POWER SYSTEM OPERATIONS WITH PROBABILISTIC GUARANTEES

A Dissertation

by

XINBO GENG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Le Xie
Committee Members,	Chanan Singh
	P. R. Kumar
	Steven L. Puller
Head of Department,	Miroslav M. Begovic

December 2019

Major Subject: Electrical Engineering

ABSTRACT

This study is motivated by the fact that uncertainties from deepening penetration of renewable energy resources have posed critical challenges to the secure and reliable operations of future electrical grids. Among various tools for decision making in uncertain environments, this study focuses on chance-constrained optimization, which provides explicit probabilistic guarantees on the feasibility of optimal solutions. Although quite a few methods have been proposed to solve chance-constrained optimization problems, there is a lack of comprehensive review and comparative analysis of the proposed methods. In this work, we provide a detailed tutorial on existing algorithms and a survey of major theoretical results of chance-constrained optimization theory. Data-driven methods, which are not constrained by any specific distributions of the underlying uncertainties, are of particular interest.

Built upon chance-constrained optimization, we propose a three-stage power system operation framework with probabilistic guarantees: (1) the optimal unit commitment in the *operational planning* stage; (2) the optimal reactive power dispatch to address the voltage security issue in the hours-ahead *adjustment period*; and (3) the secure and reliable power system operation under uncertainties in *real time*.

In the day-ahead operational planning stage, we propose a chance-constrained SCUC (c-SCUC) framework, which ensures that the risk of violating constraints is within an acceptable threshold. Using the scenario approach, c-SCUC is reformulated to the scenario-based SCUC (s-SCUC) problem. By choosing an appropriate number of scenarios, we provide theoretical guarantees on the posterior risk level of the solution to s-SCUC. Inspired by the latest progress of the scenario approach on non-convex problems, we demonstrate the structural properties of general scenario problems and analyze the specific characteristics of s-SCUC. Those characteristics were exploited to benefit the scalability and computational performance of s-SCUC.

In the adjustment period, this work first investigates the benefits of look-ahead coordination of both continuous-state and discrete-state reactive power support devices across multiple control

areas. The conventional static optimal reactive power dispatch is extended to a “moving-horizon” type formulation for the consideration of spatial and temporal variations. The optimal reactive power dispatch problem is further enhanced with chance constraints by considering the uncertainties from both renewables and contingencies. This chance-constrained optimal reactive power dispatch (c-ORPD) formulation offers system operators an effective tool to schedule voltage support devices such that the system voltage security can be ensured with quantifiable level of risk.

Security-constrained Economic Dispatch (SCED) lies at the center of real-time operation of power systems and modern electricity markets. It determines the most cost-efficient output levels of generators while keeping the real-time balance between supply and demand. In this study, we formulate and solve chance-constrained SCED (c-SCED), which ensures system security under uncertainties from renewables. The c-SCED problem also serves as a benchmark problem for a critical comparison of existing algorithms to solve chance-constrained optimization problems.

DEDICATION

To my parents.

ACKNOWLEDGMENTS

I would like to express my deepest gratitudes to my advisor Dr. Le Xie. He was a passionate researcher who led me into the world of academia, and a caring advisor who consoled me when I was deeply frustrated. I am also very thankful for the opportunities to visit UC Berkeley and MIT during his sabbatical leave. Those two semesters at the Simons Institute for the Theory of Computing and the Laboratory for Information & Decision Systems were extremely inspiring and intellectually stimulating.

My gratitudes are extended to my committee members for their insightful comments and valuable feedback on my dissertation. In addition, I would like to thank Dr. P. R. Kumar for offering two great courses and introducing me to the fascinating world of optimization theory and learning theory. I am also thankful to Dr. Chanan Singh and Dr. Steven Puller for offering the power system reliability course and the energy market course.

My PhD experience was unique and fruitful. In the past six years, I lived in six cities, went to four universities, and became “a traveling researcher”. I was very fortunate to meet lots of great mentors and collaborators: Prof. Srinivas Shakkottai (TAMU), Prof. P. R. Kumar (TAMU), Dr. Diran Obadina (ERCOT), Dr. Bryan Palmintier (NREL), Dheepak Krishnamurthy (NREL), Dr. Jinye Zhao (ISO-NE), Dr. Slava Maslennikov (ISO-NE), Dr. Eugene Litvinov (ISO-NE), Prof. Swati Gupta (GaTech), and Prof. Line Roald (WISC).

My sincere appreciations go to many friends who supported me through these difficult years and shared many precious moments together: M. Sadegh Modarresi, Abhishek Halder, Bowen Hua, Bainan Xia, Yuqi Zhou, Haotian Zhong, Sicheng Wang, Tong Huang, Ali Khodabakhsh, Guannan Qu, Anupam Thatte, Omar Urquidez, Meng Wu, Yang Chen, Xiaowen Lai, Yang Bai, Dian Chen, Yingzhong (Gary) Gu, Jian Li, Hao Ming, Cheng Qian, Bolun Xu, Xia Miao, Xi Liu, Bin Wang, Dan Wu, Kaiqing Zhang, Yushi Tan, Haoran Xu, Ang Li, Jianghui Yu, Chao Fang, Han Deng, Zhangxin, Sean, Haibo, Rayan, Athindra, Yuanyuan, Yida&Xin, Xiangtian, Xiaolin, Dongyao&Sai, Deliang, ..., and the “Geng Gang” at Simons.

I want to thank my parents, who love, educate and support me constantly and unconditionally, even though they are thousands of miles away from me.

Special thanks go to Xiao Fu, one of the most important people in my doctoral study. You amaze and inspire me on a daily basis, and help me keep positive during tough times. I am the luckiest.

The accomplishment of this dissertation, my personal and intellectual growth throughout these years, would be impossible without these people mentioned above.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Le Xie [advisor], Professor Chanan Singh and Professor P. R. Kumar of the Department of Electrical and Computer Engineering and Professor Steven L. Puller of the Department of Economics, Texas A&M University.

Part of Chapter 3 was completed with Dr. M. Sadegh Modarresi of Texas A&M University. Chapter 4 was completed with Dr. Diran Obadina of Electric Reliability Council of Texas (ERCOT). All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was in part supported by ERCOT, in part by Power Systems Engineering Research Center (PSERC), in part by NSF ECCS-1150944, ECCS-1546682, ECCS-1839616 and OAC-1934675.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION[1]

Real-time decision making in the presence of uncertainties is a classical problem that arises in many contexts. In the context of electric energy systems, a pivotal challenge is how to operate a power grid with an increasing amount of supply and demand uncertainties. The unique characteristics of such operational problem include (1) the underlying distribution of uncertainties is largely unknown (e.g. the forecast error of demand response); (2) decisions have to be made in a timely manner (e.g. a dispatch order needs to be given by 5 minutes prior to the real-time); and (3) there is a strong desire to know the risk that the system is exposed to after a decision is made (e.g. the risk of violating transmission constraints after the real-time market clears). In response to these challenges, a class of optimization problems named "chance-constrained optimization" has received increasing attention in both operations research and practical engineering communities.

Although quite a few methods have been proposed to solve chance-constrained optimization problems, there is a lack of comprehensive review and comparative analysis of the proposed methods. This dissertation first provides a comprehensive review of existing methods to chance-constrained optimization in Chapter 2: (1) scenario approach; (2) sample average approximation; and (3) robust optimization based methods. Data-driven methods, which are not constrained by any particular distributions of the underlying uncertainties, are of particular interest. Many methods reviewed in Chapter 2 are implemented in the Matlab Toolbox ConvertChanceConstraint (ccc).

Built upon chance-constrained optimization, this dissertation proposes a three-stage power system operation framework with probabilistic guarantees. The three-stage framework also outlines the remainder of this dissertation: (1) Chapter 3 examines the optimal commitment in the *operational planning* stage; (2) Chapter 4 studies the voltage security issue in the *adjustment period*; and (3) Chapter 5 investigates the secure and reliable power system *real-time* operation under uncertainties.

Keeping the balance between supply and demand is a fundamental task in power system op-

[1] Adapted with permission from [1–4].

erational planning practices. This task becomes particularly challenging due to the deepening penetration of renewable energy resources, which induces a significant amount of uncertainties. In Chapter 3, we propose a chance-constrained Security-constrained Unit Commitment (c-SCUC) framework to tackle challenges from uncertainties of renewables. The proposed c-SCUC framework seeks cost-efficient scheduling of generators while ensuring operation constraints with guaranteed probability. Inspired by the latest progress of the scenario approach on non-convex problems, we demonstrate the structural properties of general scenario problems and reveal the salient structural properties of c-SCUC, which could significantly reduce the sample complexity required by the scenario approach and speed up computation.

The uncertainties from deepening penetration of renewable energy resources have already shown to impact not only the market operations, but also the physical operations in large power systems. It is demonstrated that deterministic modeling of wind would lead to voltage insecurity in the reality where wind fluctuates. This could render deterministic control of reactive power ineffective. As an alternative, Chapter 4 proposes a chance-constrained formulation of optimal reactive power dispatch which considers the uncertainties from both renewables and contingencies. This formulation of a chance constrained optimal reactive power dispatch (cc-ORPD) offers system operators an effective tool to schedule voltage support devices such that the system voltage security can be ensured with quantifiable level of risk. The cc-ORPD problem is a Mixed-Integer Non-Linear Programming (MINLP) problem with a joint chance constraint and is extremely challenging to solve. Using sample average approximation and linearized power flow equations, the original cc-ORPD problem is approximated as a Mixed-Integer Linear Programming (MILP) problem.

Security-constrained Economic Dispatch (SCED) lies at the center of modern electricity markets and short-term power system operations. It determines the most cost-efficient output levels of generators while keeping the real-time balance between supply and demand. Chapter 5 extends SCED using chance constraints (cc-SCED), which ensures system security under uncertainties from renewables. The proposed cc-SCED problem is solved via different algorithms reviewed in Chapter 2. Chapter 5 also presents a critical comparison of existing methods to solve chance-constrained optimization problems based on numerical simulations.

The notations in this dissertation are standard. All vectors are in the real field $\mathbf{R}$. We use $\mathbf{1}$ to represent an all-one vector of appropriate size. The transpose of a vector a is $a^\intercal$. The element-wise multiplication of the same-size vectors a and b is denoted by $a \circ b$. For instance, $[a_1; a_2] \circ [b_1; b_2] = [a_1 b_1; a_2 b_2]$. Sets are in calligraphy fonts, e.g. $\mathcal{S}$. The cardinality of a set $\mathcal{S}$ is $|\mathcal{S}|$. Removal of element i from set $\mathcal{N}$ is represented by $\mathcal{N} - i$. The essential supremum is $\operatorname{ess\,sup}$.

2. A SCHEMATIC OVERVIEW OF CHANCE-CONSTRAINED OPTIMIZATION [1]

Chance-constrained optimization (CCO) is an important tool for decision making in uncertain environments. Since its birth in 1950s, CCO has found many successful applications in various fields, e.g. economics [5], control theory [6], chemical process [7, 8], water management [9] and recently in machine learning [10–15]. Chance-constrained optimization plays a particularly important role in the context of electric power systems [16, 17], applications of CCO can be found in various time-scales of power system operations and at different levels of the system.

The first chance-constrained program was formulated in [18], then was extensively studied in the following 50 years, e.g. [11, 19–26]. Previously, most methods to solve CCO problems deal with specific families of distributions, such as log-concave distributions [26, 27]. Many novel methods appeared in the past ten years, e.g. scenario approach [6], sample average approximation [28, 29] and convex approximation [30]. Most of them are generic methods that are not limited to specific distribution families and require very limited knowledge about the uncertainties. In spite of many successful applications of these methods in various fields, there is a lack of comprehensive review and a critical comparison.

The objective of this chapter is to provide a comprehensive and up-to-date review of mathematical formulations, computational algorithms, and engineering implications of chance-constrained optimization in the context of electric power systems. In particular, this chapter focuses on the data-driven approaches to solving chance-constrained optimization *without* knowing the underlying distribution of uncertainties. This chapter also briefly mentions some critical results of an alternative approach, i.e. deriving equivalent forms of chance-constrained optimization problems for specific distributions. A more general class of problems, i.e. distributionally robust optimization or ambiguous chance constraint, is beyond the scope of this dissertation.

The main contributions of this chapter are twofold:

1. We provide a detailed tutorial on existing algorithms to solve chance-constrained programs

[1] Parts of this chapter are reprinted with permission from [1].

and a survey of major theoretical results. To the best of our knowledge, there is no such review available in the literature;

2. We implement most of the reviewed methods and develop an open-source Matlab toolbox (ConvertChanceConstraint), which is available on Github [2].

2.1 Chance-constrained Optimization

2.1.1 Introduction to Chance-constrained Optimization

We study the following chance-constrained optimization problem throughout this chapter:

$$
\begin{aligned}
\text{(CCO): } \min_{x} \;& c^{\intercal} x && \text{(2.1a)}\\
\text{s.t. } & \mathbb{P}_{\xi}\Big(f(x,\xi) \leq 0\Big) \geq 1-\epsilon && \text{(2.1b)}\\
& x \in \mathcal{X} && \text{(2.1c)}
\end{aligned}
$$

where $x \in \mathbf{R}^n$ is the decision variable and random vector $\xi \in \mathbf{R}^d$ is the source of uncertainties. Without loss of generality [3], we assume the objective function is linear in x and does not depend on ξ. Constraint (2.1b) is the *chance constraint* (or *probabilistic constraint*), it requires the inner constraint $f(x,\xi) \leq 0$ to be satisfied with high probability $1-\epsilon$. The inner constraint $f(x,\xi) : \mathbf{R}^n \times \mathbf{R}^d \to \mathbf{R}^m$ consists of m individual constraints, i.e. $f(x,\xi) = \big(f_1(x,\xi), f_2(x,\xi), \cdots, f_m(x,\xi)\big)$. The set $\mathcal{X}$ stands for the deterministic constraints. Parameter ϵ is called the *violation probability* of (CCO). Notice that $f(x,\xi)$ is random due to ξ, the probability $\mathbb{P}$ is taken with respect to ξ. Sometimes the probability is denoted by $\mathbb{P}_{\xi}$ to avoid confusion.

It is worth mentioning that CCO is closely related with the theory of risk management. For example, an individual chance constraint $\mathbb{P}(f_i(x,\xi) \leq 0) \geq 1-\epsilon_i$ can be equivalently interpreted as a constraint on the value at risk $\mathrm{VaR}(f_i(x,\xi); 1-\epsilon_i) \leq 0$. This connection can be directly seen from the definition.

[2] github.com/xb00dx/ConvertChanceConstraint-ccc

[3] Using the epigraph formulation as mentioned in [31, 32].

Definition 1 (Value at Risk)**.** Value at risk (VaR) of random variable ζ at level $1-\epsilon$ is defined as

$$\mathrm{VaR}(\zeta; 1-\epsilon) := \inf\big\{\gamma : \mathbb{P}(\zeta \leq \gamma) \geq 1-\epsilon\big\} \tag{2.2}$$

More details about this can be found in Section 2.6.3.1, [33, 34] and references therein.

CCO is closely related with two other major tools for decision making with uncertainties: stochastic programming and robust optimization. The idea of sample average approximation, which originated from stochastic programming, can be applied on chance-constrained programs (Section 2.5). Section 2.6 demonstrates the connection between robust optimization and CCO.

2.1.2 Joint and Individual Chance Constraints

Constraint (2.1b) is called a *joint chance constraint* because of its multiple inner constraints [27], i.e.

$$\mathbb{P}\Big(f_1(x,\xi) \leq 0, f_2(x,\xi) \leq 0, \cdots, f_m(x,\xi) \leq 0\Big) \geq 1-\epsilon \tag{2.3}$$

Alternatively, each one of the following m constraints is called an *individual chance constraint*:

$$\mathbb{P}\Big(f_i(x,\xi) \leq 0\Big) \leq 1-\epsilon_i,\ i = 1, 2, \cdots, m \tag{2.4}$$

Joint chance constraints typically have more modeling power since an individual chance constraint is a special case ($m = 1$) of a joint chance constraint. But individual chance constraints are relatively easier to deal with (see Section 2.2.2 and 2.6.3). There are several ways to convert individual and joint chance constraints between each other.

First, a joint chance constraint can be written as a set of individual chance constraints using Bonferroni inequality or Boole's inequality. Notice (2.3) can be represented as

$$\mathbb{P}_\xi\Big(\cup_{i=1}^m \big\{f_i(x,\xi) \geq 0\big\}\Big) \leq \epsilon. \tag{2.5}$$

Since $\mathbb{P}_\xi(\cup_{i=1}^m \{f_i(x,\xi) \geq 0\}) \leq \sum_{i=1}^m \mathbb{P}_\xi(\{f_i(x,\xi) \geq 0\})$, if $\sum_{i=1}^m \epsilon_i \leq \epsilon$, then any feasible solution to (2.4) is also feasible to (2.3). In other words, (2.4) is a *safe approximation* (see Definition 11) to (2.3) when $\sum_{i=1}^{m} \epsilon_i \leq \epsilon$. With appropriate $\{\epsilon_i\}_{i=1}^{m}$, (2.4) could be a good approximation of (2.3). However, it is usually difficult to find such $\{\epsilon_i\}_{i=1}^{m}$. Some other issues of this approach are discussed in Section 2.6.4.1.

Alternatively, a joint chance constraint (2.3) is equivalent to the following individual chance constraint:

$$\mathbb{P}_{\xi}\big(\overline{f}(x,\xi) \leq 0\big) \geq 1-\epsilon \tag{2.6}$$

where $\overline{f}(x,\xi) : \mathbf{R}^n \times \mathbf{R}^d \rightarrow \mathbf{R}$ is the pointwise maximum of functions $\{f_i(x,\xi)\}_{i=1}^{m}$ over x and ξ, i.e.

$$\overline{f}(x,\xi) := \max\Big\{f_1(x,\xi), f_2(x,\xi), \cdots, f_m(x,\xi)\Big\}. \tag{2.7}$$

However, converting $\{f_i(x,\xi)\}_{i=1}^{m}$ to $\overline{f}(x,\xi)$ could lose nice structures of the original constraint $f(x,\xi) \leq 0$ and cause more difficulties.

In this dissertation, we focus on the chance-constrained optimization problems with a *joint* chance constraint.

2.1.3 Critical Definitions and Assumptions

Theoretical results in the following sections are based on the critical definitions and assumptions below.

Definition 2 (Violation Probability)**.** Let $x^\diamond$ denote a candidate solution to (CCO), its violation probability is defined as

$$\mathbb{V}(x^\diamond) := \mathbb{P}_{\xi}\Big(f(x^\diamond,\xi) \geq 0\Big) \tag{2.8}$$

Definition 3. $x^\diamond$ is a *feasible* solution for (CCO) if $x^\diamond \in \mathcal{X}$ and $\mathbb{V}(x^\diamond) \leq \epsilon$. Let $\mathcal{F}_\epsilon$ denote the set of feasible solutions to the chance constraint (2.1b),

$$\mathcal{F}_\epsilon := \{x \in \mathbf{R}^n : \mathbb{V}(x) \leq \epsilon\} = \{x \in \mathbf{R}^n : \mathbb{P}_{\xi}\Big(f(x,\xi) \leq 0\Big) \geq 1-\epsilon\}$$

then $x^\diamond$ is *feasible* to (CCO) if $x^\diamond \in \mathcal{X} \cap \mathcal{F}_\epsilon$.

Although (CCO) seeks optimal solutions under uncertainties, it is a *deterministic* optimization problem. To better see this, (CCO) can be equivalently written as $\min_{x\in\mathcal{X}}\ c^\intercal x,$ s.t. $\mathbb{V}(x) \leq \epsilon$ or $\min_{x\in\mathcal{X}\cap\mathcal{F}} c^\intercal x$.

Definition 4. Let $o^\star$ denote the optimal objective value of (CCO). For simplicity, we define $o^\star = +\infty$ when (CCO) is infeasible and $o^\star = -\infty$ when (CCO) is unbounded. Let $x^\star$ denote the optimal solution to (CCO) if exists, and $o^\star = c^\intercal x^\star$.

Definition 5. We say a candidate solution $x^\diamond$ is *conservative* if $\mathbb{V}(x^\diamond) \ll \epsilon$ or $c^\intercal x^\diamond \gg o^\star$.

Most existing theoretical results on (CCO) are built upon the following two assumptions.

Assumption 1. *Let Ξ denote the support of random variable ξ, the distribution $\xi \sim \Xi$ exists and is fixed.*

Assumption 1 only assumes the existence of an underlying distribution, but we do not necessarily need to know it to solve (CCO). Removing assumption 1 leads to a more general class of problem named *distributionally robust optimization* or *ambiguous chance constraints*. Section 2.2.4 discusses cases with Assumption 1 removed.

Assumption 2. *(1) Function $f(x, \xi)$ is convex in x for every instance of ξ, and (2) the deterministic constraints define a convex set $\mathcal{X}$.*

The convexity assumption above makes it possible to develop theories on (CCO). However, the feasible region $\mathcal{F}_\epsilon$ of (CCO) is often non-convex even under Assumption 2. More details are presented in Section 2.2.1 and 2.2.2.

2.2 Fundamental Properties

2.2.1 Hardness

Although CCO is an important and useful tool for decision making under uncertainties, it is very difficult to solve in general. Major difficulties come from two aspects:

(D1) It is difficult to check the feasibility of a candidate solution $x^\diamond$. Namely, it is intractable to evaluate the probability $\mathbb{P}_\xi\big(f(x^\diamond, \xi) \leq 0\big)$ with high accuracy. More specifically, calculating probability involves multivariate integration, which is NP-Hard [35]. The only general method might be Monte-Carlo simulation, but it can be computationally intractable due to the curse of dimensionality.

(D2) It is difficult to find the optimal solution $x^\star$ and $o^\star$ to (CCO). Even with the convexity assumption (Assumption 2), the feasible region $\mathcal{F}_\epsilon$ of (CCO) is often non-convex except a few special cases. For example, Section 2.2.3 shows the feasible region of (CCO) with separable chance constraints is a union of cones, which is non-convex in general. Although researchers have proved various sufficient conditions on the convexity of (CCO), it remains challenging to solve (CCO) because of difficulty (D1). Most of times, however, we are agnostic about the properties of the feasible region $\mathcal{F}_\epsilon$.

Despite that fact that Assumptions 1 and 2 greatly simplify the problem and make theoretical analysis on (CCO) possible, (D1) and (D2) still exist and pose great challenges to solve (CCO).

Theorem 1 ([36, 37])**.** *(CCO) is strongly NP-Hard.*

Theorem 2 ([38])**.** *Unless P $=$ NP, it is impossible to obtain a polynomial time algorithm for (CCO) with a constant approximation ratio.*

Theorem 1 formalizes the hardness results of solving (CCO), Theorem 2 further demonstrates it is also difficult to obtain approximate solutions to (CCO): any polynomial algorithm is not able to find a solution x^* (with $o^* = c^\intercal x^*$) such that $|o^*/o^\star|$ is bounded by a constant C. In other words, any polynomial-time algorithm could be arbitrarily worse.

2.2.2 Special Cases

Although (CCO) is NP-Hard to solve in general, there are several special cases in which solving (CCO) is relatively easy. The most well-known special case is (2.9), which was first proved in [21].

$$
\begin{aligned}
&\min_{x\in\mathcal{X}}\ c^\intercal x && \text{(2.9a)}\\
&\text{s.t. } \mathbb{P}(a^\intercal x + b^\intercal \xi + \xi^\intercal D x \le e) \ge 1-\epsilon && \text{(2.9b)}
\end{aligned}
$$

Parameters $a \in \mathbf{R}^n$,$b \in \mathbf{R}^d$, $D \in \mathbf{R}^{d\times n}$ and $e \in \mathbf{R}$ are fixed coefficients. $\xi \sim \mathcal{N}(\mu, \Sigma)$ is a multivariate Gaussian random vector with mean μ and covariance Σ. Notice that (2.9b) is an *individual* chance constraint with multivariate Gaussian coefficients. Let $\Phi(\cdot)^{-1}$ denote the inverse cumulative distribution function (CDF) function of a standard normal distribution. It is easy to show that if $\epsilon \le 1/2$, (2.9) is equivalent to (2.10), which is a second order cone program (SOCP) and can be solved efficiently.

$$
\begin{aligned}
&\min_{x\in\mathcal{X}}\ c^\intercal x && \text{(2.10a)}\\
&\text{s.t. } e - b^\intercal\mu - (a + D^\intercal\mu)^\intercal x \ge \\
&\qquad \Phi^{-1}(1-\epsilon)\sqrt{(b+Dx)^\intercal\Sigma(b+Dx)} && \text{(2.10b)}
\end{aligned}
$$

(2.10) also shows the possibility of deriving equivalent reformulations of chance-constrained optimization, many analytical methods to solve chance-constrained optimization are built on this observation.

The case of log-concave distribution [26, 39, 40] is another famous special case where chance constraint is convex. There are many other sufficient conditions on the convexity of chance constraints, e.g. [41–45].

2.2.3 Feasible Region

A chance-constrained program with only right hand side uncertainties (2.11) is considered in this section. With this example, we provide deeper understandings on the non-convexity of (CCO).

$$\min_{x \in \mathcal{X}} \; c^\intercal x \tag{2.11a}$$

$$\text{s.t. } \mathbb{P}(f(x) \leq \zeta) \geq p \tag{2.11b}$$

In (2.11b), the inner function $f(x) : \mathbf{R}^n \to \mathbf{R}^m$ is deterministic. The only uncertainty is the right-hand side value, represented by a random vector $\zeta \in \mathbf{R}^m$. Chance constraints like (2.11b) are also named *separable* chance constraints (or probabilistic constraints) since the deterministic and random parts are separated. We replace $1 - \epsilon$ with p in (2.11b) to follow the convention in the existing literature.

Definition 6 (p-efficient points [46])**.** Let $p \in (0, 1)$, a point $v \in \mathbf{R}^m$ is called a *p-efficient point* of the probability function $\mathbb{P}_\zeta(\zeta \leq z)$, if $\mathbb{P}_\zeta(\zeta \leq v) \geq p$ and there is no $z \leq v$, and $z \neq v$ such that $\mathbb{P}_\zeta(\zeta \leq z) \geq p$.

Theorem 3 ([46] [26])**.** *Let $\mathcal{E}$ be the index set of p-efficient points v^i, $i \in \mathcal{E}$. Let $\mathcal{F}_p := \{x \in \mathbf{R}^n : \mathbb{P}_\zeta(f(x) \leq \zeta) \geq p\}$ denote the feasible region of (2.11b), then it holds that*

$$\mathcal{F}_p = \cup_{i \in \mathcal{E}} K_i \tag{2.12}$$

where each cone K_i is defined as $K_i := v^i + \mathbf{R}^m_+$, $i \in \mathcal{E}$.

Theorem 3 shows the geometric properties of (CCO). The finite union of convex sets need not to be convex, therefore the feasible region of (CCO) is generally non-convex.

Remark 1. Many methods to solve (CCO) (e.g. [24, 47, 48]) start with a partial or complete enumeration of p-efficient points. However, the number of p-efficient points could be astronomic or even infinite. See [26, 46] and references therein for the finiteness results of p-efficient points and complete theories and algorithms on p-efficient points.

2.2.4 Ambiguous Chance Constraints

Ambiguous chance constraint is a generalization of chance constraints,

$$\mathbb{P}_{\xi\sim P}\Big(f(x,\xi)\leq 0\Big)\geq 1-\epsilon,\ \forall P\in\mathcal{P}. \tag{2.13}$$

It requires the inner chance constraint $f(x,\xi)\leq 0$ holds with probability $1-\epsilon$ for any distribution P belonging to a set of pre-defined distributions $\mathcal{P}$.

Ambiguous chance constraints are particularly useful in the cases where only partial knowledge on the distribution P is available, e.g. we know only that P belongs a given family of $\mathcal{P}$. However, it is generally more difficult to solve ambiguous chance constraints, and the theoretical results rely on different assumptions of uncertainties. This chapter only reviews solutions to CCO, studies on ambiguous chance constraints are beyond the scope of this dissertation.

2.3 An Overview of Solutions to CCO

This chapter concentrates on solutions to (CCO) with the following properties: (i) dealing with both difficulties (D1) and (D2) mentioned in Section 2.2.1; (ii) utilizing information from data (only) without making suspicious assumptions on the distribution of uncertainties; and (iii) possessing rigorous guarantees on the feasibility and optimality of returned solutions. Section 2.3.1-2.3.3 explain these three properties in detail. Section 2.3.4 provides an overview of methods with the properties above.

2.3.1 Classification of Solutions

Existing methods on (CCO) can be roughly classified into four categories [49]:

(C1) When both difficulties (D1) and (D2) in Section 2.2.1 are absent, (CCO) is convex and the probability $\mathbb{P}(f(x,\xi)\leq 0)$ is easy to calculate. The only known case in this category is the individual chance constraint (2.9) with Gaussian distributions, which might be the only special case of (CCO) that can be easily solved;

(C2) When (D1) is absent but (D2) is present, it is relatively easy to calculate $\mathbb{P}(f(x,\xi)\leq 0)$ (e.g.

finite distributions with not too many realizations). As shown in Theorem 3, the feasible region of (CCO) could be non-convex and solutions typically rely on integer programming and global optimization [49];

(C3) When (D1) is present but (D2) is absent, (CCO) is proved to be convex but remains difficult to solve because of the difficulty (D1) in calculating probabilities. This case often requires approximating the probability via simulations or specific assumptions. All examples mentioned in Section 2.2.2 except (2.9) belong to this category.

(C4) When both difficulties (D1) and (D2) are present, it is almost impossible to find the optimal solution $x^\star$ and $o^\star$. All existing methods attempt to obtain approximate solutions or sub-optimal solutions and construct upper and lower bounds on the true objective value $o^\star$ of (CCO).

Methods associated with (C1)-(C3) are briefly mentioned in Section 2.2, the remaining part of this chapter presents more general and powerful methods in category (C4).

2.3.2 Prior Knowledge

In order to solve (CCO), a reasonable amount of prior knowledge on the underlying distribution $\xi \sim \Xi$ is necessary. Figure 2.1 illustrates three categories of prior knowledge:

(K1) We know the exact distribution $\xi \sim \Xi$ thus have *complete knowledge* on the underlying distribution;

(K2) We know partially on the distribution (e.g. multivariate Gaussian distribution with bounded mean and variance) and thus have *partial knowledge*;

(K3) We have a finite dataset $\{\xi^i\}_{i=1}^N$, this is another case of *partial knowledge*.

It can be seen that prior information in (K2) is a strict subset of (K1), also by sampling we can construct a dataset in (K3) from the exact distribution in (K1). It seems (K1) is the best starting point to solve (CCO). However, *probability distributions are not known in practice, they are just*

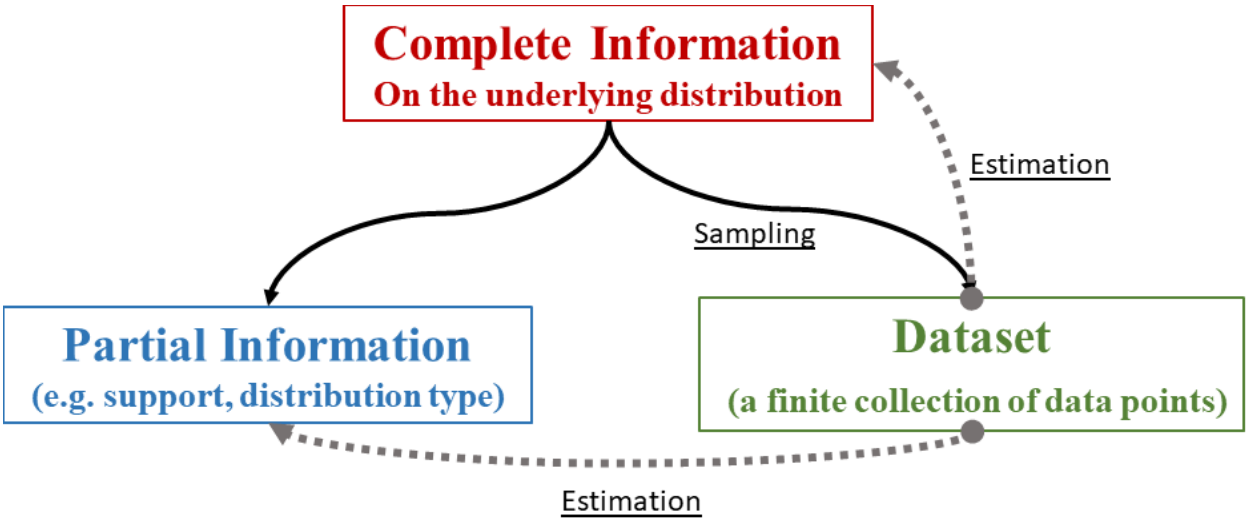


Figure 2.1: Different Knowledge Levels to Solve (CCO), reprinted with permission from [1].

models of reality and exist only in our imagination. What exists in reality is *data*. Therefore (K3) is the most practical case and becomes the focus of this dissertation. Almost all the data-driven methods to solve (CCO) are based on the following assumption.

Assumption 3. *The samples (scenarios)* ξ^i *(*$i = 1, 2, \cdots, N$*) in the* dataset $\{\xi^i\}_{i=1}^N$ *are independent and identically distributed (i.i.d.).*

2.3.3 Theoretical Guarantees

In this chapter, we concentrate on the theoretical aspects of the reviewed methods. In particular, we pay special attention to *feasibility guarantees* and *optimality guarantees*.

Given a candidate solution $x^\diamond$ to (CCO), the first and possibly most important thing is to check its *feasibility*, i.e. if $\mathbb{V}(x^\diamond) \leq \epsilon$. Although (D1) demonstrates the difficulty in calculating $\mathbb{V}(x^\diamond)$ with high accuracy, there are various feasibility guarantees that either estimate $\mathbb{V}(x^\diamond)$ or provide upper bound on $\mathbb{V}(x^\diamond)$. The feasibility results can be classified into two categories: a-priori and a-posteriori guarantees. The *a-priori* ones typically provide prior conditions on (CCO) and the dataset $\{\xi^i\}_{i=1}^N$, the feasibility of the corresponding solution $x^\diamond$ is guaranteed *before* obtaining $x^\diamond$. Examples of this type include Corollary 1, Theorem 6,13 and 11. As the name suggests, the *a-posteriori* guarantees make effects *after* obtaining $x^\diamond$. The a-posteriori guarantees are constructed based on the observations of the structural features associated with $x^\diamond$. Examples include Theorem 7 and Proposition 1.

Given a candidate solution $x^\diamond$ and the associated objective value $o^\diamond = c^\intercal x^\diamond$, another important question to be answered is about the *optimality* gap $|o^\diamond - o^\star|$. Although finding $o^\star$ is often an impossible mission because of difficulty (D2), bounding from below on $o^\star$ is relatively easier. Sections 2.4.5 and 2.5.4 dedicate to algorithms of constructing lower bounds $\underline{o} \leq o^\star$.

2.3.4 A Schematic Overview

A schematic overview of solutions to (CCO) and their relationships are presented in Figure 2.2. Akin methods are plotted in similar colors, and links among two circles indicate the connection of the two methods. The tree-like structure of Figure 2.2 illustrates the hierarchical relationship of the reviewed methods. Key references of each method are also provided. The root node of Figure 2.2 is the "ambiguous chance constraint" or distributionally robust optimization (DRO), which is the parent node of "chance-constrained optimization". This indicates that DRO contains CCO as a special case. Similarly, for example, node "scenario approach" has three child nodes "prior", "posterior" and "sampling and discarding", this indicates the scenario approach has three major variations.

As shown in Figure 2.2, CCO is a special case of ambiguous chance constraints where the set of distributions $\mathcal{P}$ is a singleton (Section 2.2.4). Therefore methods to solve ambiguous chance constraints can be applied on chance constraints as well. The methods and algorithms to solve CCO are the main focus of this chapter, we will briefly mention the connection if some methods are related with ambiguous chance constraints.

Figure 2.2 also outlines this chapter, which dedicates to a review and tutorial on chance-constrained optimization. We summarize key results on the basic properties (Section 2.2), three main approaches to solving chance-constrained optimization problems, scenario approach (Section 2.4), sample average approximation (Section 2.5) and robust optimization (RO) based methods (Section 2.6).

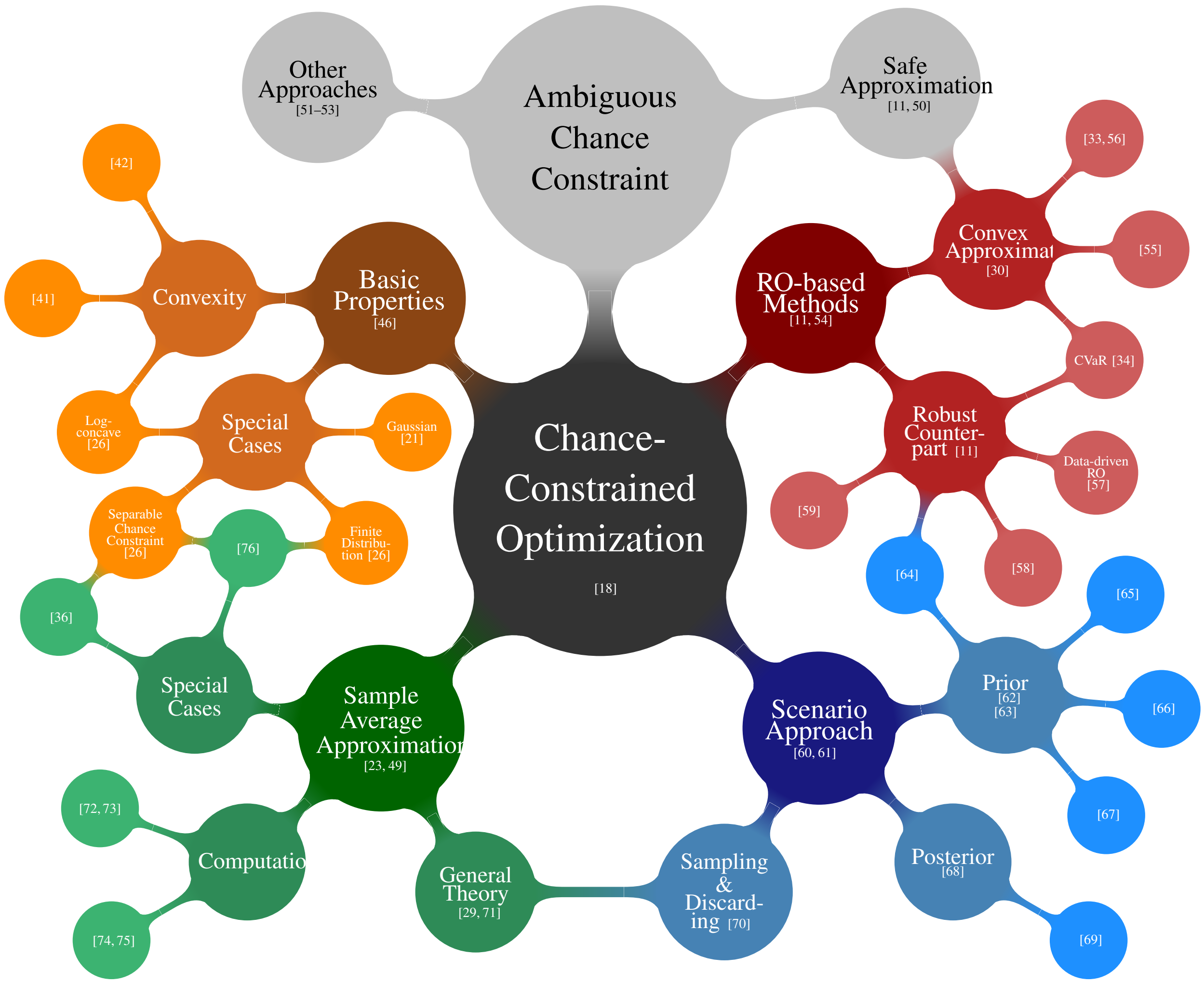


Figure 2.2: A Schematic Overview of Existing Methods and Algorithms to Solve Chance-constrained Optimization Problems, reprinted with permission from [1].

2.4 Scenario Approach

2.4.1 Introduction to Scenario Approach

Scenario approach utilizes a dataset with N scenarios $\{\xi^i\}_{i=1}^N$ to approximate the chance-constrained program (2.1) and obtains the following *scenario problem* $(\text{SP})_N$:

$$(\text{SP})_N\colon \min_{x\in\mathcal{X}} \; c^\intercal x \tag{2.14a}$$

$$\text{s.t. } f(x,\xi^1) \le 0, \cdots, f(x,\xi^N) \le 0 \tag{2.14b}$$

SP_N seeks the optimal solution x_N^* which is feasible for all N scenarios. The scenario approach is a very simple yet powerful method. The most attractive feature of the scenario approach is its generality. It requires nothing except the convexity of constraints $f(x, \xi)$ and $\mathcal{X}$. It is purely data-driven and makes no assumption on the underlying distribution.

Remark 2. SP_N is a random program. Both its optimal objective value o_N^* and optimal solution x_N^* depend on the random samples $\{\xi^i\}_{i=1}^N$, therefore they are random variables. In consequence, $\mathbb{V}(x_N^*)$ is also a random variable. Let $\mathcal{N} := \{1, 2, \cdots, N\}$ denote the index set of scenarios. The optimal objective value of SP_N is denoted by $o^*(\mathcal{N})$ to emphasize its dependence on the random samples.

Theoretical results of the scenario approach are built upon the following assumption in addition to Assumptions 1, 2 and 3.

Assumption 4 (Feasibility and Uniqueness [62]). *Every scenario problem (SP)$_N$ is feasible, and its feasibility region has a non-empty interior. Moreover, the optimal solution x_N^* of (SP)$_N$ exists and is unique.*

If there exist multiple optimal solutions, the tie-break rules in [60] can be applied to obtain a unique solution.

Remark 3 (Sample Complexity N). We first provide some intuition on the scenario approach. When solving (SP)$_N$ with a very large number of scenarios, the solution x_N^* will be robust to almost every realization of ξ, thus the violation probability goes to zero. Although x_N^* is a feasible solution to (CCO) as $N \to +\infty$, it is overly conservative because $\mathbb{V}(x^*) \approx 0 \ll \epsilon$. On the other hand, using too few scenarios for SP_N might result in infeasible solutions x_N^* to (CCO). Notice that N is the only tuning parameter in the scenario approach, the most important question in the scenario approach theory is: *what is the right sample complexity N?* Namely, what is the smallest N such that $\mathbb{V}(x_N^*) \leq \epsilon$ (with high probability)? Rigorous answers to the sample complexity question are built upon the structural properties of SP_N.

2.4.2 Structural Properties of the Scenario Problem

Among N scenarios in the dataset $\{\xi^i\}_{i=1}^N$, there are some important scenarios having direct impacts on the optimal solution x_N^*.

Definition 7 (Support Scenario [60])**.** Scenario ξ^i is a *support scenario* for (SP)$_N$ if its removal changes the solution of (SP)$_N$. The set of *support scenarios* of (SP_N) is denoted by $\mathcal{S}$.

Theorem 4 ([60,63])**.** *Under Assumption 2, the number of support scenarios in SP$_N$ is at most n, i.e. $|\mathcal{S}| \leq n$.*

Theorem 4 is built upon Helly's theorem and Radon's theorem [77] in convex analysis. For non-convex problems, the number of support scenarios could be greater than the number of decision variables n. An example for non-convex problems is provided in [69].

Definition 8 (Fully-supported Problem [62])**.** A scenario problem SP$_N$ with $N \geq n$ is *fully-supported* if the number of support scenarios is exactly n. Scenario problems with $|\mathcal{S}| < n$ are referred as *non-fully-supported* problems.

Definition 9 (Non-degenerate Problem [62, 63])**.** Problem SP$_N$ is said to be *non-degenerate*, if $o^*(\mathcal{N}) = o^*(\mathcal{S})$. In other words, SP$_N$ is *non-degenerate* if the solution of (SP)$_N$ with all scenarios in place coincides with the solution to the program with only the support scenarios are kept.

2.4.3 A-priori Feasibility Guarantees

Obtaining a-priori feasibility guarantees on the solution x_N^* to SP$_N$ typically involves the following three steps:

1. Exploring the problem structure of SP$_N$ and obtain an upper bound $\overline{h}$ on the number of support scenarios;

2. Choosing a good sample complexity $N(\epsilon, \beta, \overline{h})$ using Corollary 1, Theorem 6 or Remark 4;

3. Solving the scenario problem SP$_N$ and obtain x_N^* and o_N^*.

Theorem 5 ([62]). *Under Assumption 1, 2 and 3, for a non-degenerate problem SP$_N$, it holds that*

$$\mathbb{P}^N\Big(\mathbb{V}(x_N^*) > \epsilon\Big) \leq \sum_{i=1}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{2.15}$$

The probability $\mathbb{P}^N$ is taken with respect to N random samples $\{\xi^i\}_{i=1}^N$, and the inequality is tight for fully-supported problems.

As mentioned in Remark 2, $\mathbb{V}(x_N^*)$ is a random variable, its randomness comes from drawing scenarios $\{\xi^i\}_{i=1}^N$. For fully-supported problems, Theorem 5 shows the *exact* probability distribution of the violation probability $\mathbb{V}(x_N^*)$, i.e.

$$\mathbb{P}^N\Big(\mathbb{V}(x_N^*) > \epsilon\Big) = \sum_{i=1}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}, \tag{2.16}$$

the tail of a binomial distribution. We could use Theorem 5 to answer the sample complexity question in Remark 3.

Corollary 1 ([62]). *Given a violation probability $\epsilon \in (0,1)$ and a confidence parameter $\beta \in (0,1)$, if we choose the number of scenarios N (the smallest such N is denoted by N_{2008}) such that*

$$\sum_{i=0}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta \tag{2.17}$$

Let x_N^* denote the optimal solution to SP$_N$, it holds that

$$\mathbb{P}^N\Big(\mathbb{V}(x_N^*) \leq \epsilon\Big) \geq 1-\beta \tag{2.18}$$

In other words, the optimal solution x_N^* is a feasible solution to (CCO) with probability at least $1-\beta$.

Remark 2 states that the scenario approach is a randomized algorithm. Thus it is possible that the scenarios $\{\xi^i\}_{i=1}^N$ are drawn from a "bad" set and lead to infeasible solutions x_N^*, i.e.

$\mathbb{V}(x_N^*) \geq \epsilon$. The *confidence parameter* β denotes the risk of failure associated to the randomized solution algorithm [6], and it bounds the probability that x_N^* is infeasible.

For fully-supported problems, N_{2008} is the tightest upper bound on sample complexity, which cannot be improved. For non-fully supported problems, it turns out N_{2008} can be further tightened. An improved sample complexity bound is provided in Theorem 6 based on the definition of Helly's dimension.

Definition 10 (Helly's Dimension [63]). *Helly's dimension* of SP_N is the smallest integer h such that

$$\operatorname*{ess\,sup}_{\xi \in \Xi^N} |\mathcal{S}(\xi)| \leq h$$

holds for any finite $N \geq 1$. The essential supremum is denoted by $\operatorname{ess\,sup}$. We emphasize the dependence of support scenarios $\mathcal{S}$ on ξ by $\mathcal{S}(\xi)$.

Theorem 6 ([63]). *Let h denote the Helly's dimension for SP$_N$, under Assumption 1, 2 and 3, for a non-degenerate problem SP$_N$, it holds that*

$$\mathbb{P}^N\big(\mathbb{V}(x_N^*) > \epsilon\big) \leq \sum_{i=0}^{h-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \tag{2.19}$$

Equivalently, for a fixed confidence parameter $\beta \in (0,1)$, if the sample complexity N satisfies

$$\sum_{i=0}^{h-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta \tag{2.20}$$

then the following probabilistic guarantee holds

$$\mathbb{P}^N\big(\mathbb{V}(x_N^*) > \epsilon\big) \leq \beta \tag{2.21}$$

The only difference between Theorem 6 and Theorem 5 (and Corollary 1) is replacing n with Helly's dimension h in (2.19) and (2.20). Unfortunately, Helly's dimension is often difficult to calculate, while finding upper bounds $\overline{h}$ on Helly's dimension is usually a much easier task. Similarly we can replace h by $\overline{h}$ in (2.19) and (2.20), the same theoretical guarantees still hold because of the monotonicity of (2.19) and (2.20) in N and h. The support-rank defined in [66] is an upper bound on Helly's dimension, some other upper bounds can be obtained by exploiting the structural properties of the problem, e.g. [65].

Remark 4 (Sample Complexity Revisited). A binary search type algorithm could be used to find N_{2008}. And a looser but handy upper bound is provided in [31]:

$$N_{2009} := \frac{2}{\epsilon}\Big(\ln(\frac{1}{\beta}) + n\Big) \tag{2.22}$$

Notice n in (2.22) can be replaced by h or $\overline{h}$.

2.4.4 A-posteriori Feasibility Guarantees

When the desired violation probability ϵ is very small, the sample complexity of the a-priori guarantees grows with $1/\epsilon$ (Remark 4) and could be prohibitive. In other words, the a-priori approach is only suitable for the case where a *sufficient* amount of scenarios is always available. In many real-world applications (e.g. medical experiments, tests conducted by NASA), however, the amount of data is quite limited, and it could take months or cost a fortune to obtain a data point (experiment). Because of the limitation on the data availability, one of the most fundamental problem in data-driven decision making (e.g. system identification, quantitative finance) is to come up with good decisions or estimates with a moderate or even small amount of data. To overcome this, the scenario approach is extended towards a-posteriori feasibility guarantees.

Similar with the *a-priori* guarantees, obtaining a-posteriori guarantees typically requires taking the following three steps:

1. given dataset $\{\xi^i\}_{i=1}^N$, solve the corresponding scenario problem SP_N and obtain x_N^*;

2. find support scenarios in $\{\xi^i\}_{i=1}^N$, whose number is denoted as s_N^*;

3. calculate the posterior violation probability $\epsilon(\beta, s_N^*, N)$ using Theorem 7.

If the resulting violation probability $\epsilon(\beta, s_N^*, N)$ is greater than the acceptable level ϵ, we could repeat this process with more scenarios until reaching $\epsilon(\beta, s_N^*, N) \leq \epsilon$. If the number of available scenarios is limited, then it might be impossible to obtain a solution x_N^* such that $\mathbb{V}(x_N^*) \leq \epsilon$.

Theorem 7 (Wait-and-Judge [68])**.** *Given $\beta \in (0,1)$, for any $k = 0, 1, \cdots, n$, the polynomial equation in variable t*

$$\frac{\beta}{N+1}\sum_{i=k}^{N}\binom{i}{k}t^{i-k} - \binom{N}{k}t^{N-k} = 0 \tag{2.23}$$

has exactly one solution $\epsilon(k)$ in the interval $(0,1)$. Under Assumption 1, 2 and 3, for a non-degenerate problem, it holds that

$$\mathbb{P}^N(\mathbb{V}(x_N^*) \geq \epsilon(s_N^*)) \leq \beta \tag{2.24}$$

Theorem 7 is particularly useful in the following cases: (i) the problem is not fully-support thus difficult to calculate *a-priori* bounds on number of support scenarios; or (ii) only a moderate or small amount of data points is available, it is difficult to meet the sample complexity from the a-priori guarantees.

Given a candidate solution $x^\diamond$, the most straightforward method is to approximate $\mathbb{V}(x^\diamond)$ is by the empirical estimation $\hat{\epsilon}$ through Monte-Carlo simulation with $\hat{N}$ samples, i.e.

$$\hat{\epsilon} = \frac{1}{\hat{N}}\sum_{i=1}^{\hat{N}} \mathbb{1}_{f(x^\diamond,\xi^i)>0} = \frac{\hat{V}}{\hat{N}} \tag{2.25}$$

where $\hat{V} := \sum_{i=1}^{\hat{N}} \mathbb{1}_{f(x^\diamond,\xi^i)>0}$ is the total number of scenarios in which x_N^* is infeasible. Although (2.25) only involves $f(x^\diamond, \xi^i) > 0$ which is easy to calculate, it might require an astronomical number $\hat{N}$ to have accurate estimation $\hat{\epsilon}$ because of (D1). [30] shows a method to bound $\mathbb{V}(x^\diamond)$ from above using a dataset of a moderate size $\hat{N}$.

Proposition 1 ([30])**.** *Given a candidate solution $x^\diamond$ and $\hat{N}$ samples, let $\hat{V} := \sum_{i=1}^{\hat{N}} \mathbb{1}_{f(x^\diamond,\xi^i)>0}$*

and $1-\rho$ *be the confidence parameter.*

$$\overline{\epsilon} := \max_{\gamma \in [0,1]} \{\gamma : \sum_{i=0}^{\hat{V}} \binom{\hat{N}}{i} \gamma^i (1-\gamma)^{\hat{N}-i} \geq \rho\} \tag{2.26}$$

After finding an upper bound $\overline{\epsilon}$*, so that if* $\overline{\epsilon} \leq \epsilon$*, we may be sure that* $\mathbb{P}(\mathbb{V}(x^{\diamond}) \leq \epsilon) \geq 1-\rho$*.*

Remark 5. Proposition 1 is closely related with scenario approach but with one fundamental difference. Theorem 7 holds only for solution from scenario approach, while Proposition 1 can evaluate solutions from other methods.

2.4.5 Optimality Guarantees of Scenario Approach

Scenario approach together with order statistics can be used to construct lower bounds $\underline{o}$ on $o^\star$ of (CCO).

Proposition 2 ([30])**.** *Let* $\{\xi^{i,j}\}_{i=1}^N$ *(*$j = 1, 2, \cdots, K$*) be* K independent *datasets of size* N*. For the* j *th dataset, we solve the associated scenario problem SP*$_N$ *and calculate the optimal value* o_j^* *(*$j = 1, 2, \cdots, K$*). Without loss of generality, we assume that* $o_1^* \leq o_2^* \leq \cdots \leq o_K^*$*.*

Given $\delta \in (0,1)$*, let us choose positive integers L,N,L in such a way that*

$$\sum_{i=0}^{L-1} \binom{K}{i} (1-\epsilon)^{Ni} [1-(1-\epsilon)^N]^{K-i} \leq \delta \tag{2.27}$$

then with probability of at least $1-\delta$*, the random quantity* o_L^* *gives a lower bound for the true optimal value* $x^\star$*.*

[71] shows that appropriate N should be the order of $O(1/\epsilon)$ as $[1-(1-\epsilon)^N]^K \approx (1-\exp(-\epsilon N))^K$. Typically we choose proper values for N and K first, then find out the largest positive integer L that (2.27) holds true.

Proposition 2 turns out to be a general framework to construct lower bounds on (CCO). [71] extends the framework towards generating bounds using sample average approximation, which is introduced in Section 2.5.4.

2.5 Sample Average Approximation

2.5.1 Introduction to Sample Average Approximation

The idea of using sample average approximation to handle chance constraints first appeared in [23] and was subsequently improved with rigorous theoretical results in [29].

Let $\overline{f}(x,\xi) := \max\left\{f_1(x,\xi), \cdots, f_m(x,\xi)\right\}$, then (CCO) is equivalent to

$$\min_{x\in\mathcal{X}} c^\intercal x, \text{ s.t. } \mathbb{P}(\overline{f}(x,\xi) \leq 0) \geq 1-\epsilon.$$

Sample Average Approximation (SAA) approximates the true distribution of the random variable $\overline{f}(x,\xi)$ using the empirical distribution from N samples $\{\xi^i\}_{i=1}^N$, i.e. $\mathbb{P}(\overline{f}(x,\xi) \leq 0)$ is approximated by $\frac{1}{N}\sum_{i=1}^N \mathbb{1}_{\overline{f}(x,\xi^i)\leq 0}$.

$$\text{(SAA)}: \min_{x\in\mathcal{X}} \ c^\intercal x \tag{2.28a}$$

$$\text{s.t. } \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{\overline{f}(x,\xi^i)>0} \leq \varepsilon \tag{2.28b}$$

(SAA) is also a chance constrained optimization problem, but with two major differences from (CCO): (i) (SAA) is based on the empirical (discrete) distribution from the true distribution of ξ as in (CCO); (ii) (SAA) has the violation probability ε instead of ϵ in (CCO).

There are two critical questions to be addressed about (SAA). What is the connection of solutions of (SSA) with that of (CCO)? How to solve (SAA)? We first answer the second question in Section 2.5.2, then present the theoretical results of connecting (SAA) with (CCO).

2.5.2 Solving Sample Average Approximation

(SAA) can be reformulated as a mixed integer program (MIP) by introducing variables $z \in \{0,1\}^N$ [28, 29]. Binary variable z_i is an indicator if $\overline{f}(x,\xi) \leq 0$ is being violated in sample i, i.e.

$$z_i = \mathbb{1}_{\overline{f}(x,\xi^i)>0} \tag{2.29}$$

(2.29) can be equivalently written as $\overline{f}(x,\xi^i) \leq Mz_i$ with a sufficiently large coefficient $M \in \mathbf{R}_+$. Since $\overline{f}(x,\xi^i)$ is the maximum over m functions $\{f_j(x,\xi^i)\}_{j=1}^m$, $\overline{f}(x,\xi^i) \leq Mz_i$ implies $f_j(x,\xi^i) \leq Mz_i, j = 1,2,\cdots,m$. Then (SAA) is equivalent to (2.30), in which $\mathbf{1}_m$ is an all one vector with size m.

$$\min_{x,z} \; c^{\intercal}x \tag{2.30a}$$

$$\text{s.t. } f(x,\xi^1) - Mz_1\mathbf{1}_m \leq 0 \tag{2.30b}$$

$$\vdots$$

$$f(x,\xi^N) - Mz_N\mathbf{1}_m \leq 0 \tag{2.30c}$$

$$\frac{1}{N}\sum_{i=1}^{N} z_i \leq \varepsilon \tag{2.30d}$$

$$x \in \mathcal{X}, z_i \in \{0,1\}, i = 1,2,\cdots,N \tag{2.30e}$$

(2.30) is equivalent to (SAA) for general function $f(x,\xi)$, but formulations with big-**M** are typically weak formulations. Introducing big coefficients M might cause numerical issues as well. Stronger formulations of (SAA) are possible by exploiting the structural features of $f(x,\xi)$. A good example is chance-constrained linear program with separable probabilistic constraints:

$$\min_{x\in\mathcal{X}} c^{\intercal}x \text{ s.t.} \mathbb{P}(Tx \geq \xi) \geq 1-\epsilon,$$

with a constant matrix $T \in \mathbf{R}^{d\times n}$. By introducing auxiliary variables v, an equivalent but stronger formulation without big **M** is (2.31) [36].

$$\min_{x\in\mathcal{X}} \; c^{\intercal}x \tag{2.31a}$$

$$\text{s.t. } Tx = v \tag{2.31b}$$

$$v + \xi_i z_i \geq \xi_i, i = 1,2,\cdots,N \tag{2.31c}$$

$$\frac{1}{N}\sum_{i=1}^{N} z_i \leq \varepsilon \tag{2.31d}$$

$$z_i \in \{0,1\}, i = 1, 2, \cdots, N \tag{2.31e}$$

Various strong formulations for (SAA) can be found in [36] and references therein. (2.30) and (2.31) are mixed integer programs, some well-known techniques from integer programming theory can speed up the process of solving (SAA), e.g. adding cuts [36, 72, 73] and decompositions [74, 75].

2.5.3 Feasibility Guarantees of SAA

Various feasibility guarantees of (SAA) are proved in [29, 71], e.g. the asymptotic behavior of (SAA) and when $f(x,\xi)$ is Lipschitz continuous. In this section, we only present the Lipschitz case, which could be used for simulations in Section 5.2.

Assumption 5. *There exists $L > 0$ such that*

$$|\overline{f}(x,\xi) - \overline{f}(x',\xi)| \leq L\|x - x'\|_\infty, \ \forall x, x' \in \mathcal{X} \text{ and } \forall \xi \in \Xi. \tag{2.32}$$

Theorem 8 ([29]). *Suppose $\mathcal{X}$ is bounded with diameter D and $\overline{f}(x,\xi)$ is L-Lipschitz for any $\xi \in \Xi$ (Assumption 5). Let $\varepsilon \in [0,\epsilon), \theta \in (0, \epsilon - \varepsilon)$ and $\gamma > 0$. Then*

$$\mathbb{P}(\mathcal{F}^N_{\varepsilon,\gamma} \subseteq \mathcal{F}_\epsilon) \geq 1 - \left\lceil \frac{1}{\theta} \right\rceil \left\lceil \frac{2LD}{\gamma} \right\rceil^n \exp(-2N(\epsilon - \varepsilon - \theta)^2) \tag{2.33}$$

where the feasible region of (SAA) is defined as

$$\mathcal{F}^N_{\varepsilon,\gamma} := \{x \in \mathcal{X} : \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\overline{f}(x,\xi)+\gamma\leq 0} \geq 1 - \varepsilon\}. \tag{2.34}$$

For fixed ϵ and ε, if we choose $\theta = (\epsilon - \varepsilon)/2$ and a small number $\gamma > 0$, then Theorem 8 suggests that using

$$N \geq \frac{2}{(\epsilon-\varepsilon)^2}\big[\ln(\frac{1}{\beta}) + n\ln(\left\lceil\frac{2LD}{\gamma}\right\rceil) + \ln(\left\lceil\frac{2}{\epsilon-\varepsilon}\right\rceil)\big] \tag{2.35}$$

number of samples, solutions of (SAA) is feasible to (CCO) with high probability $1-\beta$, i.e. $\mathbb{P}(\mathcal{F}^N_{\varepsilon,\gamma} \subseteq \mathcal{F}_\epsilon) \geq 1-\beta$.

The results in Theorem 8 look quite similar to those of scenario approach (e.g. Remark 4). Indeed, (SAA) with $\varepsilon = 0$ is the same as the scenario problem SP$_N$. However, one major difference of Theorem 8 from the scenario approach theory is that: Theorem 8 holds for the feasible region of (SAA), i.e. $\mathcal{F}^N_{\varepsilon,\gamma} \subseteq \mathcal{F}_\epsilon$ with high probability. While the theory of the scenario approach only proves the property of the optimal solution x^*_N, i.e. x^*_N is feasible with high probability. Other feasible solutions to SP$_N$ do not necessarily process the properties guaranteed by the scenario approach (e.g. Theorem 5).

Although Theorem 8 provides explicit sample complexity bounds for (SAA) to obtain feasible solution, it requires some efforts to be applied, e.g. tuning parameters (ε, θ) and calculation of L and D. [70] provides a similar but more straightforward theoretical result.

Theorem 9 (Sampling & Discarding [70])**.** *If we draw N samples and discard any k of them, then use the scenario approach with the remaining $N-k$ samples. If N and k satisfy*

$$\binom{k+n-1}{k} \cdot \sum_{i=0}^{k+n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta \tag{2.36}$$

then $\mathbb{P}^N\Big(\mathbb{P}_\xi(f(x^*_{N,k},\xi) \leq 0) \geq 1-\epsilon\Big) \geq 1-\beta.$

Given parameters N, ϵ and β, we find the largest k that (2.36) holds, then the solution to (SAA) with $\varepsilon = k/N$ is feasible to (CCO) with probability at least $1-\beta$.

2.5.4 Optimality Guarantees of Sample Average Approximation

It is intuitive that if $\varepsilon > \epsilon$, then the objective values of SAA yield lower bounds to (CCO). Theorem 10 formalizes this intuition.

Theorem 10 ([29]). *Let $\varepsilon > \epsilon$ and assume that (CCO) has an optimal solution. Then*

$$\mathbb{P}\Big(\hat{o}_{\varepsilon}^{N} \leq o_{\epsilon}^{\star}\Big) \geq 1 - \exp(-2N(\varepsilon - \epsilon)^2). \tag{2.37}$$

Theorem 10 directly suggests a method to construct lower bounds on (CCO).

Proposition 3. *If we choose $\varepsilon > \epsilon$ and $N \geq \frac{1}{2(\varepsilon-\epsilon)^2}\log(\frac{1}{\delta})$, let o_{ε}^{SAA} denote the objective value of (SAA), then o_{ε} is a lower bound with probability at least $1-\delta$, i.e. $\mathbb{P}(o_{N,\varepsilon}^{*} \leq o_{\epsilon}^{\star}) \geq 1-\delta$.*

There is an alternative method using SAA to generate lower bounds of (CCO). [29] extends the framework in Proposition 2 towards SAA.

Proposition 4 ([29]). *Take K sets of N independent samples $\{\xi^{i,j}\}_{i=1}^{N}$, ($j = 1, 2, \cdots, K$). For the jth dataset $\{\xi^{i,j}\}_{i=1}^{N}$, we solve the associated (SAA) problem and calculate the associated objective value $o_{N,\varepsilon,j}^{*}$ (for simplicity o_j^* and $j = 1, 2, \cdots, K$). Without loss of generality, we assume that $o_1^* \leq o_2^* \leq \cdots \leq o_K^*$.*

Given $\delta \in (0,1)$, $\varepsilon \in [0,1)$, let us choose positive integers N,L, K ($L \leq K$) such that

$$\sum_{i=0}^{L-1} \binom{K}{i} \big[b(\varepsilon,\epsilon,N)\big]^{i} \big[1 - b(\varepsilon,\epsilon,N)\big]^{K-i} \geq \delta \tag{2.38}$$

where $b(\varepsilon,\epsilon,N) := \sum_{i=0}^{\lfloor \varepsilon N \rfloor} \binom{N}{i} \epsilon^{i}(1-\epsilon)^{N-i}$.

Then o_L^* serves as a lower bound to (CCO) with probability at least $1-\delta$.

2.6 Robust Optimization Related Methods

2.6.1 Introduction to Robust Optimization

The last category of solutions to (CCO) is closely related with robust optimization (RO), its typical form is shown in (2.39).

$$\text{(RC):} \min_{x \in \mathcal{X}} \; c^{\intercal} x \tag{2.39a}$$

$$\text{s.t. } f(x,\xi) \leq 0, \; \forall \xi \in \mathcal{U}_{\epsilon} \tag{2.39b}$$

(2.39) finds the optimal solution which is feasible to all realizations of uncertainties in a predefined uncertainty set $\mathcal{U}_\epsilon$. (2.39) is called the Robust Counterpart (RC) of the original problem (CCO). By constructing an uncertainty set $\mathcal{U}_\epsilon$ with proper shape and size, solutions to (RC) could be suboptimal or approximate solutions to (CCO).

Designing uncertainty sets lies at the heart of robust optimization. A good uncertainty set should meet the following two standards:

(S1) The resulting (RC) problem is computationally tractable.

(S2) The optimal solution to (RC) is not too conservative or overly optimistic.

Unfortunately, (RC) of general robust convex problems (under Assumption 2) is not always computationally tractable. For example, (RC) of a second order cone program (SOCP) with polyhedral uncertainty set is NP-Hard [54, 78, 79]. Fortunately, robust linear programs are well-studied, and (RC) of linear programs is tractable for common choices of uncertainty sets. Most tractability results of robust linear optimization are summarized in [54]. For tractable formulations of general convex RO problems, various solutions can be found in [11, 80].

For simplicity, we present solutions to the following chance-constrained linear program (CCLP) [4].

$$\min_{x\in\mathcal{X}} \ c^\mathsf{T} x \tag{2.40a}$$

$$\text{s.t. } \mathbb{P}_\xi\Big(x_0^i + \xi^\mathsf{T} x^i \le 0,\ i = 1, 2, \cdots, m\Big) \ge 1 - \epsilon \tag{2.40b}$$

and its robust counterpart

$$\min_{x\in\mathcal{X}} \ c^\mathsf{T} x \tag{2.41a}$$

$$\text{s.t. } x_0^i + \xi^\mathsf{T} x^i \le 0,\ \forall \xi \in \mathcal{U}_\epsilon,\ i = 1, 2, \cdots, m \tag{2.41b}$$

[4] A (seemingly) more general form of the linear chance constraint is $\mathbb{P}\Big(A(\xi)x \le b(\xi)\Big) \ge 1 - \epsilon$, where $A(\xi)$ and $b(\xi)$ denote affine functions of ξ. This could be equivalently represented in the form of (2.40b) by enforcing additional affine constraints [34]

In (2.40) and (2.41), decision variables are $\{x_0^i, x^i\}_{i=1}^m$, where $x_0^i \in \mathbf{R}$ and $x^i \in \mathbf{R}^n$. Uncertainties are represented by $\xi \in \mathbf{R}^d$ [5] With a little abuse of notation, we use $x = [x_0^1, x^1, \cdots, x_0^m, x^m]^\intercal$ to represent all the decision variables.

Standard (S2) is directly connected with chance constraints, we present the connection between RO and CCO in Section 2.6.2-2.6.4.

2.6.2 Safe Approximation

Almost every RO-related solution to (CCO) is based on the idea of safe approximation.

Definition 11 (Safe Approximation)**.** Let $x \in \mathcal{F}$ and $x \in \underline{\mathcal{F}}$ denote two sets of constraints. We say $\underline{\mathcal{F}}$ is a safe approximation (or inner approximation) of $\mathcal{F}$ if $\underline{\mathcal{F}} \subseteq \mathcal{F}$.

An optimization problem (SA) is called a *safe approximation* of (CCO) if $\underline{\mathcal{F}} \subseteq \mathcal{F}_\epsilon$, where $\mathcal{F}_\epsilon$ represents the feasible region of (CCO) as in Definition 3.

$$\text{(SA):} \ \min_{x \in \mathcal{X}} \ c^\intercal x \tag{2.42a}$$

$$\text{s.t. } x \in \underline{\mathcal{F}} \tag{2.42b}$$

$\underline{\mathcal{F}} \subseteq \mathcal{F}_\epsilon$ indicates that every solution to (SA) is *feasible* to (CCO). Therefore every optimal solution to (SA) is *suboptimal* to (CCO) and serves as an upper bound on (CCO).

There are two major approaches to constructing safe approximations of the chance constraint $\mathbb{P}_\xi\big(f(x,\xi) \le 0\big) \ge 1-\epsilon$: (i) constructing a function $\pi(x) \ge \mathbb{P}_\xi\big(f(x,\xi) > 0\big)$, then $\pi(x) \le \epsilon$ is a safe approximation of the chance constraint; (ii) constructing a proper uncertainty set $\mathcal{U}_\epsilon$ such that $\mathcal{F}_\epsilon \supseteq \mathcal{F}_{\mathcal{U}_\epsilon} := \{x \in \mathbf{R}^n : f(x,\xi) \le 0, \ \forall \xi \in \mathcal{U}_\epsilon\}$. Although these two approaches look quite different, Section 2.6.3.2 shows that they are closely related with each other.

We first review how to apply these two approaches to obtaining safe approximation of individual chance constraints in Section 2.6.2. Safe approximations of joint chance constraints (Section 2.6.4) are built upon the results of individual chance constraints.

[5] Notice $d = n$ in (2.40) and (2.41).

2.6.3 Safe Approximation of Individual Chance Constraints

RO has been quite successful in constructing safe approximations of individual chance constraints. A general form of individual chance-constrained programs is (2.43).

$$\min_{x \in \mathcal{X}} \; c^\intercal x \tag{2.43a}$$

$$\text{s.t. } \mathbb{P}_\xi\Big(f(x,\xi) \leq 0\Big) \geq 1-\epsilon \tag{2.43b}$$

In the individual chance constraint (2.43b), the inner function $f(x,\xi) : \mathbf{R}^n \times \mathbf{R}^d \to \mathbf{R}^1$ is a *scalar-valued* function. In Section 2.6.3, all $f(x,\xi)$ are scalar-valued functions if not specified.

Section 2.6.2 outlines two different but related approaches to constructing safe approximations. The first approach is presented in Section 2.6.3.1-2.6.3.2. The second approach is summarized in 2.6.3.3.

2.6.3.1 Convex Approximation

Convex approximation is a general framework to build safe approximations of individual chance constraints. The idea of convex approximation first appeared in [22], then was completed in [30]. The convex approximation framework is based on the concept of generating function.

Definition 12 (Generating Function)**.** A function $\phi : \mathbf{R} \to \mathbf{R}$ is called a (one-dimensional) *generating function* if it is nonnegative valued, nondecreasing, convex and satisfying the following property:

$$\phi(z) > \phi(0) = 1, \forall z > 0 \tag{2.44}$$

The idea of convex approximation starts from the following lemma.

Lemma 1. *For a positive constant $t \in \mathbf{R}_+$ and a random variable $z \in \mathbf{R}$, it holds that*

$$\mathbb{E}[\phi(t^{-1}z)] \geq \mathbb{E}[\mathbf{1}_{t^{-1}z \geq 0}] = \mathbb{P}_z(t^{-1}z \geq 0) = \mathbb{P}(z \geq 0) \tag{2.45}$$

Replace z with $f(x,\xi)$, then $\mathbb{E}[\phi(t^{-1}f(x,\xi))] \geq \mathbb{P}_\xi\Big(f(x,\xi) > 0\Big) = \mathbb{P}_\xi\Big(t^{-1}f(x,\xi) > 0\Big)$. In other words, $\mathbb{E}[\phi(t^{-1}f(x,\xi))] \leq \epsilon$ is a *safe approximation* to $\mathbb{P}_\xi\Big(f(x,\xi) \leq 0\Big) \geq 1-\epsilon$.

Theorem 11 (Convex Approximation [30])**.** *Let $\phi(\cdot)$ be a generating function, then (CA) is a safe approximation to (CCO).*

$$\textit{(CA):}\ \min_{x\in\mathcal{X}}\ c^\intercal x \tag{2.46a}$$

$$\textit{s.t.}\ \inf_{t>0}\big[t\mathbb{E}_\xi[\phi(\frac{f(x,\xi)}{t})] - t\epsilon\big] \leq 0 \tag{2.46b}$$

Under Assumption 2, (CA) is convex in x.

Remark 6. We can get rid of the strict inequality $t > 0$ by approximating it using $t \geq \delta$, where δ is very small positive number (e.g. $\delta = 10^{-4}$). Furthermore, we can show that (CA) is equivalent to (2.47), which is convex in (x,t).

$$\min_{x\in\mathcal{X},t\geq\delta}\ c^\intercal x \tag{2.47a}$$

$$\text{s.t. } t\mathbb{E}_\xi[\phi(\frac{f(x,\xi)}{t})] - t\epsilon \leq 0 \tag{2.47b}$$

Choosing a good generating functions plays a crucial role in the convex approximation framework. Choices of generating functions include: Markov bound $\phi(z) = [1+z]_+$, Chernoff bound $\phi(z) = \exp(z)$, Chebyshev bound $\phi(z) = [z+1]_+^2$ and Traditional Chebyshev bound $\phi(z) = (z+1)^2$. The *least* conservative generating function is the Markov bound $\phi(z) = [1+z]_+$ [30, 81].

Definition 13 (Conditional Value at Risk)**.** Conditional value at risk (CVaR) of a random variable z at level $1-\epsilon$ is defined as

$$\mathrm{CVaR}(z;1-\epsilon) := \inf_\gamma(\gamma + \frac{1}{\epsilon}\mathbb{E}\big[[z-\gamma]_+\big]) \tag{2.48}$$

Proposition 5 ([30, 34]). *(CA) with Markov bound $\phi(z) = [z+1]_+$ is equivalent to (2.49).*

$$\min_{x\in\mathcal{X}} \; c^\mathsf{T} x \tag{2.49a}$$

$$\textit{s.t.} \; \mathrm{CVaR}\left(f(x,\xi); 1-\epsilon\right) \le 0 \tag{2.49b}$$

Section 2.1 shows an individual chance constraint $\mathbb{P}\big(f(x,\xi) \le 0\big) \ge 1-\epsilon$ is equivalent to $\mathrm{VaR}(f(x,\xi); 1-\epsilon) \le 0$. It is well-known that $\mathrm{CVaR}(z; 1-\epsilon) \ge \mathrm{VaR}(z; 1-\epsilon)$. Therefore, $\mathrm{CVaR}(f(x,\xi); 1-\epsilon) \le 0$ implies $\mathrm{VaR}(f(x,\xi); 1-\epsilon) \le 0$. In other words, $\mathrm{CVaR}(f(x,\xi); 1-\epsilon) \le 0$ is a safe approximation to both $\mathrm{VaR}(f(x,\xi); 1-\epsilon) \le 0$ and the chance constraint (2.43b).

Remark 7 (Sample Approximation of CVaR). [33] utilizes a dataset $\{\xi^i\}_{i=1}^N$ to estimate CVaR.

$$\min_{x\in X, t} \; c^\mathsf{T} x \tag{2.50a}$$

$$\text{s.t.} \; \frac{1}{N}\sum_{i=1}^{N}[f(x,\xi^i)+t]_+ \le t\epsilon \tag{2.50b}$$

By introducing N auxiliary variables, [33] shows that (2.50) can be reformulated as a convex problem that is easy to solve. Detailed reformulation can be found in [33] and [82]. With a sufficient number of data points (N is large enough), (2.50) is a safe approximation to (CCO). However, it remains unknown about the exact requirement on the number of samples needed. The sample approximation of CVaR may not necessarily yield a safe approximation [34].

The generating function based framework in [30] was further improved and completed in [11, 50]. But the methods proposed there are mainly analytical and aim at solving distributionally robust problems, which is beyond the scope of this dissertation. More details can be found in Figure 2.2 and references therein.

2.6.3.2 CVaR-based Convex Approximation of Individual Chance Constraints

As pointed out in [30], calculating CVaR is computationally intractable. In order to obtain tractable forms of the CVaR-based convex approximation, one approach is the sample approximation in Remark 7. An alternative approach is to bound the CVaR function from above, e.g. finding a function $\pi(x) \geq \mathrm{CVaR}(f(x,\xi); 1-\epsilon)$, then $\pi(x) \leq 0$ is a safe approximation to both $\mathrm{CVaR}(f(x,\xi); 1-\epsilon) \leq 0$ and the original chance constraint (2.43). In the latter approach, the uncertainties $\xi \sim \Xi$ are partially characterized using directional deviations.

Definition 14 (Directional Deviations [83])**.** Given a random variable $\xi \in \mathbf{R}$ with zero mean, the forward deviation is defined as

$$\delta_+(\xi) := \sup_{\theta>0} \left\{ \sqrt{\frac{2\ln(\mathbb{E}[\exp(\theta\xi)])}{\theta^2}} \right\} \tag{2.51}$$

and the backward deviation is defined as

$$\delta_-(\xi) := \sup_{\theta>0} \left\{ \sqrt{\frac{2\ln(\mathbb{E}[\exp(-\theta\xi)])}{\theta^2}} \right\}. \tag{2.52}$$

Assumption 6 ([55])**.** *Let $\mathcal{W}$ denote the smallest closed convex set containing the support Ξ of ξ. We assume that the support set is a second-order conic representable set (e.g. polyhedral and ellipsoidal sets).*

Assumption 7 ([55])**.** *Assume the uncertainties $\{\xi_i\}_{i=1}^d$ are zero mean random variables, with a positive definite covariance matrix Σ. We define the following index set:*

$$\mathcal{J}_+ := \{i : \delta_+(\xi_i) < \infty\},\ \mathcal{I}_+ := \{i : \delta_+(\xi_i) = \infty\}, \tag{2.53}$$

$$\mathcal{J}_- := \{i : \delta_-(\xi_i) < \infty\},\ \mathcal{I}_- := \{i : \delta_-(\xi_i) = \infty\}. \tag{2.54}$$

For notation simplicity, we define two matrices diagonal P and Q as:

$$P := \mathrm{diag}(\delta_+(\xi_1), \cdots, \delta_+(\xi_d)),\ Q := \mathrm{diag}(\delta_-(\xi_1), \cdots, \delta_-(\xi_d)).$$

Major results developed in [55, 83] are for the individual *linear* chance constraint (2.55) with decision variables $x_0 \in \mathbf{R}, x \in \mathbf{R}^n$:

$$\mathbb{P}_\xi\Big(x_0 + \xi^\intercal x \le 0\Big) \ge 1 - \epsilon \tag{2.55}$$

Its convex approximation using CVaR (or Markov bound) is

$$t + \frac{1}{\epsilon}\mathbb{E}\big[[x_0 + \xi^\intercal x - t]_+\big] \le 0 \tag{2.56}$$

If we are able to find a function $\pi(x_0, x)$ as an upper bound on $\mathbb{E}\big[[x_0 + \xi^\intercal x]_+\big]$, then

$$t + \frac{1}{\epsilon}\pi(x_0 - t, x) \le 0 \tag{2.57}$$

is a safe approximation to (2.56).

Theorem 12. *[55] Suppose that the primitive uncertainty ξ satisfies Assumption 6 and 7. The following functions $\pi^i(x_0, x), i = 1, \cdots, 5$ are upper bounds of $\mathbb{E}_\xi\big[[x_0 + \xi^\intercal x]_+\big]$:*

$$\pi^1(x_0, x) \;:=\; \big[x_0 + \max_{\xi \in \mathcal{W}} \xi^\intercal x\big]_+ \tag{2.58}$$

$$\pi^2(x_0, x) \;:=\; x_0 + \big[-x_0 + \max_{\xi \in \mathcal{W}} (-\xi)^\intercal x\big]_+ \tag{2.59}$$

$$\pi^3(x_0, x) \;:=\; \frac{1}{2}\Big(x_0 + \sqrt{x_0^2 + x^\intercal \Sigma x}\Big) \tag{2.60}$$

$$\pi^4(x_0, x) \;:=\; \inf_{\mu > 0}\left\{\frac{\mu}{\epsilon}\exp\left(\frac{x_0}{\mu} + \frac{u^\intercal u}{2\mu^2}\right)\right\}. \tag{2.61}$$

where $u_j = \max\{x_j\delta_+(\xi_j), -x_j\delta_-(\xi_j)\}, j = 1, \cdots, n$. This bound is finite if and only if $x_j \le 0,\ \forall j \in \mathcal{I}_+$ and $x_j \ge 0,\ \forall j \in \mathcal{I}_-$.

$$\pi^5(x_0, x) \;:=\; x_0 + \inf_{\mu > 0}\left\{\frac{\mu}{\epsilon}\exp\left(-\frac{x_0}{\mu} + \frac{v^\intercal v}{2\mu^2}\right)\right\}. \tag{2.62}$$

where $v_j = \max\{-x_j\delta_+(\xi_j), x_j\delta_-(\xi_j)\}, j = 1, \cdots, n$. *This bound is finite if and only if* $x_j \geq 0,\ \forall j \in \mathcal{I}_+$ *and* $x_j \leq 0,\ \forall j \in \mathcal{I}_-$.

Remark 8. The epigraphs of $\pi^i(x_0, x),\ i = 1, \cdots, 5$ can be represented as second-order cones. Explicit representations depend on the form of $\mathcal{W}$. More details about the representation of (2.57) with different choices of $\pi^i(x_0, y)$ can be found in [55] and [82].

2.6.3.3 Constructing Uncertainty Sets

We consider the individual linear chance constraint (2.55) as in Section 2.6.3.2. The robust counterpart of (2.55) is

$$x_0 + \xi^\intercal x \leq 0,\ \forall \xi \in \mathcal{U}_\epsilon \tag{2.63}$$

Assumption 8. $\{\xi_i\}_{i=1}^d$ *are independent of each other with zero mean and take values on* $[-1, 1]^d$, *i.e.* $\mathbb{E}[\xi_i] = 0$ *and* $\xi_i \in [-1, 1]$ *for* $i = 1, 2, \cdots, d$.

Clearly, under Assumption 8, a natural choice of uncertainty set is the box $\mathcal{U}^{\text{box}} := \{\xi \in \mathbf{R}^d : -\mathbf{1} \leq \xi \leq \mathbf{1}\}$. Then $\mathcal{F}_{\mathcal{U}}^{\text{box}} := \{x \in \mathbf{R}^n : f(x, \xi) \leq 0,\ \forall \xi \in \mathcal{U}^{\text{box}}\}$ is a safe approximation to $\mathcal{F}_\epsilon$, i.e. $\mathcal{F}_{\mathcal{U}}^{\text{box}} \subseteq \mathcal{F}_\epsilon$. However, using $\mathcal{U}^{\text{box}}$ leads to $\mathbb{P}(f(x, \xi) \geq 0) = 0 \ll \epsilon$, which causes conservativeness or even infeasibility in many cases. The following choices of uncertainty sets are less conservative.

Theorem 13 ([11, 58, 59])**.** *(2.63) is a safe approximation to (2.55) if* $\mathcal{U}_\epsilon$ *is one of the following:*

$$\mathcal{U}_\epsilon^{\textit{ball}} := \left\{\xi \in \mathbf{R}^d : \|\xi\|_2 \leq \sqrt{2\ln(1/\epsilon)}\right\} \tag{2.64a}$$

$$\mathcal{U}_\epsilon^{\textit{ball-box}} := \left\{\xi \in \mathbf{R}^d : \|\xi\|_\infty \leq 1, \|\xi\|_2 \leq \sqrt{2\ln(1/\epsilon)}\right\} \tag{2.64b}$$

$$\mathcal{U}_\epsilon^{\textit{budget}} := \left\{\xi \in \mathbf{R}^d : \|\xi\|_1 \leq \sqrt{2d\ln(1/\epsilon)}\right\} \tag{2.64c}$$

And the resulting robust counterparts (RC)s are second-order cone representable (see Chapter 2 of [11] and [82]).

It turns out that constructing uncertainty set $\mathcal{U}_\epsilon$ is closely related with the convex approximation framework in Section 2.6.3.1-2.6.3.2.

Theorem 14 ([34]). *Suppose that $\pi(x_0, x)$ is a convex, closed and positively homogeneous, and is an upper bound to $\mathbb{E}_\xi\big[[x_0 + \xi^\intercal x]_+\big]$ with $\pi(x_0, 0) = x_0^+$. Then under Assumptions 6 and 7 and given $\epsilon \in (0, 1)$, it holds that for all (x_0, x) such that $\pi(x_0, x) < \infty$, we have*

$$\inf_t \Big(t + \frac{1}{\epsilon}\pi(x_0 - t, x)\Big) = x_0 + \max_{z\in\mathcal{U}_\epsilon} x^\intercal z \tag{2.65}$$

for some convex uncertainty set $\mathcal{U}_\epsilon$.

Given an upper bound $\pi(x_0, x)$ on $\mathbb{E}\big[[x_0 + \xi^\intercal x]_+\big]$ with required properties, the safe approximation (2.57) can be represented in the form of $x_0 + \max_{\xi\in\mathcal{U}_\epsilon} \xi^\intercal x$ for some $\mathcal{U}_\epsilon$. Theorem 14 only proves the existence of a corresponding uncertainty set $\mathcal{U}_\epsilon$. For the $\pi^i(x_0, x)$ functions given in Theorem 12, their corresponding uncertainty sets can be explicitly calculated.

Proposition 6 ([34]). *For the functions $\pi^i(x_0, x), i = 1, 2, \cdots, 5$ in Theorem 12, their corresponding uncertainty sets are $\mathcal{U}_\epsilon^1 \sim \mathcal{U}_\epsilon^5$ below.*

$$\mathcal{U}_\epsilon^1 := \mathcal{W}, \tag{2.66}$$

$$\mathcal{U}_\epsilon^2 := \Big\{\xi \in \mathbf{R}^d : \xi = (1 - \frac{1}{\epsilon})\zeta, \textit{ for some } \zeta \in \mathcal{W}\Big\}, \tag{2.67}$$

$$\mathcal{U}_\epsilon^3 := \Big\{\xi \in \mathbf{R}^d : \|\Sigma^{-\frac{1}{2}}\xi\|_2 \le \sqrt{\frac{1-\epsilon}{\epsilon}}\Big\} \tag{2.68}$$

$$\mathcal{U}_\epsilon^4 := \Big\{\xi \in \mathbf{R}^d : \exists s, t \in \mathbf{R}^d, \xi = s - t,\ \|P^{-1}s + Q^{-1}t\|_2 \le \sqrt{-2\ln(\epsilon)}\Big\}, \tag{2.69}$$

$$\mathcal{U}_\epsilon^5 := \Big\{\xi \in \mathbf{R}^d : \exists s, t \in \mathbf{R}^d, \xi = s - t,\ \|P^{-1}s + Q^{-1}t\|_2 \le \frac{1-\epsilon}{\epsilon}\sqrt{2\ln(\frac{1}{1-\epsilon})}\Big\}. \tag{2.70}$$

where matrices Σ, P and Q are defined in Assumptions 6 and 7.

Theorem 14 and Proposition 6 demonstrate that the two seemingly different approaches to constructing safe approximations in Section 2.6.2 are equivalent in many circumstances.

2.6.4 Safe Approximation of Joint Chance Constraints

Although RO has been successful in approximating individual chance constraints, it is rather unsatisfactory in approximating joint chance constraints [34]. We restate the joint chance constraint (2.1b) below

$$\mathbb{P}_\xi\Big(f(x,\xi) \le 0\Big) \ge 1-\epsilon. \tag{2.71}$$

Most RO-based approaches convert a joint chance constraint to several individual chance constraints, then apply the techniques in Section 2.6.3 on each individual chance constraint. Results along this line are summarized in Section 2.6.4.1. Very few approaches directly deal with joint chance constraints, these approaches are mentioned in Section 2.6.4.2.

2.6.4.1 Conversion Between Joint Chance Constraints and Individual Chance Constraints

Section 2.1.2 presents two common approaches to converting a joint chance constraint to individual chance constraints.

First, according to the Boole's inequality or Bonferroni inequality, if $\sum_{i=1}^m \epsilon_i \le \epsilon$, then the set of m individual chance constraints

$$\mathbb{P}\Big(f_i(x,\xi) \le 0\Big) \le 1-\epsilon_i,\ i=1,\cdots,m \tag{2.72}$$

is a safe approximation to the joint chance constraint $\mathbb{P}(f(x,\xi) \le 0) \le 1-\epsilon$. The main issue of this approach is the choice of $\{\epsilon_i\}_{i=1}^m$. The problem becomes intractable if taking $\{\epsilon_i\}_{i=1}^m$ as decision variables [30, 34]. It remains unclear about how to find the optimal choices of $\{\epsilon_i\}_{i=1}^m$ [6]. Obviously, this approach could be quite conservative in the following two cases: (i) the individual constraints $f_i(x,\xi),\ i=1,2,\cdots,m$ are correlated; and (ii) the choices of $\{\epsilon_i\}_{i=1}^m$ are suboptimal. [34] provides some deeper observations on the limitation of this approach: the Bonferroni's inequality could still lead to conservativeness even when (i) the individual chance constraints (2.72) are independent; and (ii) the optimal choices of $\{\epsilon_i\}_{i=1}^m$ are found. In other words, (2.72) is only a

[6] Most people simply choose $\epsilon_i = \epsilon/m$ [30, 83], which could be quite conservative if m is a large number.

safe approximation at best, it may not be equivalent to (2.1b) even with optimal $\{\epsilon_i\}_{i=1}^m$.

The second approach is to define the pointwise maximum of functions $\{f_i(x,\xi)\}_{i=1}^m$ over x and ξ, i.e.

$$\overline{f}(x,\xi) := \max\Big\{f_1(x,\xi),\cdots,f_m(x,\xi)\Big\}.$$

then the joint chance constraint $\mathbb{P}(f(x,\xi) \leq 0) \geq 1-\epsilon$ is equivalent to the individual chance constraint $\mathbb{P}_\xi\big(\overline{f}(x,\xi) \leq 0\big) \geq 1-\epsilon$. The advantage of this approach is that it does not require parameter tuning or induce additional conservativeness. In some cases, e.g. scenario approximation of CVaR in Remark 7, this could lead to formulations that are easy to solve [82]. However, in most cases, the structure of $\overline{f}(x,\xi)$ is too complicated to apply the techniques in Section 2.6.3.

2.6.4.2 Other Approaches

There might be only three RO-related approaches that directly deal with joint chance constraints. The first approach is robust conic optimization (see Chapter 5-11 of [11]). The inner constraint $f(x,\xi) \leq 0$ is written as a conic inequality, then tractable safe approximations of the robust conic inequality are derived and solved. This approach can model a majority of optimization problems under uncertainties. However, the main limitation is that the resulting robust counterparts are not tractable in many circumstances.

The second approach [34] generalizes the CVaR-based convex approximation in Theorem 12 and Proposition 6. It proposes a safe approximation to the joint chance constraint (2.1b), and the safe approximation is second-order cone representable. The performance of this approach depends on the choice of a few tuning parameters. Although it is difficult to find the optimal setting, [34] designed an algorithm that is guaranteed to improve the choice of parameters. [34] also shows that it is possible to combine all the $\pi^i(x_0,x)$ functions in Theorem 12 together to reduce conservativeness.

The third approach directly dealing with joint chance constraints is the *data-driven robust optimization* proposed in [57]. It shows that by running different hypothesis tests on datasets, it is possible to construct different uncertainty sets that lead to safe approximations of the joint chance constraint (2.1b) with high probability. It is worth noting that the theoretical results in [57] holds for non-convex functions $f(x, \xi)$, albeit the resulting (RC) is very likely to be computationally intractable.

2.7 ConvertChanceConstraint (CCC): A Matlab Toolbox

Most existing optimization solvers cannot directly solve (CCO). All reviewed methods in Section 2.4-2.6 translate (CCO) to forms that can be recognized and solved by optimization solvers, e.g. SAA converts (CCO) to a mixed integer program (MIP), which can be solved by Gurobi. When solving a chance-constrained program, a typical approach is to write the converted formulation (e.g. the MIP of SAA) in the compact format that a solver recognizes then rely on the solver to get optimal solutions. This approach is unnecessarily repetitive as it needs to be repeated by different researchers on different problems. In addition, different solvers often take various input formats, thus this typical approach is limited to one specific solver. To overcome these issues, an interface or toolbox that automatically converts (CCO) to suitable forms for a variety of solvers is needed.

The remaining part of this subsection introduces the open-source Matlab toolbox *ConvertChance-Constraint* (CCC), which is developed to automate the process of converting chance constraints. CCC is written in Matlab, one of the most popular tools in engineering and many other fields. In consideration of flexibility in modeling and compatibility with existing solvers, CCC is built on YALMIP [84], a modeling language for optimization in Matlab. CCC is open-source on Github [7], other researchers and engineers could freely use, modify and improve it.

Figure 2.3 illustrates the logic flow when using CCC to solve and analyze a chance-constrained program. The problem is first formulated in the language of Matlab and YALMIP, then the chance constraint is modeled using the *prob()* function defined in CCC. After receiving the problem formulation and specified method to use (e.g. scenario approach), CCC translates the chance constraint to the formulation that YALMIP could understand. Then YALMIP interfaces with various solvers and further translates the problem for a specific solver. After optimization solver returns the optimal

[7] https://github.com/xb00dx/ConvertChanceConstraint-ccc

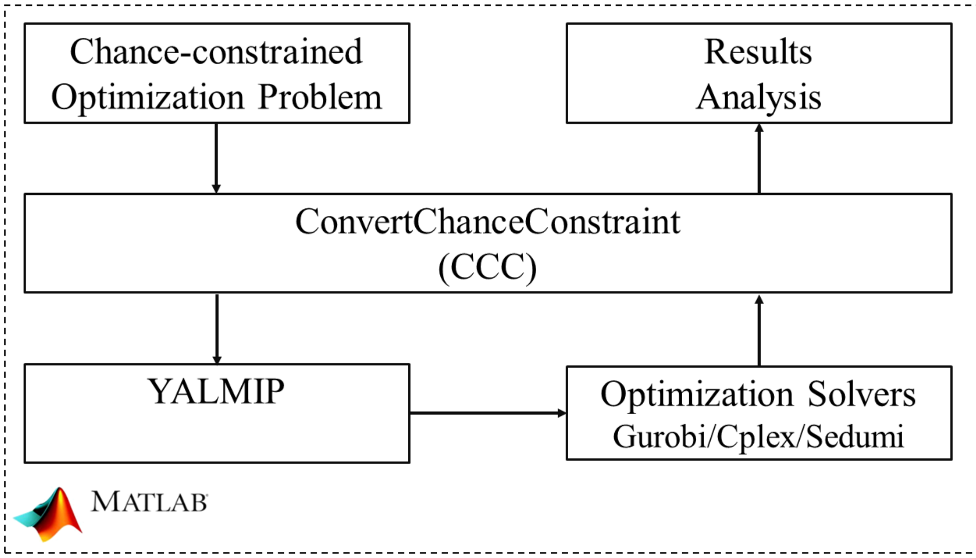


Figure 2.3: Solving and Analyzing a Chance-constrained Program via CCC, reprinted with permission from [1].

solution, CCC provides a few functions for result analysis, e.g. checking out-of-sample violation probability, calculating the posterior guarantees of the scenario approach.

Figure 2.4 presents the structure and main functions of CCC. Three major methods to solve (CCO) are implemented: scenario approach, sample average approximation and robust optimization related methods. The implementation of RO-related methods is based on the robust optimization module [85] of YALMIP. As illustrated in Figure 2.3 and 2.4, CCC is interfaced via YALMIP with most existing optimization solvers, e.g. Cplex [86], Gurobi [87], Mosek [88] and Sedumi [89].

2.8 Applications in Power Systems

A pivotal task in modern power system operation is to maintain the real-time balance of supply and demand while ensuring the system is low-cost and reliable. This pivotal task, however, faces critical challenges in the presence of rapid growth of renewable energy resources. Chance-constrained optimization, which explicitly models the risk that the system is exposed to, is a suitable conceptual framework to ensure the security and reliability of a power system under uncertainties.

There is a large body of literature adopting CCO for power system applications. Figure 2.5

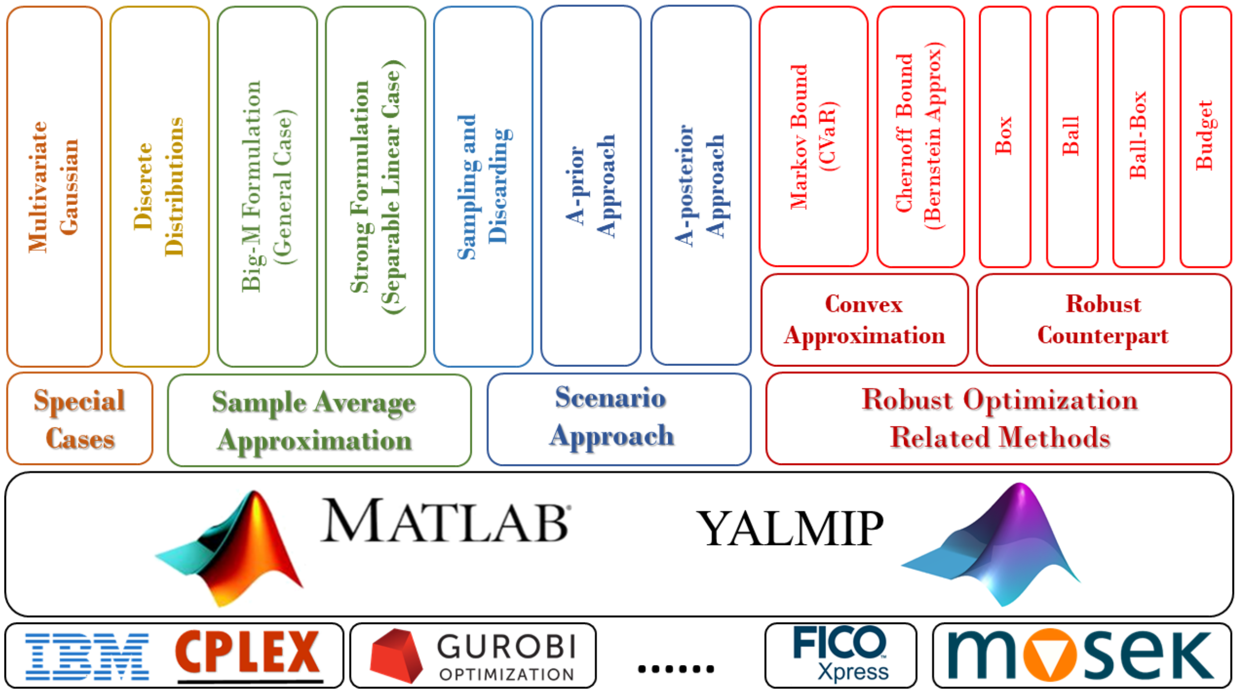


Figure 2.4: Structure and Main Functions of ConvertChanceConstraint, reprinted with permission from [1].

presents some existing applications of CCO in power systems. In the following chapters, we introduce three important applications of CCO in power systems: security-constrained economic dispatch (SCED) (Chapter 5), security-constrained unit commitment (SCUC) (5). and generation and transmission expansion.

Figure 2.5 also presents a feed-forward decision making framework for power system operations. The feed-forward framework partitions the overall decision making process into several time segments. The longer-term decisions (e.g. generation expansion) are fed into shorter-term decision making processes (e.g. unit commitment). The shorter-term decisions (e.g. generation commitment from SCUC) have direct impacts on real-time operations (e.g. dispatch results in SCED). As time draws closer to the actual physical operation, information gets much sharper and the prediction about future could be significantly improved [90].

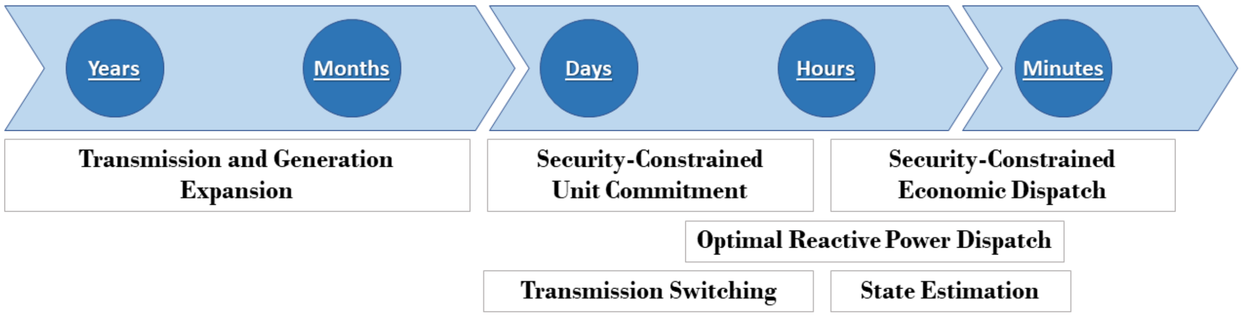


Figure 2.5: Representative Feed-forward Decisions Made in Power System Planning and Operation, reprinted with permission from [1].

2.9 Summary

This chapter presents a comprehensive review on the fundamental properties, key theoretical results, and three classes of algorithms for chance-constrained optimization. An open-source MATLAB toolbox ConvertChanceConstraint is developed to automate the process of translating chance constraints to compatible forms for mainstream optimization solvers. This chapter also briefly reviews three major applications of chance-constrained optimization in power systems. More applications of chance-constrained optimization in power systems are presented in Chapter 3, 4 and 5.

Table 2.1: Power System Applications of Chance-constrained Optimization, reprinted with permission from [1].

	Methods	Expansion	SCUC	SCED	Other Applications
Det-Equiv	Gaussian	[91] [92] [93] [94]	[95–97]	[98–112]	[113, 114]
Scenario Approach	a-priori	-	[115, 116]	[105, 117–124]	[125]
	a-posteriori	-	[115, 116, 126]	[123, 124]	-
SAA	-	[127]	[17, 128–132]	[124]	-
RO-based Approach	RLO	-	[133]	[124]	-
	Convex Approximation	-	-	[124, 134–136]	-
Others	-	[137, 138]	[139, 140]	[102, 107, 110, 118, 141, 142]	-

3. SECURITY-CONSTRAINED UNIT COMMITMENT WITH PROBABILISTIC GUARANTEES[1]

3.1 Motivation and Related Work

Security-constrained Unit commitment (SCUC) is one of the most important decisions made in power system operational planning. The SCUC problem seeks the most cost-efficient on/off decisions and dispatch schedules for generators, considering various security constraints such as generation and transmission capacity limits under contingencies.

SCUC is a decision making problem in uncertain environments by its nature. Conventional SCUC problems ensure the system is secured for a number of outages in generation, transmission, or other elements within the system. As the generation portfolio is shifting towards renewable resources, SCUC, a crucial part of power system day-ahead scheduling, needs to evolve to address the flexibility concerns.

Stochastic optimization (SO) and *robust* optimization (RO) are two common approaches for decision making under uncertainties. Both SO and RO have been successfully applied in various areas. SO relies on probabilistic models to depict uncertainties and often optimizes the objective function in the presence of randomness. SO has found many successful applications in power system operations and planning problems. For instance, references [143–145] formulate and solve the *stochastic unit commitment* problem, which minimizes the expected commitment and dispatch costs. RO takes an alternative approach, in which the uncertainty model is set-based and typically deterministic [54]. Recently, researchers in [146] formulated and solved the *robust unit commitment* problem, which minimizes the commitment and dispatch costs for the worst case in a predefined uncertainty set.

This chapter provides a perspective of solving SCUC in uncertain environments through the lens of chance-constrained optimization (CCO), which is akin to both SO and RO [1]. The main distinction between CCO and SO/RO is the *chance constraint* (see (3.1b) and (3.2b) in Section

[1]Reprinted with permission from [2].

3.2), which explicitly considers the feasibility of solutions under uncertainties. Various formulations of chance-constrained SCUC (c-SCUC) have been proposed, e.g. [16, 17, 96, 97, 128–132]. As mentioned in [1], chance-constrained optimization problems can be solved using the scenario approach, sample average approximation, or robust optimization based techniques. We take c-SCUC as an example. It is solved via sample average approximation in [17, 97, 128–132] and via robust optimization based techniques in [115, 126, 133][2].

The scenario approach is a well-known algorithm to solve CCO problems [60, 62, 63]. It was mainly targeted at convex problems (see Assumption 11), whereas SCUC is non-convex by nature due to on/off commitment decisions. Consequently, the scenario approach was considered *not applicable* for c-SCUC. An extended version of the scenario approach was proposed recently in [69], which makes it applicable for non-convex problems such as SCUC.

Our previous paper [148] might be the first attempt to apply the scenario approach on unit commitment[3]. However, the formulation therein is greatly simplified by ignoring some critical constraints such as transmission capacities. Enabled by this limiting assumption, [148] shows that the original scenario approach remains applicable in spite of the non-convexities from binary decision variables. Nonetheless, its main limitation is that the nice results in [148] only hold in the absence of transmission capacity constraints. We significantly improve [148] by considering additional security constraints such as line flow limits in the presence of uncertainties, and provides theoretical analysis on the results of the scenario approach.

The main contributions of this chapter are threefold. (1) We contribute to the non-convex scenario approach theory by proving salient structural properties of non-convex scenario problems, which extends the classical results for convex scenario problems published in [63]. (2) We formulate c-SCUC, which is later reformulated to scenario-based SCUC (s-SCUC) and solved via the scenario approach. To the best of our knowledge, we are the first to solve c-SCUC using the

[2]The method used in [115, 126] is based on [64]. It utilizes the sample complexity bound by an earlier version of the scenario approach [147], however, it is more closely related with robust optimization. Furthermore, the results in in [115, 126] might be overly conservative, since the sample complexity bound by [147] could be significantly tightened by in-depth analysis of the scenario approach, see Theorem 15 for more details.

[3]We call it unit commitment instead of SCUC because no security constraints are considered.

scenario approach while considering critical constraints such as transmission limits. (3) We design efficient algorithms to explore the structural properties of s-SCUC, which enables rigorous guarantees on the optimal solution returned by the scenario approach.

The remainder of this chapter is organized as follows. Section 3.2 summarizes the key results of the scenario approach for both convex and non-convex problems. Section 3.3 proves the structural properties of non-convex scenario problems. Section 3.4 formulates chance-constrained SCUC, which is solved via the scenario approach. Numerical results and discussions are in Section 3.5 and 3.6, respectively. Section 3.7 presents the concluding remarks. All proofs are available in Appendix A and on arXiv [2], the detailed settings of test systesm are in Appendix B.

3.2 Introduction to the Scenario Approach

This section first provides a brief introduction to chance-constrained optimization. Section 3.2.2 presents the main results of the scenario approach for convex problems. Recent progress in the scenario approach for non-convex problems are summarized in Section 3.2.3.

3.2.1 Chance-constrained Optimization

Chance-constrained optimization is a major approach for decision making in an uncertain environment. Since its birth in 1950s [18], chance-constrained optimization has been widely studied and successfully applied in various fields [1]. A typical formulation of chance-constrained optimization is presented below.

$$\min_{x} \; c^{\intercal} x \tag{3.1a}$$

$$\text{s.t. } \mathbb{P}_{\xi}\Big(f(x,\xi) \leq 0\Big) \geq 1-\epsilon \tag{3.1b}$$

$$g(x) \leq 0 \tag{3.1c}$$

We could write (3.1) in a more compact form by defining $\mathcal{X}_\xi := \{x \in \mathbf{R}^n : f(x,\xi) \leq 0\}$ and $\chi := \{x \in \mathbf{R}^n : g(x) \leq 0\}$

$$\min_{x\in\chi} \; c^\intercal x \tag{3.2a}$$

$$\text{s.t. } \mathbb{P}_\xi\Big(x \in \mathcal{X}_\xi\Big) \geq 1-\epsilon \tag{3.2b}$$

Without loss of generality, we assume that the objective is a linear function of decision variables $x \in \mathbf{R}^d$ [31]. Random vector $\xi \in \Xi$ denotes the source of uncertainties and Ξ is the support of ξ. Deterministic constraints (3.1c) are represented by set χ in (3.2). Constraint (3.1b) or (3.2b) is the *chance constraint*. The chance constraint (3.2b) requires the the inner constraint $x \in \mathcal{X}_\xi$ to be satisfied with probability at least $1-\epsilon$, where the violation probability ϵ is typically a small number (e.g. $1\%, 5\%$). In (3.2b), the set $\mathcal{X}_\xi$ depends on the realization of ξ and the probability is taken with respect to ξ.

Researchers have proposed many methods to solve chance-constrained optimization problems, e.g. sample average approximation [71], convex approximation [30], and scenario approach [60, 62, 63]. A detailed review and tutorial on chance-constrained optimization is in [1]. Compared with other methods, the scenario approach has many advantages such as computationally efficient and are applicable for a broad range of optimization problems.

3.2.2 The Scenario Approach for Convex Problems

The scenario approach (sometimes referred as scenario approximation) is one of the well-known solutions to chance-constrained optimization, but its strength is not well-understood until recently [1]. The scenario approach utilizes N independent and identically distributed (i.i.d.) scenarios $\mathcal{N} := \{\xi^1, \xi^2, \cdots, \xi^N\}$ to convert the chance-constrained program (3.1) to the *scenario problem* below:

$$\text{SP}(\mathcal{N}): \; \min_x \quad c^\intercal x \tag{3.3a}$$

$$\text{s.t. } f(x,\xi^1) \leq 0 \qquad : \mu^1 \tag{3.3b}$$

$$\vdots$$

$$f(x, \xi^N) \leq 0 \qquad : \mu^N \tag{3.3c}$$

$$g(x) \leq 0 \qquad : \lambda \tag{3.3d}$$

The scenario problem SP($\mathcal{N}$) seeks the optimal solution $x^*_{\mathcal{N}}$ that is feasible for all N scenarios. The Lagrangian multiplier associated with the ith scenario constraint $f(x, \xi^i) \leq 0$ is denoted by $\mu^i \in \mathbf{R}^m$. We can write the scenario problem SP($\mathcal{N}$) in a similar way with (3.2) by defining $\mathcal{X}_i := \{x \in \mathbf{R}^n : f(x, \xi^i) \leq 0\}$.

$$\text{SP}(\mathcal{N}): \ \min_{x \in \mathcal{X}} \ c^\intercal x \tag{3.4a}$$

$$\text{s.t. } x \in \cap_{i=1}^N \mathcal{X}_i \tag{3.4b}$$

Definition 15 (Violation Probability)**.** The *violation probability* of a candidate solution $x^\diamond$ is defined as the probability that $x^\diamond$ is infeasible:

$$\mathbb{V}(x^\diamond) := \mathbb{P}_\xi\big(x^\diamond \notin \mathcal{X}_\xi\big). \tag{3.5}$$

The scenario approach theory aims at answering the following *sample complexity* question: what is the smallest sample size N such that $x^*_{\mathcal{N}}$ is feasible (i.e. $\mathbb{V}(x^*_{\mathcal{N}}) \leq \epsilon$) to the original chance-constrained program (3.2)? Reference [62, 63] provide in-depth analysis based on the concept of support scenarios.

Definition 16 (Support Scenario [62, 63])**.** Scenario ξ^i is a *support scenario* for the scenario problem SP($\mathcal{N}$) if its removal changes the solution of SP($\mathcal{N}$).

Let $x^*_{\mathcal{N}}$ and $x^*_{\mathcal{N}-i}$ stand for the optimal solution to scenario problems SP($\mathcal{N}$) and SP($\mathcal{N} - i$), respectively. Then scenario ξ^i is a support scenario if $c^\intercal x^*_{\mathcal{N}-i} < c^\intercal x^*_{\mathcal{N}}$. We use $\mathcal{S}(\mathcal{N})$ ($\mathcal{S}$ in short) to represent the set of all support scenarios of SP($\mathcal{N}$).

Definition 17 (Non-degenerate Scenario Problem [62,63]). Let $x^*_{\mathcal{N}}$ and $x^*_{\mathcal{S}}$ be the optimal solutions to the scenario problems SP($\mathcal{N}$) and SP($\mathcal{S}$), respectively. The scenario problem SP($\mathcal{N}$) is said to be *non-degenerate*, if $c^\intercal x^*_{\mathcal{N}} = c^\intercal x^*_{\mathcal{S}}$.

Assumption 9 (Non-degeneracy [62, 63]). *For every N, the scenario problem SP($\mathcal{N}$) is non-degenerate with probability 1 with respect to scenarios $\mathcal{N} = \{\xi^1, \xi^2, \cdots, \xi^N\}$.*

Assumption 10 (Feasibility [62]). *Every scenario problem SP($\mathcal{N}$) is feasible, and its feasibility region has a non-empty interior. The optimal solution $x^*_{\mathcal{N}}$ of SP($\mathcal{N}$) exists.*

Definition 18 (Helly's Dimension [63]). Helly's dimension of the scenario problem SP($\mathcal{N}$) is the least integer h that $h \geq \operatorname{ess\,sup}_{\mathcal{N} \subseteq \Xi^N} |\mathcal{S}(\mathcal{N})|$ holds for any finite $N \geq 1$, where $|\mathcal{S}(\mathcal{N})|$ is the number of support scenarios.

Theorem 15 presents one of the most important results in the scenario approach theory, which is based on the non-degeneracy and feasibility assumptions.

Theorem 15 (Exact Feasibility [62, 63]). *Under Assumptions 9 (non-degeneracy) and 10 (feasibility), let $x^*_{\mathcal{N}}$ be the optimal solution to the scenario problem SP($\mathcal{N}$), it holds that*

$$\mathbb{P}^N\Big(\mathbb{V}(x^*_{\mathcal{N}}) > \epsilon\Big) \leq \sum_{i=1}^{h-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{3.6}$$

The probability $\mathbb{P}^N$ is taken with respect to N random scenarios $\mathcal{N} = \{\xi^i\}_{i=1}^N$, and h is the Helly's dimension of SP($\mathcal{N}$).

Stronger results without the feasibility assumption are in [62, 63]. Based on Theorem 15, the scenario approach answers the sample complexity question in Corollary 2.

Corollary 2 (Sample Complexity [62,63]). *Under Assumptions 9 (non-degeneracy) and 10 (feasibility), given a violation probability $\epsilon \in (0, 1)$ and a confidence parameter $\beta \in (0, 1)$, if we choose the smallest number of scenarios N such that*

$$\sum_{i=0}^{h-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta, \tag{3.7}$$

then it holds that

$$\mathbb{P}^N\Big(\mathbb{V}(x^*_{\mathcal{N}}) \leq \epsilon\Big) \geq 1-\beta, \tag{3.8}$$

where $x^*_{\mathcal{N}}$ is the optimal solution to SP($\mathcal{N}$), and h is the Helly's dimension of SP($\mathcal{N}$) ($0 \leq h \leq N$).

The scenario approach is essentially a randomized algorithm to solve chance-constrained optimization problems. The randomness of the scenario approach comes from drawing i.i.d. scenarios. The confidence parameter β quantifies the risk of failure due to drawing scenarios from a "bad" set. Corollary 2 shows that by choosing a proper number of scenarios, the corresponding optimal solution $x^*_{\mathcal{N}}$ is feasible (i.e. $\mathbb{V}(x^*_{\mathcal{N}}) \leq \epsilon$) with confidence at least $1-\beta$.

Assumption 11 (Convexity)**.** *The deterministic constraint $g(x) \leq 0$ is convex, and the random constraint $f(x,\xi)$ is convex in x for every instance of ξ. In other words, the sets χ and $\mathcal{X}_i$s in* (3.4) *are convex.*

Theorem 16 ([60, 63])**.** *Under Assumption 10 and 11, the number of support scenarios $|\,\mathcal{S}\,|$ for SP($\mathcal{N}$) is at most n. In other words, $h \leq n$, where n is the number of decision variables $x \in \mathbf{R}^n$ and h is Helly's dimension.*

For convex scenario problems SP($\mathcal{N}$), we could replace h by n in Theorem 15 and Corollary 2. This leads to the classical results of the scenario approach in [60, 62, 63].

Remark 9 (Towards Non-convexity). Theorem 15 and Corollary 2 do not assume convexity of $f(x,\xi)$ and $g(x)$. In theory, Theorem 15 and Corollary 2 are applicable for non-convex scenario problems if a feasible non-convex SP($\mathcal{N}$) is proved to be non-degenerate with probability 1 (e.g. [148]). In practice, however, the scenario approach was considered *not applicable* for non-convex problems. Comprehensive analysis are presented in Section 3.2.3.

3.2.3 The Scenario Approach for Non-convex Problems

The scenario approach was considered *not applicable* for non-convex problems for the following three reasons: (1) non-convexity causes degeneracy; (2) non-trivial bounds on $|\,\mathcal{S}\,|$ may not exist for non-convex SP($\mathcal{N}$); and (3) it is computationally intractable to find optimal solutions.

First, degeneracy is a common issue for non-convex problems, e.g. the scenario-based SCUC problem in Section 3.4.3. Since the non-degeneracy assumption 9 lies at the heart of the scenario approach theory, almost all results in the literature are for non-degenerate problems.

Second, it is almost impossible to prove non-trivial and practical bounds on the number of support scenarios $|\,\mathcal{S}\,|$ for non-convex problems. Reference [69] presents one extreme case, in which every scenario is a support scenario thus $|\,\mathcal{S}\,| = N$ [4]. In addition, a loose bound typically leads to an astronomical sample complexity N, which make the scenario approach unpractical. For instance, loose bounds on $|\,\mathcal{S}\,|$ for scenario-based unit commitment will require $10^3 \sim 10^4$ times more scenarios than necessary [148].

Furthermore, the most attractive feature of convex optimization is that any *local* minimum is a *global* minimum. And there exist a broad family of efficient algorithms that compute global optimal solutions for convex problems. Hence, $x^*_{\mathcal{N}}$ in Section 3.2.2 refers to the *global* optimal solution by default. It is worth noting that $x^*_{\mathcal{N}}$ is solely determined by the scenario problem SP($\mathcal{N}$) and it is *not* algorithm-dependent.

For non-convex problems, however, it is often computationally intractable to find global optimal solutions. There are many algorithms that are capable of finding *local* optimal solutions in a relatively short time. Therefore, it is more reasonable and practical to analyze the characteristics of *local* solutions for non-convex scenario problems. Algorithm $\mathbb{A} : \Xi^N \rightarrow \mathbf{R}^n$ stands for the process of finding solutions to SP($\mathcal{N}$), e.g. primal-dual interior-point method. We use $\text{opx}_{\mathbb{A}}(\mathcal{N})$ to represent a (possibly suboptimal) solution to SP($\mathcal{N}$) obtained via algorithm $\mathbb{A}$. The corresponding optimal objective value is denoted by $\text{opt}_{\mathbb{A}}(\mathcal{N})$. The subscript $\mathbb{A}$ emphasizes the fact that the solution is algorithm-dependent. And we use $\text{SP}_{\mathbb{A}}(\mathcal{N})$ to represent a scenario problem solved by algorithm $\mathbb{A}$.

Consequently, the scenario approach was considered *not applicable* for non-convex problems until very recently. By removing the non-degeneracy assumption and analyzing any feasible solutions of non-convex scenario problems, reference [69] develops a general theory for the scenario

[4] Using the trivial bound $|\,\mathcal{S}\,| \leq N$, Theorems 15 and 17 provide guarantees $\mathbb{P}(\mathbb{V}(x^*_{\mathcal{N}}) > \epsilon) \leq 1$, which is useless.

approach. This subsection summarizes its key results.

Identical to the convex case in Section 3.2.2, the scenario approach converts (3.2) to the scenario problem (3.4) using N scenarios $\mathcal{N} = \{\xi^1, \xi^2, \cdots, \xi^N\}$ for non-convex problems. The sets χ and $\mathcal{X}_\xi$ here could be non-convex.

Definition 19 (Invariant Set)**.** Let $\text{opt}_{\mathbb{A}}(\mathcal{M})$ be the optimal value of SP$(\mathcal{M})$ found by algorithm $\mathbb{A}$ for a scenario problem SP$(\mathcal{M})$. A set of scenarios $\mathcal{I}$ is an invariant (scenario) set for $\text{SP}_{\mathbb{A}}(\mathcal{N})$ if $\text{opt}_{\mathbb{A}}(\mathcal{I}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$.

The concept of invariant set is an extension of support scenarios for (possibly degenerate) non-convex scenario problems. A trivial invariant set is $\mathcal{I} = \mathcal{N}$. Algorithm $\mathbb{B} : \Xi^N \to \mathcal{I}$ represents the process of finding non-trivial invariant sets. Examples of Algorithm $\mathbb{B}$ can be found in Section 3.3 and Appendix A.1.

Theorem 17 (Posterior Guarantees for Non-convex Scenario Problems [69][5])**.** *Suppose Assumption 10 (feasibility) holds true and $\beta \in (0, 1)$ is given. Algorithm $\mathbb{A}$ solves the scenario problem SP$(\mathcal{N})$ and obtains an optimal solution $opx_{\mathbb{A}}(\mathcal{N})$. Algorithm $\mathbb{B}$ finds an invariant set $\mathcal{I}$ of cardinality $|\,\mathcal{I}\,|$. The following probabilistic guarantee holds*

$$\mathbb{P}^N\Big(\mathbb{V}\big(\,opx_{\mathbb{A}}(\mathcal{N})\big) \le \epsilon(N, |\,\mathcal{I}\,|, \beta)\Big) \ge 1 - \beta,$$

where the function $\epsilon(k, N, \beta)$ is defined as

$$\epsilon(N, k, \beta) := \begin{cases} 1 & \text{if } k = N, \\ 1 - \left(\frac{\beta}{N\binom{N}{k}}\right)^{\frac{1}{N-k}} & \text{otherwise.} \end{cases} \tag{3.9}$$

Lemma 2. *The $\epsilon(N, k, \beta)$ function defined in* (3.9) *has the following properties: (1) $\epsilon(N, k, \beta)$ is monotonically decreasing in β; (2) $\epsilon(N, k, \beta)$ is monotonically increasing in k; (3) $\epsilon(N, k, \beta)$ is monotonically decreasing in N.*

[5]Theorem 17 is a simplified version of the main result in [69], the feasibility assumption 10 is a simplified version of the admissible assumption in [69].

In order to achieve an ϵ-level solution with confidence $1-\beta$, Lemma 2 shows that the least conservative result (i.e. smallest sample complexity N) is achieved with the invariant set of minimal cardinality, which is defined as the essential set.

Definition 20 (Essential Set [63])**.** A set of scenarios $\mathcal{E} \subseteq \mathcal{N}$ is an *essential* (scenario) set for $\text{SP}_{\mathbb{A}}(\mathcal{N})$ if

$$\mathcal{E} := \arg\min\{|\,\mathcal{E}\,| : \text{opt}_{\mathbb{A}}(\mathcal{E}) = \text{opt}_{\mathbb{A}}(\mathcal{N}), \mathcal{E} \subseteq \mathcal{N}\}. \tag{3.10}$$

In other words, $\mathcal{E}$ is an invariant set of minimal cardinality.

One key step in the non-convex scenario approach is designing algorithms $\mathbb{B}$ to search for essential sets. Section 3.3 reveals the structure of general non-convex scenario problems, which lays the cornerstone for algorithms to obtain essential sets. Section 3.3 also gives one example of designing more efficient algorithms by exploiting the structural properties of specific problems.

3.3 Structural Properties of General Scenario Problems

Searching for essential sets is an important step in the non-convex scenario approach. However, the only known general algorithm to obtain essential sets is enumerating all 2^N possibilities by solving 2^N non-convex problems. This implies that searching for essential sets is in general computationally prohibitive. Section 3.3.1 first demonstrates the structural properties for general non-convex scenario problems, and proves a few special cases that finding essential sets is relatively easier. Section 3.3.2 reveals the connection between non-convex and convex scenario problems. Motivated by the structure of security-constrained unit commitment, Section 3.3.3 illustrates an efficient algorithm to track down essential sets for two-stage scenario problems.

3.3.1 Non-convex Scenario Problems

Instead of solving 2^N non-convex problems to obtain essential sets, there are two ideas to track down invariant sets with small cardinalities (not necessarily essential): (1) removing each scenario and checking if the objective changes, this idea leads to the definition of support sets; (2) removing scenarios one by one, until the scenario set cannot be further reduced, this leads to the definition of irreducible set.

Definition 21 (Support Scenario of $\text{SP}_{\mathbb{A}}(\mathcal{N})$)**.** Scenario $\xi^i \in \mathcal{N}$ is a *support scenario* for the scenario problem $\text{SP}_{\mathbb{A}}(\mathcal{N})$ if its removal changes the solution $\text{opt}_{\mathbb{A}}(\mathcal{N})$ of $\text{SP}_{\mathbb{A}}(\mathcal{N})$. The set of support scenarios (support set in short) is denoted by $\mathcal{S}_{\mathbb{A}}$.

Definition 22 (Irreducible Set)**.** A scenario set $\mathcal{R} \subseteq \mathcal{N}$ for $\text{SP}_{\mathbb{A}}(\mathcal{N})$ is *irreducible*, if (1) it is invariant, i.e. $\text{opt}_{\mathbb{A}}(\mathcal{R}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$; and (2) $\text{opt}_{\mathbb{A}}(\mathcal{R} - s) < \text{opt}_{\mathbb{A}}(\mathcal{R}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$ for any $s \in \mathcal{R}$.

Assumption 12 (Monotonicity)**.** *Let $\mathbb{A} : \Xi^N \to \mathbf{R}^n$ be an algorithm to obtain the optimal solution of a scenario problem SP($\mathcal{N}$), whose optimal objective value is represented by $opt_{\mathbb{A}}(\mathcal{N})$. We assume that the algorithm $\mathbb{A}$ always satisfies $opt_{\mathbb{A}}(\mathcal{M}) \leq opt_{\mathbb{A}}(\mathcal{N})$ if $\mathcal{M} \subseteq \mathcal{N}$.*

Assumption 12 is indeed a weak assumption. Considering two scenario problems SP($\mathcal{N}$) and SP($\mathcal{M}$) with $\mathcal{M} \subseteq \mathcal{N}$. Because the optimal solution to SP($\mathcal{N}$) will be always feasible to SP($\mathcal{M}$), algorithm $\mathbb{A}$ could use $\text{opx}_{\mathbb{A}}(\mathcal{N})$ as a starting point and obtain solution $\text{opx}_{\mathbb{A}}(\mathcal{M})$ that is not worse than $\text{opx}_{\mathbb{A}}(\mathcal{N})$.

Lemma 3 (Modified Lemma 2.10 of [63])**.** *Suppose algorithm $\mathbb{A}$ satisfies Assumption 12. Let $\mathcal{I}$ be any* invariant *set for a (possibly non-convex) scenario problem $SP_{\mathbb{A}}(N)$ and $\mathcal{S}$ stands for its support set, then $\mathcal{S} \subseteq \mathcal{I}$. Since any essential set $\mathcal{E}$ or irreducible set $\mathcal{R}$ is also invariant, then $\mathcal{S} \subseteq \mathcal{E}$ and $\mathcal{S} \subseteq \mathcal{R}$.*

Lemma 3 reveals the key relationship among the support set, essential and irreducible sets, and it lays the foundation of more important observations in Corollary 3 and 4. Lemma 3 is a generalized version of Lemma 2.10 in [63], which proved similar results for convex scenario problems. The importance of Lemma 3 is to show that the key assumption for such structural properties is the monotonicity of algorithm $\mathbb{A}$, instead of convexity (Assumption 11 in [63]).

For general (non-convex) scenario problems, the support set, essential set and irreducible set are different. Under certain circumstances, these three concepts are interchangeable. Such circumstances are depicted by an extended definition of non-degeneracy for non-convex scenario problems.

Definition 23 (Non-degeneracy of $\text{SP}_{\mathbb{A}}(\mathcal{N})$)**.** For a general scenario problem $\text{SP}_{\mathbb{A}}(\mathcal{N})$, let $\mathcal{N}$ stand for the set of all N scenarios and $\mathcal{S}$ denote the support (scenario) set. The scenario problem $\text{SP}_{\mathbb{A}}(\mathcal{N})$ is said to be *non-degenerate*, if $\text{opt}_{\mathbb{A}}(\mathcal{N}) = \text{opt}_{\mathbb{A}}(\mathcal{S})$.

Corollary 3. *Consider a (possibly non-convex) scenario problem $SP_{\mathbb{A}}(\mathcal{N})$ and an algorithm $\mathbb{A}$ satisfying Assumption 12. If $SP_{\mathbb{A}}(\mathcal{N})$ is non-degenerate, then (1) it has a unique essential set $\mathcal{E} = \mathcal{S}$; and (2) it has a unique irreducible set $\mathcal{R} = \mathcal{S}$.*

Corollary 4. *Consider a (possibly non-convex) scenario problem $SP_{\mathbb{A}}(\mathcal{N})$ and an algorithm $\mathbb{A}$ satisfying Assumption 12. The following three statements are equivalent with each other: (1) $SP_{\mathbb{A}}(\mathcal{N})$ is non-degenerate; (2) $SP_{\mathbb{A}}(\mathcal{N})$ has a unique irreducible set $\mathcal{R}$; and (3) $SP_{\mathbb{A}}(\mathcal{N})$ has a unique essential set $\mathcal{E}$.*

Corollaries 3 and 4 provide key insights in designing an efficient algorithm $\mathbb{B}$. For non-convex problems, even if Assumption 9 does not always hold, $\text{SP}_{\mathbb{A}}(\mathcal{N})$ might be non-degenerate in many instances (e.g. s-SCUC is non-degenerate in 192 out of 200 instances in Section 3.5.4). For those non-degenerate scenario problems, Corollary 3 and 4 show that we are able to find the essential set by solving only N instead of 2^N non-convex problems. Section 3.3.3 shows that the computational burden to obtain essential sets can be further reduced by exploiting the structure of specific problems.

Remark 10 (Finding Essential Sets for Non-degenerate Problems)*.* When a scenario problem is non-degenerate, we can obtain the (unique) essential set by searching for the support set or irreducible set (Corollary 3). Algorithms of finding an irreducible set (Algorithm 3 in Appendix A.1) or the support set (e.g. Algorithm 1) are based on definition. More discussions on finding the support set are in Remark 11.

3.3.2 Convex Scenario Problems

For convex scenario problems $\text{SP}(\mathcal{N})$, any local minimum is a global minimum. And a broad range of algorithms to look for global optimal solutions exist. In the convex setting, we assume any algorithm $\mathbb{A}$ returns global optimal solutions to $\text{SP}(\mathcal{N})$ by default. In Section 3.2.2 and 3.3.2, we replace $\mathrm{opx}_{\mathbb{A}}(\mathcal{N})$ and $\mathrm{opt}_{\mathbb{A}}(\mathcal{N})$ by $x^*_{\mathcal{N}}$ and $c^\intercal x^*_{\mathcal{N}}$, respectively. We also remove subscripts $\mathbb{A}$ since the definition of support set, invariant set and essential set for convex problems are no longer algorithm-dependent.

Lemma 4 (Monotonicity)**.** *Let $x^*_{\mathcal{N}}$ and $x^*_{\mathcal{M}}$ stand for the global optimal solution to the (convex) scenario problems SP$(\mathcal{N})$ and SP$(\mathcal{M})$, respectively. Then $c^\intercal x^*_{\mathcal{M}} \leq c^\intercal x^*_{\mathcal{N}}$ if $\mathcal{M} \subseteq \mathcal{N}$.*

Because $x^*_{\mathcal{N}}$ is always feasible to SP$(\mathcal{M})$ and $x^*_{\mathcal{M}}$ is globally optimal, it is obvious that $c^\intercal x^*_{\mathcal{M}} \leq c^\intercal x^*_{\mathcal{N}}$. Lemma 4 shows that any algorithm obtaining global optimal solutions will automatically satisfy Assumption 12. Therefore, all results in Section 3.3.1 hold for convex scenario problems. It is worth mentioning that similar results for convex problems were first proved in [63]. Section 3.3.1 can be regarded as an extension of classical results in [63] towards non-convex scenario problems.

Remark 11 (Finding Support Scenarios For Convex Problems). The first algorithm of searching for support scenarios (for both convex and non-convex scenario problems, Algorithm 2 in Appendix A.1) is based on definition, i.e. checking if the removal of a scenario changes the optimal solution. Algorithm 2 requires solving N scenario problems. In many cases (especially in power system applications, e.g. [123]), it is observed that the support scenarios are only a small subset of all N scenarios, i.e. $|\,\mathcal{S}\,| \ll |\,\mathcal{N}\,|$. This observation indicates the dual solution $\mu^1, \mu^2, \cdots, \mu^N$ to SP$(\mathcal{N})$ is often sparse. Lemma 5 formalizes this observation and provides an approach to narrow down the range of searching for support scenarios. Built upon Lemma 5, Algorithm 1 only requires solving $\sim |\,\mathcal{S}\,|$ scenario problems, which is much more efficient than Algorithm 2 since $|\,\mathcal{S}\,| \ll |\,\mathcal{N}\,|$.

Lemma 5. *Consider a non-degenerate and convex scenario problem SP$(\mathcal{N})$ which has at least one* strictly *feasible solution. If ξ^i is a support scenario ($i \in \mathcal{S}$), then $\|\mu^{i,*}\| > 0$, where $\mu^{i,*} \in \mathbf{R}^m$ is the optimal dual solution of SP$(\mathcal{N})$. In other words, let $\mathcal{M} := \{i \in \mathcal{N} : \|\mu^{i,*}\| > 0\}$, then $\mathcal{S} \subseteq \mathcal{M}$.*

Algorithm 1 Finding Support Scenarios Using Dual Variables, reprinted with permission from [2].

Compute the primal and dual solutions $x^*_{\mathcal{N}}$ and $\mu^{i,*}$ $(i = 1, 2, \cdots, N)$ by solving SP$(\mathcal{N})$

Let $\mathcal{M} = \{i \in \mathcal{N} : \|\mu^{i,*}\| > 0\}$. Set $\mathcal{S} \leftarrow \emptyset$.

for $i \in \mathcal{M}$ **do**

  Solve SP$_{\mathcal{M}-i}$ and compute $x^*_{\mathcal{M}-i}$

  if $c^\intercal x^*_{\mathcal{M}-i} < c^\intercal x^*_{\mathcal{N}} (= c^\intercal x^*_{\mathcal{M}})$ **then**

    $\mathcal{S} \leftarrow \mathcal{S} + i$

  end if

end for

In many cases, Algorithm 1 only needs to solve the dual problem of SP$(\mathcal{N})$, it may not be necessary to solve the primal solution $x^*_{\mathcal{N}}$. We use $x^*_{\mathcal{N}}$ in Algorithm 1 mainly for the purpose of notation simplicity.

3.3.3 Two-stage Scenario Problems

Section 3.3.1 shows that searching for essential sets can be relatively easier when a scenario problem is non-degenerate. However, finding a support set or irreducible set still requires solving N non-convex problems. Motivated by SCUC, we show that more efficient algorithms are possible by exploiting the structure of specific problems. We study the following two-stage scenario problem in this subsection.

$$\min_{y \in \mathcal{Y}} c_y^\intercal y + \min_{\substack{x \in \mathcal{X} \\ (x,y) \in \mathcal{H}}} c_x^\intercal x \tag{3.11a}$$

$$\text{s.t. } x \in \cap_{i=1}^N \mathcal{U}_i \tag{3.11b}$$

Constraints on the first-stage variables y and the second-stage variables x are denoted by $y \in \mathcal{Y}$ and $x \in \mathcal{X}$, respectively. Constraint $(x, y) \in \mathcal{H}$ represents the constraints coupling variables x and y in both stages. Set $\mathcal{U}_i$ stands for the constraints corresponding to the ith scenario ξ^i.

Problem (3.11) is an abstract form of s-SCUC in Section 3.4. Two key features of the two-stage scenario problem are: (1) the non-convexity *only* comes from constraints $y \in \mathcal{Y}$ (e.g. binary variables in SCUC), all other constraints ($\mathcal{X}, \mathcal{H}, \mathcal{U}_i$) are convex; (2) uncertainties only exist in the second stage.

Let (x^*, y^*) be a (possibly local) optimal solution that algorithm $\mathbb{A}$ returns. Given $y = y^*$, the second stage problem is convex by setting:

$$\min_{\substack{x \in \mathcal{X} \\ (x, y^*) \in \mathcal{H}}} \quad c_x^\intercal x \tag{3.12a}$$

$$\text{s.t. } x \in \cap_{i=1}^{N} \mathcal{U}_i \tag{3.12b}$$

Lemma 6. *(1) Let $\hat{\mathcal{S}}$ represent the set of support scenarios of* (3.12) *and $\mathcal{S}$ denote the support set for the two-stage problem* (3.11)*, then $\hat{\mathcal{S}} \subseteq \mathcal{S}$; (2) If $\hat{\mathcal{S}}$ is invariant for* (3.11)*, i.e. opt$_{\mathbb{A}}(\hat{\mathcal{S}}) =$ opt$_{\mathbb{A}}(\mathcal{N})$, then the two-stage scenario problem SP$_{\mathbb{A}}(\mathcal{N})$ is non-degenerate.*

Corollaries 3 and 4 demonstrate many nice properties of non-degenerate scenario problems. Lemma 6 gives a criteria of checking if the two-stage problem (e.g. s-SCUC) is non-degenerate. This lemma lays the foundation of Algorithm 4 to search for essential sets of (3.11). The main idea of Algorithm 4 is to first find the support scenarios of the second-stage problem (3.12), then verify if SP($\mathcal{N}$) is degenerate using Lemma 6. In Section 3.5.4, it turns out that s-SCUC is non-degenerate in 96% of cases, thus Algorithm 4 could obtain essential sets of s-SCUC (in Section 3.5.4) in a much shorter time.

3.4 Security-Constrained Unit Commitment with Probabilistic Guarantees

3.4.1 Nomenclature

The number of loads, generators, wind farms, transmission lines, contingencies, and snapshots are denoted by n_d, n_g, n_w, n_l, n_k and n_t, respectively.

$k \in \{0, 1, \cdots, n_k\}$ contingency index

$t \in \{0, 1, \cdots, n_t\}$ time (snapshot) index

$\iota \in \{t+1, \cdots, n_t\}$ additional time (snapshot) index in constraints (3.13j) and (3.13k)

Binary decision variables (at time t):

$z^t \in \{0,1\}^{n_g}$ generator on/off states (commitment)

$u^t \in \{0,1\}^{n_g}$ generator i is on if $u_i^t = 1$

$v^t \in \{0,1\}^{n_g}$ generator i is off if $v_i^t = 1$

Continuous decision variables (at time t, contingency k):

$g^{t,k} \in \mathbf{R}^{n_g}$ generation output

$r^t \in \mathbf{R}^{n_g}$ reserve

Parameters and constants:

$a^k \in \{0,1\}^{n_g}$ generator availability in contingency k

$\alpha_k \in \mathbf{R}_+$ weight of contingency k

$c_g \in \mathbf{R}^{n_g}$ generation costs

$c_z \in \mathbf{R}^{n_g}$ no load cost

$c_r \in \mathbf{R}^{n_g}$ reserve costs

$c_u \in \mathbf{R}^{n_g}$ startup cost

$c_v \in \mathbf{R}^{n_g}$ shutdown cost

$\hat{d}^t \in \mathbf{R}^{n_d}$ load forecast (time t)

$\tilde{d}^t \in \mathbf{R}^{n_d}$ load forecast error (time t)

$\hat{w}^t \in \mathbf{R}^{n_w}$ wind forecast (time t)

$\tilde{w}^t \in \mathbf{R}^{n_w}$ wind forecast error (time t)

$\overline{g} \in \mathbf{R}^{n_g}$ generation upper bounds

$\underline{g} \in \mathbf{R}^{n_g}$ generation lower bounds

$\overline{\gamma} \in \mathbf{R}^{n_g}$ ramping upper bounds

$\underline{\gamma} \in \mathbf{R}^{n_g}$ ramping lower bounds

$\underline{u}_i \in \mathbf{R}_+$ minimum on time for generator i

$\underline{v}_i \in \mathbf{R}_+$ minimum off time for generator i

3.4.2 Deterministic Security-constrained Unit Commitment

Deterministic security-constrained unit commitment (d-SCUC) (3.13) seeks optimal commitment and startup/shutdown decisions (z^t, u^t, v^t), generation and reserve schedules $(g^{t,k}, r^t)$ for a horizon of time steps, typically $24 \sim 36$ hours. The d-SCUC problem is being solved as a crucial part of the day-ahead market operation. Security constraints ensures the reliability of the power system after an unexpected event occurs.

$$\min_{z,u,v,g,r} \sum_{t=1}^{n_t} \left(c_z^\intercal z^t + c_u^\intercal u^t + c_v^\intercal v^t + c_r^\intercal r^t + \sum_{k=0}^{n_k} \alpha_k c_g^\intercal g^{t,k} \right) \tag{3.13a}$$

$$\text{s.t. } \mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal \hat{w}^t \geq \mathbf{1}^\intercal \hat{d}^t \tag{3.13b}$$

$$\underline{f} \leq H_g^{t,k} g^{t,k} + H_w^{t,k} \hat{w}^{t,k} - H_d^{t,k} \hat{d}^{t,k} \leq \overline{f} \tag{3.13c}$$

$$a^k \circ \underline{\gamma} \leq g^{t,k} - g^{t-1,k} \leq a^k \circ \overline{\gamma} \tag{3.13d}$$

$$a^k \circ (g^{t,0} - r^t) \leq g^{t,k} \leq a^k \circ (g^{t,0} + r^t) \tag{3.13e}$$

$$k \in [0, n_k], t \in [1, n_t]$$

$$\underline{g} \circ z^t \leq g^{t,0} \leq \overline{g} \circ z^t \tag{3.13f}$$

$$\underline{g} \circ z^t \leq g^{t,0} - r^t \leq g^{t,0} + r^t \leq \overline{g} \circ z^t \tag{3.13g}$$

$$z^{t-1} - z^t + u^t \geq 0 \tag{3.13h}$$

$$z^t - z^{t-1} + v^t \geq 0 \tag{3.13i}$$

$$t \in [1, n_t]$$

$$z_i^t - z_i^{t-1} \leq z_i^\iota,\ i \in [1, n_g] \tag{3.13j}$$

$$\iota \in [t+1, \min\{t + \underline{u}_i - 1, n_t\}], t \in [2, n_t]$$

$$z_i^{t-1} - z_i^t \leq 1 - z_i^\iota,\ i \in [1, n_g] \tag{3.13k}$$

$$\iota \in [t+1, \min\{t + \underline{v}_i - 1, n_t\}], t \in [2, n_t]$$

The objective of (3.13) is to minimize total operation costs, including no-load costs $c_z^\intercal z^t$, startup costs $c_u^\intercal u^t$, shutdown costs $c_v^\intercal v^t$, generation costs $c_g^\intercal g^{t,k}$ and reserve costs $c_r^\intercal s^t$. Security constraints ensure: enough supply to meet demand (3.13b), transmission line flow within limits (3.13c), generation levels within ramping limits (3.13d) and capacity limits (3.13f) in any contingency k. Constraints (3.13e) and (3.13g) are about the relationship between generation and reserve in any contingency k. Constraints (3.13h)-(3.13i) are the logistic constraints about commitment status, startup and shutdown decisions. Minimum on/off time constraints for all generators are in (3.13j)-(3.13k). Constraints (3.13d)-(3.13g) also guarantee the consistency of generation levels $g^{t,k}$ with commitment decisions z^t and generator availability a^k in contingency k [148].

The d-SCUC formulation utilizes the expected wind generation and load forecast, it does not take the uncertainties from wind and load into consideration. We propose an improved formulation of d-SCUC using chance constraints, which explicitly guarantee the system security with a tunable level of risk ϵ with respect to uncertainties.

$$\mathbb{P}_{\tilde{w} \times \tilde{d}} \Big(\mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal (\hat{w}^t + \tilde{w}^t) \geq \mathbf{1}^\intercal (\hat{d}^t + \tilde{d}^t), \tag{3.14a}$$

$$\underline{f} \leq H_g^{t,k} g^{t,k} + H_w^{t,k} (\hat{w}^t + \tilde{w}^t) - H_d^{t,k} (\hat{d}^t + \tilde{d}^t) \leq \overline{f}, \tag{3.14b}$$

$$k \in [0, n_k], t \in [1, n_t] \Big) \geq 1 - \epsilon$$

The formulation of chance-constrained Security-constrained Unit Commitment (c-SCUC) is presented below. Instead of using expected load $\hat{d}^t$ as in (3.13), we consider loads d^t as forecast $\hat{d}^t$

plus a random forecast error $\tilde{d}^t$ (i.e. $d^t = \hat{d}^t + \tilde{d}^t$).

$$\begin{aligned}
\min\ & (3.13a) \\
\text{s.t. } & (3.13b)(3.13c)(3.13d)(3.13e)(3.13f)(3.13g)(3.13h)(3.13i)(3.13j)(3.13k) \\
& (3.14a)(3.14b)
\end{aligned}$$

Comparing with d-SCUC, the only difference of c-SCUC is the addition of the chance constraint (3.14a). The chance constraint guarantees there will be enough supply to meet the net demand with probability no less than $1 - \epsilon$.

To reveal the structure of c-SCUC, we define the sets below:

$$\mathcal{B} := \big\{(z, u, v) : \text{(3.13h), (3.13i), (3.13j), (3.13k)}\big\} \tag{3.15a}$$

$$\mathcal{C} := \big\{(g, r) : \text{(3.13b), (3.13c), (3.13d), (3.13e)}\big\} \tag{3.15b}$$

$$\mathcal{H} := \big\{(z, g, r) : \text{(3.13f), (3.13g)}\big\} \tag{3.15c}$$

$$\mathcal{U} := \big\{(g) : \text{(3.14a)(3.14b)}\big\} \tag{3.15d}$$

Then c-SCUC can be succinctly represented as:

$$\begin{aligned}
\min_{z,u,v,g,r}\ & (3.13a) \\
\text{s.t. } & (z, u, v) \in \mathcal{B} \\
& (g, r) \in \mathcal{C},\ (z, g, r) \in \mathcal{H} \\
& \mathbb{P}\big(g \in \mathcal{U}\big) \geq 1 - \epsilon
\end{aligned}$$

Sets $\mathcal{B}$ and $\mathcal{C}$ stand for the *deterministic* constraints for binary and continuous variables, respectively. Set $\mathcal{H}$ represents the hybrid constraints related with both continuous and binary variables. Set $\mathcal{U}$ denotes all constraints related with uncertainties. Using the scenario approach, c-SCUC is converted to the scenario-based SCUC (s-SCUC) problem below:

$$\min_{(z,u,v)\in\mathcal{B}} \sum_{t=1}^{n_t}\Big(c_z^\intercal z^t + c_u^\intercal u^t + c_v^\intercal v^t\Big) +$$
$$\min_{(z,g,r)\in\mathcal{H}} \sum_{t=1}^{n_t}\Big(c_r^\intercal r^t + \sum_{k=0}^{n_k}\alpha_k c_g^\intercal g^{t,k}\Big)$$
$$\text{s.t. } (g,r)\in\mathcal{C}$$
$$g\in\cap_{i=1}^{N}\mathcal{U}_i$$

Remark 12 (Structural Properties of SCUC). SCUC is a two-stage optimization problem by nature, it has the following nice properties. Firstly, the non-convexity *only* exists in the first stage, i.e. $y\in\mathcal{Y}$. Given a first-stage solution y, the second stage is a simple linear program. Secondly, uncertainties come from renewables in the operation stage (only in the second stage). Based on the nice structural properties above, Section 3.3.3 shows that we are able to track down essential sets by solving two MILPs and $\sim|\mathcal{S}|$ linear programs.

3.4.3 Degeneracy of s-SCUC

This section presents an example to show that s-SCUC could be degenerate in many cases, which violates Assumption 9. Therefore almost all results of the classical scenario approach are not applicable. For s-SCUC, theoretical guarantees are only possible through the non-convex scenario approach in Section 3.2.3.

We use a 3-bus system to illustrate the degeneracy of s-SCUC. Configurations of the 3-bus system are in [2]. In order to visualize the feasible region of s-SCUC, we simplify the problem by (1) only considering one snapshot ($n_t = 1$) and ignoring initial status (thus no u, v variables); (2) removing reserve constraints (no r variables). By doing this, there are only four decision variables left: z_1, z_2, g_1, g_2. The on/off states z_1, z_2 can be inferred from values of g_1 and g_2, therefore the feasible region of the simplified s-SCUC can be visualized on the (g_1, g_2)-plane.

Using Definition 17, showing the degeneracy of s-SCUC includes three steps: (1) obtaining the optimal solution to SP($\mathcal{N}$); (2) finding all support scenarios $\mathcal{S}$ of SP($\mathcal{N}$); and (3) checking if

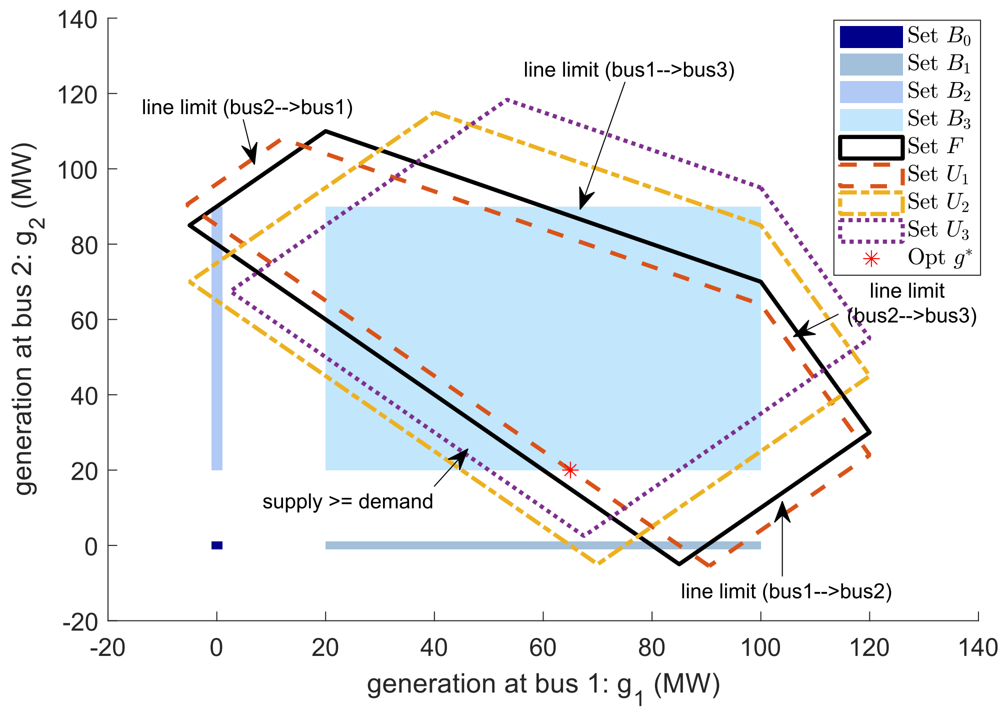


Figure 3.1: An illustrative example that s-SCUC is degenerate (3-bus system), illustration of the feasible region with constraints of all scenarios ($\mathcal{U}_1$,$\mathcal{U}_2$,$\mathcal{U}_3$), reprinted with permission from [2].

the optimal solution of SP($\mathcal{N}$) is the same as SP($\mathcal{S}$). Fig. 3.1 first visualizes constraints $\mathcal{B}_0 \sim \mathcal{B}_3$, which represents the region of 4 possible generator on/off status (e.g. $\mathcal{B}_1 : z_1 = 1, z_2 = 0$, $\mathcal{B}_3 : z_1 = 1, z_2 = 1$). The black solid lines denote constraints (3.13b), (3.13c) and (3.13f) using forecast values (d-SCUC). The red, yellow and purple dotted lines are three sets ($\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3$) of constraints corresponding to three scenarios. Given the setting that generator 1 is much cheaper than generator 2, we can easily eyeball the optimal solution with all constraints presented, marked by the red $*$. Next, we observe that removing scenario 1 ($\mathcal{U}_1$, red lines) changes the optimal solution, while removing scenario scenario 2 ($\mathcal{U}_2$, yellow lines) or scenario 3 ($\mathcal{U}_3$, purple lines) makes no difference. Thus scenario 1 is the only support scenario. Finally, we examine the scenario problem with only support scenarios presented. Fig. 3.2 shows that the optimal solution becomes the red $\diamond$ with only scenario 1, which is clearly different than the optimal solution in Fig. 3.1. Hence, this instance of s-SCUC is degenerate.

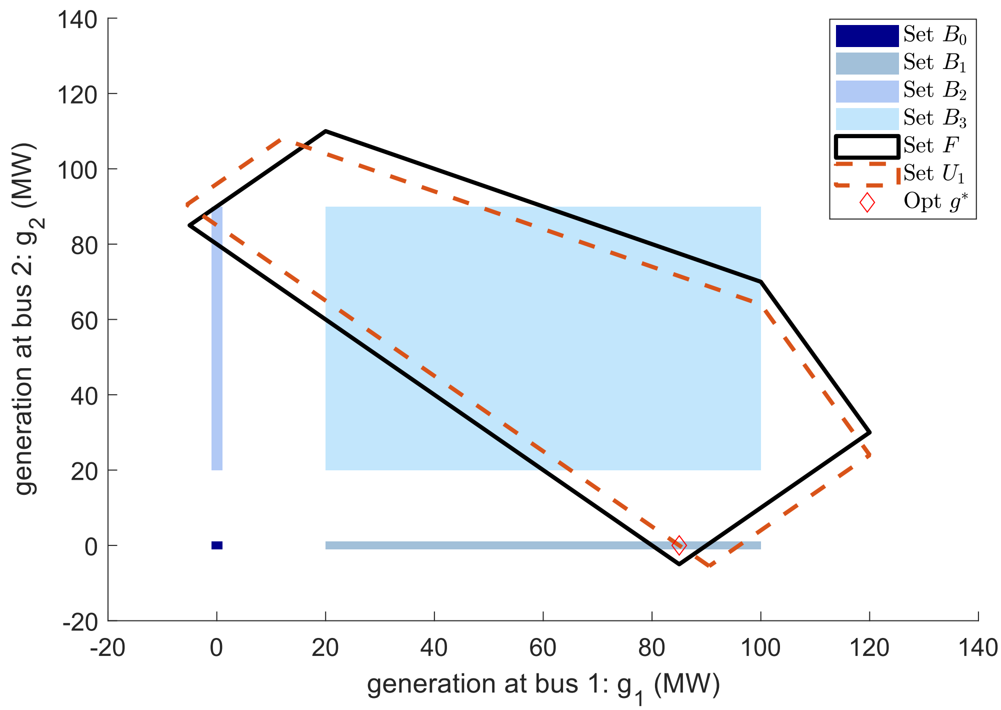


Figure 3.2: An illustrative example that s-SCUC is degenerate (3-bus system), illustration of the feasible region with only support scenarios ($\mathcal{U}_1$), reprinted with permission from [2].

3.5 Case Study

3.5.1 Settings of the 118-bus System

Numerical simulations were conducted on a modified 118-bus, 184-line, 54-generator, 24-hour system [149]. Most settings are identical as [149], except 5 wind farms are added to the system as in [150]. The s-SCUC problems were solved using 64 GB memory on the Hera server (hera.ece.tamu.edu), provided by Texas A&M University. The mathematical models for s-SCUC was formulated using YALMIP [84] on Matlab R2019a and solved using Gurobi v8.1 [87].

After obtaining a solution $\text{opx}_{\mathbb{A}}(\mathcal{N})$ to s-SCUC, Theorem 17 provides an upper bound $\epsilon(N, |\mathcal{I}|, \beta)$ on the *actual* violation probability $\mathbb{V}(\text{opx}_{\mathbb{A}}(\mathcal{N}))$. The theoretical guarantee $\epsilon(N, |\mathcal{I}|, \beta)$ is referred as posterior ϵ in the numerical results. The actual violation probability $\mathbb{V}(\text{opx}_{\mathbb{A}}(\mathcal{N}))$ is estimated by the out-of-sample violation probability $\hat{\epsilon}$, using an independent set of 10^6 scenarios.

To quantify the randomness of the scenario approach, for each sample complexity $N = 100, 200, \cdots, 1000$, we solve the corresponding s-SCUC problems on 10 independent sets of scenarios (i.e. 10 Monte-Carlo simulations). Results in both Fig. 3.3 and 3.4 show the average, maximum and minimum values in 10 Monte-Carlo simulations.

3.5.2 Cost vs Security: a trade-off

Fig. 3.3 shows the out-of-sample violation probability $\hat{\epsilon}$ and objective value (total cost). The shadowed area shows the max-min values in 10 Monte-Carlo simulations, and the solid line is the average value of 10 independent simulations. It is shown that the system risk level (violation probability) is reduced by 83% (from $\sim 30\%$ to $\sim 5\%$) by $\sim 1.1\%$ increase in total system costs. Similar observations were found in [1, 123, 148].

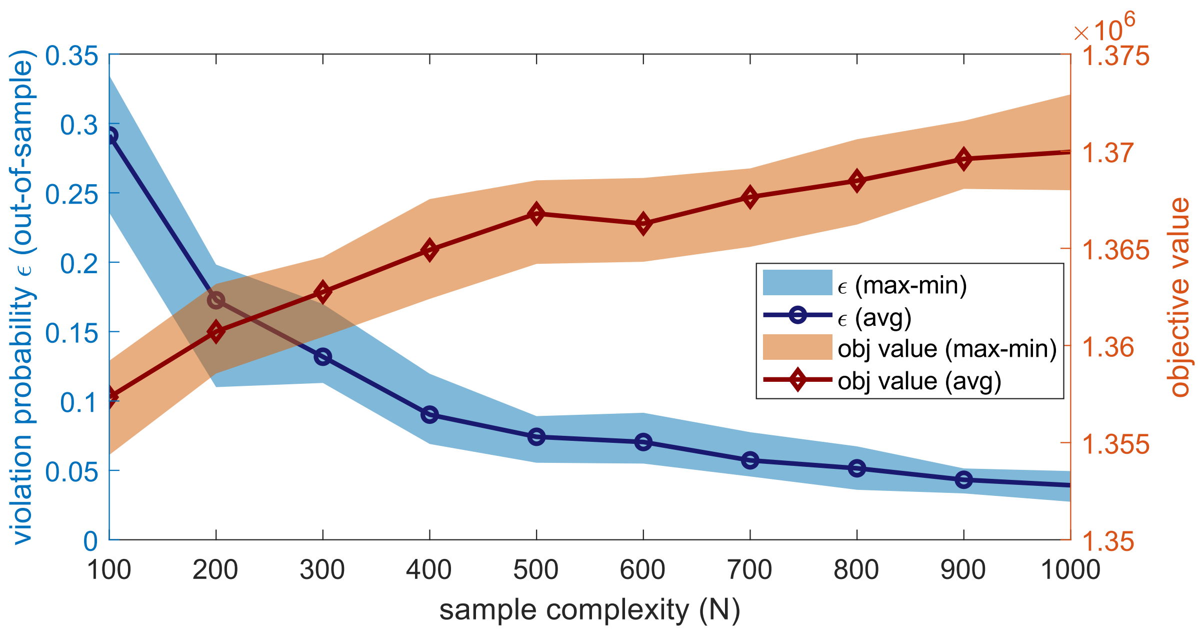


Figure 3.3: Cost vs Security: a Trade-off, reprinted with permission from [2].

3.5.3 Violation Probability

Fig. 3.4 presents the out-of-sample violation probability $\hat{\epsilon}$ and theoretical guarantees (posterior ϵ provided by Theorem 17). Since the cardinality of essential sets differ for each scenario problem (Fig. 3.5), the posterior guarantee ϵ is a band instead of a line. As illustrated in Fig. 3.4, the actual violation probability (approximated by $\hat{\epsilon}$) is bounded by the theoretical guarantees. This verifies the correctness of Theorem 17. The conservative ratio is $2 \sim 4$ (e.g. when out-of-sample $\hat{\epsilon}$ is $\sim 5\%$, Theorem 17 gives an upper bound $10\% \sim 20\%$).

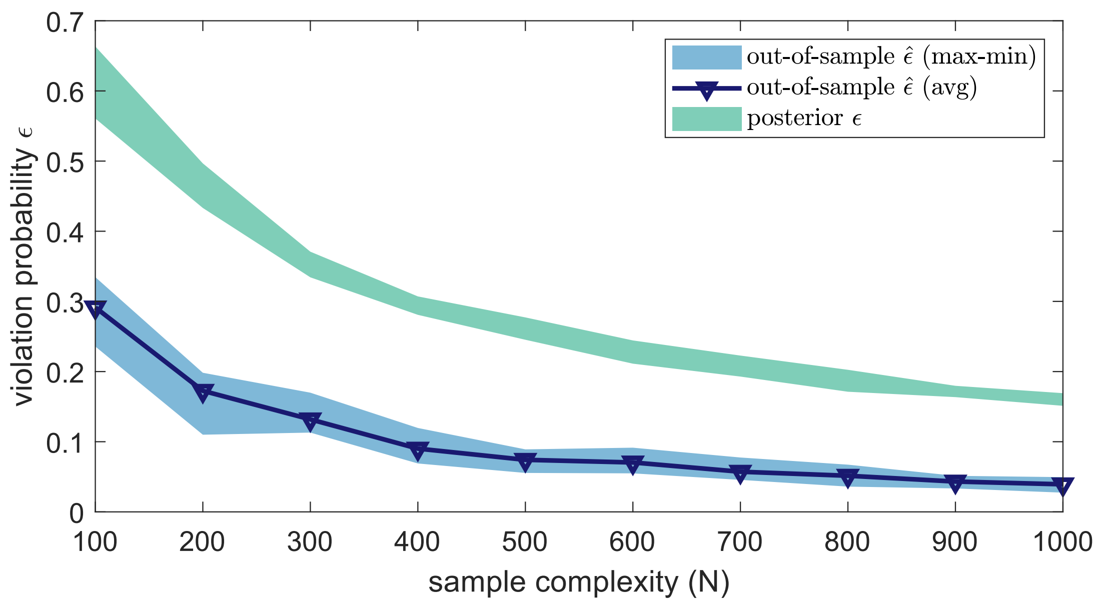


Figure 3.4: Out-of-sample Violation Probabilities and Theoretical Guarantees, reprinted with permission from [2].

3.5.4 Searching for Essential Sets for s-SCUC

s-SCUC was observed to be non-degenerate in 192 out of 200 simulations[6]. In other words, in 96% cases, we are able to find an essential set by solving $5 \sim 35$ linear programs and 2 mixed

[6]We conducted 10 simulation for 10 different sample complexities $(100, 200, \cdots, 1000)$ under two different settings: with/without $N-1$ contingencies, both include transmission constraints.

integer linear programs. It takes from 4934 seconds ($N = 100$) to 6847 seconds ($N = 1000$) to solve one MILP (s-SCUC). When searching for support scenarios for the second-stage problem (a linear program), it takes $281 \sim 388$ seconds to solve one LP. For those 8 out of 200 simulations, it takes an extra 20 hours to find an irreducible set using Algorithm 3. This computation time can be greatly reduced by tricks such as choosing appropriate starting points[7].

3.6 Discussions

3.6.1 Cardinality of Essential Sets

Fig. 3.5 compares the cardinalities of essential sets for three cases: (a) c-SCUC with $N-1$ contingencies but without transmission constraints, results of case (a) are obtained from [148]); (b) c-SCUC with transmission constraints but without $N-1$ contingencies; and (c) c-SCUC with both transmission constraints and $N-1$ contingencies.

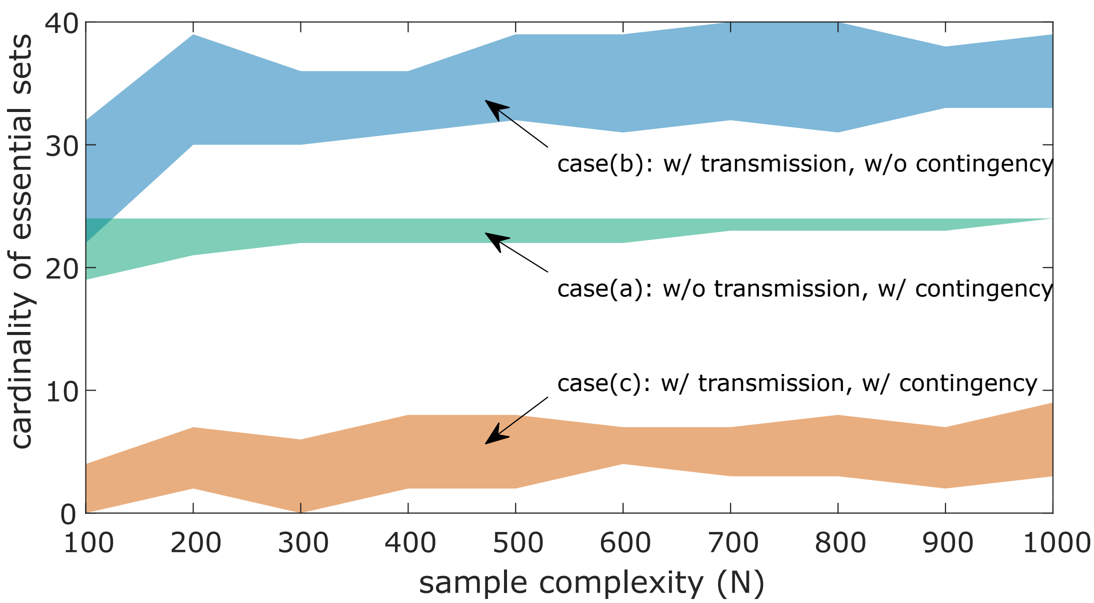


Figure 3.5: Cardinality of Essential Sets, reprinted with permission from [2].

[7]For example, when removing scenarios s and t consecutively in Algorithm 3, the solution $\text{opt}_{\mathbb{A}}(\mathcal{N} - s)$ is feasible to $\text{SP}(\mathcal{N} - s - t)$ thus can serve as a good starting point.

Case (a) is the simplest, in [148] we show that the scenario problem for unit commitment satisfies the non-degeneracy assumption 9, and the cardinality of essential sets is bounded by the number of snapshots n_t, i.e. $|\,\mathcal{S}\,| \leq n_t = 24$ in Fig. 3.5. Case (b) and (c) include transmission capacity constraints. As demonstrated in Section 3.4.3, s-SCUC could be degenerate with transmission constraints. Theoretically speaking, the cardinality of essential sets might be unbounded for non-convex problems. As observed in Fig. 3.5, the cardinality of essential sets ($30 \sim 40$ in case 2, $0 \sim 10$ in case 3) is greatly smaller than the number of decision variables (e.g. about 4000 binary variables and around 75000 continuous decision variables). This observation implies that the number of scenarios N required could be much smaller than expected.

Another interesting observation is that including $N-1$ contingency constraints reduces $|\,\mathcal{E}\,|$. This observation has two implications. First, $N-1$ contingency constraints not only protect the system from unexpected device failures, they also help reduce the impacts of uncertainties from renewables. Second, including $N-1$ contingency constraints could help reduce sample complexity. Similar with the observations in [123], this observation indicates that the scenario approach might be of practical use.

3.6.2 From Posterior to Prior Guarantees

Theorem 17 gives *posterior* guarantees on the quality of solutions, namely, we calculate $\epsilon(N,k,\beta)$ *after* obtaining the solution opx$(\mathcal{N})$. Lemma 2 proves that the $\epsilon(N,k,\beta)$ function in (3.9) is monotone in N and k. This implies that we can obtain prior guarantees. In other words, if the cardinality of essential sets is proved to be at most h ($|\,\mathcal{E}\,| \leq h$), then we can find the smallest $\hat{N}$ such that

$$\epsilon \geq 1 - \Big(\frac{\beta}{\hat{N}\binom{\hat{N}}{h}}\Big)^{\frac{1}{\hat{N}-h}} \tag{3.16}$$

holds for given ϵ and β. Then the solution $\text{opx}_{\mathbb{A}}(\mathcal{N})$ to the scenario problem using $\hat{N}$ scenarios has the guarantee $\mathbb{P}(\mathbb{V}(\text{opx}_{\mathbb{A}}(\mathcal{N})) \leq \epsilon) \geq 1-\beta$. This prior guarantee holds *before* solving the scenario problem with $\hat{N}$ scenarios. If a rigorous bound h on $|\,\mathcal{E}\,|$ can be proved, then there is no need to numerically search for essential sets. This is particularly attractive compared with posterior guarantees.

3.7 Summary

This chapter solves chance-constrained SCUC via the scenario approach and obtains rigorous theoretical guarantees on the solution. We demonstrate the structural properties of (possibly non-convex) general scenario problems. To obtain the tightest theoretical guarantees for chance-constrained SCUC, we design efficient algorithms to search for essential sets by exploiting the salient structures of SCUC. Numerical results on an IEEE benchmark system show that the essential scenario set is only a small subset of all scenarios. This implies that we can obtain relatively robust solutions (i.e. small ϵ) using only a moderate number of scenarios. Furthermore, we observe that some power engineering practices (e.g. $N - 1$ criteria) can help us reduce the number of scenarios needed while maintaining the same level of risk.

Future work includes reducing conservativeness by improving the complexity bound in Theorem 17 and investigating the performance of the (non-convex) scenario approach on larger-scale realistic systems.

4. LOOK-AHEAD OPTIMAL REACTIVE POWER DISPATCH WITH PROBABILISTIC GUARANTEES[1]

4.1 Look-ahead Optimal Reactive Power Dispatch

Renewable energy, which is becoming a major component of power resources, is highly variable with limited predictability. In many regions such as Texas, renewables (primarily wind) are located far from the load centers. The long distance between renewable generation and loads, compounded with the fact that such renewables are highly variable, lead to increasing concerns of maintaining voltage security for operations.

This chapter is motivated by the need to coordinate these reactive power support devices (RPSDs) in order to ensure voltage security with deep renewable penetration. The particular set of RPSDs discussed in this chapter include both continuous-state devices (e.g. SVCs) and discrete-state devices (e.g. capacitor banks). In practice, many of these devices are operated by multiple transmission owners. The decision of each control area is often not coordinated. This, in turn, leads to unnecessary frequent operations of RPSDs in large power systems. With advances in sensing, communication, and computing, the forecast of renewables in the near-term operation is improving. Such improvement could be leveraged by the system operator for more economic scheduling of RPSDs to ensure security. In this chapter, we examine the potential of look-ahead operation of the RPSDs across large geographical areas with high renewables.

There has been a substantial body of literature examining the impact of renewable penetration on system voltage profiles. Reference [151] presents four case studies of European countries and discuss system stability issues due to fluctuating renewables. Many papers formulate an optimal reactive power dispatch (ORPD) problem to solve the voltage problem induced by high penetration of wind power [152–154]. The ORPD problem often aims at finding optimal settings of current installed RPSDs to ensure system voltage constraints [155].

There are in general two families of voltage constraints: (1) voltage *stability* constraints; and

[1]Reprinted with permission from [3, 4].

(2) voltage *security* constraints. *Stability* of a power system refers to the continuance of intact operation following a disturbance. It depends on the operating condition and the nature of the physical disturbance [156]. Voltage stability constraints often require a particular form of voltage stability index larger or smaller than a threshold. Some typical voltage stability index include the distance to the nose point of the PV curve [157] and smallest singular value of the power flow Jacobian matrix [158]. *Security* of a power system refers to the degree of risk in its ability to survive imminent disturbances (contingencies) without interruption of customer service. It relates to robustness of the system to imminent disturbances and, hence, depends on the system operating condition as well as the contingent probability of disturbances [156]. Voltage security constraints typically require the voltage magnitudes within desired ranges under a set of plausible contingency scenarios [159].

As an effort to mitigate the impacts of renewables on voltage stability and security, many papers suggest the use of RPSDs [152]. However, [155] points out two major shortcomings of current ORPD literatures:

1. Because of the low computational burden, some ORPD problems only determine the schedules of continuous RPSDs or the discrete decision variables are relaxed to be continuous. This lack of coordination might cause troubles and it is necessary to formulate a comprehensive coordination framework including all the devices of the system;

2. Most of the proposed approaches focus on a single snapshot coordination of RPSDs. The past or future states of the system are not taken into account. Given the increasing intertemporal variability from renewables, it becomes more and more important to establish a look-ahead framework to schedule all the RPSDs. Detailed models of the operation costs of RPSDs are necessary in a look-ahead framework.

There has also been efforts developing multi-period coordination of voltage support devices. Reference [155] formulates a multi-objective mixed integer nonlinear programming (MINLP) problem and utilizes generalized Bender's decomposition to find out the optimal switching pattern of discrete voltage controllers and ensuring the voltage security of the system. Reference [160] proposes a three-stage coordination framework to minimize total lines loss and number of control actions. The problems solved in both [155] and [160] are MINLP problems.

In this section, we formulate the problem of look-ahead coordination of RPSDs in line with [155]. The proposed framework considers the operation cost of discrete control devices and transmission losses and guarantees the system voltage security with respect to $N-1$ contingencies. Instead of solving a MINLP problem as in [155] and [160], we solve the linearized problem, which provides insights into the operations of RPSDs and is much more computationally efficient and provides insights into the operations.

The rest of this sec is organized as follows: Section 4.1.1 formulates the look-ahead coordination of RPSDs as a MINLP problem. The main difficulties and linearized problem are presented in Section 4.1.2. Case studies and further discussions are provided in Section 4.1.3. Section 4.1.5 examines the impacts of wind uncertainties.

4.1.1 Look-ahead ORPD

Independent System Operators (ISOs) have the overall responsibility of monitoring and maintaining voltage security over its footprint, which often is comprised of multiple control areas. As an example, the Electric Reliability Council of Texas (ERCOT) manages its voltage issues with multiple entities.

ISOs typically perform voltage security screening studies focusing on the next few hours or days by looking at the estimates of voltage obtained from the supervisory control and data acquisition (SCADA) system. If any bus voltage goes beyond the predetermined limits, ISOs will assess the real-time voltage stability and determine necessary corrective operations.

In practice, the control of reactive power support devices such as capacitor banks are often decided at multiple transmission service providers (i.e. control areas). Given their limited information about the entire system, it often leads to unnecessarily volatile switches of these capacitor banks.

In Section 4.1.1.1, we propose a coordination framework for the operation of reactive devices.

The ISO coordinates all the reactive power devices and figure out the most efficient solution to guarantee the security of the system.

4.1.1.1 Problem Formulation

$$\min \quad \sum_{t=1}^{T} \Big(h_B(Q_B[t]) + \lambda[t] \sum_{t=0}^{T} \omega^c P_L^c[t] \Big) \tag{4.1a}$$

$$\text{s.t. } P^c[t] = A_G^c(P_G[t] + \eta^c P_\delta^c[t]) + A_W P_W[t] - A_D P_D[t] \tag{4.1b}$$

$$Q^c[t] = A_G^c Q_G^c[t] + A_C Q_C[t] + A_B Q_B[t] - A_D Q_D[t] \tag{4.1c}$$

$$P_\delta^c[t] = \mathbf{1}^\intercal (A_D P_D[t] - A_G^c P_G[t] - A_W P_W[t]) \tag{4.1d}$$

$$P_i^c[t] = \sum_{j=1}^{n_b} |V_i^c[t]||V_j^c[t]||Y_{ij}| \cos(\theta_i^c[t] - \theta_j^c[t] - \phi_{ij}) \tag{4.1e}$$

$$Q_i^c[t] = \sum_{j=1}^{n_b} |V_i^c[t]||V_j^c[t]||Y_{ij}| \sin(\theta_i^c[t] - \theta_j^c[t] - \phi_{ij}) \tag{4.1f}$$

$$P_L^c[t] = \sum_{\substack{l=1 \\ l:i\sim j}}^{n_l} g_l \big(|V_i[t]|^2 + |V_j[t]|^2 - 2|V_i[t]||V_j[t]| \cos(\theta_i - \theta_j) \big) \tag{4.1g}$$

$$|V^c|^- \le |V^c[t]| \le |V^c|^+ \tag{4.1h}$$

$$Q_C^- \le Q_C[t] \le Q_C^+ \tag{4.1i}$$

$$Q_B[t] \in \{0, Q_B^+\} \tag{4.1j}$$

$$(S_G^-)^2 \le (P_G[t] + \eta^c P_\delta^c[t])^2 + (Q_G^c[t])^2 \le (S_G^+)^2 \tag{4.1k}$$

$$i, j = 1, 2, \cdots, n_b, \quad c = 0, 1, 2, \cdots, n_c, t = 1, 2, \cdots, T$$

Problem (4.1) aims at finding the most smoothed and economic operation schedule of the reactive power support devices (RPSDs) in the upcoming T snapshots while ensuring voltage security in n_c contingency scenarios. Variables with $[t]$ are in snapshot t, and those with superscript ^{c} belong to contingency scenario c. Decision variables include the operating states of *discrete* RPSDs (e.g. shunt capacitors) $Q_B[t]$, operating states of *continuous* RPSDs (e.g. SVCs) $Q_C[t]$ and the voltage set-points of generators (i.e voltage magnitudes $|V^c[t]|$ of PV buses). The reactive generation $Q_G^c[t]$ is controlled by the automatic voltage regulators (AVRs) to maintain their bus voltages at desired levels.

Eqn. (4.1b)-(4.1c) are the real and reactive power balance equations at each bus. $P^c[t]$ and $Q^c[t]$ denote the real and reactive nodal injection. $A_B \in \mathbf{R}^{n_b \times n_B}$, $A_C \in \mathbf{R}^{n_b \times n_C}$, $A_D \in \mathbf{R}^{n_b \times n_D}$, $A_G^c \in \mathbf{R}^{n_b \times n_g}$ and $A_W \in \mathbf{R}^{n_b \times n_W}$ are adjacency matrices. If component k is connected with bus i, then $A_\cdot(i,k) = 1$; otherwise $A_\cdot(i,k) = 0$. Eqn. (4.1e)-(4.1f) are the power flow equations. $|V_i^c[t]|$ and $\theta_i^c[t]$ are the voltage magnitudes and angles of bus i ($i = 1, 2, \cdots, n_b$). $Y_{ij}\angle\phi_{ij}$ represents the component related with line (i,j) (bus i to bus j) in the admittance matrix Y.

There are n_c contingency scenarios[2] being considered in Problem (4.1). More specifically, we focus on the $N-1$ contingency of losing generators[3]. The contingency of losing one generator is modeled through the adjacency matrix of generators A_G^c. Let A_G^0 denote the adjacency matrix in the normal condition. The impacts of losing generator g (in scenario c) is equivalent with setting the gth column of A_G^0 to be zeros, and the new matrix is denoted by A_G^c ($c = 1, 2, \cdots, n_c$).

The real power imbalance $P_\delta^c[t]$ due to contingency c is proportionally allocated to each generator (i.e. $P_G[t] + \eta^c P_\delta^c[t]$). The participating factor η^c is a pre-defined vector and determined by the characteristics of generators. It is worth mentioning that $\mathbf{1}^\intercal \eta^c = 1$ and $\mathbf{1}^\intercal A_G^c \eta^c = 1$, this guarantees the post-contingency balance of real power:

$$\begin{aligned} & \mathbf{1}^\intercal \left(A_G^c (P_G[t] + \eta^c P_\delta^c[t]) + A_W P_W[t] \right) \\ = \; & \mathbf{1}^\intercal A_G^c P_G[t] + \mathbf{1}^\intercal A_W P_W[t] + \mathbf{1}^\intercal A_G^c \eta^c \mathbf{1}^\intercal (A_D P_D[t] - A_G^c P_G[t] - A_W P_W[t]) \\ = \; & \mathbf{1}^\intercal A_D P_D \end{aligned} \tag{4.2}$$

Eqn. (4.1h) depicts the voltage security constraints. Voltage magnitudes will be maintained within predetermined ranges $[|V^c|^-, |V^c|^+]$ for each contingency scenario c. In this section, we

[2] c is the index of contingency scenarios. For simplicity, we use $c = 0$ represents the normal operating condition.

[3] It could be extended towards the cases of possible line failures. In that case, we need to modify the Y matrix to model the scenarios of losing transmission lines. For simplicity, we only focus on losing generators.

use $[0.95, 1.05]$ for normal operation analysis ($c = 0$) and $[0.9, 1.1]$ for contingency analysis ($c = 1, 2, \cdots, n_c$) [161].

Other constraints include the capacity of devices (Eqn. (4.1i), (4.1j)[4]). Eqn. (4.1k) represents the generation capacity limits [162, 163]. Since $P_G[t]$, η^c and $P_\delta^c[t]$ are all parameters, Eqn. (4.1k) is equivalent with the following linear inequality:

$$Q_G^{c-} \leq Q_G^c[t] \leq Q_G^{c+} \tag{4.3}$$

4.1.1.2 On Objective Function

The objective function Eqn. (4.1a) is time-coupled and includes cost of line losses and the operation costs of the RPSDs. The cost of line losses is evaluated at the market energy price $\lambda[t]$. The operation cost of discrete RPSDs $h_B(Q_B[t])$ is proportional to the number of switchings[5]:

$$h_B(Q_B[t]) = \sum_{i=1}^{n_B} \sum_{t=1}^{T} \pi_i Q_{B_i}^+ \cdot |x_{B_i}[t] - x_{B_i}[t-1]| \tag{4.4}$$

$$= \sum_{i=1}^{n_B} \sum_{t=1}^{T} \pi_i Q_{B_i}^+ \cdot (x_{B_i}[t] - x_{B_i}[t-1])^2 \tag{4.5}$$

π_i is the unit operation cost of device i for switching one time. According to [164], π_i is evaluated by the total unit installation cost divided by the total number of switchings in its lifetime. Typical values of π_i is 0.41\$/(MVar·times). For a 50MVar capacitor bank, its operation cost is $\pi Q_{B_i}^+ = 0.41 \times 50 = 20.5$ \$/times [164].

More discussions on choosing proper objective functions are provided in Section 4.1.4.1.

[4]For simplicity, we assume the discrete-state devices only have two/binary states: on or off, and use subscript $_B$.

[5]Eqn. (4.4) to (4.5) is due to the fact that $x_B[t]$ is binary. We prefer Eqn. (4.5) because of its quadratic form and positive definiteness.

4.1.2 Linearized Look-ahead ORPD

4.1.2.1 Computational Complexity

Formulation (4.1) is a Mixed Integer Non-Linear Programming (MINLP) problem. It is challenging to solve and there is no guarantee on the global optimal solution. The intractability issue is mainly due to the non-linearity of power flow equations (Eqn. (4.1e)-(4.1f)). A common approach to tackle this issue is to linearize power flow equations [165–167]. The non-linear relationship between P, Q and $|V|, \theta$ is approximated by a linear sensitivity matrix A:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = A \begin{bmatrix} \Delta\theta \\ \Delta|V| \end{bmatrix} \tag{4.6}$$

Many different approaches based on different forms of matrix A have been proposed, e.g. [164,165, 168–170]. Among all the proposed approaches, power flow Jacobian matrix is the most popular choice and has relative good linearization accuracy [165].

4.1.2.2 Linearization

We follow the idea in [168–170], the original problem (Problem (4.1)) is linearized to be Problem (4.10) using power flow Jacobian matrix.

Let $t = 0$ denote current snapshot, and $\cdot[0]$ (e.g. $P_G[0]$, $Q_G[0]$, $Q_B[0]$, $Q_C[0]$) denote the current operating states of the devices. Based on the power flow solutions of current snapshot $t = 0$, we can calculate the power flow Jacobian matrix, and the power flow equations Eqn. (4.1e)-(4.1f) can be approximated as:

$$\begin{bmatrix} P[t] - P[0] \\ Q[t] - Q[0] \end{bmatrix} \approx \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial |V|} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial |V|} \end{bmatrix}_{t=0} \begin{bmatrix} \theta[t] - \theta[0] \\ |V[t]| - |V[0]| \end{bmatrix} \tag{4.7}$$

Similarly, the line losses $P_L^c[t]$ can be approximated by:

$$P_L[t] \;\; \approx \;\; P_L[0] + \Delta P_L[t] \tag{4.8}$$

$$\Delta P_L[t] = \begin{bmatrix} \frac{\partial P_L}{\partial \theta} & \frac{\partial P_L}{\partial |V|} \end{bmatrix}_{t=0} \begin{bmatrix} \theta[t] - \theta[0] \\ |V[t]| - |V[0]| \end{bmatrix} \tag{4.9}$$

Problem (4.10) is obtained by replacing Eqn. (4.1e)-(4.1f) with Eqn. (4.7).

$$\min \quad \sum_{t=1}^{T} \Big(h_B(Q_B[t]) + \lambda[t] \sum_{c=0}^{n_c} w^c (P_L[0] + \Delta P_L^c[t]) \Big) \tag{4.10a}$$

$$\text{s.t. Eqn. } (4.1b), (4.1c), (4.1d) \tag{4.10b}$$

$$\begin{bmatrix} \Delta P^c[t] \\ \Delta Q^c[t] \end{bmatrix} = \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial |V|} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial |V|} \end{bmatrix} \begin{bmatrix} \Delta \theta^c[t] \\ \Delta |V^c[t]| \end{bmatrix} \tag{4.10c}$$

$$\Delta P_L^c[t] = \begin{bmatrix} \frac{\partial P_L}{\partial \theta} & \frac{\partial P_L}{\partial |V|} \end{bmatrix} \begin{bmatrix} \Delta \theta^c[t] \\ \Delta |V^c[t]| \end{bmatrix} \tag{4.10d}$$

$$\Delta P^c[t] = P^c[t] - P[0] \tag{4.10e}$$

$$\Delta Q^c[t] = Q^c[t] - Q[0] \tag{4.10f}$$

$$\Delta |V^c[t]| = |V^c[t]| - |V[0]| \tag{4.10g}$$

$$\Delta P_L^c[t] = P_L[t] - P_L[0] \tag{4.10h}$$

$$P[0] = A_G P_G[0] + A_W P_W[0] - A_D P_D[0] \tag{4.10i}$$

$$Q[0] = A_G Q_G[0] + A_C Q_C[0] + A_B Q_B[0] - A_D Q_D[0] \tag{4.10j}$$

$$\text{Eqn. } (4.1h), (4.1i), (4.1j), (4.3). \tag{4.10k}$$

$$\Delta \theta^- \leq \Delta \theta^c[t] \leq \Delta \theta^+ \tag{4.10l}$$

$$\Delta |V|^- \leq \Delta |V^c[t]| \leq \Delta |V|^+ \tag{4.10m}$$

$$c = 1, 2, \cdots, n_c, t = 1, 2, \cdots, T$$

Since this linear approximation is only accurate within the neighborhood of current operating point, Eqn. (4.10l) and (4.10m) are necessary. After solving Problem (4.10), its solution will be verified by solving the power flow equations of each contingency scenario and each snapshot.

4.1.3 Case Studies

In this section we provide case studies on the IEEE RTS 24-bus system [171].

4.1.3.1 Settings

We make the following changes to the IEEE RTS 24-bus system:

- one wind farm with capacity 250MW is added to bus 4;
- there is a synchronous condenser at bus 14 in the original system; three capacitor banks are added to bus 1, 10 and 15 with capacities 50MVar, 50MVar and 100MVar. As discussed in Section 4.1.1.2, the operation costs of the three capacitor banks are: 20.5\$/times, 20.5\$/times and 41\$/times.
- the load profiles and generation of wind farms are shown in Fig. 4.1. Different colors represent different bus numbers.
- each one of the two 400MW generators at bus 18 and 21 is replaced with four 100MW generators, other settings of the generators remain the same as the original 24-bus system.
- the market energy price $\lambda[t], t = 1, 2, \cdots, T$ is assumed to be 60\$/MWh and the operation interval is 15 minutes.
- all contingency scenarios have equal weights ω^c.

Given the wind and load profiles in Fig. (4.1), T security-constrained economic dispatch problems are solved to get the real generation schedule $P_G[t]$, which is shown in Fig. (4.2).

Matpower 6.0 [172] is used to calculate the power flow Jacobian matrix, line loss sensitivity matrix and power flow solutions. The test system is simulated using Matlab R2016b on a PC with Intel i7-2600 8-core CPU@3.40GHz and 16GB RAM memory.

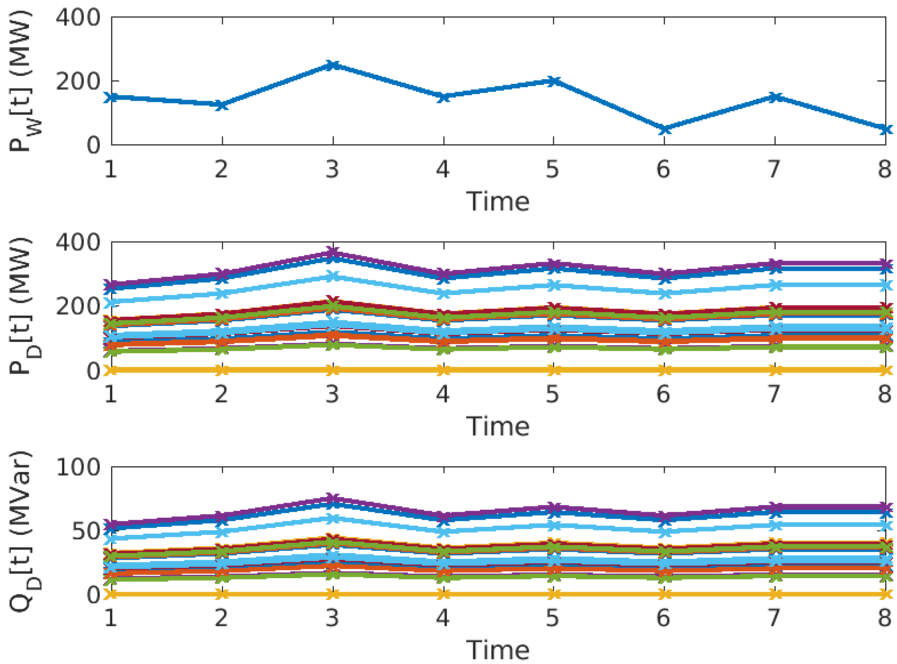


Figure 4.1: Wind And Load Profiles, reprinted with permission from [3].

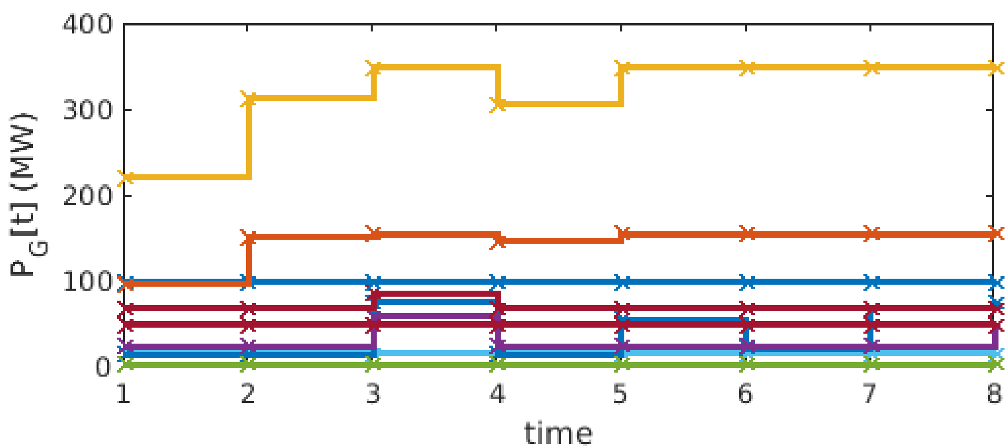


Figure 4.2: Generation Schedule, reprinted with permission from [3].

4.1.3.2 Simulation Results

Problem (4.10) is solved using Gurobi 7.0 [87]. The barrier method found the optimal solution (with 0.0% gap) in 3.79 seconds. The optimal total cost is \$6411.3, and the optimal schedules of RPSDs and voltage set-points of generators are presented in Fig. (4.3).

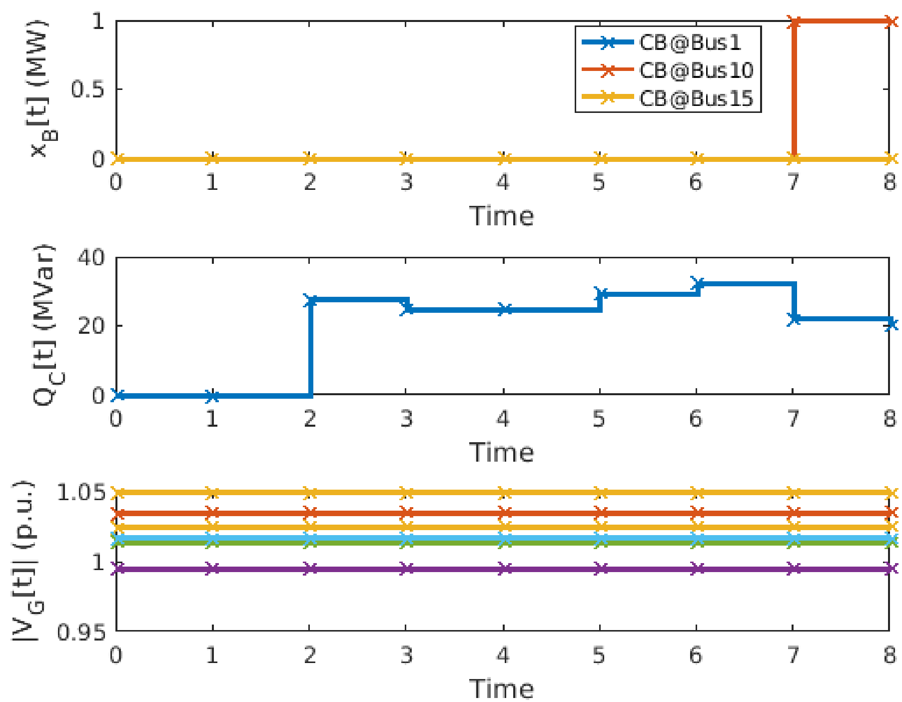


Figure 4.3: Optimal Schedules of RPSDs and Voltage Set-points of Generators, reprinted with permission from [3].

4.1.3.3 Validity of the Solution

Instead of solving the original MINLP problem (4.1) (as in [155]), we seek solutions to the approximated problem (4.10). To examine the validity of the solution to problem (4.10), we input the optimal schedules of RPSDs (Fig. (4.3)) to $T \times n_c$ power flow equations and solve them. This section compares the solution to problem (4.10) and the AC power flow solutions.

4.1.3.4 Feasibility

Since Problem (4.10) is not an accurate representation of the power system model, there is *no guarantee* that the solution to Problem (4.10) is always *feasible* to Problem (4.1). For this case study, we examine the voltage magnitudes in all the n_c scenarios and T snapshots. All the voltage magnitudes are within desired ranges ($[0.95, 1.05]$ and $[0.9, 1.1]$).

For large wind variations or critical contingencies, it is *possible* that the solution to Problem (4.10) is not feasible to Problem (4.1). One solution to this issue is to apply more conservative bounds on voltages and device capacities when solving Problem (4.10).

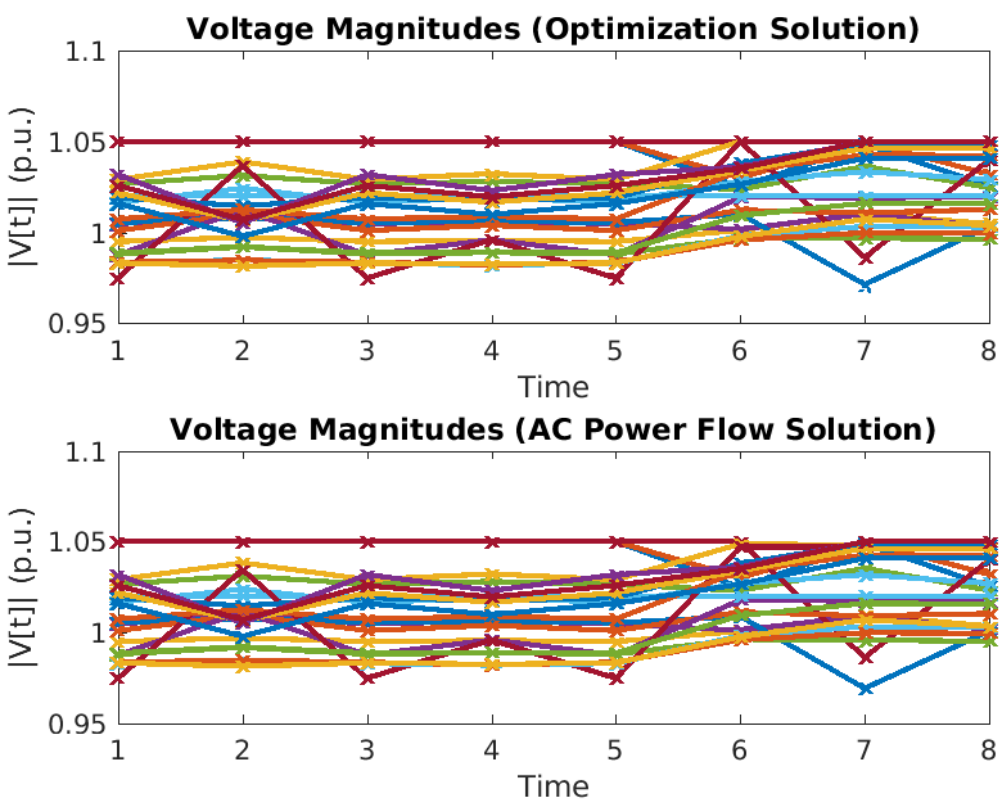


Figure 4.4: Comparison of Voltage Profiles (Normal Conditions), reprinted with permission from [3].

4.1.3.5 Accuracy

Since voltage magnitudes are critical for system security assessment, we measure the largest distance between the voltage profiles from the optimal solution and those of the power flow solutions. Fig. (4.4) provides a qualitative comparison of the voltage profiles. Quantitative analysis is presented in Table. 4.1. Let $|V^c[t]| \in \mathbf{R}^{n_b}$ denote the voltage magnitudes of scenario c at time t from AC power flow solution, and $|\hat{V}^c[t]|$ denotes the voltage magnitudes from solving Problem (4.10). The distance between $|V^c[t]|$ and $|\hat{V}^c[t]|$ is defined as:

$$d(|V^c|, |\hat{V}^c|) = \max_{c,t} \| |V^c[t]| - |\hat{V}^c[t]| \|_1 \tag{4.11}$$

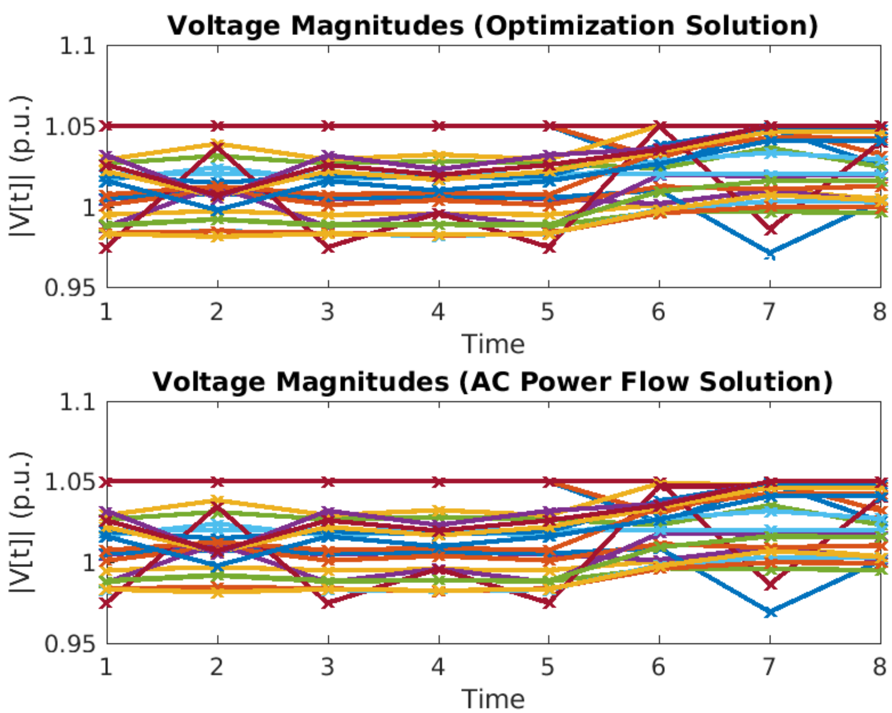


Figure 4.5: Comparison of Voltage Profiles (Contingency), reprinted with permission from [3].

Table 4.1 presents the largest errors $d(|V^c|, |\hat{V}^c|)$ in normal condition ($c = 0$) and contingency scenarios ($c = 1, 2, \cdots, 38$). All the errors are less than 0.5% in normal condition and less than 1% in all contingency scenarios. It is worth mentioning that the 38th contingency scenario often leads to the largest approximation error. The reason is that the 38th contingency scenario represents losing the largest generator in the system, which causes a substantial deviation from normal operating point.

Table 4.1: Comparison of Voltage Magnitudes ($d(|V^c|, |\hat{V}^c|)$), reprinted with permission from [3]

Time	Normal Condition	Contingency Scenario	Worst Scenario
$t = 1$	0.0010	0.0017	$c = 38$
$t = 2$	0.0015	0.0085	$c = 38$
$t = 3$	0.0040	0.0072	$c = 38$
$t = 4$	0.0011	0.0084	$c = 38$
$t = 5$	0.0023	0.0075	$c = 10$
$t = 6$	0.0041	0.0079	$c = 5$
$t = 7$	0.0024	0.0077	$c = 9$
$t = 8$	0.0042	0.0083	$c = 38$

From simulation results (Table. 4.1 and Fig. 4.4), the approximation seems satisfying. We want to emphasize that there is *no guarantee* on the approximation accuracy until rigorous theoretical analysis is conducted.

4.1.4 Discussions

4.1.4.1 On Objective Functions

Different from most of the literatures, where minimizing losses is the only objective, we include the operation costs of the control devices. There have been some studies on modeling the cost related with reactive power. For example, [164] provides a method to calculate the operation cost of reactive generations and continuous RPSDs (e.g. compensators). [173] provides more comprehensive modeling of the cost of discrete control devices.

4.1.4.2 *On the Impacts of Operation Costs of RPSDs*

By setting $\pi_i = 0$ for $i = 1, 2, \cdots, n_B$, the objective of Problem (4.1) becomes the same as most of the literatures: only minimizing the total line losses. The optimal solution to Problem (4.10) with $\pi = 0$ is shown in Fig. (4.6). Comparing Fig. (4.3) with Fig. (4.6), we found that some unnecessary operations of RPSDs may happen when their operation costs are not being considered.

The cost of line losses for the IEEE 24-bus system is about 60\$/MWh$\times 55$MW $\times 0.25$h $= 825$\$. And the cost of switching a 100MVar capacitor bank once is around 40 dollars, which is about 5% of the line loss costs. With lower energy price, the operation cost will possess a higher portion of the overall cost. Only minimizing line losses might lead to frequent switchings of RPSDs, which could increase the overall costs and lead to suboptimal solutions.

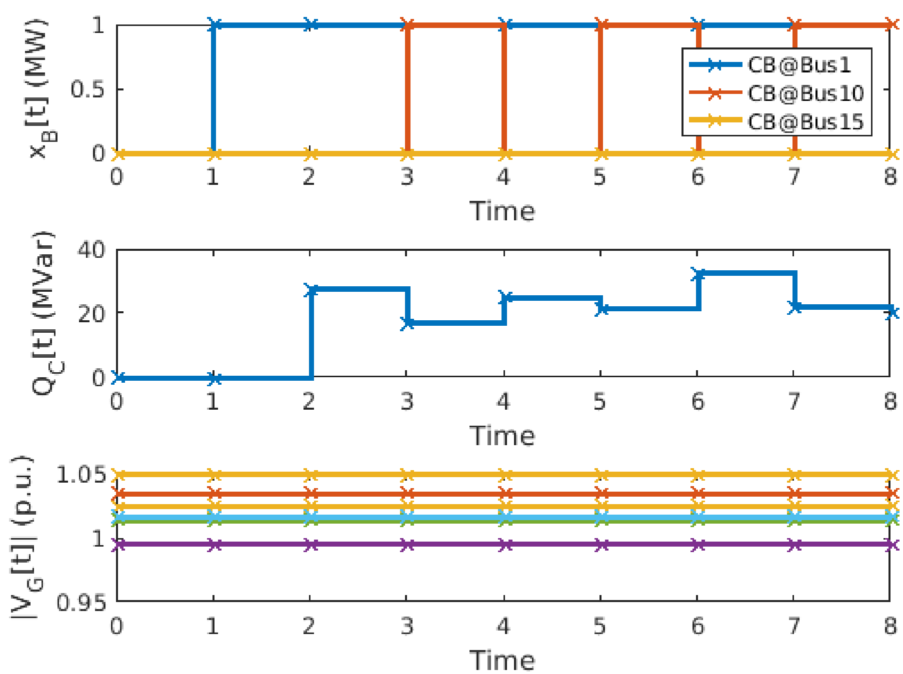


Figure 4.6: Optimal Solution (No Operation Costs), reprinted with permission from [3].

4.1.5 Estimate the Impacts of Wind Uncertainties

We focus on the influences of temporal variations of wind generations in the look-ahead coordination of RPSDs. The uncertainty of wind is another critical issue to be discussed. Because of the non-linearity of power flow equations, the relationship between voltage magnitudes and wind fluctuations is also non-linear. In this section, we approximate this non-linear relationship using the modified Jacobian matrix, and simulation results show that this approximation is quite accurate.

4.1.6 A Linear Approximation

To estimate voltage magnitude changes, we need to reformulate the power flow Jacobian matrix Eqn. (4.25). Let y_1 denote the *control* variables related with real/reactive power (i.e. ΔP of PV buses, ΔP and ΔQ of PQ buses), and y_2 denote the *state* variables (i.e. ΔQ of PV buses, ΔP and ΔQ of slack buses):

$$y_1 := \begin{bmatrix} \Delta P_{PV} \\ \Delta P_{PQ} \\ \Delta Q_{PQ} \end{bmatrix}, y_2 := \begin{bmatrix} \Delta P_{REF} \\ \Delta Q_{REF} \\ \Delta Q_{PV} \end{bmatrix} \tag{4.12}$$

Let x_1 represent the *control* variables related with voltages (i.e. $\Delta|V|$ of PV buses, $\Delta|V|$ and $\Delta\theta$ of slack buses), and x_2 denote the *state* variables (i.e. $\Delta\theta$ of PV buses, $\Delta|V|$ and $\Delta\theta$ of PQ buses):

$$x_1 := \begin{bmatrix} \Delta|V|_{PV} \\ \Delta|V|_{REF} \\ \Delta\theta_{REF} \end{bmatrix}, x_2 := \begin{bmatrix} \Delta|V|_{PQ} \\ \Delta\theta_{PV} \\ \Delta\theta_{PQ} \end{bmatrix}. \tag{4.13}$$

x_1 and y_1 represent the variables we can *directly* control in the power flow equations, the state variables x_2 and y_2 are implicitly determined by x_1 and y_1. If the power flow equation could be approximated as Eqn. (4.14), where all the *control* variables are on the right-hand side and the *state* variables are on the left-hand side, then we can easily estimate the changes of *state* variables from the changes of *control* variables.

$$\begin{bmatrix} y_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ x_1 \end{bmatrix} \tag{4.14}$$

Unfortunately, the original Jacobian matrix in Eqn. (4.7) has the form[6]:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.15}$$

which mixes the control and state variables on both sides of the equation.

By substituting y_2 and x_2 in Eqn. (4.15) using Eqn. (4.14), we get:

$$(J_{12}C - I)y_1 + (J_{11} + J_{12}D)x_1 = 0 \tag{4.16}$$

$$(A - J_{22}C)y_1 + (B - J_{21} - J_{22}D)x_1 = 0 \tag{4.17}$$

Therefore:

$$A = J_{22}J_{12}^{-1}, \quad B = J_{21} + J_{22}J_{12}^{-1}J_{11}, \quad C = J_{12}^{-1}, \quad D = J_{12}^{-1}J_{11} \tag{4.18}$$

4.1.6.1 Simulation Results

Two approaches are compared: (1) calculate voltage magnitude changes by solving power flow equations; and (2) estimate the changes using the method in Section 4.1.6.

We modify the settings of wind generations in Section 4.1.3 and focus on only one snapshot ($t = 4$). The wind uncertainties here is mainly due to its unpredictability. Reference [174] analyzes the wind data from ERCOT and concludes that the Cauchy distribution is a better choice than the Gaussian distribution, Beta distribution or Weibull distribution to fit the persistence forecast errors of wind generation. In this study, the wind forecast error (%) is modeled as a Cauchy distribution

[6]Please notice that J_{12} and J_{21} are square matrices, and J_{11} and J_{22} are rectangular matrices.

with location parameter $x_0 = 0$ and scale parameter $\gamma = 5\%$. With the optimal solutions (states of RPSDs and voltage set-points of generators) solved in Section 4.1.3, 1000 wind scenarios are generated and 1000 corresponding power flow problems are solved. Fig. 4.7 presents the scatter plot and estimated probability density functions of wind generation and voltage magnitudes of bus 4 (the location of wind farm). The blue dots in Fig. 4.7 are obtained by solving power flow equations, and the red dots are our linear approximations using the modified Jacobian matrix in Eqn. (4.14). This linear approximation is quite accurate, it has error less than 0.1% with moderate wind fluctuations ($\pm 30\%$).

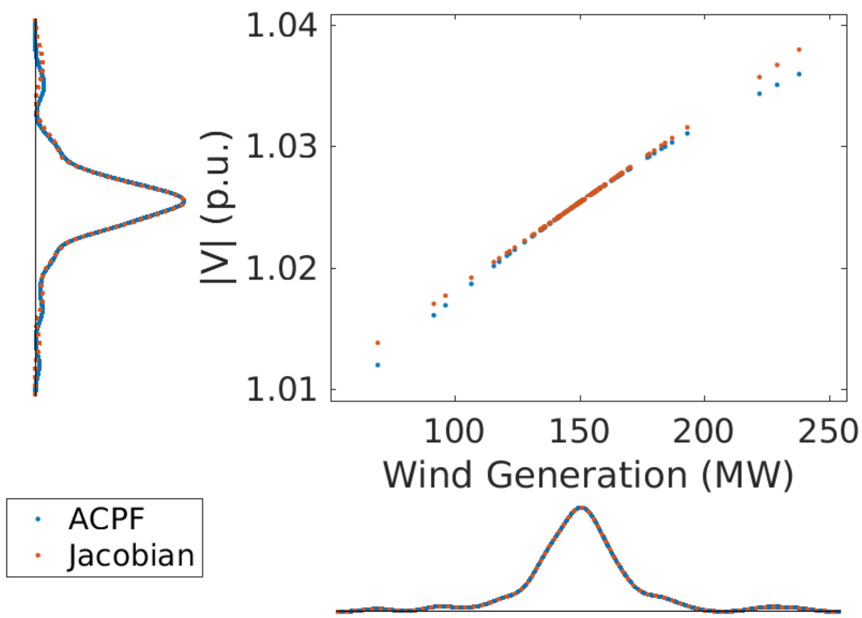


Figure 4.7: Impacts of Wind Uncertainties, reprinted with permission from [3].

As shown in Fig. 4.7, wind uncertainties result in uncertainties of voltage magnitudes. With deeper penetration of renewables or a heavy-loaded system, voltage security issue due to wind fluctuations could be quite severe. Since Problem (4.1) and (4.10) are deterministic optimization problems, they are not able to handle the voltage security issue with wind uncertainties. We need to formulate the ORPD problem using stochastic optimization frameworks such as robust optimization and stochastic programming.

4.2 Chance-constrained Optimal Reactive Power Dispatch

The high variability and limited predictability of renewables impose new challenges on the secure and reliable operation of power systems. There has been a substantial amount of literatures showing that deep penetration of renewables could jeopardize the security and reliability of power systems [152–154]. For example, the rapid increase and stochastic nature of renewables might lead to voltage issues, which could be severe when a stressed system is lack of reactive support. An Optimal Reactive Power Dispatch (ORPD) problem is often formulated for better voltage profiles [152–154]. The ORPD problem aims at finding optimal settings of current installed Reactive Power Support Devices (RPSDs) such as SVCs and Capacitor Banks to ensure system voltage constraints [155]. Although numerous papers have studied the ORPD problem, most of them adopt a deterministic formulation and uncertainties from wind are ignored.

In this chapter, we propose a framework for optimal reactive power dispatch considering joint uncertainties from wind and contingencies. The proposed framework is built upon chance-constrained programming, which is a natural and efficient tool for decision making in an uncertain environment.

4.2.1 Chance Constrained Programming

Problem (4.19) is the typical form of a single-stage chance-constrained program (CCP):

$$\min_{x} \quad c^{\intercal}x \tag{4.19a}$$

$$\text{s.t. } Ax \geq b \tag{4.19b}$$

$$\mathbb{P}_{\omega}\Big(G(\omega)x \leq h(\omega)\Big) \geq 1-\epsilon \tag{4.19c}$$

$$x \in \mathbf{R}^n$$

Problem (4.19) aims at finding a cost-minimizing strategy while satisfying a set of deterministic and probabilistic constraints. Without loss of generality [6], we assume the objective takes linear form $c^\intercal x$. Decision variables are denoted by x, and Eqn. (4.19b) is the *deterministic* constraint on x. Uncertainties appear as variable $\omega \in \mathbf{R}^m$, and the chance constraint Eqn. (4.19c) requires the inner constraint $G(\omega)x \leq h(\omega)$ to be satisfied with probability at least $1 - \epsilon$.

CCPs are often challenging to solve for the following two reasons: (1) the feasible region of a CCP is usually non-convex [11]; and (2) it is NP-hard to accurately calculate the probability in the chance-constraint [35]. There are four typical methods to get approximately optimal solutions to CCPs: (1) deriving a *deterministic equivalent* optimization problem [20, 175]; (2) *convex approximation* [11]; (3) *scenario approach* [6]; and (4) *sample average approximation* [72, 74, 176]. Because the cc-ORPD problem is a MINLP problem, sample average approximation, which is a favorable choice to handle integer variables in CCPs, is selected to solve cc-ORPD in this chapter. More details on sample average approximation is provided in Section 4.2.3.3.

4.2.2 Chance-constrained Programs in Power Systems

There are many applications of CCPs on power system problems: chance-constrained DCOPF (cc-DCOPF) [102, 122, 123, 177, 178], chance-constrained Unit Commitment (cc-UC) [135, 179], using chance-constrained programming to handle contingencies in power systems [118, 132]. In this chapter, we formulate a chance-constrained Optimal Reactive Power Dispatch (cc-ORPD) problem to address the voltage security issue induced by the deep penetration of renewables and potential contingencies. The cc-ORPD problem is unique in the following three aspects: (1) It is built upon a more accurate model of power system (i.e. AC power flow) rather than the simplified DC power flow model, which appears in most of literatures [102, 132, 135, 177, 179]. (2) The cc-ORPD problem considers the optimal operation of both continuous and discrete state voltage support devices. While in [180], only continuous-state devices (e.g. SVCs) are being considered. (3) The cc-ORPD problem ensures voltage security with respect to the joint distribution of contingencies and wind uncertainties. Whereas most literatures handling contingencies via CCPs [118, 132, 179] are based on DC power flow model. As a result, they are fundamentally incapable of addressing voltage-related issues.

The remainder of this section is organized as follows: Section 4.2.3 discusses the impacts of wind uncertainties on voltage security. Section 4.2.3.3 introduces the sample average approximation approach to solve CCPs. Motivated by the discussion in Section 4.2.3, we formulate a cc-ORPD problem in Section 4.2. Section 4.2 also elaborates how to derive a computationally tractable form of the cc-ORPD problem via sample average approximation. Case studies are presented in Section 4.2.5.

4.2.3 Impacts of Wind Uncertainties on Voltage Security

4.2.3.1 Wind Farm Modeling

The wind farm is often modeled as a negative real load or pure real power generator in most literatures. While at most Independent System Operators (ISOs) in the US, wind farms are required to provide some reactive support to reduce voltage issues. In this section, the wind farm is modeled as a negative load with constant power factor 0.95. Let $P_W \in \mathbf{R}^{|\mathcal{W}|}$ and $Q_W \in \mathbf{R}^{|\mathcal{W}|}$ denote the forecast value of a set of wind farms $\mathcal{W}$. And $\xi \in \mathbf{R}^{|\mathcal{W}|}$ represents the forecast errors of wind farms, $\xi \in \Xi$ is a random variable with underlying distribution Ξ. The actual output of wind farm w is $(P_{W,w} + jQ_{W,w})(1+\xi_w), \forall w \in \mathcal{W}$ and also random. In this chapter, we assume the underlying distribution Ξ is unknown but fixed. We also assume that the power factor is maintained at 0.95 for any wind fluctuations.

4.2.3.2 A Linear Approximation

Reference [3] shows that the voltage magnitudes of PQ buses become uncertain with wind fluctuations ξ. Fig. 4.8 presents the voltage magnitudes with respect to wind uncertainties in a modified IEEE 24-bus system [3]. The blue curve in Fig. 4.8 is obtained by solving a series of power flow equations, which is computationally expensive. Reference [3] proposes an approximation method using power flow Jacobian matrix to estimate the voltage magnitude changes to wind fluctuations. The red curve in Fig. 4.8 is calculated using the approximation method in [3]. Although the relationship between voltage magnitudes and wind fluctuation is fundamentally non-linear, Fig. 4.8

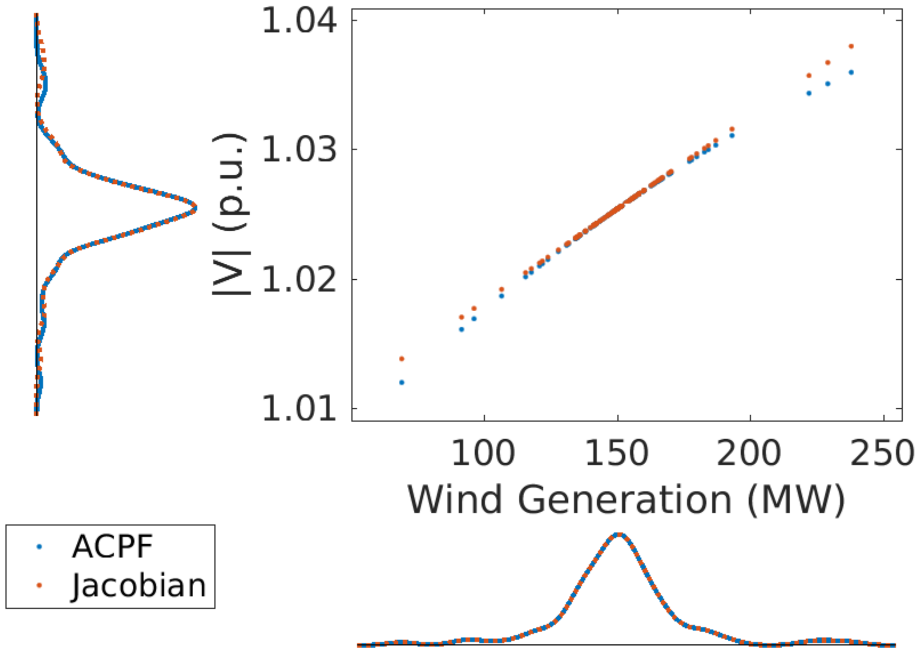


Figure 4.8: Impacts of Wind Uncertainties on Voltage Magnitudes, reprinted with permission from [4].

shows that we can get satisfying approximation using linearized power flow equations.

4.2.3.3 Sample average approximation

Given a two-stage chance-constrained program:

$$\min_{x,y(\omega)} c^\intercal x + F[y(\omega)] \tag{4.20a}$$

$$\text{s.t. } Ax \geq b \tag{4.20b}$$

$$\mathbb{P}_\omega\Big(G(\omega)x + L(\omega)y(\omega) \leq h(\omega)\Big) \geq 1 - \epsilon \tag{4.20c}$$

$$x \in \mathbf{R}_+^{n_1} \times \mathbf{Z}_+^{n_2}, y(\omega) \in \mathbf{R}_+^{n_3}$$

The first stage variable x could take both continuous and integer values. Notice that the second stage variable y depends on the realization of variable ω, thus it is denoted by $y(\omega)$.

With the well-known "sample average approximation" approach [72, 74, 176], Problem (4.20) could be reformulated as a *deterministic* Mixed $0-1$ Integer Program:

$$\min_{x,y^k,z_k} c^\intercal x + F[y^k] \tag{4.21a}$$

$$\text{s.t. } Ax \geq b \tag{4.21b}$$

$$G(\omega^k)x + L(\omega^k)y^k - Mz_k \leq h^k \tag{4.21c}$$

$$\sum_{k=1}^{N} \pi^k z_k \leq \epsilon \tag{4.21d}$$

$$x \in \mathbf{R}_+^{n_1} \times \mathbf{Z}_+^{n_2}, y(\omega^k) \in \mathbf{R}_+^{n_3}, z^k \in \{0,1\}$$

M is a sufficiently large coefficient and N scenarios are drawn from Ω: $\omega^1, \omega^2, \cdots, \omega^N \in \Omega$. The key idea of sample average approximation is quite simple: for scenario ω^k, if $z_k = 0$, then Eqn. (4.21c) becomes $G(\omega^k)x + L(\omega^k)y^k \leq h^k$; if $z_k = 1$, then Eqn. (4.21c) becomes $-M \leq h^k$, which is always true if M is large enough. In essence, $z_k = 0$ indicates the constraint is retained and $z_k = 1$ indicates violations are allowed for scenario ω^k. The chance constraint $\mathbb{P}_\omega(\dots) \geq 1-\epsilon$ is approximated by Eqn. (4.21d).

4.2.4 Chance-constrained Optimal Reactive Power Dispatch

4.2.4.1 Deterministic Optimal Reactive Power Dispatch

Our previous work [3] solved a look-ahead (deterministic) optimal reactive power dispatch (LA-det-ORPD) problem with voltage security constraints. Problem (4.22) is a simplified version (only one snapshot) of the LA-det-ORPD problem in [3].

$$\min \quad h_B(Q_B) + h_C(Q_C) + \lambda \sum_{c=0}^{n_c} \gamma^c P_L^c \tag{4.22a}$$

$$\text{s.t. } P^c = A_G^c(P_G + \eta^c P_\delta^c) + A_W P_W - A_D P_D, \forall c \tag{4.22b}$$

$$Q^c = A_G^c Q_G^c + A_C Q_C + A_B Q_B - A_D Q_D, \forall c \tag{4.22c}$$

$$P_\delta^c = \mathbf{1}^\intercal(A_D P_D - A_G^c P_G - A_W P_W), \forall c \tag{4.22d}$$

$$P_i^c = \sum_{j=1}^{n_b} |V_i^c||V_j^c||Y_{ij}| \cos(\theta_i^c - \theta_j^c - \phi_{ij}), \forall c, i \tag{4.22e}$$

$$Q_i^c = \sum_{j=1}^{n_b} |V_i^c||V_j^c||Y_{ij}| \sin(\theta_i^c - \theta_j^c - \phi_{ij}), \forall c, i \tag{4.22f}$$

$$P_L^c = \sum_{l=1, l:i\sim j}^{n_l} g_l\big(|V_i|^2 + |V_j|^2 - 2|V_i||V_j|\cos(\theta_i - \theta_j)\big), \forall c \tag{4.22g}$$

$$|V^c|^- \leq |V^c| \leq |V^c|^+ \tag{4.22h}$$

$$Q_B \in \{0, Q_B^+\}, \quad Q_C^- \leq Q_C \leq Q_C^+ \tag{4.22i}$$

$$Q_G^- \leq Q_G^c \leq Q_G^+ \tag{4.22j}$$

$$i, j = 1, 2, \cdots, n_b, \quad c = 0, 1, 2, \cdots, n_c$$

The objective of Problem (4.22) is to minimize the operation costs of RPSDs and transmission losses while ensuring voltage security in n_c contingency scenarios. All variables with superscript ^{c} belong to contingency scenario c[7]. In this chapter, we focus on the $N-1$ contingency of losing generators[8], which are modeled by the adjacency matrix of generators A_G^c. Let A_G^0 be the adjacency matrix in the normal operating condition (i.e. no contingency), A_G^c is obtained by setting the cth column of A_G^0 to zeros.

The decision variables in Problem (4.22) include the operating states of *discrete* RPSDs Q_B (e.g. shunt capacitors), those of *continuous* RPSDs Q_C (e.g. SVCs) and the voltage set-points of generators (i.e voltage magnitudes $|V^c|$ of PV buses). Eqn. (4.22e) and Eqn. (4.22f) are the nodal power balance constraints, P^c (Q^c) is the nodal real (reactive) power injection into the network. $A_B \in \mathbf{R}^{n_b \times n_B}$, $A_C \in \mathbf{R}^{n_b \times n_C}$, $A_D \in \mathbf{R}^{n_b \times n_D}$, $A_G^c \in \mathbf{R}^{n_b \times n_g}$ and $A_W \in \mathbf{R}^{n_b \times n_W}$ are adjacency matrices of related components. If component k is connected with bus i, then $A_\cdot(i,k) = 1$, otherwise $A_\cdot(i,k) = 0$. Alternating Current (AC) power flow equations are depicted in Eqn. (4.22e)

[7]For simplicity, the normal operating condition is denoted by $c = 0$.

[8]Since transmission line failures change the system topology thus the Y matrix in Eqn. (4.22e) and Eqn. (4.22f), we could simply modify the Y matrix to be Y^c to model the cases of losing transmission lines. For simplicity, we only focus on generator contingencies in this chapter.

and Eqn. (4.22f). $Y_{ij}\angle\phi_{ij} \in \mathbb{C}$ is associated with line (i, j) (from bus i to bus j) in the admittance matrix Y.

Losing generators causes significant real power imbalance P_δ^c, we adopt the *affine control* [102] scheme to proportionally allocate P_δ^c to each generator (i.e. $P_G + \eta^c P_\delta^c$). This guarantees the balance of real power after contingency [3, 102].

Eqn. (4.22g) calculates the real power losses and Eqn. (4.22h) is the *voltage security* constraints, which typically require the voltage magnitudes within desired ranges under a set of plausible contingency scenarios [159]. In this section, we use $[0.95, 1.05]$ for normal operation analysis ($c = 0$) and $[0.9, 1.1]$ for contingency analysis ($c = 1, 2, \cdots, n_c$). Eqn. (4.22i) and Eqn. (4.22j) are the capacity constraints for RPSDs and generators.

4.2.4.2 Chance-constrained ORPD

Motivated by the discussion in Section 4.2.3, we formulate a chance-constrained Optimal Reactive Power Dispatch (cc-ORPD) problem to ensure the voltage security of the system with respect to wind uncertainties $\xi \in \Xi$ and contingencies $c \in \mathcal{C}$. The cc-ORPD problem (Problem (4.23)) enhances the det-ORPD problem by adding a *joint* chance constraint Eqn. (4.23e). The violation probability ϵ in Eqn. (4.23e) *explicitly* quantifies the potential risk of voltage insecurity given the joint distribution of wind and contingencies $\mathcal{C} \times \Xi$.

$$\min \quad h_B(Q_B) + h_C(Q_C) + \lambda \mathbb{E}_{\mathcal{C}\times\Xi}\big[P_L(c,\xi)\big] \tag{4.23a}$$

$$\text{s.t. } P = A_G(c)P_G - A_G(c)\eta(c)P_\delta^c - A_D P_D + A_W \text{diag}(P_W)(1+\xi) \tag{4.23b}$$

$$Q = A_G(c)Q_G + A_C Q_C + A_B Q_B - A_D Q_D$$

$$+ A_W \text{diag}(Q_W)(1+\xi) \tag{4.23c}$$

$$\text{Power Flow Equations: Eqn.(4.22e), (4.22f), (4.22g)} \tag{4.23d}$$

$$\mathbb{P}_{\mathcal{C}\times\Xi}\Big(|V(c)|^- \le |V(c,\xi)| \le |V(c)|^+ \text{ for PQ buses}$$

$$\text{and } Q_G^- \le Q_G(c,\xi) \le Q_G^+\Big) \ge 1-\epsilon \tag{4.23e}$$

$$|V(c)|^- \leq |V| \leq |V(c)|^+ \text{ for PV buses} \tag{4.23f}$$

$$Q_B \in \{0, Q_B^+\}, \quad Q_C^- \leq Q_C \leq Q_C^+ \tag{4.23g}$$

$$i, j = 1, 2, \cdots, n_b, \quad c = 0, 1, 2, \cdots, n_c$$

The cc-ORPD problem is a two-stage chance-constrained programming problem. The *first-stage* variables are the operating states of RPSDs (Q_B and Q_C) and the voltage set points of generators (i.e. voltage magnitudes of PV buses). The *second-stage* variables include the nodal injection (P and Q), power imbalance P_δ, total line losses P_L, reactive generation Q_G, as well as the voltage magnitudes and angles of PQ buses ($|V|$ and θ). Since the parameters A_G^c and η^c depend on the contingency c, we change the notation to $A_G(c)$ and $\eta(c)$ for better understanding. Please notice that Eqn. (4.23b)-(4.23d) are equality constraints associated with random variable c and ξ, therefore the second-stage variables (e.g. P and P_L) also become random variables[9].

The cc-ORPD problem is very challenging to solve for the following three reasons: (1) some decision variables are binary, thus the feasible region of cc-ORPD is naturally non-convex; (2) the power flow equations are non-linear equations, which further increase the difficulty of solving cc-ORPD; and (3) the chance constraint Eqn. (4.23e) induces computationally intractable issues as discussed in Section 4.2.1.

The third difficulty could be handled via the sample average approximation approach introduced in Section 4.2.3.3. Given a set of scenarios $s^1, s^2, \cdots, s^{|\mathcal{S}|}$, where $\mathcal{S} = \mathcal{C} \times \Xi$ and each scenario $s^i = (c, \xi)^i \in \mathcal{S}$. We introduce binary variables $z^i \in \{0, 1\}$ for each scenario $s^i = (c, \xi)^i$. The chance-constraint in cc-ORPD could be re-written as a set of *deterministic* inequality constraints with binary variables z^i. Because we want to ensure the voltage security for all contingency scenarios $\mathcal{C}$, instead of drawing scenarios $(c, \xi)^i$ from $\mathcal{C} \times \Xi$, we draw samples $\xi^1, \xi^2, \cdots$ only from Ξ, and combine them with n_c contingency scenarios utilizing the fact that the generator contingency c and wind uncertainties ξ are independent. More specifically, let π_c denote the proba-

[9]More rigorous notations should denote the second-state variables are functions of c and ξ (e.g. $P(c, \xi)$ and $P_L(c, \xi)$). To avoid verbose notations, we only emphasize this in the chance constraint Eqn. (4.23e).

bility that contingency c happens, and ξ^k ($k = 1, 2, \cdots, N$) are the wind scenarios. The cc-ORPD problem is reformulated as Problem (4.24), where variables with superscripts ^{c,k} are associated with contingency c and wind scenario ξ^k.

$$\min \quad h_B(Q_B) + h_C(Q_C) + \lambda \sum_{c=0}^{n_c} \gamma^{c,k} \frac{1}{N} \sum_{s=1}^{N} P_L^{c,k}(P_W^s) \tag{4.24a}$$

$$\text{s.t. } P^{c,k} = A_G^c P_G - A_G^c \eta^c P_\delta^{c,k} - A_D P_D + A_W \text{diag}(P_W)(1 + \xi^k), \forall c, k \tag{4.24b}$$

$$Q^{c,k} = A_G^c Q_G^{c,k} + A_C Q_C + A_B Q_B - A_D Q_D + A_W \text{diag}(Q_W)(1 + \xi^k), \forall c, k \tag{4.24c}$$

$$P_\delta^{c,k} = \mathbf{1}^\intercal A_G^c P_G - \mathbf{1}^\intercal \overline{P_G} + P_W^\intercal \xi^k, \forall c, k \tag{4.24d}$$

$$P_i^{c,k} = \sum_{j=1}^{n_b} |V_i^{c,k}||V_j^{c,k}||Y_{ij}| \cos(\theta_i^{c,k} - \theta_j^{c,k} - \phi_{ij}), \forall c, s, i \tag{4.24e}$$

$$Q_i^{c,k} = \sum_{j=1}^{n_b} |V_i^{c,k}||V_j^{c,k}||Y_{ij}| \sin(\theta_i^{c,k} - \theta_j^{c,k} - \phi_{ij}), \forall c, s, i \tag{4.24f}$$

$$P_L^{c,k} = \sum_{l=1}^{n_l} g_l \big(|V_i^{c,k}|^2 + |V_j^{c,k}|^2 - 2|V_i^{c,k}||V_j^{c,k}| \cos(\theta_i^{c,k} - \theta_j^{c,k})\big), \forall c, k \tag{4.24g}$$

$$|V^{c,k}| - M z_{c,k} \leq |V^{c,k}|^+, \forall c, k \tag{4.24h}$$

$$|V^{c,k}| + M z_{c,k} \geq |V^{c,k}|^-, \forall c, k \tag{4.24i}$$

$$Q_G^{c,k} - M z_{c,k} \leq Q_G^+, \forall c, k \tag{4.24j}$$

$$Q_G^{c,k} + M z_{c,k} \geq Q_G^-, \forall c, k \tag{4.24k}$$

$$Q_B \in \{0, Q_B^+\}, \quad Q_C^- \leq Q_C \leq Q_C^+ \tag{4.24l}$$

$$\sum_{k=1}^{N} \frac{1}{N} \sum_{c=0}^{n_c} \pi_c z_{c,k} \leq \epsilon \tag{4.24m}$$

$$i, j = 1, 2, \cdots, n_b, \quad c = 0, 1, 2, \cdots, n_c, k = 1, 2, \cdots, N$$

4.2.4.3 Linearized cc-ORPD

Problem (4.24) is a Mixed Integer Non-Linear Programming (MINLP) problem, which is still computationally intractable. But the major difficulty here comes from the non-linear power flow equations. As shown in Section 4.2.3.2, we could obtain satisfying approximations via linearized power flow equations. Thus Eqn. (4.24e) and (4.24f) are linearized with respect to a known operating point (e.g. power flow solutions of a previous snapshot). Our future works include exploring other possible approaches to handle non-linearity of power flow equations (e.g. convex relaxation). Problem (4.27) is obtained by replacing Eqn. (4.24e)-(4.24f) with Eqn. (4.25). It is a Mixed Integer Linear Programming problem and is reliably solvable with commercial solvers.

$$\begin{bmatrix} P-\bar{P} \\ Q-\bar{Q} \end{bmatrix} \approx \begin{bmatrix} \frac{\partial P}{\partial \theta} & \frac{\partial P}{\partial |V|} \\ \frac{\partial Q}{\partial \theta} & \frac{\partial Q}{\partial |V|} \end{bmatrix}_{\bar{P},\bar{Q},|\bar{V}|,\bar{\theta}} \times \begin{bmatrix} \theta-\bar{\theta} \\ |V|-|\bar{V}| \end{bmatrix} \tag{4.25}$$

$$P_L-\bar{P}_L \approx \begin{bmatrix} \frac{\partial P_L}{\partial \theta} & \frac{\partial P_L}{\partial |V|} \end{bmatrix}_{\bar{P},\bar{Q},|\bar{V}|,\bar{\theta}} \times \begin{bmatrix} \theta-\bar{\theta} \\ |V|-|\bar{V}| \end{bmatrix} \tag{4.26}$$

$$\min \quad h_B(Q_B)+h_C(Q_C)+\lambda\sum_{c=0}^{n_c}\gamma^{c,k}\sum_{s=1}^{N}P_L^{c,k}(P_W^s) \tag{4.27a}$$

$$\text{s.t. Eqn. } (4.24b),(4.24c),(4.24d) \tag{4.27b}$$

$$\text{Eqn. (4.25), (4.26)} \tag{4.27c}$$

$$\text{Eqn. } (4.24h),(4.24i),(4.24j),(4.24k),(4.24l),(4.24m)$$

$$\Delta|V|^- \leq |V^{c,k}|-\overline{|V|} \leq \Delta|V|^+ \tag{4.27d}$$

$$\Delta|\theta|^- \leq |\theta^{c,k}|-\overline{|\theta|} \leq \Delta|\theta|^+ \tag{4.27e}$$

$$i,j=1,2,\cdots,n_b, \quad c=0,1,2,\cdots,n_c$$

4.2.5 Case Study

4.2.5.1 Settings

Case studies are conducted on a modified IEEE 24-bus system [3]. There are 38 contingencies considered in the case study, each one represents the scenario of losing one generator at a PV bus[10]. We assume the probability of the normal operating condition is $\pi_0 = 90\%$, and each contingency happens with equal probability, i.e. $\pi_c = 10\%/38 = 0.26\%$. By tuning the probabilities π_cs and ϵ, we could achieve a balance between a more economic system and a more secure system. The wind uncertainty ξ is assumed to be Gaussian $\xi \sim \mathcal{N}(0, 5\%)$, from which 100 scenarios ξ^k are drawn and plugged in Problem (4.27). It is worth mentioning that solving Problem (4.27) solely relies on the scenarios ξ^k, it does not require any prior knowledge on the underlying distribution.

4.2.5.2 Simulation Results

Problem (4.27) was solved via Matlab2016b and Gurobi 7.5 on a Desktop with Intel i7-2600 8-core CPU@3.40GHz and 16GB RAM memory. Gurobi found the optimal solution with 0.0% gap in 330 seconds. The optimal objective value is $\$1668.13$. Fig. 4.9 demonstrates the optimal voltage set points of generators and the voltage magnitudes of PQ buses in the normal operating condition. The voltage magnitudes of bus 4 and bus 14 are fluctuating due to wind uncertainties, while some buses (e.g. bus 17, 19 and 20) remain almost the same voltage magnitudes.

Besides the optimal solution to the cc-ORPD problem, we are also interested in the actual violation probability $\hat{\epsilon}$. Let $\bar{\epsilon}$ denote the expected violation probability: $\bar{\epsilon} := \sum_{k=1}^{N} \frac{1}{N} \sum_{c=0}^{n_c} \pi_c z_{c,k}^*$, where $z_{c,k}^*$ is from the optimal solution to Problem (4.27). It is obvious that $\bar{\epsilon} \leq \epsilon$. Let $\hat{\epsilon}$ denote the actual "out-of-sample" violation probability:

$$\hat{\epsilon} := \sum_{k=1}^{\hat{N}} \frac{1}{\hat{N}} \sum_{c=0}^{n_c} \pi_c \mathbf{1}_{Q_G^{c,k} \notin [Q_G^-, Q_G^+] \text{ or } |V^{c,k}| \notin [|V^c|^-, |V^c|^+]} \tag{4.28}$$

where $\mathbf{1}_{\text{conditions}}$ is the indicator function. We generate an independent set of $\hat{N}$ scenarios and cal-

[10] If there is only one generator at the PV bus, losing the generator will make it to a PQ bus. For simplicity, we replace it with two generators with half capacities.

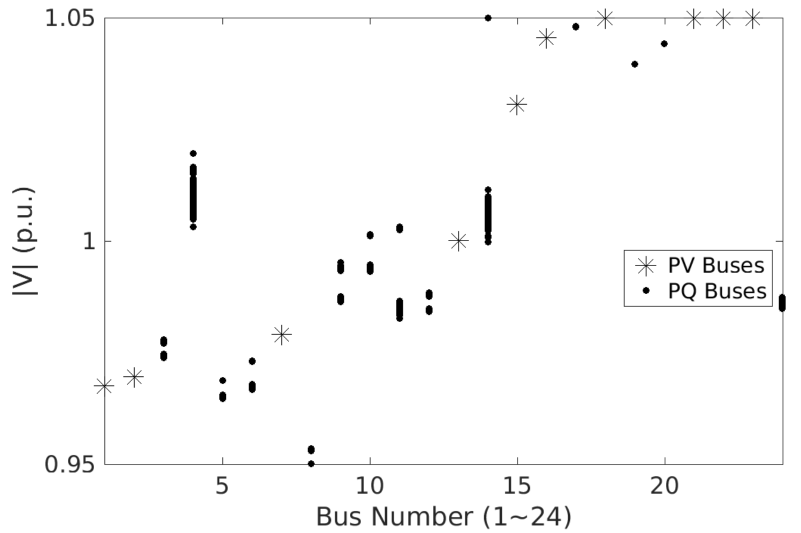


Figure 4.9: Voltage Magnitudes in the Normal Operating Condition, reprinted with permission from [4].

culate the voltage magnitudes and reactive power generations using linearized power flow equations [3] or solving the power flow equations.

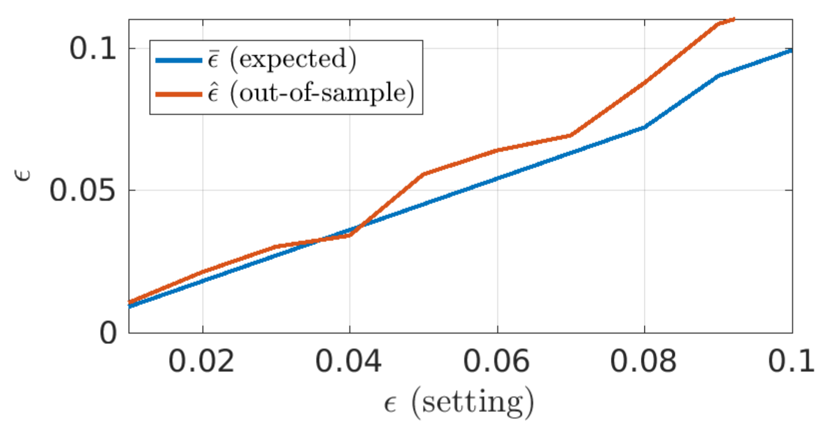


Figure 4.10: Violation Probabilities, reprinted with permission from [4].

The blue curve in Fig. 4.10 is the expected violation probability $\bar{\epsilon}$ from the optimal solution z^*. And the red line $\hat{\epsilon}$ is calculated on $\hat{N} = 100$ scenarios using linearized power flow equations [3].

The out-of-sample violation probability $\hat{\epsilon}$ is very close to $\bar{\epsilon}$. With a larger number of scenarios embedded in Problem (4.27), the expected $\bar{\epsilon}$ and actual $\hat{\epsilon}$ will be closed to the violation probability ϵ in the chance constraint.

We also compare the results of cc-ORPD (Problem (4.27)) with det-ORPD (Problem (4.22)). With a little sacrifice on the total cost, the cc-ORPD could ensures voltage security with probability 98.8%. While the results of det-ORPD lead to voltage magnitudes lower than the desired lower bound $|V^c|^-$. In the results of det-ORPD, we even observe undesirable low voltage magnitudes in the normal operating condition, which results in the large violation probability in Table 4.2.

Table 4.2: Comparison: det-ORPD vs cc-ORPD, reprinted with permission from [4].

	det-ORPD	**cc-ORPD** ($\epsilon = 0.01$)
Objective	1610.2	1668.1
$\hat{\epsilon}$	52.1%	1.2%

4.3 Summary

In this chapter, we first present a look-ahead coordination framework for reactive power support devices across multiple control areas to ensure voltage security. The proposed framework considers the global needs on reactive support and determines efficient cooperation schedules of the reactive devices. Given the fact that the original problem is non-convex, we reformulate the problem using power flow Jacobian matrix. Case studies demonstrate that the approximation is satisfying and the solution ensues the voltage security given all the contingency scenarios.

Based on the look-ahead coordination framework for reactive power devices, this chapter further proposes a chance-constrained formulation of optimal power reactive dispatch to schedule RPSDs considering uncertainties from wind and contingencies. The cc-ORPD problem is reformulated as a computationally solvable form using sample average approximation and linearized power flow equations. Case studies demonstrate the effectiveness of the proposed cc-ORPD framework. Future works include analyzing the performance gap between linearized solution and the global optimal solution, investigating convex relaxations of power flow equations and utilizing improved versions of the sample average approximation [72, 74].

5. SECURITY-CONSTRAINED ECONOMIC DISPATCH WITH PROBABILISTIC GUARANTEES[1]

5.1 Security-Constrained Economic Dispatch (SCED)

5.1.1 Deterministic SCED

Security-constrained Economic Dispatch (SCED) lies at the center of modern electricity markets and short-term power system operations. It determines the most cost-efficient output levels of generators while keeping the real-time balance between supply and demand. Different variations of the SCED problem are all based on the direct current optimal power flow (DCOPF) problem. We present a typical form of DCOPF with wind generation.

$$\text{(det-DCOPF)}: \min_{g} \; c(g) \tag{5.1a}$$

$$\text{s.t } \mathbf{1}^{\intercal} g = \mathbf{1}^{\intercal} d - \mathbf{1}^{\intercal} \hat{w} \tag{5.1b}$$

$$f = H_g g + H_w \hat{w} - H_d d \tag{5.1c}$$

$$\underline{f} \leq f \leq \overline{f} \tag{5.1d}$$

$$\underline{g} \leq g \leq \overline{g} \tag{5.1e}$$

The decision variables are generation output levels $g \in \mathbf{R}^{n_g}$. The objective of (det-DCOPF) is to minimize total generation cost $c(g)$, while ensuring total generation equates total *net* demand [2] (5.1b). Constraints include transmission line flow limits (5.1c)-(5.1d) and generation capacity limits (5.1e). Transmission line flows $f \in \mathbf{R}^{n_l}$ are calculated using (5.1c), in which H is the power transfer distribution factor (PTDF) matrix, and $H_g \in \mathbf{R}^{n_l \times n_g}$ ($H_d \in \mathbf{R}^{n_l \times n_d}$,$H_w \in \mathbf{R}^{n_l \times n_w}$) denotes the submatrix formed by the columns of H corresponding to generators (loads, wind farms). (5.1) utilizes the expected wind generation or wind forecast $\hat{w}$, we refer to (5.1) as *deterministic DCOPF* (det-DCOPF) since no uncertainties are being considered.

[1]Parts of this chapter are reprinted with permission from [1].

[2]Wind generation is treated as negative loads.

5.1.2 Chance-constrained SCED

Many researchers advance (det-DCOPF) towards a chance-constrained formulation with wind uncertainties. A representative formulation is (5.2), which appears in a majority of the existing literatures, e.g. [102, 118].

(cc-DCOPF):

$$\min_{g,\eta} \; c(g) \tag{5.2a}$$

$$\text{s.t } \mathbf{1}^\intercal g = \mathbf{1}^\intercal d - \mathbf{1}^\intercal \hat{w} \tag{5.2b}$$

$$f(\hat{w}, \tilde{w}) = H_g(g - \mathbf{1}^\intercal \tilde{w}\eta) - H_d d + H_w(\hat{w} + \tilde{w}) \tag{5.2c}$$

$$\mathbb{P}_{\tilde{w}}\Big(\underline{f} \leq f(\hat{w}, \tilde{w}) \leq \overline{f} \text{ and } \underline{g} \leq g - \mathbf{1}^\intercal \tilde{w}\eta \leq \overline{g}\Big) \geq 1 - \epsilon \tag{5.2d}$$

$$\mathbf{1}^\intercal \eta = 1 \tag{5.2e}$$

$$\underline{g} \leq g \leq \overline{g} \tag{5.2f}$$

$$-\mathbf{1} \leq \eta \leq \mathbf{1} \tag{5.2g}$$

Unlike (det-DCOPF) using wind forecast $\hat{w}$, chance-constrained SCED (cc-DCOPF) explicitly models wind generation as a random vector $w \in \mathbf{R}^{n_w}$. The wind generation $w = \hat{w} + \tilde{w}$ is decomposed into two components: the *deterministic* wind forecast value $\hat{w} \in \mathbf{R}^{n_w}$ and the *uncertain* forecast error $\tilde{w} \in \mathbf{R}^{n_w}$. To guarantee the real-time balance of supply and demand, (cc-DCOPF) introduces an affine control policy $\eta \in [-1, 1]^{n_g}$ to proportionally allocate total wind fluctuations $\mathbf{1}^\intercal \tilde{w}$ to each generator. It is easy to verify that constraints (5.2b) and (5.2e) imply the supply-demand balance in the presence of wind uncertainties, i.e.

$$\mathbf{1}^\intercal(g - \mathbf{1}^\intercal \tilde{w}\eta) = \mathbf{1}^\intercal d - \mathbf{1}^\intercal(\hat{w} + \tilde{w}), \tag{5.3}$$

The affine policy vector $\eta \in \mathbf{R}^{n_g}$ is sometimes referred as participation factor or distribution vector [118]. The (joint) chance constraint (5.2d) constrains the transmission flow and generation within their capacities with high probability $1 - \epsilon$ in the presence of wind uncertainties.

For simplicity, we only account for the major source of uncertainties (i.e. wind) in the real-time. Many references provides more complicated formulation of (cc-DCOPF), e.g. considering joint uncertainties from load and wind [107, 142], and contingencies of potential generator or transmission line outages [104].

There exist a few different but similar formulations of (cc-DCOPF). In general, policies of any form could help balance supply with demand under uncertainties. The affine policy in (cc-DCOPF) is the simplest choice and lead to optimization problems that are easy to solve. There are other papers applying different forms of policies, e.g. [100] introduces a matrix form of the affine policy $\Upsilon \in \mathbf{R}^{n_g \times n_w}$, which specifies the corrective control of each generator on each wind farm. (cc-DCOPF) is a single snapshot dispatch problem, it is straightforward to extend it to a multi-period or look-ahead dispatch problem [118,123]. Many papers evaluate the impacts of new elements in modern power systems, such as demand response [122, 181], ambient temperatures and meteorological quantities [117], and frequency control [103,181].

Although DC power flow equations have been widely accepted in modern power system operations and planning, it is only a linear approximation of the alternating current (AC) version, which is a more accurate model of the underlying physical laws. Many efforts have been made to solve the chance-constrained AC optimal power flow (cc-ACOPF) problem, e.g. [182–185]. Major difficulties to solve cc-ACOPF come from the non-convexity of AC power flow equations. It remains as an open question that how to ensure the feasibility of the non-convex AC power flow equations under uncertainties.

5.1.3 Solving cc-DCOPF

Table 2.1 summarizes various methods to solve (cc-DCOPF). The most popular one consists of two steps: (i) decomposing the joint chance constraint (5.2d) into individual ones $\mathbb{P}_\xi(f_i(x,\xi) \leq 0) \geq 1 - \epsilon_i, i = 1, 2, \cdots, m$; (ii) deriving the deterministic equivalent form of each individual chance constraint by making the Gaussian assumption. More technical details of this method are in Section 2.2.2. This method is taken by many researchers for its simplicity and computationally tractable reformulation. Although the Gaussian assumption enjoys the law of large numbers, it is often an approximation or even doubtful assumption. For example, [174] shows that the wind forecast error is better represented by Cauchy distributions instead of Gaussian ones. The first step of this method is to decompose a joint chance constraint $\mathbb{P}_\xi(f(x,\xi) \leq 0) \geq 1-\epsilon$ into individual ones. As discussed in Section 2.1.2 and 2.6.4.1, this step often introduces conservativeness because of the limitation of Bonferroni inequality. The level of conservativeness could be significant when the number of constraints m is large, which is typically the case in power systems.

The scenario approach is another commonly-accepted method. It provides rigorous guarantees on the quality of the solution and does not assume the distribution is Gaussian or any particular type. Most papers adopting the scenario approach apply the *a-priori* guarantees (e.g. Theorem 5 and 6) on (cc-DCOPF) and verify the a-posteriori feasibility of solutions through Monte-Carlo simulations (2.25). One common observation is that the solution x_N^* is often quite conservative, i.e. $\mathbb{V}(x_N^*) \ll \epsilon$. One major source of conservativeness is the loose sample complexity bounds N [3]. Since (cc-DCOPF) is convex, Theorem 4 states that the number of decision variables n is an upper bound of the number of support scenarios $|\mathcal{S}|$ or Helly's dimension h. This upper bound, as pointed out in [123], is indeed very loose. [123] reported only ~ 5 support scenarios for a chance-constrained look-ahead SCED problem with thousands of decision variables. By exploiting the structural features of (cc-DCOPF), the sample complexity bound N can be significantly improved. Unfortunately, only [123] and [122] followed this path to reduce conservativeness.

There are also many papers utilizing the robust optimization related methods to solve (cc-DCOPF). [133] constructs uncertainty sets with the help of probabilistic guarantees in [58]. References [135, 136] incorporate the convex approximation framework and compare different choices of generating functions $\phi(z)$ on (cc-DCOPF). Although there are no explicit forms of chance constraints in [134], the CVaR-oriented approach therein can be interpreted as solving cc-DCOPF using convex approximation with the choice of Markov bound.

Most papers in Table 2.1 aim at finding suboptimal solutions to (cc-DCOPF). However, it

[3] Many papers still utilize the first sample complexity bound proved in [60], which was significantly tightened in [62] and following works [63].

is somewhat surprising to note that none of them estimates how suboptimal the solution is via approaches like Proposition 2 or 4. Almost all the papers evaluate the a-posteriori feasibility by Monte-Carlo simulations with a huge sample size. Methods like Proposition 1 would be more attractive when data is limited, which is closer to the reality.

5.2 Numerical Simulations

5.2.1 Simulation Settings

Chance-constrained DCOPF (5.2) serves as a benchmark problem for a critical comparison of solutions to (CCO). We provide numerical solutions of cc-DCOPF on two test systems: a 3-bus system and the IEEE 24-bus RTS test system.

The 3-bus system is a modified version of the 3-bus system in [186]. The major difference is the removal of the load at bus 2 and the synchronous condensor at bus 3 in order to visualize the feasible region and the space of uncertainties. The original 3-bus system "*case3sc.m*" is available in the Matpower toolbox [172]. The modified system in this section can be found in the examples of CCC [4]. For simplicity, we only consider uncertainties of loads, which is modeled as Gaussian variables with 5% standard variation.

The 24-bus system in this section is a modified version of the IEEE 24-bus RTS benchmark system [171]. The transmission line capacities are set to be 60% of the original capacities. We conduct two sets of simulations on the 24-bus system with different distributions of uncertainties. The first one is similar with the 3-bus case, nodal loads are modeled as independent Gaussian variables with 5% standard deviation. The second one models the errors of nodal load forecasts as independent beta-distributed random variables, with parameters $\alpha = 25.2414$ and $\beta = 25.2692$ [5].

Ten Monte-Carlo simulations are conducted on every method to examine the randomness of solutions. For the 3-bus case, each Monte-Carlo simulation uses 100 i.i.d samples to solve cc-DCOPF. 2048 points are used in each run to solve (cc-DCOPF) of the 24-bus system. The returned solutions are evaluated on an independent set of 10^4 points.

[4] github.com/xb00dx/ConvertChanceConstraint-ccc/tree/master/examples

[5] This setting of beta distribution is from [174], and scaled from $[0, 1]$ to $[-18\%, 18\%]$.

We use Gurobi 7.10 [87] to get results of scenario approach and sample average approximation. Cplex 12.8 is used to solve (CCO) with robust counterpart and convex approximation.

5.2.2 Simulation Results

We solve cc-DCOPF on the 3-bus system with eight different methods: (1) scenario approach with prior guarantees, (SA:prior, Corollary 1); (2) scenario approach with posterior guarantees (SA:posterior, Theorem 7); (3) sample average approximation, where N and ε are chosen based on the sampling and discarding Theorem (SAA:s&d, Theorem 9); (4-7) Robust counterpart with different uncertainty sets specified in Theorem 13: box (RC:box), ball (RC:ball), ball-box (RC:ball-box) and budget (RC:budget) uncertainty sets; (8) convex approximation with Markov bound (CA:markov, Theorem 11 and Proposition 5).

We first examine the feasibility of the returned solutions from eight algorithms. Figure 5.2 and 5.3 show the out-of-sample violation probabilities $\hat{\epsilon}$ versus desired ϵ in the setting. The green dashed lines in Figure 5.2 and 5.3 denote the ideal case where $\hat{\epsilon} = \epsilon$. Any points above the green dashed line indicate infeasible solutions that $\mathbb{V}(x) > \epsilon$. Clearly all methods return feasible solutions (with high probability) to (CCO). From Figure 5.2, sample average approximation and convex approximation are less conservative than other methods. However, it is worth noting that when ϵ is small (e.g. 10^{-2}), the data-driven approximation of CVaR (Proposition 5) does not necessarily give a safe approximation to (CCO) [34]. The robust counterpart methods are typically $10 \sim 100$ times more conservative than other methods, as illustrated in the comparison of Figure 5.3a with Figure 5.3b. The conservativeness could be significantly reduced by better construction of uncertainty sets, e.g. [34, 57]. Among four different choices of uncertainty sets, the ball-box set is the least conservative one, which combines the advantages of the ball and box uncertainty sets.

Figure 5.3-5.4 present the results of the 24-bus system with Gaussian distributions. Simulation results of the beta distribution are in Figure 5.5-5.7. Observations from Figure 5.5-5.7 are similar with the case of Gaussian distributions. Every method behaves more conservative in the case of beta distributions than the case of Gaussian distributions. It is worth noting that the RO-based methods (RC:box, RC:ball, RC:ball-box in Figure 5.6) are so conservative that lead to zero

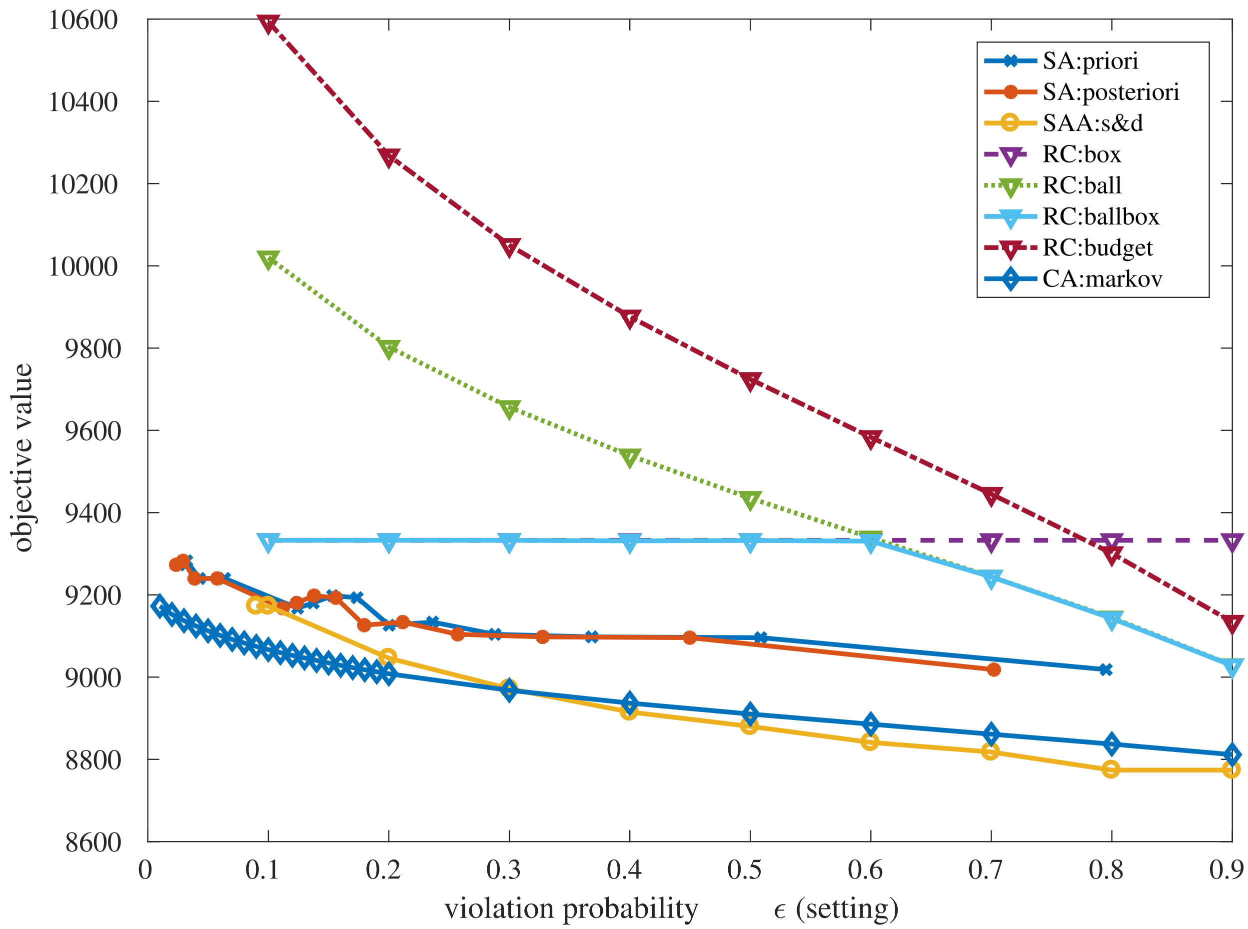
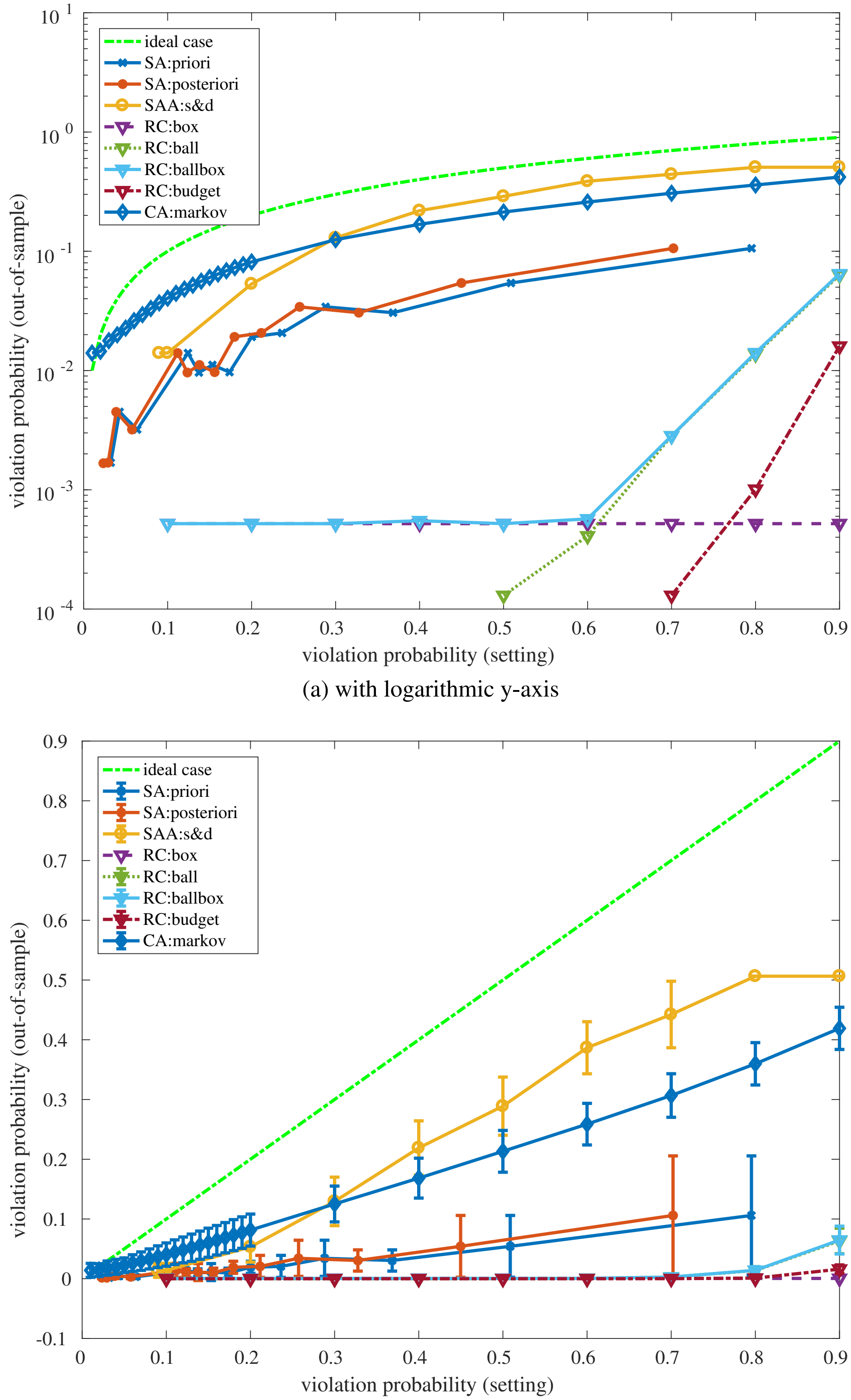


Figure 5.1: Objective Values (cc-DCOPF of the 3-bus System), reprinted with permission from [1].

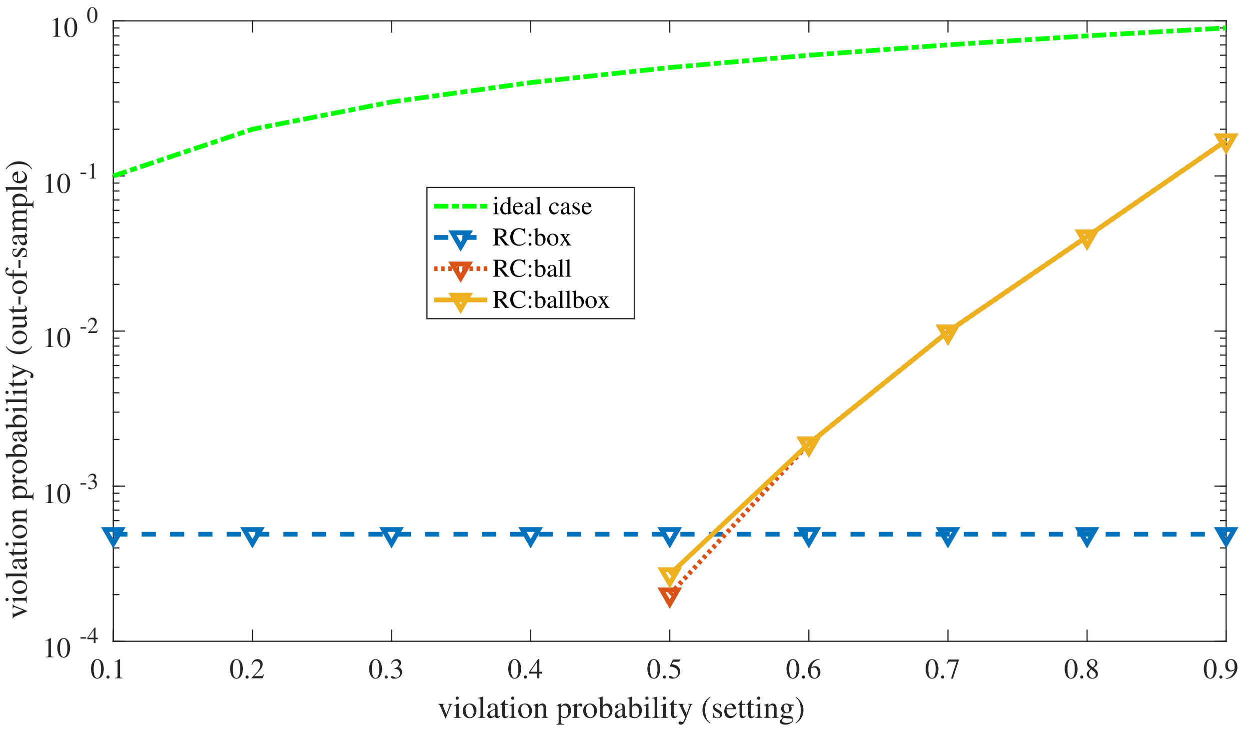
empirical violation probability $\hat{\epsilon}$.



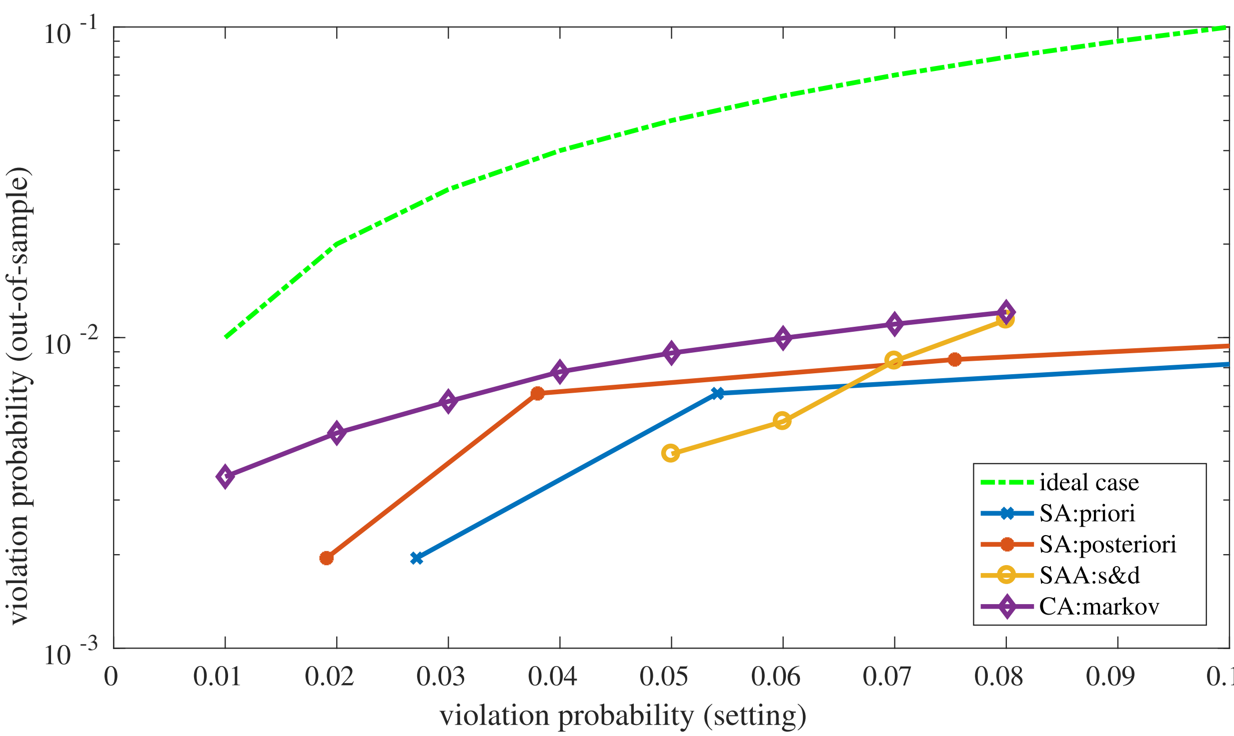
(a) with logarithmic y-axis

(b) with error bars showing standard deviations

Figure 5.2: Violation Probabilities (cc-DCOPF of the 3-bus System), reprinted with permission from [1].

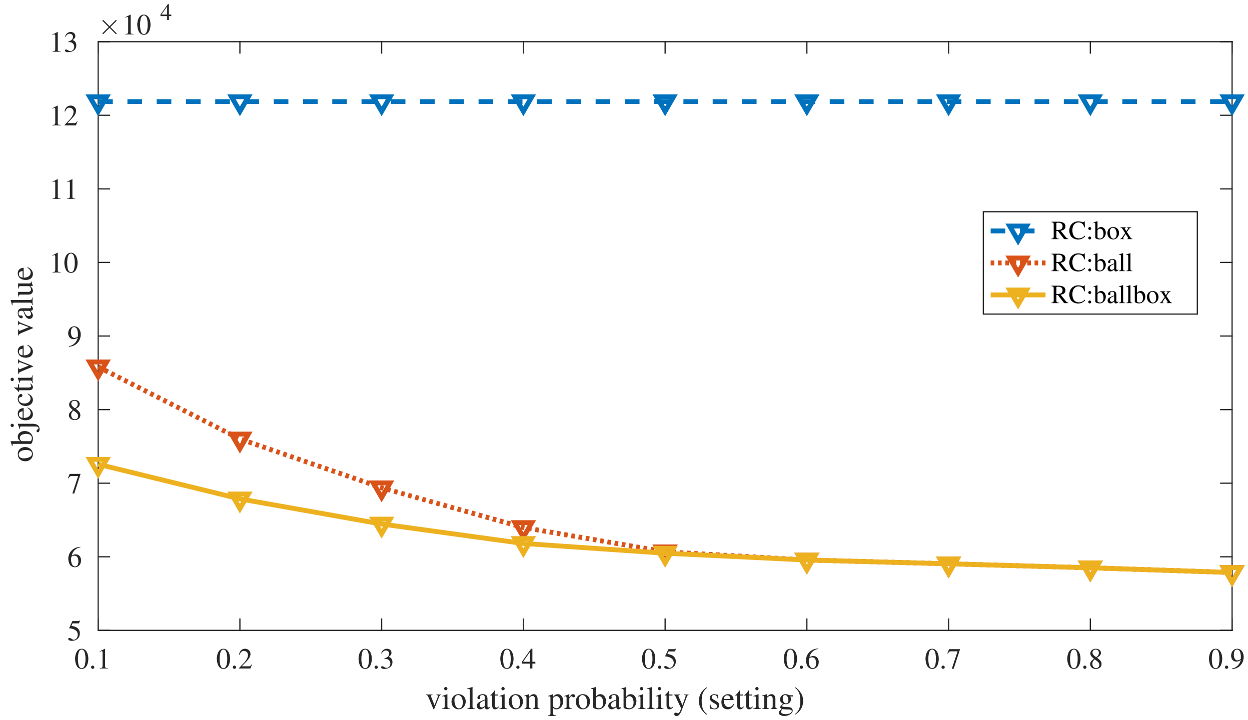


(a) robust counterpart methods

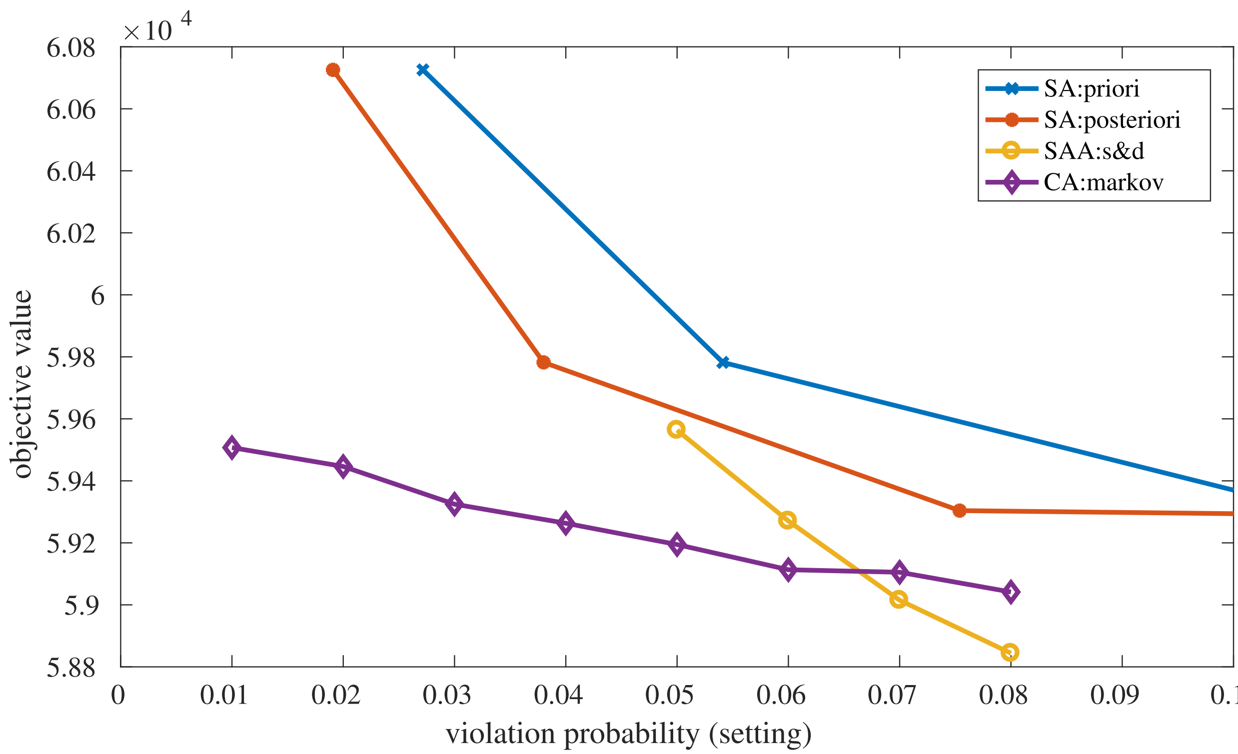


(b) other methods

Figure 5.3: Violation Probabilities (cc-DCOPF of the 24-bus System, Gaussian Distributions), reprinted with permission from [1].



(a) robust counterpart methods



(b) other methods

Figure 5.4: Objective Values (cc-DCOPF of the 24-bus System, Gaussian Distributions), reprinted with permission from [1].

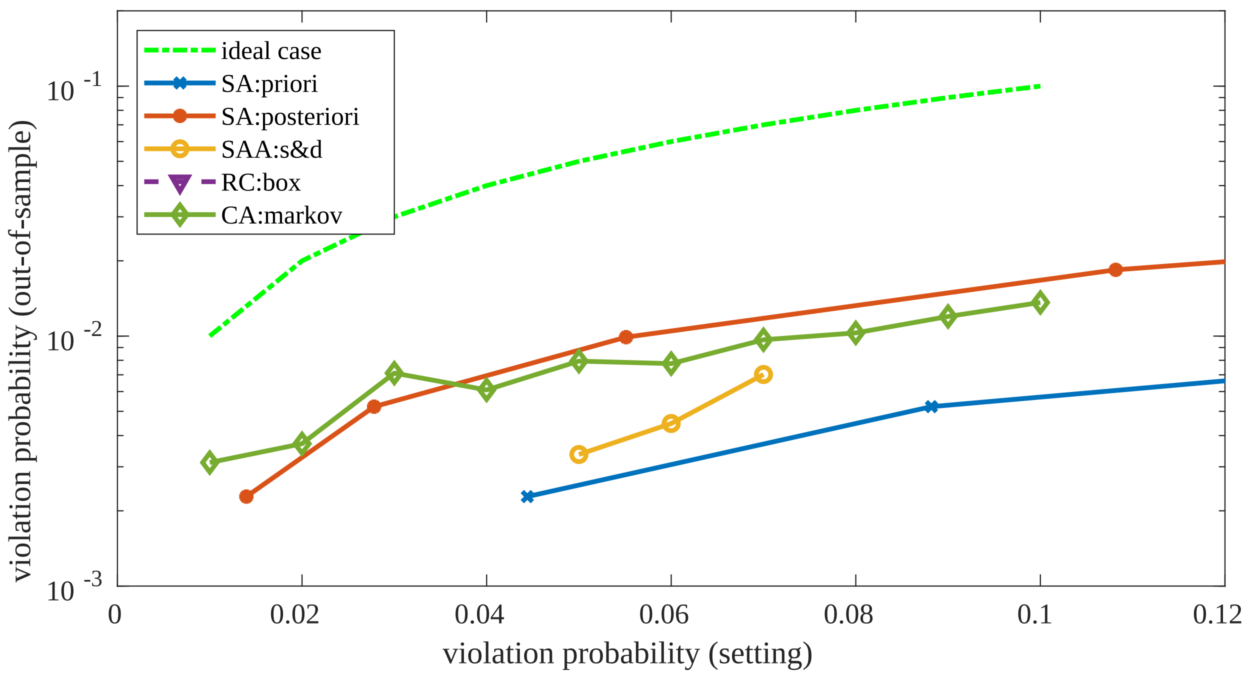


Figure 5.5: Violation Probabilities in Logarithmic Scale (cc-DCOPF of the 24-bus System, Beta Distributions), reprinted with permission from [1].

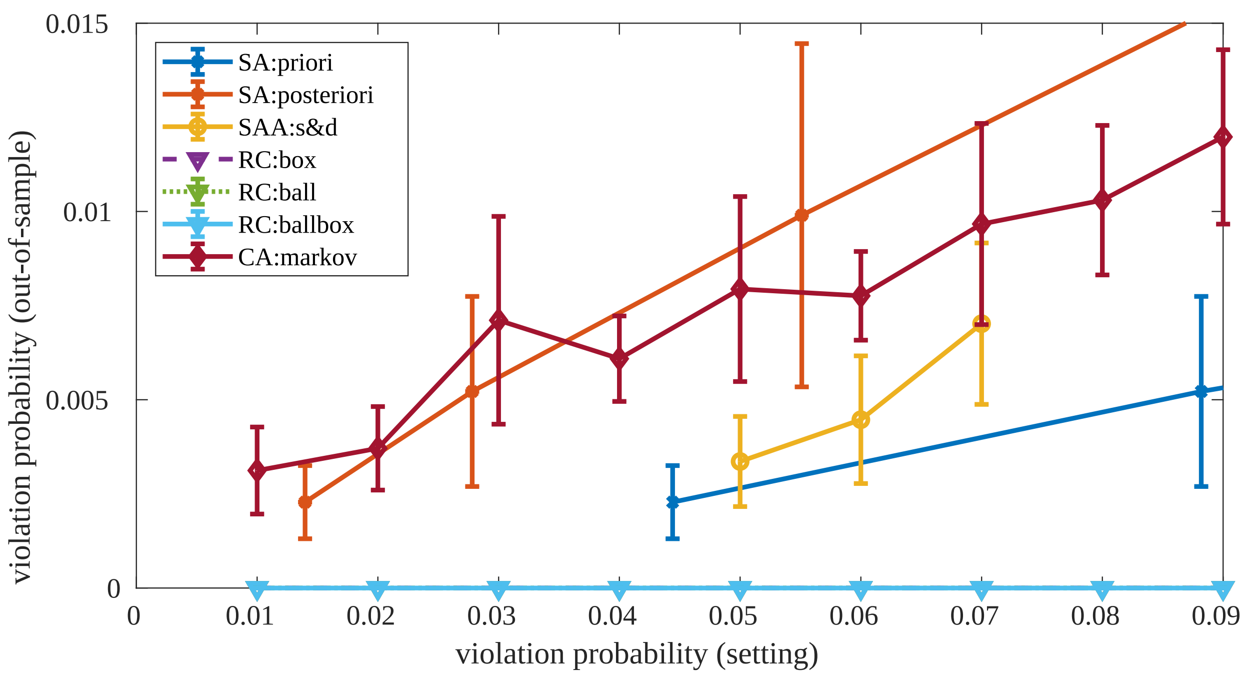
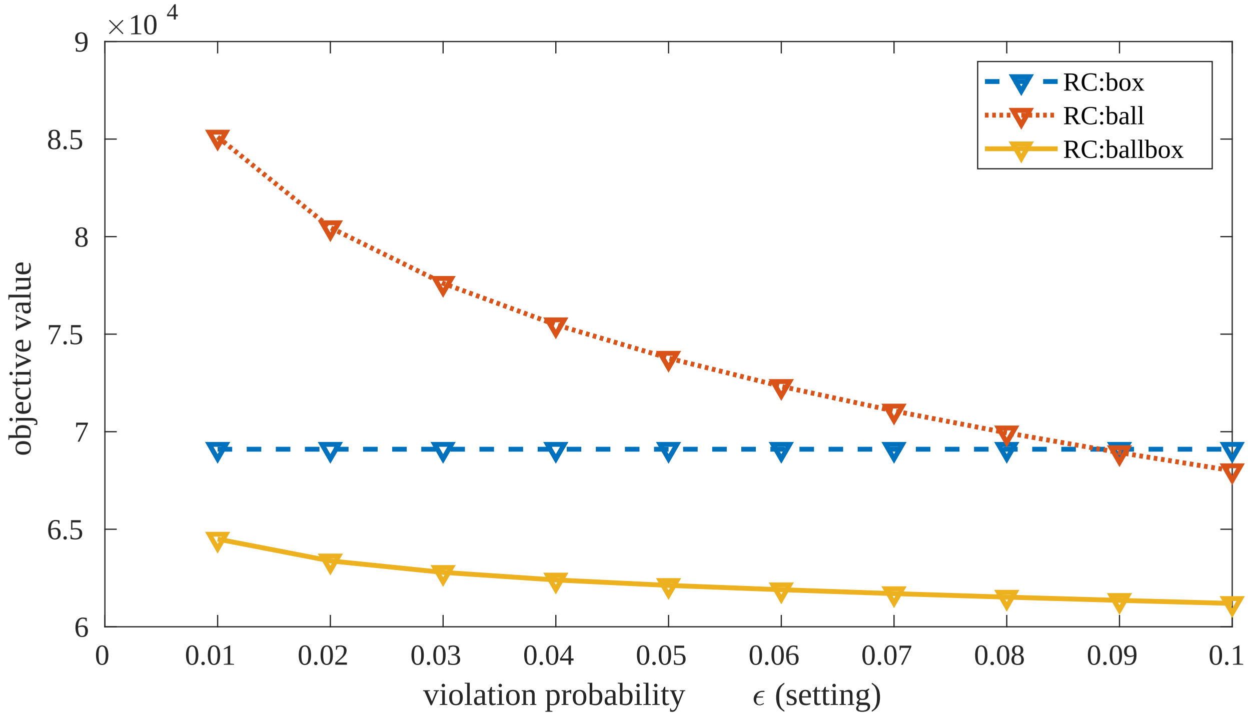
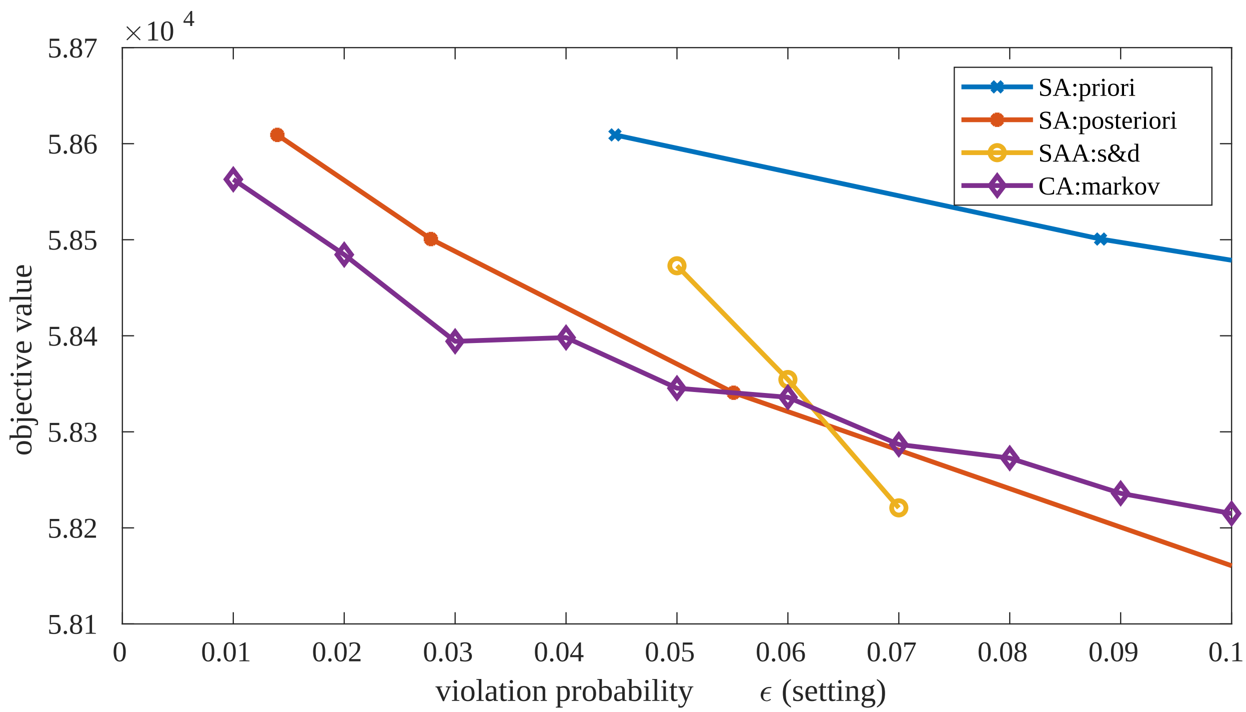


Figure 5.6: Violation Probabilities with error bars showing standard deviations (cc-DCOPF of the 24-bus System, Beta Distributions), reprinted with permission from [1].



(a) robust counterpart methods



(b) other methods

Figure 5.7: Objective Values (cc-DCOPF of the 24-bus System, Beta Distributions), reprinted with permission from [1].

6. CONCLUSIONS AND FUTURE WORK

We conclude this dissertation and propose several directions for future research in this chapter.

6.1 Summary

This dissertation first presents a comprehensive review on the fundamental properties, key theoretical results, and three categories of algorithms for chance-constrained optimization. An open-source MATLAB toolbox ConvertChanceConstraint is developed to automate the process of translating chance constraints to compatible forms for mainstream optimization solvers. Accordingly, we propose a three-stage framework for power system operations with probabilistic guarantees.

In the day-ahead operational planning stage, we formulate the chance-constrained unit commitment problem and solve it via the scenario approach. We show that the structural property of unit commitment makes the scenario approach applicable in the presence of non-convexity. It substantially reduces the necessary number of scenarios and could be further exploited to reduce the computational requirement to solve the problem.

In the intra-day adjustment period, we formulate the chance-constrained optimal power reactive dispatch problem to schedule reactive power support devices considering uncertainties from renewables and contingencies. The cc-ORPD problem is reformulated using sample average approximation and linearized power flow equations. Case studies demonstrate the effectiveness of the proposed cc-ORPD framework.

In the real-time operation stage, we formulate and solve the chance-constrained security-constrained economic dispatch problem. The cc-SCED problem also serves as a benchmark problem for a critical comparison of existing algorithms to solve chance-constrained programs on IEEE benchmark systems.

6.2 Future Work

Many interesting directions are open for future research.

In terms of theoretical investigation, an analytical comparison of existing solutions to chance-constrained optimization is necessary to substantiate the fundamental insights obtained from numerical simulations. As discussed in Chapter 2, chance-constrained optimization could be generalized towards ambiguous chance constraints or distributionally robust optimization (DRO) problems. Generalizations of the results in Chapter 2 towards DRO would be an important part of future work. When working on [82] and Chapter 2, we realized that many methods reviewed in Chapter 2 originated from statistical learning (e.g. the scenario approach). Some famous algorithms (e.g. [10, 187–190]) in machine learning can be interpreted from the viewpoint of chance-constrained optimization. And chance-constrained optimization can be combined with many machine learning algorithms (e.g. [13,191]) The connection between chance-constrained optimization and statistical learning (and machine learning) would be a very attractive direction to be explored.

In terms of applications, chance-constrained optimization in electric energy systems could go beyond operational planning practices. For example, it would be worth investigating into the economic interpretation of market power issues through the lens of chance-constrained optimization. Many other power system decision making processes under uncertainties (especially in the distribution networks) could be enhanced using chance-constrained optimization. Another interesting direction is dealing with non-convexity in chance-constrained optimization. Alternating current (AC) power flow equations, which lay the foundation for many power system applications, are a set of non-linear equations and often brings non-convexities into power system optimization problems. Chapter 4 demonstrates one solution to deal with non-convexities brought about by the power flow equations. More in-depth and rigorous analysis on AC power flow equations (e.g. convex restriction [192–194] or relaxation [195–197]) are necessary.

REFERENCES

[1] X. Geng and L. Xie, “Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization,” *Annual Reviews in Control*, 2019.

[2] X. Geng, L. Xie, and M. S. Modarresi, “A General Scenario Theory for Security-Constrained Unit Commitment with Probabilistic Guarantees,” *arXiv preprint arXiv:1910.07672*, 2019.

[3] X. Geng, L. Xie, and D. Obadina, “Voltage Security Constrained Look-ahead Coordination of Reactive Power Support Devices with High Renewables,” in *Proceedings of the 19th International Conference on Intelligent System Applications to Power Systems (ISAP)*, 2017.

[4] X. Geng, L. Xie, and D. Obadina, “Chance-constrained Optimal Reactive Power Dispatch,” in *In Proceedings of 2018 Power and Energy Society General Meeting (PESGM)*, IEEE, 2018.

[5] M. E. Yaari, “Uncertain lifetime, life insurance, and the theory of the consumer,” *The Review of Economic Studies*, vol. 32, no. 2, pp. 137–150, 1965.

[6] G. C. Calafiore, M. C. Campi, *et al.*, “The scenario approach to robust control design,” *IEEE Transactions on Automatic Control*, 2006.

[7] N. V. Sahinidis, “Optimization under uncertainty: state-of-the-art and opportunities,” *Computers & Chemical Engineering*, vol. 28, no. 6-7, pp. 971–983, 2004.

[8] R. Henrion, P. Li, A. Möller, M. C. Steinbach, M. Wendt, and G. Wozny, “Stochastic optimization for operating chemical processes under uncertainty,” in *Online optimization of large scale systems*, pp. 457–478, Springer, 2001.

[9] J. Dupačová, A. Gaivoronski, Z. Kos, and T. Szantai, “Stochastic programming in water management: A case study and a comparison of solution techniques,” *European Journal of*

Operational Research, vol. 52, no. 1, pp. 28–44, 1991.

[10] H. Xu, C. Caramanis, and S. Mannor, "Robust regression and lasso," in *Advances in Neural Information Processing Systems*, pp. 1801–1808, 2009.

[11] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009.

[12] C. Caramanis, S. Mannor, and H. Xu, "14 Robust Optimization in Machine Learning," *Optimization for machine learning*, p. 369, 2012.

[13] A. Ben-Tal, S. Bhadra, C. Bhattacharyya, and J. Saketha Nath, "Chance constrained uncertain classification via robust optimization," *Mathematical Programming*, vol. 127, pp. 145–173, Mar. 2011.

[14] S. Sra, S. Nowozin, and S. J. Wright, *Optimization for machine learning*. Mit Press, 2012.

[15] V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization: An overview," *European journal of operational research*, vol. 235, no. 3, pp. 471–483, 2014.

[16] U. A. Ozturk, M. Mazumdar, and B. A. Norman, "A solution to the stochastic unit commitment problem using chance constrained programming," *IEEE Trans. Power Syst*, 2004.

[17] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst*, 2012.

[18] A. Charnes, W. W. Cooper, and G. H. Symonds, "Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil," *Management Science*, 1958.

[19] A. Charnes and W. W. Cooper, "Chance-constrained programming," *Management science*, vol. 6, no. 1, pp. 73–79, 1959.

[20] A. Charnes and W. W. Cooper, "Deterministic equivalents for optimizing and satisficing under chance constraints," *Operations research*, 1963.

[21] S. Kataoka, "A stochastic programming model," *Econometrica: Journal of the Econometric Society*, pp. 181–196, 1963.

[22] J. Pintér, "Deterministic approximations of probability inequalities," *Zeitschrift für Operations-Research*, vol. 33, no. 4, pp. 219–239, 1989.

[23] S. Sen, "Relaxations for probabilistically constrained programs with discrete random variables," in *System Modelling and Optimization*, pp. 598–607, Springer, 1992.

[24] A. Prekopa, B. Vizvari, and T. Badics, "Programming under probabilistic constraint with discrete random variable," in *New trends in mathematical programming*, pp. 235–255, Springer, 1998.

[25] A. P. Ruszczynski and A. Shapiro, *Stochastic programming*, vol. 10. Elsevier Amsterdam, 2003.

[26] A. Prékopa, "Stochastic programming," 1995.

[27] B. L. Miller and H. M. Wagner, "Chance constrained programming with joint constraints," *Operations Research*, vol. 13, no. 6, pp. 930–945, 1965.

[28] A. Ruszczyński, "Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra," *Mathematical Programming*, vol. 93, no. 2, pp. 195–215, 2002.

[29] J. Luedtke and S. Ahmed, "A sample approximation approach for optimization with probabilistic constraints," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 674–699, 2008.

[30] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM Journal on Optimization*, 2006.

[31] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, 2009.

[32] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[33] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[34] W. Chen, M. Sim, J. Sun, and C.-P. Teo, "From CVaR to uncertainty set: Implications in joint chance-constrained optimization," *Operations research*, vol. 58, no. 2, pp. 470–485, 2010.

[35] L. Khachiyan, "The problem of calculating the volume of a polyhedron is enumerably hard," *Russian Mathematical Surveys*, 1989.

[36] J. Luedtke, S. Ahmed, and G. L. Nemhauser, "An integer programming approach for linear programs with probabilistic constraints," *Mathematical programming*, vol. 122, no. 2, pp. 247–272, 2010.

[37] F. Qiu, S. Ahmed, S. S. Dey, and L. A. Wolsey, "Covering linear programming with violations," *INFORMS Journal on Computing*, vol. 26, no. 3, pp. 531–546, 2014.

[38] S. Ahmed and W. Xie, "Relaxations and approximations of chance constraints under finite distributions," *Mathematical Programming*, vol. 170, no. 1, pp. 43–65, 2018.

[39] A. Prékopa, "Logarithmic concave measures with application to stochastic programming," *Acta Scientiarum Mathematicarum*, vol. 32, pp. 301–316, 1971.

[40] A. Prékopa, K. Yoda, and M. M. Subasi, "Uniform quasi-concavity in probabilistic constrained stochastic programming," *Operations Research Letters*, vol. 39, no. 3, pp. 188–192, 2011.

[41] C. Lagoa, "On the convexity of probabilistically constrained linear programs," in *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 1, pp. 516–521, IEEE, 1999.

[42] G. C. Calafiore and L. El Ghaoui, "On distributionally robust chance-constrained linear programs," *Journal of Optimization Theory and Applications*, vol. 130, no. 1, pp. 1–22, 2006.

[43] R. Henrion and C. Strugarek, "Convexity of chance constraints with independent random variables," *Computational Optimization and Applications*, vol. 41, no. 2, pp. 263–276, 2008.

[44] R. Henrion and C. Strugarek, "Convexity of chance constraints with dependent random variables: the use of copulae," in *Stochastic Optimization Methods in Finance and Energy*, pp. 427–439, Springer, 2011.

[45] W. Van Ackooij, "Eventual convexity of chance constrained feasible sets," *Optimization*, vol. 64, no. 5, pp. 1263–1284, 2015.

[46] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[47] P. Beraldi and A. Ruszczyński, "The probabilistic set-covering problem," *Operations Research*, vol. 50, no. 6, pp. 956–967, 2002.

[48] M. Kress, M. Penn, and M. Polukarov, "The minmax multidimensional knapsack problem with application to a chance-constrained problem," *Naval Research Logistics (NRL)*, vol. 54, no. 6, pp. 656–666, 2007.

[49] S. Ahmed and A. Shapiro, "Solving chance-constrained stochastic programs via sampling and integer programming," *Tutorials in Operations Research*, vol. 10, pp. 261–269, 2008.

[50] A. Nemirovski, "On safe tractable approximations of chance constraints," *European Journal of Operational Research*, vol. 219, no. 3, pp. 707–718, 2012.

[51] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming*, pp. 1–52, 2015.

[52] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Mathematical Programming*, vol. 158, no. 1-2, pp. 291–327, 2016.

[53] A. Shapiro, "Distributionally robust stochastic programming," *SIAM Journal on Optimization*, vol. 27, no. 4, pp. 2258–2275, 2017.

[54] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

[55] W. Chen and M. Sim, “Goal-driven optimization,” *Operations Research*, vol. 57, no. 2, pp. 342–357, 2009.

[56] R. T. Rockafellar and S. Uryasev, “Conditional value-at-risk for general loss distributions,” *Journal of banking & finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[57] D. Bertsimas, V. Gupta, and N. Kallus, “Data-driven robust optimization,” *Mathematical Programming*, 2018.

[58] D. Bertsimas and M. Sim, “The price of robustness,” *Operations research*, vol. 52, no. 1, pp. 35–53, 2004.

[59] A. Ben-Tal and A. Nemirovski, “Robust solutions of uncertain linear programs,” *Operations research letters*, vol. 25, no. 1, pp. 1–13, 1999.

[60] G. Calafiore and M. C. Campi, “Uncertain convex programs: randomized solutions and confidence levels,” *Mathematical Programming*, 2005.

[61] M. C. Campi and S. Garatti, *Introduction to the Scenario Approach*. SIAM, 2018.

[62] M. C. Campi and S. Garatti, “The exact feasibility of randomized solutions of uncertain convex programs,” *SIAM Journal on Optimization*, 2008.

[63] G. C. Calafiore, “Random convex programs,” *SIAM Journal on Optimization*, 2010.

[64] K. Margellos, P. Goulart, and J. Lygeros, “On the road between robust optimization and the scenario approach for chance constrained optimization problems,” *IEEE Transactions on Automatic Control*, 2014.

[65] X. Zhang, S. Grammatico, G. Schildbach, P. Goulart, and J. Lygeros, “On the sample size of random convex programs with structured dependence on the uncertainty,” *Automatica*, vol. 60, pp. 182–188, 2015.

[66] G. Schildbach, L. Fagiano, and M. Morari, “Randomized solutions to convex programs with multiple chance constraints,” *SIAM Journal on Optimization*, vol. 23, no. 4, pp. 2479–2501, 2013.

[67] F. Ramponi, "Consistency of the Scenario Approach," *SIAM Journal on Optimization*, vol. 28, no. 1, pp. 135–162, 2018.

[68] M. Campi and S. Garatti, "Wait-and-judge scenario optimization," *Mathematical Programming*, pp. 1–35, 2016.

[69] M. C. Campi, S. Garatti, and F. A. Ramponi, "A general scenario theory for non-convex optimization and decision making," *IEEE Transactions on Automatic Control*, 2018.

[70] M. C. Campi and S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality," *Journal of Optimization Theory and Applications*, vol. 148, no. 2, pp. 257–280, 2011.

[71] B. Pagnoncelli, S. Ahmed, and A. Shapiro, "Sample average approximation method for chance constrained programming: theory and applications," *Journal of optimization theory and applications*, 2009.

[72] M. W. Tanner and L. Ntaimo, "IIS branch-and-cut for joint chance-constrained stochastic programs and application to optimal vaccine allocation," *European Journal of Operational Research*, 2010.

[73] S. Küçükyavuz, "On mixing sets arising in chance-constrained programming," *Mathematical programming*, vol. 132, no. 1-2, pp. 31–56, 2012.

[74] B. Zeng, Y. An, and L. Kuznia, "Chance constrained mixed integer program: Bilinear and linear formulations, and Benders decomposition," *Mathematical Programming*, 2017.

[75] B. Zeng and Y. An, "Solving bilevel mixed integer program by reformulations and decomposition," *Optimization online*, pp. 1–34, 2014.

[76] S. Ahmed, "Relaxations and Approximations of Chance Constraints," July 2018.

[77] R. T. Rockafellar, *Convex analysis*. Princeton university press, 2015.

[78] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of operations research*, vol. 23, no. 4, pp. 769–805, 1998.

[79] A. Ben-Tal, A. Nemirovski, and C. Roos, "Robust solutions of uncertain quadratic and conic-quadratic problems," *SIAM Journal on Optimization*, vol. 13, no. 2, pp. 535–560, 2002.

[80] D. Bertsimas and M. Sim, "Tractable approximations to robust conic optimization problems," *Mathematical Programming*, vol. 107, no. 1-2, pp. 5–36, 2006.

[81] H. Föllmer and A. Schied, *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.

[82] X. Geng and L. Xie, "Data-driven Decision Making with Probabilistic Guarantees (Part I): A Schematic Overview of Chance-constrained Optimization," *arXiv:1903.10621*, 2019.

[83] X. Chen, M. Sim, and P. Sun, "A robust optimization perspective on stochastic programming," *Operations Research*, vol. 55, no. 6, pp. 1058–1071, 2007.

[84] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *In Proceedings of the CACSD Conference*, (Taipei, Taiwan), 2004.

[85] J. Löfberg, "Automatic robust convex programming," *Optimization methods and software*, vol. 27, no. 1, pp. 115–129, 2012.

[86] I. I. CPLEX, "V12. 1: User's Manual for CPLEX," *International Business Machines Corporation*, vol. 46, no. 53, p. 157, 2009.

[87] I. Gurobi Optimization, *Gurobi Optimizer Reference Manual*. 2016.

[88] A. Mosek, *The MOSEK optimization toolbox for MATLAB manual*. Version, 2015.

[89] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[90] L. Xie, P. M. Carvalho, L. A. Ferreira, J. Liu, B. H. Krogh, N. Popli, and M. D. Ilic, "Wind integration in power systems: Operational challenges and possible solutions," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 214–232, 2011.

[91] A. P. Sanghvi, I. H. Shavel, and R. M. Spann, "Strategic planning for power system reliability and vulnerability: an optimization model for resource planning under uncertainty," *IEEE Transactions on Power Apparatus and Systems*, no. 6, pp. 1420–1429, 1982.

[92] J. l. López, K. Ponnambalam, and V. H. Quintana, "Generation and transmission expansion under risk using stochastic programming," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1369–1378, 2007.

[93] M. Mazadi, W. Rosehart, O. Malik, and J. Aguado, "Modified chance-constrained optimization applied to the generation expansion problem," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1635–1636, 2009.

[94] M. Manickavasagam, M. F. Anjos, and W. D. Rosehart, "Sensitivity-based chance-constrained generation expansion planning," *Electric Power Systems Research*, vol. 127, pp. 32–40, 2015.

[95] X. Ding, W.-J. Lee, W. Jianxue, and L. Liu, "Studies on stochastic unit commitment formulation with flexible generating units," *Electric Power Systems Research*, 2010.

[96] D. Pozo and J. Contreras, "A chance-constrained unit commitment with an $n-k$ security criterion and significant wind generation," *IEEE Trans. Power Syst*, 2013.

[97] H. Wu, M. Shahidehpour, Z. Li, and W. Tian, "Chance-constrained day-ahead scheduling in stochastic power system operation," *IEEE Trans. Power Syst*, 2014.

[98] R. Bent, D. Bienstock, and M. Chertkov, "Synchronization-aware and algorithm-efficient chance constrained optimal power flow," in *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*, pp. 1–11, IEEE, 2013.

[99] D. Bienstock, M. Chertkov, and S. Harnett, "Robust modeling of probabilistic uncertainty in smart grids: Data ambiguous chance constrained optimum power flow," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pp. 4335–4340, IEEE, 2013.

[100] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4742–4751, 2013.

[101] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," in *PowerTech (POWERTECH), 2013 IEEE Grenoble*, pp. 1–6, IEEE, 2013.

[102] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.

[103] B. Li and J. L. Mathieu, "Analytical reformulation of chance-constrained optimal power flow with uncertain load control," in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–6, IEEE, 2015.

[104] L. Roald, S. Misra, M. Chertkov, and G. Andersson, "Optimal power flow with weighted chance constraints and general policies for generation control," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*, pp. 6927–6933, IEEE, 2015.

[105] Y. Zhang, S. Shen, and J. L. Mathieu, "Data-driven optimization approaches for optimal power flow with uncertain reserves from load control," in *American Control Conference (ACC), 2015*, pp. 3013–3018, IEEE, 2015.

[106] M. Lubin, Y. Dvorkin, and S. Backhaus, "A robust approach to chance constrained optimal power flow with renewable generation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3840–3849, 2016.

[107] M. Doostizadeh, F. Aminifar, H. Ghasemi, and H. Lesani, "Energy and reserve scheduling under wind power uncertainty: An adjustable interval approach," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2943–2952, 2016.

[108] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1626–1637, 2017.

[109] L. Roald, S. Misra, A. Morrison, and G. Andersson, "Optimized risk limits for stochastic optimal power flow," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, pp. 4476–4483, IEEE, 2017.

[110] Z. Wang, C. Shen, F. Liu, X. Wu, C.-C. Liu, and F. Gao, "Chance-constrained economic dispatch with non-Gaussian correlated wind power uncertainty," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4880–4893, 2017.

[111] M. Vrakopoulou, B. Li, and J. L. Mathieu, "Chance Constrained Reserve Scheduling Using Uncertain Controllable Loads Part I: Formulation and Scenario-based Analysis," *IEEE Transactions on Smart Grid*, 2019.

[112] B. Li, M. Vrakopoulou, and J. L. Mathieu, "Chance Constrained Reserve Scheduling Using Uncertain Controllable Loads Part II: Analytical Reformulation," *IEEE Transactions on Smart Grid*, 2019.

[113] J. López, D. Pozo, J. Contreras, and J. Mantovani, "A convex chance-constrained model for reactive power planning," *International Journal of Electrical Power & Energy Systems*, vol. 71, pp. 403–411, 2015.

[114] J. F. Franco, M. J. Rider, and R. Romero, "Robust multi-stage substation expansion planning considering stochastic demand," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2125–2134, 2016.

[115] K. Margellos, V. Rostampour, M. Vrakopoulou, M. Prandini, G. Andersson, and J. Lygeros, "Stochastic unit commitment and reserve scheduling: A tractable formulation with probabilistic certificates," in *Control Conference (ECC), 2013 European*, IEEE, 2013.

[116] X. Geng, L. Xie, and L. Xie, "Security-constrained Unit Commitment with Probabilistic Guarantees," *IEEE Transactions on Control of Network Systems (in preparation)*, 2019.

[117] M. A. Bucher, M. Vrakopoulou, and G. Andersson, "Probabilistic N- 1 security assessment incorporating dynamic line ratings," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, pp. 1–5, IEEE, 2013.

[118] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, “A probabilistic framework for reserve scheduling and security assessment of systems with high wind power penetration,” *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3885–3896, 2013.

[119] M. Vrakopoulou, M. Katsampani, K. Margellos, J. Lygeros, and G. Andersson, “Probabilistic security-constrained ac optimal power flow,” in *PowerTech (POWERTECH)*, IEEE, 2013.

[120] L. Roald, M. Vrakopoulou, F. Oldewurtel, and G. Andersson, “Risk-constrained optimal power flow with probabilistic guarantees,” in *Power Systems Computation Conference (PSCC), 2014*, pp. 1–7, IEEE, 2014.

[121] L. Roald, M. Vrakopoulou, F. Oldewurtel, and G. Andersson, “Risk-based optimal power flow with probabilistic guarantees,” *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 66–74, 2015.

[122] H. Ming, L. Xie, M. Campi, S. Garatti, and P. Kumar, “Scenario-based Economic Dispatch with Uncertain Demand Response,” *IEEE Transactions on Smart Grid*, 2017.

[123] M. S. Modarresi, L. Xie, M. Campi, S. Garatti, A. Caré, A. Thatte, and P. Kumar, “Scenario-based Economic Dispatch with Tunable Risk Levels in High-renewable Power Systems,” *IEEE Transactions on Power Systems*, 2018.

[124] X. Geng and L. Xie, “Data-driven Decision Making with Probabilistic Guarantees (Part II): Applications of Chance-constrained Optimization in Power Systems,” *arXiv*, 2019.

[125] P. Yang and A. Nehorai, “Joint optimization of hybrid energy storage and generation capacity with renewable energy,” *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1566–1574, 2014.

[126] K. Hreinsson, M. Vrakopoulou, and G. Andersson, “Stochastic security constrained unit commitment and non-spinning reserve allocation with performance guarantees,” *International Journal of Electrical Power & Energy Systems*, 2015.

[127] Y. Zhang, J. Wang, Y. Li, and X. Cao, "Chance-constrained Transmission Expansion Planning with Guaranteed Wind Power Utilization," (Chicago, IL), 2017.

[128] Q. Wang, J. Wang, and Y. Guan, "Price-based unit commitment with wind power utilization constraints," *IEEE Trans. Power Syst*, 2013.

[129] C. Zhao, Q. Wang, J. Wang, and Y. Guan, "Expected value and chance constrained stochastic unit commitment ensuring wind power utilization," *IEEE Trans. Power Syst*, 2014.

[130] W.-S. Tan and M. Shaaban, "A hybrid stochastic/deterministic unit commitment based on projected disjunctive milp reformulation," *IEEE Trans. Power Syst*, 2016.

[131] A. Bagheri, C. Zhao, and Y. Guo, "Data-driven chance-constrained stochastic unit commitment under wind power uncertainty," in *Power & Energy Society General Meeting, 2017 IEEE*, IEEE, 2017.

[132] Y. Zhang, J. Wang, B. Zeng, and Z. Hu, "Chance-Constrained Two-Stage Unit Commitment under Uncertain Load and Wind Power Output Using Bilinear Benders Decomposition," *IEEE Trans on Power Systems*, 2017.

[133] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Trans. Power Syst*, 2012.

[134] Y. Zhang and G. B. Giannakis, "Robust optimal power flow with wind integration using conditional value-at-risk," in *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pp. 654–659, IEEE, 2013.

[135] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on convex approximations of chance constraints," in *Power Systems Computation Conference (PSCC)*, IEEE, 2014.

[136] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on conditional value at risk and distributional robustness," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 116–125, 2015.

[137] N. Yang and F. Wen, “A chance constrained programming approach to transmission system expansion planning,” *Electric Power Systems Research*, vol. 75, no. 2-3, pp. 171–177, 2005.

[138] J. Qiu, Z. Y. Dong, J. Zhao, Y. Xu, F. Luo, and J. Yang, “A risk-based approach to multi-stage probabilistic transmission network planning,” *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4867–4876, 2016.

[139] G. Martínez and L. Anderson, “A risk-averse optimization model for unit commitment problems,” in *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, IEEE, 2015.

[140] Z. Wu, P. Zeng, X.-P. Zhang, and Q. Zhou, “A solution to the chance-constrained two-stage stochastic program for unit commitment with wind energy integration,” *IEEE Trans. Power Syst*, 2016.

[141] D. Ke, C. Chung, and Y. Sun, “A novel probabilistic optimal power flow model with uncertain wind power generation described by customized Gaussian mixture model,” *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 200–212, 2016.

[142] T. Mühlpfordt, T. Faulwasser, L. Roald, and V. Hagenmeyer, “Solving optimal power flow with non-Gaussian uncertainties via polynomial chaos expansion,” in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, pp. 4490–4496, IEEE, 2017.

[143] S. Takriti, J. R. Birge, and E. Long, “A stochastic model for the unit commitment problem,” *IEEE Trans. Power Syst*, 1996.

[144] L. Wu, M. Shahidehpour, and T. Li, “Stochastic Security-Constrained Unit Commitment,” *IEEE Transactions on Power Systems*, 2007.

[145] Q. P. Zheng, J. Wang, and A. L. Liu, “Stochastic optimization for unit commitment—A review,” *IEEE Trans. Power Syst*, 2015.

[146] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, “Adaptive robust optimization for the security constrained unit commitment problem,” *IEEE Trans. Power Syst*, 2013.

[147] T. Alamo, R. Tempo, and A. Luque, "On the sample complexity of randomized approaches to the analysis and design under uncertainty," in *American Control Conference (ACC), 2010*, IEEE, 2010.

[148] X. Geng and L. Xie, "Chance-constrained Unit Commitment via the Scenario Approach," in *Proceedings of the 51st North American Power Symposium*, 2019.

[149] Illinois Institute of Technology, "IEEE 118-bus, 54-unit, 24-hour system," 2004.

[150] I. Peña, C. B. Martinez-Anido, and B.-M. Hodge, "An extended IEEE 118-bus test system with high renewable penetration," *IEEE Trans. Power Syst*, 2018.

[151] P. B. Eriksen, T. Ackermann, H. Abildgaard, P. Smith, W. Winter, and J. R. Garcia, "System operation with high wind penetration," *IEEE Power and Energy Magazine*, vol. 3, no. 6, pp. 65–74, 2005.

[152] F. Dong, B. H. Chowdhury, M. L. Crow, and L. Acar, "Improving voltage stability by reactive power reserve management," *IEEE transactions on Power Systems*, vol. 20, no. 1, pp. 338–345, 2005.

[153] W. D. Rosehart, C. A. Canizares, and V. H. Quintana, "Multiobjective optimal power flows to evaluate voltage security costs in power networks," *IEEE Transactions on power systems*, 2003.

[154] X. Wang, G. Ejebe, J. Tong, and J. Waight, "Preventive/corrective control for voltage stability using direct interior point method," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 878–883, 1998.

[155] A. Rabiee and M. Parniani, "Voltage security constrained multi-period optimal reactive power flow using benders and optimality condition decompositions," *IEEE Transactions on Power Systems*, 2013.

[156] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, and others, "Definition and classification of power system

stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, 2004.

[157] T. V. Menezes, L. C. da Silva, and V. F. da Costa, "Dynamic VAR sources scheduling for improving voltage stability margin," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 969–971, 2003.

[158] B. Gao, G. Morison, and P. Kundur, "Voltage stability evaluation using modal analysis," *IEEE transactions on power systems*, vol. 7, no. 4, pp. 1529–1542, 1992.

[159] O. Obadina and G. Berg, "Var planning for power system security," *IEEE Trans on Power Systems*, 1989.

[160] S. S. Sharif and J. H. Taylor, "Dynamic optimal reactive power flow," in *American Control Conference, 1998. Proceedings of the 1998*, vol. 6, pp. 3410–3414, IEEE, 1998.

[161] The Electricity Reliability Council of Texas (ERCOT), "ERCOT System Operating Limit Methodology for Planning and Operations Horizon," tech. rep.

[162] C. J. Mozina, M. Reichard, and others, "Coordination of generator protection with generator excitation control and generator capability," in *PPIC 2008.*, pp. 62–76, IEEE, 2008.

[163] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*, vol. 7. McGraw-hill New York, 1994.

[164] J. W. Lamont and J. Fu, "Cost analysis of reactive power support," *IEEE transactions on Power Systems*, vol. 14, no. 3, pp. 890–898, 1999.

[165] A. G. Expósito, J. M. Ramos, J. R. Macias, and Y. C. Salinas, "Sensitivity-based reactive power control for voltage profile improvement," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 937–945, 1993.

[166] Z. Yang, H. Zhong, Q. Xia, and C. Kang, "A novel network model for optimal power flow with reactive power and network losses," *Electric Power Systems Research*, vol. 144, pp. 63–71, 2017.

[167] Z. Yang, A. Bose, H. Zhong, N. Zhang, Q. Xia, and C. Kang, "Optimal Reactive Power Dispatch with Accurately Modeled Discrete Control Devices: A Successive Linear Approximation Approach," *IEEE Transactions on Power Systems*, 2016.

[168] J. Qiu and S. Shahidehpour, "A new approach for minimizing power losses and improving voltage profile," *IEEE transactions on power systems*, vol. 2, no. 2, pp. 287–295, 1987.

[169] N. Deeb and S. Shahidehpour, "Linear reactive power optimization in a large power network using the decomposition approach," *IEEE Transactions on power systems*, vol. 5, no. 2, pp. 428–438, 1990.

[170] K. Abdul-Rahman and S. Shahidehpour, "A fuzzy-based optimal reactive power control," *IEEE Transactions on power systems*, vol. 8, no. 2, pp. 662–670, 1993.

[171] C. Grigg, P. Wong, R. Billinton, and others, "The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee," *IEEE Transactions on power systems*, vol. 14, no. 3, pp. 1010–1020, 1999.

[172] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst*, 2011.

[173] Y.-j. Zhang and Z. Ren, "Optimal reactive power dispatch considering costs of adjusting the control devices," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1349–1356, 2005.

[174] B.-M. Hodge and M. Milligan, "Wind power forecasting error distributions over multiple timescales," in *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–8, IEEE, 2011.

[175] A. Prékopa, *Stochastic programming*. Springer, 2013.

[176] D. R. Morgan, J. W. Eheart, and A. J. Valocchi, "Aquifer remediation design under uncertainty using a new chance constrained programming technique," *Water resources research*, vol. 29, no. 3, pp. 551–561, 1993.

[177] M. Vrakopoulou, J. L. Mathieu, and G. Andersson, "Stochastic optimal power flow with uncertain reserves from demand response," in *Hawaii International Conference on System Sciences (HICSS),*, 2014.

[178] M. S. Modarresi, L. Xie, and C. Singh, "Reserves from Controllable Swimming Pool Pumps: Reliability Assessment and Operational Planning," in *Hawaii International Conference on System Sciences (HICSS)*, 2018.

[179] R. Fernández-Blanco, Y. Dvorkin, and M. A. Ortega-Vazquez, "Probabilistic security-constrained unit commitment with generation and transmission contingencies," *IEEE Trans. Power Syst*, 2017.

[180] Z. Hu, X. Wang, and G. Taylor, "Stochastic optimal reactive power dispatch: Formulation and solution method," *International Journal of Electrical Power & Energy Systems*, 2010.

[181] Y. Zhang, S. Shen, and J. Mathieu, "Distributionally Robust Chance-Constrained Optimal Power Flow with Uncertain Renewables and Uncertain Reserves Provided by Loads," *IEEE Transactions on Power Systems*, 2017.

[182] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A Probabilistic Framework for Reserve Scheduling and N - 1 Security Assessment of Systems With High Wind Power Penetration," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3885–3896, 2013.

[183] L. Roald and G. Andersson, "Chance-Constrained AC Optimal Power Flow: Reformulations and Efficient Algorithms," *IEEE Transactions on Power Systems*, 2018.

[184] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex Relaxations of Chance Constrained AC Optimal Power Flow," *arXiv preprint arXiv:1702.08372*, 2017.

[185] E. D. Anese, K. Baker, and T. Summers, "Chance-constrained AC optimal power flow for distribution systems with renewables," *IEEE Transactions on Power Systems*, 2017.

[186] B. C. Lesieutre, D. K. Molzahn, A. R. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *Communication,*

Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on, pp. 1492–1499, IEEE, 2011.

[187] H. Xu, C. Caramanis, and S. Mannor, "Robustness and regularization of support vector machines," *Journal of Machine Learning Research*, vol. 10, no. Jul, pp. 1485–1510, 2009.

[188] H. Xu, C. Caramanis, and S. Mannor, "Robust regression and lasso," *IEEE Transactions on Information Theory*, 2010.

[189] H. Xu and S. Mannor, "Robustness and generalization," *Machine learning*, vol. 86, no. 3, pp. 391–423, 2012.

[190] D. Bertsimas and M. S. Copenhaver, "Characterization of the equivalence of robustification and regularization in linear and matrix regression," *European Journal of Operational Research*, vol. 270, no. 3, pp. 931–942, 2018.

[191] W. Xie and X. Deng, "The CCP selector: Scalable algorithms for sparse ridge regression from chance-constrained programming," *arXiv preprint arXiv:1806.03756*, 2018.

[192] H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Constructing Convex Inner Approximations of Steady-State Security Regions," *IEEE Transactions on Power Systems*, 2018.

[193] D. Lee, H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Convex Restriction of Power Flow Feasibility Set," *arXiv preprint arXiv:1803.00818*, 2018.

[194] B. Cui and X. A. Sun, "Solvability of Power Flow Equations Through Existence and Uniqueness of Complex Fixed Point," *arXiv preprint arXiv:1904.08855*, 2019.

[195] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, 2012.

[196] S. H. Low, "Convex relaxation of optimal power flow—Part I: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

[197] S. H. Low, "Convex relaxation of optimal power flow—Part II: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, 2014.

APPENDIX A

ALGORITHMS AND PROOFS

A.1 Algorithms

Algorithm 2 Find the Support Set $\mathcal{S}$ of SP($\mathcal{N}$).

```
Compute x*_N by solving SP(N).
Set S ← ∅.
for i ∈ N do
    Solve the scenario problem SP_{N−i} and compute x*_{N−i}.
    if cᵀx*_{N−i} < cᵀx*_N then
        S ← S + i.
    end if
end for
```

Algorithm 3 Find an Irreducible Set $\mathcal{I}$ of $\text{SP}_{\mathbb{A}}(\mathcal{N})$.

```
Compute opx_A(N) by solving SP_A(N). Set I ← N.
for i ∈ N do
    Compute opx_A(I − i) by solving SP(I − i).
    if opt_A(I − i) = opt_A(N) then
        I ← I − i.
    end if
end for
```

Algorithm 4 For the two-stage scenario problem (3.11).

```
Solve SP_A(N) and compute the solution (x*, y*).
Fix y = y*, find support scenarios S of the second-stage problem (3.12), e.g. using Algorithm 1.
if opt_A(S) = opt_A(N) then
    SP_A(N) is non-degenerate and S is the essential set.
else
    SP_A(N) is degenerate, the best we can find is an irreducible set, e.g. using Algorithm 3.
end if
```

A.2 Proofs

Proof of Lemma 2. Monotonicity in β is obvious. To prove 2), we show that $\ln(\frac{1-\epsilon(N,k,\beta)}{1-\epsilon(N,k+1,\beta)}) \geq 0$ for fixed values of (N, β). For simplicity, we use $\epsilon(k)$ to represent $\epsilon(N, k, \beta)$.

$$(N-k-1)\ln(\frac{1-\epsilon(k)}{1-\epsilon(k+1)}) = \frac{1}{N-k}\ln(\frac{N\binom{N}{k}}{\beta}) + \ln(\frac{N-k}{k})$$

Clearly, $\ln(\frac{1-\epsilon(k)}{1-\epsilon(k+1)}) \geq 0$ if $N \geq 2k$. We now show that it also holds for the case of $N \leq 2k$.

$$(N-k-1)\ln(\frac{1-\epsilon(k)}{1-\epsilon(k+1)}) = \frac{1}{N-k}\ln(\frac{N\binom{N}{k}}{\beta(\frac{k}{N-k})^{N-k}}) = \frac{1}{N-k}\ln(\frac{N}{\beta}\frac{\binom{N}{N-k}}{(\frac{N}{N-k}-1)^{N-k}})$$
$$\geq \frac{1}{N-k}\ln(\frac{(\frac{N}{N-k})^{N-k}}{(\frac{N}{N-k}-1)^{N-k}}) = \ln(\frac{N}{k}) \geq 0$$

The last line uses the well-known lower bound on binomial coefficients $\binom{N}{k} \geq (\frac{N}{k})^k$ and the fact that $\beta \in (0, 1)$ and $1 \leq k \leq N$.

Similarly, we prove 3) by showing $\ln(\frac{1-\epsilon(N+1,k,\beta)}{1-\epsilon(N,k,\beta)}) \geq 0$ for fixed values of (k, β). It is easy to verify this is true for the cases $N = k$ and $N = k+1$. The remainder of the proof shows that this is also true for the case $N > k+1$. For simplicity, we show that $(N-k+1)(N-k)\ln(\frac{1-\epsilon(N+1,k,\beta)}{1-\epsilon(N,k,\beta)}) \geq$

0.

$$\begin{aligned}&(N-k+1)(N-k)\ln(\frac{1-\epsilon(N+1,k,\beta)}{1-\epsilon(N,k,\beta)})\\&=(N-k)\ln(\frac{\beta}{(N+1)\binom{N+1}{k}})-(N-k+1)\ln(\frac{\beta}{N\binom{N}{k}})\\&=\ln(\frac{1}{\beta})+\ln(N)+(N-k)\ln(\frac{N(N-k+1)}{(N+1)^2})+\ln(\binom{N}{k})\end{aligned}$$

We notice that $\ln(N)$, $\ln(\binom{N}{k})$ and $\ln(\frac{N(N-k+1)}{(N+1)^2}) = \ln((1-\frac{1}{N+1})(1-\frac{k}{N+1}))$ are monotonically increasing with N, therefore $(N-k+1)(N-k)\ln(\frac{1-\epsilon(N+1,k,\beta)}{1-\epsilon(N,k,\beta)}) \geq \ln(\frac{k}{\beta}) > 0$, i.e. $\epsilon(N+1,k,\beta) \leq \epsilon(N,k,\beta)$. □

Proof of Lemma 3 [63]. For the purpose of contradiction, we assume that there is a scenario $s \in \mathcal{S}$ but $s \notin \mathcal{I}$. According to the definition of support scenarios, $\text{opt}_{\mathbb{A}}(\mathcal{N}-s) < \text{opt}_{\mathbb{A}}(\mathcal{N})$. However, Assumption 12 claims that removing scenarios will not increase the optimal objective value and $\mathcal{I} \subseteq \mathcal{N}-s$, we have $\text{opt}_{\mathbb{A}}(\mathcal{N}-s) \geq \text{opt}_{\mathbb{A}}(\mathcal{I}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$, which causes a contradiction. □

Proof of Lemma 5. We first write out the Lagrange dual function $D(\mu,\lambda)$ of SP($\mathcal{N}$):

$$D(\mu,\lambda) = \inf_x \Big(c^\intercal x + \sum_{\iota=1}^{N}(\mu^\iota)^\intercal f(x,\xi^\iota) + \lambda^\intercal g(x)\Big) \tag{A.1}$$

The Lagrange dual problem is $\max_{\mu,\lambda} D(\mu,\lambda),$ s.t. $\mu \geq 0, \lambda \geq 0$. By assumption, we know that SP($\mathcal{N}$) has a strictly feasible solution, thus Slater's condition holds and $D(\mu^*_{\mathcal{N}},\lambda^*_{\mathcal{N}}) = c^\intercal x^*_{\mathcal{N}}$ by strong duality. We then consider the Lagrange dual problem of SP($\mathcal{N}-i$). The dual solution to SP($\mathcal{N}-i$) is denoted by $\lambda^*_{\mathcal{N}-i}$ and $\mu^*_{\mathcal{N}-i} = \{\mu^{1,*}_{\mathcal{N}-i}\cdots,\mu^{i-1,*}_{\mathcal{N}-i},\mu^{i+1,*}_{\mathcal{N}-i},\cdots,\mu^{N,*}_{\mathcal{N}-i}\}$.

If ξ^i is *not* a support scenario, then $c^\intercal x^*_{\mathcal{N}} = c^\intercal x^*_{\mathcal{N}-i}$, thus $D(\mu^*_{\mathcal{N}},\lambda^*_{\mathcal{N}}) = c^\intercal x^*_{\mathcal{N}} = c^\intercal x^*_{\mathcal{N}-i} = D(\mu^*_{\mathcal{N}-i},\lambda^*_{\mathcal{N}-i})$ by Slater's condition and strong duality.

We could assign $\mu^{\iota,*}_{\mathcal{N}} = \mu^{\iota,*}_{\mathcal{N}-i}$ for $\iota \neq i$ and let $\mu^{i,*}_{\mathcal{N}} = 0$. Clearly this is one optimal solution to the dual problem of SP($\mathcal{N}$). The uniqueness of this solution is due to the non-degeneracy of SP($\mathcal{N}$) by assumption. Thus $\|\mu^*_{\mathcal{N},i}\| = 0$. □

Proof of Lemma 6. We first prove (1). The case that $\hat{\mathcal{S}} = \emptyset$ is obvious. For the case that $\hat{\mathcal{S}}$ contains at least one scenario $s \in \hat{\mathcal{S}}$. Solving the 2nd stage problem with s removed gives a different optimal solution $\hat{x}$ with $c_x^\intercal \hat{x} < c_x^\intercal x^*$. Clearly $(\hat{x}, y^*)$ is a feasible solution to SP($\mathcal{N} - s$), with

$$c_y^\intercal y^* + c_x^\intercal \hat{x} < c_y^\intercal y^* + c_x^\intercal x^* \tag{A.2}$$

therefore s is a support scenario for SP($\mathcal{N}$) and $\hat{\mathcal{S}} \subseteq \mathcal{S}$.

We then prove (2). By Assumption 12, we know that $\text{opt}_{\mathbb{A}}(\hat{\mathcal{S}}) \leq \text{opt}_{\mathbb{A}}(\mathcal{S}) \leq \text{opt}_{\mathbb{A}}(\mathcal{N})$ since $\hat{\mathcal{S}} \subseteq \mathcal{S} \subseteq \mathcal{N}$. If $\hat{\mathcal{S}}$ is invariant, i.e. $\text{opt}_{\mathbb{A}}(\hat{\mathcal{S}}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$, then $\text{opt}_{\mathbb{A}}(\mathcal{N}) \leq \text{opt}_{\mathbb{A}}(\hat{\mathcal{S}}) \leq \text{opt}_{\mathbb{A}}(\mathcal{S}) \leq \text{opt}_{\mathbb{A}}(\mathcal{N})$ gives $\text{opt}_{\mathbb{A}}(\mathcal{S}) = \text{opt}_{\mathbb{A}}(\mathcal{N})$, therefore SP($\mathcal{N}$) is non-degenerate. □

Proof of Corollary 3. We first prove (1), that is SP($\mathcal{N}$) has a unique essential set if it is non-degenerate (similar with the proof of Lemma 2.11 in [63])). From Lemma 3, an essential set can be written as $\mathcal{E} = \mathcal{S} \cup \mathcal{Y}$ where $\mathcal{Y} \subseteq (\mathcal{N} - \mathcal{S})$. The support set $\mathcal{S}$ is invariant because of the non-degeneracy of SP($\mathcal{N}$) by assumption. Since $\mathcal{E}$ is the invariant set of minimal cardinality, we can let $\mathcal{Y} = \emptyset$ and $\mathcal{S}$ is the essential set. The support set $\mathcal{S}$ is unique by definition, this implies the uniqueness of the essential set $\mathcal{E}$ for non-degenerate SP($\mathcal{N}$).

We then prove (2). Lemma 3 shows that $\mathcal{S} \subseteq \mathcal{R}$, we only need to show $\mathcal{R} \subseteq \mathcal{S}$ when $\text{SP}_{\mathbb{A}}(\mathcal{N})$ is non-degenerate. For the purpose of contradiction, we assume there exists $s \in \mathcal{R}$ but $s \notin \mathcal{S}$. By hypothesis ($s \notin \mathcal{S}$), we have $\mathcal{S} \subseteq \mathcal{R} - s$ (Lemma 3). The monotonicity assumption 12 gives $\text{opt}_{\mathbb{A}}(\mathcal{S}) \leq \text{opt}_{\mathbb{A}}(\mathcal{R} - s)$. Since $\mathcal{R}$ is irreducible, we have $\text{opt}_{\mathbb{A}}(\mathcal{R} - s) < \text{opt}_{\mathbb{A}}(\mathcal{R})$. $\text{SP}_{\mathbb{A}}(\mathcal{N})$ is non-degenerate and $\mathcal{R}$ is invariant gives $\text{opt}_{\mathbb{A}}(\mathcal{R}) = \text{opt}_{\mathbb{A}}(\mathcal{N}) = \text{opt}_{\mathbb{A}}(\mathcal{S})$. Combining the results above, we have

$$\text{opt}_{\mathbb{A}}(\mathcal{S}) \leq \text{opt}_{\mathbb{A}}(\mathcal{R} - s) < \text{opt}_{\mathbb{A}}(\mathcal{R}) = \text{opt}_{\mathbb{A}}(\mathcal{N}) = \text{opt}_{\mathbb{A}}(\mathcal{S}), \tag{A.3}$$

which is clearly a contradiction. Therefore $\mathcal{S} = \mathcal{R}$. □

Lemma 7. *Consider a (possibly non-convex) scenario problem $SP_{\mathbb{A}}(\mathcal{N})$ and an algorithm $\mathbb{A}$ sat-*

isfying Assumption 12. Suppose k is not a support scenario for SP$_{\mathbb{A}}(\mathcal{N})$, then

$$\mathcal{S}(\mathcal{N}) \subseteq \mathcal{S}(\mathcal{N} - k) \tag{A.4}$$

Note: SP$(\mathcal{N} - k)$ could have more support scenarios than SP$(\mathcal{N})$.

Proof of Lemma 7. $k \notin \mathcal{S}$ and $s \in \mathcal{S}$ give opt$(\mathcal{N} - k)$ = opt$(\mathcal{N})$ and opt$(\mathcal{N} - s) <$ opt$(\mathcal{N})$, respectively. Assumption 12 shows opt$(\mathcal{N} - k - s) \leq$ opt$(\mathcal{N} - s)$. Hence, it holds that

$$\begin{aligned} \text{opt}(\mathcal{N} - k - s) \leq \text{opt}(\mathcal{N} - s) < \text{opt}(\mathcal{N}) = \text{opt}(\mathcal{N} - k), \\ \forall s \in \mathcal{S}(\mathcal{N}), \end{aligned} \tag{A.5}$$

then s is a support scenario for SP$(\mathcal{N} - k)$. Therefore $\mathcal{S}(\mathcal{N}) \subseteq \mathcal{S}(\mathcal{N} - k)$. □

Proof of Corollary 4. (1) ⇒ (2) is proved in Corollary 3. And (2) ⇒ (3) is obvious, since the essential set $\mathcal{E}$ is irreducible. If there is only one irreducible set, then it is the essential set.

Lastly, we prove (3) ⇒ (1). We prove SP$(\mathcal{N})$ being degenerate implies the essential set is not unique (equivalent with the statement that SP$(\mathcal{N})$ is non-degenerate if it has a unique essential set). Suppose SP$(\mathcal{N})$ is degenerate, i.e. opt$(\mathcal{S}) <$ opt$(\mathcal{N})$. Consider an essential set $\mathcal{E} = \mathcal{S} \cup \mathcal{T}$ (Lemma 3), where $\mathcal{T}$ is non-empty and $k \in \mathcal{T}$. Consider the scenario problem SP$(\mathcal{N} - k)$, and opt$(\mathcal{N} - k)$ = opt$(\mathcal{N})$ because $k \notin \mathcal{S}$. We also know that $\mathcal{S}$ is contained in any essential set of SP$(\mathcal{N} - k)$ by Lemma 7, i.e. $\mathcal{E}(\mathcal{N} - k) = \mathcal{S} \cup \hat{\mathcal{T}}$. And $\hat{T}$ has to be non-empty [1]. Then opt$(\mathcal{S} \cup \hat{\mathcal{T}})$ = opt$(\mathcal{N} - k)$ = opt$(\mathcal{N})$, therefore $\mathcal{S} \cup \hat{\mathcal{T}}$ must contain at least one essential set that is different from $\mathcal{S} \cup \mathcal{T}$ (because $k \in \mathcal{T}$ and $k \notin \hat{\mathcal{T}}$). Therefore SP$(\mathcal{N})$ has more than one essential set when it is degenerate. □

[1] Otherwise opt$(\mathcal{S})$ = opt$(\mathcal{E}(\mathcal{N} - k))$ = opt$(\mathcal{N} - k)$ = opt$(\mathcal{N})$, which contradicts with the hypothesis that SP$(\mathcal{N})$ is degenerate.

APPENDIX B

CHANCE-CONSTRAINED UNIT COMMITMENT[1]

The deterministic Unit Commitment formulation utilizes the expected wind generation and load forecast, it does not take the uncertainties from wind and load into consideration. We propose an improved formulation of d-UC using chance constraints, which guarantee the system security with a tunable level of risk ϵ with respect to uncertainties.

$$\begin{aligned} \min_{z,u,v,g,r} \quad & (3.13a) \\ \text{s.t.} \quad & (3.13b)(3.13d)(3.13e)(3.13f)(3.13g)(3.13h)(3.13i)(3.13j)(3.13k) \\ & \mathbb{P}_{\tilde{w}\times\tilde{d}}\Big(\mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal(\hat{w}^t + \tilde{w}^t) \geq \mathbf{1}^\intercal(\hat{d}^t + \tilde{d}^t), \\ & \qquad k \in [0, n_k], t \in [1, n_t]\Big) \geq 1 - \epsilon \end{aligned} \tag{B.1a}$$

Problem (B.1) is the formulation of chance-constrained Unit Commitment (c-UC). Instead of using expected load $\hat{d}^t$ as in (3.13), we consider loads d^t as forecast $\hat{d}^t$ plus a random forecast error $\tilde{d}^t$ (i.e. $d^t = \hat{d}^t + \tilde{d}^t$).

Comparing with d-UC, the only difference of c-UC is the addition of the chance constraint (B.1a). The chance constraint guarantees there will be enough supply to meet the net demand in any contingency case at any time

$$\mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal(\hat{w}^t + \tilde{w}^t) \geq \mathbf{1}^\intercal(\hat{d}^t + \tilde{d}^t), k \in [0, n_k], t \in [1, n_t] \tag{B.2}$$

with probability no less than $1 - \epsilon$.

To reveal the structures of c-UC, we define the sets below:

$$\mathcal{B} := \big\{(z, u, v) : (3.13h), (3.13i), (3.13j), (3.13k)\big\} \tag{B.3a}$$

[1] Reprinted with permission from [148]

$$\mathcal{C} := \big\{(g,r) : (3.13b), (3.13d), (3.13e)\big\} \tag{B.3b}$$

$$\mathcal{H} := \big\{(z,g,r) : (3.13f), (3.13g)\big\} \tag{B.3c}$$

$$\mathcal{U} := \big\{(g) : (B.2)\big\} \tag{B.3d}$$

Then c-UC can be succinctly represented as:

$$\min_{z,u,v,g,r} \ (3.13a)$$

$$\text{s.t. } (z,u,v) \in \mathcal{B} \tag{B.4a}$$

$$(g,r) \in \mathcal{C} \tag{B.4b}$$

$$(z,g,r) \in \mathcal{H} \tag{B.4c}$$

$$\mathbb{P}\big(g \in \mathcal{U}\big) \geq 1 - \epsilon \tag{B.4d}$$

Sets $\mathcal{B}$ and $\mathcal{C}$ stand for the *deterministic* constraints for binary and continuous variables, respectively. Set $\mathcal{H}$ represents the hybrid constraints related with both continuous and binary variables. Set $\mathcal{U}$ represents all constraints related with uncertainties.

Remark 13. The non-convexity of unit commitment comes from binary variables (z, u, v). Clearly as shown in (B.4), non-convexity (i.e. set $\mathcal{B}$ and $\mathcal{H}$) only exists in deterministic constraints, and uncertain constraints $\mathcal{U}$ are only related with continuous variables. This observation plays a critical role in analyzing the structural properties of s-UC in Lemma 8 and Corollary 5.

B.1 Solving c-UC via the Scenario Approach

B.1.1 Scenario-based Unit Commitment

As explained in Chapter 3, the scenario approach reformulates (3.2) to a scenario problem (3.4) using N scenarios. For the unit commitment problem, we denote the set of N scenarios as $\mathcal{N} = \{(\tilde{d}^1, \tilde{w}^1), (\tilde{d}^2, \tilde{w}^2), \cdots, (\tilde{d}^N, \tilde{w}^N)\}$. Each load and wind scenario is a time series of length n_t: $\tilde{d}^i = (\tilde{d}^{1,i}, \cdots, \tilde{d}^{n_t,i})$, $\tilde{w}^i = (\tilde{w}^{1,i}, \cdots, \tilde{w}^{n_t,i})$. Then we define the set $\mathcal{U}_i$ corresponding to scenario i:

$$\mathcal{U}_i := \big\{g : \mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal(\hat{w}^t + \tilde{w}^{t,i}) \\ \geq \mathbf{1}^\intercal(\hat{d}^t + \tilde{d}^{t,i}), t \in [1, n_t], k \in [0, n_k]\big\} \tag{B.5}$$

The scenario problem for c-UC can be written as

$$\begin{aligned} \min_{z,u,v,g,r} \ & (3.13a) \\ \text{s.t.} \ & (B.4a), (B.4b), (B.4c) \\ & g \in \cap_{i=1}^N \mathcal{U}_i \end{aligned} \tag{B.6a}$$

Problem (B.6) is referred as s-UC in the remainder of this paper.

B.1.2 Structural Properties of s-UC

For notation simplicity, we define ι^t as the index of the scenario with the largest *net* demand forecast error at time t:

$$\iota^t := \arg_i \max \Big\{\mathbf{1}^\intercal\tilde{d}^{t,1} - \mathbf{1}^\intercal\tilde{w}^{t,1}, \cdots, \mathbf{1}^\intercal\tilde{d}^{t,N} - \mathbf{1}^\intercal\tilde{w}^{t,N}\Big\}, \tag{B.7}$$

and define $\overline{\mathcal{S}} := \{\iota^1, \iota^2, \cdots, \iota^{n_t}\}$. Clearly there might be repetitive scenario indices in $\iota^1, \iota^2, \cdots, \iota^{n_t}$, i.e. $|\overline{\mathcal{S}}| \leq n_t$.

Lemma 8. *When $n_t = 1$, s-UC has at most one support scenario. The support scenario is the one with the largest* net *demand forecast error, i.e. $(\tilde{d}^{1,\iota^1}, \tilde{w}^{1,\iota^1})$ if the number of support scenarios is not zero.*

Proof. Let ι^1 be the scenario index defined in (B.7), clearly $\mathcal{U}_{\iota^1} = \cap_{i=1}^N \mathcal{U}_i$, which implies that the removal of any scenario other than ι^1 will not change the feasible region. According to Definition 16, all other scenarios except ι^1 cannot be a support scenario. Therefore s-UC with $n_t = 1$ has at most one support scenario. □

Corollary 5. *For s-UC (B.6), let $\mathcal{S}$ denote the set of its support scenarios, then $\mathcal{S} \subseteq \overline{\mathcal{S}}$, which indicates $|\mathcal{S}| \leq |\overline{\mathcal{S}}| \leq n_t$.*

Proof. Let $\mathcal{U}_i^t = \left\{ g^t : \mathbf{1}^\intercal g^{t,k} + \mathbf{1}^\intercal(\hat{w}^t + \tilde{w}^{t,i}) \geq \mathbf{1}^\intercal(\hat{d}^t + \tilde{d}^{t,i}), k \in [0, n_k] \right\}$, then $\mathcal{U}_i = \mathcal{U}_i^1 \times \mathcal{U}_i^2 \cdots \times \mathcal{U}_i^{n_t}$. According to Lemma 8, $\cap_{i=1}^N \mathcal{U}_i^t$ has at most one support scenario, which is indexed by ι^t. Applying Lemma 8 for all n_t snapshots, we can see that the set $\overline{\mathcal{S}}$ contains all candidates to support scenarios, thus $\mathcal{S} \subseteq \overline{\mathcal{S}}$ and $|\mathcal{S}| \leq |\overline{\mathcal{S}}| \leq n_t$. □

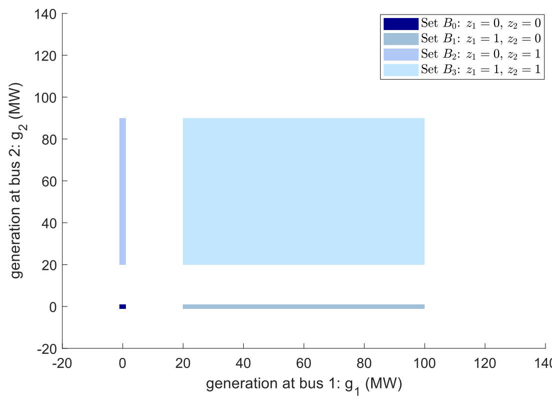
The intuition behind Lemma 8 and Corollary 5 is illustrated in Fig. B.1. Fig. B.1 visualizes the constraints and feasible region (g_1, g_2) of the 2-generator, 3-bus and 3-line system in [2]. Four blue regions ($\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$) stand for four possible on/off states of 2 generators. For example, $\mathcal{B}_2$ shows the case in which generator 1 is off ($z_1 = 0$) and generator 2 is on ($z_2 = 1$). The black solid lines represent the determine constraints $\mathcal{C}$. Three dashed/dotted lines denote three constraints ($\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3$) of three scenarios. Since the scenario constraint (B.5) is only about supply and demand (transmission limits are not included), the feasible region of s-UC is clearly defined by the scenario with the *largest net demand* ($\mathcal{U}_1$ in Fig. B.1), which is the support scenario of s-UC. The scenario with the largest net demand at each snapshot is a candidate for support scenarios (Lemma 8), therefore there are at most n_t candidates for support scenarios (Corollary 5).

B.1.3 Sample Complexity for s-UC

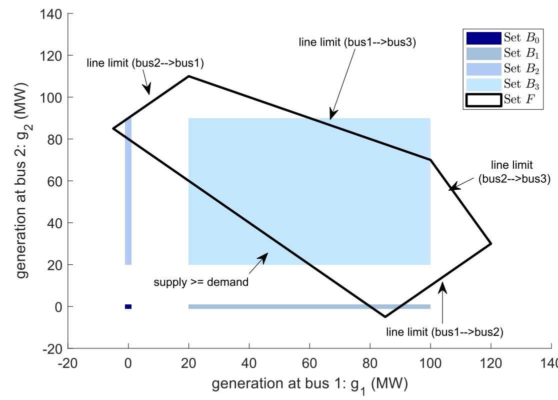
Corollary 5 shows that $|\mathcal{S}| \leq n_t$ for s-UC, then we can use the results in Corollary 1 to calculate the number of scenarios to achieve the desired security level $1-\epsilon$ with confidence $1-\beta$. Table B.1 presents the sample complexity (number of scenarios) needed with various ϵ levels for the 118-bus system in Section B.2.1.

Although unit commitment is non-convex because of the binary variables (z, u, v). It is in general difficult to estimate the number of support scenarios $|\mathcal{S}|$ a-priori. Without exploiting the structural properties of s-UC as in Corollary 5, the best bound [2] might be the number of decision variables $|\mathcal{S}| \leq n$, which is $4n_g n_t + n_g n_t n_k = 75168$ for the 118-bus system . Table B.1 also

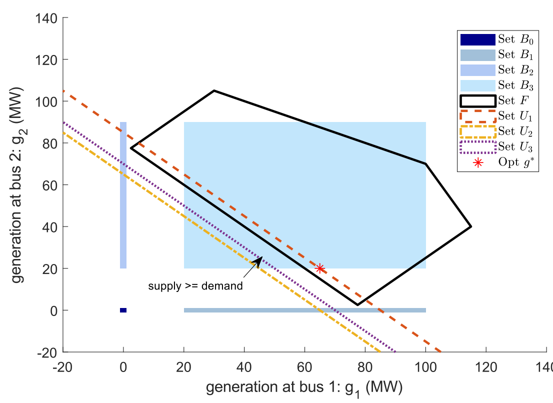
[2]Before revealing the structure of s-UC in Corollary 5, n is not an upper bound on $|\mathcal{S}|$ because s-UC is non-convex. But n is the best bound we could hope for using the results in Theorem 5 and Corollary 1.



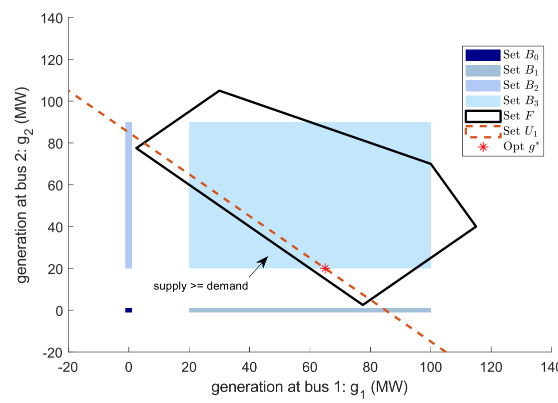
(a) Constraints $\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ representing 4 possible states of 2 generators.



(b) Deterministic constraints.



(c) Scenario constraints $\mathcal{U}$, with all scenarios ($\mathcal{U}_1$,$\mathcal{U}_2$,$\mathcal{U}_3$).



(d) Scenario constraints $\mathcal{U}$, with only support scenario ($\mathcal{U}_1$).

Figure B.1: Illustration of Lemma 8 and Corollary 5 using the 2-generator, 3-bus system in [2].

presents the sample complexity using $|\mathcal{S}| \leq 75168$. As shown in Table B.1, Corollary 5 *greatly* reduces the number of scenarios from some astronomical numbers in the case of $|\mathcal{S}| \leq 75168$. Another attractive observation is that the results in Corollary 5 holds regardless of the system size.

Table B.1: Sample Complexity for s-UC (Case118, $\beta = 10^{-4}$), adapted with permission from [148].

Violation probability ϵ	0.3	0.2	0.1	0.075	0.05	0.025	0.01
Sample Complexity N (when $\|\mathcal{S}\| \leq 24$)	143	221	455	610	921	1853	4650
Sample Complexity N (when $\|\mathcal{S}\| \leq 75168$)	253416	380419	761394	1015370	1523320	3047161	7618678

All settings of the 3-bus system can found in Table B.2.

Table B.2: Settings of the 3-bus System, reprinted with permission from [2].

Line Data					Generator Data				
Line No.	From Bus	To bus	Reactance (p.u.)	Capacity (MW)	Gen No.	Bus	Min	Max	Marginal Cost
1	1	2	1.0	20	1	1	20	100	1
2	1	3	1.0	100	2	2	20	90	100
3	2	3	1.0	100					
Load Data (MW)					Wind Data (MW)				
Bus	Forecast	Error 1	Error 2	Error 3	Bus	Forecast	Error 1	Error 2	Error 3
3	110	11	-30	-35	2	30	6	-15	-25

B.2 Case Study

B.2.1 Settings of the 118-bus System

We solve the unit commitment problem of an 118-bus system with 54 generators ($n_g = 54$) in 24 hours ($n_t = 24$) under 54 possible generator failure contingencies ($n_k = 54$). The test system is a modified version of the 118-bus system in [150]. The modified 118-bus system includes 5 wind farms at different locations.

The numerical simulation was conducted on a desktop with Intel Core i7-2600 CPU@3.40GHz and 16GB of memory. Matpower and YALMIP were used to formulate the c-UC problem in MatlabR2018a. The c-UC problem was converted to s-UC via ConvertChanceConstraint in [1], then solved using Gurobi 8.10 till the MIP gap is smaller than 0.01%.

B.2.2 Numerical Results

We solve the s-UC problem with different number of scenarios N. Given N, we conduct 10 independent Monte-Carlo simulations to examine the randomness of the scenario approach. Another independent test dataset of 10^4 points was used to evaluate the out-of-sample violation probability ϵ of the solution to s-UC.

Figure B.2 demonstrates the optimal objective values and out-of-sample ϵ with different number of scenarios. As the scenario approach theory suggests, with an increasing number of scenarios, the system risk level ϵ decreases. Figure B.2 also shows that with 0.96% of cost increase (from 1.356×10^6 to 1.369×10^6), the system risk ϵ is reduced from 19% to 2%.

Figure B.3 plots two violation probabilities. The blue solid curve illustrates the average empirical ϵ (evaluated on the test dataset of 10^4 points), the shaded area shows the largest and smallest violation probabilities in 10 Monte-Carlo runs. The dotted green lines plots the guaranteed ϵ by combining Theorem 5 with Corollary 5. Figure B.3 shows that the scenario approach is applicable on the unit commitment problem, despite its non-convexity. Furthermore, Figure B.3 also demonstrates the value of Corollary 5. Without showing that $|\mathcal{S}| \leq n_t$ as in Corollary 5, Theorem 5 is only able to provide useless guarantees (e.g. $\epsilon \leq 0.999999$ when using 1000 scenarios).

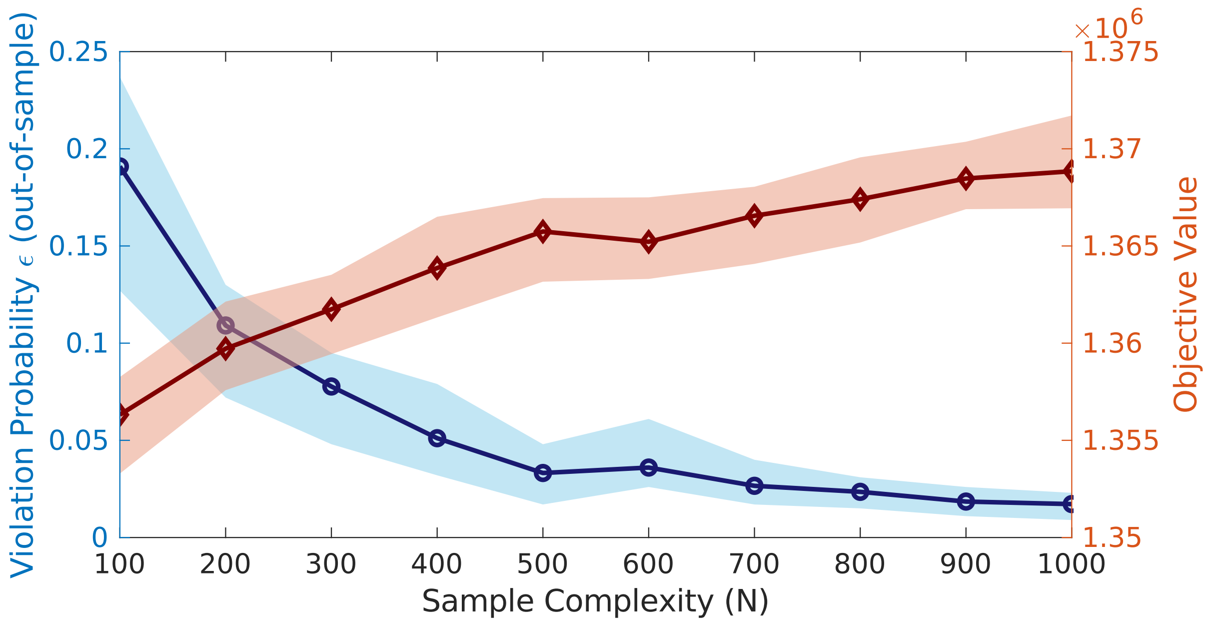


Figure B.2: Key Results of s-UC with Different Sample Complexity, reprinted with permission from [148].

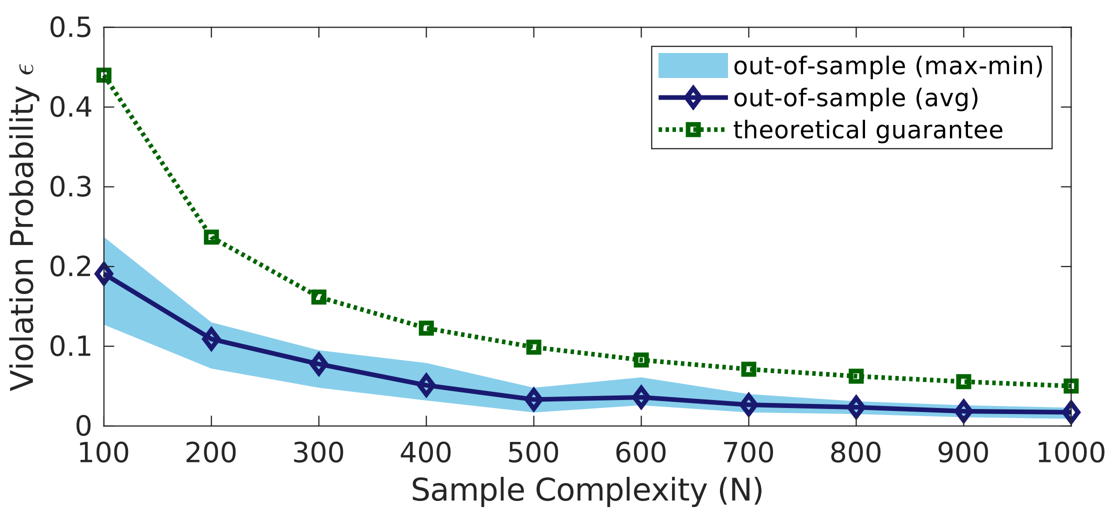


Figure B.3: Theoretical and Empirical Violation Probabilities ϵ, reprinted with permission from [148].

Table B.3: Number of Support Scenarios, reprinted with permission from [148].

N	100	200	300	400	500	600	700	800	900	1000
$\lvert\mathcal{S}\rvert$ (min)	19	21	22	22	22	22	23	23	23	24
$\lvert\mathcal{S}\rvert$ (max)	24	24	24	24	24	24	24	24	24	24

Due to the non-convexity from the binary decision variables, the scenario approach was considered not applicable on the unit commitment problem previously. One main contribution of this paper is to show the potential of the scenario approach on non-convex problems like unit commitment. By exploring the structural properties of s-UC, Section B.1 shows that the scenario approach could still provide rigorous guarantees on the quality of solutions, as in the convex case. This is all based on Lemma 8 and Corollary 5. Table B.3 shows the maximum and minimum number of support scenarios in 10 Monte Carlo runs of each given sample complexity N. This verifies the correctness of Corollary 5.

B.2.3 Scenario Reduction

When the desired risk level ϵ is very small, the scenario approach might require a large number of scenarios. This will directly cause memory and computation issues in numerical simulations. Corollary 5 turns out to be quite helpful in improving the computational performance. Corollary 5 shows that a majority of the scenarios have no impacts on the final solution and thus can be reduced. Then s-UC only needs to be solved with at most $n_t = 24$ scenarios, which can be easily identified as mentioned in Section B.1.2. We compare the results of using 1000 scenarios with those of using identified 24 (out of 1000) scenarios. Although the optimal solution is slightly different due to a few identical generators, the difference in the objective value is less than 10^{-6}.

B.2.4 Adding Security Constraints

The main limitation of this paper is not considering possible security constraints such as transmission line limits. The nice results in Corollary 5 holds only in the absence of a transmission network. We also applied the scenario approach on chance-constrained SCUC. Numerical results show that the number of support scenarios could be more than $n_t = 24$, but this number does not increase too much (e.g. $30 \sim 50$ for the 118-bus system with 186 lines). However, we are yet not able to prove nice results as in Corollary 5. This is one critical part of our ongoing works and beyond the scope of this paper.

B.3 Summary

This paper is a first step towards a practical and rigorous day-ahead decision making framework in uncertain environments. We formulate the chance-constrained unit commitment problem and solve it via the scenario approach. We show that the number of support scenarios in the unit commitment problem is at most n_t. This structural property makes the scenario approach applicable in the presence of non-convexity. It substantially reduces the necessary number of scenarios and could be further exploited to reduce the computational requirement to solve the problem. Future work will extend the results towards security-constrained unit commitment.