UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

MATHEMATICAL SCIENCES

Stochastic and Robust Models for Optimal Decision Making in Energy

Arash Mostajeran Gourtani

Thesis for the degree of Doctor of Philosophy

September 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES
MATHEMATICAL SCIENCES

Doctor of Philosophy

STOCHASTIC AND ROBUST MODELS FOR OPTIMAL DECISION MAKING IN ENERGY

by Arash Mostajeran Gourtani

The focus of the work in this thesis is to develop stochastic and robust optimisation models for decision making problems in the energy industry, most notably for (i) the medium-term trading strategy of a dominant producer in the electricity market, (ii) the short-term unit commitment in power generation, and (iii) the long-term facility location problem. These models support decision makers in incorporating the future uncertainty and, in the case of the first model, the market competition into their investment, planning, and operational decisions.

The methodological contribution of the thesis offers several novel insights into the above decision problems. In the context of the medium-term trading strategy of a dominant producer, a multi-objective two-stage bilevel stochastic model is proposed in which the dominant producer aims at maximizing the expected market share and the expected profits in a pool-based market. The model is reformulated first as a multi-objective stochastic mathematical program with equilibrium constraints and then as a mixed-integer linear programming problem. Numerical test results are reported through a medium size case study based on the Italian electricity market. The analysis of the Pareto frontier solution illustrates the trade-off between the producer's conflicting interest in maximizing the expected profit and the expected market share. It was also concluded that the dominant producer can substantially increase expected profits and/or expected market share by behaving strategically when offering power production to the market.

In the unit commitment problem, a two-stage stochastic and distributional robust model is proposed to deal with day ahead wind uncertainty. The robust problem is formulated

using two proposed uncertainty sets; the first one is based on a mixture of distributions and the second one is based on the first order moment approach. Both robust models are then reformulated as semi-infinite programs and solved as mixed integer linear programs using sampling methods. Some numerical results are presented and the results conclude that, although the robust solutions may lose the potential of utilizing the wind power in high wind climate, they perform much better in a low wind climate as compared to the two-stage non-robust stochastic solutions that do not consider the uncertainty of the distribution.

Finally, in the facility location problem with stochastic demand, a two-stage distributionally robust model is proposed to tackle the issue of incomplete information on the true distribution of the uncertainty. The uncertainty set is constructed using the moment information associated with the distribution of the random demands. Two numerical methods are proposed based on the available moment information. Specifically, we first formulate the robust problem as a semi-infinite program for the case that only the first moment information is given. The semi-infinite program is then solved by approximation using a linear decision rule, CVaR and Monte Carlo sampling. In the second method, we formulate the robust problem as a semi-definite program on the basis of the first and the second moments which is then solved by using a constraint generation algorithm. Numerical results suggest that the distributionally robust solutions offer the flexibility in hedging against uncertainty compared to the deterministic and the stochastic non-robust solutions.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Arash Mostajeran Gourtani, declare that the thesis entitled *Stochastic and Robust Models for Optimal Decision Making in Energy* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I con rm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other quali cation at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: Gourtani et al.[60]

Signed:...

Date:..

Acknowledgements

Firstly, I wish to thank my supervisors Prof. Huifu Xu and Dr. Tri-Dung Nguyen for their continuous support, patience and encouragement throughout my research. Their invaluable advice contributed greatly to the accomplishment of this work.

I am also very grateful to my good friend and colleague Dr. David Pozo for his encouragement and his input in numerous discussions regarding the technical aspects of this project. I would also like to thank all my colleagues at School of Mathematics for their support. In particular, Edwin deserves to be explicitly mentioned for sharing enumerable intellectually stimulating discussions and entraining debates, and much-needed co ee breaks.

My special thanks goes to Anca, for sharing all the good and the bad times during my PhD and for being reassuring and supportive in the latter ones.

Finally, a wholehearted thanks goes to my family and all the friends whom I feel very fortunate to have and without whose support I would never have been able to make it through my PhD.

I would like to dedicate this thesis to my parents. Without their support and belief in me I would never have had a chance to undertake this work let alone complete it.

Chapter 1

Introduction

1.1 Optimal decision making in energy

This thesis aims to analyse some of the challenges brought by the presence of uncertainty in decision making and planning problems in energy systems. The sector is often characterized by large investments in assets with long life cycles, volatile market prices, and the need for climate change mitigation. In addition, energy markets have been deregulated over the last decades. As a result, consumers, producers, and traders, as well as system operators and regulators must make decisions whose outcomes are heavily influenced by uncertain factors. In this chapter, we set to discuss some of the main sources of uncertainty in energy systems that include, but are not limited to the following:

Policy and market uncertainty
Global concerns over climate change and energy security as well as the rising importance of the role of energy as a key driver of economic growth have resulted in the ever increasing intervention of policy makers in energy markets. Consequently, governments and international institutions use a wide range of regulatory interventions to control the energy sector performance. These interventions include price controls, competition and market access rules, restrictive trade practice controls, and technical and environmental performance management.

For instance, recognizing the importance of competition in energy markets, many countries over the past three decades have implemented infrastructure reforms such as deregulation and privatization to improve efficiency in the market. Likewise, in order to tackle global warming, green policies have been introduced to force the energy sector to look for new energy production methods that place a smaller burden on the environment. For example, the European Union has already introduced climate objectives to reduce CO2 emissions in the EU by 20 percent from the 1990 level by 2020. The EU is also committed to increase the share of renewable energies to 20 percent of the total energy consumption.

Thus, the full or partial implementation of such policies, and more importantly, the e ectiveness of such interventions could result in greater levels of uncertainty in both long-term and short-term decisions.

Demand and supply uncertainty

According to the International Energy Agency, world energy demand could be doubled by 2050 (compared with that in 2009). The economic climate can have a signi cant impact on the energy demand as evidenced by the wide swings in demand and prices brought about by the 2008 global recession. Another major driver of energy demand is technological advancement. For example, the e ective adoption of smart grid, the improvement in storage technology, the integration of electric vehicles and the deployment of demand response may have signi cant implications for demand patterns. Thus, the trends in economics, geopolitical circumstances, and technological breakthroughs are major sources of uncertainty in energy demand.

On the other hand, the energy supply projections are demand-driven. Therefore, factors that will have an impact on the demand side will also impact the supply side. Technological developments, new policies, and changing prospects of fuel supply (e.g., recent shale gas discoveries) and fuel prices may in uence the choice of generation options and the generation mix in the future. In some cases, social and local acceptability of energy infrastructure projects is an important factor as well. For instance, due to the implementation of the green policy and of nancial incentives such as feed-in-tari s, the deployment of renewable resources has expanded rapidly in recent years. However, the inherent variability of renewable energies coupled with the lack of e cient storage facilities has an adverse e ect on the reliability and cost-e ectiveness of generation output.

System reliability

In order to protect energy systems against unexpected events such as disturbances, contingencies, attacks, and natural disasters, adequate security policies and practices need to be designed and implemented. Consequently, in the long term, sufficient and well-located investment is needed to maintain the power system adequacy. In the short-term, access to reliable supplies and their efficient use are required to ensure that generation equipment operates reliably and predictably.

Decision making horizon in energy systems

The variety of models used in energy system planning can be classified according to the planning horizon. Short-term planning typically deals with problems with horizons of one week or shorter, such as unit commitment and economic dispatch. Medium-term planning decisions cover a 1 to 3 years period and could be interpreted as a more strategic vision for short-term problems. The generation scheduling and strategic market participation could be classified as medium-term problems. Long-term planning is often involved with investment decision problems spanning more than 10 years in the future. Capacity expansion and facility location are some examples of such long-term planning problems. The uncertainties and risk factors associated with the planning problems could also be classified by the time horizon. For example, long-term investment decisions should account for unforeseeable parameters such as economic growth, political climate and technological advancement. On the other end of the spectrum, short-term operational decisions are subject to risk exposure due to demand and supply uncertainties. The time horizon for some of the planning decisions in energy systems and their associated uncertainties are illustrated in Figure 1.1.

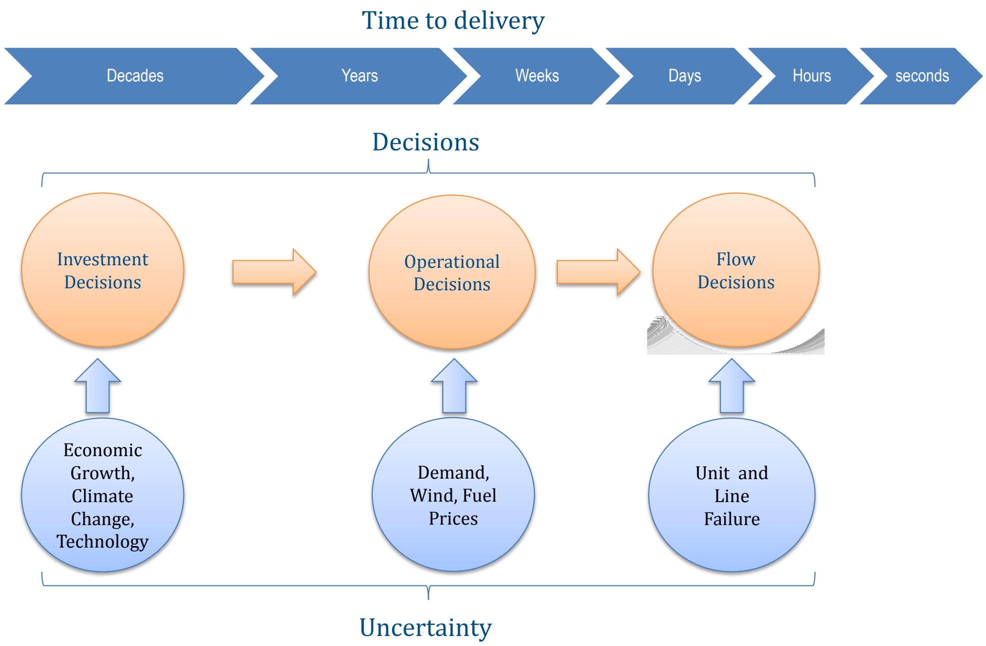


Figure 1.1: Decision making horizon in energy systems

In this thesis, we investigate three separate but closely related problems in the context of energy applications. Speci cally, we develop a game theoretic model for the medium-term generation scheduling problem with exogenous stochastic demand, a short-term unit commitment problem with exogenous stochastic wind supply, and a long-term facility location problem under future demand uncertainty. Each problem is presented in its own, free standing chapter. In the remainder of this section, a summary of each problem is provided.

1.1.1 Medium-term energy production planning problem

Over the past three decades, electricity industry has undergone considerable structural changes as governments have worked to promote competition, reliability and fair prices for consumers. In many countries, deregulation has resulted in the formation of a centralized dispatch and pricing mechanism, called the electricity pool. The day-ahead (spot) trading of the electricity in pool-based markets is carried out using a sealed-bid auction. In a pool-based auction, an independent system operator (ISO) processes bids from generators and retailers, typically on hourly basis, and determines the market price and power dispatch on the basis of a social bene t maximization framework. The ISO clears the market by finding the equilibrium clearing price of the auction based on the submitted bids. All subsequent trades are settled at this price. When the transaction of power is carried out in zones of a transmission network, the netting of imbalances determines zonal prices, power dispatch and power flows between zones.

Despite efforts to increase competitiveness in electricity market, the presence of producers with market power, in some cases, results in imperfect competition. In this work, we consider an electricity market where a single generator has a dominant position and competes with a number of smaller price taking producers. Moreover, we consider the market to consist of different zones interconnected by capacitated transmission lines. We study the medium-term (e.g., one year ahead) trading strategy of the dominant producer in the day-ahead market aiming to maximize its expected market share and expected profit simultaneously. Since the uncertainty of electricity demand becomes more significant over mid-term time horizon, it is sensible to consider the demand as stochastic.

In order to model this problem, we propose a multi-objective two-stage bilevel stochastic programming framework. At the first stage and upper level, the dominant producer aims at maximizing its expected market share and profit. At the second stage and lower level, the ISO determines the dispatches and power flows on an hourly basis, after the realization of uncertainty in market demand, by solving an optimization problem which aims at maximizing the total social welfare. Figure 1.2 illustrates the structure of the proposed model.

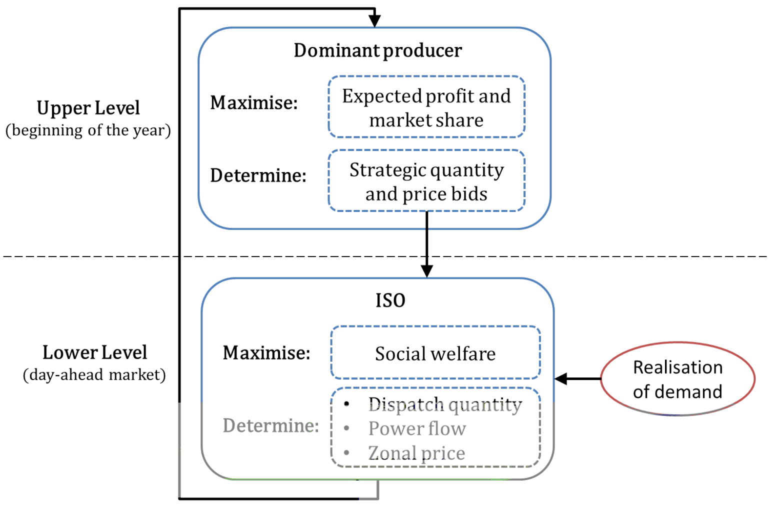


Figure 1.2: Bilevel model structure for medium term trading strategy of the dominant producer

The dominant producer can use a number of alternative bidding strategies to excercise its market power. We used three of the most common strategies to implement the model:

- making a strategic bid on power generation quantities as in the *Cournot model* by retaining some generation capacity.
- making a strategic price bid as in the *Bertrand model* by setting a price above its marginal generation cost.
- making a strategic bid of both price and quantity as in the *supply function model*.

1.1.2 Short-term unit commitment problem

One of the most crucial short-term operational decisions in power system is the unit commitment (UC) which involves finding the least-cost commitment and dispatch schedule for the generation resources in order to meet the demand. In deregulated pool-based electricity markets, the UC decisions are made by an independent system operator (ISO) on an hourly basis for a time horizon of one day (24 hours) to one week ahead. Some of the challenging factors that ISO faces in maintaining a reliable and cost-effective operation of the system are as follows:

- *Demand*: On a day-ahead basis, the ISO determines the optimal generation schedule taking into account the estimated load. In addition, a reserve capacity is also scheduled so that unanticipated deviation from the load forecast can be dealt with during the actual operating day. This excess reserve incurs additional cost which need to be accounted for in the cost minimisation objective function.
- *Technical limits*: The operation of conventional generating units is subject to physical limitations which has an impact on the scheduling decisions. For example, the power output level of a generator should be within an operating range and cannot be changed too rapidly.
- *Security*: In determining the UC decisions, the ISO needs to take into account the risk of system disturbances such as line and generator outages. The risk of such unplanned system disturbances can be mitigated through adequate contingency planning. For instance, the outage of any single system component (or predefined

set of components) should not cause a cascading outage of the system that leads to a total or partial blackout. A system that is resistant to the outage of any one component is said to be $n - 1$ secure.

- *Wind power*: Power generation from renewable sources has increased rapidly in recent years. Amongst the sources of renewable energy, wind power has a prominent position. However, due to the intermittent nature of the wind, the integration of wind power into the existing power network is one of the major challenges for the system operators. Therefore, ISO needs to take into account the wind uncertainty when consider the short-term UC problem to avoid imbalances between scheduled production and demand.

To address the above issues, we develop a two-stage unit commitment model which provides solutions that are robust against the distributional uncertainty of the wind generation. Moreover, it includes the operational characteristics of the power system including the reserve scheduling, security criteria and technical constraints.

1.1.3 Long-term facility location problem

The long-term planning projects in energy industry are often involved with strategic investment decisions that are capital intensive and non-repetitive. One of the important problems that often arises in long-term investment planning is the optimal location of facilities. That is to locate facilities and allocate customers to the facility so as to minimise the total investment cost and future service cost (e.g. distance travelled). Consider, an example, where a gas producer has an obligation to deliver certain amounts of gas at certain points in the network at certain times. The producer is aware that at times there are uctuations in demand, or interruptions in its production or transportation systems. By having storage facilities near the delivery points, the producer can reduce the transportation cost or the chance of failing to deliver. A graphical representation of the facility location problem from the gas producer's point of view is shown in Figure 1.3.

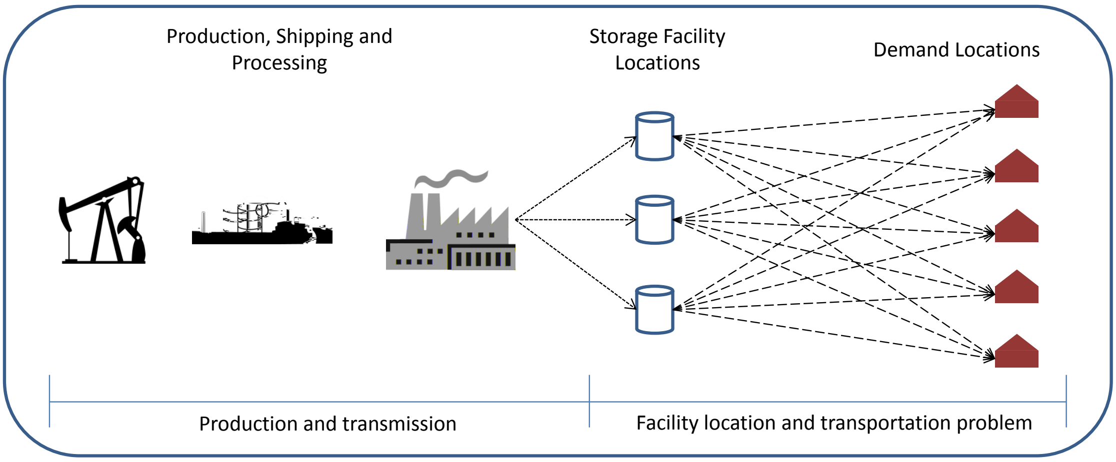


Figure 1.3: Natural gas supply chain

Whether it be building new wind farms, or constructing new gas storage facilities, the starting point for facility location analysis is the projection of the future demand. This is important since the decision for locating the facility has to be made at present and hence is subject to risks arising from uncertainties in future demands and operations of the established facilities. We propose a two-stage robust model to deal with the facility location problem where the customer demand constitute considerable uncertainty and complete information on the distribution of uncertainty is unavailable.

1.2 Optimisation under uncertainty

One of the fundamental assumptions in classical optimisation is that all data are known and of deterministic nature. However, in many real-life problems, decisions are often taken in the face of the uncertainty and the consequences of such decisions cannot fully be determined until at a later stage. Neglecting such uncertainty may have adverse e ects on the quality of the solutions. For instance, the deterministic optimal solutions may reveal to be non-optimal or even infeasible, as a result of small deviations from nominal data [13]. As a result, the eld of optimisation under uncertainty has a well-established presence in the operational research literature and a considerable progress has been made in this area over the past few decades. Within this area, two of the main modelling frameworks are stochastic programming and robust optimisation.

1.2.1 Stochastic programming

Inspired by the probability theory and statistics, the Stochastic Programming (SP) is a family of optimisation techniques, that aims to include future uncertainty into decision making. The term *stochastic*, as opposed to *deterministic*, refers to the presence of random data in the mathematical programming formulation of the problem. The field was pioneered in 1950s by the work of Dantzig and Beale ([39]-[12]), who developed the first mathematical programming models consisting of the actions which were followed by the observation and reaction, or so called *recourse decision*.

The stochastic programming approach is typically used for decision problems where some decisions need to be made today, whilst the important information only becomes available in the future and after the decision has been made. This makes the stochastic programming a highly relevant approach to deal with the planning and investment problems in energy, transportation, finance and supply chain networks.

SP problems are inherently difficult to solve and their complexity increases exponentially to the number of realizations of the uncertain parameters [51]. The classic approach to model and solve SP problems was focused on the representation of uncertainty via finite scenarios. The scenarios were generated by considering possible realizations of the uncertain variables and their associated probabilities. The additional complexity introduced by second-stage decision variables and their constraints has motivated the development of more efficient solution methods. These include the exploitation of SP model structure in developing Dantzing-Wolfe decomposition [40] and the Benders' decomposition [18]. A more recent and well-known approach to deal with the large number of scenarios in the stochastic problems by approximation is the sample average approximation (SAA) also called sample path method in the stochastic optimisation literature. There has been extensive literature on SAA; see the works of Robinson [90], Shapiro [97] and more recently Xu and Meng [78], who have investigated the convergence and applied the SAA method to approximately solve two-stage stochastic programs. The basic idea of SAA is to generate random samples of realizations of random variables and consequently the expected value function is approximated by the corresponding sample average function.

1.2.2 Robust optimisation

One of the fundamental issues in stochastic programming is the representation of uncertainty. The distribution of the underlying uncertainty parameters (random variables) are often assumed to be known in stochastic programming models. However, such assumption might not be realistic in practical cases since the precise information on the probability distribution might not be fully available. The implication of using erroneous information is two-fold; rstly, the deviation of the \assumed" distribution from the \true" distribution could lead to a suboptimal decision and secondly the obtained solution strategy could be infeasible. *Robust optimisation* provides an alternative non-probabilistic paradigm for decision making problems under uncertainty in which the distribution of the uncertain parameters is unknown except for its support. The support set contains all possible realizations of the unknown parameters and is often referred to as the *uncertainty set* in the robust optimisation literature. In robust optimisation framework, instead of seeking to immunize the solution in some probabilistic sense, the decision-maker constructs a solution that is optimal for any realization of uncertain parameters within their uncertainty set. The roots of robust optimisation can be found in the eld of robust control and in the work of Soyster [105] in the early 1970s, in which every uncertain parameter in convex programming problems was taken equal to its worst-case value within a set. Robust optimisation problems tend to be more di -cult to solve than stochastic optimization problems because of their minimax structure. This issue was addressed in the later works of Ben-Tal and Nemirovski [15, 16, 17] and independently in El-Ghaoui and Lebret [53] and El-Ghaoui et. al.[54]. In their works, the uncertain parameters were restricted to belong to ellipsoidal uncertainty sets, which removes the most unlikely outcomes from consideration and yields tractable mathematical programming problems. Bertsimas and Sim [25, 26] and Bertsimas et. al. [22] have proposed a robust optimization approach based on polyhedral uncertainty sets, which preserves the class of problems under analysis, e.g., the robust counterpart of a linear programming problem remains a linear programming problem, and thus has advantages in terms of tractability in large-scale settings. This approach can also be connected to the decision maker's attitude towards uncertainty and risk, providing guidelines to construct the uncertainty set from the historical realizations of the random variables using data-driven optimization [19].

1.2.3 Distributional robust optimisation

An alternative method which aims to create a balance between the conservatism of traditional robust approach and the strong distributional assumption in stochastic optimisation is *distributionally robust optimisation*. This approach is particularly useful when the information on the distribution of underlying uncertainty is not fully available (due to lack of historic data, estimation error or diverse views on uncertainty). In this approach the decision maker uses the available information to constructs a set of distributions in which the true probability distribution is assumed to be contained. The optimal decision is then taken w.r.t. the worst-case probability distribution within the set of distributions. Since its introduction by Scarf [94], this approach has attracted extensive research. by Zackova [122], Dupacova [48, 50], and more recently by Shapiro and Kleywegt [100] and Shapiro and Ahmed [98]. Over the past few years, it has gained substantial popularity through further contributions by Bertsimas and Popescu [23], Betsimas et al [20], Goh and Sim [56], Zhu and Fukushima [124], Goldfarb and Iyengar [57], Delage and Ye [43] and Xu et al [118], to name a few, which cover a wide range of topics ranging from numerical tractability to applications in operations research, nance, engineering and computer science.

The construction of distributional set is based on the form and level of available information about the probability distribution of the random parameters and there are a number of approaches investigated in the literature (see [49] and reference therein). In this thesis, we focus on two approaches for the construction of the distributional set. One is to use the moments of the distribution, and the other is to use a mixture of a set of known probability distributions. One of the attractive features of such approaches is that the duality theory can be used to reformulate the problem as a semi-in nite problem (or semi-de nite program under certain conditions).

1.3 Structure of the thesis

The general aim of this thesis is to develop mathematical models for optimisation problems under uncertainty and the corresponding solution methods tailored to each framework. These models are then applied to relevant problems in energy sector. The thesis follows a paper-based approach and Table 1.1 lists papers, on which the thesis is based.

Chapter	Publication
Chapter 2	\Medium-Term Trading Strategy of a Dominant Electricity Producer", Gourtani et al.[60]
Chapter 3	\Robust Unit Commitment with $n-1$ Security Criteria", Gourtani et al. [58]
Chapter 4	\A Distributionally Robust Optimisation Approach for Two-Stage Facility Location Problem", Gourtani et al.[59]

Table 1.1: List of research papers

Chapter 2, provides a stochastic bi-level programming framework to model the future demand uncertainty and spot market participation from a dominant producer perspective, who is optimising his medium-term strategic decisions.

Chapter 3, present a stochastic and distributionally robust model for the day-ahead unit commitment problem under wind supply uncertainty and security criteria.

Chapter 4, considers a robust framework for the facility location problem taking into account the future uncertainty of demand.

Finally, Chapter 5, summarises the research contributions of this thesis. It also presents the limitations of this work and suggests directions for future research.

Chapter 2

Medium-Term Trading Strategy of a Dominant Electricity Producer

Chapter Abstract

This chapter presents a multi-objective two-stage bilevel stochastic programming framework for a dominant electricity producer to determine an optimal trading strategy in a deregulated electricity spot market in a medium-term time horizon. At the first stage and upper level, the dominant producer aims at maximizing its expected market share and profit, while taking into account the trade-off between the two objectives. At the second stage and lower level, the independent system operator (ISO) determines the dispatches and power flows on an hourly basis after realization of uncertainty in market demand, by solving an optimisation problem which aims at maximizing the total social welfare. Through utilizing Karush-Kuhn-Tucker conditions, the lower level problem is formulated as a complementarity problem and subsequently the dominant producer's optimal decision making problem as a two-stage Stochastic Mathematical Problem with Equilibrium Constraints (SMPEC). To solve the SMPEC, we reformulate the SMPEC as a Mixed Integer Linear Program (MILP) by representing the complementarity constraints as a system of mixed integer linear inequalities with binary variables. Numerical tests results are reported through a medium size case study based on Italian electricity market.

2.1 Introduction

Over the past three decades, the electricity industry in many parts of the world has been restructured by the introduction of wholesale electricity markets. The way in which these markets are implemented varies from one country to another, but they all seek to provide electricity to consumers at a competitive price at the same time as giving appropriate signals for investment and new entry (see Chao and Huntington [37] (1998), Stoft [107] for more information on wholesale electricity markets). Pool-based auction is one of the most common electricity markets where an independent system operator (ISO) processes bids from generators and retailers [103] [66] and determines the market clearing price and power dispatch on the basis of a social bene t maximization framework. The ISO clears the market by nding the equilibrium clearing price of the auction based on the submitted bids. All subsequent trades are settled at this price. When the transaction of power is carried out in zones of a transmission network, the netting of imbalances determines zonal prices, power dispatch and power ows between zones.

Various optimisation and game theoretic models have been proposed to study generator's optimal bidding behavior and market competition. Supply function equilibrium (SFE) is one of them. The concept of SFE is proposed by Klemperer and Meyer [70] to derive a Nash supply function equilibrium in an oligopoly where every player faces uncertainty in demand. The model is applied to the British spot market by Green and Newbery [62], where a supply function represents a generator's one day ahead supply schedule (a stack of price and quantities in increasing order) and the uncertainty describes the daily time-varying demand. Since then, the SFE model has been widely used to study bidding behavior in a single node electricity spot market, see for instances Bolle [27], Baldick and Hogan [8], Rudkevich [93], Anderson and Xu [4].

The equilibrium program with equilibrium constraints (EPEC) is another important model. While SFE focuses on the generators' optimal bidding strategy and the resulting market equilibrium, EPEC models look into the impact of ISO dispatch mechanism and network constraints. For instances, Hobbs, Metzler and Pang [64] investigated an oligopolistic electricity market with several dominant generators located in an electric power network. Generators submit their bids to an independent system operator (ISO) and aim at maximizing their own pro ts while taking into account the competitor's reactions. By reformulating the ISO problem as a complementarity problem through the rst order optimality conditions, they developed an EPEC model where the equilibrium constraints represent the optimality conditions. Yao, Oren and Adler [121] extended the research by considering a stochastic EPEC model (SEPEC) to study the generator's strategic behaviors in a spot market with a two settlement system and network constraints in USA where the stochastic model is used to re ect a day-ahead demand uncertainty. Further research in this direction can be found in Henrion and Römisch [63], Xu and Zhang [120], Surowiec [111], Ehrenmann and Neuho [52], Zhang, Xu and Wu [123] and references therein.

In this chapter we consider an electricity market where a single generator has a dominant position. While this kind of market operation is no longer a common practice, it does exist in some countries or regions that are in the process of market deregulation, due to historical or geographical reasons. Indeed, our work is inspired by the recent research of Vespucci et al [114] on the Italian electricity markets. Instead of considering a game theoretic model such as SFE or EPEC, here we propose a multi-objective two-stage bilevel stochastic programming framework for a dominant electricity producer to determine an optimal trading strategy in a deregulated electricity spot market in a medium-term time horizon: at the rst stage and upper level, the dominant producer aims at maximizing its expected market share and pro t and at the second stage and lower level, the ISO determines the dispatches and power ows on an hourly basis, after realization of uncertainty in market demand, by solving an optimisation problem which aims at maximizing the total social welfare. A distinctive feature of our model is that the dominant generator has two objectives rather than just pro t maximization as in many previous work. This is important since producers often seek to maximize their market share to increase their market power. Furthermore, the levels of pro t could be limited by regulatory constraints such as price caps. Moreover, instead of studying optimal bidding strategies in a one day ahead market, we look into the mid-term strategies for a generator's production schedule. This makes the stochastic model particularly relevant as the uncertainty of demand becomes more signi cant over mid-term time horizon.

The main contributions of this chapter are as follows. We propose a two-stage two-objective stochastic bilevel model for studying optimal mid-term strategies of a dominant producer whose decision is based on maximizing the expected market share and expected pro t. Since the two objectives are not consistent, we investigate the frontier which represents the trade-o between the expected market share and the expected pro ts and observe that under step-wise offers and inelastic demand, this frontier is a convex curve. In order to solve the mathematical model, we reformulate it as a two-stage stochastic mixed-integer linear programming problem and carry out the analysis on the basis of three alternative bidding strategies that the dominant producer can possibly adopt: Cournot bids, Bertrand bids and Supply Functions bids. The proposed mathematical model and the numerical methodology are then applied to analyze the Italian electricity market.

The chapter is organized as follows. Section 2.2 presents a generic multi-objective bilevel optimisation framework under uncertainty. The model is further developed in Section 2.3 with specific details for it to be applied to determining the optimal bidding strategies that set out strategic production schedules in mid-term time horizon. A mathematical reformulation as a single objective MPEC is proposed and further reformulated as a mixed integer linear program in Section 2.4. In Section 2.5, we carry out numerical tests on the proposed model and numerical methodology with real data from the Italian electricity market and finally we draw conclusions in Section 2.6.

2.2 A general multi-objective stochastic bilevel programming model

Bilevel programming models are often adopted to describe the interaction between several agents who are in a hierarchical relationship in an oligopolistic market. We present a generic two-stage multi-objective bilevel stochastic programming model as follows:

$$\max_{x,y(\cdot)} \quad [\ _1(x, y(x,\cdot)), \ldots, \ _n(x, y(x,\cdot))] \tag{2.2.1}$$

$$\text{s.t.} \quad x \in X, \tag{2.2.2}$$

where for almost every $\xi \in$

$$y(x,\xi) \text{ solves } \left\{ \begin{array}{l} \min\limits_{y} \quad f(x,y,\xi) \\ s.t. \\ \quad h(x,y,\xi) = 0, \\ \quad g(x,y,\xi) \geq 0, \\ \quad y \in Y. \end{array} \right\} \tag{2.2.3}$$

The model has interesting practical interpretations: at the upper level and first stage, an agent (leader) needs to make a strategic decision, represented by decision vector x, on its investment or production schedule for the future with several objectives represented by functions ${}_1(x, y(x,\cdot)), \ldots, {}_n(x, y(x,\cdot))$, before the realization of uncertainty. Here we use random vector $\xi(\omega)$ to represent the uncertainty. In doing so, it anticipates the reaction from the other agent (follower) $y(x,\xi)$ which solves an optimisation problem (2.2.3) after the uncertainty is realized and the leader's decision x is observed. Typical objectives for the leader are overall expected profit, expected market share, some risk measures such as standard deviation and Conditional Value-at-Risk.

In the case when the follower's problem (2.2.3) has a unique solution, the two-stage bilevel program is well defined. However, if problem (2.2.3) has multiple optimal solutions, then the leader's anticipation of the follower's reaction $y(x,\xi)$ may depend on the leader's attitude: optimistic, pessimistic or indifferent. If it is optimistic, the leader would consider a positive reaction $y(x,\xi)$ from the follower which would be desirable for his/her own utility, and if it is pessimistic, then the leader would take into account the reaction $y(x,\xi)$ from follower that would be undesirable for his/her own utility.

From a numerical perspective, bilevel stochastic programming problems are in general NP-hard, i.e. no numerical scheme exists that allows solving the problem in polynomial time [44].

The leader's problem is a multi-objective program which may not have a solution that maximizes all objectives. A popular way in the literature of multi-objective optimisation is to consider Pareto optimal solutions. A decision vector x^* is said to be a *Pareto optimal solution* to the problem (2.2.1)-(2.2.3) if and only if there does not exist another $x \in X$ such that ${}_i(x, y(x,\cdot)) \geq {}_i(x^*, y(x^*,\cdot))$ for all i and ${}_j(x, y(x,\cdot)) > {}_j(x^*, y(x^*,\cdot))$ for at least one j. The set of Pareto optimal solutions define a *Pareto efficient frontier*.

The stochastic multi-objective bilevel programming model is an extension of a single objective stochastic Stackelberg leader followers game. For a detailed discussion on the latter, see [42] and [117].

Note also that the two-stage bilevel multi-objective stochastic optimisation problem is generally nonconvex. In the literature of bilevel programming, various techniques have been proposed to deal with the bilevel structure. One of the most well known approaches is to reformulate the lower-level problem as a complementarity problem or a variational inequality problem through Karush-Kuhn-Tucker (KKT) conditions. This is justi ed if the lower level problem is convex in x for almost every ξ. Consequently, we can rewrite (2.2.1)-(2.2.3) as a multi-objective stochastic mathematical program with equilibrium constraints (SMPEC):

$$\max_{x,y(\cdot),\lambda(\cdot),\mu(\cdot)} \quad [\ \ _1(x, y(x,\cdot)), \ldots, \ \ _n(x, y(x,\cdot))] \tag{2.2.4}$$

$$\text{s.t.} \quad \nabla_y L(x, y(x,\xi), \lambda(x,\xi), \mu(x,\xi)) = 0 \tag{2.2.5}$$

$$h(x, y, \xi) = 0 \tag{2.2.6}$$

$$0 \leq \lambda(x,\xi) \perp g(x, y(x,\xi), \xi) \geq 0 \tag{2.2.7}$$

$$x \in X, \ y(x,\xi) \in Y, \ a.e.\ \xi \in \tag{2.2.8}$$

where

$$L(x, y(x,\xi), \lambda(\xi), \mu(\xi)) = f(x, y, \xi) + \mu(x,\xi)h(x, y, \xi) + \lambda(x,\xi)g(x, y, \xi)$$

is the Lagrange function of the lower-level problem, and $\mu(x,\xi)$ and $\lambda(x,\xi)$ are the Lagrange multipliers associated with the lower-level constraints.

The complementary constraint (2.2.7) does not satisfy any classical constraint quali - cation for nonlinear optimisation problem such as the linear independence constraint quali cation (LICQ) and the Mangasarian-Fromovitz constraint quali cation (MFCQ). A lot of research has been carried out to address the issue including NLP-regularization [95], partial penalization [74] and mixed integer programming reformulation [55]. We will come back to this in Section 2.4 when we develop numerical methods for our speci c stochastic bilevel programming problem to be developed in the next section.

2.3 The mathematical model of the dominant producer's problem

In this section, we consider an electricity market with a large scale (dominant) producer and a number of smaller producers behaving as a competitive fringe. The market is divided in zones, interconnected by capacitated transmission networks. Both the dominant producer and its competitors own a number of power generation units across the zones and the power generated in each zone can be transmitted to other zones subject to the capacity constraint of transmission lines. The constraint may lead to the differentiation of spot prices across the network and this makes the location of a generation unit highly crucial in profitability, particularly when the market demand is inelastic.

We consider a day-ahead spot market with an Independent System Operator which oversees power trading and determines dispatch quantities, zonal prices and power flows across the network on an hourly basis, given zonal demands and power producers' bids for every generation plant. The ISO's decision is based on a framework which maximizes the total social welfare.

The objective of the dominant producer is to set strategic quantity bids for each of its generation unit on an hourly basis over a certain time horizon which maximizes its market share while managing the risk of the profit falling below a predefined level. This differs from many models in the literature where a generator often bases its decision on profit maximization. Since the levels of profit could be limited by regulatory constraints such as price caps.

This type of spot market was first considered by Vespucci et al [114] on the basis of the Italian power market. They developed a deterministic bilevel programming model for the dominant producer's decision making problem where the producer sets out an annual optimal power production schedule on an hourly basis at the upper level with an anticipation of ISO's social benefit maximization based power dispatch mechanism at the lower level. It is assumed that the dominant producer knows, for every hour of the year, the zonal hourly demands and the competitors' bids (minimum price requested, which equals to the constant generation marginal cost, and maximum offered quantity). The dominant producer determines first the quantities that the ISO would accept for each plant in the system (both his and competitors' generation units) if the dominant producer does not exert market power, i.e. the dominant producers bid prices are equal to the marginal costs, which are assumed to be constant. Outputs of this first step are the accepted quantities of both its generation units and of the competitors' units. In order to guarantee the profit level allowed by the system (equality constraint), the dominant producer tries to modify the "perfect competition" solution. It is assumed that, among all solutions that guarantee the profit level agreed by the system, the dominant producer prefers the one corresponding to the largest market share. This solution is also preferred by the system, because the larger the dominant producer's market share, the lower the zonal prices, since the dominant producers' production withdrawal activates more costly competitors' bids.

The perfect competition solution can be modified in two possible ways.

1. Consider a zone z in which the dominant producer is not indispensable to satisfy the hourly zonal demand (i.e. there is some competitor's capacity still available either in zone z or in zones connected with zone z by non-saturated transmission links). In this case the dominant producer has to reduce his own production, in order to activate more costly competitors' bids (for competitors' units whose dispatch is less than the plant capacity), which results in a higher hourly zonal price set by the Market Operator. If the reduction of the dominant producer's production is such that every competitors' bid is either completely accepted or completely rejected, then the clearing price is not uniquely determined, as it may be any value between the bid price of the last accepted bid and the bid price of the first rejected bid. In this case the model determines the clearing price as the value (between the bid price of the last accepted bid and the bid price of the first rejected bid) that allows the dominant producer to exactly obtain the prefixed profit. In the model there is not a specific representation of the bid price of the dominant producer. Indeed, only the quantities to be offered are determined by the model. Anyway, the dominant producer gets a usable information about the bid price he has to associate to the quantity offered by him: indeed he can bid at any price not greater than the clearing price determined by the model.

2. In the hours and zones in which the dominant producer is indispensable to satisfy the hourly zonal demand, the dominant producer cannot reduce his own production, therefore he will be able to offer his own production at a higher bid price. This "higher" bid price would be the zonal price-cap, if the model required annual profit maximization. But, analogously to the previous case, the model determines the clearing price as the value, belonging to the interval between the bid price of the last accepted competitors' bid and the price-cap, that allows the dominant producer to exactly obtain the prefixed profit. Therefore the dominant producer gets from the model the information that he has to offer his own production at the clearing price determined by the model.

Summarizing, in the hours under case 1 a strategy on quantities is used, while in the hours under case 2 a strategy on prices is used. The model determines the best combination of the two strategies, in terms of market share maximisation. Note that, since a prefixed profit level must be achieved, it can happen that in some hours no actions are taken by the dominant producer, that is the perfect competition solution is used, even if market power could be exerted.

While the model captures the main features of the Italian power market operation, it does not address potential uncertainties in market demand particularly in a one year ahead planning. This motivates us to consider a stochastic version of the model by explicitly taking into account the fluctuation of zonal demand on hourly basis. Moreover, instead of setting the market share as a single objective, we consider an additional objective which maximizes the expected profit and this allows us to analyze the trade-off between the two quantities.

2.3.1 Notation

Throughout this section and the rest of the chapter, a detailed mathematical formulation of the dominant producer's decision making problem is developed, a numerical method is proposed for its solution and the results of some numerical tests are presented and discussed. In order to present the mathematical model, the following notation is introduced.

⋆ **Sets**

Z set of zones across the network, indexed by z

L set of transmission lines connecting the zones, indexed by l

K set of power units owned by the dominant producer, indexed by k

J set of power units owned by competitors, indexed by j

T set of representative time periods within the time horizon, indexed by t

S set of electricity demand scenario, indexed by s

K_z set of power units owned by the dominant producer in zone z

J_z set of power units owned by other producers in zone z

I set of dominant producer's objectives, indexed by i

⋆ **Parameters**

- w_t total number of hours in time period t
- D_{zts} load demand in zone z, time t and scenario s
- p_s probability of scenario s
- $\overline{f}_l$ power ow limit in line l
- φ_{lz} power transfer distribution factor (PTDF) for zone z and line l
- $\overline{q}_{jt}$ generation capacity of the competitor producers' plant j at time t
- c_{jt} unitary generation cost of plant j at time t
- b_{jt} associated bid price of competitor producers' plant j at time t
- $\overline{Q}_{kt}$ generation capacity of the dominant producer's plant k at time t
- C_{kt} unitary generation cost of plant k at time t
- β_i weight associated with the dominant producer's objective i

⋆ **Functions**

- $^{MS}(\mathbf{q})$ expected annual market share of dominant producer
- $^{PR}(\mathbf{q}, \pi)$ expected annual pro t of dominant producer

⋆ **Variables**

- B^{str}_{kt} associated bid price of the dominant producer's plant k at time t
- B^{str}_{kt} associated bid price of the dominant producer's plant k at time t
- Q^{str}_{kt} associated bid quantity of the dominant producer's plant k at time t
- Q_{kts} production of dominant producer's plant k, at time t, in scenario s
- q_{jts} production of competitors' plant j, at time t, in scenario s
- r_{zts} net power ow in and out of zone z, at time t, in scenario s
- π_{zts} shadow price associated with zonal balance constraints (spot price) in zone z, at time t, in scenario s
- α^+_{kts} dual variable of the upper bound to production of dominant producer's power unit k at time t in scenario s
- α^-_{kts} dual variable of the lower bound to production of dominant producer's power unit k at time t in scenario s
- λ^+_{jts} dual variable of the upper bound to production of competitors' power unit j at time t in scenario s
- λ^-_{jts} dual variable of the lower bound to production of competitors' power unit j at time t in scenario s
- μ_{ts} dual variable associated to the constraint that states negligibility of power loss at time t in scenario s
- η^+_{lts} dual variable of the upper bound to power ow on transmission link l at time t in scenario s
- η^-_{lts} dual variable of the lower bound to power ow on transmission link l at time t in scenario s
- $\theta^{\alpha^+}_{kts}$ binary variable associated to the upper bound to production of dominant producer's power unit k at time t in scenario s
- $\theta^{\alpha^-}_{kts}$ binary variable associated to the lower bound to production of dominant producer's power unit k at time t in scenario s
- $\theta^{\lambda^+}_{jts}$ binary variable associated to the upper bound to production of competitors' power unit j at time t in scenario s
- $\theta^{\lambda^-}_{jts}$ binary variable associated to the lower bound to production of competitors' power unit j at time t in scenario s
- $\theta^{\eta^+}_{lts}$ binary variable associated to the upper bound to power ow on transmission link l at time t in scenario s
- $\theta^{\eta^-}_{lts}$ binary variable associated to the lower bound to power ow on transmission link l at time t in scenario s

2.3.2 The ISO's decision making problem

We start our mathematical formulation with the ISO's decision making problem. The market clearing process in a day-ahead spot market works as follows. Upon the realization of the actual demand, the ISO carries out a sealed bid auction in which the electricity producers submit their generation bids in the form of hourly price-quantity stacks. Figure 2.1 illustrates the spot market auction process.

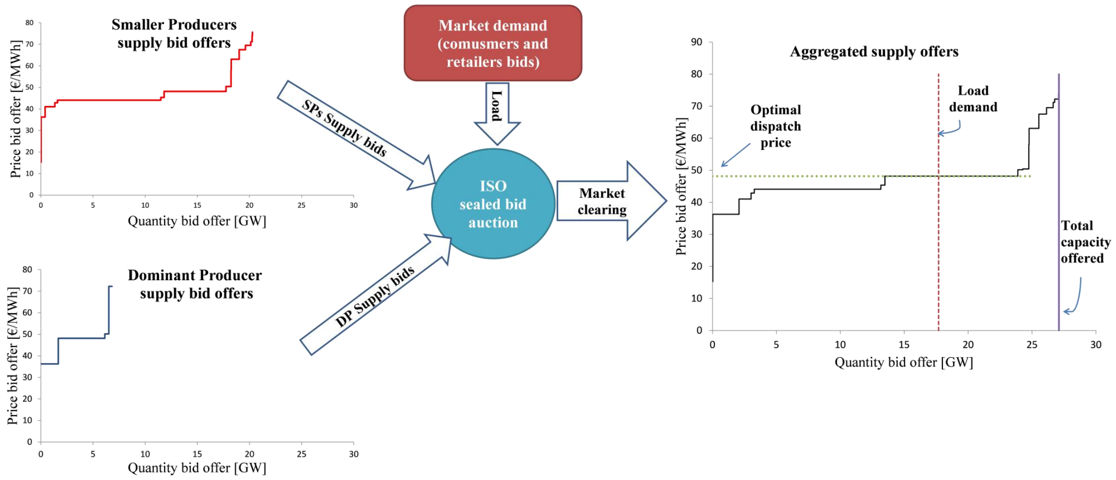


Figure 2.1: Spot (day-ahead) market clearing process

For every hour t in the time horizon, the price-quantity bid made by the dominant producer for his own power plant k is $(B_{kt}^{str}, Q_{kt}^{str})$, where B_{kt}^{str} denotes the strategic price set by the leader at the beginning of the year, subject to price-cap or marginal generation cost, and Q_{kt}^{str} denotes the strategic quantity, subject to the power plant generation capacity constraint. The price-quantity bid for competitors' power plant j is $(b_{jt}, \overline{q}_{jt})$: competitors are assumed not to be strategic players, i.e. the bid quantity is the power plant capacity and the bid price is the power plant marginal cost.

Market demand is assumed to be inelastic and subject to some random shocks. In every hour t and after the realization of uncertainty in demand, i.e. realization of scenario s, the ISO determines the dispatch schedule, zonal prices and power ows by solving the following minimization problem:

$$\min_{Q,q,r} \quad \sum_{k\in K} B_{kt}^{str}\cdot Q_{kts} + \sum_{j\in J} b_{jt}\cdot q_{jts} \tag{2.3.1}$$

$$\text{s.t.} \quad \sum_{k\in K_z} Q_{kts} + \sum_{j\in J_z} q_{jts} + r_{zts} = D_{zts}, \qquad \forall z\in Z, \quad \pi_{zts}, \tag{2.3.2}$$

$$\sum_{z\in Z} r_{zts} = 0, \qquad \mu_{ts}, \tag{2.3.3}$$

$$-\overline{f}_l \leq \sum_{z\in Z} \varphi_{lz}\cdot r_{zts} \leq \overline{f}_l, \qquad \forall l\in L, \quad \eta_{lts}^-, \eta_{lts}^+, \tag{2.3.4}$$

$$0 \leq q_{jts} \leq \overline{q}_{jt}, \qquad \forall j\in J, \quad \lambda_{jts}^-, \lambda_{jts}^+, \tag{2.3.5}$$

$$0 \leq Q_{kts} \leq Q_{kt}^{str}, \qquad \forall k\in K, \quad \alpha_{kts}^-, \alpha_{kts}^+. \tag{2.3.6}$$

Here the objective is to determine Q_{kts}, q_{jts} and r_{zts} so as to maximize the social welfare, i.e. minimize the total cost, with market demand assumed to be inelastic and independent of the market price in scenario s. The bid prices set by the dominant producer and by the competitors are B_{kt}^{str} and b_{jt} respectively, and they are given as parameters to the ISO. The bid quantities set by the dominant producer and by the competitors are Q_{kt}^{str} and $\overline{q}_{jt}$ respectively and they are also given as parameters to the ISO.

This minimization problem is subject to a number of technical and economic constraints. Equality (2.3.2) is a power balance constraint which ensures that at any time the total demand in each zone is met by the total production plus power ows in and out of each zone. Constraint (2.3.3) means that there is no power loss across the whole network. We make this assumption for the simplicity of discussion. Constraints (2.3.4) mean that the power ow parameterized by the PTDF at each transmission line is subject to the capacity limit of each line. Constraint (2.3.5) says that the competitors' overall dispatch for plant j must not exceed their generation capacity. The last constraints (2.3.6) stipulate that the dominant producer's total dispatch for the plant k must not exceed its strategic bid quantities set at the beginning of the year.

The ISO's decision making problem is a deterministic linear programming problem which can be reformulated as a mixed complementarity problem through the rst order optimality conditions:

$$\sum_{j \in J_z} q_{jts} + \sum_{k \in K_z} Q_{kts} + r_{zts} - D_{zts} = 0, \qquad \forall z \in Z, \tag{2.3.7}$$

$$\sum_{z \in Z} r_{zts} = 0, \tag{2.3.8}$$

$$\mu_{ts} + \sum_{l \in L} \left(\eta^+_{lts} - \eta^-_{lts}\right) \cdot \varphi_{lz} - \pi_{zts} = 0, \qquad \forall z \in Z, \tag{2.3.9}$$

$$0 \leq \eta^-_{lts} \perp \overline{f}_l + \sum_{z \in Z} \varphi_{lz} \cdot r_{zts} \geq 0, \qquad \forall l \in L, \tag{2.3.10}$$

$$0 \leq \eta^+_{lts} \perp \overline{f}_l - \sum_{z \in Z} \varphi_{lz} \cdot r_{zts} \geq 0, \qquad \forall l \in L, \tag{2.3.11}$$

$$b_{jt} - \pi_{zts} + \lambda^+_{jts} - \lambda^-_{jts} = 0, \qquad \forall j \in J_z, z \in Z, \tag{2.3.12}$$

$$0 \leq \lambda^-_{jts} \perp q_{jts} \geq 0, \qquad \forall j \in J, \tag{2.3.13}$$

$$0 \leq \lambda^+_{jts} \perp \overline{q}_{jt} - q_{jts} \geq 0, \qquad \forall j \in J, \tag{2.3.14}$$

$$B^{str}_{kt} - \pi_{zts} + \alpha^+_{kts} - \alpha^-_{kts} = 0, \qquad \forall k \in K_z, z \in Z, \tag{2.3.15}$$

$$0 \leq \alpha^-_{kts} \perp Q_{kts} \geq 0, \qquad \forall k \in K, \tag{2.3.16}$$

$$0 \leq \alpha^+_{kts} \perp Q^{str}_{kt} - Q_{kts} \geq 0, \qquad \forall k \in K, \tag{2.3.17}$$

where π_{zts}, η^-_{lts}, η^+_{lts}, λ^-_{jts}, λ^+_{jts}, α^-_{kts} and α^+_{kts} are Lagrange multipliers. Note that in solving the optimisation problem, ISO determines the dispatch quantities Q_{kts} and q_{jts} for the dominant producer at plant k and the competitive producers' at plant j, the net power ow in/out r_{zts} at zone z. The market clearing price at zone z is the dual variable (shadow price) of the balance constraints, i.e. the maximum price a load is willing to pay for a quantity of energy in that zone, which corresponds to Lagrange multiplier π_{zts}.

2.3.3 The dominant producer's decision making problem

The dominant producer has two main objectives: the expected market share and the expected pro t. There could be other objectives, represented by risk-aversion measures, but we will not consider them for the simplicity of discussion.

To simplify the hourly demand levels within a xed time horizon, say one year, we divide the horizon into T time periods indexed by $t = \{1, \cdots, T\}$. All the hours within a given time period assumed to have the same level of demand. We represent the number of hours in each time period t by w_t. For example we may divide a year into 4 time periods, weekday peak, weekday o -peak, weekend peak and weekend o -peak, and assume that in each period the demand level is identical.

The dominant producer's total production in time period t in scenario s is $w_t \cdot \sum_{k\in K} Q_{kts}$, where w_t denotes the number of single hours in the period. The total expected market share is de ned as the total expected power dispatched over the time horizon and can be expressed as:

$$^{MS}(\mathbf{q}) := \sum_{s\in S} p_s \sum_{t\in T} w_t \sum_{k\in K} Q_{kts}, \tag{2.3.18}$$

where $\mathbf{q} := \{Q_{kts} : k \in K, t \in T, s \in S\}$.

Likewise, we can derive the total annual expected pro t:

$$^{PR}(\mathbf{q}, \pi) := \sum_{s\in S} p_s \sum_{t\in T} w_t \left(\sum_{z\in Z} \sum_{k\in K_z} \pi_{zts} \cdot Q_{kts} - \sum_{k\in K} C_{kt} \cdot Q_{kts} \right), \tag{2.3.19}$$

where $\pi = \{\pi_{zts} : z \in Z, t \in T, s \in S\}$.

Since the decision problem of the dominant producer is of the two-objective nature, the objective function is de ned as:

$$\max \Big(\; ^{MS}(\mathbf{q}), \; ^{PR}(\mathbf{q}, \pi) \Big). \tag{2.3.20}$$

The presence of more than one objective necessitates the application of special optimisation procedures to optimize them, either simultaneously or iteratively. We use a simultaneous approach so-called the utility function method (also known as weighting function method); in which a utility function is de ned for each of the objectives according to their relative importance. More speci cally, each objective i is multiplied by a scalar $\beta_i \in [0, 1]$ such that $\sum_{i\in I} \beta_i = 1$. The value of β_i indicate the relative utility or the weight assigned to the corresponding objective i. In our problem, the dominant producer has two objectives and therefore we can de ne $\beta_1 = \beta$ and $\beta_2 = (1 - \beta)$. The total utility function can then be written as the weighted sum of both objective functions as follow:

$$\max \Big\{ \beta \cdot \; ^{MS}(\mathbf{q}) + (1 - \beta) \cdot \; ^{PR}(\mathbf{q}, \pi) \Big\}. \tag{2.3.21}$$

Since the two objectives may be con icting, by varying the parameter β value we get a frontier of optimal values and Pareto optimal solutions.

The aim of the dominant producer is to nd the optimal bidding strategy in the spot market, that is, the stack of quantity-price $(Q^{str}_{kt}, B^{str}_{kt})$. In a deregulated market where producers do not have a market power, they tend to bid in with their available production capacity $(\overline{Q}_{kt}, \overline{q}_{jt})$ and their true marginal generation cost (C_{kt}, c_{jt}). However, in an oligopolistic market where producers have market power, they have incentives to in uence the market price by retaining some of their generation capacities.

In the context of this research, we consider three di erent of the most common ways in which the dominant producer may exercise its market power:

- making a strategic bid on power generation quantities as in the Cournot model by retaining some generation capacity;
- making a strategic price bid as in the Bertrand model by setting a price above its marginal generation cost and below the price cap π;
- making a strategic bid of both price and quantity as in the supply function model.

This leads to three di erent optimisation models: the Cournot model, the Bertrand model and the supply function model.

The $Cournot\ model$ can be mathematically described as

$$\begin{aligned} \max_{Q^{str}_{kt}} \quad & \beta \cdot \ {}^{MS}(\mathbf{q}) + (1-\beta) \cdot \ {}^{PR}(\mathbf{q}, \pi) \\ \text{s.t.} \quad & 0 \leq Q^{str}_{kt} \leq \overline{Q}_{kt}, \ \forall k \in K, t \in T, \\ & \text{complementarity constraints (2.3.7)} - \text{(2.3.17)}, \ \forall t \in T, s \in S, \end{aligned} \tag{2.3.22}$$

where the price bid B^{str}_{kt} is xed to the marginal cost of generation C_{kt} of the dominant generating units.

The $Bertrand\ model$ can be presented as

$$\begin{aligned} \max_{B^{str}_{kt}} \quad & \beta \cdot \ {}^{MS}(\mathbf{q}) + (1-\beta) \cdot \ {}^{PR}(\mathbf{q}, \pi) \\ \text{s.t.} \quad & C_{kt} \leq B^{str}_{kt} \leq \pi, \ \forall k \in K, t \in T, \\ & \text{complementarity constraints (2.3.7)} - \text{(2.3.17)}, \ \forall t \in T, s \in S, \end{aligned} \tag{2.3.23}$$

where the quantity bid Q_{kt}^{str} is xed to the total available capacity $\overline{Q}_{kt}$ of the dominant generating units.

Consider the ISO problem. Since it is a linear convex program, by the strong duality theorem, the optimal dual value equals the optimal primal value, i.e.

$$\begin{aligned}\sum_{k\in K} B_{kt}^{str} Q_{kts} + \sum_{j\in J} b_{jt} q_{jts} &= \sum_{z\in Z} \pi_{zts} D_{zts} - \sum_{k\in K} \alpha_{kts}^{+} Q_{kt}^{str} - \sum_{j\in J} \lambda_{jts}^{+} \overline{q}_{jt} \\ &\quad - \sum_{l\in L} \eta_{lts}^{+} \overline{f}_l - \sum_{l\in L} \eta_{lts}^{-} \overline{f}_l. \end{aligned} \tag{2.4.1}$$

Multiplying both sides of equation (2.3.15) by Q_{kts} and summing over $z \in Z$ and $k \in K$, we obtain

$$\sum_{k\in K} B_{kt}^{str} Q_{kts} - \sum_{z\in Z} \sum_{k\in K_z} \pi_{zts} Q_{kts} + \sum_{k\in K} \alpha_{kts}^{+} Q_{kts} - \sum_{k\in K} \alpha_{kts}^{-} Q_{kts} = 0. \tag{2.4.2}$$

On the other hand, the complementarity conditions (2.3.16) and (2.3.17) imply that

$$\begin{aligned} \alpha_{kts}^{-} Q_{kts} &= 0, \\ \alpha_{kts}^{+} (Q_{kt}^{str} - Q_{kts}) &= 0, \end{aligned}$$

or, equivalently,

$$\alpha_{kts}^{+} Q_{kts} = \alpha_{kts}^{+} Q_{kt}^{str}.$$

Using the two relations above, we can rewrite (2.4.2) as

$$\sum_{k\in K} \alpha_{kts}^{+} Q_{kt}^{str} = \sum_{z\in Z} \sum_{k\in K_z} \pi_{zts} Q_{kts} - \sum_{k\in K} B_{kt}^{str} Q_{kts}.$$

By substituting $\sum_{k} \alpha_{kts}^{+} Q_{kt}^{str}$ into (2.4.1), we obtain

$$\begin{aligned}\sum_{k\in K} B_{kt}^{str} Q_{kts} + \sum_{j\in J} b_{jt} q_{jts} &= \sum_{z\in Z} \pi_{zts} D_{zts} - \sum_{l\in L} \eta_{lts}^{+} \overline{f}_l - \sum_{l\in L} \eta_{lts}^{-} \overline{f}_l \\ &\quad - \left(\sum_{z\in Z} \sum_{k\in K_z} \pi_{zts} Q_{kts} - \sum_{k\in K} B_{kt}^{str} Q_{kts} \right) - \sum_{j\in J} \lambda_{jts}^{+} \overline{q}_{jt}. \end{aligned}$$

and through some cancellations, we arrive at

$$\sum_{z\in Z} \sum_{k\in K_z} \pi_{zts} Q_{kts} = \sum_{z\in Z} \pi_{zts} D_{zts} - \sum_{j\in J} b_{jt} q_{jts} - \sum_{j\in J} \lambda_{jts}^{+} \overline{q}_{jt} - \sum_{l\in L} (\eta_{lts}^{+} + \eta_{lts}^{-}) \overline{f}_l, \tag{2.4.3}$$

where the right-hand side is a linear function of the decision variables. Based on this, the objective function of the three SMPECs can be formulated as follows

$$
\begin{aligned}
{}_{\beta}(\mathbf{q},\pi) :=& \beta \ {}^{MS}(\mathbf{q}) + (1-\beta) \ {}^{PR}(\mathbf{q},\pi) \\
=& \beta \left(\sum_{s\in S} p_s \sum_{t\in T} w_t \sum_{k\in K} Q_{kts} \right) + (1-\beta) \left(\sum_{s\in S} p_s \sum_{t\in T} w_t \left(\sum_{z\in Z} \pi_{zts} D_{zts} \right.\right. \\
& \left.\left. - \sum_{j\in J} b_{jt} q_{jts} - \sum_{j\in J} \lambda^{+}_{jts} \overline{q}_{jt} - \sum_{l\in L} (\eta^{+}_{lts} + \eta^{-}_{lts}) \overline{f}_l - \sum_{k\in K} Q_{kts} C_k \right)\right),
\end{aligned}
$$

and the ISO's KKT equivalent constraints (2.3.7){(2.3.17), for all $t \in T$ and $s \in S$, as

$$\sum_{k\in K_z} Q_{kts} + \sum_{j\in J_z} q_{jts} + r_{zts} - D_{zts} = 0, \qquad \forall z \in Z, \tag{2.4.4}$$

$$\sum_{z\in Z} r_{zts} = 0, \tag{2.4.5}$$

$$B^{str}_{kt} - \pi_{zts} + \alpha^{+}_{kts} - \alpha^{-}_{kts} = 0, \qquad \forall k \in K_z, z \in Z, \tag{2.4.6}$$

$$b_{jt} - \pi_{zts} + \lambda^{+}_{jts} - \lambda^{-}_{jts} = 0, \qquad \forall j \in J_z, z \in Z, \tag{2.4.7}$$

$$\mu_{ts} + \sum_{l\in L} \left(\eta^{+}_{lts} - \eta^{-}_{lts}\right) \varphi_{lz} - \pi_{zts} = 0, \qquad \forall z \in Z, \tag{2.4.8}$$

$$0 \le Q^{str}_{kt} - Q_{kts} \le BM^{Q}(1-\theta^{\alpha^{+}}_{kts}), \qquad \forall k \in K, \tag{2.4.9}$$

$$0 \le \alpha^{+}_{kts} \le BM^{\alpha}\theta^{\alpha^{+}}_{kts}, \qquad \forall k \in K, \tag{2.4.10}$$

$$0 \le Q_{kts} \le BM^{Q}(1-\theta^{\alpha^{-}}_{kts}), \qquad \forall k \in K, \tag{2.4.11}$$

$$0 \le \alpha^{-}_{kts} \le BM^{\alpha}\theta^{\alpha^{-}}_{kts}, \qquad \forall k \in K, \tag{2.4.12}$$

$$0 \le \overline{q}_{jt} - q_{jts} \le BM^{q}(1-\theta^{\lambda^{+}}_{jts}), \qquad \forall j \in J, \tag{2.4.13}$$

$$0 \le \lambda^{+}_{jts} \le BM^{\lambda}\theta^{\lambda^{+}}_{jts}, \qquad \forall j \in J, \tag{2.4.14}$$

$$0 \le q_{jts} \le BM^{q}(1-\theta^{\lambda^{-}}_{jts}), \qquad \forall j \in J, \tag{2.4.15}$$

$$0 \le \lambda^{-}_{jts} \le BM^{\lambda}\theta^{\lambda^{-}}_{jts}, \qquad \forall j \in J, \tag{2.4.16}$$

$$0 \le \overline{f}_l - \sum_{z\in Z} \varphi_{lz} r_{zts} \le BM^{f}(1-\theta^{\eta^{+}}_{lts}), \qquad \forall l \in L, \tag{2.4.17}$$

$$0 \le \eta^{+}_{lts} \le BM^{\eta}\theta^{\eta^{+}}_{lts}, \qquad \forall l \in L, \tag{2.4.18}$$

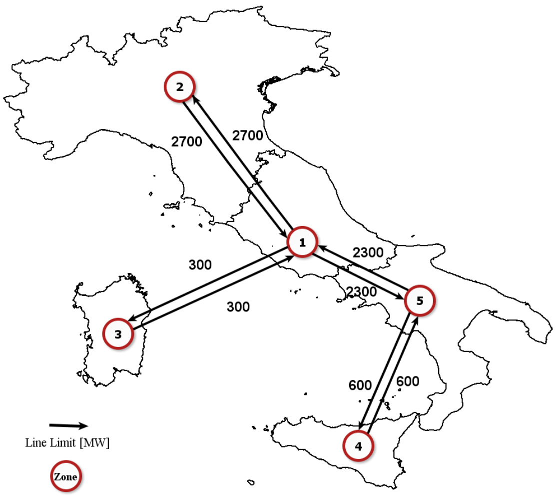
$$0 \le \overline{f}_l + \sum_{z\in Z} \varphi_{lz} r_{zts} \le BM^{f}(1-\theta^{\eta^{-}}_{lts}), \qquad \forall l \in L, \tag{2.4.19}$$

$$0 \le \eta^{-}_{lts} \le BM^{\eta}\theta^{\eta^{-}}_{lts}, \qquad \forall l \in L. \tag{2.4.20}$$

The proposed SMPEC is stated as a MILP model solvable with commercial solvers. The complementary equations (2.3.10){(2.3.17) are replaced by the linearized equations (2.4.9){(2.4.20) by introducing the binary variables $\theta^{\alpha+}_{kts}$, $\theta^{\alpha-}_{kts}$, $\theta^{\lambda+}_{jts}$, $\theta^{\lambda-}_{jts}$, $\theta^{\eta^+}_{lts}$, $\theta^{\eta^-}_{lts}$, and the su cient large constants BM^Q, BM^α, BM^q, BM^λ, BM^f and BM^η. This complementary constraints linearization is based on [55], and constitutes an exact reformulation.

2.5 Numerical tests

The proposed models have been applied to the Italian electricity market, using data related to the year 2011. The market consists of 5 zones, interconnected by 4 capacitated transmission links (Figure 2.2).



l	$\overline{f}_l$
Line	Capacity[GW]
$1 \to 2$	2.70
$1 \to 3$	0.30
$1 \to 5$	2.30
$4 \to 5$	0.60

$$\varphi_{lz} = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Figure 2.2: Italian market zones, transmission line capacities and PTDF matrix

The electricity producers consist of a dominant producer, with higher production capacity and market power, and a number of small producers which behave as a competitive fringe. They own 23 and respectively 41 generating units across the zones and are capable of transferring their generations to other zones, subject to demand level and capacity of the lines. The aggregated zonal generation capacities and zonal average marginal cost of generation are summarized in Table 2.1. We can see that the dominant producer holds around 34.6% of the total installed generation capacity and has a lower average marginal cost than the competitors in all zones apart from zone 5.

Data	Dominant producer			Competitors	
zones	Max Cap	Avg Cost	% of Installed Capacity	Max Cap	Avg Cost
Zone 1	8.08	47.59	60.20	5.34	48.13
Zone 2	6.82	43.81	25.16	20.29	47.39
Zone 3	0.86	53.66	29.78	2.02	58.89
Zone 4	2.36	55.24	54.67	1.96	73.96
Zone 5	4.21	69.43	24.96	12.65	45.70
Total	**22.33**		**34.57**	**42.26**	

Table 2.1: Aggregated zonal installed generation capacity [GW], average production cost [EURs/MWh] and zonal installed capacity rate for the dominant producer [%]

The demand for electricity is assumed to be inelastic and stochastic. Historical (2011) hourly load data is used as a basis to simplify the annual demand vector of 8760 hours into 8 representative demand levels or periods $\{t_1, \cdots, t_8\}$ as shown in Figure 2.3.

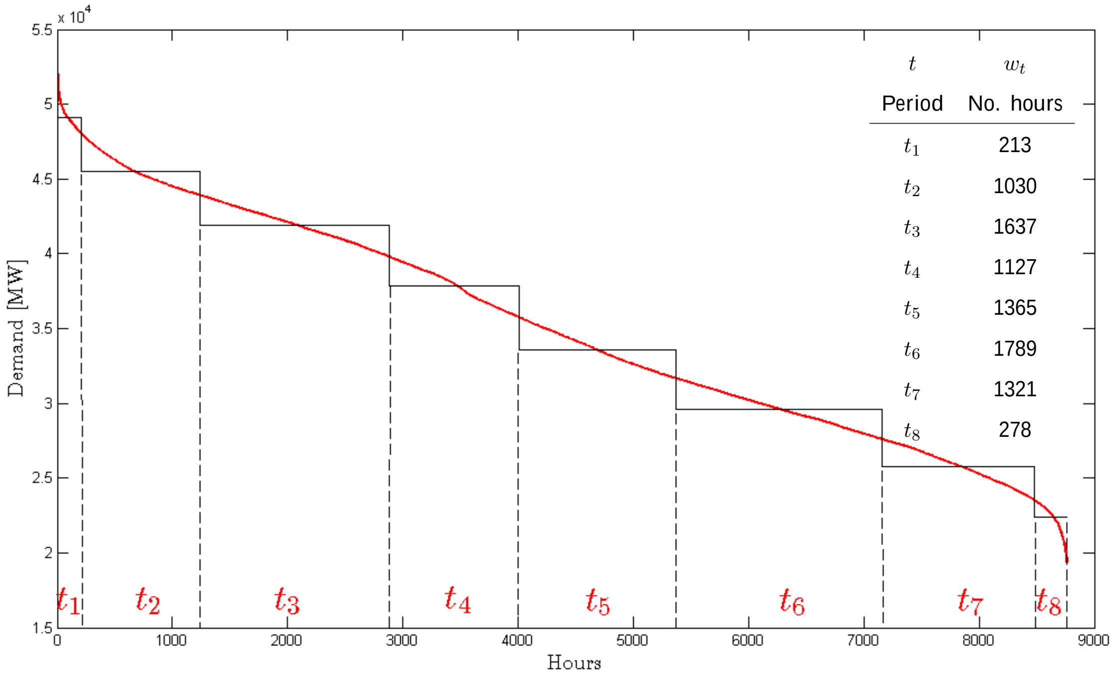


Figure 2.3: Representative load level as 8 time periods

The uncertainty in demand is represented via three demand scenarios $\{s_1, s_2, s_3\}$. For each time period, the average demand scenario s_2 is based on the load data from the year 2011 and the lower and higher scenarios for the demand, s_1 and s_3, are generated by scaling down and up the average value. The probability of each scenario is defined as p_s = [0.25 0.5 0.25]. The hourly demand scenarios for each zone for the peak time period (hours), t_1 are reported in Table 2.2.

		Zones					
	Period t_1	z_1	z_2	z_3	z_4	z_5	Total
Demand Scenarios	s_1 (Low)	13.05	26.19	1.89	3.05	3.96	48.13
	s_2 (Avg)	13.31	26.73	1.93	3.11	4.04	49.12
	s_3 (High)	13.58	27.26	1.97	3.17	4.12	50.10

Table 2.2: Zonal demand scenarios [GW] for the peak load period t_1

2.5.1 Results

In order to model the influence of the dominant producer in the Italian electricity system and analyze the consequences of its strategic behavior, we solve the model with four different approaches: (i) Cournot strategic bidding (ii) Bertrand Strategic bidding (iii) supply function strategic bidding (iv) no strategic behavior, the so called "base case", in which the dominant producer submits the capacity and marginal cost of generation of its units as its quantity and price bids. When bidding strategically, the dominant producer modifies its quantity and/or price offers in the spot market in order to increase its profit and market share.

Although in the Italian electricity market bid prices are allowed to be at price-cap, some form of restriction is usually imposed by the Regulatory Authorities, in order to avoid the over exploitation of the market. In Vespucci et al [114] the restriction is considered to allow the dominant producer to obtain a predetermined annual profit level. As a consequence of this restriction, the dominant producer tends not to bid at the cap price and not to withdraw all available capacity. In this work the maximum annual profit constraint is approximated by imposing the following restrictions:

1. the dominant producer cannot bid at a price which is more than 30% of the power plant marginal costs, i.e., $B_{kt}^{str} \leq 1.3 \cdot C_{kt}$;

2. the dominant producer o ers at least the 80% of its total installed capacity, i.e., $\sum_k Q_{kt}^{str} \geq 0.8 \cdot \sum_k \overline{Q}_{kt}$.

For each bidding approach we carried out numerical tests with 5 di erent values of β between 0 and 1, in order to analyze the trade-o between the objectives. When β is 0, the problem reduces to pro t maximization. In Figure 2.4, we show the results of the four cases. Note that the frontier points are joined for the sake of clarity. The dominant producer's expected market share is represented as its percentage share of total generation in the market. In the base case, the frontier reduce to a single point, as it is a static vision of the spot market, where both the dominant producer and competitors' bids are xed.

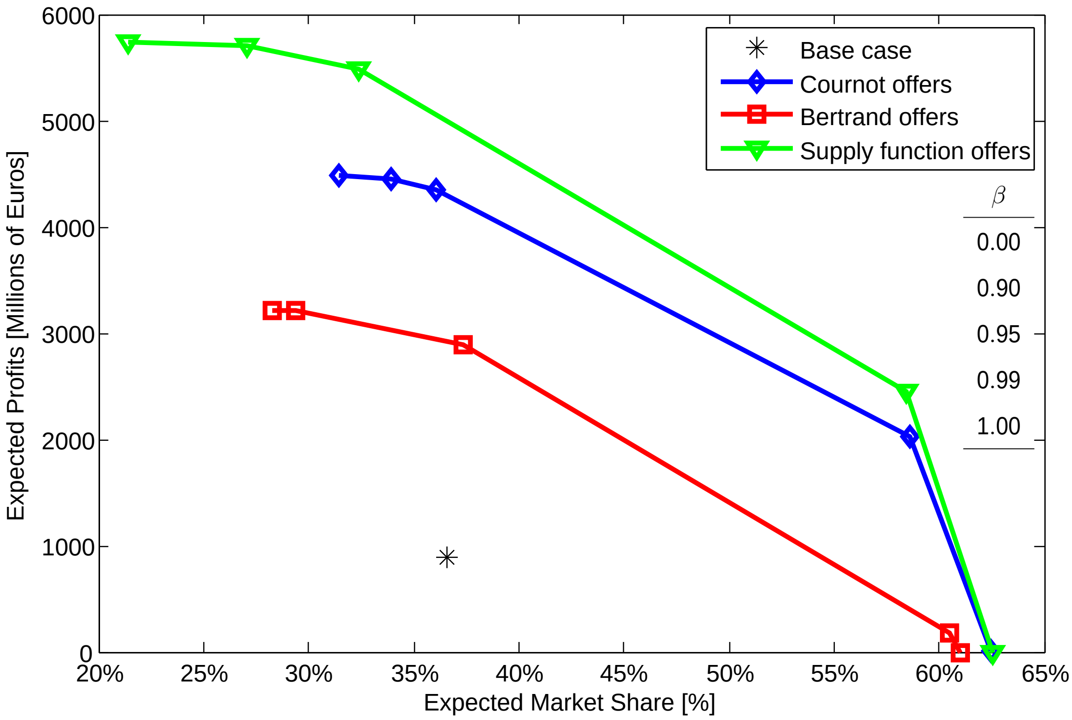


Figure 2.4: Expected pro t vs. expected market share frontier

By constructing the Pareto frontier, we can see that the dominant producer can signi cantly increase the expected pro t and/or the expected market share by bidding strategically. There is also, as expected, a clear trade-o between the two objectives in the sense that by increasing the expected market share of production, the expected profit falls and vice versa. It is interesting to compare the outcomes of the three strategic bidding approaches. By bidding strategic prices and quantities (supply function offers), the dominant producer can obtain higher profit than either bidding strategic quantities only (Cournot offers) or bidding strategic prices only (Bertrand offers), for the same levels of the market share. This can be explained by the fact that in supply function offers the dominant producer can optimize both price and quantity bids at the same time.

In order to illustrate the effect of dominant producer's strategic bidding on zonal prices, network flows and dispatch quantities, during peak hours (t_1) and high demand scenario (s_3), we compare the base case with the supply function approach for the value β = 0.95 in Figures 2.5 and 2.6. We can observe that, when bidding in a supply function manner, the dominant producer offers quantities that are less than its capacity and at higher prices than its marginal production cost across all zones in which the competitors' aggregate capacity is not sufficient to meet the demand (zones 1 to 4). The only exception is in zone 5, where the competitors' have a lower average marginal production cost than the dominant producer and also much higher production capacity (Table 2.1), which would satisfy the demand even at high levels. This would prevent the dominant producer to exert its market power. Therefore, the dominant producer offers all its capacity as quantity bid and at the same price as its marginal cost of generation in zone 5 in all demand scenarios, as shown in Figures 2.5 to 2.10.

The high demand for energy and the dominant producer' production withdrawal in zones 1 to 3 result in the dispatched quantities in these zones reaching the capacity levels for all producers ($Q^{str}_{kts} - Q_{kts}$ = 0 and $\overline{q}_{jts} - q_{jts}$ = 0). Mathematically, both λ^{+}_{jts} and α^{+}_{kts} in the complementarity constraints (2.3.12)-(2.3.17) take positive values, whereas λ^{-}_{jts} = α^{-}_{kts} = 0, which results in the zonal price π_{zts} to reach the cap level. On the other hand, the dispatch price remains low in zone 5, due to cheaper and excessive generation by the competitors and some of this cheap generation being transferred to zone 4 (reaching the capacity limit of the transmission line), and subsequently the zonal price remains well below the price cap.

In the base case, as opposed to the supply function approach, during the high demand period the zonal prices remain low and closer to the marginal generation cost. We can observe that the saturation of the transmission links leads to price differentiation in zones 4 and 5. Since the objective is to maximize both the expected pro t and the expected market share (with respect to β) the dominant producer's dispatch varies in terms of quantities when compared to the base case solution as β varies.

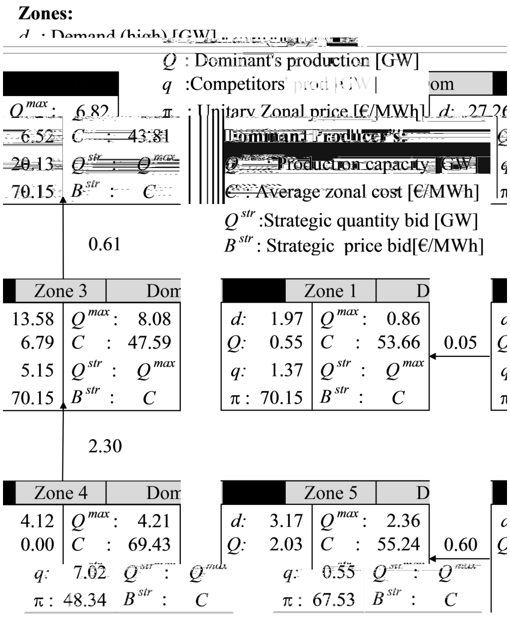


Figure 2.5: Base case solution for t_1, s_3

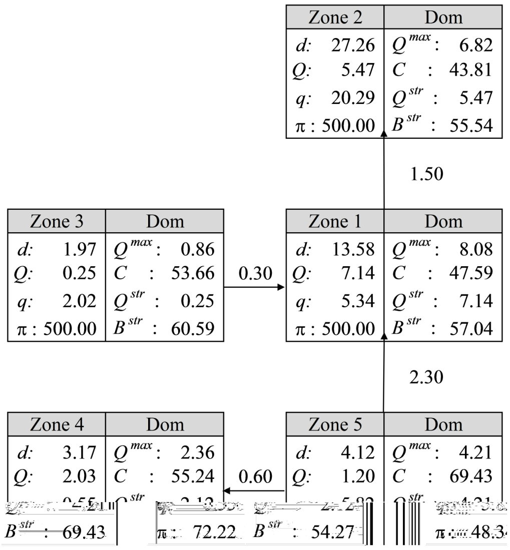


Figure 2.6: Supply function solution for t_1, s_3 and β = 0.95

The o -peak (t_8) and low demand (s_1) solutions of the same models are reported in Figures 2.7 and 2.8, which demonstrate that at lower demand level the dominant producer's strategic bids for quantity and price o er are closer to the base case, namely the capacity of its units and the marginal generation cost. Therefore, during low demand periods, the dominant producer is not able to fully exert market power, which leads to lower zonal prices.

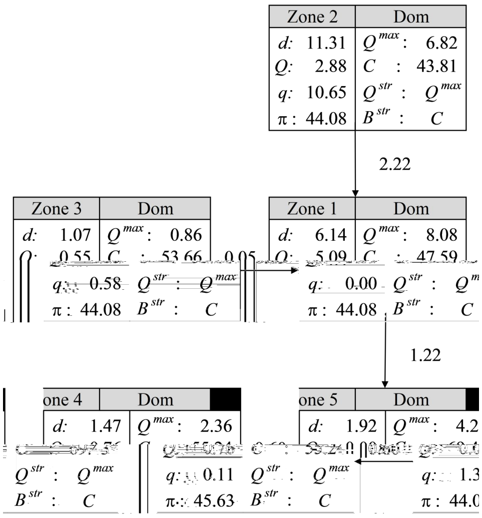


Figure 2.7: Base case solution for t_8, s_1

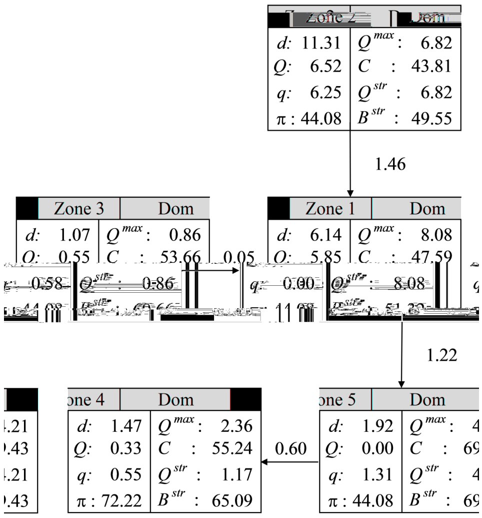


Figure 2.8: Supply function solution for t_8, s_1 and β = 0.95

Finally, in order to examine the system output sensitivity to the dominant producer's preferences, represented by the value of the weight β in the objective), we compare the supply function model at an average level demand period (t_5, s_2) for two extreme values of β, namely 0, corresponding to maximizing expected pro t only, and 1, corresponding to maximizing market share only. The results are reported in Figures 2.9 and 2.10.

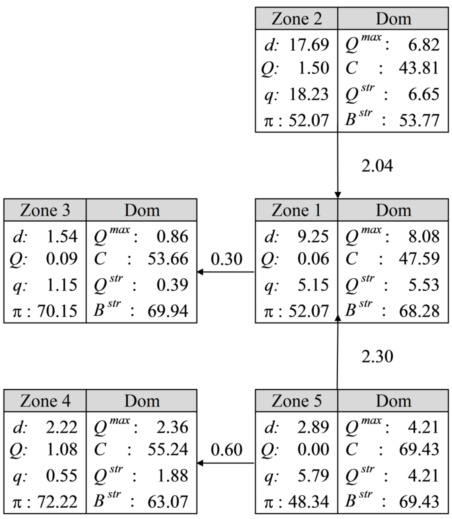


Figure 2.9: Supply function solution for t_5, s_2 and β = 0

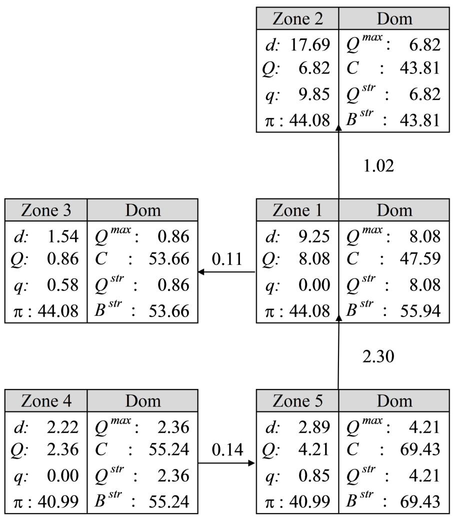


Figure 2.10: Supply function solution for t_5, s_2 and β = 1

It can be observed that when the dominant producer aims at maximizing expected profit only, the strategic bid prices are higher and the strategic bid quantities are lower, in order to force higher zonal prices and to obtain higher profits. On the other hand, when the objective is to maximize the market share, dominant producer bids the capacity of its units at lower prices, in order to fully utilize the generation capacity and obtain a higher market share.

All case studies have been solved using CPLEX 11 under GAMS. We have used a Dell PowerEdge R910 x64 computer with 4 processors at 1.87 GHz and 32 GB of RAM. Table 2.3 shows the running times and computational complexity required for solving the problems. The second to fifth columns show CPU times and computational complexity for each case study and only one point of the Pareto frontier. The CPU times shown in the second row are given by the average of all Pareto points of the frontier.

	Base Case	Cournot	Bertrand	Supply function
Avg. CPU time for 1 point of the pareto frontier	0.56 sec	123 min	19 sec	145 min
# of binary variables	0	3264	3264	3264
# of positive variables	1536	4984	4984	5168
# of free continuous variables	120	264	264	264
# of inequality constraints	1728	9976	10160	10344
# of equality constraints	144	1800	1800	1800

Table 2.3: CPU times and computational complexity of base case, Cournot, Bertrand and Supply function strategic offers

2.6 Conclusion

This chapter presents a multi-objective two-stage bilevel stochastic model for a dominant producer aiming at maximizing the expected market share and the expected profits in a pool-based markets for a mid-term horizon. The model is reformulated first as a multi-objective SMPEC and then as a mixed-integer linear programming problem suitable for application in large scale systems.

Three different strategic approaches have been proposed for the dominant producer, in order to achieve both objectives in a pool-based market: offering strategically the power generation (Cournot model); price bid (Bertrand model); and both price and quantity (supply function model).

The proposed model is then solved and the Pareto frontier is obtained. The Pareto frontier illustrates the trade-off between the dominant producer's conflicting interest in maximizing the expected profit and expected market share.

By analyzing the simulation results we draw the following conclusions.

- ⋆ The dominant producer can substantially increase expected profits and/or expected market share by behaving strategically when offering power production to the ISO.
- ⋆ The expected profit is higher when bidding in supply function manner than bidding only strategic price or quantity for the same levels of the market share.
- ⋆ In peak load demand hours, the dominant producer can exert market power by bidding higher energy prices and withdrawing some of its generation capacity, which leads to higher spot market prices. Conversely, in low demand periods the dominant producer has little influence on market clearing prices and hence the profit is similar to the base case.
- ⋆ The MILP formulation gives rise to computational efficiency, specially in the case of Bertrand approach (due to smaller solution space) .

Finally, the proposed two-objective model can be varied by looking at various other objectives such as risk measures, capacity expansion/investment, production scheduling and maintenance of generation units. We leave these topics for our future work.

Chapter 3

Robust Unit Commitment Problem

Chapter Abstract

The short term unit commitment and reserve scheduling decisions are made in the face of increasing supply side uncertainty in power systems. This has been mainly caused by a higher penetration of renewable generation that is encouraged and enforced by the market and policy makers. In this chapter, we propose a two-stage stochastic and distributionally robust modelling framework for the unit commitment problem under supply uncertainty. Based on the availability of the information on the distribution of the random supply, we consider two speci c models: i) a moment model where the mean values of the random supply variables are known, and ii) a mixture distribution model where the true probability distribution lies within the convex hull of a nite set of known distributions. In each case, we reformulate these models through dualization which leads to a semi-in nite program in the former case and a one-stage stochastic program in the latter case. We solve the reformulated models using sampling method and sample average approximation respectively. We also establish exponential rate of convergence of the optimal value when the randomization scheme is applied to discretize the semi-in nite constraints. The proposed robust unit commitment models are applied to an illustrative case study and numerical test results are reported in comparison with the two-stage non-robust stochastic programming model.

3.1 Introduction

The recent increase in the deployment of renewable energy resources such as wind power has a great impact on the short-term operational and long-term investment decisions in power systems due to their non-dispatchability and intermittent nature. In the short-term, the higher penetration of wind power and the lack of efficient storage facilities has an adverse effect on the stability of generation output. One of the most crucial decision problems that are affected by the short-term supply uncertainty is the unit commitment (UC) problem [112, 83]. The objective of the UC problem is to minimize the generation cost by determining the hourly unit commitment and the reserve schedule for the day-ahead given the demand and wind forecasts.

Classical models for the UC problem are often deterministic and consider supply and demand for electricity in the day ahead to be known in advance. Whilst the demand forecast for the day ahead can be reasonably estimated, the high reliance of the generation output on the unreliable wind power potentially makes optimal solutions of a deterministic model to be heavily infeasible or non-optimal under realized supply [80]. Probabilistic and robust optimisation models provide an alternative approach to incorporate the increased uncertainties associated with the wind and load forecasts into power system operations. In this sense, approaches to account for the uncertainty in renewable energy generation in the UC problem fall into three categories: (two-stage) stochastic programming, chance-constrained programming, and robust optimisation.

The first approach of two-stage stochastic optimisation [99] has been used widely for solving the UC problem [113, 84, 116], where energy and reserve generation are jointly scheduled to meet demand under stochastic wind supply. Chance-constrained programming is also proposed to deal with the jointly energy and reserve scheduling UC where one or several constraints must be satisfied with a given probability [86, 87]. One of the key assumptions in two-stage stochastic programming and chance-constrained programming is that the decision maker has complete information on the distribution of the uncertain parameters. However, limited predictability and high volatility of the renewable supply make this assumption non realistic.

In the third approach, classical robust optimisation, the distribution of the uncertain parameters is unknown except for its support. The support set contains all possible realizations of the unknown parameter and is often referred to as the *uncertainty set* [105, 26, 17]. From this, an equivalent deterministic problem can be derived. For the basic concept and a thorough survey of robust optimisation, we refer the interested readers to [71, 1, 13]. In the context of the UC problem, references [67, 68, 21] provide a robust optimisation formulation and an adaptive robust optimisation to address the wind power and demand uncertainty, respectively.

The min-max robust optimisation is often criticized for not utilizing partial information on the distribution of the uncertainty. *Distributionally robust optimisation* is part of the robust modeling framework where no assumption on the true probability distribution is made. But there is some knowledge of the underlying probability distribution.

There is an extensive research published on UC models where energy and reserve are scheduled together. Table 3.1 summarizes some of those references that are more closely related to the models proposed in this chapter.

Reference	Multiperiod	Network	Security	Model	Uncertainty	Available information
[28]	no	no	$n-2$	Det. with prob. constr.	contingencies	probability of outage
[5]	no	yes	$n-1$	Det. MILP	contingencies	set of plausible outages
[30]	no	yes	$n-1$	TS-SP and MILP ref.[y]	contingencies	probability of outage
[31, 32]	yes	yes	$n-1$	TS-SP and MILP ref.	contingencies	probability of outage
[29]	yes	no	$n-1$	TS-SP and MILP ref.	contingencies	probability of outage
					demand and wind	demand and wind pdf
[108]	yes	no	$n-K$	Det. with WC[x] contingency	contingencies	set of plausible outages
[88]	yes	no	$n-1$	TS-SP and MILP ref.	demand and wind	wind and demand pdf
[69]	yes	yes	$n-1$	TS-SP and MILP ref.	contingencies	probability of outage
[21]	yes	yes	$n-1$	Det. ARO[z]	demand	demand bound uncertainty
[86]	yes	no	$n-K$	CC-SP[♦] and MILP ref.	demand and wind	demand and wind pdf
Current work	yes	no	$n-1$	distributionally robust	wind	distributional information

$*\rightarrow$ Two-stage stochastic programming $\quad \dagger \rightarrow$ Mixed integer linear programming reformulation $\quad \S \rightarrow$ Worst-case

$\ddagger \rightarrow$ Adaptive robust optimisation $\quad \blacklozenge \rightarrow$ Chance-constrained stochastic programming

Table 3.1: A survey of UC models with security criteria

For a comparison between the existing literature and the proposed distributionally robust UC model in this chapter, we provide some criteria for classi cation as follows:

- The first classification criterion is based on the uncertainty sources and their available information. Contingencies, demand and wind production are some common sources of uncertainty in the UC problem with security criterion. Contingency events are usually considered as scenarios to include into a deterministic or stochastic UC constraints. Then, a bunch of post-contingency power flow operation equations should be included in the problem as presented in [28, 30, 5, 31, 32, 88, 69] to model that the system remain stable under the loss of one or more unit or line. Probability of the contingencies may be known (e.g. in [28, 30, 88, 69]) or may not be known (e.g. in [5, 108, 86]). References [29, 88, 86] include stochastic demand and wind with full knowledge of probability distributions of the uncertainties. Reference [21] includes stochastic demand with partial information, where the uncertainty set is defined as box uncertainty with a budget constraint.

  In this chapter, we model wind production as an uncertainty parameter with partial information on its probability distribution. Specifically, we propose two uncertainty sets: the first one is based on the first moment information and the second one is based on a family set of known distributions.

- The second classification criterion is based on the system reliability considerations. Current reliability policy and associated security standards in power systems mainly focus on contingencies caused by outage of transmission or generation assets. The purpose of security criteria is to keep the system stable in case of one ($n-1$ criterion) or more ($n-K$ criteria) outages of a generating unit or line, where reserve planning is justified to compensate for possible outages.

  The $n-1$ criterion has been extensively applied to UC problem [5, 30, 31, 32, 29, 88, 69]. Some studies extend the security criterion up to K simultaneous contingencies [108, 86]. However, more strict criteria increases the complexity of the model and its tractability.

  We propose a model where a $n-1$ security criterion is included in the sense that if one unit is loss, the demand is met with the scheduled reserve from the first stage under any wind scenario.

- The third criterion is based on the modeling of contingencies. Outages can be treated as deterministic parameters and the preventive actions are taken pre-contingency and through inclusion of deterministic constraints. These security constraints will ensure that enough resources for the normal operation of the system in the event of a contingency (see [5]). On the other hand, outages can be treated as stochastic parameters, then the objective function includes expected value of the second-stage corrective costs. Contingencies probabilities should be known for this model, and the set of post-contingencies equations, one for each outage, are included into the model (see [29, 31]). Note that, when others sources of uncertainty exists, post-contingencies equations should be extended for each scenario, which may leads into an intractable problem. Because of that, some authors limit the scenarios of contingencies to an umbrella of credible contingencies [5, 30, 31, 32, 69, 29]. Another approach to deal with security criteria is posing an optimisation problem to determinate the worst contingency/contingencies. The works [108] and [86] propose a worst-case optimisation problem embedded into a deterministic and chance-constrained UC model, respectively.

In this chapter, we formulate the security criterion as a deterministic constraint for the worst-case outage. We show that this worst-case outage is always the worst for any wind scenario. In this sense, we do not need to add up post-contingency equations for each plausible outage (all generating units) and each wind scenario. There is no cost term in the objective function to account for the extra cost of the corrective actions in the event of a contingency. This is a reasonable approach because contingencies has a very low probability of occurrence.

- A fourth criterion is based on the model formulation and solution approach. References [28, 5] propose a deterministic modelling framework, [29, 31] formulate a two-stage optimisation problem, [86] formulates a chance-constrained problem, and [108, 21] present a robust optimisation framework. Most of these models are solved by driving their deterministic counterparts and later reformulating them as MILPs.

  In this chapter, we propose a distributionally robust optimisation formulation which is then reformulated as an MILP using duality theory and sampling.

Paper [21] proposes a two-stage deterministic robust optimisation UC model, where the uncertainty set is de ned through a deterministic set. The solution of the proposed adaptive robust model provides immunity against all realizations of the uncertain data within the deterministic uncertainty set. However, this robust model does not take into account the distributional information of the random variables. In contrast, our model accounts for the partial/available information on the probability distribution of the uncertain data. Furthermore, in [21], the rst stage decision variables consist of the on/o commitment variables and the second stage solves an economic dispatch, therefore, they do not need to schedule reserve for the second-stage uncertainty deviations (North American's UC outlook). However, in our model, the rst stage decision variables are the set of on/o decisions and scheduled energy and reserve (European's UC outlook). Scheduled reserve is used in the second stage as corrective actions to meet the demand under wind production deviations and/or the outage of one generating unit.

Despite the fact that there is a rich body of literature focusing on two-stage stochastic and robust UC with endogenous reserve scheduling and wind generation models, a distributionally robust UC approach has not been presented yet. The main contributions of this chapter are summarized as follows:

- We propose a distributionally robust UC model to deal with day ahead wind uncertainty and robust $n-1$ security criteria.
- We consider two uncertainty sets based on the available information to model the stochastic wind. This leads to two speci c models: i) a moment model where the mean values of the random supply variables are known, and ii) a mixture distribution model where the true probability distribution lies within the convex hull of a nite set of known distributions.
- We reformulate the robust model with moment condition as a semi-in nite program through duality. Moreover, we develop a randomization scheme for solving the Lagrange dual of the robust optimisation problem. To show the convergence of the randomized problem, we consider a general mathematical program with semi-in nite constrains and establish exponential rate of convergence of the optimal value when the randomization scheme is applied to discretize the semi-in nite constraints.
- We also reformulate the mixture model as a one-stage stochastic program through duality and develop a practical solution method based on SAA to reformulate the problem as an MILP.

- We provide numerical results for an illustrative case study. We analyse and compare both approaches against a two-stage (non-robust) stochastic UC model. Specifically, we analyse the sensibility of the solutions to the variation of the mean and covariance.

The remainder of this chapter is organized as follows. In Section 3.2, deterministic UC model is introduced first, then, a two-stage UC model is presented under wind generation uncertainty and finally a distributionally robust UC is presented. In Section 3.3 and 3.4 describe the robust UC solution methodologies based on uncertainty generated with moments information and mixture of distributions, respectively. Section 3.5 presents several case studies comparing the proposed approaches. The main conclusions are summarized in Section 3.6.

3.2 Mathematical formulation

In the most basic form, the unit commitment problem involves the system operator to find the optimal schedule and the generation level for a set of conventional generating units over a planning horizon in order to meet the demand. Under a deterministic framework, the demand and supply parameters are assumed to be known and given in advance. There is a vast literature on deterministic unit commitment models, see [83] for a comprehensive survey. Consider a set $\mathcal{I} = \{1, \cdots, I\}$ of conventional generating units, indexed by i, and let $\mathcal{T} = \{1, \cdots, T\}$ be the set of time periods in the planning horizon (i.e. 24 hours of the day ahead), indexed by t. The deterministic demand for each time period is given as d_t. The system operator aims to minimize the total generation cost by planning a schedule consisting of the on/off decision for each generator in each time period defined by binary variables

$$u_{it} = \begin{cases} 1, & \text{if generator } i \text{ is on in time period } t, \\ 0, & \text{otherwise.} \end{cases}$$

The energy dispatched by the generator i at time t is defined by continuous variable q_{it}. Each generator $i \in \mathcal{I}$ has a fixed on cost of c_i^f and, if on, it has a unit generation cost of c_i^l. The upper and lower generation capacity of generator i is given by $\overline{q}_i$ and $\underline{q}_i$

respectively. The basic deterministic UC problem is then formulated as follows

$$
\begin{aligned}
\min_{\mathbf{u},\mathbf{q},\mathbf{r}} \quad & \sum_i \sum_t \left[c_i^f u_{it} + c_i^l q_{it}\right] & & \\
\text{s.t.} \quad & \sum_i q_{it} = d_t, & \forall\, t \in \mathcal{T}, & \quad (3.2.1)\\
& q_{it} \leq \overline{q}_i u_{it}, & \forall\, t \in \mathcal{T}, i \in \mathcal{I}, & \quad (3.2.2)\\
& q_{it} \geq \underline{q}_i u_{it}, & \forall\, t \in \mathcal{T}, i \in \mathcal{I}, & \quad (3.2.3)\\
& q_{it} \in \mathbb{R}^+,\ u_{it} \in \{0, 1\}, & \forall\, t \in \mathcal{T}, i \in \mathcal{I}, & \quad (3.2.4)
\end{aligned}
$$

where the balance constraint (3.2.1) ensures the forecasted demand is met at all time periods and constraints (3.2.2) and (3.2.3) impose the upper and lower generation limits of every unit for all time periods.

In the remaining of this section, we rst introduce the two-stage stochastic unit commitment problem (Sto-UC) taking into account the wind uncertainty and reliability criteria. Since it is not often possible to have complete information on the probability distribution of the wind supply, we extend the stochastic framework into a two-stage distributionally robust model. Most of the notations used in this chapter are introduced in this section.

3.2.1 Two-stage stochastic UC with uncertain net load

The unit commitment and generation scheduling problems involve inherent uncertainties stemming from the short-term volatility of demand and unpredictability of wind power. The recent progress in the eld of stochastic programming makes it an attractive approach for modelling the UC problem under uncertainty. Research carried out in [36, 112, 45] were amongst the rst that formulate the UC problem as a two-stage stochastic program. Under the stochastic framework, adequate reserves are scheduled in advance and used to hedge against the uncertainty in future demand and/or supply.

In what follows, we rst de ne operational characteristics of the power system including the reserve scheduling, the reliability criteria and the ramping constraints. We then present the UC problem under the wind generation uncertainty as a two-stage stochastic program. Assuming that the day-ahead demand load is known, we denote the stochastic net load with a random vector $\mathbf{l}(\xi)$ which is given by the di erence between the known demand $\mathbf{d}$ and stochastic wind power generation $\mathbf{w}(\xi)$ at any time period

$$l_t(\xi) = d_t - w_t(\xi), \qquad \forall\, t \in \mathcal{T}.$$

For the convenience of notation, we use $\mathbf{l}(\xi)$ and $\boldsymbol{\xi}$ interchangeably, i.e. both l_t and ξ_t refer to the stochastic net load at time period t. Mathematically $\boldsymbol{\xi}$ is a random vector de ned over measurable space ($, \mathcal{F})$ with sigma-algebra. We use to denote the support set of $\boldsymbol{\xi}$. Obviously in this context, we can assume that is a compact set (bounded and closed).

The rst-stage (*here and now*) decisions are determined prior to realization of uncertain net load $\boldsymbol{\xi}$. These include the on/o decision variable u_{it}, for generator i and time period t, the energy dispatched variable q_{it}, for the generator i at time t, and the up and down scheduled reserves for generator i at time t denoted as r_{it}^{up}, r_{it}^{dw} respectively. The upper and lower limits for reserve up/down for generator i are given by $\overline{r}_{it}^{up}/\overline{r}_{it}^{dw}$ and $\underline{r}_{it}^{up}/\underline{r}_{it}^{dw}$ respectively. The unitary cost of scheduling reserve up $c_i^{r,up}$ and down $c_i^{r,dw}$ for generator i are given as input of the problem.

The second stage (*wait and see*) after the realization of the uncertainty in net load $\boldsymbol{\xi}$ includes decision variables on the actual up and down deployed reserves, denoted by $\hat{r}_{it}^{up}(\xi), \hat{r}_{it}^{dw}(\xi)$. The unitary cost for the actual deployment of the reserve up and down are given by $\hat{c}_i^{up}$ and $\hat{c}_i^{dw}$. To avoid unbalance in the supply and demand for energy in the second stage, we introduce additional auxiliary variables for load shedding and wind spillage. The load shedding variable is denoted by $S_t(\xi)$ and represents the excess demand which can not be met by the total generation output at time period t and has to be shed at high penalty cost of c_t^{ls} . On the other hand, the wind spillage variable, denoted by $W_t(\xi)$, is equal to the excess wind power that can not be utilized upon the realization of the net load. Wind spillage incurs an unitary opportunity cost of c_t^{ws}. Figure 3.1 gives a graphical interpretation of the scheduling up and down reserves to meet the demand under any net load deviation.

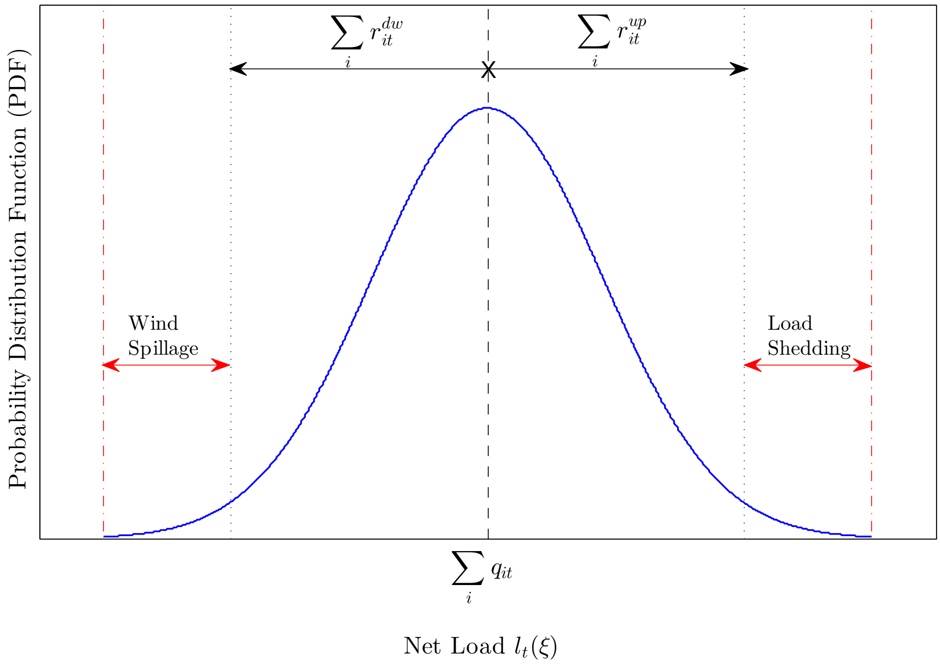


Figure 3.1: Net load distribution and scheduled production and reserve for time t

Figure 3.1 gives a graphical interpretation of the scheduling up and down reserves to meet the demand under any net load deviation. Note that for illustration purpose, it was assumed that the net load distribution $l_t(\xi)$ function follows a normal distribution.

Note that for illustration purpose, it was shown in Figure 3.1 that the net load distribution $l_t(\xi)$ function follows a normal distribution. However, in practice net load could follow other distributions.

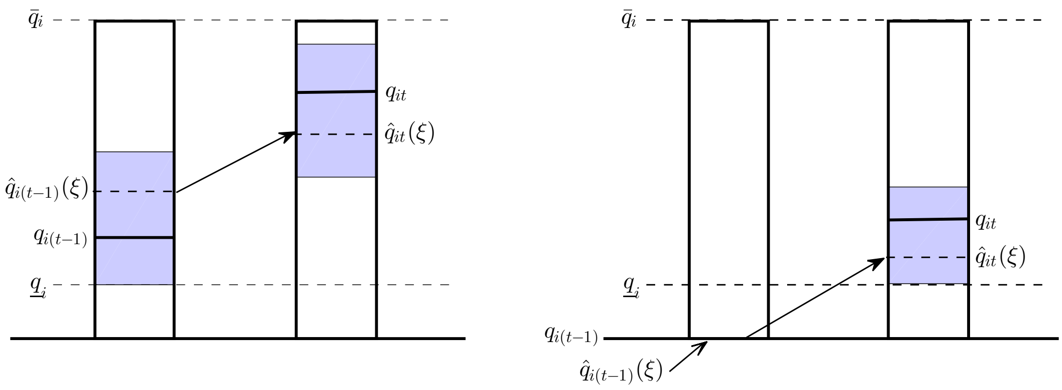
Ramp constraints

For conventional generation units, it is important to take into account the short-term dynamics of generation output for the consecutive periods. Such requirements are often referred to as ramp constraints which limit the maximum increase or decrease of generated power from one time period to the next, re ecting the thermal and mechanical inertia that has to be overtaken in order for the generating unit to increase or decrease its output.

In a two-stage stochastic framework, we need to de ne the ramp constraints for any possible realization of the net load ξ. At the second stage and after the realization of uncertain net load, the scheduled energy q_{it} does not change and the generating units adapt their production to accommodate the realized net load. To do this, up and down reserves ($\hat{r}^{up}_{it}(\xi)$, $\hat{r}^{dw}_{it}(\xi)$) are used. We denote the actual power output of unit i at time t and scenario ξ as $\hat{q}_{it}(\xi)$ which is given by

$$\hat{q}_{it}(\xi) = q_{it} + \hat{r}^{up}_{it}(\xi) - \hat{r}^{dw}_{it}(\xi), \qquad \forall t, i, \xi.$$

Note that, up and down reserves should not be simultaneously positives. We consider the ramps constraints for all four combinations of on-o decisions between any two consecutive periods $t-1$ and t as follows.



(a) $u_{i(t\ 1)}$ = 1 and u_{it} = 1 (b) $u_{i(t\ 1)}$ = 0 and u_{it} = 1

Figure 3.2: Ramp limitation

1. If $u_{i(t-1)}$ = 1 and u_{it} = 1, the generating unit i is coupled for both hours and ramp limitations in hour t are bounded by the ramp up and down rates, RU and RD respectively (see Figure 3.2a). The ramp equations will then be as follows:

$$\hat{q}_{it}(\xi) - \hat{q}_{i(t-1)}(\xi) \leq RU_i, \qquad \forall t, i, \xi, \tag{R1}$$

$$\hat{q}_{i(t-1)}(\xi) - \hat{q}_{it}(\xi) \leq RD_i, \qquad \forall t, i, \xi. \tag{R2}$$

2. If $u_{i(t-1)}$ = 0 and u_{it} = 1, generating unit i starts up at the beginning of period t and hence the limitation for hour t should be the starting up ramp SU (see Figure

3.2b). The ramp equations will then be as follows

$$\hat{q}_{it}(\xi) - \hat{q}_{i(t-1)}(\xi) \leq SU_i, \qquad \forall t, i, \xi, \tag{R1}$$

$$\hat{q}_{i(t-1)}(\xi) - \hat{q}_{it}(\xi) \leq RD_i, \qquad \forall t, i, \xi. \tag{R2}$$

Note that, (R2) always holds because $\hat{q}_{i(t-1)}(\xi)$ = 0, $\hat{q}_{it}(\xi) \geq$ 0 and $RD_i \geq$ 0.

3. If $u_{i(t-1)}$ = 1 and u_{it} = 0, generating unit i shutdowns at the beginning of time period t and hence limitation during the hour t

Having de ned the ramp constraints for all combinations of on/o decisions between two time period, we can generalize them in relations to decision variable u_{it} as follow:

$$
\begin{aligned}
&\hat{q}_{it}(\xi) - \hat{q}_{i(t-1)}(\xi) \le RU_i u_{i(t-1)} + SU_i(1 - u_{i(t-1)}), && \forall t, i, \xi, \\
&\hat{q}_{i(t-1)}(\xi) - \hat{q}_{it}(\xi) \le RD_i u_{it} + SD_i(1 - u_{it}), && \forall t, i, \xi.
\end{aligned}
$$

Reliability and $n-1$ security constraint

To ensure the system reliability under the failure of up to one scheduled generator, we consider the worst-case scenario that the system could face; That is, for any given time period, the generator with the highest total scheduled generation and actual up reserve to fail under the lowest level of wind (highest net load) available. This can be modeled as

$$\sum_i (q_{it} + r^{up}_{it}) - \max_i\{q_{it} + \hat{r}^{up}_{it}(\xi)\} + S_t(\xi) \ge d_t - \min_\xi\{w_t(\xi)\}, \qquad \forall t, \xi. \tag{3.2.5}$$

Let us denote the worst (lowest) realization of wind output at time t by

$$\underline{\xi}_t = \arg\min_\xi\{w_t(\xi)\},$$

and let $S_t(\underline{\xi})$ denote the corresponding load shedding for this realization. Since the actual up reserve used has to be less than the scheduled up reserve, i.e. $\hat{r}^{up}_{it}(\xi) \le r^{up}_{it}$ for every scenario ξ, the inequality (3.2.5) can be rewritten as

$$\sum_i (q_{it} + r^{up}_{it}) - \max_i\{q_{it} + r^{up}_{it}\} + S_t(\underline{\xi}) \ge d_t - w_t(\underline{\xi}), \qquad \forall t. \tag{3.2.6}$$

Lets denote the upper limit of total generation and reserve schedules at time t as follow

$$Q_t = \sum_i (q_{it} + r^{up}_{it}), \qquad \forall t.$$

The inequality (3.2.6) can be rewritten as

$$Q_t - (q_{it} + r^{up}_{it}) + S_t(\underline{\xi}) \ge l_t(\underline{\xi}), \qquad \forall t, i.$$

Two-stage stochastic model

The resulting mathematical model for the two-stage stochastic UC problem with $n-1$ security criterion and ramping constraints is given as follows

$$
\text{(Sto-UC)} \quad \min_{\mathbf{u},\mathbf{q},\mathbf{r}} \quad \sum_i \sum_t \Big[\underbrace{c_i^f u_{it} + c_i^l q_{it}}_{\text{Generation Cost}} + \underbrace{c_i^{r,up} r_{it}^{up} + c_i^{r,dw} r_{it}^{dw}}_{\text{Reserves Cost}} \Big] + \underbrace{\mathbb{E}_P[g(\mathbf{u},\mathbf{q},\mathbf{r},\boldsymbol{\xi})]}_{\text{Expected variation cost}}
$$

$$
\begin{aligned}
\text{s.t.} \quad & q_{it} + r_{it}^{up} \leq \overline{q}_i u_{it}, \ \forall t, i, && (3.2.7)\\
& q_{it} - r_{it}^{dw} \geq \underline{q}_i u_{it}, \ \forall t, i, && (3.2.8)\\
& \underline{r}_i^{up} u_{it} \leq r_{it}^{up} \leq \overline{r}_i^{up} u_{it}, \ \forall t, i, && (3.2.9)\\
& \underline{r}_i^{dw} u_{it} \leq r_{it}^{dw} \leq \overline{r}_i^{dw} u_{it}, \ \forall t, i, && (3.2.10)\\
& q_{it}, r_{it}^{dw}, r_{it}^{up} \in \mathbb{R}^+, \ u_{it} \in \{0,1\}, \ \forall t, i, && (3.2.11)
\end{aligned}
$$

where $\mathbb{E}_P$ denotes the mathematical expectation taken w.r.t. the distribution of $\boldsymbol{\xi}$. over probability space $(P, \ , \mathcal{F})$, and $g(\mathbf{u},\mathbf{q},\mathbf{r},\boldsymbol{\xi})$ is the optimal value of the second-stage problem and is defined as

$$
g(\mathbf{u},\mathbf{q},\mathbf{r},\boldsymbol{\xi}) = \min_{\hat{\mathbf{r}},\mathbf{W},\mathbf{S}} \quad \sum_i \sum_t \left\{ \hat{c}_i^{up} \hat{r}_{it}^{up}(\xi) + \hat{c}_i^{dw} \hat{r}_{it}^{dw}(\xi) \right\} + \sum_t \left\{ c_t^{ls} S_t(\xi) + c_t^{ws} W_t(\xi) \right\}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_i \hat{q}_{it}(\xi) + S_t(\xi) - W_t(\xi) = l_t(\xi), \ \forall t, && (3.2.12)\\
& 0 \leq \hat{r}_{it}^{up}(\xi) \leq r_{it}^{up}, \ \forall t, i, && (3.2.13)\\
& 0 \leq \hat{r}_{it}^{dw}(\xi) \leq r_{it}^{dw}, \ \forall t, i, && (3.2.14)\\
& Q_t - (q_{it} + r_{it}^{up}) + S_t(\underline{\xi}) \geq l_t(\underline{\xi}), \ \forall t, i, && (3.2.15)\\
& \hat{q}_{it}(\xi) - \hat{q}_{i(t-1)}(\xi) \leq RU_i u_{i(t-1)} + SU_i(1 - u_{i(t-1)}), \ \forall t, i, && (3.2.16)\\
& \hat{q}_{i(t-1)}(\xi) - \hat{q}_{it}(\xi) \leq RD_i u_{it} + SD_i(1 - u_{it}), \ \forall t, i. && (3.2.17)
\end{aligned}
$$

In the first stage, constraints (3.2.7) and (3.2.8) represent the generation limits including the up and down scheduled reserves. The equations (3.2.9) and (3.2.10) bound the minimum and maximum reserves to schedule.

In the second stage, constraint (3.2.12) represents the energy balance for each hour t and net load wind scenario ξ. Constraints (3.2.13) and (3.2.14) ensure that the actual up and down reserves used are within the limits of the nominal reserve scheduled in the rst stage. Furthermore, equation (3.2.15) represents the $n-1$ reliability constraint and ensures that the demand is met under the failure of up to one generating unit. Finally, constraints (3.2.16) and (3.2.17) represent the ramp constraint as de ned above.

3.2.2 Distributionally robust UC problem

Under the stochastic programming framework (Sto-UC), we assume that the \true" probability distribution P of the random wind variables is known. In practice, however, such distribution is often unknown or estimated through partial information and subjective judgements. One of the possible ways to deal with this issue is to use available information to construct a set of distributions, denoted as $\mathcal{P}$, in which the true probability distribution is assumed to be contained. The robust optimisation approach to the two-stage stochastic problem with respect to this ambiguity aims to make decisions that is optimal for the worst probability distribution from $\mathcal{P}$. The corresponding mathematical formulation is

$$\begin{aligned} \text{(R-UC)} \qquad & \min_{x} \quad c^T x + \sup_{P\in\mathcal{P}} \mathbb{E}_P[g(x,\boldsymbol{\xi})] \\ & \text{s.t.} \quad x \in \mathcal{X}, \end{aligned} \tag{3.2.18}$$

where $g(x,\boldsymbol{\xi})$ is the optimal value of the second-stage problem

$$\begin{aligned} g(x,\boldsymbol{\xi}) = \ & \min_{y} \quad h^T y \\ & \text{s.t.} \quad y \in \mathcal{Y}(x,\boldsymbol{\xi}), \end{aligned} \tag{3.2.19}$$

where $\mathcal{Y}(x,\boldsymbol{\xi})$ is the second-stage feasible set which is dependent on the rst-stage decision x and the realization $\boldsymbol{\xi}$. For convenience in notation we refer to the set of the rst-stage decision variables as $x = (\mathbf{u},\mathbf{q},\mathbf{r})$ and the feasible domain of x in the rst stage to be $\mathcal{X}$. Likewise we refer to the second-stage variables as $y = (\hat{\mathbf{r}},\mathbf{W},\mathbf{S})$. Furthermore, we refer to the rst-stage cost parameters as $c = (\mathbf{c}^f,\mathbf{c}^l,\mathbf{c}^{r,up},\mathbf{c}^{r,dw})$ and the second-stage cost parameters as $h = (\hat{\mathbf{c}}^{up},\hat{\mathbf{c}}^{dw},\mathbf{c}^{ls},\mathbf{c}^{ws})$.

In the literature of robust optimisation, (3.2.18) is known as a distributionally robust optimisation problem where the robustness is taken w.r.t. the probability distribution rather than the scenarios of the random vector $\boldsymbol{\xi}$. The set $\mathcal{P}$ is called ambiguity set.

The first-stage distributionally robust decisions are related to the energy and reserve scheduling and on/off commitment decisions. The second-stage distributionally robust UC decisions are related to the corrective actions to meet the net load. Corrective actions could be preventive as deploying the already scheduled reserve at the first stage or emergency corrective actions such as load shedding or wind spillage. UC solutions from the first stage remain feasible for any realization of the wind uncertainty that belongs to a family of distributions $\mathcal{P}$.

The uncertainty set $\mathcal{P}$ can be defined in various ways depending on the availability of information. We consider two approaches; one is to use the moment information where moments of $\boldsymbol{\xi}$ is assumed known. The other is to use mixture distribution with a set of perceived distributions.

3.3 A semi-infinite programming approach

In this section, we investigate the robust unit commitment problem where the first moment conditions of the underlying random wind, and therefore the net load, variables are known. In the case of uncertain wind, the exact probability distribution is often unknown and partial information such as the first moment condition is more readily available in practical situations. Let $\boldsymbol{\mu} \in \mathbb{R}^T$ denote the mean of $\boldsymbol{\xi}$. We consider the ambiguity $\mathcal{P}$ being defined as follows:

$$\mathcal{P} = \{P \in \mathscr{P} : \mathbb{E}_P[\boldsymbol{\xi}] = \boldsymbol{\mu}\}, \tag{3.3.1}$$

where $\mathscr{P}$ denotes the set of all probability measures of measurable space $(\ , \mathscr{B})$ induced by $\boldsymbol{\xi}$. For each fixed $x \in \mathcal{X}$, we consider the worst expected value of $g(x, \boldsymbol{\xi})$ over the ambiguity set $\mathcal{P}$:

$$H(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[g(x, \boldsymbol{\xi})]. \tag{3.3.2}$$

Using the moment conditions, we can write $H(x)$ as the optimal value of the following maximization problem

$$\begin{aligned} H(x) = \sup_{P \in \mathscr{M}^+} \quad & \int_{\Xi} g(x, \boldsymbol{\xi}) P(d\boldsymbol{\xi}), \\ \text{s.t.} \quad & \int_{\Xi} \xi_t P(d\boldsymbol{\xi}) = \mu_t, \forall t = 1, \cdots, T, \\ & \int_{\Xi} P(d\boldsymbol{\xi}) = 1, \end{aligned} \tag{3.3.3}$$

where $\mathscr{M}^+$ denotes the set of all non-negative finite measures on measurable space $(\Xi, \mathscr{B})$, μ_t is the t^{th} component of $\boldsymbol{\mu}$. Problem (3.3.3) is a typical form of a classical moment problem. We refer interested readers to monograph [73] for a comprehensive discussion of the historical background of the latter. Note that mathematically (3.3.3) is a semi-infinite linear program because the optimisation space is finite dimensional while the number of constraints is infinite. In order to deal with difficulties associated to solving such infinite dimensional problem duality theory is often used; see for example [91]. Here we follow [96, Proposition 3.1] to derive the Lagrange dual associated with the moment problem (3.3.3).

Proposition 3.3.1. *For a given $x \in \mathcal{X}$, the Lagrange dual of problem (3.3.2) is*

$$\begin{aligned} H_D(x) := \min_{\boldsymbol{\alpha}} \quad & \alpha_0 + \sum_{t=1}^{T} \alpha_t \mu_t \\ \text{s.t.} \quad & g(x, \boldsymbol{\xi}) \leq \alpha_0 + \sum_{t=1}^{T} \alpha_t \xi_t, \forall \boldsymbol{\xi} \in \Xi, \end{aligned} \tag{3.3.4}$$

where $\alpha_t \in \mathbb{R}$, $t = 0, 1, \cdots, T$, denotes the dual variables corresponding to moment problem constraints and $\Xi \subset \mathbb{R}^T$ is the support set of $\boldsymbol{\xi}$, and $H_D(x) = H(x)$.

Proof: The derivation of the dual formulation can be found in [99]. Here we include some details for completeness. Let

$$\begin{aligned} \mathcal{L}(x, P, \alpha) := & \int_{\Xi} g(x, \xi) P(d\xi) + \alpha_0 \Big(1 - \int_{\Xi} 1 P(d\xi)\Big) + \sum_{t=1}^{T} \alpha_t \Big(\mu_t - \int_{\Xi} \xi_t P(d\xi)\Big) \\ = & \int_{\Xi} \left(g(x, \xi) - \alpha_0 - \sum_{t=1}^{T} \alpha_t \xi_t \right) P(d\xi) + \alpha_0 + \sum_{t=1}^{T} \alpha_t \mu_t, \end{aligned}$$

where $\alpha_t \in \mathbb{R}$ for $t = 0, 1, \cdots, T$. The Lagrange dual of (3.3.3) is

$$H_D(x) = \min_{\boldsymbol{\alpha}} \sup_P \mathcal{L}(x, P, \alpha). \tag{3.3.5}$$

If there exist an $\boldsymbol{\xi}^* \in$ such that

$$g(x, \boldsymbol{\xi}^*) - \alpha_0 - \sum_{t=1}^{T} \alpha_t \xi_t^* > 0.$$

Then we can take $P := \sigma I(\xi^*)$, where $\sigma > 0$ is a positive constant and $I(\cdot)$ takes value 1 at $\boldsymbol{\xi}^*$ and 0 otherwise. Hence

$$\int_{\Xi} \left(g(x, \xi) - \alpha_0 - \sum_{t=1}^{T} \alpha_t \xi_t \right) P(d\xi) = \sigma \left(g(x, \boldsymbol{\xi}^*) - \alpha_0 - \sum_{t=1}^{T} \alpha_t \xi_t^* \right).$$

The R.H.S of the equation above can be arbitrarily large as σ can take any positive value. Consequently, it can be veri ed that

$$\max_{P \in \mathcal{P}} \mathcal{L}(x, P, \alpha) := \begin{cases} \alpha_0 + \sum_{t=1}^{T} \alpha_t \mu_t, & \text{if} \quad g(x, \boldsymbol{\xi}) - \alpha_0 - \sum_{t=1}^{T} \alpha_t \xi_t \leq 0, \ \forall \boldsymbol{\xi}, \\ +\infty, & \text{otherwise}, \end{cases}$$

which yields (3.3.4).

To complete the proof, let us now show $H(x) = H_D(x)$. Observe rst that for any xed $x \in \mathcal{X}$, since is compact, $H(x)$ is nite. Moreover, with the compactness of , it follows by [109, Remark 2] that $\mathcal{P}$ is compact (closed and tight) and thereby the set

$$\{\mathbb{E}_P[g(x, \boldsymbol{\xi})], \mathbb{E}_P[\boldsymbol{\xi}_t] - \mu_t : P \in \mathcal{P}\}$$

is a compact set. By [96, Proposition 3.1], $H(x) = H_D(x)$. □

Using Proposition 3.3.1, we can reformulate (R-UC) as

$$\begin{aligned} \text{(SIP-UC)} \qquad & \min_{x, \alpha_0, \boldsymbol{\alpha}} \quad c^T x + \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\mu} \\ & \text{s.t.} \quad x \in \mathcal{X}, \\ & \qquad\quad g(x, \boldsymbol{\xi}) \leq \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\xi}, \forall \boldsymbol{\xi} \in \ , \end{aligned} \tag{3.3.6}$$

under the moment conditions. Moreover, it is easy to verify that (3.3.6) is equivalent to the following program which incorporate the details of the second-stage problem

$$\begin{aligned} \min_{x,y(.),\alpha_0,\boldsymbol{\alpha}} \quad & c^T x + \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\mu} \\ \text{s.t.} \quad & x \in \mathcal{X}, \\ & h^T y(\boldsymbol{\xi}) \leq \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\xi}, \forall \boldsymbol{\xi} \in \Xi, \\ & y(\boldsymbol{\xi}) \in \mathcal{Y}(x, \boldsymbol{\xi}), \forall \boldsymbol{\xi} \in \Xi. \end{aligned} \tag{3.3.7}$$

3.3.1 Sample approximation scheme

One of the well-known solution approaches for semi-infinite programs is random discretization. The basic idea is to construct a tractable sub-problem by considering a randomly drawn finite subset of constraints. The approach has been shown to be numerically efficient and it has been widely applied to various stochastic and robust programs, see [33]. In a more recent development, Anderson et al [2] proposed a CVaR approximation scheme to a semi-infinite constraint system and then apply the well known sample average approximation to the CVaR (of the constraint function), see also Liu and Xu [76] where the approach is applied to mathematical programs with semi-infinite complementarity constraints.

Let $\xi^1, \cdots, \xi^S$ be random variables which are independent and follow identical distribution to that of $\boldsymbol{\xi}$. Let $\mathcal{S} = \{1, \cdots, S\}$. We consider the following discretization problem as an approximation to (3.3.7):

$$\begin{aligned} \min_{x,y(\xi^s):s\in\mathcal{S},\alpha_0,\boldsymbol{\alpha}} \quad & c^T x + \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\mu} \\ \text{s.t.} \quad & x \in \mathcal{X}, \\ & h^T y(\xi^s) \leq \alpha_0 + \boldsymbol{\alpha}^T \xi^s, \forall s \in \mathcal{S}, \\ & y(\xi^s) \in \mathcal{Y}(x, \xi^s), \forall s \in \mathcal{S}, \end{aligned} \tag{3.3.8}$$

where ξ^s is a realization of $\boldsymbol{\xi}$ for $s \in \mathcal{S}$. In what follows, we show the convergence of the optimal value of (3.3.8) to its true counterpart as the sample size increases. To this end, we consider an equivalent form of (3.3.8) which is presented in terms of the optimal value function of the second-stage problem:

$$\begin{aligned} \min_{x,\alpha_0,\boldsymbol{\alpha}} \quad & c^T x + \alpha_0 + \boldsymbol{\alpha}^T \boldsymbol{\mu} \\ \text{s.t.} \quad & x \in \mathcal{X}, \\ & g(x, \xi^s) \leq \alpha_0 + \boldsymbol{\alpha}^T \xi^s, \forall s \in \mathcal{S}. \end{aligned} \tag{3.3.9}$$

A clear bene t of the formulation above is that the decision variables and the number of constraints are independent of the sample and this will particularly facilitate the convergence.

To minimize the dependence on the speci c details of the objective and constraints functions of problem (3.3.9) for the convergence analysis and also for the purpose of potential applications of the convergence result, we consider the following general optimisation problem

$$\begin{aligned} \min_{x \in X} \quad & \psi(x) \\ \text{s.t.} \quad & f(x, \xi) \leq 0, \forall \xi \in \quad , \end{aligned} \tag{3.3.10}$$

where X is a compact set in a nite dimensional space, ψ and f are continuous functions which map from $\mathbb{R}^n$ and $\mathbb{R}^n \times \mathbb{R}^k$ to $\mathbb{R}$ respectively, ξ is parameter which takes values over a compact set .

Let $\xi^1, \cdots, \xi^N$ be independent and identically distributed random variables following continuous distribution with support set . We consider the discretized problem:

$$\begin{aligned} \min_{x \in X} \quad & \psi(x) \\ \text{s.t.} \quad & f(x, \xi^i) \leq 0, i = 1, \cdots, N. \end{aligned} \tag{3.3.11}$$

Let v and v_N denote respectively the optimal values of program (3.3.10) and program (3.3.11).

Lemma 3.3.1. *Assume that (a) ψ is Lipschitz continuous, (b) is a compact set, (c) ξ*

is continuous random variable with identical distribution[1] *for* ξ^i, $i = 1, \cdots, N$ *and there exist positive constants* K *and* τ *independent of* x *such that for each* $x \in X$*, there exist* $\alpha_0(x) < f^*(x) := \max_{y\in\Xi} f(x,y)$ *with*

$$1 - F_x(\alpha) \geq K\left(f^*(x) - \alpha\right)^\tau, \ \textit{for all } \alpha \in (\alpha_0(x), f^*(x)), \tag{3.3.12}$$

where F_x *denotes the cumulative distribution function of* $f(x,\xi)$ [2]*, (d)* f *is Lipschitz continuous in* x *with integrable Lipschitz modulus (w.r.t. the distribution of* ξ*). Then for any positive number* ϵ*, there exists positive constants* $C(\epsilon)$ *and* $\beta(\epsilon)$ *such that*

$$\text{Prob}(|v_N - v| \geq \epsilon) \leq C(\epsilon)e^{-\beta(\epsilon)N} \tag{3.3.13}$$

for N *sufficiently large.*

Proof: The thrust of the proof is to use CVaR and its sample average approximation to approximate the semi-in nite constraint of (3.3.10) which is in line with the convergence analysis carried out in [2]. However, there are a few important distinctions: (i) the convergence here is for the randomization scheme (3.3.11) rather than the sample average approximation of the CVaR approximation of the semi-in nite constraints, (ii) the underlying functions in the objective and constraints are not necessarily convex and (iii) the decision vector may consist of some integer variables.

Let

$$(x) := \sup_{\xi\in\Xi} f(x,\xi) \qquad \text{and} \qquad {}_N(x) := \sup_{i=1,\cdots,N} f(x,\xi^i),$$

let $\mathcal{F}$ and $\mathcal{F}^N$ denote the feasible set of problem (3.3.10) and problem (3.3.11) respectively. Then

$$\mathcal{F} = \{x : \ (x) = 0\} \qquad \text{and} \qquad \mathcal{F}^N = \{x : \ {}_N(x) = 0\}.$$

[1] Although ξ is a deterministic parameter here, we may regard it as a random variable and by writing $f(x,\xi) \leq 0$ we mean that for every realization of ξ, the inequality holds.

[2] Note that ξ could be any random variable which follows a continuous distribution with support set Ξ and the cumulative distribution function satisfying (3.3.12).

Moreover, since $_N(x) \leq \ (x)$, $\mathcal{F} \subset \mathcal{F}^N$. For $\beta \in (0,1)$, let

$$\text{CVaR}_\beta(f(x,\xi)) := \sup_\eta \left\{\eta + \frac{1}{1-\beta}\int_{y\in Y}(f(x,y)-\eta)_+\rho(y)dy\right\}$$

and

$$^N_\beta(x) := \sup_\eta \left\{\eta + \frac{1}{(1-\beta)N}\sum_{j=1}^N (f(x,\xi^j)-\eta)_+\right\}$$

where $\rho(\cdot)$ denotes the density function of the random variable ξ, $(a)_+ = \max(0,a)$ for $a \in \mathbb{R}$. In the literature, $\text{CVaR}_\beta(f(x,\xi))$ is known as conditional value at risk and $^N_\beta(x)$ is its sample average approximation, see [92, 2]. It is well known that the maximum w.r.t. η in the above formulation is achieved at a nite η. In other words, we may restrict the maximum to be taken within a closed interval $[-a,a]$ for a su ciently large, see [92]. It is easy to verify that

$$^N_\beta(x) \leq \ _N(x) \leq \ (x). \tag{3.3.14}$$

We proceed the rest of the proof in four steps.

Step 1. By the de nition of CVaR,

$$\text{CVaR}_\beta(f(x,\xi)) \leq \ (x).$$

for any $\beta \in (0,1)$, see [2]. Moreover, under condition (c), it follows by [2, Theorem 2.1] that

$$|\text{CVaR}_\beta(f(x,\xi)) - \ (x)| \leq \frac{1}{K^{1/\tau}}\frac{\tau}{1+\tau}(1-\beta)^{1/\tau}. \tag{3.3.15}$$

Therefore by driving β to 1, we obtain

$$\sup_{x\in X}|\text{CVaR}_\beta(f(x,\xi)) - \ (x)| \to 0.$$

Step 2. Using the inequalities (3.3.14), we have

$$\begin{aligned} |\ _N(x) - \ (x)| &\leq |\ ^N_\beta(x) - \ (x)| \\ &\leq |\ ^N_\beta(x) - \text{CVaR}_\beta(f(x,\xi))| + |\text{CVaR}_\beta(f(x,\xi)) - \ (x)| \end{aligned} \tag{3.3.16}$$

Let ϵ be a small positive number. By (3.3.15), we may set β su ciently close to 1 such that

$$\sup_{x \in X} |\mathrm{CVaR}_\beta\,(f(x,\xi)) - \quad (x)| \le \frac{\epsilon}{2}.$$

On the other hand, since   is compact and f is Lipschitz continuous in x with integrable modulus, by virtue of [102, Theorem 5.1], there exist positive constants $C(\epsilon)$ and $\alpha(\epsilon)$ such that

$$\begin{aligned}
&\mathrm{Prob}\Big(\sup_{x\in X} | \quad {}_\beta^N(x) - \mathrm{CVaR}_\beta\,(f(x,\xi))\,| \ge \epsilon/2\Big) \\
&\le \mathrm{Prob}\Big(\frac{1}{1-\beta}\sup_{x\in X}\sup_{\eta\in[-a,a]}\left|\frac{1}{N}\sum_{j=1}^{N}(f(x,\xi^j)-\eta)_+ - \mathbb{E}_P[(\eta - f(x,\xi))_+]\right| \ge \epsilon/2\Big) \\
&\le C(\epsilon)e^{-\alpha(\epsilon)N} \qquad (3.3.17)
\end{aligned}$$

for N su ciently large. Here in the rst inequality, we are using the fact that the maximum w.r.t. η is achieved in $[-a,a]$ for some appropriate positive constant a. See similar discussions in [119]. Therefore

$$\mathrm{Prob}\Big(| \quad {}_N(x) - \quad (x)| \ge \epsilon\Big) \le \mathrm{Prob}\Big(|\mathrm{CVaR}_\beta^N\,(f(x,\xi)) - \mathrm{CVaR}_\beta\,(f(x,\xi))\,| \ge \epsilon/2\Big) \le C(\epsilon)e^{-\alpha(\epsilon)N}.$$

Step 3. For small positive number δ, let

$$R(\delta) = \min_x \Big\{ \quad (x) : d(x,\mathcal{F})\Big\} \ge \delta,$$

where $d(x,\mathcal{F})$ denotes the distance from point x to set $\mathcal{F}$. Obviously $R(\delta) > 0$, it is monotonically increasing and $R(\delta) \to 0$ as $\delta \downarrow 0$. It is easy to observe that

$$d(x,\mathcal{F}) \ge \epsilon \Longleftrightarrow \quad (x) \ge R(\epsilon).$$

Let

$$\mathbb{D}(\mathcal{F}^N,\mathcal{F}) := \sup_{x\in\mathcal{F}^N} d(x,\mathcal{F})$$

and

$$\mathbb{H}(\mathcal{F}^N,\mathcal{F}) := \max\big(\mathbb{D}(\mathcal{F}^N,\mathcal{F}),\mathbb{D}(\mathcal{F},\mathcal{F}^N)\big),$$

which is the Hausdor distance between $\mathcal{F}^N$ and $\mathcal{F}$. Since both $\mathcal{F}^N$ and $\mathcal{F}$ are bounded, the Hausdor distance is well de ned. Moreover, since $\mathcal{F} \subset \mathcal{F}^N$, $\mathbb{H}(\mathcal{F}^N,\mathcal{F}) = \mathbb{D}(\mathcal{F}^N,\mathcal{F})$.

In what follows, we estimate Prob$\Big(\mathbb{D}(\mathcal{F}^N, \mathcal{F}) \geq \epsilon\Big)$. For any $x_N \in F^N$, since $\ _N(X_N)$ = 0, then

$$(x_N) \geq R(\delta) \Longleftrightarrow \quad (x_N) - \quad _N(x_N) \geq R(\delta).$$

Therefore

$$\begin{aligned} \text{Prob}\Big(\mathbb{D}(\mathcal{F}^N, \mathcal{F}) \geq \delta\Big) \quad &\leq \quad \text{Prob}\left(\sup_{x \in \mathcal{F}^N} | \quad (x)| \geq R(\delta)\right) \\ &\leq \text{Prob}\left(\sup_{x \in X} | \quad (x) - \quad _N(x)| \geq R(\delta)\right). \end{aligned} \tag{3.3.18}$$

Step 4. Let $x^* \in \mathcal{F}$ and $x^N \in \mathcal{F}^N$ be the optimal solutions to (3.3.10) and (3.3.11). Then by the Lipschitz continuity of ψ,

$$|v^N - v^*| = |\psi(x^N) - \psi(x^*)| \leq L\|x^N - x^*\| \leq L\mathbb{H}(\mathcal{F}^N, \mathcal{F})$$

where L denotes the Lipschitz modulus of ψ. By selecting δ such that $R(\delta/L) < \epsilon$, we arrive at

$$\text{Prob}(|v^N - v^*| \geq \epsilon) \leq \text{Prob}\Big(\mathbb{H}(\mathcal{F}^N, \mathcal{F}) \geq \delta/L\Big) \leq \text{Prob}\left(\sup_{x \in X} | \quad (x) - \quad _N(x)| \geq \epsilon\right).$$

The rest follows from Step 2. The proof is complete. □

With Lemma 3.3.1, we are ready to state the convergence of problem (3.3.9).

Theorem 3.3.1. *Let ϑ and ϑ_N denote the optimal value of (SIP-UC) and (3.3.9) respectively. Assume that ξ follows a uniform distribution. Then for any positive number ϵ, there exists positive constants $C(\epsilon)$ and $\beta(\epsilon)$ such that*

$$\text{Prob}(|\vartheta_N - \vartheta| \geq \epsilon) \leq C(\epsilon)e^{-\beta(\epsilon)N} \tag{3.3.19}$$

for N sufficiently large.

Proof: It su ces to verify the conditions of Lemma 3.3.1. Conditions (a) and (b) are obvious since the objective function is linear and is compact problem (3.3.9). Condition (c) is satis ed because $g(x, \xi)$ also follows a uniform distribution for each xed x and the cumulative distribution function of $g(x, \xi)$ is a linear function. Let us verify condition (d). It follows by [115, 81] (see also [75, Lemma 4.3]) that $g(x, \xi)$ is Lipschitz continuous w.r.t. x and ξ and since is compact, $g(x, \xi)$ is Lipschitz continuous in x with integrable Lipschitz modulus. The proof is complete. $\square$

In Theorem 3.3.1, we assume that ξ follows a uniform distribution. It might be interesting to show the conclusion when ξ follows a general continuous distribution with positive density function in the interior of , that is, $g(x, \xi)$ satisfies (3.3.12), we leave this for our future work.

The detailed formulation for the sample approximation of (SIP-UC) problem is given by mixed integer linear program below

$$
\begin{aligned}
\min_{\mathbf{u},\mathbf{q},\mathbf{r},\hat{\mathbf{r}}(.),\boldsymbol{\alpha}} \quad & c_i^f u_{it} + c_i^l q_{it} + c_i^{r,up} r_{it}^{up} + c_i^{r,dw} r_{it}^{dw} + \alpha_0 + \sum_{t=1}^{T} \alpha_t \mu_t \\
\text{s.t.} \quad & q_{it} + r_{it}^{up} \le \overline{q}_i u_{it}, && \forall t, i, \\
& q_{it} - r_{it}^{dw} \ge \underline{q}_i u_{it}, && \forall t, i, \\
& \underline{r}_i^{up} u_{it} \le r_{it}^{up} \le \overline{r}_i^{up} u_{it}, && \forall t, i, \\
& \underline{r}_i^{dw} u_{it} \le r_{it}^{dw} \le \overline{r}_i^{dw} u_{it}, && \forall t, i, \\
& q_{it}, r_{it}^{dw}, r_{it}^{up} \in \mathbb{R}^+, \ u_{it} \in \{0, 1\}, && \forall t, i, \\
& \sum_i \sum_t \{\hat{c}_i^{up} \hat{r}_{it}^{up}(\xi^s) + \hat{c}_i^{dw} \hat{r}_{it}^{dw}(\xi^s)\} \\
& \quad + \sum_t \{c_t^{ls} S_t(\xi^s) + c_t^{ws} W_t(\xi^s)\} \le \alpha_0 + \sum_{t=1}^{T} \alpha_t \xi_t^s, && \forall s, \\
& \sum_i \hat{q}_{it}(\xi^s) + S_t(\xi^s) - W_t(\xi^s) = l_t(\xi^s), && \forall t, s, \\
& 0 \le \hat{r}_{it}^{up}(\xi^s) \le r_{it}^{up}, && \forall t, i, s, \\
& 0 \le \hat{r}_{it}^{dw}(\xi^s) \le r_{it}^{dw}, && \forall t, i, s, \\
& Q_t - (q_{it} + r_{it}^{up}) + S_t(\underline{\xi}) \ge l_t(\underline{\xi}), && \forall t, i, \\
& \hat{q}_{it}(\xi^s) - \hat{q}_{i(t-1)}(\xi^s) \le RU_i u_{i(t-1)} + SU_i(1 - u_{i(t-1)}), && \forall t, i, s, \\
& \hat{q}_{i(t-1)}(\xi^s) - \hat{q}_{it}(\xi^s) \le RD_i u_{it} + SD_i(1 - u_{it}), && \forall t, i, s,
\end{aligned}
$$

where the first 5 constraints are the first-stage constraints, and the last 6 constraints represent the remaining of constraints in problem (3.3.8).

3.4 A mixture distribution approach

In the absence of complete information on the underlying distribution of random variables, the decision maker could integrate information obtained through various channels to construct a mixture probability distribution. The use of mixture distribution in the context of robust optimisation could be traced back to [85] and more recently in [124] for portfolio optimisation problems.

To de ne the ambiguity set corresponding to the robust problem (3.2.18), let $P_j, j = 1, \cdots, L$ be a set of probability measures such that $\mathbb{E}_{P_j}[g(x, \boldsymbol{\xi})]$ is well de ned for $j = 1, \cdots, L$. The ambiguity set under mixture distribution can then be de ned as follows

$$\mathcal{P} := \Big\{ \sum_{j=1}^{L} \gamma_j P_j : \quad \sum_{j=1}^{L} \gamma_j = 1, \quad \gamma_j \geq 0,\ \forall j = 1, \cdots, L \Big\},$$

where γ_j denotes the weight of distribution j. The probability distributions $P_1, \cdots, P_L$ are assumed to be known and true probability distribution is assumed to be in their convex hull. For any realization of the distribution $P := \sum_j \gamma_j P_j$ we have

$$\mathbb{E}_P[g(x, \boldsymbol{\xi})] = \sum_{j=1}^{L} \gamma_j \mathbb{E}_{P_j}[g(x, \boldsymbol{\xi})].$$

Under mixture distribution, the inner maximization problem in (3.2.18) can then be rewritten as follows

$$\begin{aligned} H(x) = \ & \sup_{\gamma} \ \sum_{j=1}^{L} \gamma_j \mathbb{E}_{P_j}[g(x, \boldsymbol{\xi})] \\ & \text{s.t.} \ \sum_j \gamma_j = 1, \\ & \qquad \gamma_j \geq 0, \qquad \forall j = 1, \cdots, L. \end{aligned} \tag{3.4.1}$$

Let λ be the dual variable corresponding to the first constraint in (3.4.1), the dual of the above problem is

$$\begin{aligned} H_D(x) = \quad & \min_{\lambda} \quad \lambda \\ & \text{s.t.} \quad \mathbb{E}_{P_j}[g(x,\boldsymbol{\xi})] \leq \lambda, \quad \forall j = 1, \cdots, L. \end{aligned} \tag{3.4.2}$$

Proposition 3.4.1. *In the case when $\boldsymbol{\xi}$ has a finite discrete distribution with the support set containing a finite number of values $\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^N$ and for a given first-stage decision x, model (3.4.2) is equivalent to*

$$\begin{aligned} \min_{y^1,\cdots,y^N,\lambda} \quad & \lambda \\ s.t. \quad & \sum_{k=1}^{N} p_j^k (h^T y^k) \leq \lambda, \forall j = 1, \cdots, L, \\ & y^k \in \mathcal{Y}(x, \boldsymbol{\xi}^k), \forall k = 1, \cdots, N, \end{aligned} \tag{3.4.3}$$

where p_j^k is the probability measure of P_j in scenario k.

Proof: If $\boldsymbol{\xi}$ has a discrete distribution with a finite number of values $\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^N$, then model (3.4.2) can be written as

$$\begin{aligned} \min_{\lambda} \quad & \lambda \\ \text{s.t.} \quad & \sum_{k=1}^{N} p_j^k g(x, \boldsymbol{\xi}^k) \leq \lambda, \quad \forall j = 1, \cdots, L, \end{aligned}$$

where $g(x, \boldsymbol{\xi}^k)$ refers to the second-stage problem

$$\begin{aligned} g(x, \boldsymbol{\xi}^k) = \quad & \min_{y} \quad h^T y \\ & \text{s.t.} \quad y \in \mathcal{Y}(x, \boldsymbol{\xi}^k). \end{aligned}$$

Let us denote the optimal solution of the above problem as

$$\hat{y}^k = \arg\min_{y \in \mathcal{Y}(x,\boldsymbol{\xi}^k)} h^T y, \forall k,$$

then we have $g(x, \boldsymbol{\xi}^k) = h^T \hat{y}^k$. Let us de ne $\hat{\lambda} = \sum_{k=1}^{N} p_j^k (h^T \hat{y}^k)$. It is clear that $(\hat{\lambda}, \hat{y})$ is a feasible solution for the problem (3.4.3). Let $(\bar{\lambda}, \bar{y}^k : k = 1, \cdots, N)$ be an optimal solution for problem(3.4.3) then $h^T \bar{y}^k \geq h^T \hat{y}^k, \forall k$ holds by the de nition of $\hat{y}^k$. Thus, we have

$$\sum_{k=1}^{N} p_j^k (h^T \bar{y}^k) \geq \sum_{k=1}^{N} p_j^k (h^T \hat{y}^k), \quad \forall j.$$

Therefore

$$\min_j \left\{ \sum_{k=1}^{N} p_j^k (h^T \bar{y}^k) \right\} \geq \min_j \left\{ \sum_{k=1}^{N} p_j^k (h^T \hat{y}^k) \right\},$$

and hence $\bar{\lambda} \geq \hat{\lambda}$. Since $(\hat{\lambda}, \hat{y})$ is a feasible solution of (3.4.3), it must also be an optimal solution. □

Thus, when $\boldsymbol{\xi}$ has a nite discrete distribution, the original robust problem (3.2.18) can be rewritten as

$$
\begin{aligned}
\text{(Mix-UC)} \qquad & \min_{x, y(.), \lambda} && c^T x + \lambda \\
& \text{s.t.} && x \in \mathcal{X}, \\
& && \sum_{k=1}^{N} p_j^k (h^T y^k) \leq \lambda, \forall j = 1, \cdots, L, \\
& && y^k \in \mathcal{Y}(x, \boldsymbol{\xi}^k), \forall k = 1, \cdots, N.
\end{aligned}
\tag{3.4.4}
$$

Note that when P_j follows a continuous distribution, it might be di cult to compute the expected value of the functions in the constraints in model (3.4.2) with respect to probability distribution. A well known approach to resolve this issue is to use sample average approximation (SAA). For a xed j, let $\boldsymbol{\xi}_j^1, \cdots, \boldsymbol{\xi}_j^{N_j}$ denote independent and identically random sampling of $\boldsymbol{\xi}$ from the probability distribution p_j then $\mathbb{E}_{p_j}[g(x, \boldsymbol{\xi})]$

can be approximated as

$$\frac{1}{N_j}\sum_{k=1}^{N_j} g(x, \boldsymbol{\xi}_j^k).$$

The SAA of the problem (3.4.4) can be written as

$$\begin{aligned}
\min_{x,y(.),\lambda} \quad & c^T x + \lambda \\
\text{s.t.} \quad & x \in \mathcal{X}, \\
& \frac{1}{N_j}\sum_{k=1}^{N_j} h^T y_j^k \le \lambda, \forall j = 1, \cdots, L, \\
& y_j^k \in \mathcal{Y}(x, \boldsymbol{\xi}_j^k), \forall k = 1, \cdots, N_j, j = 1, \cdots, L.
\end{aligned} \tag{3.4.5}$$

The detailed formulation for the Mix-UC problem is as follows

$$\begin{aligned}
\min_{\mathbf{u},\mathbf{q},\mathbf{r},\hat{\mathbf{r}}(.),\mathbf{W},\mathbf{S},\lambda} \quad & c_i^f u_{it} + c_i^l q_{it} + c_i^{r,up} r_{it}^{up} + c_i^{r,dw} r_{it}^{dw} + \lambda \\
\text{s.t.} \quad & q_{it} + r_{it}^{up} \le \overline{q}_i u_{it}, && \forall t, i, \\
& q_{it} - r_{it}^{dw} \ge \underline{q}_i u_{it}, && \forall t, i, \\
& \underline{r}_i^{up} u_{it} \le r_{it}^{up} \le \overline{r}_i^{up} u_{it}, && \forall t, i, \\
& \underline{r}_i^{dw} u_{it} \le r_{it}^{dw} \le \overline{r}_i^{dw} u_{it}, && \forall t, i, \\
& q_{it}, r_{it}^{dw}, r_{it}^{up} \in \mathbb{R}^+, \ u_{it} \in \{0, 1\}, && \forall t, i, \\
& \frac{1}{N_j}\sum_{k=1}^{N_j} \Big[\sum_i \sum_t \{\hat{c}_i^{up} \hat{r}_{it}^{up}(\xi_j^k) + \hat{c}_i^{dw} \hat{r}_{it}^{dw}(\xi_j^k)\} \\
& \qquad + \sum_t \{c_t^{ls} S_t(\xi_j^k) + c_t^{ws} W_t(\xi_j^k)\}\Big] \le \lambda && \forall j, \\
& \sum_i \hat{q}_{it}(\xi_j^k) + S_t(\xi_j^k) - W_t(\xi_j^k) = l_t(\xi_j^k) && \forall t, k, j, \\
& 0 \le \hat{r}_{it}^{up}(\xi_j^k) \le r_{it}^{up} && \forall t, i, k, j, \\
& 0 \le \hat{r}_{it}^{dw}(\xi_j^k) \le r_{it}^{dw} && \forall t, i, k, j, \\
& Q_t - (q_{it} + r_{it}^{up}) + S_t(\underline{\xi}) \ge l_t(\underline{\xi}), && \forall t, i, \\
& \hat{q}_{it}(\xi_j^k) - \hat{q}_{i(t-1)}(\xi_j^k) \le RU_i u_{i(t-1)} + SU_i(1 - u_{i(t-1)}), && \forall t, i, k, j, \\
& \hat{q}_{i(t-1)}(\xi_j^k) - \hat{q}_{it}(\xi_j^k) \le RD_i u_{it} + SD_i(1 - u_{it}), && \forall t, i, k, j.
\end{aligned}$$

Note that the SAA approach to Mix-UC results in a mixed integer linear program (MILP).

3.5 Case study

In this section, we carry out some numerical experiments to evaluate the proposed methods. To facilitate the exposition and reading, we develop a list of models and their corresponding abbreviations in Table3.3.

abbreviations	Problem	Method	Known information on uncertainty
Sto-UC	(3.2.7)-(3.2.17)	Stochastic UC	Probability distribution (P)
SIP-UC	(3.3.6)	Robust UC with semi-in nite formulation	First moments (μ)
Mix-UC	(3.4.4)	Robust UC with mixture distribution	A set of known distributions

Table 3.3: Abbreviations and reference of the methodologies

3.5.1 Data

We consider an illustrative case study based on a system with 10 generating unit. The data for the generators is based on [86] and includes the cost and limitation of the generation and reserve utilization as summarised in Table 3.4.

Unit	Costs [$/MW]					Capacities [MW]		
	Fixed [$] cost	Variable cost	Reserve up/down (scheduled)	Reserve up (actual)	Reserve down (actual)	Minimum output	Maximum output	Reserve up/down
i	c_i^f	c_i^l	$c_i^{r,up}$/ $c_i^{r,dw}$	$\hat{c}_i^{up}$	$\hat{c}_i^{dw}$	$\underline{q}_i$	$\overline{q}_i$	$\overline{r}_i^{up}/\overline{r}_i^{dw}$
1	2,550	16.19	1.80	17.81	-14.57	150	455	153
2	2,550	17.26	1.92	18.99	-15.53	150	455	153
3	1,300	16.60	1.84	18.26	-14.94	70	180	55
4	1,300	16.50	1.83	18.15	-14.85	70	180	55
5	1,620	19.70	2.19	21.67	-17.73	50	165	58
6	800	22.26	2.47	24.49	-20.03	30	90	30
7	850	27.74	3.08	30.51	-24.97	40	85	23
8	550	25.92	2.88	28.51	-23.33	20	60	20
9	550	27.27	3.03	30.00	-24.54	20	60	20
10	550	27.79	3.09	30.57	-25.01	20	60	20

Table 3.4: Ten unit system: generation and reserve cost and capacities

Stochastic base case

The hourly power demand is assumed to be known and under the two-stage stochastic setting (Sto-UC) the wind output is assumed to follow a multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \quad)$ with mean $\boldsymbol{\mu}$, standard deviation σ, correlation C and covariance $\quad = \sigma C \sigma^T$. The hourly demand, wind output mean values $\boldsymbol{\mu}$, and the standard deviations σ for 24 hours period are given in Table 3.5. The corresponding mean values for the stochastic net load is also presented in the Figure 3.3.

Hour	Demand	Wind		Hour	Demand	Wind	
t	d_t	Mean (μ_t)	SD(σ_t)	t	d_t	Mean (μ_t)	SD(σ_t)
1	1127	282	42.3	13	2254	564	126.8
2	1208	302	47.2	14	2093	523	121
3	1369	342	55.6	15	1932	483	114.7
4	1530	383	64.5	16	1691	423	103
5	1610	403	70.4	17	1610	403	100.6
6	1771	443	80.2	18	1771	443	113.5
7	1852	463	86.8	19	1932	483	126.8
8	1932	483	93.6	20	2254	564	151.4
9	2093	523	104.7	21	2093	523	143.9
10	2254	564	116.2	22	1771	443	124.5
11	2335	584	124	23	1449	362	104.1
12	2415	604	132.1	24	1288	322	94.6

Table 3.5: Hourly demand and mean/standard deviation of the wind

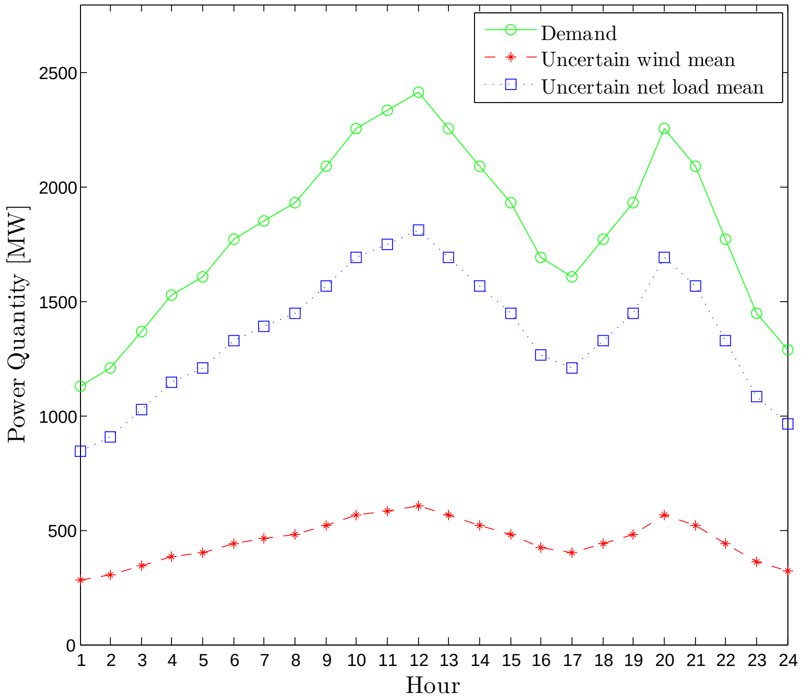


Figure 3.3: Nominal values of demand and mean wind

The correlation matrix is C is based on the hourly Danish wind output data[3] for the year 2013 and is given by

$$C = \left(\begin{array}{cccccccccccccccccccccccc}
1 & 0.994 & 0.979 & 0.962 & 0.942 & 0.913 & 0.88 & 0.844 & 0.812 & 0.778 & 0.744 & 0.705 & 0.665 & 0.629 & 0.597 & 0.566 & 0.54 & 0.51 & 0.479 & 0.462 & 0.449 & 0.428 & 0.399 & 0.372 \\
0.994 & 1 & 0.994 & 0.982 & 0.965 & 0.937 & 0.905 & 0.87 & 0.838 & 0.806 & 0.774 & 0.735 & 0.697 & 0.66 & 0.627 & 0.595 & 0.568 & 0.535 & 0.504 & 0.486 & 0.471 & 0.449 & 0.42 & 0.391 \\
0.979 & 0.994 & 1 & 0.995 & 0.982 & 0.957 & 0.927 & 0.893 & 0.862 & 0.829 & 0.797 & 0.759 & 0.719 & 0.682 & 0.648 & 0.616 & 0.587 & 0.553 & 0.522 & 0.505 & 0.489 & 0.467 & 0.439 & 0.411 \\
0.962 & 0.982 & 0.995 & 1 & 0.994 & 0.975 & 0.949 & 0.919 & 0.888 & 0.856 & 0.823 & 0.784 & 0.744 & 0.706 & 0.672 & 0.64 & 0.611 & 0.576 & 0.544 & 0.527 & 0.511 & 0.487 & 0.459 & 0.43 \\
0.942 & 0.965 & 0.982 & 0.994 & 1 & 0.992 & 0.974 & 0.949 & 0.922 & 0.892 & 0.86 & 0.821 & 0.781 & 0.741 & 0.705 & 0.673 & 0.643 & 0.608 & 0.576 & 0.558 & 0.54 & 0.515 & 0.486 & 0.454 \\
0.913 & 0.937 & 0.957 & 0.975 & 0.992 & 1 & 0.993 & 0.976 & 0.953 & 0.926 & 0.895 & 0.858 & 0.817 & 0.776 & 0.74 & 0.707 & 0.677 & 0.641 & 0.609 & 0.59 & 0.57 & 0.542 & 0.511 & 0.477 \\
0.88 & 0.905 & 0.927 & 0.949 & 0.974 & 0.993 & 1 & 0.993 & 0.976 & 0.953 & 0.924 & 0.888 & 0.847 & 0.805 & 0.768 & 0.735 & 0.705 & 0.669 & 0.636 & 0.616 & 0.594 & 0.563 & 0.53 & 0.494 \\
0.844 & 0.87 & 0.893 & 0.919 & 0.949 & 0.976 & 0.993 & 1 & 0.992 & 0.974 & 0.947 & 0.913 & 0.873 & 0.831 & 0.795 & 0.762 & 0.731 & 0.696 & 0.663 & 0.642 & 0.619 & 0.586 & 0.55 & 0.511 \\
0.812 & 0.838 & 0.862 & 0.888 & 0.922 & 0.953 & 0.976 & 0.992 & 1 & 0.992 & 0.969 & 0.939 & 0.902 & 0.862 & 0.827 & 0.796 & 0.766 & 0.731 & 0.699 & 0.678 & 0.652 & 0.616 & 0.577 & 0.536 \\
0.778 & 0.806 & 0.829 & 0.856 & 0.892 & 0.926 & 0.953 & 0.974 & 0.992 & 1 & 0.99 & 0.967 & 0.935 & 0.899 & 0.866 & 0.836 & 0.806 & 0.77 & 0.735 & 0.711 & 0.683 & 0.644 & 0.604 & 0.562 \\
0.744 & 0.774 & 0.797 & 0.823 & 0.86 & 0.895 & 0.924 & 0.947 & 0.969 & 0.99 & 1 & 0.991 & 0.969 & 0.94 & 0.91 & 0.882 & 0.854 & 0.819 & 0.78 & 0.754 & 0.723 & 0.682 & 0.641 & 0.598 \\
0.705 & 0.735 & 0.759 & 0.784 & 0.821 & 0.858 & 0.888 & 0.913 & 0.939 & 0.967 & 0.991 & 1 & 0.991 & 0.971 & 0.946 & 0.922 & 0.895 & 0.861 & 0.822 & 0.794 & 0.761 & 0.719 & 0.677 & 0.635 \\
0.665 & 0.697 & 0.719 & 0.744 & 0.781 & 0.817 & 0.847 & 0.873 & 0.902 & 0.935 & 0.969 & 0.991 & 1 & 0.993 & 0.976 & 0.957 & 0.933 & 0.901 & 0.863 & 0.835 & 0.801 & 0.759 & 0.719 & 0.679 \\
0.629 & 0.66 & 0.682 & 0.706 & 0.741 & 0.776 & 0.805 & 0.831 & 0.862 & 0.899 & 0.94 & 0.971 & 0.993 & 1 & 0.994 & 0.98 & 0.961 & 0.933 & 0.899 & 0.871 & 0.837 & 0.796 & 0.757 & 0.719 \\
0.597 & 0.627 & 0.648 & 0.672 & 0.705 & 0.74 & 0.768 & 0.795 & 0.827 & 0.866 & 0.91 & 0.946 & 0.976 & 0.994 & 1 & 0.995 & 0.981 & 0.958 & 0.928 & 0.901 & 0.867 & 0.827 & 0.789 & 0.753 \\
0.566 & 0.595 & 0.616 & 0.64 & 0.673 & 0.707 & 0.735 & 0.762 & 0.796 & 0.836 & 0.882 & 0.922 & 0.957 & 0.98 & 0.995 & 1 & 0.994 & 0.977 & 0.952 & 0.927 & 0.894 & 0.855 & 0.818 & 0.783 \\
0.54 & 0.568 & 0.587 & 0.611 & 0.643 & 0.677 & 0.705 & 0.731 & 0.766 & 0.806 & 0.854 & 0.895 & 0.933 & 0.961 & 0.981 & 0.994 & 1 & 0.993 & 0.974 & 0.951 & 0.92 & 0.882 & 0.847 & 0.814 \\
0.51 & 0.535 & 0.553 & 0.576 & 0.608 & 0.641 & 0.669 & 0.696 & 0.731 & 0.77 & 0.819 & 0.861 & 0.901 & 0.933 & 0.958 & 0.977 & 0.993 & 1 & 0.991 & 0.973 & 0.944 & 0.908 & 0.875 & 0.845 \\
0.479 & 0.504 & 0.522 & 0.544 & 0.576 & 0.609 & 0.636 & 0.663 & 0.699 & 0.735 & 0.78 & 0.822 & 0.863 & 0.899 & 0.928 & 0.952 & 0.974 & 0.991 & 1 & 0.992 & 0.969 & 0.937 & 0.905 & 0.876 \\
0.462 & 0.486 & 0.505 & 0.527 & 0.558 & 0.59 & 0.616 & 0.642 & 0.678 & 0.711 & 0.754 & 0.794 & 0.835 & 0.871 & 0.901 & 0.927 & 0.951 & 0.973 & 0.992 & 1 & 0.989 & 0.964 & 0.935 & 0.906 \\
0.449 & 0.471 & 0.489 & 0.511 & 0.54 & 0.57 & 0.594 & 0.619 & 0.652 & 0.683 & 0.723 & 0.761 & 0.801 & 0.837 & 0.867 & 0.894 & 0.92 & 0.944 & 0.969 & 0.989 & 1 & 0.99 & 0.969 & 0.944 \\
0.428 & 0.449 & 0.467 & 0.487 & 0.515 & 0.542 & 0.563 & 0.586 & 0.616 & 0.644 & 0.682 & 0.719 & 0.759 & 0.796 & 0.827 & 0.855 & 0.882 & 0.908 & 0.937 & 0.964 & 0.99 & 1 & 0.992 & 0.975 \\
0.399 & 0.42 & 0.439 & 0.459 & 0.486 & 0.511 & 0.53 & 0.55 & 0.577 & 0.604 & 0.641 & 0.677 & 0.719 & 0.757 & 0.789 & 0.818 & 0.847 & 0.875 & 0.905 & 0.935 & 0.969 & 0.992 & 1 & 0.992 \\
0.372 & 0.391 & 0.411 & 0.43 & 0.454 & 0.477 & 0.494 & 0.511 & 0.536 & 0.562 & 0.598 & 0.635 & 0.679 & 0.719 & 0.753 & 0.783 & 0.814 & 0.845 & 0.876 & 0.906 & 0.944 & 0.975 & 0.992 & 1
\end{array}\right)$$

[3] Available online at `http://energinet.dk`

SIP formulation

In the SIP formulation, we assume that only the first moment of the uncertain wind power is known and is given by μ. Based on this assumption, we implement the first proposed model and formulate the problem as a robust SIP (SIP-UC). In solving the SIP problem, we limit the support set of random hourly wind to 150 values generated from a multivariate uniform distribution.

Mixture distribution formulation

In the second approach of mixture distribution, we consider a case where the information on the uncertain net load is received through various resources, e.g. advise from a group of experts. While each alternative net load model provides a specific distribution and parameters such as the mean and the covariance, there is no consensus among the decision makers on which model contains the true distribution. Therefore, instead of relying on a particular expert model, we use the mixture model (Mix-UC) to combine all these potential distributions. Specifically, we assume that we are given three different distributions for the stochastic net load, shown in Table 3.6. We assume that three distributions are equally relevant and therefore they have equal weights in the construction of the mixture distribution. We construct the uncertainty set by drawing 50 random samples from each distribution.

Distribution	Mean	Covariance	Weight
Multivariate Normal	0.8μ		$\frac{1}{3}$
Multivariate Normal	1.2μ		$\frac{1}{3}$
Multivariate Uniform	μ		$\frac{1}{3}$

Table 3.6: Potential distributions for wind power

Computational results

The models have been implemented on an Intel Core Duo with processor at 3.33 GHz and 8 GB of RAM memory using CPLEX 12.5 under MATLAB R2012b. Table 3.7

shows the size of each problem as well as the running time for solving each problem.

Model	Scenarios (#)	Constraints (#)	Total variables (#)	Binary variables (#)	CPU time (s)
Sto-UC	150	222,000	80,160	240	10,601
Mix-UC	150	222,003	80,161	240	11,012
SIP-UC	150	222,150	80,185	240	6,021

Table 3.7: Problems size and computational times

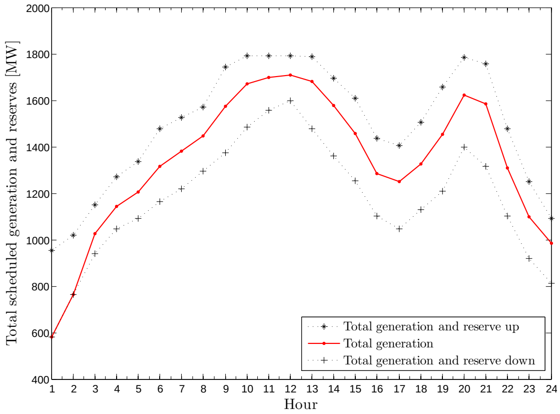
3.5.2 Numerical results

The proposed robust methods have been applied to the problem described above. The quality of robust solutions has been assessed through comparison with results achieved through the stochastic programming method. The cumulative rst-stage decisions as well as the second-stage and total costs for each solution is presented in Table 3.8, for detailed rst stage decisions please refer to Appendix A.2. It can be observed that both robust solutions have a higher total expected cost than the stochastic solution.

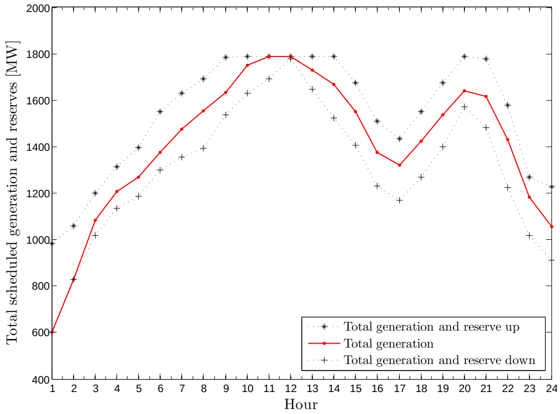
Model	First-stage decisions [Total MW]			Cost [$]		
	Generation $\sum_t \sum_i q_{it}$	Reserve up $\sum_t \sum_i r_{it}^{up}$	Reserve down $\sum_t \sum_i r_{it}^{dw}$	First-stage	Second-stage	Total expected
Sto-UC	32,148.7	3,723.4	3,906.5	823,339	48,339	871,678
Mix-UC	33,893.0	3,154.2	2,595.2	866,506	103,463	969,969
SIP-UC	32,758.3	5,306	7,328.4	885,219	191,057	1,076,276

Table 3.8: Stochastic versus robust solutions

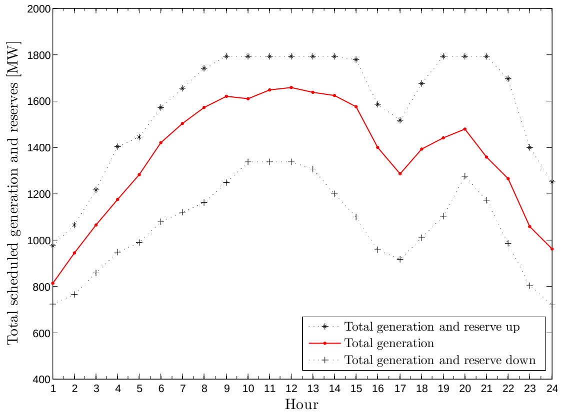
The higher rst-stage cost of the Mix-UC and SIP-UC solutions are as a result of the hedging strategy against the ambiguity of the distribution of the underlying net load uncertainty. In other words, robust models result in additional rst-stage generation or reserve scheduling. The hourly cumulative rst-stage generation and reserve levels are presented in Figure 3.4. It can be observed that, in the case of the Mix-UC solution, the hourly pattern of total generation is generally higher than the base Sto-UC solution. Additionally, the SIP-UC solution provides a greater exibility for the second stage through a higher level of up and down reserves. It can also be observed that the hourly patterns of the generation and the reserve schedule are similar to the net load quantities and at the peak hours the maximum generation capacity of generators are scheduled either in term of generation (Mix-UC) or generation with reserve (Sto-UC and SIP-UC).



(a) Sto-UC



(b) Mix-UC



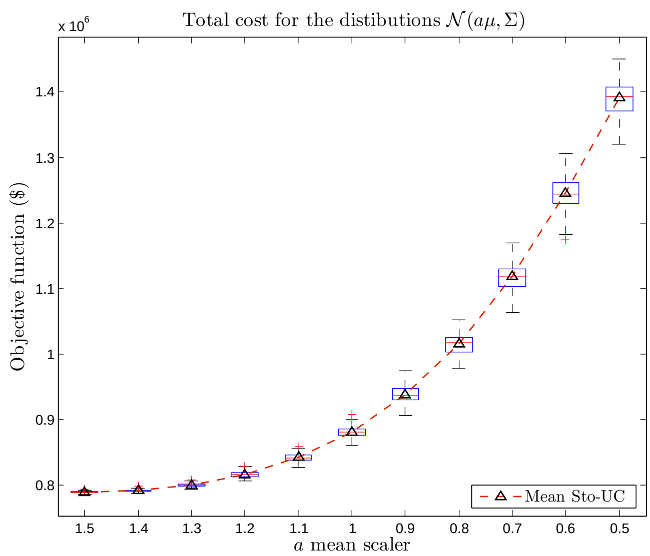
(c) SIP-UC

Figure 3.4: UC rst-stage solutions

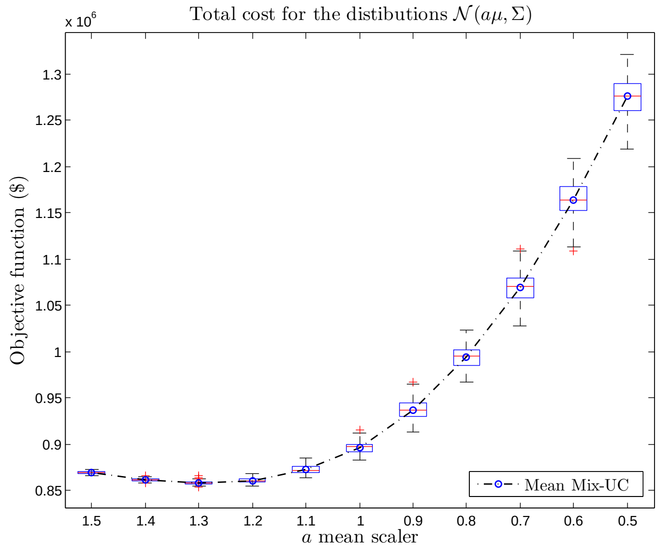
Sensitivity of solutions to variation in mean and covariance

The rst-stage decisions of the UC models determines the exibility of the solution to changes in the actual realization of the wind distribution in the second stage. i.e. we expect the robust solutions to provide greater exibility if the actual distribution of the uncertainty was di erent from the assumed distribution. To compare the performance of the Sto-UC model, MIX-UC model, and SIP-UC model, we analyse the e ect of deviation of the distribution parameters such as mean and covariance from the nominal value. In doing so, we consider a two-stage stochastic structure for each instance, in which the rst stage unit commitment and reserve schedules are xed as the rst-stage solutions of the Sto-UC model, MIX-UC model, and SIP-UC model. In each instance, we then solve the second-stage problem using a distribution with di erent means or covariances for the uncertain wind.

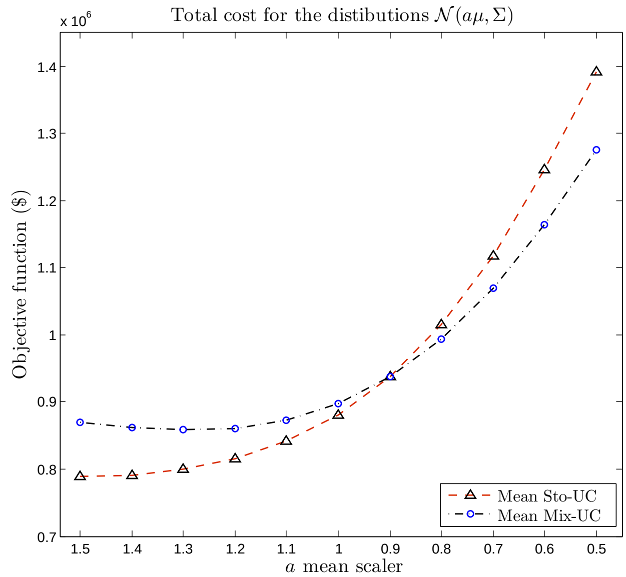
We rst study the sensitivity of solution to changes in the actual mean of the distribution, we compare the Sto-UC and Mix-UC for instances with the following wind distributions: $\mathcal{N}(a\boldsymbol{\mu}, \ ), a = \{0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5\}$. For each instance, the second-stage problem is solved by drawing 200 i.i.d samples from the corresponding distribution. Furthermore, we carry 100 independent runs for each instance. The solutions of 100 runs for each instance are summarised as a box plot and are presented in Figure 3.5a-3.5b. The mean values of objective function (total expected cost) for the 100 runs of each instance are also shown in Figure 3.5c.



(a) Sto-UC



(b) Mix-UC



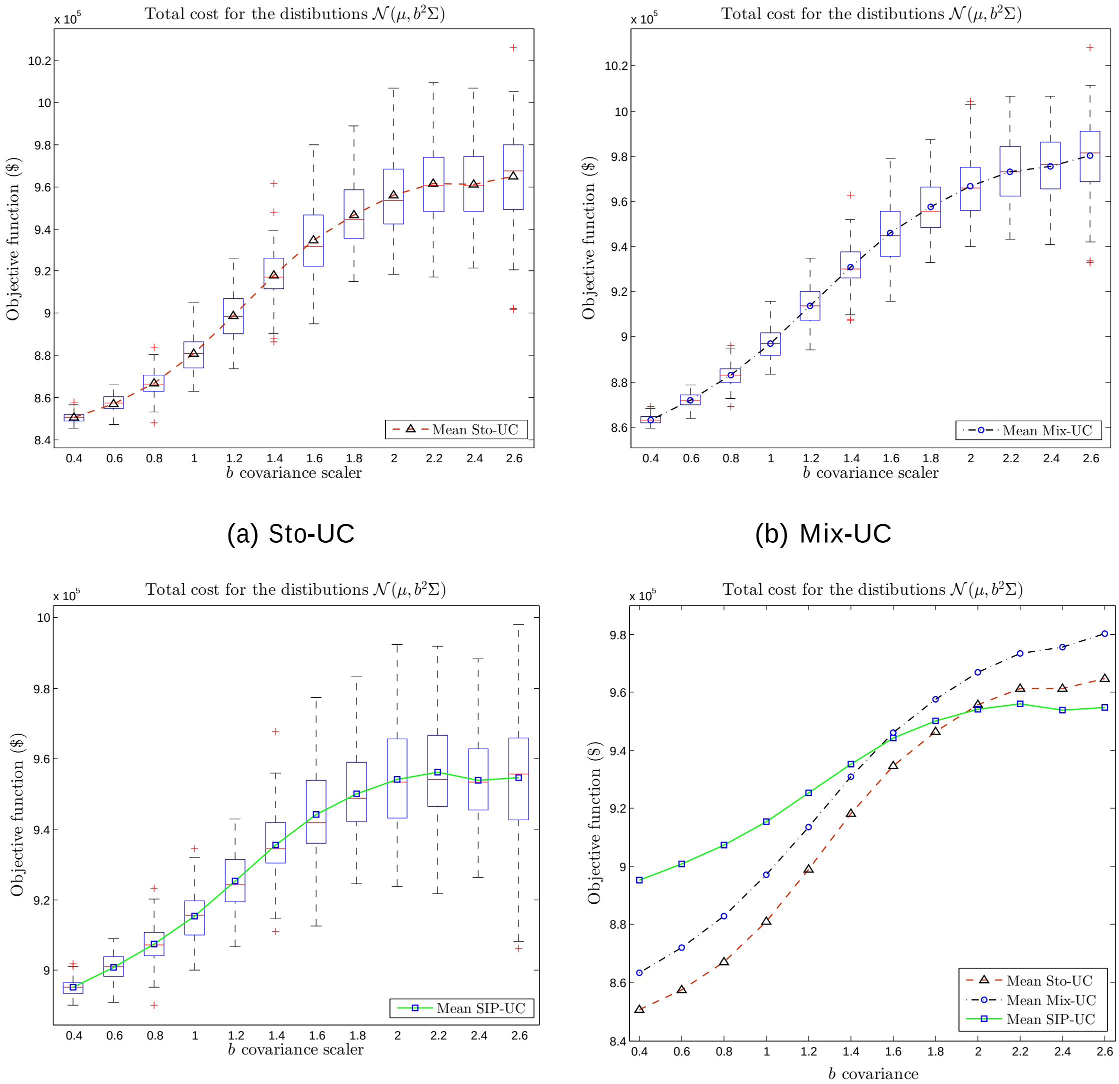
(c) Mean objective values for 100 runs

Figure 3.5: Sensitivity of solutions to variation in mean

It can be observed that when the actual wind distribution is $\mathcal{N} = (\mu, \;)$, i.e. when $a = 1$, the Sto-UC solution performs better than Mix-UC solution as expected since the assumed distribution in Sto-UC coincides with the actual distribution. The Sto-UC solution also performs better for all the instances with the mean of distribution greater than μ, i.e. when the mean of wind level is greater than expected in Sto-UC model.

This is due to the ability of the system operator to dispatch higher levels of energy using the excess wind power rather than utilizing the costly reserve. On the other hand, the Mix-UC solution performs better when the mean of the wind distribution is less than 90% of the anticipated value in Sto-UC model. In other word the Mix-UC robust solution has a lower total cost than Sto-UC solution when the wind output is less than what decision maker assumed under the Sto-UC model.

In the second set of sensitivity test instances, we compare the performance of Sto-UC, Mix-UC, SIP-UC solutions to the possible changes in the covariances of the wind distribution.



(a) Sto-UC

(b) Mix-UC

(c) SIP-UC

(d) Mean objective values for 100 runs

Figure 3.6: Sensitivity of solutions to variation in covariance

We specifically consider the following distributions for the second-stage uncertain wind: $\mathcal{N}(\boldsymbol{\mu}, (b)^2 \quad)$, where $b = \{0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2, 2.2, 2.4, 2.6\}$. Similar to previous tests, we draw 200 i.i.d samples from the corresponding distribution in each instance and repeat each instance for 100 runs. The corresponding box plots for test instances for all three models are presented in Figures 3.6a-3.6c. Furthermore, the mean value of the objective function for 100 test in each instance and each model is shown in Figure 3.6d. It can be observed that SIP-UC solution has the least sensitivity to change in covariance of the distribution, since the only available information for SIP-UC model was the first moment condition of the distribution and there was no assumption on the covariance of the distribution. However, this additional flexibility comes at a cost and it can be observed that the SIP-UC is more conservative and costly for instances with covariance coefficients $b < 2$, when compared to the Sto-UC solutions. For the instances with $b \geq 2$ the SIP-UC performs better than both Sto-UC and Mix-UC solutions.

The sensitivity of Mix-UC and Sto-UC solutions to changes in covariance are very similar and the difference between the two curve is almost unchanged across the test instances. This is due to the covariance assumptions in construction of the mixture model, i.e. covariance of all distribution used to construct the mixture model was equal to .

We have also constructed the mixture distribution using the following distributions: $\mathcal{N}(\mu, (0.8)^2 \quad), \mathcal{N}(\mu, \quad), \mathcal{N}(\mu, (1.2)^2 \quad)$. The solution of this mixture model was very similar to the Sto-UC model.

3.6 Conclusion

In this chapter, we presented a two-stage distributionally robust model that provides a novel and practical approach to deal with the uncertainty of the distribution of random wind output in the unit commitment problem. The model includes the technical ramping constraints as well as reliability condition against failure of up to one generating unit. The robustness takes into account the available information on uncertainty in two alternative ways. First, we assume that only the first moment information of the random wind is given and use duality theory to formulate the problem as a linear semi-infinite program. The SIP model is then solved using sampling. Second, we assume that the information on probability distribution of uncertain wind is received through various sources and construct a mixture model to include these into decision making. The mixture model is also reformulated using duality theory and solved through sample average approximation approach. Empirical tests have been carried out using an illustrative UC case study, taken from the literature, in order to illustrate the performance of the proposed robust models. The robust UC solutions may lose the potential of utilizing the wind power in high wind climate, however, they perform much better in a low wind climate as compared to the solutions that do not consider the uncertainty of the distribution (Sto-UC).

Chapter 4

Robust Facility Location Problem

Chapter Abstract

In this chapter, we consider a facility location problem where customer demand constitutes considerable uncertainty and complete information on the distribution of the uncertainty is unavailable. We formulate the optimal decision problem as a two-stage mixed integer programming problem: an optimal selection of facility locations in the first stage and an optimal decision on the operation of each facility in the second stage. A distributionally robust optimisation framework is proposed to hedge risks arising from incomplete information on the distribution of the uncertainty. Specifically, by exploiting the moment information, we construct a set of distributions which contains the true distribution and the optimal decision is based on the worst distribution from the set. We then develop two numerical schemes for solving the distributionally robust facility location problem: a semi-infinite programming approach which exploits moments of certain reference random variables, and a semi-definite programming approach which utilizes the mean and correlation of the underlying random variables describing the demand uncertainty. In the semi-infinite programming approach, we apply the well-known linear decision rule approach to the robust dual problem and then approximate the semi-infinite constraints through CVaR. We provide numerical tests to demonstrate the computation and properties of the robust solutions.

4.1 Introduction

The classic discrete facility location problem (FLP) involves selecting a subset of facility locations within a finite set of available locations and assigning customers to the selected facilities with the aim to minimize the combined facility setup cost and transportation cost. In the most basic form, the discrete FLPs consist of allocating p facilities to a given list of candidate locations. In these so-called p-median problems the fixed setup cost of all candidate sites are assumed to be equal. The objective function is only to minimize the total service cost to the customers, i.e. the transportation cost. Under the non-homogeneity the facilities' setup cost, the p-median problem can be extended to uncapacitated facility location problems (UFLP) in which the setup cost is also added to the objective function. The UFLP assumes that facilities can serve an unlimited amount of demand. However, in many practical problems, facilities have capacity constraints and this leads to an important family of FLPs called capacitated facility location problems (CFLP) in which the closest-assignment criteria is not sufficient for the optimality of the solution. The p-median, UFLP, and CFLP have been the subject of extensive research and interested readers might refer to [41, 89, 79, 11, 106] for some comprehensive reviews.

The aforementioned models share certain characteristics such as single-period planning horizon, single product and facility type, and deterministic parameters (i.e., demands, supplies, and costs). However, the deterministic assumption is one of the major drawbacks in coping with many real-world problems. The strategic decision on facility setup are often capital intensive, non-repetitive and spanning over a long time horizon. The decision has to be made at present and hence is subject to risks arising from uncertainties in demands and operations of the established facilities. Hedging the risk, therefore, becomes a vital component of the decision making process. The facility location problem under uncertainty has attracted considerable attention recently, see for examples, reviews in [82, 104]. Two major frameworks used to model uncertainty in the facility location problems are stochastic optimisation and robust optimisation.

In the first framework, stochastic optimisation has been a well-known mathematical method for finding optimal decision under uncertainty over the past few decades. A key assumption in this approach is that the decision maker has complete information on the distribution of the uncertainty, through either empirical data or subjective judgement.

However, in some circumstances, this might turn out to be di cult if not possible when a strategic decision has to be made well in advance of the realization of the uncertainty.

In the second approach, the robust optimisation framework, no assumption is made on the probability distribution of the uncertainty. The traditional proposed measure of robustness is the minmax cost approach in which the cost associated with the worst case scenario is minimized. Some of the examples of the minmax regret approach can be found in [6, 7, 38].

A feasible way to address the issue of distributional uncertainty in stochastic optimisation is to use the available data to construct a set of distributions which contains the true distribution of the uncertainty and make an optimal decision on the basis of the worst distribution from the set. This approach is known as distributionally robust optimisation which was proposed by Scarf [94] and has now been extensively studied over the past few decades. How to construct the set of distributions depends on the available information and there is no consensus on that. A popular way is to use moments of some random variables which may consists of the mean, variation and covariance. In this chapter, we propose a distributionally robust optimisation model for the capacitated facility location problem. We then propose two numerical schemes, namely a semi-de nite program and a semi-in nite program to solve the distributionally robust optimisation model.

The remainder of this chapter is organized as follows. In Section 4.2, we formally describe the deterministic model, a two-stage stochastic model and a disrtibutionally robust formulation of the stochastic model. We then proceed discussions on numerical schemes in the following two consecutive sections for solving the robust model depending on the availability of information on the distribution of demands: a semi-in nite programming (SIP) scheme in Section 4.3 and a semi-de nite programming (SDP) scheme in Section 4.4. Finally in Section 4.5, we report numerical test results of the two schemes.

4.2 Facility location models

We rst introduce the classic deterministic capacitated facility location (D-FLP) problem. By taking into account the future uncertainty we extend it to a two-stage stochastic facility location (S-FLP) modelling framework, which then forms the basis for developing a distributionally robust facility location (R-FLP) model.

4.2.1 Deterministic facility location model

There is a vast literature on the deterministic facility location problem. A good review of the related literature is carried out by Owen and Daskin [82] and Daskin [41]. Suppose there are up to n facilities to be opened in a set of possible locations $I = \{1, \cdots, n\}$, indexed by i. Let $J = \{1, \cdots, m\}$ j. The cation decision variable z_i is de ned as

$$z_i = \begin{cases} 1 & \text{if the facility } i \text{ is opened,} \\ 0 & \text{otherwise,} \end{cases}$$

and continuous assignment variable x_{ij} determines the service quantity,Td438 0r the trans-

j.

d420(if)-402(op)-28(ened,Td420(has)-403(a)-402(service)-403(capac-)]TJ -242.063 -20.324 Td [(it)28(y)]TJ/F23 10.9091 Tf 17.2

-329(is)-329(assumed)-328(that)-329(the)-329()]TJ7mer demand and the y

service cost of facility i j is

wn,Td348(e.g.)-480(transp)-28(ortation)-345(costs)-345(are)-345(prop)-28(ortional)-345(to)]TJ -72.371 -20.324 Td [(the)-393(

demand and supply capacity constraints. The deterministic FLP is formulated as follows

$$
\begin{aligned}
\text{(D-FLP)} \qquad \min_{\mathbf{z},\mathbf{x},\mathbf{w}} \quad & \sum_i b_i z_i + \sum_{i,j} c_{ij} x_{ij} + \sum_j C w_j \\
\text{s.t.} \quad & \sum_i x_{ij} + w_j \geq d_j, \ \forall\, j \in J, && (4.2.1)\\
& \sum_j x_{ij} \leq z_i s_i, \ \forall\, i \in I, && (4.2.2)\\
& x_{ij} \geq 0, \ \forall\, j \in J, \ i \in I, && (4.2.3)\\
& w_j \geq 0, \ \forall\, j \in J, && (4.2.4)\\
& z_i \in \{0, 1\}, \ \forall\, i \in I. && (4.2.5)
\end{aligned}
$$

Balance constraint (4.2.1) ensures that the demands of all customers are met. Constraint (4.2.2) prevents the service level assigned to each facility to exceed its capacity and also ensures that the customers cannot be served by un-built facilities (i.e. when $z_i = 0$, x_{ij} must be equal to zero). Finally, constraints (4.2.3) and (4.2.5) enforce the non-negativity of the service quantities and the binary nature of the facility allocating decisions respectively.

4.2.2 Two-stage stochastic model

The facility location problems involve uncertainties that stem from unpredictability of demand, supply and service costs. Since the facility location decisions are irreversible and capital intensive, it is vital to take into account the future uncertainties when the facility location decisions are made. Louveaux [77] rst introduced a two-stage stochastic program with recourse for solving simple plant location problems and p-median problems where uncertainties in demand, production and transportation costs are considered. To extend deterministic model described in subsection 4.2.1 to a stochastic setting, customer demand is assumed to be stochastic with a known probability distribution. Instead of having a deterministic demand vector $\mathbf{d}$, we use the notation $\mathbf{d}(\boldsymbol{\xi})$ for the stochastic demands that depend on a random vector $\boldsymbol{\xi}$. For convenience in notation, we use $\mathbf{d}(\boldsymbol{\xi})$ and $\boldsymbol{\xi}$ interchangeably, i.e. both $d_j(\boldsymbol{\xi})$ and ξ_j refer to the stochastic demand of customer j. The objective of the two-stage problem is to minimize the sum of xed investment cost of allocating the facilities and the expected future transportation costs. The resulting mathematical model is given as follows

$$\text{(S-FLP)} \qquad \begin{aligned} \min_{\mathbf{z}} \quad & \sum_i b_i z_i + \mathbb{E}\left[g(\mathbf{z}, \boldsymbol{\xi})\right] \\ \text{s.t.} \quad & z_i \in \{0, 1\}, \ \forall\, i \in I, \end{aligned} \tag{4.2.6}$$

where $g(\mathbf{z}, \boldsymbol{\xi})$ is the optimal value of the second stage transportation problem

$$\begin{aligned} g(\mathbf{z}, \boldsymbol{\xi}) = \min_{x,w} \quad & \sum_{i,j} c_{ij} x_{ij} + \sum_j C w_j, \\ \text{s.t.} \quad & \sum_i x_{ij} + w_j \geq d_j(\boldsymbol{\xi}), \ \forall\, j \in J, \\ & \sum_j x_{ij} \leq z_i s_i, \ \forall\, i \in I, \\ & x_{ij} \geq 0, \ \forall\, i \in I, \ j \in J, \\ & w_j \geq 0, \ \forall\, j \in J. \end{aligned} \tag{4.2.7}$$

The decision on $\mathbf{z}$ in the first stage determines the location of new facilities to be built, before the realization of the uncertain demand $\mathbf{d}(\boldsymbol{\xi})$, while the second stage specifies the allocation of transportation resources after the demand is realized. The optimal value $g(\mathbf{z}, \boldsymbol{\xi})$ of the second stage is therefore a function of the first stage variables $\mathbf{z}$ and a realization of uncertain demand. The expected transportation cost is based on the probability distribution of random vector $\boldsymbol{\xi}$, which is assumed to be bounded with a support set .

4.2.3 Distributionally robust facility location model

One of the major difficulties which often arise in facility location problems is the lack of complete information on the probability distribution of future customer demand. We consider a setting where there might be limited information on the probability distribution P of the random parameters $\boldsymbol{\xi}$. Suppose that we are able to construct a set of probability measures[1], denoted by $\mathcal{P}$, which contains the true probability distribution P. In order to hedge against ambiguity of the true distribution, we may consider a robust model where the optimal decision on locations of the new facilities to be built and the operation of all facilities is based on the worst distribution from the ambiguity set. The corresponding distributionally robust optimisation problem can be formulated as follows

$$\begin{aligned} \text{(R-FLP)} \qquad & \min_{\mathbf{z}} \quad \mathbf{b}^T\mathbf{z} + \sup_{P\in\mathcal{P}} \mathbb{E}_P\left[g(\mathbf{z},\boldsymbol{\xi})\right] \\ & \text{s.t.} \quad \mathbf{z} \in \{0,1\}^n. \end{aligned} \tag{4.2.8}$$

In practice, there are various ways to construct the uncertainty set. One approach is through the moments of the underlying probability distribution, which may consist of the mean, variation and covariance. Another way is to use a mixture distribution, which is a weighted distribution that can harmonize different potential distributions. In this chapter, based on the available moment information, we investigate two formulations for the proposed distributionally robust model in (4.2.8). In the first approach, we assume that only the first moment information is specified and we reformulate the problem as a semi-infinite program (SIP). In the second approach, we assume the first and second moments of the unknown distribution are given and on this basis we reformulate model (4.2.8) as a semi-definite program (SDP). We then develop appropriate numerical methods to solve these problems.

4.3 A semi-infinite approach

In this section we consider the case when the uncertainty set is defined through the first moment

$$\mathcal{P} = \{P \in \mathscr{P} : \mathbb{E}_P[\boldsymbol{\xi}] = \boldsymbol{\mu}\}, \tag{4.3.1}$$

[1]Throughout the chapter, we use interchangeably the terminology probability measure and probability distribution.

where $\boldsymbol{\mu}$ is the mean value of the random demand $\boldsymbol{\xi}$. Let us reconsider the inner maximization problem associated with the robust problem (4.2.8)

$$H(\mathbf{z}) = \sup_{P \in \mathcal{P}} \mathbb{E}_P\left[g(\mathbf{z}, \boldsymbol{\xi})\right].$$

We can derive the dual formulation with respect to the moment condition as

$$\begin{aligned} H_D(\mathbf{z}) = \quad & \min_{\lambda_0, \boldsymbol{\lambda}} \quad \lambda_0 + \boldsymbol{\lambda}^T \boldsymbol{\mu} \\ & \text{s.t.} \quad g(\mathbf{z}, \boldsymbol{\xi}) \leq \lambda_0 + \boldsymbol{\xi}^T \boldsymbol{\lambda}, \ \forall \boldsymbol{\xi} \in \ , \end{aligned} \tag{4.3.2}$$

where $\subset \mathbb{R}^m$ is the support set of $\boldsymbol{\xi}$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\lambda_0 \in \mathbb{R}$ are the dual variables associated with the moment constraints and the normalization constraint respectively. Since the support set is infinite, the dual problem (4.3.2) is a linear semi-infinite programming problem (SIP). It is important to note that (4.3.2) is a deterministic semi-infinite programming problem. If is structured, e.g., polyhedral or semi-algebraic and g is linear or quadratic w.r.t. $\boldsymbol{\xi}$, then through the well known S-lemma, the semi-infinite system can be represented as an SDP, see for instance [125]. Here, we don't assume any special structure as such. To avoid duality gap, we assume that the regularity conditions specified in [99] hold. Specifically, we assume the dual problem has a non-empty and bounded set of optimal solutions and also the support set is convex and compact. The second stage maximization problem in (4.2.8) can therefore be replaced by its dual as follows

$$\begin{aligned} \text{(R-SIP)} \qquad & \min_{\mathbf{z}, \lambda_0, \boldsymbol{\lambda}} \quad \mathbf{b}^T \mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T \boldsymbol{\mu} \\ & \text{s.t.} \quad \mathbf{z} \in \{0, 1\}^n, \\ & \qquad\ \ g(\mathbf{z}, \boldsymbol{\xi}) \leq \lambda_0 + \boldsymbol{\xi}^T \boldsymbol{\lambda}, \ \forall \boldsymbol{\xi} \in \ , \end{aligned} \tag{4.3.3}$$

or equivalently

$$
\begin{aligned}
\min_{\mathbf{z},\mathbf{x}(\cdot),\mathbf{w}(\cdot),\lambda_0,\boldsymbol{\lambda}} \quad & \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\
\text{s.t.} \quad & \mathbf{z} \in \{0,1\}^n, \\
& \mathbf{c} \bullet \mathbf{x}(\boldsymbol{\xi}) + Ce^T\mathbf{w}(\boldsymbol{\xi}) \leq \lambda_0 + \boldsymbol{\xi}^T\boldsymbol{\lambda}, \ \forall \boldsymbol{\xi} \in \quad , \\
& \Big(\mathbf{x}(\boldsymbol{\xi}), \mathbf{w}(\boldsymbol{\xi})\Big) \in \mathcal{G}(\mathbf{z}, \boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \quad ,
\end{aligned}
\tag{4.3.4}
$$

where $\mathbf{x}(\boldsymbol{\xi}) \in \mathbb{R}^{n\times m}$ and $\mathbf{w}(\boldsymbol{\xi}) \in \mathbb{R}^m$ are optimal transportation decisions for each fixed $\mathbf{z}$ and for each realization of $\boldsymbol{\xi}$, and where $\mathbf{c}$ is the matrix of transportation cost coefficients, C is the cost of serving customers from the external source, and $\mathbf{e} \in \mathbb{R}^m$ is a vector of all ones. Moreover, $\mathcal{G}(\mathbf{z}, \boldsymbol{\xi})$ is the feasible regions associated to second stage problem (4.2.7).

4.3.1 Linear decision rule approximation

One of the main challenges in solving the semi-infinite problem above is the dependence of the second stage transportation variables $\mathbf{x}(\boldsymbol{\xi})$ and $\mathbf{w}(\boldsymbol{\xi})$ on random variable $\boldsymbol{\xi}$. These \adjustable" variables are often referred to as *decision rules*, and their presence could often complicate the solution procedure. Formally, a decision rule $\mathbf{x}(\boldsymbol{\xi})$ can be defined as a vector valued function, mapping the random variables $\boldsymbol{\xi} \in \mathbb{R}^m$ with support set into the decisions. The decision rule problem can be interpreted as identifying the best decision $\mathbf{x}(\boldsymbol{\xi}) \in \quad \subset \mathcal{X}$ once $\boldsymbol{\xi}$ is observed, where $\mathcal{X}$ denotes the set of all the mapping from to $\mathbb{R}^{n\times m}$ and a subset of $\mathcal{X}$.

One of the tractable approximation schemes to deal with the decision rules is to restrict their feasible set to the ones that have a functionality affine relation with the uncertain random variables (that are affine functions of the uncertain data). This approach was proposed by Ben-Tal et al. [14] and was extended in [101] and [72] to develop tractable numerical procedure for stochastic programming problems. Here, we take the initiative to apply the linear decision rule (LDR) approximation to problem (4.3.4), that is, we impose the dependence of transportation decisions on the random demand to follow linear functions:

$$\begin{aligned} \mathbf{x}(\boldsymbol{\xi}) &= \mathbf{X}\boldsymbol{\xi} + \mathbf{x^0}, \\ \mathbf{w}(\boldsymbol{\xi}) &= \mathbf{W}\boldsymbol{\xi} + \mathbf{w^0}, \end{aligned}$$

where $\mathbf{X} \in \mathbb{R}^{(nm\times m)}, \mathbf{W} \in \mathbb{R}^{m\times m} \in \mathbb{R}^{n\times m}, \mathbf{x^0} \in \mathbb{R}^{n\times m}$, and $\mathbf{w^0} \in \mathbb{R}^{m}$. Consequently, problem (4.3.4) can be approximated as

$$\begin{aligned} \text{(R-LDR)} \qquad & \min_{\mathbf{z},\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda}} && \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\ & \text{s.t.} && \mathbf{z} \in \{0,1\}^n, \\ & && \mathbf{c} \bullet (\mathbf{X}\boldsymbol{\xi}) + \mathbf{x^0} + C\mathbf{e}^T(\mathbf{W}\boldsymbol{\xi}) + \mathbf{w^0} \le \lambda_0 + \boldsymbol{\xi}^T\boldsymbol{\lambda}, \ \forall \boldsymbol{\xi} \in \quad , \\ & && \left(\mathbf{X}\boldsymbol{\xi} + \mathbf{x^0}, \mathbf{W}\boldsymbol{\xi} + \mathbf{w^0}\right) \in \mathcal{G}(\mathbf{z},\boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \quad . \end{aligned} \tag{4.3.5}$$

Note that here we are slightly abusing the notation in this formulation: $\boldsymbol{\xi}$ should be understood as a parameter rather than a random variable. Indeed, it represents a realization of the random vector $\boldsymbol{\xi}$. The optimal value of the LDR approximation problem will generate an upper bound on the optimal value of original robust problem (4.3.4).

4.3.2 Conditional value at risk approximation

Having de ned the LDR formulation of the original robust semi-in nite problem, we approximate the rst semi-in nite constraint with Conditional Value at Risk (CVaR) and then approximate the latter through Monte Carlo sampling to reduce the number of constraints. One of the main advantages of using CVaR is that it converts the semi-in nite number of constraints into a single constraint. A recent study [3] has shown promising performance of CVaR approximation in dealing with semi-in nite problems. The CVaR method has been extensively used in stochastic programming for approximating the chance constraints and we may refer the readers to [110, 65] for more details.

In the case of our 0 G 0 0 T414roximating and $\mathcal{Q} := (\mathbf{X}, \mathbf{x^0}, \mathbf{W}, \mathbf{w^0}, \lambda_0, \boldsymbol{\lambda})$. The semi-in nite constraint of (4.3.5) can be written as

$$\sup_{\boldsymbol{\xi}\in\Xi} h(\mathcal{Q}, \boldsymbol{\xi}) \leq 0,$$

Let $\boldsymbol{\xi}$ be a random vector with support . Then $\sup_{\boldsymbol{\xi}\in\Xi} h(\mathcal{Q}, \boldsymbol{\xi})$ can be approximated by CVaR of $h(\mathcal{Q}, \boldsymbol{\xi})$, which is de ned as

$$\text{CVaR}_\beta(h(\mathcal{Q}, \boldsymbol{\xi})) = \min_{\eta\in\mathbb{R}} \quad _\beta(\mathcal{Q}, \eta),$$

where

$$_\beta(\mathcal{Q}, \eta) = \eta + \frac{1}{1-\beta} \int_{\tilde{\boldsymbol{\xi}}\in\Xi} \big(h(\mathcal{Q}, \boldsymbol{\xi}) - \eta\big)_+ \, P(d\boldsymbol{\xi}),$$

$(\tau)_+ = \max(0, \tau)$, and P denotes the distribution of $\boldsymbol{\xi}$.

Note that the set consisting of the values η is a nonempty, closed and bounded interval. Consequently, the CVaR approximation of the semi-in nite constraint in problem (4.3.5) can be expressed as follows:

$$\min_{\eta\in\mathbb{R}} \left(\eta + \frac{1}{1-\beta} \mathbb{E}\left[\big(h(\mathcal{Q}, \boldsymbol{\xi}) - \eta\big)_+ \right] \right) \leq 0. \tag{4.3.6}$$

It is important to distinguish the expectation $\mathbb{E}[\cdot]$ here from the expectation $\mathbb{E}[\cdot]$ in (S-FLP). The former should be understood as a mathematical expectation taken w.r.t. any distribution of any random variable $\boldsymbol{\xi}$ with support set . In other words here the $\boldsymbol{\xi}$ does not have to be identical to the $\boldsymbol{\xi}$ in (S-FLP). For example, we may set $\boldsymbol{\xi}$ as a random variable with uniform distribution over . Of course, the selection of $\boldsymbol{\xi}$ and its distribution will a ect the quantity of CVaR of h and the rate of approximation to its essential supremum.

Under some mild conditions, we can show that the error arising from the approximation scheme does not have signi cant impact on the optimal value, see [3]. By replacing the constraint in original LDR problem, we can write the CVaR approximation problem as

$$
\begin{aligned}
\text{(R-CVaR)} \qquad & \min_{\eta,\mathbf{z},\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda}} && \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\
& \text{s.t.} && \mathbf{z} \in \{0,1\}^n, \\
& && \eta + \frac{1}{1-\beta}\mathbb{E}\Big[\Big(h(\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda},\boldsymbol{\xi}) - \eta\Big)_+\Big] \le 0, \\
& && \Big(\mathbf{X}\boldsymbol{\xi} + \mathbf{x^0}, \mathbf{W}\boldsymbol{\xi} + \mathbf{w^0}\Big) \in \mathcal{G}(\mathbf{z},\boldsymbol{\xi}),\ \forall \boldsymbol{\xi} \in \ .
\end{aligned}
\tag{4.3.7}
$$

Under the assumption that LDR problem (4.3.5) satisfies the Slater constraint qualification, the optimal solution of CVaR approximation problem (4.3.7) converges to optimal solution of LDR problem as $\beta \to 1$.

Discretization through sampling

One of the well-known solution approaches for semi-infinite programs is random discretization. The basic idea is to construct a tractable sub-problem by considering a randomly drawn finite subset of constraints and hence enlarging the solution set. Calafiore and Campi [33, 34] investigated this approach and used Monte Carlo sampling (often referred to as sample average approximation) to approximate the convex problems consisting of linear objectives and semi-infinite constraints. They showed that the resulting randomized solution fails to satisfy only a small proportion of the original constraints for a sufficiently large sample size. An explicit bound on the measures of the original constraints that may be violated by the randomized solution is derived. The approach has been shown to be numerically efficient and it has been widely applied to various stochastic and robust programs, we refer interested readers to [35, 97] and references therein.

In this chapter, we apply the Monte Carlo sampling approach respectively to the original semi-infinite problem (4.3.4), its LDR approximation (4.3.5) and the CVaR formulation (4.3.7).

To construct the sample space, let $\mathcal{K} = \{1, \cdots, K\}$ denote the finite set of sample indices and $\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^K$ to be an independent and identically distributed (i.i.d) sampling of $\boldsymbol{\xi}$.

We may construct the discretized approximation of problem (4.3.4) as follows

$$
\begin{aligned}
\min_{\mathbf{z},x(\xi^k),w(\xi^k):k\in\mathcal{K},\lambda_0,\boldsymbol{\lambda}} \quad & \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\
\text{s.t.} \quad & \mathbf{z} \in \{0,1\}^n, \\
& \mathbf{c} \bullet \mathbf{x}(\boldsymbol{\xi}^k) + C\mathbf{e}^T\mathbf{w}(\boldsymbol{\xi}^k) \le \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\xi}^k,\ \forall k \in \mathcal{K}, \\
& \Big(\mathbf{x}(\boldsymbol{\xi}^k), \mathbf{w}(\boldsymbol{\xi}^k)\Big) \in \mathcal{G}(\mathbf{z}, \boldsymbol{\xi}^k),\ \forall k \in \mathcal{K}.
\end{aligned}
\tag{4.3.8}
$$

Similarly the discretized LDR problem (4.3.5) can be formulated as

$$
\begin{aligned}
\min_{\mathbf{z},\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda}} \quad & \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\
\text{s.t.} \quad & \mathbf{z} \in \{0,1\}^n, \\
& \mathbf{c} \bullet (\mathbf{X}\boldsymbol{\xi}^k) + \mathbf{x^0} + C\mathbf{e}^T(\mathbf{W}\boldsymbol{\xi}^k) + \mathbf{w^0} \le \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\xi}^k,\ \forall k \in \mathcal{K}, \\
& \Big(\mathbf{X}\boldsymbol{\xi}^k + \mathbf{x^0}, \mathbf{W}\boldsymbol{\xi}^k + \mathbf{w^0}\Big) \in \mathcal{G}(\mathbf{z}, \boldsymbol{\xi}^k),\ \forall k \in \mathcal{K},
\end{aligned}
\tag{4.3.9}
$$

and nally, we apply the sample average approximation (SAA) scheme to CVaR approximation problem (4.3.7) as follows

$$
\begin{aligned}
\min_{\eta,\mathbf{z},\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda}} \quad & \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\
\text{s.t.} \quad & \mathbf{z} \in \{0,1\}^n, \\
& \eta + \frac{1}{(1-\beta)K}\sum_{k=1}^{K}\left[\Big(h(\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda},\boldsymbol{\xi}^k) - \eta\Big)_+\right] \le 0, \\
& \Big(\mathbf{X}\boldsymbol{\xi}^k + \mathbf{x^0}, \mathbf{W}\boldsymbol{\xi}^k + \mathbf{w^0}\Big) \in \mathcal{G}(\mathbf{z}, \boldsymbol{\xi}^k),\ \forall k \in \mathcal{K}.
\end{aligned}
\tag{4.3.10}
$$

Compared to (4.3.8), the CVaR approximation scheme allows one to take a few samples at the tail rather than the extreme one, and in that way smooth up or stabilize the numerical computation. In the case of CVaR formulation, we replace the CVaR constraint with the equivalent system of linear inequalities by introducing additional positive dummy variables $\theta^1, \cdots \theta^k$ as follows

$$\begin{cases} \eta + \dfrac{1}{(1-\beta)K} \displaystyle\sum_{k=1}^{K} \theta^k \leq 0, \\ h(\mathbf{X}, \mathbf{x^0}, \mathbf{W}, \mathbf{w^0}, \lambda_0, \boldsymbol{\lambda}, \boldsymbol{\xi}^k) - \eta \leq \theta^k, \ \forall k \in \mathcal{K}, \\ \theta^k \geq 0, \ \forall k \in \mathcal{K}, \end{cases} \tag{4.3.11}$$

the substitution results in

$$\begin{aligned} \min_{\theta,\eta,\mathbf{z},\mathbf{X},\mathbf{x^0},\mathbf{W},\mathbf{w^0},\lambda_0,\boldsymbol{\lambda}} \quad & \mathbf{b}^T\mathbf{z} + \lambda_0 + \boldsymbol{\lambda}^T\boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{z} \in \{0,1\}^n, \\ & \eta + \frac{1}{(1-\beta)K} \sum_{k=1}^{K} \theta^k \leq 0, \\ & h(\mathbf{X}, \mathbf{x^0}, \mathbf{W}, \mathbf{w^0}, \lambda_0, \boldsymbol{\lambda}, \boldsymbol{\xi}^k) - \eta \leq \theta^k, \ \forall k \in \mathcal{K}, \\ & \left(\mathbf{X}\boldsymbol{\xi}^k + \mathbf{x^0}, \mathbf{W}\boldsymbol{\xi}^k + \mathbf{w^0}\right) \in \mathcal{G}(\mathbf{z}, \boldsymbol{\xi}^k), \ \forall k \in \mathcal{K}, \\ & \theta^k \geq 0, \ \forall k \in \mathcal{K}. \end{aligned} \tag{4.3.12}$$

The reformulation will e ectively address the non-smoothness caused by the $(.)_+$ operation but at the cost of introducing additional variables and constraints. It is worthwhile to do that here as the latter formulation will result in an overall MILP. A detailed formulation of problems (4.3.8), (4.3.9) and (4.3.12) is provided in Appendix A.1.

4.4 A semi-definite programming approach

In this section, we investigate (R-FLP) with the rst and second moment of the underlying random variables. Throughout this chapter, let $\mathscr{P}$ denote the set of all probability measures of ξ. We consider the following ambiguity set

$$\mathcal{P} = \left\{ P \in \mathscr{P} : \mathbb{E}_P[\boldsymbol{\xi}] = \boldsymbol{\mu}, \quad \mathbb{E}_P[\boldsymbol{\xi}\boldsymbol{\xi}^T] = \mathbf{Q} \right\}, \tag{4.4.1}$$

where $\boldsymbol{\mu}$ denotes the mean and $\mathbf{Q}$ the second moment, both of which are assumed to be known. Let

$$H(\mathbf{z}) = \sup_{P \in \mathcal{P}} \mathbb{E}_P \left[g(\mathbf{z}, \boldsymbol{\xi}) \right]. \tag{4.4.2}$$

Problem (4.4.2) is related to the classical problem of moments. Here, instead of finding a feasible distribution $P \in \mathcal{P}$, we want to find one which maximizes the expected value of $g(\mathbf{z}, \boldsymbol{\xi})$. For a discussion on the background of the problem of moments the interested reader is referred to [73]. Let $\mathscr{M}^+$ denote the set of all non-negative finite measures on measurable space $(\Xi, \mathscr{B})$. Then

$$\begin{aligned}
H(\mathbf{z}) := \quad & \sup_{P \in \mathscr{M}^+} \quad \int_\Xi g(\mathbf{z}, \boldsymbol{\xi}) P(d\boldsymbol{\xi}) \\
& \text{s.t.} \qquad \int_\Xi \xi_i \xi_j P(d\boldsymbol{\xi}) = Q_{ij}, \ \forall i, j = 1, \cdots, m, \\
& \qquad\qquad \int_\Xi \xi_i P(d\boldsymbol{\xi}) = \mu_i, \ \forall i = 1, \cdots, m, \\
& \qquad\qquad \int_\Xi P(d\boldsymbol{\xi}) = 1,
\end{aligned} \tag{4.4.3}$$

where Q_{ij} denotes the $(ij) - th$ component of Q, and μ_j the $j - th$ component of μ. The problem above is a semi-infinite linear program. Duality theory is often used to deal with the difficulty of solving such infinite dimensional problems; see for example, [91]. For the duality theory applied in the case of moment problems we refer the readers to [96]. Let $\mathbb{S}^{m \times m}$ denote the space of m by m real matrices. Let $\mathbf{Y} \in \mathbb{S}^{m \times m}$, $\mathbf{y} \in \mathbb{R}^m$ and $y_0 \in \mathbb{R}$ denote the dual variables associated with the moment constraints. We can then write the dual of problem (4.4.3) as follows

$$\begin{aligned}
H_D(\mathbf{z}) = \quad & \min_{\mathbf{Y}, \mathbf{y}, y_0} \quad \mathbf{Q} \bullet \mathbf{Y} + \boldsymbol{\mu}^{\mathbf{T}} \mathbf{y} + y_0 \\
& \text{s.t.} \qquad \boldsymbol{\xi}^T \mathbf{Y} \boldsymbol{\xi} + \boldsymbol{\xi}^T \mathbf{y} + y_0 \geq g(\mathbf{z}, \boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \mathbb{R}^m,
\end{aligned} \tag{4.4.4}$$

where $\mathbf{Q} \bullet \mathbf{Y}$ denote the Frobenius inner product of matrices $\mathbf{Q}$ and $\mathbf{Y}$. The weak duality condition, $H(\mathbf{z}) \leq H_D(\mathbf{z})$, can be shown easily (see [24] for the proof).

For strong duality results to hold, we make the following assumption

Assumption 4.4.1. *The linear matrix inequality* $\mathbf{Q} - \boldsymbol{\mu}\boldsymbol{\mu}^T \succ 0$ *hold, where* $A \succ 0$ *means* A *is positive definite.*

Note that, for $(\boldsymbol{\mu}, \mathbf{Q})$ to be valid rst and second moments of some random variable, it is necessary to have condition $\{\mathbf{Q} - \boldsymbol{\mu}\boldsymbol{\mu}^T \succeq 0\}$ satis ed. Assumption 4.4.1 is slightly stronger as we replace the $\succeq$ sign by the $\succ$ sign. This is needed for technicality reason in proving the strong duality result which is formally stated in the following proposition

Proposition 4.4.1. *Under Assumption 4.4.1,* $H(\mathbf{z}) = H_D(\mathbf{z})$.

Proof: Let us de ne $\mathcal{H} = \left\{(M, v) \mid M = M^T,\ M \succ vv^T\right\}$ and let $X \in \mathbb{R}^n$ be a random vector which follows the standard multivariate normal distribution. For any $(M, v) \in \mathcal{H}$, there exists a symmetric matrix W such that $W^2 = (M - vv^T)$. We can then construct a random variable $\boldsymbol{\xi} = WX + v$ which has a mean value of v and a covariance matrix of $(M - vv^T)$, i.e.,

$$\begin{aligned}
&\int_{\Xi} \xi_i \xi_j P(d\boldsymbol{\xi}) = M_{ij},\ \forall i, j = 1, \cdots, m, \\
&\int_{\Xi} \xi_i P(d\boldsymbol{\xi}) = v_i,\ \forall i = 1, \cdots, m, \qquad (4.4.5) \\
&\int_{\Xi} P(d\boldsymbol{\xi}) = 1.
\end{aligned}$$

In other words, using any choice of $(M, v) \in \mathcal{H}$ to replace $(\mathbf{Q}, \boldsymbol{\mu})$ in the R.H.S of (4.4.5) would lead to a feasible $\boldsymbol{\xi}$. In addition, we can show that $\mathcal{H}$ is an open set. Therefore, under Assumption 4.4.1, i.e. $(\mathbf{Q}, \boldsymbol{\mu}) \in \mathcal{H}$, we also have $(\mathbf{Q}, \boldsymbol{\mu})$ to belong to the interior of $\mathcal{H}$. As a result, there exists a neighbourhood $\mathcal{B}$ small enough around $(\mathbf{Q}, \boldsymbol{\mu})$ such that if we replace the R.H.S of (4.4.5) by any $(M, v) \in \mathcal{B}$, the system of equalities (4.4.5) is still feasible (for some di erent $\boldsymbol{\xi}$). This is the su cient condition for having strong duality result to hold, i.e. $H(\mathbf{z}) = H_D(\mathbf{z})$, as stated in [96, Proposition 3.4]. $\square$

From the strong duality result, problem (4.2.8) can be equivalently written as

$$\begin{aligned} \min_{\mathbf{z},\mathbf{Y},\mathbf{y},y_0} \quad & \mathbf{b}^T\mathbf{z} + \mathbf{Q} \bullet \mathbf{Y} + \boldsymbol{\mu}^{\mathbf{T}}\mathbf{y} + y_0 \\ \text{s.t.} \quad & \boldsymbol{\xi}^T\mathbf{Y}\boldsymbol{\xi} + \boldsymbol{\xi}^T\mathbf{y} + y_0 \geq g(\mathbf{z},\boldsymbol{\xi}),\ \forall \boldsymbol{\xi} \in \quad , \\ & \mathbf{z} \in \{0,1\}^n. \end{aligned} \tag{4.4.6}$$

Let us now write down the dual of the transportation problem described in problem (4.2.7)

$$\begin{aligned} g_D(\mathbf{z},\boldsymbol{\xi}) = \max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \quad & \sum_j \alpha_j\xi_j - \sum_i \beta_i(s_i z_i) \\ \text{s.t.} \quad & \alpha_j - \beta_i \leq c_{ij},\ \forall\ i \in I,\ j \in J, \\ & \alpha_j \leq C,\ \forall\ j \in J, \\ & \alpha_j, \beta_i \geq 0,\ \forall\ i \in I,\ j \in J, \end{aligned}$$

where α_j, $j \in J$, are dual variables associated with demand constraints and β_i, $i \in I$, are dual variables associated with supply constraints in model (4.2.7). Observe that problem (4.2.7) satis es Mangasarian-Fromovitz constraint quali cation (MFCQ). Thus the Lagrange multipliers of the problem are bounded and there exists a positive number C' such that the problem above is equivalent to the following

$$\begin{aligned} \text{(DTP)} \qquad g_D(\mathbf{z},\boldsymbol{\xi}) = \max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \quad & \sum_j \alpha_j\xi_j - \sum_i \beta_i(s_i z_i) \\ \text{s.t.} \quad & \alpha_j - \beta_i \leq c_{ij},\ \forall\ i \in I,\ j \in J, \\ & \alpha_j \leq C,\ \forall\ j \in J, \\ & \alpha_j, \beta_i \geq 0,\ \forall\ i \in I,\ j \in J, \\ & \beta_j \leq C', \forall j \in J. \end{aligned}$$

To ease the notation, let $\gamma_\alpha := \boldsymbol{\alpha}$, $\gamma_\beta := \boldsymbol{\beta}$, and $\boldsymbol{\gamma} := (\gamma_\alpha, \gamma_\beta)$. Let denote the feasible set of (DTP). It is easy to observe that is a polyhedral in $\mathbb{R}^{|I|+|J|}$ where $|I|$ and $|J|$

denote the cardinality of the index set I and J respectively.

Proposition 4.4.2. *is a bounded polyhedral with a finite number of vertices.*

Proof: It is easy to observe that is a bounded polyhedral. The second part of assertion follows from Balinski [10]. $\square$

Let $\{\gamma^1, \cdots, \gamma^N\}$ denote the set of vertices. Using the notation introduced above, we can rewrite (DTP) in a neater form

$$\begin{aligned} g_D(\mathbf{z}, \boldsymbol{\xi}) = \quad & \max_{\boldsymbol{\gamma}(\mathbf{z})} \quad \boldsymbol{\xi}^T \gamma_\alpha - {\gamma_\beta}^T (\mathbf{s} \circ \mathbf{z}) \\ & \text{s.t.} \quad \boldsymbol{\gamma} \in \quad , \end{aligned} \tag{4.4.7}$$

where $(\mathbf{s} \circ \mathbf{z})$ denotes an m-dimensional vector with components $s_i z_i$ for $i \in I$. Combining (4.4.6) and (DTP), we can recast the robust facility location problem (4.2.8) as a semi-de nite program through the following proposition.

Proposition 4.4.3. *Let $\mathcal{P}$ be defined as in (4.4.1) and $\equiv \mathrm{I\!R}^m$. Under Assumption 4.4.1, the two-stage distributionally robust facility location problem (4.2.8) can be reformulated as the following semi-definite optimization problem:*

$$\text{(R-SDP)} \qquad \begin{aligned} & \min_{\mathbf{z}, \mathbf{Y}, \mathbf{y}, y_0} \quad \mathbf{b}^T \mathbf{z} + \mathbf{Q} \bullet \mathbf{Y} + \boldsymbol{\mu}^{\mathbf{T}} \mathbf{y} + y_0 \\ & \quad s.t. \qquad \mathbf{z} \in \{0, 1\}^n, \\ & \qquad\qquad \begin{bmatrix} y_0 + {\gamma_\beta}^T (\mathbf{s} \circ \mathbf{z}) & \frac{1}{2}(\mathbf{y} - \gamma_\alpha)^T \\ \frac{1}{2}(\mathbf{y} - \gamma_\alpha) & \mathbf{Y} \end{bmatrix} \succeq 0, \ \forall \boldsymbol{\gamma} \in \{\gamma^1, \cdots, \gamma^N\}. \end{aligned} \tag{4.4.8}$$

Here and later on we write $M \succeq 0$ for matrix M being positive semi-definite.

Proof: It follows from Proposition 4.4.1 that, under Assumption 4.4.1, $H(\mathbf{z}) = H_D(\mathbf{z})$ and problem (4.2.8) can be reformulated as problem (4.4.6). Note that the reformulation still involves $g(\mathbf{z}, \boldsymbol{\xi})$. Since $g(\mathbf{z}, \boldsymbol{\xi})$ and $g_D(\mathbf{z}, \boldsymbol{\xi})$ are primal and dual LPs of each other and since both of them are feasible (i.e. by setting $\mathbf{x} = 0$, $\omega_j = d_j, \forall j \in J$ in the primal and $\alpha = \beta = 0$ in the dual), strong duality result holds and we have $g(\mathbf{z}, \boldsymbol{\xi}) = g_D(\mathbf{z}, \boldsymbol{\xi})$.

Thus, we can replace the second stage transportation problem through its dual and rewrite the constraint of problem (4.4.6) as:

$$\boldsymbol{\xi}^T\mathbf{Y}\boldsymbol{\xi} + \boldsymbol{\xi}^T\mathbf{y} + y_0 \geq \max_{\boldsymbol{\gamma}\in\Gamma}\left\{\boldsymbol{\xi}^T\boldsymbol{\gamma}_\alpha - {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z})\right\}, \quad \forall\boldsymbol{\xi}\in\mathbb{R}^m. \tag{4.4.9}$$

Since Γ is bounded with a finite set of extreme points $\{\boldsymbol{\gamma}^1, \cdots, \boldsymbol{\gamma}^N\}$, the maximizer of the LP on the R.H.S of (4.4.9) occurs at one of the extreme points. Thus, (4.4.9) can be equivalently rewritten as

$$\boldsymbol{\xi}^T\mathbf{Y}\boldsymbol{\xi} + \boldsymbol{\xi}^T\mathbf{y} + y_0 \geq \boldsymbol{\xi}^T\boldsymbol{\gamma}_\alpha - {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z}), \quad \forall\boldsymbol{\xi}\in\mathbb{R}^m,\ \forall\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^1, \cdots, \boldsymbol{\gamma}^N\}.$$

A simple rearrangement yields

$$\min_{\boldsymbol{\xi}}\left\{\boldsymbol{\xi}^T\mathbf{Y}\boldsymbol{\xi} + \boldsymbol{\xi}^T(\mathbf{y} - \boldsymbol{\gamma}_\alpha) + y_0 + {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z})\right\} \geq 0, \quad \forall\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^1, \cdots, \boldsymbol{\gamma}^N\}. \tag{4.4.10}$$

We can show that inequality (4.4.10) holds if and only if

$$\begin{bmatrix} y_0 + {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z}) & \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma}_\alpha)^T \\ \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma}_\alpha) & \mathbf{Y} \end{bmatrix} \succeq 0, \quad \forall\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^1, \cdots, \boldsymbol{\gamma}^N\}. \tag{4.4.11}$$

Here, it is very clear that (4.4.11) implies (4.4.10). For the reverse, suppose that (4.4.11) does not hold, i.e. there exists γ and $(q_0, \boldsymbol{q})$ such that

$$\begin{bmatrix} q_0 & \boldsymbol{q}^T \end{bmatrix} \begin{bmatrix} y_0 + {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z}) & \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma}_\alpha)^T \\ \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma}_\alpha) & \mathbf{Y} \end{bmatrix} \begin{bmatrix} q_0 \\ \boldsymbol{q} \end{bmatrix} < 0.$$

We then need to show that (4.4.10) does not hold neither. If $q_0 \neq 0$, then we can construct $\boldsymbol{\xi} = \boldsymbol{q}/q_0$ and obtain $\boldsymbol{\xi}^T\mathbf{Y}\boldsymbol{\xi} + \boldsymbol{\xi}^T(\mathbf{y} - \boldsymbol{\gamma}_\alpha) + y_0 + {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z}) < 0$ which means (4.4.10) does not hold. If $q_0 = 0$, we have $\boldsymbol{q}^T\mathbf{Y}\boldsymbol{q} < 0$. We can then construct $\boldsymbol{\xi} = \delta\boldsymbol{q}$ with sufficiently large δ such that $\delta^2\boldsymbol{q}^T\mathbf{Y}\boldsymbol{q} + \delta\boldsymbol{q}^T(\mathbf{y} - \boldsymbol{\gamma}_\alpha) + y_0 + {\boldsymbol{\gamma}_\beta}^T(\mathbf{s}\circ\mathbf{z}) < 0$ which also means (4.4.10) does not hold.

Finally, we can replace the constraint in (4.4.4) with the SDP constraint (4.4.11) and obtain the SDP (4.4.8).

□

Notice that the derivation from inequality (4.4.10) to inequality (4.4.11) requires $\equiv$ $\mathrm{I\!R}^m$. In practice, there is often some information on the bounds of the uncertain parameters. For example the customer demand cannot take negative values. In order to handle the inde nite set of constraints that appears in problem (4.4.6) for this case, we will approximate the inde nite constraint with a nite set of semi-de nite constraints as shown next.

Suppose the support set is speci ed as $= \prod_{j \in J} {}_j$, where ${}_j = [\underline{\xi}_j, \bar{\xi}_j]$ for all $j \in J$ and $\underline{\boldsymbol{\xi}}$ and $\bar{\boldsymbol{\xi}}$ are some given lower and upper bounds. For technicality reason in proving strong duality, we make an assumption that there exists a random vector X with support set such that $E[X] = 0$ and $E[XX^T] = I$, where $I \in \mathrm{I\!R}^{m \times m}$ is the identity matrix.

Under this new assumption and Assumption 4.4.1, we can show that the strong duality result still holds where the proof is very similar to that of Proposition 4.4.1. The only di erence is in the way we construct the random variable X (i.e., instead of choosing a multi-variate normal random variable, we choose X such that $E[X] = 0$ and $E[XX^T] = I$). It is noted that, the new assumption can be relaxed further by a proper scaling of the random variables $\boldsymbol{\xi}$.

Once we have derived the strong duality result, problem (4.4.4) become

$$\begin{aligned} H_D(\mathbf{z}) = \min_{\mathbf{Y}, \mathbf{y}, y_0} \quad & \mathbf{Q} \bullet \mathbf{Y} + \boldsymbol{\mu}^{\mathbf{T}} \mathbf{y} + y_0 \\ \text{s.t.} \quad & \boldsymbol{\xi}^T \mathbf{Y} \boldsymbol{\xi} + \boldsymbol{\xi}^T \mathbf{y} + y_0 \geq g(\mathbf{z}, \boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \ , \end{aligned} \tag{4.4.12}$$

where the only di erence compared to problem (4.4.4) is that we have replaced the semi-in nite constraint from $\{\forall \boldsymbol{\xi} \in \mathrm{I\!R}^n\}$ to $\{\forall \boldsymbol{\xi} \in \ \}$. Inequality (4.4.10) now become

$$\min_{\underline{\boldsymbol{\xi}} \leq \boldsymbol{\xi} \leq \bar{\boldsymbol{\xi}}} \left[\boldsymbol{\xi}^T \mathbf{Y} \boldsymbol{\xi} + \boldsymbol{\xi}^T (\mathbf{y} - \gamma_\alpha) + y_0 + \boldsymbol{\gamma}_\beta{}^T (\mathbf{s} \circ \mathbf{z}) \right] \geq 0, \quad \forall \gamma \in \{\gamma^1, \cdots, \gamma^N\}, \tag{4.4.13}$$

which is equivalent to

$$\phi(\gamma) \geq 0, \ \forall \gamma \in \{\gamma^1, \cdots, \gamma^N\}, \tag{4.4.14}$$

where $\phi(\gamma)$ is the optimal value of the following program

$$\begin{aligned}
\min_{\boldsymbol{\xi}} & \left\{ \begin{bmatrix} 1 & \boldsymbol{\xi} \end{bmatrix} \begin{bmatrix} y_0 + \boldsymbol{\gamma_\beta}^T(\mathbf{s} \circ \mathbf{z}) & \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha})^T \\ \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha}) & \mathbf{Y} \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{\xi} \end{bmatrix} \right\} \\
\text{s.t.} & \quad \begin{bmatrix} 1 & \boldsymbol{\xi} \end{bmatrix} V_j \begin{bmatrix} 1 \\ \boldsymbol{\xi} \end{bmatrix} \leq 0, \qquad \forall j \in J,
\end{aligned}$$

where $V_j = \begin{bmatrix} \bar{\xi}_j \underline{\xi}_j & v_j^T \\ v_j & I_j \end{bmatrix}$, with I_j denoting an $m \times m$ matrix with all elements being 0 except and 1 at (j, j), and v_j is an m-dimensional vector with all components are equal to zero except for the jth element which is equal to $-(\underline{\xi}_j + \bar{\xi}_j)/2$.

The S-lemma [46] provides a su cient condition for the non-negativity of the quadratic objective function over the quadratic inequalities corresponding to the bounds. In other words, for the conditions (4.4.14) to be satis ed for each $\gamma \in$, it su ces that there exists $\mathbf{h} \geq 0$ such that

$$\begin{bmatrix} y_0 + \boldsymbol{\gamma_\beta}^T(\mathbf{s} \circ \mathbf{z}) & \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha})^T \\ \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha}) & \mathbf{Y} \end{bmatrix} + \sum_{j \in J} h_j V_j \succeq 0,$$

where h_j denotes the jth component of vector $\mathbf{h}$. Consequently, the SDP problem (4.4.8) can be reformulated as

$$\begin{aligned}
\min_{\mathbf{z}, \mathbf{Y}, \mathbf{y}, y_0, \mathbf{h}} \quad & \mathbf{b}^T\mathbf{z} + \mathbf{Q} \bullet \mathbf{Y} + \boldsymbol{\mu}^{\mathbf{T}}\mathbf{y} + y_0 \\
\text{s.t.} \quad & \mathbf{z} \in \{0, 1\}^n, \\
& \mathbf{h} \geq 0, \\
& \begin{bmatrix} y_0 + \boldsymbol{\gamma_\beta}^T(\mathbf{s} \circ \mathbf{z}) & \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha})^T \\ \frac{1}{2}(\mathbf{y} - \boldsymbol{\gamma_\alpha}) & \mathbf{Y} \end{bmatrix} + \sum_{j \in J} h_j V_j \succeq 0, \ \forall \boldsymbol{\gamma} \in \{\boldsymbol{\gamma}^1, \cdots, \boldsymbol{\gamma}^N\}.
\end{aligned} \tag{4.4.15}$$

Remark: Problem (4.4.15) is very similar to problem (4.4.8) except for the newly introduced decision variable $\mathbf{h}$. If we restrict $\mathbf{h} = 0$, then problem (4.4.15) is exactly the same with problem (4.4.8). Each $\mathbf{h} > 0$ essentially enlarges the feasible domain of $(\mathbf{Y}, \mathbf{y}, y_0)$ in problem (4.4.8) to a larger feasible domain of $(\mathbf{Y}, \mathbf{y}, y_0)$ in problem (4.4.15).

From computational perspective, problems (4.4.8) and (4.4.15) are complex to solve for two main reasons.

- The problem constitutes N semi-definite constraints where N is the number of vertices of . Balinski [10] shows that for the case of DTP polyhedra, N is finite but it grows exponentially as the problem size increases. This means the number of semi-definite constraints increases at exponential rate with the increase of the problem size.
- Due to the binary decision variables $\mathbf{z}$, the problem is NP-hard. This is well known in deterministic FLPs.

In the following subsections, we address these numerical challenges. Concerning the first issue of having many SDP constraints, we propose a constraint generation (CG) algorithm. We first fix the binary facility location variables $\mathbf{z}$ and focus on solving problem (4.4.15) using the constraint generation method, which involves solving a series of SDP programs with progressively higher number of constraints until the optimal solution is found. At limit, the problem reaches the full complexity of problem (4.4.15). In practice, however, it might converge to an optimal solution much sooner. To deal with the second issue of having binary variables, we use a genetic algorithm (GA), which utilized the CG algorithm for computing the fitness function, to search for the optimal facility location variables.

4.4.1 Constraint generation algorithm

We start implementing the CG algorithm by treating the decision variables $\mathbf{z}$ as a parameter, i.e. setting it to a fixed value. This will reduce problem (4.4.15) to an SDP program with a linear objective function and a finite number of constraints (denoted by the set $\mathcal{R}$) corresponding to some extreme points of the DTP problem. The enumeration of all extreme points can be numerically cumbersome since that could be very large. A less complex sub-problem can be constructed by selecting a subset of constraints $\mathcal{R}_{sub} \subseteq \mathcal{R}$. However, the solution obtained might violate some constraints of the original problem. The next step is to identify such violated constraints which are then added to the sub-problem in an iterative manner. This procedure is applied until the relaxed problem solution is feasible and the global optimal solution is reached. The detail of such procedure is explained in Algorithm 1:

Algorithm 1 Constraint generation algorithm

Step 0. **Sub-problem initiation:** set $k := 1$; $\mathcal{R}_{sub} = \{\gamma^1, \cdots, \gamma^p\}$ for $p \ll N$.

Step 1. **Sub-problem solution:** solve problem (4.4.15) with the subset of constraints $\mathcal{R}_{sub}$ to obtain $\mathbf{Y}^{(k)}$, $\mathbf{y}^{(k)}$, $y_0^{(k)}$ and $\mathbf{h}^{(k)}$.

Step 2. **Feasibility:** verify feasibility of the solution $\mathbf{Y}^{(k)}$, $\mathbf{y}^{(k)}$ and $y_0^{(k)}$ through Algorithm 2 and if:

Step 2.1 $(\mathbf{Y}^{(k)}, \mathbf{y}^{(k)}, y_0^{(k)})$ is infeasible, then add the violating constraint to $\mathcal{R}_{sub}$, set $k = k + 1$ and go to *Step* 1.

Step 2.2 otherwise, optimal solution found, stop the algorithm.

Note that *Step* 2, on checking the feasibility of $(\mathbf{Y}^{(k)}, \mathbf{y}^{(k)}, y_0^{(k)})$ to the original problem, we also identify a violating constraint if that was not the case. To nd a violating constraint, we need to solve the following problem:

$$\min_{\gamma\in\Gamma} \quad \begin{bmatrix} y_0^{(k)} + {\gamma_\beta}^T(\mathbf{s}\circ\mathbf{z}) & \frac{1}{2}(\mathbf{y}^{(k)} - \gamma_\alpha)^T \\ \frac{1}{2}(\mathbf{y}^{(k)} - \gamma_\alpha) & \mathbf{Y}^{(k)} \end{bmatrix} + \sum_{j\in J} h_j^{(k)} V_j \not\succeq 0, \qquad (4.4.16)$$

which can be done by solving

$$\min_{\gamma\in\Gamma}\left(\min_{\underline{\boldsymbol{\xi}}\,\leq\,\boldsymbol{\xi}\,\leq\,\bar{\boldsymbol{\xi}}}\left[\boldsymbol{\xi}^T\mathbf{Y}^{(k)}\boldsymbol{\xi} + \boldsymbol{\xi}^T(\mathbf{y}^{(k)} - \gamma_\alpha) + y_0^{(k)} + {\gamma_\beta}^T(\mathbf{s}\circ\mathbf{z})\right]\right), \qquad (4.4.17)$$

and checking whether this is greater than or equal to zero. For each xed γ, the inner minimization problem of (4.4.17) is a quadratic programming problem. Let us denote its objective by

$$W = \boldsymbol{\xi}^T \mathbf{Y}^{(k)} \boldsymbol{\xi} + \boldsymbol{\xi}^T (\mathbf{y}^{(k)} - \gamma_\alpha) + y_0^{(k)} + {\gamma_\beta}^T (\mathbf{s} \circ \mathbf{z}). \tag{4.4.18}$$

The feasibility assessment procedure is nested within the CG algorithm and explained in detail in Algorithm 2. Although CG is not an exact Algorithm, it provides an e cient method to deal with the non-convexity of problem (4.4.17).

Algorithm 2 Feasibility veri cation procedure

Step (A) **Initialization**: set $r = 0$, generate random $\boldsymbol{\xi}^{(0)}$ and set $\boldsymbol{\xi} = \boldsymbol{\xi}^{(r)}$.

Step (B) **Extreme point calculation**: solve the DTP (4.4.7) to get $\gamma^{(r)}$.

Step (C) **Violation verification**: use $\boldsymbol{\xi}^{(r)}$ and $\gamma^{(r)}$ to calculate

$$W^{(r)} = {\boldsymbol{\xi}^{(r)}}^T \mathbf{Y}^{(k)} \boldsymbol{\xi}^{(r)} + {\boldsymbol{\xi}^{(r)}}^T (\mathbf{y}^{(k)} - \gamma_\alpha^{(r)}) + y_0^{(k)} + {\gamma_\beta^{(r)}}^T (\mathbf{s} \circ \mathbf{z}),$$

and if:

Step (C.1) $W^{(r)} < 0$, conclude $(\mathbf{Y}^{(k)}, \mathbf{y}^{(k)}, y_0^{(k)})$ is infeasible solution,
stop Algorithm 2
(and go back to *Step* 2.1 of Algorithm 1 to add the violating constraint.)

Step (C.2) otherwise, if $(W^{(r-1)} - W^{(r)}) \geq 10^{-6}$:

Step (C.2.1) use $\gamma_\alpha^{(r)}$ and $\gamma_\alpha^{(r)}$ to compute $\boldsymbol{\xi}^*$ through solving the quadratic problem (the inner minimization problem of (4.4.17)), set $r = r + 1$, $\boldsymbol{\xi}^{(r)} = \boldsymbol{\xi}^*$ and go to *Step*.(B).

Step (C.2.2) otherwise, feasible solution reached, go to *Step* 2.2 of Algorithm 1, i.e. stop the CG algorithm.

4.4.2 Genetic algorithm (GA)

The constraint generation algorithm described in subsection 4.4.1 evaluates the objective function of the robust facility location problem (4.2.8) for each xed $\mathbf{z}$. We need to search for the optimal value of the facility location variable $\mathbf{z}$. The original SDP problem can be rewritten as a mixed integer non-linear programming problem

$$\begin{array}{ll} \min\limits_{\mathbf{z}} & \mathbf{b^T z} + H_D(\mathbf{z}) \\ \text{s.t.} & \mathbf{z} \in \{0, 1\}^n. \end{array}$$

Since the MINLP problem above is NP-complete, we propose a heuristic method to solve the problem. More speci cally, we use a Genetic Algorithm (GA), which is a probabilistic search method that mimics the biological model of natural selection. The GA algorithm applies the principle of \survival of the ttest" to a population of potential solutions to produce progressively better solutions over the generations. The detail GA is explained in Algorithm 3.

Algorithm 3 Genetic Algorithm

Step 0. **Population initialization**: randomly generate a chromosome population of size $\mathcal{N}$: $\mathbf{z}^1, \cdots \mathbf{z}^{\mathcal{N}}$,

Step 1. **Evaluation**: for each chromosome calculate the tness, using the solution of SDP problem $H_D(\mathbf{z})$ as the parameter,

Step 2. **Verification**: is the termination criterion is met?

Step 2.1 if yes, stop.

Step 2.2 if no proceed to next step,

Step 3. **Selection**: a pair of \parent" solutions are selected from population for breeding,

Step 4. **Crossover**: with the cross over probability of p^c, exchange parts of the two selected parents and create two o spring solutions,

Step 5. **Mutation**: with the mutation probability of p^m, randomly change the gene values of the two o spring solutions,

Step 6. **New population**: create a new population using the new o springs and check if the new population have the size $\mathcal{N}$?

Step 6.1 if yes, go to *Step* 1.

Step 6.2 if no, go to *Step* 3.

In order to illustrate the synchronization of the GA and CG approaches, the process of interaction between the two algorithms is shown in Figure.4.1

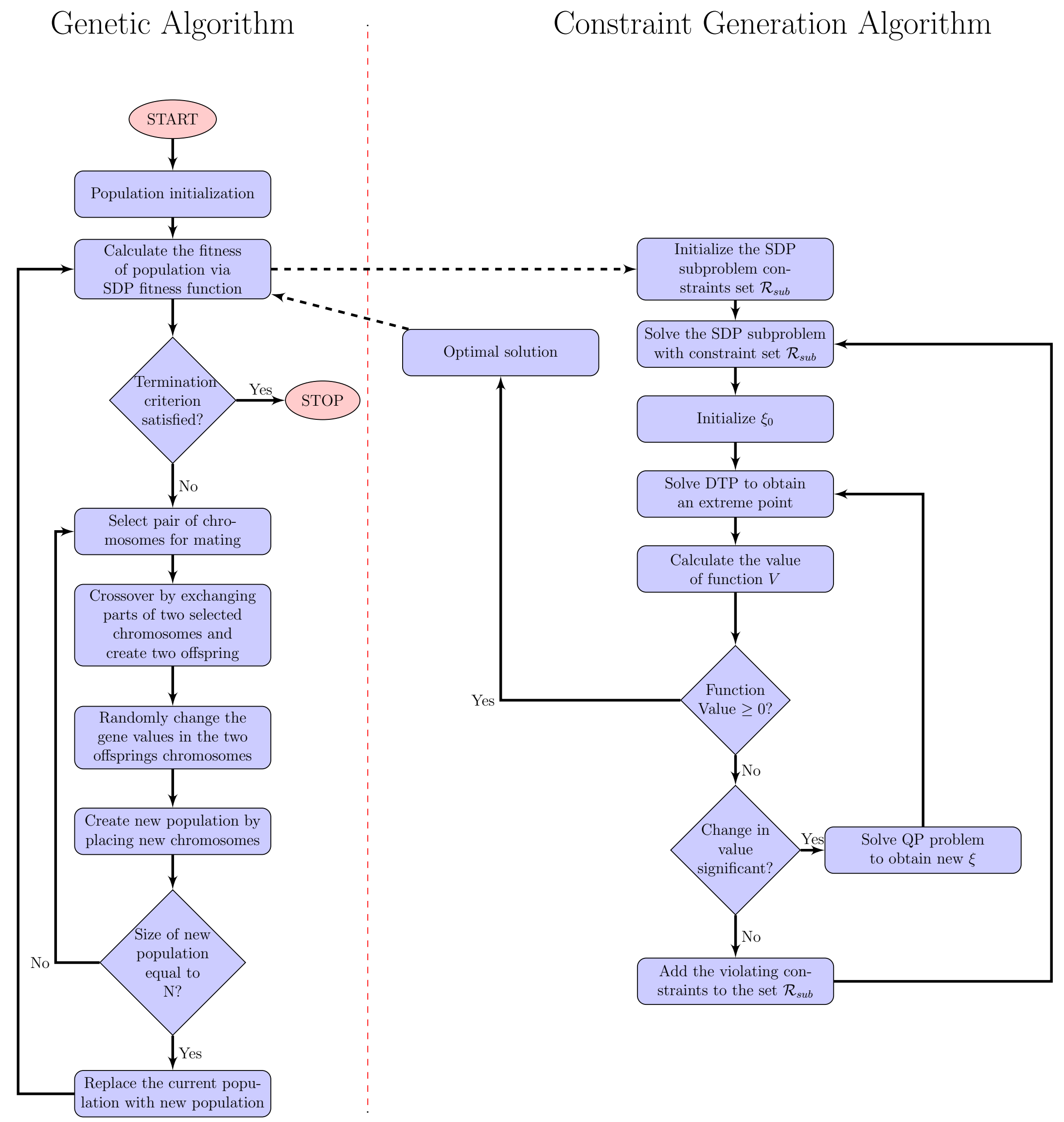


Figure 4.1: A hybrid algorithm

4.5 Computational results

In this section we report the numerical experiments performed to evaluate the proposed methodologies. For convenience in recapping these methods, Table 4.1 provides a quick reference on their abbreviations and the key di erences among them.

Abbreviation	Problem	Methods	Known information on demand
D-FLP	(4.2.1)	Deterministic FLP	Actual values ($\mathbf{d}$)
S-FLP	(4.2.6)	Stochastic FLP	Probability distribution (P)
R-SIP	(4.3.4)	Robust FLP with semi-in nite formulation	
R-LDR	(4.3.5)	LDR approximation of R-SIP problem	First moment ($\boldsymbol{\mu}$)
R-CVaR	(4.3.7)	CVaR approximation of R-LDR problem	
R-SDP	(4.4.15)	Robust FLP with semi-de nite formulation	First and second moments ($\boldsymbol{\mu}, \mathbf{Q}$)

Table 4.1: Abbreviations and reference of methodologies used

4.5.1 A small case study

We study a small scale facility location problem to illustrate the quality of solutions obtained from the proposed solution methods. We randomly generate a facility location problem with 4 demand nodes and 3 potential locations to build facilities. The transportation costs are assumed to be proportional to the distances between the customers and the facilities. Figure 4.2 shows the network layout of this problem.

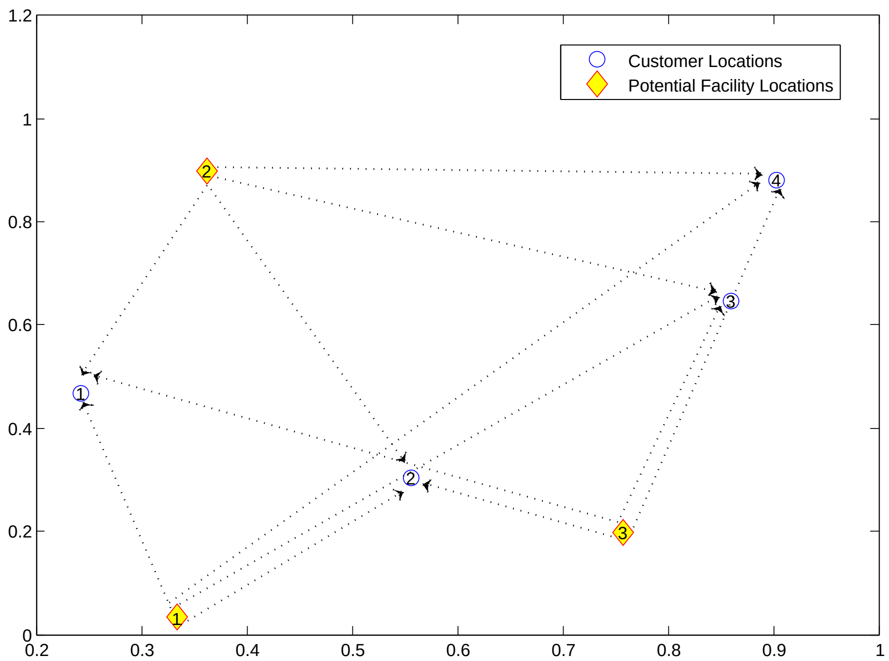


Figure 4.2: The network of facility location problem

The transportation cost, fixed investment cost and the capacity of each potential facility are given in Table 4.2.

Supply points	Demand points				Supply capacity	Fixed initial cost
	dem_1	dem_2	dem_3	dem_4		
sup_1	14	12	21	25	200	2000
sup_2	14	18	16	16	300	3200
sup_3	17	10	14	19	254	3700

Table 4.2: Transportation and investment costs and capacity of the facilities

We assume that customer demands are unknown prior to the construction of facilities and we are only given the first moment information of the demand with $\boldsymbol{\mu} = (150, 150, 100, 100)$. Moreover, in the SDP formulation of the problem, we are also provided with the second moment information of the uncertain demand

$$\mathbf{Q} = \begin{pmatrix} 22669.34 & 22511.07 & 15038.4 & 15026.78 \\ 22511.07 & 22551.57 & 15000.53 & 14988.58 \\ 15038.4 & 15000.53 & 10045.48 & 10013.84 \\ 15026.78 & 14988.58 & 10013.84 & 10031.04 \end{pmatrix}$$

4.5.1.1 Robust SDP formulation

Under the assumption of available first and second moments, we formulate the problem as a robust SDP (R-SDP). Since the problem size is small, we first consider all combinations of the possible facility location decisions (2^3 possible combinations of z). For each potential solution, the full R-SDP problem (4.4.15) is constructed by including all of the extreme points of the dual transportation polytopes (obtained by using the signature method in [9]). The full R-SDP problem is then solved for each potential solution and the total cost of each decision is compared in Table 4.3. Note that the very high cost of the \infeasible" solutions is due to high cost of using the external facility when demand exceeds the capacity of facilities built. It can be observed that solution number 4, i.e. $\mathbf{z} = (0, 1, 1)$ has the least cost and therefore is optimal.

Facility Decision	Possible FL decisions							
	sol_1	sol_2	sol_3	sol_4	sol_5	sol_6	sol_7	sol_8
z_1	0	0	0	**0**	1	1	1	1
z_2	0	0	1	**1**	0	0	1	1
z_3	0	1	0	**1**	0	1	0	1
Cost	100,132.1	56,012.3	47,851.8	**13,723.6**	64,639.2	21,349.9	13,829.1	15,507.8

Table 4.3: Worst expected cost associated with each possible facility location decision

In order to illustrate the performance of the proposed constraint generating algorithm for the R-SDP formulation, we rst xed the facility decisions to $\mathbf{z} = (1, 1, 1)$. In each run, the problem was initiated by randomly selecting one of the SDP constraints corresponding one of the extreme points of the dual transportation problem. We record the objective value after each iteration of the constraint generation process. Figure 4.3 summarizes the objective value after each iteration of algorithm (by taking the average over 100 runs, each with a di erent starting point). It can be observed that the convergence of the constraint generation algorithm to the optimal value takes place after around 15 iterations.

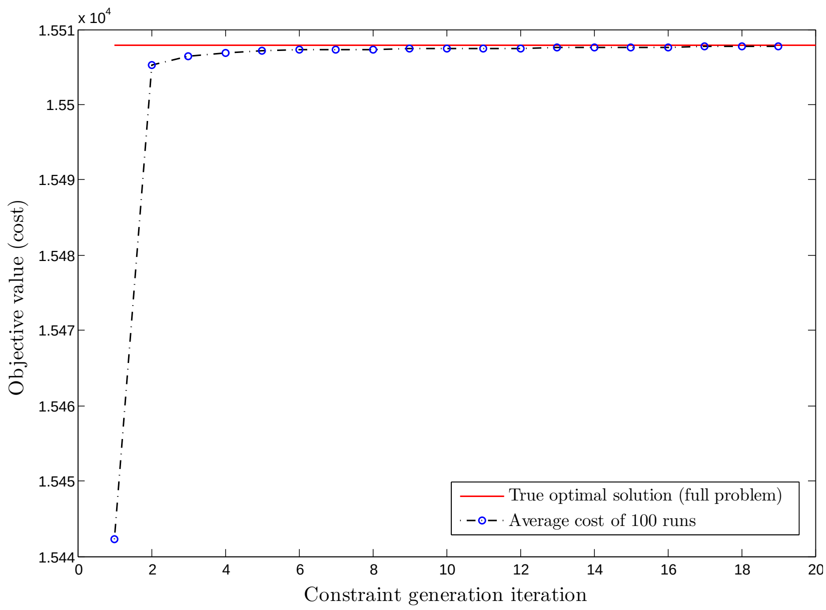


Figure 4.3: Convergence of the constraint generation algorithm solutions

After finding the optimal solution for a fixed facility location decision $\mathbf{z}$, the next step of the proposed solution method is to find the optimal $\mathbf{z}$ using GA. However, this is not necessary for this small instance as it has only 8 possible solutions.

4.5.1.2 Robust SIP formulation

Let us assume that we are only given the first moment information $\boldsymbol{\mu}$ and the second moment Q is unknown. We implement the second proposed method and formulate the FLP as a semi-infinite program. For assessing the quality of LDR approximation (R-LDR) of the \true" robust SIP (R-SIP) solution, we limit the support set of the random demand for each customer to 2000 values generated from the uniform distribution $\mathcal{U}(0, 250)$. We then solve the full R-SIP and its R-LDR approximation over this support set. As shown in Table 4.4, R-LDR solution provides an upper bound approximation with 0.9% deviation from the original R-SIP solution. For the comparison purposes, we also solve the deterministic version of the instance by assuming that the demand values are knows and given by $\boldsymbol{\mu}$. The solution from deterministic problem (D-FLP) is then benchmarked against the robust solutions in Table 4.4.

Models	Optimal FLP decisions			Total cost
	z_1	z_2	z_3	
D-FLP	1	1	0	12,300.00
R-SDP	0	1	1	13,723.61
R-SIP	1	1	1	15,641.42
R-LDR	1	1	1	15,781.42

Table 4.4: Deterministic versus robust solutions

The deterministic solution is to install just enough capacity to meet the predicted (assumed to be known) demand values by locating facilities 1 and 2. In other words, D-FLP solution provides no flexibility for possible variation in future demand. The robust solutions on the other hand, offers to install the facilities with a higher total capacity at a higher total cost (i.e. constructing facilities 2 and 3 in R-SDP case and all facilities in R-SIP case), to hedge against the risk of not meeting the customer demand.

In the next step, we implement the proposed CVaR approximation (R-CVaR) of the R-LDR solution. Using the same support set of 2000 values, we solved the R-CVaR approximation with various β values. The results are presented in Figure 4.4. It can be observed that the CVaR solution approximates the R-LDR optimal solution consistently and without any error for $\beta \geq 0.7$.

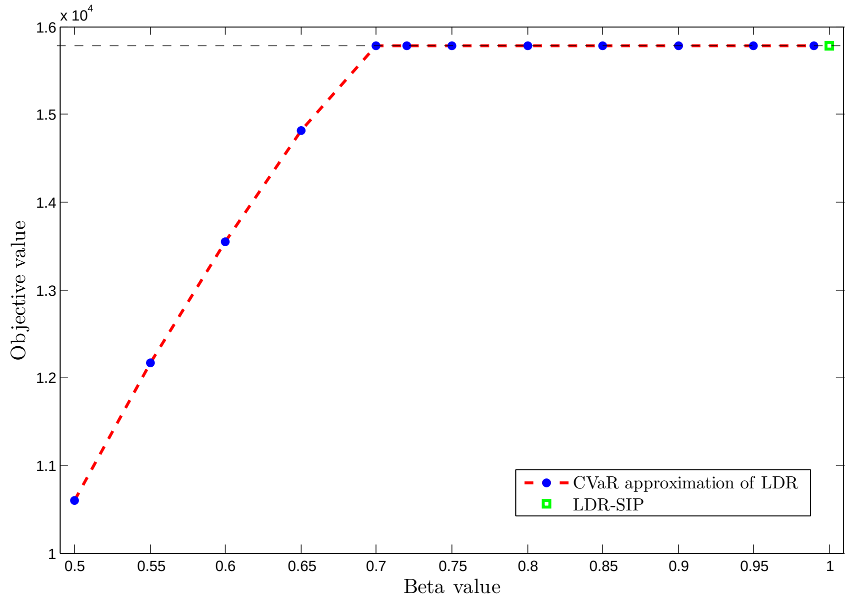
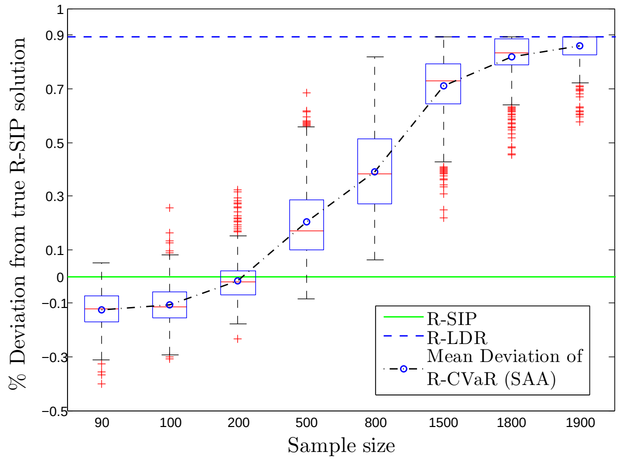


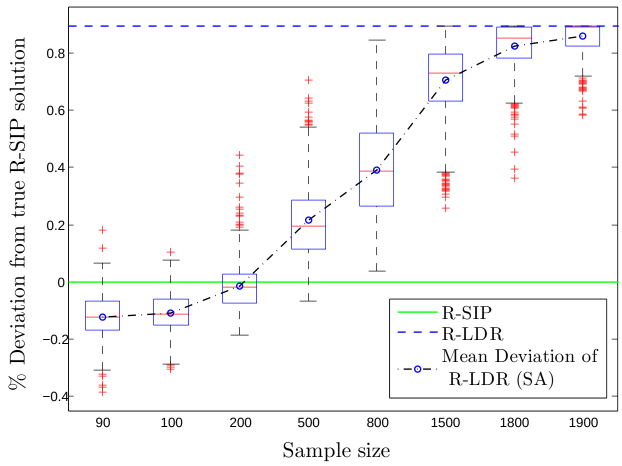
Figure 4.4: CVaR approximation of R-LDR problem for various β values

4.5.1.3 Discretization through sampling

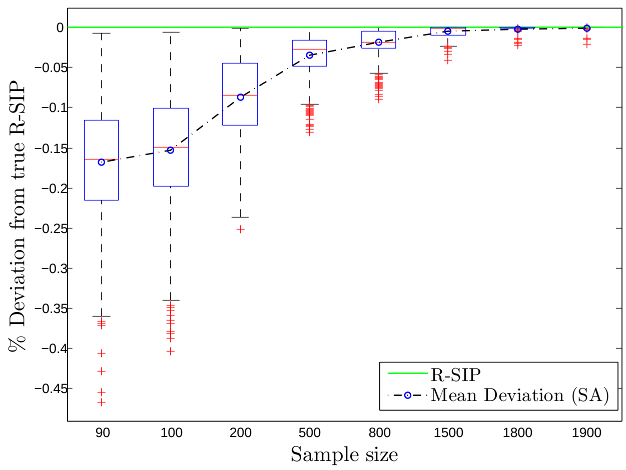
The complexity of the R-LDR and R-CVaR schemes increases by the number of scenarios. As described in subsection 4.3.2, one way to resolve this issue is to use sample approximation (SA) for R-LDR and sample average approximation for R-CVaR. These involve drawing i.i.d samples from the underlying distribution of the uncertainty. In the case of the rst instance, we have chosen the samples from the same support set used to run the full R-SIP tests. To assess the quality of sample approximation, we solve R-SIP, R-DLR and R-CVaR using various sample sizes. For each sample size, we carried out 1000 independent runs. The β value for all of the CVaR instances was set to 0.99. The normalized deviation of the sample approximations of all problems from the true robust (full R-SIP) solution is shown in Figure 4.5 for various choices of the sample sizes.



(a) SAA of R-CVaR



(b) SA of R-LDR



(c) SA of R-SIP

Figure 4.5: Normalized deviation of approximation problems from the true robust solution

It can be observed that, for each sample size, the mean deviation of R-CVaR and R-LDR solutions from the true robust solution are very similar. They range from -0.4% to 0.9% of the true solution (here 0.9% deviation means that the approximate solution is 0.9% higher than the true value of the robust solution). We can also see that the sample approximation of R-CVaR and R-LDR converge to the full R-LDR solution as sample size increases. Furthermore, the sampling method applied to R-SIP provides a very good approximation of the true solution even for small sample sizes.

Figure 4.6 shows the average computation times for 1000 runs of each sample size.

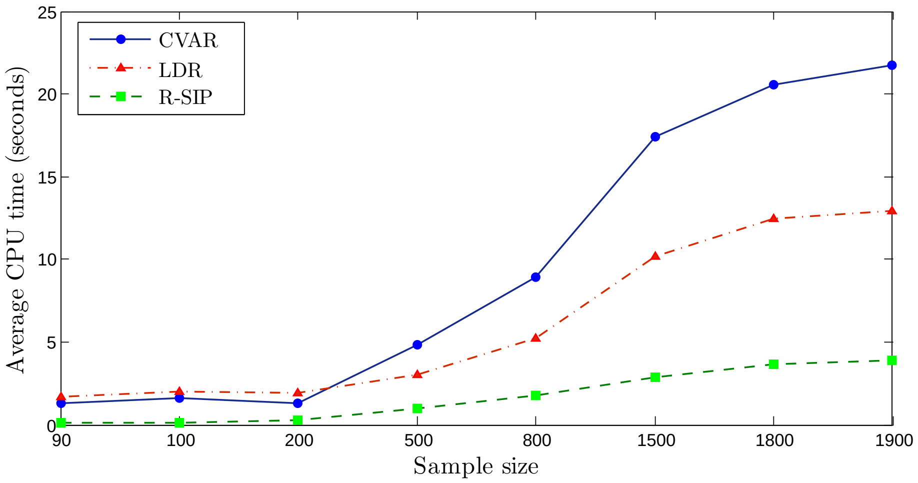


Figure 4.6: Average computation time

4.5.2 Medium size test instances

In this section, we consider a set of larger facility location problems. Test instances are selected from those existing in the literature. We have modi ed and used p_1 to p_6 test problems presented by D az and Fernandez [47] for the single-source capacitated facility location problem [2]. These networks consist of 10 potential facility locations and 20 customer demand points. We have used the given demand values in test instances as the rst moment of the customer demand distribution. The second moment matrix for the SDP formulation was randomly generated.

[2]available at http://www-eio.upc.edu/~elena/sscplp/index.html

4.5.2.1 SDP formulation

We implement the constraint generation and GA algorithms for solving the medium size instances in MATLAB R2012b utilising SeDuMi 1.3 and Cplex LP solvers. GA population size is set to 40 with stopping criteria of a maximum of 30 generations or 10 Stall generations. The iterative method and the solution of problem p_1 is illustrated in Figure 4.7. As is shown in this gure, the algorithm converges in 10 iterations and halts after 15 iterations.

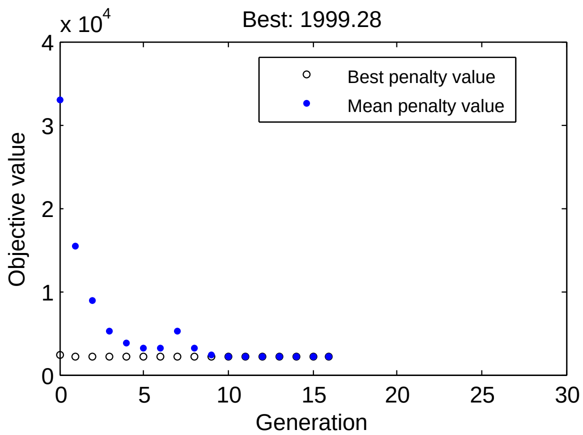


Figure 4.7: CGGA convergence and solution for test p_1

The test results for instances p_1 to p_6 are presented in the Table 4.5. In order to verify the performance of GA, we also ran the CG problem for every combination of $\mathbf{z}$ and found the optimal solution with minimum cost. In all 6 tests the solution obtained from GA matched the optimal solution[3].

Test	Optimal robust decisions										SDP solution	CPU (s)	D-FLP
	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}			
p_1	1	1	0	1	1	0	1	1	1	1	1999.28	4692	1810.85
p_2	1	1	0	1	0	1	1	1	1	1	4380.12	3517	4060.41
p_3	1	1	1	0	1	1	1	0	1	1	6118.17	5076	5744.54
p_4	1	0	1	1	0	1	1	1	0	1	7293.38	5708	6982.16
p_5	1	1	0	0	1	1	1	1	1	0	4603.65	3518	4417.26
p_6	1	1	1	1	1	1	1	0	1	0	2296.99	3069	2172.59

Table 4.5: Robust facility location solutions

[3] Despite the consistent performance of GA in finding the optimal solution in these instances, in theory an optimal solution cannot be guaranteed. However, in practice, the “local” solutions obtained using GA are often of high quality.

For comparison purposes, we also include the deterministic solutions of the instances in the table above. As expected, all of the robust solutions have a higher costs than deterministic facility decisions as a result of installing higher total capacities. Although the lower cost of D-FLP solutions comes at the expense of non- exibility of the facility decisions against the uctuation in future customer demand.

Furthermore, we analyze the robustness of these solutions by implementing the stochastic version of the problems in which the distribution of demand uncertainty is assumed to be known. For this purpose, we have selected the test instance p_5. Let us denote the rst moment and the covariance matrix corresponding to this test instant by $(\boldsymbol{\mu},\ )$. We assume that the probability distribution P is known and given by a multivariate uniform distribution with parameters $(\boldsymbol{\mu},\ )$. The resulting two-stage stochastic problem, given by (4.2.6), is then solved via SAA and the optimal facility location solution Z^*_{STOC} is obtained. We then test the performance of Z^*_{STOC} against the robust facility location solution, denoted by Z^*_{ROB}, by solving problem (4.4.15) for xed $\mathbf{z}$ and sampling $\boldsymbol{\xi}$ from various multivariate distributions with $\boldsymbol{\mu}$ as the rst moment and as the covariance matrix.

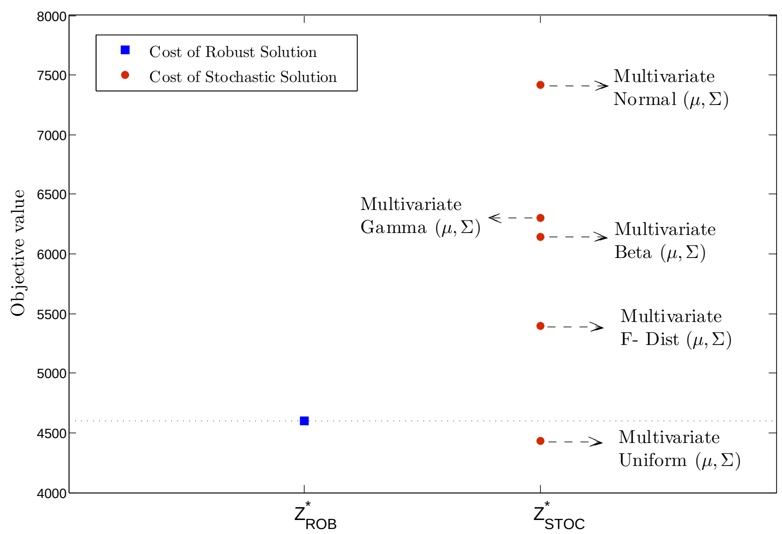


Figure 4.8: Performance of robust solution Z^*_{ROB} against stochastic solution Z^*_{STOC}

This comparison in illustrated in Figure 4.8 and it can be observed that the robust solution is not sensitive to change in probability distribution whilst the performance of stochastic solution can be highly a ected by the choice of the probability distribution of the underlying uncertainty. In other word, without taking robust measures, the stochastic performance might be very poor if the assumed distribution is wrong. This can be observed in Figure 4.8; the stochastic solution outperforms the robust solution if the actual distribution is uniform (as assumed). However, if the actual distribution is not uniform, e.g. in case of the normal distribution, the stochastic solution could result in a much higher cost than the robust solution.

4.5.2.2 SIP formulation

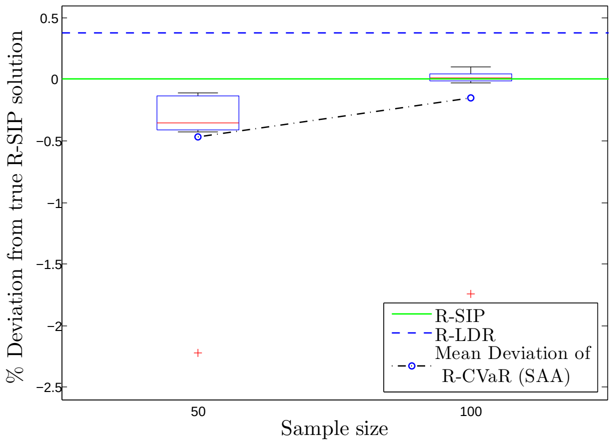
We now reconsider the problem instance p_5 and assume that the only available information on demand uncertainty is the first moment of the distribution $\boldsymbol{\mu}$. The R-SIP framework is then used to construct and solve this problem. As before, we limit the support set of the random variable $\boldsymbol{\xi}$ to 200 uniformly generated random values. For comparison purposes, we summarize the full R-SIP solution to this problem along with those from D-FLP, S-FLP and full R-SDP versions of the problem in the Table 4.6 below.

Model	Available information	Optimal FLP decisions										Total cost	Installed capacity
		z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}		
D-FLP	$\mathbf{d}$	1	1	0	0	0	1	1	1	1	1	4417.26	458
S-FLP	P	1	1	0	1	1	1	1	0	1	0	4495.33	462
R-SDP	$(\boldsymbol{\mu}, \mathbf{Q})$	1	1	0	0	1	1	1	1	1	0	4603.65	481
R-SIP	$\boldsymbol{\mu}$	1	1	0	0	1	1	1	1	1	1	4938.95	511

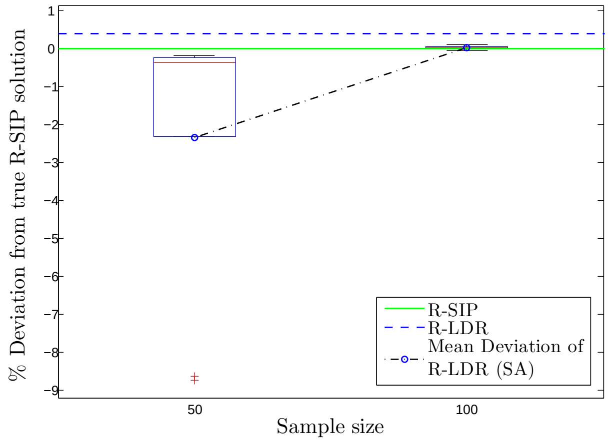
Table 4.6: FLP solutions for various methodologies

It can be observed that R-SIP offers a more flexible solution by installing a higher level of supply capacity. This, of course, comes at the expense of a higher total expected cost. Also, as we include progressively less information on the uncertainty in our models, the solution becomes more and more robust (higher levels of supply capacity installed), which is intuitively sensible.

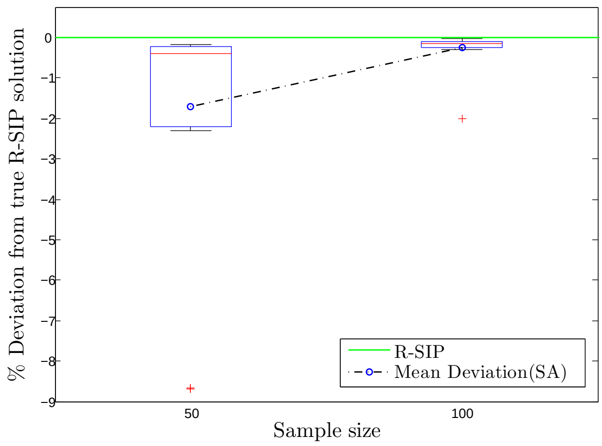
Having solved the full R-SIP p_5, we consider sample approximation of the R-SIP, R-LDR and R-CVaR. Various sample sizes are used for each problem and repeat each test for 100 times. The normalized deviation from the full R-SIP solution is computed and presented in Figure 4.9.



(a) SAA of R-CVaR



(b) SA of R-LDR



(c) SA of R-SIP

Figure 4.9: Deviation of approximation solutions from the true robust solution

It can be observed that in all models, the sampling scheme results in approximation of the true robust solution (R-SIP) with a very low error (with less than 3% mean deviation). Moreover, in the case of smaller sample sizes, the CVaR model preforms much better than the other two models and has the mean deviation of less than 0.5%.

4.6 Conclusion

In this chapter we consider a capacitated facility location problem with customer demand uncertainty and propose a two-stage distributionally robust model for the problem to tackle the issue of incomplete information on the true distribution of the uncertainty. We constructed the uncertainty set using the moments information associated with the distribution of the random demands. Two numerical methods are proposed based on the available moment information. Specifically, we first formulate the robust problem as a semi-infinite program for the case that only the first moment information is given. The semi-infinite program is then solved by approximation using a linear decision rule, CVaR and Monte Carlo sampling. Moreover, we formulate the robust problem as a semi-definite program on the basis of the first and the second moments which is then solved by using a constraint generation algorithm. Finally, we carry out numerical tests for a small instance and also some medium-sized instances taken from the literature. In each case, the distributionally robust solutions offer the flexibility in hedging against uncertainty compared to the deterministic and the stochastic solutions.

In the future, it would be interesting to study the possibility of extending the results and methodologies presented in this chapter to include uncertainty in supply, multi-stage problems and the competition in the market. Also, we would like to explore the problem structure to enhance the solution algorithms for a better performance in large scale instances. Another open direction is to apply the proposed methodologies and numerical schemes to the practical problems with similar structure and characteristics such as uncertain the supply and demand. Some examples related to energy industry could include the wind farm site location problem and liquefied natural gas storage facility location problems.

Chapter 5

Summary and Conclusions

This thesis has presented new stochastic and robust optimisation and equilibrium modelling frameworks as well as solution schemes for decision making problems under uncertainty; those that were motivated by planning and investment problems in the energy industry.

In this chapter, a summary of the results of the work in chapters 2, 3 and 4 is given followed by a discussion on the limitations of the developed models and the potential avenues for further research.

5.1 Medium-term trading strategy of a dominant electricity producer

The second chapter of the thesis investigated the medium-term strategic behaviour of a dominant electricity in a pool-based day-ahead auction. A two-stage bi-level stochastic model was presented to take into account the demand uncertainty and the competition from a number of smaller producers. The pro t and the market share maximization objectives of the dominant producers were also incorporated in the modelling framework. The reformulation of the model resulted in a multi-objective stochastic mathematical program with equilibrium constraints (SMPEC). In order to make this model applicable to large scale systems the SMPEC model was then formulated as a mixed-integer linear programming problem. The applicability of the proposed model was illustrated through a realistic case study based on the Italian power system. By analysing the simulation it was concluded that the dominant producer can substantially increase expected profits and/or expected market share by behaving strategically when offering power production to the ISO. Additionally, the expected profit is higher when bidding in supply function manner than bidding only strategic price or quantity for the same levels of the market share. Moreover, we found out that in peak load demand hours, the dominant producer can exert market power by bidding higher energy prices and withdrawing some of its generation capacity, which leads to higher spot market prices. Conversely, in low demand periods the dominant producer has little influence on market clearing prices and hence the profit is similar to the base case.

5.2 Robust unit commitment problem

In the third chapter of the thesis, a two-stage stochastic and distributionally robust model was developed for the unit commitment problem under the supply uncertainty. The model accounts for the technical constraints such as ramping requirement and unexpected generation outages of up to one generating unit ($n - 1$ security criteria). We considered two robust; the first one is based on the available moment information and the second one consider a mixture distribution on the random supply. The robust model with moment condition was reformulated as a semi-infinite program (SIP-UC) through duality and solved via a sampling scheme. The mixture distribution model (Mix-UC) is also reformulate using duality theory and solved using sample average approximation technique. The proposed modelling frameworks were tested on a medium size case study and the obtained solutions were compared against the two-stage non-robust stochastic formulation (Sto-UC). By analysing the solutions it was concluded that, despite having higher expected cost, both robust solutions offer more flexibility in hedging against the uncertainty than the Sto-UC solution by having a higher levels of reserve schedule in the first stage. Furthermore, the SIP-UC solution offers more flexibility than Mix-UC solution due to lower specificity of assumptions on the distribution of the net load uncertainty in the SIP-UC model.

5.3 Robust facility location problem

The facility location decisions are irreversible and capital intensive and therefore it is vital to take into account the future uncertainties, such as demand, when the facility location decisions are made. In Chapter 4, a two-stage robust model was proposed to include the distributional uncertainty associated with the stochastic demand into long-term facility location decisions. The available moment information were used to construct two formulations for the robust problem. In the first formulation, the first moment conditions were used to formulate the robust problem as a semi-infinite program. In the second formulation, the first and the second moment conditions were exploited to formulate the robust problem as a semi-definite program. The semi-infinite program is solved by approximation using a linear decision rule, CVaR and Monte Carlo Sampling. Moreover, a constraint generation algorithm is developed for solving the semi-definite program. The performance of the proposed models and numerical methods were investigated through some medium size test instances taken from the literature. Based on the numerical results we can conclude that both of the robust formulations result in more conservative but more flexible facility location solutions compared to the traditional deterministic and non-robust stochastic solutions. Moreover, the semi-infinite model solution offered a more flexible solution by installing a higher level of supply capacity. This, of course, resulted in a higher total expected cost. Therefore, it was concluded that including progressively less information on uncertainty in the models, the obtained solutions become more flexible against the risk of demand uncertainty.

5.4 Future research

The model and the solution methods presented in each chapter of this thesis contributed to the existing literature. However, these models contain some limitations. To go beyond these limitations, we suggest the following direction of future research for each model:

In the case of the proposed model for the medium-term trading strategy of dominant producer:

- The proposed multi-objective model can be extended by looking at various other aspects such as risk measures, capacity expansion/investment, production scheduling and maintenance of generation units.

- In our analysis, the demand for electricity is assumed to be inelastic. However, the integration of demand response technologies in the future smart grid could lead to demand elasticity. It would be interesting to investigate the e ect of short-term demand elasticity in the market clearing process.

- The e ect of renewable power on market prices could be introduced in the model.

- As more demand scenarios are incorporated in the stochastic framework, the problem size increases and so does the complexity of solving it. Developing decomposition techniques to exploit the structure of this problem could be explored.

- Finally, modelling competition between producers in the framework of Equilibrium Problem with Equilibrium Constraints (EPEC) would be particularly interesting.

In the case of the robust unit commitment problem :

- The recent advances in development of smart-grid could lead to a more responsive demand. For instance, when the energy prices and the reserve prices in the system tend to increase due to contingencies, customers may consider ramping down their demand to avoid possible blackouts. It would be interesting to study the impact of responsive (elastic) demand on the short-term scheduling decisions.

- In this work, we considered a pool-based market with a single node. The model can be extended to include a network of zones interconnected with transmission lines. Further contingencies such as transmission line failures could also be included.

- Finally, the security criteria could be extended to $(n - K)$ scheme to provide additional reliability in the operation of the system.

Finally, in the robust facility location problem:

- In the proposed model, we assumed that only the demand is stochastic and the other model parameters such as supply and transportation costs were assumed to be deterministic. In reality, such parameters could also be subject to uncertainty and it would be interesting to investigate a possible extension of the robust models to include additional uncertainties.

- Also, we would like to explore the problem structure to enhance the solution algorithms for a better performance in large scale instances.
- Another open direction is to apply the proposed methodologies and numerical schemes to practical problems with similar structure and characteristics such as uncertain supply and demand. Some examples related to the energy industry could include the wind farm site location problem and the lique ed natural gas storage facility location problem.

Appendix A

Appendix

A.1 Expanded formulations for the facility location problem

R-SIP problem (SAA):

$$
\begin{aligned}
\min_{z,x(.),w,\lambda_0,\lambda} \quad & \sum_i b_i z_i + \lambda_0 + \sum_{j=1}^{m} \lambda_j \mu_j \\
\text{s.t.} \quad & \sum_i z_i \le n, \\
& z_i \in \{0,1\},\ \forall\ i \in I, \\
& \sum_{i,j} c_{ij} x_{ij}(\xi^k) + \sum_j c w_j(\xi^k) \le \lambda_0 + \sum_{j=1}^{m} \lambda_j \xi_j^k,\ \forall k \in K, \\
& \sum_i x_{ij}(\xi^k) + w_j(\xi^k) \ge \xi_j^k,\ \forall\ j \in J,\ k \in K, \\
& \sum_j x_{ij}(\xi^k) \le z_i s_i,\ \forall\ i \in I,\ k \in K, \\
& x_{ij}(\xi^k) \ge 0,\ \forall\ i \in I,\ j \in J,\ k \in K, \\
& w_j(\xi^k) \ge 0,\ \forall\ j \in J,\ k \in K.
\end{aligned}
$$

R-LDR problem (SAA):

$$
\begin{aligned}
\min_{z,X,\bar{x},W,\bar{w},\lambda_0,\lambda} \quad & \sum_i b_i z_i + \lambda_0 + \sum_{j=1}^{m} \lambda_j \mu_j, \\
\text{s.t.} \quad & \sum_i z_i \le n \\
& z_i \in \{0,1\}, \ \forall \ i \in I, \\
& \sum_{i,j} c_{ij} \Big[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \Big] + \sum_j c \Big[\sum_{l=1}^{m} W_j^l \xi_l^k + w_j \Big] \le \lambda_0 + \sum_{j=1}^{m} \lambda_j \xi_j^k, \ \forall k \in K, \\
& \sum_i \left[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \right] + \sum_{l=1}^{m} W_j^l \xi_l^k + w_j \ge \xi_j^k, \ \forall j \in J, k \in K, \\
& \sum_i \left[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \right] \le z_i s_i, \ \forall i \in I, k \in K, \\
& \sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \ge 0, \ \forall \ i \in I, \ j \in J, \ k \in K, \\
& \sum_{l=1}^{m} W_j^l \xi_l^k + w_j \ge 0, \ \forall \ j \in J, \ k \in K.
\end{aligned}
$$

R-CVaR problem (SAA):

$$
\begin{aligned}
\min_{\theta,\eta,z,X,\bar{x},W,\bar{w},\lambda_0,\lambda} \quad & \sum_i b_i z_i + \lambda_0 + \sum_{j=1}^{m} \lambda_j \mu_j, \\
\text{s.t.} \quad & \sum_i z_i \le n \\
& z_i \in \{0,1\}, \ \forall \ i \in I, \\
& \boldsymbol{\eta} + \frac{1}{(1-\beta)\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \theta^k \le 0, \\
& \sum_{i,j} c_{ij} \Big[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \Big] + \sum_j c \Big[\sum_{l=1}^{m} W_j^l \xi_l^k + w_j \Big] - \lambda_0 - \sum_{j=1}^{m} \lambda_j \xi_j^k - \boldsymbol{\eta} \le \theta^k, \ \forall k \in K, \\
& \sum_i \left[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \right] + \sum_{l=1}^{m} W_j^l \xi_l^k + w_j \ge \xi_j^k, \ \forall j \in J, k \in K, \\
& \sum_i \left[\sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \right] \le z_i s_i, \ \forall i \in I, k \in K, \\
& \sum_{l=1}^{m} X_{ij}^l \xi_l^k + x_{ij} \ge 0, \ \forall \ i \in I, \ j \in J, \ k \in K, \\
& \sum_{l=1}^{m} W_j^l \xi_l^k + w_j \ge 0, \ \forall \ j \in J, \ k \in K, \\
& \theta^k \ge 0, \ \forall k \in K.
\end{aligned}
$$

A.2 Detailed solutions

	Unit	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
Unit commitment and Generation	1	201.8	317.9	431.3	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455
	2	190.2	281.3	311.3	338.1	353.3	377.3	398.1	431.7	449.4	455	455	455	455	455	411.5	366.6	355.9	365.4	383.8	455	452.5	410.3	357.7	357.7
	3	86.7	0	104.3	168.6	180	180	180	180	180	180	180	180	180	180	180	176.3	165.9	180	180	180	180	145	105	0
	4	101.9	163.9	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	179.9	172.5
	5	0	0	0	0	0	69	107.5	115	115	158	165	165	157.5	136.3	125.6	107.5	94.4	107.5	115	115	107.5	75	0	0
	6	0	0	0	0	37.8	52.8	60	60	60	60	81.8	80	70	60	45	0	0	39.6	60	60	60	45	0	0
	7	0	0	0	0	0	0	0	0	47.7	62.5	62.5	62.5	60	0	0	0	0	0	0	58	60	0	0	0
	8	0	0	0	0	0	0	0	26.2	36.2	40	40	46.9	43.5	40	30	0	0	0	26	40	30	0	0	0
	9	0	0	0	0	0	0	0	0	26.5	40	40	43.6	40	40	30	0	0	0	26	40	30	0	0	0
	10	0	0	0	0	0	0	0	0	26	40	40	42	40	30	0	0	0	0	26.7	40	30	0	0	0
	Total	580.6	763.1	1026.9	1141.7	1206.1	1314.1	1380.6	1447.9	1575.8	1670.5	1699.3	1710	1681	1576.3	1457.1	1285.4	1251.2	1327.5	1452.5	1623	1585	1310.3	1097.6	985.2
Nominal reserve up	1	124.9	101.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	136.5	138.3	120	116.9	101.7	77.7	56.9	23.3	5.6	0	0	0	0	0	43.5	88.4	99.1	89.6	71.2	0	2.5	44.7	97.3	97.3
	3	55	0	3	11.4	0	0	0	0	0	0	0	0	0	0	0	3.7	14.1	0	0	0	0	35	55	0
	4	55	16.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	7.5
	5	0	0	0	0	0	57.5	57.5	50	50	7	0	0	7.5	28.7	39.4	57.5	39.9	57.5	50	50	57.5	57.5	0	0
	6	0	0	0	0	30	30	30	30	30	30	8.2	10	20	30	30	0	0	30	30	30	30	30	0	0
	7	0	0	0	0	0	0	0	0	22.5	22.5	22.5	22.5	22.5	0	0	0	0	0	0	22.5	22.5	0	0	0
	8	0	0	0	0	0	0	0	20	20	20	20	13.1	16.5	20	20	0	0	0	20	20	20	0	0	0
	9	0	0	0	0	0	0	0	0	20	20	20	16.4	20	20	20	0	0	0	20	20	20	0	0	0
	10	0	0	0	0	0	0	0	0	20	20	20	18	20	20	0	0	0	0	13.3	20	20	0	0	0
	Total	371.4	256	123	128.3	131.7	165.2	144.4	123.3	168.1	119.5	90.7	80	106.5	118.7	152.9	149.6	153.1	177.1	204.5	162.5	172.5	167.2	152.4	104.8
Nominal reserve down	1	0	0	15.7	29.5	28.3	36	37.4	29	39.6	35.2	23.3	29.1	27.5	23.7	25.8	17	11.5	22.7	47.5	40.9	33.9	73.8	95.8	65.6
	2	0	0	19.1	11.3	33.1	48.3	72.3	93.1	126.7	128.4	57	0	79.2	139.2	63	19.1	61.6	50.9	60.4	78.8	145.1	63.1	28.2	52.7
	3	0	0	0	0	51.3	55	45.4	21.7	12.7	0	0	0	0	0	55	55	45.5	55	55	55	39.6	0	0	0
	4	0	0	51.7	55	0	3.8	0	0	0	0	0	0	0	0	4.3	55	41.3	51.1	55	49.4	0	55	55	55
	5	0	0	0	0	0	0	0	0	0	0	43	46.6	35.1	0	21.3	36.4	44.4	16.3	7.5	0	0	0	0	0
	6	0	0	0	0	2.5	5.9	7.2	8.9	0	0	0	18.2	16.3	10	15	0	0	2.2	0	0	0	15	0	0
	7	0	0	0	0	0	0	0	0	7.7	22.5	17.5	17.5	20	0	0	0	0	0	0	0	20	0	0	0
	8	0	0	0	0	0	0	0	0	10	0	0	0	6.9	13.1	10	0	0	0	6	0	10	0	0	0
	9	0	0	0	0	0	0	0	0	0	0	0	0	3.6	20	10	0	0	0	6	0	10	0	0	0
	10	0	0	0	0	0	0	0	0	4	0	0	0	16	10	0	0	0	0	6.7	0	10	0	0	0
	Total	0	0	86.5	95.8	115.2	149	162.3	152.7	200.7	186.1	140.8	111.4	204.6	216	204.4	182.5	204.3	198.2	244.1	224.1	268.6	206.9	179	173.3

Table A.1: Sto-UC rst-stage solution: nominal generation, reserve up and down quantities

	Unit	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
Unit commitment and Generation	1	222.1	352.5	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455
	2	206.8	303.4	344.8	384	384.2	386.6	419.5	455	455	455	455	455	455	455	454.3	401.1	396.8	431.5	455	455	455	436.6	368.9	287.8
	3	70	0	102.7	165.5	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	145
	4	104	170.9	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	168.1
	5	0	0	0	0	69	107.5	115.2	116.9	126.2	165	165	165	165	147.2	131.1	115.8	107.5	112.8	115	115	110.6	75	0	0
	6	0	0	0	0	0	40.7	60	61.1	61.1	90	90	90	80	70	60	45	0	40	60	75.8	66.5	45	0	0
	7	0	0	0	0	0	0	0	0	57.5	63.9	85	85	65	60	0	0	0	0	0	59.6	60	0	0	0
	8	0	0	0	24	0	26	39.2	40	40	60	60	60	50	40	30	0	0	26	40	40	40	30	0	0
	9	0	0	0	0	0	0	26.6	40	40	60	60	60	50	40	30	0	0	0	26	40	40	30	0	0
	10	0	0	0	0	0	0	0	27.1	40	40	60	60	50	40	30	0	0	0	25	40	30	0	0	0
	Total	602.9	826.8	1082.5	1208.5	1268.2	1375.8	1475.5	1555.1	1634.8	1748.9	1790	1790	1730	1667.2	1550.4	1376.9	1319.3	1425.3	1536	1640.4	1617.1	1431.6	1183.9	1055.9
Nominal reserve up	1	152.5	102.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	117.5	119.7	110.2	71	70.8	68.4	35.5	0	0	0	0	0	0	0	0.7	53.9	58.2	23.5	0	0	0	18.4	86.1	125
	3	55	0	9	14.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35
	4	55	9.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11.9
	5	0	0	0	0	57.5	57.5	49.8	48.1	38.8	0	0	0	0	17.8	33.9	49.2	57.5	52.2	50	50	54.4	57.5	0	0
	6	0	0	0	0	0	30	30	28.9	28.9	0	0	0	10	20	30	30	0	30	30	14.2	23.5	30	0	0
	7	0	0	0	0	0	0	0	0	22.5	21.1	0	0	20	22.5	0	0	0	0	0	22.5	22.5	0	0	0
	8	0	0	0	20	0	20	20	20	20	0	0	0	10	20	20	0	0	20	20	20	20	20	0	0
	9	0	0	0	0	0	0	20	20	20	0	0	0	10	20	20	0	0	0	20	20	20	20	0	0
	10	0	0	0	0	0	0	0	20	20	20	0	0	10	20	20	0	0	0	20	20	20	0	0	0
	Total	380	231.3	119.2	105.5	128.3	175.9	155.3	137	150.2	41.1	0	0	60	120.3	124.6	133.1	115.7	125.7	140	146.7	160.4	145.9	86.1	171.9
Nominal reserve down	1	0	0	12.5	21	7.3	0.9	3.4	34	12.3	61.4	60.4	12.3	51.8	27.7	32.1	17.5	13.6	31.5	9.3	38.7	37.9	48	14.5	56.4
	2	0	0	43	49.2	73.1	74.8	81.6	73.9	63.7	34.2	0	0	0	50.8	71.4	43	80.3	90.2	122.2	30.6	50	113.3	40.8	43.4
	3	0	0	0	0	0	0	0.9	0	0	0	0	0	0	0	0	36.4	0	0	0	0	0	17.9	55	0
	4	0	0	10.1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55	46.2
	5	0	0	0	0	0	0	12.4	3.7	1.9	11.2	0	0	21.7	14.5	0	34.7	57.5	27.4	4.4	0	0	0	0	0
	6	0	0	0	0	0	0	3.5	18.4	1.1	1.1	0	0	10	0	10	15	0	0	0	0	0	8.4	0	0
	7	0	0	0	0	0	0	0	0	0	12.5	18.9	0	0	20	0	0	0	0	0	0	20	0	0	0
	8	0	0	0	4	0	2	13.8	18.5	0	0	0	0	0	10	10	0	0	6	0	0	0	10	0	0
	9	0	0	0	0	0	0	6.6	13.1	20	0	19.1	0	0	10	10	0	0	0	0	0	18.7	10	0	0
	10	0	0	0	0	0	0	0	0	0	0	0	0	0	11.9	10	0	0	0	0	0	6.7	0	0	0
	Total	0	0	65.6	75.2	80.4	77.7	122.2	161.6	99	120.4	98.4	12.3	83.5	144.9	143.5	146.6	151.4	155.1	135.9	69.3	133.3	207.6	165.3	146

Table A.2: Mix-UC rst-stage solution: nominal generation, reserve up and down quantities

	Unit	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
Unit commitment and Generation	1	219.9	352	427.2	455	453.3	455	455	455	455	455	455	455	455	455	455	455	455	455	455	455	398.1	398.1	398.1	398.1
	2	202	262.7	305	303.3	306.6	378	434.5	415.2	436.7	428	455	455	455	455	445.2	389.4	333.8	305	305	305	305	305	303.6	283.1
	3	105	156.7	153.2	162.5	180	170.9	165	180	180	180	180	180	180	172.5	165	165	164.8	164.8	164.8	180	153	134.8	134.8	134.8
	4	103.8	172.9	180	180	180	180	180	180	180	180	180	180	180	180	165	165	157.5	155.7	155.7	180	145	145	145	145
	5	50	0	0	72.9	115	115	115	115	124.1	124.1	132.4	132.4	123.7	117.8	115	115	115	115	115	115.1	115	107.5	75	0
	6	30	0	0	0	45	60	60	60	60	60	60	70.7	60	60	60	60	60	60	60	60	60	45	0	0
	7	40	0	0	0	0	0	0	45	62.5	62.5	62.5	62.5	62.5	62.5	60	0	0	45	62.5	62.5	62.5	47.6	0	0
	8	20	0	0	0	0	30	40	40	40	40	40	40	40	40	40	28.7	0	30	40	40	40	30	0	0
	9	20	0	0	0	0	30	32	40	40	40	40	40	40	40	40	20	0	30	40	40	40	30	0	0
	10	20	0	0	0	0	0	21.7	40	40	40	40	40	40	40	30	0	0	30	40	40	40	20	0	0
	Total	810.7	944.3	1065.4	1173.7	1279.9	1418.9	1503.2	1570.2	1618.3	1609.6	1644.9	1655.6	1636.2	1622.8	1575.2	1398.1	1286.1	1390.5	1438	1477.6	1358.6	1263	1056.5	961
Nominal reserve up	1	40.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	56.9	56.9	56.9	56.9
	2	23	89.3	122.2	151.7	146.7	77	20.5	39.8	18.3	27	0	0	0	0	9.8	65.6	121.2	150	150	150	150	150	151.4	152.5
	3	0	23.3	26.8	17.5	0	9.1	15	0	0	0	0	0	0	7.5	15	15	15.2	15.2	15.2	0	27	45.2	45.2	45.2
	4	9.7	7.1	0	0	0	0	0	0	0	0	0	0	0	0	15	15	22.5	24.3	24.3	0	35	35	35	35
	5	25	0	0	57.5	15.5	50	50	50	40.9	40.9	32.6	32.6	41.3	47.2	50	50	50	50	50	49.9	50	57.5	52.4	0
	6	15	0	0	0	0	15	30	30	30	30	30	19.3	30	30	30	30	20.6	30	30	30	30	30	0	0
	7	20	0	0	0	0	0	0	15	22.5	22.5	22.5	22.5	22.5	22.5	22.5	0	0	15	22.5	22.5	22.5	17.5	0	0
	8	10	0	0	0	0	0	10	20	20	20	20	20	20	20	20	10.9	0	0	20	20	20	20	0	0
	9	10	0	0	0	0	0	18	12	20	20	20	20	20	20	20	0	0	0	20	20	20	20	0	0
	10	10	0	0	0	0	0	8.3	1.7	20	20	20	20	20	20	20	0	0	0	20	20	20	0	0	0
	Total	163.2	119.7	149	226.7	162.2	151.1	151.8	168.5	171.7	180.4	145.1	134.4	153.8	167.2	202.3	186.5	229.5	284.5	352	312.4	431.4	432.1	340.9	289.6
Nominal reserve down	1	0	0	0	0	0	0	0	3.9	0	0	0	0	0	0	58	93.5	88.9	127	67.2	0	36.2	93.1	93.1	93.1
	2	52	112.7	152.5	152.5	152.5	122.1	152	144.9	110.2	131.7	123	150	150	150	150	140.2	84.4	28.8	0	0	0	7.8	75.8	89
	3	35	55	55	50.1	55	55	55	55	55	21.3	42.1	40.7	52.8	55	55	55	55	54.8	54.8	54.8	16	43	24.8	24.8
	4	1.7	11.7	1.7	0	10.4	55	55	55	36.3	0	0	0	0	47.8	55	55	55	47.5	45.7	35.2	0	35	35	35
	5	0	0	0	22.9	57.5	57.5	57.5	57.5	57.5	9.1	32.3	17.4	17.4	57.5	57.5	57.5	57.5	57.5	57.5	0.3	22.4	57.5	25	0
	6	0	0	0	0	15	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	15	0	0
	7	0	0	0	0	0	0	0	5	22.5	22.5	22.5	22.5	22.5	22.5	20	0	0	5	22.5	22.5	22.5	7.6	0	0
	8	0	0	0	0	0	10	20	20	20	20	20	20	20	20	20	8.7	0	10	20	20	20	10	0	0
	9	0	0	0	0	0	10	12	20	20	20	20	20	20	20	20	0	0	10	20	20	20	10	0	0
	10	0	0	0	0	0	0	1.7	20	20	20	20	20	20	20	10	0	0	10	20	20	20	0	0	0
	Total	88.7	179.4	209.2	225.5	290.4	339.6	383.2	411.3	371.5	274.6	309.9	320.6	332.7	422.8	475.5	439.9	370.8	380.6	337.7	202.8	187.1	279	253.7	241.9

Table A.3: SIP-UC rst-stage solution: nominal generation, reserve up and down quantities

References

[1] H. Aissi, C. Bazgan, and D. Vanderpooten. Min{max and min{max regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research*, 197(2):427{438, 2009.

[2] E. Anderson, H. Xu, and D. Zhang. Con dence levels for cvar risk measures and minimax limits.

[3] E. Anderson, H. Xu, and D. Zhang. CVaR approximations for minimax and robust convex optimization. *Optimisation Online*, 2013.

[4] E. J. Anderson and H. Xu. Supply function equilibrium in electricity spot markets with contracts and price caps. *Journal of Optimization Theory and Applications*, 124(2):257{283, 2005.

[5] J. M. Arroyo and F. D. Galiana. Energy and reserve pricing in security and network-constrained electricity markets. *IEEE Transactions on Power Systems*, 20(2):634{643, 2005.

[6] I. Averbakh and O. Berman. Minimax regret p-center location on a network with demand uncertainty. *Location Science*, 5(4):247{254, 1997.

[7] I. Averbakh and O. Berman. Minmax regret median location on a network under uncertainty. *INFORMS Journal on Computing*, 12(2):104{110, 2000.

[8] R. Baldick and W. W. Hogan. *Capacity constrained supply function equilibrium models of electricity markets: stability, non-decreasing constraints, and function space iterations*. Citeseer, 2001.

[9] M. L. Balinski and F. J. Rispoli. Signature classes of transportation polytopes. *Mathematical programming*, 60(1-3):127{144, 1993.

[10] M. L. Balinski and A. Russako . Faces of dual transportation polyhedra. In *Mathematical Programming at Oberwolfach II*, pages 1{8. Springer, 1984.

[11] W. J. Baumol and P. Wolfe. A warehouse-location problem. *Operations Research*, 6(2):252{263, 1958.

[12] E. M. Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 173{184, 1955.

[13] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton Univ Pr, 2009.

[14] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351{376, 2004.

[15] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769{805, 1998.

[16] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1{13, 1999.

[17] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical programming*, 88(3):411{424, 2000.

[18] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238{252, 1962.

[19] D. Bertsimas and D. B. Brown. Constructing uncertainty sets for robust linear optimization. *Operations research*, 57(6):1483{1495, 2009.

[20] D. Bertsimas, X. V. Doan, K. Natarajan, and C.-P. Teo. Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research*, 35(3):580{602, 2010.

[21] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems*, 28(1):52{63, 2013.

[22] D. Bertsimas, D. Pachamanova, and M. Sim. Robust linear optimization under general norms. *Operations Research Letters*, 32(6):510{516, 2004.

[23] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780{804, 2005.

[24] D. Bertsimas and J. Sethuraman. Moment problems and semidefinite optimization. In *Handbook of semidefinite programming*, pages 469{509. Springer, 2000.

[25] D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical programming*, 98(1-3):49{71, 2003.

[26] D. Bertsimas and M. Sim. The price of robustness. *Operations research*, 52(1):35{53, 2004.

[27] F. Bolle. Supply function equilibria and the danger of tacit collusion: the case of spot markets for electricity. *Energy economics*, 14(2):94{102, 1992.

[28] F. Bouffard and F. D. Galiana. An electricity market with a probabilistic spinning reserve criterion. *IEEE Transactions on Power Systems*, 19(1):300{307, 2004.

[29] F. Bouffard and F. D. Galiana. Stochastic security for operations planning with significant wind power generation. In *Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1{11. IEEE, 2008.

[30] F. Bouffard, F. D. Galiana, and J. M. Arroyo. Umbrella contingencies in security-constrained optimal power flow. In *15th Power Systems Computation Conference, PSCC*, volume 5, 2005.

[31] F. Bouffard, F. D. Galiana, and A. J. Conejo. Market-clearing with stochastic security-part I: formulation. *IEEE Transactions on Power Systems*, 20(4):1818{1826, 2005.

[32] F. Bouffard, F. D. Galiana, and A. J. Conejo. Market-clearing with stochastic security-part II: Case studies. *IEEE Transactions on Power Systems*, 20(4):1827{1835, 2005.

[33] G. Calafiore and M. C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25{46, 2005.

[34] G. C. Cala ore and M. C. Campi. The scenario approach to robust control design. *Automatic Control, IEEE Transactions on*, 51(5):742{753, 2006.

[35] M. C. Campi and S. Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of Optimization Theory and Applications*, 148(2):257{280, 2011.

[36] P. Carpentier, G. Gohen, J.-C. Culioli, and A. Renaud. Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Transactions on Power Systems*, 11(2):1067{1073, 1996.

[37] H. Chao, H. G. Huntington, et al. *Designing competitive electricity markets*. Kluwer Academic Publishers Boston, MA, 1998.

[38] E. Conde. Minmax regret location{allocation problem on a network under uncertainty. *European journal of operational research*, 179(3):1025{1039, 2007.

[39] G. B. Dantzig. Linear programming under uncertainty. *Management science*, 1(3-4):197{206, 1955.

[40] G. B. Dantzig and A. Madansky. On the solution of two-stage linear programs under uncertainty. In *Proceedings of the fourth berkeley symposium on mathematical statistics and probability*, volume 1, pages 165{176. University of California Press, Berkeley, 1961.

[41] M. S. Daskin. *Network and discrete location: models, algorithms, and applications*. John Wiley & Sons, 2011.

[42] D. De Wolf and Y. Smeers. A stochastic version of a stackelberg-nash-cournot equilibrium model. *Management Science*, 43(2):190{197, 1997.

[43] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595{612, 2010.

[44] S. Dempe. *Foundations of bilevel programming*. Springer, 2002.

[45] D. Dentcheva and W. R•misch. *Optimal power generation under uncertainty via stochastic programming*. Springer, 1998.

[46] K. Derinkuyu and M. C. P nar. On the s-procedure and some variants. *Mathematical Methods of Operations Research*, 64(1):55{77, 2006.

[47] J. D az and E. Fernandez. A branch-and-price algorithm for the single source capacitated plant location problem. *Journal of the Operational Research Society*, pages 728{740, 2002.

[48] J. Dupacova. The minimax approach to stochastic programming and an illustrative application. *Stochastics: An International Journal of Probability and Stochastic Processes*, 20(1):73{88, 1987.

[49] J. Dupacova. Stochastic programming: minimax approach stochastic programming: Minimax approach. In *Encyclopedia of Optimization*, pages 3778{3782. Springer, 2009.

[50] J. Dupacova. Uncertainties in minimax stochastic programs. *Optimization*, 60(10-11):1235{1250, 2011.

[51] M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423{432, 2006.

[52] A. Ehrenmann and K. Neuho . A comparison of electricity market designs in networks. *Operations research*, 57(2):274{286, 2009.

[53] L. El-Ghaoui and H. Lebret. Robust solutions to least-square problems to uncertain data matrices. *Sima Journal on Matrix Analysis and Applications*, 18:1035{1064, 1997.

[54] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semide nite programs. *SIAM Journal on Optimization*, 9(1):33{52, 1998.

[55] J. Fortuny-Amat and B. McCarl. A representation and economic interpretation of a two-level programming problem. *Journal of the operational Research Society*, 32(9):783{792, 1981.

[56] J. Goh and M. Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4-part-1):902{917, 2010.

[57] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1{38, 2003.

[58] A. Gourtani, T.-D. Nguyen, D. Pozo, and H. Xu. Robust unit commitment with $n-1$ security criterion. 2014.

[59] A. Gourtani, T.-D. Nguyen, and H. Xu. A distributionally robust optimization approach for two-stage facility location problems. 2014.

[60] A. Gourtani, D. Pozo, M. T. Vespucci, and H. Xu. Medium-term trading strategy of a dominant electricity producer. *Energy Systems*, 5(2):323{347, 2014.

[61] S. Goyal. Improving vam for unbalanced transportation problems. *The Journal of the Operational Research Society*, 35(12):1113{1114, 1984.

[62] R. J. Green and D. M. Newbery. Competition in the british electricity spot market. *Journal of political economy*, pages 929{953, 1992.

[63] R. Henrion and W. R•misch. On m-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling. *Applications of Mathematics*, 52(6):473{494, 2007.

[64] B. F. Hobbs, C. B. Metzler, and J.-S. Pang. Strategic gaming analysis for electric power systems: An mpec approach. *Power Systems, IEEE Transactions on*, 15(2):638{645, 2000.

[65] L. J. Hong, Y. Yang, and L. Zhang. Sequential convex approximations to joint chance constrained programs: A monte carlo approach. *Operations Research*, 59(3):617{630, 2011.

[66] M. D. Ilic, F. Galiana, and L. Fink. *Power systems restructuring: engineering and economics*, volume 448. Springer, 1998.

[67] R. Jiang, J. Wang, and Y. Guan. Robust unit commitment with wind power and pumped storage hydro. *IEEE Transactions on Power Systems*, 27(2):800{810, 2012.

[68] R. Jiang, M. Zhang, G. Li, and Y. Guan. Two-stage network constrained robust unit commitment problem. *European Journal of Operational Research*, 2013.

[69] E. Karangelos and F. Bou ard. Towards full integration of demand-side resources in joint forward energy/reserve electricity markets. *IEEE Transactions on Power Systems*, 27(1):280{289, 2012.

[70] P. D. Klemperer and M. A. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica: Journal of the Econometric Society*, pages 1243{1277, 1989.

[71] P. Kouvelis and G. Yu. *Robust discrete optimization and its applications*, volume 14. Springer, 1997.

[72] D. Kuhn, W. Wiesemann, and A. Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177{209, 2011.

[73] H. J. Landau. *Moments in mathematics*, volume 37. American Mathematical Soc., 1987.

[74] G. Liu, J. Ye, and J. Zhu. Partial exact penalty for mathematical programs with equilibrium constraints. *Set-Valued Analysis*, 16(5-6):785{804, 2008.

[75] Y. Liu, W. Römisch, and H. Xu. Quantitative stability analysis of stochastic generalized equations. *SIAM Journal on Optimization*, 2014.

[76] Y. Liu and H. Xu. Entropic approximation for mathematical programs with robust equilibrium constraints. *SIAM Journal on Optimization*, 2014.

[77] F. Louveaux. Discrete stochastic location models. *Annals of Operations research*, 6(2):21{34, 1986.

[78] F. Meng and H. Xu. Exponential convergence of sample average approximation methods for a class of stochastic mathematical programs with complementarity constraints. *Journal of Computational Mathematics*, 24(6), 2006.

[79] P. B. Mirchandani and R. L. Francis. *Discrete location theory*. 1990.

[80] A. Nemirovski. Lectures on robust convex optimization. *class notes, Georgia Inst. of Technol., Fall*, 2009.

[81] F. Nozicka. *Theorie der linearen parametrischen Optimierung*, volume 24. Akademie-Verlag, 1974.

[82] S. H. Owen and M. S. Daskin. Strategic facility location: A review. *European Journal of Operational Research*, 111(3):423{447, 1998.

[83] N. P. Padhy. Unit commitment { a bibliographical survey. *IEEE Transactions on Power Systems*, 19(2):1196{1205, 2004.

[84] A. Papavasiliou, S. S. Oren, and R. P. O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197{2206, 2011.

[85] D. Peel and G. J. McLachlan. Robust mixture modelling using the t distribution. *Statistics and computing*, 10(4):339{348, 2000.

[86] D. Pozo and J. Contreras. A chance-constrained unit commitment with an $n - K$ security criterion and signi cant wind generation. *IEEE Transactions on Power Systems*, 28(3):2842{2851, 2013.

[87] W. Qianfan, G. Yongpei, and W. Jianhui. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Transactions on Power Systems*, 27(1):206{215, 2012.

[88] J. F. Restrepo and F. D. Galiana. Assessing the yearly impact of wind power through a new hybrid deterministic/stochastic unit commitment. *IEEE Transactions on Power Systems*, 26(1):401{410, 2011.

[89] C. S. ReVelle and H. A. Eiselt. Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165(1):1{19, 2005.

[90] S. M. Robinson. Analysis of sample-path optimization. *Mathematics of Operations Research*, 21(3):513{528, 1996.

[91] R. T. Rockafellar. *Conjugate Duality and Optimization*, volume 16. SIAM, 1974.

[92] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21{42, 2000.

[93] A. Rudkevich. Supply function equilibrium in power markets: learning all the way. *TCA technical paper*, (1299-1702), 1999.

[94] H. Scarf, K. Arrow, and S. Karlin. A min-max solution of an inventory problem. *Studies in the mathematical theory of inventory and production*, 10:201{209, 1958.

[95] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918{936, 2001.

[96] A. Shapiro. On duality theory of conic linear problems. In *Semi-infinite programming*, pages 135{165. Springer, 2001.

[97] A. Shapiro. Monte carlo sampling methods. *Handbooks in operations research and management science*, 10:353{425, 2003.

[98] A. Shapiro and S. Ahmed. On a class of minimax stochastic programs. *SIAM Journal on Optimization*, 14(4):1237{1249, 2004.

[99] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on stochastic programming: modeling and theory*, volume 9. SIAM, 2009.

[100] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic problems. *Optimization Methods and Software*, 17(3):523{542, 2002.

[101] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, pages 111{146. Springer, 2005.

[102] A. Shapiro and H. Xu. Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. *Optimization*, 57(3):395{418, 2008.

[103] G. B. Sheble. *Computational auction mechanisms for restructured power industry operation*, volume 500. springer, 1999.

[104] L. V. Snyder. Facility location under uncertainty: a review. *IIE Transactions*, 38(7):547{564, 2006.

[105] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154{1157, 1973.

[106] R. Sridharan. The capacitated plant location problem. *European Journal of Operational Research*, 87(2):203{213, 1995.

[107] S. Stoft. Power system economics. *Journal of Energy Literature*, 8:94{99, 2002.

[108] A. Street, F. Oliveira, and J. M. Arroyo. Contingency-constrained unit commitment with $n - K$ security criterion: A robust optimization approach. *IEEE Transactions on Power Systems*, 26(3):1581{1590, 2011.

[109] H. Sun and H. Xu. Asymptotic convergence analysis for distributional robust optimization and equilibrium problems. 2013.

[110] H. Sun, H. Xu, and Y. Wang. Asymptotic analysis of sample average approximation for stochastic optimization problems with joint chance constraints via conditional value at risk and di erence of convex functions. *Journal of Optimization Theory and Applications*, pages 1{28, 2012.

[111] T. M. Surowiec. *Explicit stationarity conditions and solution characterization for equilibrium problems with equilibrium constraints*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2010.

[112] S. Takriti, J. R. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497{1508, 1996.

[113] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley. Unit commitment for systems with signi cant wind penetration. *IEEE Transactions on Power Systems*, 24(2):592{601, 2009.

[114] M. T. Vespucci, M. Innorta, and G. Cervigni. A mixed integer linear programming model of a zonal electricity market with a dominant producer. *Energy Economics*, 2012.

[115] D. W. Walkup, R. J.-B. Wets, et al. Lifting projections of convex polyhedra. *Pacific Journal of Mathematics*, 28(2):465{475, 1969.

[116] J. Wang, M. Shahidehpour, and Z. Li. Security-constrained unit commitment with volatile wind power generation. *IEEE Transactions on Power Systems*, 23(3):1319{1327, 2008.

[117] H. Xu. An MPCC approach for stochastic stackelberg{nash{cournot equilibrium. *Optimization*, 54(1):27{57, 2005.

[118] H. Xu, C. Caramanis, and S. Mannor. A distributional interpretation of robust optimization. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 552{556. IEEE, 2010.

[119] H. Xu and D. Zhang. Smooth sample average approximation of stationary points in nonsmooth stochastic optimization and applications. *Mathematical Programming*, 119(2):371{401, 2009.

[120] H. Xu and D. Zhang. Stochastic nash equilibrium problems: sample average approximation and applications. *Computational Optimization and Applications*, 55(3):597{645, 2013.

[121] J. Yao, S. S. Oren, and I. Adler. Two-settlement electricity markets with price caps and cournot generation rms. *european journal of operational research*, 181(3):1279{1296, 2007.

[122] J. Zackova. On minimax solutions of stochastic linear programming problems. *Časopis pro pěstování matematiky*, 91(4):423{430, 1966.

[123] D. Zhang, H. Xu, and Y. Wu. A two stage stochastic equilibrium model for electricity markets with two way contracts. *Mathematical Methods of Operations Research*, 71(1):1{45, 2010.

[124] S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations research*, 57(5):1155{1168, 2009.

[125] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167{198, 2013.