OPTIMIZATION UNDER UNCERTAINTY MODELS IN POWER SYSTEMS OPERATIONS

By

SADRA BABAEI

Bachelor of Science in Industrial Engineering
Iran University of Science and Technology
Tehran, Iran
2009

Master of Science in Industrial Engineering
Amirkabir University of Technology
Tehran, Iran
2012

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2019

OPTIMIZATION UNDER UNCERTAINTY MODELS IN POWER SYSTEMS OPERATIONS

Dissertation Approved:

Dr. Chaoyue Zhao

Dissertation Advisor

Dr. Austin Buchanan

Dr. Farzad Yousefian

Dr. Yuanxiong Guo

Dedicated to my

beloved parents, Usef and Soheila,

beloved wife, Elaheh,

and beloved son, Adrian

The dedication reflects the views of the author and is not endorsed by committee members or Oklahoma State University.

ACKNOWLEDGMENTS

Foremost, I would like to express my most sincere gratitude to my advisor Dr. Chaoyue Zhao, who has been both a friend and mentor to me during my Ph.D studies. I have benefited greatly from her expertise, careful guidance, continuous encouragement, and untiring support. She is one of the smartest people I know. I could not have imagined having a better advisor during this journey.

I am grateful to my doctoral committee members Drs. Austin Buchanan, Farzad Yousefian, and Yuanxiong Guo for their constructive comments and suggestions and guide on me. I also deeply appreciate their flexibility to support my compact timeline. In addition, I would like to thank my collaborators Drs. Ruiwei Jiang and Lei Fan for their helpful discussions, advices, and suggestions.

My sincere thanks must also go to all faculties at IEM department, especially Profs. Sunderesh Heragu and Balabhaskar Balasundaram for their valuable suggestions and kind supports during my Ph.D and their assistance in my job search process. I would like to acknowledge all IEM administrative staff, especially Laura Brown, for their effort and help in patiently answering my questions.

I sincerely appreciate the financial support by grants from the National Science Foundation (ECCS-1610935 and CMMI-1662589).

Many thanks to my amazing friends and colleagues at Oklahoma State University, who have been encouraging me over these wonderful years in Stillwater. There is no enough space

The acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

to name all of you! You guys know who you are.

I would like to dedicate this work to my parents Usef Babaei and Soheila Moradian for their emotionally and financially supports, unconditional trust, and endless patience throughout my life. I hope that I have made you proud. I also highly appreciate my brother, Edris, for his encouragement.

Finally, and above everything, I give a special feeling of gratitude to my lovely wife, Elaheh, who makes my life worthwhile. This work cannot be done without her extremely support and countless sacrifices. My acknowledgement Section would be incomplete without thanking to my four-month-old son, Adrian, who is my inspiration to be stronger. I love you to the moon and back.

The acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: SADRA BABAEI

Date of Degree: July, 2019

Title of Study: OPTIMIZATION UNDER UNCERTAINTY MODELS IN POWER SYSTEMS OPERATIONS

Major Field: INDUSTRIAL ENGINEERING AND MANAGEMENT

Abstract: Uncertainty is a critical issue in many power system problems. While distributed energy resources (DERs) like solar panels and wind turbines are exciting energy sources in meeting the nations increasing energy demand, backing up the electricity grid in the event of outages, and peak shaving in the case of high demand charges, they also introduce new difficulties to the operation of power systems. One of the primary hurdles is the stochasticity of renewable energy generation caused by variations in day-to-day weather. If not properly addressed, this can lead to rolling blackouts and other detrimental outcomes in the grid. In addition, modern energy infrastructure is highly vulnerable to increasingly severe weather conditions. Because of inherently unpredictable weather conditions and intricacy of power systems, evaluating and mitigating the underlying risk of power system interruption are highly demanding for the system operators. This introduces new degrees of uncertainty that must be accounted for by power production facilities and system operators.

This study explores reformulations as well as approximation approaches to derive innovative decision-making under uncertainty models in the power system management. In particular, using techniques in Stochastic Programming, Robust Optimization, and Distributionally Robust Optimization, different uncertainty management schema are developed for protecting power grids from adversarial environments and accommodating renewable energy resources in the optimization of power system operations to provide resilient, reliable and cost-effective daily power generation scheduling.

More specifically, we begin with developing an incentive-based coordination mechanism between a wind energy supplier and a conventional energy supplier to hedge against the risks of electricity market price and wind power generation. Then, we address the energy management problem of a portfolio of DERs, a virtual power plant (VPP), to characterize and evaluate the standard attributes/parameters in the VPPs bid submitted to the energy market. Finally, we propose a data-driven model to assist the system operators to reduce the impacts of random component failures. In particular, a distributionally robust model is devised for designing a distribution power system to withstand the risk of disruptions imposed by natural disasters.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Nowadays optimization models and techniques play an essential role in power system management. From long-term network configuration planning to short-term hourly scheduling and operation, mathematical models and algorithms provide practical guidance for the decision maker. One of the primary challenges in most of the power system problems is the ubiquitous presence of the uncertainty. This uncertainty can be arisen from a number of sources, ranging from high integration of intermittent renewable generations to electric component failures, which potentially expose the system to safety issues and economic losses. Therefore, mitigating the uncertainty is an important research topic in order to ensure satisfactory performance of a power system. This dissertation introduces innovative mathematical models and enhanced algorithms to support decision making under uncertainty in the power system management. The rest of this Introduction Chapter will first provide the motivation behind the problems studied, and then briefly review the traditional and more recent advancements in the optimization under uncertainty field, and finally outline an overview of this dissertation.

1.1 Motivation of this dissertation

Currently most of the deregulated wholesale electricity markets in the United States employ a two-settlement mechanism including day-ahead and real-time markets for electricity transactions, plus an auxiliary service market to ensure system reliability. In the day-ahead market, energy suppliers and consumers submit their generation biddings and consumption offers, respectively, to the independent system operator (ISO), who is responsible for maintaining

the balance between the supply and demand at every moment in the time. Day-ahead market usually includes 24 hourly auctions that take place one day in advance. The market participants first submit their biddings and offers for the entire 24 hours, and then by using a market clearing procedure, the ISO clears the day-ahead market for all 24 periods to obtain the locational marginal prices (LMP) for each specific hour as the trading basis so that the supplies are compensated and demands are charged by LMPs.

Real-time market is occurred some minutes before the actual power delivery by the suppliers. This market also comprises 24 hours. However, the ISO clears the real-time market in a hour-by-hour basis. That means the market participants submit their biddings and offers for a specific time period, the LMP is settled for this period, then they submit their biddings and offers for the next period, and this alternative procedure continues for the whole 24 hours. Different markets follow different policies for the settlement of the real-time market. In general, the energy suppliers will be usually paid at the real-time price for their positive deviations (power supplies in the real-time are higher than the scheduled ones in day-ahead market), and they will be usually charged for their negative deviations (power supplies in the real-time are lower than the scheduled ones in day-ahead market).

Traditional structure of the power system only permits to the large fossil and nuclear power suppliers to participate in the market. However, the need for sustainable electricity has led to significant changes to this structures. According to the Global Wind Energy Council [1], the global cumulative installed wind power capacity has dramatically increased from 24.4 GW in the year 2001 to 539.6 GW in 2017. There is a similar pattern for the solar power, indicating that the total installations soared to 405.3 GW from 1.6 GW in the same range of years [2]. As a consequences of these advancements, many ISOs in the U.S. such as CAISO and MISO allow for renewable energy suppliers to participate in the market. While renewable resources are an exciting source of energy with many economic and environmental benefits, as a nascent technology, they also introduce new challenges to the operation of energy systems.

One of the primary difficulties is the variability of renewable energy generation caused by fluctuations in day-to-day weather. These variations make it difficult for the ISO to balance load consumption and power generation in power grids, which in turn, may trigger serious issues. The shortage of energy, on one hand, may put the security and reliability of the energy supply into jeopardy, and on the other hand, brings penalty cost to the renewable energy suppliers, which can increase the cost of the renewable energy as a result. Chapter 2 of this dissertation aims to find a remedy for a wind energy supplier to secure itself with respect to the wind power uncertainties in the operating day in order to be reliably integrated into the system operations.

Furthermore, due to the increasing penetration of Distributed Energy Resources (DERs), power system operators face significant challenges of ensuring the effective integration of DERs. Distributed energy resources (DERs) are a precious portion of the resource mix to establish reliable management of the green electricity grid. DERs can play essential roles in meeting the nation's increasing energy demand, backing up the electricity grid in the event of outages, and peak shaving in the case of high demand charges. Contrary to conventional generators that lie on the transmission systems and are dispatched by the system operator, DERs are spreading across the distribution systems and because of their small capacities, they are invisible to the system. Furthermore, due to their lack of controllability over their outputs, DERs are either excluded from participation in the wholesale energy market or they are not able to participate in a cost-effective manner [3]. With the current growing demand for electrical energy, reliability of supply has emerged as a major concern. To maintain a reliable grid, it is essential that all of the energy suppliers satisfy their commitment schedules. However, the outputs of DERs are usually subject to uncertainty, which makes them as inefficient and unreliable generation resources. Traditionally, the approach of connecting DERs in most cases is based on a so-called fit and forget regime, meaning that DERs are designed to fulfill a worst-case situation (like peak load) which may only occur limited times

during a year [4]. This conservative approach prevents an effective share of electricity supply from DERs' side. To address the above issues, the concept of virtual power plant to aggregate a collection of DREs and describe them as an equivalent of a larger power generator has been explored in recent years [5]. Basically, the VPP is a representation of a portfolio of heterogeneous DERs, including flexible loads, distributed generators (DGs) and energy storage facilities that aggregate their capacities and participate in the wholesale energy market as a single entity. Both DERs' owners and distribution system benefit from VPP. On one hand, it enables DER owners to be visible by the system operator similar to a transmission-connected generator and optimize their operation in the most cost-effective manner. On the other hand, VPP enables system operators to efficiently utilize available capacity and fulfill security standards. However, it is very challenging to evaluate some physical properties such as ramping rates and capacity limits of the VPP as one generation resource, since different DERs may have different physical constraints. In addition, submitting inaccurate parameters in the ISOs' market clear engine can threaten the grid operation. This challenge calls for new models for power system operations, which is the topic of Chapter 3 of this dissertation.

Another source of the uncertainty that can highly impact the performance of the power system is the random failure of the system components. Recently U.S. has witnessed repeated severe power outages due to natural disasters such as hurricane Sandy [6] and tropical storm Irene [7]. Only between years of 2003–2012, nearly 679 weather-related power outages happened in the U.S. and each influenced more than 50,000 customers [8]. Unfortunately, the severity and frequency of natural disasters have been trending upwards. For example, in the last ten years, the U.S. has suffered from seven of the ten most costly storms in its history [9]. The growing threat from natural disasters calls for a better planning of the power grids to improve system resiliency. According to the report [10] by the President's Council of Economic Advisers and the U.S. Department of Energy, nearly 80–90% of outages in the power system occur along the distribution systems, and often lead to interruptions of power

supply to end customers. Practically, a distribution system is operated in a radial topology so as to make the design and protection coordination as simple as possible. Despite its simplicity, any contingency in the distribution system can interrupt the continuity of power supply to all customers downstream the on-contingency area. Hence, topology designing of the distribution system is a critical task for the reliability, economic operation, and resilience of distribution systems, and this topic will be explored in the Chapter 4.

1.2 Background on optimization under uncertainty methods

Traditional deterministic decision-making models assume a perfect knowledge of the system, i.e., accurate values for the system parameters. However, such precise information is rarely available in many real-life applications, including power system problems, for a variety of reasons ranging from simply measurement errors to uncertain information about the future. Any incorrect input data can potentially yield to infeasible strategies or reveal undesirable performance when implemented. Optimization under uncertainty problems are characterized by the necessity of taking actions without fully understanding what their consequences will be in the future. However, they may offer multiple openings for taking corrective actions later to mitigate the unfavorable effects as more and more observations are unfolded. Depending on the nature of the problem and the uncertainty involved, conventional methods to deal with uncertainty can be generally classified in two streams: Stochastic programming and Robust optimization.

According to the Stochastic Programming (SP) approach, the uncertain events in a problem are captured by random variables that are governed by a "known" probability distribution. Under the standard two-stage stochastic programming paradigm, the decision variables are divided in two parts. The first stage decisions are those that are made before the realizations of the random parameters. In the second stage once the uncertain quantities are revealed, a set of recourse decisions are made in response to the observed uncertain events.

A classical two-stage stochastic model can be written as follows:

$$\min_{\mathbf{x}\in\mathcal{X}} \ \mathbf{c}^T\mathbf{x} + E_{\mathbb{P}}\ [\mathcal{Q}(\mathbf{x},\boldsymbol{\xi})], \tag{1.1}$$

where $\mathcal{X}$ is the feasible region of the first stage decisions, and $\mathcal{Q}(\mathbf{x},\boldsymbol{\xi})$ is the optimal value of the second stage problem:

$$\mathcal{Q}(\mathbf{x},\boldsymbol{\xi}) = \min_{\mathbf{y}} \ \mathbf{q}^T\mathbf{y} \tag{1.2}$$

$$\text{s.t.} \ \ \mathbf{W}\mathbf{y} = \mathbf{h} - \mathbf{T}\mathbf{x} \tag{1.3}$$

$$\mathbf{y} \geq 0, \tag{1.4}$$

where $\boldsymbol{\xi} = (\mathbf{q}, \mathbf{T}, \mathbf{W}, \mathbf{h})$ represents the random information. The objective of a SP model is to determine the optimal first stage decisions in such a way that the total first stage cost and the second stage expected cost is minimized.

Having the theory of probability and stochastic processes as its strong pillars, stochastic programming has solid mathematical foundations, which proves its flexibility and usefulness in a variety of applications. Despite its widespread applicabilities, there are still major challenges associated with the SP framework. One essential criticism casting on such programs is the assumption of perfectly knowing the distribution of random parameters. Although it is possible to drive a distribution that best fits the historical samples, in order to approximate the true distribution of data, the solutions can be misleading and away from true optimal solutions. Even worse, in some circumstances, the historical data are either unavailable or untrustworthy. Robust optimization (RO) has emerged as an attractive optimization framework to deal with problems with uncertain parameters. The basic idea behind the RO is to define a so-called uncertainty region for the uncertain parameters and then solve the problem in such a way that the optimal solution remains feasible for all possible realizations

of uncertain parameters vary within the prescribed uncertainty set. A typical two-stage adaptive robust optimization problem is in the following form:

$$\min_{\mathbf{x}\in\mathcal{X}} \max_{\xi\in\mathcal{W}} \mathbf{c}^T\mathbf{x} + \mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}) \tag{1.5}$$

where $\mathcal{W}$ is the uncertainty set of $\boldsymbol{\xi}$. An important feature of the above formulation is that the second stage decisions can be adjusted depending on the realization of $\boldsymbol{\xi}$, whereas in static robust models the decisions are already determined before the realization of $\boldsymbol{\xi}$.

The main advantages of the RO approach are that it requires only minimal information to construct the deterministic uncertainty set $\mathcal{W}$, and it guarantees that the obtained first stage decisions are feasible for most outcomes of the uncertain parameters. However, this approach faces the challenge of providing over conservative solutions as the objective function is being minimized respect to the worst-case scenario $\boldsymbol{\xi}$ in the uncertainty set $\mathcal{W}$, which might rarely happen in practice.

In recent years the field of optimization under uncertainty has undergone major advances for leveraging observations of random variables as direct inputs to the mathematical programming problems. As a newly emerged framework, Distributionally Robust Optimization (DRO) aims at facilitating this goal. More specifically, in the DRO model, instead of considering the randomness of uncertain parameters, an unknown probability distribution is considered and characterized by learning from the available historical data. In other words, compatible with the decision maker's prior information, the DRO models consider a set of probability distributions of the uncertain parameters (termed ambiguity set) using certain statistical inferences (e.g., moment information). Then, the strategy is to find a solution that is optimal with respect to the worst-case probability distribution within the ambiguity set. The mathematical model of the distributionally robust optimization can be described as

follows:

$$\min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbb{P}\in\mathcal{P}} \mathbf{c}^T\mathbf{x} + E_{\mathbb{P}}[\mathcal{Q}(\mathbf{x},\boldsymbol{\xi})] \tag{1.6}$$

where $\mathcal{P}$ is defined as the ambiguity set of the distribution $\mathbb{P}$.

DRO approach has several appealing benefits that make it outstanding in data-driven applications compared to other methods (like stochastic programming and Robust optimization). For instance, even though we access to a vast amount of data in data-driven environments, it is possible that we are not able to find any particular distribution to fit in the data pattern. Even if we can infer the true distribution from data, SP approach usually reveals poor out-of-sample performance. In particular, if we tailor an SP for a given dataset (with a known true distribution) and evaluate the optimal solution on a different dataset (with the same distribution), the result may not be satisfactory. This phenomenon is called optimizer's curse in the literature [11]. The main reason of optimizer's curse in SP is that SP requires a two-phase decision making process for solving the underlying problem. In the first phase, the historical data serves as an input to estimate the probability distribution of the uncertain parameter. The goal of this phase is usually to achieve a minimum prediction error. Then in the second phase, the estimated distribution serves as an input to the optimization problem that aims at finding optimal decisions. Hence, the distribution is not calibrated for the optimization problem. However, DRO approach does not decompose the decision making process into two (estimation and optimization) phases. Indeed, DRO models can implicitly find an estimated distribution that is tailored for the optimization problem. As another advantage, DRO models are often tractable (depending on the choice of ambiguity set) even if the corresponding SP is not. For example, assuming that the probability distribution is continuous, the corresponding SP may not be efficiently solvable since it requires computing multidimensional integral in the objective function, which is NP-hard in general. However,

as we show in Chapter 3, the DRO model is tractable in this case for our problem.

In addition, contrary to the classical RO that is basically a distribution-free approach and minimizes the total cost based on a "worst-case scenario", DRO models account for distributional knowledge through the ambiguity set, and minimize the total expected cost based on a "worst-case distribution" over a set of probability distributions (ambiguity set). Hence, DRO models are less conservative than RO. In addition, as more information included in the ambiguity set (e.g., the support set, mean, variance, and temporal correlations considered in our approach), the DRO models exclude more pathological distributions that are far away from the true distribution, and thus result in less conservative solutions.

1.3 Overview of this dissertation

In this dissertation, we develop different concepts in the field of optimization under uncertainty to mitigate the corresponding potential undesirable consequences. In particular, the following essential research questions are addressed:

- How to recover the energy shortage of a renewable energy supplier when participating in the electricity market?
- How to facilitate the integration of distributed energy resources into the wholesale electricity market by transforming all the information into one standard bid?
- How to configure a distribution power system to withstand against the risk of random contingencies?

The first question is explored in Chapter 2. This Chapter proposes an incentive-based coordination mechanism between a wind energy supplier (WES) and a conventional energy supplier (CES) to achieve a Pareto improvement. To comply with its day-ahead schedules and hedge against the intermittent wind energy generation, the WES is allowed to outsource

a backup power capacity from the CES via making a bilateral contract. However, unanimous agreement cannot always be achieved since each party plays on its own interest. We employ the concept of swing option contracts to further encourage the suppliers to reach an agreement of the contract. On one hand, the WES can leverage the uncertainty of wind output by covering possible energy shortage from the CES. On the other hand, the CES can optimally allocate its energy capacity by participating into the electricity market and offering capacity to backup the shortage of energy from the WES. The bidding problem for each supplier is formulated as a multi-stage stochastic programming model, with the objective of maximizing the expected profit while maintaining a low level of risk. Unlike the traditional two-stage approach, the proposed multi-stage model can effectively capture the impact of rebidding process in the real-time market. We incorporate Conditional-Value-at-Risk as a risk measure to characterize the effect of risk perception of suppliers on their bidding decisions. Meanwhile, a game theory based approach is developed to obtain the contract items between the suppliers. Implementation results on real cases are provided to illustrate the effectiveness of our proposed framework.

The second question is addressed in Chapter 3, where we develop a model to evaluate the physical characteristics of the VPP, i.e., its maximum capacity and ramping capabilities, given the uncertainty in wind power output and load consumption. The proposed model is based on a distributionally robust optimization approach that utilizes moment information (e.g., mean and covariance) of the unknown parameter. We reformulate the model as a mixed binary second-order conic program and develop a separation framework to address it. We first solve a two-stage problem and then benchmark it with a multi-stage case. Case studies are conducted to show the performance of the proposed approach.

The third question is investigated in Chapter 4, where a distributionally robust optimization (DRO) model is proposed for designing the topology of a new distribution system facing random contingencies (e.g., imposed by natural disasters). The proposed DRO model

optimally configures the network topology and integrates distributed generation to effectively meet the loads. Moreover, we take into account the uncertainty of contingency. Using the moment information of distribution line failures, we construct an ambiguity set of the contingency probability distribution, and minimize the expected amount of load shedding with regard to the worst-case distribution within the ambiguity set. As compared with a classical robust optimization model, the DRO model explicitly considers the contingency uncertainty and so provides a less conservative configuration, yielding a better out-of-sample performance. We recast the proposed model to facilitate the column-and-constraint generation algorithm. We demonstrate the out-of-sample performance of the proposed approach in numerical case studies.

CHAPTER II

INCENTIVE-BASED COORDINATION MECHANISM FOR RENEWABLE AND CONVENTIONAL ENERGY SUPPLIERS

2.1 Problem description and literature review

Renewable energy, especially wind energy, has been increasing penetration into power systems. Projections revealed by National Renewable Energy Laboratory reflect that nearly 80% of the electricity will be served by renewable resources to satisfy hourly-based demands in every region of U.S. by 2050 [12]. To promote the growth of renewable energy generation, many European countries and U.S. have established incentive policies like renewable portfolio standard in their electricity markets. As a consequence of these regulations, several U.S. Independent System Operators (ISOs), such as MISO and ERCOT, allow for the participation of renewable power suppliers, particularly Wind Energy Suppliers (WESs), in their electricity markets [13]. A WES confronts with two kinds of uncertainties when it submits energy bids to the market: price and wind power output uncertainties. If the WES is not able to deliver what it commits to the electricity market due to the wind output intermittency, it will face considerable penalty costs for its energy shortage. There are a few options for the WES to mitigate the intermittency or to cover the shortage, but they all have their own limitations.

The first option for the WES is to utilize energy storage resources such as pumped-storage units [14], batteries [15], and air compressed [16] to mitigate its energy shortage risk. However, the energy storage capacities are usually very expensive, and the high costs hinder large-volume installations. The second remedy approach is to purchase energy from

the energy market in the form of ancillary services. The main drawback of this strategy is that the system operator does not guarantee the provision of enough power to cover the WES's shortage, due to uncertainties from both the supply and the demand sides. Moreover, more deployments of ancillary services will increase the market clearing price of these services and lead to higher total costs [17]. The hybrid wind-conventional system like wind-thermal ([18, 19]) is examined as another approach. [19] proposes a wind-thermal coordinated trading mechanism for the day-ahead energy market. The problem is formulated as a two-stage stochastic optimization model to maximize the total profit when wind plants are included in the generation portfolio of strategic producers. [20] investigates the expansion planning strategy of quick start generators, like gas-fired generators with flexible minimum up/down and large ramping capabilities, to accommodate the fluctuations of wind generation. However, these generators are often associated with significant operating costs. Another solution for the WES to mitigate wind power output uncertainty is to outsource a backup capacity from a Conventional Energy Supplier (CES) ([21, 22]). For instance, Wartsila Corporation delivered 203 MW gas power plant near San Antonio, Texas to South Texas Electric Cooperative to provide backup power for their customers in 65 counties, where an increasing penetration of wind power brought challenges to the grid stability [23]. In [22], a trading strategy between WES and CES based on a two-stage stochastic programming model is introduced. However, the designed contractual terms are not flexible enough in the sense that the power delivery from CES to WES is unconditional. More particularly, because of wind uncertainty and aversion to charging imbalance penalties, WES is willing to arrange flexible orders from CES. That is, WES would prefer purchasing backup power from the market directly to executing the contract with CES if the contract price is higher than the market price. But the contract proposed in [22] obligates WES to trade the capacity level predetermined in the contract.

Existing literatures on optimal bidding strategies for the WES mostly formulate the problem as a two-stage stochastic programming model (see e.g., [24], [25], [26], [27], among

others). According to the two-stage stochastic programming model, in the first stage, the WES submits its bidding in the day-ahead (DA) market before the wind power output and the prices of DA and real-time (RT) markets become known. While in the second stage, the WES decides on its transactions in the RT market once the wind output and DA prices are available and only RT prices are unknown. The RT bidding for all hours of the operating day are made simultaneously at the beginning of the day. Hence, this two-stage approach does not allow any flexibility in real-time decisions after revealing the realization of uncertain parameters during each time period. Moreover, this approach fails to appropriately represent current practices of dependencies among successive periods of wind power outputs since it assumes that the wind output scenarios for the entire operating day are available at the beginning of the day. Recently, a two-step procedure approach is proposed in [28] for bidding strategy of a WES and the Conditional-Value-at-Risk employed to manage the risk of its profit. In the first step, the bidding strategy for the DA market is decided. Then in the second step, once the actual scheduling in the DA market is identified, the WES derives the bidding strategy in the RT market for each hour of the day separately. Meanwhile, the WES can update the scenarios as new information is observed. However, this study neglects the impacts of hourly-based RT decisions on the DA biddings.

As the preliminary study of this research, a coordination mechanism between the CES and WES is designed in [29], where the optimal bidding strategy problem is modeled using the traditional two-stage stochastic approach. To the best of our knowledge, this is the first study that formulates the optimal bidding and contract strategy for a wind energy supplier as a multi-stage stochastic programming model. In our model, the real-time decisions are made in each period on a period basis according to the revealed real-time price and available wind output in that period. In this way, the suppliers have the opportunity to update their real-time decisions as time progresses and more information about the uncertain parameters becomes available. Furthermore, the proposed approach uses an incentive-based mechanism

to compensate the WES's energy shortage. Particularly, a two-part structured bilateral contract is developed to allow Pareto improvements to both sides in response to the uncertain market changes. Our designed bilateral contract between the CES and WES is characterized by three main parameters: backup capacity level, reservation price, and execution price. Backup capacity level is the amount of energy that CES committed to deliver to WES upon its request. Reservation price is an allowance paid by WES to CES for reserving one unit of capacity. The execution price is paid by WES to CES for actually using a unit of capacity. This trading mechanism is similar to the swing option contracts ([30, 31, 32, 33]). In order to obtain the contract parameters between these two suppliers, we employ a game theory framework, which can guarantee the maximum achievable profits. The main contributions of this chapter can be listed as follows.

1. We develop a multi-stage stochastic programming to assist both suppliers to optimally submit their bids in the day-ahead and real-time markets. By capturing dependencies between electricity prices as well as wind outputs in consecutive time periods and allowing the suppliers to rebid in the real-time market, our model gives more efficient solutions than the two-stage model.

2. Using swing option contracts, our proposed framework provides flexible contracts to motivate the suppliers to reach an agreement. Meanwhile, it achieves maximum achievable profits for two parties considering operational limitations imposed on the suppliers.

3. We incorporate Conditional-Value-at-Risk (CVaR) into the proposed multi-stage model to provide a useful tool for the suppliers with various risk attitudes, from risk-neutral to risk-averse. In other words, our model generates optimal bidding strategies based on the suppliers' conservativeness levels.

The remainder of this chapter is organized as follows. In Section 2.3, we describe the

market clearing process in the two-settlement electricity market. In Section 2.4, we derive multi-stage stochastic programming models to describe the optimal bidding strategies for both CES and WES. Furthermore, we investigate the solution methodology and procedure for exploring Nash equilibrium in Section 2.5. In Section 2.6, we provide case studies and conduct computational experiments. Finally, we conclude this study in Section 2.7.

2.2 Nomenclature

A. Sets

$\mathcal{G}$	Set of generators.
$\mathcal{I}$	Set of day-ahead price realizations.
$\mathcal{K}$	Set of whole scenarios including market prices and wind outputs.
$\mathcal{N}^+(n)$	Set of child nodes of node n.
$\mathcal{N}_t$	Set of nodes in the scenario tree at time period t.
$\mathcal{T}$	Set of time periods.

B. Parameters

$F_c(.)$	Cost with respect to the generation level.
L_g	Minimum power output of generator g.
RD_g	Ramp-down rate for generator g.
RU_g	Ramp-up rate for generator g.
SD_g	Shut-down cost for generator g.
SU_g	Start-up cost for generator g.
U_g	Maximum power output of generator g.
W^{Max}	Installed capacity of the wind farm.
α	Confidence level.
γ	Risk preference parameter.
μ_k	Probability of occurrence of the whole scenario (including market

prices and wind outputs) k.

π_{tn}^{i} Probability of n^{th} realization of real-time parameters (wind output and real-time market price) in scenario tree at time period t corresponding to the day-ahead price realization i.

ρ_i Probability of scenario i occurrence for the day-ahead price.

C. Decision Variables

$\mathrm{Ct}^{\mathrm{C},DA}$ Cost of CES from DA market.

$\mathrm{Ct}^{\mathrm{C},RT}$ Cost of CES from RT market.

$\mathrm{Ct}^{\mathrm{W},DA}$ Cost of WES from DA market.

$\mathrm{Ct}^{\mathrm{W},RT}$ Cost of WES from RT market.

$\mathrm{Rv}^{\mathrm{C},DA}$ Revenue of CES from DA market.

$\mathrm{Rv}^{\mathrm{C},RT}$ Revenue of CES from RT market.

$\mathrm{Rv}^{\mathrm{W},DA}$ Revenue of WES from DA market.

$\mathrm{Rv}^{\mathrm{W},RT}$ Revenue of WES from RT market.

o_{tg} Binary variable to indicate if generator g is on at time period t.

S_t Backup capacity level at time period t.

s_{tg} Backup capacity provided by the generator g at time period t.

u_{tg} Binary variable to indicate if generator g is started up at time period t.

v_{tg} Binary variable to indicate if generator g is shut down at time period t.

w_t Unit execution price at time period t.

x_{tgn}^{i} Real-time generation amount at time period t by generator g in node n of the scenario tree corresponding to the day-ahead price realization i.

y_{tg}^{DA} Power offered by the conventional generator g in day-ahead market at time period t.

$y_{tgn}^{RT,i}$ Power offered by the conventional generator g in real-time market at time period

t in node n of the scenario tree corresponding to the day-ahead price realization i.

$z_{tn}^{C,i}$ Wind power output curtailed by the wind supplier in real-time market at time period t in node n of the scenario tree corresponding to the day-ahead price realization i.

z_{it}^{DA} Power offered by the wind supplier in day-ahead market for time period t and scenario i.

$z_{tn}^{RT,i}$ Power offered by the wind supplier in real-time market at time period t in node n of scenario tree corresponding to the day-ahead price realization i.

$z_{tn}^{S,i}$ Power purchased by the wind supplier from real-time market at time period t in node n of the scenario tree corresponding to the day-ahead price realization i.

$z_{tn}^{U,i}$ Power utilized by the wind supplier from bilateral contract at time period t in node n of scenario tree corresponding to the day-ahead price realization i.

δ_k Auxiliary variable in scenario k for linearizing of conditional-value-at-risk.

θ_t Unit reservation price at time period t.

ξ Auxiliary variable for calculating conditional-value-at-risk.

D. Random Parameters

W_{tn}^{i} Wind power production at time period t in node n of the scenario tree corresponding to the day-ahead price realization i.

β_t^i Proportion of backup capacity that is actually utilized by wind supplier at time period t corresponding to the day-ahead price realization i.

κ_{tn}^{i} Real-time market price at time period t in node n of the scenario tree corresponding to the day-ahead price realization i.

λ_{ti} Day-ahead market price at time period t and scenario i.

2.3 Market framework

In U.S. electricity markets such as MISO and CAISO, market participants can submit energy bids in both day-ahead and real-time markets [13]. The day-ahead market is a one-time bidding process, while the real-time market consists of sequentially bidding processes for each hour. Hence, knowing the price of previous hours, market participants have the opportunity to update the bid prior to the deadline of each operating hour. Market participants can submit their bids to the ISO in different modes. We investigate a paradigm in which the suppliers submit offers in the form of self-scheduling. In the self-scheduling mode [34], the supplier is responsible for its commitment and generation level for each time period and it only submits energy quantities to the ISO, instead of quantity-price bid pairs. In this study, we assume that CES and WES can establish a bilateral contract. This agreement can provide a hedging mechanism for both suppliers. That is, the CES is allowed to allocate its energy capacity by submitting bids to the day-ahead market and real-time market, and backing up the WES's energy shortage, based on the market price and the contract price. Similarly, the WES can also participate in the markets by buying or selling its energy and covering its energy shortage with the backup capacity provided by the CES. On one hand, this mechanism creates an incentive for the WES to accommodate wind output uncertainty and to avoid energy shortage penalties. On the other hand, pre-committing capacity to WES via a bilateral contract yields to a wise utilization of capacity and a recovery of underlying costs for the CES in the presence of unknown market price.

2.3.1 Uncertainty characterization

We consider three uncertainty parameters in our model: day-ahead (DA) market price, real-time (RT) market price, and wind energy production. A multi-stage scenario tree is developed to describe the possible realizations of uncertain parameters. The multi-stage

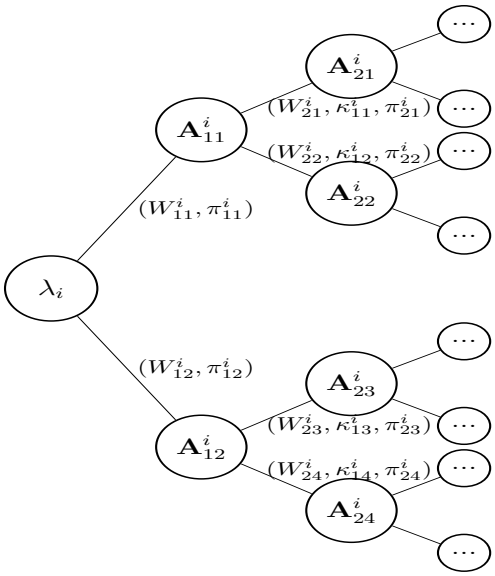

Figure 2.1: Multi-stage scenario tree for a realized day-ahead market

scenario tree can effectively capture the correlation of wind outputs as well as RT market prices among different time periods. For example, Fig. 2.1 illustrates a scenario tree for the WES. Each node, except the root node, corresponds to a real-time decision and each branch corresponds to a realization of random parameters. The root node can be interpreted as a realized DA price. For the i^{th} observation of the DA price, the wind energy for the first hour with its corresponding probability is realized at branches (W^i_{11}, π^i_{11}) and (W^i_{12}, π^i_{12}). Then, the WES decides on the amount of energy transactions with CES or market operator at nodes $\mathbf{A}^i_{11}$ and $\mathbf{A}^i_{12}$ to offset its possible energy deficit. Branches $(W^i_{21}, \kappa^i_{11}, \pi^i_{21})$, $(W^i_{22}, \kappa^i_{12}, \pi^i_{22})$, $(W^i_{23}, \kappa^i_{13}, \pi^i_{23})$, $(W^i_{24}, \kappa^i_{14}, \pi^i_{24})$ represent four samples $(W^i_{21}, W^i_{22}, W^i_{23}, W^i_{24})$ of the wind outputs in the second hour and four samples $(\kappa^i_{11}, \kappa^i_{12}, \kappa^i_{13}, \kappa^i_{14})$ of the RT prices in the first hour with their associated probabilities $(\pi^i_{21}, \pi^i_{22}, \pi^i_{23}, \pi^i_{24})$. Correspondingly, decisions for the second hour are made at nodes $\mathbf{A}^i_{21}$, $\mathbf{A}^i_{22}$, $\mathbf{A}^i_{23}$, $\mathbf{A}^i_{24}$. This procedure will continue until the last operating hour. Note here that the RT decisions corresponding to each node n are made after observing the realizations of wind outputs and RT prices along the path from the root node to the node n. Thus, the uncertainties at each period in node n are wind outputs for

the next periods as well as RT prices for the current and oncoming periods.

2.3.2 Assumptions

We make the following assumptions about the model:

- Both suppliers' strategies and payoff functions are public information. In other words, each supplier has complete knowledge about the strategies and payoffs of the other supplier, but not the decisions.
- The operation cost for the WES is negligible.
- Both WES and CES are considered to be price-takers in the day-ahead and real-time markets. This means both suppliers have no market-power in the energy markets and therefore, their offers have no impact on the market clearing price. This is a reasonable assumption since we assume both WES and CES hold small shares of generation compared to the total generation in the market.
- All bids will be accepted in the market. In other words, the suppliers can bid a low price along with their generation quantities to ensure that their bids will be accepted.

2.3.3 Decision sequence

The market timeline and decision making process in our framework can be summarized as follows:

- The WES and CES sign a bilateral contract.
- One day prior to the operating day, the WES and CES submit their bids into the DA market for all hours simultaneously.
- The ISO clears the DA market and releases the DA price to the suppliers.

- Closing to its actual energy delivery for time period t in the operating day, the WES identifies its wind output for period t.
- The suppliers submit their biddings and offerings in the RT market for period t. Additionally, the WES determines how much energy to be utilized from its reserved capacity in the contract for period t.
- The ISO clears the RT market for period t and announces the price for this hour.

Steps 2.3.3-2.3.3 are repeated until the last time period in the operating day. Fig. 2.2 illustrates the above decision sequence.

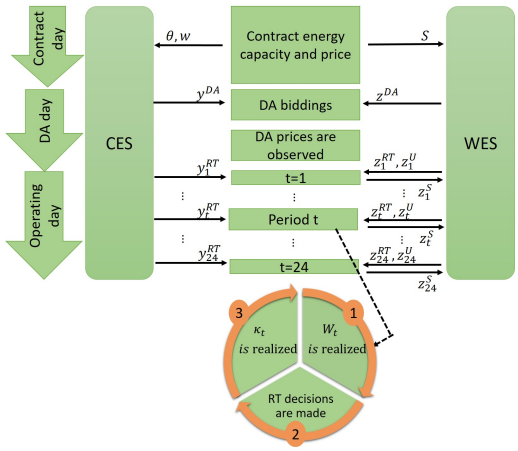

Figure 2.2: The market timeline and decision making process

2.4 Mathematical formulation

In this section, we formulate the optimal bidding problems for both conventional and wind energy suppliers, which can participate in the wholesale electricity markets and make a bilateral contract with each other. We adopt a multi-stage stochastic programming approach to model the self-scheduling process for both CES and WES in the day-ahead and real-time markets. Since stochastic programming is inherently a risk-neutral approach, we incorporate Conditional-Value-at-Risk (CVaR) as a risk measure in our model to manage the financial

risk of the suppliers. By definition, with respect to a specified confidence level α, CVaR_α is the conditional expectation of profits below the $(1-\alpha)$-percentile of the profit distribution. The $(1-\alpha)$-percentile of the profit distribution is known as Value-at-Risk (VaR), which is the largest value that guarantees the profit falls below that value only with a small probability $(1-\alpha)$. Mathematically speaking, CVaR is evaluated by the following equation:

$$\text{CVaR}_\alpha = \mathbb{E}\Big[\text{profit}\Big|\text{profit} \leq \text{VaR}_\alpha\Big] \tag{2.1}$$

For a scenario-based stochastic optimization model, CVaR can be calculated by the following linear optimization problem [35]:

$$\begin{aligned} \text{CVaR}_\alpha = \max_{\delta \geq 0, \xi} & \left\{ \xi - \frac{1}{1-\alpha} \sum_{i \in \mathcal{I}} \tau_i \delta_i \right\} \\ s.t. \;\; & \delta_i \geq \xi - X(i), \quad \forall i, \end{aligned} \tag{2.2}$$

where optimal ξ represents VaR, $X(i)$ is the i^{th} scenario of profit with its associated probability τ_i, and δ_i is an auxiliary variable indicating the difference between VaR and the scenario profit, which is positive if the scenario profit is less than VaR, and is zero otherwise. Then, we formulate the problem for both CES and WES as follows.

2.4.1 Conventional energy supplier

The CES is considered to operate and schedule a number of thermal units. For CES, DA decisions are unit reservation price, unit execution price, online or offline status of generators, start up and shut down decisions, and energy offered to the DA market. The RT decisions at different time periods corresponding to each node in the scenario tree are energy offered to the RT market and energy outputs. Given backup capacity level S provided by WES, the problem of identifying the best bidding and contract strategy for the conventional supplier

can be formulated as follows:

$$\Pi^C(S) = \max \;\; \mathbb{E}[\mathrm{Rv}^{C,DA}] + \mathbb{E}[\mathrm{Rv}^{C,RT}] -$$

$$\mathrm{Ct}^{C,DA} - \mathbb{E}[\mathrm{Ct}^{C,RT}] + \gamma_C \, \mathrm{CVaR}_\alpha \tag{2.3}$$

$$s.t. \;\; \mathbb{E}[\mathrm{Rv}^{C,DA}] = \sum_{t\in\mathcal{T}} S_t\theta_t + \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\sum_{i\in\mathcal{I}} \rho_i\lambda_{ti} y^{DA}_{tg}, \tag{2.4}$$

$$\mathbb{E}[\mathrm{Rv}^{C,RT}] = \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\sum_{i\in\mathcal{I}} \rho_i \Big[\sum_{n\in\mathcal{N}_t} \sum_{j\in\mathcal{N}^+(n)} \pi^i_{tj}\kappa^i_{tj} y^{RT,i}_{tgn} \Big] +$$

$$\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} \rho_i \beta^i_t S_t w_t, \tag{2.5}$$

$$\mathrm{Ct}^{C,DA} = \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} (SU_g u_{tg} + SD_g v_{tg}), \tag{2.6}$$

$$\mathbb{E}[\mathrm{Ct}^{C,RT}] = \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\sum_{i\in\mathcal{I}} \rho_i \Big[\sum_{n\in\mathcal{N}_t} F_c(x^i_{tgn}) \Big], \tag{2.7}$$

$$\mathrm{CVaR}_\alpha = \xi - \frac{1}{1-\alpha} \sum_{k\in\mathcal{K}} \mu_k \delta_k, \tag{2.8}$$

$$L_g o_{tg} \le y^{DA}_{tg} + \mathrm{s}_{tg} \le U_g o_{tg}, \quad \forall t, g \tag{2.9}$$

$$-o_{(t-1)g} + o_{tg} - u_{tg} \le 0, \quad \forall t, g \tag{2.10}$$

$$o_{(t-1)g} - o_{tg} - v_{tg} \le 0, \quad \forall t, g \tag{2.11}$$

$$\sum_{g\in\mathcal{G}} \mathrm{s}_{tg} = S_t, \quad \forall t \tag{2.12}$$

$$\sum_{g\in\mathcal{G}} x^i_{tgn} = \sum_{g\in\mathcal{G}} y^{DA}_{tg} + \sum_{g\in\mathcal{G}} y^{RT,i}_{tgn} + \beta^i_t S_t, \quad \forall t, i, n \tag{2.13}$$

$$L_g o_{tg} \le y^{DA}_{tg} + y^{RT,i}_{tgn} + \mathrm{s}_{tg} \le U_g o_{tg}, \quad \forall t, g, n \tag{2.14}$$

$$L_g o_{tg} \le x^i_{tgn} \le U_g o_{tg}, \quad \forall t, g, i, n \tag{2.15}$$

$$x^i_{tgn} - x^i_{(t-1)gn} \le RU_g o_{(t-1)g} +$$

$$U_g(1 - o_{(t-1)g}), \quad \forall t, g, n \tag{2.16}$$

$$x^i_{(t-1)gn} - x^i_{tgn} \le RD_g o_{tg} + U_g(1 - o_{tg}), \;\; \forall t, g, n \tag{2.17}$$

$$\xi - \Big[\mathrm{Rv}^{C,DA}_k + \mathrm{Rv}^{C,RT}_k - \mathrm{Ct}^{C,DA}_k - \mathrm{Ct}^{C,RT}_k \Big] \le \delta_k, \;\; \forall k \tag{2.18}$$

$$u_{tg}, v_{tg}, o_{tg} \in \{0, 1\},$$

$$\theta_t, w_t, y_{tg}^{DA}, y_{tgn}^{RT,i}, x_{tgn}^{i}, \delta_k \geq 0, \quad \forall t, g, n, i, k. \tag{2.19}$$

The objective for CES is to maximize the expected profit while maintaining a reasonable level of risk. The risk preference parameter γ_C allows us to make a balance between the expected profit and CVaR, and as a result, to generate different bidding strategies. When the value of γ_C is equal to zero, the CES is totally risk-neutral. That means, the CES maximizes its expected profit while it ignores the risk of profit. As the value of γ_C increases, the CES becomes more risk-averse, in the sense that it maximizes both the expected profit and CVaR. Maximizing CVaR is intended to increase the average profit of worst scenarios that encounter with very low probabilities. If the value of γ_C is large enough, the CES only maximizes CVaR to ensure that a minimum level of profit is obtained with a high probability α. The first two terms in the objective function (2.3) express the revenues of CES and the following two terms indicate the costs of CES from DA and RT markets, respectively. The expected DA revenue is calculated in (2.4), which includes DA incomes from reserving capacity for WES in the bilateral contract and DA bidding. The expected RT profit in (2.5) results from RT bidding and providing capacity to WES. We assume that from the historical data, the RT utilized capacity by WES can be estimated as a fraction (i.e., β) of the backup capacity. The DA cost in (2.6) is associated with the generators start-up/shut-down costs. The RT cost in (2.7) pertains to the generation cost, which is approximated by a m-piece piecewise linear function. Constraints (2.9), (2.14), and (2.15) enforce the generation capacity on each thermal unit. (2.10) and (2.11) represent unit start-up and shut down constraints, respectively. Constraints (2.12) ensure that the CES provides the required backup capacity to WES. Power balance constraints are expressed in (2.13). Constraints (2.16) and (2.17) impose ramping rate limits for each unit. Finally, constraints (2.18) calculate CVaR.

2.4.2 Wind energy supplier

Regarding WES, the DA decisions are backup capacity level and energy offered to the DA market, while the RT decisions associated with each node in the scenario tree are energy offered to the RT market, energy purchased from real-time market, energy utilized from the bilateral contract, and wind outputs that are curtailed. Given unit reservation price θ and execution price w provided by CES, the problem of finding the optimal bidding and contract strategy for the wind supplier can be defined as follows:

$$\Pi^W(\theta, w) = \max \ \ \mathbb{E}[\mathrm{Rv}^{\mathrm{W},DA}] + \mathbb{E}[\mathrm{Rv}^{\mathrm{W},RT}] - \mathrm{Ct}^{\mathrm{W},DA} - \mathbb{E}[\mathrm{Ct}^{\mathrm{W},RT}] + \gamma_W \ \mathrm{CVaR}_\alpha \tag{2.20}$$

$$s.t. \quad \mathbb{E}[\mathrm{Rv}^{\mathrm{W},DA}] = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} \rho_i \lambda_{ti} z_t^{DA}, \tag{2.21}$$

$$\mathbb{E}[\mathrm{Rv}^{\mathrm{W},RT}] = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} \rho_i \Bigg[\sum_{n\in\mathcal{N}_t} \sum_{j\in\mathcal{N}^+(n)} \pi_{tj}^i \kappa_{tj}^i z_{tn}^{RT,i} \Bigg], \tag{2.22}$$

$$\mathrm{Ct}^{\mathrm{W},DA} = \sum_{t\in\mathcal{T}} \theta_t S_t, \tag{2.23}$$

$$\mathbb{E}[\mathrm{Ct}^{\mathrm{W},RT}] = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} \rho_i \Bigg[\sum_{n\in\mathcal{N}_t} \sum_{j\in\mathcal{N}^+(n)} \pi_{tj}^i \kappa_{tj}^i z_{tn}^{S,i} \Bigg] + \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} \rho_i \Bigg[\sum_{n\in\mathcal{N}_t} w_t z_{tn}^{U,i} \Bigg], \tag{2.24}$$

$$\mathrm{CVaR}_\alpha = \xi - \frac{1}{1-\alpha} \sum_{k\in\mathcal{K}} \mu_k \delta_k, \tag{2.25}$$

$$0 \le z_t^{DA} \le W^{Max}, \quad \forall t \tag{2.26}$$

$$W_{tn}^i = z_t^{DA} + z_{tn}^{RT,i} + z_{tn}^{C,i} - z_{tn}^{S,i} - z_{tn}^{U,i}, \quad \forall t, n, i \tag{2.27}$$

$$0 \le z_{tn}^{U,i} \le S_t, \quad \forall t, n, i \tag{2.28}$$

$$\xi - \Big[\mathrm{Rv}_k^{\mathrm{W},DA} + \mathrm{Rv}_k^{\mathrm{W},RT} - \mathrm{Ct}_k^{\mathrm{W},DA} - \mathrm{Ct}_k^{\mathrm{W},RT} \Big] \le \delta_k, \ \forall k \tag{2.29}$$

$$S_t, z_t^{DA}, z_{tn}^{RT,i}, z_{tn}^{C,i}, z_{tn}^{U,i}, z_{tn}^{S,i}, \delta_k \ge 0, \quad \forall t, n, i, k. \tag{2.30}$$

Similar to CES, the objective for WES is to maximize the expected profit and CVaR. The first two components in (2.20) represent revenues of WES from DA and RT markets, respectively, and the following two components represent costs of WES from both markets. The DA and RT revenues in (2.21) and (2.22) come from selling energy to the corresponding markets. The DA cost (2.23) is caused by the reserved capacity from CES and the RT cost (2.24) stems from actually utilizing the reserved capacity from CES and purchasing energy from the RT market. Constraints (2.26) limit the power that WES can trade in the DA market. Constraints (2.27) indicate power balance. That is, the total realized wind output should be equal to the amount of energy offered in the DA as well as RT markets, and wind curtailment minus the amount of energy purchased from CES and RT market. Constraints (2.28) bound the amount of power transaction in the contract with CES. Finally, constraints (2.29) evaluate CVaR.

2.5 Solution methodology

Game theory is a powerful framework for analyzing strategic decision situations where the payoff of each individual decision maker relies on the decision of other decision makers. Recent publications in power market area pay more attention to game theory as it conceivably supports competition in the market ([36, 37, 38]). Market participants are always seeking to know whether they are better off by cooperation or by non-cooperation in competitive markets. A necessary condition for the non-cooperative game is that a binding commitment about price fixing and quantity fixing has to be made in such a way that all participants can benefit from it. The suppliers can play either a pure or mixed strategy game [39]. In the pure game, each supplier can choose only a particular strategy from its strategy set. Unlike the pure games, the mixed strategy games allow for choosing multiple strategies based on an assigned probability distribution. One concern associated with the mixed strategy is that it is not clear how an energy supplier would actually implement a mixed strategy. Accordingly,

this is not a good fit for our problem setting, since in this study, the suppliers should come up with one certain single contract and one plan to optimize their market participation based upon this certain contract. Considering multiple contracts each associated with a probability distribution can cause implementation issues. Therefore, we will focus on the pure strategy game in this study. Notice that the considered game is a nonzero-sum game since the sum of the suppliers objective functions is not zero (even after scaling and translation). Nash equilibrium is a concept solution used in game theory to describe an equilibrium where no participants has any incentive to unilaterally change its own strategy. To demonstrate the mathematical procedure of finding Nash equilibrium, let us assume $S_t^1, S_t^2, ..., S_t^M$ be finite discretization of the set $\mathscr{S}$ of capacity strategies and $(\theta_t^1, w_t^1), (\theta_t^2, w_t^2), ..., (\theta_t^J, w_t^J)$ be finite discretization of the set $\mathscr{P}$ of price contract terms (θ, w). Let us also assume that $\phi^C(\theta, w \mid S)$ indicates the total profit of CES corresponding to its price contract decisions (θ, w), given backup capacity S, and $\phi^W(S \mid \theta, w)$ represents the total profit of WES corresponding to its backup capacity decision S, given price contract strategy (θ, w). If (S^*, θ^*, w^*) is a Nash equilibrium, then none of the suppliers can profitably stray from the strategy (S^*, θ^*, w^*). The algorithm for finding Nash equilibrium is depicted in a flow chart in Fig. 2.3. Since both sets $\mathscr{S}$ and $\mathscr{P}$ include finite discrete elements and we delete dominated points, the algorithm will terminate in finite number of iterations. Meanwhile, if no Nash equilibrium does exist, then the suppliers do not adopt a contract.

2.6 Case study

In this section, we consider one wind plant and one thermal plant including two units. The installed capacity of the wind plant and thermal plant are 100 MW and 130 MW, respectively. We use the historical market price data during January 1-November 1, 2014 from the MISO-Michigan hub. We exclude weekend data to preclude any weekly seasonality. DA and RT price samples are generated using the procedure proposed in [40], and the wind

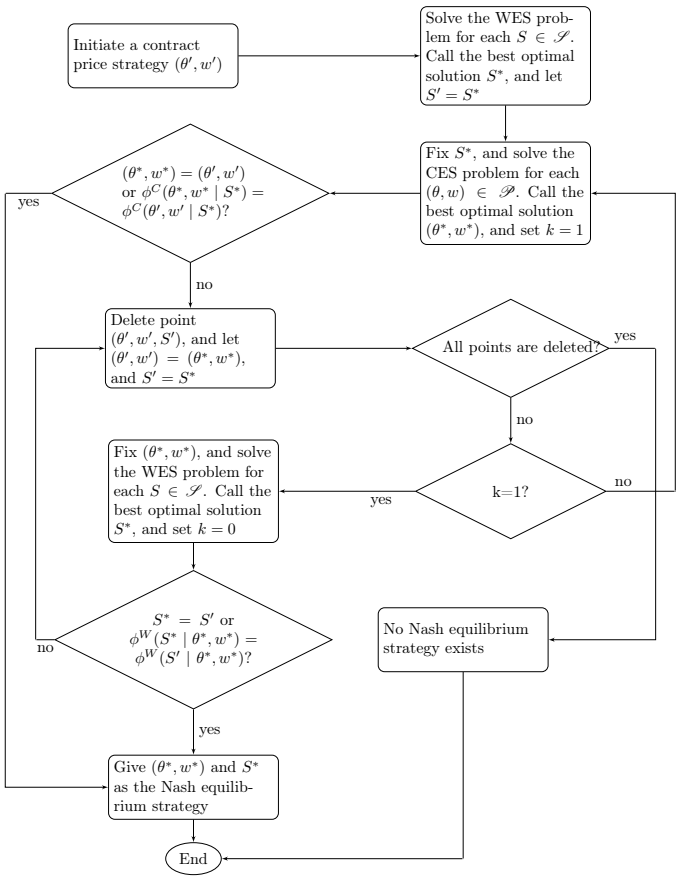

Figure 2.3: Flow chart of solution methodology

output samples are simulated using the method applied in [41]. Moreover, for multi-stage scenario tree construction and real-time scenario reduction, we employ the method proposed in [42]. We conduct several experiments with different scenario sizes. Unless state otherwise, the day-ahead and real-time sample sizes are 50, and the corresponding multi-stage real-time scenario tree includes 932 nodes. The confidence level is set at $\alpha = 0.9$ to calculate CVaR. All of the experiments are implemented in C++ and solved with CPLEX 12.6 on a computer with Intel Xeon 3.2 GHz and 8 GB memory. We concentrate more on the behavior of WES in the case studies. In the following part, we first verify the effectiveness of the proposed trading mechanism by giving several numerical examples. Second, we compare our multi-stage stochastic programming model with the two-stage stochastic model.

2.6.1 Effects of signing contract

To assess the proposed method, out-of-sample simulations are carried out in daily time steps over a 13-week horizon. The expected weekly profits of the suppliers are compared in two different scenarios as shown in Fig. 2.4. The first scenario (no contract) is that the suppliers participate in the energy market only. The second (potential contract) is that the suppliers have the option to transact with each other in addition to participating in the market. From Fig. 2.4, we can see that the expected weekly profits obtained in the second scenario are always superior to those obtained in the first scenario. This means both CES and WES benefit from conducting the contract. Yet the differences in some weeks (e.g., weeks 3 and 12) are smaller since the contract is not exercised for some days during the corresponding weeks. To better understand the details of transaction mechanism, in the following we concentrate on a specific day of the planning horizon, in which the contract is signed.

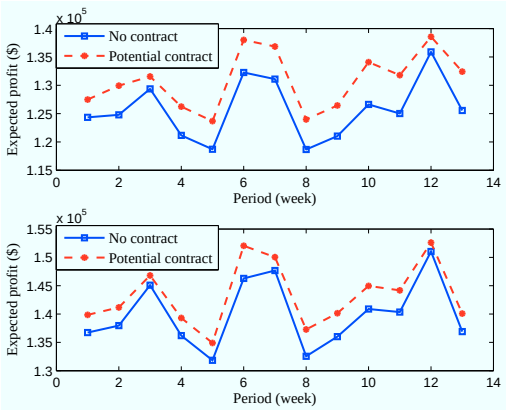

Figure 2.4: Expected weekly profits for WES (above) and CES (below)

The expected hourly profits of the WES are illustrated in Fig. 2.5. It can be observed that having transaction with the CES brings more benefits to the WES as its total increased profit is $3085, with increment rate of 3.1%. We notice that the CES also finds the bilateral contract viable with totally 2.5% increment rate. Fig. 2.5 also shows that the WES mostly gets advantage from the contract at time period 4. The reason is that the volatile RT market price becomes pretty high at this period and thus the WES inclines to utilize its backup

capacity from the CES instead of buying energy from the RT market with higher price. On

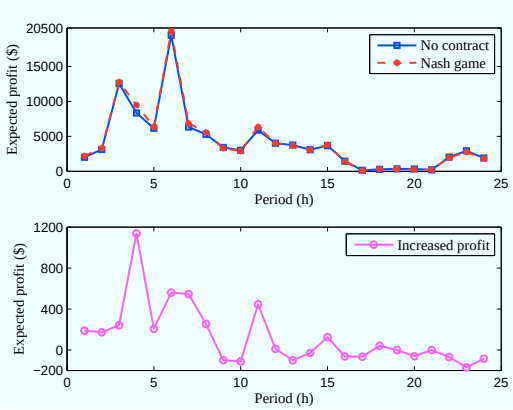

Figure 2.5: Expected hourly profits (above) and increased hourly profits (below) from participating in both energy market and bilateral contract for WES

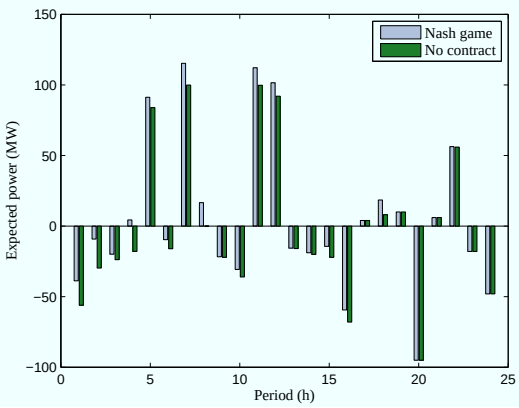

Figure 2.6: Expected hourly power traded with RT market for WES

the other hand, we observe that the profit of the first scenario (not engaging in the contract) is slightly greater than the second one (Nash game case) during some periods. For instance in time period 23, the execution price w_{23} is greater than the RT market price. Thus, the WES has no tendency to exercise the contract. However, since it has already paid $\theta_{23}S_{23}$ for reserving the backup capacity, adopting the contract yields to less profit for the WES at this period. It is worthy to mention that for some periods that the WES has no wish to use the backup, like time period 23. Therefore, our two-part price bilateral contract (swing option contract) provides a more cost-effective solution compared to the traditional one-part

price contract (forward contract). The reason is that our contract represents the right, but not the obligation, to purchase power at the prearranged execution price. Hence, when the market price is less than the execution price, the WES does not exercise the contract, but purchase the energy from the market with the market price to recover the shortage, plus a reservation price (reservation price usually contains a small portion of the execution price) paid to CES for reserving the capacity. However, the forward contract obligates the WES to purchase from the CES, though the contract price is higher than the market price. Therefore, the forward contract might be more costly to the WES. In other words, by considering the reservation price in the contract, the risk is more diversified between the suppliers and the WES is further motivated to sign the contract.

Fig. 2.6 indicates the expected energy traded in the RT market for each time periods. For periods that the expected DA price is higher than the RT price, the WES decides to bid into the DA market as much as possible and then recover its energy shortage by trading with CES and RT market. In this case, the bilateral contract provides a precious opportunity for WES to purchase less expensive power from the RT market. Moreover, for the periods that the RT price is higher than the DA price, the WES prefers to assign all of its capacity into the RT market. In this case, the contract allows the WES to utilize the backup capacity with low price and sell it to the RT market with higher price. For other periods like time periods 23 and 24, we can readily identify that the power traded in the RT market for Nash game and no contract cases are the same. The reason is that it is not economically justifiable for the WES to use its backup capacity. Considering above discussion, it can be clearly observed that our devised bilateral contract provides a flexible tool for the WES to secure itself with respect to the uncertainties in the operating day.

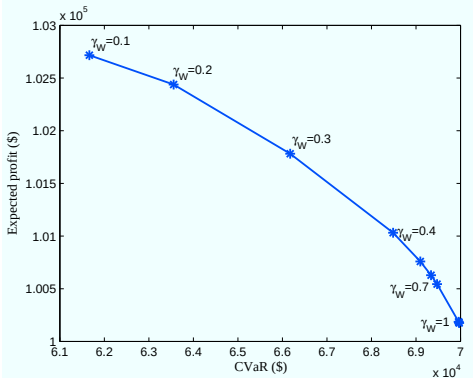

Figure 2.7: Expected profit versus CVaR for WES considering different risk perceptions

2.6.2 Effects of risk perception

Fig. 2.7 shows the efficient frontier for the WES. Efficient frontier plays a crucial role for the WES to resolve the tradeoff between the expected profit and risk. For a low risk-averse behavior of WES ($\gamma_W = 0.1$), the value of expected profit is \$102716 with CVaR of \$61663. By moving to a high risk-averse case ($\gamma_W = 1$), the value of CVaR grows by 13.5% at the cost of just 2.4% reduction in the expected profit. This is a compelling result since by a small decline in expected profit, the risk of profit volatility is significantly reduced.

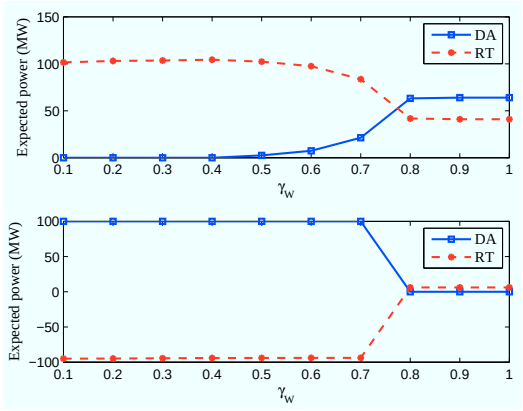

Figure 2.8: Expected power traded by WES in DA and RT markets for time period 12 (above) and time period 20 (below)

Fig. 2.8 illustrates the impact of various risk attitudes on the amount of power that WES trades with the DA and RT markets for time periods 12 and 20. We choose these periods

as they reveal two distinct properties. The wind output is highly volatile in time period 12 and the expected RT price is greater than the DA price. However, the wind output has less fluctuations in time period 20 and the DA price is greater than the RT price. Considering time period 12, as the WES becomes more risk-averse (the value of γ_W increases), it offers more power in the DA market and less in the RT market to hedge against the rise in RT price. In contrast, the WES has tendency to trade less power in the DA and more in the RT markets for time period 20. By this strategy, the WES lowers its profit volatility in the hope that it can sell its generation in the RT market at a reasonable high price.

2.6.3 Effects of multi-stage programming

In order for evaluate the performance of our multi-stage stochastic programming model, we compare the results with the two-stage stochastic case. We carry out case studies for 4 different days, with DA price samples of two sizes $(50, 100)$, and RT price as well as wind output samples of two sizes $(50, 100)$. According to the Table 2.2, the expected profits attained by the multi-stage model are always superior respect to those attained by the two-stage model. This is because the multi-stage solution comes up with more flexibility in RT decisions with respect to the uncertain parameter realizations. However, there is a tradeoff between flexibility and computational efficiency, when using the multi-stage model.

2.7 Conclusion

We develop a multi-stage stochastic programming model for the wind and conventional energy suppliers to optimize their bidding strategies in the both day-ahead and real-time markets. Our proposed model provides an opportunity to the wind supplier to update its real-time decisions as time progresses when more information about wind outputs and real-time market prices become available. In addition, using option contract with Nash equilibrium framework, an incentive-based trading mechanism is investigated to help the wind energy supplier recover

Table 2.2: Comparing objective values for multi-stage and two-stage stochastic programming

DA		50				100			
RT		50		100		50		100	
Instance		Two	Multi	Two	Multi	Two	Multi	Two	Multi
Day 1	profit($)	147587	148018	144620	144694	138872	139283	139429	139450
	time(s)	10.29	84.73	10.05	129.11	9.6	211.38	11.94	333.16
Day 2	profit($)	140049	140348	137870	138105	113004	113263	112537	112720
	time(s)	12.21	87.55	11.17	108.72	10.59	195.71	11.16	255.91
Day 3	profit($)	129962	143016	130792	144321	136225	149022	134446	147613
	time(s)	23.04	92.18	19.48	110.39	20.19	205.14	23.11	342.1
Day 4	profit($)	130506	131204	128988	129767	117397	118312	116864	117710
	time(s)	11.62	81.27	13.77	162.13	19.45	198.21	20.1	232.93

its energy deficit and at the same time bring the conventional energy supplier to obtain more profits. The computational results verify that our proposed approach is effective in accommodating wind and price uncertainties. Finally, as the future work of this chapter, our framework can be extended to the case with one conventional supplier and several renewable suppliers, and the case with several conventional suppliers and one renewable supplier.

CHAPTER III

A DATA-DRIVEN MODEL OF VIRTUAL POWER PLANTS IN DAY-AHEAD UNIT COMMITMENT

3.1 Problem description and literature review

With the influx of distributed energy resources (DERs), passive power networks are going through a transformation, from a centralized to a decentralized scheme, to enhance the flexibility and reliability of the system by offering more resources to the grid operator [43, 44]. However, since each individual distributed energy resource has small capacities and is lack of controllability over its outputs, small distributed energy resource is either excluded from participating in the wholesale energy market or it is not able to participate in a cost-effective manner. Currently, Independent System Operators (ISOs) have limited control over the DERs connected to the grid, because most of these DERs are invisible to ISOs. Moreover, due to the computational limitation, ISOs are not able to simultaneously co-optimize schedules of conventional generators and a huge number of DERs across the grid. Accordingly, the development of an entity in the wholesale market as a market participant to represent and operate these DERs, becomes an important approach to facilitate the utilization of renewable energy resources and modernization of electricity grid. As an aggregator, the virtual power plant (VPP) [5] acts as an intermediary between the DERs and ISO, and allows small DERs to be pooled and actively participate in the wholesale energy markets. Such a bundled entity alleviates the ISO from having to obtain additional reserves or other ancillary service products for mitigating renewables intermittency. The concept of VPP is successfully implemented in

Belgium and Netherlands [45], and is exploited in Germany and United Kingdom [46]. In July 2018, the first trial of VPP initiated by Tesla, has been successfully set up in South Australia, which is expected to generate 250 MW of solar power and 650 MWh of battery storage capacity [47]. Moreover, [48] investigates the method to estimate the operation and cost of VPP and report the study of a VPP project on an 11 KV system in Brixton by using the data from UK Power Network, and [49] reports the studies of VPPs across cities in the iURBAN project. However, although via aggregation, the VPP enlarges the visibility of DER units to ISOs in consideration of market participation and operation, it is still challenging to transform all the information of aggregated DERs into one bid with standard attributes such as maximum capacity, ramping restrictions, so that it can fit ISOs' bidding offer schema and the wholesale market's mechanism. In addition, submitting inaccurate parameters in the ISOs' market clear engine can jeopardize the grid operation. In this study, we provide a framework to characterize and evaluate the standard attributes/parameters in the VPP's bid submitted to the ISO that can optimize its entire portfolio as an aggregation of distributed energy resources.

Existing literature on the operation and scheduling of the VPP mostly seeks to optimize the dispatching of resources within the VPP so as to maximize its profitability. References [50, 51] consider the self-scheduling operation of the VPP, which means that the VPP only submits energy quantities to ISO, instead of quantity-price bid pairs. Price-based unit commitment models are proposed in [52, 53] to develop the bidding strategy for a VPP that behaves as a price-taker in the market, since the VPP is assumed to be a relatively small entity that has less impact on the market clearing price compared to the other market participants. In [54], a large scale of wind farm-energy storage system is studied as a price maker. The behavior of rival participants is taken into account using the residual demand curve. In [55] and [56], a bilevel problem for the optimal bidding strategy of VPP is addressed, in which the upper level aims to maximize the profit of VPP, and the lower level calculates

the ISO day-ahead market clearing price. In [50], VPPs are allowed to establish bilateral contracts to hedge against the volatility of the electricity market. The idea of introducing VPPs as ancillary service providers is investigated in [57, 58]. A detailed literature review about the scheduling problem of VPP is presented in [59].

Another key direction of studying VPP focuses on managing the uncertain parameters like renewable generation output and load consumption. Stochastic programming (SP) has been extensively utilized for this purpose [60, 61]. Using this approach, the uncertain parameters are characterized by a set of scenarios based on the estimated probability distribution. For example, [62] studies a two-stage stochastic programming model for the optimal offering strategy of a VPP in the day-ahead and balancing markets. The uncertain wind power and market price are represented by a set of equi-probable scenarios. A major obstacle of SP is that fixing a particular probability distribution of the uncertain parameter may yield to biased solutions with unreliable out-of-sample performance. Moreover, SP usually suffers from the curse of dimensionality, which means that the computational difficulty surges exponentially in the number of scenarios, and makes it impractical to solve large scale problems. As an alternative approach, robust optimization (RO) aims at constructing an uncertainty set to characterize the uncertain parameter, and allows the uncertain parameter to run adversely within the constructed uncertainty set to guarantee the feasibility of the optimal solution. For instance, in [63, 64], confidence bounds are constructed for the uncertain wind and market price, and robust bidding strategy models are proposed for a VPP consisting of price-responsive demands, wind power plants, and storage units. However, the RO approach is criticized as its over-conservativeness since it ignores the probabilistic nature of unknown parameters and the solution is solely based on the worst-case scenario.

To cope with the limitations of stochastic and robust optimization approaches, distributionally robust (DR) optimization models have been developed (see e.g., [65], [66]). According to this approach, the probability distribution of the uncertain parameters is itself subject to

uncertainty. In fact, the probability distribution is merely known to be within an ambiguity set, which can be characterized using certain statistical properties (e.g., estimation of mean and covariance). To guarantee the robustness of the approach, DR approach finds a solution that minimizes the worst-case expected cost over the ambiguity set. Unlike the traditional SP that exploits a collection of representative scenarios based on an estimated probability distribution to characterize the uncertain parameters, and thereby has no robustness to the error of distribution estimation, DR models release the assumption on any particular distribution. Therefore, this approach can accommodate the estimation error on the distribution due to the noisiness and incompleteness of the data, and also avoid the computational prohibition of scenario enumerations. Furthermore, contrary to the classical RO that is basically a distribution-free approach and minimize the total cost based on a worst-case scenario, DR models account for distributional knowledge through the ambiguity set, and minimizes the total expected cost based on a worst-case distribution over a set of probability distributions (ambiguity set). Hence, DR models trigger to less conservative solutions.

Because of these advantages, DR models have been successfully applied in power system problems under uncertainties, including contingency-constrained unit commitment [67], unit commitment with wind power integration [68], reserve scheduling [69], and optimal power flow [70]. However, most of the works fail to consider the physical limits (i.e., the support space) for the uncertain parameters, which is essentially critical in many applications. For example, renewable generators (e.g., wind turbines) have a limited capacity that cannot be exceeded from. The resulted solutions in these works, though restricted by the set of the distributions, are more conservative since the uncertain parameter itself can take very large positive or negative value even with its distribution still being within the range. Modeling of dependencies among the uncertain parameters (e.g., renewable energy outputs) is another crucial feature that is usually captured by including covariance matrix in the ambiguity set. Despite importance, there are limited studies in this regard. [69] stipulates that the

covariance of uncertain renewable outputs exactly matches the empirical covariance obtained from data and formulates a semi-definite program, and [71] bounds the correlation between pairs of wind farms generations and proposes a second-order conic program.

This study presents a two-stage DR model for the VPPs participation in the wholesale market. The VPP is a profit-driven entity that participates in the day-ahead wholesale market by submitting the cost curve and other related parameters like its maximum capacity and ramping limits to ISO. The goal is to find these standard attributes in the VPP's bid submitted to the ISO that can optimize its entire portfolio. The ISO collects such bidding information from all participants and run the reliability unit commitment to decide about the generation amount of each participant. Thus, the optimal solutions, i.e., the bidding parameter information of VPP will be served as an input for the ISO's unit commitment run. Furthermore, we consider two uncertainties in the model, i.e., the uncertain renewable energy output, and the unknown energy cleared by the ISO. We represent these two uncertainties with one parameter: the virtual net load, which is defined as the MW cleared by the ISO minus the renewable generation. It is assumed that in the first stage (day-ahead market), the VPP determines its total capacity and ramping limits to be reported to ISO that can optimize its entire portfolio by considering the physical constraints and the uncertain virtual net load. The second stage will give the first-stage a recourse, so that in the real-time operating day, after knowing the virtual net load, the VPP is able to supply enough power as it reported, by controlling its generation level of its conventional generators and power storage level. Using available moment information such as the empirical mean and covariance matrix of VPP's virtual net load that are learned from the data, we construct a second-order conic (SOC) representable ambiguity set for the unknown probability distribution, and reformulate the DR problem as a second-order conic programming (SOCP), which is efficiently solvable by off-the-shelf solvers like CPLEX. The objective of the model is to minimize the worst-case expected total cost over all probability distributions of the virtual net load in the ambiguity

set. The conservativeness of the model can be adjusted based on the preference of the VPP operator. That is, if the VPP operator utilizes more information about the virtual net load data, the ambiguity set becomes smaller and the model becomes less conservative accordingly. On the contrary, if the VPP operator ignores the moment information on probability distribution and just utilizes the boundary information of the virtual net load (i.e., upper and lower bound), the proposed model is reduced to a traditional robust optimization model. Moreover, the proposed ambiguity set is able to capture the temporal dependencies among different time periods in the virtual net load profile by considering the covariance matrix.

A more realistic approach to model the VPP's problem is to allow the sequential revelation of the uncertain virtual net load and restrict the dispatch decisions to only hinge on the virtual net load observed up to the current time period. This restriction is called the non-anticipativity of dispatch decisions and the resulting formulation describes a multi-stage model. The main advantage of the multi-stage model over the two-stage approach is that the former framework caters for a dynamic decision making, where the VPP operator has the opportunity to update its knowledge about uncertain outcomes as they unfold in periods. Using linear decision rules, we extend the two-stage model to the multi-stage case, where recourse decisions take the form of a linear function of uncertain virtual net loads and a set of auxiliary variables.

The contributions of this study are summarized as follows:

- We propose an innovative distributionally robust optimization model to help VPPs to optimally characterize the parameters in their bidding offers to ISO for the reliability unit commitment run.

- The DR model can effectively manage the intrinsic uncertainty arising from virtual net load consumption. An tractable reformulation of the proposed DR model is derived,

which can be implemented effectively by off-the-shelf solvers.

- To better capture the nonanticipativity of the uncertainty, we extend the two-stage case to a multi-stage DR problem by using linear decision rules, and benchmark it with the two-stage DR model.

The rest of the chapter is organized as follows. In Section 3.3, we formulate the two-stage and multi-stage DR models, and describe the ambiguity set of virtual net load probability distributions. In Section 3.4, we derive a separation framework to address the DR models. Finally, in Section 3.5, we demonstrate the effectiveness of two-stage and multi-stage models through several case studies.

3.2 Nomenclature

A. Sets

$\mathcal{N}^g$	Set of generators.
$\mathcal{T}$	Set of time periods.

B. Parameters

SU_i	Start-up cost for generator i.
SD_i	shut-down cost for generator i.
NL_i	No load cost for generator i.
CU_i^g	Maximal generation capacity for generator i.
CD_i^g	Minimal generation capacity for generator i.
UT_i	Minimum up-time for generator i.
DT_i	Minimum down-time for generator i.
RU_i	Ramp-up limit for generator i.
$\overline{\mathrm{RU}}_i$	Start-up ramp-up limit for generator i.
RD_i	Ramp-down limit for generator i.

$\overline{\mathrm{RD}}_i$ Shut-down ramp-down limit for generator i.

L_t^{in} Load level inside the VPP in time period t.

$C_i^g(.)$ Fuel cost function of generator i.

x_0^s Initial storage level.

η^{s+} Discharging efficiency of storage unit.

η^{s-} Charging efficiency of storage unit.

CU_t^{s+} Maximum discharging level of storage unit in time period t.

CU_t^{s-} Maximum charging level of storage unit in time period t.

CU^s Capacity of storage unit.

λ_t^C Capacity price in time period t.

C. First Stage Decision Variables

y_{it}^+ Binary variable to indicate if generator i is started up in time period t.

y_{it}^- Binary variable to indicate if generator i is shut down in time period t.

y_{it}^o Binary variable to indicate if generator i is on in time period t.

y_t^C Capacity offering by a VPP to the ISO for time period t.

y_t^{RU} Ramp-Up limit offering by a VPP to the ISO for time period t.

y_t^{RD} Ramp-Down limit offering by a VPP to the ISO for time period t.

D. Second Stage Decision Variables

x_{it}^g Power produced by generator i in time period t.

x_t^{s+} Amount of power discharged by storage unit in time period t.

x_t^{s-} Amount of power absorbed by storage unit in time period t.

x_t^{vpp} Total power generation of a VPP in time period t.

E. Random parameter

ξ_t Virtual net load consumption in time period t.

3.3 Mathematical formulation

We consider a VPP consisting of conventional generators, a wind farm, an energy storage unit, and non-flexible customers. Before submitting its bids to the day-ahead market, the VPP utilizes the moment information of virtual net load to construct an ambiguity set of virtual net load probability distribution. Using the ambiguity set and the physical characters of each conventional generators, the VPP determines the offering information needed to be submitted to the ISO, i.e., capacity, ramping limits, and cost curve in such a way that it minimizes the worst-case expected cost. We present the DR model as follows:

$$\min_{\mathbf{y}} \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}^g}(\mathrm{SU}_i y_{it}^{+} + \mathrm{SD}_i y_{it}^{-} + \mathrm{NL}_i y_{it}^{o}) - \sum_{t\in\mathcal{T}} \lambda_t^C y_t^C + \max_{\mathbb{P}\in\mathcal{D}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})], \tag{3.1a}$$

$$s.t.\ \ y_{it}^{o} - y_{i,t-1}^{o} = y_{it}^{+} - y_{it}^{-},\ \forall i\in\mathcal{N}^g,\ t\in\mathcal{T}, \tag{3.1b}$$

$$-y_{i,t-1}^{o} + y_{it}^{o} - y_{ik}^{o} \le 0, \quad \forall i\in\mathcal{N}^g,\ t\in\mathcal{T},\ 1\le k-(t-1)\le \mathrm{UT}_i, \tag{3.1c}$$

$$y_{i,t-1}^{o} - y_{it}^{o} + y_{ik}^{o} \le 1, \quad \forall i\in\mathcal{N}^g,\ t\in\mathcal{T},\ 1\le k-(t-1)\le \mathrm{DT}_i, \tag{3.1d}$$

$$y_t^C \le \sum_{i\in\mathcal{N}^g} \mathrm{CU}_i^g y_{it}^{o} - L_t^{in},\ \forall t\in\mathcal{T}, \tag{3.1e}$$

$$y_t^{RU} \le \sum_{i\in\mathcal{N}^g} (\mathrm{RU}_i y_{i,t-1}^{o} + \overline{\mathrm{RU}}_i y_{it}^{+}),\ \forall t\in\mathcal{T}, \tag{3.1f}$$

$$y_t^{RD} \le \sum_{i\in\mathcal{N}^g} (\mathrm{RD}_i y_{it}^{o} + \overline{\mathrm{RD}}_i y_{it}^{-}),\ \forall t\in\mathcal{T}, \tag{3.1g}$$

$$y_{it}^{o}, y_{it}^{+}, y_{it}^{-} \in \{0,1\},\ \ \forall i\in\mathcal{N}^g, t\in\mathcal{T}, \tag{3.1h}$$

where $\mathbf{y} := (\mathbf{y}^o, \mathbf{y}^+, \mathbf{y}^-, \mathbf{y}^C, \mathbf{y}^{RU}, \mathbf{y}^{RD})$ denotes first-stage decisions, $\mathcal{D}$ indicates the ambiguity set of virtual net load probability distribution, and $Q(\mathbf{y},\boldsymbol{\xi})$ represents the operating cost for a given first-stage decision $\mathbf{y}$ and realized virtual net load $\boldsymbol{\xi}$, and it can be calculated as

follows:

$$Q(\mathbf{y}, \boldsymbol{\xi}) = \min_{\mathbf{x}} \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}^g} C_i^g(x_{it}^g), \tag{3.2a}$$

$$s.t.\ \ x_t^{\text{vpp}} = \sum_{i\in\mathcal{N}^g} x_{it}^g + x_t^{s+} - x_t^{s-},\ \forall t \in \mathcal{T}, \tag{3.2b}$$

$$\xi_t \le x_t^{\text{vpp}} \le y_t^C,\ \ \forall t \in \mathcal{T}, \tag{3.2c}$$

$$x_t^{\text{vpp}} - x_{t-1}^{\text{vpp}} \le y_t^{RU},\ \ \forall t \in \mathcal{T}, \tag{3.2d}$$

$$x_{t-1}^{\text{vpp}} - x_t^{\text{vpp}} \le y_t^{RD},\ \ \forall t \in \mathcal{T}, \tag{3.2e}$$

$$\text{CD}_i^g y_{it}^o \le x_{it}^g \le \text{CU}_i^g y_{it}^o,\ \forall i \in \mathcal{N}^g, t \in \mathcal{T}, \tag{3.2f}$$

$$x_{it}^g - x_{i,t-1}^g \le \text{RU}_i y_{i,t-1}^o + \overline{\text{RU}}_i y_{it}^+,\ \forall i \in \mathcal{N}^g, t \in \mathcal{T}, \tag{3.2g}$$

$$x_{i,t-1}^g - x_{it}^g \le \text{RD}_i y_t^o + \overline{\text{RD}}_i y_{it}^-,\ \forall i \in \mathcal{N}^g, t \in \mathcal{T}, \tag{3.2h}$$

$$0 \le \eta^{s+} x_t^{s+} \le \text{CU}_t^{s+},\ \ \forall t \in \mathcal{T}, \tag{3.2i}$$

$$0 \le \eta^{s-} x_t^{s-} \le \text{CU}_t^{s-},\ \ \forall t \in \mathcal{T}, \tag{3.2j}$$

$$0 \le x_0^s + \sum_{j\in[1:t]} (\eta^{s-} x_j^{s-} - \frac{1}{\eta^{s+}} x_j^{s+}) \le \text{CU}^s, \forall t \in \mathcal{T}, \tag{3.2k}$$

where $\mathbf{x} := (\mathbf{x}^g, \mathbf{x}^{s+}, \mathbf{x}^{s-}, \mathbf{x}^{\text{vpp}})$ denotes the second-stage decisions, including the conventional generation amount, storage charging/discharging amount and total generation level. In the above formulation (3.1)–(3.2), the objective function (3.1a) is to minimize the worst-case expected total net cost, i.e., total cost minus the potential revenue by selling the capacity to the grid. The total cost includes start-up, shut-down, no load, and fuel costs. The potential revenue is assumed to be linearly depending on the value of offered capacity to ISO. Note that we do not put the potential revenue from selling power to the customers in the objective since it is a constant value in our model. Constraints (3.1b) indicate the commitment relationship among y_{it}^+, y_{it}^-, and y_{it}^o. Constraints (3.1c) and (3.1d) represent minimum up-time and minimum down-time limits, respectively. Constraints (3.1e) define an upper bound for

the total capacity offered by VPP to ISO, which is the total available capacity, minus the load inside the VPP. Constraints (3.1f) and (3.1g) characterize the ramp-up and ramp-down limits offered by VPP to ISO. The objective of formulation (3.2) is to minimize the total operating cost, while respecting dispatch and storage related constraints. More precisely, $C_i^g(x_{it}^g)$ represents the quadratic fuel cost function corresponding to the generation level x_{it}^g, and it can be estimated by a N-piece-wise linear function as follows:

$$x_{it}^c \geq \delta_i^n y_{it}^o + \varrho_i^n x_{it}^g, \ \forall n = 1, ..., N, \ i \in \mathcal{N}^g, \ t \in \mathcal{T}. \tag{3.3}$$

Constraints (3.2b) calculate the total power generation by the VPP. Constraints (3.2c) describe that the total generation amount of VPP should satisfy the virtual net load, and also should not exceed the capacity level. Constraints (3.2d) and (3.2e) describe ramp-up and ramp-down limits for the VPP. Constraints (3.2f)-(3.2h) describe generation as well as ramping limits for the conventional generators inside the VPP. Constraints (3.2i)-(3.2j) describe upper bounds for the amount of discharge and charge levels. Finally, constraints (3.2k) enforce energy storage limits for the storage unit. After solving the above models

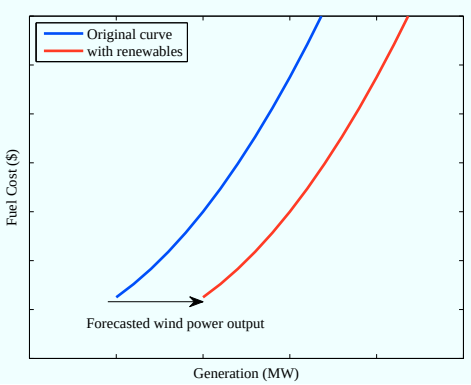

Figure 3.1: An example of cost curve

(3.1)–(3.2), the first-stage decisions including the capacity y_t^C and ramping limits y_t^{RU} and y_t^{RD} of the VPP can be found. As for the cost curve, since the storage unit is only responsible for smoothing out the scheduled generation levels by absorbing power when the demand

is low and generating power in peak demand hours, it does not impact on the cost curve. Therefore, the cost curve submitted to the ISO will have the similar shape as the cost curve of conventional generators, except a shift because of the wind power output, as illustrated in Fig. 3.1. The reason for shifting the curve is that the conventional generators will be started up only when the wind power output is not enough to met the demand.

3.3.1 Abstract formulation

For notation brevity, we recast the problem into the following compact matrix formulation:

$$\min_{\mathbf{y}\in\mathcal{Y}} \mathbf{c}^T\mathbf{y} + \max_{\mathbb{P}\in\mathcal{D}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})], \tag{3.4a}$$

with

$$Q(\mathbf{y},\boldsymbol{\xi}) = \min_{\mathbf{x}}\{\mathbf{d}^T\mathbf{x} :\ \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{x} \geq \mathbf{h} - \mathbf{M}\boldsymbol{\xi}\}, \tag{3.4b}$$

where $\mathbf{c}^T\mathbf{y}$ indicates commitment costs minus capacity value, $\mathcal{Y}$ represents the first-stage constraints (3.1b)–(3.1h), $\mathbf{d}^T\mathbf{x}$ indicates the operating cost, and $\mathbf{A}\mathbf{y}+\mathbf{B}\mathbf{x} \geq \mathbf{h}-\mathbf{M}\boldsymbol{\xi}$ represents the second-stage constraints (3.2a)–(3.2k) for a fixed commitment and offering decision $\mathbf{y}$ and virtual net load realization $\boldsymbol{\xi}$.

3.3.2 Ambiguity set construction

We construct an ambiguity set $\mathcal{D}$ of the virtual net load probability distribution by using its moment information. This ambiguity set can capture the dynamics of virtual net load evolution over time periods. Another main feature of this ambiguity set is that since it is a second-order conic representable set [72], it can be reformulated as a second-order conic program, which can be solved by many commercial solvers. More precisely, the ambiguity set

is constructed as:

$$\mathcal{D} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{\mathtt{T}}) \middle| \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \Omega) = 1, \\ E_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu}, \\ E_{\mathbb{P}}[(\boldsymbol{\xi} - \boldsymbol{\mu})^2] \leq \boldsymbol{\gamma}, \\ E_{\mathbb{P}}\Big[\big(\sum_{i=k}^{t}(\xi_i - \mu_i)\big)^2\Big] \leq \bar{\gamma}_{kt}, \forall k \leq t, \end{array} \right\} \tag{3.5}$$

where $\mathcal{P}_0(.)$ is the set of all probability distributions, $\mathtt{T} = |\mathcal{T}|$ is the number of time periods, Ω is the support space of $\boldsymbol{\xi}$ and defined by $\Omega = [\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}]$, $\boldsymbol{\mu}$ is the estimation of the mean value of uncertain virtual net load $\boldsymbol{\xi}$, and parameters $\boldsymbol{\gamma}$ and $\bar{\boldsymbol{\gamma}}$ are used to adjust the conservativeness of the optimal solution, which can capture the VPP operator's risk attitude. A risk-averse VPP operator may select larger values for these parameters to enrich the ambiguity set with more distributions, and thus arrive at a more robust solution. On the other hand, a risk-prone VPP operator may tend to choose smaller values of $\boldsymbol{\gamma}$ and $\bar{\boldsymbol{\gamma}}$ to exclude pathological distributions with the moment information far away from the sample ones, and thus obtain less conservative solutions. In the above set (3.5), the first constraint incorporates a range for the virtual net load. The second constraint ensures that $\boldsymbol{\xi}$, given by any distribution in $\mathcal{D}$, has the same mean as the empirical mean, while the third constraint bound the variance of $\boldsymbol{\xi}$. Finally, the last constraint captures the virtual net load correlations across time periods. Indeed, this constraint ensures that the variance of sum of the virtual net loads during the time window $[k, t]$ is bounded by $\bar{\gamma}_{kt}$. In practice, the ambiguity set (3.5) is characterized by parameters $\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}, \boldsymbol{\mu}, \boldsymbol{\gamma}$, and $\bar{\boldsymbol{\gamma}}$, which can be learned from historical data. Consider N data samples $\{\boldsymbol{\xi}^\ell\}_{\ell=1}^N$ of $\boldsymbol{\xi}$ such that $\boldsymbol{\xi}^\ell = [\xi_1, \xi_2, ..., \xi_{\mathtt{T}}]$, $\underline{\boldsymbol{\xi}}$ and $\overline{\boldsymbol{\xi}}$ can be set to be .05- and .95-quantiles of random virtual net load, respectively, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ can be chosen as the sample mean $\boldsymbol{\mu} = N^{-1}\sum_{\ell=1}^N \boldsymbol{\xi}^\ell$ and sample variance $\boldsymbol{\gamma} = N^{-1}\sum_{\ell=1}^N (\boldsymbol{\xi}^\ell - \boldsymbol{\mu})^2$. $\bar{\gamma}_{kt}$ can be determined by summing up specific elements in the covariance matrix of $\boldsymbol{\xi}$, i.e., $\bar{\gamma}_{kt} = \mathbf{f}_{kt}^T \boldsymbol{\Sigma} \mathbf{f}_{kt}$, where $\mathbf{f}_{kt}$ is a vector with i^{th} element equals to one if i is within the time window $[k, t]$, and zero

otherwise, and $\mathbf{\Sigma}$ denotes sample covariance matrix $\mathbf{\Sigma} = N^{-1}\sum_{\ell=1}^{N}(\boldsymbol{\xi}^{\ell} - \boldsymbol{\mu})(\boldsymbol{\xi}^{\ell} - \boldsymbol{\mu})^{T}$.

As explained in [73], in order to derive a tractable reformulation for the DR problems with the ambiguity set (3.5), it is equivalent to work with the following lifted ambiguity set:

$$\overline{\mathcal{D}} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{\mathrm{T}\times\mathrm{T}\times\frac{\mathrm{T}(\mathrm{T}+1)}{2}}) \;\middle|\; \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \overline{\Omega}) = 1, \\ E_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu}, \\ E_{\mathbb{P}}[\mathbf{u}] \leq \boldsymbol{\gamma} \\ E_{\mathbb{P}}[v_{kt}] \leq \bar{\gamma}_{kt}, \ \forall k \leq t, \end{array} \right\} \tag{3.6}$$

where $\overline{\Omega}$ is the lifted support set and defined as follows:

$$\overline{\Omega} = \left\{ (\boldsymbol{\xi}, \mathbf{u}, \mathbf{v}) \;\middle|\; \begin{array}{l} \underline{\boldsymbol{\xi}} \leq \boldsymbol{\xi} \leq \overline{\boldsymbol{\xi}}, \\ (\boldsymbol{\xi} - \boldsymbol{\mu})^2 \leq \mathbf{u}, \\ \left(\sum_{i=k}^{t}(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)\right)^2 \leq v_{kt}, \ \forall k \leq t, \end{array} \right\} \tag{3.7}$$

That is, $\mathcal{D}$ includes set of marginal distributions of $\boldsymbol{\xi}$, where the joint distribution of $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ is in $\overline{\mathcal{D}}$.

3.3.3 Multi-stage formulation

The main assumption of the two-stage model is that all of the real-time dispatch decisions are made simultaneously at the beginning of the operating day. However, the unit commitment problem is inherently sequential. That means the uncertain virtual net load is revealed as the time progresses, and dispatch decisions at each period are made after knowing the realization of uncertain parameters up to that period (non-anticipativity enforcement). In other words, the VPP operator first observes the uncertain virtual net load of the first period $\boldsymbol{\xi}_1$, and then takes real-time decisions $\mathbf{x}_1(\boldsymbol{\xi}_1)$ of the first period. Subsequently, the virtual net load of second period $\boldsymbol{\xi}_2$ is realized, and then the VPP operator takes real-time decisions $\mathbf{x}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$

of the second period accordingly. This alternating process continues over the entire T periods. We present the multi-stage DR model as follows:

$$\min_{\mathbf{y}\in\mathcal{Y},\mathbf{x}(.)} \quad \left(\mathbf{c}^T\mathbf{y} + \max_{\mathbb{P}\in\overline{\mathcal{D}}} E_{\mathbb{P}}\Big[\mathbf{d}^T\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]})\Big]\right) \tag{3.8a}$$

$$s.t. \;\; \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}) \geq \mathbf{h} - \mathbf{M}\boldsymbol{\xi}_{[t]}, \;\; \forall(\boldsymbol{\xi}, \mathbf{u}) \in \overline{\Omega}, \tag{3.8b}$$

where $\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]})$ denotes that the dispatch decision at time period t is a function of the uncertain virtual net load as well as the auxiliary random variable associated with the lifted ambiguity set realized up to time period t, and $\min_{\mathbf{x}(.)}$ can be interpreted as optimizing over the policies, i.e., functions of random variables. Solving problem (3.8) is challenging since dispatch decisions are general functions of all past uncertain parameter realizations, instead of a finite vector of decision variables [74]. An effective approach on addressing this fully adaptive problem is to apply the linear decision rule technique, which restricts the dispatch decisions to be linear functions of the uncertain parameters ([68, 75, 76]). More precisely, we define the following policies for decision variables $x^g_{it}, x^{s+}_{it}, x^{s-}_{it}, x^{\text{vpp}}_{it}$:

$$x_{it}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}) = x^0_{it} + x^1_{it}\xi_t + x^2_{it}u_t. \tag{3.9}$$

Notice that linear coefficients $(\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2)$ are considered as decision variables in the problem and they are defined separately for decision variable $x^g_{it}, x^{s+}_{it}, x^{s-}_{it}, x^{\text{vpp}}_{it}$. Furthermore, in these policies it is assumed that the second-stage decision variables are affine functions of both primary and auxiliary random parameters. [73] has shown that this enhanced linear decision rule can significantly improve the computational results compared with the one restricting to the primary random parameters only. In addition, with the decision rule defined in (3.9), the nonanticipativity requirement can be met automatically.

3.4 Solution methodology

In this section, we develop solution methods for the two-stage DR problem (3.4) and multi-stage DR problem (3.8), respectively.

3.4.1 Solution method for two-stage DR

First, we dualize the second-stage worst-case expectation problem $\max_{\mathbb{P}\in\overline{\mathcal{D}}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})]$ and obtain the following dual problem:

$$\min_{\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\alpha} \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{3.10a}$$

$$s.t.\ \ \eta \geq F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}), \tag{3.10b}$$

with

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} Q(\mathbf{y},\boldsymbol{\xi}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{3.10c}$$

where $\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}$ are dual variables associated with the constraints in the ambiguity set (3.6). Note that we can add a slack variable in constraint (3.2c) to ensure the feasibility of the second-stage problem and in return consider a penalty cost for under-generation in the objective function. Since the second-stage linear problem $Q(\mathbf{y},\boldsymbol{\xi})$ is feasible and bounded, strong duality holds and $Q(\mathbf{y},\boldsymbol{\xi})$ can be replaced by its dual formulation. Hence, the subproblem (3.10c) can be equivalently formulated as follows:

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) =$$

$$\max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \max_{\pi\in\Pi} \boldsymbol{\pi}^T(\mathbf{h} - \mathbf{M}\boldsymbol{\xi} - \mathbf{B}\mathbf{y}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{3.11}$$

where $\Pi = \{\boldsymbol{\pi} \geq 0 : \boldsymbol{\pi}^T\mathbf{B} = \mathbf{d}\}$ is the set of dual variables of $Q(\mathbf{y},\boldsymbol{\xi})$. Since the feasible regions Π and $\overline{\Omega}$ are separable, the optimal solution of problem (3.11) occurs at extreme points of these regions. Therefore, if we denote all the extreme points of Π as $\{\boldsymbol{\pi}_i^*\}_{i=1}^I$ and exchange

the order of two maximization operations, i.e., $\max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}}$ and $\max_{\pi\in\Pi}$, $F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha})$ can be further reformulated as follows:

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{\pi_i,\forall i} \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \boldsymbol{\pi}_i^{*T}(\mathbf{h}-\mathbf{M}\boldsymbol{\xi}-\mathbf{B}\mathbf{y}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}. \tag{3.12}$$

Since the inner maximization problem in (3.12) is bounded with non-empty interior, conic duality can be applied. Taking dual of inner maximization problem in (3.12) leads to (details are provided in the Appendix A):

$$\begin{aligned}
F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{\pi_i,\forall i}\ & \boldsymbol{\pi}_i^{*T}(\mathbf{h}-\mathbf{B}\mathbf{y}) \\
& + \min_{\boldsymbol{\psi}}\ \ \overline{\boldsymbol{\xi}}^T\bar{\boldsymbol{\tau}} - \underline{\boldsymbol{\xi}}^T\tilde{\boldsymbol{\tau}} - \boldsymbol{\mu}^T\bar{\boldsymbol{\theta}} + \frac{1}{2}\mathbf{1}^T(\hat{\boldsymbol{\theta}}-\tilde{\boldsymbol{\theta}}) \\
& - \sum_{t\in\mathcal{T}}\sum_{k\in[1:t]}\sum_{i\in[k:t]} \mu_i\bar{\phi}_{kt} + \frac{1}{2}\mathbf{1}^T(\hat{\boldsymbol{\phi}}-\tilde{\boldsymbol{\phi}}), && (3.13a)\\
s.t.\ \ & \tilde{\tau}_t - \bar{\tau}_t + \bar{\theta}_t + \bar{\phi} = \mathbf{e}_t{}^T(\mathbf{M}^T\boldsymbol{\pi}_i^* + \boldsymbol{\lambda}),\ \ \forall t\in\mathcal{T}, && (3.13b)\\
& \frac{1}{2}(\tilde{\theta}_t + \hat{\theta}_t) = \beta_t,\ \ \forall t\in\mathcal{T}, && (3.13c)\\
& \frac{1}{2}(\tilde{\phi}_{kt} + \hat{\phi}_{kt}) = \alpha_{kt},\ \ \forall t\in\mathcal{T}, k\le t, && (3.13d)\\
& \sqrt{\bar{\theta}_t^2 + \tilde{\theta}_t^2} \le \hat{\theta}_t,\ \ \forall t\in\mathcal{T}, && (3.13e)\\
& \sqrt{\bar{\phi}_{kt}^2 + \tilde{\phi}_{kt}^2} \le \hat{\phi}_{kt},\ \ \forall t\in\mathcal{T}, k\le t, && (3.13f)
\end{aligned}$$

where $\mathbf{1}$ is a vector whose elements are all one, $\boldsymbol{e}_t$ is an unit vector whose t^{th} component is one and zero otherwise, and $\boldsymbol{\psi} := \{\bar{\boldsymbol{\tau}},\tilde{\boldsymbol{\tau}},\bar{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}},\bar{\boldsymbol{\phi}},\tilde{\boldsymbol{\phi}},\hat{\boldsymbol{\phi}}\}$ are dual variables associated with the constraints in the inner maximization, i.e., constraints in $\overline{\Omega}$ (3.7). We rewrite the inner minimization problem in (3.13) in a compact form: $\min\{\mathbf{b}^T\boldsymbol{\psi} : \mathbf{E}\boldsymbol{\psi} \succeq_\kappa \mathbf{f}\boldsymbol{\alpha}+\mathbf{g}\boldsymbol{\beta}+\mathbf{p}\boldsymbol{\lambda}+\mathbf{q}\boldsymbol{\pi}^*\}$, where $\succeq_\kappa$ denotes the generalized inequality respect to some cone κ. Now, by considering the second-stage dual formulation (3.10) and replacing $F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha})$ in (3.10c) with formulation

(3.13), the original DR optimization problem (3.4) is equivalent to:

$$\min_{\mathbf{y},\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\psi}_i} \mathbf{c}^T\mathbf{y} + \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{3.14a}$$

$$s.t.\ \ \eta \geq \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}_i, \quad \forall i = 1, \ldots, I, \tag{3.14b}$$

$$\mathbf{E}\boldsymbol{\psi}_i \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}_i^*, \quad \forall i = 1, \ldots, I, \tag{3.14c}$$

$$\mathbf{y} \in \mathcal{Y}, \tag{3.14d}$$

where $\boldsymbol{\psi}_i$ is a vector of decisions corresponding to the extreme point $\boldsymbol{\pi}_i^*$. Notice here that in (3.14b), we release the maximization operation in (3.13a) since it is equivalent to restrict the constraint (3.14b) holds for every extreme point $\boldsymbol{\pi}_i^*$, and we release the minimization operation in (3.13a) since it is equivalent to the existence of a feasible solution of constraint (3.14b) with constraint (3.14c). The above problem (3.14) has an appropriate structure to apply a two-level decomposition algorithm. More precisely, in the kth iteration, we solve the following master problem:

$$\min_{\mathbf{y},\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\psi}_i} \mathbf{c}^T\mathbf{y} + \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{3.15a}$$

$$s.t.\ \ \eta \geq \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}_i, \quad \forall i \leq k, \tag{3.15b}$$

$$\mathbf{E}\boldsymbol{\psi}_i \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}_i^*, \quad \forall i \leq k, \tag{3.15c}$$

$$\mathbf{y} \in \mathcal{Y}, \tag{3.15d}$$

to obtain the first-stage decisions. Then, the master problem (3.15) can be augmented iteratively (i.e., adding new variables $\boldsymbol{\psi}_i$ and the corresponding cuts) using the information provided by the subproblem (3.11). The more detailed procedure for solving the problem (3.14) is summarized in Algorithm 1. Observe that since the feasible region Π includes finite number of extreme points, the Algorithm 1 converges in finite number of steps. This is shown in the proposition 1.

Algorithm 1: Solution procedure for problem (3.14)

Initialize iteration index $i = 0$, and set $\mathcal{I} = \emptyset$.
repeat
  Solve the master problem (3.15). Let $(\mathbf{y}, \eta, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ be the optimal solution.
  Solve the subproblem problem $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ in (3.11) using the Algorithm 2.
  Let $\boldsymbol{\pi}^*_{i+1}$ be the optimal solution.
  $\mathcal{I} = \mathcal{I} \cup \{\boldsymbol{\pi}^*_{i+1}\}$.
  Define the new variable $\boldsymbol{\psi}_{i+1}$ and the corresponding constraints (3.15b)–(3.15c).
  $i = i + 1$.
until $\eta \geq F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$.
Output: $(\mathbf{y}, \eta, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ is the optimal solution.

Proposition 1. *Algorithm 1 converges to the optimal solution of problem* (3.14) *in a finite number of iterations.*

Proof. We rewrite the inner minimization problem in (3.13) as: $\min\{\mathbf{b}^T\boldsymbol{\psi} : \psi \in \Psi\}$, where $\Psi = \{\psi : \mathbf{E}\boldsymbol{\psi} \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}^*\}$. Accordingly, F can be written as:

$$F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) = \max_{\pi_i, \forall i} \min_{\psi \in \Psi} \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}, \tag{3.16}$$

where $\{\boldsymbol{\pi}^*_i\}_{i=1}^I$ represents the set of extreme points of Π. Since Π is a polyhedron, $\{\boldsymbol{\pi}^*_i\}_{i=1}^I$ is a finite set, i.e., its cardinality is equal to I. In the kth iteration of Algorithm 1, we have that:

$$\eta = \max_{\pi_i, i \leq k} \min_{\psi \in \Psi} \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}. \tag{3.17}$$

Comparing (3.16) and (3.17), it is obvious that $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) \geq \eta$. Therefore, the stopping criterion for the algorithm is $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) \leq \eta$. We will report $(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ as the optimal first-stage decisions when this criterion is achieved. In the worst-case condition, the Algorithm explores all the extreme points of Π, i.e. it reaches at iteration I. Therefore, the stopping criterion will be achieved for at most I iterations, which means that Algorithm 1 converges to the optimal solution in a finite number of iterations. □

Moreover, notice that the Algorithm 1 is a natural extension of the column-and-constraint generation algorithm proposed in [77], where the polyhedron uncertainty set is extended to a SOC set. A similar extension is proposed in [78] to solve the robust AC optimal power flow problem.

To solve the subproblem, we need to evaluate $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ at each iteration of Algorithm 1, which is a second-order conic problem with a quadratic objective function. We employ the Alternative Separation Heuristic (ASH) to solve this problem. The basic idea behind this framework is to obtain the optimal $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ with a fixed $\boldsymbol{\pi}$, and then fixed the obtained $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ to find the optimal $\boldsymbol{\pi}$ of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$. This back and forth procedure continues until the optimality gap is no more than a predefined level. Since the feasible region Π is a polyhedron, the ASH converges to a KKT point of (3.11) in a finite number of iterations [79, 78]. The ASH framework is summarized in Algorithm 2.

Algorithm 2: Alternative Separation Heuristic for solving subproblem (3.11)

Pick a $\widehat{\boldsymbol{\pi}} \in \Pi$, and optimality gap $\widehat{\epsilon}$
repeat
  Fix $\boldsymbol{\pi} = \widehat{\boldsymbol{\pi}}$ and solve (3.11). Let $(\widehat{\boldsymbol{\xi}}, \widehat{\mathbf{u}}, \widehat{\mathbf{v}})$ be the optimal solution with objective value Υ_1.
  Fix $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v}) = (\widehat{\boldsymbol{\xi}}, \widehat{\mathbf{u}}, \widehat{\mathbf{v}})$ and solve (3.11). Let $\widehat{\boldsymbol{\pi}}$ be the optimal solution with objective value Υ_2.
until $|\Upsilon_1 - \Upsilon_2| \leq \widehat{\epsilon}$.
Output: Υ_1 is the estimation of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ with solution $\widehat{\boldsymbol{\pi}}$.

3.4.2 Solution method for multi-stage DR

In order to solve the multi-stage DR model (3.8) using linear decision rules (3.9), we can apply a duality-based approach to reformulate the problem. More specifically, similar to the previous approach we start from dualizing the worst-case expectation problem $\max_{\mathbb{P} \in \overline{\mathcal{D}}} E_{\mathbb{P}}[Q(\mathbf{y}, \boldsymbol{\xi})]$.

Thus, (3.8) is equivalent to:

$$\min_{\mathbf{y}\in\mathcal{Y},\mathbf{x}^0,\mathbf{x}^1,\mathbf{x}^2} \mathbf{c}^T\mathbf{y} + \min_{\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\alpha} \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{3.18a}$$

$$s.t.\ \ \eta \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \mathbf{d}^T\mathbf{x}(\boldsymbol{\xi}_{[t]},\mathbf{u}_{[t]}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{3.18b}$$

$$\mathbf{A}\mathbf{y} - \mathbf{h} \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} -\mathbf{B}\mathbf{x}(\boldsymbol{\xi}_{[t]},\mathbf{u}_{[t]}) - \mathbf{M}\boldsymbol{\xi}. \tag{3.18c}$$

By substituting (3.9) into the above formulation, constraints (3.18b) and (3.18c) become:

$$\eta \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \mathbf{d}^T(\mathbf{x}^0 + \boldsymbol{\xi}^T\mathbf{x}^1 + \mathbf{u}^T\mathbf{x}^2) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{3.19a}$$

$$\mathbf{A}\mathbf{y} - \mathbf{h} \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} -\mathbf{B}(\mathbf{x}^0 + \boldsymbol{\xi}^T\mathbf{x}^1 + \mathbf{u}^T\mathbf{x}^2) - \mathbf{M}\boldsymbol{\xi}. \tag{3.19b}$$

Note that constraints (3.19a) and (3.19b) have the same structure. In the following, we derive a reformulation for the constraint (3.19a). Similar approach can be done for constraint (3.19b). Applying strong duality, (3.19a) can be written as:

$$\eta - \mathbf{d}^T\mathbf{x}^0 \geq \min_{\boldsymbol{\chi}\in\Xi(\mathbf{d},\mathbf{x}^1,\mathbf{x}^2)} \mathbf{g}^T\boldsymbol{\chi}, \tag{3.20}$$

where $\boldsymbol{\chi}$ and $\Xi(\mathbf{d},\mathbf{x}^1,\mathbf{x}^2)$ are the corresponding dual variables and dual feasible region of the constraints in $\overline{\Omega}$ with respect to $(\boldsymbol{\xi},\mathbf{u},\mathbf{v})$, and $\mathbf{g}$ is also derived from the definition of $\overline{\Omega}$. Note that since the second-order cone is self-dual, feasible region Ξ has a similar structure to $\overline{\Omega}$. Relation (3.20) is further equivalent to the existence of $\boldsymbol{\chi} \in \Xi(\mathbf{d},\mathbf{x}^1,\mathbf{x}^2)$ such that $\eta - \mathbf{d}^T\mathbf{x}^0 \geq \mathbf{g}^T\boldsymbol{\chi}$. Thus, using duality-based approach, constraints (3.19a)–(3.19b) can be reformulated as a finite number of linear and second-order conic constraints. As a result, the original problem (3.18) can be reformulated as a single minimization problem involving new

added variables and constraints.

3.5 Case study

In this section, we first apply the two-stage DR model based on a real data set to provide the operation of the VPP. Second, we test the performance of two-stage and multi-stage DR models for various simulated data sets. All the experiments are implemented in the C++ language with CPLEX 12.6 on a computer with Intel Xeon 3.2 GHz and 8 GB memory.

3.5.1 Data preparation

We consider a VPP including three conventional generators, a wind farm, a storage unit, and a set of loads that all are located in a single bus in the system. The characteristics of the generators are shown in Tables 3.2–3.3. The load and wind power outputs are collected from PJM market website [80] and scaled down to fit the test system. The mean value together, upper and lower limits of the virtual net load profile are depicted in Fig. 3.2. We assume that the capacity and efficiency of the storage unit on both absorbing and generating electricity are 100MW and 0.9, respectively. The load consumption inside the VPP and the capacity price are randomly simulated from intervals $[4, 20]$MWs and $\$[5, 30]$/MW, respectively. The penalty cost for load shedding is set to be \$6000/MWh. In addition, for the comparative studies in subsections 3.5.4 and 3.5.3, we consider various numerical settings to evaluate the performance of DR models. More precisely, we randomly generate 5,000 samples of virtual net load consumption $\boldsymbol{\xi}$ such that each ξ_t follows a normal distribution with mean $\mu_t \in [50, 80]$, standard deviation $\sigma_t = \mu_t \varepsilon$ with $\varepsilon \in \{0.1, 0.2\}$, and correlation coefficient $\zeta \in \{0, 0.25, 0.5, 0.75, 1\}$. These 5,000 samples comprise the support set Ω, which will be used later to build the ambiguity set.

Table 3.2: Generator data

Gen	Lower (MW)	Upper (MW)	Min. down (h)	Min. up (h)	Ramp (MW/h)
G1	5	20	1	1	10
G2	5	50	2	2	25
G3	50	100	2	3	50

Table 3.3: Fuel data

Gen	a (MBtu)	b (MBtu/ MWh)	c (MBtu/ MW^2h)	Start Up Fuel (MBtu)	Fuel Price ($/MBtu)
G1	31.67	29.24	0.0697	40	1
G2	58.81	22.94	0.0098	60	1
G3	50	6	0.0004	100	1

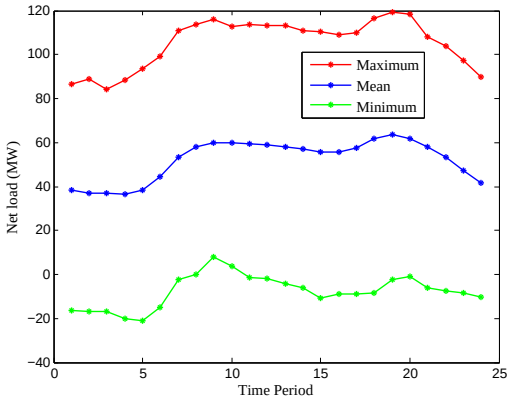

Figure 3.2: Virtual net load profile

3.5.2 VPP offering parameters

We solve the two-stage DR model to determine the VPP offering parameters. The offering capacity at each time period is reported in Fig. 3.3. The optimal ramp-up and ramp-down parameters are obtained 75MW/h for periods 19, 20, and 21, and 85MW/h for other periods. As it can be seen in Fig. 3.3, there is a significant reduction in offering capacity at time periods 18 and 20. The reason is that the capacity price at these periods suddenly drops to its minimum level. To recover this unexpectedly low capacity price, the generator $G1$, which is more expensive than $G2$ and $G3$, is turned off at these periods. On the other hand, the

VPP offers higher level of capacity at time periods like 1,2, and 22 since the inside load is low at these periods. In order to investigate the performance of the DR model in the real time operations, we fixed the optimal unit commitment and offering decisions and solve the dispatch problem for a realized virtual net load. The total power generation by conventional generators, storage operation, and the realized virtual net load are shown in Fig. 3.4. We can observe that the storage unit absorbs the power during the periods that the virtual net load is low (e.g., periods 5–6), and generates the power during the periods that the virtual net load is high (e.g., 8–10). More specifically, the storage unit absorbs 70MW power at period 5 (valley demand hour) when the wind power output is higher than the grid demand, and generates 43MW at period 17 to preclude from load shedding when the expensive generator $G1$ is offline at this period. Thus, we can conclude that the storage unit significantly contributes to the VPP flexibility.

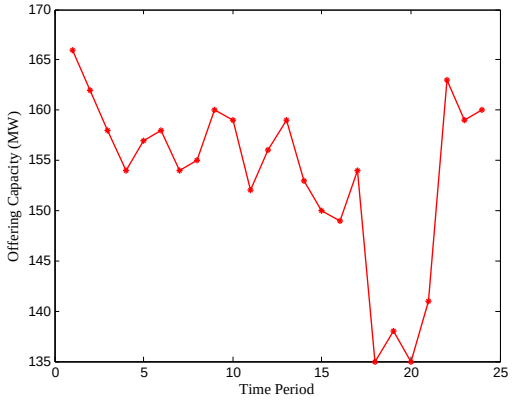

Figure 3.3: Capacity offered to ISO

3.5.3 Comparing with robust optimization

In this subsection, we compare the two-stage DR (TDR) model with the RO approach. For this purpose, since the VPP operator has limited information of Ω, we build the ambiguity set of virtual net load probability distribution by choosing 50 samples from Ω, and calculating

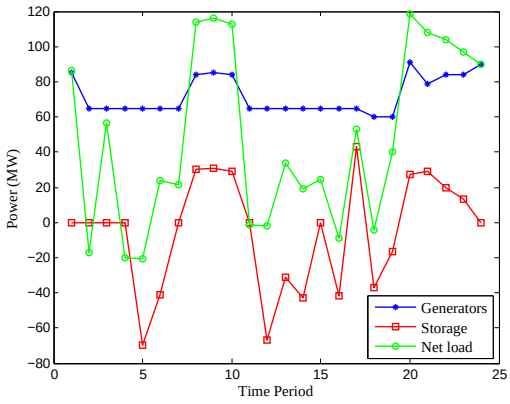

Figure 3.4: Optimal dispatch decisions

the parameters $\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}, \boldsymbol{\mu}, \boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}$, as explained in subsection 3.3.2. After solving the corresponding TDR and RO models, we fix the optimal first-stage decision $\mathbf{y}$, and run the second-stage problem $Q(\mathbf{y}, \boldsymbol{\xi})$ for all 5,000 samples in Ω. We report the average operating cost in Table 3.4. The results verify that TDR model provides less conservative solutions compared to the RO approach. In particular, TDR model can save up to \$532. Moreover, as the value of ε increases, i.e., the size of support set is expanded, the difference between TDR and RO models reduces, since the TDR model becomes more conservative.

Table 3.4: Two-stage DR vs. two-stage RO

	$\varepsilon = 0.1$		$\varepsilon = 0.2$	
ζ	TDR	RO	TDR	RO
$\zeta = 0$	19169	19701	24114	24461
$\zeta = 0.25$	19367	19715	24949	24956
$\zeta = 0.5$	19215	19730	24701	25041
$\zeta = 0.75$	19290	19801	24168	24512
$\zeta = 1$	19378	19887	23876	24378

3.5.4 Comparing with multi-stage model

In this subsection, we benchmark the multi-stage DR model (MDR) with the two-stage DR model (TDR), and report the profit values (i.e., the opposite value of the objective) and CPU

time in Table 3.5. We can observe that the MDR model provides a lower profit, which is because that the MDR model enforces nonanticipativity constraints. That is, as compared with MDR model, the TDR model assumes the virtual net load consumption information throughout the entire time periods is known at the beginning of the operating day, which offers more flexibility in the dispatch decisions.

Table 3.5: Multi-stage DR vs. two-stage DR

		$\varepsilon = 0.1$		$\varepsilon = 0.2$	
ζ		TDR	MDR	TDR	MDR
$\zeta = 0$	profit($)	34370	28421	27470	25200
	time(s)	71	39	82	44
$\zeta = 0.25$	profit($)	34364	28485	16583	14625
	time(s)	92	35	101	59
$\zeta = 0.5$	profit($)	31189	25435	18550	17431
	time(s)	165	66	144	72
$\zeta = 0.75$	profit($)	34548	28382	28688	23723
	time(s)	89	56	80	59
$\zeta = 1$	profit($)	32090	26952	19795	11006
	time(s)	109	79	111	85

Table 3.6: Two-stage DR vs. two-stage RO for a 6-conventional-generator case

Time series model	TDR	RO
$\text{ARIMA}(1,1,1)(1,0,0)_{24}$	30492	32152
$\text{ARIMA}(2,1,2)(0,1,1)_{24}$	30844	32297
$\text{ARIMA}(2,1,2)(1,1,1)_{24}$	30579	32348

3.5.5 Computational results for a complicated system

In this subsection, we evaluate the performance of our framework for the case that the random samples do not necessarily follow a standard distribution such as the multivariate normal distribution. For this purpose, the stochastic process describing the virtual net load behavior is captured through parsimonious time series models. Following the Box-Jenkins'

procedure of the model identification [81] and using Bayesian information criterion, we choose three competing seasonal ARIMA models and simulate samples accordingly. In the case studies, we also assume that the VPP includes six conventional generators to represent a more complicated system. We follow the procedure in the subsection 3.5.3 to compare the TDR with RO using samples generated from time series models. As shown in Table 3.6, TDR model leads to less conservative results than those of RO in the out-of-sample simulation. This demonstrates efficiency of our approach even when the underlying uncertainty distribution is not normal. Furthermore, Table 3.7 compares profit values obtained in TDR and MDR approaches for the complicated system. We can also see that TDR model yields to more profits at the expense of having a full knowledge of realized virtual net load throughout the entire scheduling horizon. This confirms our observation from the previous subsection 3.5.4.

Table 3.7: Multi-stage DR vs. two-stage DR for a 6-conventional-generator case

Time series model		TDR	MDR
$\text{ARIMA}(1,1,1)(1,0,0)_{24}$	profit($)	49708	42890
	time(s)	265	104
$\text{ARIMA}(2,1,2)(0,1,1)_{24}$	profit($)	47993	41344
	time(s)	201	110
$\text{ARIMA}(2,1,2)(1,1,1)_{24}$	profit($)	48770	42351
	time(s)	234	131

CHAPTER IV

DISTRIBUTIONALLY ROBUST DISTRIBUTION NETWORK CONFIGURATION UNDER RANDOM CONTINGENCY

4.1 Problem description and literature review

The distribution network planning is widely investigated in existing literature and broadly categorized in three parts: distribution configuration planning ([82, 83]), distribution reconfiguration and self-healing planning ([84, 85]), and distribution reinforcement and expansion planning ([86, 87]). The main objective of distribution configuration planning is to design a new system to meet the demand in the most cost-effective and reliable way. Distribution reconfiguration and self-healing planning aim at improving or recovering of network functionality by altering the topological structure of the network. In particular, a self-healing process is brought up when a contingency occurs in the system. Distribution reinforcement and expansion planning involve enhancing the resilience of the network to protect against possible damages or expanding current facilities to increase reliability. This chapter focuses on the distribution network configuration part. As we borrow some ideas from the self-healing and reinforcement planning literature, we also briefly review the relevant works in these domains.

Existing mathematical models of distribution network configuration involve various design variables that usually include the location [88] and size [89] of equipments like substations and feeders. As the penetration of distributed generation (DG) resources grows, the location and sizing of the DG units has also received increasing attention in the literature (see, e.g., [90], [91], [92, 93]). The network topology is another important design variable (see,

e.g., [94, 95, 96], [97]). [94] proposes an optimal network topology design that minimizes investment and variable costs associated with power losses and reliability. [95] considers network reconfiguration and maintains a radial network topology by ensuring that the node-incidence matrix has non-zero determinant. [96] explicitly incorporates the radiality constraints in the distribution system configuration model and considers the integration of DG units. None of the above works incorporate the possibility of contingency occurrences in the planning stage.

Most of existing planning models in the literature incorporate contingencies in a post-outage recovery formulation that identifies an optimal network reconfiguration and promptly restores the system. [84] studies a comprehensive framework for the distribution system in both normal operation and self-healing modes. In the normal operation mode, the objective is to minimize the operation costs. When a contingency happens, the system enters the self-healing mode by sectionalizing the on–outage zone into a set of self-supplied microgrids (MGs) to pick up the maximum amount of loads. [85] develops a systematic framework including planning and operating stages for a smart distribution system. In the planning stage, the goal is to construct self-sufficient MGs using various DGs and storage units. In the operating stage, a new formulation that incorporates both emergency reactions and system restoration is addressed for carrying out optimal self-healing control actions. [98] proposes a graph-theoretic distribution system restoration algorithm to find an optimal network reconfiguration after multiple contingencies arise in the system, where the MGs are modeled as virtual feeders and the distribution system is modeled as a spanning tree. All of the above works are under the premise that the contingencies have already been located and then we perform system reconfiguration to enhance its reliability. In contrast, this study considers the stochasticity of the contingency (e.g., caused by natural disasters).

Existing distribution reinforcement planning models consider stochastic contingencies and carry out pre-event enhancement activities including vegetation management, pole refurbish-

ments, and undergrounding of power lines [6]. [86] presents a two-stage robust optimization model for optimally allocating DG resources and hardening lines before the upcoming natural disasters. A new uncertainty set for contingency occurrences is developed to capture the spatial and temporal dynamic of hurricanes. [87] proposes a new tri-level optimization approach to mitigate the impacts of extreme weather events on the distribution system, with the objective of minimizing hardening investment and the worst-case load shedding cost. An infrastructure fragility model is exploited by considering a time-varying uncertainty set of disastrous events. Even though the above works adopt realistic uncertainty sets for modeling the contingency, challenges still exist for the robust optimization approaches. Indeed, they completely neglect the probabilistic characteristics of the contingency. Accordingly, the robust optimization approaches may only focus on the worst-case contingency and yield over-conservative solutions.

To tackle these challenges, distributionally robust (DR) models have been proposed [65]. The DR models consider a set of probability distributions of the uncertain parameters (termed ambiguity set) using certain statistical characteristics (e.g., moments). Then, we search for a solution that is optimal with respect to the worst-case probability distribution within the ambiguity set. DR models have been applied on various power system problems, such as unit commitment [68], reserve scheduling [99], congestion management [100], and transmission expansion planning [101].

To the best of our knowledge, this chapter conducts the first study of DR models for distribution network configuration when facing contingency. Our main contributions include: (a) by incorporating the contingency probability distribution, our DR model is able to capture the contingencies with lower probability but high impacts, two key features of natural disaster-induced outages; (b) we recast the DR model as a two-stage robust optimization formulation that facilitate the column-and-constraint generation algorithm (see Proposition 2); (c) solving the DR model yields a worst-case contingency distribution, which can be used

(e.g., in simulation models) to examine other topology configuration/re-configuration policies facing random contingency (see Proposition 3); (d) numerical case studies demonstrate the better out-of-sample performance of our DR model.

The remainder of the chapter is organized as follows. In Section 4.3, we describe the DR formulation including the network configuration, the restoration process, and the ambiguity set of contingency probability distribution. In Section 4.4, we derive an equivalent reformulation and employ the column-and-constraint generation framework to solve the problem. In Section 4.5, we conduct case studies and analyze the computational results. Finally, we conclude this Chapter in Section 4.6.

4.2 Nomenclature

A. Sets

$\mathcal{T}$	Set of time periods.
$\mathcal{N}$	Set of nodes.
$\mathcal{E}$	Set of power lines.

B. Parameters

B_y	Available budget for power line constructions.
B_w	Available number (budget) of distributed generators for allocation.
N_z	Maximum number of affected power lines during the contingency.
c_{mn}	Construction cost of line (m, n).
ϕ_{mn}	Resistance of the power line (m, n).
η_{mn}	Reactance of the power line (m, n).
K_{mn}	Upper limit of active power flow in line (m, n).
R_{mn}	Upper limit of reactive power flow in line (m, n).
D^p_{nt}	Active power load at node n in time t.
D^q_{nt}	Reactive power load at node n in time t.

C_n^p Active power capacity of substation or distributed generation unit at node n.

C_n^q Reactive power capacity of substation or distributed generation unit at node n.

ν^{max} Upper bound of voltage.

ν^{min} Lower bound of voltage.

V_0 Reference voltage value.

$\mu_{mn,t}^{max}$ Upper bound of failure rate in line (m,n) in time t.

τ_{mn}^{rst} Minimum restoration time of line (m,n) during contingency.

C. First-stage Decision Variables

y_{mn} Binary variable for network configuration; equals 1 if line (m,n) is constructed, 0 otherwise.

w_n Binary variable; equals 1 if the distributed generation unit is placed at node n, 0 otherwise.

f_{mn} Fictitious flow across line (m,n) for configuring the network.

$\mathbf{g}$ Vector of first stage decision variables including y_{mn}, w_n, and f_{mn}.

$\boldsymbol{\beta}, \gamma$ Dual variables in the reformulation of the distributionally robust model.

D. Second-stage Decision Variables

$p_{mn,t}$ Active power flow across line (m,n) in period t.

$q_{mn,t}$ Reactive power flow across line (m,n) in period t.

x_{nt}^p Active power generation at node n in period t.

x_{nt}^q Reactive power generation at node n in period t.

ν_{nt} Voltage magnitude at node n in period t.

s_{nt} Load shedding at node n in period t.

$\boldsymbol{\pi}$ Dual variables in subproblem reformulation.

$\mathbf{u}$ Vector of second-stage decision variables including $p_{mn,t}$, $q_{mn,t}$, x_{nt}^p, x_{nt}^q, ν_{nt}, and s_{nt}.

E. Random Parameter

$z_{mn,t}$ Bernoulli random variable; equals 0 if line (m, n) is affected in period t, 1 otherwise.

4.3 Mathematical model

We propose a distributionally robust optimization model for a distribution network facing random contingency. The model involves two stages. In the first stage, we form a set of radially configured networks, each energized by a substation within the network. In addition, we allocate a set of available DGs in the system. Then, the contingency launches a set of disruptions to the system to inflict damages. In the second stage, we take restoration actions to minimize the load shedding by rescheduling the output of substations and DGs.

4.3.1 Distribution network configuration

We plan to establish a distribution system in a new community without existing facilities. In this community, only the locations of loads and substations are identified. It is assumed that the substations are connected to a higher-level substation in the grid. Let graph $G = (\mathcal{N}, \mathcal{E})$ represent the distribution network, where $\mathcal{N}$ denotes the set of nodes and $\mathcal{E}$ denotes the set of distribution lines that can be constructed. Also, assume that substations are located in the set $\mathcal{R} \subset \mathcal{N}$. In the devised network configuration, the distribution system consists of a set of radial networks in the sense that each load bus is connected to a substation directly or via other nodes. In other words, we construct a spanning forest with $|\mathcal{R}|$ components, each

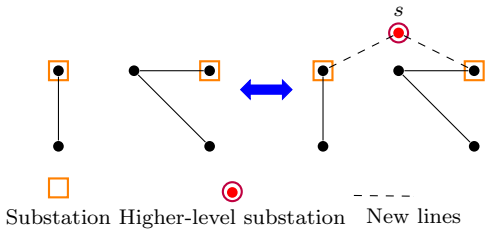

Figure 4.1: Example of a spanning tree representation

rooted at one substation. For this purpose, we add a new higher-level node s to graph G

and connect it to all substation nodes, i.e., nodes in $\mathcal{R}$. We call the new graph $G' = (\mathcal{N}', \mathcal{E}')$. Now constructing a spanning forest rooted in $\mathcal{R}$ is equivalent to constructing a spanning tree of this new graph G', where all newly added lines (i.e., $\mathcal{E}'\backslash\mathcal{E}$) are included in the tree (see Fig. 4.1 for an example). To formulate the spanning tree, we employ the single commodity formulation [102] as follows:

$$\sum_{n|(s,n)\in\mathcal{E}'} f_{sn} = |\mathcal{N}'| - 1, \tag{4.1a}$$

$$\sum_{m|(m,n)\in\mathcal{E}'} f_{mn} - \sum_{m|(n,m)\in\mathcal{E}'} f_{nm} = 1, \quad \forall n \in \mathcal{N}'\backslash s, \tag{4.1b}$$

$$\sum_{(m,n)\in\mathcal{E}'} y_{mn} = |\mathcal{N}'| - 1, \tag{4.1c}$$

$$f_{mn} \leq (|\mathcal{N}'| - 1) y_{mn}, \quad \forall (m,n) \in \mathcal{E}', \tag{4.1d}$$

$$y_{mn} = 1, \quad \forall (m,n) \in \mathcal{E}'\backslash\mathcal{E}, \tag{4.1e}$$

$$f_{mn} \geq 0, \quad y_{mn} \in \{0,1\}, \quad \forall (m,n) \in \mathcal{E}'. \tag{4.1f}$$

We remark that f_{mn} does not represent the power flow along the line (m,n). Instead, it represents fictitious flow to mathematically guarantee that the distribution network is radial. Constraint (4.1a) indicates that there must be $|\mathcal{N}'| - 1$ arcs leaving the root node s in order to form a spanning tree. Constraints (4.1b) ensure the connectivity of the spanning tree. Constraint (4.1c) specifies that, in the constructed spanning tree, the number of connected lines should be one unit less than the number of nodes. Constraints (4.1d) designate that the capacity of fictitious flow on each line should be no more than the total number of connected lines. Constraints (4.1e) indicate that all substations should be connected to the higher-level node s.

Furthermore, we consider the budget constraints on the number of available DG units for

installation and the total construction costs, as stated in (4.1g) and (4.1h), respectively:

$$\sum_{n\in\mathcal{N}} w_n \leq B_w, \tag{4.1g}$$

$$\sum_{(m,n)\in\mathcal{E}} c_{mn} y_{mn} \leq B_y. \tag{4.1h}$$

4.3.2 Post-contingency restoration process

In this study, we model the whole restoration process using a number of corrective actions to minimize the load shedding. We adopt the well-studied linearized approximation of the DistFlow model (see, e.g., [86, 103]) to formulate power flow in the distribution system after the contingency. According to this model, active and reactive power balance flows at each bus are expressed as follows:

$$\sum_{k|(n,k)\in\mathcal{E}} p_{nk,t} = p_{mn,t} - D^p_{nt} + x^p_{nt} + s_{nt},$$
$$\forall n \in \mathcal{N}, \quad \forall (m,n) \in \mathcal{E}, \quad \forall t \in \mathcal{T}, \tag{4.2a}$$

$$\sum_{k|(n,k)\in\mathcal{E}} q_{nk,t} = q_{mn,t} - D^q_{nt} + x^q_{nt},$$
$$\forall n \in \mathcal{N}, \quad \forall (m,n) \in \mathcal{E}, \quad \forall t \in \mathcal{T}. \tag{4.2b}$$

According to the linearized DistFlow model, the relationship of voltage level between any pair of adjacent nodes is characterized by the following constraints:

$$\nu_{nt} y_{mn} = \nu_{mt} y_{mn} - (\phi_{mn} p_{mn,t} + \eta_{mn} q_{mn,t})/V_0,$$
$$\forall m,n \in \mathcal{N} | (m,n) \in \mathcal{E}, \quad \forall t \in \mathcal{T}. \tag{4.2c}$$

Moreover, the voltage level at each node should be within a permissible range:

$$\nu^{min} \leq \nu_{nt} \leq \nu^{max}, \quad \forall n \in \mathcal{N}, \quad \forall t \in \mathcal{T}. \tag{4.2d}$$

Additionally, if line (m, n) is not constructed in the configuration stage or constructed but disrupted during the contingency, the power flow on line (m, n) should be zero. These restrictions are described by the following constraints:

$$0 \leq p_{mn,t} \leq K_{mn} z_{mn,t} y_{mn}, \ \forall (m, n) \in \mathcal{E}, \forall t \in \mathcal{T}, \tag{4.2e}$$

$$0 \leq q_{mn,t} \leq R_{mn} z_{mn,t} y_{mn}, \ \forall (m, n) \in \mathcal{E}, \forall t \in \mathcal{T}. \tag{4.2f}$$

In our proposed framework, each radial network is rooted at a node where the substation is placed. Moreover, the DG units can supply power not only to their neighboring loads but also to all nodes in the connected network. The active and reactive power capacity of the substations and DGs are described by the following constraints:

$$0 \leq x^p_{nt} \leq C^p_n, \quad \forall n \in \mathcal{R}, \quad \forall t \in \mathcal{T}, \tag{4.2g}$$

$$0 \leq x^q_{nt} \leq C^q_n, \quad \forall n \in \mathcal{R}, \quad \forall t \in \mathcal{T}, \tag{4.2h}$$

$$0 \leq x^p_{nt} \leq w_n C^p_n, \quad \forall n \in \mathcal{N} \backslash \mathcal{R}, \quad \forall t \in \mathcal{T}, \tag{4.2i}$$

$$0 \leq x^q_{nt} \leq w_n C^q_n, \quad \forall n \in \mathcal{N} \backslash \mathcal{R}, \quad \forall t \in \mathcal{T}. \tag{4.2j}$$

Finally, the unsatisfied active demand at each node should be no more than the active demand at that node:

$$0 \leq s_{nt} \leq D^p_{nt}, \quad \forall n \in \mathcal{N}, \quad \forall t \in \mathcal{T}. \tag{4.2k}$$

4.3.3 Ambiguity set of contingency

Different approaches have been proposed in the literature to deal with the uncertainty of contingency. Stochastic programming (SP) is well-known for modeling contingency due to natural disasters (see, e.g., [104, 105, 106]). Using statistical methods, SP estimates the joint probability distribution of contingency and then generates a set of scenarios to represent the stochastic contingency in decision making. The major drawback of this approach is that the underlying probability distribution often cannot be estimated accurately, and the computational effort significantly increases as the number of contingency scenarios increases. Robust optimization (RO) is another well-known approach to cope with the uncertainty of contingency (see, e.g., [86, 87]). Applied on the distribution network configuration problem, RO identifies the most critical contingencies by solving the following bilevel model:

$$\max_{\mathbf{z}\in\mathfrak{D}(\mathbf{g})} Q(\mathbf{g},\mathbf{z}) \tag{4.3a}$$

$$\text{s.t. } \mathfrak{D}(\mathbf{g}) = \Bigg\{ \sum_{(m,n)\in\mathcal{E}} (1 - z_{mn,t}) \leq N_z,\ \forall t \in \mathcal{T}, \tag{4.3b}$$

$$1 - z_{mn,t} \leq y_{mn},\ \forall (m,n) \in \mathcal{E},\ \forall t \in \mathcal{T}, \tag{4.3c}$$

$$z_{mn,t+\tau} \leq z_{mn,t},\ \forall (m,n) \in \mathcal{E},\ \forall \tau \leq \tau^{\text{rst}}_{mn} \Bigg\}, \tag{4.3d}$$

where,

$$Q(\mathbf{g},\mathbf{z}) = \min_{\mathbf{u}\in\mathcal{H}(\mathbf{g},\mathbf{z})} \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} s_{nt}, \tag{4.4a}$$

$$\text{s.t. } \mathcal{H}(\mathbf{g},\mathbf{z}) = \Big\{ \mathbf{u} : \text{Constraints (4.2a)-(4.2k)} \Big\}, \tag{4.4b}$$

where, $\mathbf{g} := (\mathbf{y}, \mathbf{w}, \mathbf{f})$ indicates network configuration and DG allocation decision variables, $\mathbf{u} := (\mathbf{p}, \mathbf{q}, \mathbf{x}^p, \mathbf{x}^q, \boldsymbol{\nu}, \mathbf{s})$ denotes the post-contingency decision variables, and $Q(\mathbf{g},\mathbf{z})$ represents the minimum load shedding for given topology $\mathbf{g}$ and contingency $\mathbf{z}$. Moreover, $\mathfrak{D}(\mathbf{g})$ specifies

the set of all possible contingency scenarios. We assume by constraints (4.3b) that the number of simultaneous line outages is bounded by N_z, which can be calibrated based on reliability analyses of distribution lines during contingency (see, e.g., [107]). Constraints (4.3c) designate that only constructed lines can be affected, i.e., $z_{mn,t}$ is set to be one whenever y_{mn} equals zero. However, variables $z_{mn,t}$ only appear in constraints (4.2e)–(4.2f), whose right-hand sides equal zero if $y_{mn} = 0$, regardless of the value of $z_{mn,t}$. Hence, we can relax constraints (4.3c) without loss of optimality. Constraints (4.3d) model the minimum restoration time of failing distribution lines. As discussed in Section 4.1, the RO model may only focus on the worst-case contingency (i.e., $\mathbf{z} \in \mathfrak{D}(\mathbf{g})$ that maximizes $Q(\mathbf{g}, \mathbf{z})$ in (4.3a)) and yield over-conservative topology design and/or DG allocation.

To overcome the challenges of the classical stochastic and robust approaches, we propose a DR framework considering a family of joint probability distributions of contingency based on the moment information of the random parameters (see, e.g., [65, 68, 108]). More specifically, we define the ambiguity set as follows:

$$\mathscr{D} = \left\{ \mathbb{P} \in \mathcal{P}\Big(\mathfrak{D}(\mathbf{g})\Big) : 0 \leq E_{\mathbb{P}}[1 - \mathbf{z}] \leq \boldsymbol{\mu}^{max} \right\}, \tag{4.5}$$

where $\mathcal{P}\Big(\mathfrak{D}(\mathbf{g})\Big)$ consists of all probability distributions on a sigma-field of $\mathfrak{D}(\mathbf{g})$. Constraints (4.5) imply that the marginal probability of each line (m, n) not working during time unit t has an upper limit $\boldsymbol{\mu}^{max}$. We note that, although $\mathscr{D}$ models the contingency of new distribution lines, the distributional information (e.g., $\boldsymbol{\mu}^{max}$ and $\mathfrak{D}(\mathbf{g})$) can be calibrated based on reliability analyses of distribution lines (see, e.g., [107]). Accordingly, we consider the following DR model:

$$\max_{\mathbb{P} \in \mathscr{D}} E_{\mathbb{P}}[Q(\mathbf{g}, \mathbf{z})]. \tag{4.6}$$

Here, instead of considering the worst-case scenario of contingency as in the RO model, we consider the worst-case distribution of contingency and the corresponding expected load shedding. Hence, our approach, though still risk-averse, is less conservative than the RO approach.

4.3.4 Distributionally robust optimization model

Our distributionally robust optimization model aims to find an optimal distribution system configuration to minimize the load shedding under random contingency:

$$\min_{\mathbf{g}\in\mathcal{G}} \max_{\mathbb{P}\in\mathscr{D}} E_{\mathbb{P}}[Q(\mathbf{g},\mathbf{z})], \tag{4.7a}$$

$$\text{s.t.} \quad \mathcal{G} = \Big\{ \mathbf{g} : \text{Constraints (4.1a)-(4.1h)} \Big\}. \tag{4.7b}$$

In above formulation, the objective function (4.7a) aims to minimize the worst-case expected load shedding $Q(\mathbf{g},\mathbf{z})$.

4.4 Solution methodology

In this section, we first derive reformulations of the worst-case expectation model (4.6) and the DR model (4.7a)–(4.7b), respectively. Then, we describe a solution approach based on the column-and-constraint generation (CCG) framework. Finally, we derive the worst-case distribution of contingency.

4.4.1 Problem reformulation

Proposition 2. *For fixed* $\mathbf{g} \in \mathcal{G}$*, we have*

$$\max_{\mathbb{P}\in\mathscr{D}} E_{\mathbb{P}}[Q(\mathbf{g},\mathbf{z})] =$$

$$\min_{\boldsymbol{\beta}\geq 0} \max_{\mathbf{z}\in\mathfrak{D}(\mathbf{g})} \left\{ Q(\mathbf{g},\mathbf{z}) + \sum_{t\in\mathcal{T}} \sum_{(m,n)\in\mathcal{E}} (\mu^{max}_{mn,t} + z_{mn,t} - 1)\beta_{mn,t} \right\},$$

where $\boldsymbol{\beta}$ represent dual variables associated with constraints (4.5).

Proof. We rewrite $\max_{\mathbb{P}\in\mathbb{D}} E_{\mathbb{P}}[Q(\mathbf{g},\mathbf{z})]$ as:

$$\max_{\mathbb{P}\in\mathbb{D}} E_{\mathbb{P}}[Q(\mathbf{g},\mathbf{z})] = \max_{\mathbb{P}} \int_{\mathfrak{D}(\mathbf{g})} Q(\mathbf{g},\mathbf{z}) d\mathbb{P}, \tag{4.8a}$$

$$\text{s.t.} \quad \int_{\mathfrak{D}(\mathbf{g})} d\mathbb{P} = 1, \tag{4.8b}$$

$$\int_{\mathfrak{D}(\mathbf{g})} (1 - z_{mn,t}) d\mathbb{P} \leq \mu^{max}_{mn,t}, \forall (m,n) \in \mathcal{E},\ \forall t \in \mathcal{T}. \tag{4.8c}$$

The feasible region of the problem (4.8a)–(4.8c) has an interior point. In other words, there exists a $\hat{\mathbb{P}}$ that satisfies constraint (4.8b) at equality and constraint (4.8c) strictly. For example, we can set $\hat{\mathbb{P}}$ to be the probability distribution solely supported on the scenario that no contingency arises in the system, i.e., $z_{mn,t} = 1, \forall (m,n) \in \mathcal{E}, t \in \mathcal{T}$. Thus, the Slater's condition holds between the problem (4.8a)–(4.8c) and the following dual formulation:

$$\min_{\boldsymbol{\beta}\geq 0,\gamma} \gamma + \sum_{t\in\mathcal{T}} \sum_{(m,n)\in\mathcal{E}} \mu^{max}_{mn,t} \beta_{mn,t}, \tag{4.9}$$

s.t.

$$\gamma + \sum_{t\in\mathcal{T}} \sum_{(m,n)\in\mathcal{E}} (1 - z_{mn,t})\beta_{mn,t} \geq Q(\mathbf{g},\mathbf{z}),\ \forall \mathbf{z} \in \mathfrak{D}(\mathbf{g}). \tag{4.10}$$

where γ and $\boldsymbol{\beta}$ are dual variables associated with constraints (4.8b) and (4.8c), respectively. In the dual formulation, we observe that the optimal γ should satisfy

$$\gamma = \max_{\mathbf{z}\in\mathfrak{D}(\mathbf{g})} \left\{ Q(\mathbf{g},\mathbf{z}) - \sum_{t\in\mathcal{T}} \sum_{(m,n)\in\mathcal{E}} (1 - z_{mn,t})\beta_{mn,t} \right\}. \tag{4.11}$$

Substituting γ from (4.11) to the objective function (4.9) completes the proof. □

By Proposition 2 and combining two minimizations, we obtain the following equivalent reformulation of formulation (4.7a)–(4.7b):

$$\min_{\boldsymbol{\beta}\geq 0,\mathbf{g}\in\mathcal{G}} \max_{\mathbf{z}\in\mathfrak{D}(\mathbf{g})} \min_{\mathbf{u}\in\mathcal{H}(\mathbf{g},\mathbf{z})} \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} s_{nt} + \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} (\mu^{max}_{mn,t} + z_{mn,t} - 1)\beta_{mn,t}. \tag{4.12}$$

Therefore, the DR model (4.7a)–(4.7b) is transformed into the classical robust optimization problem (4.12).

4.4.2 Column-and-constraint generation framework

We employ the CCG framework [77] to solve the problem (4.12). We describe the master problem in the r^{th} iteration of the CCG framework as follows:

$$\min_{\boldsymbol{\beta}\geq 0,\mathbf{g}\in\mathcal{G},\lambda,\mathbf{u}^j} \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} (\mu^{max}_{mn,t} - 1)\beta_{mn,t} + \lambda \tag{4.13a}$$

$$\text{s.t.} \quad \lambda \geq \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} s^j_{nt} + \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} z^j_{mn,t}\beta_{mn,t},$$

$$\forall \mathbf{z}^j \in \mathcal{F}, \quad \forall j = 1, ..., r, \tag{4.13b}$$

$$\mathbf{u}^j \in \mathcal{H}(\mathbf{g}, \mathbf{z}^j), \quad \forall \mathbf{z}^j \in \mathcal{F}, \quad \forall j = 1, ..., r, \tag{4.13c}$$

where $\mathcal{F} \subseteq \mathfrak{D}(\mathbf{g})$. In the CCG framework, set $\mathcal{F}$ is iteratively augmented by incorporating more scenarios. Note that, the master problem is a relaxation of the original problem, in which the set of contingency $\mathfrak{D}(\mathbf{g})$ consists of all possible scenarios satisfying constraints (4.3b) (note that, as discussed in Section 4.3.3, we have relaxed (4.3c) without loss of optimality). Therefore, solving the master problem (4.13a)–(4.13c) yields a lower bound for that optimal value of (4.12). In contrast, the following subproblem yields an upper bound:

$$\max_{\mathbf{z}\in\mathfrak{D}(\hat{\mathbf{g}})} \min_{\mathbf{u}\in\mathcal{H}(\hat{\mathbf{g}},\mathbf{z})} \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} s_{nt} + \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} \hat{\beta}_{mn,t} z_{mn,t}, \tag{4.14}$$

where decisions $\hat{\mathbf{g}}$ and $\hat{\boldsymbol{\beta}}$ are obtained from solving the master problem (4.13a)–(4.13c). Note that $(\hat{\mathbf{g}}, \hat{\boldsymbol{\beta}})$ is feasible to the problem (4.12). Hence, the optimal objective value of (4.14), plus constant $\sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}}(\mu^{max}_{mn,t}-1)\hat{\beta}_{mn,t}$, is an upper bound for (4.12). Moreover, since the inner minimization problem of (4.14) is always feasible and bounded (a trivial solution is when all loads are shed), we take the dual of this minimization problem with strong duality and convert the bilevel subproblem (4.14) into the following single-level bilinear maximization problem:

$$
\begin{aligned}
&\max_{\mathbf{z}\in\mathfrak{D}(\mathbf{g}),\boldsymbol{\pi}} \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} \hat{\beta}_{mn,t} z_{mn,t} - \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} D^p_{nt}\pi^1_{nt} \\
&- \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} D^q_{nt}\pi^2_{nt} + \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} K_{mn}\pi^3_{mn,t} z_{mn,t} y_{mn} \\
&+ \sum_{t\in\mathcal{T}}\sum_{(m,n)\in\mathcal{E}} R_{mn}\pi^4_{mn,t} z_{mn,t} y_{mn} \\
&+ \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}\setminus\mathcal{R}} w_n C^p_n \pi^5_{nt} + \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}\setminus\mathcal{R}} w_n C^q_n \pi^6_{nt} \\
&+ \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{R}} C^p_n \pi^7_{nt} + \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{R}} C^q_n \pi^8_{nt} \\
&+ \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} v^{max}\pi^{10}_{nt} - \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} v^{min}\pi^{11}_{nt} \\
&+ \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} D^p_{nt}\pi^{12}_{nt} \qquad (4.15a)
\end{aligned}
$$

$$
\begin{aligned}
&\text{s.t. } \pi^3_{mn,t} + \pi^1_{mt} - \pi^1_{nt} + \frac{\phi_{mn}}{V_0}\pi^9_{nt} \le 0, \\
&\qquad \forall m,n\in\mathcal{N} | (m,n)\in\mathcal{E}, \quad \forall t\in\mathcal{T}, \qquad (4.15b) \\
&\qquad \pi^4_{mn,t} + \pi^2_{mt} - \pi^2_{nt} + \frac{\eta_{mn}}{V_0}\pi^9_{nt} \le 0, \\
&\qquad \forall m,n\in\mathcal{N} | (m,n)\in\mathcal{E}, \quad \forall t\in\mathcal{T}, \qquad (4.15c) \\
&\qquad -\pi^1_{nt} + \pi^5_{nt} \le 0, \quad \forall n\in\mathcal{N}\setminus\mathcal{R}, \quad \forall t\in\mathcal{T}, \qquad (4.15d) \\
&\qquad -\pi^2_{nt} + \pi^6_{nt} \le 0, \quad \forall n\in\mathcal{N}\setminus\mathcal{R}, \quad \forall t\in\mathcal{T}, \qquad (4.15e) \\
&\qquad -\pi^1_{nt} + \pi^7_{nt} \le 0, \quad \forall n\in\mathcal{R}, \quad \forall t\in\mathcal{T}, \qquad (4.15f)
\end{aligned}
$$

$$-\pi^2_{nt} + \pi^8_{nt} \leq 0, \quad \forall n \in \mathcal{R}, \quad \forall t \in \mathcal{T}, \tag{4.15g}$$

$$\pi^{10}_{nt} - \pi^{11}_{nt} + \pi^9_{jt} - \sum_{i|(n,i)\in\mathcal{E}} \pi^9_{it} \leq 0,$$
$$j|(j,n) \in \mathcal{E}, \quad \forall n \in \mathcal{N}, \quad \forall t \in \mathcal{T}, \tag{4.15h}$$

$$-\pi^1_{nt} + \pi^{12}_{nt} \leq 1, \quad \forall n \in \mathcal{N}, \quad \forall t \in \mathcal{T}, \tag{4.15i}$$

$$\pi^1_{nt}, \pi^2_{nt}, \pi^9_{nt} \text{ are free and other variables are nonpositive,}$$

where $\boldsymbol{\pi}$ represent dual variables pertaining to constraints (4.2a)–(4.2k). Note that bilinear terms $\boldsymbol{\pi}\mathbf{z}$ in the objective function (4.15a) can be linearized using the McCormick method [109], which recasts the problem (4.15a)–(4.15i) as a mixed-integer linear program and facilitates efficient off-the-shelf solvers like CPLEX. The CCG framework is summarized as follows:

Step 0: Initialization. Pick an optimality gap ϵ. Set LB $= -\infty$, UB $= +\infty$, set of contingencies $\mathcal{F} = \emptyset$, and iteration index $r = 1$.

Step 1: Solve the master problem (4.13a)–(4.13c), obtain the optimal value objMP and optimal configuration decisions $\hat{\mathbf{g}}^r$ and $\hat{\boldsymbol{\beta}}^r$, and update LB = objMP.

Step 2: Solve the subproblem (4.15a)–(4.15i), obtain the optimal value objSP and an optimal contingency scenario $\hat{\mathbf{z}}^r$. Update UB $= \min\{\text{UB}, \text{objSP} + \sum_{t\in\mathcal{T}} \sum_{(m,n)\in\mathcal{E}} (\mu^{max}_{mn,t} - 1)\hat{\beta}_{mn,t}\}$, and $\mathcal{F} = \mathcal{F} \cup \{\hat{\mathbf{z}}^r\}$.

Step 3: If Gap $= (\text{UB} - \text{LB})/\text{LB} \leq \epsilon$, then terminate and output $\hat{\mathbf{g}}^r$ as an optimal solution; otherwise, update $r = r + 1$ and go to the next step.

Step 4: Create second-stage variables $\mathbf{u}^r$ and the corresponding constraints $\mathbf{u}^r \in \mathcal{H}(\mathbf{g}, \hat{\mathbf{z}}^r)$. Add them to the master problem and go to Step 1.

An important by-product of the CCG framework is the worst-case contingency probability distribution, which is formalized in the following proposition.

Proposition 3. *Suppose that the CCG framework terminates at the R^{th} iteration with optimal solutions $(\hat{\beta}^R, \hat{g}^R, \hat{\lambda}^R, \{\hat{u}^j\}_{j=1,\dots,R})$. Then, if we resolve formulation* (4.13a)–(4.13c) *with variables g and u^j fixed at $\hat{g}^R$ and $\hat{u}^j$, respectively, then the dual optimal solutions associated with constraints* (4.13b)*, denoted as $\{\psi_j\}_{j=1,\dots,R}$, characterize the worst-case contingency probability distribution, i.e., $\mathbb{P}\{\mathbf{z} = \mathbf{z}^j\} = \psi_j$, $\forall j = 1, \dots, R$.*

Proof. With variables g and u^j fixed at $\hat{g}^R$ and $\hat{u}^j$, respectively, we take the dual of formulation (4.13a)–(4.13c) to obtain:

$$\max_{\psi \geq 0} \quad \sum_{j=1}^{R} \psi_j \Big(\sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} s_{nt}^j \Big) \tag{4.16a}$$

$$\text{s.t.} \quad \sum_{j=1}^{R} \psi_j (1 - z_{mn,t}^j) \leq \mu_{mn,t}^{max},$$

$$\forall t \in \mathcal{T}, \quad \forall (m,n) \in \mathcal{E}, \tag{4.16b}$$

$$\sum_{j=1}^{R} \psi_j = 1. \tag{4.16c}$$

By constraints (4.16b)–(4.16c), $\{\psi_j\}_{j=1,\dots,R}$ characterize a probability distribution supported on scenarios $\{\mathbf{z}^j\}_{j=1,\dots,R}$ such that $\mathbb{P}\{\mathbf{z} = \mathbf{z}^j\} = \psi_j$, $\forall j = 1, \dots, R$. As the CCG framework terminates at the R^{th} iteration and by the strong duality of linear programming, formulation (4.16a)–(4.16c) is equivalent to the worst-case expectation formulation (4.6), i.e., these two formulations yield the same optimal value. It follows that $\{\psi_j\}_{j=1,\dots,R}$ characterize the worst-case contingency probability distribution. □

4.5 Case study

To evaluate the effectiveness of our approach, we conduct three case studies. In the first study, the distribution network includes 33 nodes, 3 substations, and 2 DG units for allocation. The 3 substations are located at nodes 1, 11, and 25, respectively. In the second study, the

system contains 69 nodes, 4 substations, and 3 DG units for allocation. The substations are located at nodes 1, 13, 39, and 61, respectively. In the third study, the system has 123 nodes, 5 substations, and 5 DG units for allocation. The substations are located at nodes 1, 23, 57, 81, and 100, respectively. The active and reactive power capacities of the DGs are assumed to be 100KW and 50KVar, respectively. In both studies, we consider 24 hours in the post-contingency restoration, i.e., $\mathcal{T} = \{1, \ldots, 24\}$. The active and reactive power loads at each node are randomly generated from intervals $[30, 200]$KW and $[5, 100]$KVars, respectively. The construction costs for distribution lines are randomly generated from intervals proportional to their length. Overall, the construction costs are within the interval $\$[40, 100] \times 10^4$. The contingency status for distribution lines is assumed to follow independent Bernoulli distributions with different failure probabilities that vary within the interval $[0, 0.01]$. Unless stated otherwise, we set the construction budget B_y and the maximum number of affected lines N_z to be $\$1770 \times 10^4$ and 3 for the 33-node system, and $\$4480 \times 10^4$ and 4 for the 69-node system, and $\$8580 \times 10^4$ and 5 for the 123-node system, respectively. All case studies are implemented in C++ with CPLEX 12.6 on a computer with Intel Xeon 3.2 GHz and 8 GB memory.

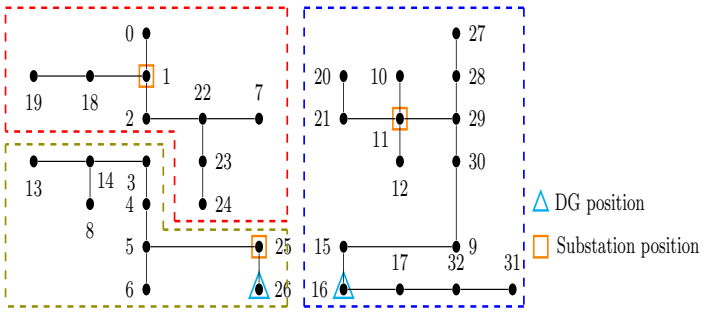

Figure 4.2: Optimal configuration for the 33-node distribution system

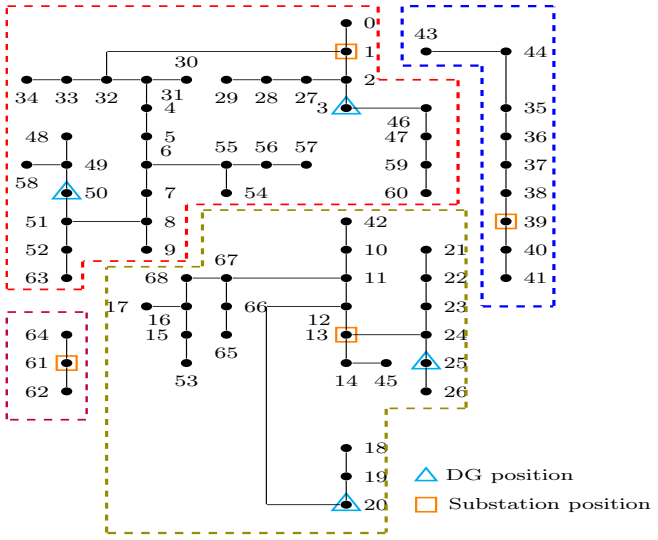

Figure 4.3: Optimal configuration for the 69-node distribution system

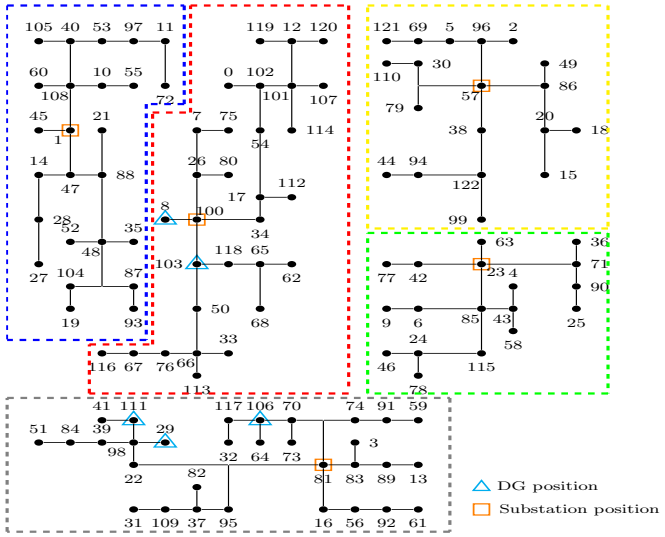

Figure 4.4: Optimal configuration for the 123-node distribution system

4.5.1 Optimal distribution network configuration

We report optimal configurations for the 33-node, the 69-node, and 123-node distribution systems in Figs. 4.2, 4.3, and 4.4, respectively. We also compare our DR model with the RO model. For comparison purposes, we fix configuration decisions obtained by each model and then simulate the load shedding using randomly generated contingencies. Table 4.2

Table 4.2: Comparison of load shedding

	DR model				Robust model			
Nodes	WCD (kW)	WCS (kW)	Sim (kW)	Time (s)	WCD (kW)	WCS (kW)	Sim (kW)	Time (s)
33	1655	2535	1451	91	1921	2352	1648	179
69	4297	5119	3594	173	4570	4998	4014	372
123	9704	10359	8771	293	9943	10185	9370	503

reports the expected load shedding under the worst-case contingency distribution (WCD), the load shedding under the worst-case contingency scenario (WCS), the average load shedding under a randomly simulated contingency distribution within $\mathscr{D}$ (Sim), and the computational time of both models. The results verify that the DR approach yields lower load sheddings under both worst-case distribution and randomly simulated distributions. In particular, our approach leads to more than 11%, 10%, and 6% reduction in average load shedding under the randomly simulated contingency distribution and more than 13%, 5%, and 3% reduction under the worst-case distribution for the 33-node, 69-node, and 123-node distribution systems, respectively. For the worst-case contingency scenario, RO model triggers less load shedding, which was expected, because RO optimizes the system configuration with respect to the worst-case contingency scenario. In addition, the CPU seconds taken to solve the test instances demonstrate the efficacy of the proposed solution approach. To further verify the efficacy, we replicate the experiments on 10 randomly generated instances. For the 33-node system, the average and maximum number of iterations the CCG algorithm takes to converge are 7.6 and 12, respectively; for the 69-node system, the average and maximum number of iterations are 8.4 and 14, respectively; and for the 123-node system the average and maximum number of iterations are 6.4 and 10, respectively.

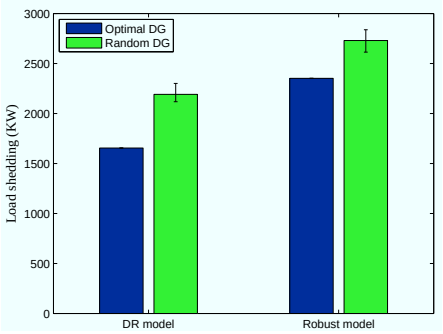

Figure 4.5: Comparisons of optimal and random DG allocation in the 33-node distribution system

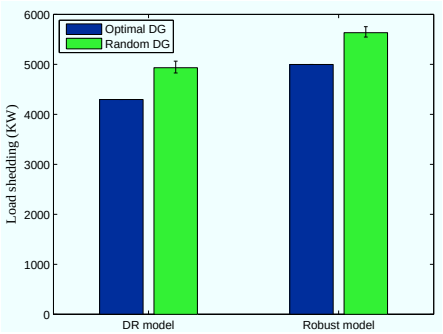

Figure 4.6: Comparisons of optimal and random DG allocation in the 69-node distribution system

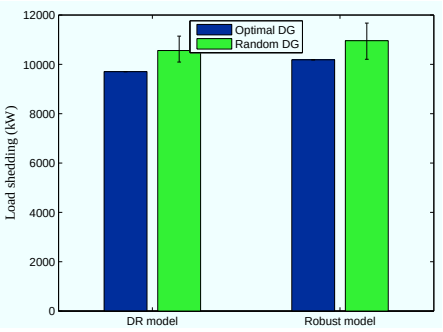

Figure 4.7: Comparisons of optimal and random DG allocation in the 123-node distribution system

4.5.2 On the value of optimal DG allocation

We conduct a set of experiments to evaluate the value of optimally allocating DG units in the distribution system. In Figs. 4.5, 4.6, and 4.7 we compare the level of load shedding when DG units are optimally located with the case when DG units are randomly deployed.

For "optimal DG", we solve the DR and RO models. For "random DG", we first randomly place DGs and then solve both models to configure the distribution system. We perform the experiments for 5 times and report the average values to mitigate the randomness. From Figs. 4.5 and 4.6, we observe that locating DGs properly can significantly decrease the load shedding. This is because when the distribution system is affected by contingencies, the loads in islanded zones can be effectively picked up by the existing DG resources. As a result, better DG allocation significantly enhances the system resiliency.

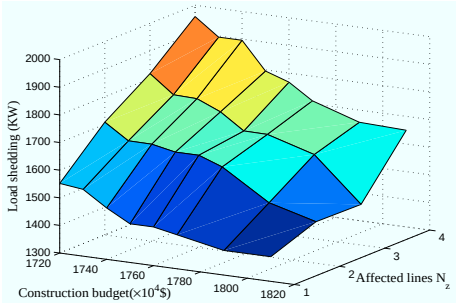

Figure 4.8: Average load shedding under various line construction budget and affected lines for the 33-node distribution system

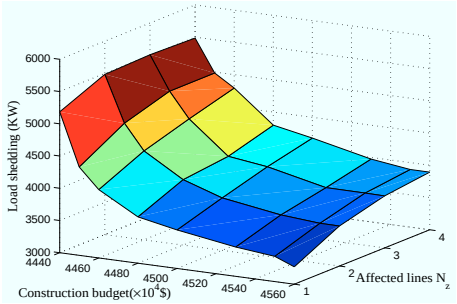

Figure 4.9: Average load shedding under various line construction budget and affected lines for the 69-node distribution system

4.5.3 Impact of construction and contingency budgets

In Figs. 4.8, 4.9, and 4.10, we depict the amounts of expected load shedding under various line construction budgets (i.e., B_y) and contingency budgets (i.e., N_z). From these two figures, we observe that load shedding reduces as B_y increases and as N_z decreases, i.e., as

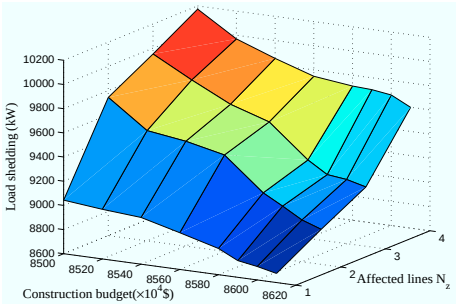

Figure 4.10: Average load shedding under various line construction budget and affected lines for the 123-node distribution system

we allow the contingency to affect less power lines in the DR model. This is intuitive. In addition, we observe that load shedding is sensitive to the construction budget. For example, by increasing the budget from $\$4440\times10^4$ to $\$4480\times10^4$ when $N_z = 4$ in Fig. 4.10, the load shedding decreases from 5485KW to 4297KW, which means that a 0.9% budgetary rise translates into a 21.6% load shedding reduction. Furthermore, we observe that the impact of construction budget is marginally diminishing. For example, increasing the budget from $\$4500\times10^4$ to $\$4560\times10^4$ (i.e., by 1.3%) results in a 6.7% load shedding reduction. This observation highlights the necessity of implementing a cost-effective distribution configuration planning.

4.5.4 Worst-case contingency distribution

The worst-case contingency distribution for the 69-node distribution system is reported in Table 4.3. We select a subset of representative scenarios to display and omit other scenarios with smaller probability values. From this table, we observe that the contingency probabilities for different power lines are highly heterogeneous. This provides the system operator a guideline on the system vulnerability and a meaningful contingency probability distribution that can be used in other vulnerability analyses.

Table 4.3: Worst-case contingency distribution for the 69-node system

Scenario	Affected lines	Probability
1	6-15,13-14,34-35,39-40	0.0031
2	5-26,6-15,12-13,38-39	0.0025
3	1-2,5-26,38-39,39-40	0.0021
4	12-13,13-14,38-39,61-62	0.0002

4.6 Conclusion

In this Chapter, we studied a DRO approach that configures a distribution system considering probabilistic characterizations of the contingencies (e.g., the outage probability of each distribution line). Out-of-sample simulations demonstrated that, comparing with the traditional robust optimization approach, the proposed DRO approach provides less conservative configurations that can reduce post-disaster load shedding. Moreover, the computational results highlighted the values of optimal allocation of DG units and construction budget in the distribution system resiliency. This study can be extended in the following directions: (i) incorporating distributed energy resources (such as solar panels and storage units) into the system via optimized location and sizing, (ii) incorporating network reconfiguration after the contingencies take place, (iii) adopting ac power flow in the post-contingency restoration process, and (iv) considering the contingencies of other components.

CHAPTER V

SUMMARY

Uncertainty is a persistent issue in many power system problems. Renewable energy has maintained increasing penetration into the power grid due to its environment friendly nature and economic benefits; However, many utilities express concerns about integration of wind energy technologies into the power grids due to the variability and intermittency nature of the renewables and their impacts on reliable operation of the grid. Moreover, system components are always exposed to failure due to a variety of reasons like natural disaster occurrences, which are totally unforeseeable with severe potential consequences. Theses uncertainties crucially challenge current practices for the power system operation and management. In this dissertation, several optimization under uncertainty models are developed to better manage the inherent stochasticity in three important problems in the power system area. More specifically, we begin with proposing a multi-stage stochastic programming formulation for the wind power supplier to hedge against the risk of energy shortage. Then, we develop robust optimization and more advanced data-driven models for the aggregated distributed energy resources and distribution network designing problems. In particular, we propose distributionally robust formulations, for which the distribution of the random variable varies within a given ambiguity set. For the ambiguity set construction, we utilize statistical inferences like variance and covariance information. Possible future directions are as follows. Due to the sequential dynamics of decisions in many power system problems (like bidding process in real-time market), a natural way to model this procedure is using Markov Decision Process (MDP). Using the MDP setting, the decision maker (e.g, the virtual power plant

operator) can be considered as a goal-seeking agent that has no knowledge of the external world and interact with the uncertain environment (e.g., ISO and individual DERs) so as to maximize a numerical performance measure. Reinforcement algorithms like Q-learning and SARSA can be employed to solve the resulting MDP.

BIBLIOGRAPHY

[1] "Global wind statistics 2017," http://gwec.net/wp-content/uploads/vip/GWEC_PRstats2017_EN-003_FINAL.pdf, 2017.

[2] "Global market outlook for solar power," http://www.solarpowereurope.org/wp-content/uploads/2018/09/Global-Market-Outlook-2018-2022.pdf, 2018.

[3] G. Chalkiadakis, V. Robu, R. Kota, A. Rogers, and N. R. Jennings, "Cooperatives of distributed energy resources for efficient virtual power plants," in *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, 2011, pp. 787–794.

[4] D. Pudjianto, C. K. Gan, V. Stanojevic, M. Aunedi, P. Djapic, and G. Strbac, "Value of integrating distributed energy resources in the UK electricity system," in *IEEE Power and Energy Society General Meeting*, 2010, pp. 1–6.

[5] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Renewable Power Generation*, vol. 1, no. 1, pp. 10–16, 2007.

[6] A. M. Salman, Y. Li, and M. G. Stewart, "Evaluating system reliability and targeted hardening strategies of power distribution systems subjected to hurricanes," *Reliability Engineering & System Safety*, vol. 144, pp. 319–333, 2015.

[7] L. Che, M. Khodayar, and M. Shahidehpour, "Only connect: Microgrids for distribution system restoration," *IEEE Power and Energy Magazine*, vol. 12, no. 1, pp. 70–81, 2014.

[8] E. O. of the President, "Economic benefits of increasing electric grid resilience to weather outages-august 2013."

[9] D. T. Ton and W. P. Wang, "A more resilient grid: The us department of energy joins with stakeholders in an R&D plan," *IEEE Power and Energy Magazine*, vol. 13, no. 3, pp. 26–34, 2015.

[10] "President's Council of Economic Advisers and the U.S. Department of Energy, Economic Benefits of Increasing Electric Grid Resilience to Weather Outages," Tech. Rep., 2013.

[11] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

[12] T. Mai, D. Sandor, R. Wiser, and T. Schneider, "Renewable electricity futures study. executive summary," National Renewable Energy Laboratory (NREL), Golden, CO., Tech. Rep., 2012.

[13] MISO, "Energy and operating reserves," https://www.misoenergy.org/Training/MarketParticipantTraining/Pages/200LevelTraining.aspx,2015., 2015.

[14] J. Garcia-Gonzalez, R. M. R. de la Muela, L. M. Santos, and A. M. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 460–468, 2008.

[15] Q. Jiang and H. Wang, "Two-time-scale coordination control for a battery energy storage system to mitigate wind power fluctuations," *IEEE Transactions on Energy Conversion*, vol. 28, no. 1, pp. 52–61, 2013.

[16] H. Daneshi and A. Srivastava, "Security-constrained unit commitment with wind generation and compressed air energy storage," *IET Generation, Transmission & Distribution*, vol. 6, no. 2, pp. 167–175, 2012.

[17] D. J. Maggio, "Impacts of wind-powered generation resource integration on prices in the ercot nodal market," in *IEEE Power and Energy Society General Meeting*, 2012, pp. 1–4.

[18] V. S. Pappala, I. Erlich, K. Rohrig, and J. Dobschinski, "A stochastic model for the optimal operation of a wind-thermal power system," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 940–950, 2009.

[19] A. T. Al-Awami and M. A. El-Sharkawi, "Coordinated trading of wind and thermal energy," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 277–287, 2011.

[20] S. Kamalinia and M. Shahidehpour, "Generation expansion planning in wind-thermal power systems," *IET generation, transmission & distribution*, vol. 4, no. 8, pp. 940–951, 2010.

[21] Y. Zhou, T. Liu, and C. Zhao, "Backup capacity coordination with renewable energy certificates in a regional electricity market," *IISE Transactions*, vol. 50, no. 8, pp. 711–719, 2018.

[22] T. Dai and W. Qiao, "Trading wind power in a competitive electricity market using stochastic programing and game theory," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 805–815, 2013.

[23] "Wartsila to deliver 203 MWe Gas Power Plant Near San Antonio, Texas to South Texas Electric Cooperative." [Online]. Available: http://www.wartsila.com/en/gas-power-plant-to-south-texas-electric-cooperative.

[24] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Short-term trading for a wind power producer," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 554–564, 2010.

[25] H. M. I. Pousinho, V. M. F. Mendes, and J. P. d. S. Catalão, "A risk-averse optimization model for trading wind energy in a market environment under uncertainty," *Energy*, vol. 36, no. 8, pp. 4935–4942, 2011.

[26] M. E. Khodayar and M. Shahidehpour, "Stochastic price-based coordination of intrahour wind energy and storage in a generation company," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 554–562, 2013.

[27] A. A. S. de la Nieta, J. Contreras, and J. I. Munoz, "Optimal coordinated wind-hydro bidding strategies in day-ahead markets," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 798–809, 2013.

[28] L. Baringo and A. J. Conejo, "Offering strategy of wind-power producer: A multi-stage risk-constrained approach," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1420–1429, 2016.

[29] S. Babaei, C. Zhao, and T. Liu, "Incentive-based coordination mechanism for backup renewable energy investment," in *IEEE Power Engineering Society General Meeting*, Chicago, 2017, pp. 1–5.

[30] R. M. Kovacevic and G. C. Pflug, "Electricity swing option pricing by stochastic bilevel optimization: a survey and new approaches," *European Journal of Operational Research*, vol. 237, no. 2, pp. 389–403, 2014.

[31] T. Kluge, "Pricing swing options and other electricity derivatives," Ph.D. dissertation, University of Oxford, 2006.

[32] F. Cheng, M. Ettl, G. Y. Lin, M. Schwarz, and D. D. Yao, "Flexible supply contracts via options," *IBM TJ Watson Research Center Working Paper*, 2003.

[33] S.-J. Deng and S. S. Oren, "Electricity derivatives and risk management," *Energy*, vol. 31, no. 6-7, pp. 940–953, 2006.

[34] H. Y. Yamin and S. M. Shahidehpour, "Risk and profit in self-scheduling for gencos," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 2104–2106, 2004.

[35] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[36] H. Wu, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "A game theoretic approach to risk-based optimal bidding strategies for electric vehicle aggregators in electricity markets with variable wind energy resources," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 374–385, 2016.

[37] Q. Xu, N. Zhang, C. Kang, Q. Xia, D. He, C. Liu, Y. Huang, L. Cheng, and J. Bai, "A game theoretical pricing mechanism for multi-area spinning reserve trading considering wind power uncertainty," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1084–1095, 2016.

[38] D. Chattopadhyay and T. Alpcan, "A game-theoretic analysis of wind generation variability on electricity markets," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2069–2077, 2014.

[39] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. Siam, 1999, vol. 23.

[40] A. Botterud, Z. Zhou, J. Wang, R. J. Bessa, H. Keko, J. Sumaili, and V. Miranda, "Wind power trading under uncertainty in lmp markets," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 894–903, 2012.

[41] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1319–1327, 2008.

[42] H. Heitsch and W. Römisch, "Scenario tree modeling for multistage stochastic programs," *Mathematical Programming*, vol. 118, no. 2, pp. 371–406, 2009.

[43] S. Burger, J. P. Chaves-Ávila, C. Batlle, and I. J. Pérez-Arriaga, "A review of the value of aggregators in electricity systems," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 395–405, 2017.

[44] H. Nosair and F. Bouffard, "Energy-centric flexibility management in power systems," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5071–5081, 2016.

[45] B. Willems, "Physical and financial virtual power plants." [Online]. Available: https://ssrn.com/abstract=808944

[46] "Reducing the cost of system intermittency using demand side control measures," 2006.

[47] G. Parkinson, "Tesla builds case for 250mw virtual power plant after first trial success," https://reneweconomy.com.au/tesla-builds-case-for-250mw-virtual-power-plant-after-first-trial-success-63950/, 2018.

[48] D. Pudjianto, G. Strbac, and D. Boyer, "Virtual power plant: managing synergies and conflicts between transmission system operator and distribution system operator control objectives," *CIRED-Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 2049–2052, 2017.

[49] M. Oates, A. Melia, V. Ferrando *et al.*, "Energy balancing accross cities: Virtual power plant prototype and iurban case studies," *Entrepreneurship and Sustainability Issues*, vol. 4, no. 3, pp. 351–363, 2017.

[50] H. Pandžić, I. Kuzle, and T. Capuder, "Virtual power plant mid-term dispatch optimization," *Applied Energy*, vol. 101, pp. 134–141, 2013.

[51] R. M. Lima, A. Q. Novais, and A. J. Conejo, "Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. an adaptive robust optimization approach," *European Journal of Operational Research*, vol. 240, no. 2, pp. 457–475, 2015.

[52] E. Mashhour and S. M. Moghaddas-Tafreshi, "Bidding strategy of virtual power plant for participating in energy and spinning reserve markets–part I : Problem formulation," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 949–956, 2011.

[53] M. Peik-herfeh, H. Seifi, and M. K. Sheikh-El-Eslami, "Two-stage approach for optimal dispatch of distributed energy resources in distribution networks considering virtual power plant concept," *International Transactions on Electrical Energy Systems*, vol. 24, no. 1, pp. 43–63, 2014.

[54] H. Ding, P. Pinson, Z. Hu, J. Wang, and Y. Song, "Optimal offering and operating strategy for a large wind-storage system as a price maker," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4904–4913, 2017.

[55] E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Optimal offering strategy of a virtual power plant: A stochastic bi-level approach," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 794–806, 2016.

[56] L. Ju, Z. Tan, J. Yuan, Q. Tan, H. Li, and F. Dong, "A bi-level stochastic scheduling optimization model for a virtual power plant connected to a wind–photovoltaic–energy storage system considering the uncertainty and demand response," *Applied Energy*, vol. 171, pp. 184–199, 2016.

[57] P. Moutis, P. S. Georgilakis, and N. D. Hatziargyriou, "Voltage regulation support along a distribution line by a virtual power plant based on a center of mass load modeling," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3029–3038, 2018.

[58] D. Koraki and K. Strunz, "Wind and solar power integration in electricity markets and distribution networks through service-centric virtual power plants," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 473–485, 2018.

[59] S. M. Nosratabadi, R.-A. Hooshmand, and E. Gholipour, "A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 341–363, 2017.

[60] R. M. Lima, A. J. Conejo, S. Langodan, I. Hoteit, and O. M. Knio, "Risk-averse formulations and methods for a virtual power plant," *Computers & Operations Research*, vol. 96, no. 2, pp. 350–373, 2018.

[61] S. R. Dabbagh and M. K. Sheikh-El-Eslami, "Risk assessment of virtual power plants offering in energy and reserve markets," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3572–3582, 2016.

[62] H. Pandžić, J. M. Morales, A. J. Conejo, and I. Kuzle, "Offering model for a virtual power plant based on stochastic programming," *Applied Energy*, vol. 105, pp. 282–292, 2013.

[63] M. Rahimiyan and L. Baringo, "Strategic bidding for a virtual power plant in the day-ahead and real-time markets: A price-taker robust optimization approach," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2676–2687, 2016.

[64] M. Shabanzadeh, M.-K. Sheikh-El-Eslami, and M.-R. Haghifam, "The design of a

risk-hedging tool for virtual power plants via robust optimization approach," *Applied Energy*, vol. 155, pp. 766–777, 2015.

[65] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations research*, vol. 58, no. 3, pp. 595–612, 2010.

[66] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.

[67] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 94–102, 2018.

[68] P. Xiong, P. Jirutitijaroen, and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 39–49, 2017.

[69] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 289–300, 2016.

[70] Y. Zhang, S. Shen, and J. L. Mathieu, "Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1378–1388, 2017.

[71] F. Alismail, P. Xiong, and C. Singh, "Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 536–544, 2018.

[72] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Siam, 2001, vol. 2.

[73] D. Bertsimas, M. Sim, and M. Zhang, "Adaptive distributionally robust optimization," *Management Science*, 2018.

[74] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, no. 2, pp. 351–376, 2004.

[75] A. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.

[76] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1678–1688, 2017.

[77] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

[78] A. Lorca and X. A. Sun, "The adaptive robust multi-period alternating current optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1993–2003, 2018.

[79] H. Konno, "A cutting plane algorithm for solving bilinear programs," *Mathematical Programming*, vol. 11, no. 1, pp. 14–27, 1976.

[80] PJM. [Online]. Available: http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx

[81] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "Arima-based time series model

of stochastic wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 667–676, 2010.

[82] J. Moreira, E. Miguez, C. Vilacha, and A. F. Otero, "Large-scale network layout optimization for radial distribution networks by parallel computing," *IEEE transactions on Power Delivery*, vol. 26, no. 3, pp. 1946–1951, 2011.

[83] D. Kumar and S. Samantaray, "Design of an advanced electric power distribution systems using seeker optimization algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 196–217, 2014.

[84] Z. Wang and J. Wang, "Self-healing resilient distribution systems based on sectionalization into microgrids," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3139–3149, 2015.

[85] S. A. Arefifar, Y. A.-R. I. Mohamed, and T. H. EL-Fouly, "Comprehensive operational planning framework for self-healing control actions in smart distribution grids," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4192–4200, 2013.

[86] W. Yuan, J. Wang, F. Qiu, C. Chen, C. Kang, and B. Zeng, "Robust optimization-based resilient distribution network planning against natural disasters," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2817–2826, 2016.

[87] S. Ma, B. Chen, and Z. Wang, "Resilience enhancement strategy for distribution systems under extreme weather events," *IEEE Transactions on Smart Grid*, 2016.

[88] A. Samui, S. Singh, T. Ghose, and S. Samantaray, "A direct approach to optimal feeder routing for radial distribution system," *IEEE Transactions on Power Delivery*, vol. 27, no. 1, pp. 253–260, 2012.

[89] A. Navarro and H. Rudnick, "Large-scale distribution planning–part I: Simultaneous network and transformer optimization," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 744–751, 2009.

[90] K. Mahmoud, N. Yorino, and A. Ahmed, "Optimal distributed generation allocation in distribution systems for loss minimization," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 960–969, 2016.

[91] B. R. Pereira, G. R. M. da Costa, J. Contreras, and J. R. S. Mantovani, "Optimal distributed generation and reactive power allocation in electrical distribution systems," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 975–984, 2016.

[92] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 360–370, 2010.

[93] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2541–2551, 2011.

[94] E. Míguez, J. Cidrás, E. Díaz-Dorado, and J. L. García-Dornelas, "An improved branch-exchange algorithm for large-scale distribution network planning," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 931–936, 2002.

[95] A. Y. Abdelaziz, F. Mohamed, S. Mekhamer, and M. Badr, "Distribution system reconfiguration using a modified tabu search algorithm," *Electric Power Systems Research*, vol. 80, no. 8, pp. 943–953, 2010.

[96] M. Lavorato, J. F. Franco, M. J. Rider, and R. Romero, "Imposing radiality constraints

in distribution system optimization problems," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 172–180, 2012.

[97] S. A. El Batawy and W. G. Morsi, "Optimal secondary distribution system design considering rooftop solar photovoltaics," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1662–1671, 2016.

[98] J. Li, X.-Y. Ma, C.-C. Liu, and K. P. Schneider, "Distribution system restoration with microgrids using spanning tree search," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 3021–3029, 2014.

[99] Q. Bian, H. Xin, Z. Wang, D. Gan, and K. P. Wong, "Distributionally robust solution to the reserve scheduling problem with partial information of wind power," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2822–2823, 2015.

[100] F. Qiu and J. Wang, "Distributionally robust congestion management with dynamic line ratings," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2198–2199, 2015.

[101] A. Bagheri, J. Wang, and C. Zhao, "Data-driven stochastic transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3461–3470, 2017.

[102] T. L. Magnanti and L. A. Wolsey, "Optimal trees," *Handbooks in operations research and management science*, vol. 7, pp. 503–615, 1995.

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

[104] H. Gao, Y. Chen, S. Mei, S. Huang, and Y. Xu, "Resilience-oriented pre-hurricane

resource allocation in distribution systems considering electric buses," *Proceedings of the IEEE*, 2017.

[105] A. Arab, A. Khodaei, S. K. Khator, K. Ding, V. A. Emesih, and Z. Han, "Stochastic pre-hurricane restoration planning for electric power systems infrastructure," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 1046–1054, 2015.

[106] E. Yamangil, R. Bent, and S. Backhaus, "Designing resilient electrical distribution grids," *arXiv preprint arXiv:1409.4477*, 2014.

[107] Y. Sa, "Reliability analysis of electric distribution lines," Ph.D. dissertation, McGill University, 2002.

[108] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 94–102, 2018.

[109] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part iconvex underestimating problems," *Mathematical Programming*, vol. 10, no. 1, pp. 147–175, 1976.

[110] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1-3, pp. 193–228, 1998.

APPENDIX A

Detailed explanation on deriving model (3.13)

Consider the following second-order conic program:

$$\begin{aligned} &\min f^T x \\ &s.t. \;\; ||A_i x + b_i|| \leq c_i^T + d_i, \quad i = 1, \ldots, N, \end{aligned}$$

where $x \in \mathbb{R}^n$ is the optimization variable, and the problem parameters are $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{(n_i-1)n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, and the norm we use in the constraints is the Euclidean norm, i.e., $||y|| = \sqrt{y^T y}$. The dual of this problem is as follows [110]:

$$\begin{aligned} &\max -\sum_{i=1}^{N} (b_i^T z_i + d_i w_i) \\ &s.t. \;\; \sum_{i=1}^{N} (A_i^T z_i + c_i w_i) = f, \\ &||z_i|| \leq w_i, \quad i = 1, \ldots, N, \end{aligned}$$

where $z_i \in \mathbb{R}^{n_i-1}$ and $w \in \mathbb{R}^N$ are dual optimization variables. By applying this primal-dual relationship, we can get the dual formulation of (3.12). In order to apply it, it suffices that we reformulate constraints (3.7) as a set of second-order conic constraints. Consider the following constraint:

$$(\boldsymbol{\xi} - \boldsymbol{\mu})^2 \leq \mathbf{u}$$

We can rewrite it as follow:

$$\sqrt{(\boldsymbol{\xi}-\boldsymbol{\mu})^2+\left(\frac{\mathbf{u}-1}{2}\right)^2} \leq \frac{\mathbf{u}+1}{2},$$

or equivalently:

$$\left|\left|\begin{pmatrix}\boldsymbol{\xi}-\boldsymbol{\mu}\\ \frac{\mathbf{u}-1}{2}\end{pmatrix}\right|\right| \leq \frac{\mathbf{u}+1}{2},$$

which is a second-order conic constraint. We can apply the same technique for the other quadratic constraints in (3.7) and get the dual formulation of (3.12).

VITA

Sadra Babaei

Candidate for the Degree of

Doctor of Philosophy

Dissertation: OPTIMIZATION UNDER UNCERTAINTY MODELS IN POWER SYSTEMS OPERATIONS

Major Field: Industrial Engineering and Management

Biographical:

Personal Data: Born in Hamedan, Iran.

Education:

- Completed the requirements for Doctor of Philosophy in Industrial Engineering and Management at Oklahoma State University, Stillwater, Oklahoma in July, 2019.
- Received the M.S. degree from Amirkabir University of Technology in Industrial Engineering, Iran, Tehran, 2012
- Received the B.S. degree from Iran University of Science and Technology in Industrial Engineering, Iran, Tehran, 2009

Selected Publications:

- S. Babaei, C. Zhao, and L. Fan, "A Data-Driven Model of Virtual Power Plants in Day-Ahead Unit Commitment", Accepted for publication in IEEE Transactions on Power Systems, 2019
- S. Babaei, C. Zhao, L. Fan, and T. Liu, "Incentive-Based Coordination Mechanism for Renewable and Conventional Energy Suppliers", IEEE Transactions on Power Systems, vol. 34, no. 3., pp. 1761-1770, 2018
- S. Babaei, R. Jiang, and C. Zhao, "Distributionally Robust Distribution Network Configuration under Random Contingency", submitted to IEEE Transactions on Power Systems