TUNING-BASED ALGORITHMS FOR CHANCE-CONSTRAINED OPTIMIZATION IN POWER SYSTEMS APPLICATIONS

by

Ashley Mimi Hou

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Electrical and Computer Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2023

Date of final oral examination: 12/12/2022

The dissertation is approved by the following members of the Final Oral Committee:
Line A. Roald, Asst. Professor, Electrical and Computer Engineering, UW-Madison
Bernard Lesieutre, Professor, Electrical and Computer Engineering, UW-Madison
James Luedtke, Professor, Industrial and Systems Engineering, UW-Madison
Dimitris Papailiopoulos, Assoc. Prof., Electrical and Computer Engineering, UW-Madison
Yury Maximov, Staff Scientist, Los Alamos National Laboratory

To my family, for their unconditional love and support.

ACKNOWLEDGMENTS

The past five years have been some of the most challenging, yet rewarding years I have experienced. I could not have made it through without the immense amount of kindness, support, and encouragement I received from all the wonderful people I encountered along the way. While I am indebted to far more people than I can name here, I would like to thank a few people that have been fundamental to shaping my experience in Madison.

First and foremost, I would like to express my deepest gratitude to my advisor Line Roald. I am immensely grateful that four years ago, Line took a chance on me as a student with no power systems background. Her endless patience, vast wealth of knowledge, infectious curiosity, and compassionate mentorship has expertly guided me throughout my PhD. I am further appreciative of her commitment to cultivating a healthy, positive, and collaborative lab environment, which has made these past few years of exploring research such an enjoyable experience.

I am also especially thankful to my committee members Bernie Lesieutre, Jim Luedtke, Dimitris Papailiopoulos, and Yury Maximov for spending the time to review this thesis and provide insightful comments and constructive feedback.

Thank you to all the faculty that I have had the pleasure of interacting with throughout my time at UW-Madison. Moreover, I would be remiss not to mention the plethora of teachers and professors that paved the way for my post-graduate education.

I have been extremely fortunate to have been a part of a research group as welcoming and supportive as the WISPO group. I appreciate the warm welcome I received from my senior lab mates Adria, Jonathan, David, and Sogol me as I transitioned into working on power systems. Thank you to my collaborator Kshitij for teaching me all about distribution grids. To my office mates Noah, Jiaqi, Yasmine, and Aditiya, as well as fellow WISPO members Xue, Julia, Maitreyee, and many more, thank you for all the post-Power Hour lunch outings, bike rides, and afternoons at the terrace throughout the years. A very special thanks goes to Joe and Sofia for all the laughter and late night (mis)adventures.

Doing a PhD would be an impossible endeavor without the support of great friends. Thank you to Audrey, Heng, Ravi, Stephen, and Alex for many GSA happy hours, hiking trips, and lots of commiseration about the trials and tribulations of graduate school in between. I would like to thank my roommates at Brittingham, Stephen, Lilly, Ariel, and Kyle, for many family dinners and summer evenings out on the porch. I am grateful to Madeline and Marina for keeping me sane outside of school. A particular thank you goes to my number one fan Aga, who never fails to brighten my day and with whom I look forward to continuing our Madison adventures in New York. To my friends beyond Madison, especially Vicky, Vivian, Jess, Deepti, Karolina, Tonya, and Tiff, you have all made me feel so loved and supported despite being cities, states, and continents away.

Thank you all for always being a phone call away and a place to escape to whenever I needed it.

Of course, I must thank Kyle for his endless encouragement and constant support throughout the ebbs and flows of graduate school. I am grateful to have navigated the past few years alongside of you (and George) and am excited for our next chapter together.

Last, but certainly not least, I would like to express my sincerest gratitude to my family. I am deeply indebted to my parents for their unconditional love, their unwavering support, and the sacrifices they have made in order to provide me and my brother with the opportunities we have had. Thank you for instilling in us the importance of education, dedication, and hard work.

ABSTRACT

As climate change unfolds, the path to a more sustainable future necessitates the development and effective integration of renewable energy sources, such as wind and solar generation, into existing power grids. However, the inherent intermittency, variability, and unpredictability of renewable generation pose critical challenges to the reliability of power systems operation. This dissertation addresses the technical considerations of explicitly incorporating these uncertainties into power systems operational planning via the development of chance-constrained optimal power flow (OPF) methods.

Chance-constrained models guarantee system security by limiting the violation probability of uncertain constraints. Existing methods often face a trade off between solution quality, feasibility, and computational tractability. This thesis addresses this trade off by proposing a data-driven tuning-based algorithm that obtains desirable solutions to chance-constrained problems by using feedback from sample-based evaluations to iteratively adapt simple, approximate optimization models. We further develop a theoretical framework for obtaining probabilistic feasibility guarantees for general tuning-based stochastic optimization methods by proposing a two-step tuning methodology that adds an a posteriori solution verification step. Not only does our method retain computational tractability by decoupling the solving of the optimization problem from the consideration of the uncertainty, but it also provides guarantees on solution feasibility without introducing excess conservatism or relying on assumptions on the uncertainty distribution.

Our tuning method is applied to two variants of the OPF problem. We identify approximate models, solution evaluation procedures, and tuning schemes suitable for the structure and properties of each problem setting. We first consider a DC linearized formulation of OPF for dispatching generation in transmission grids under renewable generation and load forecast uncertainty. Here, existing analytical chance constraint reformulations are leveraged and a bisection search is used to iteratively tune an identified single-dimensional safety parameter. In a case study, we observe the method achieves non-conservative solutions that meet desired violation probabilities for not only the single chance-constrained case, but also the significantly more challenging joint chance-constrained case.

We then shift focus to the distribution grid setting, where uncertainty arises from distributed energy resources (DERs) and load variability. In this setting, the three-phase unbalanced AC OPF problem is used to find optimal reactive power set points for solar PV inverters while minimizing voltage unbalance. For the single-chance constrained problem, we adapt the aforementioned approximate model and bisection search heuristic to this more complicated setting, where the uncertainty enters the constraints non-linearly and

implicitly. The case studies use realistic solar PV and load data to demonstrate that while the algorithm is able to enforce single chance constraints, the use of a single-dimensional tuning parameter is insufficiently flexible for the joint chance-constrained case. We subsequently propose a risk allocation-based scheme that enables multi-dimensional tuning. Empirical results demonstrate that the method successfully limits the joint violation probability as well as achieves probabilistic feasibility guarantees when used as a component of the two-step tuning methodology.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Sets

$\mathcal{A}$	Set of active constraints
$\mathcal{D}$	Set of loads in a power network
$\mathcal{G}$	Set of conventional generators in a power network
$\mathcal{L}$	Set of lines in a power network
$\mathcal{N}$	Set of buses or nodes in a power network
$\mathcal{U}_\xi$	Robust optimization uncertainty set
$\mathcal{V}$	Set of violated constraints
Ω	Uncertainty set (Chapters 6, 7)
$\Phi = \{a, b, c\}$	Set of phases
$\Xi = \{\xi^{(1)}, \ldots, \xi^{(N)}\}$	Set of N uncertainty samples

Optimization variables

$\theta^{\phi}_{i,\omega}, \boldsymbol{\Theta}_{i,\omega}$	Voltage angle corresponding to uncertain power injections
$\theta^{\phi}_{i}, \boldsymbol{\Theta}_{i}$	Voltage angle corresponding to average power injections
$p^{\phi}_{G,0,\omega}, \mathbf{P}_{G,0,\omega}$	Active power injection at substation corresponding to uncertain power injections
$p^{\phi}_{G,0}, \mathbf{P}_{G,0}$	Active power injection at substation corresponding to average power injections
$q^{\phi}_{G,0,\omega}, \mathbf{Q}_{G,0,\omega}$	Reactive power injection at substation corresponding to uncertain power injections
$q^{\phi}_{G,0}, \mathbf{Q}_{G,0}$	Reactive power injection at substation corresponding to average power injections
$q^{\phi}_{G,i}, \mathbf{Q}_{G,i}$	Reactive power generation of solar PV

$\boldsymbol{v}^{+}_{\mathbf{d}l,\omega}, \boldsymbol{v}^{+}_{\mathbf{q}l,\omega}$	Real and imaginary components of positive sequence voltage phasor $\boldsymbol{v}^{+}_{l,\omega}$ corresponding to uncertain power injections
$\boldsymbol{v}^{-}_{\mathbf{d}l,\omega}, \boldsymbol{v}^{-}_{\mathbf{q}l,\omega}$	Real and imaginary components of negative sequence voltage phasor $\boldsymbol{v}^{-}_{l,\omega}$ corresponding to uncertain power injections
$\boldsymbol{v}^{+}_{\mathbf{d}l}, \boldsymbol{v}^{+}_{\mathbf{q}l}$	Real and imaginary components of positive sequence voltage phasor $\boldsymbol{v}^{+}_{l,\omega}$ corresponding to average power injections
$\boldsymbol{v}^{-}_{\mathbf{d}l}, \boldsymbol{v}^{-}_{\mathbf{q}l}$	Real and imaginary components of negative sequence voltage phasor $\boldsymbol{v}^{-}_{l,\omega}$ corresponding to average power injections
$\mathbf{p}_g$	Decision variable for the chance-constrained DC OPF problem
$\mathbf{x}$	Decision variable for general chance-constrained problems
$\underline{\nu}^{\phi}_{q,i}, \overline{\nu}^{\phi}_{q,i}$	Dual variables associated with the lower and upper reactive power constraints
$\underline{\nu}^{\phi}_{v,i}, \overline{\nu}^{\phi}_{v,i}$	Dual variables associated with the lower and upper voltage magnitude constraints

Random variables

$\delta p^{\phi}_{G,i,\omega}, \delta P_{G,i,\omega}$	Deviation of uncertain active power generation of solar PV from average value
$\delta p^{\phi}_{L,i,\omega}, \delta P_{L,i,\omega}$	Deviation of uncertain active power load demand from average value
$\hat{E}_j(\mathbf{x}, \Xi)$	Empirical violation probability of solution $\mathbf{x}$ w.r.t. $j \in \mathcal{C}$ evaluated on samples Ξ
$\hat{E}_{max}(\mathbf{x}, \Xi)$	Maximum or worst case violation probability amongst constraints $j \in \mathcal{C}$ evaluated on samples Ξ
$\hat{E}_{joint}(\mathbf{x}, \Xi)$	Joint violation probability evaluated on samples Ξ
$\omega \in \Omega$	Uncertainty (Chapters 6, 7)
$\Omega_\xi \in \mathbb{R}$	Total power mismatch (sum of all uncertainty in system)
$\xi \in \mathbb{R}^m$	Uncertainty
$E_j(\mathbf{x})$	Violation probability of solution $\mathbf{x}$ w.r.t. $j \in \mathcal{C}$
$p^{\phi}_{G,i,\omega}, P_{G,i,\omega}$	Uncertain active power generation of solar PV
$p^{\phi}_{L,i,\omega}, P_{L,i,\omega}$	Uncertain active power load demand

$p^{\phi}_{i,\omega}, P_{i,\omega}$	Uncertain active power injection
$q^{\phi}_{L,i,\omega}, Q_{L,i,\omega}$	Uncertain reactive power load demand
$q^{\phi}_{i,\omega}, Q_{i,\omega}$	Uncertain reactive power injection
$Y_j(\mathbf{x}, \xi) \in \{0, 1\}$	Indicator variable of violation of constraint $j \in \mathcal{C}$
$Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)}))$	Random variable indicating that both uncertainty realizations $\xi^{(i)}, \xi^{(j)}$ lead to the same outcome (constraint violation or not)

Constants and parameters

$\alpha \in [0, 1]^n$	Vector of generator participation factors
$\beta \in \mathbb{R}_+$	Confidence level for un-equal margin case
$\delta_{SA} \in \mathbb{R}_+$	Scenario approach confidence level
$\delta_{tune} \in \mathbb{R}_+$	Tuning step confidence level
$\delta_t \in \mathbb{R}_+$	Verification step confidence level
$\epsilon \in [0, 1]$	Desired chance constraint violation probability for a single chance constraint or a joint chance constraint
$\epsilon_j \in [0, 1]$	Desired chance constraint violation probability of the chance constraint on constraint $j \in \mathcal{C}$
$\eta > 1$	Relaxation multiplier for risk allocation algorithm
$\eta \in \mathbb{R}_+$	Bisection search convergence threshold
$\mu_\xi \in \mathbb{R}^m$	Expectation or mean of random variable ξ
$\overline{\lambda}^{\phi}_{q,i}, \overline{\lambda}_q$	Upper reactive power constraint tightening
$\overline{\lambda}^{\phi}_{v,i}, \overline{\lambda}_v$	Upper voltage constraint tightening
$\overline{p}^{\phi}_{G,i}, \overline{P}_{G,i}$	Average active power generation of solar PV
$\overline{p}^{\phi}_{L,i}, \overline{P}_{L,i}$	Average active power load demand
$\overline{p}^{\phi}_{i,\omega}, \overline{P}_{i,\omega}$	Average active power injection
$\overline{q}^{\phi}_{G,i}, \underline{q}^{\phi}_{G,i}$	Upper and lower PV inverter limits w.r.t. average generation
$\overline{q}^{\phi}_{L,i}, \overline{Q}_{L,i}$	Average reactive power load demand

$\overline{q}^{\phi}_{i,\omega}, \overline{Q}_{i,\omega}$	Average reactive power injection
$\Sigma_{\xi} \in \mathbb{R}^{m} \times \mathbb{R}^{m}$	Covariance matrix of random variable ξ
$\sigma^{\phi}_{q,i}$	Estimated standard deviation of reactive power
$\sigma^{\phi}_{v,i}$	Estimated standard deviation of voltage magnitude
$\tau \geqslant 01$	Step size multiplier for risk allocation algorithm
$\underline{\epsilon}^{\phi}_{q,i}, \overline{\epsilon}^{\phi}_{q,i}$	Risk allocations for the upper and lower reactive power constraint
$\underline{\epsilon}^{\phi}_{v,i}, \overline{\epsilon}^{\phi}_{v,i}$	Risk allocations for the upper and lower voltage magnitude constraint
$\underline{\lambda}^{\phi}_{q,i}, \underline{\Lambda}_{q}$	Lower reactive power constraint tightening
$\underline{\lambda}^{\phi}_{v,i}, \underline{\Lambda}_{v}$	Lower voltage constraint tightening
$\underline{v}, \overline{v}$	Voltage magnitude limits
$c \in [0, 1]$	Constant for relating sample set size and conservativeness for tuning
$d \in \mathbb{R}^{n}_{+}$	Vector of loads
G, B	Nodal conductance and susceptance matrices of nodal admittance matrix Y
$k \in \mathbb{R}_{+}$	Additional margin for tuning (e.g., $t_{tune} = t + k$
$M \in \mathbb{R}^{l \times n}$	Matrix of power transfer distribution factors (PTDFs)
$p^{min}_{g}, p^{max}_{g} \in \mathbb{R}^{n}_{+}$	Vectors of upper and lower generation limits
$p^{max}_{l} \in \mathbb{R}^{l}_{+}$	Vectors line limits
$s \in \mathbb{R}_{+}$	Tuning parameter
$s_{min}, s_{max} \in \mathbb{R}_{+}$	Upper and lower limits on the tuning parameter
$t \in \mathbb{R}_{+}$	Verification margin
$t_{tune} \in \mathbb{R}_{+}$	Tuning margin
$\gamma^{\phi}_{L,i}, \Gamma_{L,i}$	Power ratio of load demand

Functions and probability distributions

$\hat{F}^{\phi}_{q,i}(\cdot)$	Empirical distribution of reactive power limit
$\hat{F}^{\phi}_{v,i}(\cdot)$	Empirical distribution of voltage magnitude

$\lambda_j : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$	Constraint tightening function for constraint $j \in \mathcal{C}$
$\Phi^{-1}(\cdot)$	Inverse cumulative distribution of a Gaussian random variable
$C(\cdot), S(\cdot)$	Cosine and sine components of branch angle matrix
$c : \mathbb{R}^n \to \mathbb{R}$	Cost function for generation
$f(\cdot)$	AC power flow equations (short-hand)
$f : \mathbb{R}^n \to \mathbb{R}$	Objective function of the general chance-constrained problem
$F_\xi(\cdot)$	Cumulative distribution function of the random variable ξ
$g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$	Inner constraint of the (single) chance constraint
$g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$	Inner constraint of multiple single chance constraint or joint chance constraint

Abbreviations

AC	Alternating current
CC OPF	Chance-constrained OPF
CCP	Chance-constrained problem
CDF	Cumulative distribution function
DC	Direct current
DER	Distributed energy resource
JCC	Joint chance constraint
OPF	Optimal power flow
PV	Photovoltaic
SCC	Single chance constraint
VUF	Voltage unbalance factor

LIST OF PUBLICATIONS

[GHR22] Kshitij Girigoudar, Ashley M. Hou, and Line A. Roald. "Chance-Constrained AC Optimal Power Flow for Unbalanced Distribution Grids". In: *11th Bulk Power Systems Dynamics and Control Symposium (IREP 2022)*. IEEE. 2022.

[HR20] Ashley M. Hou and Line A. Roald. "Chance Constraint Tuning for Optimal Power Flow". In: *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE. 2020, pp. 1–6.

[HR22a] Ashley M. Hou and Line A. Roald. "Data-Driven Tuning for Chance Constrained Optimization: Analysis and Extensions". In: *TOP* (2022). ISSN: 1134-5764, 1863-8279. DOI: 10.1007/s11750-022-00639-z.

[HR22b] Ashley M. Hou and Line A. Roald. "Data-Driven Tuning for Chance-Constrained Optimization: Two Steps Towards Probabilistic Performance Guarantees". In: *IEEE Control Systems Letters* 6 (2022), pp. 1400–1405. DOI: 10.1109/LCSYS.2021.3096826.

1 INTRODUCTION

The integration of renewable energy is a necessary step towards building more sustainable energy systems. However, renewable energy sources such as wind turbines or solar photovoltaics (PVs) are often intermittent in nature and likely to deviate from their forecasted generation, posing reliability and security challenges under current operational paradigms. Thus, as renewable generation continues to increase, it becomes imperative to account for *uncertainty* in power systems operation. This thesis focuses on addressing uncertainty in the optimal power flow (OPF) problem, which is an optimization problem fundamental to power systems operations and planning.

1.1 MOTIVATION

The optimal power flow (OPF) problem is widely used by power systems operators and researchers to make decisions about power systems operations. While there exists a variety of formulations with different decision variables, objective functions, and constraints, the OPF problem generally involves determining the generation dispatch that minimizes the total cost of generation subject to physical constraints on power flow and other operational limits such as generator or line flow capacity. Some variation of OPF is solved for several applications, such as day-ahead market clearing and real-time operation. This thesis concentrates on the power system operation and operational planning time scale, which encompasses decisions made between a few days to a few minutes ahead of real time power delivery.

In practice, OPF is formulated as a *deterministic* problem, assuming forecasts of uncertain variables and ignoring forecast errors. However, as the electric grid is facing higher levels of uncertainty resulting from the increased penetration of renewable energy generation and increased load variability, it is becoming more critical to *incorporate uncertainty modeling in power systems decision making*. If uncertainty is ignored or mis-characterized, the system runs the risk of violating operational constraints. Unfortunately, attempting to secure the system against all possible uncertainty realizations may come with increased operational costs. There exists a vast breadth of literature incorporating uncertainty in OPF models, primarily differing in terms of how the uncertainty is represented, which impacts the trade-off between system security and operational cost. Common methods include (but not limited to) two- and multi-stage stochastic formulations [Zha+17; KLR20; LS18; JKK15], robust [LDB16; Soa+18; CWZ18] and distributionally robust [ZSM17; MD18; XA18] methods, and chance-constrained techniques [ZL11; Vra+13a; Roa+13; Cao+13; BCH14; Sum+14; MGL14; Roa+15b; RA17; Roa+17; WRM18; Ven+18; Müh+19; LVM19].

In this work, we consider chance-constrained formulations of OPF. Chance constraints are an intuitive way to account for uncertainties as they require that constraints hold probabilistically, with a user-chosen acceptable violation probability parameter. Via this parameter, the aforementioned trade-off between cost (i.e., the cost of the generation dispatch) and security (i.e., the allowed level of constraint violations) can be explicitly specified. Chance-constrained models have also been observed to align with industry practices, such as for probabilistic reserve dimensioning [AZ15]. Unfortunately, chance-constrained problems are generally intractable as formulated and require reformulation of the probabilistic constraint to be solved. While certain problem classes and uncertainty distributions allow for tight (i.e., exact) reformulations, these methods generally introduce some level of sub-optimality or conservativeness to the resulting solution or rely on limiting distributional assumptions that may not be realistic in practical settings (e.g., Gaussian uncertainty) [Roa+13; BCH14].

This thesis aims to tackle the trade off between *solution quality, feasibility, and computational tractability* for chance-constrained problems via the use of data-driven tuning-based algorithms. In particular, we focus on methods that use feedback from uncertainty data to automatically adapt a tractable reformulation of the chance-constrained problem. Since the uncertainty resulting from renewable generation and load variability may not necessarily follow known or parameterizable distributions, we focus on the setting where *no distributional assumptions* are made on the uncertainty. Uncertainty information is instead incorporated solely via available data. Unlike many existing data-driven methods, our proposed tuning-based approach decouples the consideration of the uncertainty from the solving of the optimization problem, rather than incorporating the uncertainty samples directly within the problem formulation. As a result, our method enables the ability to learn from uncertainty data without risking computational tractability.

We further highlight that our tuning framework is highly flexible and can be applied to a wide range of problem formulations, including problems with complicated, non-convex structures that many existing chance-constrained methods cannot tackle. With respect to power systems applications, this is particularly important in settings where the full, non-linear, non-convex AC power flow equations must be used, such as when working with distribution grids.

1.2 CONTRIBUTIONS

The overall objective of this thesis is to address challenges associated with incorporating uncertainty in OPF problems for operational planning in transmission and distribution grids. The main contributions are threefold:

(i) **Proposal of the tuning algorithm for solving chance-constrained optimization problems**: This work proposes an iterative, tuning-based methodology for chance-

constrained optimization problems. The approach combines the use of approximate optimization models with feedback from uncertainty data to automatically refine model parameters, resulting in more desirable solutions. A particular advantage of the tuning-based method is that each algorithm component (i.e., approximate model, evaluation method, and tuning strategy) can be adapted to suit various problems structures or applications. As a result, this method is suitable for general non-convex, large-scale problems as well as problems with unknown or non-parametrizable uncertainty distributions, which are often computationally challenging to solve.

(ii) **Development of theoretical guarantees for tuning-based algorithms**: While tuning-based algorithms for stochastic optimization have empirically performed well and often have clear computational advantages, many approaches lack theoretical guarantees for the resulting solutions. In particular, methods typically do not take into consideration issues that may arise from the iterative re-use of samples, which may lead to overfitting to the input samples. To address this gap, this thesis develops a theoretical framework for obtaining probabilistic feasibility guarantees for general tuning-based stochastic optimization methods. We propose a two-step methodology, where the tuning-based algorithm generates a solution and an a posteriori solution evaluation is performed to provide feasibility results on the solution. Subsequently, we analyze how parameters used in the two-step methodology can be chosen a priori such that feasible solutions are likely to be obtained.

(iii) **Application of tuning-based methods to OPF case studies**: We propose practical implementations of the tuning method that can be used to find solutions for two variants of the OPF problem: the DC OPF problem for generation dispatch in transmission grids under uncertainty from renewable generation and load variability and the three-phase unbalanced AC OPF problem for determining solar PV reactive power set points that minimize voltage unbalance in distribution grids subject

to uncertainty from solar PV active generation and load variability. We identify appropriate approximation models, evaluation methodologies, and tuning schemes tailored to each problem setting and application.

The approximate models used for all OPF problems in this thesis reformulate constraints into a deterministic version that consists of the nominal constraint (i.e., the constraint without uncertainty) and a tightening term that functions as a security margin against the uncertainty. The calculation of the constraint tightenings and iterative tuning vary based on the problem structure and application. For the single and joint chance-constrained DC OPF and the single chance-constrained three-phase unbalanced AC OPF problems, the constraint tightenings are parameterized by a single tuning parameter and updated via a bisection search. For the joint chance-constrained three-phase unbalanced AC OPF problem, a more sophisticated approach involving adjusting a multi-dimensional parameter using the idea of risk allocation is taken.

For all applications, we provide numerical case study results and analyze the quality of the resulting solutions in terms of operational cost, system security, and theoretical feasibility results as well as assess computational aspects of the algorithm performance, such as scalability and data requirements.

1.3 THESIS ORGANIZATION

The following is an overview of the content of each chapter of this thesis:

Chapter 2 presents an overview of foundational topics that this thesis is built upon. We review requisite background information on the OPF problem and chance-constrained optimization. We present an overview of existing chance-constrained optimization litera-

ture, focusing specifically on applications to power systems operation.

Chapter 3 develops a data-driven tuning framework for chance-constrained optimization. The method combines the use of an approximation optimization problem with a posteriori sample-based evaluations to obtain non-conservative solutions to both single and joint chance-constrained problems. We propose a bisection-search based tuning algorithm, which is applied to the single and joint chance-constrained DC OPF problems for transmission networks in the case study.

Chapter 4 proposes a two-step tuning algorithm that combines a solution generation step, which uses the tuning algorithm to obtain a solution to the chance-constrained problem, with an a posteriori solution verification step, which allows us to obtain an a posteriori probabilistic guarantee on solution feasibility. We develop theory towards obtaining an a priori probabilistic guarantee. The chapter concludes with case study applying the method to joint chance-constrained DC OPF.

Chapter 5 analyzes the two-step tuning algorithm from Chapter 4 to characterize settings in which the algorithm can be best applied. We empirically assess the validity of the assumptions that underpin the theoretical developments in Chapter 4. The chapter additionally derives sample complexity requirements for the algorithm and provides extended numerical simulation results for joint chance-constrained DC OPF that explore the scalability of the algorithm.

Chapter 6 adapts the tuning algorithm for application to distribution networks. We formulate a three-phase unbalanced single chance-constrained AC OPF model that minimizes voltage unbalance and propose an adaption of the bisection search-based tuning algorithm to suit this problem setting. Detailed numerical results using real solar PV and

load data are presented that demonstrate the algorithm's ability to meet desired violation probabilities and investigate how properties of the uncertainty data can influence resulting solutions.

Chapter 7 highlights the limitations of solving joint chance-constrained problems using the original bisection-search based tuning method. A multi-dimensional tuning scheme based on the idea of risk allocation is proposed and implemented for a joint chance-constrained version of the three-phase AC OPF problem. In the case study, numerical results demonstrate the ability of this multi-dimensional tuning scheme in obtaining feasible solutions for joint chance-constrained AC OPF.

Chapter 8 summarizes the contributions of this thesis and provide an outlook on future research directions that build upon the ideas presented in this thesis.

2 BACKGROUND

In this chapter, we first present notation used throughout this thesis. We then provide a brief review of relevant background from power systems and stochastic optimization as well as a literature review of existing related methods.

2.1 NOTATION

Let $\mathbb{R}$ be the set of real numbers, $\mathbb{C}$ be the set of complex numbers, and $\mathbb{Z}$ be the set of integers. Let $[a, b]$ be the set of real numbers on the interval from a to b. Let $[n]$, where $n \in \mathbb{Z}^+$ denote the set $\{1, \ldots, n\}$. Other sets are generally denoted with upper case letters in calligraphic font (e.g., $\mathcal{N}$), with a couple exceptions. Notably, Ξ is used to denote the set of uncertainty samples. For vectors, lower case subscripts x_i generally refer to the i-th element of the vector x. Single lower case subscripts (e.g., x_i) are typically used as indices for decision variables, generators, buses, etc. Double lower case subscripts (e.g., x_{ij}) are typically used as indices for for transmission lines, denoting a connection from bus i to j. **Bold font** is used to denote optimization decision variables. All other notation will be introduced as it appears throughout the thesis.

2.2 CHANCE-CONSTRAINED OPTIMIZATION

Chance-constrained optimization offers a powerful tool for modeling decision making problems in the face of uncertainty. Chance constraints were first introduced by [CCS58] as a method to solve optimization problems under uncertainty. Subsequently, [MW65] proposed a joint chance-constrained formulation and [Pré70] proposed a non-linear formulation. Since its conception, chance-constrained optimization has been widely adopted to tackle stochastic problems in numerous domains, including finance [KPU02; GOO03; PAS09a], transportation [BOW11; DFL18], network design [SL13; NMK22], and healthcare [Duq+20] to name a few.

Chance constraints intuitively model the trade-off between solution optimality and system reliability by specifying an acceptable probability of violation, often denoted using parameter $\epsilon \in [0, 1]$. This parameter is chosen by the decision maker before the optimization problem is solved and generally set to low value (i.e., $\epsilon = 0.05$ or $\epsilon = 0.01$). A smaller ϵ indicates a lower tolerance to constraint violations (typically at the cost of obtaining a higher resulting objective value), while a higher ϵ indicates a willingness to accept higher risk.

In general, chance-constrained optimization problems (CCP) are formulated as

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \tag{2.1a}$$

$$\text{s.t. } \mathbb{P}_{\xi}\left(g(\mathbf{x}, \xi) \leqslant 0\right) \geqslant 1 - \epsilon. \tag{2.1b}$$

The optimization decision variable is denoted by vector $\mathbf{x} \in X$, where $X \subseteq \mathbb{R}^n$ is the deterministic feasible region. The uncertainty is denoted by random vector $\xi \in \mathbb{R}^m$ and has a cumulative distribution function of $F_{\xi}(\cdot)$. The function $f : \mathbb{R}^n \to \mathbb{R}$ represents the objective function to be minimized. The function $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ represents an

uncertain constraint that is required to hold with probability at least $1 - \epsilon$. The overall constraint (2.1b) is known as a chance constraint or probabilistic constraint.

2.2.1 *Single and joint chance constraints*

There exist two varieties of chance constraints: *single* or individual chance constraints (SCCs) and *joint* chance constraints (JCCs).

Single chance constraints secure constraints individually, with separate acceptable violation parameters for each probabilistic constraint. Thus, they have the advantage of allowing the decision maker to specify different violation probabilities for constraints that may have varying levels of restrictiveness. Optimization problems can have one single chance constraint, as in the above formulation (2.1), or multiple single chance constraints. Let $\mathcal{C}$ denote the set of constraints in the optimization problem, i.e., $\mathcal{C} = \{g_1(\mathbf{x}, \xi) \leqslant 0, \ldots, g_K(\mathbf{x}, \xi) \leqslant 0\}$. The single chance-constrained problem (SCCP) with multiple single chance constraints can be written

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \tag{2.2a}$$

$$\text{s.t. } \mathbb{P}_\xi \left(g_j(\mathbf{x}, \xi) \leqslant 0 \right) \geqslant 1 - \epsilon_j, \quad \forall j \in \mathcal{C}, \tag{2.2b}$$

where $g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are the individual constraints required to each hold with desired violation probability $\epsilon_j \in [0, 1]$. In this thesis, without loss of generality, we limit consideration to the case where all single chance constraints have the same desired violation probability ϵ.

In contrast, joint chance constraints require that all inner constraints are *simultaneously enforced* with a single acceptable violation probability ϵ. Examples include simultaneously enforcing constraints over several time steps (e.g., limiting violations across an entire multi-time period or rolling horizon), enforcing constraints across all buses in a network,

or simply enforcing a two-sided constraint across a single bus (e.g., jointly enforcing the upper and lower voltage magnitude constraints).

The joint chance-constrained problem (JCCP) can be written

$$\min_{\mathbf{x}\in X} f(\mathbf{x}) \tag{2.3a}$$

$$\text{s.t. } \mathbb{P}_{\xi}\left(g_j(\mathbf{x},\xi) \leqslant 0, \quad \forall j \in \mathcal{C}\right) \geqslant 1-\epsilon. \tag{2.3b}$$

For notational convenience, by defining

$$g(\mathbf{x},\xi) := \max_{j\in\mathcal{C}} g_j(\mathbf{x},\xi), \tag{2.4}$$

the formulation (2.3) can be represented using the same notation as (2.1).

The primary difference between the single and joint chance-constrained formulations is the placement of the qualifier $\forall j \in \mathcal{C}$ in (2.2) and (2.3). To illustrate the difference, consider a line flow constraint with an upper and lower limit. The probability that the lower limit alone is satisfied may include uncertainty realizations for which the the upper limit is not satisfied (and vice versa). In the single chance-constrained formulation, we want both constraint to have violation probabilities below their respective ϵ_j thresholds. In this case, even when both single chance constraints hold probabilistically, individual realizations of the uncertainty may satisfy one constraint and fail to satisfy the other. In some cases, we may not want to allow such scenarios, but rather desire that *both* the upper and lower limits are satisfied under a given individual uncertainty realization. This is what the joint formulation allows us to do.

Joint formulations are desirable in many applications as they provide stronger guarantees on overall system security, e.g., a system is deemed secure if and only if all constraints are simultaneously met. However, they are significantly more challenging to solve, primarily owning to the interdependence between system constraints, which can result in many approaches achieving highly conservative solutions.

We note that we do not make any structural assumptions on the constraint function(s) $g(\mathbf{x}, \xi)$ (i.e., how variables $\mathbf{x}$ and ξ enter into the function). As stated, this problem formulation captures a very broad range of optimization problems. Any assumptions required for the methods presented in this thesis will be stated in their corresponding chapters.

2.2.2 *Uncertainty assumptions*

The tractability of chance-constrained problems customarily depends on the amount of information available about the distribution of the uncertain variable. In general, such knowledge falls under three categories:

(i) *Complete distribution knowledge*: The exact distribution of the uncertainty is known.

(ii) *Partial distributional knowledge*: Partial information about the uncertainty distribution is know. For example, the moments (e.g., mean and variance) are known or certain characteristics of the distribution are know (e.g., unimodality).

(iii) *No distributional knowledge*: No information about the uncertainty distribution is known. In this case, samples of the uncertainty are typically available (also known as uncertainty scenarios). These samples can be obtained from, e.g., historical data and may not necessarily be parameterized by a known probability distribution.

The work presented in this thesis falls under the last case, where *we do not assume any distributional knowledge* and only assume the availability or ability to draw a set of i.i.d. samples of the uncertainty ξ. We denote this sample set by $\Xi := \{\xi^{(1)}, \ldots, \xi^{(N)}\}$.

Recall that the main sources of uncertainty considered in this thesis are due to renewable generation fluctuations and load variability, which are both values that can vary in a continuous, unbounded range. We can either define the uncertainty ξ as (i) the

deviations from the forecasted values or (ii) the renewable generation and load values themselves. Both interpretations are used in this thesis.

2.3 CHANCE-CONSTRAINED OPTIMAL POWER FLOW

This thesis focuses on the optimal power flow (OPF) problem, which is an optimization problem that minimizes an objective (e.g., the total cost of conventional electricity generation) while enforcing physical limits (e.g., power balance limits and transmission capacity) and engineering constraints (e.g., conventional generation limits). Many variations of the OPF problem exist, with formulations differing depending on the usage and application of the problem.

For transmission grid operations, the objective of OPF is often to minimize the total cost of conventional, dispatchable active power generation. DC linearized approximations and single-phase equivalents are typically sufficient to represent the power flow, which allow for more computational tractability. On the other hand, at the distribution grid level, a greater emphasis is placed on power quality issues. Often alternative objectives such as minimizing power losses or voltage unbalance are considered [ABL17]. Since distribution feeders are inherently unbalanced and exhibit high R/X ratios, it becomes imperative to consider the full, three-phase unbalanced, non-linear, non-convex AC power flow equations. The resulting optimization problem is non-convex, often of a much larger size, and significantly more challenging to solve. In fact, OPF, in general, has been shown to be NP-hard [LL11]. The formulations of OPF considered in this thesis will be described in detail in the chapters corresponding to their usage. The DC OPF problem for transmission grid operation is presented in Chapter 3 and the three-phase unbalanced AC OPF problem for distribution grid operation is presented in Chapter 6.

Traditionally, OPF is cast as a deterministic problem, using point forecasts of uncertain variables such as renewable generation or load. Such an approach can fail dramatically if forecast errors are large, resulting in situations where, for example, there is an excess amount of generation that causes line flows to significantly exceed line ratings. Chance constraints have been proposed as an effective means to endogenously account for these uncertainty sources and prevent such situations by imposing an upper bound on the violation probability for constraints. Not only are they an intuitive means to ensure system security, they align well with industry practices and have been adopted for applications such as probabilistic reserve dimensioning [AZ15].

Much of the existing literature on chance-constrained OPF has focused on modeling and solving problems formulated with the DC power flow equations, using both single chance constraints [ZL11; Roa+13; Old+13] as well as joint chance constraints [BT17; BB18; Pen+20]. However, there is an increasing need for methods that incorporate the full AC power flow equations, which is critical in applications such as voltage control or distribution grid optimization. The main challenge here is that uncertainty now enters the problem both non-linearly as well as implicitly. Methods for solving AC OPF have primarily focused on formulations using single-chance constraints [ZL11; Vra+13b; Roa+17], with only a limited amount of work considering the joint chance-constrained AC OPF problem [BB18; BB19]. We further note that the majority of methods for AC OPF rely on linearizations or approximations of the AC power flow equations as well as often use single-phase equivalents of the three-phase unbalanced equations. As a result, these methods may not be applicable to distribution grid applications.

To interpret chance constraint violations in power systems applications, we note a distinction between ”soft” and ”hard” constraints. ”Soft” constraints are when the decision maker wants the constraint to hold generally but is willing to allow a non-trivial chance of violation. This can be beneficial in scenarios where allowing a constraint to be violated would sufficiently increase the optimality of the resulting solution. An example

of ”soft” constraints are line limits, which can typically endure short periods of being overloaded but prolonged periods of overloading can cause problems such as line damage or fires. On the other hand, ”hard” constraints are constraints that cannot be violated, for example, due to violating the physics of the system. An example of a hard constraint is the apparent power limits of PV inverters, which typically include control systems that protect the device from overloading (as this would damage the inverter). For such constraints, a violation can be interpreted as a situation where the notion of a ”violation probability” corresponds to a situation in which the system model no longer accurately reflect system operations (i.e., if the inverter no longer provide the expected amount of power, this might impact line flows and other quantities in ways that our model does not capture).

2.4 OVERVIEW OF EXISTING METHODS

Chance-constrained problems are difficult to solve due to two fundamental issues. First, merely evaluating the probabilistic constraint (2.1b) involves multi-dimensional integration (and may not be possible when the underlying probabilistic distribution is unknown or partially unknown, as in our setting). Second, the feasible region of the chance-constrained problem is generally non-convex, even in cases where the underlying deterministic problem itself is convex. It has been shown that chance-constrained problems, in general, are NP-hard to solve [LAN10]. Moreover, it is also NP-hard to simply check the feasibility of a solution [AS08].

Much of the existing literature on CCPs focuses on developing tractable reformulations or approximations of chance-constrained problems that can achieve high quality (as measured by optimality), feasible solutions and can be tractably and efficiently solved. In this section, we review several common approaches for CCPs, focusing particularly on

methods applied to power systems operation and those relevant to the work presented in this thesis. We refer the reader to [GX19; Roa+23] for a more complete overview of stochastic methods in power systems.

2.4.1 *Analytical reformulation approaches*

A common technique is the formulation of deterministic analytical reformulations of chance constraints [Roa+13; BCH14; Wu+14; Roa+15b; LM15]. When the deterministic reformulation is solved, the resulting solution is guaranteed to be feasible to the original chance constraint. Under certain assumptions (e.g., when the distribution of the uncertainty is known), these reformulations can be tight. However, such reformulations typically come at the expense of strong distributional assumptions, which may not necessarily hold in many applications [DH17].

Moment-based analytical reformulation

In the following, we review a commonly used analytical reformulation applicable in the special case where the uncertainty follows a Gaussian distribution and the chance constraint is linear in ξ, i.e., the inner constraint be written as

$$g(\mathbf{x}, \xi) := a(\mathbf{x}) + b(\mathbf{x})^\top \xi - c, \tag{2.5}$$

where $a(\mathbf{x}) \in \mathbb{R}$, $b(\mathbf{x}) \in \mathbb{R}^m$, and $c \in \mathbb{R}$ are parameters. Let the uncertainty follow $\xi \sim \mathcal{N}(\mu_\xi, \Sigma_\xi)$, where $\mathcal{N}(\mu_\xi, \Sigma_\xi)$ denotes the multivariate Gaussian distribution with mean $\mu_\xi \in \mathbb{R}^m$ and covariance $\Sigma_\xi \in \mathbb{R}^{m \times m}$. The chance constraint on (2.5) can be tightly reformulated as the deterministic constraint

$$a(\mathbf{x}) + b(\mathbf{x})^\top \mu_\xi + \Phi^{-1}(1-\epsilon)\|b(\mathbf{x})^\top \Sigma_\xi^{1/2}\|_2 \leqslant c, \tag{2.6}$$

where $\Phi^{-1}(1-\epsilon)$ is the inverse cumulative distribution (CDF) of the Gaussian distribution. In other words, if we solve the optimization problem with the above deterministic constraint in place of the chance constraint, the resulting solution $\mathbf{x}^*$ satisfies the chance constraint with exactly $1-\epsilon$ probability. The term $\Phi^{-1}(1-\epsilon)$ is commonly known as the "safety parameter". If $b(\mathbf{x})$ is a constant and not a function of $\mathbf{x}$ or linear in $\mathbf{x}$, the resulting constraints are linear or second-order cone constraints, respectively, and can be efficiently solved by commercial solvers.

In this case, since the distribution of ξ is presumed to be known, the safety parameter s can be directly calculated. Suitable safety parameter values have been proposed for a variety of distributions [Roa+13; Roa+15b; LVM19]. However, when the uncertainty distribution is only partially known, the safety parameters typically only admit reformulations that are not tight. When no distributional knowledge is known, it unclear how to determine a suitable s parameter.

2.4.2 *Sample-based approaches*

In contrast to analytical reformulations, sample-based approaches do not require strong distributional or structural assumptions. Furthermore, they are applicable to both single and joint chance constraints. The most common sample-based approach is the scenario approach [CC06; CGR15], which solves the chance-constrained problem by replacing instances of the random variable with samples. Theoretical guarantees and sample complexity bounds for scenario approaches have been well studied [CC05; CG08; CGP09; Cal10]. Unfortunately, they often suffer from tractability issues resulting from the large sample size requirements needed to achieve desired feasibility guarantees [CGP09; ATL10; VLM19]. In particular, the required number of samples often scales with the number of decision variables, rendering high-dimensional problems intractable due the the prohibitively large sample requirements [CGP09]. Furthermore, the aforementioned

results focus on convex problems, with few recent works providing theoretical results for non-convex problems [CGR18; GVH20; GXM22].

In general, scenario approaches result in very conservative solutions, i.e., solutions with lower violation probability and higher cost than required, because each constraint must be enforced for every uncertainty realization. Furthermore, solutions can vary dramatically with different sample sets and performance cannot be improved with more samples. Both these issues have been often observed in applications to OPF [Vra+13a; RA17; LVM19]. Methods that allow for discarding of some sampled constraints been developed to reduce conservativeness and improve optimality [CG11; Min+17].

Other sample-based approaches for chance-constrained problems include the sample average approximation (SAA) approach [LA08; PAS09b]. The SAA method replaces the uncertain variable with samples, but does not require all sampled constraints to be satisfied. Instead, an optimal solution is found by identifying a subset of samples where violations are allowed. Several works have established probabilistic feasibility results [LA08]. While these methods can be less conservative than scenario approaches, they are typically recast as mixed-integer problems with limited numerical tractability [Rus02; ZSM15; KJ21]. Recent work [Pen+19] has investigated the use of smooth non-linear continuous approximations, but does not offer optimality guarantees.

2.4.3 *Conservative approximation approaches*

Chance-constrained problems can also be solved via the use of convex inner approximations of the non-convex chance constraint. The most well known method involves replacing the chance constraint with the conditional value-at-risk (CVaR) approximation [NS07; Sum+15], which typically yields a sub-optimal solution. Alternative convex approximations include ALSO-X [Ahm+17], which has been demonstrated to always

out-perform the CVaR approximation and is able to return optimal solutions to the chance-constrained problem under certain conditions [JX22].

2.4.4 *Robust optimization approaches*

Chance-constrained problems are also closely related to robust optimization problems, which requires finding a solution such that all constraints are satisfied for all uncertainty realizations within a set [BEN09], i.e.,

$$\min_{\mathbf{x}\in X} f(\mathbf{x}) \tag{2.7a}$$

$$\text{s.t. } g(\mathbf{x}, \xi) \leqslant 0, \qquad \forall \xi \in \mathcal{U}_\xi, \tag{2.7b}$$

where $\mathcal{U}_\xi$ is a pre-specified uncertainty set. These formulations are desirable when security against all uncertainty realizations in a set is necessary, e.g., if we know the uncertainty varies within an interval and want to ensure the solution is robust to *all* realizations in this interval (rather than *nearly all*, in the chance-constrained case). Distributionally robust methods extend this to settings where constraints must hold for all distributions in an ambiguity set [LDB16; ZSM17; XA17; CKW22].

Compared to chance-constrained models, robust models are generally more computationally tractable, provided that the underlying uncertainty sets $\mathcal{U}_\xi$ satisfy mild convexity and computability assumptions [BEN09]. Several works have reformulated robust and distributionally robust models into tractable convex problems [CE06; ZKR13; XA17; XA18]. However, robust methods often result in overly conservative solutions due to the requirement that constraints must satisfy even the worst-case scenario in an uncertainty set [ZG13; Pan+15].

Chance-constrained problems can be interpreted via a robust optimization lens. In particular, consider the constraint (2.5) from Section 2.4.1. This derived deterministic

reformulation (2.6) is equivalent to using a solving the robust problem (2.7) with an ellipsodal uncertainty set, i.e.,

$$\mathcal{U}_\xi := \{\xi : \|\xi\|_2 \leqslant \Phi^{-1}(1-\epsilon)\}. \tag{2.8}$$

The challenge of using such methods is identifying a suitable uncertainty set that ensures feasibility of the chance constraint and admits a tractable problem [BBC11]. Data-driven methods have been proposed to model uncertainty sets [LS15; Dua+18] as well as methods to adjust the uncertainty set [Jab13].

2.4.5 *Methods for joint chance-constrained problems*

In many applications, enforcing joint chance constraints can be more desirable than only enforcing single chance constraints as they allow for guarantees on the *overall* system security. Unfortunately, joint constraints are notoriously difficult to handle. The sample-based approaches discussed in Section 2.4.2 offer a way to solve joint chance-constrained problems directly. In particular, the SAA can be used to solve JCCPs [LA08] and has been applied to the joint chance-constrained OPF problem in [Por+22].

Several other methods for single chance-constrained problems, such as the analytical reformulations previously discussed in Section 2.4.1, do not easily apply to joint chance-constrained problems. Instead, a common technique used is to decompose the joint chance constraint into single chance constraints and using the union bound (also known as Boole's inequality [Boo54] or the Bonferroni inequalities) to create a conservative upper bound. Methods have attempted to decrease conservatism by optimizing the risk allocations [OW08]. Several works have proposed alternative approaches that can achieve tighter bounds [BT17], including via the use statistical learning [BB19]. However, these methods still remain fairly conservative. Heuristic iterative methods have also been proposed

[JHS20; VT11], but lack theoretical guarantees for joint chance constraint satisfaction. Alternative methods for joint chance-constrained problem, such as distributionally robust methods [CKW22], etc., have also been proposed.

Scenario approaches can also be directly used to solve joint chance-constrained problems. However, in power systems applications, the optimization problems of interested are often non-convex (e.g., unit commitment of AC OPF formulations) and the existing a priori feasibility guarantees of the scenario approach often do not rigorously hold. While analogous a priori bounds have been derived for certain non-convex problem classes such as mixed-integer programs [ESL14] or problems with non-convex objectives [Gra+15], the results are often conservative. Furthermore, there are no a priori results applicable to general non-convex problems [CGR18]. Instead, recent literature has sought to establish a posteriori bounds for general non-convex problems [CGR18] as well as for unit commitment [GXM22], and AC OPF problems [MVH16].

3 TUNING-BASED ALGORITHM FOR CHANCE-CONSTRAINED OPTIMIZATION

In this chapter, we propose a data-driven parameter tuning method that utilizes results from sample-based tests to achieve a desired violation probability for both single and joint chance-constrained optimization problems. The motivating example for the tuning method is the optimal power flow problem. We present a proof-of-concept case study, where we apply our method to solving the chance-constrained DC optimal power flow problem in the transmission grid setting.

The material presented in this chapter is based on the publication [HR20].

3.1 MOTIVATION AND RELATED WORK

Traditional solution methods for chance-constrained problems typically require uncertainty information to be directly incorporated in a detailed problem formulation, which can significantly increase the problem size and resulting computational complexity. Typically, there is a trade-off between the computational tractability (which is reduced as more information is taken into account) and the quality of the resulting solution (which is improved with more information). An alternative is to consider a tuning-based approach, which utilize a posteriori sample-based evaluations to iteratively used to refine optimization model formulations. Tuning-based approaches enables us to learn from the uncertainty data without directly incorporating it within the problem formulation, thus

providing a truly data-driven approach to stochastic optimization. These two contrasting methods are illustrated in Figure 3.1

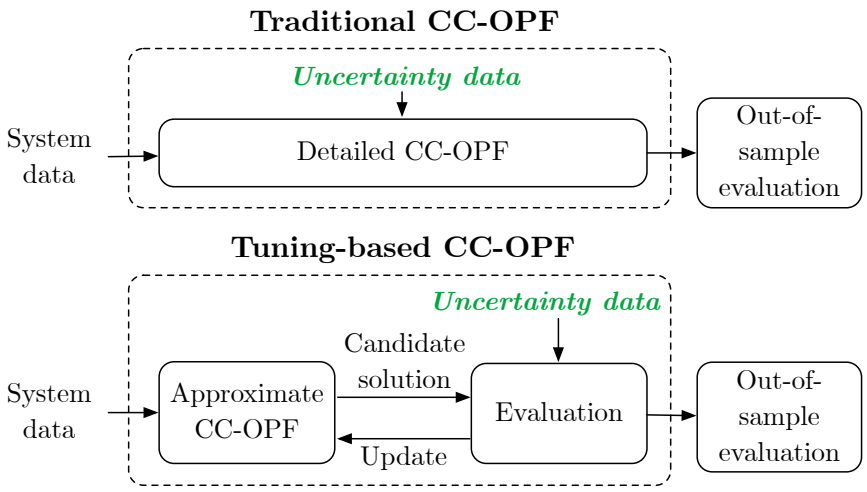


Figure 3.1: Traditional solution methods for CC OPF (top) include uncertainty information directly in a detailed problem formulation. The tuning method (bottom) uses a simple, approximate problem formulation which is iteratively updated using uncertainty data.

Several existing works [Old+15; RA17; Pen+19; LQ19; MMD20; GVH20] have used parameter tuning to adapt chance-constrained or stochastic formulations on the go. In particular, [Old+13; Old+15] perform online updates to a conservative chance-constrained formulation based on observed empirical violation levels. [LQ19] addresses the theory of tuning safety parameters for a generic chance-constrained problem while still maintaining probabilistic guarantees in a limited data regime. In [PLW19], an iterative tuning process was proposed to more accurately determine the value of a safety parameter in joint chance-constrained OPF. The tuning process in [PLW19] can be characterized as *fine-tuning*, where smaller adjustments are made to an accurate model. Methods have also directly tightened constraints [RA17; MR18] or varied the targeted violation probability in analytical reformulations [OW08; JHS20].

In this chapter, we propose an tuning method tuning method that iterates between solving an approximate optimization problem to obtain a candidate solution, evaluating this candidate solution using available samples of the uncertainty, and updating the

approximate problem based on the results of the evaluation. The decoupling the solving of the optimization problem from the consideration of the uncertainty samples can can help preserve computational tractability of the method (e.g., in contrast to the scenario approach, which includes samples in the problem formulation and can become computationally challenging to solve if the number of samples is large).

We emphasize that this method can be generally applicable to many chance-constrained problems, so long as there exists a suitable approximate problem that is computationally efficient to solve and can be parameterized by a tunable parameter. Therefore, although the optimization problem must be solved at each iteration of the tuning process, our method may be less computationally intense, yet more accurate, than alternative methods. Furthermore, our proposed method is applicable to both the single and joint chance-constrained problem. We note that we do not assume any distribution assumptions on the uncertainty, only that we have available a set of samples $\Xi = \{\xi^{(1)}, \ldots, \xi^{(N)}\}$. Moreover, we do not impose any structural assumptions on the chance constraints or the objective function.

3.2 TUNING-BASED ALGORITHM

In this section, we detail the three main components of the tuning algorithm: (i) the formulation of the approximate optimization model, (ii) the sample-based evaluation step, and (iii) the tuning step. In each of these descriptions, we present examples of each component that are used in the case study in this chapter as well as generally throughout this thesis. We note that these components can be formulated in a myriad of ways, depending on the optimization problem and application.

3.2.1 *Approximate optimization problem*

In the first step, we obtain a *candidate solution* by solving an approximate optimization model. In general, the model capture salient features of the problem setting, incorporate a suitable uncertainty representation, and contain parameters that are practical for tuning. Adjusting the tuning parameters should ideally result in solutions of varying cost and/or degrees of conservativenss. Finally, because this approximate problem will be solved several times during the iterative process, it must be computationally inexpensive and easily solved with commercial solvers.

There are several options available for approximate optimization models. We propose an approximation where the chance constraints are replaced with deterministic constraints consisting of the nominal estimated version of the inner constraint of the chance constraint (i.e., g_j) and a tightening term. The tightening term is parameterized by a *tuning parameter* s, which is iteratively adjusted based on a posteriori sample-based evaluations, and can be determined based on the set of available samples Ξ. Here, we choose to use a single dimensional tuning parameter $s \in \mathbb{R}_+$. While using separate safety parameters for each constraint would give us additional degrees of freedom and possibly more optimal solutions, it is a non-trivial extension due to the large number of constraints and resulting high-dimensional tuning challenge. This idea of multi-dimensional tuning is explored in Chapter 7. We further note that this approximate problem can be used for finding feasible solutions to both single and joint chance-constrained cases, even when we do not explicitly decompose the desired joint violation probability ϵ into single joint violation probabilities (i.e., via the Bonferroni bound as discussed in Section 2.4.5).

The most general formulation of the approximate model is as follows:

$$\min_{\mathbf{x}\in X} f(\mathbf{x}) \tag{3.1a}$$

$$\text{s.t. } g_j(x, \hat{\mu}_\xi) + \lambda_j(\mathbf{x}, s, \Xi) \leqslant 0, \quad \forall j \in \mathcal{C}. \tag{3.1b}$$

where $\hat{\mu}_\xi$ is an estimate of the mean of the uncertainty, calculated using samples Ξ, i.e., $\hat{\mu}_\xi := \frac{1}{N}\sum_{i=1}^{N} \xi^{(i)}$. The nominal constraint $g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is the inner constraint of the chance constraint, with the estimated mean $\hat{\mu}_\xi$ used in place of the uncertainty term ξ. The constraint tightening, denoted using $\lambda_j : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^{N\times n} \to \mathbb{R}$, is a function of the tuning parameter $s \in \mathbb{R}_+$, the decision variable $\mathbf{x}$, and the set of uncertainty samples Ξ. Our goal is to tune the safety parameter such that our resulting solution is feasible to the original problem (2.1) but avoids being overly conservative. We note that this deterministic approximation does not explicitly take into account uncertainty information outside the sampled data (i.e., it does not rely on any distributional knowledge stemming from distributional assumptions). Furthermore, the approximation secures against uncertainties solely via the tightening term. A higher tightening results in a constraint that is more robust to uncertainties, but also leads to a more costly solution.

We note that while there are several options for defining λ_j, there are a few considerations to take:

(i) We need to determine what (and how much) information from the sample set to use in formulation the tightening. In this chapter, we propose that the tightening is a function of the empirical mean $\hat{\mu}_\xi$ and empirical covariance $\hat{\Sigma}_\xi$, estimated from the sample set. We propose that the tuning of s will account for possible estimation error.

(ii) We need to consider the structure of λ_j, which will impact the ease of solving the approximate optimization problem. Recall that the resulting approximate problem needs to be able to be quickly solved by a commercial solver. Thus, it would be

desirable for the reformulated constraints to be, e.g., linear or second order cone constraints.

(iii) There exists a clear relationship between the tuning parameter s and the conservativeness of the corresponding solution. Ideally, as the tuning parameter s increase, the constraints should tighten, resulting in a lower violation probability of the constraint and higher objective value.

As an example, we propose a formulation of a simple approximate problem based on the moment-based analytical reformulation presented in Section 2.4.1. Let us consider a single chance constraint, where the inner constraint g_j is linearly separable in the decision variable $\mathbf{x}$ and uncertainty ξ and the uncertainty only occurs in the left-hand-side of the constraint, i.e.,

$$\mathbb{P}_\xi\left(a_j(\mathbf{x}) + b_j^\top \xi - c_j \leqslant 0\right) \geqslant 1 - \epsilon. \tag{3.2}$$

The moment-based analytical reformulation is the following linear constraint

$$a_j(\mathbf{x}) + b_j^\top \hat{\mu}_\xi - c_j + s\|b_j^\top \hat{\Sigma}_\xi^{1/2}\|_2 \leqslant 0, \tag{3.3}$$

where s is the safety parameter, which we take as the tuning parameter. We can see that the nominal constraint corresponds to $g_j(\mathbf{x}, \hat{\mu}_\xi) = a_j(\mathbf{x}) + b_j^\top \hat{\mu}_\xi - c_j$ and the constraint tightening is $\lambda_j(s, \Xi) := s\|b_j^\top \Sigma_\xi^{1/2}\|_2$. Notice that $\|b_j^\top \hat{\Sigma_\xi}^{1/2}\|_2$ is an estimate of the standard deviation of the nominal constraint. Thus, this formulation essentially corresponds to the moment-based analytical reformulation (2.6), where the safety parameter is the tuning parameter.

Furthermore, we highlight that in this case where the chance constraints are bilinear, the standard deviation estimate term $\|b_j^\top \hat{\Sigma_\xi}^{1/2}\|_2$ is not dependent on the decision variables. Thus, the tightening $\lambda_j(s, \Xi)$ is monotonic in s. Since the reformulation admits a

linear program, the feasible space of the deterministic problem would decrease as the constraints are tightened, resulting in a higher objective value.

Robust optimization perspective

We can consider the above analytical reformulation in a more general viewpoint by using connections to robust optimization. Recall the robust formulation 2.7 from Section 2.4.4. The above constraint reformulation (3.1) can be interpreted as a robust constraint with an ellipsoidal uncertainty set, i.e., the constraint is equivalent to the set

$$g(\mathbf{x}, \xi) \leqslant 0, \quad \forall \xi \in \mathcal{U}_\xi := \{\xi : \|\xi\|_2 \leqslant s\}, \tag{3.4}$$

where the tuning parameter s controls the size of the uncertainty set [BEN09]. Intuitively, by choosing s such that the uncertainty set $\mathcal{U}_\xi$ contains sufficient probability mass, a solution satisfying the set of constraints (3.4) will satisfy the original chance constraint with probability at least $1 - \epsilon$.

3.2.2 *Sample-based evaluation*

We solve the approximate problem (3.1) to obtain a candidate solution $\mathbf{x}$. We then use sample-based evaluations to assess the feasibility of the solution with respect to a constraint $g(\mathbf{x}, \xi) \leqslant 0$. We define the random variable $Y(\mathbf{x}, \xi)$ to be an indicator of whether the constraint is violated, i.e.,

$$Y(\mathbf{x}, \xi) := \begin{cases} 0 & \text{if } g(\mathbf{x}, \xi) \leqslant 0 \\ 1 & \text{otherwise.} \end{cases} \tag{3.5}$$

Thus, we can write the *violation probability* of solution $\mathbf{x}$ with respect to constraint $g(\mathbf{x}, \xi) \leqslant 0$ as

$$E(\mathbf{x}) := \mathbb{E}_{\xi}[Y(\mathbf{x}, \xi)] = \Pr(g(\mathbf{x}, \xi) > 0). \tag{3.6}$$

However, since we do not make any distributional assumptions on ξ (and, in many applications, the distribution is unknown), we cannot directly evaluate $E(\mathbf{x})$. Therefore, we approximate it by using a set of i.i.d. samples Ξ. We evaluate the indicator random variable $Y(\mathbf{x}, \xi^{(i)})$ for each sample $\xi^{(i)} \in \Xi$. For a solution $\mathbf{x}$ and i.i.d. sample $\xi^{(i)}$, the realizations of $Y(\mathbf{x}, \xi^{(i)})$ are i.i.d. as well. The *empirical violation probability* of solution $\mathbf{x}$ with respect to constraint $g(\mathbf{x}, \xi) \leqslant 0$ is then defined as

$$\hat{E}(\mathbf{x}, \Xi) := \frac{1}{N} \sum_{i=1}^{N} Y_j(\mathbf{x}, \xi^{(i)}). \tag{3.7}$$

For single chance-constrained problems with multiple single chance constraints, each constraint $j \in \mathcal{C}$ has an associated empirical violation probability that is evaluated. We use subscripts j to denote the indicator random variables, violation probability, and empirical violation probability of constraint $j \in \mathcal{C}$, i.e., $Y_j(\mathbf{x}, \xi)$, $E_j(\mathbf{x})$, and $\hat{E}_j(\mathbf{x}, \Xi)$. We note that for the tuning process, since we need to ensure that *all* the single chance constraints have a violation probability of below ϵ, we need to calculate the maximum or worst case violation probability amongst all constraints, i.e.,

$$\hat{E}_{\max}(\mathbf{x}, \Xi) := \max_{j \in \mathcal{C}} \hat{E}_j(\mathbf{x}, \Xi) \leqslant \epsilon. \tag{3.8}$$

For the joint chance-constrained problem, we calculate the empirical violation probability of the joint constraint $g(\mathbf{x}, \xi) = \max_{j \in \mathcal{C}} g_j(\mathbf{x}, \xi)$, which, for clarity purposes, we denote using $\hat{E}_{\text{joint}}(\mathbf{x}, \Xi)$.

The solution $\mathbf{x}$ is *feasible* for chance-constrained problem (2.1) if its violation probability is less than or equal to the desired acceptable violation probability ϵ, i.e.,

$$\mathrm{E}(\mathbf{x}) \leqslant \epsilon. \tag{3.9}$$

For single chance-constrained problems with multiple chance constraints, we require the worst case violation probability $\mathrm{E}_{\max}(\mathbf{x})$ must be less than ϵ. For joint chance constrained problems, we require the joint violation probability $\mathrm{E}_{\text{joint}}(\mathbf{x})$ must be less than ϵ.

Since we are unable to obtain the true violation probabilities of the constraints, we then instead check what we call the *empirical feasibility criterion*,

$$\hat{\mathrm{E}}(\mathbf{x}, \Xi) \leqslant \epsilon. \tag{3.10}$$

Again, the empirical worst case violation probability $\hat{\mathrm{E}}_{\max}(\mathbf{x}, \Xi)$ or empirical joint violation probability $\hat{\mathrm{E}}_{\text{joint}}(\mathbf{x}, \Xi)$ is used for the SCCP with multiple chance constraints or JCCP, respectively. If this criteria satisfied, we declare $\mathbf{x}$ to be *empirically feasible*.

3.2.3 *Tuning*

In the tuning step, we adjust the tightenings by adjusting the tuning parameter s based on the result of the evaluation step. If we observe that the solution does not satisfy the empirical feasibility criterion, we tighten the reformulated constraint to promote more conservative solutions in future iterations. Conversely, if the solution does satisfy the criterion, we relax the reformulated constraint, leading to lower cost and less conservative solutions in future iterations.

There are several ways to perform this tightening scheme. We propose a simple heuristic method based on a bisection search to determine the least conservative value

of s that provides a feasible solution to the chance-constrained problem. We denote this value by s^*, and the corresponding solution by $\mathbf{x}^*$. The algorithm is as follows:

(i) *Initialization*: Initialize the tuning algorithm by setting the iteration count to $k = 0$ and determining suitable upper and lower bounds for the tuning parameter s to ensure that $s_{\min} \leqslant s^* \leqslant s_{\max}$:

For the lower bound, we use $s^0_{\min} = 0$, which corresponds to the case where no uncertainty is considered in the generation and line flow constraints. Because s has in inverse monotonic correlation with the observed empirical violation probability $\hat{E}_G(\tilde{x}, \Xi)$, we expect that the solution to the approximate problem (3.1) corresponding to $s = 0$ will have the highest possible empirical violation probability for any $s \geqslant 0$.

For the upper bound, we use the Cantelli inequality [Roa+15b] since we assume availability of the first and second moments μ_ξ and Σ_ξ. We set $s^0_{\max} = \sqrt{(1-\epsilon)/\epsilon}$.

Define tuning parameter to be the midpoint of the two bounds, $s^0 = (s_{\max} - s_{\min})/2$.

(ii) *Solve approximate problem*: Solve the approximate problem (3.1) with s^k to obtain a candidate solution $\mathbf{x}^k$.

(iii) *Check feasibility*: Determine whether the candidate solution $\mathbf{x}^k$ satisfies the empirical feasibility criterion (3.10).

(iv) *Update approximation*: If $\mathbf{x}^k$ does not satisfy the feasibility criterion, tighten the constraint by setting the lower bound $s_{\min} = s^k$, which will increase the value of s in future iterations and thus promote more conservative solutions with higher cost.

Conversely, if $\mathbf{x}^k$ is feasible, relax the reformulation by setting the upper bound $s_{\max} = s^k$. This will decrease the value of s in future iterations, thus leading to lower cost, less conservative solutions.

(v) *Check convergence*: If $|s_{max} - s_{min}| \leqslant \eta$, where η is a pre-specified tolerance level, stop the iterations and return the lowest cost solution $\mathbf{x}^*$ satisfying the empirical feasibility criterion.

Otherwise, update the iteration count $k = k + 1$, update the tuning parameter $s^k = (s_{max} - s_{min})/2$ and go to Step 2.

Convergence of the bisection search is guaranteed to occur in $\lfloor \log_2 \left((s^0_{max} - s^0_{min})/\eta \right) \rfloor$ iterations.

3.3 CASE STUDY: APPLICATION TO CC DC OPF

We apply the tuning algorithm to a chance-constrained formulation of the optimal power flow (OPF) problem to handle uncertainties resulting renewable generation and load variability. We first present the formulation of the chance-constrained DC optimal power flow (CC DC OPF) with both single and joint chance constraints, which is based on [BCH14; Vra+13a; Roa+15b]. Then, we describe the details of the application of the tuning algorithm, including a detailed formulation of the approximate optimization problem, and present numerical simulation results for the IEEE RTS96 24-bus system to demonstrate that our method is computationally efficient and enforces chance constraints without over-conservatism.

3.3.1 *Power systems modeling*

We represent a power system using an undirected graph $G = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of buses with $n = |\mathcal{N}|$ and $\mathcal{L}$ is the set of lines with $l = |\mathcal{L}|$. Without loss of generality, we assume that each bus has one generator, $g \in \mathcal{G} \subseteq \mathbb{R}^n$, one load $d \in \mathcal{D} \subseteq \mathbb{R}^n$, and one uncertainty source denoted by random variable $\xi \in \Xi \subseteq \mathbb{R}^n$. Multiple (or zero) generators,

loads, or uncertainty sources are handled by summation (or setting the respective elements to zero).

The optimization problem aims to minimize total power generation costs while enforcing physical system constraints on power flow balance and generator and line limits. The decision variable is controllable (i.e., dispatchable) generation, represented using the variable $\mathbf{p}_g \in \mathbb{R}^n$. We consider the DC linearized approximation of the power flow equations, where we assume voltages are constant at magnitude 1 per unit, angle differences are small, and the system is lossless.

Uncertainty modeling

We consider the uncertainty, denoted $\xi \in \mathbb{R}^n$, in this problem to be fluctuation from the forecasted values of renewable generation and load uncertainty. Let the covariance of the fluctuations be denoted by $\Sigma_\xi = \mathrm{Cov}[\xi] \in \mathbb{R}^{n\times n}$. We do not assume that the uncertainty sources are independent, allowing the off-diagonals of the covariance matrix to be non-zero. We denote the total power mismatch resulting from uncertainty sources across all buses using

$$\Omega_\xi = \sum_{i\in\mathcal{N}} \xi_i. \tag{3.11}$$

Power balance and generation control

Power systems operations requires power production and consumption to be balanced. The total power mismatch Ω_ξ must be balanced by adjustments in controllable generation. We model these adjustments using an affine control policy based on actions of the automatic generation control (ACG), where Ω_ξ is divided amongst generators according

to participation factors $\alpha \in [0,1]^n$ [Vra+13a]. We assume each generator contributes according to its maximum nominal output, i.e.,

$$\alpha_i = \frac{p_{g,i}^{max}}{\sum_{j \in \mathcal{G}} p_{g,j}^{max}}, \quad \forall i \in \mathcal{G}, \tag{3.12}$$

where $p_{g,i}^{max}$ is the maximum generation limit of generator $i \in \mathcal{G}$. Thus, for each generator $i \in \mathcal{G}$, the actual generation (including the adjustment) is $\mathbf{p}_{g,i} - \alpha_i \Omega_\xi$. The total power balance can be enforced with the following constraint

$$\sum_{i \in \mathcal{N}} \mathbf{p}_{g,i} - \alpha_i \Omega_\xi - d_i + \xi_i = 0. \tag{3.13}$$

Because $\sum_{i \in \mathcal{G}} \alpha_i = 1$ guarantees that any forecast deviation ξ is automatically balanced out with an equal adjustment in generation, it is sufficient to guarantee power balance for $\xi = 0$, i.e.,

$$\sum_{i \in \mathcal{N}} \mathbf{p}_{g,i} - d_i = 0. \tag{3.14}$$

Power flows

We use a linear DC approximation to represent the power flow on lines $ij \in \mathcal{L}$ connecting buses $i, j \in \mathcal{N}$. We define $M \in \mathbb{R}^{l \times n}$ to be the matrix of power transfer distributions factors (PTDFs) [CWWoo], which relates the changes in active power flow to power injections at buses. Power flow on line $ij \in \mathcal{L}$ can be expressed as

$$p_{ij} = M_{(ij,\cdot)}(\mathbf{p}_g - \alpha \Omega_\xi + \xi - d), \tag{3.15}$$

where $M_{(ij,\cdot)}$ is the row of M corresponding to line ij.

3.3.2 *Problem formulation*

We formulate the single and joint chance-constrained OPF problems using the power systems model described above. The problems are formulated as a single period problem, where the time-step corresponding to the period can be, e.g., 15 minutes.

Objective function

Our objective in this problem is to minimize the total generation cost of the scheduled generation $\mathbf{p}_g$. Generation costs are modeled using a quadratic cost function,

$$c(\mathbf{p}_g) = \sum_{i \in \mathcal{G}} \left(c_{2,i} \mathbf{p}_{g,i}^2 + c_{1,i} \mathbf{p}_{g,i} + c_{0,i} \right), \tag{3.16}$$

where $c_{2,i}, c_{1,i}$ and $c_{i,0}$ are the quadratic, linear and constant cost coefficients corresponding to generator $i \in \mathcal{G}$.

Single chance-constrained OPF problem

The OPF problem with single chance constraints (SCC OPF) require each constraint to be satisfied individually with separate acceptable violation probabilities. In this case, we

assume that all chance constraints are required to hold with the same violation probability $\epsilon \in [0, 1]$. We formulate the problems as follows

$$\min_{\mathbf{p}_g} \quad c(\mathbf{p}_g) \tag{3.17a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} \mathbf{p}_{g,i} - d_i = 0 \tag{3.17b}$$

$$\mathbb{P}_\xi\left(\mathbf{p}_{g,i} - \alpha_i \Omega_\xi \leqslant p_{g,i}^{max}\right) \geqslant 1 - \epsilon, \qquad \forall i \in \mathcal{G}, \tag{3.17c}$$

$$\mathbb{P}_\xi\left(\mathbf{p}_{g,i} - \alpha_i \Omega_\xi \geqslant p_{g,i}^{min}\right) \geqslant 1 - \epsilon, \qquad \forall i \in \mathcal{G}, \tag{3.17d}$$

$$\mathbb{P}_\xi\left(M_{(ij,\cdot)}(\mathbf{p}_g - \alpha\Omega_\xi + \xi - d) \leqslant p_{l,ij}^{max}\right) \geqslant 1 - \epsilon, \qquad \forall ij \in \mathcal{L}, \tag{3.17e}$$

$$\mathbb{P}_\xi\left(M_{(ij,\cdot)}(\mathbf{p}_g - \alpha\Omega_\xi + \xi - d) \geqslant -p_{l,ij}^{max}\right) \geqslant 1 - \epsilon, \qquad \forall ij \in \mathcal{L}. \tag{3.17f}$$

Constraint (3.17b) is the power balance constraint. The generation limits are enforced by (3.17c) and (3.17d), where $p_g^{max}, p_g^{min} \in \mathbb{R}_+^n$ are the upper and lower generation limits. The line flows are enforced by (3.17e) and (3.17f), where $p_{l,ij}^{max} \in \mathbb{R}_+^l$ denotes the maximum line flow.

We note that the generator chance constraints (3.17c), (3.17d) depend only on the total power mismatch Ω_ξ, which is a scalar random variable. As a result, all generators will adjust their generation output up or down in *perfect correlation*.

Here, an important note is that equality constraints are used to represent power balance (i.e., the condition that the generated power must equal the consumed power at all times). Power balance constraints must be satisfied at all times in order to adhere to standard laws of physics. However, in the context of chance-constrained problems, this is only possible if the values of the decision variables adapt as the random variables change. In this problem, these changes are represented using an affine control policy.

Joint chance-constrained OPF problem

The formulation with a joint chance constraint requires all constraints to be simultaneously enforced with a single acceptable violation probability, ϵ. The joint chance-constrained OPF problem (JCC OPF) can be formulated as

$$\min_{\mathbf{p}_g} \quad c(\mathbf{p}_g) \tag{3.18a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} \mathbf{p}_{G,i} - d_i = 0 \tag{3.18b}$$

$$\mathbb{P}_\xi \left(\begin{array}{l} \mathbf{p}_{g,i} - \alpha_i \Omega_\xi \leqslant p_{g,i}^{max}, \quad \forall i \in \mathcal{G}, \\ \mathbf{p}_{g,i} - \alpha_i \Omega_\xi \geqslant p_{g,i}^{min}, \quad \forall i \in \mathcal{G}, \\ M_{(ij,\cdot)}(\mathbf{p}_g - \alpha \Omega_\xi + \xi - d) \leqslant p_{l,ij}^{max}, \quad \forall ij \in \mathcal{L}, \\ M_{(ij,\cdot)}(\mathbf{p}_g - \alpha \Omega_\xi + \xi - d) \geqslant p_{l,ij}^{max}, \quad \forall ij \in \mathcal{L}. \end{array} \right) \geqslant 1 - \epsilon. \tag{3.18c}$$

3.3.3 *Application of tuning method*

Approximate problem formulation

Because the constraints in the CC OPF problem are linearly separable in the decision variables $\mathbf{p}_g$ and the uncertainty ξ, we can use the reformulation detailed by (3.3) applied to the single chance constraints. The resulting approximate problem is as follows

$$\min_{\mathbf{p}_g} \; c(\mathbf{p}_g) \tag{3.19a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} \mathbf{p}_{g,i} - d_i = 0 \tag{3.19b}$$

$$\mathbf{p}_{g,i} \leqslant p_{g,i}^{max} - s\|\alpha_i \mathbb{1}_{1,n} \Sigma_\xi^{1/2}\|_2 \qquad \forall i \in \mathcal{G} \tag{3.19c}$$

$$\mathbf{p}_{g,i} \geqslant p_{g,i}^{min} + s\|\alpha_i \mathbb{1}_{1,n} \Sigma_\xi^{1/2}\|_2 \qquad \forall i \in \mathcal{G} \tag{3.19d}$$

$$M_{(ij,\cdot)}(\mathbf{p}_g - d) \leqslant p_{l,ij}^{max} - s\|M_{(ij,\cdot)}(I - \alpha \mathbb{1}_{1,m})\Sigma_\xi^{1/2}\|_2 \qquad \forall ij \in \mathcal{L} \tag{3.19e}$$

$$M_{(ij,\cdot)}(\mathbf{p}_g - d) \geqslant -p_{l,ij}^{max} + s\|M_{(ij,\cdot)}(I - \alpha \mathbb{1}_{1,m})\Sigma_\xi^{1/2}\|_2 \qquad \forall ij \in \mathcal{L}. \tag{3.19f}$$

As previously discussed, we propose that by accurately tuning the safety parameter s, this approximate reformulation can be carried over to solve the joint chance-constrained version of the problem as well.

Evaluation

After solving the approximate optimization problem (3.19) and obtaining a candidate solution $\mathbf{p}_g$, we evaluate the solution as detailed in Section 3.2.2. For SCC OPF, we evaluate the violation probabilities of all single chance constraints and take the maximum (i.e., worst case) violation probability amongst all constraints (3.8), denoted $\hat{E}_{max}(\mathbf{p}_g, \Xi)$, to use in the tuning process. For JCC OPF, we evaluate the empirical joint violation probability, denoted $\hat{E}_{joint}(\mathbf{p}_g, \Xi)$, according to (3.10) to use in the tuning process.

Tuning

We follow the same bisection tuning steps described in Section 3.2.3.

3.3.4 *Test system*

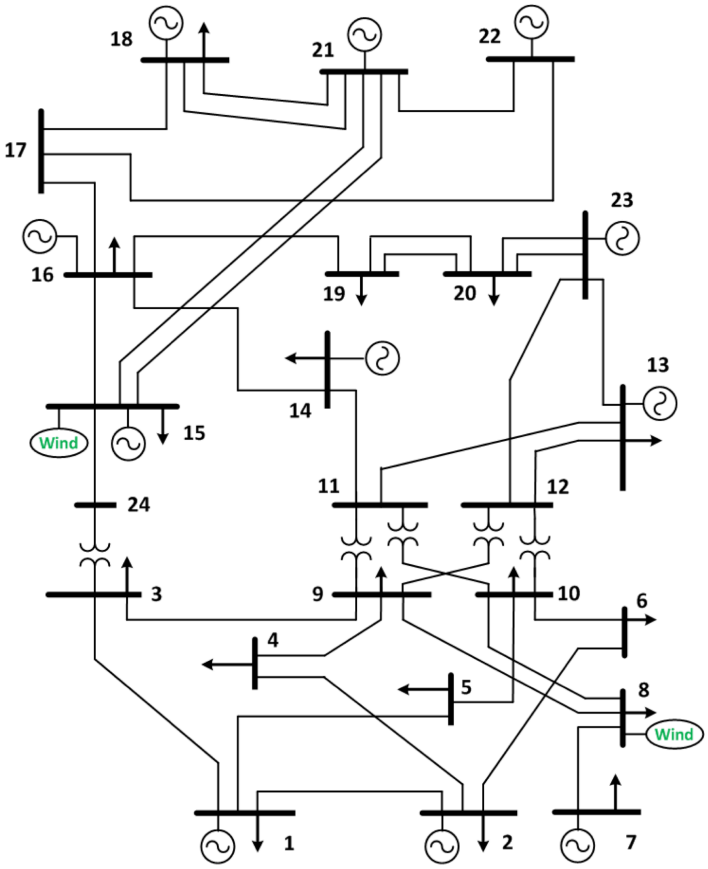


Figure 3.2: Modified IEEE RTS96 24-bus system with uncertainty fluctuations at bus 8 and bus 15.

We evaluate our method on the IEEE RTS96 24-bus system [Gri+99], illustrated in Figure 3.2. We make the following modifications to the existing system parameters: (i) All line capacities are reduced by 70%, (ii) the minimum output is set to zero on all dispatchable generators, and (iii) the maximum output is doubled on all generators.

Tuning input			In-sample results						Out-of-sample	
Problem	Uncertainty	ϵ	Iterations	Cost	s	s_{true}	$\hat{E}_{max}$	$\hat{E}_{joint}$	$\hat{E}_{max}$	$\hat{E}_{joint}$
SCC OPF	Gaussian	0.1	10.7	42201.6	1.3012	1.2816	0.1000	0.0976	0.2960	0.2947
		0.05	9.4	42376.1	1.6676	1.6449	0.0501	0.0483	0.1585	0.1577
		0.01	9.6	42709.0	2.3624	2.3263	0.0100	0.0095	0.0338	0.0338
	Non-Gaussian	0.1	10.4	42799.6	1.3376	-	0.1001	0.1007	0.3031	0.3044
		0.05	9.6	43105.4	1.6677	-	0.0501	0.0495	0.1597	0.1609
		0.01	8.9	43680.4	2.2844	-	0.0100	0.0095	0.0274	0.0275
JCC OPF	Gaussian	0.1	15.6	42485.6	1.8971	-	0.0307	0.0296	0.1001	0.1001
		0.05	14.5	42632.9	2.2054	-	0.0149	0.0141	0.0501	0.0500
		0.01	12.1	42918.5	2.8014	-	0.0032	0.0027	0.0100	0.0100
	Non-Gaussian	0.1	14.4	43284.4	1.8585	-	0.0316	0.0307	0.1001	0.1000
		0.05	14.6	43507.0	2.1008	-	0.0167	0.0161	0.0500	0.0501
		0.01	13.3	43924.0	2.5538	-	0.0042	0.0038	0.0100	0.0101

Table 3.1: Results for the bisection tuning method for SCC OPF and JCC OPF using Gaussian and non-Gaussian uncertainty data. Each value in the table is averaged over 20 replications.

We add uncertainty sources to represent load variability on buses 8 and 15 and synthetically generate samples that follow Gaussian and non-Gaussian distributions. For Gaussian data, samples are generated from a multivariate Gaussian distribution with zero mean, standard deviations of 9.4 MW (bus 8) and 13.1 MW (bus 15), and a correlation coefficient of $\rho = 0.2$. For non-Gaussian data, samples are generated by mixing the following distributions: (i) a zero-mean multivariate Gaussian with standard deviations of 7 MW and 14 MW and a correlation coefficient of 0.5, (ii) a zero-mean multivariate Gaussian with standard deviations of both 6 MW and a correlation coefficient of 0.1, and (iii) a uniform distribution on the interval $[-30, 30]$.

3.3.5 *Numerical results*

We consider four different cases: SCC OPF with Gaussian data, SCC OPF with non-Gaussian data, JCC OPF with Gaussian data, and JCC OPF with non-Gaussian data. For all experiments, we use desired violation probabilities $\epsilon = \{0.1, 0.05, 0.01\}$. We run 20 replications with tolerance $\eta = 10^{-4}$. We use $|\Xi| = 10,000$ samples of the uncertainty ξ during the tuning process. We perform an additional out-of-sample single and joint violation probability evaluation using $|\Xi_{\text{oos}}| = 100,000$ samples of the uncertainty. All simulations are implemented in Julia with JuMP [DHL17] and PowerModels [Cof+18], solved using Gurobi [Gur21], and run on a standard laptop computer.

We present all results in Table 3.1, which shows the number of iterations, cost of generation of the resulting set point, safety parameter values s, and the in- and out-of-sample maximum (i.e., worst case) violation probability across all single chance constraints, $\hat{E}_{\max}(\mathbf{p}_g, \Xi)$ and $\hat{E}_{\max}(\mathbf{p}_g, \Xi_{\text{oos}})$, and the joint chance constraint violation probability, $\hat{E}_{\text{joint}}(\mathbf{p}_g, \Xi)$ and $\hat{E}_{\text{joint}}(\mathbf{p}_g, \Xi_{\text{oos}})$. All numbers in the table represent average values across the 20 replications. For the case where the SCC OPF problem with Gaussian uncertainty data, the "true" s value is also reported by evaluating $s_{\text{true}} = \Phi^{-1}(1 - \epsilon)$.

Proof-of-concept: SCC OPF with Gaussian data

To demonstrate the accuracy and efficiency of our method, we perform a proof of concept study by evaluating the performance of tuning on SCC OPF using Gaussian uncertainty data. Note that because we use a single tuning parameter for all constraints, the violation probability of each of the constraints may vary. We choose to tune according to the worst case violation probability. However, because the uncertainty here is Gaussian, the violation should theoretically be the same across all constraints if the same s is used. In this case, we can directly calculate the true safety parameter value by evaluating $s_{\text{true}} = \Phi^{-1}(1-\epsilon)$. We test whether our method identifies a parameter $s \approx s_{\text{true}}$.The desired epsilon values $\epsilon = \{0.1, 0.05, 0.01\}$ correspond to true safety parameter values $s_{\text{true}} = \{1.2816, 1.6449, 2.3263\}$.

The numerical results are shown in the top section of Table 3.1. We observe our algorithm terminates when the worst case empirical violation probability $\hat{E}_{\max}(\mathbf{p}_g, \Xi)$ very close to ϵ. Furthermore, using $10,000$ samples to evaluate the violation probability in the tuning step is sufficient to achieve out-of-sample violation probabilities that are very close to the desired values. Finally, we conclude that the final tuning parameter s value is close to the true parameter value s_{true}, verifying the accuracy of our method.

An interesting but subtle observation is that, for every ϵ, the tuning parameter values obtained with the algorithm are typically always slightly greater (i.e., more conservative) than s_{true}. This is true not only for the average s presented in Table 3.1, but for *each* of the 20 algorithm replications. To explain this behavior, we look at the results of a single run for the desired violation probability $\epsilon = 0.1$ in more detail, as shown in Figure 3.3 (right). This figure plots the empirical violation probability for all the active constraints in the problem, which includes several generator maximum and minimum limits (aggregated into one bar because they are perfectly correlated and hence have the same violation probability), as well as three line limits. The light blue line shows the desired violation probability $\epsilon = 0.1$, while the blue and grey bars represent the empirical

violation probability of each constraints, as observed in the tuning and out-of-sample evaluation, respectively.

We see that the tuning algorithm chooses the value of the tuning parameter s to ensure that the worst case single violation probability $\hat{E}_{max}(\mathbf{p}_g, \Xi) \leqslant 0.1$, leading to a conservative result. In fact, because we are using Gaussian data, we can calculate the true violation probability corresponding to the tuned value of s, $\epsilon_s = 1 - \Phi(s) \approx 0.095$. We observe that the out-of-sample violation probabilities of each constraint, illustrated with the grey bars, are all close to this true violations probability.

SCC-OPF with non-Gaussian data

We next study how the tuning method performs for SCC-OPF using non-Gaussian data. We observe that the number of iterations is not impacted by the type of uncertainty data, and that the algorithm also terminates with $\hat{E}_{max}(\mathbf{p}_g, \Xi) \approx \epsilon$ for the non-Gaussian data. We again observe that $10,000$ samples in the tuning algorithm is sufficient to obtain an $\hat{E}_{max}(\mathbf{p}_g, \Xi_{oos})$ that is very close to the desired values.

Figure 3.3 (right) shows the result for a single run in more detail. We observe larger variations in the in- and out-of-sample violation probabilities amongst the constraints compared to using Gaussian data. Because we do not have the Gaussian distribution assumption, it is no longer true that the same safety parameter will lead to the same true violation probability for all constraints. Therefore, the bars representing the in- and out-of-sample violation probabilities have greater variability that in the Gaussian case.

JCC-OPF with Gaussian and non-Gaussian data

We investigate the behavior of the algorithm for JCC-OPF. We use the same case study set up and the same Gaussian and non-Gaussian distributions as above, but tune the parameter s to achieve a desired *joint* violation probability of $\epsilon = \{0.05, 0.1, 0.15\}$. The results are shown in the bottom half of Table 3.1.

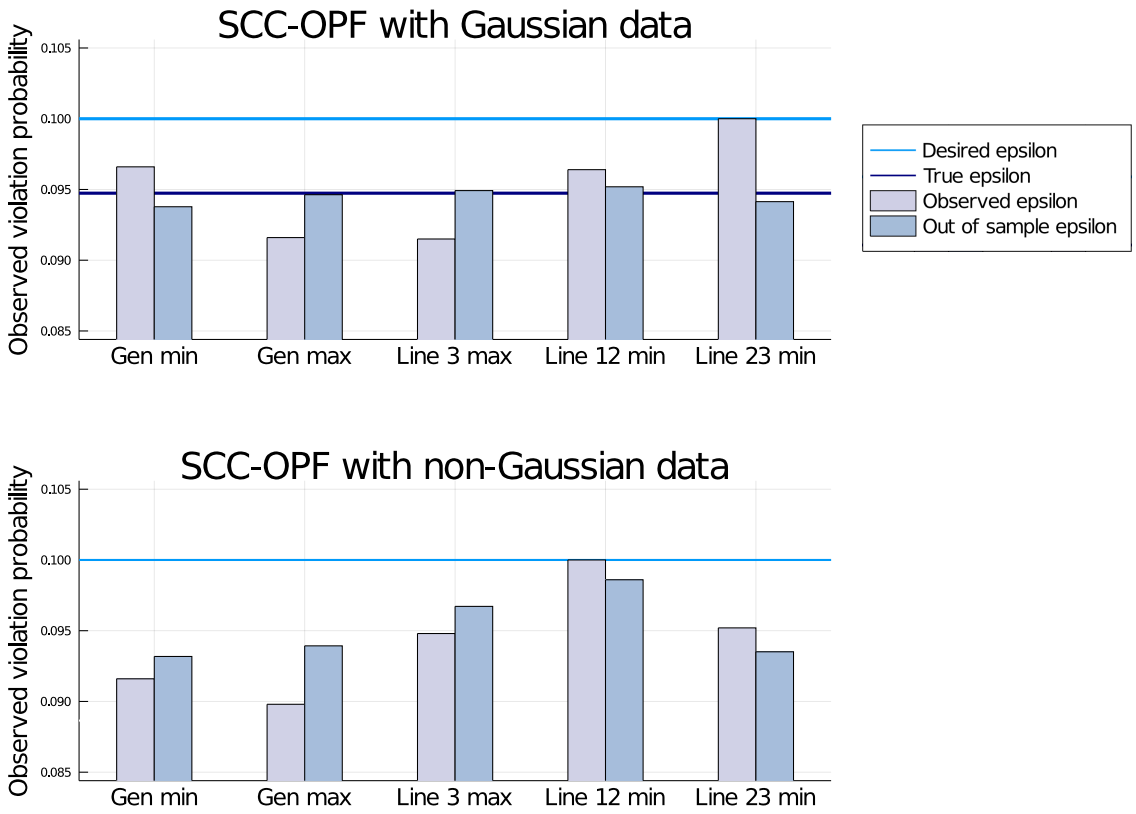


Figure 3.3: Empirical violation probability for the active constraints in the SCC OPF problem, with Gaussian (left) and non-Gaussian (right) data. The active constraints includes several generator maximum and minimum limits (represented by one bar because they are perfectly correlated and hence have the same violation probability), as well as three line limits. The blue and grey bars show the empirical violation probability of each constraint, as observed in the tuning and out-of-sample evaluation respectively. The light blue line shows ϵ, while the dark blue line is $\epsilon_s = 1 - \Phi(s)$.

We observe that the number of iterations remains similar, but is slightly higher compared to the number of iterations for the single chance-constrained case. We see that the algorithm manages to determine an s value that meets the joint violation probability $\hat{E}_{\text{joint}}(\mathbf{p}_g, \Xi)$ exactly.

To gain some more insight into the solutions for the JCC OPF, Figure 3.4 plots the individual and joint constraint violation levels for JCC OPF with Gaussian and non-Gaussian data. We again observe that the violation probability of individual constraints are distributed more evenly in the case of Gaussian data, while the non-Gaussian data lead to larger variations despite all constraints sharing the same s. For both data types, the

sum of the violation probabilities of the individual chance constraints exceeds the joint violation probability, i.e., $\sum_{j\in\mathcal{C}} \hat{E}_j(\mathbf{p}_g, \Xi) \geqslant \hat{E}_{joint}(\mathbf{p}_g, \Xi)$. This indicates that our algorithm is able to account for some of the correlation between different constraints, which is not done with methods using, e.g., Boole's inequality, to decompose joint constraints into individual chance constraints. As a result, we are able to obtain a less conservative solution.

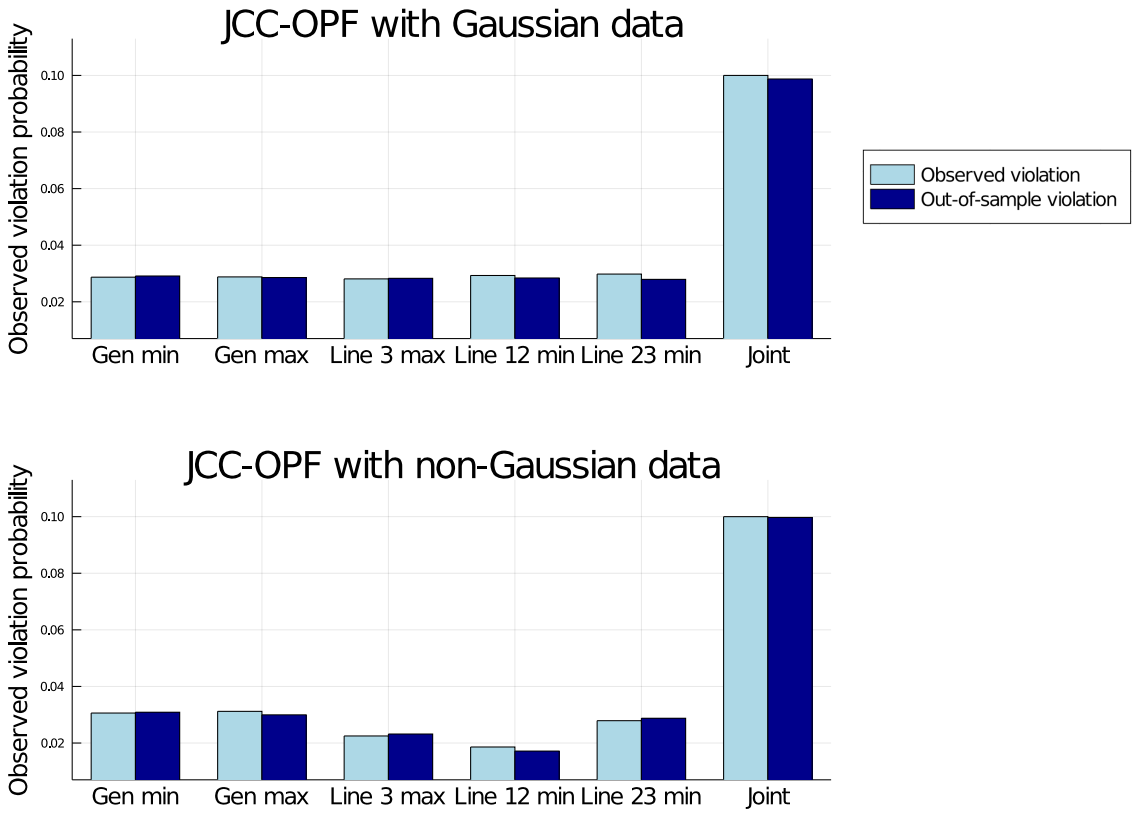


Figure 3.4: Empirical violation probability for individual active constraints in the JCC OPF problem, as well as the joint violation probability. We show results for both Gaussian data (left) and non-Gaussian data (right). The grey and blue bars represent the empirical violation probability of each constraints, as observed in the tuning and out-of-sample evaluation, respectively.

3.4 CONCLUSION

In this chapter, we proposed a data-driven tuning method for solving single and joint chance-constrained problems that iterates between solving an approximate optimization model and using a posteriori sample-based evaluations to adjust the model formulation. The key idea is to formulate a numerically tractable approximation of the chance-constrained problem and identify suitable tuning parameters which can be tightened to obtain a more secure (but more costly) solution or relaxed to obtain a less secure (but less costly) solution. Out method does not make any distributional assumptions on the uncertainty, but assumes access to a set of representative samples of the uncertainty.

In general, the tuning methods of interest consist of three major components: (i) a computationally lightweight approximate optimization model parameterized by a tunable parameter, (ii) an evaluation step, which evaluates candidate solutions using available uncertainty samples, and (iii) a tuning step. There are many ways of defining the approximated problem and tuning parameters as well as many possible strategies to adapt the parameters in each iteration. We discuss considerations for how to determine suitable versions of each component and an example algorithm, which uses an approximate model derived from existing analytical chance-constrained reformulations and a bisection search tuning scheme. This algorithm formulation allows us to utilize information from samples without directly including them in the optimization problem itself, preserving the computational tractability of the method.

In the case study, we formulate a single and joint chance-constrained version of the DC OPF problem and evaluate our proposed bisection search-based tuning algorithm on a modified version of the IEEE RTS 96 24-bus system. Our results indicate that the tuning algorithm is able to obtain solutions that meet the desired violation probability exactly for both the single and joint chance-constrained formulations, when either Gaussian or non-Gaussian uncertainty data is considered.

4 TUNING WITH PROBABLISTIC GUARANTEES

In this chapter, we propose a method to obtain feasibility guarantees for tuning-based algorithms, such as the one proposed in Chapter 3, by using a two-step approach. This approach involves a *solution generation step*, where we generate a solution by tuning the problem parameters, and a *solution verification step*, where we use a posteriori evaluations to obtain probabilistic guarantees on the resulting solution. Our main contribution is to analyze the relationships between the empirical feasibility criterion applied in the tuning step and the rigorous feasibility criterion applied in the verification step, and discuss how to design the tuning step such that the algorithm is likely to return a feasible solution. Our results are not tailored to a specific tuning method, but can be used with any tuning-based solution generation method.

The material presented in this chapter is based on the publication [HR22].

4.1 MOTIVATION AND CHALLENGES

While tuning-based algorithms for chance-constrained optimization have been empirically observed to perform well [Old+15; RA17; PLW19; MMD20; GVH20; HR20], they often do not provide feasibility guarantees for the final solution. A particular shortcoming of existing tuning methods, including the algorithm proposed in Chapter 3, is that often they do not consider how re-solving a problem multiple times using the same data may compromise the probabilistic guarantees on the resulting solution. This is primarily due to two inaccuracies associated with empirically approximating the violation probability in

an iterative manner. We note that these challenges are not unique to our tuning approach, but also affect other methods that rely on tuning to obtain solutions to chance constrained problems.

We propose a two-step tuning algorithm that can be used to provide probabilistic feasibility guarantees on solutions attained from a wide range of problem formulations and tuning strategy variations. The advantages include:

(i) *Few limitations on problem structure*: The use of simple approximate models that do not provide feasibility guarantees allows us to *handle problems with undesirable or complicated structure* (e.g., non-convex optimization problems).

(ii) *Computational tractability*: While simpler approximate formulations may be less accurate, as they does not capture as granular system detail, they are much more *computationally tractable* and allow us to take advantage of the efficiency and scalability of commercial, off-the-shelf solvers. The possible model inaccuracy or mis-match can be compensated for by leveraging large amounts of data for tuning.

(iii) *Decoupling uncertainty evaluation and optimization*: Unlike other methods, we do not incorporate the uncertainty data within the problem itself but rather perform an evaluation of the uncertainty after obtaining a candidate solution. This allows us to leverage large amounts of data without increasing the computational complexity of the optimization problem being solved. Our method is constructed such that the more data we have access to, the less conservative the solutions we can obtain.

In order to obtain feasibility guarantees for tuning-based algorithms for chance-constrained problems, we must address the two following main sources of inaccuracy associated with empirically approximating the violation probability in an iterative manner.

4.1.1 *Challenge 1: Finite number of samples*

The first challenge arises because we use a finite number of samples, and the empirical violation probability $\hat{E}(\mathbf{x}, \Xi)$ may differ from the true violation probability $E(\mathbf{x})$. Therefore, even if a solution satisfies $\hat{E}(\mathbf{x}, \Xi) \leqslant \epsilon$, we cannot conclude $E(\mathbf{x}) \leqslant \epsilon$. To combat this challenge, we can use Hoeffding's inequality to characterize the relationship between the true and empirical violation probability of a solution.

Theorem 4.1. *(Hoeffding's inequality [Hoe63]) For bounded i.i.d. random variables* $Y_1, \ldots, Y_N \in [0, 1]$ *and* $t \geqslant 0$,

$$\mathbb{P}\left(\mathbb{E}[Y] < \frac{1}{N}\sum_{i=1}^{N} Y_i + t\right) > 1 - \exp(-2Nt^2). \tag{4.1}$$

Let δ denote the desired confidence level. By choosing

$$t \geqslant \sqrt{\frac{1}{2N}\ln\frac{1}{\delta}}, \tag{4.2}$$

we guarantee with confidence $1-\delta$ that the expectation is upper bounded by the sample mean plus margin t.

We can apply this to candidate solutions $\mathbf{x}$ obtained from . For a solution $\mathbf{x}$ and set of i.i.d. samples Ξ, let $Y_i = Y(\mathbf{x}, \xi^{(i)})$ and $\mathbb{E}[Y] = \mathbb{E}_\xi[Y(\mathbf{x}, \xi)]$. Note that $Y(\mathbf{x}, \xi^{(i)}) \in [0, 1]$ and $\mathbb{E}_\xi[Y(\mathbf{x}, \xi)] = E(\mathbf{x})$. Thus, given a desired confidence level $1-\delta$ and margin t chosen according to (4.2), we evaluate the empirical feasibility $\hat{E}(\mathbf{x}, \Xi)$ and check if

$$\hat{E}(\mathbf{x}, \Xi) + t \leqslant \epsilon \tag{4.3}$$

holds. If satisfied, by Hoeffding's inequality, $\mathbf{x}$ is declared to be empirically feasible to the chance-constrained problem with confidence $1-\delta$.

4.1.2 *Challenge 2: Non-i.i.d. realizations of* $Y(\mathbf{x}, \xi^{(i)})$

The second less obvious, but more complex challenge is related to the fact that we evaluate $\hat{E}(\mathbf{x}, \Xi)$ multiple times on the same set of samples, Ξ. After obtaining the first candidate solution $\mathbf{x}_1$, the evaluated indicator random variables $Y(\mathbf{x}_1, \xi^{(i)})$ are i.i.d. because the uncertainty samples $\xi^{(i)}$ are i.i.d. uncertainty samples. However, in each subsequent iteration $j > 1$, because s is adjusted based on the empirical violation probability of the previous solution, $\mathbf{x}_j$ is a function of the previous solution $\mathbf{x}_{j-1}$ as well as the samples. Consequently, the random variables $Y(\mathbf{x}_j, \xi^{(i)})$ are no longer i.i.d. for all subsequent iterations $j > 1$. Therefore, the empirical violation probability estimated in these iterations $\hat{E}(\mathbf{x}_1, \Xi)$ is not an unbiased estimator of $E(\mathbf{x})$. Consequently, in iterations $j > 1$, we cannot apply Hoeffding's inequality to assess feasibility.

Note that this i.i.d. issue could be resolved by drawing a new sample set in every iteration. However, this would require much more data and may pose convergence issues for the algorithm. We therefore limit our analysis to algorithms where the same sample set is used throughout for tuning.

4.2 TWO-STEP TUNING APPROACH

These aforementioned challenges associated with tuning chance-constrained formulations are not unique to our tuning approach proposed in Chapter 3, but also affect other methods that rely on tuning to obtain solutions to chance constrained problems. In the following, we propose a method to address these drawbacks.

To address the issue of non-i.i.d. sample data in the tuning process, we propose a two-step method to recover feasibility guarantees for tuning-based methods. This two-step method uses two i.i.d. drawn sample sets $\Xi_1 = \{\xi^{(1)}, \ldots, \xi^{(N_1)}\}$ and $\Xi_2 = \{\xi^{(1)}, \ldots, \xi^{(N_2)}\}$,

where $N_1 + N_2 = N$, where N is the total number of available samples. The algorithm consists of: (i) a *solution generation step* that utilizes a tuning method to obtain a solution to the chance-constrained problem (2.1) using the first sample set Ξ_1, and (ii) an a posteriori *solution verification step* to guarantee solution feasibility using the second sample set Ξ_2.

4.2.1 *Solution generation step*

In the *solution generation step*, we run a tuning algorithm, but instead of using the empirical feasibility criterion (3.10), we define a new empirical feasibility criterion that modifies our original criterion (3.10) by including a *tuning margin* $t_{\text{tune}} \in (0, \epsilon]$

$$\hat{E}(\mathbf{x}, \Xi_1) + t_{\text{tune}} \leqslant \epsilon. \tag{4.4}$$

We run the tuning algorithm to termination using this criterion and return the lowest cost solution that satisfies (4.4), denoted using $\mathbf{x}^*$. Note that even if we chose t_{tune} according to (4.2), this solution will not have the probabilistic feasibility guarantee given by Hoeffding's inequality (Theorem 4.1). This is a result of Challenge 2 (4.1.2), i.e., that the indicator random variables $Y(\mathbf{x}^*, \xi^{(i)})$ evaluated on the resulting solution $\mathbf{x}^*$ are not i.i.d. and therefore do not satisfy the assumptions required for Hoeffding's inequality.

4.2.2 *Solution verification step*

In the *solution verification step*, we use the other sample set Ξ_2 to perform an a posteriori verification of the feasibility criterion for the solution $\mathbf{x}^*$. We choose a desired confidence level $\delta_t \in [0, 1]$ and *verification margin* $t \in \mathbb{R}_+$ according to Hoeffding's inequality (4.2),

based on the sample size $N_2 = |\Xi_2|$. We then check whether the solution x^* satisfies the following criterion

$$\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon. \tag{4.5}$$

If the criterion is not satisfied, we discard the solution and declare that the algorithm did not find a solution. Otherwise, if satisfied, we can declare the solution as feasible to the original chance constraint (2.1) with probability at least $1 - \delta_t$. This gives rise to the following *a posteriori probabilistic guarantee*.

Claim 4.2. *(**A posteriori feasibility guarantee**, [HR22].) Let $\mathbf{x}^*$ be a solution obtained from the two-step tuning algorithm. If $\mathbf{x}^*$ satisfies the a posteriori solution verification criterion (4.5)*

$$\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon,$$

where $t \geqslant \sqrt{\frac{1}{2N} \ln \frac{1}{\delta_t}}$, then the solution $\mathbf{x}^*$ is feasible to the original chance constraint (2.1) with probability at least $1 - \delta_t$, i.e.,

$$\mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon \mid \hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \right) \geqslant 1 - \delta_t. \tag{4.6}$$

Proof. The proof follows directly from Hoeffding's inequality [Hoe63]. □

We note that this a posteriori feasibility result applies to any solution $\mathbf{x}^*$, regardless of the method used to obtain it. However, an issue that arises is that in the event that the verification feasibility criterion (4.5) is not satisfied by the solution $\mathbf{x}^*$, we cannot "go back" and choose a different t_{tune}. As a result, $\mathbf{x}^*$ would not have any feasibility guarantees with respect to the original chance constraint (2.1b). Thus, the main remaining challenge is *how to choose the tuning margin* t_{tune} before we begin the solution generation step such that the solution will likely satisfy (4.5). By relating these two margins, we are able to take a step towards a priori feasibility guarantees for the tuning-based algorithms. This is analyzed in the subsequent section.

4.3 CHOOSING THE TUNING MARGIN

To support our subsequent analysis, we first provide a few basic claims and assumptions.

First, consider the probability that a solution $\mathbf{x}^*$ obtained from the tuning process, which will satisfies the tuning feasibility criterion (4.4) by design, will satisfy the a posteriori verification feasibility criterion (4.5).

Claim 4.3. *Let Ξ_1 and Ξ_2 be the tuning and verification sample sets, respectively. A solution $\mathbf{x}^*$ obtained from the solution generation step satisfies the following inequality*

$$\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\epsilon\middle|\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\leqslant\epsilon\right)\geqslant\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\right). \quad (4.7)$$

Proof. The left hand side denotes the probability in which the verification criterion $\hat{E}(\mathbf{x},\Xi_2)+t\leqslant\epsilon$ holds for a solution $\mathbf{x}^*$ given that the tuning criterion holds $\hat{E}(\mathbf{x},\Xi_2)+t_{tune}\leqslant\epsilon$. We can partition this into the situation in which: (i) $\hat{E}(\mathbf{x},\Xi_2)+t\leqslant\hat{E}(\mathbf{x},\Xi_2)+t_{tune}\leqslant\epsilon$ or (ii) $\hat{E}(\mathbf{x},\Xi_2)+t_{tune}\leqslant\hat{E}(\mathbf{x},\Xi_2)+t_{tune}\leqslant\epsilon$. Therefore, we obtain the following upper bound

$$\begin{aligned}
&\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\epsilon\middle|\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\leqslant\epsilon\right)\\
&=\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\leqslant\epsilon\right)+\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\leqslant\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\epsilon\right)\\
&\geqslant\mathbb{P}_\xi\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\hat{E}(\mathbf{x}^*,\Xi_1)+t_{tune}\right),
\end{aligned}$$

where the third line results from the non-negativity of probabilities. □

The results detailed in the remainder of the section rely on the following assumption:

Assumption 1. *(**Approximately unbiased estimator**.) We assume that it is a good approximation to treat $\hat{E}(\mathbf{x}^*,\Xi_1)$ as an unbiased estimator of the violation probability ϵ of $\mathbf{x}^*$.*

In other words, we assume that $\mathbf{x}^*$ is not dependent on Ξ_1, which implies that $\hat{E}(\mathbf{x}^*, \Xi_1)$ is an unbiased estimator of the violation probability ϵ of $\mathbf{x}^*$. We recognize that this assumption contradict the issue we highlighted with "Challenge 2" in Section 4.1.2 and therefore the following results do not rigorously provide a priori feasibility guarantees for the two-step tuning algorithm. However, we emphasize that the a posteriori feasibility guarantee (Claim 4.2) holds regardless of whether this Assumption 1 is satisfied.

Furthermore, we conjecture (and later empirically verify) that the bias of $\hat{E}(\mathbf{x}^*, \Xi_1)$ is typically small and use Assumption 1 in our analysis. The subsequent results provide guidance on how to a priori choose the tuning and verification margins such that the a posteriori feasibility guarantee will be likely to hold in practice and takes a step towards providing an a priori guarantee on solution feasibility for this two-step tuning algorithm.

In the following, we analyze three choices of tuning and verification margins, t_{tune} and t, and how they impact the probability of passing the a posteriori verification test (4.5). In the first two cases, we assume the samples are split equally between the two steps, i.e., $N_1 = N_2 = N/2$, where N is the total number of samples available. In the third case, we let $N_1 \neq N_2$, which results in margins that are parameterized the number of samples, i.e., $t_{\text{tune}}(N_1)$ and $t(N_2)$.

4.3.1 *Equal margins*

In the first scenario, we use an equal number of samples and equal margins in each step, i.e., $t_{\text{tune}} = t$.

Claim 4.4. *(**Equal margins**.) Let the margins be $t_{tune} = t$ and the desired violation probability be ϵ. Under Assumption 1, the probability that a solution $\mathbf{x}^*$ satisfying the tuning feasibility criterion (4.4) will satisfy the a posteriori feasibility criterion (4.5) is at least 0.5, i.e.,*

$$\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*, \Xi_2) + t > \epsilon \middle| \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} > \epsilon\right) \geqslant 0.5. \tag{4.8}$$

Proof.

$$\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*, \Xi_2) + t > \epsilon \middle| \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} > \epsilon\right) \tag{4.9}$$

$$\geqslant \mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*, \Xi_2) + t > \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}}\right) \tag{4.10}$$

$$= \mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*, \Xi_2) > \hat{E}(\mathbf{x}^*, \Xi_1)\right) \tag{4.11}$$

$$= 0.5. \tag{4.12}$$

The first inequality results from Claim 4.3. The first equality holds due to $t = t_{\text{tune}}$ and the second equality holds from Assumption 1. □

Note that if Assumption 1 is inaccurate, the second equality will not hold. The result will depend on which direction that tuning biases the evaluated empirical violation probability.

4.3.2 *Unequal margin*

In many cases, we want a higher probability that 0.5 of passing the verification feasibility criterion (4.5). Recognize that it does not make sense for the tuning margin to be less than the verification margin, $t_{\text{tune}} < t$, since we would end up with a lower probability than 0.5. Instead, we define $t_{\text{tune}} = t + k$, where $k \in \mathbb{R}_+$ can be thought of as an additional margin that increases the probability of passing the a posteriori feasibility check.

Claim 4.5. *(**Unequal margins**.) Under Assumption 1, if we choose* $t_{tune} \geqslant t + k$ *with*

$$k \geqslant \sqrt{\frac{1}{N_1} \ln \frac{1}{\beta}}, \tag{4.13}$$

where $\beta \in [0, 1]$*, a solution* $\mathbf{x}^*$ *from the tuning step will satisfy the a posteriori feasibility criterion with* $1 - \beta$ *confidence, i.e.,*

$$\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \middle| \hat{E}(\mathbf{x}^* \Xi_1) + t_{tune} \leqslant \epsilon\right) \geqslant 1 - \beta. \tag{4.14}$$

Proof. First, we define the random variable

$$Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)})) := Y(\mathbf{x}^*, \xi^{(i)}) - Y(\mathbf{x}^*, \xi^{(j)}) \tag{4.15}$$

to represent whether two i.i.d. samples $\xi^{(i)}$ and $\xi^{(j)}$ both lead to violations or satisfaction of the constraint $g(\mathbf{x}^*, \xi) \leqslant 0$, i.e., $Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)})) = 0$ if both samples lead to the same outcome and $Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)})) = 1$ otherwise. If $\xi^{(i)}$ and $\xi^{(j)}$ are i.i.d. samples drawn form the same distribution and $\mathbf{x}^*$ is not dependent on either sample, then by Assumption 1,

$$\mathbb{E}_{\xi}\left[Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)}))\right] = \mathbb{E}_{\xi}\left[Y(\mathbf{x}^*, \xi^{(i)})\right] - \mathbb{E}_{\xi}\left[Y(\mathbf{x}^*, \xi^{(j)})\right] = 0. \tag{4.16}$$

Note that when Assumption 1 is not satisfied, then $\mathbb{E}_{\xi}[Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)}))] \neq 0$ and can be interpreted as the expected bias of $\hat{E}(\mathbf{x}^*, \Xi_1)$ as an estimator of $E(\mathbf{x}^*)$.

Next, we can bound the left hand side of 4.14 as follows

$$
\begin{aligned}
&\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\epsilon\middle|\hat{E}(\mathbf{x}^*,\Xi_1)+t_{\text{tune}}\leqslant\epsilon\right)\\
&\quad\geqslant\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x}^*,\Xi_2)+t\leqslant\hat{E}(\mathbf{x}^*,\Xi_1)+t_{\text{tune}}\right)\\
&\quad\geqslant\mathbb{P}_{\xi}\left(\hat{E}(\mathbf{x},\Xi_2)-\hat{E}(\mathbf{x},\Xi_1)\leqslant k\right)\\
&\quad=\mathbb{P}_{\xi}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}Y(\mathbf{x}^*,\xi^{(i)})-\frac{1}{N_1}\sum_{j=1}^{N_1}Y(\mathbf{x}^*,\xi^{(j)})\leqslant k\right)\\
&\quad=\mathbb{P}_{\xi}\left(\frac{1}{N_1}\sum_{i=1,j=1}^{N_1}Z(\mathbf{x}^*,\xi^{(i)},\xi^{(j)})\leqslant k\right)\\
&\quad=\mathbb{P}_{\xi}\left(\frac{1}{N_1}\sum_{i=1,j=1}^{N_1}Z(\mathbf{x}^*,\xi^{(i)},\xi^{(j)})-\mathbb{E}[Z(\mathbf{x}^*,\xi^{(i)},\xi^{(j)})]\leqslant k\right)\\
&\quad\geqslant 1-\exp(N_1k^2).
\end{aligned}
$$

The first inequality follows from Claim 4.3 and the second inequality follows from the definition of $t_{\text{tune}}=t+k$. The next equalities result from (3.7), (4.15), and (4.16). The final inequality results from Hoeffding's inequality [Hoe63], since $k>0$ and random variables $Z(\mathbf{x}^*,\xi^{(i)},\xi^{(j)}))$ are i.i.d. and bounded on $[-1,1]$. By choosing k according to (4.13), we obtain the $1-\beta$ confidence level. □

Therefore, from Claim 4.5, by choosing $k=\sqrt{(1/N_1)\ln(1/\beta)}$, a solution $\mathbf{x}^*$ generated from the tuning step will satisfy the a posteriori feasibility criterion with $1-\beta$ confidence.

Remark 4.6. *With $k=0$, we obtain $\beta=1$, which indicates that we have no confidence that our solution $\mathbf{x}^*$ will satisfy the a posteriori feasibility check. However, Claim 4.4 shows that $k=0$ actually provides a feasible solution with probability 0.5. This shows that the bound in Claim 4.5 is very conservative.*

4.3.3 *Unequal sample split*

We can also split the available samples unequally between the two steps, i.e., $N_1 \neq N_2$. Because $N_1 \neq N_2$, the bound in Claim 4.5 is not applicable due to the use of random variable $Z(\mathbf{x}^*, \xi^{(i)}, \xi^{(j)})$ in the derivation.

In this case, when defining the margins according to (4.2), the margins will be parameterized by the sample sizes, i.e.,

$$t_{\text{tune}}(N_1) = \sqrt{\frac{1}{2N_1} \ln \frac{1}{\delta}} \tag{4.17}$$

$$t(N_2) = \sqrt{\frac{1}{2N_2} \ln \frac{1}{\delta}}. \tag{4.18}$$

When $N_2 > N_1$, we expect less variance in the verification step evaluation of $E(\mathbf{x}^*, \Xi_2)$ than in the evaluation of $E(\mathbf{x}^*, \Xi_1)$. This is reflected in having a smaller margin in the verification step than in the tuning step, with $t(N_2) < t_{\text{tune}}(N_1)$. As a result, using more samples in the verification step has a similar effect as the use of the additional margin k in Section 4.3.2 and should increase the confidence that a solution from tuning will satisfy the verification feasibility criterion. On the other hand, when $N_2 < N_1$, the margins follow $t(N_2) > t_{\text{tune}}(N_1)$. Thus, similarly to the previous results, we expect a low probability that a solution would satisfy the a posteriori feasibility criterion.

This difference has a similar effect as the margin k in the previous section and will increase the confidence that the solution will satisfy the a posteriori feasibility criterion. However, we expect this effect to be even more pronounced because $E(\mathbf{x}^*, \Xi_2)$ will follow a distribution which is more concentrated and has less variability than $E(\mathbf{x}^*, \Xi_1)$ due to the higher number of samples.

4.4 A PRIORI FEASIBILITY GUARANTEE

Under Assumption 1, we provide a *heuristic guarantee* that the solution $\mathbf{x}^*$ obtained with our two-step method will be feasible both in the verification step and to the original chance-constrained problem (2.1).

Claim 4.7. *Under Assumption 1 and for margins* $t_{tune} = t + k$, $t \geqslant \sqrt{\frac{1}{2N_1} \ln \frac{1}{\delta_t}}$, *and* $k \geqslant \sqrt{\frac{1}{N_1} \ln \frac{1}{\beta}}$, *the probability that a solution* $\mathbf{x}^*$ *obtained from the two-step tuning process is feasible to the original chance-constrained problem* (2.1) *is at least* $(1-\delta_t)(1-\beta)$, *i.e.,*

$$\mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon | \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) \geqslant (1-\delta_t)(1-\beta).$$

Proof. First, we note that the conditional event $\hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon$ (i.e., the event in which the solution $\mathbf{x}^*$ satisfies the tuning criterion (4.4)) can be partitioned into the events where (i) $\mathbf{x}^*$ satisfies the tuning criterion *and* the verification criterion or (ii) $\mathbf{x}^*$ satisfies the tuning criterion *but not* the verification criterion. Using this, we apply the law of total probability to obtain (4.19b).

$$\begin{aligned}
&\mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) && (4.19a)\\
&\quad = \mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \left(\left(\hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) \cap \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \right) \right) \right) \\
&\qquad\qquad \cdot \mathbb{P}_\xi \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) && (4.19b)\\
&\qquad + \mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \left(\left(\hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) \cap \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \geqslant \epsilon \right) \right) \right) \\
&\qquad\qquad \cdot \mathbb{P}_\xi \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \geqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) \\
&\geqslant \mathbb{P}_\xi \left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \left(\left(\hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right) \cap \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \right) \right) \right) && (4.19c)\\
&\qquad\qquad \cdot \mathbb{P}_\xi \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon \right).
\end{aligned}$$

The inequality in (4.19c) is obtained by dropping the second part of (4.19b).

The first term in (4.19c) can be bounded as follows

$$\mathbb{P}_\xi\left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \left(\left(\hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \epsilon\right) \cap \left(\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon\right)\right)\right) \tag{4.20a}$$

$$= \mathbb{P}_\xi\left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \epsilon\right) \tag{4.20b}$$

$$+ \mathbb{P}_\xi\left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon\right)$$

$$\geqslant \mathbb{P}_\xi\left(E(\mathbf{x}^*) \leqslant \epsilon \,\middle|\, \hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \epsilon\right) \tag{4.20c}$$

$$\geqslant \mathbb{P}_\xi\left(E(\mathbf{x}^*) \leqslant \hat{E}(\mathbf{x}^*, \Xi_2) + t\right) \tag{4.20d}$$

$$\geqslant 1 - \delta_t. \tag{4.20e}$$

Here, the equality holds because the event where both the tuning and verification criterion hold happens when either $\hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \epsilon$ or $\hat{E}(\mathbf{x}^*, \Xi_2) + t \leqslant \hat{E}(\mathbf{x}^*, \Xi_1) + t_{\text{tune}} \leqslant \epsilon$. The second inequality holds since the event in (4.20d) is a subset of the event in (4.20c). □

Remark 4.8. *The above proof relies on Claim 4.5, which uses Assumption 1. However, making this assumption allows us to apply Hoeffding's inequality to* $\hat{E}(\mathbf{x}^*, \Xi_1)$ *and obtain a bound on the probability that* $\mathbf{x}^*$ *is feasible to the chance-constrained problem*

$$\mathbb{P}_\xi(E(\mathbf{x}^*) \leqslant \epsilon) \tag{4.21a}$$

$$= \mathbb{P}_\xi(E(\mathbf{x}^*) \leqslant \epsilon | \hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon) \mathbb{P}_\xi(\hat{E}(\mathbf{x}^*, \Xi_1) + t_{tune} \leqslant \epsilon) \tag{4.21b}$$

$$\geqslant (1 - \delta_t)(1 - \beta)(1 - \delta_{tune}). \tag{4.21c}$$

While this above bound does not rigorously hold due to the use of Assumption 1, the resulting confidence level can be used to guide our choice of margin parameters.

	Parameters				Feasible solutions		Cost			
	$1-\beta$	t_{tune}	t	k	Algorithm	Out-of-sample	Avg	Avg (feas)	Min	Max
Equal margins	-	0.0215	0.0215	-	497/1000	1000/1000	42,740		42,698	42,790
Unequal margins	0.05	0.0247	0.0215	0.0032	340/400	400/400	42,763	42,766	42,726	42,810
	0.25	0.0291	0.0215	0.0076	398/400	400/400	42,798	42,796	42,737	42,849
	0.5	0.0332	0.0215	0.0118	400/400	400/400	42,835	42,835	42,775	42,894
	0.9	0.0429	0.0215	0.0213	400/400	400/400	42,972	42,973	42,890	43,053
	0.95	0.0459	0.0215	0.0245	400/400	400/400	43,057	43,056	42,955	43,184

Table 4.1: Results using an even sample split $N_1 = N_2 = 5,000$, with equal tuning margins $t_{tune} = t$ and unequal tuning margins $t_{tune} = t + k$.

Sample split		Parameters		Feasible solutions		Cost				
N_1	N_2	t_{tune}	t	Algorithm	Out-of-sample	Avg	St dev	Min	Max	Avg feas
7,000	3,000	0.0181	0.0277	1/100	100/100	42,720	11.9	42,698	42,785	42,153
4,000	6,000	0.024	0.0196	93/100	100/100	42,755	18.3	42,710	42,802	42,758
1,000	9,000	0.048	0.016	100/100	100/100	43,160	97.9	42,900	43,447	42,739

Table 4.2: Solutions obtained with an uneven sample split $N_1 \neq N_2$ and margins $t_{tune}(N_1)$ and $t(N_2)$.

4.5 CASE STUDY I: TUNING MARGINS

The following case study assesses how the choice of margins used in the solution generation and verification steps affects algorithm performance in terms of solution feasibility. We look at the cases where the margins are equal $t_{\text{tune}} = t$, the margins are defined by $t_{\text{tune}} = t + k$, and the margins are defined by an un-even sample split. All simulations in this section are performed on the IEEE RTS96 24-bus system [Gri+99], with the same modifications and non-Gaussian synthetic uncertainty data as described in Section 3.3.4. For the solution generation step, the same approximate optimization model, solution evaluation process, and bisection tuning scheme as presented in Section 3.2 is used.

4.5.1 *Equal margins*

We first examine the case with equal tuning and verification margins $t_{\text{tune}} = t = 0.0215$. We use a desired $\epsilon = 0.05$, a confidence level $\delta = 0.01$, and assume availability of $N_1 = N_2 = 5,000$ samples. We run 1,000 replications of the algorithm (assuming a new set of $N_1 + N_2 = 10,000$ samples in each replication) and present the results in the top section of Table 4.1.

Our algorithm results in solutions that satisfy the a posteriori verification criterion for 497/1000 replications. This is consistent with our theoretical result (4.8), indicating that the bias of $\hat{E}_G(x^*, \Xi_1)$ is small. In the out-of-sample evaluation, we observe that 1000/1000 replications result in solutions that satisfy $\hat{E}(\mathbf{x}^*, \Xi_{\text{oos}}) \leqslant \epsilon$, which highlights the conservatism of Hoeffding's bound.

4.5.2 *Unequal margins*

We next investigate how using an extra tuning margin increases the probability of obtaining solutions that satisfy the test criterion (4.5) in the verification step. Here, we again define the verification margin $t = 0.0215$ according to (4.2) but choose the tuning margin to be $t_{\text{tune}} = t + k$ with $k = \sqrt{(1/N_1) \ln 1/\beta}$, where $1 - \beta$ is the confidence level that a solution generated from tuning satisfies the a posteriori check. We use confidence values $1 - \beta = \{0.05, 0.25, 0.5, 0.9, 0.95\}$ and perform 400 replications of the algorithm. The results are reported in the bottom section of Table 4.1. We also perform an out-of-sample evaluation of the solution x^* to approximate the true violation probability of x^*, denoted by $\hat{E}(\mathbf{x}^*, \Xi_{\text{oos}})$. We compare the result with the desired ϵ to estimate whether $\mathbf{x}^*$ is feasible to original JCC-OPF problem (reported in the "out-sample" row).

In Table 4.1, for $1 - \beta \geqslant 0.5$, all replications result in solutions that satisfy the a posteriori check. For confidence values $1 - \beta = \{0.05, 0.25\}$, the actual proportion of solutions passing the a posteriori check is 0.85 and 0.995, respectively. Our main observation is that the portion of replications satisfying the a posteriori criterion (4.5) is thus significantly higher than $1 - \beta$, indicating that our bound is conservative. In particular, we see that for $1 - \beta \geqslant 0.5$, all replications result in solutions that satisfy the a posteriori check. For confidence values $1 - \beta = \{0.05, 0.25\}$, the actual proportion of solutions passing the a posteriori check is 0.85 and 0.995, respectively.

In terms of the cost of solutions resulting from the different confidence guarantees, we observe that the cost does varies, but only slightly with different levels of confidence, with a maximum difference of less than 1%. This indicates that we can obtain solutions with higher confidence guarantees but little increase in total cost.

Finally, in the out-of-sample evaluation, for all confidence values and all runs, the algorithm results in solutions satisfying $\hat{E}(\mathbf{x}^*, \Xi_{\text{oos}}) \leqslant \epsilon$, indicating that our method does result in conservative solutions.

4.5.3 *Un-even sample split*

We look at the case where $N_1 \neq N_2$, which results in different tuning and verification margins, as given by (4.17). Table 4.2 shows the choices of N_1 and N_2, the margins, and algorithm outcomes. We first observe that when $N_1 > N_2$, and subsequently $t_{\text{tune}}(N_1) < t(N_2)$, only 1/100 replications attain a final solution that passes the a posteriori check. In contrast, when $N_1 < N_2$ and $t_{\text{tune}}(N_1) > t(N_2)$, a very high fraction of the solutions pass the a posteriori feasibility check (93/100 and 100/100, respectively). This demonstrates the benefits of using more samples in the verification step and validates our expectations.

Comparing the cost of the solutions obtained with different sample splits, we observe that using a higher number of samples in the tuning step leads to a lower average cost and lower variance. This is as expected, because (i) the tuning margin $t_{\text{tune}}(N_1)$ is smaller and (ii) using more samples result in less variance in the empirical violation evaluation. However, as discussed above, these lower cost solutions also have a lower probability of passing the feasibility check in the verification stage because $t(N_2)$ is larger.

We next compare the solutions obtained with the different sample splits in Table 4.2 with solutions obtained with the different margins k in Table 4.1. We observe that solutions obtained with sample split $N_1 = 4000, N_2 = 6000$ provides *both* lower average cost *and* higher feasibility, as compared with solutions obtained with confidence $1 - \beta = 0.05$. On the other hand, a sample split $N_1 = 1000, N_2 = 9000$ provide solutions with higher average cost than any of the solutions with similar levels of feasibility $1 - \beta = \{0.05, 0.9, 0.95\}$. This indicates that if we choose the sample split carefully, we can obtain better solutions than by introducing margin k.

4.6 CASE STUDY II: PERFORMANCE COMPARISON

Next, we compare the performance of our two-step tuning algorithm against the scenario approach on various test systems. We investigate the performance and scalability of the two-step tuning algorithm by comparing its performance against the scenario approach applied to the IEEE RTS96 24-bus network and using uncertainty data as described in Section 3.3.4.

The scenario approach replaces the original problem with a reformulation that enforces feasibility of all constraints for N_{SA} i.i.d. uncertainty samples. The scenario approach problem formulation used is presented in Appendix A.1. Given n decision variables, we use

$$N_{SA} \geqslant \frac{1}{\epsilon}\left(n + \ln\left(\frac{1}{\delta_{SA}}\right) + \sqrt{2n \ln \frac{1}{\delta_{SA}}}\right) \tag{4.22}$$

samples for the scenario approach, which provides an a priori guarantee that the solution satisfies the joint chance constraint with confidence at least $1 - \delta_{SA}$ [ATL10].

We use parameters $\epsilon = 0.05$, $\delta_{tune} = \delta_t = \beta = 0.001$, and $\delta_s = 0.003$. In order to ensure a fair comparison between the methods, for the scenario approach, we use a confidence level of

$$1 - \delta_{SA} = (1 - \delta_t)(1 - \beta)(1 - \delta_{tune}) = 0.002. \tag{4.23}$$

With $n = 33$ decision variables in the IEEE RTS96 test case, the scenario approach requires $N_{SA} = 1,168$ samples to achieve confidence $1 - \delta_s$. For the two-step tuning approach, we use $N_1 = N_2 = 25,000$, resulting in parameters $t_{tune} = 0.0201$, $t = 0.0083$, and $k = 0.0118$.

In Table 4.3, we compare the average and standard deviation of the solution costs across 100 replications, as well as the empirical violation evaluated on an out-of-sample set

	Cost		Violation probability	
	Average	St dev	Average	St dev
Tuning method	42,731	5.2	0.0297	0.0012
Scenario approach	43,144	103.5	0.0039	0.0020

Table 4.3: Comparison of tuning algorithm with the scenario approach. The average and standard deviations of the cost (objective) and out-of-sample violation probability for 100 algorithm replications are reported.

with $100,000$ samples for each replication. With these parameters, all solutions generated by the two-step tuning method satisfy the a posteriori feasibility test (4.5), and are thus feasible with probability $1 - \delta = 0.999$.

We observe that the two-step tuning algorithm achieves about a 1% lower average cost than the scenario approach and a much smaller standard deviation in solution costs across replications. The average out-of-sample violation probabilities are 0.0297 and 0.0039 for solutions obtained from the two-step tuning method and scenario approach, respectively, and again our tuning method results in lower variance than the scenario approach. From this, we see that solutions obtained from tuning are *significantly* less conservative, with lower cost and higher out-of-sample empirical violation. The two-step tuning approach also provides more consistent solutions, as indicated by the lower standard deviations, as a consequence of its ability to leverage the information available in the larger set of samples. Moreover, the lower standard deviations can also indicate that the tuning method may provide solutions that are less dependent on the sample set, as a result of not using the uncertainty samples within the optimizaiton problem itself.

Although the solutions of the two-step tuning method are more desireable, the method suffers from the drawback that it requires significantly higher number of samples compared to the scenario approach (in this case, $50,000$ samples compared to $1,190$) to ensure the tuning and verification margins are small enough compared to ϵ. On the other hand, while the samples required for the tuning method is independent of the size of the

optimization problem, the number of samples needed for the scenario approach scales with the number of decision variables, which may be problematic for problems with large amounts of decision variables, e.g., AC OPF. Furthermore, the scenario approach generally cannot use additional samples even when available, as a large number of samples can make the resulting optimization problem intractable and/or the resulting solution overly conservative.

4.7 CONCLUSION

In this chapter, we first highlight two significant challenges associated with obtaining feasibility guarantees for tuning-based algorithms of chance-constrained problems, such as the algorithm proposed in Chapter 3. To combat these challenges, a two-step tuning framework is proposed, where a tuning-based solution generation step is followed by a solution verification step that uses an out-of-sample uncertainty set to provide probabilistic feasibility guarantees on the resulting solution. The solution verification step allows us to obtain an *a posteriori* feasibility guarantee on solutions identified using tuning algorithms, including the bisection search-based tuning algorithm from Chapter 3.

We then provide theoretical results that give guidance on how to adapt the feasibility criteria used in the solution generation and verification steps to increase the probability that the resulting solutions be feasible in the verification step. While these results rely on a limiting assumption which, the analysis takes us a step towards providing *a priori* probabilistic feasibility guarantees for tuning based methods.

Finally, we apply our two-step tuning method to the joint chance-constrained DC OPF problem, obtaining numerical results supporting our theoretical conclusions. We further demonstrate that tuning is able to obtain less conservative solutions compared to the scenario approach.

5 TUNING WITH PROBABLISTIC GUARANTEES: EXTENSIONS AND ANALYSIS

This chapter extends the results of the two-step tuning algorithm presented in the previous chapter. We use theoretical analyses and numerical examples to analyze the strengths and limitations of the two-step tuning algorithm in order to characterize settings where our tuning algorithm can be best applied.

This chapter aims to address the following open challenges:

(i) In Chapter 4, we obtained a *heuristic* a priori feasibility result for solutions obtained via tuning. Due to the reliance on the assumption that the empirical violation probability in the solution generation step is an approximately unbiased estimator (Assumption 1), we were unable to derive a rigorous a priori feasibility guarantee. In this chapter, we provide a methodology to empirically test the validity of this assumption in practice. We subsequently demonstrate that for the joint chance-constrained OPF problem, using the bisection search tuning procedure and moment-based reformulation proposed in Chapter 3, the assumption is mostly mild.

(ii) We next examine the relationship between sample set size and the expected solution conservativeness for the tuning algorithm. The a priori feasibility result (Claim 4.7) of Chapter 4 implied that with access to more data, we would be able to obtain higher quality, less conservative solutions by using lower margins. This is in contrast to alternative sample-based methods for solving chance-constrained problems discussed in Section 2.4.2, where increasing sample sizes may lead to over-conservativeness or computational intractability. We first derive the required sample

complexity of the two-step tuning method and characterize the relationship between the number of samples used in the algorithm and the expected conservativeness of the resulting solution. We demonstrate that for large scale optimization problems with large amounts of decision variables, our tuning method is more sample efficient than a comparable method, the scenario approach.

(iii) Finally, we extend the numerical performance results presented in Section 4.6 by applying the tuning algorithm to solve chance-constrained DC OPF for larger test systems, namely the IEEE 118 bus system [CWW00] and Polish Winter Peak test case with 2,383 buses [ZMT11]. We demonstrate the scalability of the tuning algorithm by benchmarking the algorithm run time and solution optimality and feasibility against the scenario approach and CVaR inner approximation [NS07] solved using the sample average approximation method.

5.1 ANALYSIS OF APPROXIMATELY UN-BIASED ASSUMPTION

In this section, we propose two methods that can be used to assess whether the approximately unbiased estimator assumption (Assumption 1) is appropriate to make in practical use cases.

5.1.1 *Empirical distribution assessment*

Let $\mathbf{x}^*$ be a solution obtained from a single run of the tuning algorithm. Let $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ denote the sets of indicator random variables obtained from evaluating solution

$\mathbf{x}^*$ on the tuning sample set Ξ_1 and the verification sample set Ξ_2, respectively. Let N_1 and N_2 be the number of samples in each corresponding sample set.

We expand on the result from previous work [HR22], which assessed the bias of estimator $\hat{E}(\mathbf{x}^*, \Xi_1)$ by plotting and analyzing the distribution of

$$Z(\mathbf{x}^*, \Xi_1, \Xi_2) := \hat{E}(\mathbf{x}^*, \Xi_1) - \hat{E}(\mathbf{x}^*, \Xi_2). \tag{5.1}$$

Claim 5.1. *In the case where Assumption 1 holds, we claim that*

$$\mathbb{E}[Z(\mathbf{x}^*, \Xi_1, \Xi_2)] = 0. \tag{5.2}$$

Proof. Assumption 1 implies that the indicator random variables in sets $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ follow the same Bernoulli distribution, i.e., the Bernoulli parameter and expected value of the indicator random variables are the same. It follows that

$$\begin{aligned}
\mathbb{E}[Z(\mathbf{x}^*, \Xi_1, \Xi_2)] &= \mathbb{E}[\hat{E}(\mathbf{x}^*, \Xi_1) - \hat{E}(\mathbf{x}^*, \Xi_2)] && (5.3)\\
&= \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} Y(\mathbf{x}^*, \xi_1^{(i)}) - \frac{1}{N}\sum_{i=1}^{N} Y(\mathbf{x}^*, \xi_2^{(i)})\right] && (5.4)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbb{E}[Y(\mathbf{x}^*, \xi_1^{(i)})] - \mathbb{E}[Y(\mathbf{x}^*, \xi_2^{(i)})]\right) && (5.5)\\
&= 0. && (5.6)
\end{aligned}$$

□

Claim 5.2. *In the case where where Assumption 1 holds, we claim that $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ follows a normal distribution with zero mean.*

Proof. By Assumption 1, $\hat{E}(\mathbf{x}^*, \Xi_1)$ and $\hat{E}(\mathbf{x}^*, \Xi_2)$ are sample means both estimating the true violation probability $E(\mathbf{x}^*)$, i.e., they are calculated as the sum of random variables drawn from the same distribution. As a consequence of the central limit theorem, with a

sufficiently large number of samples, the sample means are normally distributed. Thus, $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ is normally distributed since the subtraction of two normal distributions also follows a normal distribution. □

Since we only have access to the samples $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$, we are limited to empirically assessing whether the distribution of $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ is a normally distributed with zero mean. To do so, we can run a large number of replications of the two-step tuning algorithm to obtain an empirical distribution of $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ and inspect whether it is normally distributed with zero mean. In the case that $\hat{E}(\mathbf{x}^*, \Xi_1)$ is biased, we would expect to see a non-zero mean and/or a non-normal distribution (e.g., a non-symmetric distribution). The case study presented at the end of the section provides a methodology for performing this assessment.

5.1.2 *Hypothesis testing*

We propose a more rigorous alternative to assessing the bias of $\hat{E}(\mathbf{x}^*, \Xi_1)$ via the use of statistical hypothesis testing. Let use p_1 and p_2 to denote the Bernoulli random variable parameters for sets $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$. We can perform a Wald test [Wal45] to determine whether there is statistically significant evidence that the two sets of samples are drawn from *different* Bernoulli distribution (i.e., $p_1 \neq p_2$). We define our null hypothesis

$$H_0 : p_1 = p_2$$

as the case where the Bernoulli parameters are equal. This corresponds to the interpretation that the sample sets $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ follow the same distribution, implying

that $\hat{E}(\mathbf{x}^*, \Xi_1)$ and $\hat{E}(\mathbf{x}^*, \Xi_2)$ are both unbiased estimators of the violation probability of solution $\mathbf{x}^*$. Our alternative hypothesis

$$H_1 : p_1 \neq p_2$$

corresponds to the case that the Bernoulli parameters are not equal. In this case, there is statistically significant evidence that the sample sets $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ do not follow the same distribution. Because we know that $\hat{E}(\mathbf{x}^*, \Xi_2)$ is an unbiased estimator of the violation probability, this result would imply that $\hat{E}(\mathbf{x}^*, \Xi_1)$ is *not* an unbiased estimator of the violation probability of $\mathbf{x}^*$.

To perform this hypothesis test, we first calculate Wald statistic using the sample sets $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ as follows

$$W := \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p_1}(1-\hat{p_1}) + \hat{p_2}(1-\hat{p_2})}{N}}}, \tag{5.7}$$

where $\hat{p_1} := \hat{E}(\mathbf{x}^*, \Xi_1)$ and $\hat{p_2} := \hat{E}(\mathbf{x}^*, \Xi_2)$ are the evaluated empirical means of the respective sample sets. Then, for a specified significance level $0 < \alpha < 1$, we reject the null hypothesis H_0 in favor of the alternative hypothesis H_1 if

$$|W| > z_{\alpha/2}, \tag{5.8}$$

where $z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2)$ is the inverse cumulative standard normal distribution evaluated at $1 - \alpha/2$. The significance level α denotes the probability we reject the null hypothesis given the null hypothesis is true. If (5.8) is not satisfied, then we fail to reject H_0. In the case that $\hat{E}(\mathbf{x}^*, \Xi_1)$ is an unbiased estimator of the violation probability, we would expect to fail to reject H_0, indicating that there is no statistical significance that the two sample sets come from different underlying Bernoulli distributions.

5.1.3 *Case study: Application to CC DC OPF*

We use the methodology presented in Section 5.1 to assess whether the approximate unbiased estimator assumption (Assumption 1) is valid for our implementation of the tuning algorithm for the chance-constrained DC OPF problem with a joint chance constraint (3.18). We use the same approximate optimization model, solution evaluation process, and bisection tuning scheme as presented in Section 3.2 for the solution generation step and the a posteriori verification step as detailed in Section 4.2.2.

Test system

We use the IEEE 118 bus system [CWW00] with the following modifications from [Roa+17]: (i) the maximum generation capacity and load are scaled by a factor of 1.25, (ii) the minimum generation capacity is set to zero, (iii) all loads fluctuate around their forecasted consumption with a standard deviation of 10% of their forecasted consumption.

We set the desired violation probability of the chance constraint to $\epsilon = 0.05$, which is a value commonly used in power systems applications. To ensure a high likelihood that the resulting solution is feasible to the original chance constraint, we use confidence parameters $\delta_{\text{tune}} = \delta_t = \beta = 0.001$. Using Claim 4.7, we calculate the corresponding margin values t_{tune} and t.

Empirical distribution assessment

We investigate the cases where we use samples sizes $N = 2500$, $N = 5000$, and $N = 7500$ in each step of the tuning process. We perform 250 independent replications of the algorithm, each using new draws of the sample sets Ξ_1 and Ξ_2.

We compute the empirical distributions of $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ based on data from all of the 250 runs of the algorithm for each of the above cases. The resulting histograms are shown in Figure 5.1. In Table 5.1, we report the mean, median, and 10% and 90% quantiles

of each distribution. We can observe that for all sample sizes, the mean is negative with a very small magnitude and the median is zero, demonstrating there exist a very slight skew to the left. This indicates that the bisection search-based tuning imposes a slightly conservative bias on the empirical violation probability $\hat{E}(\mathbf{x}^*, \Xi_1)$. However, we note that this bias is rather trivial given that the magnitude of the mean of $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ (representing the mean difference in the estimated empirical violation probabilities) is less than 0.01% compared to the the magnitude of the empirical violation probabilities. Further, the median value of $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ is zero, indicating that $\hat{E}(\mathbf{x}^*, \Xi_1)$ is larger than $\hat{E}(\mathbf{x}^*, \Xi_2)$ half of the time (as we would expect if they came from the same distribution. As we increase the sample size, while we observe a decrease in the standard deviation, we see than the mean and median remain relatively constant.

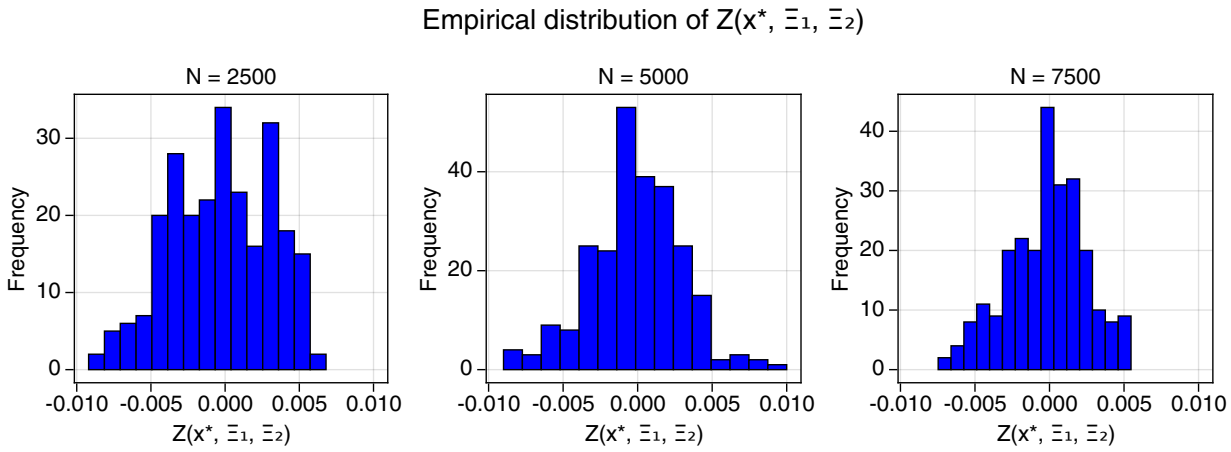


Figure 5.1: Plotted is the distribution of $Z(\mathbf{x}^*, \Xi_1, \Xi_2) = \hat{E}(\mathbf{x}^*, \Xi_1) - \hat{E}(\mathbf{x}^*, \Xi_2)$ obtained over 250 algorithm replications. Shown are the cases where sample sizes of $N = \{2500, 5000, 7500\}$ are used for the tuning and verification sets. A slight leftward skew in the distribution can be observed, demonstrating a slight negative bias in $Z(\mathbf{x}^*, \Xi_1, \Xi_2)$ for all cases.

Hypothesis testing

Next, we perform the hypothesis test methodology detailed in Section 5.1.2. This hypothesis testing can be done for each individual run, leaving us with 250 different instances where the hypothesis was either rejected or not rejected. We report the results in Table

	$N = 2500$	$N = 5000$	$N = 7500$
Mean	-0.0002	-0.0001	-0.0002
St. dev.	0.0034	0.0031	0.0026
Median	0.0000	0.0000	0.0000
10% quantile	-0.0044	-0.0038	-0.0040
90% quantile	0.0044	0.0036	0.0029

Table 5.1: Summary statistics for the values of random variable $Z(\mathbf{x}^*, \Xi_1, \Xi_2) = \hat{E}(\mathbf{x}^*, \Xi_1) - \hat{E}(\mathbf{x}^*, \Xi_2)$ obtained over 250 algorithm replications for the cases where $N = \{2500, 5000, 7500\}$ samples are used for the tuning and verification sets.

	$N = 2500$	$N = 5000$	$N = 7500$
Fail to reject H_0	235/250	237/250	237/250
Reject H_0	15/250	13/250	13/250

Table 5.2: Results from performing a hypothesis test to assess whether the sets of indicator random variables $Y(\mathbf{x}^*, \Xi_1)$ and $Y(\mathbf{x}^*, \Xi_2)$ obtained in the solution generation and verification steps come from the same underlying Bernoulli distribution. Shown are whether the 250 algorithm replications result in failing to reject the null hypothesis H_0 (same distribution) or rejecting the null hypothesis H_0 (different distribution) for the cases when sample sets of size $N = \{2500, 5000, 7500\}$ are used.

5.2. We observe that for all sample sizes, the majority of replications fail to reject the null hypothesis. This indicates that for the majority (but not all) of algorithm runs, there is no statistical significance that the two sample sets come from different underlying distributions, meaning that for those runs we cannot conclude $\hat{E}(\mathbf{x}^*, \Xi_1)$ is biased. However, in about 5% of the instances, the hypothesis is rejected. This implies that in some cases, depending on the uncertainty samples drawn, the bisection search my result in overfitting to the tuning samples and subsequently biasing $\hat{E}(\mathbf{x}^*, \Xi_1)$. We also observe that as the sample size increases, there is a slight increase in the number of replications that fail to reject the null hypothesis, indicating that incorporating more data may limit the possibility of overfitting to the tuning samples.

5.2 SAMPLE COMPLEXITY AND SCALABILITY ANALYSIS

In this section, we explore the sample complexity required for the tuning method under various choices of parameters (such as the desired violation probability ϵ and confidence levels δ_{tune}, δ_t, and β) and number of decision variables. We discuss these results for general problems, as well as provide specific results for the chance-constrained DC OPF problem.

5.2.1 *Minimum number of samples*

We first provide a bound on the sample complexity of the two-step tuning algorithm.

Claim 5.3. *The minimum number of samples* N *required for each step of the two-step tuning algorithm is*

$$N \geqslant \max\left\{\frac{1}{2\epsilon^2}\ln\frac{1}{\delta_{tune}}, \frac{1}{\epsilon^2}\left(\frac{1}{2}\ln\frac{1}{\delta_t}+\ln\frac{1}{\beta}+\sqrt{2\ln\frac{1}{\delta_t}\ln\frac{1}{\beta}}\right)\right\}. \tag{5.9}$$

Proof. First, in the solution generation step, our goal is to adjust the tuning parameter such that the resulting solution satisfies the empirical violation criterion (4.4). We also note that, by definition, the lowest empirical violation probability we can achieve is zero, i.e., $\hat{E}(\mathbf{x}^*, \Xi_1) \geqslant 0$ and $\hat{E}(\mathbf{x}^*, \Xi_2) \geqslant 0$, which implies the following upper bounds on the margins

$$\epsilon \geqslant t_{\text{tune}}, \tag{5.10}$$

$$\epsilon \geqslant t. \tag{5.11}$$

The algorithm construction and a priori feasibility guarantee (Claim 4.7) provides a lower bound on the tuning margin t_{tune}, implying

$$\epsilon \geqslant t_{\text{tune}} \geqslant \sqrt{\frac{1}{2N} \ln \frac{1}{\delta_{\text{tune}}}}. \tag{5.12}$$

Claim 4.7 also requires that $t_{\text{tune}} \geqslant t + k$ and provides lower bounds on the verification margin t and parameter k, implying

$$\epsilon \geqslant t_{\text{tune}} \geqslant \sqrt{\frac{1}{2N} \ln \frac{1}{\delta_t}} + \sqrt{\frac{1}{N} \ln \frac{1}{\beta}}. \tag{5.13}$$

We can then solve equations (5.13) and (5.12) to obtain the required number of samples N in terms of parameters ϵ, δ_{tune}, δ_t, and β. From equation (5.13), we obtain

$$N \geqslant \frac{1}{\epsilon^2} \left(\frac{1}{2} \ln \frac{1}{\delta_t} + \ln \frac{1}{\beta} + \sqrt{2 \ln \frac{1}{\delta_t} \ln \frac{1}{\beta}} \right). \tag{5.14}$$

From equation (5.12), we obtain

$$N \geqslant \frac{1}{2\epsilon^2} \ln \frac{1}{\delta_{\text{tune}}}. \tag{5.15}$$

Combining these two expressions, we obtain the bound (5.9) on the minimum number of samples required for each of the two steps of the tuning algorithm. □

We observe that the number of samples required has an inverse quadratic dependence on the desired violation probability ϵ, which may be prohibitive if ϵ is required to be extremely small.

5.2.2 *Relationship between sample size and solution conservativeness*

The above result corresponds to the case where $t_{tune} = \epsilon$, meaning the maximum tuning margin is used and the tuning algorithm will aim to find a solution with empirical violation probability close to zero, i.e., $\hat{E}(\mathbf{x}^*, \Xi_1) \leqslant \epsilon - t_{tune} = 0$. Clearly, this will result in overly conservative solutions. To obtain less conservative solutions, we would like to use a smaller tuning margin, which consequently requires a larger number of samples.

Let us define a constant $c \in [0, 1]$. We can define the tuning margin as a function of c and the desired violation probability ϵ using

$$t_{tune} := c\epsilon. \tag{5.16}$$

By choosing an appropriate constant c, we are effectively able to control the violation probability of the solution resulting from tuning. In particular, the choice of t_{tune} would result in tuning a solution to have an empirical violation probability of

$$\hat{E}(\mathbf{x}^*, \Xi_1) \leqslant \epsilon - t_{tune} = \epsilon(1 - c). \tag{5.17}$$

Using (5.16), we can obtain the following sample complexity bound, analogous to Claim 5.3:

Claim 5.4. *Using a tuning margin satisfying $t_{tune} = c\epsilon$, where $c \in [0, 1]$ is a constant, results in the following sample complexity bound:*

$$N \geqslant \max\left\{ \frac{1}{2c^2\epsilon^2} \ln \frac{1}{\delta_{tune}}, \frac{1}{c^2\epsilon^2}\left(\frac{1}{2} \ln \frac{1}{\delta_t} + \ln \frac{1}{\beta} + \sqrt{2 \ln \frac{1}{\delta_t} \ln \frac{1}{\beta}} \right) \right\}. \tag{5.18}$$

Proof. The proof follows similarly to that of Claim 5.3. □

We note that this sample complexity bound has an inverse quadratic relationship with the choice of c. We also observe there exists an inverse relationship between the expected conservativeness of the resulting solution and the number of required samples. E.g., with a small c, we would expect to obtain a solution with a violation probability closer to ϵ, meaning a less conservative solution, but a larger sample size requirement (and vice versa). The benefit of this relationship is that it allows us to obtain improved solutions with increased access to data. Furthermore, unlike methods that directly incorporate samples in the problem formulation itself (e.g., the scenario approach), including more samples in our algorithm is not computationally prohibitive because the data is only used for an a posteriori evaluation of the generated solutions.

5.2.3 *Comparison to the scenario approach*

We compare the number of samples required for our method to the number of samples required for the scenario approach. For the solution of the scenario approach to satisfy the chance constraint (2.1b) with confidence level $(1-\delta_s) \in [0,1]$, we require at least

$$N_{SA} \geqslant \frac{1}{\epsilon}\left(n + \ln\frac{1}{\delta_s} + \sqrt{2n\ln\frac{1}{\delta_s}}\right), \tag{5.19}$$

samples, where n represents the number of decision variables [ATL10]. In order to compare this sample complexity against that of the tuning approach, we use a confidence level of $\delta_s := 1-(1-\delta_t)(1-\beta)(1-\delta_{tune})$.

We can immediately observe that the number of samples required for the scenario approach scales with only $1/\epsilon$, compared to $1/\epsilon^2$ required for the tuning approach. Thus, in the case where the desired violation ϵ is very close to zero, we expect that the scenario approach will require significantly less samples than the tuning approach. On the other hand, while the sample complexity of the scenario approach scales linearly with the

number of decision variables n, the sample complexity of our method does not have such a dependence. This is particularly advantageous for large scale optimization problems that have a very high number of decision variables. Finally, unlike our result for the tuning method (5.18), the sample complexity result for the scenario approach does not have an explicit relation to conservativeness. However, we observe that the scenario approach typically results in more conservative solutions as the number of samples used increases. This is because the scenario approach is formulated such that the resulting solution must be secure against all uncertainty samples drawn, including potential "worst-case samples".

5.2.4 *Numerical comparison of sample complexity*

We perform a numerical comparison of the number of required samples for the tuning method to that required for the scenario approach, with the formulation presented in Appendix A.1. We examine sample complexity as a function of desired violation ϵ and as a function of the number of decision variables n in the problem. We use confidence parameters $\delta_{\text{tune}} = \delta_t = \beta = 0.001$ for the tuning method. For the scenario approach, we choose $\delta_s = 1 - (1 - \delta_t)(1 - \beta)(1 - \delta_{\text{tune}}) = 0.003$.

Sample complexity as a function of ϵ

In Figure 5.2, we plot the sample complexity of the tuning approach for values $c = 1$, $c = 2/3$, and $c = 1/2$ and the scenario approach against desired violation probabilities ranging across $0 \leqslant \epsilon \leqslant 0.10$. We plot cases for the number of decision variables $n = \{10, 100, 1000, 10000, 100000\}$.

For problems with less decision variables, such as $n = 10$ and $n = 100$ (plotted in the yellow and green lines, respectively), the scenario approach requires significantly less samples than the tuning approach (dark blue line) across all values of c, particularly

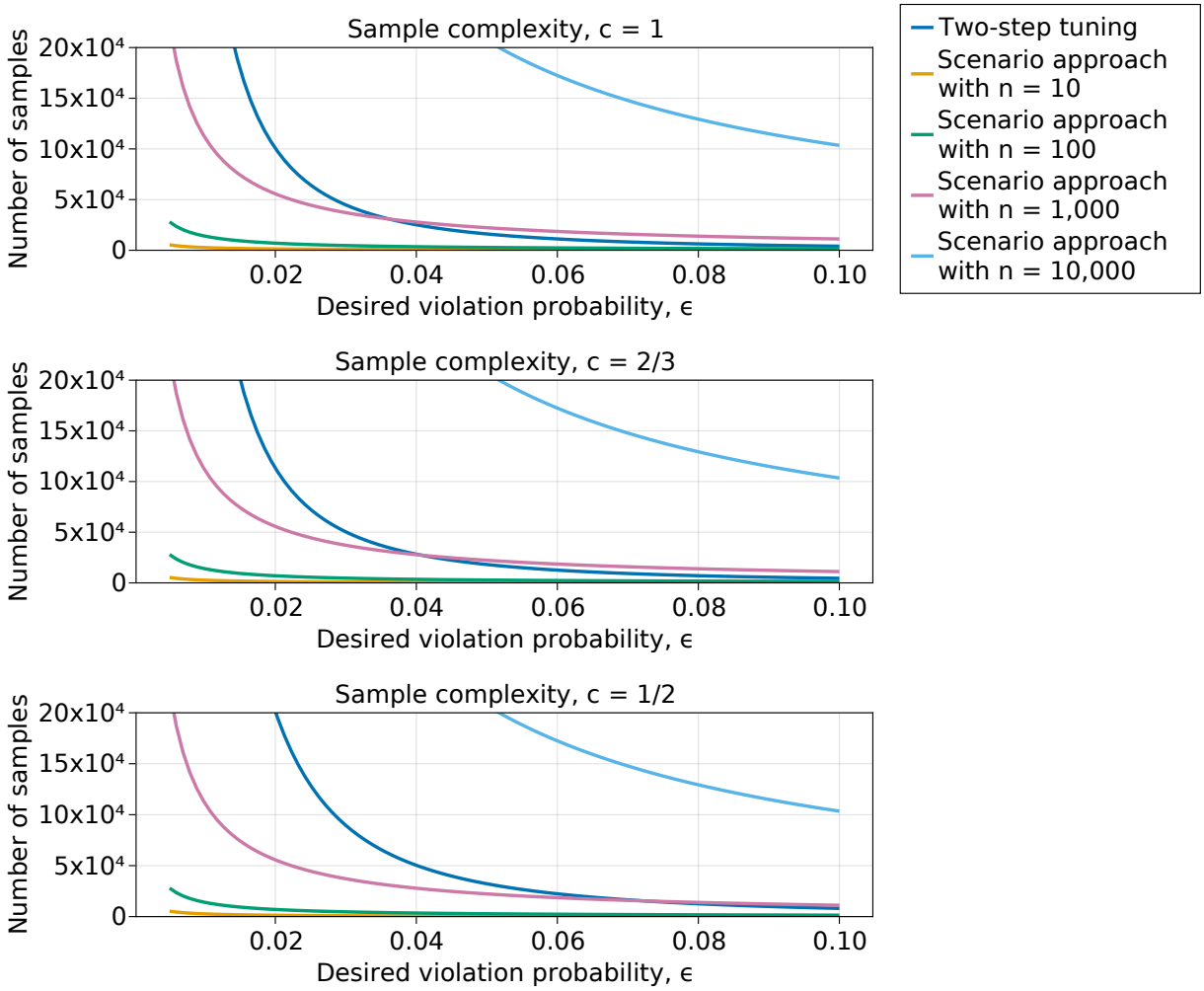


Figure 5.2: We compare the sample complexity of the tuning approach to the scenario approach for desired violation probability values $0 \leqslant \epsilon \leqslant 0.10$ and tuning approach constants $c = 1, 2/3, 1/2$. We look at cases where the numbers of decision variables are $n = 10, 100, 10000$, and 10000. Observe that the tuning approach (dark blue line) always requires less samples that the scenario approach when there is a large number of decision variables (i.e., $n = 10,000$, plotted in light blue). With less decision variables, the tuning approach will be comparable to the scenario approach in number of samples when ϵ is high.

for small values of ϵ. However, as ϵ grows, we see that the number of samples required for the tuning process begins to approach that of the scenario approach. For increased problem sizes, we see that in certain cases, the tuning approach is more sample efficient that the scenario approach. For example, for $n = 10,000$ decision variables, for all c values and plotted values of ϵ, the tuning method requires significantly less samples than the scenario approach.

Furthermore, we also see that as c decreases, the sample complexity of the tuning approach increases relative the that of the scenario approach (as evidence by the dark blue curve moving leftward). For example, if we look at a problem with $n = 1000$ decision variables, for the $c = 1$ case, for values $\epsilon > 0.0357$, the scenario approach requires more samples. As c decreases, the minimum violation probability ϵ in which the scenario approach begins to require more samples than tuning increases. For $c = 2/3$ and $c = 1/2$, we see that the violation probability must be $\epsilon > 0.0405$ and $\epsilon > 0.0722$, respectively.

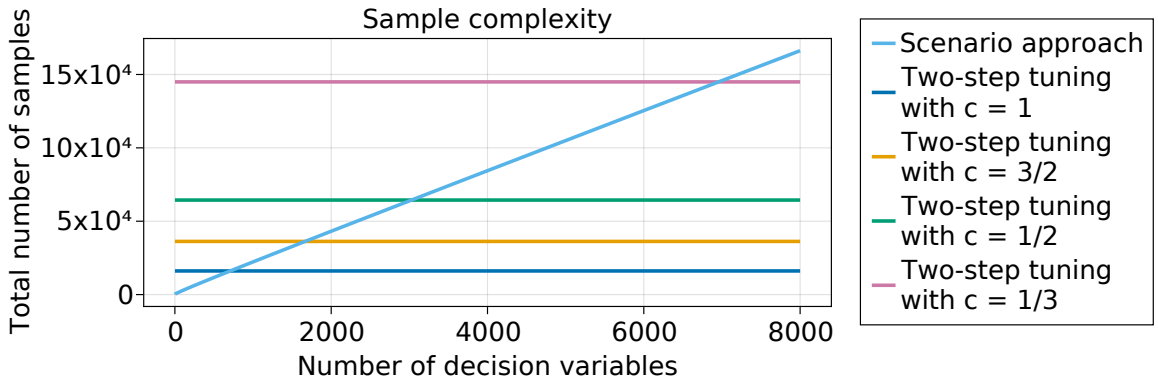


Figure 5.3: The sample complexity plotted against the number of decision variables for the scenario approach and tuning approaches with $c = \{1, 1.5, 2, 3\}$. The parameters $\epsilon = 0.05$, $\delta_{tune} = \delta_t = \beta = 0.001$, and $\delta_s = 0.003$ are used.

Sample complexity as a function of c

Next, we assess the size at which an optimization problem may be solved with less samples using the tuning method in comparison to the scenario approach. In Figure 5.3, we plot the sample complexity against the number of decision variables while keeping the desired violation fixed at $\epsilon = 0.05$. We use the sample confidence values as above ($\delta_{tune} = \delta_t = \beta = 0.001$ and $\delta_s = 0.003$). We calculate the margin values t_{tune} and t using Claim 4.7. For the tuning approach, we plot the sample complexity for $c = \{1, 2/3, 1/2, 1/3\}$. We see that using the tuning method with $c = 1$, the number of samples required will be less than that required for the scenario approach when $n > 706$. For $c = 2/3$, $c = 1/2$, and $c = 1/3$, this becomes true at $n > 1,667$, $n > 3,028$, and $n > 6,957$, respectively.

We can conclude that when the optimization has a large number of decision variables, the two-step tuning method will be more sample efficient in comparison to the scenario approach, even if a lower c value is chosen.

5.2.5 *Case study: Sample complexity and scalability for CC DC OPF*

This case study assesses the sample complexity necessary for applying the two-step tuning algorithm to chance-constrained DC OPF with joint chance constraints (3.18) and compares against a baseline given by the scenario approach. For the tuning approach, we use the same parameters as in Section 5.2.4, i.e., $\delta_{\text{tune}} = \delta_t = \beta = 0.001$ for the tuning method and $\delta_s = 1 - (1 - \delta_t)(1 - \beta)(1 - \delta_{\text{tune}}) = 0.003$ for the scenario approach.

For chance-constrained DC OPF, the number of decision variables n is simply the number of conventional dispatchable generators in the system. As evidenced in Section 5.2, unless the system being analyzed is very large, with the number of generators at least on the order of thousands, the number of decision variables may not reach the point where the tuning method becomes more sample efficient than the scenario approach. For example, the IEEE 118 bus system and Polish test system only have 19 and 327 generators, respectively. Using the scenario approach to solve chance-constrained DC OPF for these cases would require less samples than using the tuning approach for any desired level of conservativeness $c \in [0, 1]$.

Looking at other problem variations of optimal power flow, the number of decision variables can rapidly increase. For example, consider the AC formulation of CC OPF (AC CC-OPF) presented in [RA17], which has several additional decision variables, including the voltage magnitude and voltage angle variables for each bus and the active and reactive generation variables for each generation. The resulting number of decision variables becomes $n = 2|\mathcal{N}| + 2|\mathcal{G}|$, where $\mathcal{N}$ and $\mathcal{G}$ represent the sets of buses and generators,

		DC formulation		AC formulation	
	Buses	n	N_{SA}	n	N_{SA}
IEEE 118 bus	118	19	794	274	6,724
Polish test case	2,383	327	7,889	5420	113,534
C2FEN03411 Industry	3,411	969	21,619	8,760	181,697
C2FEN04601 Synthetic	4,601	408	9,654	10,018	207300
C2FEN10480 Synthetic	10,480	777	17,556	22,514	460,626
C2FEN16955 Industry	16,955	1,868	40,423	37,646	766,264
C2FEN22700 Industry	22,723	4,062	85,701	53,570	1.09×10^6
C2FEN31777 Synthetic	31,777	4,663	98,032	72,880	1.47×10^6

Table 5.3: For the various synthetic and industry test cases found in [ARP] and [ZMT11], for the DC and AC formulations of CC OPF, we calculate the number of decision variables n and minimum number of samples required for the scenario approach N_{SA}. We use parameters $\epsilon = 0.05$ and $\delta_s = 0.003$. In comparison, we note that for constant $c = 1/2$ and parameters $\delta_t = \delta_{tune} = \beta = 0.001$, the number of samples required for each step of the tuning approach is $32,209$ (equal to 64,418 samples total). Thus, we see that for the AC formulations of all test cases except the IEEE 118 bus case, using the tuning method would require less samples.

respectively. As a result, solving AC CC OPF using the scenario approach may require *substantially* more decision variables than the tuning method.

In Table 5.3, for various synthetic and industry test cases found in [ARP] and [ZMT11], for both the DC and AC problem formulations, we calculate the number of decision variables n and minimum number of samples required for the scenario approach N_{SA}. We see that in all cases, compared to the DC formulation, the AC formulation results in a much greater number of decision variables (at least one, if not two, orders of magnitude greater). Furthermore, we can compare this against the tuning approach with constant $c = 1/2$, which would require $64,418$ total samples ($N = 32,209$ samples for each step). We observe that for several test cases, using the scenario approach to solve the DC formulations may require less samples than using the tuning method. However, when the AC formulations are considered, the tuning approach generally requires significantly lower numbers of samples than the scenario approach.

In Figure 5.4, we plot the sample requirements for the AC OPF formulation of the test cases listed in Table 5.3. For the tuning approach, we plot the number of samples required for $c = 1, 2/3, 1/2, 1/3$ using the dark blue, orange, green, and pink horizontal lines, respectively. The blue diagonal line shows the required number of samples for the scenario approach. The number of decision variables for each test case is indicated using the vertical dotted lines. We can see that for all test cases, all versions of the tuning algorithm require less samples than the scenario approach. We can also observe that for the two largest test cases (the 22,700 and 31,777 bus cases), the number of samples required for the scenario approach is an order of magnitude greater than the number required for the tuning algorithm even for small c values. We consequently argue that for solving an AC OPF problem, where the number of decision variables can be very large, our tuning algorithm has a competitive edge against the scenario approach.

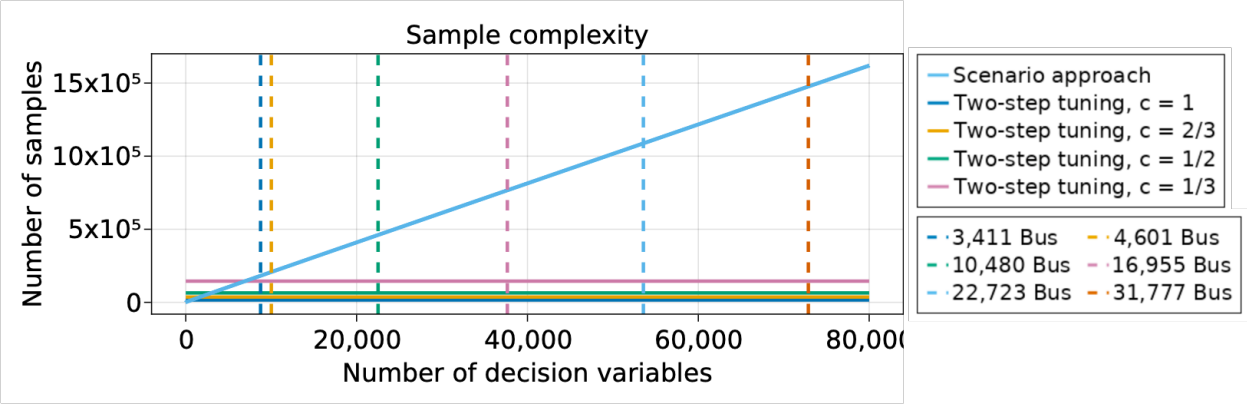


Figure 5.4: The sample complexity plotted against the number of decision variables for the scenario approach and tuning approaches with constants $c = \{1, 2/3, 1/2, 1/3\}$ for selected CC OPF test cases using the AC formulation (plotted using the vertical dotted lines). We see that for formulations of OPF with a very large number of decision variables, such as the test cases plotted, the tuning algorithm requires significantly less samples than the scenario approach.

5.3 CASE STUDY: PERFORMANCE ASSESSMENT FOR CC DC OPF

In this section, we investigate the scalability and performance of the tuning approach, extending the numerical results from the previous chapter 4.6 to larger scale networks, namely the IEEE 118 bus test case [CWW00] and the Polish Winter Peak test case [ZMT11]. We compare the optimality, conservativeness, and run times of our method against those obtained with the scenario approach and the CVaR inner approximation [NS07], detailed in Appendix A.2.

For the tuning method and scenario approach, we use the confidence parameters $\epsilon = 0.05$, $\delta_{\text{tune}} = \delta_t = \beta = 0.001$, and $\delta_s = 0.003$. For the scenario approach, we calculate the number of samples required based on these parameters using (5.19).

We note that, in general, the CVaR reformulation is a convex approximation and known to be a safe (inner) approximation of the chance-constrained problem. However, since CVaR-constrained problems are computationally intractable as-is due to the expected value term in the constraint, we use a sample average approximation with the sample set $\Xi_{\text{CVaR}} = \{\xi^{(1)}, \ldots, \xi^{(N)}\}$ to obtain a tractable formulation [WA08]. This sample-based method can introduce approximation error or inaccuracy and may not necessarily yield a solution that is guaranteed to be feasible to the underlying chance-constrained problem. As a result, we note the results obtained with the CVaR reformulation do not have theoretical guarantees in the same way as our tuning method and scenario approach. For simplicity, for the sample average approximation of the CVaR approximation, we use the same number of samples as used in the scenario approach. The exact formulation of the CVaR method we implement is presented in Appendix A.2. For simplicity, for the sample average approximation, we use the same number of samples as used in the scenario approach.

5.3.1 *IEEE 118 bus test case*

We use the same modified IEEE 118 bus system [CWW00] as detailed in Section 5.1.3. The maximum generation capacity and load are scaled by a factor of 1.25. The minimum generation capacity is set to zero. We assume that all loads fluctuate around their forecasted consumption with a standard deviation of 10% of their forecasted consumption.

In Table 5.4, we report the following results: the number of samples used, the solution cost, the solution's in-sample a posteriori evaluated violation probability, the solution's out-of-sample violation probability evaluated on $100,000$ i.i.d. samples, and the run times for the tuning approach using $c = \{1, 2/3, 1/2\}$ and the scenario approach. For the tuning approach, we also report the tuning parameter s of the solution and the run time for each algorithm component.

We see that for all choices of c, the scenario approach does require significantly less samples than the tuning approach. When $c = 1$, we see that the tuning approach results in slightly more costly and conservative solutions than both the scenario approach and CVaR approximation. However, as we decrease c and increase the number of samples used, the tuning algorithm is able to achieve lower cost, less conservative solutions than both methods. We also note a corresponding decrease in the tuning parameter s (i.e., a relaxing of the constraint tightenings), which allows for these less conservative results.

For the tuning method, we see that the run time increases roughly linearly with the number of samples required. In particular, the increase in run time is driven by the increased computation time of the sample-based evaluations, while the computation time for solving the reformulated problem remains rather constant. We additionally note that because the tuning approach requires solving the reformulation several times in order to perform the binary search on the tuning parameter, its run time is slower than that of the other compared methods, which only require solving a (larger) optimization problem once. However, we note that the run times of the tuning approach are not prohibitively

	Tuning			S.A.	CVaR
	$c = 1$	$c = 2/3$	$c = 1/2$		
Samples	16,106	36,236	64,420	1,698	1,698
Cost	116,424	116,405	116,403	116,416	116,413
Tuning parameter, s	3.502	2.452	2.244	-	-
Violation (in-sample)	0.0000	0.0156	0.0241	0.0000	0.0002
Violation (out-sample)	0.0001	0.0160	0.0245	0.0010	0.0016
Total run time	16.2 s	34.4 s	60.9 s	5.1 s	5.1 s
Solution generation	13.6 s	32.1 s	58.0 s	-	-
Solving reformulation	1.3 s	1.2 s	1.3 s	-	-
Sample-based evaluation	1.5 s	3.8 s	6.0 s	-	-
Iterations	10	10	10	-	-
Solution verification	1.6 s	3.5 s	6.0 s	-	-

Table 5.4: Results of the two-step tuning algorithm (Tuning) with $c = \{1, 2/3, 1/2\}$, the scenario approach (S.A.), and the CVaR approximation (CVaR) for the IEEE 118 bus system.

long, despite the use of significantly more samples. Furthermore, properly adjusting the tolerance parameter of the binary search may yield comparable performance results with a lower number of iterations, resulting in a lower overall run time.

5.3.2 *Polish Winter Peak test case with 2,383 buses*

We perform the same analysis for the Polish Winter Peak test case with 2,383 buses, provided with Matpower 5.1 [ZMT11]. We modify the test case according to [Roa+17] as follows: The maximum generation capacity is scaled by a factor of 2.0 and the minimum generation capacity is set to zero. The line limits are scaled by a factor of 2.5. All loads above 25 MW are assumed to have Gaussian uncertainty with a standard deviation of 20% of the forecasted consumption with zero correlation. In Table 5.5, we report the same results as in the previous section for the Polish test case.

In this case, while the number of samples used for the scenario approach (and, thus, the CVaR method) is larger in comparison to that used for the previously assessed IEEE

	Tuning			S.A.	CVaR
	$c = 1$	$c = 2/3$	$c = 1/2$		
Samples	16,106	36,236	64,420	7,889	7,889
Cost	701,019	700,894	700,880	701,030	701,006
Tuning parameter, s	3.986	2.388	2.242	-	-
Violation (in-sample)	0.0000	0.0168	0.0245	0.0000	0.0001
Violation (out-sample)	0.0003	0.0159	0.0251	0.00004	0.00027
Total run time	1 h 17 m	1 hr 23 m	1 h 43 m	15 m	16 m
Solution generation	1 h 16 m	1 h 21 m	1 hr 39 m	-	-
Solving reformulation	6 m	6 m	6 m	-	-
Sample-based evaluation	1 m	2 m	4 m	-	-
Iterations	10	10	10	-	-
Solution verification	1 m	2 m	4 m	-	-

Table 5.5: Results of the two-step tuning algorithm with $c = \{1, 2/3, 1/2\}$, the scenario approach, and the CVaR approximation for the Polish test system.

118 bus case, we still see that the scenario approach uses less samples than the tuning approach. However, we observe that for all values of c, the tuning approach results in lower costs as well as higher out-of-sample violation probabilities compared to either the scenario approach or CVaR methods. We similarly see that as we increase the amount of data we use in the tuning algorithm (i.e., by decreasing c), we can obtain solutions that are more optimal and less conservative.

We note that in both the Polish test case and in the IEEE 118 bus system, the difference in the cost of solutions amongst all algorithm runs investigated is marginal despite a notable difference in the observed violation probabilities. This may be attributed to the fact that in both test cases, the uncertainty fluctuations of the load are relatively small in comparison to the forecasted consumption. Moreover, the number of active line constraints in the reformulations is very small in comparison to the total number of constraints in the system. Therefore, using test cases with greater amounts of uncertainty or reduced the line limits will likely result in greater cost reductions as the violation probability increases.

In the Polish test case, we additionally see that the computation time of the two step tuning method is significantly longer than the scenario approach and CVaR method. As discussed in the previous section, this can mainly be attributed to the iterative re-solving of the approximate problem. Looking at the run time breakdown for the tuning approach, as we increase the number of samples used (i.e., decrease c), we see that the increase in run time is primarily driven by the increase in time required to perform the sample-based evaluations in the solution generation step and the solution verification step. However, the computation time needed to solve the reformulated problem at each step remains constant regardless of the number of samples used, and is about 3 times lower than the time needed to solve the scenario approach or CVaR optimization problems. As mentioned in the previous section, we can decrease the run time of the tuning method by adjusting the tolerance parameter of the binary search, reducing the required number of iterations. Furthermore, the run time of the sample-based evaluation and verification steps can be decrease via parallelization.

5.4 CONCLUSION

In this chapter, we extended theoretical and numerical results of the two-step tuning method for joint chance-constrained optimization problems, presented in the previous chapter.

We proposed a methodology using statistical hypothesis testing to assess the validity of the approximately unbiased estimator assumption (Assumption 1) in practical applications. We demonstrated that for the chance-constrained DC OPF problem, this assumption is mostly mild, indicating that the previously established a priori feasibility guarantee will likely hold for this setting. However, the validity of the assumption may vary for applications which have different underlying problem structures, use different

approximate problem reformulations, or implement different tuning schemes. In particular, the choice of method used in the solution generation step may play a large role in the level of bias observed due to the possibility of overfitting to the tuning samples. This means that the assessment of Assumption 1 has to be repeated for different application. We also note that, regardless of the validity of the assumption, any solution which passes the a posteriori verification step still achieves the a posteriori feasibility guarantee.

We then derived sample complexity results for the two-step tuning method, which determine the amount of data needed to run the method and highlight a relation between the number of samples required for the tuning process and the approximate conservativeness of the resulting solution. We observe that the tuning algorithm is more sample efficient than the scenario approach in applications where the number of decision variables is very large (e.g., on the order of thousands). Consequently, for power systems applications, the tuning method may be particularly attractive for problems that consider the AC formulation of the power flow equations or in situations where there is a large number of generators (e.g., when distributed energy resources are considered). We also emphasize that beyond requiring more samples, the standard scenario approach does not provide a priori guarantees for non-convex problems such as AC OPF, whereas our method can still be applied.

Finally, our numerical results for the chance-constrained DC OPF problem indicate our method is able to accommodate the usage of large sample sizes without compromising numerical tractability. For the IEEE 118 bus system and Polish test case, not only are we are able to obtain less conservative and lower cost solutions than the scenario approach and CVaR approximation by leveraging larger sample sizes, we are able to do so without a substantial increase in run time. As a result, we postulate that in data-rich environments, the tuning approach is a preferable to the scenario approach.

6 TUNING FOR DISTRIBUTION NETWORKS

In this section, we shift our focus from transmission grids to distribution grids. The high penetration of distributed energy resources (DERs) can result in increased power injection variability and uncertainty in distribution grids. As a result, developing methods that ensure security against uncertainty arising from uncontrollable device behaviors (e.g., rooftop solar PV systems and electric vehicles) is critical. Unfortunately, a direct application of the tuning method is not possible because of the unique challenges that distribution grids pose, primarily the necessity of using the full, non-linear, non-convex AC power flow equations.

The methods presented in this chapter was done in collaboration with Kshitij Girigoudar, and was published in [GHR22]. In particular, the three-phase AC OPF model formulation was contributed by Kshitij and based on [GR21]. The main contribution of the chapter is the adaption of the tuning algorithm to the three-phase AC OPF model and the corresponding numerical results. The numerical results presented in this chapterare different from and extend extend the results published in [GHR22].

We note that the notation in this chapter (along with the Chapter 7) is self-contained, unless an equation from Chapters 2 through 5 is explicitly referenced.

6.1 MOTIVATION, CHALLENGES, RELATED WORKS

Approaches for stochastic OPF with uncertain renewable energy generation and load have been well studied in the context of transmission systems [ZL11; Roa+13; Sum+15; Sum+14;

Lia+13]. However, methods developed for transmission systems often leverage DC power flow representations and single-phase equivalents to represent the system, which are not applicable to distributions grids, who are inherently unbalanced and exhibit high R/X ratios. Furthermore, while transmission operations focus primarily on congestion management and system balancing, distribution utilities focus on managing voltage magnitudes, voltage unbalance, and other power quality issues for end customers.

As a result, there is a range of stochastic OPF models that have been developed specifically for distribution grids. This includes robust and distributionally robust methods [CWZ18; Soa+18; MD18], stochastic approximation techniques [Kek+15], and chance-constrained formulations [Kar+]. These models focus primarily on managing or reducing voltage magnitude violations [CWZ18; MD18] while minimizing objectives such as cost of energy [Soa+18; Kar+; MD18], deviation from a desired power withdrawal at the substation [MD18], or losses [Kek+15; CWZ18; Kar+]. They often leverage power flow formulations that rely on a radial network topology [San+16; Ber+18; Jab06; Gan+14; Ker06], include approximate representations of unbalance [Arn+16; DZG13; LL11; ZDL17], and sometimes take advantage of iterative solution algorithms such as forward-backward sweep [KAH18; FZA16; KAH19]. Many of these formulations are linear approximations [San+16; Arn+16; DBS17] or convex relaxations [Jab06; Gan+14; DZG13; LL11; ZDL17], which may converge to solutions that are not actually AC feasible. Methods that do consider the full, non-linear, non-convex AC power flow formulation [Soa+18] or guarantee convergence to a solution that satisfies those equations [Kar+] typically only consider balanced, single-phase systems.

A particular challenge for stochastic and data-driven methods in distribution grids is the characterization of the uncertainty. Assumptions that may be true when considering large numbers of customers (e.g., the law of large numbers) no longer apply to the small number of consumers in distribution feeders, where both load and DER power injections are hard to accurately forecast and exhibit large variability over time. The most

flexible way of modeling these distributions is using scenarios. However, the quality of the solution may be sensitive to the choice of scenarios, and therefore choosing an appropriate set of samples (e.g., based on historical data) is a challenging yet important aspect that often is overlooked. To remedy this, we investigate how the number of data points and choice of scenario sets impact the solution quality. Using realistic load and solar PV data from Pecan Street [Pec], we assess how the solution changes as we use either entire days of data or randomly sampled data points across multiple days.

6.2 THREE-PHASE UNBALANCED CC AC OPF

In this section, we formulate a chance-constrained version of the three-phase unbalance AC optimal power flow (AC OPF) for distribution grids. The distribution grid model and the deterministic OPF model is based on [GR21], but was extended to account for uncertainty in collaboration with Kshitij Girigoudar.

6.2.1 *Notation*

We represent complex phasors using the following notation:

$$v = \underbrace{|v|\angle\theta}_{\text{polar form}} = \underbrace{v_d + jv_q}_{\text{rectangular form}}, \tag{6.1}$$

where $|v|$ and θ represent the magnitude and angle components of the polar form, v_d and v_q represent the real and imaginary components of the rectangular form, and $j = \sqrt{-1}$. The complex conjugate of v is denoted by v^*. We represent the element-wise product of two vectors using $\odot$.

All scalar values are denoted using small letters, while all vector counterparts and matrices are denoted using capital letters. Distribution grid parameters such as voltage magnitude, power, and admittance are expressed using per-unit quantities.

We consider a three-phase distribution grid with one slack bus and a set of remaining three-phase buses, denoted $\mathcal{N}$, with $n = |\mathcal{N}|$. Let $\Phi = \{a, b, c\}$ represent the set of phases. The superscripts a, b, and c, denote that the value corresponds with the respective phase. The distribution substation is chosen as the slack node (i.e., reference node for the voltage angle measurement) with index $i = 0$.

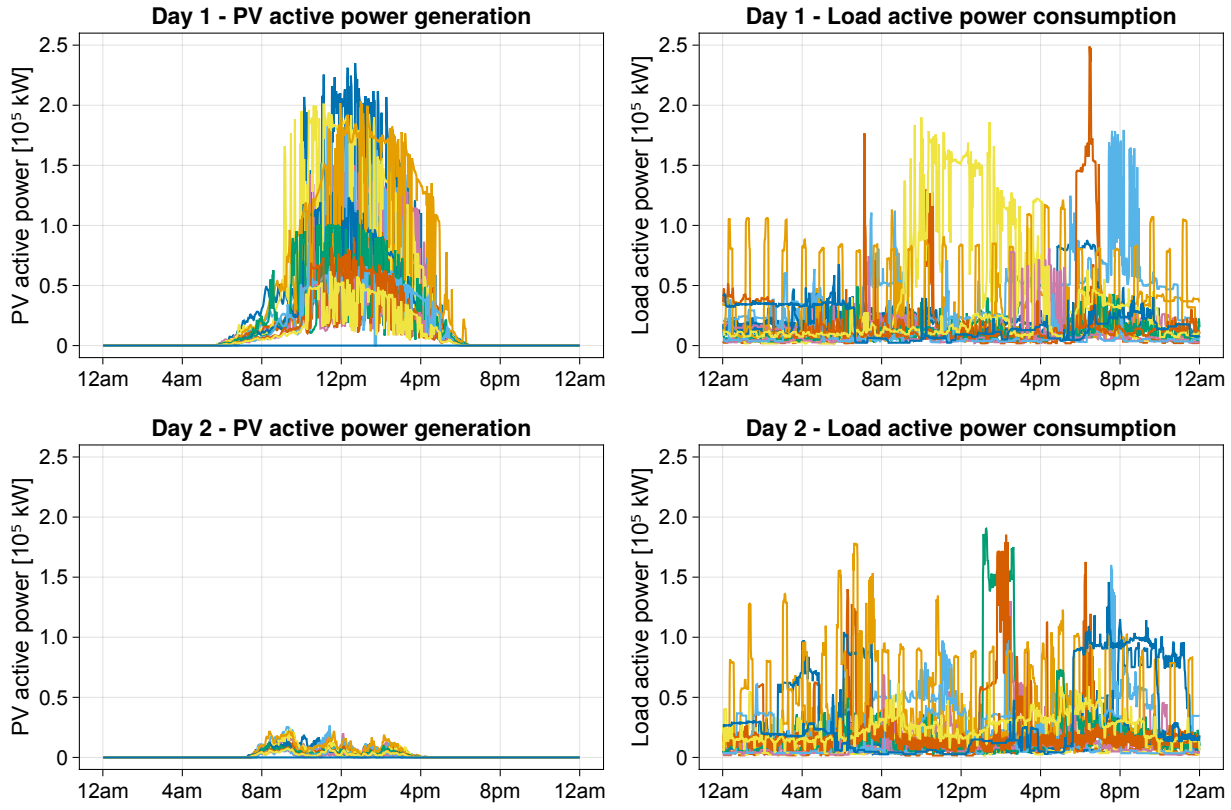


Figure 6.1: One-minute resolution Pecan Street data for 15 houses for two different days. The top shows the PV active power generation and the bottom shows the load active power demand. Day 1 (top) is be observed to have high levels of solar PV active power generation and likely represents a "sunny" day, while Day 2 (bottom) is observed to have much lower levels of solar PV generation and likely represents a "cloudy" day.

6.2.2 *Uncertainty modeling*

Due to the small number of households served by different parts of a distribution feeder, we do not observe the same smoothing effect as in transmission grids [Ker06]. Figure 6.1 shows residential solar PV and load measurements of a single house for a single day obtained from Pecan Street's data set. We can see that both profiles are highly variable and may not necessarily fit standard probability distributions. Moreover, both load and solar PV are very challenging to accurately forecast [Swe+20].

We instead represent the PV generation and load consumption of individual houses using random variables and directly leverage historical data (which may, in many cases, can be obtained with a delay by utilities using smart meters [Ali+15]. We let $\Omega := \{\omega^{(1)}, \ldots, \omega^{(N)}\}$ denote the full uncertainty data set, where each $\omega^{(i)}$ represents a grouping of a single realization of PV generation and load consumption at all houses in the network, i.e., $\omega^{(i)} \in \mathbb{R}^{nx3} \times \mathbb{R}^{nx3} \times \mathbb{R}^{nx3}$. We use a subscript ω to indicate dependency on a given uncertainty realization ω (i.e., any variable with the subscript is a random variable).

Modeling of active power generation from solar PV

We assume that the utility does not perform active power curtailment and therefore consider the active power generation of the solar PV inverters as uncontrollable random variables. We denote the three-phase active power generation at bus $i \in \mathcal{N}$ under uncertainty realization ω using

$$P_{G,i,\omega} = \begin{bmatrix} p^a_{G,i,\omega} \\ p^b_{G,i,\omega} \\ p^c_{G,i,\omega} \end{bmatrix} \in \mathbb{R}^3 \tag{6.2}$$

We denote the average active PV generation across all realizations in the sample set Ω using

$$\bar{P}_{G,i} = \begin{bmatrix} \bar{p}^{a}_{G,i} \\ \bar{p}^{b}_{G,i} \\ \bar{p}^{b}_{G,i} \end{bmatrix} \in \mathbb{R}^3, \tag{6.3}$$

where each entry of the vectors is calculated by taking the sample average across all $\omega \in \Omega$, i.e.,

$$\bar{p}^{\phi}_{G,i} = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} p^{\phi}_{G,i,\omega}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{6.4a}$$

We can similarly can decompose the uncertain active PV generation into the sum of an average term and fluctuating deviation term, which can be calculated using

$$\delta P_{G,i,\omega} := P_{G,i,\omega} - \bar{P}_{G,i} \in \mathbb{R}^3. \tag{6.5a}$$

Modeling of loads

For each bus $i \in \mathcal{N}$, we denote the three-phase active and reactive power consumption under the uncertainty realization ω using

$$P_{L,i,\omega} := \begin{bmatrix} p^{a}_{L,i,\omega} \\ p^{b}_{L,i,\omega} \\ p^{c}_{L,i,\omega} \end{bmatrix} \in \mathbb{R}^3, \qquad Q_{L,i,\omega} := \begin{bmatrix} q^{a}_{L,i,\omega} \\ q^{b}_{L,i,\omega} \\ q^{c}_{L,i,\omega} \end{bmatrix} \in \mathbb{R}^3, \tag{6.6}$$

If reactive power measurements are not available, we assume the loads operate with a constant power factor $pf_i^\phi \in \mathbb{R}, \forall\ \phi \in \Phi$, resulting in the following reactive power consumption

$$Q_{L,i,\omega} = \Gamma_{L,i} \odot P_{L,i,\omega} = \begin{bmatrix} \gamma_{L,i}^a \\ \gamma_{L,i}^b \\ \gamma_{L,i}^c \end{bmatrix} \odot P_{L,i,\omega}, \qquad \forall i \in \mathcal{N}, \omega \in \Omega. \tag{6.7}$$

where the entries of the constant factor $\Gamma_{L,i} \in \mathbb{R}^3$ are computed using

$$\gamma_{L,i}^\phi = \sqrt{\frac{1 - {pf_i^\phi}^2}{{pf_i^\phi}^2}}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{6.8}$$

We define the average and fluctuating deviation terms of the active and reactive load similarly to the active generation using

$$\bar{P}_{L,i} = \begin{bmatrix} \bar{p}_{L,i}^a \\ \bar{p}_{L,i}^b \\ \bar{p}_{L,i}^b \end{bmatrix} \in \mathbb{R}^3, \qquad \bar{Q}_{L,i} = \begin{bmatrix} \bar{q}_{L,i}^a \\ \bar{q}_{L,i}^b \\ \bar{q}_{L,i}^c \end{bmatrix} \in \mathbb{R}^3, \tag{6.9}$$

and

$$\delta P_{L,i,\omega} := P_{L,i,\omega} - \bar{P}_{L,i} \in \mathbb{R}^3 \tag{6.10a}$$

$$\delta Q_{L,i,\omega} := Q_{L,i,\omega} - \bar{Q}_{L,i} \in \mathbb{R}^3. \tag{6.10b}$$

6.2.3 *Reactive power control*

We assume that all solar PV inverters are equipped with smart inverters and provide opportunities for reactive power control as outlined in the IEEE Standard 1547-2018 [PS18].

In this chapter, we assume that the inverter is operating in the constant reactive power mode, where each inverter provides reactive power according to a given set-point,

$$\mathbf{Q}_{G,i} = \begin{bmatrix} \mathbf{q}^{a}_{G,i} \\ \mathbf{q}^{b}_{G,i} \\ \mathbf{q}^{c}_{G,i} \end{bmatrix} \in \mathbb{R}^3. \tag{6.11}$$

This set-point is provided by the utility (without accounting for, e.g., local voltage measurements) and serves as a decision variable in our problem. Our goal is to identify a suitable set-point $\mathbf{Q}_{G,i}$ that will remain safe for all uncertainty realizations ω and contributes to minimizing voltage unbalance while ensuring voltage magnitudes remain within their limits and PV inverter active and reactive power remains within apparent power limits. Because the active power generation $p^{\phi}_{G,i,\omega}$ is uncertain, we enforce the apparent power limits using chance constraints, which require that the constraint holds with high probability. For a single-phase PV inverter connected to $\phi \in \Phi$ at bus $i \in \mathcal{N}$ with apparent power rating $|s^{\phi}_{G,i}| \in \mathbb{R}$, the nominal inverter reactive power $\mathbf{q}^{\phi}_{G,i}$ is constrained by the two sided limit

$$(\mathbf{q}^{\phi}_{G,i})^2 + (p^{\phi}_{G,i,\omega})^2 \leqslant |s^{\phi}_{G,i}|. \tag{6.12}$$

This is equivalent to the following one sided constraints

$$\mathbf{q}^{\phi}_{G,i} \leqslant \sqrt{|s^{\phi}_{G,i}| - (p^{\phi}_{G,i,\omega})^2} \tag{6.13a}$$

$$\mathbf{q}^{\phi}_{G,i} \geqslant -\sqrt{|s^{\phi}_{G,i}| - (p^{\phi}_{G,i,\omega})^2} \tag{6.13b}$$

$$\tag{6.13c}$$

Adding the chance constraints, we obtain

$$\mathbb{P}_\omega\left(\mathbf{q}^{\phi}_{G,i} \leqslant \sqrt{|s^{\phi}_{G,i}| - (p^{\phi}_{G,i,\omega})^2}\right) \geqslant 1-\epsilon \tag{6.14a}$$

$$\mathbb{P}_\omega\left(\mathbf{q}^{\phi}_{G,i} \geqslant -\sqrt{|s^{\phi}_{G,i}| - (p^{\phi}_{G,i,\omega})^2}\right) \geqslant 1-\epsilon \tag{6.14b}$$

where $\epsilon \in [0,1]$ is the acceptable violation probability. Here, we use individual chance constraints for each single-phase connection and choose to use the same ϵ parameter across all inverter constraints.

We assume the chance constraints enforcing the apparent power limit of the PV inverter is *hard constraints*. If the reactive power exceeds these limits, the PV inverter becomes overloaded and would not operate as expected. In this case, the inverter would not provide the reactive power according to the set point, but rather trigger protections that can, e.g., disconnect the inverter or cap the reactive power generation at the limit dictated by the apparent rating of the inverter and the active power generation. Thus, a violation of these constraints (6.14) should be interpreted as the probability that an inverter is not able to provide the desired reactive power to the grid.

6.2.4 *Distribution grid model*

We briefly give an overview of the distribution grid model used, and refer the read to [GR21] for a detailed formulation.

The active and reactive power generation variables at the substation, which is chosen as the slack node with index $i = 0$, are denoted

$$P_{G,0} = \begin{bmatrix} p^a_{G,0} \\ p^b_{G,0} \\ p^c_{G,0} \end{bmatrix} \in \mathbb{R}^3, \qquad Q_{G,0} = \begin{bmatrix} q^a_{G,0} \\ q^b_{G,0} \\ q^c_{G,0} \end{bmatrix} \in \mathbb{R}^3. \tag{6.15}$$

Without loss of generality, we assume all buses have three phases. For single and two-phase nodes, we set the entries corresponding to the missing nodes to zero. The resulting total number of single-phase connections in the distribution grid is $3(n+1)$. We assume that there is one solar PV inverter and one load at each single-phase connection of every node $i \in \mathcal{N}$. If any node connected to a phase has no source or load, we set the corresponding entries to zero. We further note that our model does not make any assumptions on radiality, and therefore can be applied to distribution grids that are operated in non-radial configurations.

Grid parameters

Following [GR20], the critical distribution grid components are modeled using an overall nodal admittance matrix

$$Y = G + jB \in \mathbb{C}^{3(n+1)\times 3(n+1)}, \tag{6.16}$$

where $G, B \in \mathbb{R}^{3(n+1)\times 3(n+1)}$ denote the nodal conductance and susceptance matrices, respectively.

Nodal power injections

The active power injection $P_{i,\omega} \in \mathbb{R}^3$ at each node $i \in \mathcal{N}$ depend on the uncertainty realization ω and is defined as the difference between the active power generation and the active load consumption,

$$P_{i,\omega} := P_{G,i,\omega} - P_{L,i,\omega} = (\bar{P}_{G,i} + \delta P_{G,i,\omega}) - (\bar{P}_{L,i}\delta P_{L,i,\omega}). \tag{6.17}$$

Similarly, the reactive power injection $Q_{i,\omega} \in \mathbb{R}^3$ at node $i \in \mathcal{N}$ is defined as the difference between the controllable reactive power generation of solar PV inverters $\mathbf{Q_G}_{,i}$ and the reactive power load $Q_{L,i,\omega}$,

$$Q_{i,\omega} := \mathbf{Q_G}_{,i} - Q_{L,i,\omega} = \mathbf{Q_G}_{,i} - (\bar{Q}_{L,i} - \delta Q_{L,i,\omega}) \tag{6.18}$$

The average active and reactive power injections at each node $i \in \mathcal{N}$ can be expressed by

$$\bar{P}_i = \bar{P}_{G,i} - \bar{P}_{L,i} \tag{6.19}$$

$$\bar{Q}_i = \mathbf{Q_G}_{,i} - \bar{Q}_{L,i}. \tag{6.20}$$

The only controllable active power generator in the network is the substation and therefore it maintains the active power balance by supplying the difference between the active power generation and load for all uncertainty realizations as well as any additional power needed to cover power losses. Since the substation has no load, the active power injection is equal to the active power injection. The active power generation at the substation $\mathbf{P_G}_{,0,\omega} \in \mathbb{R}^3$ is taken as a decision variable in the optimization problem.

Similarly, the substation also guarantees reactive power balance, covering the difference between production, consumption, and losses. The reactive power injection at the substation is also taken as a decision variable $\mathbf{Q_G}_{,0,\omega} \in \mathbb{R}^3$.

Voltage representation

The three-phase OPF is implemented in the polar coordinate frame using the phase-to-neutral voltage magnitude and angle variables at every node $i \in \mathcal{N}$. The voltage

magnitudes and angles corresponding to the average power injections $\bar{P}_i$, $\bar{Q}_i$ are denoted by

$$|\mathbf{V}_i| = \begin{bmatrix} |\mathbf{v}_i^a| \\ |\mathbf{v}_i^b| \\ |\mathbf{v}_i^c| \end{bmatrix} \in \mathbb{R}^3, \qquad \mathbf{\Theta}_i = \begin{bmatrix} \theta_i^a \\ \theta_i^b \\ \theta_i^c \end{bmatrix} \in \mathbb{R}^3, \tag{6.21}$$

respectively. As the active and reactive power injections change, the voltages change as well. The voltage magnitude and angle under an uncertainty realization ω is denoted by

$$|\mathbf{V}_{i,\omega}| = \begin{bmatrix} |\mathbf{v}_{i,\omega}^a| \\ |\mathbf{v}_{i,\omega}^b| \\ |\mathbf{v}_{i,\omega}^c| \end{bmatrix} \in \mathbb{R}^3, \qquad \mathbf{\Theta}_i = \begin{bmatrix} \theta_{i,\omega}^a \\ \theta_{i,\omega}^b \\ \theta_{i,\omega}^c \end{bmatrix} \in \mathbb{R}^3, \tag{6.22}$$

The distribution substation is considered as the reference for voltage angle measurements. We assume the voltage is independent of ω and fixed at the substation, i.e.,

$$|\mathbf{V}_0| \angle \mathbf{\Theta}_0 = \begin{bmatrix} 1\angle 0^\circ \\ 1\angle -120^\circ \\ 1\angle 120^\circ \end{bmatrix}. \tag{6.23}$$

Since voltage magnitudes at all buses $i \in \mathcal{N}$ (i.e., not the substation) have a dependence on the uncertainty, we enforce the voltage magnitude limits using the following single chance constraints

$$\mathbb{P}_\omega \left(|\mathbf{v}_{i,\omega}^\phi| \leqslant \overline{v} \right) \geqslant 1 - \epsilon, \qquad \forall i \in \mathcal{N}, \phi \in \Phi, \tag{6.24}$$

$$\mathbb{P}_\omega \left(|\mathbf{v}_{i,\omega}^\phi| \geqslant \underline{v} \right) \geqslant 1 - \epsilon, \qquad \forall i \in \mathcal{N}, \phi \in \Phi, \tag{6.25}$$

where $\underline{v}$ and $\overline{v} \in \mathbf{R}$ are the lower and upper voltage magnitude bounds and $\epsilon \in [0, 1]$ is the acceptable violation probability. These voltage magnitude constraints can be considered

as *soft constraints*, where a constraint violation indicates an under- or over-voltage. The violations of soft constraints is acceptable if the duration and magnitude are small. Moreover, the acceptable violation probabilities can vary for each constraint, but we choose to use the same probability ϵ for all chance constraints on voltage magnitude constraints in this chapter.

Power flow

We use the following shorthand to represent the AC power flow equations at all nodes $i \in \{0, \mathcal{N}\}$ for all uncertainty realizations $\omega \in \Omega$,

$$f(|\mathbf{V}_{i,\omega}|, \boldsymbol{\Theta}_{i,\omega}, P_{i,\omega}, Q_{i,\omega}) = 0, \qquad \forall i \in \{0, \mathcal{N}\},\ \omega \in \Omega. \tag{6.26}$$

The equations are functions of the voltage magnitude $|\mathbf{V}_{i,\omega}|$ and voltage angle $\boldsymbol{\Theta}_{i,\omega}$ as well as the active power $P_{i,\omega}$ and nodal reactive power $Q_{i,\omega}$.

This shorthand represents the following power balance equations

$$P_{i,\omega} = |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0,\mathcal{N}\}} \left[G_{ik} \odot C\left(\boldsymbol{\Theta}_{ik,\omega}+\right) + B_{ik} \odot S\left(\boldsymbol{\Theta}_{ik,\omega}\right) \right] |\mathbf{V}_{k,\omega}| \tag{6.27a}$$

$$Q_{i,\omega} = |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0,\mathcal{N}\}} \left[G_{ik} \odot C\left(\boldsymbol{\Theta}_{ik,\omega}+\right) - B_{ik} \odot S\left(\boldsymbol{\Theta}_{ik,\omega}\right) \right] |\mathbf{V}_{k,\omega}| \tag{6.27b}$$

where $G_{ik}, B_{ik} \in \mathbb{R}^{3\times3}$ represent the real and imaginary sub-matrices of the nodal admittance matrix Y for a three-phase branch ik and $C(\boldsymbol{\Theta}_{ik,\omega})$ and $S(\boldsymbol{\Theta}_{ik,\omega})$ represent the cosine and sine components of the branch angle matrix $\boldsymbol{\Theta}_{ik,\omega} \in \mathbb{R}^{3\times3}$. For a detailed treatment of the power flow equations, we refer the reader to [GR20].

6.2.5 *Objective function*

In this chapter, we choose to minimize voltage unbalance as our OPF objective. We utilize the IEC standard 61000-2-2 [02] commonly referred to as Voltage Unbalance Factor (VUF) to define voltage unbalance. For a three-phase node l, the square of VUF for the uncertainty realization ω is defined by

$$\mathrm{VUF}_{l,\omega}^2 = \frac{|v_{l,\omega}^-|^2}{|v_{l,\omega}^+|}^2 = \frac{(\mathbf{v}_{\mathbf{d}l,\omega}^-)^2 + (\mathbf{v}_{\mathbf{q}l,\omega}^-)^2}{(\mathbf{v}_{\mathbf{d}l,\omega}^+)^2 + (\mathbf{v}_{\mathbf{q}l,\omega}^+)^2}, \tag{6.28}$$

where $\mathbf{v}_{\mathbf{d}l,\omega}^-, \mathbf{v}_{\mathbf{q}l,\omega}^-$ and $\mathbf{v}_{\mathbf{d}l,\omega}^+, \mathbf{v}_{\mathbf{q}l,\omega}^+$ are the rectangular form representation of negative sequence voltage $v_{l,\omega}^-$ and positive sequence voltage $v_{l,\omega}^+$, respectively. The terms are non-linear functions of the voltage variables $\mathbf{V}_{l,\omega}, \mathbf{\Theta}_{l,\omega}$ and defined using

$$\mathbf{v}_{\mathbf{d}l,\omega}^- = \Re\{v_{l,\omega}^-\}, \qquad \mathbf{v}_{\mathbf{q}l,\omega}^- = \Im\{v_{l,\omega}^-\}, \tag{6.29a}$$

$$\mathbf{v}_{\mathbf{d}l,\omega}^+ = \Re(v_{l,\omega}^+), \qquad \mathbf{v}_{\mathbf{q}l,\omega}^+ = \Im(v_{l,\omega}^+), \tag{6.29b}$$

where

$$v_{l,\omega}^- = |\mathbf{v}_{l,\omega}^a| \angle \theta_{l,\omega}^a + |\mathbf{v}_{l,\omega}^b| \angle (\theta_{l,\omega}^b - 120^\circ) + |\mathbf{v}_{l,\omega}^c| \angle (\theta_{l,\omega}^c + 120^\circ), \tag{6.29c}$$

$$v_{l,\omega}^+ = |\mathbf{v}_{l,\omega}^a| \angle \theta_{l,\omega}^a + |\mathbf{v}_{l,\omega}^b| \angle (\theta_{l,\omega}^b + 120^\circ) + |\mathbf{v}_{l,\omega}^c| \angle (\theta_{l,\omega}^c - 120^\circ). \tag{6.29d}$$

We refer the reader to [GR20] and [GHR22] for full details on the formulation.

6.2.6 *Problem formulation*

The optimization problem minimizes the squared VUF (6.28) subject to the power balance constraints (6.26) and constraints on the voltage magnitude (6.24) and inverter limits (6.14).

The optimization decision variables are the voltage magnitudes at all nodes $|\mathbf{V}| \in \mathbb{R}^{3(n+1)}$, voltage angles at all nodes $\mathbf{\Theta} \in \mathbb{R}^{3(n+1)}$, active power generation at the substation $\mathbf{P_{G,0}} \in \mathbb{R}^3$, and reactive power generation at all nodes $\mathbf{Q_G} \in \mathbb{R}^{3(n+1)}$.

As with the DC OPF case, either single or joint chance constraints can be used. In this chapter, we focus on the single chance-constrained case, which is formulated as follows:

$$\min_{\mathbf{P_{G,0}}, \mathbf{Q_G}, |\mathbf{V}|, \mathbf{\Theta}} \sum_{\omega \in \Omega} \sum_{l \in \mathcal{N}} \text{VUF}_{l,\omega}^2 \tag{6.30a}$$

$$\text{s.t.} \quad f\left(|\mathbf{V}_{i,\omega}|, \mathbf{\Theta}_{i,\omega}, P_{i,\omega}, Q_{i,\omega}\right) = 0, \qquad \forall\, i \in \{0, \mathcal{N}\},\ \omega \in \Omega, \tag{6.30b}$$

$$\mathbb{P}_\omega\left(|\mathbf{v}_{i,\omega}^{\phi}| \leqslant \overline{v}\right) \geqslant 1 - \epsilon, \qquad \forall\, \phi \in \Phi,\ i \in \mathcal{N}, \tag{6.30c}$$

$$\mathbb{P}_\omega\left(|\mathbf{v}_{i,\omega}^{\phi}| \geqslant \underline{v}\right) \geqslant 1 - \epsilon, \qquad \forall\, \phi \in \Phi,\ i \in \mathcal{N}, \tag{6.30d}$$

$$\mathbb{P}_\omega\left(\mathbf{q}_{\mathbf{G},i}^{\phi} \leqslant \sqrt{|s_{G,i}^{\phi}| - (p_{G,i,\omega}^{\phi})^2}\right) \geqslant 1 - \epsilon, \qquad \forall\, \phi \in \Phi,\ i \in \mathcal{N}, \tag{6.30e}$$

$$\mathbb{P}_\omega\left(\mathbf{q}_{\mathbf{G},i}^{\phi} \geqslant -\sqrt{|s_{G,i}^{\phi}| - (p_{G,i,\omega}^{\phi})^2}\right) \geqslant 1 - \epsilon, \qquad \forall\, \phi \in \Phi,\ i \in \mathcal{N}, \tag{6.30f}$$

$$|\mathbf{V}_0| \angle \mathbf{\Theta}_0 = \begin{bmatrix} 1\angle 0^\circ & 1\angle -120^\circ & 1\angle 120^\circ \end{bmatrix}^\top, \tag{6.30g}$$

In general, the chance constraints are interpreted as the likelihood that the distribution grid operation will need to take extra control actions in real time to protect the system (e.g., from inverter overloading or over- and under-voltages). By choosing high acceptable violation probabilities ϵ the system is at a higher risk of insecure operation as it may necessitate frequent deployment of real-time controls, which are not always available. However, choosing low violation probabilities is expensive, but makes system operation safer and less stressful for the operator [RA17].

We emphasize that both the objective function and the constraints are highly non-linear in the decision variables. Moreover, there is no explicit way in which we can represent how the uncertainty propagates throughout the system and thus no straightforward analytical reformulation can be applied (as done in the DC OPF case presented in Chapter

3). In the following, we detail how we formulate an approximate model suitable for this problem that can allow for the application of the tuning method proposed in Chapter 3.

6.3 APPLICATION OF TUNING METHOD

In this section, we apply the tuning method detailed in Section 3.2.3 to the three-phase CC AC OPF problem. This problem is significantly more challenging than the CC DC OPF problem considered in Chapters 4 and 5 because the inner constraints of the chance constraints are non-linear in the optimization decision variables as well as the uncertainty. Moreover, the uncertainty also enters the problem constraints implicitly (e.g., via power losses). As a result, we cannot use the tuning approach exactly as detailed in Section 3.2.3, but need to make some modifications to adapt the method to this problem.

We note that this non-linear, non-convex problem settings highlights the usefulness of the tuning method. We are able to use the full non-linear, non-convex AC power flow equations, without needing to perform any explicit linearizations or approximations. Because we deliberately choose to use an approximate reformulation that is simple to solve, the resulting algorithm can also take advantage of the computational efficiency of commercial solvers. Furthermore, the tuning method allows us to *decouple* the solving of the three-phase AC OPF problem from the consideration of the uncertainty, which is completely captured via the tightening terms. However, the main challenge of using this style of deterministic reformulation is accurately determining and updating the values of the uncertainty margins. Unlike in the DC OPF case, there is not an explicit analogous relationship to analytical reformulations of the problem or robust optimization interpretation, which we can use to formulate the tightening. The remainder of this section details the main components of the iterative approach.

6.3.1 *Approximate model formulation*

As the approximate problem formulation, we choose to use a generalized reformulation based on [RA17].

Power flow

Rather than enforcing the power flow equations in (6.27) for all uncertainty realizations, which may be computationally prohibitive to do, we enforce a single set of power flow equations at each node using the average power injections $\bar{P}_i$, $\bar{Q}_i$ and corresponding voltage variables $|\mathbf{V}_i|$, $\mathbf{\Theta}_i$. The shorthand representation of the power balance constraint becomes

$$f(|\mathbf{V}_i|, \mathbf{\Theta}_i, \overline{P}_i, \overline{Q}_i) = 0, \qquad \forall i \in \{0, \mathcal{N}\}. \tag{6.31}$$

Inverter limits

We replace the inverter limits (6.14) with deterministic box constraints, consisting of a nominal limit with a tightening term. The upper and lower limits for a PV inverter connected to phase $\phi \in \Phi$ at node $i \in \mathcal{N}$ are calculated using (6.13), with the average generation $\overline{p}^{\phi}_{G,i}$ in place of the random PV active generation $p_{G,i,\omega}$, i.e.,

$$\overline{q}^{\phi}_{G,i} := \sqrt{|s^{\phi}_{G,i}|^2 - (\overline{p}^{\phi}_{G,i})^2}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.32a}$$

$$\underline{q}^{\phi}_{G,i} := -\sqrt{|s^{\phi}_{G,i}|^2 - (\overline{p}^{\phi}_{G,i})^2}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{6.32b}$$

We denote the tightenings for the upper and lower limits of a single-phase connection using $\overline{\lambda}^{\phi}_{q,i}$ and $\underline{\lambda}^{\phi}_{q,i} \in \mathbb{R}^{+}$. For notational convenience, the following vectors are used express the tightenings for all buses and phases:

$$\underline{\Lambda}_q = \left[\left[\underline{\lambda}^{\phi}_{q,i}\right]^{\top}_{\phi\in\Phi}\right]^{\top}_{i\in\mathcal{N}}, \qquad \overline{\Lambda}_q = \left[\left[\overline{\lambda}^{\phi}_{q,i}\right]^{\top}_{\phi\in\Phi}\right]^{\top}_{i\in\mathcal{N}}.$$

These tightenings can be pre-calculated since the inverter reactive power limits depend only on the uncertainty samples, which are known a priori.

For example, for the upper inverter limit, we evaluate the limit $\overline{q}^{\phi}_{G,i,\omega}$ for each uncertainty sample $\omega \in \Omega$ using Eq. (6.32a) with the sampled value $p_{G,i,\omega}$ in place of the average value $\overline{p}^{\phi}_{G,i}$. By calculating this limit under all uncertainty realizations, we obtain an *e*mpirical distribution for the upper reactive power generation limit, which we denote as $\hat{F}^{\phi}_{q,i}(\cdot)$. The empirical distribution for the lower reactive power generation limit would simply be the negative, i.e., $-\hat{F}^{\phi}_{q,i}(\cdot)$. We then find the desired quantile of the empirical distribution and use them to directly calculate the constraint tightenings. For the chance constraint on the upper reactive power limit to hold, we require the upper limit to be set to $\hat{F}^{\phi}_{q,i}(\epsilon)$, which is the ϵ quantile of the empirical distribution of the upper reactive power limit. We can observe that the appropriate constraint tightening to ensure the chance constraint holds corresponds to the difference between the nominal inverter lower limit $\underline{q}^{\phi}_{G,i}$ and $\hat{F}^{\phi}_{q,i}(\epsilon)$.

Thus, the upper and lower constraint tightenings for reactive power limits are calculated as follows:

$$\overline{\lambda}^{\phi}_{q,i} = \overline{q}^{\phi}_{G,i} - \hat{F}^{\phi}_{q,i}(\epsilon), \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.33}$$

$$\underline{\lambda}^{\phi}_{q,i} = -\hat{F}^{\phi}_{q,i}(1-\epsilon) - \underline{q}^{\phi}_{G,i}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{6.34}$$

This process is also illustrated in Figure 6.2. We note that these tightenings are independent of our decision variables, and as a result they can be calculated before we

start solving the optimization problem. Thus, once these tightening values are calculated, they remain fixed throughout the entirety of the iterative algorithm and are not iteratively tuned.

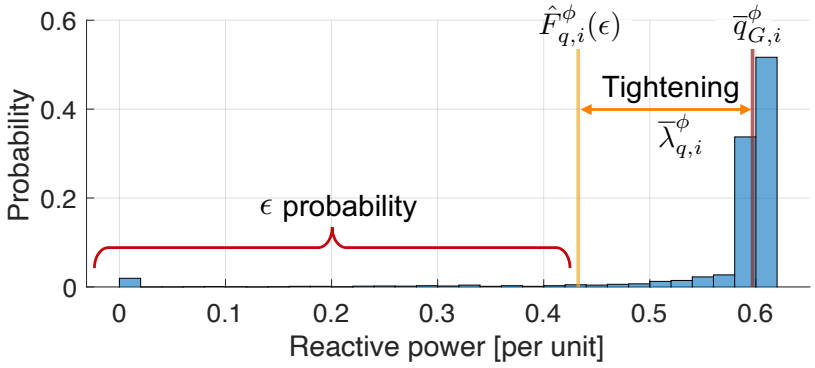


Figure 6.2: The histogram represents the empirical distribution of an example inverter reactive power upper limit, $\hat{F}^{\phi}_{q,i}(\cdot)$. The constraint tightening $\overline{\lambda}^{\phi}_{q,i}$, illustrated with the orange line, is given by the distance between the nominal limit $\overline{q}^{\phi}_{G,i}$ (brown line) and the ϵ quantile of the empirical distribution of the upper reactive power limit $\hat{F}^{\phi}_{q,i}(\epsilon)$ (yellow line).

Voltage limits

We similarly replace the chance constraints on the upper and lower voltage magnitude constraints (6.24) with the corresponding nominal limits and tightenings, denoted by $\overline{\lambda}^{\phi}_{v,i}, \underline{\lambda}^{\phi}_{v,i} \in \mathbb{R}^{+}$, respectively. The vector representations are

$$\underline{\lambda}_V = \left[\left[\underline{\lambda}^{\phi}_{V,i}\right]_{\phi\in\Phi}\right]^{\top}_{i\in\mathcal{N}}, \qquad \overline{\lambda}_V = \left[\left[\overline{\lambda}^{\phi\in\Phi}_{V,i}\right]_{\phi\in\Phi}\right]^{\top}_{i\in\mathcal{N}}.$$

In contrast to the inverter limits, the voltage constraint tightenings *need to be iteratively calculated and updated* within the tuning algorithm since they require obtaining a solution from the approximate problem in order to be determined.

Objective function

We calculate the deterministic VUF by replacing the voltage variables $\mathbf{V}_{l,\omega}, \mathbf{\Theta}_{l,\omega}$ with their counterparts $|\mathbf{V}_l|, \mathbf{\Theta}_l$ that correspond to the average power injections in (6.28) and (6.29).

Reformulated problem

The resulting approximate optimization problem obtained from using the reformulation described above is deterministic and takes the following form:

$$\min_{\mathbf{P}_{G,0}, \mathbf{Q}_G, |\mathbf{V}|, \mathbf{\Theta}} \quad \sum_{l \in \mathcal{N}} \mathrm{VUF}_l^2 \tag{6.35a}$$

$$\text{s.t.} \quad f(|\mathbf{V}_i|, \mathbf{\Theta}_i, \bar{P}_i, \bar{Q}_i) = 0, \qquad \forall i \in \{0, \mathcal{N}\} \tag{6.35b}$$

$$|\mathbf{v}_i^\phi| \leqslant \overline{v} - \overline{\lambda}_{v,i}^\phi, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.35c}$$

$$|\mathbf{v}_i^\phi| \geqslant \underline{v} + \underline{\lambda}_{v,i}^\phi, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.35d}$$

$$\mathbf{q}_{G,i}^\phi \leqslant \overline{q}_{G,i}^\phi - \overline{\lambda}_{q,i}^\phi, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.35e}$$

$$\mathbf{q}_{G,i}^\phi \geqslant \underline{q}_{G,i}^\phi + \underline{\lambda}_{q,i}^\phi, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.35f}$$

$$\mathbf{V}_0 \angle \mathbf{\Theta}_0 = \begin{bmatrix} 1\angle 0^\circ & 1\angle -120^\circ & 1\angle 120^\circ \end{bmatrix}^\top. \tag{6.35g}$$

Let $\mathbf{X} := (\mathbf{P}_{G,0}, \mathbf{Q}_G, |\mathbf{V}|, \mathbf{\Theta})$ denote a solution to the reformulated problem (6.35).

6.3.2 *Solution evaluation*

After we solve the approximate problem (6.35) and obtain a candidate solution $\mathbf{X}$, we use the available uncertainty samples to evaluate the conservativeness of the voltage constraint tightenings by estimating the empirical violation probability of the solution, using the same methodology as in Section 3.2.2.

As an example, let us take the upper voltage magnitude constraint for node $i \in \mathcal{N}$ connected to phase $\phi \in \Phi$ and samples $p^{\phi}_{G,i,\omega}$ and $p^{\phi}_{L,i,\omega}$. We define the indicator random variable

$$Y^{\phi}_{\overline{v},i}(\mathbf{X}, p^{\phi}_{G,i,\omega}, p^{\phi}_{L,i,\omega}) := \begin{cases} 0 & \text{if } |\mathbf{v}^{\phi}_{i,\omega}| \leqslant \overline{v}, \\ 1 & \text{otherwise,} \end{cases} \tag{6.36}$$

which evaluates to 0 if the upper voltage constraint holds and 1 otherwise. Then, the empirical violation probability for this constraint is calculated by averaging over all the indicator random variables calculating for each sample $\omega \in \Omega$, i.e.,

$$\hat{E}^{\phi}_{\overline{v},i}(\mathbf{X}) := \frac{1}{|\Omega|} \sum_{\omega \in \Omega} Y^{\phi}_{\overline{v},i}(\mathbf{X}, p^{\phi}_{G,i,\omega}, p^{\phi}_{L,i,\omega}). \tag{6.37}$$

We similarly evaluate the empirical violation probabilities for the voltage lower limits $\hat{E}^{\phi}_{\underline{v},i}(\mathbf{X})$, inverter upper limits $\hat{E}^{\phi}_{\overline{q},i}(\mathbf{X})$, and inverter lower limits $\hat{E}^{\phi}_{\underline{q},i}(\mathbf{X})$.

For a candidate solution $\mathbf{X}$, the worst case empirical violation probability across a single constraint type is defined as the maximum empirical violation probability observed across all constraints of that type. For example, the worst case empirical violation probability for the voltage upper limits is defined as follows

$$\hat{E}^{\max}_{\overline{v}}(\mathbf{X}) = \max_{i \in \mathcal{N}, \phi \in \Phi} \left\{ \hat{E}^{\phi}_{\overline{v},i}(\mathbf{X}) \right\}. \tag{6.38}$$

Furthermore, the worst case empirical violation probability across *all* voltage constraints (both the upper and lower limits) is defined as

$$\hat{E}^{\max}_{v}(\mathbf{X}) = \max \left\{ \hat{E}^{\max}_{\overline{v}}(\mathbf{X}), \hat{E}^{\max}_{\underline{v}}(\mathbf{X}) \right\}. \tag{6.39}$$

Since in this chapter we are only considering the single chance-constrained problem, we perform tuning based on the worst case empirical violation probability amongst all voltage constraints.

6.3.3 *Tuning algorithm*

This solution algorithm employs a tuning-based approached adapted from Chapter 3. There are several new challenges associated with the reformulated problem that arise from the use of the AC representation.

First, it is not possible to exactly define an explicit reformulation for the voltage constraint tightenings as the uncertainty implicitly enters in the voltage constraints. Instead, we take a similar approach to what is done in Chapter 3 and define a constraint tightenings parameterized by an adjustable parameter. However, unlike in Chapter 3, we cannot explicitly define the standard deviation of the nominal constraint and have to estimate it using samples.

Second, the use of the bisection search for this problem may be problematic due to the lack of monotonicity. In the reformulation of the DC OPF model 3.19 in Chapter 3, the constraint tightenings consisted of the tuning parameter s multiplied by a *constant* standard deviation term (i.e., the tightening did not dependent on any decision variables). As a result, the tightening would increase monotonically with s. Since the reformulation was a linear program, the feasible space of the reformulated optimization problem would decrease as the constraints were tightened, resulting in a higher objective value. Moreover, the empirical violation probability would monotonically decrease as s increased since the constraints were linear in the uncertainty. Here, these relationships no longer hold because the reformulation is a non-linear optimization problem and the uncertainty enters the constraints non-linearly and implicitly. While the monotonicity principle no longer

holds, we still propose that the bisection search can still be a useful heuristic for tuning s to find high quality solutions.

To perform tuning for the three-phase AC OPF problem, we similarly define the tightenings as parameterized by a single-dimensional tuning parameter, denoted $s \in \mathbb{R}^+$, and use a bisection search to adjust s based on the evaluated worst case empirical violation probability amongst the all voltage constraints $\hat{E}_v^{\max}(\mathbf{X})$.

The voltage magnitude constraint tightenings are defined as the product of the tuning parameter s and an estimation of the standard deviation of the nominal constraint, resulting in symmetric tightenings for the upper and lower voltage magnitude constraints. The tightenings take the following form:

$$\overline{\lambda}_{v,i}^{\phi} = \underline{\lambda}_{v,i}^{\phi} := s \cdot \sigma_{v,i}^{\phi}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{6.40}$$

where $\sigma_{v,i}^{\phi} \in \mathbb{R}^+$ represents the estimated standard deviation of the voltage magnitude at node $i \in \mathcal{N}$ connected to phase $\phi \in \Phi$. We determine this estimate before the commencement of the algorithm and keep it constant throughout the tuning. To obtain this estimate, we first solve the deterministic reformulation (6.35) using voltage constraint tightenings set to zero $\overline{\Lambda}_v = \underline{\Lambda}_v = 0$ and inverter limit tightenings $\overline{\Lambda}_q, \underline{\Lambda}_q$ initialized according to Section 6.3.1. After obtaining an operating point $\mathbf{X}^{(0)}$, we run power flows using the samples $\omega \in \Omega$ to obtain an empirical distribution of the voltage $f_{v,i}^{\phi,(0)}(\cdot)$, similarly to what is done in Section 6.3.1. The standard deviation of the voltage empirical distribution is taken as the estimated standard deviation $\sigma_{v,i}^{\phi}$.

The operating point $\mathbf{X}^{(0)}$ is also used to initialize the upper bound $s_{\max}^{(0)}$ and lower bound $s_{\min}^{(0)}$ of the tuning parameter. We aim to initialize these bounds such that the mid-point of the bounds (which will be the initial tuning parameter value $s^{(0)}$ since we

are performing a bisection search) will approximate taking the ϵ quantile. The initial bounds are as follows:

$$s_{\min}^{(0)} = 0, \tag{6.41a}$$

$$s_{\max}^{(0)} = \frac{2}{\sigma_{v,i}^{\phi}} \cdot \left(\max_{i \in \mathcal{N}, \phi \in \Phi} \left\{ |\mathbf{v}_i^{\phi,(0)}| - f_{v,i}^{\phi,(0)}(\epsilon_v) \right\} \right). \tag{6.41b}$$

The lower bound corresponds to the case where all tightenings are zero. The upper bound calculates the maximum difference between the voltage magnitude operating point and ϵ_v quantile across all single phase connections. The resulting value is divided by the standard deviation estimate and doubled.

The tuning method comprises of the following steps:

(i) *Initialize:* Set the iteration count to $k = 0$ and calculate the inverter limit tightenings $\overline{\Lambda}_q, \underline{\Lambda}_q$ as described in Section 6.3.1. Initialize the tuning parameter bounds $s_{\min}^{(0)}, s_{\max}^{(0)}$ according to (6.41) and calculate the initial value of the tuning parameter by taking the mid-point of the bounds

$$s^{(0)} = (s_{\max}^{(0)} - s_{\min}^{(0)})/2 + s_{\min}^{(0)}. \tag{6.42}$$

Calculate $\overline{\Lambda}_v^{(0)}, \underline{\Lambda}_v^{(0)}$ according to (6.40).

(ii) *Solve approximate problem:* Solve (6.35) using tightenings $\overline{\Lambda}_v^{(k)}, \underline{\Lambda}_v^{(k)}, \overline{\Lambda}_q, \underline{\Lambda}_q$ to obtain candidate operating point $\mathbf{X}^{(k+1)}$.

(iii) *Solution evaluation:* Perform the empirical violation probability evaluation described in Section 6.3.2 to calculate the worst case empirical violation probability $\hat{E}_v^{\max}(\mathbf{X})$ as defined in (6.39).

(iv) *Update tightenings:*

If $\hat{E}_v^{\max}(\mathbf{X}^{(k+1)}) \leqslant \epsilon_v$, decrease s by setting $s_{\max}^{(k+1)} = s^{(k)}$ to obtain a less conservative solution.

If $\hat{E}_v^{\max}(\mathbf{X}^{(k+1)}) > \epsilon_v$, increase s by setting $s_{\min}^{(k+1)} = s^{(k)}$ to obtain a more conservative solution.

Obtain a new tuning parameter $s^{(k+1)} = (s_{\max}^{(k+1)} - s_{\min}^{(k+1)})/2 + s_{\min}^{(k+1)}$ and update the tightenings according to (6.40).

(v) *Check convergence:* Terminate if the worst case violation probability is within a tolerance level $\eta \in \mathbb{R}^+$ of the desired ϵ_v, i.e, $|\hat{E}_v^{\max}(\mathbf{X}^{(k+1)}) - \epsilon_v| \leqslant \eta$, or if the upper and lower bounds have converged within tolerance $\eta_s \in \mathbb{R}^+$, i.e. $|s_{\max}^{(k)} - s_{\min}^{(k)}| \leqslant \eta_s$.

Return the solution with the lowest objective that is empirically feasible to the chance constraint, i.e.,

$$\hat{E}_v^{\max}(\mathbf{X}^{(k+1)}) \leqslant \epsilon_v. \tag{6.43}$$

Otherwise, increase iteration count $k = k + 1$, and go back to step (ii).

The bisection search is guaranteed to converge within $\lfloor \log_2(s_{\max}^{(0)} - s_{\min}^{(0)})/\eta_s \rfloor$ iterations.

6.4 CASE STUDY OVERVIEW

In the case studies, we perform simulations on the IEEE 13-node radial distribution feeder [Sch+17] using realistic uncertainty data for solar PV generation and load demand to assess the performance of the tuning algorithm. Our goal is to find reactive power set-points for the inverters that are valid for the whole day, i.e., we assume that the distribution of possible load and solar PV values is representative for power injection

profiles across the whole day. In this section, we describe the test system and the data sets used to obtain the results in the subsequent sections.

The optimization problem and algorithms are implemented in Julia [Bez+17] using JuMP [DHL17] with the solver Ipopt [WB06].

6.4.1 *Test system*

As the test system, we use the modified IEEE 13-bus feeder [Sch+17], which consists of 15 houses represented as single-phase connections at seven nodes as shown in Figure 6.3. We assume each house has a single-phase rooftop solar PV system. The inverter ratings of the houses are set to the maximum observed solar PV value of the corresponding house.

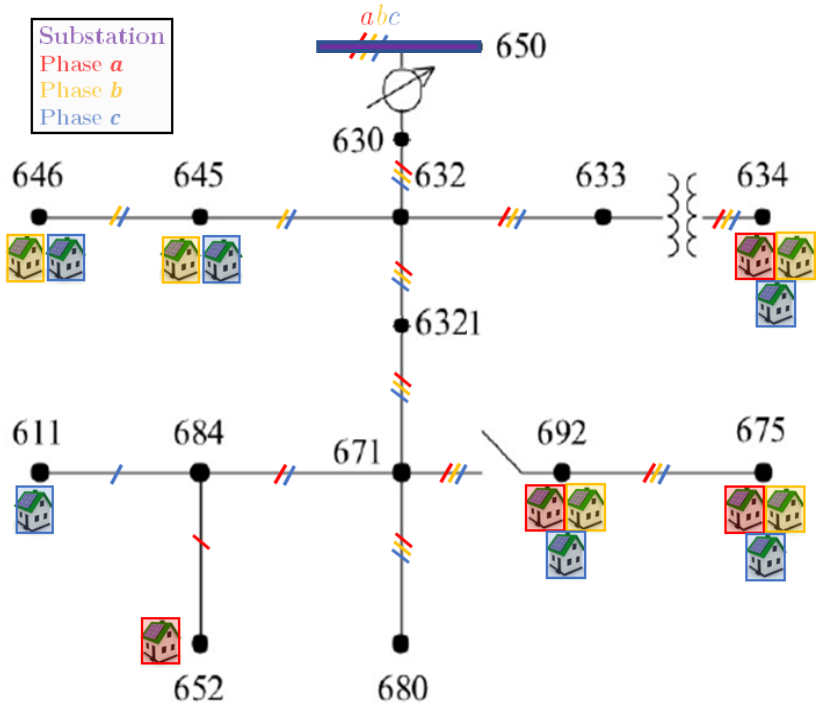


Figure 6.3: The IEEE 13-bus radial distribution test feeder, modified to include 15 residential houses.

6.4.2 *Uncertainty data*

We use data from Pecan Street's Dataport [Pec], which contains 6 months (184 days) of data collected from residential homes in New York between May 2019 and October 2019. The data comprises of 1-minute resolution measurements of rooftop solar PV generation and profiles for 30 unique loads, including, e.g., electric vehicle charging, HVAC, and appliance use.

The data set consists of 25 houses total, with 14 houses with non-zero solar PV measurements. We assign the solar PV and load profiles of these 14 houses along with an additional house with no solar PV measurements (but with load measurements) to the houses in our test system. In order for our data to match the existing loads in test feeder used, we scale both the solar PV and load by a factor of 20. We take the maximum solar PV output for each house as the maximum inverter rating.

We randomly split the available 184 days into two sets, each with 92 days, one set used for the tuning algorithm (i.e., as in-sample data) and one set used for an out-of-sample evaluation (if the simple tuning algorithm detailed in Chapter 3 is used) or for the a posteriori verification step (if the two-step algorithm detailed in Chapter 4 is used). We denote these two full sets of samples using $\Xi^{\text{full}}_{\text{tune}}$ and $\Xi^{\text{full}}_{\text{eval}}$.

In-sample data

To obtain the data sets used in-sample for the tuning algorithm, we sample from $\Xi^{\text{full}}_{\text{tune}}$, which contains 92 full days of data (corresponding to $132,480$ total samples) using two types of sampling methods:

(i) *Random samples*: We randomly draw N samples from the entire pool of $132,480$ total samples. This data set assumes that the data from the full data set represents the probability distribution of the data and draws i.i.d. samples from this set. This does

not guarantee that all time steps in the day are represented, but is likely to represent a wider variety of operating conditions as it includes data from more days.

(ii) *Full day samples*: We randomly choose N_{days} days from the full set of 92 days, which corresponds to $N = N_{\text{days}} \cdot 1440$ samples. This data set corresponds to the common practice of using data from a set of "representative days". In this case, there may be strong correlations among subsequent time steps, but all time steps will be represented in the data used in the algorithms.

Out-of-sample data

We form data sets for out-of-sample evaluations using the other 92 day set, $\Xi^{\text{full}}_{\text{eval}}$. For the out-of-sample evaluation of the simple tuning algorithm, we similarly draw sets using random sample or full day samples. For the a posteriori verification step of the two-step algorithm, we create the data set using random sampling since the set must consist of i.i.d. samples.

6.4.3 *Investigations*

To examine the performance of the tuning method applied to the distribution grid setting, we perform a variety of case studies, summarized as follows:

(i) We investigate how *different data sampling procedures* can impact the performance of the tuning method. We analyze the performance of the tuning algorithm with the full day and random samples across several different algorithm replications, and aim to assess which sampling procedure is most suitable.

(ii) We provide a *detailed results* for one algorithm replication, looking at the resulting constraint tightenings, inverter and voltage set points, and constraint violations evaluated on in- and out-of-sample data.

(iii) We investigate the algorithm performance when only *partial day* samples are used. Since solar PV generation is highly temporally dependent, rather than finding a single inverter set-point for the entire day, it may be advantageous to determine set-points for smaller periods of the day.

			Voltage magnitude			Reactive power			
		Total VUF	Max upper	Max lower	Joint	Max upper	Max lower	Joint	Total joint
Random, N = 2880	In-sample	6.432	0.049	0.032	0.132	0.05	0.05	0.121	0.195
	Out-of-sample	6.355	0.056	0.04	0.154	0.072	0.056	0.134	0.217
Random, N = 5760	In-sample	6.470	0.05	0.032	0.131	0.05	0.05	0.123	0.195
	Out-of-sample	6.401	0.054	0.038	0.15	0.07	0.055	0.134	0.213
Random, N = 11520	In-sample	6.452	0.05	0.031	0.13	0.05	0.05	0.123	0.193
	Out-of-sample	6.386	0.054	0.039	0.15	0.07	0.055	0.134	0.214
Full day, $N_{day} = 2$	In-sample	6.562	0.044	0.032	0.105	0.05	0.05	0.104	0.162
	Out-of-sample	6.552	0.054	0.036	0.135	0.084	0.048	0.134	0.208
Full day, $N_{day} = 4$	In-sample	6.327	0.05	0.035	0.127	0.05	0.05	0.124	0.198
	Out-of-sample	6.250	0.061	0.043	0.159	0.088	0.056	0.145	0.230
Full day, $N_{day} = 8$	In-sample	6.242	0.048	0.034	0.129	0.05	0.05	0.125	0.195
	Out-of-sample	6.243	0.059	0.045	0.169	0.084	0.072	0.150	0.237

Table 6.1: The in- and out-of-sample results from applying the tuning method to the three-phase single chance-constrained AC OPF problem. The table reports the total VUF calculated on the samples and the worst-case, the maximum (i.e., worst case) and joint violation probabilities across the voltage magnitude constraints, the maximum and joint violation probabilities across the reactive power constraints, and the overall joint violation probability. Highlighted in blue are the constraints in which the worst-case violation probabilities meets the desired $\epsilon = 0.05$ exactly. Highlighted in red are the out-of-sample violation probabilities that exceed the desired $\epsilon = 0.05$.

6.5 CASE STUDY: DATA SAMPLING

We perform numerical experiments to assess the performance of the tuning method using random samples and full day samples. For both sampling methods, we choose data set sizes of $N = \{2880, 5760, 11520\}$, corresponding to $N_{day} = \{2, 4, 8\}$ days. We perform 10 algorithm replications, each using an independent sample draw from the overall full data set. We set the individual chance constraint violation probability to be $\epsilon = 0.05$ for all constraints. The resulting solutions are evaluated using the same out-of-sample data set consisting of $N_{days} = 45$ full days.

Table 6.1 presents the worst case (maximum) in- and out-of-sample violation probabilities for each constraint type (i.e., voltage magnitude upper and lower constraints and reactive power upper and lower constraints), the joint violation probabilities across the voltage constraints, the reactive power constraints, and all constraints, and total VUF for solutions obtained using tuning on the six sample sets. We report the VUF value calculated using the actual uncertainty samples, not the VUF calculated in the objective function of the reformulated problem (see Section 6.3.1). The values in the table are the average across the 10 replications. We highlight in blue the cases where the worst-case in-sample violation probability of the constraint exactly meets the desired $\epsilon = 0.05$. We highlight in red the cases where there worst case out-of-sample violation probability of the constraint exceeds the desired the desired $\epsilon = 0.05$.

We can see that for all six sample sets, the tuning algorithm is able to ensure that all individual chance constraints have an in-sample empirical violation probability of less than or equal to $\epsilon = 0.05$. However, the out-of-sample evaluated violation probabilities often exceed the desired $\epsilon = 0.05$. In particular, the some of the upper reactive power limits and lower voltage magnitude limits exceed violation probabilities of $\epsilon = 0.05$. Tuning using full day data tend to give higher out-of-sample violation probabilities. To

ensure that the out-of-sample values are below the desired ϵ threshold, we can apply the two-step tuning method. This is explored in Chapter 7.

An interesting observation is that the total VUF of solutions obtained from using randomly sampled data is relatively consistent amongst the three sample sizes. However, the total VUF of solutions obtained from using full days of samples *decreases* as the number of days (and samples) increase. Moreover, when at least $N_{day} = 4$ days of samples are used, the total VUF of the solutions is lower than that of solutions obtained from randomly sampled data.

A final observation is that the sum of the individual violation probabilities is substantially larger than joint violation probability, indicating that there are correlations between the individual constraints. In Chapter 7, we also delve into solving the joint chance-constrained version of this problem.

6.6 CASE STUDY: SINGLE REPLICATION RESULTS

The subsequent sections discuss results of a single replication of the tuning algorithm, using a sample set consisting of $N = 2880$ random samples.

6.6.1 *Constraint tightenings*

Figure 6.4 illustrates the voltage magnitude set-points and constraints tightenings across all single-phase buses (left) and PV inverter reactive power set-points and constraint tightenings across all inverters (right). The nominal constraints are plotted with a dashed red line, the tightened constraints are plotted with blue lines, and the set points are plotted with yellow dots. The tightening is the difference between the red dashed and the blue lines.

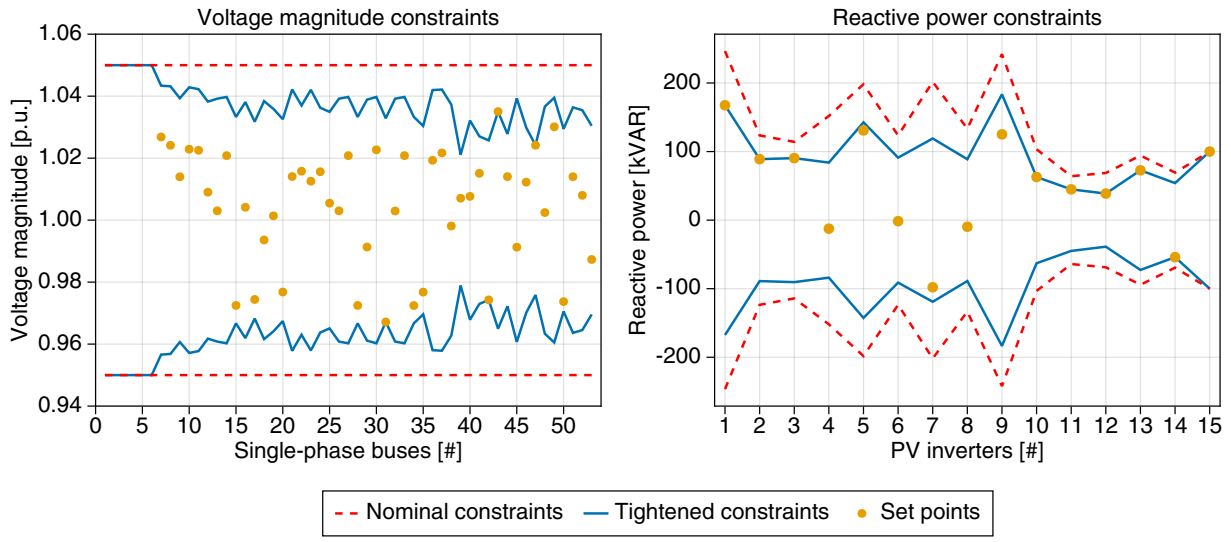
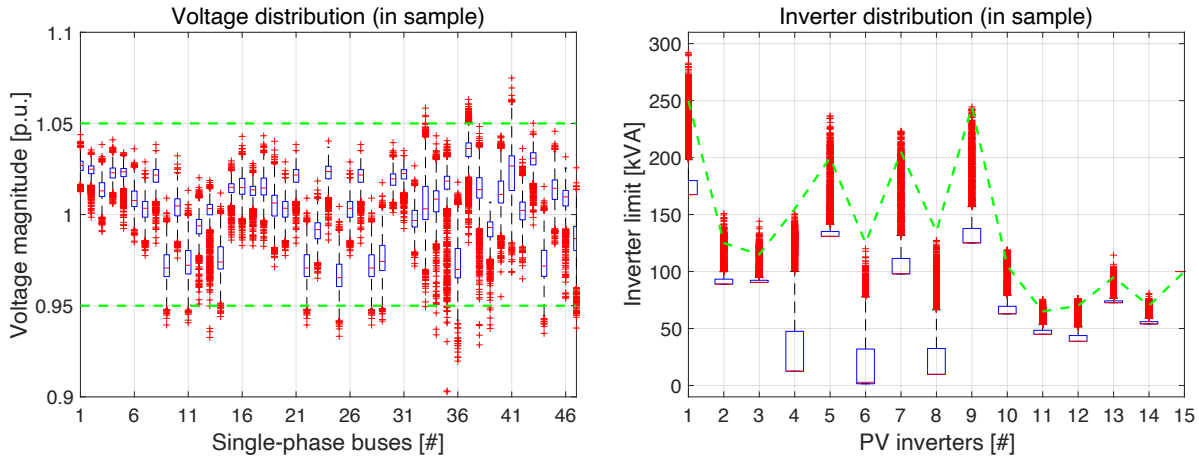


Figure 6.4: Voltage magnitude and reactive power nominal constraints (red dashed line), tightened constraints (blue solid line), and set points (yellow dots).

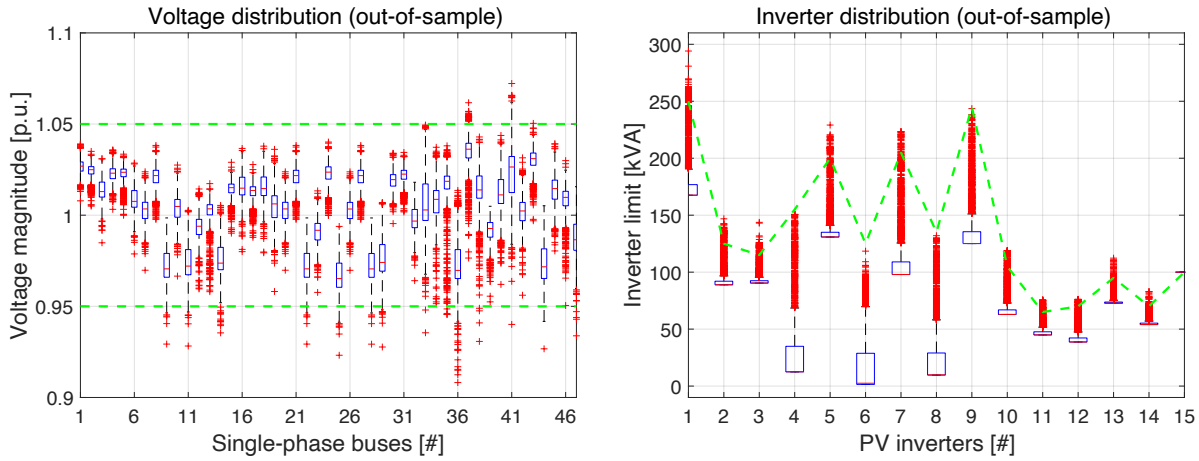
We first observe that the tightenings for the upper and lower voltage constraints are symmetric. This is as expected, because the voltage tightening of the tuning-based method is given by the product between the standard deviation of the voltage magnitude at each node (which is the same for the upper and lower bound) and the tuning parameter s (which is shared among all voltage constraints) and is thus the same for both the upper and lower bound. We further recognize that most constraints are not active (i.e., not binding). In particular, only 4 voltage constraints are active. With the inverter constraints, we see that a much larger fraction of the constraints are active, with many set points either at the maximum or minimum allowable reactive power injections.

6.6.2 *Constraint violations by node and inverter*

We next compare the performance of the methods in terms of constraint violations for the same replication discussed above. To assess the differences in performance, we compute the voltage magnitude and inverter apparent power for each in-sample and out-of-sample realization. Figures 6.5 show the box and whisker plots of the in sample



(a) Distribution calculated on in-sample data.



(b) Distribution calculated on an out-of-sample set consisting of $N_{day} = 2$ full days.

Figure 6.5: Box and whisker plots of the distributions of the voltage magnitude (left) and inverter apparent power (right) calculated by running power flow simulations on in-sample (top) and out-of-sample (bottom) data sets using the reactive power set point obtained from tuning. The green dashed lines represent the lower and upper voltage magnitude limits and the maximum inverter apparent power limits.

(Figure 6.5a) and out-of-sample (Figure 6.5b) voltage magnitude (left) and reactive power (right) distributions. The out-of-sample evaluation was performed on a data set consisting of $N_{day} = 2$ randomly drawn full days.

We see that both the in- and out-of-sample distributions look relatively comparable, indicating that the in-sample distribution includes a representative set of samples. The figures also display that there are more under-voltages than over-voltages. Moreover,

we can observe that compared to the in-sample voltage distribution, the out-of-sample distribution has outlier points that violate the lower voltage limit with greater magnitude.

6.6.3 *Constraint violations by time of day*

We analyze how the constraint violations vary across a single day. Figure 6.6 shows the number of single-phase buses with voltage magnitude constraint violations (top) and PV inverters with reactive power constraint violations (bottom) per time step evaluated on an out-of-sample day using a set point obtained from tuning on $N = 2880$ random samples. We see that most violations for both types of constraints occur around noon, which is when PV active generation is the highest (as illustrated at the beginning of the chapter in Figure 6.1). However, there are a few violations of the lower voltage constraints in the early morning and evening.

Figure 6.7 plots the maximum violation *magnitude* across all buses and inverters for each time step. The violation magnitude is shown as a percentage of the constraint. For the reactive power constraints, we note that these percentages appear very large. This is not necessarily because the violation magnitudes are excessively high, but because some reactive power constraints at certain parts of the day (particularly around noon) are very small (i.e., at or close to zero) due to the high amount of active power generation from the PV inverters. We see that, similarly to in Figure 6.6, violations mainly occur around noon, when solar PV generation is the highest. For the voltage magnitude constraints, we see that the lower limit violations are of a considerably smaller magnitude than those of the upper limits. For the reactive power constraints we see that, in general, the upper limits have higher violation magnitudes, with the exception of the lower limit violation around 10am.

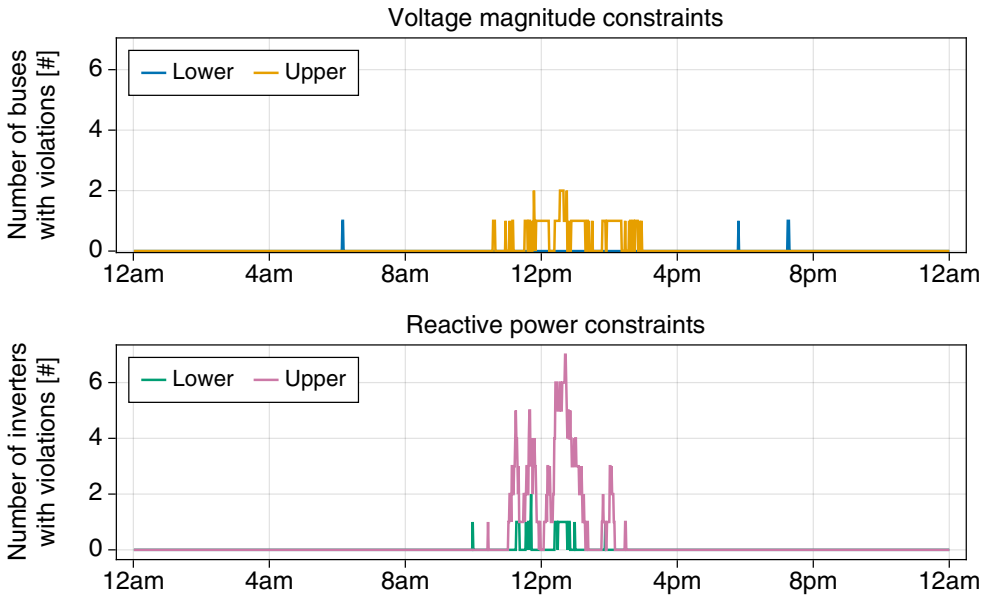


Figure 6.6: The number of single-phase buses with voltage magnitude constraint violations (top) and PV inverters with reactive power constraint violations (bottom) that occur in a single out-of-sample evaluation day.

6.7 CONCLUSION

In this chapter, we solve the single chance-constrained three-phase unbalanced AC OPF problem to determine solar PV inverter reactive power set points that minimize voltage unbalance. We adapt the bisection-based tuning algorithm from Chapter 3 to this distribution grid setting. A notable challenge in formulating an suitable approximate optimization model for the AC OPF setting is that the uncertainty enters into the constraints non-linearly and implicitly. We use a similar reformulation of the chance constraints as in Chapter 3, where the deterministic constraints consist of the nominal constraints tightened with uncertainty margins that are defined as the product of a tuning parameter and an estimate of the standard deviation of the nominal constraint. Unlike in the DC OPF case, there is not an explicit representation of the standard deviation, so it is estimated by running power flow simulations using the uncertainty samples. Similarly, the evaluation

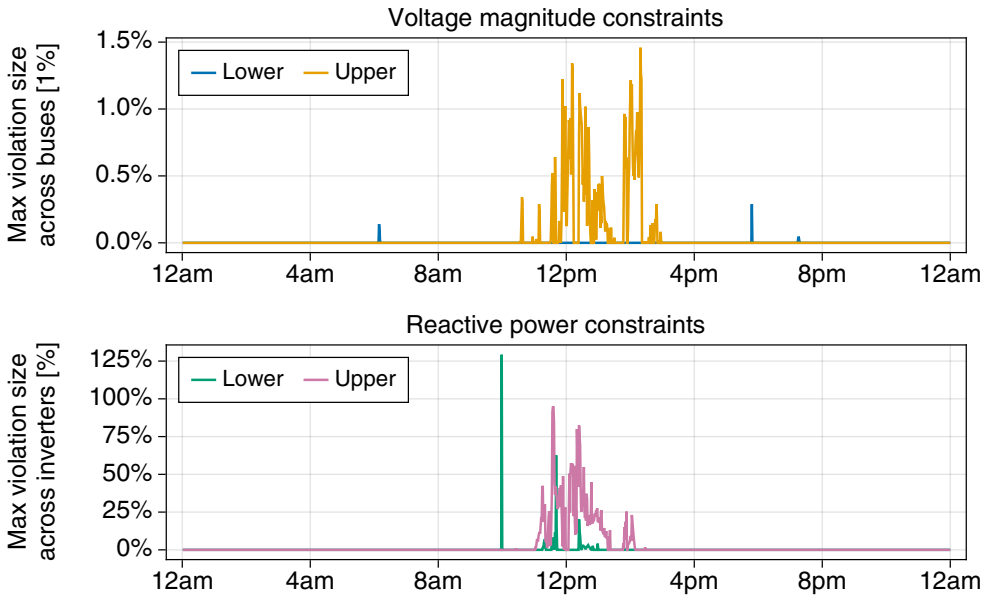


Figure 6.7: The maximum violation magnitude (as a percentage of the constraint limit) across all buses and inverters for each time step.

step of the tuning algorithm also uses power flow simulations to determine worst case empirical violation probability across constraints to use for the bisection search tuning.

Numerical simulations were performed on the IEEE 13-bus test feeding using real residential load and PV data. We first compared the performance of the algorithm using different methods for creating the in-sample data set used as input to the tuning algorithm (i.e., randomly sampled data points and randomly chosen full days of data) as well as different data set sizes. We see that the bisection search-based tuning method is able to obtain the desired violation probability ϵ in the in-sample evaluation. However, the empirical violation probabilities in the out-of-sample evaluation often slightly exceed the desired ϵ.

7 RISK ALLOCATION METHODS

In previous chapters, we have performed tuning by iteratively adjusting a tuning parameter via a bisection search. However, this single dimensional tuning approach fixes the tuning parameter to be uniform across all chance constraints, which may be insufficiently flexible in certain problem formulations, particularly when joint chance constraints are considered. In this chapter, we instead explore *multi-dimensional* tuning methods that separately adjust several tuning parameter. Our motivating example is the joint chance-constrained version of the three-phase unbalanced AC OPF problem for distribution grid control presented in Chapter 6.

We begin by reviewing a commonly used method to solve joint chance-constrained problems that decomposes the joint chance constraint into single chance constraints. The joint violation probability is "shared" equally among the single constraints, giving rise to uniform individual violation probabilities or *risk levels*. We subsequently discuss how this uniform risk allocation is typically insufficient for the OPF problems we deal with, as the lack of consideration of correlations between the constraints can result in overly conservative or infeasible solutions. We review existing literature on methods for determining non-uniform risk level allocations for joint chance-constrained problems and discuss their limitations. Finally, we propose a risk allocation-based tuning algorithm, inspired by [OW08] and [JHS20], that iteratively tunes a multi-dimensional vector representing the risk levels assigned to each individual constraint in order to find a solution that satisfies the desired joint violation probability. We conclude by presenting numerical results applying this method to the joint chance-constrained three-phase AC OPF problem.

The notation in this chapter (along with the Chapter 6) is self-contained, unless an equation from Chapters 2 through 5 is explicitly referenced.

7.1 BACKGROUND, MOTIVATION, AND CHALLENGES

Several solution techniques for single chance-constrained problems do not easily extend to joint constrained problems. As a result, a common solution method involves decomposing the joint chance constraint into single chance constraints and invoking conservative upper bounds on the single chance constraints to ensure that the joint violation probability does not exceed the desired limit. Consider the following joint chance constraint

$$\mathbb{P}_{\xi}\big(g_j(x,\xi) \leqslant 0, \quad \forall j \in \mathcal{C}\big) \geqslant 1-\epsilon, \tag{7.1}$$

where the constraint is required to hold with at least $1-\epsilon$ probability. The joint constraint can be decomposed into the following single chance constraints,

$$\Pr(g_j(x,\xi) \leqslant 0) \geqslant 1-\epsilon_j, \qquad \forall j \in \mathcal{C}, \tag{7.2}$$

where each ϵ_j represents the acceptable violation probability of the corresponding single chance constraints. Let $K = |\mathcal{C}|$ be the number of single chance constraints from the decomposition of a joint chance constraint. We refer to a chosen assignment of violation probabilities for single chance constraints $\epsilon_1, \ldots \epsilon_K$ as a *risk allocation* and denote it using ϵ_{RA} (to distinguish it from the desired joint violation probability ϵ).

A popular method to determine a suitable risk allocation is to invoke the union bound (also known as Boole's inequality [Boo54] or the Bonferroni inequalities [Car36]). The intuition behind this method is as follows. The joint chance constraint (7.1) can

equivalently be written as ensuring the violation probability is less than the desired violation ϵ, i.e.,

$$\mathbb{P}_\xi\big(g_j(x,\xi) > 0, \;\; j \in \mathcal{C}\big) < \epsilon. \tag{7.3}$$

Applying the inclusion-exclusion principle, the left hand side of (7.3), which represents the joint violation probability, can be written

$$\mathbb{P}_\xi\left(\bigcup_{j\in\mathcal{C}}\{g_j(x,\xi) > 0\}\right) = \sum_{j\in\mathcal{C}}\mathbb{P}_\xi\big(g_j(x,\xi) > 0\big) \tag{7.4a}$$

$$- \sum_{j<l}\mathbb{P}_\xi\big(\{g_j(x,\xi) > 0\}\cap\{g_l(x,\xi) > 0\}\big) \tag{7.4b}$$

$$+ \ldots + (-1)^{K-1}\mathbb{P}_\xi\left(\bigcap_{j\in\mathcal{C}}\{g_j(x,\xi) > 0\}\right) \tag{7.4c}$$

The first term denotes the sum of the violation probabilities across individual constraints, while the subsequent terms account for the cases where multiple individual constraints are jointly violated. Let $P_{\text{overlap}} \in [0,1]$ denote the overall aggregate of these probabilities,

$$P_{\text{overlap}} := \sum_{j\in\mathcal{C}}\mathbb{P}_\xi\big(g_j(x,\xi) > 0\big) - \mathbb{P}_\xi\left(\bigcup_{j\in\mathcal{C}}\{g_j(x,\xi) > 0\}\right). \tag{7.5}$$

Rather than considering the entire expression (7.4), the union bound truncates (7.4) to obtain the following inequality

$$\mathbb{P}_\xi\left(\bigcup_{j\in\mathcal{C}}\{g_j(x,\xi) > 0\}\right) \leqslant \sum_{j\in\mathcal{C}}\mathbb{P}_\xi\big(g_j(x,\xi) > 0\big). \tag{7.6}$$

By choosing a risk allocation ϵ_{RA} that satisfies

$$\epsilon_1 + \cdots + \epsilon_K \leqslant \epsilon, \tag{7.7}$$

and solving the corresponding single chance-constrained problem (7.2), the joint chance constraint (7.1) is guaranteed to hold with probability at least $1 - \epsilon$. We can see that the union bound removes the P_{overlap} term, which can make the application of the bound conservative. In particular, it has been shown that even by using an optimally chosen risk allocation (i.e., where the individual violation probabilities are assigned to constraints in the most optimal way), the union bound still leads to conservative solutions [Che+10]. In essence, P_{overlap} can be interpreted as a metric for the conservativeness of using a method that obtains solutions via performing an upper bound such as (7.6).

Often, the risk allocation ϵ_{RA} is not optimized but rather the violation probabilities for the single chance constraints are chosen uniformly [Bla06], i.e.,

$$\epsilon_j = \epsilon / K, \quad \forall j \in \mathcal{C}. \tag{7.8}$$

However, this uniform risk allocation results in further conservativeness because it *does not account for potential correlations between the individual constraints*. It was shown in [XAJ19] that, in general, optimizing the Bonferroni approximation is NP-hard, but there are sufficient conditions leading to tractability. Several other works have also attempted to optimize the risk allocation using techniques such as ellipsoidal relaxations [Van04] and particle control [Bla06], but have nevertheless yielded conservative results.

In the power systems applications studied in this thesis, it is particularly important to account for these correlations, as some constraints are highly correlated and only a small subset of constraints tend to be close to their limit. Thus, the number of constraints experiencing any violations is typically small. For example, in the DC OPF problem (3.3.2) studied in Chapters 3 and 4, we noted that the generator outputs were correlated (i.e., all generators adjust their output together). In the AC OPF problem (6.30) studied in the previous chapter, we take the solar PV generation at residential houses as a source of uncertainty. In this setting, we expect there to be some degree of spatial correlation

between the solar PV generation at each house if they are geographically located close together. As a result, if, for example, there is a lot of active power generation from the solar PV system at one house causing the upper voltage magnitude limit to be violated, then we expect the same to happen at nearby houses. In the time series plots from the previous chapter (Figure 6.6), we can see that constraints, particularly reactive power constraints, are often jointly violated. Consequently, for our applications, it is critical to seek alternative risk allocation methods that take into account correlations between constraints and constraint violations.

We can draw a link between using uniform risk allocations to our tuning method using a single tuning parameter. If we consider, e.g., the approximate problem model formulated based on the moment-based analytical reformulation (3.3), we see that the the tuning parameter s takes the place of a safety parameter defined by $F^{-1}(1-\epsilon)$, where ϵ is the acceptable violation probability of the single chance constraint and the inverse cumulative distribution of the uncertainty $F^{-1}(\cdot)$. Using the same tuning parameter s across all individual constraints of a problem can be essentially thought of as choosing a uniform risk allocation across constraints. In other words, the process of choosing a non-uniform risk allocation can be thought of as a form of *multi-dimensional tuning*. In general, multi-dimensional tuning is challenging due to the large number of constraints present in OPF problems (particularly in the three-phase unbalance AC formulation considered for distribution grids). The tuning problem must attempt to search for an optimal risk allocation in high dimensional space.

In this chapter, we propose an multi-dimensional tuning scheme based on the idea of risk allocation for the *joint* chance-constrained version of the three-phase unbalanced AC OPF problem considered in Chapter 6. In the case study, we compare the risk allocation-based tuning method to the bisection search-based algorithm.

7.2 LITERATURE REVIEW

There are numerous works in the literature that focus on reducing conservativeness when solving joint chance-constrained problems via decomposition into single chance constraints. The problem of risk allocation was first introduced in [OW08], which proposed a two-stage optimization problem where the upper stage problem finds an optimal risk allocation and the lower stage problem uses the risk allocation to solve the JCCP reformulated as individial constraints. To solve the upper stage problem, the paper proposed the *iterative risk allocation* (IRA) algorithm, where individual risk levels are allocated based on tightening inactive constraints (which are less likely to be violated) and relaxing active constraints (which are more likely to be violated). While this method is often less conservative than using a uniform risk allocation, it still maintains satisfaction of Boole's inequality, which guarantees satisfaction of the joint chance constraint but can lead to conservatism.

Several works have adapted the two-stage risk allocation framework. Notably, [VT11] uses a bisection search is used to scale a uniform risk allocation ϵ/K such that the resulting optimization problem admits a solution with an empirical joint violation probability close to the desired violation ϵ. While the risk allocation no longer needs to satisfy Boole's inequality, the use of a uniform risk across all individual constraints may lead to a conservative solution by not taking into account that certain constraints are more or less likely to be violated or that allowing higher violation probabilities for some constraints can reduce cost.

Another line of work has attempted to derive tighter versions of Boole's inequality [BT17] or use machine learning methods to improve upon Boole's inequality [BB18; BB19]. The latter two methods use sample-based simulations to estimate the violation probability of different combinations of individual constraints, and incorporates these probabilities into the bound (7.6) to achieve a tighter bound. A similar approach is

taken in [JHS20], but instead allows for non-uniform risk allocations, allowing it to achieve less conservative solutions than uniform allocations. Here, the proposed adaptive risk allocation strategy assigns risk levels to individual constraint proportional to their empirical violation probability.

The scenario approach can also be used to directly solve non-convex joint chance-constrained problems. While there exists an extensive body of work on theoretical guarantees and sample complexity bounds for convex problems, there do not exist analogous results broadly applicable to general non-convex problems [CGR18]. In particular, to our knowledge, there does not exist any a priori theoretical results for the scenario approach applicable for the three-phase unbalanced AC OPF problem considered in this chapter.

7.3 RISK ALLOCATION TUNING ALGORITHM

Inspired by [OW08] and [JHS20], we propose an algorithm for risk allocation for the joint chance-constrained version of the three-phase AC OPF problem from Chapter 6. We use the a similar deterministic reformation of the single chance-constrained problem as presented in Section 6.35, but with slightly different constraint tightening definitions. Here, the tuning parameter becomes the vector of risk allocations. The resulting candidate solution is evaluated, then the constraints are updated using a new tuning scheme based on risk allocation.

In the following, we present the joint chance-constrained problem formulation and provide details on the three components: (i) the formulation of the approximate optimization model, (ii) the sample-based solution evaluation step, and (iii) the tuning scheme.

7.3.1 *Problem formulation*

We use the same uncertainty modeling, reactive power control, and distribution grid model as in Chapter 6. The three-phase AC joint chance-constrained OPF problem is formulated as follows

$$\min_{\mathbf{P}_{G,0},\mathbf{Q}_G,|\mathbf{V}|,\mathbf{\Theta}} \sum_{\omega\in\Omega}\sum_{l\in\mathcal{N}} \mathrm{VUF}_{l,\omega}^2 \tag{7.9a}$$

$$\text{s.t.} \quad f\left(|\mathbf{V}_{i,\omega}|, \mathbf{\Theta}_{i,\omega}, P_{i,\omega}, Q_{i,\omega}\right) = 0, \qquad \forall i \in \{0,\mathcal{N}\},\ \omega\in\Omega, \tag{7.9b}$$

$$\mathbb{P}_\omega \left(\begin{array}{ll} |\mathbf{v}_{i,\omega}^\phi| \leqslant \overline{v}, & \forall \phi\in\Phi, i\in\mathcal{N} \\ |\mathbf{v}_{i,\omega}^\phi| \geqslant \underline{v}, & \forall \phi\in\Phi, i\in\mathcal{N} \\ \mathbf{q}_{G,i}^\phi \leqslant \sqrt{|s_{G,i}^\phi| - (p_{G,i,\omega}^\phi)^2}, & \forall \phi\in\Phi, i\in\mathcal{N} \\ \mathbf{q}_{G,i}^\phi \geqslant -\sqrt{|s_{G,i}^\phi| - (p_{G,i,\omega}^\phi)^2}, & \forall \phi\in\Phi, i\in\mathcal{N} \end{array} \right) \geqslant 1-\epsilon, \tag{7.9c}$$

$$|\mathbf{V}_0|\angle\mathbf{\Theta}_0 = \begin{bmatrix} 1\angle 0^\circ & 1\angle -120^\circ & 1\angle 120^\circ \end{bmatrix}^\top, \tag{7.9d}$$

where $\mathbf{X} = (\mathbf{P}_{G,0}, \mathbf{Q}_G, |\mathbf{V}|, \mathbf{\Theta})$ denotes the decision variables.

7.3.2 *Approximate problem*

We use the same approximate problem as formulated in (6.35), where the power flow equations are enforced for the average power injections and the chance constraints are reformulated as their deterministic counterparts, tightened by a tuneable parameter. However, unlike in Chapter 6, we do not pre-calculate the inverter tightenings and do not define the voltage magnitude tightenings as the product of parameter s and the estimated standard deviation. Instead, the tightenings for both types of constraints will be iteratively determined via the risk allocation throughout the tuning process.

Constraint tightening calculation

Let $\underline{\epsilon}^{\phi}_{q,i}$, $\overline{\epsilon}^{\phi}_{q,i}$, $\underline{\epsilon}^{\phi}_{v,i}$, and $\overline{\epsilon}^{\phi}_{v,i}$ denote the risk allocations for the lower reactive power, upper reactive power, lower voltage magnitude, and upper voltage magnitude constraints at bus i, connected to phase ϕ. For notational convenience, we use $\epsilon_{\text{RA}} = (\underline{\epsilon}^{\phi}_{q,i}, \overline{\epsilon}^{\phi}_{q,i}, \underline{\epsilon}^{\phi}_{v,i}, \overline{\epsilon}^{\phi}_{v,i})$ to denote all the risk allocations for the problem.

For the inverter constraints, given a risk allocation ϵ_{RA}, we determine the corresponding constraint tightening using the same process as in Section 6.3.1, summarized here for ease of the reader. We use the uncertainty samples $\omega \in \Omega$ to obtain the empirical distributions for the upper and lower reactive power limits, denoted $\hat{F}^{\phi}_{q,i}(\cdot)$ and $-\hat{F}^{\phi}_{q,i}(\cdot)$. Then, given a risk allocation ϵ_{RA}, we take the corresponding quantile of the empirical distribution of the constraint and calculate the constraint tightening to be the difference between the nominal limit and the evaluated quantile, i.e.,

$$\overline{\lambda}^{\phi}_{q,i} = \overline{q}^{\phi}_{G,i} - \hat{F}^{\phi}_{q,i}(\overline{\epsilon}^{\phi}_{q,i}), \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{7.10a}$$

$$\underline{\lambda}^{\phi}_{q,i} = -\hat{F}^{\phi}_{q,i}(1 - \underline{\epsilon}^{\phi}_{q,i}) - \underline{q}^{\phi}_{G,i}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{7.10b}$$

For the voltage magnitude constraints, a similar process is performed. The main difference is that to calculate the empirical distributions, we need to obtain a set point from the optimization problem. We use the inverter set point $\mathbf{Q_G}$ from the previous iteration of the optimization problem and uncertainty samples $\omega \in \Omega$ to run power flow simulations to obtain an empirical distribution for the voltage magnitude at every single-phase bus, denoted $\hat{F}^{\phi}_{v,i}(\cdot)$. We then calculate the tightenings as

$$\underline{\lambda}^{\phi}_{v,i} = |\mathbf{v}^{\phi}_i| - \hat{F}^{\phi}_{v,i}(\underline{\epsilon}^{\phi}_{v,i}), \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{7.11a}$$

$$\overline{\lambda}^{\phi}_{v,i} = \hat{F}^{\phi}_{v,i}(1 - \overline{\epsilon}^{\phi}_{v,i}) - |\mathbf{v}^{\phi}_i|, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{7.11b}$$

7.3.3 *Solution evaluation*

We use the same sample-based evaluation as in Section 6.3.2, but instead evaluate the joint violation probability of the candidate solution $\hat{E}_{\text{joint}}(\mathbf{X})$. In particular, we define the indicator random variable $Y(\mathbf{X}, P_{G,\omega}, P_{L,\omega})$ to evaluate to 0 if all inner constraints of the joint constraint (7.9c) holds and 1 otherwise. The joint empirical violation probability is defined as

$$\hat{E}_{\text{joint}}(\mathbf{X}) := \frac{1}{\Omega} \sum_{\omega \in \Omega} Y(\mathbf{X}, P_{G,\omega}, P_{L,\omega}). \tag{7.12}$$

7.3.4 *Tuning scheme*

The most significant difference between this risk allocation tuning algorithm and the previously proposed algorithm in Section 6.3 lies in the tuning scheme. Instead of tuning a parameter s, we directly update the risk allocations $\epsilon_{\text{RA}} = (\underline{\epsilon}^{\phi}_{q,i}, \overline{\epsilon}^{\phi}_{q,i}, \underline{\epsilon}^{\phi}_{v,i}, \overline{\epsilon}^{\phi}_{v,i})$ in each iteration, inspired by the iterative risk allocation algorithm [OW08].

We incorporate two main ideas to guide the algorithm tuning heuristics: (i) We use the observations from the case studies in Chapter 6 that only a fraction of constraints are active and/or violated and only modify these constraints. (ii) We use the general idea from [JHS20] of allocating the allowable risk (or violation probability) for each constraint according to where it is needed the most in order to reduce the overall cost of the problem. In particular, we incorporate violation probability information as well as information from the dual variables from each constraint to determine which constraints have the most impact in minimizing the objective. We detail the tuning process and intuition behind each step in the following.

Relaxing constraints

When the candidate solution is evaluated to be too conservative (i.e., $\hat{E}(\mathbf{X})_{\text{joint}} \leqslant \epsilon$, we seek to relax the constraints by *increasing the risk allocation* which corresponds to decreasing the constraint tightenings. In general, we note that the goal here is more to decrease the objective value (i.e., total VUF) to obtain a more optimal solution rather simply reaching the targeted violation probability ϵ. As long as the violation probability is below the ϵ threshold, the primary goal is obtaining as low of an objective as possible.

To this end, we observe that increasing the risk allocation ϵ_{RA} (and thus relaxing the constraint) on active constraints with non-zero dual variables is likely to reduce the objective, while keeping keeping the joint violation probability within bounds. This is because the dual variables associated with each constraint serve as an indicator as to which constraints have the most impact on the objective value.

Let $\underline{\nu}^{\phi}_{q,i}$, $\overline{\nu}^{\phi}_{q,i}$, $\underline{\nu}^{\phi}_{v,i}$, and $\overline{\nu}^{\phi}_{v,i}$ denote the dual variables associated with the lower reactive power, upper reactive power, lower voltage magnitude, and upper voltage magnitude constraints at bus i, connected to phase ϕ, respectively. Let $\mathcal{A} \subseteq \mathcal{C}$ denote the set of active constraints.

We let total amount of risk chosen to be allocated amongst all constraints be the difference between the desired and empirical joint violation probabilities, i.e., $\epsilon - \hat{E}_{\text{joint}}(\mathbf{X})$, which would be positive in this case. The update rule for active constraint $j \in \mathcal{A}$ is as follows

$$\epsilon^{(k+1)}_{\text{RA},j} = \epsilon^{(k)}_{\text{RA},j} + \left(\epsilon - \hat{E}_{\text{joint}}(\mathbf{X}^{(k+1)})\right) \cdot \frac{\nu^{(k+1)}_j}{\sum_{l \in \mathcal{A}} \nu^{(k+1)}_l} \cdot \tau, \tag{7.13}$$

where $\tau \geqslant 0$ is a constant parameter. To increase the step size (and rate of convergence), we can use a larger τ value.

Tightening constraints

On the other hand, when the candidate solution is not conservative enough and has too high of a violation probability (i.e., $\hat{E}_{joint}(\mathbf{X}) \geqslant \epsilon$), we seek to tighten the constraints by *decreasing the risk allocation* and subsequently increasing the constraint tightenings. We similarly allocate $\epsilon - \hat{E}_{joint}(\mathbf{X})$, which is negative in this case, amongst all the constraints. Since we observe that there are violated constraints that are in-active, we instead choose to tighten the *violated constraints* proportionally to the size of their violation.Let $\mathcal{V} \subseteq \mathcal{C}$ denote the set of violated constraints.

The update rule for violated constraint $j \in \mathcal{C}$ is as follows

$$\epsilon_{RA,j}^{(k+1)} = \epsilon_{RA,j}^{(k)} + \left(\epsilon - \hat{E}_{joint}(\mathbf{X}^{(k+1)})\right) \cdot \frac{\hat{E}_j(\mathbf{X}^{(k+1)})}{\sum_{l \in \mathcal{V}} \hat{E}_l(\mathbf{X}^{(k+1)})} \cdot \tau. \tag{7.14}$$

Infeasiblility

Finally, in the case that a risk allocation causes the optimization problem to become infeasible, we first determine which constraints are infeasible by running a power flow using the resulting reactive power set point from the solver. Then, we relax the infeasible constraints by increasing their corresponding risk allocation by a multiple of $\zeta > 1$, i.e., for infeasible constraint $j \in \mathcal{C}$, the update rule is

$$\epsilon_{RA,j}^{(k+1)} = \epsilon_{RA,j}^{(k)} \cdot \zeta. \tag{7.15}$$

Initialization

We initialize the risk allocation $\epsilon_{RA}^{(0)}$ to a uniform risk allocation that admits a feasible solution to the optimization problem. One choice could be to use the uniform allocation that satisfies the Bonferroni inequality, i.e. (7.8). However, in our problem setting, due to the large number of constraints, the risk allocated to each constraint becomes very small

and the resulting optimization problem is feasible. Instead, we choose a larger uniform risk allocation. We also note it is fine to set the initial risk allocations of each constraint to the desired joint violation ϵ, but the algorithm may require more iterations to converge to the desired risk allocation.

Convergence

The algorithm converges when the empirical joint violation probability is within a tolerance level $\eta \in \mathbb{R}^+$ of the desired joint violation probability ϵ, i.e.,

$$\hat{E}_{\text{joint}}(\mathbf{X}) \leqslant \epsilon. \tag{7.16}$$

We note that it is possible for solutions that have similar empirical joint violation probabilities to have different risk allocations and different costs. Therefore, we choose to return the solution with the lowest objective value that has an empirical joint violation probability less than the desired ϵ. We note that due to the non-linearity in the system, the solution with the lowest empirical violation probability may not be the one with the lowest objective value. If the tuning algorithm is unable to find a solution that has a low enough empirical joint violation probability, we return the solution with the lowest empirical joint violation.

7.3.5 *Algorithm*

The algorithm proceeds as follows:

(i) *Initialize*: Set the iteration count to $k = 0$ and initialize a uniform risk allocation, $\epsilon_{\text{RA}}^{(0)} = (\underline{\epsilon}_{q,i}^{\phi}, \overline{\epsilon}_{q,i}^{\phi}, \underline{\epsilon}_{v,i}^{\phi}, \overline{\epsilon}_{v,i}^{\phi})$.

Calculate the initial reactive power constraint tightenings $\overline{\Lambda}_q^{(0)}, \underline{\Lambda}_q^{(0)}$ according to (7.10) using the initial risk allocation. The approximate problem (6.35) is solved

using $\overline{\Lambda}_q^{(0)}, \underline{\Lambda}_q^{(0)}$ and voltage magnitude tightenings all set to zero to obtain an initial operating point $\mathbf{X}^{(0)}$, which is used to calculate the initial voltage magnitude tightening $\overline{\Lambda}_v^{(0)}, \underline{\Lambda}_v^{(0)}$ according to (7.11).

(ii) *Solve approximate problem*: Solve the approximate problem (6.35) using the tightenings $\overline{\Lambda}_q^{(k)}, \underline{\Lambda}_q^{(k)}, \overline{\Lambda}_v^{(k)}, \underline{\Lambda}_v^{(k)}$ to obtain a candidate operating point $\mathbf{X}^{(k+1)}$ and dual variables associated with each constraint $\underline{\nu}_{q,i}^{\phi}$, $\overline{\nu}_{q,i}^{\phi}$, $\underline{\nu}_{v,i}^{\phi}$, $\overline{\nu}_{v,i}^{\phi}$.

If the approximate problem is infeasible, relax the infeasible constraints by setting

$$\epsilon_{RA,j}^{(k+1)} = \epsilon_{RA,j}^{(k)} \cdot \zeta. \tag{7.17}$$

(iii) *Solution evaluation*: Calculate the joint violation probability of the candidate solution $\hat{E}_{joint}(\mathbf{X}^{(k+1)})$ as described in Section 7.3.3.

(iv) *Update risk allocation*:

If $\hat{E}_{joint}(\mathbf{X}^{(k+1)}) < \epsilon$: Relax *active* constraints $j \in \mathcal{A}$ proportional to their dual variables

$$\epsilon_{RA,j}^{(k+1)} = \epsilon_{RA,j}^{(k)} + \left(\epsilon - \hat{E}_{joint}(\mathbf{X}^{(k+1)})\right) \cdot \frac{\nu_j^{(k+1)}}{\sum_{l \in \mathcal{A}} \nu_l^{(k+1)}} \cdot \tau, \tag{7.18}$$

If $\hat{E}_{joint}(\mathbf{X}^{(k+1)}) > \epsilon$: Tighten *violated* constraints $j \in \mathcal{V}$ proportional to their violation

$$\epsilon_{RA,j}^{(k+1)} = \epsilon_{RA,j}^{(k)} + \left(\epsilon - \hat{E}_{joint}(\mathbf{X}^{(k+1)})\right) \cdot \frac{\hat{E}_j(\mathbf{X}^{(k+1)})}{\sum_{l \in \mathcal{V}} \hat{E}_l(\mathbf{X}^{(k+1)})} \cdot \tau. \tag{7.19}$$

(v) *Update tightenings*: Calculate the new constraint tightening $\overline{\Lambda}_q^{(k+1)}, \underline{\Lambda}_q^{(k+1)}, \overline{\Lambda}_v^{(k+1)}, \underline{\Lambda}_v^{(k+1)}$ using the updated risk allocation $\epsilon_{RA}^{(k+1)}$

(vi) *Check convergence*: Terminate if the empirical joint violation probability is within a tolerance level $\eta \in \mathbb{R}^+$ of the desired ϵ, i.e, $|\hat{E}_{joint}(\mathbf{X}^{(k+1)}) - \epsilon| \leqslant \eta$.

Return the solution with the lowest objective that is empirically feasible to the chance constraint, i.e.,

$$\hat{E}_{\text{joint}}(\mathbf{X}^{(k+1)}) \leqslant \epsilon. \tag{7.20}$$

Otherwise, increase iteration count $k = k + 1$, and go back to step (ii).

7.4 CASE STUDY OVERVIEW

In the case studies, we perform simulations on the IEEE 13-node radial distribution feeder [Sch+17] using uncertainty data for solar PV generation and load demand from Pecan Street [Pec], similar to the case used in Chapter 6.

The investigations performed are summarized as follows:

(i) We first apply the bisection search tuning method from Chapter 6 to the joint chance-constrained problem. The results demonstrate that tuning a single parameter does not allow us to achieve a low enough joint violation probability and motivates the necessity of multi-dimensional tuning.

(ii) We provide results using the risk allocation tuning method and demonstrate that is able to attain the desired joint violation probability.

(iii) We use the risk allocation algorithm in conjunction with the *two-step tuning* framework described in Chapter 4 to obtain theoretical guarantees.

We note that, in this thesis, we do not provide benchmark numerical results with alternative methods for solving joint chance constrained problems. This is primarily due to the fact that many existing methods do not easily extend to non-convex problems, particularly problems of the size and complexity of the three-phase unbalanced AC

OPF problem considered in this chapter. For example, for the test system we use in our case studies, when applying the uniform decomposition based on Boole's inequality the risk allocation results in an infeasible optimization problem due to the large number of problem constraints. Alternatively, if we consider using the scenario approach, it is unclear how many samples are required for our optimization problem to achieve the desired probabilistic feasibility guarantee. Thus, to ensure a reliable solution, a large number of scenarios may need to be included, which may be computationally prohibitively due to our large problem size. We leave further numerical results, including benchmarking against alternative joint chance-constrained methods, for future work.

7.5 CASE STUDY: SINGLE-DIMENSIONAL TUNING

We first assess the performance of applying bisection-search tuning method described in Section 6.3.3 to the joint chance-constrained problem and demonstrate that it is insufficiently flexible to attain low enough joint violation probabilities.

7.5.1 *Reactive power constraint tightening*

To apply the bisection-search based tuning algorithm, we need to make a few modifications to existing method. We use the same reformulated problem as described in (6.35). Previously, the reactive power constraint tightening were pre-calculated to the desired quantile, which cannot be done in the joint chance-constrained case because we need to adjust the violation probability for individual constraints as part of the tuning process. Instead, we iteratively tune *both* the voltage magnitude and reactive power constraints which are defined as the product of the tuning parameter s and the estimated standard

deviations of the empirical distribution $\sigma^{\phi}_{q,i}, \sigma_{v,i} \in \mathbb{R}^{+}$ for the reactive power and voltage magnitude constraints. This gives us the following expressions

$$\overline{\lambda}^{\phi}_{q,i} = \underline{\lambda}^{\phi}_{q,i} = s \cdot \sigma^{\phi}_{q,i}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}, \tag{7.21}$$

$$\overline{\lambda}^{\phi}_{v,i} = \underline{\lambda}^{\phi}_{v,i} = s \cdot \sigma^{\phi}_{v,i}, \qquad \forall \phi \in \Phi, i \in \mathcal{N}. \tag{7.22}$$

We use a single tuning parameter s across all constraints. As previously mentioned, this can be interpreted as using a uniform risk allocation across all constraints.

7.5.2 *Numerical results*

We perform the bisection-based algorithm for desired joint violation probabilities of $\epsilon = \{0.05, 0.10, 0.15\}$. For the in-sample tuning data, we use $N = 2880$ random samples, as done in Section 6.5. We perform the out-of-sample evaluation using the same data set consisting of $N_{\text{days}} = 45$ full days as used in Section 6.5. We perform 10 algorithm replications.

Table 7.1 shows the in- and out-of-sample VUF, joint violation probabilities (across all constraints, across all voltage magnitude constraints, and across all reactive power constraints), and worst-case violation probabilities for each constraint type (voltage magnitude upper and lower constraints and reactive power upper and lower constraints). All values are the averaged values on the 10 algorithm replications.

We can see that the algorithm is completely unable to obtain solutions with in- or out-of-sample joint violation probabilities close to $\epsilon = \{0.05, 0.10\}$. The algorithm comes close in the $\epsilon = 0.15$ case. We see that the empirical violation probabilities of the reactive power constraints are significantly higher than those of the voltage magnitude constraints in all cases, indicating that a single tuning parameter does not allow enough flexibility for

	In/out sample	Overall		Voltage magnitude			Reactive power		
		VUF	Joint	Upper	Lower	Joint	Upper	Lower	Joint
$\epsilon = 0.05$	In	7.9757	0.1739	0.0070	0.0119	0.0382	0.0786	0.0782	0.1502
	Out	7.8941	0.1810	0.0066	0.0107	0.0336	0.0914	0.0838	0.1595
$\epsilon = 0.10$	In	7.8294	0.1779	0.0089	0.0126	0.0428	0.0791	0.0801	0.1530
	Out	7.7484	0.1854	0.0091	0.0115	0.0406	0.0929	0.0856	0.1621
$\epsilon = 0.15$	In	7.8076	0.1649	0.0079	0.0089	0.0361	0.0788	0.0805	0.1462
	Out	7.7261	0.1742	0.0077	0.0072	0.0343	0.0923	0.0854	0.1570

Table 7.1: Results from applying the bisection search tuning method to three-phase joint chance-constrained AC OPF problem using a tuning data set consisting of $N = 2880$ random samples. The results are the averaged values from 10 algorithm replications. The out-of-sample evaluation was performed on $N_{days} = 45$ full days of data.

both of the different constraint types. We likely need a tighter parameter for the reactive power constraints than for the voltage magnitude constraints.

7.5.3 *Parameter sweeping*

We investigate this further by sweeping over tuning parameter values and plotting the empirical violation and VUF values of the resulting solutions from the reformulation. Figure 7.1 shows the empirical joint violation probabilities across the voltage magnitude constraints (blue), across the reactive power constraints (yellow), and across all constraints (green) on the left and the total VUF on the right.

Figure 7.1(a) sweeps across parameter values $s = [1, 2.6]$, where both the voltage magnitude and reactive power constraint tightenings are adjusted according to the s value. We note that when $s > 2.6$, the optimization problem becomes infeasible. We observe the overall joint violation probability (green) only reaches a minimum of $\hat{E}_{joint}(\mathbf{X}) = 0.1611$ (occurring when $s = 2.4$). This high joint violation probability is primarily driven by the high reactive power constraint violation probabilities (yellow). The joint reactive power violation probability never decreases past $\hat{E}_{q,joint}(\mathbf{X}) = 0.1326$ (occurring when $s = 2.6$),

while the joint voltage magnitude violation probability (blue) reaches a minimum of $\hat{E}_{v,\text{joint}}(\mathbf{X}) = 0.0316$ when $s = 2.4$. Consequently, for the AC OPF problem we study, when only using a single dimensional tuning parameter, the lowest possible joint violation probability we can obtain is $\hat{E}_{\text{joint}}(\mathbf{X}) = 0.1611$.

To give an example of how multi-dimensional tuning can help improve the solution, we use separate tuning parameters for the voltage magnitude constraints and reactive power constraints. Let the respective parameters be denoted s_{Vm} and s_{Qg}. Figure 7.1(b) fixes $s_{Vm} = 2.4$ (since it is the tuning parameter value that minimized the joint violation above) and sweep across $s_{Qg} = [3, 4.6]$. The minimum joint violation probability is now $\hat{E}_{\text{joint}}(\mathbf{X}) = 0.1257$, occurring when $s_{Qg} = 3.6$. However, we see that the reactive power joint violation (yellow) monotonically decreases as s_{Qg} increases (until $s_{Qg} = 4.6$, after which, the problem becomes infeasible). However, the voltage magnitude joint violation increases as s_{Qg} increases, resulting in an overall joint violation probability that is nearly constant around 0.145. Subsequently, in Figure 7.1(c), we see if we can obtain a lower joint violation probability by fixing $s_{Qg} = 4.6$ and sweeping the voltage constraint tuning parameter across $s_{Vm} = [1, 2.4]$. At $s_{Vm} = 2.1$, the joint violation probability is $\hat{E}_{\text{joint}}(\mathbf{X}) = 0.1094$.

Overall, we see that the tightenings for the two types of constraints (i.e., the voltage magnitude constraints and the reactive power constraints) behave differently enough that we cannot use the same tuning parameter across them. However, when we extend our tuning even to a simple two-dimensional case (where we tune each type of constraint separately) we can obtain much better results.

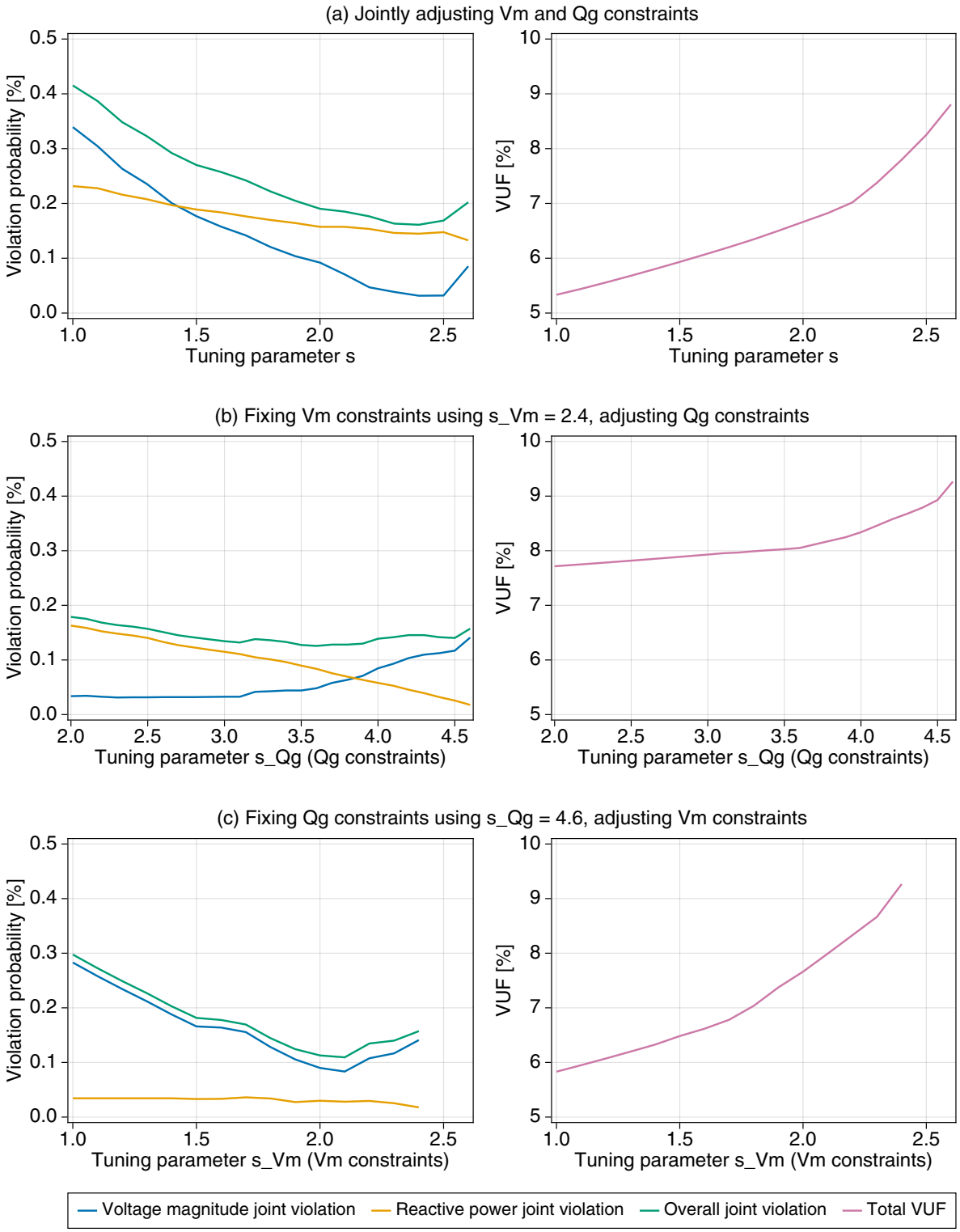


Figure 7.1: Plots showing the empirical joint violation probability and total VUF for various tuning parameter values s.

7.6 CASE STUDY: APPLICATION OF RISK ALLOCATION TUNING

We apply the risk allocation tuning method described in Section 7.3. We assess the algorithm performance for desired violation probability values $\epsilon = \{0.05, 0.10, 0.15\}$. For the in-sample data, we use the random sampling method with $N = 2880$ samples to obtain an uncertainty data set as done in Section 6.4.2. The out-of-sample evaluation is performed using the out-of-sample data set consisting of $N_{\text{days}} = 45$ full days. We perform 10 replications of the algorithm. We use algorithm parameters $\tau = 2$ (to increase the algorithm step size) and $\zeta = 1.2$ (the factor by which infeasible constraints are relaxed). To initialize the risk allocation, we choose a uniform risk allocation that results in a feasible solution. In this case, we set the initial risk allocation to be 0.0125 for all constraints. We note that this is much higher than what we would choose if we tried use the Bonferroni bound (which resulted in an infeasible solution).

Table 7.2 presents the in- and out-of-sample VUF, joint empirical violation probabilities (across all constraints, voltage magnitude constraints, and reactive power constraints), and worst-case violation probabilities for each constraint type (voltage magnitude upper and lower constraints, reactive power upper and lower constraints), and P_{overlap}, averaged over 10 algorithm replications.

7.6.1 *Performance evaluation*

From Table 7.2, we can see the risk allocation tuning algorithm is unable to obtain a solution empirically feasible to the joint chance constraint with the desired in-sample joint violation probability in the case where $\epsilon = 0.05$. However, when $\epsilon = \{0.1, 0.15\}$, the risk allocation tuning algorithm is able to obtain in-sample empirical joint violation

ϵ	In/out sample	Overall			Voltage magnitude			Reactive power		
		VUF	Joint	$P_{overlap}$	Upper	Lower	Joint	Upper	Lower	Joint
0.05	In	8.1083	0.0870	0.0861	0.0114	0.0084	0.0491	0.0125	0.0041	0.0509
	Out	8.0354	0.0994	0.1194	0.0159	0.0097	0.0539	0.0246	0.0048	0.0634
0.1	In	7.4354	0.1000	0.1111	0.0112	0.0119	0.0628	0.0125	0.0112	0.0536
	Out	7.361	0.1132	0.1634	0.0164	0.0149	0.0693	0.0252	0.0163	0.0673
0.15	In	6.4571	0.1488	0.1945	0.0185	0.0557	0.1083	0.0126	0.0114	0.0535
	Out	6.3774	0.1660	0.25854	0.0218	0.0580	0.1190	0.0244	0.0170	0.0658

Table 7.2: Results from applying the *risk allocation* tuning method to three-phase joint chance-constrained AC OPF problem using a tuning data set consisting of $N = 2880$ random samples. The results are the averaged values from 10 algorithm replications. The out-of-sample evaluation was performed on $N_{days} = 45$ full days of data.

probabilities exactly equal to ϵ. The out-of-sample violation probabilities are slightly larger than the in-sample values. In comparison, from the bisection search-based tuning algorithm results presented in Table 7.1, the algorithm was unable to obtain the desired ϵ for any of the values. Thus, the results highlight that the risk allocation tuning algorithm indeed is able to obtain lower joint violation probabilities.

When looking at the VUF values from the risk allocation method, we see that the VUF increases as the desired violation probability ϵ increases, indicating that we obtain higher VUF solutions as we tighten the constraints more, as expected. Interestingly, we see that the out-of-sample VUF values on average lower than the in-sample values (despite the out-of-sample joint violation probability being larger on average).

7.6.2 *Algorithm convergence*

Figure 7.2 plots the total VUF (blue), joint violation probability (green), and violation probability overlap $P_{overlap}$ (pink) for each iteration of a single replication of the risk allocation tuning algorithm using a tuning sample set of $N = 2880$ random samples, targeting a desired $\epsilon = 0.15$. We can see that the VUF starts off high, as the initial

iterations result in solutions with lower than desired empirical joint violation (i.e., too conservative solutions). In iterations 1 to 6, the constraints are relaxed and the VUF drops. Then, the algorithm iterates between relaxing and tightening constraints until it converges to the desired $\epsilon = 0.15$.

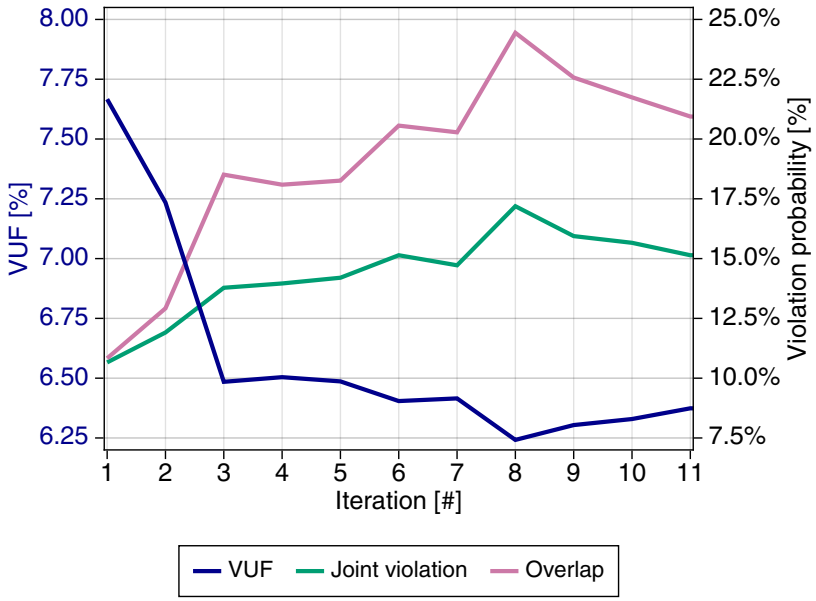


Figure 7.2: The total VUF (blue), in-sample empirical joint violation probability (green), and the aggregation of joint violating probabilities $P_{overlap}$ (pink) across all iterations of a single replication of the risk allocation algorithm, using $N = 2880$ and $\epsilon = 0.15$.

7.6.3 *Constraint tightenings*

We next compare the constraint tightenings and OPF set points obtained with the two different tuning methods using $N = 2440$ tuning samples and $\epsilon = 0.15$. Shown in Figure 7.3 are the voltage magnitude set points and constraint tightenings across all single-phase buses (left) and reactive power set points and constraint tightenings across all PV inverters (right) for single algorithm replications. The nominal constraints are plotted with the red dashed line, the tightened constraints are plotted in blue (bisection search) and green

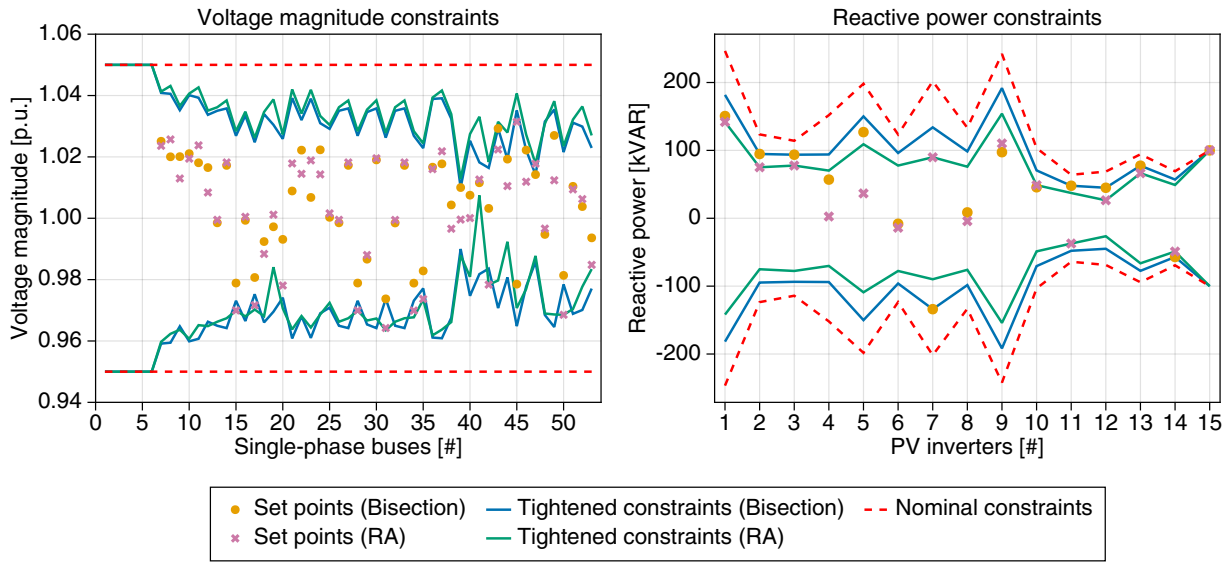


Figure 7.3: Comparison of the constraint tightenings and OPF solutions resulting from the risk allocation tuning method and the bisection-search tuning method for single algorithm replications using $N = 2440$ samples and $\epsilon = 0.15$.

(risk allocation), and the OPF set points are plotted with yellow (bisection search) and pink (risk allocation).

As expected, we can see that the risk allocation algorithm tightenings (green) are not symmetric, unlike the tightenings for the bisection search algorithm (blue). However, we do see that the constraint tightenings for both the voltage magnitude and reactive power constraints are relatively similar for several buses and inverters and generally follow the same pattern.

Another observation for the risk allocation algorithm is that most constraints are not active (i.e., not binding). In particular, only a small fraction of upper voltage magnitude constraints are active, while a larger fraction of the lower voltage magnitude constraints are active. We see that over half of the upper reactive power constraints are active, while only two of the lower reactive power constraints are active.

7.6.4 *Risk allocation results*

In Figure 7.4, we plot the resulting risk allocation from same algorithm replication in the previous results, where a tuning set of $N = 2880$ samples and desired $\epsilon = 0.15$ is used. Immediately, we can see that the risk allocation is non-uniform. In this case, the initial risk allocation of 0.0125 resulted in a solution that was too conservative (i.e., had an empirical violation probability less than $\epsilon = 0.15$), so the algorithm mostly had to relax constraints by increasing the risk allocations of active constraints.

We can see that most constraints have risk allocations that remain at the initial 0.0125. This is as expected since, as seen in Figure 7.3, most constraints are not active. Furthermore, in Figure 7.4, we observe not only constraint relaxations, but tightenings as well. The voltage constraints are mostly relaxed, with some constraints (e.g., the upper constraint at buses 44, 46, and 47 and the lower constraint at bus 31) being allocated very high risk allocations relative to the initial 0.0125. On the other hand, for the reactive power constraints, while the upper constraint at inverter 15 is given a higher risk allocation, indicating that it is being relaxed, several upper and lower constraints end up with a very small risk allocation (meaning they have been tightened). This is also consistent with results shown in Figure 7.2, where we see that initially the problem is too conservative and is relaxed, but then iterates between tightening and relaxing to obtain the desired $\epsilon = 0.15$ joint violation.

Figure 7.5 plots the empirical violation probabilities of each constraint. We can see that most constraints do not have violations. The constraints that do have violations correlate with the constraints that are relaxed (i.e., with lower risk allocations) in Figure 7.4.

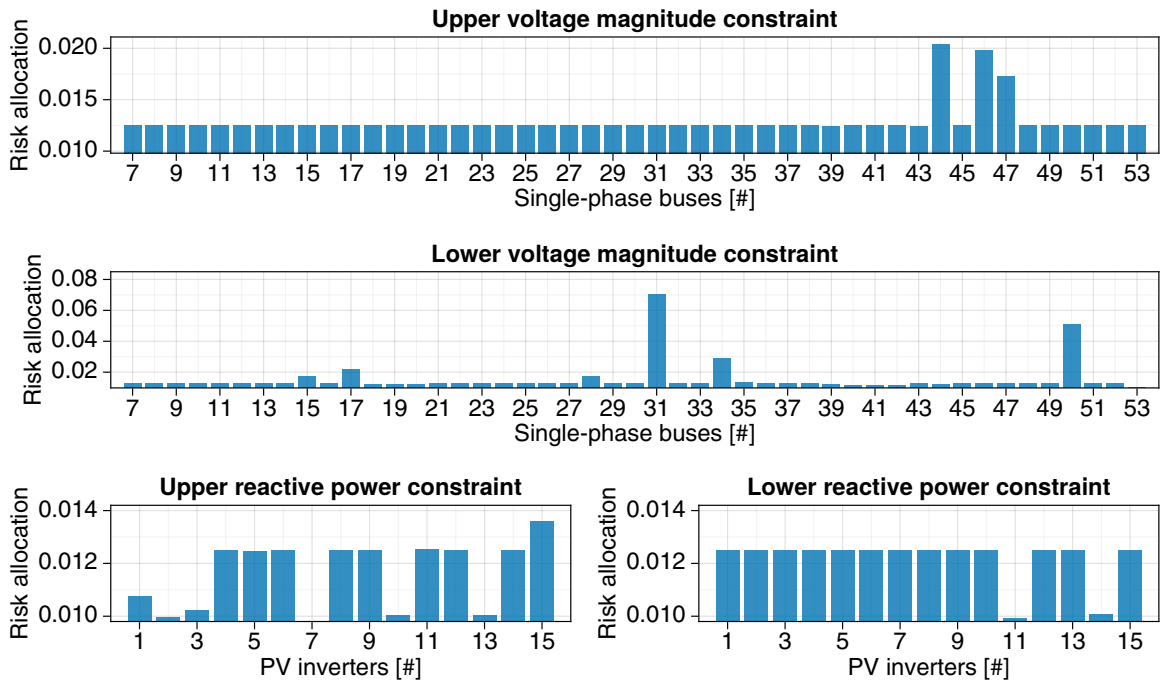


Figure 7.4: The final risk allocation for all constraints of a single replication of the risk allocation algorithm using $N = 2880$ samples and $\epsilon = 0.15$.

7.6.5 *Correlations between individual constraints*

Next, we assess the importance of accounting for the difference between the sum of the individual violation probabilities and the joint violation probabilities, which was quantified by P_{overlap} in equation (7.5).

We highlight that the P_{overlap} values shown in Table 7.2 are large, relative to the desired joint violation probability for all ϵ values considered. This indicates that the sum of the individual violation probabilities is much larger than the joint violation probability, which implies that individual constraints are frequently jointly violated and that the risk allocation algorithm is able to account for these correlations between individual constraints. Furthermore, the large P_{overlap} values underscore the conservativeness of the union bound for the AC OPF problem we consider. In fact, if we apply the union bound

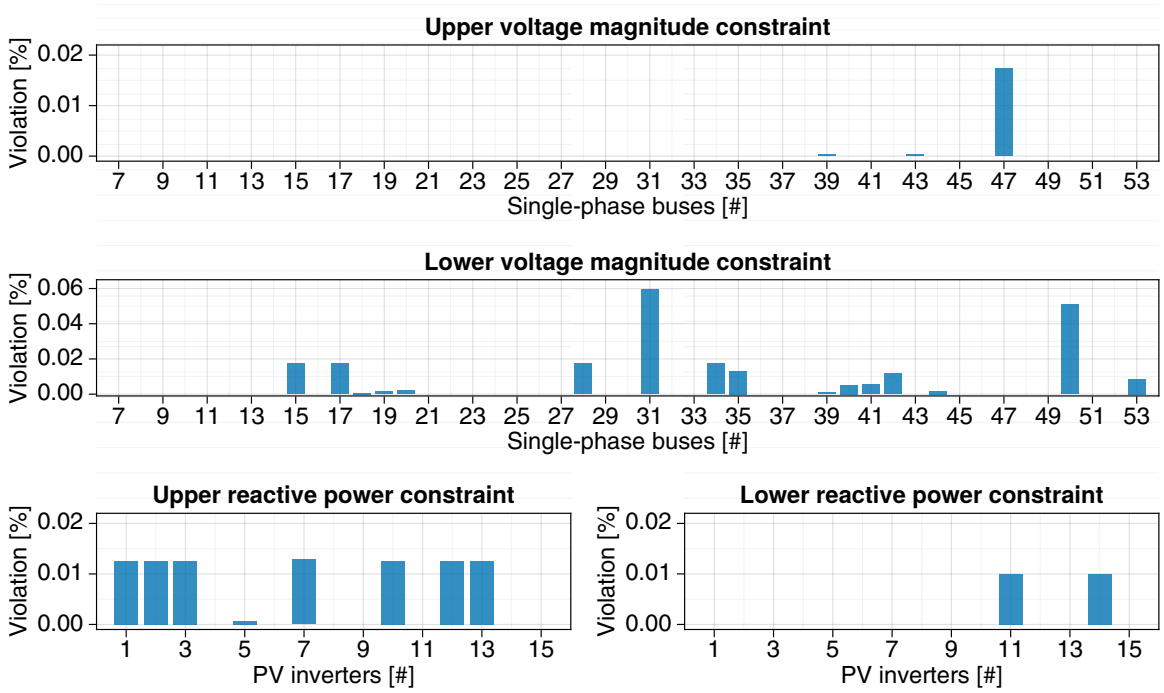


Figure 7.5: The final violation probabilities of each constraint of a single replication of the risk allocation algorithm using $N = 2880$ samples and $\epsilon = 0.15$.

and allocate the overall risk ϵ uniformly for this problem, the resulting optimization problem becomes infeasible.

Looking at the joint violation across the voltage magnitude constraints and across the reactive power constraints, we can see that the sum of these values is only slightly larger than the joint violation across all constraints. This indicates that the occurrence of joint violations are typically within the constraint types themselves. This can be observed in Figure 7.6, which plots the number of violated constraints at each time step during an *out-of-sample day*, with a set point obtained from the risk allocation tuning algorithm run on a sample set of $N = 2880$ tuning samples and $\epsilon = 0.15$. The top figure shows number of single-phase buses with voltage magnitude lower constraint violations (blue) and upper constraint violations (yellow) at each time step. The bottom figure shows the number of PV inverters with reactive power lower constraint violations (green) and upper constraint violations (pink) at each time step. We see that several joint violations (with violations in both voltage magnitude and reactive power constraints) occur around noon.

However, in general, voltage magnitude constraint violations are more spread throughout the day, while reactive power constraints only occur around noon during this day.

In Figure 7.7, we plot the magnitude of the maximum constraint violations (across all single-phase buses or inverters) for each time step in an out-of-sample day using the same set point as in the above. Similarly to the set point obtained from the bisection search shown in Figure 6.7, we see that the magnitude of the voltage magnitude violations are very small (less that 1% of the limit). For the reactive power constraints, we note that these percentages appear very large because some reactive power constraints at certain parts of the day (particularly around noon) are at or close to zero due to the high amount of active power generation from the PV inverters.

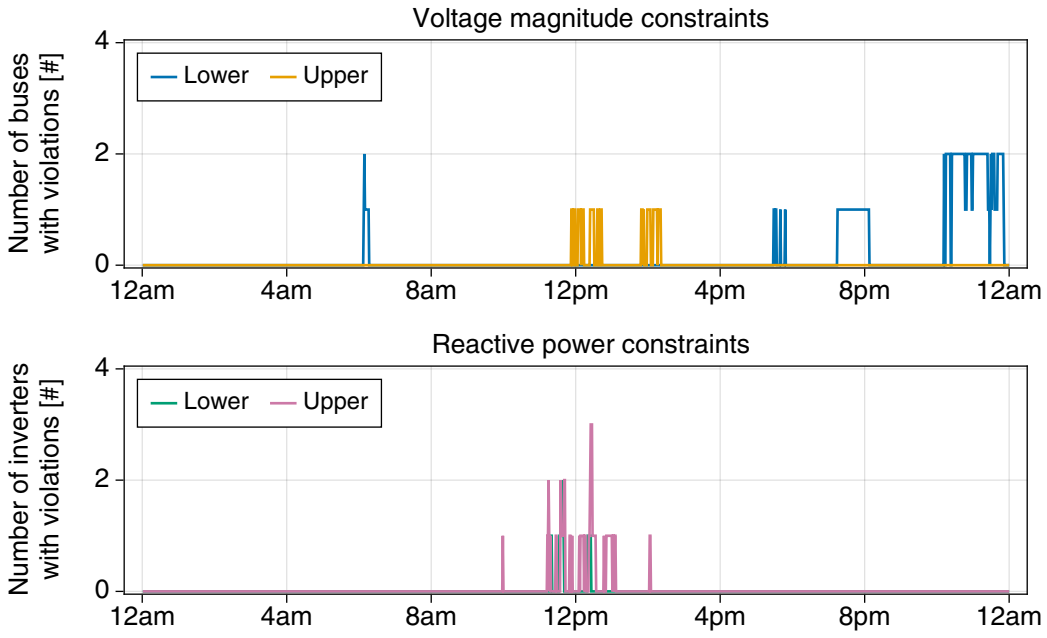


Figure 7.6: Plot of the number of constraint violations that occur across a single out-of-sample evaluation day using a set point from the risk allocation tuning algorithm run with $N = 2880$ tuning samples and $\epsilon = 0.15$.

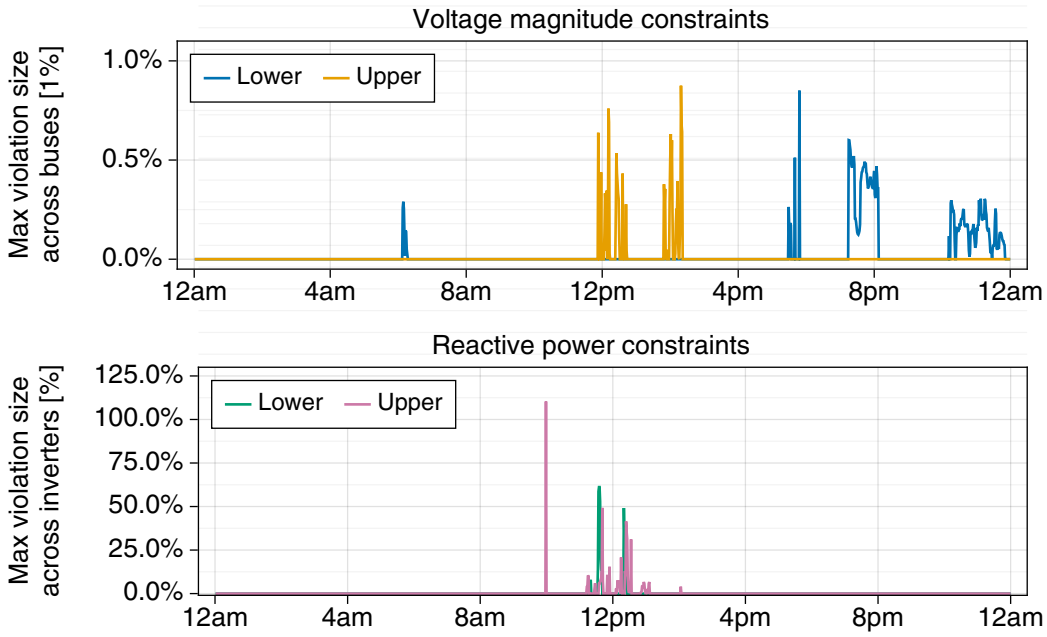


Figure 7.7: Plot of the maximum constraint violation across each constraint type as a percentage of the nominal constraint occurring at each time step using a set point from the risk allocation tuning algorithm run with $N = 2880$ tuning samples and $\epsilon = 0.15$.

7.7 CASE STUDY: APPLICATION OF TWO-STEP TUNING

In this section, we combine the risk allocation tuning algorithm with the two-step methodology proposed in Chapter 4 to obtain theoretical guarantees on the joint chance-constrained three-phase AC OPF problem.

We use the same test system as detailed in Section 6.4.1 and target a desired $\epsilon = 0.15$. We use confidence parameters $\delta_{\text{tune}} = \delta_t = \beta = 0.001$. We choose $c = 1/3$, which gives

			Voltage magnitude			Reactive power		
	Total VUF	Total joint	Upper	Lower	Joint	Upper	Lower	Joint
Generation	7.5269	0.0999	0.0114	0.0114	0.0622	0.0125	0.0114	0.0546
Verification	7.5185	0.1063	0.0141	0.0112	0.0646	0.0211	0.0133	0.0633

Table 7.3: Two-step tuning results for the AC OPF problem, using $\epsilon = 0.15$, $c = 1/3$, and $N = 72471$. The results are the averaged values from 10 algorithm replications.

a tuning margin $t_{\text{tune}} = \epsilon/3 = 0.05$. As a result, the tuning algorithm targets a joint violation of $\epsilon - t_{\text{tune}} = 0.15 - 0.5 = 0.1$.

Following the sample complexity bound (5.9) evaluated for $\delta_{\text{tune}} = 0.001$ and $\epsilon = 0.1$, we use $N = 72471$ randomly drawn uncertainty samples in each step. This results in requiring $t > \sqrt{\frac{1}{2N} \ln \frac{1}{\delta_t}} \approx 0.00690$ and $k > \sqrt{\frac{1}{N} \ln \frac{1}{\beta}} \approx 0.00976$. Therefore, we choose the verification margin to be $t = 0.007$. We note that the tuning margin $t_{\text{tune}} = 0.05$ satisfies the requirement of $t_{\text{tune}} \geqslant t + k$ and therefore Claim 4.5 applies.

Recall that in the solution generation step (i.e., application of the tuning algorithm), we required the final solution $\mathbf{X}^*$ to pass the following empirical feasibility criterion

$$\hat{E}_{\text{joint}}(\mathbf{X}^*) + t_{\text{tune}} \leqslant \epsilon \qquad \Leftrightarrow \qquad \hat{E}_{\text{joint}}(\mathbf{X}^*) \leqslant 0.1. \tag{7.23}$$

In the solution verification step, the a posteriori feasibility guarantee (Claim 4.2) states that a solution $\mathbf{X}^*$ is feasible to the original chance constraint (7.9c) with probability at least $1 - \delta_t$ if $\mathbf{X}^*$ satisfies the a posteriori solution verification criterion. The criterion requires the empirical violation probability of $\mathbf{X}^*$ plus the verification margin t be less than the desired violation ϵ, i.e.,

$$\hat{E}(\mathbf{X}^*, \Xi_2) + t \leqslant \epsilon \qquad \Leftrightarrow \qquad \hat{E}(\mathbf{X}^*, \Xi_2) \leqslant 0.143. \tag{7.24a}$$

Table 7.3 shows the VUF and violation probabilities of both the solution generation step (in-sample) and the a posteriori verification step (out-of-sample), averaged across 10 algorithm replications. The table reports the worst case (i.e., maximum across the constraint type) violation probabilities of the upper and lower voltage magnitude and reactive power constraints, the joint voltage magnitude and joint reactive power violation probabilities, and overall joint violation probability.

The risk allocation tuning algorithm is able to obtain an average in-sample joint violation probability of $\epsilon - t_{\text{tune}} \leqslant 0.1$ in the solution generation step. However, when

looking at the replication results, we see that only 4/10 replications have a joint violation probability of equal to or less than 0.1. The remaining 6/10 replications all achieve an empirical in-sample joint violation of 0.1001. Although this violation probability is slightly too high, it is within the desired tolerance level of the algorithms. Thus, all solutions are declared feasible for the solution generation step and verified using the verification step. Furthermore, we note that regardless if (7.23) holds, if the verification feasibility criterion (7.24) is satisfied, the a posteriori feasibility guarantee (Claim 4.2) is still applicable.

In the solution verification step, we see that not only does the average a posteriori evaluated joint violation probability shown in Table 7.3 satisfy the verification criterion (7.24), but all algorithm replications do as well. In fact, the highest a posteriori evaluated empirical violation probability is 0.1076. Thus, solutions from all replications can be declared to be feasible to the joint chance-constrained problem with confidence level $1 - \delta_t = 0.999$ via Claim 4.2.

The a priori feasibility result (Claim 4.7) provides that we should have a confidence level of

$$(1 - \delta_t)(1 - \beta) = (1 - 0.001)(1 - 0.001) = 0.998 \tag{7.25}$$

that a solution obtained from the two-step tuning process (i.e., a solution that satisfies the tuning criterion) is feasible to the original chance-constrained problem. Given that all algorithm replications that satisfied the tuning criterion also satisfied the verification criteria, our empirical results align with this a priori claim.

7.8 CONCLUSION

This chapter proposes a multi-dimensional tuning method for joint chance-constrained optimization based on the idea of risk allocation. The method is used to find solar PV

inverter reactive power set points via the joint chance constrained three-phase unbalanced AC OPF problem formulated in Chapter 6.

In the case study, we first demonstrate that the single-dimensional bisection search-based tuning algorithm is unable to obtain OPF solutions with the desired joint violation probability. Then, in contrast, the risk allocation-based tuning method is shown to be able to achieve solutions with the desired joint violation probability. Finally, we combine the risk allocation algorithm with the two-step tuning methodology and establish that we are able to get probabilistic feasibility guarantees on the solution.

8 CONCLUSION

This dissertation introduced a flexible data-driven tuning framework for solving chance-constrained optimization problems that balances optimality, feasibility, and computational tractability. Variations of the tuning approach were designed to solve OPF problems adapted for transmission and distribution grid operations and operational planning, in settings where the distribution of uncertainty is unknown. The following summarizes the content of each chapter, highlights some of the main takeaways of the thesis, and identifies directions for future work.

8.1 SUMMARY

Chapter 3 proposed a generalized tuning methodology for chance-constrained problems that combines a simple, computationally lightweight approximate optimization model with feedback from sample-based evaluations to iteratively adapt model parameters. An example algorithm employing an approximate model (inspired by the moment-based analytical chance constraint reformulation) and a bisection search on a single-dimensional safety parameter is presented. It is used in the case study to determine low cost generation schedules under uncertainty due to renewable generation and load variability for transmission grid operation via solving the single and joint chance-constrained DC OPF problems.

Chapter 4 devised a two-step framework that allows tuning methods, such as the algorithm proposed in Chapter 3, to achieve rigorous a posteriori probabilistic feasibility

guarantees. The relationship between the tuning and verification steps is analyzed to provide a theoretical basis for an algorithm design that *a priori* asserts a feasible solution will be returned with high probability. The validity of the assumptions used in deriving the a priori results of the two-step tuning method is further examined in Chapter 5, along with the algorithm's sample complexity and scalability to larger network sizes. Case studies in Chapters 4 and 5 apply the two-step method to joint chance-constrained DC OPF for transmission grids and provide numerical results that validate the theoretical developments.

Shifting focus to the distribution grid setting, Chapter 6 adapted the bisection search-based tuning algorithm to solve a single chance-constrained formulation of the three-phase unbalanced AC OPF problem used to find reactive power set points for PV inverters that minimize the total voltage unbalance. Case studies explore how to effectively sample realistic solar PV and load data in the tuning process and demonstrate that the algorithm is able to achieve solutions with the desired violation probability.

To more effectively solve joint chance-constrained problems, an alternative tuning heuristic that allows for multi-dimensional tuning is proposed in Chapter 7. The risk allocation-based algorithm decomposes the joint chance constraint into single chance constraints and iteratively adjusts the risk levels (i.e., acceptable violation probabilities) of the individual constraints using sample-based evaluations as well as solution characteristics. Simulation results demonstrate this multi-dimensional method performs favorably to alternative methods, such as the bisection-search algorithm and the usage of the Bonferroni bound. The risk allocation tuning algorithm is further combined with the two-step methodology and is shown to be able to find solutions that are feasible to the joint chance constraint with high probability.

8.2 CONCLUSIONS

We highlight some of the overarching findings and conclusions of this thesis below.

(i) *Balancing solution optimality, feasibility, and computational tractability*: In the case studies throughout the thesis, we see that, given a fixed number of samples, to ensure a greater probability of feasibility, we need to use higher margin terms, which results in less optimal solutions. On the other hand, we can decrease the margin (and increase optimality) without impacting the feasibility guarantees by incorporating more samples into the tuning process. However, the increase in the required number of samples may limit computational tractability.

(ii) *Decoupling of the optimization problem from the uncertainty evaluation*: Unlike other methods, we do not incorporate the uncertainty data within the problem itself but rather perform an evaluation of the uncertainty after obtaining a candidate solution. This allows us to leverage large amounts of data without increasing the computational complexity of the optimization problem being solved.

(iii) *Use of approximate optimization models*: The use of approximate models that do not have to provide probabilistic guarantees allows us to *handle problems with undesirable or complicated structure* (e.g., non-convex optimization problems). While a simpler approximate formulation may be less accurate, these models are much more *computationally tractable* and allow us to take advantage of the efficiency and scalability of commercial, off-the-shelf solvers. We make up for the possible model inaccuracy by leveraging large amounts of data for tuning.

(iv) *Correlations between constraints need to be considered*: For joint chance-constrained problems, particularly in power systems applications, we have observed that there are significant correlations between constraints as observed by the large overlap term

values in Chapter 7. As a result, we need to consider methods that take into account these correlations in order to obtain solutions that are not overly conservative.

8.3 FUTURE WORK

This dissertation proposed data-driven tuning algorithms for the OPF problem under uncertainty, with the ability to obtain probabilistic feasibility guarantees. Building on the proposed framework, we present a few directions for future work.

8.3.1 *Identifying new approximate problem formulations*

The advantage of using tuning-based algorithms is that it relies on the use of a simple, computationally lightweight optimization problem that is iteratively updated. A key challenge addressed in this thesis is the formulation of approximate models best suited for the optimization problems being solved. As mentioned before, the approximate model should capture salient features of the problem setting (e.g., transmission or distribution grid model), incorporate an suitable uncertainty representation, and contain parameters that are amenable to tuning.

There are several avenues for formulating alternative approximate problem models:

(i) In this thesis, we use an approximate model with constraint tightenings defined as the tuning parameter multiplied by an estimate of the standard deviation of the nominal constraint or defined as a quantile of the empirical distribution. In Section 3.2.1, we see that the former choice of tightening is analogous to tuning the size of an ellipsoidal uncertainty set. In future work, we can investigate the use of different uncertainty sets, such as polyhedral uncertainty sets [BEN09], which may provide a more accurate representation of the uncertainty distribution.

(ii) We can explore the use of alternative risk measures, which give rise to different probabilistic constraint formulations. For example, in the case studies in Sections 6.6.3 and 7.6.5, we see that the magnitude of violation of the constraints varies amongst the constraint type and uncertainty realization. We can use alternative risk measures that account for the violation magnitude *as well as* the violation probability. Examples may include the use of CVaR constraints [NS07] or weighted chance constraints [Roa+15a]. Certain risk measures can also be linked to different uncertainty sets, as certain classes of risk measures have been shown to be equivalent to some uncertainty sets used in robust optimization [BB09].

(iii) We can also study the impact of the grid model on solution performance. In Chapters 6 and 7, for distribution grid tuning, we use the full non-linear, non-convex three-phase power flow formulation, which may be challenging to solve for large scale, realistic distribution grids. Most existing techniques used to reduce computational complexity for distribution grid optimization are based on linear approximations and convex relaxations of the power balance equations [GR21]. We can consider using different approximate power flow models for solving the optimization problem, while still using the full three-phase, non-convex distribution grid model in the evaluation step.

8.3.2 *Alternative tuning schemes*

This thesis proposes two different tuning schemes: a simple single-dimensional bisection search-based tuning algorithm and a more sophisticated multi-dimensional risk allocation-based tuning algorithm. We can also investigate alternative updating rules, such as using a gradient-descent style scheme for tuning. Furthermore, we can look to incorporate different sources of information, such as information on the correlations

between constraints, into the tuning scheme. In general, the most suitable tuning scheme to use will depend on the choice of approximate model formulation.

A COMPARISON METHODS FOR CHANCE-CONSTRAINED DC OPF

A.1 SCENARIO APPROACH REFORMULATION

The scenario approach [CGP09; CC06] is a common approach used to solve chance-constrained problems. While it requires no assumptions on the underlying distribution of the uncertainty ξ, it relies on the availability of uncertainty samples or scenarios $\Xi_{SA} = \{\xi^{(1)}, \dots, \xi^{(N_{SA})}\}$ to transform probabilistic constraints into deterministic constraints. The chance-constrained problem is reformulated by enforcing that all constraints must hold for all scenarios, i.e.,

$$\min_{x} \quad \sum_{i \in \mathcal{N}} c_{2,i} x_i^2 + c_{1,i} x_i + c_{0,i} \tag{A.1a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} x_i - d_i = 0, \tag{A.1b}$$

$$\Omega_\xi^{(k)} := \sum_{i \in \mathcal{N}} \xi_i^{(k)} \qquad \forall k \in [N_{SA}] \tag{A.1c}$$

$$x_i - \alpha_i \Omega_\xi^{(k)} \leqslant p_{G,i}^{max}, \qquad \forall i \in \mathcal{N},\ k \in \Xi_{SA} \tag{A.1d}$$

$$x_i - \alpha_i \Omega_\xi^{(k)} \geqslant p_{G,i}^{min}, \qquad \forall i \in \mathcal{N},\ k \in \Xi_{SA} \tag{A.1e}$$

$$M_{(ij,\cdot)}(x - \alpha \Omega_\xi^{(k)} + \xi^{(k)} - d) \leqslant p_{ij}^{max}, \qquad \forall ij \in \mathcal{L},\ k \in \Xi_{SA} \tag{A.1f}$$

$$M_{(ij,\cdot)}(x - \alpha \Omega_\xi^{(k)} + \xi^{(k)} - d) \geqslant p_{ij}^{min}, \qquad \forall ij \in \mathcal{L},\ k \in \Xi_{SA} \tag{A.1g}$$

Here, the objective function and nodal balance constraint remain the same as in the deterministic problem. Constraint (A.1c) defines $\Omega_\xi^{(k)}$ as the total uncertainty summed

across all nodes $\mathcal{N}$ for sample draw $k \in [N_{SA}]$. Constraints (A.1d) - (A.1g) enforce the generator and line constraints for each uncertainty realization.

A.2 CVAR REFORMULATION

The CVaR reformulation [NS07] is a conservative (inner) convex approximation of the joint chance-constrained problem. To solve the CVaR reformulation, we use the sample average approximation approach [WA08] with a sample set $\Xi_{CVaR} = \{\xi^{(1)}, \ldots, \xi^{(N)}\}$, which results in the following deterministic approximate problem:

$$\min_{x,\beta \leqslant 0, s} \quad \sum_{i \in \mathcal{N}} c_{2,i} x_i^2 + c_{1,i} x_i + c_{0,i} \tag{A.2a}$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} x_i - d_i = 0 \tag{A.2b}$$

$$\Omega_\xi^{(k)} = \sum_{i \in \mathcal{N}} \xi_i^{(k)}, \qquad \forall k \in [N] \tag{A.2c}$$

$$x_i - \alpha_i \Omega_\xi^{(k)} - p_{G,i}^{max} \leqslant s_k, \qquad \forall i \in \mathcal{N}, \forall k \in [N] \tag{A.2d}$$

$$p_{G,i}^{min} - x_i + \alpha_i \Omega_\xi^{(k)} \leqslant s_k, \qquad \forall i \in \mathcal{N}, \forall k \in [N] \tag{A.2e}$$

$$M_{(ij,\cdot)}(x - \alpha \Omega_\xi^{(k)} + \xi^{(k)} - d) - p_{ij}^{max} \leqslant s_k, \qquad \forall ij \in \mathcal{L}, \forall k \in [N] \tag{A.2f}$$

$$p_{ij}^{min} - M_{(ij,\cdot)}(x - \alpha \Omega_\xi^{(k)} + \xi^{(k)} - d) \leqslant s_k, \qquad \forall ij \in \mathcal{L}, \forall k \in [N] \tag{A.2g}$$

$$\frac{1}{N} \sum_{k=1}^{N} s_k - (1 - \epsilon)\beta \leqslant 0 \tag{A.2h}$$

$$s_k \geqslant \beta \qquad \forall k \in [N]. \tag{A.2i}$$

The objective function, nodal balance constraint, and definition of $\Omega_k^{(k)}$ remain the same as for the scenario approach. Here, we introduce non-negative auxiliary variables s_k for each uncertainty realization $k = 1, \ldots, N$ to denote violations in the generator and line

flow constraints (A.2d)-(A.2g). Auxiliary variable $\beta \leqslant 0$ and constraints (A.2h) and (A.2i) result directly from the definition of a CVaR constraint [NS07].

BIBLIOGRAPHY

[02] *IEC 61000-2-2, EMC – Part 2-2: Environment – Compatibility Levels for Low Frequency Conducted Disturbances and Signalling in Public Low-Voltage Power Supply Systems*. 2002.

[ABL17] Hamdi Abdi, Soheil Derafshi Beigvand, and Massimo La Scala. "A review of optimal power flow studies applied to smart grids and microgrids". In: *Renewable and Sustainable Energy Reviews* 71 (2017), pp. 742–766.

[Ahm+17] Shabbir Ahmed et al. "Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs". In: *Mathematical Programming* 162.1 (2017), pp. 51–81.

[Ali+15] Arash Alimardani et al. "Distribution system state estimation based on nonsynchronized smart meters". In: *IEEE Transactions on Smart Grid* 6.6 (2015), pp. 2919–2928.

[Arn+16] Daniel B Arnold et al. "Optimal dispatch of reactive power for voltage regulation and balancing in unbalanced distribution systems". In: *IEEE PES General Meeting*. IEEE. 2016, pp. 1–5.

[ARP] ARPA-E. "ARPA-E Grid Optimization Competition Datasets". In: (). URL: https://gocompetition.energy.gov/datasets/c2finalevent.

[AS08] Shabbir Ahmed and Alexander Shapiro. "Solving chance-constrained stochastic programs via sampling and integer programming". In: *State-of-the-art decision-making tools in the information-intensive age*. Informs, 2008, pp. 261–269.

[ATL10] T. Alamo, R. Tempo, and A. Luque. "On the sample complexity of randomized approaches to the analysis and design under uncertainty". In: *Proceedings of the 2010 American Control Conference*. 2010, pp. 4671–4676. DOI: 10.1109/ACC.2010.5531078.

[AZ15] Farzaneh Abbaspourtorbati and Marek Zima. "The Swiss reserve market: Stochastic programming in practice". In: *IEEE Transactions on Power Systems* 31.2 (2015), pp. 1188–1194.

[BB09] Dimitris Bertsimas and David B Brown. "Constructing uncertainty sets for robust linear optimization". In: *Operations research* 57.6 (2009), pp. 1483–1495.

[BB18] Kyri Baker and Andrey Bernstein.
"JOINT CHANCE CONSTRAINTS REDUCTION THROUGH LEARNING IN ACTIVE DISTRIBUTION NETWORKS".
In: *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. 2018, pp. 922–926. DOI: 10.1109/GlobalSIP.2018.8646440.

[BB19] Kyri Baker and Andrey Bernstein. "Joint Chance Constraints in AC Optimal Power Flow: Improving Bounds Through Learning".
In: *IEEE Transactions on Smart Grid* 10.6 (2019), pp. 6376–6385.
DOI: 10.1109/TSG.2019.2903767.

[BBC11] Dimitris Bertsimas, David B Brown, and Constantine Caramanis.
"Theory and applications of robust optimization".
In: *SIAM review* 53.3 (2011), pp. 464–501.

[BCH14] Daniel Bienstock, Michael Chertkov, and Sean Harnett.
"Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty". In: *SIAM Review* 56.3 (2014), pp. 461–495.

[BEN09] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski.
Robust optimization. Vol. 28. Princeton University Press, 2009.

[Ber+18] Andrey Bernstein et al. "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models".
In: *IEEE Trans. Power Systems* 33.6 (2018), pp. 5832–5843.

[Bez+17] J. Bezanson et al. "Julia: A Fresh Approach to Numerical Computing".
In: *SIAM Review* 59.1 (2017), pp. 65–98. DOI: 10.1137/141000671.
eprint: https://doi.org/10.1137/141000671.
URL: https://doi.org/10.1137/141000671.

[Bla06] Lars Blackmore. "A probabilistic particle control approach to optimal, robust predictive control".
In: *AIAA Guidance, Navigation, and Control Conference and Exhibit*. 2006, p. 6240.

[Boo54] George Boole. *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities*. Dover Publications, 1854.

[BOW11] Lars Blackmore, Masahiro Ono, and Brian C Williams.
"Chance-constrained optimal path planning with obstacles".
In: *IEEE Transactions on Robotics* 27.6 (2011), pp. 1080–1094.

[BT17] Kyri Baker and Bridget Toomey. "Efficient Relaxations for Joint Chance Constrained AC Optimal Power Flow". en.
In: *Electric Power Systems Research* 148 (July 2017), pp. 230–236.
ISSN: 03787796. DOI: 10.1016/j.epsr.2017.04.001.

[Cal10] Giuseppe Carlo Calafiore. "Random convex programs".
In: *SIAM Journal on Optimization* 20.6 (2010), pp. 3427–3464.

[Cao+13] Yijia Cao et al. "Chance-Constrained Optimization-Based Unbalanced Optimal Power Flow for Radial Distribution Networks". In: *IEEE Transactions on Power Delivery* 28.3 (2013), pp. 1855–1864. DOI: 10.1109/TPWRD.2013.2259509.

[Car36] E Carlo. "Bonferroni. Teoria statistica delle classi e calcolo delle probabilita". In: *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8.3-62 (1936), p. 4.

[CC05] Giuseppe Calafiore and Marco C Campi. "Uncertain convex programs: randomized solutions and confidence levels". In: *Mathematical Programming* 102.1 (2005), pp. 25–46.

[CC06] Giuseppe C Calafiore and Marco C Campi. "The scenario approach to robust control design". In: *IEEE Transactions on automatic control* 51.5 (2006), pp. 742–753.

[CCS58] Abraham Charnes, William W Cooper, and Gifford H Symonds. "Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil". In: *Management science* 4.3 (1958), pp. 235–263.

[CE06] Giuseppe Carlo Calafiore and Laurent El Ghaoui. "On distributionally robust chance-constrained linear programs". In: *Journal of Optimization Theory and Applications* 130.1 (2006), pp. 1–22.

[CG08] Marco C Campi and Simone Garatti. "The exact feasibility of randomized solutions of uncertain convex programs". In: *SIAM Journal on Optimization* 19.3 (2008), pp. 1211–1230.

[CG11] Marco Campi and Simone Garatti. "A Sampling-and-Discarding Approach to Chance-Constrained Optimization: Feasibility and Optimality". In: *Journal of Optimization Theory and Applications* 148 (Feb. 2011), pp. 257–280. DOI: 10.1007/s10957-010-9754-6.

[CGP09] Marco C Campi, Simone Garatti, and Maria Prandini. "The scenario approach for systems and control design". In: *Annual Reviews in Control* 33.2 (2009), pp. 149–157.

[CGR15] Marco C Campi, Simone Garatti, and Federico A Ramponi. "Non-convex scenario optimization with application to system identification". In: *IEEE Conference on Decision and Control*. IEEE. 2015, pp. 4023–4028.

[CGR18] Marco Claudio Campi, Simone Garatti, and Federico Alessandro Ramponi. "A General Scenario Theory for Nonconvex Optimization and Decision Making". In: *IEEE Transactions on Automatic Control* 63.12 (2018), pp. 4067–4078. DOI: 10.1109/TAC.2018.2808446.

[Che+10] Wenqing Chen et al. "From CVaR to uncertainty set: Implications in joint chance-constrained optimization". In: *Operations research* 58.2 (2010), pp. 470–485.

[CKW22] Zhi Chen, Daniel Kuhn, and Wolfram Wiesemann.
"Data-driven chance constrained programs over Wasserstein balls".
In: *Operations Research* (2022).

[Cof+18] Carleton Coffrin et al. "PowerModels.jl: An Open-Source Framework for Exploring Power Flow Formulations".
In: *2018 Power Systems Computation Conference (PSCC)*. June 2018, pp. 1–8.
DOI: 10.23919/PSCC.2018.8442948.

[CWWoo] R.D. Christie, B. F. Wollenberg, and I. Wangensteen.
"Transmission Management in the Deregulated Environment".
In: *Proceedings of the IEEE* 88.2 (Feb. 2000), pp. 170–195.

[CWZ18] Xin Chen, Wenchuan Wu, and Boming Zhang.
"Robust Capacity Assessment of Distributed Generation in Unbalanced Distribution Networks Incorporating ANM Techniques".
In: *IEEE Transactions on Sustainable Energy* 9.2 (2018), pp. 651–663.
DOI: 10.1109/TSTE.2017.2754421.

[DBS17] Emiliano Dall'Anese, Kyri Baker, and Tyler Summers. "Chance-constrained AC optimal power flow for distribution systems with renewables".
In: *IEEE Transactions on Power Systems* 32.5 (2017), pp. 3427–3438.

[DFL18] Thai Dinh, Ricardo Fukasawa, and James Luedtke.
"Exact algorithms for the chance-constrained vehicle routing problem".
In: *Mathematical Programming* 172.1 (2018), pp. 105–138.

[DH17] Andre Luiz Diniz and René Henrion. "On probabilistic constraints with multivariate truncated Gaussian and lognormal distributions".
In: *Energy Systems* 8.1 (2017), pp. 149–167.

[DHL17] Iain Dunning, Joey Huchette, and Miles Lubin.
"JuMP: A Modeling Language for Mathematical Optimization".
In: *SIAM Review* 59.2 (2017), pp. 295–320. DOI: 10.1137/15M1020575.

[Dua+18] Chao Duan et al. "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric".
In: *IEEE Transactions on Power Systems* 33.5 (2018), pp. 4924–4936.

[Duq+20] Daniel Duque et al.
"Timing social distancing to avert unmanageable COVID-19 hospital surges".
In: *Proceedings of the National Academy of Sciences* 117.33 (2020),
pp. 19873–19878.

[DZG13] Emiliano Dall'Anese, Hao Zhu, and Georgios B Giannakis.
"Distributed optimal power flow for smart microgrids".
In: *IEEE Trans. Smart Grid* 4.3 (2013), pp. 1464–1475.

[ESL14] Peyman Mohajerin Esfahani, Tobias Sutter, and John Lygeros.
"Performance bounds for the scenario approach and an extension to a class of non-convex programs".
In: *IEEE Transactions on Automatic Control* 60.1 (2014), pp. 46–58.

[FZA16] Philipp Fortenbacher, Martin Zellner, and Göran Andersson. "Optimal sizing and placement of distributed storage in low voltage networks". In: *2016 Power Systems Computation Conference (PSCC)*. IEEE. 2016, pp. 1–7.

[Gan+14] Lingwen Gan et al. "Exact convex relaxation of optimal power flow in radial networks". In: *IEEE Trans. Automatic Control* 60.1 (2014), pp. 72–87.

[GHR22] Kshitij Girigoudar, Ashley M. Hou, and Line A. Roald. "Chance-Constrained AC Optimal Power Flow for Unbalanced Distribution Grids". In: *11th Bulk Power Systems Dynamics and Control Symposium (IREP 2022)*. IEEE. 2022.

[GOO03] Laurent El Ghaoui, Maksim Oks, and Francois Oustry. "Worst-case value-at-risk and robust portfolio optimization: A conic programming approach". In: *Operations research* 51.4 (2003), pp. 543–556.

[GR20] Kshitij Girigoudar and Line A Roald. "On the impact of different voltage unbalance metrics in distribution system optimization". In: *Electric Power Systems Research* 189 (2020), p. 106656.

[GR21] Kshitij Girigoudar and Line A Roald. "Linearized Three-Phase Optimal Power Flow Models for Distribution Grids with Voltage Unbalance". In: *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE. 2021, pp. 4214–4221.

[Gra+15] Sergio Grammatico et al. "A scenario approach for non-convex control design". In: *IEEE Transactions on Automatic Control* 61.2 (2015), pp. 334–345.

[Gri+99] C. Grigg et al. "The IEEE Reliability Test System-1996". en. In: 14.3 (Aug./1999), pp. 1010–1020. ISSN: 08858950. DOI: 10.1109/59.780914.

[Gur21] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2021. URL: https://www.gurobi.com.

[GVH20] Sijia Geng, Maria Vrakopoulou, and Ian A Hiskens. "Optimal Capacity Design and Operation of Energy Hub Systems". In: *Proceedings of the IEEE* (2020).

[GX19] Xinbo Geng and Le Xie. *Data-driven Decision Making with Probabilistic Guarantees (Part 1): A Schematic Overview of Chance-constrained Optimization*. 2019. arXiv: 1903.10621 [math.OC].

[GXM22] Xinbo Geng, Le Xie, and M. Sadegh Modarresi. "Computing Essential Sets for Convex and Nonconvex Scenario Problems: Theory and Application". In: *IEEE Transactions on Control of Network Systems* 9.1 (2022), pp. 269–281. DOI: 10.1109/TCNS.2021.3100404.

[Hoe63] Wassily Hoeffding. "Probability Inequalities for Sums of Bounded Random Variables". In: *Journal of the American Statistical Association* 58.301 (1963), pp. 13–30. DOI: 10.1080/01621459.1963.10500830.

[HR20] Ashley M. Hou and Line A. Roald.
"Chance Constraint Tuning for Optimal Power Flow". In: *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*.
IEEE. 2020, pp. 1–6.

[HR22] Ashley M. Hou and Line A. Roald.
"Data-Driven Tuning for Chance-Constrained Optimization: Two Steps Towards Probabilistic Performance Guarantees".
In: *IEEE Control Systems Letters* 6 (2022), pp. 1400–1405.
DOI: 10.1109/LCSYS.2021.3096826.

[Jab06] Rabih A Jabr. "Radial distribution load flow using conic programming".
In: *IEEE Trans. power systems* 21.3 (2006), pp. 1458–1459.

[Jab13] Rabih A. Jabr. "Adjustable Robust OPF With Renewable Energy Sources".
In: *IEEE Transactions on Power Systems* 28.4 (2013), pp. 4742–4751.
DOI: 10.1109/TPWRS.2013.2275013.

[JHS20] Mengshuo Jia, Gabriela Hug, and Chen Shen.
Iterative Decomposition of Joint Chance Constraints in OPF. 2020.
arXiv: 2010.11746 [eess.SY].

[JKK15] Rabih A. Jabr, Sami Karaki, and Joe Akl Korbane.
"Robust Multi-Period OPF With Storage and Renewables".
In: *IEEE Transactions on Power Systems* 30.5 (2015), pp. 2790–2799.
DOI: 10.1109/TPWRS.2014.2365835.

[JX22] Nan Jiang and Weijun Xie. "ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs".
In: *Operations Research* (2022).

[KAH18] Stavros Karagiannopoulos, Petros Aristidou, and Gabriela Hug.
"A centralised control method for tackling unbalances in active distribution grids". In: *Power Systems Computation Conference (PSCC)*. 2018.

[KAH19] Stavros Karagiannopoulos, Petros Aristidou, and Gabriela Hug.
"Data-Driven Local Control Design for Active Distribution Grids Using Off-Line Optimal Power Flow and Machine Learning Techniques".
In: *IEEE Transactions on Smart Grid* 10.6 (2019), pp. 6461–6471.
DOI: 10.1109/TSG.2019.2905348.

[Kar+] Stavros Karagiannopoulos et al.
"Operational planning of active distribution grids under uncertainty". In.

[Kek+15] Vassilis Kekatos et al.
"Stochastic Reactive Power Management in Microgrids With Renewables".
In: *IEEE Transactions on Power Systems* 30.6 (2015), pp. 3386–3395.
DOI: 10.1109/TPWRS.2014.2369452.

[Ker06] William H Kersting. *Distribution system modeling and analysis*.
CRC press, 2006.

[KJ21] Simge Küçükyavuz and Ruiwei Jiang. "Chance-constrained optimization: A review of mixed-integer conic formulations and applications". In: *arXiv preprint arXiv:2101.08746* (2021), p. 2.

[KLR20] Rohit Kannan, James Luedtke, and Line Roald. "Stochastic DC Optimal Power Flow With Reserve Saturation". In: *Power Systems Computation Conference (PSCC)* (2020).

[KPU02] Pavlo Krokhmal, Jonas Palmquist, and Stanislav Uryasev. "Portfolio optimization with conditional value-at-risk objective and constraints". In: *Journal of risk* 4 (2002), pp. 43–68.

[LA08] James Luedtke and Shabbir Ahmed. "A sample approximation approach for optimization with probabilistic constraints". In: *SIAM Journal on Optimization* 19.2 (2008), pp. 674–699.

[LAN10] James Luedtke, Shabbir Ahmed, and George L. Nemhauser. "An Integer Programming Approach for Linear Programs with Probabilistic Constraints". In: *Mathematical Programming* 122.2 (Apr. 2010), pp. 247–272. ISSN: 1436-4646. DOI: 10.1007/s10107-008-0247-4.

[LDB16] Miles Lubin, Yury Dvorkin, and Scott Backhaus. "A Robust Approach to Chance Constrained Optimal Power Flow With Renewable Generation". In: *IEEE Transactions on Power Systems* 31.5 (2016), pp. 3840–3849. DOI: 10.1109/TPWRS.2015.2499753.

[Lia+13] Jiaqi Liang et al. "Two-level dynamic stochastic optimal power flow control for power systems with intermittent renewable generation". In: *IEEE Transactions on Power Systems* 28.3 (2013), pp. 2670–2678.

[LL11] Javad Lavaei and Steven H Low. "Zero duality gap in optimal power flow problem". In: *IEEE Trans. Power Sys.* 27.1 (2011), pp. 92–107.

[LM15] Bowen Li and Johanna L. Mathieu. "Analytical reformulation of chance-constrained optimal power flow with uncertain load control". In: *2015 IEEE Eindhoven PowerTech*. 2015, pp. 1–6. DOI: 10.1109/PTC.2015.7232803.

[LQ19] Henry Lam and Huajie Qian. "Combating Conservativeness in Data-Driven Optimization under Uncertainty: A Solution Path Approach". In: *arXiv preprint arXiv:1909.06477* (2019).

[LS15] Alvaro Lorca and Xu Andy Sun. "Adaptive Robust Optimization With Dynamic Uncertainty Sets for Multi-Period Economic Dispatch Under Significant Wind". In: *IEEE Transactions on Power Systems* 30.4 (2015), pp. 1702–1713. DOI: 10.1109/TPWRS.2014.2357714.

[LS18] Álvaro Lorca and Xu Andy Sun. "The Adaptive Robust Multi-Period Alternating Current Optimal Power Flow Problem".
In: *IEEE Transactions on Power Systems* 33.2 (2018), pp. 1993–2003.
DOI: 10.1109/TPWRS.2017.2743348.

[LVM19] Bowen Li, Maria Vrakopoulou, and Johanna L. Mathieu.
"Chance Constrained Reserve Scheduling Using Uncertain Controllable Loads Part II: Analytical Reformulation".
In: *IEEE Transactions on Smart Grid* 10.2 (2019), pp. 1618–1625.
DOI: 10.1109/TSG.2017.2773603.

[MD18] Robert Mieth and Yury Dvorkin. "Data-Driven Distributionally Robust Optimal Power Flow for Distribution Systems".
In: *IEEE Control Systems Letters* 2.3 (2018), pp. 363–368.
DOI: 10.1109/LCSYS.2018.2836870.

[MGL14] Kostas Margellos, Paul Goulart, and John Lygeros.
"On the road between robust optimization and the scenario approach for chance constrained optimization problems".
In: *IEEE Transactions on Automatic Control* 59.8 (2014), pp. 2258–2263.

[Min+17] Hao Ming et al.
"Scenario-based economic dispatch with uncertain demand response".
In: *IEEE Transactions on Smart Grid* 10.2 (2017), pp. 1858–1868.

[MMD20] Ilyes Mezghani, Sidhant Misra, and Deepjyoti Deka.
"Stochastic AC optimal power flow: A data-driven approach".
In: *Electric Power Systems Research* 189 (2020), p. 106567.

[MR18] Daniel K Molzahn and Line A Roald. "Towards an AC optimal power flow algorithm with robust feasibility guarantees".
In: *2018 Power Systems Computation Conference (PSCC)*. IEEE. 2018, pp. 1–7.

[Müh+19] Tillmann Mühlpfordt et al.
"Chance-constrained ac optimal power flow: A polynomial chaos approach".
In: *IEEE Transactions on Power Systems* 34.6 (2019), pp. 4806–4816.

[MVH16] Jennifer F Marley, Maria Vrakopoulou, and Ian A Hiskens. "An AC-QP optimal power flow algorithm considering wind forecast uncertainty".
In: *2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*. IEEE. 2016, pp. 317–323.

[MW65] Bruce L Miller and Harvey M Wagner.
"Chance constrained programming with joint constraints".
In: *Operations Research* 13.6 (1965), pp. 930–945.

[NMK22] Nilay Noyan, Merve Meraklı, and Simge Küçükyavuz.
"Two-stage stochastic programming under multivariate risk constraints with an application to humanitarian relief network design".
In: *Mathematical Programming* 191.1 (2022), pp. 7–45.

[NS07] Arkadi Nemirovski and Alexander Shapiro. "Convex Approximations of Chance Constrained Programs". en. In: *SIAM Journal on Optimization* 17.4 (Jan. 2007), pp. 969–996. ISSN: 1052-6234, 1095-7189. DOI: 10.1137/050622328.

[Old+13] Frauke Oldewurtel et al. "Adaptively constrained stochastic model predictive control for closed-loop constraint satisfaction". In: *2013 American Control Conference*. IEEE. 2013, pp. 4674–4681.

[Old+15] Frauke Oldewurtel et al. "Adaptively constrained stochastic model predictive control applied to security constrained optimal power flow". In: *American Control Conference*. 2015, pp. 931–936.

[OW08] Masahiro Ono and Brian C Williams. "Iterative risk allocation: A new approach to robust model predictive control with a joint chance constraint". In: *2008 47th IEEE Conference on Decision and Control*. IEEE. 2008, pp. 3427–3432.

[Pan+15] Hrvoje Pandžić et al. "Toward cost-efficient and reliable unit commitment under uncertainty". In: *IEEE Transactions on Power Systems* 31.2 (2015), pp. 970–982.

[PAS09a] Bernardo K Pagnoncelli, Shabbir Ahmed, and Alexander Shapiro. "Computational study of a chance constrained portfolio selection problem". In: *Journal of Optimization Theory and Applications* 142.2 (2009), pp. 399–416.

[PAS09b] Bernardo K Pagnoncelli, Shabbir Ahmed, and Alexander Shapiro. "Sample average approximation method for chance constrained programming: theory and applications". In: *Journal of optimization theory and applications* 142.2 (2009), pp. 399–416.

[Pec] Pecan Street Inc. *Dataport: The world's largest energy data resource*. URL: https://dataport.pecanstreet.org/.

[Pen+19] Alejandra Pena-Ordieres et al. "DC Optimal Power Flow with Joint Chance Constraints". In: *arXiv preprint arXiv:1911.12439* (2019).

[Pen+20] Alejandra Pena-Ordieres et al. "Dc optimal power flow with joint chance constraints". In: *IEEE Transactions on Power Systems* (2020).

[PLW19] Alejandra Peña-Ordieres, James R Luedtke, and Andreas Wächter. "Solving Chance-Constrained Problems via a Smooth Sample-Based Nonlinear Approximation". In: *arXiv preprint arXiv:1905.07377* (2019).

[Por+22] Álvaro Porras et al. "Tight and Compact Sample Average Approximation for Joint Chance Constrained Optimal Power Flow". In: *arXiv preprint arXiv:2205.03370* (2022).

[Pré70] András Prékopa. "On probabilistic constrained programming". In: *Proceedings of the Princeton symposium on mathematical programming*. Vol. 113. Princeton, NJ. 1970, p. 138.

[PS18] Distributed Generation Photovoltaics and Energy Storage.
"IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces".
In: *IEEE Std* (2018), pp. 1547–2018.

[RA17] Line Roald and Göran Andersson. "Chance-constrained AC optimal power flow: Reformulations and efficient algorithms".
In: *IEEE Transactions on Power Systems* 33.3 (2017), pp. 2906–2918.

[Roa+13] Line Roald et al. "Analytical Reformulation of Security Constrained Optimal Power Flow with Probabilistic Constraints". en.
In: *2013 IEEE Grenoble Conference*. Grenoble, France: IEEE, June 2013, pp. 1–6.
ISBN: 978-1-4673-5669-5. DOI: 10.1109/PTC.2013.6652224.

[Roa+15a] Line Roald et al. "Optimal power flow with weighted chance constraints and general policies for generation control".
In: *2015 54th IEEE conference on decision and control (CDC)*. IEEE. 2015, pp. 6927–6933.

[Roa+15b] Line Roald et al. "Security constrained optimal power flow with distributionally robust chance constraints".
In: *arXiv preprint arXiv:1508.06061* (2015).

[Roa+17] Line Roald et al. "Corrective Control to Handle Forecast Uncertainty: A Chance Constrained Optimal Power Flow".
In: *IEEE Transactions on Power Systems* 32.2 (2017), pp. 1626–1637.
DOI: 10.1109/TPWRS.2016.2602805.

[Roa+23] Line A. Roald et al. "Power systems optimization under uncertainty: A review of methods and applications".
In: *Electric Power Systems Research* 214 (2023), p. 108725. ISSN: 0378-7796.
DOI: https://doi.org/10.1016/j.epsr.2022.108725.
URL: https://www.sciencedirect.com/science/article/pii/S0378779622007842.

[Rus02] Andrzej Ruszczyński. "Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra".
In: *Mathematical Programming* 93.2 (2002), pp. 195–215.

[San+16] Michael D Sankur et al. "A linearized power flow model for optimization in unbalanced distribution systems". In: *arXiv preprint arXiv:1606.04492* (2016).

[Sch+17] KP Schneider et al. "Analytic considerations and design basis for the IEEE distribution test feeders".
In: *IEEE Transactions on Power Systems* 33.3 (2017), pp. 3181–3188.

[SL13] Yongjia Song and James R Luedtke. "Branch-and-cut approaches for chance-constrained formulations of reliable network design problems".
In: *Mathematical Programming Computation* 5.4 (2013), pp. 397–432.

[Soa+18] Tiago Soares et al. “Active Distribution Grid Management Based on Robust AC Optimal Power Flow”.
In: *IEEE Transactions on Smart Grid* 9.6 (2018), pp. 6229–6241.
DOI: 10.1109/TSG.2017.2707065.

[Sum+14] Tyler Summers et al. “Stochastic optimal power flow based on convex approximations of chance constraints”.
In: *2014 Power Systems Computation Conference*. IEEE. 2014, pp. 1–7.

[Sum+15] Tyler Summers et al. “Stochastic optimal power flow based on conditional value at risk and distributional robustness”. In: *International Journal of Electrical Power & Energy Systems* 72 (2015), pp. 116–125.

[Swe+20] Conor Sweeney et al. “The future of forecasting for renewable energy”.
In: *Wiley Interdisciplinary Reviews: Energy and Environment* 9.2 (2020), e365.

[Van04] Dennis Harald Van Hessem. “Stochastic inequality constrained closed-loop model predictive control–with application to chemical process operation”.
In: (2004).

[Ven+18] Andreas Venzke et al.
“Convex Relaxations of Chance Constrained AC Optimal Power Flow”.
In: *IEEE Transactions on Power Systems* 33.3 (2018), pp. 2829–2841.
DOI: 10.1109/TPWRS.2017.2760699.

[VLM19] Maria Vrakopoulou, Bowen Li, and Johanna L. Mathieu.
“Chance Constrained Reserve Scheduling Using Uncertain Controllable Loads Part I: Formulation and Scenario-Based Analysis”.
In: *IEEE Transactions on Smart Grid* 10.2 (2019), pp. 1608–1617.
DOI: 10.1109/TSG.2017.2773627.

[Vra+13a] Maria Vrakopoulou et al.
“A Probabilistic Framework for Reserve Scheduling and ${\rm N}-1$ Security Assessment of Systems With High Wind Power Penetration”. en.
In: *IEEE Transactions on Power Systems* 28.4 (Nov. 2013), pp. 3885–3896.
ISSN: 0885-8950, 1558-0679. DOI: 10.1109/TPWRS.2013.2272546.

[Vra+13b] Maria Vrakopoulou et al.
“Probabilistic security-constrained AC optimal power flow”.
In: *2013 IEEE Grenoble Conference*. IEEE. 2013, pp. 1–6.

[VT11] Michael P. Vitus and Claire J. Tomlin.
“On feedback design and risk allocation in chance constrained control”.
In: *2011 50th IEEE Conference on Decision and Control and European Control Conference*. 2011, pp. 734–739. DOI: 10.1109/CDC.2011.6160721.

[WA08] Wei Wang and Shabbir Ahmed. “Sample average approximation of expected value constrained stochastic programs”.
In: *Operations Research Letters* 36.5 (2008), pp. 515–519.

[Wal45] A. Wald. "Sequential Tests of Statistical Hypotheses".
In: *The Annals of Mathematical Statistics* 16.2 (1945), pp. 117–186.
ISSN: 00034851. URL: http://www.jstor.org/stable/2235829.

[WB06] A. Wächter and L. T. Biegler.
"On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming".
In: *Math. Programming* 106.1 (2006), pp. 25–57.

[WRM18] Tillmann Weisser, Line Roald, and Sidhant Misra.
"Chance-constrained optimization for non-linear network flow problems".
In: *arXiv preprint arXiv:1803.02696* (2018).

[Wu+14] Hongyu Wu et al. "Chance-Constrained Day-Ahead Scheduling in Stochastic Power System Operation".
In: *IEEE Transactions on Power Systems* 29.4 (2014), pp. 1583–1591.
DOI: 10.1109/TPWRS.2013.2296438.

[XA17] Weijun Xie and Shabbir Ahmed. "Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation".
In: *IEEE Transactions on Power Systems* 33.2 (2017), pp. 1860–1867.

[XA18] Weijun Xie and Shabbir Ahmed. "On deterministic reformulations of distributionally robust joint chance constrained optimization problems".
In: *SIAM Journal on Optimization* 28.2 (2018), pp. 1151–1182.

[XAJ19] Weijun Xie, Shabbir Ahmed, and Ruiwei Jiang. "Optimized Bonferroni approximations of distributionally robust joint chance constraints".
In: *Mathematical Programming* (2019), pp. 1–34.

[ZDL17] Changhong Zhao, Emiliano Dall'Anese, and Steven H Low. "Convex relaxation of OPF in multiphase radial networks with delta connection".
In: *Proceedings of the 10th IREP*. 2017, pp. 0885–8950.

[ZG13] Chaoyue Zhao and Yongpei Guan.
"Unified stochastic and robust unit commitment".
In: *IEEE Transactions on Power Systems* 28.3 (2013), pp. 3353–3361.

[Zha+17] Yiling Zhang et al.
"Two-stage distributionally robust optimal power flow with flexible loads".
In: *2017 IEEE Manchester PowerTech*. 2017, pp. 1–6.
DOI: 10.1109/PTC.2017.7981202.

[ZKR13] Steve Zymler, Daniel Kuhn, and Berç Rustem. "Distributionally robust joint chance constraints with second-order moment information".
In: *Mathematical Programming* 137.1 (2013), pp. 167–198.

[ZL11] Hui Zhang and Pu Li. "Chance constrained programming for optimal power flow under uncertainty".
In: *IEEE Transactions on Power Systems* 26.4 (2011), pp. 2417–2424.

[ZMT11] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education". In: *IEEE Transactions on Power Systems* 26.1 (2011), pp. 12–19. DOI: 10.1109/TPWRS.2010.2051168.

[ZSM15] Yiling Zhang, Siqian Shen, and Johanna L Mathieu. "Data-driven optimization approaches for optimal power flow with uncertain reserves from load control". In: *2015 American Control Conference (ACC)*. IEEE. 2015, pp. 3013–3018.

[ZSM17] Yiling Zhang, Siqian Shen, and Johanna L. Mathieu. "Distributionally Robust Chance-Constrained Optimal Power Flow With Uncertain Renewables and Uncertain Reserves Provided by Loads". In: *IEEE Transactions on Power Systems* 32.2 (2017), pp. 1378–1388. DOI: 10.1109/TPWRS.2016.2572104.