ and σ^2) of the*

Table 5.1. Line impedance parameters

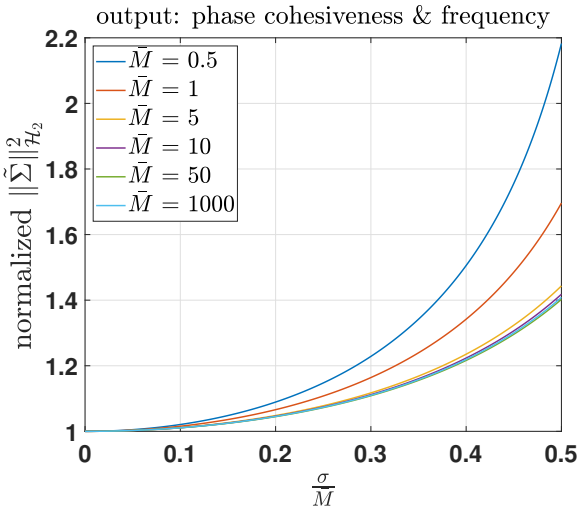
line $\mathcal{E}(i,j)$	(1,2)	(2,3)	(3,4)	(4,1)
r_{ij}	0.4	0.5	0.6	0.28
x_{ij}	0.386	0.294	0.596	0.474



(a) phase cohesiveness output



(b) frequency output



(c) phase cohesiveness & frequency output $\kappa = 10$

Figure 5.2. Comparison of the squared $\mathcal{H}_2$ norm of the stochastic system $\tilde{\Sigma}$ with three outputs under various values of $\bar{M}$ and inertia variance, quantified by $\frac{\sigma}{\bar{M}}$, for damping ratio $\beta = 1$. The results are normalized by the $\mathcal{H}_2$ norm with only additive noise, i.e., $\sigma^2 = 0$, [cf. [211, 249]].

original distribution M in terms of parameters (e.g., $\hat{M}$ and $\hat{\sigma}^2$) of the inverse distribution M^{-1}. An approximate formula. commonly known as the δ-method (based on a Taylor series expansion [26]), can be used to estimate the mean and variance of M^{-1} considering variations around the mean value $\bar{M}$ [194]. This yields an approximation of the distribution parameters of M^{-1} give by $\hat{M}^{-1} \approx \bar{M}^{-1}$, $\hat{\sigma}^2 \approx \frac{\sigma^2}{\bar{M}^4}$. The $\mathcal{H}_2$ norm of the system (5.7) *with the phase cohesiveness output is then approximated by*

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 \approx \frac{1}{\bar{M}} \operatorname{Tr}\left(\left[\left(\frac{2\beta}{\bar{M}} - \frac{\sigma^2\beta^2}{\bar{M}^4}\right) I - \frac{\sigma^2}{\bar{M}^3}\tilde{L}_B\right]^{-1} \tilde{L}_G \tilde{L}_B^{-1}\right).$$

The $\mathcal{H}_2$ norm (5.7) *for the other outputs can be similarly approximated. We emphasize that this approximation is only to facilitate interpretation; it is possible but less intuitive to work with the inverse inertia distribution.*

Fig.5.2(a) shows the $\mathcal{H}_2$ norm for the phase cohesiveness output with increasing variance of inertia disturbances σ^2. In contrast to the additive noise case, where the $\mathcal{H}_2$ norm of (5.4) is independent of the system inertia [249, 211], the analytical (Theorem 1) and numerical results in Fig.5.2(a) demonstrate that the $\mathcal{H}_2$ norm is strongly dependent on the nominal system inertia and its distribution. The results in Fig.5.2(a) also demonstrate that a low-inertia power grid with larger inertia disturbance incur more resistive power losses in the presence of persistent disturbances or transient events. Fig.5.2(b)-Fig.5.2(c) shows that the $\mathcal{H}_2$ norms of systems with frequency output and phase cohesiveness & frequency output are also increasing functions of mean and variance of the system inertia disturbance. The system will suffer a larger stability degradation in the presence of stronger system inertia disturbance. The results from Fig.5.2(a)-Fig.5.2(c) also indicate that a power system with lower inertia is less robust to system inertia disturbances. In particular, the degradation of performance and stability margin is more severe in a power grid with less rotational inertia.

The analysis and numerical results indicate that the grid topology (via the Laplacians) plays an essential role in system stability. The first moment stability criteria (mean stability)

requires the second smallest eigenvalue of Laplacian matrix L_B to strictly larger than zero. Interestingly, the mean-square stability condition requires an upper bound on the *largest* eigenvalue of Laplacian matrix L_B

$$\lambda_{\max}(L_B) < \frac{2\beta - \hat{M}\hat{\sigma}^2\beta^2}{\hat{\sigma}^2\hat{M}^2}.$$

The mean and mean-square stability criteria provide a theoretical boundary on grid topology in terms of system coefficients and the distribution parameters of inertia disturbance (mean and variance). The mean-square stability condition indicates that highly connected power grids are more sensitive to multiplicative noise, which contrasts with the mean stability condition, where increasing algebraic connectivity improves stability robustness.

5.5 Conclusion and Outlooks

In this chapter, we proposed a stochastic swing equation with both multiplicative and additive noise to study low- and variable-inertia power system through the system $\mathcal{H}_2$ norm. The $\mathcal{H}_2$ norm can be computed by solving a generalized Lyapunov equation. For grids with homogeneous inertia and damping ratio, we derived an analytical expression of the $\mathcal{H}_2$ norm for various performance outputs. In contrast to the additive case, inertia variations may cause second-moment instability, even when the state mean is stable. Further, the performance metrics always depends on physical properties (via the nominal inertia, damping ratio, inertia distribution parameters) and network structure (via the susceptance matrix). Numerical results also indicate that the low-inertia grids are vulnerable to large system inertia disturbances.

Ongoing works and potential future research directions include 1) more detailed stability analysis based on various performance metrics; 2) further numerical and analytical discussion of the stochastic swing equations and 3) design of optimal controllers for low- and variable-inertia grids with stochastic system inertia.

Appendix

Proof. (of Theorem 3). The squared $\mathcal{H}_2$ norm of (5.7) is given by

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \mathrm{Tr}\left(\tilde{B}^\intercal \tilde{Q} \tilde{B}\right), \tag{5.12}$$

where $\tilde{Q}$ is the observability Gramian, which can be obtained by solving the generalized Lyapunov equation

$$\tilde{A}_0^\intercal \tilde{Q} + \tilde{Q}\tilde{A}_0 + \sum_{i=1}^{N} \hat{\sigma}_i^2 \tilde{A}_i^\intercal \tilde{Q} \tilde{A}_i = -\tilde{C}^\intercal \tilde{C}.$$

The system matrices include the homogeneous inertia $\hat{M}$, inertia variance $\hat{\sigma}^2$, and damping coefficient β. We substitute $\tilde{B}$ in (5.12) and partition the Gramian as $\tilde{Q} = \begin{bmatrix} \tilde{Q}_1 & \tilde{Q}_0 \\ \tilde{Q}_0^\intercal & \tilde{Q}_2 \end{bmatrix}$. Due to the structure of the system matrices, the squared $\mathcal{H}_2$ norm of system $\tilde{\Sigma}$ becomes

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}^2}\,\mathrm{Tr}\left(\tilde{Q}_2\right). \tag{5.13}$$

Expanding the generalized Lyapunov equation yields

$$\begin{aligned} &\begin{bmatrix} -\frac{1}{\hat{M}}\tilde{L}_B\tilde{Q}_0^\intercal, & -\frac{1}{\hat{M}}\tilde{L}_B\tilde{Q}_2 \\ \tilde{Q}_1 - \frac{\beta}{\hat{M}}\tilde{Q}_0^\intercal, & \tilde{Q}_0 - \frac{\beta}{\hat{M}}\tilde{Q}_2 \end{bmatrix} + \begin{bmatrix} -\frac{1}{\hat{M}}\tilde{Q}_0\tilde{L}_B, & \tilde{Q}_1 - \frac{\beta}{\hat{M}}\tilde{Q}_0 \\ -\frac{1}{\hat{M}}\tilde{Q}_2\tilde{L}_B, & \tilde{Q}_0^\intercal - \frac{\beta}{\hat{M}}\tilde{Q}_2 \end{bmatrix} \\ &\qquad + \hat{\sigma}^2 \begin{bmatrix} \tilde{L}_B\tilde{Q}_2\tilde{L}_B, & \beta\tilde{L}_B\tilde{Q}_2 \\ \beta\tilde{Q}_2\tilde{L}_B, & \beta^2\tilde{Q}_2 \end{bmatrix} = \begin{bmatrix} -\tilde{J}, & 0 \\ 0, & -\tilde{K} \end{bmatrix}, \end{aligned} \tag{5.14}$$

and the diagonal blocks are

$$-\frac{1}{\hat{M}}\tilde{L}_B\tilde{Q}_0^\intercal - \frac{1}{\hat{M}}\tilde{Q}_0\tilde{L}_B + \hat{\sigma}^2\tilde{L}_B\tilde{Q}_2\tilde{L}_B = -\tilde{J}, \tag{5.15a}$$

$$\tilde{Q}_0 - \frac{\beta}{\hat{M}}\tilde{Q}_2 + \tilde{Q}_0^\intercal - \frac{\beta}{\hat{M}}\tilde{Q}_2 + \hat{\sigma}^2\beta^2\tilde{Q}_2 = -\tilde{K}. \tag{5.15b}$$

Since $\tilde{L}_B$ is nonsingular, the above two equations can be rearranged as

$$-\tilde{L}_B\tilde{Q}_0^\intercal\tilde{L}_B^{-1} - \tilde{Q}_0 + \hat{\sigma}^2\hat{M}\tilde{L}_B\tilde{Q}_2 = -\hat{M}\tilde{J}\tilde{L}_B^{-1} \tag{5.16a}$$

$$\tilde{Q}_0 + \tilde{Q}_0^\intercal + \left(\hat{\sigma}^2\beta^2 - \frac{2\beta}{\hat{M}}\right)\tilde{Q}_2 = -\tilde{K}. \tag{5.16b}$$

Adding (5.16a) to (5.16b) and multiplying by -1 gives

$$-\tilde{Q}_0^\intercal + \tilde{L}_B\tilde{Q}_0^\intercal\tilde{L}_B^{-1} + \tilde{P}\tilde{Q}_2 = \hat{M}\tilde{J}\tilde{L}_B^{-1} + \tilde{K}, \tag{5.17}$$

where $\tilde{P} = \left[\left(\frac{2\beta}{\hat{M}} - \hat{\sigma}^2\beta^2\right)I - \hat{\sigma}^2\hat{M}\tilde{L}_B\right]$. Then multiplying by $\tilde{P}^{-1}$ and taking the trace gives

$$\begin{aligned}\mathrm{Tr}\left(\tilde{P}^{-1}\left[\tilde{L}_B\tilde{Q}_0^\intercal\tilde{L}_B^{-1} - \tilde{Q}_0^\intercal\right]\right) + \mathrm{Tr}\left(\tilde{Q}_2\right)& \\ = \mathrm{Tr}\left[\tilde{P}^{-1}\left(\hat{M}\tilde{J}\tilde{L}_B^{-1} + \tilde{K}\right)\right].&\end{aligned}$$

We will show that the first term in the above equation is zero. We define $\tilde{P} = \left[aI - b\tilde{L}_B\right]$, where $a = \frac{2\beta}{\hat{M}} - \hat{\sigma}^2\beta^2$ and $b = \hat{\sigma}^2\hat{M}$. It can be seen that $\tilde{P}^{-1}$ and $\tilde{L}_B$ commute by expanding the term $\left[I - \frac{b}{a}\tilde{L}_B\right]^{-1}$ in a Neumann series

$$\begin{aligned}\tilde{P}^{-1}\tilde{L}_B = \left[aI - b\tilde{L}_B\right]^{-1}\tilde{L}_B &= \frac{1}{a}\left[I - \frac{b}{a}\tilde{L}_B\right]^{-1}\tilde{L}_B \\ &= \frac{1}{a}\sum_{k=0}^{\infty}\left(\frac{b}{a}\tilde{L}_B\right)^k\tilde{L}_B = \tilde{L}_B\tilde{P}^{-1}.\end{aligned} \tag{5.18}$$

Since $\tilde{P}^{-1}\tilde{L}_B = \tilde{L}_B\tilde{P}^{-1}$, it follows that

$$\mathrm{Tr}\left(\tilde{P}^{-1}\left[\tilde{L}_B\tilde{Q}_0^\intercal\tilde{L}_B^{-1} - \tilde{Q}_0^\intercal\right]\right) = 0, \tag{5.19}$$

$$\mathrm{Tr}\left(\tilde{Q}_2\right) = \mathrm{Tr}\left[\tilde{P}^{-1}\left(\hat{M}\tilde{J}\tilde{L}_B^{-1} + \tilde{K}\right)\right]. \tag{5.20}$$

Finally, substituting $\mathrm{Tr}(\tilde{Q}_2)$ into (5.13) leads to

$$\|\tilde{\Sigma}\|_{\mathcal{H}_2}^2 = \frac{1}{\hat{M}^2}\,\mathrm{Tr}\left[\tilde{P}^{-1}\left(\hat{M}\tilde{J}\tilde{L}_B^{-1} + \tilde{K}\right)\right], \tag{5.21}$$

which concludes the proof. □

CHAPTER 6

STOCHASTIC DYNAMIC PROGRAMMING FOR WIND FARM POWER MAXIMIZATION [1]

Wind plants can increase annual energy production with advanced control algorithms by coordinating the operating points of individual turbine controllers across the farm. It remains a challenge to achieve performance improvements in practice because of the difficulty of utilizing models that capture pertinent complex aerodynamic phenomena while remaining amenable to control design. We formulate a multi-stage stochastic optimal control problem for wind farm power maximization and show that it can be solved analytically via dynamic programming. In particular, our model incorporates state- and input-dependent multiplicative noise whose distributions capture stochastic wind fluctuations. The optimal control policies and value functions explicitly incorporate the moments of these distributions, establishing a connection between wind flow data and optimal feedback control. We illustrate the results with numerical experiments.

6.1 Introduction

Wind energy is an important component of future energy systems to meet growing energy demands. As wind power continues to account for a larger portion of the world-wide energy portfolio, the optimal operation of wind farms offers both challenges and opportunities to further improve performance at the levels of single turbines, wind farms, and power grids. Due to nonlinear aerodynamic interaction through wakes and unpredictable wind variations,

[1]This chapter is based on work supported by the National Science Foundation (NSF) under grant CMMI-1728605, and partially funded by The University of Texas at Dallas Office of Research through the SCI program.

Chapter 6 in part is a reprint of material published in:

future optimal control strategies for wind farms will require sophisticated models to capture and manage *stochastic* wind fluctuations.

Maximizing the wind power capture has been discussed in the scope of wind turbines [203, 284, 204, 196] and wind farms [235, 140, 146, 244, 181, 241, 150, 132, 38, 41, 47, 46, 64, 233, 105, 110, 106, 229, 139, 289, 185, 206]. In Region 2 operation (below-rated wind speed), the wind plant is operated to maximize the power output. In this regime, there are inherent tradeoffs between the wake of upstream turbines and the power extracted from downstream turbines. Due to this aerodynamic coupling, maximizing total power of wind farms cannot be achieved by myopically maximizing the power output for each individual wind turbine in the array [245]. Therefore, depending on layout and wind conditions, it may be essential to have a coordinated control framework for wind farms to determine the optimal control strategy for each wind turbine to improve annual energy production.

Many challenges and related solutions for wind farm power maximization have been highlighted and discussed in [203]. Recent control strategies for optimal operation have been proposed using both model-based [41, 105, 110, 229, 140, 146, 244, 181, 241, 150, 132, 38], and model-free strategies [64, 185, 206, 106, 289]. Model-based strategies provide solutions that typically have faster response times than model-free approaches. However, the models used for control design can deviate from actual wind field and turbine characteristics in practice, which can limit the effectiveness of model-based control strategies. The reader is referred to the introduction in [64], and the references therein, for further discussion on model-based and model-free strategies for wind plant power maximization.

In this chapter, we focus on wind power maximization in Region 2. The work presented here generalizes the simple actuator disk model (ADM) utilized in [229] to a stochastic version and pose a multi-stage stochastic optimal control problem for wind farm power maximization. The stochastic actuator disk model balances complexity and tractability by incorporating unsteady aerodynamic phenomena into the distributions of random variables

in the model. Estimates of the statistics of these distributions can then be exploited in the control algorithm to improve overall efficiency of the farm in the presence of stochastic wind flow.

Our main contributions are as follows:

- We formulate a multi-stage stochastic optimal control problem for wind farm power maximization and show that it can be analytically solved via dynamic programming. In particular, our model generalizes that of [229] by incorporating state- and input-dependent multiplicative noises to capture the uncertain wake effects of wind turbines. The stochastic version of the ADM relaxes a strong assumption of a deterministic ADM, such as steady wind over the rotor disk. In contrast to existing work, the proposed stochastic multi-stage formulation allows us to maximize the wind farm power by explicitly incorporating information about the probability distributions of wind fluctuations into control decisions.

- By solving the proposed multi-stage stochastic optimization, we show that the optimal feedback control policies for the turbines are linear with respect to upstream wind velocity, but in contrast to [229], the optimal gain coefficients depend explicitly on the statistics of the multiplicative noises, which can be estimated from high-fidelity wind flow simulations or experimental data. This provides a direct connection between statistical properties of the unsteady wind flow physics and the optimal feedback control of wind farms. We also show that for the stochastic ADM with both multiplicative and additive noise, the optimal policies are nonlinear.

The framework, while elementary for real-world applications, illustrates a rigorous process for incorporating flow statistics into the wind farm power maximization problem. The dependence of control solutions on the statistics of the wind fluctuations makes intuitive

sense, as one cannot expect a single control algorithm to be optimal under a range of unsteady wind conditions. In future work, we will extend the stochastic approach presented in this chapter to more representative, yet tractable, models of the flow physics and loads as done in [233].

6.2 Problem Formulation

Our model is a generalization of the one in [229], which utilizes the actuator disk model (ADM) [48, 184]. Let P denote the power extracted by an ideal turbine rotor, let F denote the force done by the wind on the rotor, let V_0 denote the free stream upwind velocity, let V denote the wind velocity at the disk, and let V_1 denote the far wake velocity. The ADM model is then

$$P = FV, \tag{6.1a}$$

$$F = \rho A(V_0 - V_1)V, \tag{6.1b}$$

$$V = V_0 - u, \tag{6.1c}$$

$$V_1 = V_0 - 2u, \tag{6.1d}$$

where ρ is the air density, A is the rotor swept area, and $u \geq 0$ is the reduction in air velocity between the free stream and the rotor plane, which can be interpreted as a control input. In practice, u can be controlled by adjusting the angular rotor speed or the collective blade pitch angle.

Deterministic Model: We consider a one-dimensional cascade of wind turbines, illustrated in Fig. 6.1. We assume that the wind direction is along the row of turbines and is not varying. The ADM model given in (6.1) can be written in state-space form by letting x_k and x_{k+1} denote the wind velocity upstream and downstream of the k-th turbine (i.e., $x_k = V_k$ in (6.1d), for $k = 0, 1$). The scalar control input for the k-th turbine is denoted by u_k, which is the controllable wind velocity deficit at the rotor disk, and y_k is an output to

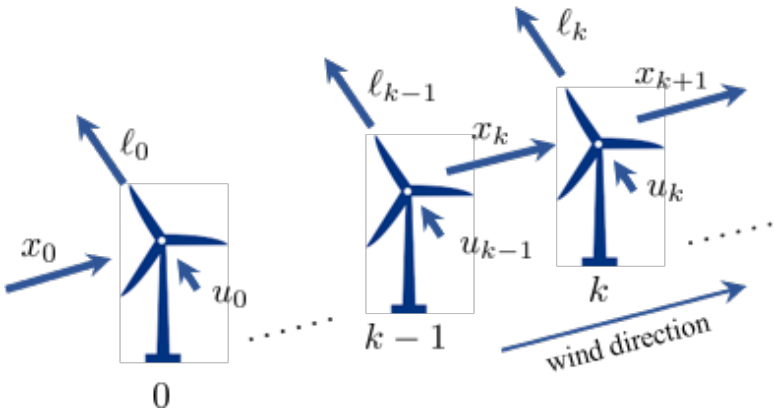


Figure 6.1. A cascade of N wind turbines; $k = 0$ indicates the most upstream location.

estimate the power extraction of turbine k (i.e., $y_k = V_k - u_k$ in (6.1c)). Then the velocity V_{k+1} in the far wake of the rotor (6.1d) and the rotor effect at the disk in velocity (6.1c) can be written as below in (6.2a) and (6.2b). The power extraction of the k-th wind turbine using ADM model (6.1) in state-space expression is given in (6.2c)

$$x_{k+1} = x_k - 2u_k, \tag{6.2a}$$

$$y_k = x_k - u_k, \tag{6.2b}$$

$$p_k(y_k, u_k) = 2\rho A y_k^2 u_k, \tag{6.2c}$$

where the control input is constrained by $u_k \in [0, \frac{1}{2}x_k]$ so that the wind velocity in the far wake remains positive. To simplify the notation, we eliminate the constant in (6.2c) and come to the constant-free turbine power function $\ell(x_k, u_k)$, which will serve as a stage cost in our subsequent multi-stage optimal control problem

$$\ell(x_k, u_k) = (x_k - u_k)^2 u_k. \tag{6.3}$$

Note that this function is jointly *cubic* in the state and control input. Further details of this model may be found in [229].

Stochastic Model: The simple model described above captures basic wind farm turbine interactions. But it fails to capture stochastic wind fluctuations that are also relevant

to optimizing the total power output. High fidelity computational fluid dynamic models offer extreme detail of flows but are cumbersome to incorporate into high-level operational decision making. Therefore, we consider here a stochastic extension of the deterministic actuator disk model above that can capture more complex phenomena, such as stochastic wind fluctuations, while remaining computationally tractable.

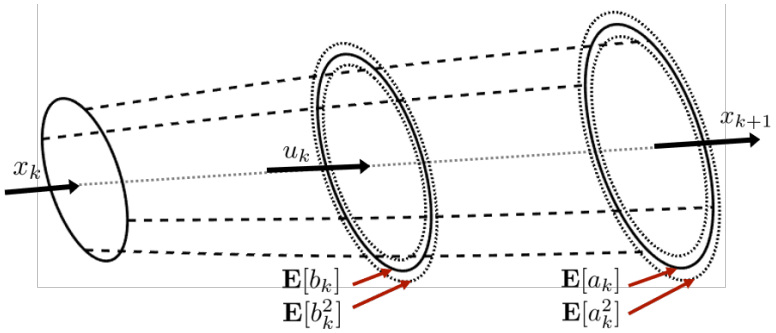


Figure 6.2. Stochastic actuator disk model and stream-tube diagram for wind power extraction. The solid and dashed lines indicate the wind field mean and associated stochastic variations, respectively, which relate to the moments of the multiplicative variations parameters a_k and b_k.

$$\psi_k = -\frac{3Q_{k+1}\Sigma_{b,k}\mu_{a,k} - 2 + \sqrt{(3Q_{k+1}\Sigma_{b,k}\mu_{a,k} - 2)^2 - 3(Q_{k+1}\Gamma_{b,k} + 1)(3Q_{k+1}\Sigma_{a,k}\mu_{b,k} + 1)}}{3(Q_{k+1}\Gamma_{b,k} + 1)}, \tag{6.4a}$$

$$Q_k = (1 - \psi_k)^2\psi_k + Q_{k+1}\left(\Gamma_{a,k} + \Gamma_{b,k}\psi_k^3 + 3\Sigma_{b,k}\mu_{a,k}\psi_k^2 + 3\Sigma_{a,k}\mu_{b,k}\psi_k\right). \tag{6.4b}$$

The stochastic actuator disk model is given by

$$x_{k+1} = a_k x_k + b_k u_k, \tag{6.5}$$

where $a_k \sim \mathcal{P}_{a,k}$ is a state multiplicative random variable and $b_k \sim \mathcal{P}_{b,k}$ is an input multiplicative random variable. The model is illustrated in Fig. 6.2. We assume that the random variables a_k and b_k are independent for all k and independent of each other. This model captures stochastic wind fluctuations. In particular, the multiplicative noises a_k and b_k provide

a simple model for the inherent stochasticity of far wake recovery. We assume that moments up to order three of each of the distributions $\mathcal{P}_{a,k}, \mathcal{P}_{b,k}$ are known (or can be estimated from high-fidelity simulation or experimental data). For the state mean dynamics to match the deterministic model (6.2a), we can set $\mathbf{E}[a_k] = 1$, $\mathbf{E}[b_k] = -2$.

6.3 Stochastic Optimal Control for Wind Power Maximization

The objective of the operator is to select control inputs $u_0, ..., u_{N-1}$ to maximize the aggregate power of the wind turbine cascade given by the sum of (6.3) over all turbines. However, since in the stochastic model the states (and therefore the power outputs) are random variables, we maximize the *expected* aggregate power and search for closed-loop feedback control policies that specify control inputs as a function of the state x_k. In particular, we seek to solve the multi-stage stochastic optimal control problem

$$\max_{\pi_0,\ldots,\pi_{N-1}} \quad \mathbf{E}\sum_{k=0}^{N-1}(x_k - u_k)^2 u_k, \tag{6.6}$$

where the decision variables $\pi_k(\cdot)$ are the control policies (i.e., $u_k = \pi_k(x_k)$), and the expectation is taken with respect to the random variable sequences a_k, b_k. As in [229], we will show that the optimal policies are linear and the optimal value functions are *cubic*. In contrast to [229], the parameters of both the optimal policies and value functions depend on the moments of the distribution of the random variables in the model. We have the following main result.

Theorem 4. *Consider the wind farm power maximization problem for a cascade of N identical turbines modeled with the stochastic actuator disk model* (6.5), (6.6). *Let x_0 denote the free stream velocity entering the cascade. The distributions of a_k and b_k are described by their raw moments up to third order, namely their means $\mu_{a,k}$, $\mu_{b,k}$, second (raw) moments $\Sigma_{a,k}$, $\Sigma_{b,k}$ and third (raw) moments $\Gamma_{a,k}$, $\Gamma_{b,k}$. Under these assumptions, the optimal feedback*

control policies are linear in the state and given by

$$u_k^* = \pi^*(x_k) = \psi_k x_k, \quad k = 1, ..., N-1, \tag{6.7}$$

where the gain coefficients ψ_k *are given in* (6.4a) *and the backwards recursion* (6.4b) *for* $k = N-1, ..., 0$ *with initialization* $Q_N = 0$*. The maximum power produced by the wind farm as a function of initial upstream wind velocity is given by*

$$P_0^*(x_0) = 2\rho A Q_0 x_0^3, \tag{6.8}$$

where ρ *is the air density,* A *is the rotor swept area, and* Q_0 *is the initial value of the backwards recursion* (6.4b) *with* $Q_N = 0$*.*

Proof. The dynamic programming algorithm [24, 33] for solving stochastic optimal control problems is given by the recursion

$$\begin{aligned} G_k^*(x_k) &= \max_{u_k \in [0, \frac{1}{2}x_k]} \mathbf{E}\left\{\ell(x_k, u_k) + G_{k+1}^*\left(x_{k+1}\right)\right)\right\}, \\ \pi^*(x_k) &= \arg \max_{u_k \in [0, \frac{1}{2}x_k]} \mathbf{E}\left\{\ell(x_k, u_k) + G_{k+1}^*\left(x_{k+1}\right)\right)\right\}, \end{aligned} \tag{6.9}$$

where $G_k^*(x_k)$ represents the optimal (normalized) wind farm power from turbine k as a function of the state x_k, with initialization $G_N^*(x_k) = 0$. We first solve the last tail subproblem at $k = N-1$ with $G_N^*(x) = 0$. We have

$$\frac{\partial \ell(x_{N-1}, u_{N-1})}{\partial u_{N-1}} = (x_{N-1} - u_{N_1})(x_{N-1} - 3u_{N-1}) = 0,$$

for which the policy $u_{N-1}^* = \frac{1}{3}x_{N-1}$ is the unique maximizer and satisfies the constraint $u_{N-1} \in [0, \frac{1}{2}x_{N-1}]$. Substituting this optimal policy back into the value expression yields the optimal power function

$$G_{N-1}^*(x_{N-1}) = \frac{4}{27}x_{N-1}^3.$$

Note that this function is a cubic in the state. Accordingly, we parameterize the optimal power functions as $G_k^*(x_k) = Q_k x_k^3$ and consider a general step in the backward recursion. To obtain the optimal policy, we define the function inside the maximization operation

$$G_k(x_k, u_k) := (x_k - u_k)^2 u_k + Q_{k+1} \mathbf{E}\left[(a_k x_k + b_k u_k)^3\right]. \tag{6.10}$$

Expanding the second term and taking the expectation by utilizing the (raw) moment information from the distributions of a_k and b_k, and then taking the partial derivative of $G_k(x_k, u_k)$ with respect to u_k gives a quadratic polynomial in u_k. As above, one of the roots of this polynomial corresponds to the unique maximizing input, which is a *linear* function of the state. Carrying out the algebra yields

$$u_k^* = \pi^*(x_k) = \psi_k x_k, \tag{6.11}$$

where the gain parameters ψ_k are given in (6.4a). Note that the optimal policies all satisfy the constraints on u_k. To obtain a backwards recursion for the value function coefficients Q_k, we substitute $u_k^* = \psi_k x_k$ back into (6.10)

$$\begin{aligned} G_k^*(x_k, u_k^*) &= Q_k x_k^3 \\ &= (x_k - u_k^*)^2 u_k^* + Q_{k+1} \mathbf{E}\left[(a_k x_k + b_k u_k^*)^3\right]. \end{aligned} \tag{6.12}$$

Since u_k^* is linear in x_k, the optimal value functions are cubic in the state. Matching the coefficients on both sides of (6.12), we come to (6.4b). Eq. (6.8) follows from (6.12) for $k = 0$, (6.6) and (6.2c), which concludes the proof. □

Remark 9. *(Optimal policies and value functions with central moments.) The random variables a_k and b_k can also be described by their higher-order central moments, namely their variances $\sigma_{a,k}^2$, $\sigma_{b,k}^2$ and skewnesses $\gamma_{a,k}$, $\gamma_{b,k}$. The optimal linear state feedback control policies can also be written in terms of central moments instead of raw moments by using*

$$\Sigma = \sigma^2 + \mu^2, \quad \Gamma = \sigma^3 \gamma + 3\sigma^2 \mu + \mu^3. \tag{6.14}$$

$$\pi^*_{N-2}(x) =$$
$$-\frac{\Delta_k + \sqrt{\Delta^2_{N-2} - 3(Q_{N-1}\Gamma_{b,N-2}+1)\big[\,(3Q_{N-1}\Sigma_{a,N-2}\mu_{b,N-2}+1)\,x^2 + 3Q_{N-1}\Sigma_{c,N-2}\mu_{b,N-2}\big]}}{3(Q_{N-1}\Gamma_{b,N-2}+1)},$$
where, $\Delta_{N-2} = (3Q_{N-1}\Sigma_{b,N-2}\mu_{a,N-2} - 2)x.$ (6.13)

Corollary 2. *Under the assumptions of Theorem 1, we define the efficiency η_ℓ of the ℓ-th sub-array*[2] *by*

$$\eta_\ell := \mathbf{E}\left[\frac{P_\ell}{\frac{1}{2}\rho A x_\ell^3}\right], \tag{6.15}$$

where x_ℓ is the free stream velocity entering the subarray cascaded turbines from ℓ to $N-1$ and P_ℓ denotes the aggregated power from the ℓ-th subarray of wind turbines. The optimal efficiency η^*_ℓ of the l-th sub-array has the form

$$\eta^*_\ell = 4Q_\ell, \quad \forall \ell \in \{0, \ldots, N-1\}, \tag{6.16}$$

which is achieved with the optimal control sequence $u^*_\ell, \ldots, u^*_{N-1}$, where Q_ℓ is calculated from (6.4b).

Proof. The maximum power produced by the $N-\ell$ turbines is

$$P^*_\ell = 2\rho A Q_\ell x^3_\ell, \tag{6.17}$$

under the optimal control sequence $u^*_\ell, \ldots, u^*_{N-1}$ with Q_ℓ computed via (6.4b). We substitute the optimal power (6.17) into (6.15) and obtain (6.16), which concludes the proof. □

Next, we consider a stochastic actuator disk model with both multiplicative and additive noise, which allows a more general description of uncertainty in wind fluctuations. Interestingly, in contrast to classical linear quadratic problems, when additive noise is included the

[2]The efficiency η_ℓ defined here quantifies the energy extraction of sub-array ℓ compared to energy in the wind entering the sub-array. Note that due to aerodynamic wake coupling, it is possible for the optimal efficiency of the sub-array to exceed the efficiency obtained by independently setting individual turbine induction factors to achieve the single-turbine Betz limits.

optimal policies are no longer linear in general, and so the optimal value functions are no longer cubic. This highlights a computational limitation with this more general model that makes the approach more difficult to implement in practice.

Theorem 5. *(Stochastic actuator disk model with additive noise.) Consider the stochastic ADM* (6.5) *with additive noise*

$$x_{k+1} = a_k x_k + b_k u_k + c_k, \tag{6.18}$$

where $c_k \sim \mathcal{P}_c$ is a zero-mean additive random variable with second moment $\Sigma_{c,k}$ and third moment $\Gamma_{c,k}$. In the penultimate tail subproblem, the optimal policy has the nonlinear form

$$\pi^*_{N-2}(x) = \delta x + \sqrt{\alpha + \beta x^2}$$

for some constants δ, α, and β; the exact expression is given in (6.13)*. As a result, the corresponding optimal value function at turbine location $N-2$ is non-cubic, and so the remaining optimal policies and value functions are nonlinear and non-cubic, respectively.*

Proof. Consider again the dynamic programming recursion (6.9). Since $G^*_N(x) = 0$, the last tail subproblem is identical to that in Theorem 1, so that $G^*_{N-1}(x_{N-1}) = \frac{4}{27}x^3_{N-1}$. Consider now the penultimate tail subproblem for $k = N-2$

$$G_{N-2}(x,u) = (x-u)^2 u + \frac{4}{27}\mathbf{E}\left[(a_k x + b_k u + c_k)^3\right]. \tag{6.19}$$

Taking the expectation of the second term by utilizing the (raw) moments of a_k, b_k and c_k, and then taking the partial derivative with respect to u and setting to zero yields a quadratic optimality condition in u. Carrying out some algebra as above, it turns out that the roots of this polynomial are no longer linear in the state, in contrast to the results in Theorem 1. The optimal control policy is thus a *nonlinear* function of state of the form $\pi^*_{N-2}(x) = \delta x + \sqrt{\alpha + \beta x^2}$ for some constants δ, α, and β. The exact expression for the

maximizing control input derived from the quadratic optimality condition is given in (6.13). It can also be seen that when the additive noise variance $\Sigma_{c,k}$ is zero (i.e., the additive noise is absent since it also has zero mean), then $\alpha = 0$ and we recover the linear policy of of Theorem 1 since $x \geq 0$. Finally, these observations also lead to the conclusion that none of the remaining optimal policies and value functions are linear and cubic, respectively, and will in fact become increasingly complicated as the recursion proceeds backward toward the beginning of the array. □

6.4 Numerical Experiments

To illustrate our results, we consider a cascade with $N = 10$ identical turbines to analyze the performance of the optimal gain sequence $\{\psi_0, \ldots, \psi_9\}$ for the proposed stochastic actuator disk model. As is commonly done in the literature [48, 184], we refer to these gains as *induction factors*. The stochastic model parameters a_k and b_k are all independent of each other and spatially homogeneous ($\mu_{a,k} = \mu_a$, $\mu_{b,k} = \mu_b$, $\sigma_{a,k} = \sigma_a$, $\sigma_{b,k} = \sigma_b$, $\gamma_{a,k} = \gamma_a$ and $\gamma_{b,k} = \gamma_b$, $\forall k$)[3].

Fig. 6.3 illustrates the optimal induction factor sequence $\{\psi_0, \ldots, \psi_9\}$ and Fig. 6.4 depicts the optimal efficiency η_ℓ^* under different standard deviation values of the *input-dependent* multiplicative noise b_k. The induction factors in Fig. 6.3 are normalized by 1/3, which is the value achieving the Betz limit for a single isolated turbine [48]. We set the mean value μ_a to 1, and the skewness to zero. Fig. 6.4 demonstrates that the optimal array efficiency improves with increasing variance on b_k. This result is intuitively reasonable, in the sense that higher variability of the velocity deficits in the far wake may lead to increased power extraction. We speculate that this multiplicative stochastic perturbation on the velocity deficit may provide a mathematically simple way of capturing physical phenomena

[3]To have clearer interpretation of our results, we discuss the results in the terms of central moments. No additive noise is considered in this section.

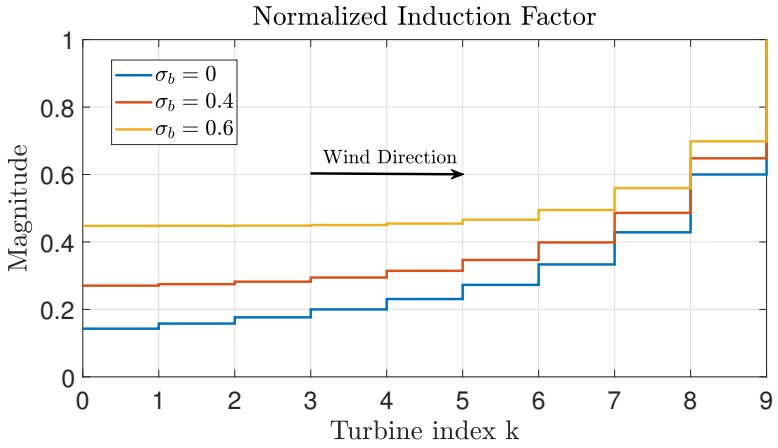


Figure 6.3. Normalized induction factors defined as $\frac{\psi_k}{1/3}$ for deterministic model ($\mu_a = 1, \mu_b = -2$) and stochastic model with various values of input-dependent multiplicative noise standard deviation ($\mu_a = 1, \sigma_a = 0, \mu_b = -2, \sigma_b > 0, \gamma_a = 0$ and $\gamma_b = 0$).

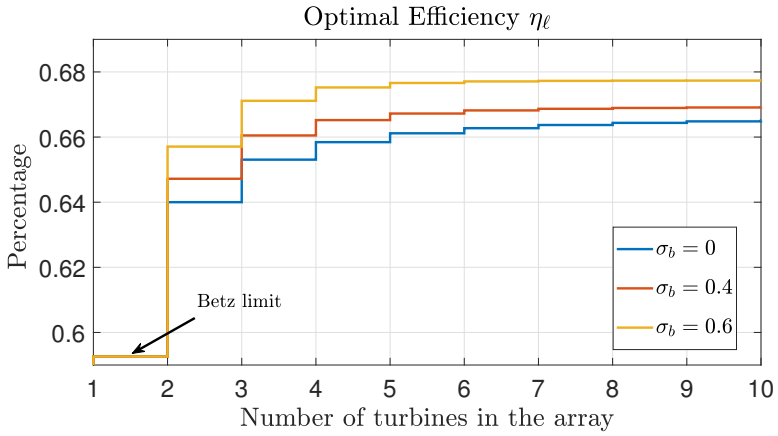


Figure 6.4. Comparison of optimal efficiency η_ℓ for deterministic model ($\mu_a = 1, \mu_b = -2$) and stochastic model with various values of input-dependent multiplicative noise standard deviation ($\mu_a = 1, \sigma_a = 0, \mu_b = -2, \sigma_b > 0, \gamma_a = 0$ and $\gamma_b = 0$).

such as mixing or entrainment, which are known to promote energy extraction [257, 234] Note also that the case $\sigma_b = 0$ reproduces the results for the deterministic ADM model in [229]. It should be noted that as the standard deviation σ_b is increased, the induction factors increase. That is, the leading upstream turbines are working more as the multiplicative noise is increased; which again is consistent with the conventional wisdom that the more turbulent the wind is the closer the turbines should be to their isolated optimum set point [63]

Figs 6.5 and 6.6 provide the optimal induction factor sequence and efficiency under different standard deviation values of the *state-dependent* noise on a_k, and without input-dependent noise (i.e., b_k is fixed and constant for all k). Both figures demonstrate that the optimal induction factor sequence from the stochastic actuator disk model also increases the efficiency and improves performance under larger variations.

To match the expected wind velocity of the conventional deterministic ADM, the mean value of the *state-dependent* noise should be set to unity (i.e., $\mu_a = 1$) [229]. However, having $\mu_a = 1$ together with a non-zero variance in the state-dependent multiplicative noise a_k leads to null optimal induction factors for leading upstream turbines, since in this case the model essentially predicts that additional energy will be injected into the wake further downstream. This indicates that the parameters in the stochastic ADM should be carefully calibrated based on measured data in order to capture appropriate (possibly heterogeneous) spatio-temporal flow variations and obtain reasonable control policies for the array. To appropriately incorporate stochasticity of the wind flow, we set the mean value of a_k to $\mu_a = 0.99$, and vary the standard deviation σ_a to describe statistical fluctuations. The key observation is that regardless of the value of μ_a, the proposed approach improves efficiency with increasing variance by exploiting statistical knowledge of wind field fluctuations and incorporating this information into optimal control policies for wind farm power maximization.

The stochastic actuator disk model of a wind farm with cascaded wind turbines captures stochastic wind fluctuations. By definition, the optimal control laws derived from stochastic

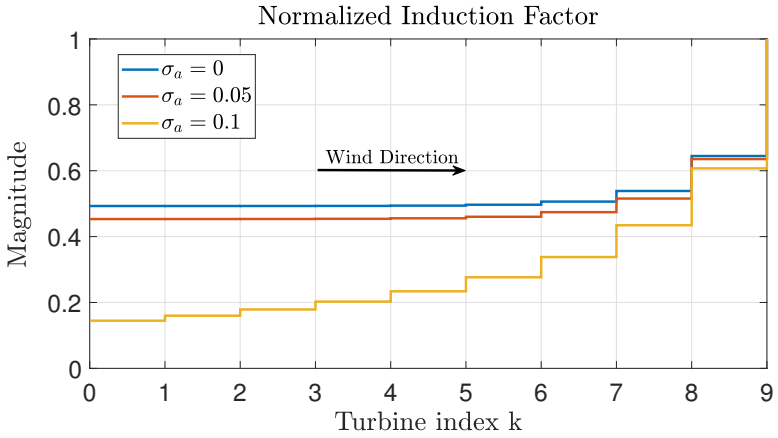


Figure 6.5. Normalized induction factor defined as $\frac{\psi_k}{1/3}$ for deterministic model ($\mu_a = 0.99, \mu_b = -2$) and for stochastic model with various values of state-dependent multiplicative noise standard deviation ($\mu_a = 0.99, \sigma_a > 0, \mu_b = -2, \sigma_b = 0, \gamma_a = 0$ and $\gamma_b = 0$).

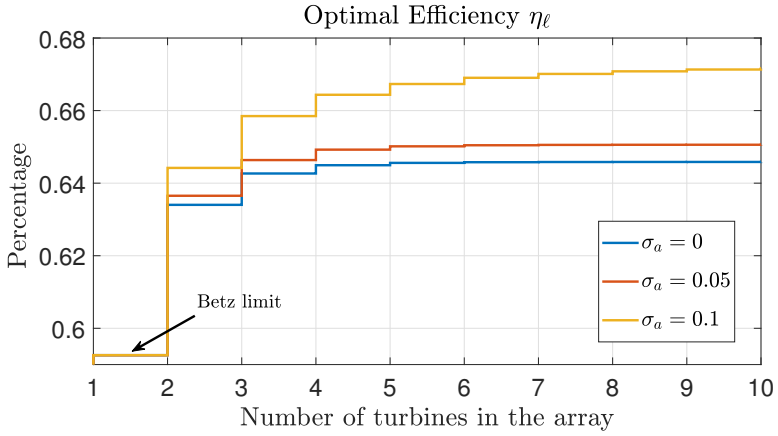


Figure 6.6. Comparison of optimal efficiency η_ℓ for deterministic model ($\mu_a = 0.99, \mu_b = -2$) and stochastic model with various values of state-dependent multiplicative noise standard deviation ($\mu_a = 0.99, \sigma_a > 0, \mu_b = -2, \sigma_b = 0, \gamma_a = 0$ and $\gamma_b = 0$).

dynamic programming achieve superior performance to laws derived from a deterministic model of the same complexity, allowing the turbines to recognize and react to the particular wind field characteristics. Data derived directly from measurements or simulations can be incorporated directly into the control law to improve the aerodynamic efficiency a wind farm for specific wind fluctuation statistics. It is worth emphasizing that more work is necessary to incorporate tractable noise models that are consistent with the flow physics; this chapter is a first step in this direction.

6.5 Conclusions and Outlooks

We have formulated a multi-stage stochastic optimal control problem for maximizing the power output of a one dimensional wind farm array and shown that it can be solved analytically via dynamic programming. The optimal control policies depend explicitly on the statistics of multiplicative noise, which can be related to stochastic wind fluctuations.

Our results provide an initial step toward defining a wind farm control strategy that tractably incorporates statistical knowledge of stochastic wind fluctuations. However, there remain several lines of future work that can extend the present results in various ways to more fully understand the possibilities and limits for maximizing annual energy production. Our future work will involve

(a) utilizing more realistic wake models;

(b) estimating necessary statistics from high-fidelity numerical simulations and experimental data;

(c) performance evaluation of the policies on high-fidelity models, which may improve the results in [234];

(d) considering more realistic array geometries;

(e) exploring computationally efficient approximation of nonlinear optimal control strategies; if needed.

CHAPTER 7

SUMMARY AND CONCLUSIONS

In this dissertation, we discuss the challenges of having stochastic optimization and optimal control in complex networked systems with large variation uncertainties. An important challenge is how to design noise-aware control strategies with limited information of uncertainties, which is able to successfully tradeoff the system performance, operation efficiency, robustness and complexity. To address this problem, it is crucial to incorporate the probability distribution information into performance analysis and the optimal control policies for various objectives.

In Chapter 2 and Chapter 3, we formulated a data-based distributionally robust stochastic optimal control problem to have optimal scheduling in power systems and water networks. The proposed optimal control strategies are completely based on finite training dataset of uncertainties (i.e., renewables and water demands), instead of assuming the prescribed probability distributions. The main contribution of our approach is that the conservativeness of optimal decisions is explicitly controllable, which successfully tradeoffs the efficiency of system operations, risk of operational constraint violations, inherent sampling errors within training dataset and the out-of-sample performance.

In Chapter 4, we presented an optimal power flow solver in power systems with state estimation feedback to promote the practical implementation of existing OPF formulations for the extreme large-scale distribution networks (i.e., 11,000-nodes), in the case without reliable and efficient measurement infrastructure. We analytically demonstrated the convergence of the proposed gradient-based algorithm and numerically demonstrated the robust performance to the measurement noises and estimation errors.

In Chapter 5 and Chapter 6, we considered the linear systems with multiplicative noises to incorporate the probability distribution information of modelling errors for performance

analysis and optimal feedback controller design. A closed-form mean-square stability criteria was established for low-inertia power grids, as function of network topology, moments information of multiplicative noises and system parameters. The optimal control sequence were developed for wind farm power maximization, which explicitly connected the wind flow data and optimal feedback control.

In general, this dissertation developed a data-based noise-aware optimal control framework via the distributionally robust optimization and the feedback control techniques. The optimal scheduling and feedback control policies were established for the vital municipal networked infrastructures (i.e., power systems, water networks and wind farms), based on the limited probability distribution information extracted from the finite sampling dataset of uncertainties.

REFERENCES

[1] (2012). Global energy forecasting competition 2012 wind energy competition.

[2] (2014). National solar radiation databased (NSRDB).

[3] Abido, M. (2002). Optimal power flow using particle swarm optimization. *International Journal of Electrical Power & Energy Systems 24*(7), 563–571.

[4] Abur, A. and A. G. Exposito (2004). *Power system state estimation: theory and implementation*. CRC press.

[5] Ahmad, F., A. Rasool, E. Ozsoy, R. Sekar, A. Sabanovic, and M. Elitaş (2018). Distribution system state estimation-a step towards smart grid. *Renewable and Sustainable Energy Reviews 81*, 2659–2671.

[6] Ahmadyar, A. S., S. Riaz, G. Verbič, A. Chapman, and D. J. Hill (2018). A framework for assessing renewable integration limits with respect to frequency performance. *IEEE Trans. on Power Systems 33*(4), 4444–4453.

[7] Ahmadyar, A. S., S. Riaz, G. Verbič, J. Riesz, and A. Chapman (2016, Sept.). Assessment of minimum inertia requirement for system frequency stability. In *IEEE International Conference on Power System Technology*, pp. 1–6.

[8] Al-Hemairi, H. and R. Shakir (2006). Minimizing leakage rates in water distribution networks through optimal valves settings. In *World Environmental and Water Resource Congress 2006: Examining the Confluence of Environmental and Water Concerns*, pp. 1–13.

[9] Alsac, O. and B. Stott (1974, May). Optimal load flow with steady-state security. *IEEE Transactions on Power Apparatus and Systems* (3), 745–751.

[10] Anderson, E. J. and K. H. Al-Jamal (1995). Hydraulic-network simplification. *Journal of water resources planning and management 121*(3), 235–240.

[11] ApS, M. (2017). *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*

[12] Archibald, T. W. and S. E. Marshall (2018). Review of mathematical programming applications in water resource management under uncertainty. *Environmental Modeling & Assessment 23*(6), 753–777.

[13] Arnold, L. (1974). Stochastic differential equations. *New York*.

[14] Arrow, K. J., S. Karlin, H. E. Scarf, et al. (1958). Studies in the mathematical theory of inventory and production.

[15] Bai, Z. and D. Skoogh (2006). A projection method for model reduction of bilinear dynamical systems. *Linear algebra and its applications 415*(2-3), 406–425.

[16] Baker, K., E. Dall'Anese, and T. Summers (2016, Sept.). Distribution-agnostic stochastic optimal power flow for distribution grids. In *IEEE North American Power Symposium*, pp. 1–6.

[17] Baker, K., G. Hug, and X. Li (2017). Energy storage sizing taking into account forecast uncertainties and receding horizon operation. *IEEE Transactions on Sustainable Energy 8*(1), 331–340.

[18] Baldick, R., B. H. Kim, C. Chase, and Y. Luo (1999, Aug.). A fast distributed implementation of optimal power flow. *IEEE Transactions on Power Systems 14*(3), 858–864.

[19] Bank, J. and J. Hambrick (2013). Development of a high resolution, real time, distribution-level metering system and associated visualization, modeling, and data analysis functions. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO.

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

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

[22] Baran, M. E. and F. F. Wu (1989c). Optimal capacitor placement on radial distribution systems. *IEEE Transactions on Power Delivery 4*(1), 725–734.

[23] Beck, H.-P. and R. Hesse (2007). Virtual synchronous machine. In *9th IEEE International Conference on Electrical Power Quality and Utilisation*, pp. 1–6.

[24] Bellman, R. (1954). The Theory of Dynamic Programming. Technical report, RAND Corp Santa Monica CA.

[25] Ben-Tal, A., D. Den Hertog, and J.-P. Vial (2015). Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical programming 149*(1-2), 265–299.

[26] Benaroya, H., S. M. Han, S. Han, and M. Nagurka (2005). *Probability models in engineering and science*. CRC press.

[27] Benner, P. and T. Damm (2011). Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. *SIAM Journal on Control and Optimization 49*(2), 686–711.

[28] Bergen, A. R. and D. J. Hill (1981). A structure preserving model for power system stability analysis. *IEEE Trans. on Power Apparatus and Systems* (1), 25–35.

[29] Bernstein, A. and E. Dall'Anese (2017a, Sept.). Linear power-flow models in multiphase distribution networks. In *IEEE PES Innovative Smart Grid Technologies Conference Europe*.

[30] Bernstein, A. and E. Dall'Anese (2017b). Linear power-flow models in multiphase distribution networks. In *IEEE PES Innovative Smart Grid Technologies Conference Europe*.

[31] Bernstein, A., L. Reyes-Chamorro, J.-Y. Le Boudec, and M. Paolone (2015). A composable method for real-time control of active distribution networks with explicit power setpoints. Part I: Framework. *Electric Power Systems Research 125*, 254–264.

[32] Bernstein, A., C. Wang, E. DallAnese, J.-Y. Le Boudec, and C. Zhao (2018). Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models. *IEEE Transactions on Power Systems 33*(6), 5832–5843.

[33] Bertsekas, D. P., D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas (2005). *Dynamic Programming and Optimal Control*, Volume 1. Athena scientific Belmont, MA.

[34] Bertsimas, D., X. V. Doan, K. Natarajan, and C.-P. Teo (2010). Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research 35*(3), 580–602.

[35] Bertsimas, D., V. Gupta, and N. Kallus (2017). Robust sample average approximation. *Mathematical Programming*, 1–66.

[36] Bertsimas, D., V. Gupta, and N. Kallus (2018). Robust sample average approximation. *Mathematical Programming 171*(1-2), 217–282.

[37] Bevrani, H., T. Ise, and Y. Miura (2014). Virtual synchronous generators: A survey and new perspectives. *International Journal of Electrical Power & Energy Systems 54*, 244–254.

[38] Biegel, B., D. Madjidian, V. Spudić, A. Rantzer, and J. Stoustrup (2013, Aug.). Distributed low-complexity controller for wind power plant in derated operation. In *IEEE International Conference on Control Applications*, pp. 146–151.

[39] Bienstock, D., M. Chertkov, and S. Harnett (2014). Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review 56*(3), 461–495.

[40] Bismut, J.-M. (1976). Linear quadratic optimal stochastic control with random coefficients. *SIAM Journal on Control and Optimization 14*(3), 419–444.

[41] Bitar, E. and P. Seiler (2013, June). Coordinated control of a wind turbine array for power maximization. In *American Control Conference*, pp. 2898–2904.

[42] Bolognani, S. and F. Dörfler (2015, Oct.). Fast power system analysis via implicit linearization of the power flow manifold. In *IEEE Annual Allerton Conference on Communication, Control, and Computing*, pp. 402–409.

[43] Boyd, S., L. El Ghaoui, E. Feron, and V. Balakrishnan (1994). *Linear matrix inequalities in system and control theory*, Volume 15. SIAM.

[44] Brentan, B. M., G. L. Meirelles, D. Manzi, and E. Luvizotto (2018). Water demand time series generation for distribution network modeling and water demand forecasting. *Urban Water Journal 15*(2), 150–158.

[45] Breuer, T. and I. Csiszár (2016). Measuring distribution model risk. *Mathematical Finance 26*(2), 395–411.

[46] Buccafusca, L., C. Beck, and G. Dullerud (2017, Aug.). Modeling and maximizing power in wind turbine arrays. In *IEEE Conference on Control Technology and Applications*, pp. 773–778.

[47] Buccafusca, L. and C. L. Beck (2018, Dec.). Maximizing power in wind turbine arrays with variable wind dynamics. In *IEEE Conference on Decision and Control*, pp. 2667–2672.

[48] Burton, T., D. Sharpe, and N. Jenkins (2001). *Handbook of Wind Energy*. John Wiley & Sons.

[49] Calafiore, G. C. and L. El Ghaoui (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications 130*(1), 1–22.

[50] Callaway, D. S. and I. A. Hiskens (2010). Achieving controllability of electric loads. *Proceedings of the IEEE 99*(1), 184–199.

[51] Camacho, E. F., T. Samad, M. Garcia-Sanz, and I. Hiskens (2011). Control for renewable energy and smart grids. *The Impact of Control Technology, Control Systems Society*, 69–88.

[52] Cannon, M., B. Kouvaritakis, and X. Wu (2009). Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica 45*(1), 167–172.

[53] Capitanescu, F., S. Fliscounakis, P. Panciatici, and L. Wehenkel (2012, Nov.). Cautious operation planning under uncertainties. *IEEE Transactions on Power Systems 27*(4), 1859–1869.

[54] Carpentier, J. (1962). Contribution to the economic dispatch problem. *Bulletin de la Societe Francaise des Electriciens 3*(8), 431–447.

[55] Castelletti, A., F. Pianosi, and R. Soncini-Sessa (2012). Stochastic and robust control of water resource systems: Concepts, methods and applications. In *System Identification, Environmental Modelling, and Control System Design*, pp. 383–401. Springer.

[56] Chen, S., X. Li, and X. Y. Zhou (1998). Stochastic linear quadratic regulators with indefinite control weight costs. *SIAM Journal on Control and Optimization 36*(5), 1685–1702.

[57] Chen, W., M. Sim, J. Sun, and C.-P. Teo (2010). From cvar to uncertainty set: Implications in joint chance-constrained optimization. *Operations research 58*(2), 470–485.

[58] Chiang, H.-D. (2011). *Direct methods for stability analysis of electric power systems: theoretical foundation, BCU methodologies, and applications*. John Wiley & Sons.

[59] Chiang, H.-D. and M. E. Baran (1990). On the existence and uniqueness of load flow solution for radial distribution power networks. *IEEE Transactions on Circuits and Systems 37*(3), 410–416.

[60] Chiang, H.-D., F. Wu, and P. Varaiya (1987). Foundations of direct methods for power system transient stability analysis. *IEEE Trans. on Circuits and systems 34*(2), 160–173.

[61] Christakou, K., J.-Y. LeBoudec, M. Paolone, and D.-C. Tomozei (2013, June). Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks. *IEEE Transactions on Smart Grid 4*(2), 741–750.

[62] Christie, R. D., B. F. Wollenberg, and I. Wangensteen (2000). Transmission management in the deregulated environment. *Proceedings of the IEEE 88*(2), 170–195.

[63] Ciri, U., M. Rotea, C. Santoni, and S. Leonardi (2017). Large-eddy simulations with extremum-seeking control for individual wind turbine power optimization. *Wind Energy 20*(9), 1617–1634.

[64] Ciri, U., M. A. Rotea, and S. Leonardi (2017). Model-free control of wind farms: A comparative study between individual and coordinated extremum seeking. *Renewable energy 113*, 1033–1045.

[65] Cohen, D., U. Shamir, and G. Sinai (2000). Optimal operation of multi-quality water supply systems-II: The QH model. *Engineering Optimization A35 32*(6), 687–719.

[66] Colombino, M., E. Dall'Anese, and A. Bernstein (2019). Online optimization as a feedback controller: Stability and tracking. *IEEE Transactions on Control of Network Systems*.

[67] Colombino, M., D. Groß, J.-S. Brouillon, and F. Dörfler (2019). Global phase and magnitude synchronization of coupled oscillators with application to the control of grid-forming power inverters. *IEEE Trans. on Automatic Control*.

[68] Conejo, A. J., M. Carrión, and J. M. Morales (2010). *Decision making under uncertainty in electricity markets*, Volume 1. Springer.

[69] Coppens, P., M. Schuurmans, and P. Patrinos (2020). Data-driven distributionally robust lqr with multiplicative noise. In *Learning for Dynamics and Control*, pp. 521–530. PMLR.

[70] Coulson, J., J. Lygeros, and F. Dörfler (2020). Distributionally robust chance constrained data-enabled predictive control. *arXiv preprint arXiv:2006.01702*.

[71] Curi, S., D. Groß, and F. Dörfler (2017). Control of low-inertia power grids: A model reduction approach. In *56th IEEE Annual Conference on Decision and Control*, pp. 5708–5713.

[72] Dall'Anese, E., K. Baker, and T. Summers (2017, Sept.). Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems 32*(5), 3427–3438.

[73] Dall'Anese, E., H. Zhu, and G. B. Giannakis (2013, Sept.). Distributed optimal power flow for smart microgrids. *IEEE Transactions on Smart Grid 4*(3), 1464–1475.

[74] DallAnese, E. and A. Simonetto (2016). Optimal power flow pursuit. *IEEE Transactions on Smart Grid 9*(2), 942–952.

[75] D'Ambrosio, C., A. Lodi, S. Wiese, and C. Bragalli (2015). Mathematical programming techniques in water network optimization. *European Journal of Operational Research 243*(3), 774–788.

[76] Damm, T. (2004). *Rational matrix equations in stochastic control*, Volume 297. Springer Science & Business Media.

[77] Damm, T. (2008). Direct methods and adi-preconditioned krylov subspace methods for generalized lyapunov equations. *Numerical Linear Algebra with Applications 15*(9), 853–871.

[78] D'Arco, S. and J. A. Suul (2013). Virtual synchronous machines classification of implementations and analysis of equivalence to droop controllers for microgrids. In *IEEE Grenoble PowerTech*, pp. 1–7.

[79] Dehghanpour, K., Z. Wang, J. Wang, Y. Yuan, and F. Bu (2018). A survey on state estimation techniques and challenges in smart distribution systems. *IEEE Transactions on Smart Grid 10*(2), 2312–2322.

[80] Dehghanpour, K., Y. Yuan, Z. Wang, and F. Bu (2019). A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation. *IEEE Transactions on Smart Grid*.

[81] Delage, E. and Y. Ye (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research 58*(3), 595–612.

[82] Dhople, S. V., S. S. Guggilam, and Y. C. Chen (2015). Linear approximations to ac power flow in rectangular coordinates. In *IEEE Allerton Conference on Communication, Control, and Computing*, pp. 211–217.

[83] Ding, H., Z. Hu, and Y. Song (2015). Value of the energy storage system in an electric bus fast charging station. *Applied Energy 157*, 630–639.

[84] Divya, K. and J. Østergaard (2009). Battery energy storage technology for power systemsan overview. *Electric Power Systems Research 79*(4), 511–520.

[85] Dommel, H. W. and W. F. Tinney (1968, Oct.). Optimal power flow solutions. *IEEE Transactions on Power Apparatus and Systems* (10), 1866–1876.

[86] Dörfler, F. and F. Bullo (2012). Synchronization and transient stability in power networks and nonuniform kuramoto oscillators. *SIAM Journal on Control and Optimization 50*(3), 1616–1642.

[87] Dörfler, F. and F. Bullo (2013). Kron reduction of graphs with applications to electrical networks. *IEEE Trans. on Circuits and Systems 60*(1), 150–163.

[88] Dörfler, F., J. W. Simpson-Porco, and F. Bullo (2016). Breaking the hierarchy: Distributed control and economic optimality in microgrids. *IEEE Trans. on Control of Network Systems 3*(3), 241–253.

[89] Duan, C., W. Fang, L. Jiang, L. Yao, and J. Liu (2018). Distributionally robust chance-constrained approximate ac-opf with wasserstein metric. *IEEE Transactions on Power Systems 33*(5), 4924–4936.

[90] Dugan, R. and R. Arritt (2010). The IEEE 8500-node test feeder. *Electric Power Research Institute, Palo Alto, CA, USA*.

[91] Džafić, I., M. Gilles, R. A. Jabr, B. C. Pal, and S. Henselmeyer (2013). Real time estimation of loads in radial and unsymmetrical three-phase distribution networks. *IEEE Transactions on Power Systems 28*(4), 4839–4848.

[92] El Ghaoui, L. (1995). State-feedback control of systems with multiplicative noise via linear matrix inequalities. *Systems & Control Letters 24*(3), 223–228.

[93] Elia, N. (2005). Remote stabilization over fading channels. *Systems & Control Letters 54*(3), 237–249.

[94] EPRI (2015). Summary of EPRI test circuits, ckt7. `http://solar.ucsd.edu/c/wp-content/uploads/2015/05/Readme.pdf`.

[95] Erdoğan, E. and G. Iyengar (2006, June). Ambiguous chance constrained problems and robust optimization. *Mathematical Programming 107*(1-2), 37–61.

[96] Esfahani, P. M. and D. Kuhn (2017, July). Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 1–52.

[97] Esfahani, P. M., S. Shafieezadeh-Abadeh, G. A. Hanasusanto, and D. Kuhn (2018). Data-driven inverse optimization with imperfect information. *Mathematical Programming 167*(1), 191–234.

[98] Fooladivanda, D. and J. A. Taylor (2018). Energy-optimal pump scheduling and water flow. *IEEE Transactions on Control of Network Systems 5*(3), 1016–1026.

[99] Fortenbacher, P., J. L. Mathieu, and G. Andersson (2017). Modeling and optimal operation of distributed battery storage in low voltage grids. *IEEE Transactions on Power Systems 32*(6), 4340–4350.

[100] Friedman, J. H. (2001, Oct.). Greedy function approximation: a gradient boosting machine. *The Annals of statistics*, 1189–1232.

[101] Gallego, G. and I. Moon (1993). The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society 44*(8), 825–834.

[102] Gan, L. and S. H. Low (2016). An online gradient algorithm for optimal power flow on radial networks. *IEEE Journal on Selected Areas in Communications 34*(3), 625–638.

[103] Gao, R., L. Xie, Y. Xie, and H. Xu (2018). Robust hypothesis testing using wasserstein uncertainty sets. In *Advances in Neural Information Processing Systems*, pp. 7902–7912.

[104] Gao, R., J. Xu, and H. Zhang (2017). Receding horizon control for multiplicative noise stochastic systems with input delay. *Automatica 81*, 390–396.

[105] Gebraad, P., F. Teeuwisse, J. Van Wingerden, P. A. Fleming, S. Ruben, J. Marden, and L. Pao (2016). Wind plant power optimization through yaw control using a parametric model for wake effectsa CFD simulation study. *Wind Energy 19*(1), 95–114.

[106] Gebraad, P. and J. Van Wingerden (2015). Maximum power-point tracking control for wind farms. *Wind Energy 18*(3), 429–447.

[107] Gershon, E., U. Shaked, and I. Yaesh (2001). H control and filtering of discrete-time stochastic systems with multiplicative noise. *Automatica 37*(3), 409–417.

[108] Ghaoui, L. E., M. Oks, and F. Oustry (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations research 51*(4), 543–556.

[109] Givens, C. R. and R. M. Shortt (1984). A class of wasserstein metrics for probability distributions. *The Michigan Mathematical Journal 31*(2), 231–240.

[110] Goit, J. P. and J. Meyers (2015). Optimal control of energy extraction in wind-farm boundary layers. *Journal of Fluid Mechanics 768*, 5–50.

[111] Goldfarb, D. and G. Iyengar (2003). Robust portfolio selection problems. *Mathematics of operations research 28*(1), 1–38.

[112] Gomez-Exposito, A. and A. Abur (2004). *Power System State Estimation: Theory and Implementation*. CRC press.

[113] Goryashko, A. P. and A. S. Nemirovski (2014). Robust energy cost optimization of water distribution system with uncertain demand. *Automation and Remote Control 75*(10), 1754–1769.

[114] Gravell, B., P. M. Esfahani, and T. Summers (2019). Learning robust control for lqr systems with multiplicative noise via policy gradient. *IEEE Transactions on Automatic Control, arXiv preprint arXiv:1905.13547*.

[115] Gravell, B., P. M. Esfahani, and T. Summers (2020). Robust control design for linear systems via multiplicative noise. *IFAC World Congress*.

[116] Gravell, B., K. Ganapathy, and T. Summers (2020). Policy iteration for linear quadratic games with stochastic parameters. *IEEE Control Systems Letters 5*(1), 307–312.

[117] Grosso, J., C. Ocampo-Martínez, V. Puig, and B. Joseph (2014). Chance-constrained model predictive control for drinking water networks. *Journal of Process Control 24*(5), 504–516.

[118] Grosso, J., C. Ocampo-Martínez, V. Puig, D. Limon, and M. Pereira (2014, June). Economic MPC for the management of drinking water networks. In *European Control Conference*, pp. 790–795.

[119] Grosso, J. M., P. Velarde, C. Ocampo-Martinez, J. M. Maestre, and V. Puig (2017). Stochastic model predictive control approaches applied to drinking water networks. *Optimal Control Applications and Methods 38*(4), 541–558.

[120] Guggilam, S. S., E. Dall'Anese, Y. C. Chen, S. V. Dhople, and G. B. Giannakis (2016, July). Scalable optimization methods for distribution networks with high PV integration. *IEEE Transactions on Smart Grid 7*(4), 2061–2070.

[121] Guo, Y., K. Baker, E. Dall'Anese, Z. Hu, and T. Summers (2018, June). Stochastic optimal power flow based on data-driven distributionally robust optimization. In *2018 American Control Conference*, pp. 1–7. IEEE.

[122] Guo, Y., K. Baker, E. Dall'Aness, Z. Hu, and T. Summers (2018). Data-based distributionally robust stochastic optimal power flow, Part II: Case studies. *IEEE Transactions on Power Systems*.

[123] Guo, Y., M. Rotea, and T. Summers (2020). Stochastic dynamic programming for wind farm power maximization. In *2020 American Control Conference (ACC)*, pp. 4824–4829. IEEE.

[124] Guo, Y. and T. Summers (2020). Distributionally robust stochastic optimal power flow and risk management. In *2020 IFAC World Congress*.

[125] Guo, Y. and T. H. Summers (2019, July). A performance and stability analysis of low-inertia power grids with stochastic system inertia. In *American Control Conference*, pp. 1–6.

[126] Guo, Y., X. Zhou, C. Zhao, Y. Chen, T. Summers, and L. Chen (2020, July). Solving optimal power flow for distribution networks with state estimation feedback. In *Annual American Control Conference*, pp. 1–8.

[127] Hajebrahimi, A., I. Kamwa, E. Delage, and M. Abdelaziz (2020). Adaptive distributionally robust optimization for electricity and electrified transportation planning. *IEEE Transactions on Smart Grid*.

[128] Hakobyan, A. and I. Yang (2020). Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk. *arXiv preprint arXiv:2001.04727*.

[129] Halldórsson, B. V. and R. H. Tütüncü (2003). An interior-point method for a class of saddle-point problems. *Journal of Optimization Theory and Applications 116*(3), 559–590.

[130] Harris, C. M. and D. M. Wolpert (1998). Signal-dependent noise determines motor planning. *Nature 394*(6695), 780–784.

[131] Hinrichsen, D. and A. J. Pritchard (1996). Stability radii of systems with stochastic uncertainty and their optimization by output feedback. *SIAM journal on control and optimization 34*(6), 1972–1998.

[132] Horvat, T., V. Spudić, and M. Baotić (2012, July). Quasi-stationary optimal control for wind farm with closely spaced turbines. In *IEEE International Convention MIPRO*, pp. 829–834.

[133] Hu, W., G. Niu, I. Sato, and M. Sugiyama (2018). Does distributionally robust supervised learning give robust classifiers? In *International Conference on Machine Learning*, pp. 2029–2037. PMLR.

[134] Huang, Y., W. Zhang, and H. Zhang (2006). Infinite horizon lq optimal control for discrete-time stochastic systems. In *2006 6th World Congress on Intelligent Control and Automation*, Volume 1, pp. 252–256. IEEE.

[135] Inc., C. R. (2012, August). CVX: Matlab software for disciplined convex programming, version 2.0. `http://cvxr.com/cvx`.

[136] Jabr, R. A., S. Karaki, and J. A. Korbane (2015, Sept.). Robust multi-period OPF with storage and renewables. *IEEE Transactions on Power Systems 30*(5), 2790–2799.

[137] Jiang, R. and Y. Guan (2016, July). Data-driven chance constrained stochastic program. *Mathematical Programming 158*(1-2), 291–327.

[138] Jiang, Y., R. Pates, and E. Mallada (2017). Performance tradeoffs of dynamically controlled grid-connected inverters in low inertia power systems. In *56th IEEE Annual Conference on Decision and Control*, pp. 5098–5105.

[139] Johnson, K. E. and G. Fritsch (2012). Assessment of extremum seeking control for wind farm energy production. *Wind Engineering 36*(6), 701–715.

[140] Johnson, K. E. and N. Thomas (2009, June). Wind farm control: Addressing the aerodynamic interaction among wind turbines. In *American Control Conference*, pp. 2104–2109.

[141] Jouini, T. and Z. Sun (2020). Performance analysis and optimization of power systems with spatially correlated noise. *arXiv preprint arXiv:2005.13680*.

[142] Jowitt, P. W. and G. Germanopoulos (1992). Optimal pump scheduling in water-supply networks. *Journal of Water Resources Planning and Management 118*(4), 406–422.

[143] Kantorovich, L. and G. Rubinshtein (1958). On a space of totally additive functions, vestn. *Vestn Lening. Univ 13*, 52–59.

[144] Kleinman, D. (1969a). On the stability of linear stochastic systems. *IEEE Trans. on Automatic Control 14*(4), 429–430.

[145] Kleinman, D. (1969b). Optimal stationary control of linear systems with control-dependent noise. *IEEE Transactions on Automatic Control 14*(6), 673–677.

[146] Knudsen, T., T. Bak, and M. Soltani (2009). Distributed control of large-scale offshore wind farms. In *European Wind Energy Conference and Exhibition*. Citeseer.

[147] Koller, M., T. Borsche, A. Ulbig, and G. Andersson (2013). Defining a degradation cost function for optimal control of a battery energy storage system. In *IEEE Grenoble PowerTech (POWERTECH)*, pp. 1–6. IEEE.

[148] Koshal, J., A. Nedić, and U. V. Shanbhag (2011). Multiuser optimization: Distributed algorithms and error analysis. *SIAM Journal on Optimization 21*(3), 1046–1081.

[149] Kozin, F. (1969). A survey of stability of stochastic systems. *Automatica 5*(1), 95–112.

[150] Kristalny, M. and D. Madjidian (2011, Dec.). Decentralized feedforward control of wind farms: prospects and open problems. In *IEEE Conference on Decision and Control and European Control Conference*, pp. 3464–3469.

[151] Kroposki, B., B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan (2017). Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy. *IEEE Power and Energy Magazine 15*(2), 61–73.

[152] Kundur, P. et al. (2004). Definition and classification of power system stability. *IEEE Trans. on Power Systems 19*(2), 1387–1401.

[153] Kundur, P., N. J. Balu, and M. G. Lauby (1994). *Power system stability and control*, Volume 7. New York: McGraw-hill.

[154] Kushner, H. J. (1967). Stochastic stability and control. Technical report, Brown Univ Providence RI.

[155] Lavelle, M. and T. Grose (2013, February). Water demand for energy to double by 2035. *National Geographic News*.

[156] Lee, C. and S. Mehrotra (2015). A distributionally-robust approach for finding support vector machines. *Manuscript, available at http://www. optimization-online. org/DB_HTML/2015/06/4965. html*.

[157] Li, B., R. Jiang, and J. L. Mathieu (2016, Dec.). Distributionally robust risk-constrained optimal power flow using moment and unimodality information. In *IEEE Conference on Decision and Control*, pp. 2425–2430.

[158] Li, B., R. Jiang, and J. L. Mathieu (2017). Ambiguous risk constraints with moment and unimodality information. *Mathematical Programming*, 1–42.

[159] Li, B., J. L. Mathieu, and R. Jiang (2018, June). Distributionally robust chance-constrained optimal power flow assuming log-concave distributions. In *Power systems computation conference (PSCC)*.

[160] Li, B., M. Vrakopoulou, and J. L. Mathieu (2017). Chance constrained reserve scheduling using uncertain controllable loads, Part II: Analytical reformulation. *IEEE Transactions on Smart Grid*.

[161] Li, H., J. Xu, and H. Zhang (2020). Linear quadratic regulation for discrete-time systems with input delay and colored multiplicative noise. *Systems & Control Letters 143*, 104740.

[162] Li, J., D. Li, Y. Xi, and L. Cen (2014). Formulation of mpc for multiplicative stochastic uncertainty by multi-step probabilistic sets. *IFAC Proceedings Volumes 47*(3), 10421–10426.

[163] Li, J. Y. and R. H. Kwon (2013). Portfolio selection under model uncertainty: a penalized moment-based optimization approach. *Journal of Global Optimization 56*(1), 131–164.

[164] Li, J. Y.-M. (2018). Closed-form solutions for worst-case law invariant risk measures with application to robust portfolio optimization. *Operations Research 66*(6), 1533–1541.

[165] Li, N., C. Zhao, and L. Chen (2016). Connecting automatic generation control and economic dispatch from an optimization view. *IEEE Trans. on Control of Network Systems 3*(3), 254–264.

[166] Li, P., H. Wang, and B. Zhang (2017, Aug.). A distributed online pricing strategy for demand response programs. *IEEE Transactions on Smart Grid*.

[167] Li, P. and B. Zhang (2017). Linear estimation of treatment effects in demand response: An experimental design approach. *arXiv preprint arXiv:1706.09835*.

[168] Li, W., E. Todorov, and R. E. Skelton (2005). Estimation and control of systems with multiplicative noise via linear matrix inequalities. In *Annual American Control Conference*, pp. 1811–1816.

[169] Li, Z., Q. Guo, H. Sun, and J. Wang (2016). Sufficient conditions for exact relaxation of complementarity constraints for storage-concerned economic dispatch. *IEEE Transactions on Power Systems 31*(2), 1653–1654.

[170] Liang, X., J. Xu, and H. Zhang (2017). Discrete-time lqg control with input delay and multiplicative noise. *IEEE Transactions on Aerospace and Electronic Systems 53*(6), 3079–3090.

[171] Linderoth, J., A. Shapiro, and S. Wright (2006). The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research 142*(1), 215–241.

[172] Liu, J., J. Tang, F. Ponci, A. Monti, C. Muscas, and P. A. Pegoraro (2012). Trade-offs in PMU deployment for state estimation in active distribution grids. *IEEE Transactions on Smart Grid 3*(2), 915–924.

[173] Lotfi, S. and S. A. Zenios (2018). Robust var and cvar optimization under joint ambiguity in distributions, means, and covariances. *European Journal of Operational Research 269*(2), 556–576.

[174] Louca, R. and E. Bitar (2016). Stochastic AC optimal power flow with affine recourse. In *IEEE Conference on Decision and Control*, pp. 2431–2436.

[175] Low, S. H. (2014a, March). Convex relaxation of optimal power flow, Part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems 1*(1), 15–27.

[176] Low, S. H. (2014b, June). Convex relaxation of optimal power flow, Part II: Exactness. *IEEE Transactions on Control of Network Systems 1*(2), 177–189.

[177] Lu, J. and R. E. Skelton (2002). Mean-square small gain theorem for stochastic control: discrete-time case. *IEEE Transactions on Automatic Control 47*(3), 490–494.

[178] Lubin, M., Y. Dvorkin, and S. Backhaus (2016, Sept.). A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Transactions on Power Systems 31*(5), 3840–3849.

[179] Luo, H., Z. Hu, H. Zhang, and H. Chen (2018). Coordinated active power control strategy for deloaded wind turbines to improve regulation performance in AGC. *IEEE Trans. on Power Systems*.

[180] Madani, R., S. Sojoudi, and J. Lavaei (2015, Jan.). Convex relaxation for optimal power flow problem: Mesh networks. *IEEE Transactions on Power Systems 30*(1), 199–211.

[181] Madjidian, D., K. Mårtensson, and A. Rantzer (2011, June). A distributed power coordination scheme for fatigue load reduction in wind farms. In *American Control Conference*, pp. 5219–5224.

[182] Maharjan, S., Q. Zhu, Y. Zhang, S. Gjessing, and T. Basar (2013). Dependable demand response management in the smart grid: A stackelberg game approach. *IEEE Transactions on Smart Grid 4*(1), 120–132.

[183] Mallada, E., C. Zhao, and S. Low (2017). Optimal load-side control for frequency regulation in smart grids. *IEEE Trans. on Automatic Control 62*(12), 6294–6309.

[184] Manwell, J. F., J. G. McGowan, and A. L. Rogers (2010). *Wind Energy Explained: Theory, Design and Application*. John Wiley & Sons.

[185] Marden, J. R., S. D. Ruben, and L. Y. Pao (2013). A model-free approach to wind farm control using game theoretic methods. *IEEE Transactions on Control Systems Technology 21*(4), 1207–1214.

[186] Mehrotra, S. and H. Zhang (2014). Models and algorithms for distributionally robust least squares problems. *Mathematical Programming 146*(1-2), 123–141.

[187] Mešanović, A., U. Münz, and C. Heyde (2016). Comparison of $\mathcal{H}_\infty$, $\mathcal{H}_2$, and pole optimization for power system oscillation damping with remote renewable generation. *IFAC-PapersOnLine 49*(27), 103–108.

[188] Milano, F., F. Dörfler, G. Hug, D. J. Hill, and G. Verbič (2018). Foundations and challenges of low-inertia systems. In *2018 Power Systems Computation Conference (PSCC)*, pp. 1–25. IEEE.

[189] Mohagheghi, E., A. Gabash, M. Alramlawi, and P. Li (2018). Real-time optimal power flow with reactive power dispatch of wind stations using a reconciliation algorithm. *Renewable Energy 126*, 509–523.

[190] Mohagheghi, E., A. Gabash, and P. Li (2017). A framework for real-time optimal power flow under wind energy penetration. *Energies 10*(4), 535.

[191] Mohsenian-Rad, A.-H., V. W. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia (2010). Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Transactions on Smart Grid 1*(3), 320–331.

[192] Molzahn, D. K., I. A. Hiskens, et al. (2019). A survey of relaxations and approximations of the power flow equations. *Foundations and Trends® in Electric Energy Systems 4*(1-2), 1–221.

[193] Montalvo, I., J. Izquierdo, R. Pérez, and M. M. Tung (2008). Particle swarm optimization applied to the design of water supply systems. *Computers & Mathematics with Applications 56*(3), 769–776.

[194] Mood, A., F. Graybill, and D. Boes (1974). Introduction to statistical theory.

[195] Motter, A. E., S. A. Myers, M. Anghel, and T. Nishikawa (2013). Spontaneous synchrony in power-grid networks. *Nature Physics 9*(3).

[196] Munteanu, I., A. I. Bratcu, N.-A. Cutululis, and E. Ceanga (2008). *Optimal Control of Wind Energy systems: Towards a Global Approach*. Springer Science & Business Media.

[197] Nguyen, V. A., D. Kuhn, and P. M. Esfahani (2018). Distributionally robust inverse covariance estimation: The wasserstein shrinkage estimator. *arXiv preprint arXiv:1805.07194*.

[198] Ocampo-Martinez, C., V. Puig, G. Cembrano, and J. Quevedo (2013). Application of predictive control strategies to the management of complex networks in the urban water cycle [applications of control]. *IEEE Control Systems Magazine 33*(1), 15–41.

[199] Øksendal, B. (2003). Stochastic differential equations. In *Stochastic differential equations*, pp. 65–84. Springer.

[200] Ortega-Vazquez, M. A. (2014). Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty. *IET Generation, Transmission & Distribution 8*(6), 1007–1016.

[201] Pagnier, L. and P. Jacquod (2019). Optimal placement of inertia and primary control: a matrix perturbation theory approach. *IEEE Access 7*, 145889–145900.

[202] Palensky, P. and D. Dietrich (2011). Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE Transactions on Industrial Informatics 7*(3), 381–388.

[203] Pao, L. Y. and K. E. Johnson (2009, June). A tutorial on the dynamics and control of wind turbines and wind farms. In *American Control Conference*, pp. 2076–2089.

[204] Pao, L. Y. and K. E. Johnson (2011). Control of wind turbines. *IEEE Control Systems Magazine 31*(2), 44–62.

[205] Park, J. and G. Bayraksan (2020). A multistage distributionally robust optimization approach to water allocation under climate uncertainty. *arXiv preprint arXiv:2005.07811*.

[206] Park, J., S. Kwon, and K. H. Law (2013, April). Wind farm power maximization based on a cooperative static game approach. In *2013 Active and Passive Smart Structures and Integrated Systems*, Volume 8688, pp. 86880R.

[207] Perelman, L. and A. Ostfeld (2012). Water-distribution systems simplifications through clustering. *Journal of Water Resources Planning and Management 138*(3), 218–229.

[208] Perninge, M. and C. Hamon (2013, May). A stochastic optimal power flow problem with stability constraints, Part II: The optimization problem. *IEEE Transactions on Power Systems 28*(2), 1849–1857.

[209] Piagi, P. and R. H. Lasseter (2006). Autonomous control of microgrids. In *IEEE Power Engineering Society General Meeting*.

[210] Picallo, M., S. Bolognani, and F. Dörfler (2020). Closing the loop: Dynamic state estimation and feedback optimization of power grids. *Electric Power Systems Research*.

[211] Pirani, M., J. W. Simpson-Porco, and B. Fidan (2017). System-theoretic performance metrics for low-inertia stability of power networks. In *56th IEEE Annual Conference on Decision Control*, pp. 5106–5111.

[212] Poolla, B. K., S. Bolognani, and F. Dörfler (2017). Optimal placement of virtual inertia in power grids. *IEEE Trans. on Automatic Control 62*(12), 6209–6220.

[213] Prescott, S. L. and B. Ulanicki (2003). Dynamic modeling of pressure reducing valves. *Journal of Hydraulic Engineering 129*(10), 804–812.

[214] Primadianto, A. and C.-N. Lu (2016). A review on distribution system state estimation. *IEEE Transactions on Power Systems 32*(5), 3875–3883.

[215] Purchala, K., L. Meeus, D. Van Dommelen, and R. Belmans (2005). Usefulness of DC power flow for active power flow analysis. In *IEEE PES General Meeting*, pp. 454–459.

[216] Qi, T., J. Chen, W. Su, and M. Fu (2016). Control under stochastic multiplicative uncertainties: Part i, fundamental conditions of stabilizability. *IEEE Transactions on Automatic Control 62*(3), 1269–1284.

[217] Rahbar, K., J. Xu, and R. Zhang (2015). Real-time energy storage management for renewable integration in microgrid: An off-line optimization approach. *IEEE Transactions on Smart Grid 6*(1), 124–134.

[218] Ramadesigan, V., P. W. Northrop, S. De, S. Santhanagopalan, R. D. Braatz, and V. R. Subramanian (2012). Modeling and simulation of lithium-ion batteries from a systems engineering perspective. *Journal of The Electrochemical Society 159*(3), R31–R45.

[219] Rami, M. A. and X. Y. Zhou (2000). Linear matrix inequalities, riccati equations, and indefinite stochastic linear quadratic controls. *IEEE Transactions on Automatic Control 45*(6), 1131–1143.

[220] Read, T. R. and N. A. Cressie (2012). *Goodness-of-fit statistics for discrete multivariate data*. Springer Science & Business Media.

[221] Renganathan, V., N. Hashemi, J. Ruths, and T. H. Summers (2020a). Distributionally robust tuning of anomaly detectors in cyber-physical systems with stealthy attacks. In *2020 American Control Conference (ACC)*, pp. 1247–1252. IEEE.

[222] Renganathan, V., N. Hashemi, J. Ruths, and T. H. Summers (2020b). Higher-order moment-based anomaly detection. *arXiv preprint arXiv:2011.01522*.

[223] Renganathan, V., I. Shames, and T. H. Summers (2020). Towards integrated perception and motion planning with distributionally robust risk constraints. *arXiv preprint arXiv:2002.02928*.

[224] Roald, L., S. Misra, T. Krause, and G. Andersson (2017, March). Corrective control to handle forecast uncertainty: A chance constrained optimal power flow. *IEEE Transactions on Power Systems 32*(2), 1626–1637.

[225] Roald, L., F. Oldewurtel, T. Krause, and G. Andersson (2013, June). Analytical reformulation of security constrained optimal power flow with probabilistic constraints. In *IEEE Grenoble PowerTech*, pp. 1–6.

[226] Roald, L., F. Oldewurtel, B. Van Parys, and G. Andersson (2015, Aug.). Security constrained optimal power flow with distributionally robust chance constraints. *arXiv preprint arXiv:1508.06061*.

[227] Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of risk 2*, 21–42.

[228] Rossman, L. A. et al. (2000). EPANET 2: users manual. *US Environmental Protection Agency. Office of Research and Development*.

[229] Rotea, M. A. (2014). Dynamic programming framework for wind power maximization. *IFAC Proceedings Volumes 47*(3), 3639–3644.

[230] Safaoui, S., L. Lindemann, D. V. Dimarogonas, I. Shames, and T. H. Summers (2020). Control design for risk-based signal temporal logic specifications. *IEEE Control Systems Letters*.

[231] Sampathirao, A. K., P. Sopasakis, A. Bemporad, and P. P. Patrinos (2018). GPU accelerated stochastic predictive control of drinking water networks. *IEEE Transactions on Control Systems Technology 26*(2), 551–562.

[232] Sankar, G. S., S. M. Kumar, S. Narasimhan, S. Narasimhan, and S. M. Bhallamudi (2015). Optimal control of water distribution networks with storage facilities. *Journal of Process Control 32*, 127–137.

[233] Santhanagopalan, V., M. Rotea, and G. Iungo (2018). Performance optimization of a wind turbine column for different incoming wind turbulence. *Renewable Energy 116*, 232–243.

[234] Santoni, C., U. Ciri, M. Rotea, and S. Leonardi (2015). Development of a high fidelity CFD code for wind farm control. In *American Control Conference*.

[235] Schepers, J. and S. Van der Pijl (2007). Improved modelling of wake aerodynamics and assessment of new farm control strategies. In *Journal of Physics: Conference Series*, Volume 75, pp. 012039. IOP Publishing.

[236] Schweppe, F. C. and J. Wildes (1970). Power system static-state estimation, Part I-III. *IEEE Transactions on Power Apparatus and Systems* (1), 120–135.

[237] Shi, Y., B. Xu, D. Wang, and B. Zhang (2018). Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains. *IEEE Transactions on Power Systems 33*(3), 2882–2894.

[238] Sinha, M., F. Dörfler, B. B. Johnson, and S. V. Dhople (2017). Uncovering droop control laws embedded within the nonlinear dynamics of van der pol oscillators. *IEEE Trans. on Control of Network Systems 4*(2), 347–358.

[239] Skalyga, M. and Q. Wu (2020). Distributionally robust co-optimization of energy and reserve dispatch of integrated electricity and heat system. In *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pp. 1–6. IEEE.

[240] Skworcow, P., D. Paluszczyszyn, and B. Ulanicki (2014). Pump schedules optimisation with pressure aspects in complex large-scale water distribution systems. *Copernicus Publications on behalf of the Delft University of Technology*.

[241] Soleimanzadeh, M. and R. Wisniewski (2011). Controller design for a wind farm, considering both power and load aspects. *Mechatronics 21*(4), 720–727.

[242] Song, X. and J. H. Park (2016). Linear optimal estimation for discrete-time measurement delay systems with multichannel multiplicative noise. *IEEE Transactions on Circuits and Systems II: Express Briefs 64*(2), 156–160.

[243] Song, Y., D. J. Hill, and T. Liu (2017). Network-based analysis of small-disturbance angle stability of power systems. *IEEE Trans. on Control of Network Systems*.

[244] Spudic, V., M. Jelavic, M. Baotic, and N. Peric (2010). Hierarchical wind farm control for power/load optimization. *The Science of Making Torque from Wind*.

[245] Steinbuch, M., W. De Boer, O. Bosgra, S. Peeters, and J. Ploeg (1988). Optimal control of wind power plants. *Journal of Wind Engineering and Industrial Aerodynamics 27*(1-3), 237–246.

[246] Summers, T. (2018). Distributionally robust sampling-based motion planning under uncertainty. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6518–6523. IEEE.

[247] Summers, T., J. Warrington, M. Morari, and J. Lygeros (2015a). Stochastic optimal power flow based on conditional value at risk and distributional robustness. *International Journal of Electrical Power & Energy Systems 72*, 116–125.

[248] Summers, T., J. Warrington, M. Morari, and J. Lygeros (2015b, Nov.). Stochastic optimal power flow based on conditional value at risk and distributional robustness. *International Journal of Electrical Power & Energy Systems 72*, 116–125.

[249] Tegling, E., B. Bamieh, and D. F. Gayme (2015). The price of synchrony: Evaluating the resistive losses in synchronizing power networks. *IEEE Trans. Control of Network Systems 2*(3), 254–266.

[250] Tielens, P. and D. Van Hertem (2012). Grid inertia and frequency control in power systems with high penetration of renewables.

[251] Todorov, E. and M. I. Jordan (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience 5*(11), 1226–1235.

[252] Tôrres, L. A., J. P. Hespanha, and J. Moehlis (2015). Synchronization of identical oscillators coupled through a symmetric network with dynamics: A constructive approach with applications to parallel operation of inverters. *IEEE Trans. on Automatic Control 60*(12), 3226–3241.

[253] Ulanicki, B., H. AbdelMeguid, P. Bounds, and R. Patel (2008). Pressure control in district metering areas with boundary and internal pressure reducing valves. *Water Distribution Systems Analysis 2008*, 1–13.

[254] Ulbig, A., T. S. Borsche, and G. Andersson (2014). Impact of low rotational inertia on power system stability and operation. *IFAC Proceedings Volumes 47*(3), 7290–7297.

[255] Ulbig, A., T. S. Borsche, and G. Andersson (2015). Analyzing rotational inertia, grid topology and their role for power system stability. *IFAC-PapersOnLine 48*(30), 541–547.

[256] Venzke, A., L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis (2018). Convex relaxations of chance constrained AC optimal power flow. *IEEE Transactions on Power Systems 33*(3), 2829–2841.

[257] VerHulst, C. and C. Meneveau (2014). Large eddy simulation study of the kinetic energy entrainment by energetic turbulent flow structures in large wind farms. *Physics of Fluids 26*(2), 025113.

[258] Verleye, D. and E.-H. Aghezzaf (2013). Optimising production and distribution operations in large water supply networks: A piecewise linear optimisation approach. *International Journal of Production Research 51*(23-24), 7170–7189.

[259] Verriest, E. (2008). Time variant balancing and nonlinear balanced realizations. In *Model Order Reduction: Theory, Research Aspects and Applications*, pp. 213–250. Springer.

[260] Vrakopoulou, M., M. Katsampani, K. Margellos, J. Lygeros, and G. Andersson (2013, June). Probabilistic security-constrained AC optimal power flow. In *IEEE Grenoble PowerTech*, pp. 1–6.

[261] Vrakopoulou, M., B. Li, and J. L. Mathieu (2017). Chance constrained reserve scheduling using uncertain controllable loads, Part I: Formulation and scenario-based analysis. *IEEE Transactions on Smart Grid*.

[262] Vrakopoulou, M., K. Margellos, J. Lygeros, and G. Andersson (2013, Nov.). A probabilistic framework for reserve scheduling and N-1 security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems 28*(4), 3885–3896.

[263] Wang, C. and S. Chen (2020). A distributionally robust optimization for blood supply network considering disasters. *Transportation Research Part E: Logistics and Transportation Review 134*, 101840.

[264] Wang, F. and V. Balakrishnan (2002). Robust kalman filters for linear time-varying systems with stochastic parametric uncertainties. *IEEE transactions on signal processing 50*(4), 803–813.

[265] Wang, S., A. Taha, N. Gatsis, and M. Giacomoni (2020). Receding horizon control for drinking water networks: The case for geometric programming. *IEEE Transactions on Control of Network Systems*.

[266] Wang, S., A. F. Taha, N. Gatsis, and M. Giacomoni (2019). Geometric programming-based control for nonlinear, dae-constrained water distribution networks. In *Annual American Control Conference*, pp. 1470–1475.

[267] Wang, S., A. F. Taha, N. Gatsis, L. Sela, and M. Giacomoni (2020). Probabilistic state estimation in water networks. *arXiv preprint arXiv:2002.01083*.

[268] Wang, Y., C. Ocampo-Martinez, and V. Puig (2016). Stochastic model predictive control based on gaussian processes applied to drinking water networks. *IET Control Theory & Applications 10*(8), 947–955.

[269] Wang, Y., Y. Yang, L. Tang, W. Sun, and B. Li (2020). A wasserstein based two-stage distributionally robust optimization model for optimal operation of cchp micro-grid under uncertainties. *International Journal of Electrical Power & Energy Systems 119*, 105941.

[270] Warrington, J., P. Goulart, S. Mariéthoz, and M. Morari (2013, Nov.). Policy-based reserves for power systems. *IEEE Transactions on Power Systems 28*(4), 4427–4437.

[271] Watkins Jr, D. W. and D. C. McKinney (1997). Finding robust solutions to water resources problems. *Journal of Water Resources Planning and Management 123*(1), 49–58.

[272] Wiesemann, W., D. Kuhn, and M. Sim (2014, Oct.). Distributionally robust convex optimization. *Operations Research 62*(6), 1358–1376.

[273] Willems, J. and G. Blankenship (1971). Frequency domain stability criteria for stochastic systems. *IEEE Transactions on Automatic Control 16*(4), 292–299.

[274] Winter, W., K. Elkington, G. Bareux, and J. Kostevc (2015). Pushing the limits: Europe's new grid: Innovative tools to combat transmission bottlenecks and reduced inertia. *IEEE Power and Energy Magazine 13*(1), 60–74.

[275] Wonham, W. M. (1967). Optimal stationary control of a linear system with state-dependent noise. *SIAM Journal on Control 5*(3), 486–500.

[276] Wonham, W. M. (1968). On a matrix riccati equation of stochastic control. *SIAM Journal on Control 6*(4), 681–697.

[277] Wood, A. J. and B. F. Wollenberg (2012). *Power generation, operation, and control*. John Wiley & Sons.

[278] Wozabal, D. (2012). A framework for optimization under ambiguity. *Annals of Operations Research 193*(1), 21–47.

[279] Wozabal, D. (2014). Robustifying convex risk measures for linear portfolios: A nonparametric approach. *Operations Research 62*(6), 1302–1315.

[280] Wu, F. F. (1990). Power system state estimation: A survey. *International Journal of Electrical Power & Energy Systems 12*(2), 80–87.

[281] Wu, F. F. and A. Monticelli (1985). Network observability: theory. *IEEE Transactions on Power Apparatus and Systems* (5), 1042–1048.

[282] Xiao, C., D. Sutanto, K. M. Muttaqi, and M. Zhang (2020). Multi-period data driven control strategy for real-time management of energy storages in virtual power plants integrated with power grid. *International Journal of Electrical Power & Energy Systems 118*, 105747.

[283] Xiao, N., L. Xie, and L. Qiu (2009). Mean square stabilization of multi-input systems over stochastic multiplicative channels. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 6893–6898. IEEE.

[284] Xiao, Y., Y. Li, and M. A. Rotea (2019). CART3 field tests for wind turbine region-2 operation with extremum seeking controllers. *IEEE Transactions on Control Systems Technology 27*(4), 1744–1752.

[285] Xie, W. and S. Ahmed (2018). Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Transactions on Power Systems 33*(2), 1860–1867.

[286] Xing, Y., B. Gravell, X. He, K. H. Johansson, and T. Summers (2020). Linear system identification under multiplicative noise from multiple trajectory data. In *Annual American Control Conference*.

[287] Xu, B., Y. Shi, D. S. Kirschen, and B. Zhang (2018). Optimal battery participation in frequency regulation markets. *IEEE Transactions on Power Systems*.

[288] Yang, F., Z. Wang, and Y. Hung (2002). Robust kalman filtering for discrete time-varying uncertain systems with multiplicative noises. *IEEE Transactions on Automatic Control 47*(7), 1179–1183.

[289] Yang, Z., Y. Li, and J. E. Seem (2013, Oct.). Maximizing wind farm energy capture via nested-loop extremum seeking control. In *ASME Dynamic Systems and Control Conference*, pp. 1–8.

[290] Ying, W., Z. Zhi, A. Botterud, K. Zhang, and D. Qia (2016). Stochastic coordinated operation of wind and battery energy storage system considering battery degradation. *Journal of Modern Power Systems and Clean Energy 4*(4), 581–592.

[291] Yong, T. and R. H. Lasseter (2000, July). Stochastic optimal power flow: Formulation and solution. In *IEEE Power Engineering Society Summer Meeting*, Volume 1, pp. 237–242.

[292] Yu, G., R. Powell, and M. Sterling (1994). Optimized pump scheduling in water distribution systems. *Journal of Optimization Theory and Applications 83*(3), 463–488.

[293] Zamzam, A. S., X. Fu, and N. D. Sidiropoulos (2019). Data-driven learning-based optimization for distribution system state estimation. *IEEE Transactions on Power Systems*.

[294] Zessler, U. and U. Shamir (1989). Optimal operation of water distribution systems. *Journal of Water Resources Planning and Management 115*(6), 735–752.

[295] Zhang, H., Z. Hu, Z. Xu, and Y. Song (2017). Evaluation of achievable vehicle-to-grid capacity using aggregate pev model. *IEEE Transactions on Power Systems 32*(1), 784–794.

[296] Zhang, H., L. Li, J. Xu, and M. Fu (2015). Linear quadratic regulation and stabilization of discrete-time systems with delay and multiplicative noise. *IEEE Transactions on Automatic Control 60*(10), 2599–2613.

[297] Zhang, H. and P. Li (2011, Nov.). Chance constrained programming for optimal power flow under uncertainty. *IEEE Transactions on Power Systems 26*(4), 2417–2424.

[298] Zhang, H. and J. Xu (2015). Stochastic control for ito system with state transmission delay. In *2015 American Control Conference (ACC)*, pp. 5282–5287. IEEE.

[299] Zhang, L. and J. Lam (2002). On $\mathcal{H}_2$ model reduction of bilinear systems. *Automatica 38*(2), 205–216.

[300] Zhang, Y. and J. Dong. Distributionally robust chance-constrained building load control under uncertain renewables. Technical report, Technical report, June 2020. http://www. optimization-online. org/DB_FILE .

[301] Zhang, Y. and G. B. Giannakis (2013, Oct.). Robust optimal power flow with wind integration using conditional value-at-risk. In *IEEE International Conference on Smart Grid Communications*, pp. 654–659.

[302] Zhao, C., U. Topcu, N. Li, and S. Low (2014). Design and stability of load-side primary frequency control in power systems. *IEEE Trans. on Automatic Control 59*(5), 1177–1189.

[303] Zhong, Q.-C. and T. Hornik (2012). *Control of power inverters in renewable energy and smart grid integration*, Volume 97. John Wiley & Sons.

[304] Zhou, X., E. Dall'Anese, L. Chen, and K. Baker (2017). Incentive-based voltage regulation in distribution networks. In *American Control Conference (ACC), 2017*, pp. 2732–2738. IEEE.

[305] Zhou, X., Z. Liu, Y. Guo, C. Zhao, and L. Chen (2019). Gradient-based multi-area distribution system state estimation. *arXiv preprint, arXiv:1909.11266* .

[306] Zhou, X., Z. Liu, C. Zhao, and L. Chen (2020). Accelerated voltage regulation in multi-phase distribution networks based on hierarchical distributed algorithm. *IEEE Transactions on Power Systems*.

[307] Zhu, L. and D. J. Hill (2018). Stability analysis of power systems: A network synchronization perspective. *SIAM Journal on Control and Optimization 56*(3), 1640–1664.

[308] Zimmerman, R. D. (2010). Matpower 4.0 b4 users manual. *Power Syst Eng Res Cent*, 1–105.

[309] Zimmerman, R. D., C. E. Murillo-Sánchez, R. J. Thomas, et al. (2011). Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems 26*(1), 12–19.

[310] Zohrizadeh, F., C. Josz, M. Jin, R. Madani, J. Lavaei, and S. Sojoudi (2020). A survey on conic relaxations of optimal power flow problem. *European journal of operational research*.

[311] Zymler, S., D. Kuhn, and B. Rustem (2013). Worst-case value at risk of nonlinear portfolios. *Management Science 59*(1), 172–188.

BIOGRAPHICAL SKETCH

Yi Guo was born in Wuzhong, Ninxia Province, People's Republic of China, in 1991. He received his BS degree in electrical engineering from Xi'an JiaoTong University, Xian, Shaanxi, China, in 2013 and his MS degree in electrical engineering from the University of Michigan, Dearborn, MI, USA, in 2015. He is currently a PhD candidate under the supervision of Prof. Tyler H. Summers in the Department of Mechanical Engineering at The University of Texas at Dallas. He did his summer internship at National Renewable Energy Laboratory (NREL) in 2019. His research interests are in control and optimization in networks, with applications to electric power networks, water networks, wind farms and transportation electrification. He received the Best Author Award from Journal of Power System Technology in 2016, one of the Best Reviewer Awards from IEEE Transactions on Smart Grid in 2018 and the Best Paper Award (seven awards given for the papers in the period of 2017-2019) from IEEE Transactions on Power Systems in 2020. He also received the Betty and Gifford Johnson Endowment Travel Award from UT-Dallas in 2019.

CURRICULUM VITAE

Yi Guo

Contact Information:

Department of Mechanical Engineering
The University of Texas at Dallas
800 W. Campbell Rd.
Richardson, TX 75080-3021, U.S.A.
Email: yi.guo2@utdallas.edu

Education:

PhD, Mechanical Engineering, 2016 - 2020
The University of Texas at Dallas, Richardson, TX, USA

MS, Mechanical Engineering, 2019 - 2020
The University of Texas at Dallas, Richardson, TX, USA

MS, Electrical Engineering, 2013 - 2015
The University of Michigan, Dearborn, MI, USA

BS, Electrical Engineering, 2009 - 2013
Xi'an Jiaotong University, Xi'an, Shaanxi, China

Selected Awards:

Best Paper Award, IEEE Transactions on Power Systems 2020
(7 awards given across papers published between 2017-2019)
Betty and Gifford Johnson Endowment Travel Award, UT-Dallas 2019
Best Reviewer, IEEE Transactions on Smart Grid 2018
Best Author Award, Journal of Power System Technology 2016
Nonresident Graduate Student Scholarship, U of Michigan 2013-2015
Nonresident Undergraduate Student Scholarship, U of Michigan 2012-2013
Siyuan Scholarship, Xi'an JiaoTong University 2010-2013

Research Interests:

Stochastic Optimization and Optimal Control for Complex Networked Systems; Online Optimization and Learning-based Control; Autonomous Power Grids Operation and Optimization, Multi-energy Systems; Power and Energy Systems, Water Networks, Wind Farms and Intelligent Transportation

Publications:

Journal Articles (including under review):

1. Ahmad Taha, Shen Wang, **Yi Guo**, Tyler Summers, Nikolaos Gatsis, Marcio Giacomoni and Ahmed Abokifa "Revisiting the water quality sensor placement problem: optimizing network observability and state estimation metrics", *Journal of Water Resources Planning and Management*, (under review)

2. **Yi Guo**, Shen Wang, Ahmad Taha and Tyler Summers, "Optimal pump control for water distribution networks via data-based distributional robustness", *Automatica*, (under review)

3. **Yi Guo**, Xinyang Zhou, Changhong Zhao, Lijun Chen and Tyler Summers, "Optimal power flow with state estimation in-the-loop for distribution networks", *IEEE Transactions on Power Systems*, (under review)

4. Xinyang Zhou, Zhiyuan Liu, **Yi Guo**, Changhong Zhao, Jianqiao Huang and Lijun Chen, "Gradient-based multi-area distribution system state estimation", *IEEE Transactions on Smart Grid*, 2020

5. **Yi Guo**, Kyri Baker, Emiliano Dall'Anese, Zechun Hu and Tyler Summers, "Data-based distributionally robust stochastic optimal power flow, Part I: Methodologies", *IEEE Transactions on Power Systems*, 34(2), 1483-1292, March 2019. [**Best Paper Award**]

6. **Yi Guo**, Kyri Baker, Emiliano Dall'Anese, Zechun Hu and Tyler Summers, "Data-based distributionally robust stochastic optimal power flow, Part II: Case studies", *IEEE Transactions on Power Systems*, 34(2), 1493-1530, March 2019. [**Best Paper Award**]

7. Xiaoyu Ma, **Yi Guo**, Xianan Chen, Yukai Xiang and Kun-Long Chen, "Impact of coreless current transformer position on current measurement", *IEEE Transactions on Instrumentation and Measurement,* 68(10), 3801-3809, Oct. 2019

8. Kun-Long Chen, **Yi Guo** and Xiaoyu Ma, "Contactless voltage sensor for overhead transmission lines ", *IET Generation, Transmission & Distribution*, 12(4), 957-966, Feb. 2018

9. Kun-Long Chen, Ren-Shuo Wang, **Yi Guo**, Nanming Chen and Wei-Jen Lee, "A redundancy mechanism design for hall-based electronic current transformers", *Energies*, 10(3), 312, Mar. 2017

10. **Yi Guo**, Jingwei Xiong, Shengyao Xu and Wencong Su, "Two-stage economic operation of microgrid-like electric vehicle parking deck", *IEEE Transactions on Smart Grid*, 6(3), 1703-1712, May 2016

11. **Yi Guo**, Zechun Hu, Hongcai Zhang, Wencong Su, Kaiqiao Zhan and Zhiwei Xu, "A statistical method to evaluate the capability of residential distribution network for accommodating electric vehicle charging load", *Journal of Power System Technology*, 39(9), Sept. 2015. [**Best Author Award**]

12. Jingwei Xiong, Kuilin Zhang, **Yi Guo** and Wencong Su, "Investigate the impacts of PEV charging facilities on integrated electric distribution system and electrified transportation system", *IEEE Transactions on Transportation Electrification*, 1(2), 178-187, Aug. 2015

Conference Papers

1. Chengda Ji, **Yi Guo**, Tyler Summers and Dennice Gayme, "Performance of networked systems with a finite-time single node sensor failure", *2020 IEEE Conference on Decision and Control*, Jeju Island, Republic of Korea (under review)

2. **Yi Guo** and Tyler Summers, "Distributionally robust stochastic optimal water flow and risk management", *2020 IFAC World Congress*, Berlin, German, 2020 (to appear)

3. **Yi Guo**, Mario Rotea and Tyler Summers, "Stochastic dynamic programming for wind farm power maximization", *American Control Conference*, Denver, CO, 2020 (to appear)

4. **Yi Guo**, Xinyang Zhou, Changhong Zhao, Yu Chen, Tyler Summers and Lijun Chen, "Solving optimal power flow for distribution networks with state estimation feedback", *American Control Conference*, Denver, CO, 2020 (to appear)

5. Ben Gravell, **Yi Guo** and Tyler Summers, "Sparse optimal control of networks with multiplicative noise via policy gradient", *IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys)*, Chicago, IL, 2019

6. **Yi Guo** and Tyler Summers, "A performance and stability analysis of low-inertia power grids with stochastic system inertia", *American Control Conference*, Philadelphia, USA, 2019

7. **Yi Guo**, Kyri Baker, Emiliano Dall'Aness, Zechun Hu and Tyler Summers, "Stochastic optimal power flow based on data-driven distributionally robust optimization", *American Control Conference*, Milwaukee, USA, 2018

8. Kun-Long Chen, Xiangguo Yang, Xiaoyu Ma and **Yi Guo**, "High-voltage phasing test using a contactless sensor", *3rd International Conference on Intelligent Green Building and Smart Grid*, Yilan, Taiwan, Apr. 2018

9. **Yi Guo**, Shengyao Xu and Wencong Su, "Smart management systems of plug-in electric vehicle charging services", *2015 IEEE Transportation Electrification Conference and Expo*, Dearborn, Michigan, USA. June 14-17, 2015

10. **Yi Guo**, Xuanchen Liu, Yu Yan, Ni Zhang and Wencong Su, "Economic analysis of plug-in electric vehicles parking deck with dynamic pricing", *2014 IEEE Power and Energy Society General Meeting*, National Harbor, MD, USA. July 27-31, 2014

11. **Yi Guo**, Jian Hu and Wencong Su, "Stochastic optimization for economic operation of plug-in electric vehicle charging stations at a municipal parking deck integrated with on-site renewable energy generation", *2014 IEEE Transportation Electrification Conference and Expo*, Dearborn, Michigan, USA. June 15-18, 2014

12. Xuanchen Liu, **Yi Guo** and Wencong Su, "Day-ahead resource scheduling for economic operation of residential green building with low-carbon emissions, *Fourth Annual Great Lakes Symposium on Smart Grid and the New Energy Economy*, Chicago, USA. September 22-25, 2014

Teaching Experience:

Graduate Teaching Assistant, The University of Texas at Dallas
MECH 6326: Optimal Control and Dynamic Programming, Spring 2018

Graduate Teaching Assistant, The University of Michigan
ECE 542: Introduction to Power Management and Reliability, Winter 2014
ECE 4432: Renewable Power System, Fall 2013

Services:

International Program Technical Committee 2016 - Present
International Conference on Vehicle Technology and Intelligent Transport Systems

Journal Article Reviewer
IEEE Transactions on Power Systems ∘ IEEE Transactions on Smart Grid (**2018 Best Reviewer**) ∘ IEEE Transactions on Sustainable Energy ∘ IEEE Transactions on Transportation Electrification ∘ IEEE Transactions on Industrial Informatics ∘ IEEE Access ∘ IET Generation & Transmission & Distribution ∘ Robotics and Autonomous Systems ∘ China Communication ∘ Journal of Power System Technology ∘ IEEE Control System Letters ∘ Energy Sources Part B: Economics Planning and Policy

Conference Paper Reviewer
American Control Conference ∘ IEEE Conference on Decision and Control ∘ ASME Dynamic Systems and Control Conference ∘ IEEE Texas Power and Energy Conference (TPEC) ∘ IEEE SmartGridComm ∘ IEEE Transportation Electrification Conference and Exposition (ITEC) ∘ IEEE Power & Energy Society General Meeting (PESGM)