Advances in Electric Power Systems: Robustness, Adaptability, and Fairness

by

Xu Andy Sun

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Author ....
Sloan School of Management
July 8, 2011

Certified by.. ....
Dimitris Bertsimas
Boeing Leaders for Global Operations Professor
Thesis Supervisor

Accepted by ....
Patrick Jaillet
Co-director, Operations Research Center

Advances in Electric Power Systems: Robustness, Adaptability, and Fairness

by

Xu Andy Sun

Submitted to the Sloan School of Management
on July 8, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

The electricity industry has been experiencing fundamental changes over the past decade. Two of the arguably most significant driving forces are the integration of renewable energy resources into the electric power system and the creation of the deregulated electricity markets. Many new challenges arise. In this thesis, we focus on two important ones: How to reliably operate the power system under high penetration of intermittent and uncertain renewable resources and uncertain demand; and how to design an electricity market that considers both efficiency and fairness. We present some new advances in these directions.

In the first part of the thesis, we focus on the first issue in the context of the unit commitment (UC) problem, one of the most critical daily operations of an electric power system. Unit commitment in large scale power systems faces new challenges of increasing uncertainty from both generation and load. We propose an *adaptive robust* model for the security constrained unit commitment problem in the presence of nodal net load uncertainty. We develop a practical solution methodology based on a combination of Benders decomposition type algorithm and outer approximation techniques. We present an extensive numerical study on the real-world large scale power system operated by the ISO New England (ISO-NE). Computational results demonstrate the advantages of the robust model over the traditional reserve adjustment approach in terms of economic efficiency, operational reliability, and robustness to uncertain distributions.

In the second part of the thesis, we are concerned with a geometric characterization of the performance of adaptive robust solutions in a multi-stage stochastic optimization problem. We study the notion of *finite adaptability* in a general setting of multi-stage stochastic and adaptive optimization. We show a significant role that geometric properties of uncertainty sets, such as symmetry, play in determining the power of robust and finitely adaptable solutions. We show that a class of finitely adaptable solutions is a good approximation for both the multi-stage stochastic as well as the adaptive optimization problem. To the best of our knowledge, these are the first approximation results for multi-stage problems in such generality. Moreover, the results and the proof techniques are quite general and extend to include important constraints such as integrality and linear conic constraints.

In the third part of the thesis, we focus on how to design an auction and pricing scheme for the day-ahead electricity market that achieves both economic efficiency and fairness. The work is motivated by two outstanding problems in the current practice — the uplift problem and equitable selection problem. The uplift problem is that the electricity payment determined by the electricity price cannot fully recover the production cost (especially the fixed cost) of some committed generators, and therefore the ISOs make side payments to such generators to make up the loss. The equitable selection problem is how to achieve fairness and integrity of the day-ahead auction in choosing from multiple (near) optimal solutions. We offer a new perspective and propose a family of *fairness* based auction and pricing schemes that resolve these two problems. We present numerical test result using ISO-NE's day-ahead market data. The proposed auction- pricing schemes produce a frontier plot of efficiency versus fairness, which can be used as a vaulable decision tool for the system operation.

Thesis Supervisor: Dimitris Bertsimas
Title: Boeing Leaders for Global Operations Professor

Acknowledgments

I am very grateful to my thesis supervisor Professor Dimitris Bertsimas for his motivation, guidance, encouragement, and high research and ethical standards. It is a great fortune that I have the privilege to work with him, and know him not only as an advisor but also as a friend. He is such an extraordinary man in so many aspects that he has been a constant source of inspiration to me. This experience has forever changed my perspective on research, and will have a long-lasting influence on my career.

I would also like to thank my committee members Professor Robert Freund, Professor Georgia Perakis, and Professor Vivek F. Farias for their support through the years. I am deeply indebted to Professor Robert Freund for introducing me to the field of operations research, his mentoring during my initial years at the ORC, and his unconditional support throughout my Ph.D. program. One of the key ideas of Chapter 3 in this thesis came out of my understanding of the concept of symmetry, which I learned primarily from Rob and Alex Belloni's earlier work.

The direction of my research was also greatly influenced by my interaction with ISO New England Inc., especially the Department of Business Architecture & Technology led by Eugene Litvinov. I am greatly indebted to Eugene's warm welcome of an 'outsider' into the field of electric power systems, his opennese and vision for research collaboration, and his generous support. The two summers that I spent in his group, first as a visitor, then as an intern, were pleasant and formative for my understanding of the field. It has been very enjoyable and productive to work with Tongxin Zheng and Jinye Zhao from the group. And, none of these would be possible without the generous financial support from the ISO for the past year.

I would like to thank Professor Ignacio Pérez Arriaga for teaching the course on engineering, economics, and regulation of the electric power sector, and for organizing a wonderful field trip to the ISO New England, from which I first met Eugene's group. I would also like to thank Professor Marija D. Ilić for the insightful discussion and financial support for me to attend the Seventh Annual Carnegie Mellon Conference on the Electricity Industry.

The ORC has been such a warm, congenial, and cozy place that I am sure I will miss. I would like to thank Professors Dimitris Bertsimas, Cindy Barnhart, and Patrick Jaillet for serving as co-directors during my time at the center. I would also like to thank the ORC staff, Paulette Mosley, Laura Rose, and Andrew Carvalho, and Margaret Lo from the Sloan School Dean's Office. The greatest about the ORC is the people. Countless late night discussions, chats, complaints, encouragement, and laughs made up the unforgettable memory.

Above all, no words can fully express my thankfulness to my parents. I am extremely grateful for their love, care, and faith in me. To them this thesis is dedicated.

Finally, I would like to acknowledge that this research was supported in part by a grant from ISO New England Inc..

Contents

1 Introduction **15**

1.1 Reliable Operation of the Electric Power Systems Under Supply and Demand Uncertainty . . . 15

1.2 Fairness and Efficiency in Electric Market Design . . . 16

1.3 Exciting Development in Operations Research: Robust and Adaptive Optimization . . . 18

1.4 The Structure and Contributions of the Thesis . . . 19

2 Adaptive Robust Optimization for Security Constrained Unit Commitment Problems **25**

2.1 Introduction . . . 25

2.2 The Deterministic SCUC Problem . . . 29

2.3 The Two-stage Fully Adaptive Robust SCUC Formulation . . . 34

2.3.1 The Uncertainty Set of Net Load . . . 34

2.3.2 The Adaptive Robust SCUC Formulation . . . 35

2.4 Solution Method to Solve the Adaptive Robust Model . . . 37

2.4.1 The Outer Level: Benders Decomposition Algorithm . . . 37

2.4.2 The Inner Level: Solve $R(\boldsymbol{x})$ . . . 38

2.5 Computational Experiments . . . 40

2.5.1 Cost Efficiency and the Choice of the Budget Level . . . 43

2.5.2 Reliability of Dispatch Operation . . . 44

2.5.3 Robustness Against Load Distributions . . . 46

2.6 Conclusion . . . 48

3 A Geometric Characterization of the Power of Finite Adaptability in Multistage Stochastic and Adaptive Optimization **51**

3.1 Introduction 51

3.2 Models and Preliminaries 55

3.2.1 Two-stage Optimization Models 56

3.2.2 Multi-stage Optimization Models 58

3.2.3 Examples 62

3.2.4 Finitely Adaptable Solutions 64

3.2.5 The Symmetry of a Convex Set 66

3.2.6 The Translation Factor $\rho(\boldsymbol{u}, \mathcal{U})$ 69

3.3 Our Contributions 71

3.4 Two-stage Stochastic Optimization Problem 75

3.4.1 Tightness of the Bound 77

3.4.2 An Alternative Bound 78

3.4.3 Examples: Stochasticity Gap for Specific Uncertainty Sets . . . 80

3.5 Multi-stage Stochastic Problem under RHS Uncertainty 83

3.5.1 Algorithm 84

3.5.2 Proof of Theorem 3 86

3.6 Multi-stage Adaptive Problem 89

3.6.1 Right-hand Side Uncertainty: Π^K_{Adapt} 89

3.6.2 RHS and Cost Uncertainty 90

3.6.3 An Alternative Bound 92

3.7 Extension to General Convex Cones 95

3.7.1 Stochastic Problem with Linear Conic Constraints 95

3.8 Extensions to Integer Variables 96

3.8.1 Multi-stage Stochastic Problem 97

3.8.2 Multi-stage Adaptive Optimization Problem 97

3.9 Conclusion 98

3.10 Appendix 99

3.10.1 Examples: Symmetry of Specific Sets 99

3.10.2 Proof of Theorem 6. 109

4 A Fairness-based Proposal for Electricity Market Design 113

4.1 Introduction 113

4.2 The current day-ahead market auction and pricing scheme 116

4.3 Max-Min Fairness Based Electricity Market Design 120

4.3.1 Fairness Concepts 120

4.3.2 Max-Min Fairness 121

4.3.3 Max-min Fairness Auction and Pricing Scheme 123

4.3.4 Properties of the MMF Scheme 125

4.4 β-Fairness 131

4.5 Efficient Fairness Pricing Scheme 137

4.6 The Efficiency-Fairness Tradeoff Curve: An Example Using Real-World Data 140

4.7 Conclusion 143

4.8 Appendix: Max-Min Fairness 144

5 Conclusion and Future Research 149

List of Figures

2-1 Flow chart of the proposed two-level algorithm. 40

2-2 The average dispatch costs of the AdptRob approach for normally and uniformly distributed load. 47

2-3 The average dispatch costs of the ResAdj approach for normally and uniformly distributed load. 48

3-1 Illustration of the evolution of uncertainty in a multi-stage problem. . 59

3-2 A finitely adaptable solution for a 4-stage problem. For each uncertainty set, we specify the set of corresponding solutions in the finitely adaptable solution policy. 66

3-3 Geometric illustration of $\mathbf{sym}(\boldsymbol{u}, \mathcal{U})$. 68

3-4 The figure on the left is a symmetric polytope with symmetry 1. The middle figure illustrates a standard simplex in $\mathbb{R}^m$ with symmetry $1/m$. The right figure shows the intersection of a Euclidean ball with $\mathbb{R}^m_+$, which has symmetry $1/\sqrt{m}$. 69

3-5 Geometry of the translation factor. 70

3-6 A geometric perspective on the stochasticity gap. 79

3-7 A geometric perspective on the adaptability gap. 94

4-1 Efficiency-fairness tradeoff curve. Let u be the utility vector of a solution and $\bar{u}$ be its mean. Four inequity measures are defined as variance: $\sum_i (u_i - \bar{u})^2$; sum of absolute deviations: $\sum_i \sum_j |u_i - u_j|$; reversed Jain's index: $\frac{\sum_i u_i^2}{n^2 \bar{u}^2}$; Gini index: $\frac{\sum_i \sum_j |u_i - u_j|}{2n^2 \bar{u}}$. In each plot, from the highest efficiency are the solutions of the MaxSW-MP scheme, the limiting pricing scheme, and the tradeoff scheme with $\beta = 0.3, 0.5, 0.7, 0.9$. 142

List of Tables

2.1 The average dispatch costs and total costs of the AdptRob and ResAdj for normally distributed load for $\Delta^t/N_d = 0, 0.1, \ldots, 1$. 44

2.2 The average dispatch costs and total costs of the AdptRob and ResAdj for normally distributed load $\Delta^t/\sqrt{N_d} = 0.5, 1, \ldots, 3$. 45

2.3 Standard deviation of the dispatch costs of the two approaches and their ratio for normally distributed load. 45

2.4 Penalty costs of AdptRob and ResAdj approaches for normally distributed load. 46

2.5 Standard deviation of dispatch costs for uniformly distributed load and the relative difference with those of the normally distributed load. . . 48

3.1 Symmetry and corresponding stochasticity gap for various uncertainty sets. Note that footnote (a) or (b) means the tight bound is from the symmetry bound (3.13) or the scaling bound (3.17). 111

4.1 β-fairness tradeoff schemes for the simple example. For each β, the first four utilities are the utilities of generators; the last one is the consumer utility (since both consumers have the same utility, only one is shown). 135

Chapter 1

Introduction

1.1 Reliable Operation of the Electric Power Systems Under Supply and Demand Uncertainty

Over the past decade, significant new development has taken place in the electricity industry around the world. One of the most important driving forces is the move toward alternative energy resources. Indeed, over the past decade, tremendous amount of effort and resources have been invested in developing renewable energy technologies and installing renewable energy generating plants. Many countries in the world have established alternative energy mandate. For instance, 24 states in the US have renewable portfolio standards, and the European Union has a mandate of 20% electricity generated from renewable energy resources by the year 2020. To keep the development of renewable energy sustainable and to realize its full potential, new challenges to the operations of electric power systems have to be solved. In particular, the following issue becomes critical:

- How to reliably operate the power system under high penetration of intermittent and uncertain renewable resources as well as uncertain demand.

To take full advantage of the renewable energy resources, the bulk power system needs to be operated in a reliable, resilient, and robust way. This need is particularly felt in the day-ahead power system scheduling problem, i.e., the so-called security constrained unit commitment (SCUC) problem. The SCUC is a key daily operation of any large scale electric power system in both vertically integrated industry and deregulated electricity markets. It has been one of the most challenging problems in the power system operation, due to its large scale, discrete nature, and complex constraints. Since the unit commitment decisions have to be made a day before the real-time operation, the uncertainty in the generation and demand forecast is a significant factor that complicates the planning, especially with the fast growing renewable energy penetration and demand-side price sensitive response.

The current operational practice heavily relies on deterministic decision models and heuristic, inflexible reserve requirement [45, 85], which makes it difficult and economically inefficient to adjust to the emerging system of growing uncertainty in both supply and demand. Most of the existing proposals that explicitly deal with uncertainty use the stochastic optimization approach, which requires a probability distribution of the uncertainty, and often relies on sampling discrete scenarios of the uncertainty realizations [27, 82, 68, 87, 86]. This approach has practical limitations in the application to large scale power systems. In particular, it may be difficult to identify an accurate probability distribution of the uncertainty. Also, stochastic UC solutions only provide probabilistic guarantees to the system reliability. It often requires large number of scenario samples to obtain a reasonable reliability guarantee, which results in intensive computation. The limitations of the current deterministic approach and the stochastic optimization model calls for a new methodology for unit commitment operations.

1.2 Fairness and Efficiency in Electric Market Design

Another important driving force of the significant development in the electricity industry is the creation of deregulated electricity markets. Since the 1990's, the electricity industry has undergone deregulation and restructuring in several countries such as UK, Spain, Chile, Brazil, among others, and several regions of the United States. Vertically integrated utility companies are restructured; electricity production and transmission are separated into different enterprises; Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs) are established to manage the regional transmission grids and wholesale electricity markets. Today, ten ISOs and RTOs serve two-thirds of the electricity consumers in the United States and more than half of Canada's population. Although more than two decades have passed, some fundamental issues in electricity market design still remain in active debate, attracting considerable attention from both academia and industry [79].

The day-ahead (DA) market design is one of the most difficult among electricity market design problems. In the current reality of deregulated electricity markets, the Independent System Operators (ISOs) adopt the social welfare maximization principle in solving a unit commitment problem and a marginal cost pricing scheme to determine the electricity price as the incremental cost of producing one more unit of electricity. A long noticed shortcoming of this approach is the so-called uplift problem, i.e., the electricity payment determined by the electricity price cannot fully recover the production cost (especially the fixed cost) of some committed generators, and therefore the ISOs make side payments to such generators to make up the loss.

Recent proposals have focused on augmenting the current social welfare maximization auction and marginal cost pricing scheme by imposing prices for integer decisions [67, 88], or directly using the Lagrangian multipliers of the mixed integer unit commitment problem to form the electricity price [50, 44]. One drawback of the first approach is that the introduced energy prices are sometimes hard to interpret and implement in practice, in a sense, they are similar to the uplift payment in nature. The second approach minimizes the total uplift payment, however, its definition of uplift is different from the one used in current practice, also it does not fully resolve the problem – uplift payment can still be positive. Computationally, it requires solving a Lagrangian relaxation of the mixed-integer unit commitment problem, and potentially a primal version of the unit commitment as well.

Another long standing issue in the current practice is how to choose one market clearing solution from multiple optimal or near-optimal solutions [53, 48]. We call this equitable selection problem. Although all solution candidates have the same level of system welfare, selecting one solution against others can have significant financial implications on individual generators and consumers. Current practice lacks clear-cut criteria in choosing from potentially numerous solution candidates. In fact, even computing all the solution candidates is a highly nontrivial task.

This thesis aims to provide new perspective and proposals to these issues. In our consideration, both problems are symptoms of a deeper unfairness issue: for the uplift problem, some generators that produce electricity make no profit, while other market participants, especially on the consumer side, collect considerable utility surpluses; for the second problem, choosing one solution from others can clearly cause fairness and inequity concerns among market participants. Thus, the following question seems to bear significant implication on the electricity market design problem:

- How to design an auction and pricing scheme for the day-ahead market that achieves both economic efficiency and fairness.

1.3 Exciting Development in Operations Research: Robust and Adaptive Optimization

As the electricity industry is witnessing exciting development, the operations research community has also been making significant progress, especially in the area of multistage decision making under uncertainty. Multistage optimization problems under uncertainty are prevalent in numerous scientific and engineering disciplines (including electric power systems) and have attracted researchers from diverse backgrounds. Several solution approaches have been proposed including exact and approximate dynamic programming, stochastic programming, and simulation-based methods. As early as 1970s, Soyster [80] and Falk [36] considered static linear programming problems with uncertain matrix coefficients, right-hand side, and cost coefficients. The goal was to find a solution that is feasible for all realizations of uncertainty to minimize the worst case cost. This new solution concept remained almost unnoticed until the late 1990's, when a series of papers by Ben-tal and Nemirovski [11, 12, 13], Ben-tal et. al. [1], El Ghaoui and Lebret [40], El Ghaoui et. al. [41], Bertsimas and Sim [21, 22], and Bertsimas et. al. [30] considerably extended the previous work in linear programs to much more general convex optimization problems. Robust Optimization has since become an active research field. Later, the framework is further extended by a sequence of papers by Ben-tal et. al. [10, 7, 8], in which the decision is allowed to be made after the uncertainty was revealed, thus introducing adaptable policies and Adaptive Optimization. It was shown that an arbitrary adaptable policy typically results in intractable problems, and proposed a special class of affine decision rules. In the work by Bertsimas and Caramanis [16], a hierarchy of finitely adaptable policies with increasing adaptability was proposed.

As is shown, it is an exciting time to conduct interdisciplinary research on the boundary of electric power systems and operations research. Indeed, this thesis presents some advances in the above two important directions for power systems operation and electricity market design, using the recent development in Robust and Adaptive Optimization. In particular, we develop a new methodology for the security constrained unit commitment operation that is both *robust* and *adaptable* to the uncertain supply and demand. We propose a new design for the day-ahead electricity market that explicitly considers both efficiency and *fairness*. We also present a contribution to the theory of multi-stage stochastic and adaptive optimization, where we found a tight geometric characterization of the power of finitely adaptable solutions for mutli-stage stochastic optimization problems.

1.4 The Structure and Contributions of the Thesis

The structure and the main contributions of the thesis are summarized below.

- In Chapter 2 of the thesis, we focus our attention on the security constrained unit commitment problem.

1. We formulate a two-stage adaptive robust optimization model for the security constrained unit commitment (SCUC) problem with nodal net load uncertainty. It is a first such SCUC model using adaptive robust optimization and incorporating all types of constraints considered in the real world, such as network, contingency, reserve, ramp-rate, and all other physical, inter-temporal constraints. The first-stage commitment decision is robust and the second-stage dispatch solution is fully adaptive to all variations of uncertainty. This endorses the solution with both robustness and high adaptability. Special techniques are taken to further reduce the conservativeness of the solution by controlling the size of the uncertainty region and choosing proper budget of uncertainty levels using the probability law.

2. Due to the complexity and novelty of the adaptive robust model, efficient algorithms are in great need. We develop a practical two-level solution algorithm. The algorithm combines a cutting plane type method with heurstic speed-up in the outer level with an efficient approximate algorithm in the inner level. This algorithm can be viewed as a further generalization to the traditional Generalized Benders Decomposition framework, where in our case the objective function is given implicitly by a nontrivial bilinear optimization problem. Our algorithm can be useful to many other problems, where the two-stage adaptive robust optimization model is applicable, such as capacity expansion, resource planning, and airline management.

3. We conduct extensive numerical tests on the large scale power system operated by ISO New England with all detailed constraints used in reality (reserve, transmission, ramping rate, constraint violation penalty, etc). Based on the numerical results, we compare the performance of the adaptive robust UC model to the current practice in three aspects, namely economic efficiency, real-time operation reliability, and robustness to probability distributions of the uncertainty. The numerical test demonstrates the clear advantages of the adaptive robust models in all three aspects.

- Chapter 3 of the thesis is of a theoretical nature, but was originally motivated by the above adaptive robust SCUC model.

  1. We discover a significant role that geometric quantities, such as the symmetry of the uncertainty set, play in characterizing the performance of the finite adaptability solutions in multistage stochastic optimization and adaptive optimization. In particular, in both two-stage and multi-stage models, the optimal objective value of a finitely adaptive solution is upper bounded by $(1 + \rho/s)$ times the optimal value of a multistage stochastic optimization solution, where s is the symmetry of the uncertainty set and $\rho \in [0, 1]$ relates how far the set is removed from the origin. The significance of the result is that it shows the finitely adaptive solution is a good approximation of a multistage stochastic solution, e.g., when the uncertainty set is symmetric ($s = 1$, we have a 2-approximation). While multistage stochastic optimization is essentially intractable, finitely adaptable optimization is computationally tractable.

  2. The above geometric result opens the way to characterize performance for many structured uncertainty set. We investigate and give many examples, such as budgeted uncertainty set, symmetric sets, uncertainty set motivated by Central Limit Theorem, etc. This provides a useful guideline to the modelers in choosing and designing proper uncertainty sets.

  3. We extend our results to general linear conic inequalities, thus include the important case of multistage optimization with positive semidefinite (SDP) constraints. We also extend our results to problems with integral decisions in the first-stage, which include capacity planning type problems.

- The above two parts correspond to the *robustness* and *adaptability* in the thesis title. Chapter 4 of the thesis deals with the design of an *efficient* and *fair* day-ahead (DA) electricity market.

  1. We propose and investigate the notion of β-fairness that encompasses a

full spectrum of fairness proposals. The case $\beta = 0$ corresponds to current practice of maximizing social welfare, whereas $\beta = 1$ corresponds to a solution that maximizes the minimum utility among market participants, the so-called max-min fairness. Under the max-min fairness scheme, every market participant is guaranteed a nonnegative surplus, so no side payments are needed, and thus the uplift problem is naturally solved. We broadly investigate various properties, computational methods, and economic implications of the max-min fairness scheme.

2. By varying β, we investigate the tradeoff between efficiency and fairness of the β-fairness auction and pricing scheme. We formulate a family of mixed integer optimization models that explicitly parametrizes this tradeoff. The solutions for different parameter values are used to generate an efficiency-fairness curve. Such a frontier plot illustrates a flexible tradeoff between efficiency and fairness achieved by our proposal, and thus, it can be used to facilitate the ISO's decision-making in choosing an appropriate operating point of its market. A particular revealing observation from the efficiency-fairness curve is that the current operational practice ($\beta = 0$) is not Pareto efficient, and significant improvement can be achieved by our proposed schemes.

3. We demonstrate that there is a $\beta_0 \in (0, 1)$, at which the social welfare is as high as current practice, while the fairness property is strictly better and the side payments are strictly lower than current practice. This shows that, in accordance with well formulated fairness principle, the β-fairness tradeoff scheme with $\beta = \beta_0$ automatically selects the most fair solution from potentially numerous maximum social welfare solution candidates. This provides a solution to the second issue of the current practice to achieving fairness and integrity in choosing from multiple optimal social welfare solutions.

4. We investigate the convergent behavior of β-fairness as $\beta \to 0$, and fur-

ther propose a new pricing scheme called efficient fairness (EF) scheme, which decouples from the MaxSW auction and simultaneously achieves the maximum social welfare and max-min fairness. It also enjoys easier computation (only solves one simple linear program). We also formally show that among all pricing schemes, the EF pricing scheme is the most likely to eliminate uplift payment. When uplift is unavoidable under any pricing scheme, the EF scheme minimizes the largest uplift payment.

Chapter 2

Adaptive Robust Optimization for Security Constrained Unit Commitment Problems

2.1 Introduction

Unit commitment (UC) is one of the most critical decision processes performed by system operators in deregulated electricity markets as well as in vertically integrated utilities. The objective of the UC problem is to find a unit commitment schedule that minimizes the commitment and dispatch costs of meeting the forecasted system load, taking into account various physical, inter-temporal constraints for generating resources, transmission, and system reliability requirements.

During the normal real-time operation, system operator dispatches the committed generation resources to satisfy the actual demand and reliability requirements. In the event that the actual system condition significantly deviates from the expected condition, system operator needs to take corrective actions such as committing expensive fast-start generators, voltage reduction, or load shedding in emergency situation to maintain system security. The main causes of the unexpected events come from the uncertainties associated with the load forecast error, changes of system interchange schedules, generator's failure to follow dispatch signals, and unexpected transmission and generation outages.

In recent years, higher penetration of variable generation resources (such as wind power, solar power, and distributed generators) and more price-responsive demand participation have posed new challenges to the unit commitment process, especially in the independent system operator (ISO) managed electricity markets. It becomes important for the ISOs to have an effective methodology that produces robust unit commitment decisions and ensures the system reliability in the presence of increasing real-time uncertainty.

Previous studies of uncertainty management in the UC problem can be divided into two groups. The first group commits and dispatches generating resources to meet a deterministic forecasted load, and handles uncertainty by imposing conservative reserve requirements. The second group relies on stochastic optimization techniques. The first group, so-called the reserve adjustment method, is widely used in today's power industry. Much of research along this vein, including [2, 24, 23, 43], has focused on analyzing the levels of reserve requirements based on deterministic criteria, such as a loss of the largest generator or system import change. Such an approach is easy to implement in practice. However, committing extra generation resources as reserves could be an economically inefficient way to handle uncertainty, especially when the reserve requirement is determined largely by some ad-hoc rules, rather than a systematic analysis. Also, since the UC decision only considers the expected operating condition, even with enough reserve available, the power system may still suffer capacity inadequacy when the real-time condition, such as load, deviates significantly from the expected value.

The stochastic optimization approach explicitly incorporates a probability distribution of the uncertainty, and often relies on presampling discrete scenarios of the uncertainty realizations [27, 82, 68, 87, 86]. This approach has some practical limitations in the application to large scale power systems. First, it may be difficult to identify an accurate probability distribution of the uncertainty. Second, stochastic UC solutions only provide probabilistic guarantees to the system reliability. To obtain a reasonably high guarantee requires large number of scenario samples, which results in a problem that is computationally intensive. To improve the robustness of stochastic UC solutions, Ruiz et al. [74] proposed a hybrid approach combining the reserve requirement and stochastic optimization methods. A recent work [35] proposed an interesting framework that combines uncertainty quantification with stochastic optimization. This framework could also be integrated with the robust optimization formulation proposed in this chapter.

Robust optimization has recently gained substantial popularity as a modeling framework for optimization under parameter uncertainty, led by the work [11, 12, 13, 1, 40, 41, 21, 22, 30]. The approach is attractive in several aspects. First, it only requires moderate information of the underlying uncertainty, such as the mean and the range of the uncertain data; and the framework is flexible enough that the modeler can always incorporate more probabilistic information such as covariance to the uncertainty model, when such information is available. Second, the robust model constructs an optimal solution that immunizes against all realizations of the uncertain data within a deterministic uncertainty set. Such a robustness is a desirable feature, especially when the penalty associated with infeasible solutions is very high, as the case in the power system operations. Hence, the concept of robust optimization is consistent with the risk-averse fashion in which the power systems are operated. There has been a broad variety of applications of robust optimization in engineering and management sciences (such as structural design, integrated circuit design, statistics, inventory management, to name a few. See [9] and references therein).

In this chapter, we propose a two-stage adaptive robust optimization model for the security constrained unit commitment (SCUC) problem, where the first-stage UC decision and the second-stage dispatch decision are robust against all uncertain nodal net load realizations. Furthermore, the second-stage dispatch solution has full adaptability to the uncertainty. The critical constrains such as network constraints, ramp rate constraints and transmission security constraints are incorporated into the proposed model as well. It is key to design a proper uncertainty set to control the conservatism of the robust solution. We use a special technique proposed in

[21, 22] for this purpose. We develop a practical solution method, and extensively test the method on a real-world power system. Papers [70, 89] proposed similar robust optimization UC models. However, their proposals ignored reserve constraints and did not study critical issues such as the impact of robust solutions on system efficiency, operational stability, and robustness of the UC solutions against different probability distributions of uncertainty, neither did they compare the robust UC model with the current practice in a real world power system. Our research was conducted and written independently of the works in [70, 89]. The main contributions of our work are summarized below.

1. We formulate a two-stage adaptive robust optimization model for the SCUC problem. Given a pre-specified nodal net load uncertainty set, the two-stage adaptive robust UC model obtains an "immunized against uncertainty" first-stage commitment decision and a second-stage adaptive dispatch actions by minimizing the sum of the unit commitment cost and the dispatch cost under the worst-case realization of uncertain nodal net load. The nodal net load uncertainty set models variable resources such as non-dispatchable wind generation, real-time demand variation, and interchange uncertainty. The parameters in the uncertainty set provide control over the conservatism of the robust solution.

2. We develop a practical solution methodology to solve the adaptive robust model. In particular, we design a two-level decomposition approach. A Benders decomposition type algorithm is employed to decompose the overall problem into a master problem involving the first-stage commitment decisions at the outer level and a bilinear subproblem associated with the second-stage dispatch actions at the inner level, which is solved by the outer approximation approach [32, 37]. The proposed solution method applies to general polyhedral uncertainty sets. We present a computational study that shows the efficiency of the method. Besides the unit commitment studied in this chapter, the proposed mixed-integer adaptive robust model is applicable to many two-stage decision making problems, such as robust inventory management problems, robust loca-

tion transportation problems, etc. The efficient solution methodology developed here, therefore, may be of interest to a much wider community.

3. We conduct extensive numerical experiments on the real-world large scale power system operated by the ISO New England. We study the performance of the adaptive robust model and provide detailed comparison with the current practice, the reserve adjustment approach. Specifically, we analyze the merit of the adaptive robust model from three aspects: economic efficiency, contribution to real-time operation reliability, and robustness to probability distributions of the uncertainty. The extensive computational results show that the adaptive robust UC model outperforms the current practice in all three aspects. This clearly demonstrates the advantages of the adaptive robust UC model.

This chapter is organized as follows. Section 2.2 describes the deterministic SCUC formulation. Section 2.3 introduces the two-stage adaptive robust SCUC formulation. Section 2.4 discusses the solution methodology. Section 2.5 presents computational results, including a discussion on the proper way to choose the level of conservatism in the robust model. Section 2.6 concludes with discussion.

2.2 The Deterministic SCUC Problem

For the unit commitment decision making, the system operator usually has access to a wide range of detailed data listed below, including economic data of generator's production costs or supply curve in a market setting, physical characteristics of each generator, expected load forecast, system reserve requirement, network parameters and transmission line ratings.

- N_g, N_d, N_b, N_l, T: The number of generators, loads, nodes, transmission lines, and time periods (in hours).
- $\mathcal{N}_g, \mathcal{N}_d, \mathcal{N}_b, \mathcal{N}_l, \mathcal{T}$: The corresponding sets of generators, loads, nodes, transmission lines, and time periods (in hours).
- S_i^t, G_i^t, F_i^t: Start-up, shut-down, no-load costs of generator i at time t.

- $C_i^t(\cdot)$: Variable cost of generator i at time t as a function of production levels.
- $p_i^{\max}, p_i^{\min}$: Maximum and minimum production levels of generator i (usually called Ecomax and Ecomin, respectively).
- RU_i^t, RD_i^t: Ramp-up, ramp-down rates of generator i at time t.
- $\text{MinUp}_i, \text{MinDown}_i$: Minimum-up and minimum-down times of generator i.
- $f_l^{\max}$: Flow limit on transmission line l in base case.
- $f_{l,i}^{\max}$: Flow limit on transmission line l in contingency i (i.e. line i is dropped).
- $\boldsymbol{B}_p, \boldsymbol{B}_d$: Network incidence matrices for generators and load.
- $\boldsymbol{a}_l$: Network shift factor vector for line l for the base case.
- $\boldsymbol{a}_{l,i}$: Network shift factor vector for line l in contingency i.
- $\bar{d}_j^t$: Expected demand at node j, time t.
- $\mathcal{N}_r$: The set of system reserve requirements $\mathcal{N}_r = \{\text{TMSR, T10, T30}\}$.
- $\mathcal{A}$: The set of reserve products $\mathcal{A} = \{\text{TMSR, TMNSR, TMOR}\}$.
- $\mathcal{A}_k$: The set of reserve products needed to satisfy reserve requirement $k \in \mathcal{N}_r$: $\mathcal{A}_{\text{TMSR}} = \{\text{TMSR}\}, \mathcal{A}_{\text{T10}} = \{\text{TMSR,TMNSR}\}, \mathcal{A}_{\text{T30}} = \{\text{TMSR,TMNSR,TMOR}\}$.
- $\bar{q}_{i,k}^t$: Reserve capacity of generator i, requirement $k \in \mathcal{N}_r$, time t.
- $\bar{q}_{i,k}^t$: System reserve requirement of $k \in \mathcal{N}_r$, time t.

The system operator in the current practice commits extra generation resource, called reserve, in the day-ahead scheduling. The reserve capacity will be available to the system operator in the real-time operation to prepare for unexpected loss of generators or other system disruptions. According to how fast the reserve capacity can respond to the emergency, there are three important types of reserves: ten-minute spinning reserve (TMSR), ten-minute nonspinning reserve (TMNSR), and thirty-minute operating reserve (TMOR). Other types of reserves exist, such as regulation service (automatic generation control) which responds to frequency changes in the system second by second, and supplement reserve.

The decision variables of the unit commitment problem are:

- $x_i^t \in \{0,1\}$: If generator i is on at time t, $x_i^t = 1$; otherwise $x_i^t = 0$.
- $u_i^t \in \{0,1\}$: If generator i is turned on at time t, $u_i^t = 1$; otherwise $u_i^t = 0$.
- $v_i^t \in \{0,1\}$: If generator i is turned down at time t, $v_i^t = 1$; otherwise $v_i^t = 0$.
- $p_i^t \in [0,\infty)$: Production of generator i at time t.
- $q_{i,a}^t \in [0,\infty)$: Reserve of generator i, type $a \in \mathcal{A}$, time t.

A standard deterministic UC model is formulated below [45, 85].

$$\min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{p},\boldsymbol{q}} \quad \sum_{t=1}^{T}\sum_{i=1}^{N_g} x_i^t F_i^t + u_i^t S_i^t + v_i^t G_i^t + C_i^t(p_i^t) \tag{2.1}$$

$$x_i^{t-1} - x_i^t + u_i^t \geq 0, \quad \forall i \in \mathcal{N}_g, t \in \mathcal{T}, \tag{2.2}$$

$$x_i^t - x_i^{t-1} + v_i^t \geq 0, \quad \forall i \in \mathcal{N}_g, t \in \mathcal{T}, \tag{2.3}$$

$$x_i^t - x_i^{t-1} \leq x_i^\tau,\ \forall \tau \in [t+1, \min\{t + \mathrm{MinUp}_i - 1, T\}], t \in [2,T], \tag{2.4}$$

$$x_i^{t-1} - x_i^t \leq 1 - x_i^\tau,\ \forall \tau \in [t+1, \min\{t + \mathrm{MinDw}_i - 1, T\}], t \in [2,T], \tag{2.5}$$

$$\sum_{i=1}^{N_g} p_i^t = \sum_{j=1}^{N_d} \bar{d}_j^t, \qquad \forall t \in \mathcal{T}, \tag{2.6}$$

$$-RD_i^t \leq p_i^t - p_i^{t-1} \leq RU_i^t, \quad \forall i \in \mathcal{N}_g, t \in \mathcal{T}, \tag{2.7}$$

$$-f_l^{\max} \leq \boldsymbol{a}_l'(\boldsymbol{B}_p\boldsymbol{p}^t - \boldsymbol{B}_d\boldsymbol{d}^t) \leq f_l^{\max}, \quad \forall t \in \mathcal{T}, l \in \mathcal{N}_l, \tag{2.8}$$

$$-f_{l,i}^{\max} \leq \boldsymbol{a}_{l,i}'(\boldsymbol{B}_p\boldsymbol{p}^t - \boldsymbol{B}_d\boldsymbol{d}^t) \leq f_{l,i}^{\max}, \quad \forall t \in T, l \in CT_i, i \in N_l, \tag{2.9}$$

$$p_i^{\min} \leq p_i^t + \sum_{a\in\mathcal{A}} q_{i,a}^t \leq p_i^{\max}, \quad \forall i \in \mathcal{N}_g, t \in \mathcal{T}, \tag{2.10}$$

$$\sum_{i\in\mathcal{N}_g, a\in\mathcal{A}_k} q_{i,a}^t \geq \bar{q}_k^t, \quad \forall t \in \mathcal{T}, k \in \mathcal{N}_r, \tag{2.11}$$

$$\sum_{a\in\mathcal{A}_k} q_{i,a}^t \leq \bar{q}_{i,k}^t, \quad \forall i \in \mathcal{N}_g, t \in \mathcal{T}, k \in \mathcal{N}_r. \tag{2.12}$$

Eq. (2.2) and (2.3) are logic constraints between on and off status and the turn-on and turn-off actions. In particular, a generator i is turned on at time t $u_i^t = 1$ if and only if $x_i^{t-1} = 0, x_i^t = 1$. Similarly, a generator i is turned off at time t if and only if $x_i^{t-1} = 1, x_i^t = 0$. Eq. (2.4) and (2.5) are constraints of minimum up and minimum down times for each generator, i.e. if a generator is turned on at time t, then it must remain on at least for the next $(\text{MinUp}_i - 1)$ periods, and similar for the shutdown constraint. There are multiple ways to model these constraints. The specific form that we use here follow the formulation proposed in [83]. The convex hull of the binary variable $\boldsymbol{x}$ defined by Eq. (2.4)-(2.5) is explicitly characterized in [58]. Furthermore, the convex hull of the binary variables $(\boldsymbol{x}, \boldsymbol{u})$ defined by Eq. (2.2)-(2.3) and Eq. (2.4)-(2.5) is characterized in [71]. In the software implementation of our robust unit commitment model, we use the strong formulation proposed in [71].

Eq. (2.6) is the energy balance equation that matches the system level supply and load at each time period. Eq. (2.7) is the ramp rate constraint, i.e. the speed at which a generator can increase or decrease its production level is bounded in a range. Notice that the ramping constraint is a complicating constraint that couples many consecutive time periods.

Eq. (2.8) is the transmission flow constraint for the base case, where all transmission lines are functioning. Eq. (2.9) is the transmission line constraint for the i-th contingency where transmission line i is tripped. In this situation, the network topology is changed, so are the shift factor $\boldsymbol{a}_{l,i}$ and flow limits $f_{l,i}^{\max}$. Eq. (2.8) and (2.9) are the DC approximation of the nonlinear AC power flow equations. In Appendix, we review the physical laws that govern the power flows and give a derivation of the DC approximation from the the AC model. Incorporating the nonlinear AC flow equations into the unit commitment problem would result in a nonlinear mixed-integer optimization problem, which is significantly more difficult to solve than a linear mixed-integer problem. In the literature, it has been standard to use the DC model in the unit commitment problem. It remains a challenging problem to incorporate AC power flow model into the economic dispatch problem where the commitment variables are fixed.

Eq. (2.10) is the constraint that the sum of the production output and the reserve should be within the upper and lower bounds for each generator. Eq. (2.11) describes the requirement on how much reserve the system should have reserve category $k \in \mathcal{N}_r$ at time t. Eq. (2.12) says generator i can provide at most $\bar{q}_{i,k,t}$ for reserve requirement k at time t. Since reserve also takes up the generation capacity,

The variable production cost, or the supply curve in a market setting, $C_i^t(p_i^t)$ is an increasing convex piece-wise linear function of the production output p_i^t. In the ISO-NE's day-ahead energy market, each generator is allowed to submit a supply curve of up to 20 pieces.

To simplify the notation, we present a compact matrix formulation for the rest of the chapter.

$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & \boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{F}\boldsymbol{x} \leq \boldsymbol{f}, && (2.13)\\
& \boldsymbol{H}\boldsymbol{y} \leq \boldsymbol{h}, && (2.14)\\
& \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leq \boldsymbol{g}, && (2.15)\\
& \boldsymbol{I}_u\boldsymbol{y} = \bar{\boldsymbol{d}}, && (2.16)\\
& \boldsymbol{x} \text{ binary.}
\end{aligned}$$

The binary variable $\boldsymbol{x}$ is a vector of commitment related decision including the on/off and start-up/shut-down status of each generation unit for each time interval of the commitment period, usually 24 hours in an ISO setting. The continuous variable $\boldsymbol{y}$ is a vector of dispatch related decision including the generation output, load consumption levels, resource reserve levels, and power flows in the transmission network for each time interval.

The objective function is to minimize the sum of commitment cost $\boldsymbol{c}^T\boldsymbol{x}$ (including start-up, no-load, and shut-down costs) and dispatch cost $\boldsymbol{b}^T\boldsymbol{y}$ over the planning horizon. Constraint (2.13), involving only commitment variables, contains minimum up and down, and start-up/shut-down constraints. Constraint (2.14) includes dispatch related constraints such as energy balance (equality can always be written as two opposite inequalities), reserve requirement, reserve capacity, transmission limit, and ramping constraints. Constraint (2.15) couples the commitment and dispatch decisions, including minimum and maximum generation capacity constraints. Constraint (2.16) emphasizes the fact that the uncertain nodal net loads are fixed at expected values in the deterministic model ($\boldsymbol{I}_u$ selects the components from $\boldsymbol{y}$ that correspond to uncertain resources).

2.3 The Two-stage Fully Adaptive Robust SCUC Formulation

In this section, we first discuss the uncertainty set, which is a key building block of the robust model. Then, we introduce the two-stage adaptive robust SCUC formulation and provide a detailed explanation.

2.3.1 The Uncertainty Set of Net Load

The first step to build a robust model is to construct an uncertainty set. Unlike the stochastic optimization approach, the uncertainty model in a robust optimization formulation is not a probability distribution, but rather a deterministic set. In our model, the uncertain parameter is the nodal net load. We consider the following uncertainty set of nodal net load at each time period t in the planning horizon $\mathcal{T}$,

$$\begin{aligned} \mathcal{D}^t(\bar{\boldsymbol{d}}^t, \hat{\boldsymbol{d}}^t, \Delta^t) := \Bigg\{ \boldsymbol{d}^t \in \mathbb{R}^{N_d} : \sum_{i \in \mathcal{N}_d} \frac{|d_i^t - \bar{d}_i^t|}{\hat{d}_i^t} \leq \Delta^t, \\ d_i^t \in \left[\bar{d}_i^t - \hat{d}_i^t, \bar{d}_i^t + \hat{d}_i^t\right], \forall i \in \mathcal{N}_d \Bigg\}, \end{aligned} \tag{2.17}$$

where $\boldsymbol{d}^t = (d_i^t, i \in \mathcal{N}_d)$ is the vector of uncertain net loads at time t, $\hat{d}_i^t$ is the deviation of the uncertain load from the nominal value $\bar{d}_i^t$, and the interval $\left[\bar{d}_i^t - \hat{d}_i^t, \bar{d}_i^t + \hat{d}_i^t\right]$ is the range of the uncertain d_i^t, and the inequality in (2.17) controls the total deviation of all load from their nominal values at time t. The parameter Δ^t is the "budget of uncertainty", taking values between 0 and N_d. When $\Delta^t = 0$, the uncertainty set is a singleton, i.e. $\mathcal{D}^t = \{\bar{\boldsymbol{d}}^t\}$, corresponding to the deterministic case. As Δ^t increases, the size of the uncertainty set $\mathcal{D}^t$ enlarges, which implies that larger total deviation from the expected net load is considered. Therefore, the resulting robust UC solutions are more conservative and the system is protected against a higher degree of uncertainty. In particular, when $\Delta^t = N_d$, $\mathcal{D}^t$ equals to the entire hypercube defined by the intervals for each d_i^t for $i \in \mathcal{N}_d$.

More compactly, the uncertainty set can be written as the intersection of a l_1 ball and a l_∞ ball.

$$\mathcal{D}^t(\bar{\boldsymbol{d}}^t, \hat{\boldsymbol{d}}^t, \Delta^t) = \{\boldsymbol{d}^t \in \mathbb{R}^{N_d} : \|\hat{\boldsymbol{D}}_t^{-1}(\boldsymbol{d}^t - \bar{\boldsymbol{d}}^t)\|_1 \leq \Delta^t, \|\boldsymbol{d}^t - \bar{\boldsymbol{d}}^t\|_\infty \leq \hat{\boldsymbol{d}}^t\},$$

where $\hat{\boldsymbol{D}}_t = \mathbf{diag}(\hat{d}_1^t, \ldots, \hat{d}_{N_d}^t)$.

2.3.2 The Adaptive Robust SCUC Formulation

The two-stage adaptive robust SCUC problem is formulated as follows,

$$\begin{aligned}
\min_{\boldsymbol{x}, \boldsymbol{y}(\cdot)} \quad & \left(\boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{d} \in \mathcal{D}} \boldsymbol{b}^T\boldsymbol{y}(\boldsymbol{d})\right) && (2.18)\\
\text{s.t.} \quad & \boldsymbol{F}\boldsymbol{x} \leq \boldsymbol{f},\ \boldsymbol{x} \text{ is binary},\\
& \boldsymbol{H}\boldsymbol{y}(\boldsymbol{d}) \leq \boldsymbol{h}(\boldsymbol{d}),\ \forall \boldsymbol{d} \in \mathcal{D},\\
& \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{d}) \leq \boldsymbol{g},\ \forall \boldsymbol{d} \in \mathcal{D},\\
& \boldsymbol{I}_u\boldsymbol{y}(\boldsymbol{d}) = \boldsymbol{d},\ \forall \boldsymbol{d} \in \mathcal{D},
\end{aligned}$$

where $\boldsymbol{d} = (\boldsymbol{d}^t, t \in \mathcal{T})$ and $\mathcal{D} = \prod_{i \in \mathcal{T}} \mathcal{D}^t$. The objective function has two parts, reflecting the two-stage nature of the decision. The first part is the commitment cost. The second part is the worst case second-stage dispatch cost.

In the adaptive robust model, the commitment decision $\boldsymbol{x}$ takes into account all possible realization of future net load in the uncertainty set. Such a UC solution remains feasible, thus *robust*, for any realization of the uncertain net load. In comparison, the traditional UC solution only guarantees feasibility for a single nominal net load, whereas the stochastic optimization UC solution only considers a finite set of preselected scenarios of the uncertain net load. Furthermore, in our formulation the optimal second-stage decision $\boldsymbol{y}(\boldsymbol{d})$ is a function of the uncertain net load $\boldsymbol{d}$, therefore, *fully adaptive* to any realization of the uncertainty. The functional form of $\boldsymbol{y}(\cdot)$ is determined implicitly by the optimization problem, rather than being presumed. This full adaptability models the economic dispatch procedure in the real-time operation.

The above formulation can be recast in the following equivalent form, which is easier to see the structure of the two-stage robust model.

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \left(\boldsymbol{c}^T \boldsymbol{x} + \max_{\boldsymbol{d} \in \mathcal{D}} \min_{\boldsymbol{y} \in \Omega(\boldsymbol{x}, \boldsymbol{d})} \boldsymbol{b}^T \boldsymbol{y} \right) \\ \text{s.t.} \quad & \boldsymbol{F}\boldsymbol{x} \leq \boldsymbol{f}, \ \ \boldsymbol{x} \text{ binary}, \end{aligned} \tag{2.19}$$

where $\Omega(\boldsymbol{x}, \boldsymbol{d}) = \{\boldsymbol{y} : \boldsymbol{H}\boldsymbol{y} \leq \boldsymbol{h}, \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leq \boldsymbol{g}, \boldsymbol{I}_u \boldsymbol{y} = \boldsymbol{d}\}$ is the set of feasible dispatch solutions for a fixed commitment decision $\boldsymbol{x}$ and nodal net load realization $\boldsymbol{d}$. The first stage decision is the unit commitment decision $\boldsymbol{x}$. For a fixed commitment decision, the second stage implicitly finds the economic dispatch solution for each realization of the load, and identifies the worst case load that results in the maximum dispatch cost.

It is useful to write out the dual of the dispatch problem $\min_{\boldsymbol{y} \in \Omega(\boldsymbol{x}, \boldsymbol{d})} \boldsymbol{b}^T \boldsymbol{y}$. Denote its cost by $S(\boldsymbol{x}, \boldsymbol{d})$.

$$\begin{aligned} S(\boldsymbol{x}, \boldsymbol{d}) = \max_{\boldsymbol{\varphi}, \boldsymbol{\lambda}, \boldsymbol{\eta}} \ & \boldsymbol{\lambda}^T (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{g}) - \boldsymbol{\varphi}^T \boldsymbol{h} + \boldsymbol{\eta}^T \boldsymbol{d} \\ & - \boldsymbol{\lambda}^T \boldsymbol{B} - \boldsymbol{\varphi}^T \boldsymbol{H} + \boldsymbol{\eta}^T \boldsymbol{I}_u = \boldsymbol{b}^T, \\ & \boldsymbol{\varphi} \geq \boldsymbol{0}, \boldsymbol{\lambda} \geq \boldsymbol{0}, \boldsymbol{\eta} \text{ free} \end{aligned} \tag{2.20}$$

where $\boldsymbol{\varphi}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}$ are the multipliers of the constraints (2.14), (2.15) and (2.16), respectively.

Now, the second-stage problem $\max_{\boldsymbol{d} \in \mathcal{D}} \min_{\boldsymbol{y} \in \Omega(\boldsymbol{x}, \boldsymbol{d})} \boldsymbol{b}^T \boldsymbol{y}$ is equivalent to a bilinear optimization problem given below

$$
\begin{aligned}
R(\boldsymbol{x}) = \max_{\boldsymbol{d},\boldsymbol{\varphi},\boldsymbol{\lambda},\boldsymbol{\eta}} \; & \boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{g}) - \boldsymbol{\varphi}^T\boldsymbol{h} + \boldsymbol{\eta}^T\boldsymbol{d} \\
& -\boldsymbol{\lambda}^T\boldsymbol{B} - \boldsymbol{\varphi}^T\boldsymbol{H} + \boldsymbol{\eta}^T\boldsymbol{I}_u = \boldsymbol{b}^T, \\
& d \in \mathcal{D} \\
& \boldsymbol{\varphi} \geq \mathbf{0}, \boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{\eta} \text{ free.}
\end{aligned} \tag{2.21}
$$

The uncertainty set $\mathcal{D}$ is assumed to be polyhedral. Computing $R(\boldsymbol{x})$ is non-trivial. To see this, notice that the objective function of (2.21) contains a non-concave bilinear term $\boldsymbol{\eta}^T\boldsymbol{d}$, and bilinear programs are in general NP-hard to solve. Also notice that the objective function optimized over $(\boldsymbol{\varphi}, \boldsymbol{\lambda}, \boldsymbol{\eta})$ results in a convex piece-wise linear function in $\boldsymbol{d}$ (the pieces are not given explicitly), and maximization over such a convex function is NP-hard. Throughout this chapter, we assume $R(\boldsymbol{x}) < +\infty$ for all feasible $\boldsymbol{x}$. This can be ensured by adding penalty terms in the dispatch constraints. We omit the penalty terms here for a clearer presentation.

2.4 Solution Method to Solve the Adaptive Robust Model

As analyzed in the previous section, the adaptive robust formulation (2.19) is a two-stage problem. The first-stage is to find an optimal commitment decision $\boldsymbol{x}$. The second-stage is to find the worst-case dispatch cost under a fixed commitment solution. Therefore, we naturally have a two-level algorithm. The outer level employs a Benders decomposition (BD) type cutting plane algorithm to obtain $\boldsymbol{x}$ using the information (i.e. cuts) generated from the inner level, which solves the bilinear optimization problem (2.21).

2.4.1 The Outer Level: Benders Decomposition Algorithm

The Benders decomposition algorithm is described below.

Initialization: Let $\boldsymbol{x}_0$ be a feasible first-stage solution. Solve $R(\boldsymbol{x}_0)$ defined by (2.21) to get an initial solution $(\boldsymbol{d}_1, \boldsymbol{\varphi}_1, \boldsymbol{\lambda}_1, \boldsymbol{\eta}_1)$. Set the outer level lower bound $L^{BD} = -\infty$, upper bound $U^{BD} = +\infty$ and the number of iteration $k = 1$. Choose an outer level convergence tolerance level $\epsilon (> 0)$.

Iteration $k \geq 1$:

Step 1: Solve BD master problem.

The master problem of BD is the following mixed integer program (MIP):

$$\begin{aligned} \min_{\boldsymbol{x},\alpha} \ & \boldsymbol{c}^T \boldsymbol{x} + \alpha \\ \text{s.t. } & \alpha \geq \boldsymbol{\lambda}_l^T (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{g}) - \boldsymbol{\varphi}_l^T \boldsymbol{h} + \boldsymbol{\eta}_l^T \boldsymbol{d}_l, \ \forall l \leq k, \\ & \boldsymbol{F}\boldsymbol{x} \leq \boldsymbol{f}, \boldsymbol{x} \text{ binary}. \end{aligned} \tag{2.22}$$

Let $(\boldsymbol{x}_k, \alpha_k)$ be the optimum. Set $L^{BD} = \boldsymbol{c}^T \boldsymbol{x}_k + \alpha_k$.

Step 2: Solve BD subproblem $R(\boldsymbol{x}_k)$.

We will discuss the methodology to solve $R(\boldsymbol{x}_k)$ in the next subsection. Let $(\boldsymbol{d}_{k+1}, \boldsymbol{\varphi}_{k+1}, \boldsymbol{\lambda}_{k+1}, \boldsymbol{\eta}_{k+1})$ be the optimal solution. Set $U^{BD} = \boldsymbol{c}^T \boldsymbol{x}_k + R(\boldsymbol{x}_k)$.

Step 3: Check the outer level convergence.

If $U^{BD} - L^{BD} < \epsilon$, stop and return $\boldsymbol{x}_k$. Otherwise, let $k = k + 1$, and go to step 1.

To speed up the convergence of the above BD algorithm, we find it helpful to add dispatch constraints $\Omega(\boldsymbol{x}_k, \boldsymbol{d}_k)$ to the BD master problem (2.22) at certain iteration k when U^{BD} or L^{BD} has improved slowly.

2.4.2 The Inner Level: Solve $R(\boldsymbol{x})$

An outer approximation (OA) algorithm [32, 37] is used to solve the bilinear program (2.21). Since the problem (2.21) is nonconcave, only a local optimum is guaranteed by the OA algorithm. To verify the quality of the solution, we test on different initial conditions and observe fast convergence and consistent results. The OA algorithm is described below.

Initialization: Fix the unit commitment decision $\boldsymbol{x}_k$ passed from the kth iteration of the outer level BD algorithm. Find an initial net load $\boldsymbol{d}_1 \in \mathcal{D}$. Set the inner level lower bound $L^{OA} = -\infty$, upper bound $U^{OA} = +\infty$ and number of iteration $j = 1$. Choose an inner level convergence tolerance level $\delta(> 0)$.

Iteration $j \geq 1$:

Step 1: Solve OA subproblem $S(\boldsymbol{x}_k, \boldsymbol{d}_j)$.

Solve $S(\boldsymbol{x}_k, \boldsymbol{d}_j)$, the dual of the dispatch problem defined by (2.20). Let $(\boldsymbol{\varphi}_j, \boldsymbol{\lambda}_j, \boldsymbol{\eta}_j)$ be the optimal solution. Set $L^{OA} = S(\boldsymbol{x}_k, \boldsymbol{d}_j)$. Define $L_j(\boldsymbol{d}_j, \boldsymbol{\eta}_j)$, the linearization of the bilinear term $\boldsymbol{\eta}^T\boldsymbol{d}$ with respect to $(\boldsymbol{d}_j, \boldsymbol{\eta}_j)$, as follows

$$L_j(\boldsymbol{d}_j, \boldsymbol{\eta}_j) = \boldsymbol{\eta}_j^T\boldsymbol{d}_j + (\boldsymbol{\eta} - \boldsymbol{\eta}_j)^T\boldsymbol{d}_j + (\boldsymbol{d} - \boldsymbol{d}_j)^T\boldsymbol{\eta}_j.$$

Step 2: Solve OA master problem.

Solve the linearized version of $R(\boldsymbol{x}_k)$, defined below:

$$\begin{aligned}
U(\boldsymbol{d}_j, \boldsymbol{\eta}_j) = \max_{\boldsymbol{d},\boldsymbol{\varphi},\boldsymbol{\lambda},\boldsymbol{\eta},\beta}\ & \boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{g}) - \boldsymbol{\varphi}^T\boldsymbol{h} + \beta \\
& \beta \leq L_i(\boldsymbol{d}_i, \boldsymbol{\eta}_i), \quad \forall i = 1, \ldots, j \\
& -\boldsymbol{\lambda}^T\boldsymbol{B} - \boldsymbol{\varphi}^T\boldsymbol{H} + \boldsymbol{\eta}^T\boldsymbol{I}_u = \boldsymbol{b}^T, \\
& d \in \mathcal{D}, \\
& \boldsymbol{\varphi} \geq \boldsymbol{0}, \boldsymbol{\lambda} \geq \boldsymbol{0}, \boldsymbol{\eta} \text{ free.}
\end{aligned}$$

Since the uncertainty set $\mathcal{D}$ is assumed to be polyhedral, $U(\boldsymbol{d}_j, \boldsymbol{\eta}_j)$ is a linear program. Denote $(\boldsymbol{d}_{j+1}, \boldsymbol{\varphi}_{j+1}, \boldsymbol{\lambda}_{j+1}, \boldsymbol{\eta}_{j+1}, \beta_{j+1})$ as the optimal solution. Set the inner level upper bound as $U^{OA} = U(\boldsymbol{d}_j, \boldsymbol{\eta}_j)$.

Step 3: Check the inner level convergence. If $U^{OA} - L^{OA} < \delta$, then stop and output the current solution. Otherwise, set $j = j + 1$, go to Step 1 of the OA algorithm.

The overall algorithmic framework is summarized in Fig. 2-1. In our implementation, we include all contingency transmission constraints in the BD subproblem. However, we could also use simultaneous feasibility test (SFT) procedure to generate transmission security constraints on the fly.

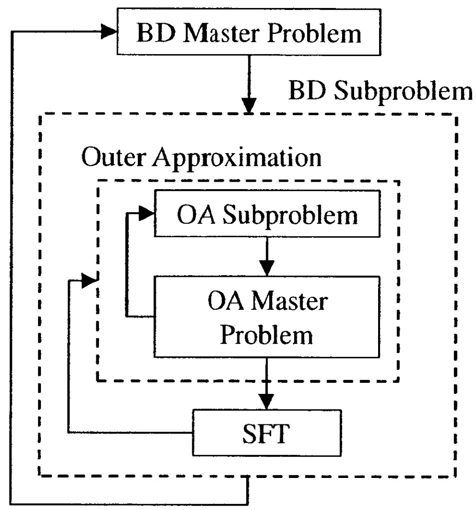

Figure 2-1: Flow chart of the proposed two-level algorithm.

2.5 Computational Experiments

In this section, we present a computational study to evaluate the performance of the adaptive robust (AdaptRob) approach and the reserve adjustment (ResAdj) approach. We test on the power system operated by the ISO New England Inc. (ISO-NE). We have the following data and uncertainty model.

The system and market data: The system has 312 generating units, 174 loads, and 2816 nodes. We select four representative transmission lines that interconnect four major load zones in the ISO-NE's system. The market data is taken from a normal winter day of the ISO-NE's day-ahead energy market. In particular, we have 24-hour data of generators' offer curves, reserve offers, expected nodal load, system reserve requirement (10-min spinning reserve, 10-min non-spinning reserve, and 30-min reserve), and various network parameters. The average total generation capacity per hour is 31,400 MW and the average forecasted load per hour is 14,136 MW.

The uncertainty model: We use the budgeted uncertainty set defined in (2.17), in which $\bar{d}_j^t$ is the nominal load given in the data. We set the range of load variation to be $\hat{d}_j^t = 0.1\bar{d}_j^t$ for each load j at time t. The budget of uncertainty Δ^t takes values in the entire range of 0 to N_d. We will discuss the proper way to choose Δ^t within this range that results in the best performance of the robust solution in the following subsection.

For the ResAdj approach, we solve the deterministic UC problems presented in Section II at the expected load level with adjusted reserve requirement. In particular, we model the reserve adjustment as follows,

$$q_{j,t} = q_{j,t}^0 + \frac{\Delta^t}{N_d} \sum_{i=1}^{N_d} \hat{d}_i^t,$$

where $q_{j,t}$ is the system reserve requirement of type j at time t, composed of the basic reserve level $q_{j,t}^0$ and an adjustment part proportional to the total variation of load. Here, the uncertainty budget Δ^t also controls the level of conservatism of the ResAdj solution.

The computational experiments proceed as follows.

1. Obtain UC solutions: Solve the ResAdj and AdaptRob UC models respectively for different uncertainty budgets: $\Delta^t \in [0, N_d]$.

2. Dispatch simulation: For each UC solution, solve the dispatch problem repeatedly for two sets of randomly generated loads, each of which has 1000 samples.

One set of randomly sampled load follows a uniform distribution in the interval $[\bar{d}_j^t - \hat{d}_j^t, \bar{d}_j^t + \hat{d}_j^t]$ for each load j at time t. The other set follows a normal distribution with mean $\bar{d}_j^t$ and standard deviation $\hat{d}_j^t/1.44$, which results in an 85 percentile of load falling between $[\bar{d}_j^t - \hat{d}_j^t, \bar{d}_j^t + \hat{d}_j^t]$ (the load is truncated for nonnegative values). Notice that samples from both sets may fall outside of the budgeted uncertainty set, especially when the budget Δ^t is close to 0. In this way, we can test the robustness of our model in the situation where the uncertainty model is not entirely accurate.

To mimic the high cost of dispatching fast-start units or load shedding in the real-time operation, we introduce a slack variable $\boldsymbol{z}$ for energy balance, reserve requirement, and transmission constraints in the dispatch simulation. If the real-time dispatch incurs any energy deficiency, reserve shortage, or transmission violation, at least one component of $\boldsymbol{z}$ will be positive. The dispatch cost is the sum of production cost and penalty cost, i.e. $\boldsymbol{c}^T\boldsymbol{p} + \boldsymbol{\kappa}^T\boldsymbol{z}$, where $\boldsymbol{\kappa}$ is set to \$5000/MWh for each component.

The proposed two-level algorithm for the two-stage adaptive robust model is implemented in GAMS. The mixed integer program and linear program in the algorithm are solved with CPLEX 12.1.0 on a PC laptop with an Intel Core(TM) 2Duo 2.50GHz CPU and 3GB memory. We set the convergence tolerance for the outer level BD algorithm to be $\epsilon = 10^{-4}$, and the convergence tolerance for the inner level algorithm to be $\delta = 10^{-3}$. The MIP gap for the BD master problem is set to 10^{-5}. The average computation time to solve the robust UC problem is 6.14 hours. The average computation time for the reserve adjustment approach is 1.65 hours. The computational results presented in the following subsections use the above tolerance levels and MIP gap. If we relax the BD convergence tolerance ϵ to 10^{-3} and the MIP gap of the BD master problem to 10^{-3}, the average computation time to solve the robust UC problem decreases to 1.46 hours with an average of 0.17% increase in terms of the worst-case total cost, which shows there may be enough room for a fined tuning of the algorithm paramters to improve the computational time without satisfying too much on solution accuracy.

We compare the performance of the reserve adjustment approach and the adaptive robust approach in three aspects: (1) The average dispatch and total costs, (2) the volatility of these costs, and (3) the sensitivity of the these costs to different probability distributions of the uncertain load. The average cost indicates the economic efficiency of the UC decision; the volatility of these costs indicates the reliability of the real-time dispatch operation under the UC decision; the third aspect indicates how robust the UC decision is against load probability distributions.

The main conclusion is that (a) by properly adjusting the budget level of the uncertainty set, the adaptive robust solution has lower average dispatch and total costs, indicating better economic efficiency of the robust approach; (b) the adaptive robust solution significantly reduces the volatility of the total costs, as well as the penalty cost in the dispatch operation; (c) the adaptive robust solution is significantly more robust to different probability distributions of load.

We want to remark that (a) is quite contrary to the general impression that the robust solution is always overconservative. In fact, by choosing a proper uncertainty model, the robust solution achieves better economic efficiency than the conventional approach. We will present detailed discussion later. The computational results for normally and uniformly distributed loads are similar in illustrating (a) and (b). We only present the results for normally distributed loads. For (c), we compare the results of the two distributions.

2.5.1 Cost Efficiency and the Choice of the Budget Level

Table 2.5.1 reports the average dispatch costs and total costs of AdptRob and ResAdj solutions for normally distributed load when the uncertainty budget Δ^t varies from 0 to N_d. We can see that the AdptRob has lower average dispatch costs for all values of Δ^t. The average total costs of the two approaches are more comparable.

To quantify the comparison, we define the *cost saving* as

$$\text{cost saving} = (\text{ResAdj cost} - \text{AdptRob cost})/(\text{ResAdj cost}).$$

For the normally distributed load, the AdptRob approach always has lower average dispatch cost than the ResAdj approach, and can save up to 2.7% or \$472k (at $\Delta = 0.1N_d$), which is a significant saving for a daily operation. The total cost saving of the AdptRob approach ranges from -0.84% (at $\Delta = 0.4N_d$) to 1.19% (at $\Delta = 0.1N_d$). Since the AdptRob approach protects the system against the worst-case load realization, in general it commits more generation resources than the reserve adjustment approach, which only plans for the expected load.

For both average dispatch and total costs, we observe that the robust solution performs best when the budget level Δ^t is relatively small, e.g. around $0.1N_d$. This

	AdptRob		ResAdj	
budget Δ^t/N_d	dispatch cost (M$)	total cost (M$)	dispatch cost (M$)	total cost (M$)
0.0	19.3530	20.8503	19.3530	20.8503
0.1	16.9852	18.7290	17.4581	18.9547
0.2	17.0513	18.8265	17.2391	18.7400
0.3	17.0949	18.8773	17.2563	18.7595
0.4	17.1448	18.9425	17.2570	18.7843
0.5	17.1583	18.9569	17.2893	18.8250
0.6	17.1705	18.9723	17.3030	18.8506
0.7	17.1719	18.9896	17.3537	18.9062
0.8	17.1715	18.9892	17.3899	18.9472
0.9	17.1655	18.9898	17.3990	18.9669
1.0	17.1652	18.9894	17.4524	19.0660

Table 2.1: The average dispatch costs and total costs of the AdptRob and ResAdj for normally distributed load for $\Delta^t/N_d = 0, 0.1, \dots, 1$.

phenomenon can be explained by the probability law, namely the central limit theorem. When a large number N_d of random loads (independent as we assume) are aggregated, the volatility of the total load scales according to $\sqrt{N_d}$. Therefore, a proper level of uncertainty budget in the uncertainty set (2.17) should be chosen as $\Delta^t \sim O(\sqrt{N_d})$. Table 2.5.1 shows the results for Δ^t in $0.5\sqrt{N_d}$ to $3\sqrt{N_d}$, equivalently Δ^t equals $0.038N_d$ to $0.227N_d$, for $N_d = 174$ in our system. We can see that, in this range of the uncertainty budget, the AdptRob has an even higher saving: the average dispatch cost saving from 1.86% or $321.2k (at $\Delta = 1.5\sqrt{N_d}$) to 6.96% or $1.27 Million (at $\Delta = 0.5\sqrt{N_d}$), and the average total cost saving from 0.34% or $64.1k (at $\Delta = 1.5\sqrt{N_d}$) to 5.48% or $1.08 Million (at $\Delta = 0.5\sqrt{N_d}$). This demonstrates that the probability law provides a proper guideline for choosing the best budget level.

2.5.2 Reliability of Dispatch Operation

The adaptive robust approach significantly reduces the volatility of the real-time dispatch costs. Table 2.5.2 shows the standard deviation (std) of the dispatch costs of the two approaches, and their ratios (ResAdj/AdptRob) for normally distributed load. We can see that the std for the reserve adjustment approach is almost an order of

	AdptRob		ResAdj	
budget $\Delta^t/\sqrt{N_d}$	dispatch cost (M$)	total cost (M$)	dispatch cost (M$)	total cost (M$)
0.5	16.9195	18.6050	18.1855	19.6837
1.0	16.9650	18.6688	17.4907	18.9942
1.5	16.9815	18.7365	17.3027	18.8006
2.0	17.0297	18.7937	17.7403	19.2415
2.5	17.0586	18.8366	17.6567	19.1618
3.0	17.0745	18.8526	18.0804	19.5889

Table 2.2: The average dispatch costs and total costs of the AdptRob and ResAdj for normally distributed load $\Delta^t/\sqrt{N_d} = 0.5, 1, \dots, 3$.

magnitude higher than that for the adaptive robust approach (8.15-14.48 times). The significant reduction in the std of the dispatch cost is closely related to the reduction in the penalty cost. Table 2.5.2 shows the penalty costs of the two approaches. Recall that the dispatch cost is the sum of the production cost and penalty cost. The penalty cost occurs whenever there is a violation in the energy balance, reserve requirement, or transmission constraints. The system operator has to take expensive emergency actions such as dispatching fast-start units or load-shedding to maintain system reliability. All these actions add volatility to the dispatch costs.

	AdptRob	ResAdj	ResAdj
budget Δ^t/N_d	std dispatch cost (k$)	std dispatch cost (k$)	/ AdptRob
0.0	1,769.5107	1,769.5107	1
0.1	47.4900	687.5098	14.48
0.2	46.3647	687.5098	8.62
0.3	45.4248	377.7901	8.32
0.4	44.2397	366.7359	8.29
0.5	44.1075	377.1875	8.55
0.6	43.9936	370.8673	8.43
0.7	43.9263	377.0631	8.58
0.8	43.9338	370.7203	8.44
0.9	43.9023	357.9338	8.15
1.0	43.9431	361.0376	8.22

Table 2.3: Standard deviation of the dispatch costs of the two approaches and their ratio for normally distributed load.

	AdptRob	ResAdj	
budget Δ^t/N_d	penalty cost (k$)	penalty cost (k$)	percent of disp. cost
0.0	2377.62	2,377.6237	12.29%
0.1	0	497.8243	2.85%
0.2	0	272.4362	1.58%
0.3	0	268.0298	1.55%
0.4	0	252.3463	1.46%
0.5	0	267.6559	1.55%
0.6	0	259.7954	1.50%
0.7	0	267.6559	1.54%
0.8	0	259.7954	1.49%
0.9	0	241.6133	1.39%
1.0	0	247.4119	1.42%

Table 2.4: Penalty costs of AdptRob and ResAdj approaches for normally distributed load.

As we observe, all three types of constraint violations (enery, reserve, and transmission) may occur for ResAdj solutions, indicating the potential ineffectiveness of a deterministic approach that only considers nominal load and system level reserves. In contrast, the AdptRob approach commits resources by taking consideration of all load realizations in the uncertainty model, as well as the transmission network. Furthermore, the robust solutions remain feasible even when the load realization is outside of the uncertainty set (as the case for normally distributed load).

In conclusion, the low volatility of the dispatch cost and the zero penalty cost of the adaptive robust approach demonstrates its operational effectiveness in reducing costly emergency actions and improving system reliability.

2.5.3 Robustness Against Load Distributions

In practice, it is a non-trivial job to accurately identify the probability distribution of the load uncertainty for each node, especially in a large-scale power system. Thus, it is important for a UC solution to have stable economic and operational performance over different distributions of the uncertain load. The simulation results show that the adaptive robust approach exhibits such desirable property. In comparison, the performance of the reserve adjustment approach can be significantly affected by the underlying probability distribution.

As shown in Fig. 2-2, the average dispatch costs of the AdptRob approach are almost the same for loads with normal and uniform distributions. The absolute difference between the two curves is between \$6.32k and \$15.80k for the entire range of $\Delta^t = 0.1N_d, \ldots, 0.9N_d$. The relative difference is between 0.037% to 0.092% (defined as (normal cost-uniform cost)/normal cost).

The ResAdj approach has a rather different picture, as shown in Fig. 2-3. The average dispatch costs are significantly affected by the load probability distribution. The absolute difference of the two curves varies between \$174.42k and \$382.26k. The relative difference is between 1.0% to 2.2%. In both measures, the difference is more than 20 times larger than that of the AdptRob approach.

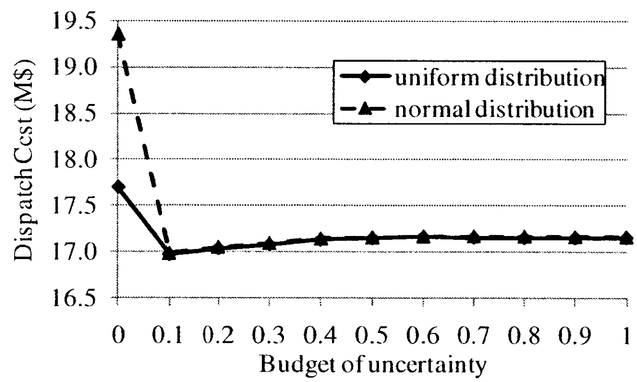

Figure 2-2: The average dispatch costs of the AdptRob approach for normally and uniformly distributed load.

We also study the effect of the load distribution on the standard deviation of dispatch costs. Table 2.5.3 shows the std of dispatch costs for loads with uniform distribution and the relative difference between the uniform distribution and the normal distribution (defined as (normal std - uniform std)/normal std, c.f. Table 2.5.2 for normal std). As shown in the table, the relative change of the std is around 18.8% for the AdptRob approach, and is around 59.6% for the ResAdj approach, which is more than three times higher. It is also interesting to note that the AdptRob approach significantly reduces the relative change of the std comparing to the deterministic

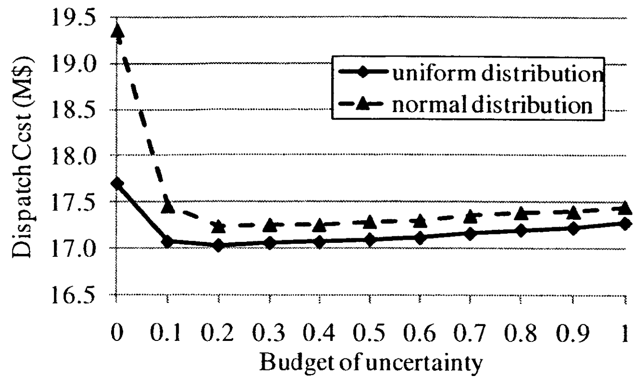

Figure 2-3: The average dispatch costs of the ResAdj approach for normally and uniformly distributed load.

approach ($\Delta = 0$) from 64.70% to around 18.8%. On the other hand, the ResAdj approach is much less effective in this respect (from 64.70% to around 59.6%).

	AdptRob		ResAdj	
budget Δ^t/N_d	std. disp. cost (k$)	Relat. diff	std. disp. cost (k$)	relat. diff
0.0	624.5939	64.70%	624.5939	64.70%
0.1	38.5824	18.76%	227.7822	66.87%
0.2	37.6948	18.70%	153.5149	61.61%
0.3	36.8956	18.78%	157.0428	58.43%
0.4	36.2507	18.06%	150.7548	58.89%
0.5	35.7960	18.84%	156.9203	58.40%
0.6	35.6687	18.92%	154.4555	58.35%
0.7	35.6179	18.91%	156.8091	58.41%
0.8	35.6067	18.95%	154.3305	58.37%
0.9	35.5761	18.97%	148.2292	58.59%
1.0	35.6092	18.97%	151.8466	57.94%

Table 2.5: Standard deviation of dispatch costs for uniformly distributed load and the relative difference with those of the normally distributed load.

2.6 Conclusion

The adaptive robust model and its solution technique presented in this chapter provide a novel and practical approach to handle uncertainties in the unit commitment process. Such a framework naturally fits into the daily reliability unit commitment process in an ISO environment. We develop a practical solution method with the goal for the method to be applicable in the real-world large scale power system operation. We conduct extensive tests on the large scale system operated by the ISO New England, and compare our model with the current reserve adjustment approach. We find that by properly setting the level of conservatism in the uncertainty model, the adaptive robust model exhibits sizable savings on both average dispatch and total costs, and significantly reduces the volatility of the dispatch cost, thus, improves the real-time reliability of the power system operation. The robust model also shows resilient performance under different probability distributions of load.

Some discussions are in order. The stochastic factors that influence the unit commitment problem are associated with both supply (e.g. wind power, etc.) and demand (e.g. demand forecast errors, price responsive demand, etc.) To address these various uncertainty elements, the proposed two-stage adaptive robust framework models the uncertainty at the individual nodal level. Therefore, the impact of resource level uncertainty on the transmission system can be evaluated. Moreover, the proposed model can be readily extended to include uncertainties related to inter-tie exchanges, system-wide and zonal level load, and interface limit de-rating.

In our model, we assume that the commitment cost function and the dispatch cost function are both linear. In reality, the production cost could be a quadratic function, the start-up cost could follow an exponential function, and the shut-down cost could be a step function. The nonlinearity in the corresponding cost functions can be modeled by the mixed-integer linear reformulation proposed in [28].

Our work focuses on polyhedral uncertainty sets. However, the model can be extended to include general convex uncertainty sets, such as ellipsoidal uncertainty sets. The framework of the proposed solution methodology, especially the outer approximation technique to solve the second-stage problem, can be generalized to handle nonlinear convex constraints.

In the current practice, the SFT runs iteratively with the unit commitment procedure by gradually adding violated transmission security constraints to the economic dispatch problem. Alternatively, as we implemented in our numerical test, we can impose a set of critical transmission security constraints in the second-stage problem without running the SFT. These critical transmission security constraints are more likely to be violated than other constraints based on the expert knowledge and historical data, and they are usually a small subset of the total transmission constraints. Therefore, this alternative approach can reduce the computation time for solving the second-stage problem.

Chapter 3

A Geometric Characterization of the Power of Finite Adaptability in Multistage Stochastic and Adaptive Optimization

3.1 Introduction

In most real world problems, several parameters are uncertain at the optimization phase and decisions are required to be made in the face of these uncertainties. Deterministic optimization is often not useful for such problems as solutions obtained through such an approach might be sensitive to even small perturbations in the problem parameters. Stochastic optimization was introduced by Dantzig [31] and Beale [3], and since then has been extensively studied in the literature. A stochastic optimization approach assumes a probability distribution over the uncertain parameters and seeks to optimize the expected value of the objective function typically. We refer the reader to several textbooks including Infanger [51], Kall and Wallace [55], Prékopa [69], Shapiro [75], Shapiro et al. [76] and the references therein for a comprehensive review of stochastic optimization.

While the stochastic optimization approach has its merits and there has been reasonable progress in the field, there are two significant drawbacks of the approach.

1. In a typical application, only historical data is available rather than probability distributions. Modeling uncertainty using a probability distribution is a choice we make, and it is not a primitive quantity fundamentally linked to the application.

2. More importantly, the stochastic optimization approach is by and large computationally intractable. Shapiro and Nemirovski [77] give hardness results for two-stage and multi-stage stochastic optimization problems where they show the multi-stage stochastic optimization is computationally intractable even if approximate solutions are desired. Dyer and Stougie [33] show that a multi-stage stochastic optimization problem where the distribution of uncertain parameters in any stage also depends on the decisions in past stages is PSPACE-hard.

To solve a two-stage stochastic optimization problem, Shapiro and Nemirovski [77] show that a sampling based algorithm provides approximate solutions given that a sufficiently large number of scenarios are sampled from the assumed distribution and the problem has a *relatively complete recourse*. A two-stage stochastic optimization problem is said to have a relatively complete recourse if for every first stage decision there is a feasible second stage recourse solution almost everywhere, i.e., with probability one (see the book by Shapiro et al. [76]). Shmoys and Swamy [78, 81] and Gupta et al. [46, 47] consider the two-stage and multi-stage stochastic set covering problem under certain restrictions on the objective coefficients, and propose sampling based algorithms that use a small number of scenario samples to construct a good first-stage decision. However, the stochastic set covering problem admits a complete recourse, i.e., for any first-stage decision there is a feasible recourse decision in each scenario, and the sampling based algorithms only work for problems with complete or relatively complete recourse.

More recently, the robust optimization approach has been considered to address optimization under uncertainty and has been studied extensively (see Ben-Tal and

Nemirovski [11, 12, 14], El Ghaoui and Lebret [34], Goldfarb and Iyengar [42], Bertsimas and Sim [21], Bertsimas and Sim [22]). In a robust optimization approach, the uncertain parameters are assumed to belong to some uncertainty set and the goal is to construct a single (static) solution that is feasible for all possible realizations of the uncertain parameters from the set and optimizes the worst-case objective. We point the reader to the survey by Bertsimas et al. [15] and the book by Ben-Tal et al. [9] and the references therein for an extensive review of the literature in robust optimization. This approach is significantly more tractable as compared to a stochastic optimization approach and the robust problem is equivalent to the corresponding deterministic problem in computational complexity for a large class of problems and uncertainty sets [15]. However, the robust optimization approach has the following drawbacks. Since it optimizes over the worst-case realization of the uncertain parameters, it may produce highly conservative solutions that may not perform well in the expected case. Moreover, the robust approach computes a single (static) solution even for a multi-stage problem with several stages of decision-making as opposed a fully-adaptable solution where decisions in each stage depend on the actual realizations of the parameters in past stages. This may further add to the conservativeness.

Another approach is to consider solutions that are fully-adaptable in each stage and depend on the realizations of the parameters in past stages and optimize over the worst case. Such solution approaches have been considered in the literature and referred to as adjustable robust policies (see Ben-Tal et al. [10] and the book by Ben-Tal et al. [9] for a detailed discussion of these policies). Unfortunately, the adjustable robust problem is computationally intractable and Ben-Tal et al. [10] introduce an affinely adjustable robust solution approach to approximate the adjustable robust problem. Affine solution approaches (or just affine policies) were introduced in the context of stochastic programming (see Gartska and Wets [39] and Rockafeller and Wets [73]) and have been extensively studied in control theory (see the survey by Bemporad and Morari [6]). Affine policies are useful due to their computational tractability and strong empirical performance. Recently Bertsimas et al. [20] show that affine policies are optimal for a single dimension multi-stage problem with box constraints and box uncertainty sets. Bertsimas and Goyal [18] consider affine policies for the two-stage adaptive (or adjustable robust) problem and give a tight bound of $O(\sqrt{\mathsf{dim}(\mathcal{U})})$ on the performance of affine policies with respect to a fully-adaptable solution where $\mathsf{dim}(\mathcal{U})$ denotes the dimension of the uncertainty set.

In this chapter, we consider a class of solutions called *finitely adaptable* solutions that were introduced by Bertsimas and Caramanis [16]. In this class of solutions, the decision-maker apriori computes a small number of solutions instead of just a single (static) solution such that for every possible realization of the uncertain parameters, at least one of them is feasible and in each stage, the decision-maker implements the best solution from the given set of solutions. Therefore, a finitely adaptable solution policy is a generalization of the static robust solution and is a middle ground between the static solution policy and the fully-adaptable policy. It can be thought of as a special case of a piecewise-affine policy where the each piece is a static solution instead of an affine solution. As compared to a fully-adaptable solution, which prescribes a solution for all possible scenarios (possibly an uncountable set), a finitely adaptable solution has only a small finite number of solutions. Therefore, the decision space is much sparser and it is significantly more tractable than the fully-adaptable solution. Furthermore, for each possible scenario, at least one of the finitely adaptable solutions is feasible. This makes it different from sampling based approaches where a small number of scenarios are sampled and an optimal decision is computed for only the sampled scenarios, while the rest of the scenarios are ignored. We believe that finitely adaptable solutions are consistent with how decisions are made in most real world problems. Unlike dynamic programming that prescribes an optimal decision for each (possibly uncountable) future state of the world, we make decisions for only a few states of the world.

We aim to analyze the performance of static robust and finitely adaptable solution policies for the two-stage and multi-stage stochastic optimization problems, respectively. We show that the performance of these solution approaches as compared to the optimal fully-adaptable stochastic solution depends on fundamental geometric properties of the uncertainty set including symmetry for a fairly general class of models. Bertsimas and Goyal [18] analyze the performance of static robust solution in two-stage stochastic problems for perfectly symmetric uncertainty sets such as ellipsoids, and norm-balls. We consider a generalized notion of symmetry of a convex set introduced in Minkowski [62], where the symmetry of a convex set is a number between 0 and 1. The symmetry of a set being equal to one implies that it is perfectly symmetric (such as an ellipsoid). We show that the performance of the static robust and the finitely adaptable solutions for the corresponding two-stage and multi-stage stochastic optimization problems depends on the symmetry of the uncertainty set.

This is a two-fold generalization of the results in [18]. We extend the results in [18] of performance of a static robust solution in two-stage stochastic optimization problems, for general convex uncertainty sets using a generalized notion of symmetry. Furthermore, we also generalize the static robust solution policy to a finitely adaptable solution policy for the multi-stage stochastic optimization problem and give a similar bound on its performance that is related to the symmetry of the uncertainty sets. The results are quite general and extend to important cases such as integrality constraints on decision variables and linear conic inequalities. To the best of our knowledge, there were no approximation bounds for the multi-stage problem in such generality.

3.2 Models and Preliminaries

In this section, we first setup the models for two-stage and multi-stage stochastic, robust, and adaptive optimization problems. We also discuss the solution concept of finitely adaptability for multi-stage problems. Then we introduce the geometric quantity of symmetry of a convex compact set and some properties that will be used. Lastly, we define a quantity called translation factor of a convex set.

3.2.1 Two-stage Optimization Models

A two-stage stochastic optimization problem, Π^2_{Stoch}, is defined as,

$$
\begin{aligned}
z_{\text{Stoch}} := \min_{\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{b})} \quad & \boldsymbol{c}^T \boldsymbol{x} + \mathbb{E}_\mu[\boldsymbol{d}^T \boldsymbol{y}(\boldsymbol{b})] \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{b}) \geq \boldsymbol{b}, \ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U} \subseteq \mathbb{R}^m_+, \\
& \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}^{n_1 - p_1}_+, \\
& \boldsymbol{y}(\boldsymbol{b}) \in \mathbb{R}^{p_2} \times \mathbb{R}^{n_2 - p_2}_+, \ \forall \boldsymbol{b} \in \mathcal{U}.
\end{aligned}
\tag{3.1}
$$

Here, the first-stage decision variable is denoted as $\boldsymbol{x}$; the second-stage decision variable is $y(\boldsymbol{b})$ for $\boldsymbol{b} \in \mathcal{U}$, where $\boldsymbol{b}$ is the uncertain right-hand side with the uncertainty set denoted as $\mathcal{U}$. The optimization in (3.1) for the second stage decisions $y(\cdot)$ is performed over the space of piecewise affine functions, since there are finitely many bases of the system of linear inequalities $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{b}) \geq \boldsymbol{b}$. Note that some of the decision variables in both the first and second stage are free, i.e., not constrained to be non-negative. A probability measure μ is defined on $\mathcal{U}$. Both $\boldsymbol{A}$ and $\boldsymbol{B}$ are certain, and there is no restriction on the coefficients in $\boldsymbol{A}, \boldsymbol{B}$ or on the objective coefficients $\boldsymbol{c}, \boldsymbol{d}$. The linear constraints $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{b}) \geq \boldsymbol{b}$ hold for μ almost everywhere on $\mathcal{U}$, i.e., the set of $\boldsymbol{b} \in \mathcal{U}$ for which the linear constraints are not satisfied has measure zero. We use the notation μ-a.e. $\boldsymbol{b} \in \mathcal{U}$ to denote this.

The key assumption in the above problem is that the uncertainty set $\mathcal{U}$ is contained in the nonnegative orthant. In addition, we make the following two technical assumptions.

1. z_{Stoch} is finite. This assumption implies that $z_{\text{Stoch}} \geq 0$. Since if $z_{\text{Stoch}} < 0$, then it must be unbounded from below as we can scale the solution by an arbitrary positive factor and still get a feasible solution.

2. For any first-stage solution $\boldsymbol{x}$, and a second-stage solution $\boldsymbol{y}(\cdot)$ that is feasible for μ almost everywhere on $\mathcal{U}$, $\mathbb{E}_\mu[\boldsymbol{y}(\boldsymbol{b})]$ exists.

These technical conditions are assumed to hold for all the multi-stage stochastic models considered in the chapter as well. We would like to note that since $\boldsymbol{A}$ is not necessarily equal to $\boldsymbol{B}$, our model does not admit relatively complete recourse.

A two-stage adaptive optimization problem, Π^2_{Adapt}, is given as,

$$\begin{aligned} z_{\text{Adapt}} := \min_{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{b})} \quad & \boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{b}\in\mathcal{U}} \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{b}) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{b}) \geq \boldsymbol{b},\ \forall \boldsymbol{b} \in \mathcal{U} \subseteq \mathbb{R}^m_+, \\ & \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}^{n_1-p_1}_+, \\ & \boldsymbol{y}(\boldsymbol{b}) \in \mathbb{R}^{p_2} \times \mathbb{R}^{n_2-p_2}_+,\ \forall \boldsymbol{b} \in \mathcal{U}. \end{aligned} \tag{3.2}$$

Note that the above problem is also referred to as an adjustable robust problem in the literature (see Ben-Tal et al. [10] and the book by Ben-Tal et al. [9]).

The corresponding static robust optimization problem, Π_{Rob}, is defined as,

$$\begin{aligned} z_{\text{Rob}} := \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{b},\ \forall \boldsymbol{b} \in \mathcal{U} \subseteq \mathbb{R}^m_+, \\ & \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}^{n_1-p_1}_+, \\ & \boldsymbol{y} \in \mathbb{R}^{p_2} \times \mathbb{R}^{n_2-p_2}_+. \end{aligned} \tag{3.3}$$

Note that any feasible solution to the robust problem (3.3) is also feasible for the adaptive problem (3.2), and thus $z_{\text{Adapt}} \leq z_{\text{Rob}}$. Moreover, any feasible solution to the adaptive problem (3.2) is also feasible for the stochastic problem (3.1) and in addition, we have

$$\boldsymbol{c}^T\boldsymbol{x} + \mathbb{E}_\mu[\boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{b})] \leq \boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{b}\in\mathcal{U}} \boldsymbol{d}^T\boldsymbol{y}(\boldsymbol{b}),$$

leading to $z_{\text{Stoch}} \leq z_{\text{Adapt}}$. Hence, we have $z_{\text{Stoch}} \leq z_{\text{Adapt}} \leq z_{\text{Rob}}$.

We would like to note that when the left hand side of the constraints, i.e., $\boldsymbol{A}$ and $\boldsymbol{B}$ are uncertain, even a two-stage, two-dimensional problem can have an unbounded gap between the static robust solution and the stochastic solution.

3.2.2 Multi-stage Optimization Models

For multi-stage problems, we consider a fairly general model where the evolution of the multi-stage uncertainty is given by a directed acyclic network $G = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of nodes corresponding to different uncertainty sets, and $\mathcal{A}$ is the set of arcs that describe the evolution of uncertainty from Stage k to $(k+1)$ for all $k = 2, \ldots, K-1$. In each Stage $(k+1)$ for $k = 1, \ldots, K-1$, the uncertain parameters $\boldsymbol{u}_k$ belong to one of the N_k uncertainty sets, $\mathcal{U}_1^k, \ldots, \mathcal{U}_{N_k}^k \subset \mathbb{R}_+^m$. We also assume that the probability distribution of $\boldsymbol{u}_k$ conditioned on the fact that $\boldsymbol{u}_k \in \mathcal{U}_j^k$ for all $k = 2, \ldots, K$, $j = 1, \ldots, N_k$ is known. In our notation, the multi-stage uncertainty network starts from Stage 2 (and not Stage 1) and we refer to the uncertainty sets and uncertain parameters in Stage $(k+1)$ using index k. Therefore, in a K-stage problem, the index $k \in \{1, \ldots, K-1\}$.

In Stage 2, there is a single node $\mathcal{U}^1$, which we refer to as the *root node*. Therefore, $N_1 = 1$. For any $k = 1, \ldots, K-1$, suppose the uncertain parameters $\boldsymbol{u}_k$ in Stage $(k+1)$, belongs to $\mathcal{U}_j^k$ for some $j = 1, \ldots, N_k$. Then for any edge from $\mathcal{U}_j^k$ to $\mathcal{U}_{j'}^{k+1}$ in the directed network, $\boldsymbol{u}_{k+1} \in \mathcal{U}_{j'}^{k+1}$ with probability $p_{j,j'}^k$, which is an observable event in Stage $(k+2)$. In other words, in Stage $(k+2)$, we can observe which state transition happened in stage $(k+1)$. Therefore, at every stage, we know the realizations of the uncertain parameters in past stages as well as the path of the uncertainty evolution in G and the decisions in each stage depend on both of these. Note that since we also observe the path of the uncertainty evolution (i.e. what edges in the directed network were realized in each stage), the uncertainty sets in a given stage need not be disjoint. This model of uncertainty is a generalization of the scenario tree model often used in stochastic programming (see Shapiro et al. [76]) where the multi-stage uncertainty is described by a tree. If the directed acyclic network G in our model is a tree and each uncertainty set is a singleton, our model reduces to a scenario tree model described in [76].

The evolution of the multi-stage uncertainty is illustrated in Figure 3-1. We would like to note that this is a very general model of multi-stage uncertainty. For instance, consider a multi-period inventory management problem, where the demand is uncertain. In each stage, we observe the realized demand and also a signal from the market about the next period demand such as the weekly retail sales index. In our model, the observed market signals correspond to the path in the multi-stage uncertainty network and the observed demand is the actual realization of the uncertain parameters. As another example, consider a multi-period asset management problem with uncertain asset returns. In each period, we observe the asset returns and also a market signal such as S&P 500 or NASDAQ indices. Again, the market signals correspond to the path in the multi-stage uncertainty network in our model and observed asset returns are the realization of uncertain parameters in the model.

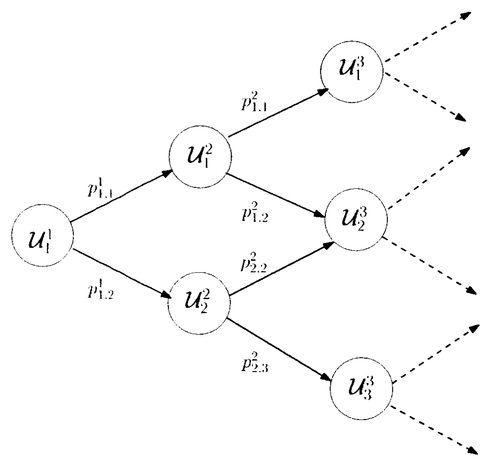

Figure 3-1: Illustration of the evolution of uncertainty in a multi-stage problem.

Let $\mathcal{P}$ denote the set of directed paths in G from the root node, $\mathcal{U}_1^1$, to node $\mathcal{U}_j^{K-1}$, for all $j = 1, \ldots, N_{K-1}$. We denote any path $\boldsymbol{P} \in \mathcal{P}$ as an ordered sequence, $(j_1, \ldots, j_{K-1})$ of the indices of the uncertainty sets in each stage that occur on $\boldsymbol{P}$. Let $\boldsymbol{\Omega}(\boldsymbol{P})$ denote the set of possible realizations of the multi-stage uncertainty from uncertainty sets in $\boldsymbol{P}$. For any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$,

$$\boldsymbol{\Omega}(\boldsymbol{P}) = \left\{ \boldsymbol{\omega} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-1}) \;\middle|\; \boldsymbol{b}_k \in \mathcal{U}_{j_k}^k,\ \forall k = 1, \ldots, K-1 \right\}. \qquad (3.4)$$

For any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$, and $k = 1, \ldots, K-1$, let $\boldsymbol{P}[k]$ denote the index sequence of Path $\boldsymbol{P}$ from Stage 2 to Stage $(k+1)$, i.e., $\boldsymbol{P}[k] = (j_1, \ldots, j_k)$. Let $\mathcal{P}[k] = \{\boldsymbol{P}[k] \mid \boldsymbol{P} \in \mathcal{P}\}$. Also, for any $\boldsymbol{\omega} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-1}) \in \boldsymbol{\Omega}(\boldsymbol{P})$, let $\boldsymbol{\omega}[k]$ denote the subsequence of first k elements of $\boldsymbol{\omega}$, i.e., $\boldsymbol{\omega}[k] = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k)$.

We define a probability measure for the multi-stage uncertainty as follows. For any $k = 1, \ldots, K-1$, $j = 1, \ldots, N_k$, let μ_j^k be a probability measure defined on $\mathcal{U}_j^k$ that is independent of the probability measures over other uncertainty sets in the network. Therefore, for any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$, the probability measure $\mu_{\boldsymbol{P}}$ over the set $\Omega(\boldsymbol{P})$ is the product measure of $\mu_{j_k}^k$, $k = 1, \ldots, K-1$. Moreover, the probability that the uncertain parameters realize from path $\boldsymbol{P}$ is given by $\prod_{k=1}^{K-2} p_{j_k, j_{k+1}}^k$. This defines a probability measure on the set of realizations of the multi-stage uncertainty, $\bigcup_{\boldsymbol{P} \in \mathcal{P}} \Omega(\boldsymbol{P})$.

We can now formulate the K-stage stochastic optimization problem Π_{Stoch}^K, where the right-hand side of the constraints is uncertain, as follows.

$$
\begin{aligned}
z_{\text{Stoch}}^K = \min \ & \boldsymbol{c}^T \boldsymbol{x} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P} \in \mathcal{P}} \left[\mathbb{E}_{\mu_{\boldsymbol{P}}} \left[\boldsymbol{d}_k^T \boldsymbol{y}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) \right] \right] \\
\text{s.t.} \ & \forall \boldsymbol{P} \in \mathcal{P}, \ \mu_{\boldsymbol{P}}\text{-a.e. } \boldsymbol{\omega} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-1}) \in \boldsymbol{\Omega}(\boldsymbol{P}) \\
& \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \boldsymbol{y}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) \geq \sum_{k=1}^{K-1} \boldsymbol{b}_k, \\
& \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}_+^{n_1 - p_1}, \\
& \boldsymbol{y}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) \in \mathbb{R}^{p_k} \times \mathbb{R}_+^{n_k - p_k}, \ \forall k = 1, \ldots, K-1,
\end{aligned}
\tag{3.5}
$$

There is no restriction on the coefficients in the constraint matrices $\boldsymbol{A}, \boldsymbol{B}_k$ or on the objective coefficients $\boldsymbol{c}, \boldsymbol{d}_k, k = 1, \ldots, K-1$. However, we require that each uncertainty set $\mathcal{U}_j^k \subseteq \mathbb{R}_+^m$ for all $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. As mentioned before, we assume that z_{Stoch}^K is finite and the expectation of every feasible multi-stage solution exists.

We also formulate the K-stage adaptive optimization problem Π^K_{Adapt} as follows.

$$
\begin{aligned}
z^K_{\text{Adapt}} = \min \ & \boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{P}\in\mathcal{P},\boldsymbol{\omega}\in\Omega(\boldsymbol{P})} \ \min_{\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]),k=1,\ldots,K-1} \sum_{k=1}^{K-1} \boldsymbol{d}_k^T\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \\
& \text{s.t. } \forall \boldsymbol{P}\in\mathcal{P},\ \forall\boldsymbol{\omega} = (\boldsymbol{b}_1,\ldots,\boldsymbol{b}_{K-1}) \in \boldsymbol{\Omega}(\boldsymbol{P}) \\
& \quad \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{B}_k\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \geq \sum_{k=1}^{K-1}\boldsymbol{b}_k, \\
& \quad \boldsymbol{x}\in\mathbb{R}^{p_1}\times\mathbb{R}_+^{n_1-p_1}, \\
& \quad \boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \in \mathbb{R}^{p_k}\times\mathbb{R}_+^{n_k-p_k},\ \forall k=1,\ldots,K-1.
\end{aligned}
\tag{3.6}
$$

We also formulate the K-stage stochastic optimization problem $\Pi^K_{\text{Stoch}(\boldsymbol{b},\boldsymbol{d})}$, where both the right-hand side and the objective coefficients are uncertain, as follows. The subscript Stoch$(\boldsymbol{b},\boldsymbol{d})$ in the subscript of the problem name denotes that both the right hand side $\boldsymbol{b}$ and the objective coefficient $\boldsymbol{d}$ are uncertain. Here, $\boldsymbol{\omega}$ denotes the sequence of uncertain right-hand side and objective coefficients realizations and for any $k = 1,\ldots,K-1$, $\boldsymbol{\omega}[k]$ denotes the subsequence of first k elements. Also, for any $\boldsymbol{P}\in\mathcal{P}$, the measure $\mu_{\boldsymbol{P}}$ is the product measure of the measures on the uncertainty sets in Path $\boldsymbol{P}$.

$$
\begin{aligned}
z^K_{\text{Stoch}(\boldsymbol{b},\boldsymbol{d})} = \min \quad & \boldsymbol{c}^T\boldsymbol{x} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}\left[\mathbb{E}_{\mu_{\boldsymbol{P}}}\left[\boldsymbol{d}_k^T\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k])\right]\right] \\
& \text{s.t. } \forall\boldsymbol{P}\in\mathcal{P}, \mu_{\boldsymbol{P}}\text{-a.e. } \boldsymbol{\omega} = \big((\boldsymbol{b}_1,\boldsymbol{d}_1),\ldots,(\boldsymbol{b}_{K-1},\boldsymbol{d}_{K-1})\big) \in \boldsymbol{\Omega}(\boldsymbol{P}) \\
& \quad \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{B}_k\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \geq \sum_{k=1}^{K-1}\boldsymbol{b}_k, \\
& \quad \boldsymbol{x}\in\mathbb{R}^{p_1}\times\mathbb{R}_+^{n_1-p_1}, \\
& \quad \boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \in \mathbb{R}^{p_k}\times\mathbb{R}_+^{n_k-p_k},\ \forall k=1,\ldots,K-1.
\end{aligned}
\tag{3.7}
$$

Also, we can formulate the K-stage adaptive problem $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$,

$$\begin{aligned} z^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})} = \min \quad & \boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{P}\in\mathcal{P},\boldsymbol{\omega}\in\Omega(\boldsymbol{P})} \min_{\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]),k=1,\ldots,K-1} \sum_{k=1}^{K-1} \boldsymbol{d}_k^T\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \\ \text{s.t. } & \forall \boldsymbol{P}\in\mathcal{P},\ \forall\boldsymbol{\omega} = \big((\boldsymbol{b}_1,\boldsymbol{d}_1),\ldots,(\boldsymbol{b}_{K-1},\boldsymbol{d}_{K-1})\big)\in\boldsymbol{\Omega} \\ & \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{B}_k\boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k]) \geq \sum_{k=1}^{K}\boldsymbol{b}_k, \\ & \boldsymbol{x}\in\mathbb{R}^{p_1}\times\mathbb{R}^{n_1-p_1}_+, \\ & \boldsymbol{y}_k(\boldsymbol{\omega}[k],\boldsymbol{P}[k])\in\mathbb{R}^{p_k}\times\mathbb{R}^{n_k-p_k}_+,\ \forall k=1,\ldots,K-1. \end{aligned} \tag{3.8}$$

In Section 3.7, we consider an extension to the case where the constraints are general linear conic inequalities and the right hand side uncertainty set is a convex and compact subset of the underlying cone. Furthermore, in Section 3.8, we also consider extensions of the above two-stage and multi-stage models where some decision variables in each stage are integer.

3.2.3 Examples

In this section, we show two classical problems that can be formulated in our framework to illustrate the applicability of our models.

Multi-period inventory management problem. We show that we can model the classical multi-period inventory management problem as a special case of (3.5). In a classical single-item inventory management problem, the goal in each period is to decide on the quantity of the item to order under an uncertain future period demand. In each period, each unit of excess inventory incurs a holding cost and each unit of backlogged demand incurs a per-unit penalty cost and the goal is to make ordering decisions such that the sum of expected holding and backorder-penalty cost is minimized.

As an example, we model a 3-stage problem in our framework. Let $\boldsymbol{\Omega}$ denote the set of demand scenarios with a probability measure μ, x denote the initial inventory, $y_k(b_1,\ldots,b_k)$ denote the backlog and $z_k(b_1,\ldots,b_k)$ denote the order quantity in Stage

$(k+1)$ when the first k-period demand is $(b_1, \dots, b_k)$. Let h_k denote the per unit holding cost and p_k denote the per-unit backlog penalty in Stage $(k+1)$. We model the 3-stage inventory management problem as follows where the decision variables are x, $y_1(b_1), y_2(b_1, b_2)$ and $z_1(b_1)$ for all $(b_1, b_2) \in \mathcal{U}$.

$$\begin{aligned}
\min \quad & \mathbb{E}_{b_1}\Big[h_1\big(x + y_1(b_1) - b_1\big) + p_1 y_1(b_1) + \mathbb{E}_{b_2|b_1}\big[h_2\big(x + z_1(b_1) + y_2(b_1, b_2) - (b_1 + b_2)\big) \\
& \quad + p_2 y_2(b_1, b_2)\big]\Big] \\
\text{s.t.} \quad & x + y_1(b_1) \geq b_1, \ \mu\text{-a.e. } (b_1, b_2) \in \Omega \\
& x + z_1(b_1) + y_2(b_1, b_2) \geq b_1 + b_2, \ \mu\text{-a.e. } (b_1, b_2) \in \Omega \\
& x, y_1(b_1), y_2(b_1, b_2), z_1(b_1) \geq 0.
\end{aligned}$$

The above formulation can be generalized to a multi-period problem in a straightforward manner. We can also generalize to multi-item variants of the problem. However, we would like to note that we can not model a capacity constraint on the order quantity, since we require that the constraints are of "greater than or equal to" form with nonnegative right-hand sides. Nevertheless, the formulation is fairly general even with this restriction.

Capacity planning under demand uncertainty. In many important applications, we encounter the problem of planning for capacity to serve an uncertain future demand. For instance, in a call center, staffing capacity decisions need to be made well in advance of the realization of the uncertain demand. In electricity grid operations, generator unit-commitment decisions are required to be made at least a day ahead of the realization of the uncertain electricity demand because of the large startup time of the generators. In a facility location problem with uncertain future demand, the facilities need to be opened well in advance of the realized demand. Therefore, the capacity planning problem under demand uncertainty is important and widely applicable.

We show that this can be modeled in our framework using the example of the facility location problem. For illustration, we use a 2-stage problem. Let $\mathcal{F}$ denote the set of facilities and $\mathcal{D}$ denote the set of demand points. For each facility $i \in \mathcal{F}$, let x_i be an integer decision variable that denotes the capacity for facility i. For each point $j \in \mathcal{D}$, let b_j denote the uncertain future demand and let $y_{ij}(\boldsymbol{b})$ denote the amount of demand of j assigned to facility i when the demand vector is $\boldsymbol{b}$. Also, let d_{ij} denote the cost of assigning a unit demand from point j to facility i and let c_i denote the per-unit capacity cost of facility i. Therefore, we can formulate the problem as follows. Let $\mathcal{U}$ be the uncertainty set for the demand and let μ be a probability measure defined on $\mathcal{U}$.

$$
\begin{aligned}
\min \quad & \sum_{i \in \mathcal{F}} c_i x_i + \mathbb{E}_\mu \left[\sum_{j \in \mathcal{D}} d_{ij} y_{ij}(\boldsymbol{b}) \right] \\
\text{s.t.} \quad & \sum_{i \in \mathcal{F}} y_{ij}(\boldsymbol{b}) \geq b_j, \ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}, \ \forall j \in \mathcal{D} \\
& x_i - \sum_{j \in \mathcal{D}} y_{ij}(\boldsymbol{b}) \geq 0, \ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}, \ \forall i \in \mathcal{F} \\
& x_i \in \mathbb{Z}_+, \ \forall i \in \mathcal{F} \\
& y_{ij}(\boldsymbol{b}) \in \mathbb{R}_+, \ \forall i \in \mathcal{F}, j \in \mathcal{D}, \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}.
\end{aligned}
$$

3.2.4 Finitely Adaptable Solutions

We consider a *finitely adaptable* class of solutions for the multi-stage stochastic and adaptive optimization problems described above. This class of solutions was introduced by Bertsimas and Caramanis [16] where the decision-maker computes a small set of solutions in each stage apriori.

A static robust solution policy specifies a single solution that is feasible for all possible realizations of the uncertain parameters. On the other extreme, a fully-adaptable solution policy specifies a solution for each possible realization of the uncertain parameters in past stages. Typically, the set of possible realizations of the uncertain parameters is uncountable which implies that an optimal fully-adaptable solution policy is a function from an uncountable set of scenarios to optimal decisions for each scenario and often suffers from the "curse of dimensionality". A finitely adaptable solution is a tractable approach that bridges the gap between a static robust solution and a fully-adaptable solution. In a general finitely adaptable solution policy, instead of computing an optimal decision for each scenario, we partition the scenarios into a small number of sets and compute a single solution for each possible set. The partitioning of the set of scenarios is problem specific and is chosen by the decision-maker. A finitely adaptable solution policy is a special case of a piecewise affine solution where the solution in each piece is a static solution.

For the multi-stage stochastic and adaptive problems (3.5) and (3.6), we consider the partition of the set of scenarios based on the realized path in the multi-stage uncertainty network. In particular, in each Stage $(k+1)$, the decision $\boldsymbol{y}_k$ depends on the path of the uncertainty realization until Stage $(k+1)$. Therefore, there are $|\mathcal{P}[k]|$ different solutions for each Stage $(k+1)$, $k = 1, \ldots, K-1$; one corresponding to each directed path from the root node to a node in Stage $(k+1)$ in the multi-stage uncertainty network. For any realization of uncertain parameters and the path in the uncertainty network, the solution policy implements the solution corresponding to the realized path. Figure 3-2 illustrates the number of solutions in the finitely adaptable solution for each stage and each uncertainty set in a multi-stage uncertainty network. In the example in Figure 3-2, there are following four directed paths from Stage 2 to Stage K $(K = 4)$: $\boldsymbol{P}_1 = (1,1,1), \boldsymbol{P}_2 = (1,1,2), \boldsymbol{P}_3 = (1,2,2), \boldsymbol{P}_4 = (1,2,3)$. We know that $\boldsymbol{P}_j[1] = (1)$ for all $j = 1, \ldots, 4$. Also, $\boldsymbol{P}_1[2] = \boldsymbol{P}_2[2] = (1,1)$ and $\boldsymbol{P}_3[2] = \boldsymbol{P}_4[2] = (1,2)$. Therefore, in a finitely adaptable solution, we have the following decision variables apart from $\boldsymbol{x}$.

$$\begin{aligned}
\text{Stage 2} &: \boldsymbol{y}_1(1) \\
\text{Stage 3} &: \boldsymbol{y}_2(1,1), \boldsymbol{y}_2(1,2) \\
\text{Stage 4} &: \boldsymbol{y}_3(1,1,1), \boldsymbol{y}_3(1,1,2), \boldsymbol{y}_3(1,2,2), \boldsymbol{y}_3(1,2,3).
\end{aligned}$$

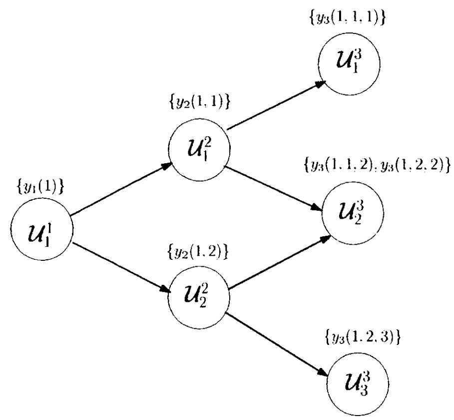

Figure 3-2: A finitely adaptable solution for a 4-stage problem. For each uncertainty set, we specify the set of corresponding solutions in the finitely adaptable solution policy.

3.2.5 The Symmetry of a Convex Set

Given a nonempty compact convex set $\mathcal{U} \subset \mathbb{R}^m$ and a point $\boldsymbol{u} \in \mathcal{U}$, we define the symmetry of $\boldsymbol{u}$ with respect to $\mathcal{U}$ as follows.

$$\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) := \max\{\alpha \geq 0 : \boldsymbol{u} + \alpha(\boldsymbol{u} - \boldsymbol{u}') \in \mathcal{U}, \ \forall \boldsymbol{u}' \in \mathcal{U}\}. \tag{3.9}$$

In order to develop a geometric intuition on $\mathbf{sym}(\boldsymbol{u}, \mathcal{U})$, we first show the following result. For this discussion, assume $\mathcal{U}$ is full-dimensional.

Lemma 1. *Let $\mathcal{L}$ be the set of lines in $\mathbb{R}^m$ passing through $\boldsymbol{u}$. For any line $\boldsymbol{l} \in \mathcal{L}$, let $\boldsymbol{u}_l'$ and $\boldsymbol{u}_l''$ be the points of intersection of $\boldsymbol{l}$ with the boundary of $\mathcal{U}$, denoted $\delta(\mathcal{U})$ (these exist as $\mathcal{U}$ is full-dimensional and compact). Then,*

a) $\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) \leq \min\left(\dfrac{\|\boldsymbol{u} - \boldsymbol{u}_l''\|}{\|\boldsymbol{u} - \boldsymbol{u}_l'\|}, \dfrac{\|\boldsymbol{u} - \boldsymbol{u}_l'\|}{\|\boldsymbol{u} - \boldsymbol{u}_l''\|}\right)$,

b) $\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) = \min\limits_{l \in \mathcal{L}} \min\left(\dfrac{\|\boldsymbol{u} - \boldsymbol{u}_l''\|}{\|\boldsymbol{u} - \boldsymbol{u}_l'\|}, \dfrac{\|\boldsymbol{u} - \boldsymbol{u}_l'\|}{\|\boldsymbol{u} - \boldsymbol{u}_l''\|}\right)$.

Proof. Let $\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) = \alpha_1$. Consider any $\boldsymbol{l} \in \mathcal{L}$ and consider $\boldsymbol{u}_l' \in \delta(\mathcal{U})$ as illustrated in Figure 3-3. Let $\boldsymbol{u}_{l,r}' = \boldsymbol{u} + \alpha_1(\boldsymbol{u} - \boldsymbol{u}_l')$. From (3.9), we know that $\boldsymbol{u}_{l,r}' \in \mathcal{U}$. Note that $\boldsymbol{u}_{l,r}'$ is a *scaled reflection* of $\boldsymbol{u}_l'$ about $\boldsymbol{u}$ by a factor α_1. Furthermore, $\boldsymbol{u}_{l,r}'$ lies on

$\boldsymbol{l}$. Therefore, $\boldsymbol{u}'_{l,r} \in \mathcal{U}$ implies that

$$\|\boldsymbol{u} - \boldsymbol{u}'_{l,r}\| \leq \|\boldsymbol{u} - \boldsymbol{u}''_l\| \Rightarrow \alpha_1 \cdot \|\boldsymbol{u} - \boldsymbol{u}'_l\| \leq \|\boldsymbol{u} - \boldsymbol{u}''_l\|,$$

which in turn implies that

$$\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) = \alpha_1 \leq \frac{\|\boldsymbol{u} - \boldsymbol{u}''_l\|}{\|\boldsymbol{u} - \boldsymbol{u}'_l\|}.$$

Using a similar argument starting with $\boldsymbol{u}''_l$ instead of $\boldsymbol{u}'_l$, we obtain that

$$\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) \leq \frac{\|\boldsymbol{u} - \boldsymbol{u}'_l\|}{\|\boldsymbol{u} - \boldsymbol{u}''_l\|}.$$

To prove the second part, let

$$\alpha_2 = \min_{l \in \mathcal{L}} \min \left(\frac{\|\boldsymbol{u} - \boldsymbol{u}''_l\|}{\|\boldsymbol{u} - \boldsymbol{u}'_l\|}, \frac{\|\boldsymbol{u} - \boldsymbol{u}'_l\|}{\|\boldsymbol{u} - \boldsymbol{u}''_l\|} \right).$$

From the above argument, we know that $\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) \leq \alpha_2$. Consider any $\boldsymbol{u}' \in \mathcal{U}$. We show that $(\boldsymbol{u} + \alpha_2(\boldsymbol{u} - \boldsymbol{u}')) \in \mathcal{U}$.

Let $\boldsymbol{l}$ denote the line joining $\boldsymbol{u}$ and $\boldsymbol{u}'$. Clearly, $\boldsymbol{l} \in \mathcal{L}$. Without loss of generality, suppose $\boldsymbol{u}'$ belongs to the line segment between $\boldsymbol{u}$ and $\boldsymbol{u}'_l$. Therefore, $\boldsymbol{u} - \boldsymbol{u}' = \gamma(\boldsymbol{u} - \boldsymbol{u}'_l)$ for some $0 \leq \gamma \leq 1$ and

$$\|\boldsymbol{u} - \boldsymbol{u}'\| \leq \|\boldsymbol{u} - \boldsymbol{u}'_l\|. \tag{3.10}$$

We know that

$$\begin{aligned}
&\alpha_2 \leq \frac{\|\boldsymbol{u} - \boldsymbol{u}''_l\|}{\|\boldsymbol{u} - \boldsymbol{u}'_l\|} \\
&\Rightarrow \alpha_2 \|\boldsymbol{u} - \boldsymbol{u}'_l\| \leq \|\boldsymbol{u} - \boldsymbol{u}''_l\| \\
&\Rightarrow \alpha_2 \|\boldsymbol{u} - \boldsymbol{u}'\| \leq \|\boldsymbol{u} - \boldsymbol{u}''_l\|,
\end{aligned}$$

where the last inequality follows from (3.10). Therefore, $\boldsymbol{u} + \alpha_2\|\boldsymbol{u} - \boldsymbol{u}'\|$ belongs to the line segment between $\boldsymbol{u}$ and $\boldsymbol{u}''_l$ and thus, belongs to $\mathcal{U}$. Therefore, $\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) \geq \alpha_2$. □

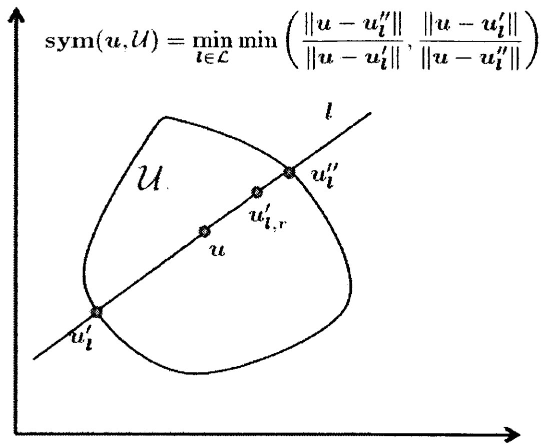

Figure 3-3: Geometric illustration of $\mathbf{sym}(\boldsymbol{u}, \mathcal{U})$.

The symmetry of set $\mathcal{U}$ is defined as

$$\mathbf{sym}(\mathcal{U}) := \max\{\mathbf{sym}(\boldsymbol{u}, \mathcal{U}) \mid \boldsymbol{u} \in \mathcal{U}\}. \tag{3.11}$$

An optimizer $\boldsymbol{u}_0$ of (3.11) is called a point of symmetry of $\mathcal{U}$. This definition of symmetry can be traced back to Minkowski [62] and is the first and the most widely used symmetry measure. Refer to Belloni and Freund [5] for a broad investigation of the properties of the symmetry measure defined in (3.9) and (3.11). Note that the above definition generalizes the notion of perfect symmetry considered by Bertsimas and Goyal [18]. In [18], the authors define that a set $\mathcal{U}$ is symmetric if there exists $\boldsymbol{u}_0 \in \mathcal{U}$ such that, for any $\boldsymbol{z} \in \mathbb{R}^m$, $(\boldsymbol{u}_0 + \boldsymbol{z}) \in \mathcal{U} \Leftrightarrow (\boldsymbol{u}_0 - \boldsymbol{z}) \in \mathcal{U}$. Equivalently, $\boldsymbol{u} \in \mathcal{U} \Leftrightarrow (2\boldsymbol{u}_0 - \boldsymbol{u}) \in \mathcal{U}$. According to the definition in (3.11), $\mathbf{sym}(\mathcal{U}) = 1$ for such a set. Figure 3-4 illustrates symmetries of several interesting convex sets.

Lemma 2 (Belloni and Freund [5]). *For any nonempty convex compact set $\mathcal{U} \subseteq \mathbb{R}^m$, the symmetry of $\mathcal{U}$ satisfies,*

$$\frac{1}{m} \leq \mathbf{sym}(\mathcal{U}) \leq 1.$$

The symmetry of a convex set is at most 1, which is achieved for a perfectly symmetric set; and at least $1/m$, which is achieved by a standard simplex defined

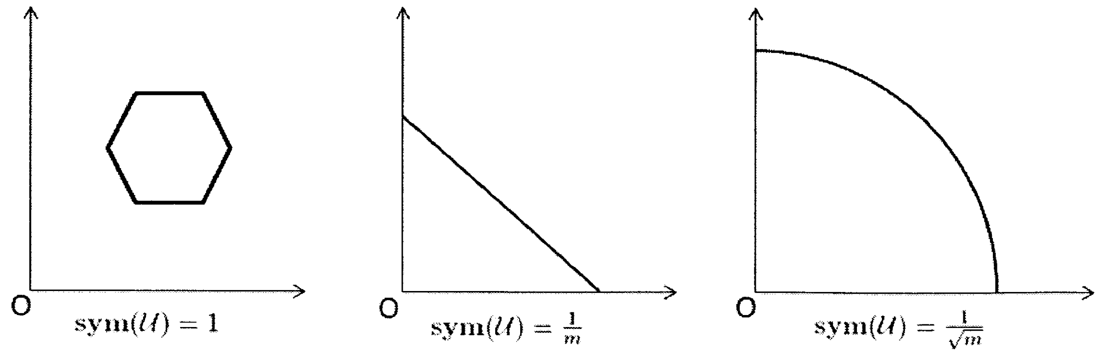

Figure 3-4: The figure on the left is a symmetric polytope with symmetry 1. The middle figure illustrates a standard simplex in $\mathbb{R}^m$ with symmetry $1/m$. The right figure shows the intersection of a Euclidean ball with $\mathbb{R}^m_+$, which has symmetry $1/\sqrt{m}$.

as $\Delta = \{\boldsymbol{x} \in \mathbb{R}^m_+ \mid \sum_{i=1}^m x_i \leq 1\}$. The lower bound follows from Löwner-John Theorem [52] (see Belloni and Freund [5]). The following lemma is used later.

Lemma 3. *Let $\mathcal{U} \subset \mathbb{R}^m_+$ be a convex and compact set such that $\boldsymbol{u}_0$ is the point of symmetry of $\mathcal{U}$. Then,*

$$\left(1 + \frac{1}{\mathbf{sym}(\mathcal{U})}\right) \cdot \boldsymbol{u}_0 \geq \boldsymbol{u}, \ \forall \boldsymbol{u} \in \mathcal{U}.$$

Proof. From the definition of symmetry in (3.11), we have that for any $\boldsymbol{u} \in \mathcal{U}$,

$$\boldsymbol{u}_0 + \mathbf{sym}(\mathcal{U})(\boldsymbol{u}_0 - \boldsymbol{u}) = (\mathbf{sym}(\mathcal{U}) + 1)\boldsymbol{u}_0 - \mathbf{sym}(\mathcal{U})\boldsymbol{u} \in \mathcal{U},$$

which implies that $(\mathbf{sym}(\mathcal{U}) + 1)\boldsymbol{u}_0 - \mathbf{sym}(\mathcal{U})\boldsymbol{u} \geq \mathbf{0}$ since $\mathcal{U} \subset \mathbb{R}^m_+$. □

3.2.6 The Translation Factor $\rho(\boldsymbol{u}, \mathcal{U})$

For a convex compact set $\mathcal{U} \subset \mathbb{R}^m_+$, we define a *translation factor* $\rho(\boldsymbol{u}, \mathcal{U})$, the *translation factor* of $\boldsymbol{u} \in \mathcal{U}$ with respect to $\mathcal{U}$, as follows.

$$\rho(\boldsymbol{u}, \mathcal{U}) = \min\{\alpha \in \mathbb{R}_+ \mid \mathcal{U} - (1 - \alpha) \cdot \boldsymbol{u} \subset \mathbb{R}^m_+\}.$$

In other words, $\mathcal{U}' := \mathcal{U} - (1-\rho)\boldsymbol{u}$ is the maximum possible translation of $\mathcal{U}$ in the direction $-\boldsymbol{u}$ such that $\mathcal{U}' \subset \mathbb{R}^m_+$. Figure 3-5 gives a geometric picture. Note that for $\alpha = 1$, $\mathcal{U} - (1-\alpha)\cdot \boldsymbol{u} = \mathcal{U} \subset \mathbb{R}^m_+$. Therefore, $0 < \rho \leq 1$. And ρ approaches 0, when the set $\mathcal{U}$ moves away from the origin. If there exists $\boldsymbol{u} \in \mathcal{U}$ such that $\boldsymbol{u}$ is at the boundary of $\mathbb{R}^m_+$, then $\rho = 1$. We denote

$$\rho(\mathcal{U}) := \rho(\boldsymbol{u}_0, \mathcal{U}),$$

where $\boldsymbol{u}_0$ is the symmetry point of the set $\mathcal{U}$. The following lemma is used later in

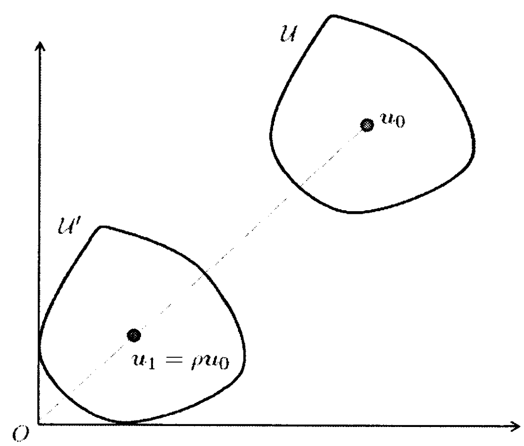

Figure 3-5: Geometry of the translation factor.

the chapter.

Lemma 4. *Let $\mathcal{U} \subset \mathbb{R}^m_+$ be a convex and compact set such that $\boldsymbol{u}_0$ is the point of symmetry of $\mathcal{U}$. Let $s = \mathbf{sym}(\mathcal{U}) = \mathbf{sym}(\boldsymbol{u}_0, \mathcal{U})$ and $\rho = \rho(\mathcal{U}) = \rho(\boldsymbol{u}_0, \mathcal{U})$. Then,*

$$\left(1 + \frac{\rho}{s}\right) \cdot \boldsymbol{u}_0 \geq \boldsymbol{u}, \; \forall \boldsymbol{u} \in \mathcal{U}.$$

Proof. Let $\mathcal{U}' = \mathcal{U} - (1-\rho)\boldsymbol{u}_0$. Let $\boldsymbol{u}_1 := \boldsymbol{u}_0 - (1-\rho)\boldsymbol{u}_0$. Also let $\boldsymbol{z} := \boldsymbol{u}_0 - \boldsymbol{u}_1 = (1-\rho)\boldsymbol{u}_0$. Figure 3-5 gives a geometric picture. Note that $\mathbf{sym}(\mathcal{U}') = \mathbf{sym}(\mathcal{U}) = s$.

From Lemma 3, we know that

$$\left(1+\frac{1}{s}\right)\boldsymbol{u}_1 \geq \boldsymbol{u}', \quad \forall \boldsymbol{u}' \in \mathcal{U}'.$$

Adding $\boldsymbol{z}$ on both sides, we have,

$$\begin{aligned}
&\left(1+\frac{1}{s}\right)\boldsymbol{u}_1 + \boldsymbol{z} \geq \boldsymbol{u}, \quad \forall \boldsymbol{u} \in \mathcal{U},\\
\Rightarrow &\left(1+\frac{1}{s}\right)\rho\boldsymbol{u}_0 + (1-\rho)\boldsymbol{u}_0 \geq \boldsymbol{u}, \quad \forall \boldsymbol{u} \in \mathcal{U},\\
\Rightarrow &\left(1+\frac{\rho}{s}\right)\boldsymbol{u}_0 \geq \boldsymbol{u}, \quad \forall \boldsymbol{u} \in \mathcal{U}.
\end{aligned}$$

□

3.3 Our Contributions

Our contributions are two-fold. We present two significant generalizations of the model and results in Bertsimas and Goyal [18], where the authors characterize the performance of static robust solutions for two-stage stochastic and adaptive optimization problems under the assumption that the uncertainty sets are perfectly symmetric.

Firstly, we generalize the two-stage results to general uncertainty sets. We show that the performance of a static robust solution for two-stage stochastic and adaptive optimization problems depends on the general notion of symmetry (3.11) of the uncertainty set. The bounds are independent of the constraint matrices and any other problem data. Our bounds are also tight for all possible values of symmetry and reduce to the results in [18] for perfectly symmetric sets.

Secondly, we consider the multi-stage extensions of the two-stage models in [18] as described above and show that a class of finitely adaptable solutions which is a generalization of the static robust solution, is a good approximation for both the stochastic and the adaptive problem. The proof techniques are very general and easily extend to the case where some of the decision variables are integer constrained and the case where the constraints are linear conic inequalities for a general convex cone. To the best of our knowledge, these are the first performance bounds for the multi-stage problem in such generality.

Our main contributions are summarized below.

Stochastic optimization. For the two-stage stochastic optimization problem under right-hand side uncertainty, we show the following bound under a fairly general condition on the probability measure,

$$z_{\text{Rob}} \leq \left(1 + \frac{\rho}{s}\right) \cdot z_{\text{Stoch}},$$

where $s = \mathbf{sym}(\mathcal{U})$ and $\rho = \rho(\mathcal{U})$ are the symmetry and the translation factor of the uncertainty set, respectively. Note that the above bound compares the cost of an optimal static solution with the expected cost of an optimal fully-adaptable two-stage stochastic solution and shows that the static cost is at most $(1 + \rho/s)$ times the optimal stochastic cost. The performance of the static robust solution for the two-stage stochastic problem can possibly be even better. For the two-stage problem, we only implement the first-stage part of the static robust solution. For the second-stage, we compute an optimal solution after the uncertain parameters (right hand side in our case) is realized. Therefore, the expected second-stage cost for this solution policy is at most the second-stage cost of the static robust solution and the total expected cost is at most $z_{\text{Rob}} \leq (1 + \rho/s) \cdot z_{\text{Stoch}}$. Since z_{Rob} is an upper bound on the expected cost of the solution obtained from an optimal static solution, the bound obtained by comparing z_{Rob} to the optimal stochastic cost is in fact a conservative bound.

For multi-stage problems, we show that a K-stage stochastic optimization problem, Π^K_{Stoch}, can be well approximated efficiently by a finitely adaptable solution. In particular, there is a finitely adaptable solution with at most $|\mathcal{P}|$ solutions that is a $(1 + \rho/s)$-approximation of the original problem where s is the minimum symmetry over all sets in the uncertainty network and ρ is the maximum translation factor of any uncertainty set. Note that when all sets are perfectly symmetric, i.e., $s = 1$, the finitely adaptable solution is a 2-approximation, generalizing the result of [18] to multi-stage. We also show that the bound of $(1 + \rho/s)$ is tight.

Note that since $0 < \rho \leq 1$ and $s \geq 1/m$,

$$\left(1 + \frac{\rho}{s}\right) \leq (m+1),$$

which shows that, for a fixed dimension, the performance bound for robust and finitely adaptable solutions is *independent* of the particular data of an instance. This is surprising since when the left-hand side has uncertainty, i.e., $\boldsymbol{A}, \boldsymbol{B}$ are uncertain, even a two-stage two-dimensional problem can have an unbounded gap between the static robust solution and the stochastic solution as mentioned earlier. This also indicates that our assumptions on the model, namely, the right-hand side uncertainty (and/or cost uncertainty) and the uncertainty set contained in the nonnegative orthant, are tight. If these assumptions are relaxed, the performance gap becomes unbounded even for fixed dimensional uncertainty.

For the case when both cost and right-hand side are uncertain in $\Pi^K_{\text{Stoch}(\boldsymbol{b},\boldsymbol{d})}$, the performance of any finitely adaptable solution can be arbitrarily worse as compared to the optimal fully-adaptable stochastic solution. This result follows along the lines of arbitrary bad performance of a static robust solution in two-stage stochastic problems when both cost and right-hand sides are uncertain as shown in Bertsimas and Goyal [18].

Adaptive optimization. We show that for a multi-stage adaptive optimization problem, Π^K_{Adapt}, where only the right-hand side of the constraints is uncertain, the cost of a finitely adaptable solution is at most $(1 + \rho/s)$ times the optimal cost of a fully-adaptable multi-stage solution, where s is the minimum symmetry of all sets in the uncertainty network and ρ is the maximum translation factor of the point of symmetry over all the uncertainty sets. This bound follows from the bound for the performance of a finitely adaptable solution for the multi-stage stochastic problem. Furthermore, if the uncertainty comes from hypercube sets, then a finitely adaptable solution with at most $|\mathcal{P}|$ solutions at each node of the uncertainty network is an optimal solution for the adaptive problem.

For the case when both cost and right-hand side are uncertain in $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$, we show that the worst-case cost a finitely adaptable solution with at most $|\mathcal{P}|$ different solutions at each node of the uncertainty network is at most $(1+\rho/s)^2$ times the cost of an optimal fully-adaptable solution.

Extensions. We consider an extension of the above multi-stage models to the case where the constraints are linear conic inequalities and the uncertainty set belongs to the underlying cone. We also consider the case where some of the decision variables are constrained to be integers. Our proof techniques are quite general and the results extend to both these cases.

For the case of linear conic inequalities and the uncertainty set contained in the underlying convex cone, we show that a finitely adaptable (static robust) solution is a $(1+\rho/s)$-approximation for the multi-stage (two-stage respectively) stochastic and adaptive problem with right hand side uncertainty. The result also holds for the adaptive problem with both the right hand side and objective coefficient uncertainty and the performance bound for the finitely adaptable (static robust) solution is $(1+\rho/s)^2$ for the multi-stage (two-stage respectively) problem.

We also consider the case where some of the decision variables are integer constrained. For the multi-stage (two-stage) stochastic problem with right hand side uncertainty, if some of the first-stage decision variables are integer constrained, a finitely adaptable (static robust respectively) solution is a $\lceil(1+\rho/s)\rceil$ approximation with respect to an optimal fully-adaptable solution. For the multi-stage adaptive problem, we can handle integer constrained decision variables in all stages unlike the stochastic problem where we can handle integrality constraints only on the first stage decision variables. We show that for the multi-stage (two-stage) adaptive problem with right hand side uncertainty and integrality constraints on some of the decision variables in each stage, a finitely adaptable (static robust respectively) solution is a $\lceil(1+\rho/s)\rceil$-approximation. For the multi-stage (two-stage) adaptive problem with both right hand side and objective coefficient uncertainty and integrality constraints on variables, a finitely adaptable (static robust respectively) solution is a $\lceil(1+\rho/s)\rceil\cdot(1+\rho/s)$-approximation.

Outline. The rest of the chapter is organized as follows. In Section 3.4, we present the performance bound of a static-robust solution that depends on the symmetry of the uncertainty set for the two-stage stochastic optimization problem. We also show that the bound is tight and present explicit bounds for several interesting and commonly used uncertainty sets. In Section 3.5, we present the finitely adaptable solution policy for the multi-stage stochastic optimization problem and discuss its performance bounds. In Section 3.6, we discuss the results for multi-stage adaptive optimization problems. In Sections 3.7 and 3.8, we present extensions of our results for the models with linear conic constraints and integrality constraints respectively.

3.4 Two-stage Stochastic Optimization Problem

In this section, we consider the two-stage stochastic optimization problem (3.1) and show that the performance of a static robust solution depends on the symmetry of the uncertainty set.

Theorem 1. *Consider the two-stage stochastic optimization problem in* (3.1)*. Let* μ *be the probability measure on the uncertainty set* $\mathcal{U} \subset \mathbb{R}^m_+$*,* $\boldsymbol{b}_0$ *be the point of symmetry of* $\mathcal{U}$*, and* $\rho = \rho(\boldsymbol{b}_0, \mathcal{U})$ *be the translation factor of* $\boldsymbol{b}_0$ *with respect to* $\mathcal{U}$*. Denote* $s = \mathbf{sym}(\mathcal{U})$*. Assume the probability measure* μ *satisfies,*

$$\mathbb{E}_\mu[\boldsymbol{b}] \geq \boldsymbol{b}_0. \tag{3.12}$$

Then,

$$z_{\text{Rob}} \leq \left(1 + \frac{\rho}{s}\right) \cdot z_{\text{Stoch}}. \tag{3.13}$$

Proof. From Lemma 4, we know that

$$\left(1 + \frac{\rho}{s}\right)\boldsymbol{b}_0 \geq \boldsymbol{b}, \quad \forall \boldsymbol{b} \in \mathcal{U}. \tag{3.14}$$

For brevity, let $\tau := (1 + \rho/s)$. Suppose $(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{b}) : \boldsymbol{b} \in \mathcal{U})$ is an optimal fully-adaptable stochastic solution. We first show that the solution, $(\tau\boldsymbol{x}, \tau\mathbb{E}[\boldsymbol{y}(\boldsymbol{b})])$, is a feasible solution to the robust problem (3.3). By the feasibility of $(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{b})$ μ-a.e.

$\boldsymbol{b} \in \mathcal{U}$), we have

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}(\tau \boldsymbol{y}(\boldsymbol{b})) \;\; \geq \;\; \tau \boldsymbol{b}, \; \mu - a.e. \boldsymbol{b} \in \mathcal{U}.$$

Taking expectation on both sides, we have

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}\big(\tau \mathbb{E}[\boldsymbol{y}(\boldsymbol{b})]\big) \geq \tau \mathbb{E}[\boldsymbol{b}] \geq \tau \boldsymbol{b}_0 \geq \boldsymbol{b}, \; \forall \boldsymbol{b} \in \mathcal{U},$$

where the second inequality follows from (3.12) and the last inequality follows from (3.14). Therefore, $(\tau \boldsymbol{x}, \tau \mathbb{E}[\boldsymbol{y}(\boldsymbol{b})])$ is a feasible solution for the robust problem (3.3) and,

$$z_{\text{Rob}} \leq \tau \cdot (\boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \mathbb{E}[\boldsymbol{y}(\boldsymbol{b})]). \tag{3.15}$$

Also, by definition, $z_{\text{Stoch}} = \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \mathbb{E}_{\boldsymbol{b}}[\boldsymbol{y}(\boldsymbol{b})]$, which implies that $z_{\text{Stoch}} \leq \tau \cdot z_{\text{Rob}}$ $\square$

For brevity, we refer to (3.13) as the symmetry bound. The following comments are in order.

1. The symmetry bound (3.13) is *independent* of the problem data $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{c}, \boldsymbol{d}$, depending only on the geometric properties of the uncertainty set $\mathcal{U}$, namely the symmetry and the translation factor of $\mathcal{U}$.

2. Theorem 1 can also be stated in a more general way, removing Assumption (3.12) on the probability measure, and use the symmetry of the expectation point. In particular, the bound (3.13) holds for $s = \mathbf{sym}(\mathbb{E}_\mu[\boldsymbol{b}], \mathcal{U})$ and $\rho = \rho(\mathbb{E}_\mu[\boldsymbol{b}], \mathcal{U})$.

   However, we would like to note that (3.12) is a mild assumption, especially for symmetric uncertainty sets. Any symmetric probability measure on a symmetric uncertainty set satisfies this assumption. It also emphasizes the role that the symmetry of the uncertainty set plays in the bound.

3. As already mentioned in Section 3.3, a small relaxation from the assumptions of our model would cause unbounded performance gap. In particular, if the

assumption, $\mathcal{U} \subset \mathbb{R}^m_+$, is relaxed, or the constraint coefficients are uncertain, the gap between z_{Rob} and z_{Stoch} cannot be bounded even in small dimensional problems. The following examples illustrate this fact.

(a) $\mathcal{U} \not\subseteq \mathbb{R}^m_+$: Consider the instance where $m = 1$ and $c = 0, d = 1, A = 0, B = 1$, the uncertainty set $\mathcal{U} = [-1, 1]$, and a uniform distribution on $\mathcal{U}$. The optimal stochastic solution has cost $z_{\text{Stoch}}(b) = 0$, while $z_{\text{Rob}}(b) = 2$. Thus, the gap is unbounded.

(b) $\boldsymbol{A}, \boldsymbol{B}$ uncertain: Consider the following instance taken from Ben-Tal et al. [9],

$$\begin{aligned} \min \quad & x \\ \text{s.t.} \quad & \begin{bmatrix} -\frac{1}{2}b \\ 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ b-2 \end{bmatrix} y(b) \geq \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \forall b \in [0, r], \\ & x, y(b) \geq 0, \forall b \in [0, r], \end{aligned}$$

where $0 < r < 1$. From the constraints, we have that $x \geq 2/(1-r)$. Therefore, the optimal cost of a static solution is at least $2/(1-r)$. However, the optimal stochastic cost is at most 4. For details, refer to [9]. When r approaches 1, the gap tends to infinity.

3.4.1 Tightness of the Bound

In this section, we show that the bound given in Theorem 1 is tight. In particular, we show that for any given symmetry and translation factor, there exist a family of instances of the two-stage stochastic optimization problem such that the bound in Theorem 1 holds with equality.

Theorem 2. *Given any symmetry $1/m \leq s \leq 1$ and translation factor, $0 < \rho \leq 1$, there exist a family of instances such that $z_{\text{Rob}} = (1 + \rho/s) \cdot z_{\text{Stoch}}$.*

Proof. For $p \geq 1$, let

$$\mathbf{B}_p^+ = \{\boldsymbol{b} \in \mathbb{R}^m_+ \mid \|\boldsymbol{b}\|_p \leq 1\}.$$

In Appendix 3.10.1, we show that

$$\mathbf{sym}(\mathbf{B}_p^+) = \left(\frac{1}{m}\right)^{\frac{1}{p}},$$

and the symmetry point is,

$$\boldsymbol{b}_0(\mathbf{B}_p^+) = \frac{1}{m^{1/p}+1}\boldsymbol{e},$$

where $\boldsymbol{b}_0(\mathcal{U})$ denotes the symmetry point for any set $\mathcal{U}$. Also, let $(\mathbf{B}_p^+)' := \mathbf{B}_p^+ + r\boldsymbol{e}$, for some $r \geq 0$. Then, given any symmetry $s \geq 1/m$ and translation factor, $\rho \leq 1$, we can find a $p \geq 1$ and $r \geq 0$ such that

$$s = \left(\frac{1}{m}\right)^{\frac{1}{p}}, \quad \rho = \frac{1}{(m^{1/p}+1)r+1}.$$

Now consider a problem instance where $\boldsymbol{A} = \mathbf{0}, \boldsymbol{B} = \boldsymbol{I}, \boldsymbol{c} = \mathbf{0}, \boldsymbol{d} = \boldsymbol{e}$, and a probability measure whose expectation is at the symmetry point of $(\mathbf{B}_p^+)'$. The optimal static robust solution is $\boldsymbol{y} = (r+1)\boldsymbol{e}$ and the optimal stochastic solution is $\boldsymbol{y}(\boldsymbol{b}) = \boldsymbol{b}$ for all $\boldsymbol{b} \in (\mathbf{B}_p^+)'$. Therefore,

$$z_{\text{Stoch}} = \left(r + \frac{1}{m^{1/p}+1}\right)m, \quad z_{\text{Rob}} = (r+1)m,$$

and,

$$\frac{z_{\text{Rob}}}{z_{\text{Stoch}}} = \frac{r+1}{\left(r + \frac{1}{m^{1/p}+1}\right)} = 1 + \frac{\rho}{s},$$

which shows the bound in Theorem 1 is tight. □

3.4.2 An Alternative Bound

In this section, we present another performance bound on the robust solution for the two-stage stochastic problems, and compare it with the symmetry bound (3.13).

For an uncertainty set $\mathcal{U}$, let $\boldsymbol{b}^h$ as $b_j^h := \max_{\boldsymbol{b}\in\mathcal{U}} b_j$. Also, suppose the probability measure μ satisfies (3.12), i.e., $\mathbb{E}_\mu[\boldsymbol{b}] \geq \boldsymbol{b}_0$, where $\boldsymbol{b}_0$ is the symmetry point of $\mathcal{U}$. Let

$$\theta_s^* := \min\{\theta : \theta \cdot \boldsymbol{b}_0 \geq \boldsymbol{b}^h\}. \tag{3.16}$$

Using an argument similar to the proof of Theorem 1, we can show that the performance gap is at most θ_s^*, where the subscript s stands for stochasticity, i.e.,

$$z_{\text{Rob}} \leq \theta_s^* \cdot z_{\text{Stoch}}. \tag{3.17}$$

Let $s = \mathbf{sym}(\boldsymbol{b}_0, \mathcal{U})$ and $\rho = \rho(\boldsymbol{b}_0, \mathcal{U})$. From Lemma (4), we also know that

$$\left(1 + \frac{\rho}{s}\right) \cdot \boldsymbol{b}_0 \geq \boldsymbol{b},\ \forall \boldsymbol{b} \in \mathcal{U}.$$

Therefore, $\theta_s^* \leq (1 + \rho/s)$. So, θ_s^* is upper bounded by the symmetry bound obtained in Theorem 1. For brevity, we refer to (3.16) as a scaling bound. A geometric picture is given in Figure 3-6. As shown in the figure, $\overline{\boldsymbol{u}}$ is obtained from scaling $\mathbb{E}_\mu[\boldsymbol{b}]$ by a factor greater than one such that $\overline{\boldsymbol{u}}$ dominates all the points in $\mathcal{U}$, and θ_s^* is the smallest such factor.

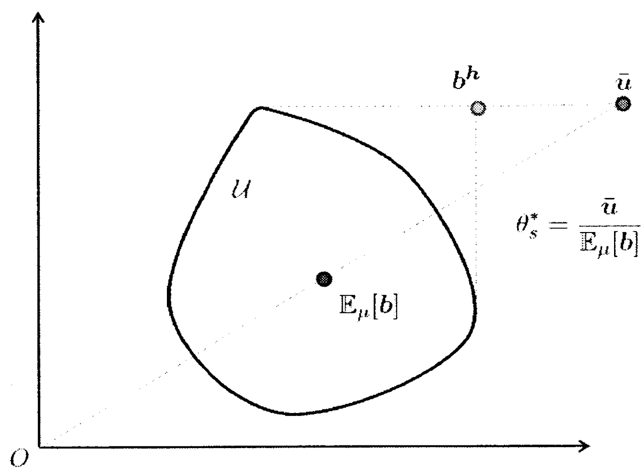

Figure 3-6: A geometric perspective on the stochasticity gap.

In the following section, we show that the scaling bound can be strictly tighter than the symmetry bound. On the other hand, the symmetry bound relates the performance of a static robust solution to key geometric properties of the uncertainty set. Furthermore, both bounds are equal for several interesting classes of uncertainty sets as discussed in the following section. In Theorem 2, we show that the symmetry bound is tight for any $\rho \leq 1$, $s \geq 1/m$ for a family of uncertainty sets.

The symmetry bound reveals several key qualitative properties of the performance gap that are difficult to see from the scaling bound. For example, for any symmetric uncertainty set, without the need to compute $\boldsymbol{b}^h$ and θ_s^*, we have a general bound of two on the stochasticity gap. Since symmetry of any convex set is at least $1/m$, the bound for any convex uncertainty set is at most $(m+1)$. Both these bounds are not obvious from the scaling bound. Most importantly, in practice the symmetry bound gives an informative guideline in designing uncertainty sets where the robust solution has guaranteed performance.

3.4.3 Examples: Stochasticity Gap for Specific Uncertainty Sets

In this subsection, we give examples of specific uncertainty sets and characterize their symmetry. Both bounds on the stochasticity gap, the symmetry bound (3.13) as well the scaling bound (3.16), are presented in Table 3.1 for several interesting uncertainty sets. In most of the cases, the scaling bound is equal to the symmetry bound. The proofs of the symmetry computation of various uncertainty sets are deferred to the Appendix.

An L_p-ball intersected with the nonnegative orthant. We define,

$$\mathbf{B}_p^+ := \{\boldsymbol{b} \in \mathbb{R}_+^m \mid \|\boldsymbol{b}\|_p \leq 1\},$$

for $p \geq 1$. In Appendix 3.10.1 we show the following.

$$\mathbf{sym}(\mathbf{B}_p^+) = \left(\frac{1}{m}\right)^{\frac{1}{p}}, \quad \boldsymbol{b}_0(\mathbf{B}_p^+) = \frac{1}{m^{\frac{1}{p}}+1}\boldsymbol{e}. \tag{3.18}$$

Therefore, if $\mathcal{U} = \mathbf{B}_p^+$ and the probability measure satisfies condition (3.12), then $z_{\text{Rob}} \leq (1 + m^{\frac{1}{p}}) z_{\text{Stoch}}$. The bound is tight as shown in Theorem 2. There are several interesting cases for $\mathbf{B}_p^+$ uncertainty sets. In particular,

1. For $p = 1$, $\mathbf{B}_p^+$ is the standard simplex centered at the origin. The symmetry is $1/m$ and stochasticity gap is $m + 1$.

2. For $p = 2$, $\mathbf{B}_p^+$ is the Euclidean ball in the nonnegative orthant. Its symmetry is $1/\sqrt{m}$, and the stochasticity gap is $1 + \sqrt{m}$.

3. For $p = \infty$, $\mathbf{B}_p^+$ is a hypercube centered at $\boldsymbol{e}/2$ and touches the origin. The symmetry is 1 and the stochasticity gap is 2.

Ellipsoidal uncertainty set. An ellipsoidal uncertainty set is defined as,

$$\mathcal{U} := \{\boldsymbol{b} \in \mathbb{R}_+^m \mid \|\boldsymbol{E}(\boldsymbol{b} - \bar{\boldsymbol{b}})\|_2 \leq 1\}, \tag{3.19}$$

where $\boldsymbol{E}$ is an invertible matrix. Since $\mathcal{U}$ is symmetric, the stochasticity gap is bounded by 2 if $\mathcal{U} \subseteq \mathbb{R}_+^m$ and the probability measure satisfies (3.12). In Appendix 3.10.1, we show that the bound can be improved to the following.

$$z_{\text{Rob}} \leq \left(1 + \max_{1 \leq i \leq m} \frac{E_{ii}^{-1}}{\bar{b}_i}\right) z_{\text{Stoch}} \leq 2 z_{\text{Stoch}}.$$

Intersection of two L_p-balls. Consider the following uncertainty set.

$$\mathcal{U} := \{\boldsymbol{b} \in \mathbb{R}_+^m \mid \|\boldsymbol{b}\|_{p_1} \leq 1, \|\boldsymbol{b}\|_{p_2} \leq r\}, \quad \text{for } 0 < r < 1 \text{ and } 1 \leq p_1 < p_2. \tag{3.20}$$

Assume the following condition holds,

$$\frac{r}{\|\boldsymbol{e}\|_{p_2}} > \frac{1}{\|\boldsymbol{e}\|_{p_1}} \Leftrightarrow r m^{\frac{1}{p_1}} > m^{\frac{1}{p_2}}, \tag{3.21}$$

which guarantees that the intersection of the two unit norm balls is non-trivial. In

Appendix 3.10.1, we show that

$$\mathbf{sym}(\mathcal{U}) = \frac{1}{rm^{\frac{1}{p_1}}}, \quad \boldsymbol{b}_0(\mathcal{U}) = \frac{r}{rm^{\frac{1}{p_1}} + 1}\boldsymbol{e}. \tag{3.22}$$

The budgeted uncertainty set. The budgeted uncertainty set is defined as,

$$\Delta_k := \left\{ \boldsymbol{b} \in [0,1]^m \;\middle|\; \sum_{i=1}^{m} b_i \leq k \right\}, \quad \text{for } 1 \leq k \leq m. \tag{3.23}$$

In Appendix 3.10.1, we show that

$$\mathbf{sym}(\Delta_k) = \frac{k}{m}, \quad \boldsymbol{b}_0(\Delta_k) = \frac{k}{m+k}\boldsymbol{e}. \tag{3.24}$$

Demand uncertainty set. We define the following set,

$$\mathbf{DU} := \left\{ \boldsymbol{b} \in \mathbb{R}^m_+ \;\middle|\; \left| \frac{\sum_{i \in S} \boldsymbol{b}_i - |S|\mu}{\sqrt{|S|}} \right| \leq \Gamma, \ \forall S \subseteq N := \{1, \ldots, m\} \right\}. \tag{3.25}$$

Such a set can model the demand uncertainty where $\boldsymbol{b}$ is the demand of m products; μ and Γ are the center and the span of the uncertain range.

The set **DU** has different symmetry properties, depending on the relation between μ and Γ. If $\mu \geq \Gamma$, the set **DU** is in fact symmetric. Intuitively, **DU** is the intersection of an L_∞ ball centered at $\mu\boldsymbol{e}$ with $(2^m - 2m)$ halfspaces that are symmetric with respect to $\mu\boldsymbol{e}$. If $\mu < \Gamma$, **DU** is not symmetric any more — part of it is cut off by the nonnegative orthant. In Appendix 3.10.1, we present a proof of the following proposition, which summarizes the symmetry property of **DU** for all the cases.

Proposition 1. *Assume the uncertainty set is* **DU***,*

1. *If $\mu \geq \Gamma$, then,*

$$\mathbf{sym}(\mathbf{DU}) = 1, \quad \boldsymbol{b}_0(\mathbf{DU}) = \mu\boldsymbol{e}. \tag{3.26}$$

2. *If* $\frac{1}{\sqrt{m}}\Gamma < \mu < \Gamma$, *then,*

$$\mathbf{sym}(\mathbf{DU}) = \frac{\sqrt{m}\mu + \Gamma}{(1+\sqrt{m})\Gamma}, \quad \boldsymbol{b}_0(\mathbf{DU}) = \frac{(\sqrt{m}\mu + \Gamma)(\mu+\Gamma)}{\sqrt{m}\mu + (2+\sqrt{m})\Gamma}\boldsymbol{e}. \tag{3.27}$$

3. *If* $0 \le \mu \le \frac{1}{\sqrt{m}}\Gamma$, *then,*

$$\mathbf{sym}(\mathbf{DU}) = \frac{\sqrt{m}\mu + \Gamma}{\sqrt{m}(\mu+\Gamma)}, \quad \boldsymbol{b}_0(\mathbf{DU}) = \frac{(\sqrt{m}\mu + \Gamma)(\mu+\Gamma)}{2\sqrt{m}\mu + (1+\sqrt{m})\Gamma}\boldsymbol{e}. \tag{3.28}$$

3.5 Multi-stage Stochastic Problem under RHS Uncertainty

In this section, we consider the multi-stage stochastic optimization problem, Π^K_{Stoch}, under right hand side uncertainty where the multi-stage uncertainty is described by a directed network as discussed earlier. We show that a finitely adaptable class of solutions is a good approximation for the fully-adaptable multi-stage problem. Furthermore, the performance ratio of the finitely adaptable solution depends on the geometric properties of the uncertainty sets in the multi-stage uncertainty network. The number of solutions at each Stage $(k+1)$ for $k = 1, \ldots, K-1$ depends on the number of directed paths in the uncertainty network from the root node to nodes in Stage $(k+1)$. Therefore, if $\mathcal{P}$ is the set of all directed paths from Stage 2 to Stage K, the total number of solutions in any stage in the finitely adaptable solution policy is bounded by $|\mathcal{P}|$.

Theorem 3. *Let* $s = \min_{k,j} \mathbf{sym}(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$ *and* $\rho = \max_{k,j} \rho(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$. *Suppose* $\mathbb{E}_{\boldsymbol{b}}[\boldsymbol{b} \mid \boldsymbol{b} \in \mathcal{U}^k_j] \ge \boldsymbol{u}_{k,j}$ *where* $\boldsymbol{u}_{k,j}$ *is the point of symmetry of* $\mathcal{U}^k_j \subseteq \mathbb{R}^m_+$ *for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Then there is a finitely adaptable solution policy that can be computed efficiently and has at most* $|\mathcal{P}|$ *solutions in each stage, where* $\mathcal{P}$ *is the set of directed paths from the root node in Stage* 2 *to nodes in Stage* K *in the multi-stage uncertainty network, such that the expected cost is at most* $(1 + \rho/s)$

times the optimal cost of Π^K_{Stoch}.

3.5.1 Algorithm

We first describe an algorithm to construct a finitely adaptable solution. In each Stage $(k+1)$ for $k = 1, \ldots, K-1$, the set of finitely adaptable solutions contains a unique solution corresponding to each directed path from the root node in Stage 2 to a node in Stage $(k+1)$. Therefore, we consider $|\mathcal{P}[k]|$ solutions in Stage $(k+1)$ each indexed by $\boldsymbol{P}[k]$ for all $\boldsymbol{P} \in \mathcal{P}$. In other words, the finitely adaptable solution is specified by the first-stage solution $\boldsymbol{x}$, and for each Stage $(k+1)$ for $k = 1, \ldots, K-1$, $\boldsymbol{y}_k(\boldsymbol{P}[k])$ for all $\boldsymbol{P}[k] \in \mathcal{P}[k]$. Recall that for any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$, the probability that the uncertain parameters realize from the path $\boldsymbol{P}$ is given by,

$$Pr(j_1, \ldots, j_{K-1}) = \prod_{k=1}^{K-2} p^k_{j_k, j_{k+1}}, \tag{3.29}$$

and the measure $\mu_{\boldsymbol{P}}$ is defined as a product measure of the measures on the uncertainty sets in Path $\boldsymbol{P}$. We first show that for any $k = 1, \ldots, K-1$, $j = 1, \ldots, N_k$, $(1+\rho/s) \cdot \boldsymbol{u}_{k,j}$ dominates all points in $\mathcal{U}^k_j$ coordinatewise.

Lemma 5. *For any* $k = 1, \ldots, K$, $j = 1, \ldots, N_k$, *for all* $\boldsymbol{u} \in \mathcal{U}^k_j$,

$$\left(1 + \frac{\rho}{s}\right) \cdot \boldsymbol{u}_{k,j} \geq \boldsymbol{u}.$$

Proof. Let $\rho_1 = \rho(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$, $s_1 = \mathbf{sym}(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$. From Lemma 4, we know that

$$\left(1 + \frac{\rho_1}{s_1}\right) \cdot \boldsymbol{u}_{k,j} \geq \boldsymbol{u},\ \forall \boldsymbol{u} \in \mathcal{U}^k_j.$$

By definition in Theorem 3, we know that $\rho_1 \leq \rho$ and $s_1 \geq s$. Therefore, for all $\boldsymbol{u} \in \mathcal{U}^k_j$,

$$\left(1 + \frac{\rho}{s}\right) \cdot \boldsymbol{u}_{k,j} \geq \left(1 + \frac{\rho_1}{s_1}\right) \cdot \boldsymbol{u}_{k,j} \geq \boldsymbol{u}.$$

□

We consider the following multi-stage problem to compute the finitely adaptable solution.

$$\begin{aligned}
z_{\mathcal{A}} = \min \quad & \boldsymbol{c}^T\boldsymbol{x} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}[\boldsymbol{d}_k^T\boldsymbol{y}_k(\boldsymbol{P}[k])] \\
\text{s.t. } & \forall \boldsymbol{P} = (j_1,\ldots,j_{K-1}) \in \mathcal{P} \\
& \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{B}_k\boldsymbol{y}_k(\boldsymbol{P}[k]) \geq \left(1+\frac{\rho}{s}\right)\cdot\sum_{k=1}^{K-1}\boldsymbol{u}_{k,j_k} \\
& \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}_+^{n_1-p_1}, \\
& \boldsymbol{y}_k(\boldsymbol{P}[k]) \in \mathbb{R}^{p_k} \times \mathbb{R}_+^{n_k-p_k},\ \forall k = 1,\ldots,K-1.
\end{aligned} \tag{3.30}$$

The number of constraints in (3.30) is equal to $|\mathcal{P}|\cdot m$, where m is the number of rows of $\boldsymbol{A}, \boldsymbol{B}_k$ for all $k = 1,\ldots,K-1$. Also, the number of decision variables in each Stage $(k+1)$ for $k = 1,\ldots,K-1$ is $|\mathcal{P}[k]|$, namely, $\boldsymbol{y}_k(\boldsymbol{P}[k])$ for all $\boldsymbol{P}[k] \in \mathcal{P}[k]$. Therefore, the solution is finitely adaptable as there are only a finite number of solutions in each stage. Furthermore, in some cases, the number of directed paths is small and polynomial in the size of the uncertainty network. For instance, if the uncertainty network is a path, we require exactly one solution at each stage in our finitely adaptable solution. If the network is a tree, then there is a single path to each uncertainty set from the root node in Stage 2 and therefore, the number of solutions in the finitely adaptable solution is exactly equal to the number of uncertainty sets in the uncertainty network. However, in general, the number of directed paths can be exponential in the input size. For example, for the case of a recombining directed network in Figure 3-1, the number of directed paths from the root node to the node j in Stage $(k+1)$, $j = 1,\ldots,N_k$ is equal to $\binom{k}{j}$, which is exponential in k, while the input size is $\mathrm{O}(k^2)$.

In the next subsection, we show that the finitely adaptable solution computed in (3.30) is a good approximation of a fully-adaptable optimal stochastic solution.

3.5.2 Proof of Theorem 3

For brevity, let $\tau = (1 + \rho/s)$. The proof proceeds as follows. We first consider a particular finitely adaptable solution feasible to (3.30), which implies that its expected cost is at least $z_{\mathcal{A}}$. We then extend this particular finitely adaptable solution to a fully-adaptable one without changing its expected cost. Finally, we show that the expected cost of the extended solution is equal to τ times the expected cost of an optimal fully-adaptable solution.

Let $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k])$ denote an optimal fully-adaptable solution for Π^K_{Stoch} for all $k = 1, \ldots, K-1$, $\boldsymbol{P} \in \mathcal{P}$, $\mu_{\boldsymbol{P}}$-a.e. $\boldsymbol{\omega} \in \Omega(\boldsymbol{P})$. For the first step of the proof, we consider the following particular finitely adaptable solution for (3.30). For all $\boldsymbol{P} \in \mathcal{P}$, $k = 1, \ldots, K-1$, let

$$\overline{\boldsymbol{y}}_k(\boldsymbol{P}[k]) = \tau \cdot \mathbb{E}_{\omega \in \Omega(\boldsymbol{P})}\left[\hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k])\right]. \tag{3.31}$$

Also, let $\overline{\boldsymbol{x}} = \tau \cdot \hat{\boldsymbol{x}}$. We show that the above finitely adaptable solution is feasible for (3.30). Consider any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$. Now,

$$\boldsymbol{A}\hat{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) \geq \sum_{k=1}^{K-1} \boldsymbol{b}_k, \mu_{\boldsymbol{P}}\text{-a.e. } \boldsymbol{\omega} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-1}) \in \Omega(\boldsymbol{P}).$$

Taking conditional expectation with respect to $\mu_{\boldsymbol{P}}$ on both sides, we have that

$$\boldsymbol{A}\hat{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \mathbb{E}_{\mu_{\boldsymbol{P}}}\left[\hat{\boldsymbol{y}}_k\left(\boldsymbol{\omega}[k], \boldsymbol{P}[k]\right)\right] \geq \sum_{k=1}^{K-1} \mathbb{E}_{\mu_{\boldsymbol{P}}}\left[\boldsymbol{b}_k\right] \geq \sum_{k=1}^{K} \boldsymbol{u}_{k,j_k}, \tag{3.32}$$

where the last inequality follows as $\mathbb{E}_{\mu_{\boldsymbol{P}}}[\boldsymbol{b}_k] \geq \boldsymbol{u}_{k,j_k}$. Therefore, for any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in$

$\mathcal{P}$,

$$\begin{aligned} \boldsymbol{A}\overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \overline{\boldsymbol{y}}_k(\boldsymbol{P}[k]) &= \boldsymbol{A}(\tau \cdot \hat{\boldsymbol{x}}) + \sum_{k=1}^{K-1} \boldsymbol{B}_k \left(\tau \cdot \mathbb{E}_{\mu_{\boldsymbol{P}}} \left[\boldsymbol{d}_k^T \hat{\boldsymbol{y}}_k\left(\boldsymbol{\omega}[k], \boldsymbol{P}[k]\right)\right]\right) \\ &= \tau \cdot \left(\boldsymbol{A}\hat{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \mathbb{E}_{\mu_{\boldsymbol{P}}} \left[\boldsymbol{d}_k^T \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k])\right] \right) \\ &\geq \tau \cdot \left(\sum_{k=1}^{K-1} \boldsymbol{u}_{k,j_k} \right), \end{aligned} \tag{3.33}$$

where (3.33) follows from (3.32). Therefore, the solution $\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}$ is a feasible solution for (3.30). Let $\overline{z}$ be the expected cost of $\overline{\boldsymbol{x}}$, $\overline{\boldsymbol{y}}$, i.e.,

$$\overline{z} = \boldsymbol{c}^T \overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}[\boldsymbol{d}_k^T \overline{\boldsymbol{y}}_k(\boldsymbol{P}[k])].$$

Clearly, $z_{\mathcal{A}} \leq \overline{z}$. For the second step of the proof, we extend the finitely adaptable solution $\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}$ to a feasible solution $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}$ for Π^K_{Stoch} as follows.

$$\begin{aligned} \tilde{\boldsymbol{x}} &= \overline{\boldsymbol{x}} \\ \tilde{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) &= \overline{\boldsymbol{y}}_k(\boldsymbol{P}[k]),\ \forall \boldsymbol{P} \in \mathcal{P},\ \forall k = 1, \ldots, K-1. \end{aligned} \tag{3.34}$$

We show that the extended solution $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}$ is a feasible solution for Π^K_{Stoch}. Consider any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$, $\boldsymbol{\omega} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-1}) \in \Omega(\boldsymbol{P})$. Therefore,

$$\begin{aligned} \boldsymbol{A}\tilde{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \tilde{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) &= \boldsymbol{A}\overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \overline{\boldsymbol{y}}_k(\boldsymbol{P}[k]) \\ &\geq \tau \cdot \sum_{k=1}^{K-1} \boldsymbol{u}_{k,j_k} \qquad (3.35) \\ &\geq \sum_{k=1}^{K-1} \boldsymbol{b}_k, \end{aligned}$$

where (3.35) follows from the feasibility of the solution $\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}$ for (3.30) and the last inequality follows from Lemma 5. Let $\tilde{z}$ be the expected cost of $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}$, i.e.,

$$\begin{aligned}
\tilde{z} &= \boldsymbol{c}^T\tilde{\boldsymbol{x}} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}\left[\mathbb{E}_{\mu_{\boldsymbol{P}}}[\boldsymbol{d}_k^T\tilde{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k])]\right] \\
&= \boldsymbol{c}^T\overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}\left[\mathbb{E}_{\mu_{\boldsymbol{P}}}[\boldsymbol{d}_k^T\overline{\boldsymbol{y}}_k(\boldsymbol{P}[k])]\right] \\
&= \boldsymbol{c}^T\overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}[\boldsymbol{d}_k^T\overline{\boldsymbol{y}}_k(\boldsymbol{P}[k])] \qquad (3.36) \\
&= \overline{z}
\end{aligned}$$

where (3.36) follows from (3.34). Now, we compare $\tilde{z}$ to the expected cost of an optimal fully-adaptable solution as follows.

$$\begin{aligned}
\tilde{z} &= \overline{z} \\
&= \boldsymbol{c}^T\overline{\boldsymbol{x}} + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}[\boldsymbol{d}_k^T\overline{\boldsymbol{y}}_k(\boldsymbol{P}[k])] \\
&= \boldsymbol{c}^T(\tau\cdot\boldsymbol{x}) + \sum_{k=1}^{K-1} \mathbb{E}_{\boldsymbol{P}\in\mathcal{P}}\left[\boldsymbol{d}_k^T\left(\tau\cdot\mathbb{E}_{\mu_{\boldsymbol{P}}}\left[\hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k])\right]\right)\right] \\
&= \tau\cdot z_{\text{Stoch}}^K, \qquad (3.37)
\end{aligned}$$

where (3.37) follows from the optimality of $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ for Π_{Stoch}^K. Combining $z_{\mathcal{A}} \leq \overline{z}$ and (3.37), we obtain that $z_{\mathcal{A}} \leq \tau\cdot z_{\text{Stoch}}^K$. □

Stochastic problem under cost and RHS uncertainty. For the multi-stage stochastic problem, $\Pi_{\text{Stoch}(\boldsymbol{b},\boldsymbol{d})}^K$ where both the objective coefficients and the right hand side are uncertain, a finitely adaptable solution performs arbitrarily worse as compared to the fully-adaptable solution. It follows from one of the results in Bertsimas and Goyal [18], where the authors show that a static-robust solution may perform arbitrarily worse as compared to an optimal fully-adaptable two-stage solution for the stochastic problem when both cost and right hand side are uncertain.

3.6 Multi-stage Adaptive Problem

In this section, we consider multi-stage adaptive optimization problems and show that a finitely adaptable solution is a good approximation to the fully-adaptable solution. Furthermore, the performance bound of a finitely adaptable solution is related to the symmetry and translation factors of the uncertainty sets as in the case of stochastic optimization problems. The approximation guarantee for the case when only the right hand sides of the multi-stage adaptive problem are uncertain, Π^K_{Adapt}, follows directly from the approximation bounds for the stochastic problem. Surprisingly, we can also show that a finitely adaptable solution is also a good approximation for the multi-stage adaptive problem when both the right hand sides and the objective coefficients are uncertain, $\Pi^K_{\mathrm{Adapt}(\boldsymbol{b},\boldsymbol{d})}$, unlike the stochastic counterpart. Our results generalize the performance of a static robust solution for two-stage adaptive optimization problem under right hand side and/or objective coefficient uncertainty.

3.6.1 Right-hand Side Uncertainty: Π^K_{Adapt}

We show that there is a finitely adaptable solution with at most $|\mathcal{P}|$ solutions for each stage and each uncertainty set, that is a good approximation of the fully-adaptable problem. Furthermore, such a finitely adaptable solution can be computed efficiently using exactly the algorithm described in Section 3.5.1. The performance bound of the finitely adaptable solution policy follows directly from its performance bound with respect to the stochastic problem in Theorem 3. Therefore, we have the following theorem.

Theorem 4. *Suppose $s = \min_{k,j} \mathbf{sym}(\mathcal{U}^k_j)$ and $\boldsymbol{u}_{k,j} \in \mathcal{U}^k_j \subseteq \mathbb{R}^m_+$ is the point of symmetry for all $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. Let $\rho = \max_{k,j} \rho(\mathcal{U}^k_j)$ for all $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. Then there is a finitely adaptable solution policy that can be computed efficiently and has at most $|\mathcal{P}|$ solutions in each stage, where $\mathcal{P}$ is the set of directed paths from the root node to any node in Stage K of the multi-stage uncertainty network, such that its worst-case cost is at most $(1+\rho/s)$ times the optimal cost of Π^K_{Adapt}.*

3.6.2 RHS and Cost Uncertainty

In this section, we consider the multi-stage adaptive optimization problem where both the right hand side and the objective coefficients are uncertain. While for the stochastic problem, the performance of a finitely adaptable solution can be arbitrarily bad with respect to an optimal stochastic solution, surprisingly, we can show that there exists a finitely adaptable solution with at most $|\mathcal{P}|$ solutions for each stage for each uncertainty set, that is a good approximation for the multi-stage problem.

Theorem 5. *Suppose* $s = \min_{k,j} \mathbf{sym}(\mathcal{U}_j^k)$ *and* $\boldsymbol{u}_{k,j} \in \mathcal{U}_j^k \subseteq \mathbb{R}_+^{m+n_k}$ *is the point of symmetry for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Also, let* $\boldsymbol{u}_{k,j}^{\boldsymbol{b}}, \boldsymbol{u}_{k,j}^{\boldsymbol{d}}$ *denote the right hand side and the objective coefficient uncertainty in* $\boldsymbol{u}_{k,j}$ *respectively. Let* $\rho = \max_{k,j} \rho(\mathcal{U}_j^k)$ *for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Then there is a finitely adaptable solution policy that can be computed efficiently and has at most* $|\mathcal{P}|$ *solutions where* $\mathcal{P}$ *is the set of directed paths from the root node to any node in Stage* K *of the multi-stage uncertainty network, such that its worst-case cost is at most* $(1+\rho/s)^2$ *times the optimal cost of* $\Pi^K_{\mathrm{Adapt}(\boldsymbol{b},\boldsymbol{d})}$.

Proof. From Lemma 4, we know that for any uncertainty set $\mathcal{U}_j^k$ in Stage $(k+1)$, $\boldsymbol{u} \leq (1+\rho/s) \cdot \boldsymbol{u}_{k,j}$ for all $\boldsymbol{u} \in \mathcal{U}_j^k$. We compute a finitely adaptable solution by solving the following multi-stage problem similar to the one used to compute a finitely adaptable solution for the stochastic counterpart.

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^T\boldsymbol{x} + \max_{\boldsymbol{P}\in\mathcal{P}} \min_{\boldsymbol{y}_k(\boldsymbol{P}[k]), k=1,\ldots,K-1} \sum_{k=1}^{K-1} \boldsymbol{d}_k^T \boldsymbol{y}_k(\boldsymbol{P}[k]) \\
\text{s.t.} \quad & \forall \boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P} \\
& \boldsymbol{A}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \boldsymbol{y}_k(\boldsymbol{P}[k]) \geq \left(1 + \frac{\rho}{s}\right) \cdot \sum_{k=1}^{K-1} \boldsymbol{u}_{k,j_k}^{\boldsymbol{b}}, \\
& \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}_+^{n_1-p_1}, \\
& \boldsymbol{y}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]) \in \mathbb{R}^{p_k} \times \mathbb{R}_+^{n_k-p_k}, \ \forall k = 1, \ldots, K-1.
\end{aligned}
\tag{3.38}
$$

Note that the above problem is similar to (3.30) except the objective function. Using an argument similar to the second part of the proof of Theorem 3 (see Section 3.5.2), we can show that a finitely adaptable solution of (3.38) can be extended to a feasible fully-adaptable solution for $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$ of the same worst-case cost.

Suppose $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}[k], \boldsymbol{P}[k]))$ for all $k = 1, \ldots, K-1$, $\boldsymbol{P} \in \mathcal{P}, \boldsymbol{\omega} \in \Omega(\boldsymbol{P})$ denote a optimal fully-adaptable solution for $\Pi^K_{\text{Adapt}}(\boldsymbol{b}, \boldsymbol{d})$. As defined earlier, for any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$, let

$$\boldsymbol{\omega}_{\boldsymbol{P}} = (\boldsymbol{u}_{1,j_1}, \ldots, \boldsymbol{u}_{K-1,j_{K-1}}), \tag{3.39}$$

Consider the following approximate solution for (3.38).

$$\tilde{\boldsymbol{x}} = \left(1 + \frac{\rho}{s}\right) \cdot \hat{\boldsymbol{x}}, \tag{3.40}$$

and for any $\boldsymbol{P} \in \mathcal{P}$, for all $k = 1, \ldots, K-1$,

$$\tilde{\boldsymbol{y}}_k(\boldsymbol{P}[k]) = \left(1 + \frac{\rho}{s}\right) \cdot \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}_{\boldsymbol{P}}[k], \boldsymbol{P}[k]). \tag{3.41}$$

For brevity, as before, let $\tau = (1 + \rho/s)$. We first need to show that the solution $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}$ is feasible for (3.38). For any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$,

$$\boldsymbol{A}\tilde{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \tilde{\boldsymbol{y}}_k(\boldsymbol{P}[k]) = \tau \cdot \left(\boldsymbol{A}\hat{\boldsymbol{x}} + \sum_{k=1}^{K-1} \boldsymbol{B}_k \hat{\boldsymbol{y}}_k(\boldsymbol{\omega}_{\boldsymbol{P}}[k], \boldsymbol{P}[k])\right) \geq \tau \cdot \left(\sum_{k=1}^{K-1} \boldsymbol{u}_{k,j_k}\right),$$

where the last inequality follows from the feasibility for $\hat{\boldsymbol{y}}$ for $\boldsymbol{\omega}_{\boldsymbol{P}}$. Thus, the solution $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}$ is feasible for (3.38).

To bound the worst-case cost of the solution, we show that for any $\boldsymbol{P} \in \mathcal{P}$, the cost of the approximate finitely adaptable solution is at most $(1+\rho/s)^2$ times the worst-case cost of the optimal fully-adaptable solution. Consider any $\boldsymbol{P} = (j_1, \ldots, j_{K-1}) \in \mathcal{P}$.

Now,

$$
\begin{aligned}
c^T\tilde{x} + \sum_{k=1}^{K-1} d_k^T \tilde{y}_k(P[k]) &= \tau \cdot \left(c^T\hat{x} + \sum_{k=1}^{K-1} d_k^T \hat{y}_k(\omega_P[k], P[k])\right) \\
&\leq \tau \cdot \left(c^T\hat{x} + \sum_{k=1}^{K-1} \left(\tau \cdot (u^d_{k,j_k})^T \hat{y}_k(\omega_P[k], P[k])\right) \right) \quad (3.42) \\
&= \tau \cdot c^T\hat{x} + \tau^2 \cdot \left(\sum_{k=1}^{K-1} (u^d_{k,j_k})^T \hat{y}_k(\omega_P[k], P[k])\right) \\
&\leq \tau^2 \cdot z^K_{\mathrm{Adapt}(b,d)}, \quad (3.43)
\end{aligned}
$$

where (3.42) follows as $d \leq \tau \cdot u^d_{k,j_k}$ for all $(b,d) \in \mathcal{U}^k_{j_k}$, $k = 1, \ldots, K-1$. Inequality (3.43) follows as $\omega_P \in \Omega(P)$ and thus, is a feasible scenario in $\Pi^K_{\mathrm{Adapt}(b,d)}$. $\square$

3.6.3 An Alternative Bound

In this section, we discuss an alternative bound for the performance of the finitely adaptable solution as compared to the optimal fully-adaptable solution similar to the one we present for the stochastic problem. For the sake of simplicity, we present this alternative bound for the adaptive problem under right hand side uncertainty, Π^K_{Adapt}. The bound extends in a straightforward manner to the case of both right hand side and cost uncertainty.

The proof of Theorem 4 (and also Theorem 5) is based on the construction of a good finitely adaptable solution from an optimal fully-adaptable solution in the following manner. For each $P \in \mathcal{P}$, we consider the scenario ω_P where in each stage the uncertain parameter realization is the point of symmetry of the corresponding uncertainty set on path P. We show that the solution for scenario ω_P scaled by a factor $(1 + \rho/s)$ is a good feasible finitely adaptable solution for all $\omega \in \Omega(P)$. Since $\omega_P \in \Omega(P)$, the cost for an optimal solution for this scenario is a lower bound on z^K_{Adapt} which implies a bound of $(1 + \rho/s)$ for the finitely adaptable solution with respect to the optimal. Now, if for some other scenario $\omega'(P) \in \Omega(P)$, a smaller scaling factor than $(1 + \rho/s)$ suffices to obtain a feasible finitely adaptable solution for all $\omega \in \Omega(P)$ for all $P \in \mathcal{P}$, this would imply a smaller bound on the performance of the finitely adaptable solution.

Following the discussion in Section 3.4.2, we define $\boldsymbol{b}^h(\mathcal{U})$ as follows. For all $j = 1, \ldots, m$, $b_j^h(\mathcal{U}) := \max_{\boldsymbol{b} \in \mathcal{U}} b_j$. Let

$$\theta(\mathcal{U}) = \min\{\theta \mid \exists \boldsymbol{b} \in \mathcal{U}, \theta \cdot \boldsymbol{b} \geq \boldsymbol{b}^h(\mathcal{U})\}.$$

Note that $\theta(\mathcal{U}) \leq (1+\rho/s)$ as $(1+\rho/s) \cdot \boldsymbol{b}_0 \geq \boldsymbol{b}^h(\mathcal{U})$, where $\boldsymbol{b}_0$ is the point of symmetry of $\mathcal{U}$, $\rho = \rho(\boldsymbol{b}_0, \mathcal{U})$, and $s = \mathbf{sym}(\boldsymbol{b}_0, \mathcal{U})$. Let $\boldsymbol{b}^1(\mathcal{U})$ denote the vector $\boldsymbol{b} \in \mathcal{U}$ that achieves the minimum value of θ. Let

$$\theta_a^* = \max_{k=1,\ldots,K, j=1,\ldots,N_k} \theta(\mathcal{U}_j^k).$$

For each $k = 1, \ldots, K-1$, $j = 1, \ldots, N_k$, $\theta(\mathcal{U}_j^k)$ is defined by a feasible uncertainty realization from $\mathcal{U}_j^k$. Therefore, scaling the solution corresponding to $\boldsymbol{b}^1(\mathcal{U}_j^k)$ by a factor $\theta_a^* \geq \theta(\mathcal{U}_j^k)$ produces a feasible finitely adaptable solution. Also, as we note above, $\theta_a^* \leq (1+\rho/s)$. Therefore, θ_a^* is upper bounded by the bound in Theorem 4. We refer to θ_a^* as a scaling bound as earlier.

We can interpret the scaling bound geometrically as follows. For any $\mathcal{U} \subseteq \mathbb{R}_+^m$, let $\theta = \theta(\mathcal{U})$, $\boldsymbol{b}^1 = \boldsymbol{b}^1(\mathcal{U})$, and $\boldsymbol{b}^h = \boldsymbol{b}^h(\mathcal{U})$. We know that

$$\theta \cdot \boldsymbol{b}^1 \geq \boldsymbol{b}^h \Rightarrow \boldsymbol{b}^1 \geq \frac{1}{\theta} \cdot \boldsymbol{b}^h.$$

Therefore, $1/\theta$ is the minimum scaling factor for $\boldsymbol{b}^h$ such that it is dominated by some point in $\mathcal{U}$ coordinate-wise. Note that $(1/\theta) \cdot \boldsymbol{b}^h$ does not necessarily belong to $\mathcal{U}$ but is always contained in,

$$\tilde{\mathcal{U}} = \{\boldsymbol{b} \in \mathbb{R}_+^m \mid \exists \boldsymbol{b}' \in \mathcal{U}, \boldsymbol{b}' \geq \boldsymbol{b}\}. \tag{3.44}$$

To see this, consider the following uncertainty set, $\mathcal{U} \subset \mathbb{R}_+^3$, where $\mathcal{U}$ is the convex hull of $\big((0,0,0), (1,0,1), (0,1,1)\big)$. We can alternatively define $\mathcal{U}$ as $\mathcal{U} = \{\boldsymbol{b} \in [0,1]^3 \mid b_3 = b_1 + b_2\}$. It is easy to observe that $\boldsymbol{b}^h = (1,1,1)$. We show that $\boldsymbol{b}^1 = (1/2, 1/2, 1) \in \mathcal{U}$

and $\theta = 2$. We know that $\theta b_j^1 \geq 1$ for all $j = 1, 2$. Therefore,

$$b_3^1 = b_1^1 + b_2^1 \geq \frac{2}{\theta}.$$

Furthermore, $b_3^1 \leq 1$ which implies that $\theta \geq 2$. For $\theta = 2$, $\boldsymbol{b}^1 = (1/2, 1/2, 1) \in \mathcal{U}$ satisfies $\theta \boldsymbol{b}^1 \geq \boldsymbol{b}^h$. Now,

$$\frac{1}{\theta} \cdot \boldsymbol{b}^h = \frac{1}{2} \cdot (1, 1, 1) \notin \mathcal{U}.$$

However, $\boldsymbol{b}^1 \geq (1/\theta)\boldsymbol{b}^h$, and $\boldsymbol{b}^1 \in \mathcal{U}$ which implies that $(1/\theta)\boldsymbol{b}^h \in \tilde{\mathcal{U}}$ as defined in (3.44). The geometric picture is given in Figure 3-7. Note that in Figure 3-7, $\boldsymbol{b}^1 = (1/\theta)\boldsymbol{b}^h$ but this is not true in general as illustrated in the above example.

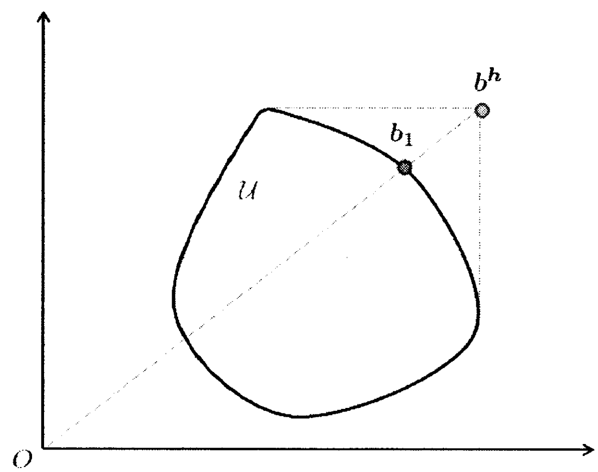

Figure 3-7: A geometric perspective on the adaptability gap.

For certain uncertainty sets, the scaling bound of θ_a^* is strictly better than the bound in Theorem 4.

1. **Hypercube**. Suppose each uncertainty set in the multi-stage uncertainty network is a hypercube. Therefore, $s = 1$. Also, suppose $\rho = 1$ where ρ is as defined in Theorem 4. The bound in Theorem 4 is $(1 + \rho/s) = 2$. On the other hand, $\theta_a^* = 1$ since $\boldsymbol{b}^h(\mathcal{U}) \in \mathcal{U}$, when $\mathcal{U}$ is a hypercube. Therefore, the two bounds are in fact different. The scaling bound implies that the finitely

adaptable solution is optimal, while the symmetry bound implies that it is only a 2-approximation for the multi-stage adaptive optimization problem.

2. **Hypersphere**. Suppose the uncertainty sets are all hyperspheres in $\mathbb{R}^m_+$, i.e., L_2-balls, with unit radius and centered at $\boldsymbol{e} = (1, 1, \ldots, 1)$. Therefore, $\rho = 1$ and also $s = 1$ which implies that the bound from Theorem 4 is 2. On the other hand, it is easy to note that the scaling bound,

$$\theta_a^* = \frac{2}{1 + 1/\sqrt{m}}.$$

In this case, the scaling bound is not significantly better than the bound in Theorem 4.

3.7 Extension to General Convex Cones

We consider extensions of our results to the case where the constraints are general linear conic inequalities and the uncertainty set belongs to the underlying cone. For simplicity, we discuss the two-stage case. The generalization also applies to the multi-stage problems.

3.7.1 Stochastic Problem with Linear Conic Constraints

We consider the following two-stage conic stochastic optimization problem.

$$\begin{aligned} z_{\text{Stoch}}^{\mathcal{K}} := \min_{\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{b})} \quad & \boldsymbol{c}^T \boldsymbol{x} + \mathbb{E}_\mu[\boldsymbol{d}^T \boldsymbol{y}(\boldsymbol{b})] \\ \text{s.t.} \quad & \mathcal{A}\boldsymbol{x} + \mathcal{B}\boldsymbol{y}(\boldsymbol{b}) \succeq_{\mathcal{K}} \boldsymbol{b}, \ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}, \\ & \boldsymbol{x} \in \mathbb{R}^{p_1} \times \mathbb{R}^{n_1 - p_1}_+, \\ & \boldsymbol{y}(\boldsymbol{b}) \in \mathbb{R}^{p_2} \times \mathbb{R}^{n_2 - p_2}_+, \end{aligned} \tag{3.45}$$

where $\mathcal{K}$ is a closed pointed convex cone in a finite dimensional space, such as the nonnegative orthant $\mathbb{R}^m_+$, the second-order cone (SOC) $\{(\boldsymbol{b}, t) \in \mathbb{R}^{m+1} : \|\boldsymbol{b}\|_2 \leq t\}$, and the semidefinite cone $\mathbb{S}^m_+$. Here $\mathcal{A}, \mathcal{B}$ are mappings from $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$ to the finite dimensional space that contains $\boldsymbol{\mathcal{K}}$, respectively. For example, if $\boldsymbol{\mathcal{K}}$ is the nonnegative orthant or the second-order cone, both $\mathcal{A}$ and $\mathcal{B}$ are matrices of the appropriate dimension. If $\boldsymbol{\mathcal{K}}$ is the semi-definite (SDP) cone, then $\mathcal{A}, \mathcal{B}$ are linear mappings defined as,

$$\mathcal{A}\boldsymbol{x} = \sum_{i=1}^{n_1} x_i \boldsymbol{A}_i, \; \mathcal{B}\boldsymbol{y} = \sum_{i=1}^{n_2} y_i \boldsymbol{B}_i,$$

where A_i and B_i are symmetric matrices. The linear conic inequality (3.45) is equivalent to the inclusion in the cone, i.e., $\mathcal{A}\boldsymbol{x} + \mathcal{B}\boldsymbol{y}(\boldsymbol{b}) \succeq_{\boldsymbol{\mathcal{K}}} \boldsymbol{b} \Leftrightarrow \mathcal{A}\boldsymbol{x} + \mathcal{B}\boldsymbol{y}(\boldsymbol{b}) - \boldsymbol{b} \in \boldsymbol{\mathcal{K}}$.

Lemma 6. *Suppose the convex, compact set $\mathcal{U} \subset \boldsymbol{\mathcal{K}}$, and $\boldsymbol{b}_0$ is the point of symmetry of $\mathcal{U}$. Then,*

$$\left(1 + \frac{1}{\mathbf{sym}(\mathcal{U})}\right) \cdot \boldsymbol{b}_0 \succeq_{\boldsymbol{\mathcal{K}}} \boldsymbol{b}, \; \forall \boldsymbol{b} \in \mathcal{U}.$$

Proof. By the definition of symmetry and the assumption that $\mathcal{U} \subset \boldsymbol{\mathcal{K}}$, we have for any $\boldsymbol{b} \in \mathcal{U}$, $\boldsymbol{b}_0 + \mathbf{sym}(\mathcal{U})(\boldsymbol{b}_0 - \boldsymbol{b}) = (\mathbf{sym}(\mathcal{U}) + 1)\boldsymbol{b}_0 - \mathbf{sym}(\mathcal{U})\boldsymbol{b} \in \boldsymbol{\mathcal{K}}$. □

Theorem 6. *Consider the two-stage stochastic optimization problem in* (3.45). *Let μ be the probability measure on the uncertainty set $\mathcal{U}$, $\boldsymbol{b}_0$ be the point of symmetry of $\mathcal{U}$, and $\rho = \rho(\boldsymbol{b}_0, \mathcal{U})$ be the translation factor of $\boldsymbol{b}_0$ with respect to $\mathcal{U}$. Denote $s = \mathbf{sym}(\mathcal{U})$. Assume the probability measure μ satisfies, $\mathbb{E}_\mu[\boldsymbol{b}] \succeq_{\boldsymbol{\mathcal{K}}} \boldsymbol{b}_0$. Then the cost of an optimal static solution is at most $(1 + \rho/s) \cdot z^{\boldsymbol{\mathcal{K}}}_{\mathrm{Stoch}}$.*

The proof of the above theorem is similar to the proof of Theorem 1. For the sake of completeness, we present the proof of Theorem 6 in Appendix 3.10.2. A similar result holds for the corresponding adaptive optimization problem as well.

3.8 Extensions to Integer Variables

We can extend our results to the case when some decision variables are integer constrained for both the stochastic as well as the adaptive optimization problems. In the case of the multi-stage stochastic optimization problem with right hand side uncertainty, we can handle integer decision variables only in the first stage. Whereas for the multi-stage adaptive problem, we can handle integer decision variables in every stage for both versions, Π^K_{Adapt} and $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$.

3.8.1 Multi-stage Stochastic Problem

We consider the multi-stage stochastic problem, Π^K_{Stoch} as defined in (3.5) with an additional constraint that some of the first-stage decision variables $\boldsymbol{x}$ are required to be integers. Even for this case, we show that a finitely adaptable solution provides a good approximation.

Theorem 7. *Consider the multi-stage stochastic problem* Π^K_{Stoch} (3.5)*, with additional integer constraints on some first stage decision variables. Suppose* $\mathbb{E}_{\boldsymbol{b}}[\boldsymbol{b} \mid \boldsymbol{b} \in U^k_j] = \boldsymbol{u}_{k,j}$ *for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Let* $s = \min_{k,j} \mathbf{sym}(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$ *and* $\rho = \max_{k,j} \rho(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$. *Then there is a finitely adaptable solution policy that can be computed efficiently and has at most* $|\mathcal{P}|$ *solutions in each stage, where* $\mathcal{P}$ *is the set of directed paths from the root node to nodes in Stage* K *in the multi-stage uncertainty network, such that the expected cost is at most* $\lceil (1 + \rho/s) \rceil$ *times the optimal cost.*

The proof of Theorem 7 is exactly similar to the proof of Theorem 3 except that we need to handle the integrality constraints in constructing a feasible finitely adaptable solution. Therefore, instead of scaling the optimal fully-adaptable solution by a factor of $(1 + \rho/s)$ to construct a feasible finitely adaptable solution,we need to scale by $\lceil (1 + \rho/s) \rceil$ to preserve the integrality constraints. This implies that the performance ratio of the finitely adaptable solution with respect to an optimal fully-adaptable solution is at most $\lceil (1+\rho/s) \rceil$. Note that the stochastic problem, with both right hand side and objective coefficients uncertainty, can not be well approximated by a finitely adaptable solution even without the integrality constraints.

3.8.2 Multi-stage Adaptive Optimization Problem

For the multi-stage adaptive problem, we can handle integer decision variables in all stages. In particular, we have the following theorems.

Theorem 8. *Consider the multi-stage adaptive problem,* Π^K_{Adapt} (3.6), *with additional integer constraints on some decision variables in each stage. Suppose* $s = \min_{k,j} \mathbf{sym}(\mathcal{U}^k_j)$ *and* $\boldsymbol{u}_{k,j} \in \mathcal{U}^k_j$ *is the point of symmetry for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Let* $\rho = \max_{k,j} \rho(\mathcal{U}^k_j)$ *for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Then there is a finitely adaptable solution policy that can be computed efficiently and has at most* $|\mathcal{P}|$ *solutions where* $\mathcal{P}$ *is the set of directed paths from the root node to any node in Stage* K *of the multi-stage uncertainty network, such that its worst-case cost is at most* $\lceil (1+\rho/s) \rceil$ *times the optimal cost.*

Theorem 9. *Consider the multi-stage adaptive problem,* $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$ (3.8), *with additional integer constraints on some decision variables in each stage. Suppose* $s = \min_{k,j} \mathbf{sym}(\mathcal{U}^k_j)$ *and* $\boldsymbol{u}_{k,j} \in \mathcal{U}^k_j$ *is the point of symmetry for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Also, let* $\boldsymbol{u}^{\boldsymbol{b}}_{k,j}, \boldsymbol{u}^{\boldsymbol{d}}_{k,j}$ *denote the right hand side and the objective coefficient uncertainty in* $\boldsymbol{u}_{k,j}$ *respectively. Let* $\rho = \max_{k,j} \rho(\boldsymbol{u}_{k,j}, \mathcal{U}^k_j)$ *for all* $j = 1, \ldots, N_k$, $k = 1, \ldots, K-1$. *Then there is a finitely adaptable solution policy that can be computed efficiently and has at most* $|\mathcal{P}|$ *solutions where* $\mathcal{P}$ *is the set of directed paths from the root node to any node in Stage* K *of the multi-stage uncertainty network, such that its worst-case cost is at most* $\lceil (1+\rho/s) \rceil \cdot (1+\rho/s)$ *times the optimal cost of* $\Pi^K_{\text{Adapt}(\boldsymbol{b},\boldsymbol{d})}$.

3.9 Conclusion

In this chapter, we propose tractable solution policies for two-stage and multi-stage stochastic and adaptive optimization problems and relate the performance of the approximate solution approaches with the fundamental geometric properties of the uncertainty set. For a fairly general stochastic optimization problem, we show that the performance of a static robust solution for the two-stage problem and a finitely adaptable solution for the multi-stage problem is related to the symmetry and translation factor of the uncertainty sets. In particular, the performance bound is $(1+\rho/s)$ where ρ is the translation factor of the uncertainty sets and s is the symmetry. We also show that the bound is tight, i.e., given any symmetry and translation factor, there exists a family of instances where the uncertainty sets have the given symmetry and translation factor, and the cost of an optimal static robust solution is exactly equal to $(1 + \rho/s)$ times the optimal stochastic cost. For most commonly used uncertainty sets, the performance bound gives quite interesting results. For instance, if the sets are perfectly symmetric, i.e., $s = 1$, the bound is less than or equal to 2. Refer to Table 3.1 for a list of examples of several interesting uncertainty sets and corresponding bounds. In any model, the uncertainty set is the modeler's choice. Our bound offers important insights in this choice of uncertainty set as well.

While we show that the stochastic problem with only right hand side uncertainty can be well approximated, the static robust solution and the finitely adaptable solution are not a good approximation for the case where both the right hand side and the objective coefficients are uncertain. However, for the adaptive optimization problem, we show that the static robust and the finitely adaptable solution are a good approximation for the two-stage and multi-stage version respectively, even when both the right hand side and the objective coefficients are uncertain. The performance bound in this case is $(1 + \rho/s)^2$ where again ρ is the translation factor of the uncertainty sets and s is the symmetry. This bound is not as strong as the bound for the stochastic problem. We also present an alternate geometric bound for this case.

3.10 Appendix

3.10.1 Examples: Symmetry of Specific Sets

Our main tool in calculating the symmetry of a convex compact set is the following proposition in Belloni and Freund [5],

Proposition 2 (Belloni and Freund [5]). *Let S be a convex body, and consider the representation of S as the intersection of halfspaces: $S = \{x \in \mathbb{R}^m | a_i^T \boldsymbol{x} \leq b_i,\ i \in I\}$ for some (possibly unbounded) index set I, and let $\delta_i^* := \max_{\boldsymbol{x} \in S}\{-a_i^T \boldsymbol{x}\}$ for $i \in I$. Then for all $\boldsymbol{x} \in S$,*

$$\mathbf{sym}(\boldsymbol{x}, S) = \inf_{i \in I} \left\{ \frac{b_i - a_i^T \boldsymbol{x}}{\delta_i^* + a_i^T \boldsymbol{x}} \right\}.$$

General L_p half-ball for $p \geq 1$.

Define an L_p half-ball as

$$\mathbf{HB}_p := \{\boldsymbol{b} \in \mathbb{R}^m \mid \|\boldsymbol{b}\|_p \leq 1, b_1 \geq 0\}.$$

The dual norm is the L_q norm with $\frac{1}{p} + \frac{1}{q} = 1$. The symmetry of L_p half-ball is summarized as follows,

Proposition 3. *The symmetry and point of symmetry of an L_p half-ball are,*

$$\mathbf{sym}(\mathbf{HB}_p) = \left(\frac{1}{2}\right)^{\frac{1}{p}}, \quad \boldsymbol{b}_0(\mathbf{HB}_p) = \frac{1}{2^{\frac{1}{p}} + 1}\boldsymbol{e}_1.$$

Proof. The L_p half-ball can be represented by halfspaces as

$$\mathbf{HB}_p = \{\boldsymbol{b} \in \mathbb{R}^m \mid -b_1 \leq 0, \boldsymbol{\pi}^T\boldsymbol{b} \leq 1, \forall \|\boldsymbol{\pi}\|_q \leq 1, \pi_1 \geq 0\}.$$

For each $\boldsymbol{\pi}$ in the dual unit ball (i.e., $\|\boldsymbol{\pi}\|_q = 1$), define

$$\delta^*(\boldsymbol{\pi}) := \max_{\|\boldsymbol{b}\|_p \leq 1, b_1 \geq 0} -\boldsymbol{\pi}^T\boldsymbol{b},$$

whose optimum $\boldsymbol{b}^*$ satisfies $b_1^* = 0, \|\boldsymbol{b}^*\|_p = 1$. Therefore we have,

$$\delta^*(\boldsymbol{\pi}) = \max_{\|\tilde{\boldsymbol{b}}\|_q = 1} -\tilde{\boldsymbol{\pi}}^T\tilde{\boldsymbol{b}} = \|\tilde{\boldsymbol{\pi}}\|_q = (1 - \pi_1^q)^{1/q},$$

where $\boldsymbol{b} = (b_1; \tilde{\boldsymbol{b}})$ and $\boldsymbol{\pi} = (\pi_1; \tilde{\boldsymbol{\pi}})$. Also define $\delta^*(\boldsymbol{e}_1) := \max_{\boldsymbol{b} \in \mathbf{HB}_p} \boldsymbol{b}_1 = 1$. Now we can compute the symmetry of $\mathbf{HB}_p$. Due to the geometry, the point of symmetry has the form $\boldsymbol{b}_0 = \alpha\boldsymbol{e}_1$. We have,

$$\mathbf{sym}(\mathbf{HB}_p) = \max_{\alpha \in [0,1]} \min\left\{ \inf_{\pi_1 \in [0,1]} \frac{1 - \pi_1\alpha}{(1 - \pi_1^q)^{\frac{1}{q}} + \pi_1\alpha}, \frac{\alpha}{1 - \alpha} \right\}.$$

The maximum is achieved when the inf term is equal to the second term, because the inf term is decreasing in α and the second term is increasing in α. The inf term has optimality condition,

$$-\alpha((1-\pi_1^q)^{\frac{1}{q}}+\pi_1\alpha)=(1-\pi_1\alpha)\bigg(\frac{-\pi_1^{q-1}}{(1-\pi_1)^{\frac{1}{p}}}+\alpha\bigg).$$

Therefore, we have,

$$\frac{\alpha}{\frac{-\pi_1^{q-1}}{(1-\pi_1)^{\frac{1}{p}}}-\alpha}=\frac{\alpha}{1-\alpha},$$

which implies that $\frac{-\pi_1^{q-1}}{(1-\pi_1)^{\frac{1}{p}}}=1$, i.e., $\pi_1=(\frac{1}{2})^{\frac{1}{q}}$, and $\alpha=\frac{1}{2^{\frac{1}{p}}+1}$. □

Intersection of an L_p ball with $\mathbb{R}_+^m$ for $p\geq 1$.

Recall that an L_p-ball intersected with the nonnegative orthant is defined as $\mathbf{B}_p^+ := \{\boldsymbol{b}\in\mathbb{R}^m \mid \|\boldsymbol{b}\|_p\leq 1, \boldsymbol{b}\geq 0\}$. We show that the symmetry and symmetry point of $\mathbf{B}_p^+$ are as defined in (3.18). The L_p half-ball can be represented by halfspaces as

$$\mathbf{B}_p^+=\{\boldsymbol{b}\in\mathbb{R}^m \mid -\boldsymbol{b}\leq\mathbf{0}, \boldsymbol{\pi}^T\boldsymbol{b}\leq 1, \forall\|\boldsymbol{\pi}\|_q\leq 1, \boldsymbol{\pi}\geq\mathbf{0}\},$$

where $\|\cdot\|_q$ is the dual-norm with $\frac{1}{p}+\frac{1}{q}=1$. Therefore, define $\delta^*(\boldsymbol{\pi}):=\max\limits_{\boldsymbol{b}\in\mathbf{B}_p^+}-\boldsymbol{\pi}^T\boldsymbol{b}=0$, and $\delta^*(\boldsymbol{e}_k):=\max\limits_{\boldsymbol{b}\in\mathbf{B}_p^+} b_k=1$. Therefore, the symmetry can be computed as,

$$\mathbf{sym}(\mathbf{B}_p^+):=\max_{\alpha\in[0,(\frac{1}{m})^{\frac{1}{p}}]}\min\left\{\inf_{\|\boldsymbol{\pi}\|_q=1,\boldsymbol{\pi}\geq\mathbf{0}}\frac{1-\alpha\boldsymbol{e}^T\boldsymbol{\pi}}{\alpha\boldsymbol{e}^T\boldsymbol{\pi}}, \frac{\alpha}{1-\alpha}\right\},$$

where we use the property that the symmetry point $\boldsymbol{b}_0=\alpha\boldsymbol{e}$ for some $\alpha\in[0,1/\|\boldsymbol{e}\|_p]$. The maximum is achieved when the inf term is equal to the second term. The inf term can be computed, since $\max\limits_{\|\boldsymbol{\pi}\|_q=1,\boldsymbol{\pi}\geq\mathbf{0}}\boldsymbol{e}^T\boldsymbol{\pi}=\|\boldsymbol{e}\|_p$. Thus, at the symmetry point, we have $\dfrac{1}{\alpha\|\boldsymbol{e}\|_p}-1=\dfrac{\alpha}{1-\alpha}$. Therefore, $\alpha=\dfrac{1}{\|\boldsymbol{e}\|_p+1}=\dfrac{1}{m^{\frac{1}{p}}+1}$, which gives the results.

Ellipsoidal uncertainty set.

An ellipsoidal uncertainty set that is contained in the nonnegative orthant can be defined as,

$$\mathcal{U} := \{\boldsymbol{b} \mid \|\boldsymbol{E}(\boldsymbol{b}-\bar{\boldsymbol{b}})\|_2 \leq 1\} \subset \mathbb{R}^m_+. \tag{3.46}$$

We assume the ellipsoid has full dimension, thus, $\bar{\boldsymbol{b}} > \boldsymbol{0}$. The symmetry of an ellipsoid is 1. But the translation factor depends on the position of the center $\bar{\boldsymbol{b}}$. The following proposition computes the translation factor.

Proposition 4. *Assume the uncertainty set $\mathcal{U}$ is defined in (3.46). The translation factor $\rho(\bar{\boldsymbol{b}}, \mathcal{U})$ is given as,*

$$\rho(\bar{\boldsymbol{b}}, \mathcal{U}) = \max_{1 \leq i \leq m} \frac{\sqrt{E_{ii}^{-1}}}{\bar{b}_i},$$

where E_{ii}^{-1} is the i-th diagonal element of the inverse matrix E^{-1}.

Proof. From the definition, the translation factor is the smallest ρ such that

$$b_i^l := \min\{b_i \| \|\boldsymbol{E}(\boldsymbol{b}-\rho\bar{\boldsymbol{b}})\|_2 \leq 1\} \geq 0, \quad \forall i = 1, 2, \ldots, m.$$

From the optimality conditions, we can get that $b_i^l = \rho\bar{b}_i - \sqrt{E_{ii}^{-1}}$, which gives the result. □

Intersection of two L_p balls with $\mathbb{R}^m_+$.

Consider the uncertainty set $\mathcal{U}$ defined in (3.20) where $1 \leq p_1 < p_2$, $0 < r < 1$, and suppose (3.21) holds. We show that the symmetry and symmetry point of $\mathcal{U}$ are as defined in (3.22).

The uncertainty set $\mathcal{U}$ can be represented by the intersection of halfspaces,

$$\mathcal{U} = \{\boldsymbol{b} \in \mathbb{R}^m \mid -\boldsymbol{b} \leq \boldsymbol{0}, \boldsymbol{\pi}^T\boldsymbol{b} \leq 1, \boldsymbol{\lambda}^T\boldsymbol{b} \leq r, \forall \|\boldsymbol{\pi}\|_{q_1} \leq 1, \boldsymbol{\pi} \geq \boldsymbol{0}, \|\boldsymbol{\lambda}\|_{q_2} \leq r, \boldsymbol{\lambda} \geq \boldsymbol{0}\},$$

where $\frac{1}{p_1}+\frac{1}{q_1}=1, \frac{1}{p_2}+\frac{1}{q_2}=1$. Compute the following quantities,

$$\begin{aligned}
\delta^*(\boldsymbol{\pi}) &:= \max_{\boldsymbol{b}\in\mathcal{U}} -\boldsymbol{\pi}^T\boldsymbol{b} = 0, \quad \forall \|\boldsymbol{\pi}\|_{q_1} \le 1, \boldsymbol{\pi} \ge \mathbf{0}, \\
\delta^*(\boldsymbol{\lambda}) &:= \max_{\boldsymbol{b}\in\mathcal{U}} -\boldsymbol{\lambda}^T\boldsymbol{b} = 0, \quad \forall \|\boldsymbol{\lambda}\|_{q_2} \le r, \boldsymbol{\lambda} \ge \mathbf{0}, \\
\delta^*(\boldsymbol{e}_k) &:= \max_{\boldsymbol{b}\in\mathcal{U}} b_k = r, \quad \forall k = 1,\ldots,m.
\end{aligned}$$

Since the set $\mathcal{U}$ is symmetric with respect to the direction $\boldsymbol{e}$, the symmetry point must have the form $\boldsymbol{b}_0 = \alpha\boldsymbol{e}$. The symmetry can be computed as,

$$\mathbf{sym}(\mathcal{U}) := \max_{\alpha\in[0,(\frac{1}{m})^{\frac{1}{p_1}}]} \min\left\{ \inf_{\|\boldsymbol{\pi}\|_{q_1}=1,\boldsymbol{\pi}\ge\mathbf{0}} \frac{1-\alpha\boldsymbol{e}^T\boldsymbol{\pi}}{\alpha\boldsymbol{e}^T\boldsymbol{\pi}}, \inf_{\|\boldsymbol{\lambda}\|_{q_2}=1,\boldsymbol{\lambda}\ge\mathbf{0}} \frac{1-\alpha\boldsymbol{e}^T\boldsymbol{\lambda}}{\alpha\boldsymbol{e}^T\boldsymbol{\lambda}}, \frac{\alpha}{r-\alpha} \right\}, \tag{3.47}$$

As is shown in the proof of the previous proposition, we have,

$$\begin{aligned}
\inf_{\|\boldsymbol{\pi}\|_{q_1}=1,\boldsymbol{\pi}\ge\mathbf{0}} \frac{1-\alpha\boldsymbol{e}^T\boldsymbol{\pi}}{\alpha\boldsymbol{e}^T\boldsymbol{\pi}} &= \frac{1}{\alpha\|\boldsymbol{e}\|_{p_1}} - 1, \\
\inf_{\|\boldsymbol{\lambda}\|_{q_2}=1,\boldsymbol{\lambda}\ge\mathbf{0}} \frac{1-\alpha\boldsymbol{e}^T\boldsymbol{\lambda}}{\alpha\boldsymbol{e}^T\boldsymbol{\lambda}} &= \frac{r}{\alpha\|\boldsymbol{e}\|_{p_2}} - 1.
\end{aligned}$$

Using condition (3.21), we know the first term is dominated by the second term for any α, therefore, the maximum in the symmetry formula (3.47) is achieved when

$$\frac{1}{\alpha\|\boldsymbol{e}\|_{p_1}} - 1 = \frac{\alpha}{r-\alpha}.$$

Therefore, $\alpha = \dfrac{r}{r\|\boldsymbol{e}\|_{p_1}+1} = \dfrac{r}{rm^{\frac{1}{p_1}}+1}$, which gives the results.

Budgeted uncertainty set.

An important type of uncertainty sets is the budgeted uncertainty set, Δ_k as defined in (3.23). We show that the symmetry and the symmetry point are as defined in (3.24).

First, we observe that the symmetry point $\boldsymbol{b}_0(\Delta_k)$ must be of the form $\boldsymbol{b}_0 = \alpha\boldsymbol{e}$, due to the geometry of Δ_k. Then, use Proposition 2, for $k \geq 1$, we have,

$$\mathbf{sym}(\Delta_k) = \max_{0 \leq \alpha \leq \frac{k}{m}} \min\left\{\frac{\alpha}{1-\alpha}, \frac{k - m\alpha}{m\alpha}\right\}.$$

By the monotonicity of $\frac{\alpha}{1-\alpha}$ and $\frac{k-m\alpha}{m\alpha}$, the maximum is achieved at $\alpha = \frac{k}{m+k}$. Thus, the symmetry point of Δ_k is $\boldsymbol{b}_0 = \frac{k}{m+k}e$ and $\mathbf{sym}(\Delta_k) = \frac{k}{m}$.

Demand uncertainty set.

Proof. Proposition 1 For $\mu \geq \Gamma$, the hypercube centered at $\mu\boldsymbol{e}$ is completely contained in the positive orthant. For each S, define

$$\delta^*(S_+) := \max_{\boldsymbol{b} \in \mathbf{DU}} -\boldsymbol{e}_S^T\boldsymbol{b} = -|S|\mu + \sqrt{|S|}\Gamma,$$

$$\delta^*(S_-) := \max_{\boldsymbol{b} \in \mathbf{DU}} \boldsymbol{e}_S^T\boldsymbol{b} = |S|\mu + \sqrt{|S|}\Gamma.$$

The symmetry of $\boldsymbol{b}$ is given as follows,

$$\mathbf{sym}(\boldsymbol{b}, \mathbf{DU}) = \min_{S \subseteq \mathbf{DU}} \left\{\frac{|S|\mu + \sqrt{|S|}\Gamma - \boldsymbol{e}_S^T\boldsymbol{b}}{-|S|\mu + \sqrt{|S|}\Gamma + \boldsymbol{e}_S^T\boldsymbol{b}}, \frac{-|S|\mu + \sqrt{|S|}\Gamma + \boldsymbol{e}_S^T\boldsymbol{b}}{|S|\mu + \sqrt{|S|}\Gamma - \boldsymbol{e}_S^T\boldsymbol{b}}\right\}.$$

Therefore, $\mathbf{sym}(\mu\boldsymbol{e}, \mathbf{DU}) = 1$, which proves (3.26).

For $\frac{1}{\sqrt{m}}\Gamma < \mu < \Gamma$, μ is in one of the intervals of $[\frac{1}{\sqrt{2}}\Gamma, \Gamma), [\frac{1}{\sqrt{3}}\Gamma, \frac{1}{\sqrt{2}}\Gamma], \ldots, (\frac{1}{\sqrt{m}}\Gamma, \frac{1}{\sqrt{m-1}}\Gamma]$. We show that (3.27) holds for $\mu \in [\frac{1}{\sqrt{2}}\Gamma, \Gamma]$. A similar argument works for other intervals.

Let us assume $\mu \in [\frac{1}{\sqrt{2}}\Gamma, \Gamma]$. Notice that $|S|\mu - \sqrt{|S|}\Gamma \geq 0$ for all $|S| \geq 2$. Therefore, the constraints in the definition of $\mathbf{DU}$ (3.25) can be written as,

$$|S|\mu - \sqrt{|S|}\Gamma \leq \sum_{i \in S} b_i \leq |S|\mu + \sqrt{|S|}\Gamma, \quad \forall |S| \geq 2.$$

But for all $|S| = 1$, since $\mu - \Gamma < 0$, we will have $0 \leq b_i \leq \mu + \Gamma$ for all $i = 1, \ldots, m$. Thus, according to Proposition 2, we can compute the following quantities for all

$|S| \geq 2$,

$$\delta^*(S_+) := \max_{\boldsymbol{b}\in\mathbf{DU}} -\boldsymbol{e}_S^T\boldsymbol{b} = -|S|\mu + \sqrt{|S|}\Gamma,$$
$$\delta^*(S_-) := \max_{\boldsymbol{b}\in\mathbf{DU}} \boldsymbol{e}_S^T\boldsymbol{b} = |S|\mu + \sqrt{|S|}\Gamma,$$

and for $|S| = 1$, we have $\delta^*(i_+) := \max_{\boldsymbol{b}\in\mathbf{DU}} -b_i = 0$, $\delta^*(i_-) := \max_{\boldsymbol{b}\in\mathbf{DU}} b_i = \mu + \Gamma$ for all $i = 1, \ldots, m$. Since $\mathbf{DU}$ is symmetric about the direction $\boldsymbol{e}$, the point of symmetry has the form $\boldsymbol{b}_0 = \alpha\boldsymbol{e}$. The range of possible α is determined by the constraint $m\mu - \sqrt{m}\Gamma \leq m\alpha \leq m\mu + \sqrt{m}\Gamma$. Now we can compute the symmetry of $\alpha\boldsymbol{e}$,

$$\begin{aligned}\mathbf{sym}(\alpha\boldsymbol{e}, \mathbf{DU}) = \min\Bigg\{ & \frac{(\mu+\Gamma)-\alpha}{\alpha}, \frac{\alpha}{(\mu+\Gamma)-\alpha}, \frac{(|S|\mu + \sqrt{|S|}\Gamma) - |S|\alpha}{(-|S|\mu + \sqrt{|S|}\Gamma) + |S|\alpha}, \\ & \frac{(-|S|\mu + \sqrt{|S|}\Gamma) + |S|\alpha}{(|S|\mu + \sqrt{|S|}\Gamma) - |S|\alpha}, \forall |S| \geq 2\Bigg\}.\end{aligned} \tag{3.48}$$

The symmetry of $\mathbf{DU}$ is given as,

$$\mathbf{sym}(\mathbf{DU}) = \max\{\mathbf{sym}(\alpha\boldsymbol{e}, \mathbf{DU}) : \alpha \in [\mu - \frac{1}{\sqrt{m}}\Gamma, \mu + \frac{1}{\sqrt{m}}\Gamma]\}.$$

Observe that $\frac{(|S|\mu+\sqrt{|S|}\Gamma)-|S|\alpha}{(-|S|\mu+\sqrt{|S|}\Gamma)+|S|\alpha}$ is a decreasing function in α, and $\frac{(-|S|\mu+\sqrt{|S|}\Gamma)+|S|\alpha}{(|S|\mu+\sqrt{|S|}\Gamma)-|S|\alpha}$ is an increasing function in α. Both functions have value 1 when $\alpha = \mu$, and decrease when $|S|$ increases. The formula (3.48) can be simplified as,

$$\begin{aligned}\mathbf{sym}(\alpha\boldsymbol{e}, \mathbf{DU}) = \min\Bigg\{ & \frac{(\mu+\Gamma)-\alpha}{\alpha}, \frac{\alpha}{(\mu+\Gamma)-\alpha}, \frac{(m\mu + \sqrt{m}\Gamma) - m\alpha}{(-m\mu + \sqrt{m}\Gamma) + m\alpha}, \\ & \frac{(-m\mu + \sqrt{m}\Gamma) + m\alpha}{(m\mu + \sqrt{m}\Gamma) - m\alpha}\Bigg\}.\end{aligned} \tag{3.49}$$

The maximum of $\mathbf{sym}(\alpha\boldsymbol{e}, \mathbf{DU})$ is achieved at the following intersection,

$$\frac{\alpha}{(\mu+\Gamma)-\alpha} = \frac{(m\mu + \sqrt{m}\Gamma) - m\alpha}{(-m\mu + \sqrt{m}\Gamma) + m\alpha},$$

which gives the symmetry and the symmetry point in (3.27). Since (3.49) also holds when μ is in any other interval, it implies that (3.27) is true for all $\frac{1}{\sqrt{m}}\Gamma < \mu < \Gamma$.

When $0 \leq \mu \leq \frac{1}{\sqrt{m}}\Gamma$, following a similar argument, we have,

$$\mathbf{sym}(\alpha \boldsymbol{e}, \mathbf{DU}) = \min_{|S|\geq 1}\left\{\frac{(|S|\mu + \sqrt{|S|}\Gamma) - |S|\alpha}{|S|\alpha}, \frac{|S|\alpha}{(|S|\mu + \sqrt{|S|}\Gamma) - |S|\alpha}\right\}.$$

Then, the maximum is achieved when $\dfrac{\alpha}{(\mu+\Gamma)-\alpha} = \dfrac{(m\mu + \sqrt{m}\Gamma) - m\alpha}{m\alpha}$, which gives the results in (3.28). □

Parallel slabs.

A parallel slab is defined as $\mathbf{PS} := \{\boldsymbol{b} \in \mathbb{R}^m_+ \mid L \leq \boldsymbol{e}^T\boldsymbol{b} \leq U\}$ for $0 \leq L \leq U$ and $U > 0$. The symmetry and symmetry point of a parallel slab is given as,

Proposition 5.

$$\mathbf{sym}(\mathbf{PS}) = \frac{1}{m - \frac{L}{U}}, \tag{3.50}$$

$$\boldsymbol{b}_0(\mathbf{PS}) = \frac{U^2}{(m+1)U - L}\boldsymbol{e}. \tag{3.51}$$

Proof. Define the following quantities,

$$\delta^*_L := \max_{\boldsymbol{b}\in\mathbf{PS}} \boldsymbol{e}^T\boldsymbol{b} = U, \quad \delta^*_U := \max_{\boldsymbol{b}\in\mathbf{PS}} -\boldsymbol{e}^T\boldsymbol{b} = -L,$$

and for each $k = 1, \ldots, m$,

$$\delta^*(\boldsymbol{e_k}) := \max_{\boldsymbol{b}\in\mathbf{PS}} \boldsymbol{e_k}^T\boldsymbol{b} = U.$$

According to the geometry of the set, the point of symmetry is of the form $\boldsymbol{b}_0 = \alpha\boldsymbol{e}$. Thus, the symmetry can be written as,

$$\mathbf{sym}(\mathbf{PS}) = \max_{\alpha\in[\frac{L}{m},\frac{U}{m}]} \min\left\{\frac{m\alpha - L}{U - m\alpha}, \frac{U - m\alpha}{m\alpha - L}, \frac{\alpha}{U-\alpha}\right\}.$$

The first two terms are inverse of each other, therefore one increasing and the other decreasing in α; the third term is increasing in α. Thus, the maximum is attained when the second term is equal to the third, i.e.,

$$\frac{U - m\alpha}{m\alpha - L} = \frac{\alpha}{U - \alpha},$$

which gives the results. $\square$

$\mathcal{U} = \mathbf{Conv}(\Delta_1, \{\boldsymbol{e}\})$.

The convex hull of the standard m-simplex $\Delta_1 = \{\boldsymbol{b} \in \mathbb{R}^m \mid \boldsymbol{e}^T\boldsymbol{b} \leq 1, \boldsymbol{b} \geq \boldsymbol{0}\}$ and a point $\boldsymbol{e} = (1, \dots, 1)^T$ has a slightly improved symmetry comparing with the simplex itself, as shown below.

Proposition 6. *The symmetry and point of symmetry of $\mathcal{U} = \mathbf{Conv}(\Delta_1, \{\boldsymbol{e}\})$ are given as,*

$$\mathbf{sym}(\mathcal{U}) = \frac{1}{m-1}, \quad \boldsymbol{b}_0(\mathcal{U}) = \frac{1}{m}\boldsymbol{e}.$$

Proof. The set $\mathcal{U}$ can be represented by halfspaces as

$$\mathcal{U} = \{\boldsymbol{b} \in \mathbb{R}^m \mid -\boldsymbol{b} \leq \boldsymbol{0}, \boldsymbol{\pi}_i^T\boldsymbol{b} \leq 1, \forall i = 1, \dots, m\},$$

where $\boldsymbol{\pi}_i$ has $(2-m)$ in the i-th entry and 1 elsewhere. For each i, define,

$$\delta_i^* = \max_{\boldsymbol{b} \in \mathcal{U}} \boldsymbol{b}_i = 1, \quad \delta(\boldsymbol{\pi}_i)^* = \max_{\boldsymbol{b} \in \mathcal{U}} -\boldsymbol{\pi}_i^T\boldsymbol{b} = m - 2.$$

By the symmetry of $\mathcal{U}$, we know the symmetry point should be on the $\boldsymbol{e}$ direction. Therefore, we have,

$$\mathbf{sym}(\mathcal{U}) = \max_{\alpha \in [0,1]} \min\left\{\frac{1 - \alpha\boldsymbol{\pi}_i^T\boldsymbol{e}}{(m-2) + \alpha\boldsymbol{\pi}_i^T\boldsymbol{e}}, \frac{\alpha}{1-\alpha}\right\}, \tag{3.52}$$

$$= \max_{\alpha \in [0,1]} \min\left\{\frac{1-\alpha}{m-2+\alpha}, \frac{\alpha}{1-\alpha}\right\}. \tag{3.53}$$

The maximum is achieved when $\frac{1-\alpha}{m-2+\alpha} = \frac{\alpha}{1-\alpha}$, which gives $\alpha = \frac{1}{m}$ and symmetry is $\frac{1}{m-1}$. $\square$

The matrix simplex.

Define the uncertainty set $\mathcal{U}$ as follows,

$$\mathcal{U} = \{\boldsymbol{B} \in \mathbb{S}_+^m \mid \boldsymbol{I} \bullet \boldsymbol{B} \leq 1\}, \tag{3.54}$$

where $\mathbb{S}_+^m$ is the cone of positive semidefinite matrices, $\boldsymbol{I}$ is the identity matrix. We use it as an example of uncertainty sets for robust SDP problems. The following proposition shows that $\mathcal{U}$ has a similar symmetry property as a simplex in $\mathbb{R}^m$.

Proposition 7. *Let $\mathcal{U}$ be defined in (3.54) and $\boldsymbol{B}_0$ be the center of symmetry of $\mathcal{U}$. We have,*

$$\mathbf{sym}(\mathcal{U}) = \frac{1}{m}, \quad \boldsymbol{B}_0 = \frac{1}{m+1}\boldsymbol{I}.$$

Proof. The set $\mathcal{U}$ can be written in the form of intersection of halfspaces,

$$\mathcal{U} = \{\boldsymbol{B} \in \mathbb{S}^m \mid \boldsymbol{I} \bullet \boldsymbol{B} \leq 1, (\boldsymbol{q}\boldsymbol{q}^T) \bullet \boldsymbol{B} \leq 1, \forall \|\boldsymbol{q}\|_2 = 1\}.$$

Then, we can compute the following quantities,

$$\delta^*(\boldsymbol{q}) = \max_{\boldsymbol{B}\in\mathcal{U}} \quad (\boldsymbol{q}\boldsymbol{q}^T) \bullet \boldsymbol{B} = \max_{\boldsymbol{p}:\|\boldsymbol{p}\|=1} (\boldsymbol{q}^T\boldsymbol{p})^2 = 1, \tag{3.55}$$

$$\delta^*(\boldsymbol{I}) = -\min_{\boldsymbol{B}\in\mathcal{U}} \quad \boldsymbol{I} \bullet \boldsymbol{B} = 0. \tag{3.56}$$

The second equality in (3.55) comes from the fact that the extreme points of $\mathcal{U}$ are rank-1 matrices $\boldsymbol{p}\boldsymbol{p}^T$ with $\boldsymbol{p}^T\boldsymbol{p} = 1$. (3.56) follows from $\boldsymbol{I} \bullet \boldsymbol{B} \geq 0$ and the minimum is achieved at $\boldsymbol{B} = \boldsymbol{0}$. Now we can compute the symmetry of $\mathcal{U}$ as,

$$\mathbf{sym}(\mathcal{U}) = \max_{\boldsymbol{B}\in\mathcal{U}} \min \left\{ \frac{1 - \boldsymbol{I} \bullet \boldsymbol{B}}{\boldsymbol{I} \bullet \boldsymbol{B}}, \inf_{\boldsymbol{q}:\|\boldsymbol{q}\|=1} \frac{(\boldsymbol{q}\boldsymbol{q}^T) \bullet \boldsymbol{B}}{1 - (\boldsymbol{q}\boldsymbol{q}^T) \bullet \boldsymbol{B}} \right\}. \tag{3.57}$$

To compute the inf term is equivalent to solving the following problem, which admits a closed form solution by Courant minimax theorem: $\lambda_1(\boldsymbol{B}) = \min_{\boldsymbol{q}:\|\boldsymbol{q}\|=1} \boldsymbol{q}^T\boldsymbol{B}\boldsymbol{q}$, where $\lambda_1(\boldsymbol{B})$ is the smallest eigenvalue of $\boldsymbol{B}$, denoted as $(\lambda_1, \ldots, \lambda_m)$ in ascending order. Then, (3.57) can be rewritten as,

$$\mathbf{sym}(\mathcal{U}) = \max_{\boldsymbol{e}^T\boldsymbol{\lambda}\leq 1, \boldsymbol{\lambda}\geq \mathbf{0}} \min\left\{\frac{1-\boldsymbol{e}^T\boldsymbol{\lambda}}{\boldsymbol{e}^T\boldsymbol{\lambda}}, \frac{\lambda_1}{1-\lambda_1}\right\}, \tag{3.58}$$

which can be reformulated as,

$$\mathbf{sym}(\mathcal{U}) = \max_{z,\boldsymbol{\lambda}}\{z \; : \; \boldsymbol{e}^T\boldsymbol{\lambda} \leq \frac{1}{z+1}, \boldsymbol{e}^T\boldsymbol{\lambda} \leq 1, \; \boldsymbol{\lambda} \geq \mathbf{0}, \lambda_i \geq \frac{z}{z+1}, \forall i = 1, \ldots, m.\}$$

From the constraints, we have

$$\frac{1}{z+1} \geq \boldsymbol{e}^T\boldsymbol{\lambda} \geq \frac{mz}{z+1} \Rightarrow z \leq \frac{1}{m}.$$

In fact, $z = 1/m$ can be achieve by the feasible solution $\lambda_i = 1/(m+1)$ for all i. This completes the proof. □

3.10.2 Proof of Theorem 6.

Proof. Theorem 6 Let $\mathcal{U}'$ be the translation of the uncertainty set $\mathcal{U}$ with the translation factor ρ, i.e. $\mathcal{U}' = \mathcal{U} - (1-\rho)\boldsymbol{b}_0$. Let $\boldsymbol{b}_1 := \boldsymbol{b}_0 - (1-\rho)\boldsymbol{b}_0$. Also let $\boldsymbol{z} := \boldsymbol{b}_0 - \boldsymbol{b}_1 = (1-\rho)\boldsymbol{b}_0$. Following a similar argument as in the proof of Theorem 1, we have

$$\left(1+\frac{1}{s}\right)\boldsymbol{b}_1 \succeq_{\mathcal{K}} \boldsymbol{b}', \quad \forall \boldsymbol{b}' \in \mathcal{U}'.$$

Translating back to the original uncertainty set $\mathcal{U}$ by adding $\boldsymbol{z}$ on both sides, we have,

$$\begin{aligned}
&\left(1+\frac{1}{s}\right)\boldsymbol{b}_1 + \boldsymbol{z} \succeq_{\mathcal{K}} \boldsymbol{b}, \quad \forall \boldsymbol{b} \in \mathcal{U}, \\
\Rightarrow &\left(1+\frac{1}{s}\right)\rho\boldsymbol{b}_0 + (1-\rho)\boldsymbol{b}_0 \succeq_{\mathcal{K}} \boldsymbol{b}, \quad \forall \boldsymbol{b} \in \mathcal{U}, \\
\Rightarrow &\left(1+\frac{\rho}{s}\right)\boldsymbol{b}_0 \succeq_{\mathcal{K}} \boldsymbol{b}, \quad \forall \boldsymbol{b} \in \mathcal{U}.
\end{aligned} \tag{3.59}$$

For brevity, let $\tau := (1+\rho/s)$. Suppose $(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{b}), \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U})$ is an optimal solution of the stochastic optimization problem (3.45). We have,

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}(\tau \boldsymbol{y}(\boldsymbol{b})) \succeq_{\mathcal{K}} \tau \boldsymbol{b},\ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}.$$

Equivalently,

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}(\tau \boldsymbol{y}(\boldsymbol{b})) - \tau \boldsymbol{b} \in \mathcal{K},\ \mu\text{-a.e. } \boldsymbol{b} \in \mathcal{U}.$$

Therefore, the expectation with respect to the probability measure μ satisfies,

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}(\mathbb{E}_{\mu}[\tau \boldsymbol{y}(\boldsymbol{b})]) - \mathbb{E}_{\mu}[\tau \boldsymbol{b}] \in \mathcal{K}.$$

Thus, we have

$$\boldsymbol{A}(\tau \boldsymbol{x}) + \boldsymbol{B}(\mathbb{E}_{\mu}[\tau \boldsymbol{y}(\boldsymbol{b})]) \succeq_{\mathcal{K}} \mathbb{E}_{\mu}[\tau \boldsymbol{b}] \succeq_{\mathcal{K}} \tau \boldsymbol{b}_0 \succeq_{\mathcal{K}} \boldsymbol{b},\ \forall \boldsymbol{b} \in \mathcal{U}.$$

The last inequality uses (3.59). Therefore, $(\tau \boldsymbol{x}, \mathbb{E}_{\mu}[\tau \boldsymbol{y}(\boldsymbol{b})])$ is a feasible static robust solution. Therefore, the cost of an optimal static robust solution is at most,

$$\boldsymbol{c}^T(\tau \boldsymbol{x}) + \boldsymbol{d}^T \mathbb{E}_{\mu}[\tau \boldsymbol{y}(\boldsymbol{b})] = \tau(\boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \mathbb{E}_{\mu}[\boldsymbol{y}(\boldsymbol{b})]).$$

Furthermore, $z^{\mathcal{K}}_{\text{Stoch}} = \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{d}^T \mathbb{E}_{\mu}[\boldsymbol{y}(\boldsymbol{b})]$, which implies that the cost of an optimal static solution is at most $\tau \cdot z^{\mathcal{K}}_{\text{Stoch}}$. $\square$

No.	Uncertainty set	Symmetry	Stochasticity gap
1	$\{\boldsymbol{b} : \lVert\boldsymbol{b}\rVert_p \le 1, \boldsymbol{b} \ge \boldsymbol{0}\}$	$\frac{1}{m^{1/p}}$	$1 + m^{\frac{1}{p}}$ (a.b)
2	$\{\boldsymbol{b} : \lVert\boldsymbol{b} - \bar{\boldsymbol{b}}\rVert_p \le 1, b_1 \ge \bar{b}_1\} \subset \mathbb{R}^m_+$ $1 \le p < \infty$	$\frac{1}{2^{1/p}}$	$1 + \max_{1\le i\le m} \bar{b}_i^{-1}$ (b)
3	$\{\boldsymbol{b} : \lVert\boldsymbol{b} - \bar{\boldsymbol{b}}\rVert_\infty \le 1, b_1 \ge \bar{b}_1\} \subset \mathbb{R}^m_+$	1	$1 + \max_{1\le i\le m} \bar{b}_i^{-1}$ (a.b)
4	$\{\boldsymbol{b} : \lVert\boldsymbol{b} - \bar{\boldsymbol{b}}\rVert_p \le 1\} \subset \mathbb{R}^m_+$	1	$1 + \max_{1\le i\le m} \bar{b}_i^{-1}$ (a.b)
5	$\{\boldsymbol{b} : \lVert\boldsymbol{E}(\boldsymbol{b} - \bar{\boldsymbol{b}})\rVert_2 \le 1\} \subset \mathbb{R}^m_+$	1	$1 + \max_{1\le i\le m} \frac{E_{ii}^{-1}}{\bar{b}_i}$ (a.b)
6	$\{\boldsymbol{b} : \lVert\boldsymbol{b}\rVert_{p_1} \le 1, \lVert\boldsymbol{b}\rVert_{p_2} \le r, \boldsymbol{b} \ge \boldsymbol{0}\}$	$\frac{1}{rm^{1/p_1}}$	$1 + rm^{\frac{1}{p_1}}$ (a.b)
7	Budgeted uncertainty set Δ_k $(1 \le k \le m)$	$\frac{k}{m}$	$1 + \frac{m}{k}$ (a.b)
8	Demand uncertainty set $\mathbf{DU}(\mu, \Gamma)$ $(\mu \ge \Gamma)$	1	$1 + \frac{\Gamma}{\mu}$ (a.b)
9	Demand uncertainty set $\mathbf{DU}(\mu, \Gamma)$ $(\frac{1}{\sqrt{m}}\Gamma < \mu < \Gamma)$	$\frac{\sqrt{m}\mu + \Gamma}{(1+\sqrt{m})\Gamma}$	$1 + \frac{(1+\sqrt{m})\Gamma}{\sqrt{m}\mu + \Gamma}$ (a.b)
10	Demand uncertainty set $\mathbf{DU}(\mu, \Gamma)$ $(0 \le \mu \le \frac{1}{\sqrt{m}}\Gamma)$	$\frac{\sqrt{m}\mu + \Gamma}{\sqrt{m}(\mu + \Gamma)}$	$1 + \frac{\sqrt{m}(\mu+\Gamma)}{\sqrt{m}\mu + \Gamma}$ (a.b)
11	$\{\boldsymbol{b} : L \le \boldsymbol{e}^T\boldsymbol{b} \le U, \boldsymbol{b} \ge \boldsymbol{0}\}$	$\frac{1}{m - \frac{L}{U}}$	$(m+1) - \frac{L}{U}$ (a.b)
12	$\mathbf{conv}(\Delta_1, \{\boldsymbol{e}\})$	$\frac{1}{m-1}$	m (a.b)
13	$\mathbf{conv}(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k)$	$\frac{1}{m} \le s \le 1$	$\max_{1\le i\le m} \frac{b_i^h}{\bar{b}_i}$ (b)
14	$\{\boldsymbol{B} \in \mathbb{S}^m_+ \mid \boldsymbol{I} \bullet \boldsymbol{B} \le 1\}$	$\frac{1}{m}$	$1 + m$ (a)

Table 3.1: Symmetry and corresponding stochasticity gap for various uncertainty sets. Note that footnote (a) or (b) means the tight bound is from the symmetry bound (3.13) or the scaling bound (3.17).

Chapter 4

A Fairness-based Proposal for Electricity Market Design

4.1 Introduction

The electricity industry has undergone significant deregulation and restructuring in several regions of the United States since the late 1990s [79]. Vertically integrated utility companies have been restructured; electricity production and transmission were separated into different enterprises; Independent System Operators (ISO) and Regional Transmission Organizations (RTO) have been established to manage the regional transmission grids and wholesale electricity markets. Today, ten ISOs and RTOs serve two-thirds of the electricity consumers in the United States and more than half of the population in Canada. Although more than a decade has passed, some fundamental issues on electricity market design still remain hotly debated in both academia and industry. Our work aims to provide new perspective and proposals to some of these issues.

We focus on the day-ahead (DA) electricity market, which exhibits some of the most difficult issues in market design. A day-ahead market schedules the supply and consumption of electricity for the next day. Two essential components of a DA market include an auction scheme that determines the production and consumption levels of all market participants, and a pricing scheme that determines the electricity price [29, 49]. Most ISOs adopt the social welfare maximization principle and a marginal cost pricing scheme (throughout the chapter, we denote them as MaxSW and MCP, respectively). In this scheme, the auction decisions are made by solving a unit commitment problem — a mixed integer optimization problem with the objective to maximize the total social welfare while satsifying various operation, reliability, and security constraints; the electricity price is determined as the incremental cost of producing one more unit of electricity, which is derived from the dual solution of a linear economic dispatch problem with fixed integer decisions.

A long noticed shortcoming of the MaxSW-MCP scheme is the so-called uplift problem [59]. The uplift problem is that the electricity payment determined by the electricity price cannot fully recover the production cost (especially the fixed cost) of some committed generators; the ISO has to arrange side payments for such generators to make up the loss.

Both academic researchers and industrial practitioners are in active search for a solution. Recent proposals have focused on augmenting the current MaxSW-MCP scheme by introducing prices for integer decisions [67, 88, 50, 44]. Enhancing market transparency and competitiveness is clearly an important goal as claimed by the existing proposals. In our opinion, the uplift problem is a symptom of what we consider a deeper fairness problem in the currently used MaxSW-MCP scheme: some generators that produce electricity make no profit, while other market participants, especially on the consumer side, collect considerable utility surpluses.

To address this inequity, we introduce and investigate the notion of β-fairness that addresses the tradeoff between social welfare and fairness. The case $\beta = 0$ corresponds to current practice of maximizing social welfare, whereas $\beta = 1$ corresponds to a solution that maximizes the minimum utility among market participants, the so-called max-min fairness. Under the max-min fairness principle, every market participant is guaranteed a nonnegative surplus, so no side payments are needed, and thus the uplift problem is naturally solved. By varying β, we investigate the tradeoff between efficiency and fairness of the auction and pricing scheme. We formulate a family of mixed integer optimization models that explicitly parametrizes this tradeoff. The solutions for different parameter values are used to generate an efficiency-fairness curve, which demonstrates a significant improvement of fairness in our proposed schemes. Such a frontier plot also illustrates a flexible tradeoff between efficiency and fairness achieved by our proposal, and thus, it can be used to facilitate the ISO's decision-making in choosing an appropriate operating point of its market. We further show that the current operational practice ($\beta = 0$) is not Pareto efficient, while there is a $\beta_0 < 1$, at which the social welfare is the same as current practice, while the fairness property is strictly better, and the side payments are strictly lower than current practices. This gives a solution to another long noticed problem of the current practice, namely achieve fairness and integrity of the auction in choosing from multiple (near) optimal solutions. We broadly investigate the properties of such β-fair solutions both theoretically and empirically.

The chapter is organized as follows. In Section 4.2, we review the auction and pricing scheme implemented in the current day-ahead electricity markets, and illustrate the uplift problem in a simple example. In Section 4.3, we introduce the concept of max-min fairness, and propose a new auction and pricing scheme based on max-min fairness. We discuss in detail the properties, computational methods, and economic implications of max-min fairness. In Section 4.4, we introduce the notion of β-fairness that interpolates between maximum welfare and max-min fairness and explore its properties. In Section 4.5, we investigate the case of β-fairness as $\beta \to 0$ and propose a new pricing scheme. We show that under this pricing scheme the social welfare is the same as current practice, while the max-min fairness is achieved, and the side payments are strictly lower than current practices. We also show such a pricing scheme is, in a sense, the most likely to eliminate uplift payment among any pricing schemes. In Section 4.6, we apply the proposed β-fairness to a large scale problem based on ISO New England day-ahead market. Finally, we conclude with some discussion of future research.

4.2 The current day-ahead market auction and pricing scheme

A day-ahead electricity market is a forward market that determines the supply and consumption of electricity for the next operating day. In such a market, generating units submit offer bids to the ISO. An offer bid is usually composed of two parts, a variable cost part which specifies the functional relation between the supply quantity and the marginal cost of production, and a fixed cost part which comprises of start-up cost, shut-down cost, and no-load cost. Other information such as the ramp rate (i.e., how fast a generator can change its output level) is also requested by the ISO for scheduling purpose. The demand side is composed of fixed demand, which has a fixed consumption level regardless of the electricity price, as well as price-sensitive demand, whose consumption level is a function of electricity price as described in a demand bid.

Using the bidding information collected from the market participants, the ISO solves an auction and pricing problem that clears the market by determining the commitment status and production levels of generators, consumption levels of demand, and the electricity price. For this purpose, most ISOs in the United States adopt the social welfare maximization auction and marginal cost pricing (MaxSW-MCP) scheme. The objective is to maximize the total utility of the system, subject to energy balance constraint, transmission network constraints, and other physical and security constraints. Then, the electricity price is determined by the shadow price of the energy balance constraint. This price is also called the uniform electricity price, since it is constant at all different locations in the network. In reality, the shadow prices of the transmission constraints are also incorporated, so that the electricity prices are the sum of the uniform price and a component that is locational dependent. In the following, we present a simplified MaxSW-MCP model, where we only consider a single period auction with no transmission constraints. We also assume all the demand bids are price sensitive. We will discuss later the implication of relaxing these limitations.

Let $\mathcal{G}$ be the set of generators, $\mathcal{D}$ be the set of consumers in the day-ahead market. Let $\underline{p}_j$ and $\overline{p}_j$ be the minimum and maximum levels of production specified in generator j's offer bid. Similarly, $\underline{d}_i$ and $\overline{d}_i$ are the minimum and maximum levels of consumption specified in consumer i's demand bid. For each generator $j \in \mathcal{G}$, the binary variable x_j is the commitment status (i.e., $x_j = 1$ if the generator j is committed to production, $x_j = 0$ otherwise), and the continuous variable p_j is the production level in mega-watt (MW). The generation cost, $C_j(p_j, x_j) = f_j x_j + c_j(p_j)$, contains both fixed cost f_j and variable cost $c_j(p_j)$ as a function of p_j. For each consumer $i \in \mathcal{D}$, d_i is the consumption level, and $B_i(d_i)$ is the bid price function. In the current bidding convention, $C_j(p_j, x_j)$ is a piecewise nondecreasing linear function of the production level p_j. Each linear piece is described by a bidding block. The consumer bid $B_i(d_i)$ is usually a linear function of consumption level d_i. The MaxSW-MCP scheme is implemented by the following constrained mixed integer optimization model,

$$
\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{p},\boldsymbol{d}} \quad & \sum_{i\in\mathcal{D}} B_i(d_i) - \sum_{j\in\mathcal{G}} C_j(p_j, x_j) \\
\text{s.t.} \quad & \sum_{j\in\mathcal{G}} p_j - \sum_{i\in\mathcal{D}} d_i = 0, \quad (\lambda), \quad \text{(energy balance)} \\
& \underline{p}_i x_i \le p_i \le \bar{p}_i x_i, \quad \forall i \in \mathcal{G}, \quad \text{(production level bounds)} \\
& \underline{d}_i \le d_i \le \bar{d}_i, \quad \forall i \in \mathcal{D}, \quad \text{(consumption level bounds)} \\
& x_i \in \{0, 1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

The electricity price λ (in dollars per mega-watt hour) is the dual variable of the energy balance constraint in the corresponding linear optimization problem, where the binary variable $\boldsymbol{x}$ is fixed at its optimal value. Using the interpretation of shadow prices, we know that the electricity price λ is the marginal cost of supplying one more unit of electricity at the optimal level of production and consumption.

The above objective function can also be written as the sum of the utilities of all generators and consumers, thus illustrating the nature of the social welfare maximization. In particular, the utility u_j of a generator $j \in \mathcal{G}$ is defined as the difference between its revenue and bid costs, i.e., $u_j = \lambda p_j - C_j(p_j, x_j)$. The utility u_i of a consumer $i \in \mathcal{D}$ is defined as the customer surplus, i.e., $u_i = B_i(d_i) - \lambda d_i$. Then, applying energy balance constraint $\sum_{j \in \mathcal{G}} p_j = \sum_{i \in \mathcal{D}} d_i$, we have

$$\sum_{i \in \mathcal{D}} B_i(d_i) - \sum_{j \in \mathcal{G}} C_j(p_j, x_j) = \sum_{i \in \mathcal{D}} (B_i(d_i) - \lambda d_i) + \sum_{j \in \mathcal{G}} (\lambda p_j - C_j(p_j, x_j))$$
$$= \sum_{i \in \mathcal{D}} u_i + \sum_{j \in \mathcal{G}} u_j = \text{Social Welfare.}$$

Thus, the MaxSW solution maximizes the total wellness of the system, and, from a cost point of view, it is regarded as economically efficient. It is worth noticing that the social welfare is independent of the electricity price. In other words, the maximum economic efficiency can be obtained with different electricity prices. The auction determines the overall efficiency of the solution, while the pricing scheme affects the distribution of utility among participants. Other pricing schemes can be used together with the MaxSW auction to achieve certain welfare distribution purpose without sacrificing the system efficiency. We will propose one such pricing scheme later in the chapter.

According to current practice, if a committed generator j has a negative utility, that is, the revenue λp_j is strictly less than the as-bid generation cost $C_j(p_j, x_j)$, then the ISO compensates such a generator by paying a side payment that covers any loss, i.e., the side payment is equal to $\max\{C_j(p_j, x_j) - \lambda p_j, 0\}$. This payment is called the uplift payment. The following simple example of four generators and two consumers illustrates the MaxSW-MCP scheme and the uplift payment.

Consumer	Bid Price	$\underline{d}$	$\overline{d}$
1	30	0	120
2	30	0	120

Generator	Var Cost	Fixed Cost	$\underline{p}$	$\overline{p}$
1	15	150	50	100
2	16	140	0	100
3	19	120	50	100
4	22	100	0	100

Throughout this chapter, the price or cost is always in \$/MWh, and the production or consumption level is in MW. The utility is measured in dollars. Under the MaxSW-MCP scheme, both consumers are served at their maximum demand level. The following table shows the electricity price, the generation levels, and the utilities. The first four utilities are the utilities of the generators, and the last two are consumers' utilities. The last column is the social welfare. As expected from the MaxSW property, this is the highest level of social welfare achievable in this example.

Scheme	Price	Production	Utilities	SW
MaxSW-MCP	16	[100,90,50,0]	[-50, -140, -270, 0, 1680, 1680]	2900

The uplift payment for the generators can be easily calculated as \$50, \$140, \$270, and \$0. All the committed generators lose money without uplift payment, and the distribution of utility is very lopsided toward consumers, who have very high surpluses. This small example illustrates that the marginal cost pricing scheme does not take into account the fixed generation cost, and the social welfare maximization could potentially provide discriminating solutions that sacrifice utilities of certain participants in order to gain an overall high utility for the system. Both of these features are critical in causing the uplift problem.

The current make-whole payment approach can be seen as a way to redistribute utility among generators and consumers. Although it is easy to implement, the make-whole payment is an out-of-market action, and seems rather ad hoc. There are several recent proposals that aim to solve the problem by a more systematic approach. For instance, certain "commitment ticket prices" were introduced for binary commitment variables in [67]; "stay-online" was considered as a type of service in addition to electricity, and dual variables of some artificial "no-load constraints" were defined as "no-load clearing price" in [88]; dual variables of the mixed integer unit commitment problem were used as electricity price in [50, 44], where it was shown that the resulted electricity price minimizes the total uplift payments defined to cover both the production costs and the lost opportunity costs (it is worth noticing that this definition of uplift payment is different from the current ISO definition.) A detailed comparison of the above three proposals can be found in [59]. All these schemes keep the social welfare maximization as the auction objective and determine the electricity price using dual variables of either a linear economic dispatch problem with additional artificial constraints, or a mixed integer unit commitment problem to reflect integer decisions in the electricity price, which can be viewed as an extension of the marginal cost pricing idea.

However as we pointed out, both the social welfare maximization auction and the marginal cost pricing scheme contribute to the presence of the uplift payment. To resolve this issue, essentially a certain form of redistribution of utility must occur among the players. A new angle to look at the problem is to consider the fairness property of the auction and pricing outcomes. In the next section, we introduce a new auction and pricing formulation based on the notion of max-min fairness. Under this scheme, the uplift problem is naturally resolved. Then, we consider the important issue of tradeoff between economic efficiency and fairness, and propose the notion of β-fairness and a family of new auction and pricing schemes.

4.3 Max-Min Fairness Based Electricity Market Design

4.3.1 Fairness Concepts

Fairness issues in resource allocation problems have been extensively studied in many areas, notably in social sciences, welfare economics, and some engineering applications such as network communication[60, 65], air traffic flow management[19, 84], and some financial management problems [56]. Many principles of resource allocation have been proposed over the years, from the oldest theory of justice by Aristotle's equitable principle that allocates the resources in proportion to some pre-existing claims or rights that the participants have, to 18th century classical utilitarianism founded by Jeremy Bentham that dictates an allocation of resources that maximizes the total utilities, to Nash's proposal [63] , which allows a transfer of resources between two participants if the gainer's utility increases by a larger percentage than the loser's utility decrease, to John Rawls's theory of justice [72] and Kalai-Smorodinsky's solution concept [54] in bargaining problems, both of which give priority to the participants that are the least well off, so to ensure the highest minimum level of utility that every participant derives. For a comprehensive review and an investigation of the price of fairness in resource allocation problems see [17].

Aristotle's equitable principle has been widely used in situations where participants have specific pre-existing rights to the resources such as in the case of profit sharing between the shareholders. This is not the case for scheduling and trading electricity in a deregulated electricity market. Therefore, Aristotle's principle is not considered further. Bentham's utilitarianism is also extensively used in resource allocation problems especially when the sum of all the utilities serves as a measure of system efficiency. In fact, as mentioned in the first section, it has been adopted as the auction scheme in ISO's electricity markets. However, the utilitarianism approach has been criticised for its fairness properties, because some participants may suffer greatly reduced utilities in order to confer a benefit to the system, as we demonstrated in the simple example in Section 4.2. Nash's comparison principle and Kalai-Smorodinsky's solution concept are both derived from rigorous axiomatic foundations, see also [17].

We next introduce the concept of max-min fairness and an optimization formulation for finding such a solution. We refer to Appendix (Section 4.8) for a formal discussion of the basic properties of max-min fairness.

4.3.2 Max-Min Fairness

Let Q be the set of all achievable utilities in a resource allocation problem with n participants. Let $\boldsymbol{u}$ be a vector of utility allocation, $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in Q$. Then, $\boldsymbol{u}$ is *lexicographically* greater than or equal to another $\boldsymbol{v} \in Q$, denoted as $\boldsymbol{u} \geq_{\text{lex}} \boldsymbol{v}$, if $u_1 \geq v_1$, or if $u_1 = v_1$ and $u_2 \geq v_2$, or if $u_1 = v_1, u_2 = v_2$ and $u_3 \geq v_3$, etc. In other words, the first component of a vector is compared first, whose order determines the order of the two vectors. If the first components are equal, the second components are compared, and so on. We denote the problem of finding a lexicographically maximum vector as $\text{lexmax}_{\boldsymbol{u} \in Q}\,(u_1, u_2, \ldots, u_n)$.

If we sort $\boldsymbol{u}$ in a nondecreasing order, and denote the sorted version as

$$(u_{(1)}, u_{(2)}, \ldots, u_{(n)}),$$

the max-min fairness solution is defined as an optimal solution to the following lexmax problem,

$$\text{lexmax}_{\boldsymbol{u} \in Q}\ (u_{(1)}, u_{(2)}, \ldots, u_{(n)}). \tag{4.1}$$

Intuitively, a max-min fairness solution first maximizes the lowest utility level among all the participants, then maximizes the second lowest utility level while ensuring every participant derives at least the lowest level utility, and so on. Notice that (4.1) involves sort operation, which is highly nonlinear. The following equivalent formulation is more convenient to work with (see Lemma 15 in the Appendix).

$$\begin{aligned} \text{lexmax}\ & (r_1 - \boldsymbol{e}^T \boldsymbol{s}_1, 2r_2 - \boldsymbol{e}^T \boldsymbol{s}_2, \ldots, nr_n - \boldsymbol{e}^T \boldsymbol{s}_n) \\ \text{s.t.}\ \ & r_i \boldsymbol{e} - \boldsymbol{s}_i \leq \boldsymbol{u}, \quad \forall i = 1, \ldots, n, \\ & \boldsymbol{s}_i \geq \boldsymbol{0}, \quad \forall i = 1, \ldots, n, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

Let $(r_1^*, \boldsymbol{s}_2^*, \ldots, r_n^*, \boldsymbol{s}_n^*)$ and $\boldsymbol{u}^*$ be an optimal solution to the above problem. Then $\boldsymbol{u}^*$ is a max-min fairness solution, and $kr_k^* - \boldsymbol{e}^T \boldsymbol{s}_k^*$ is the sum of the first k smallest utilities of $\boldsymbol{u}^*$. Thus, the sorted utilities are given as $u_{(1)}^* = r_1^* - \boldsymbol{e}^T \boldsymbol{s}_1^*$ and $u_{(k)}^* = (kr_k^* - \boldsymbol{e}^T \boldsymbol{s}_k^*) - ((k-1)r_{k-1}^* - \boldsymbol{e}^T \boldsymbol{s}_{k-1}^*)$ for all $k = 2, \ldots, n$. The above formulation suggests a simple algorithm that iteratively solves for each level of the sorted utility. In fact, it is proven that the max-min fairness solution cannot be obtained by a single optimization problem.

A simple iterative algorithm for max-min fairness:

1. For $k = 1$, solve the following problem to obtain the smallest level of utility in a

max-min fairness solution,

$$\begin{aligned} t_1^* = \max_{r_1, \boldsymbol{s}_1, \boldsymbol{u}} \quad & r_1 - \boldsymbol{e}^T \boldsymbol{s}_1 \\ \text{s.t.} \quad & r_1 \boldsymbol{e} - \boldsymbol{s}_1 \leq \boldsymbol{u}, \\ & \boldsymbol{s}_1 \geq \boldsymbol{0}, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

2. For $k \geq 2$, keep the previous $k-1$ smallest utility levels and solve for the k-th level. The variables are $r_1, \ldots, r_k, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_k$, and $\boldsymbol{u}$,

$$\begin{aligned} t_k^* = \max_{\boldsymbol{r}, \boldsymbol{s}, \boldsymbol{u}} \quad & k r_k - \boldsymbol{e}^T \boldsymbol{s}_k \\ \text{s.t.} \quad & j r_j - \boldsymbol{e}^T \boldsymbol{s}_j \geq t_j^*, \quad \forall j = 1, \ldots, k-1, \\ & r_j \boldsymbol{e} - \boldsymbol{s}_j \leq \boldsymbol{u}, \quad \forall j = 1, \ldots, k, \\ & \boldsymbol{s}_j \geq \boldsymbol{0}, \quad \forall j = 1, \ldots, k, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

3. At $k = n$, the max-min fairness solution is obtained.

It is worth mentioning that $t_1^* \leq t_2^* \leq \cdots \leq t_n^*$, as t_k^* is the sum of the first k smallest utilities.

4.3.3 Max-min Fairness Auction and Pricing Scheme

Assume N generators and M consumers participate in the market. Denote $\mathcal{G} = \{1, \ldots, N\}$ as the index set of generators, and $\mathcal{D} = \{N+1, \ldots, N+M\}$ for consumers. As discussed in Section 4.2, the ISO has bidding information of each generator's cost function C_j, production range $\underline{p}_j, \overline{p}_j$, and each consumer's value function B_i, demand range $\underline{d}_i, \overline{d}_i$. We propose the following max-min fairness auction and pricing scheme,

$$
\begin{aligned}
(MMF) \quad & \operatorname{lexmax}\,(r_1 - \boldsymbol{e}^T \boldsymbol{s}_1, 2r_2 - \boldsymbol{e}^T \boldsymbol{s}_2, \ldots, (N+M) r_{N+M} - \boldsymbol{e}^T \boldsymbol{s}_{N+M}) && (4.2)\\
\text{s.t. } & r_k - s_{ki} \leq \lambda p_i - C_i(p_i, x_i), \forall i \in \mathcal{G}, k \in \mathcal{G} \cup \mathcal{D}, && (4.3)\\
& r_k - s_{ki} \leq B_i(d_i) - \lambda d_i, \forall i \in \mathcal{D}, k \in \mathcal{G} \cup \mathcal{D}, && (4.4)\\
& \boldsymbol{s}_k \geq \boldsymbol{0}, \quad \forall k \in \mathcal{G} \cup \mathcal{D}, && (4.5)\\
& \boldsymbol{e}_p^T \boldsymbol{p} - \boldsymbol{e}_d^T \boldsymbol{d} = 0, && (4.6)\\
& \underline{p}_i x_i \leq p_i \leq \bar{p}_i x_i, \quad \forall i \in \mathcal{G}, && (4.7)\\
& \underline{d}_i \leq d_i \leq \bar{d}_i, \quad \forall i \in \mathcal{D}, && (4.8)\\
& \lambda \geq 0, x_i \in \{0,1\}, \quad \forall i \in \mathcal{G}. && (4.9)
\end{aligned}
$$

Constraint (4.6) is the energy balance equation. (4.7) and (4.8) are production and consumption limits for generators and consumers. Note that the lower consumption level of a consumer is always zero, following the current bidding practice, i.e., $\underline{d}_i = 0$ for all i. Thus, no binary variables are required for loads.

The above max-min fairness problem can be solved by applying the simple iterative algorithm outlined in the previous subsection. In particular, we have the following algorithm.

1. $k = 1$, maximize the minimum utility level: t_1^* denotes the maximized minimum utility level.

$$
\begin{aligned}
(MMF_1) \quad t_1^* = & \max t_1\\
\text{s.t. } & t_1 \leq \lambda_1 p_i - C_i(p_i, x_i), \quad \forall i \in \mathcal{G},\\
& t_1 \leq B_i(d_i) - \lambda_1 d_i, \quad \forall i \in \mathcal{D},\\
& \boldsymbol{e}_p^T \boldsymbol{p} - \boldsymbol{e}_d^T \boldsymbol{d} = 0,\\
& \underline{p}_i x_i \leq p_i \leq \bar{p}_i x_i, \quad \forall i \in \mathcal{G},\\
& \underline{d}_i \leq d_i \leq \bar{d}_i, \quad \forall i \in \mathcal{D},\\
& \lambda_1 \geq 0, x_i \in \{0,1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

2. $k = 2, \ldots, N+M$, t_k^* is the maximized k-th smallest utility level, while ensuring the first $(k-1)$ smallest utility levels are at least $t_1^*, \ldots, t_{k-1}^*$,

$$
\begin{aligned}
(MMF_k) \quad t_k^* = \max\ & kr_k - \boldsymbol{e}^T \boldsymbol{s}_k \\
\text{s.t.}\ & r_j - s_{ji} \leq \lambda_k p_i - C_i(p_i, x_i), \quad \forall i \in \mathcal{G}, \forall j = 1, \ldots, k, \\
& r_j - s_{ji} \leq B_i(d_i) - \lambda_k d_i, \quad \forall i \in \mathcal{D}, \forall j = 1, \ldots, k, \\
& jr_j - \boldsymbol{e}^T \boldsymbol{s}_j \geq t_j^*, \quad \forall j = 1, \ldots, k-1, \\
& \boldsymbol{s}_j \geq \boldsymbol{0}, \quad \forall j = 1, \ldots, k, \\
& \boldsymbol{e}_p^T \boldsymbol{p} - \boldsymbol{e}_d^T \boldsymbol{d} = 0, \\
& \underline{p}_i x_i \leq p_i \leq \bar{p}_i x_i, \quad \forall i \in \mathcal{G}, \\
& \underline{d}_i \leq d_i \leq \bar{d}_i, \quad \forall i \in \mathcal{D}, \\
& \lambda_k \geq 0, x_i \in \{0, 1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

3. $k = N + M$, the optimal solution $(\boldsymbol{x}^*, \boldsymbol{p}^*, \boldsymbol{d}^*)$ of (MMF_{N+M}) is the final max-min fairness auction decision, and the final electricity price $\lambda^* = \lambda_{N+M}^*$.

Each (MMF_k) is a mixed-integer optimization problem with nonlinear constraints caused by the terms $\lambda_k p_i$ and $\lambda_k d_i$. To solve (MMF_k), we use a two-level strategy: first fix the price variable λ_k and solve the corresponding linear MIP by a commercial solver such as CPLEX, then search over λ_k which is a one-dimensional variable. Some special properties of (MMF_k) are shown in the next subsection that help make the computation more efficient.

4.3.4 Properties of the MMF Scheme

In this subsection, we explore various properties of the proposed MMF scheme from both economic and computational perspectives. In particular, we study the effect of the MMF scheme on utility distribution and uplift payment, and the behavior of the electricity price and its implication (Theorem 10, Lemma 7, and Corollary 1). Structural properties of the subproblem (MMF_k), such as the activity of market players, the quasi-convexity and continuity of the first k smallest utility as a function of λ_k, are discussed in Lemmas 8, 9, and 10. These properties are exploited in the solution algorithm to accelerate the computation.

Theorem 10. *$t_k^* \geq 0$ for all $k = 1, \ldots, N + M$. Thus, the uplift payment is always zero.*

Proof. Observe that the zero solution, $(\lambda, \boldsymbol{x}, \boldsymbol{p}, \boldsymbol{d}, \boldsymbol{r}, \boldsymbol{s}) = \boldsymbol{0}$, is feasible for (MMF) problem, where every participant's utility is zero. This zero solution must be dominated in lexicographic order by the utilities of the max-min fairness solution. Therefore, the minimum level utility must be nonnegative, i.e., $t_1^* \geq 0$, which implies every generator and consumer has a nonnegative utility in the max-min fairness solution. Thus, each generator's production cost including both variable and fixed costs is fully covered by its electricity revenue. The uplift payment is always zero in a max-min fairness solution. □

This is a simple but important result. Among various forms of fairness, the max-min fairness is regarded as the strongest. Our results show that improving the fairness property of the auction and pricing solution is effective in resolving the uplifty payment problem. Another interesting implication is that if some consumer has a fixed demand regardless of the electricity price, then the all zero solution may not be feasible for the (MMF) problem anymore, and the uplift payment can again be positive. This indicates that increasing demand responsiveness is helpful in reducing uplift payment in the max-min fairness scheme.

Next, we characterize the electricity price in the MMF scheme. The production cost function $C_j(p_j, x_j)$ of a generator includes the variable cost and the fixed cost such as startup and no-load costs. We assume a simple case where both generators and consumers have single-block bids. Then, the utility of a generator j is $u_j = (\lambda - c_j)p_j - f_j x_j$, where c_j is the variable cost and f_j is the fixed cost. The utility of a consumer i is $u_i = (b_i - \lambda)d_i$, where b_i is consumer i's valuation of electricity. We call a committed generator or a served load an *active* participant.

Lemma 7. *The final electricity price* λ^* *satisfies*

$$\max_{j \text{ active}} \left(c_j + \frac{f_j}{p_j} \right) < \lambda^* < \min_{i \text{ active}} b_i. \tag{4.10}$$

Proof. To avoid triviality, we assume there exists at least one nonzero feasible solution to the (MMF) problem. Then the active participants in the final auction decision must have positive utilities, i.e., $(\lambda - c_j)p_j - f_j > 0$ and $(b_i - \lambda)d_i > 0$, therefore, $c_j + \frac{f_j}{p_j} < \lambda < b_i$, for all active i, j's. □

The above lemma shows that the final electricity price is strictly higher than the average cost (thus marginal cost) of any *active* generator. This 'premium' over the variable production cost provides sufficient revenue for active generators to recover their fixed cost.

Furthermore, the price is strictly lower than the valuation of electricity of any active consumer. We assume a linear valuation function $B_i(d_i) = b_i d_i$ in the lemma. However, for a more general $B_i(d_i)$, a similar relationship holds as $\lambda < \partial B_i(d_i)/\partial d_i$ for any active demand i. This clearly shows that the price sensitivity of demand strongly affects the final electricity price. In the industry, there is a consensus that increasing demand-side responsiveness will help lower the electricity price. This lemma rigorously demonstrates this effect in the max-min fairness auction and pricing scheme.

From (4.10), it may seem that the generators can simply increase their bidding costs or consumers can decrease their offer price to manipulate the price. However, because only active participants affect the price and who will be active is not known a priori, naive manipulation of price may put them in disadvantage. An analogy can be found in the traditional MaxSW-MCP market: since generators do not know in advance who will be the marginal unit, simply increasing the bid cost might make the generators not committed. It is an interesting topic to systematically study the bidding behavior in a max-min fairness market.

The following corollary of Lemma 7 shows that bids of excessively high cost or offers of excessively low valuation will not be active.

Corollary 1. *In the max-min fairness scheme, a generator is not committed if its bid cost is higher than every consumer's offer price. Similarly, a consumer is not served if its offer price is lower than every generator's bid cost.*

Proof. Because the served consumers' offer prices and the committed generators' costs are strictly separated by the electricity price shown in Lemma 7. □

Since the electricity price λ is a decision variable in the (MMF) problem, each sub-problem (MMF_k) is a bilinear mixed integer optimization problem. Solve (MMF_k) is nontrivial. In the following, we show some properties of the (MMF_k) problem that are useful for simplifying computation.

A participant that is active at one level might not be active in the following levels. However, in the following case, all participants are active in the final max-min fairness solution, and the integrality constraints in the MMF problem can be dropped.

Lemma 8. *If $t_1^* > 0$, then all participants are active, and at subsequent levels, all participants remain active.*

Proof. A participant's utility is positive if only it is active. In subsequent levels, all participants must remain active to have nonzero utility, since $t_k^* \geq t_1^* > 0$. □

On the other hand, $t_1^* = 0$ implies that, for any electricity price λ_1, there is always a participant who would have a negative utility if it becomes active. When $t_1^* = 0$, the electricity price λ_1 is not uniquely defined. Therefore, the first level k that has $t_k^* > 0$ is important in determining the final electricity price. At this level, we can discard the first $(k-1)$ players that have zero production and consumption levels, and set the rest players to be active, thus reduce the situation to the case of $t_1^* > 0$. Without loss of generality, we can analyze the property of $t_1(\lambda)$ as a function of λ, which can be writen in a more compact form as follows,

$$t_1(\lambda) = \max_{(\boldsymbol{p},\boldsymbol{d})\in Y} \min\{(\lambda - c_j)p_j - f_j, \dots, (b_i - \lambda)d_i, \dots\}, \tag{4.11}$$

where Y is the polytope of the feasible $\boldsymbol{p}$ and $\boldsymbol{d}$, defined by constraints (4.6)-(4.8).

Definition (4.11) might seem to suggest a piecewise linear structure of $t_1(\lambda)$. Interestingly, this is not the case. Simple examples show that $t_1(\lambda)$ is piecewise differentiable, but some pieces can have nonzero curvature. Function $t_1(\lambda)$ can be also written as a standard linear optimization problem with constraint matrix linearly parametrized by a single variable λ. This type of problems with general matrices are discussed in [38], which shows that $t_1(\lambda)$ could be nonconcave and even discontinuous. Fortunately, the special structure of our problem makes $t_1(\lambda)$ a continous quasi-concave function.

Lemma 9. *The function $t_1(\lambda)$ is quasi-concave in $\lambda \in [\max_i c_i, \min_j b_j]$.*

Proof. Define a function $h(\gamma, \boldsymbol{p}, \boldsymbol{d}) = \min\{\gamma_i p_i - f_i, \dots, \gamma_j d_j, \dots\}$ with domain $\gamma \geq 0, \boldsymbol{p} \geq \boldsymbol{0}, \boldsymbol{d} \geq \boldsymbol{0}$, and f_i's are nonnegative constants. Then $h(\gamma, \boldsymbol{p}, \boldsymbol{d})$ is a quasi-concave function, because the upper level set $\{(\gamma, \boldsymbol{p}, \boldsymbol{d}) \mid h(\gamma, \boldsymbol{p}, \boldsymbol{d}) \geq r, (\gamma, \boldsymbol{p}, \boldsymbol{d}) \geq \boldsymbol{0}\}$ is a convex set.

Define an affine transformation as $\gamma_i \to \gamma_i - c_i$, $\gamma_j \to b_j - \gamma_j$, $p_i \to p_i$, $d_j \to d_j$ on the same domain, i.e., $\gamma_i - c_i \geq 0$ and $b_j - \gamma_j \geq 0$, $p_i \geq 0$, $d_j \geq 0$. The resulting function $\tilde{h}(\gamma, \boldsymbol{p}, \boldsymbol{d})$ is still quasi-concave. Then, restricting all $\gamma_i = \lambda$ where $\max_i c_i \leq \lambda \leq \min_j b_j$, $\tilde{h}(\lambda, \boldsymbol{p}, \boldsymbol{d})$ is quasi-concave in $(\lambda, \boldsymbol{p}, \boldsymbol{d})$. Then, $t_1(\lambda) = \max_{(\boldsymbol{p}, \boldsymbol{d}) \in Y} \tilde{h}(\lambda, \boldsymbol{p}, \boldsymbol{d})$ is quasi-concave. $\square$

For the continuity of $t_1(\lambda)$, we have the following more general result.

Lemma 10. *Assume function $f(\boldsymbol{x}, \boldsymbol{y})$ is jointly continuous in $(\boldsymbol{x}, \boldsymbol{y}) \in dom(f)$ and C is a compact set in $\mathbb{R}^n$, then $g(\boldsymbol{x}) := \max_{\boldsymbol{y} \in C} f(\boldsymbol{x}, \boldsymbol{y})$ is continuous in $dom(g)$.*

Proof. $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathrm{dom}(g)$, we have

$$
\begin{aligned}
|g(\boldsymbol{x}_1) - g(\boldsymbol{x}_2)| &= |\max_{\boldsymbol{y} \in C} f(\boldsymbol{x}_1, \boldsymbol{y}) - \max_{\boldsymbol{y} \in C} f(\boldsymbol{x}_2, \boldsymbol{y})| \\
&\leq \max_{\boldsymbol{y} \in C} |f(\boldsymbol{x}_1, \boldsymbol{y}) - f(\boldsymbol{x}_2, \boldsymbol{y})|.
\end{aligned}
$$

Since $f(\boldsymbol{x}, \boldsymbol{y})$ is jointly continuous in $(\boldsymbol{x}, \boldsymbol{y})$ and C is compact, $f(\boldsymbol{x}, \boldsymbol{y})$ is uniformly continuous on C. Hence, $\forall \epsilon > 0$, there exists $\delta > 0$, such that $|f(\boldsymbol{x}_1, \boldsymbol{y}) - f(\boldsymbol{x}_2, \boldsymbol{y})| \leq \epsilon$

for any $\|(\boldsymbol{x}_1, \boldsymbol{y}_1) - (\boldsymbol{x}_2, \boldsymbol{y}_2)\| \leq \delta$. Therefore, $|f(\boldsymbol{x}_1, \boldsymbol{y}) - f(\boldsymbol{x}_2, \boldsymbol{y})| \leq \epsilon$, for any $\|(\boldsymbol{x}_1, \boldsymbol{y}) - (\boldsymbol{x}_2, \boldsymbol{y})\| = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leq \delta$. □

Based on our computational experience, we conjecture that $t_1(\lambda)$ is actually concave in the domain where $t_1(\lambda)$ is strictly positive, and $t_1(\lambda)$ has a unique maximizer. The quasi-concavity and continuity of $t_1(\lambda)$ makes the one dimensional maximization problem over λ easier than the case of a general nonconcave function. Using the results from [38], we can do a binary search over λ to find the maximizer.

We can compute the MMF solution of the simple example of four generators and two consumers given in Section 4.2. Both MaxSW and MMF solutions are listed below. Two consumers are served at the maximum level $\overline{d}_i$, so they are omitted from the table.

Scheme	Price	Production	Utilities	SW
MaxSW	16	[100,90,50,0]	[-50, -140, -270, 0, 1680, 1680]	2900
MMF	27.0945	[50,44.043,57.895,88.062]	[454.72,348.63, ", ", ", "]	2197.89

The notation " means the last four utilities are all equal to 348.63. In the MaxSW solution, the electricity price is set by generator 2. The utility distribution among generators and loads is quite uneven. All three active generators have negative utilities $[-50, -140, -270]$, i.e. their production costs are not covered by electricity revenues. The MMF solution has a higher electricity price \$27.0945. All generators are committed, and are making positive profit, therefore no uplift payment is needed as anticipated by Theorem 10.

In the max-min fairness scheme, utilities are more equalized between generators and consumers. An inevitable consequence of the improved fairness is the relative reduction in system efficiency. In fact, as shown in the table, the social welfare of the max-min fairness solution is \$2197.94, about 75.8% of the MaxSW solution. In general, the social welfare of a fairness solution is lower than that of an MaxSW solution. While it is desirable to improve the fairness of an auction solution, it is also important to achieve a high economic efficiency. In the next section, we propose a notion of β-fairness and a family of auction and pricing schemes, which explicitly explores the tradeoff between fairness and efficiency. The max-min fairness scheme proposed in Subsection 4.3.3 is a special realization of this family.

4.4 β-Fairness

Since the max-min fairness solution may incur low social welfare, and the MaxSW solution suffers from fairness issues, we propose a notion of β-fairness that interpolates these two extremes. The following is the definition of β-fairness in an algorithmic framework.

1. Fix a $\beta \in [0, 1)$. For $k = 1$, solve the following problem,

$$
\begin{aligned}
(P_\beta^1) \quad & \max\ (1-\beta)\left(\sum_{i\in\mathcal{D}} B_i(d_i) - \sum_{i\in\mathcal{G}} C_i(p_i, x_i)\right) + \beta t_1 \\
\text{s.t.}\ & t_1 \le \lambda_1 p_i - C_i(p_i, x_i), \quad \forall i \in \mathcal{G}, \\
& t_1 \le B_i(d_i) - \lambda_1 d_i, \quad \forall i \in \mathcal{D}, \\
& \boldsymbol{e}_p^T \boldsymbol{p} - \boldsymbol{e}_d^T \boldsymbol{d} = 0, \\
& \underline{p}_i x_i \le p_i \le \bar{p}_i x_i, \quad \forall i \in \mathcal{G}, \\
& \underline{d}_i \le d_i \le \bar{d}_i, \quad \forall i \in \mathcal{D}, \\
& \lambda_1 \ge 0, x_i \in \{0, 1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

2. For $k = 2, \ldots, N+M$, solve the following problem,

$$
\begin{aligned}
(P_\beta^k) \quad & \max\ (1-\beta)\left(\sum_{i\in\mathcal{D}} B_i(d_i) - \sum_{i\in\mathcal{G}} C_i(p_i, x_i)\right) + \beta(k r_k - \boldsymbol{e}^T \boldsymbol{s}_k) \\
\text{s.t.}\ & r_j - s_{ji} \le \lambda_k p_i - C_i(p_i, x_i), \quad \forall i \in \mathcal{G}, \forall j = 1, \ldots, k, \\
& r_j - s_{ji} \le B_i(d_i) - \lambda_k d_i, \quad \forall i \in \mathcal{D}, \forall j = 1, \ldots, k, \\
& j r_j - \boldsymbol{e}^T \boldsymbol{s}_j \ge t_j^*, \quad \forall j = 1, \ldots, k-1, \\
& \boldsymbol{s}_j \ge \boldsymbol{0}, \quad \forall j = 1, \ldots, k,
\end{aligned}
$$

$$
\begin{aligned}
&\boldsymbol{e}_p^T\boldsymbol{p} - \boldsymbol{e}_d^T\boldsymbol{d} = 0,\\
&\underline{p}_i x_i \le p_i \le \bar{p}_i x_i, \quad \forall i \in \mathcal{G},\\
&\underline{d}_i \le d_i \le \bar{d}_i, \quad \forall i \in \mathcal{D},\\
&\lambda_k \ge 0, x_i \in \{0,1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

Let R_k be the set of optimal $(r_k^*, \boldsymbol{s}_k^*)$ of the above problem.

Let $t_k^* \leftarrow \max_{(r_k^*, \boldsymbol{s}_k^*)\in R_k}\{kr_k^* - \boldsymbol{e}^T\boldsymbol{s}_k^*\}$.

The final auction and pricing solutions are obtained at stage $N + M$, denoted as $(\boldsymbol{x}_\beta^*, \boldsymbol{p}_\beta^*, \boldsymbol{d}_\beta^*, \lambda_\beta^*)$.

In the following, we first show that the solution of the above β-fairness tradeoff scheme is Pareto optimal. To prove for a general setting, we use the notation introduced in Section 4.3.2: Let Q be the set of all achievable utilities in a resource allocation problem with n participants. Let $\boldsymbol{u}$ be a vector of utility allocation, $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in Q$, and $(u_{(1)}, u_{(2)}, \ldots, u_{(n)})$ is the utility sorted in a nondecreasing order. Then, the k-th level problem in the β-fairness scheme can be written as,

$$
\begin{aligned}
(P_\beta^k) \quad \max \quad & (1-\beta)\sum_{i=1}^{n} u_i + \beta \sum_{i=1}^{k} u_{(i)}\\
\text{s.t.} \quad & \sum_{i=1}^{j} u_{(i)} \ge t_j^*, \quad j = 1, \ldots, k-1,\\
& u \in Q.
\end{aligned}
$$

Theorem 11. *Let $\boldsymbol{u}^*$ be an optimal solution of the β-fairness tradeoff scheme. Then, $\boldsymbol{u}^*$ is Pareto optimal.*

Proof. Suppose $\boldsymbol{u}^*$ is not Pareto optimal, then there exists another utility vector $\boldsymbol{u} \in Q$ such that $u_i \ge u_i^*$ for all $i = 1, \ldots, n$ and for at least one i, $u_i > u_i^*$. Then, we have $\sum_{i=1}^n u_i > \sum_{i=1}^n u_i^*$.

Now we want to show that the sorted version of $\boldsymbol{u}$ also dominates the sorted version of $\boldsymbol{u}^*$, i.e., $u_{(i)} \ge u_{(i)}^*$ for all i. Without loss of generality, assume $\boldsymbol{u}^*$ is already sorted, i.e., $u_i^* = u_{(i)}^*$ for all i. For a specific $i \in \{1, \ldots, n\}$, let j be the original position of $u_{(i)}$, that is, $u_{(i)} = u_j$, and let k be the sorted position of u_i, i.e., $u_i = u_{(k)}$. If $j \geq i$, then $u_{(i)} = u_j \geq u_j^* \geq u_i^* = u_{(i)}^*$. If $j < i$ and $k \leq i$, then $u_{(i)} \geq u_{(k)} = u_i \geq u_i^* = u_{(i)}^*$. If $j < i$ and $k > i$, then among $u_{(1)}, \ldots, u_{(i-1)}$, at least one of them is from an element in $u_{i+1}, \ldots, u_n$, i.e., there exists $a \in \{1, \ldots, i-1\}$ and $b \in \{i+1, \ldots, n\}$ such that $u_{(a)} = u_b$. Then, $u_{(i)} \geq u_{(a)} = u_b \geq u_b^* \geq u_i^* = u_{(i)}^*$. Therefore, $u_{(i)} \geq u_{(i)}^*$ for all i.

Hence, for any $j = 1, \ldots, k-1$, it holds that $\sum_{i=1}^{j} u_{(i)} \geq \sum_{i=1}^{j} u_{(i)}^* \geq t_j^*$. Thus, $\boldsymbol{u}$ is a feasible solution of the tradeoff scheme. But $\sum_{i=1}^{n} u_i + \sum_{i=1}^{j} u_{(i)} > \sum_{i=1}^{n} u_i^* + \sum_{i=1}^{j} u_{(i)}^*$, which contradicts with the optimality of $\boldsymbol{u}^*$. As a special case, this also shows the well-known fact that the max-min fairness solution (for $\beta = 1$) is Pareto optimal. □

The parameter β controls the degree of tradeoff between the social welfare and fairness. The case $\beta = 1$ is the max-min fairness scheme proposed in Subsection 4.3.3. When β approaches zero, more weight is put on the social welfare part. In particular, we have the following property of t_k^* that is analogous to the fairness solution.

Lemma 11. *t_k^* maximizes the sum of the first k smallest utilities among all the optimal solutions of (P_β^k).*

Proof. For each optimal solution $(r_k^*, \boldsymbol{s}_k^*, \boldsymbol{u}^*)$ of the problem (P_β^k), by Lemma 14 (see Appendix), $(kr_k^* - \boldsymbol{e}^T \boldsymbol{s}_k^*)$ is equal to the sum of the first k smallest utilities. Since we choose t_k^* as $\max_{(r_k^*, \boldsymbol{s}_k^*) \in R_k} \{kr_k^* - \boldsymbol{e}^T \boldsymbol{s}_k^*\}$, t_k^* is the maximum over all the optimal solutions. □

Theorem 12. *Let $Z^k(\beta)$ be the optimal value of the tradeoff scheme (P_β^k), and Z^* be the maximum social welfare of the MaxSW scheme. Then, the following limit holds for all $k = 1, \ldots, N$,*

$$\lim_{\beta \to 0} Z^k(\beta) = Z^*.$$

Proof. It suffices to show that $Z^k(\beta)$ is continuous in β at $\beta = 0$. Notice that the parameter β is only in the objective function, and does not affect the feasible region. Using perturbation theory developed in [25, Proposition 4.4], a set of sufficient conditions for the continuity of $Z^k(\beta)$ are: (a) the objective function

$$f(\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{x}, \boldsymbol{r}, \boldsymbol{s}, \beta) := (1-\beta)\left(\sum_{i \in \mathcal{D}} B_i(d_i) - \sum_{i \in \mathcal{G}} C_i(p_i, x_i)\right) + \beta(kr_k - \boldsymbol{e}^T \boldsymbol{s}_k)$$

is continuous in $(\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{x}, \boldsymbol{r}, \boldsymbol{s}, \beta)$; (b) the feasible region of Problem (P^k_β) is closed and bounded.

Condition (a) is guaranteed by the continuity of $B_i(d_i)$ and $C_i(p_i, x_i)$, especially if we only consider linear functional form for $B_i(d_i)$ and $C_i(p_i, x_i)$, i.e., $B_i(d_i) = b_i d_i, C_i(p_i, x_i) = c_i p_i + f_i x_i$. For Condition (b), the original feasible region is closed, but unbounded, because $\boldsymbol{r}, \boldsymbol{s}, \lambda$ may take arbitrarily large values. However, we can restrict their values without affecting the optimal value. The price λ can be limited as $0 \leq \lambda \leq \max_{i \in \mathcal{D}} b_i$, since the optimal price must be below the highest offer price from consumers. From the production and consumption level constraints, we know the feasible region for $(\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{x})$ is closed and bounded, combining with the boundedness of λ, it implies that the set of possible utilities is closed and bounded, independent of β.

By Lemma 11, for any fixed β, the optimal $(jr^*_j - \boldsymbol{e}^T \boldsymbol{s}^*_j)$ is equal to the sum of the first j smallest optimal utilities. Then, we can apply Lemma 14 (proved in Appendix) that $u^*_{(j)} \leq r^*_j \leq u^*_{(j+1)}$, $0 \leq s_{ji} \leq u^*_{(j+1)} - u^*_i$ for $i \in \{(1), \ldots, (j)\}$ (i.e., for all i such that u^*_i is among the first j smallest utilities), and $s_{ji} = 0$ for all other i's. Since the utility set is bounded, r^*_j and $\boldsymbol{s}^*_j$ are bounded for all $j = 1, \ldots, k$. Therefore, the entire feasible region of (P^k_β) is closed and bounded, independent of β. $\square$

The above theorem confirms with the intuition that as $\beta \to 0$, the tradeoff scheme converges to the MaxSW scheme. However, notice that this convergence is not trivial in the sense that the continuity of the objective function alone is not sufficient to guarantee the continuity of the optimal value. The compactness of the feasible region is also crucial (see discussion in [25, Section 4.1]).

β	Price	Production	Utilities	SW
1	27.1	[50, 44.0, 57.9, 88.1]	[454.7,348.7, 348.7, 348.6, 348.7]	2197.9
0.9	27.1	[50,44.0,57.9, 88.1]	[454.7, 348.7, 348.6, 348.6, 348.7]	2197.9
0.8	27.1	[50, 44.0, 57.9, 88.1]	[454.7, 348.6, 348.6, 348.6, 348.6]	2197.9
0.7	26.1	[55.9, 60.5, 83.3, 40.3]	[469.9, 469.8, 469.8, 64.7, 469.9]	2414.1
0.6	25.4	[67.0, 73.0, 99.9, 0]	[548.8,548.7, 522.6, 0, 535.2]	2717.2
0.5	25.4	[67.2, 73.2, 99.6, 0]	[550.2, 550.0, 520.1,0, 548.9]	2718.3
0.4	25.4	[67.2, 73.2, 99.7, 0]	[550.3,549.6, 520.7,0, 548.7]	2718.1
0.3	24.9	[79.5, 87.3, 73.2, 0]	[627.4, 626.7,303.3, 0, 626.2]	2809.8
0.2	25.4	[83.0, 90.7, 66.3, 0]	[712.1,712.1,303.5,0, 553.2]	2813.1
0.1	25.9	[85.5, 93.1, 61.4, 0]	[781.9, 781.9, 303.5 ,0, 492.0]	2851.4
0.069	27.45	[100, 89.9, 50.1, 0]	[1095, 889.1, 303.5, 0, 306]	2899.6
0.068	27.47	[100, 90, 50, 0]	[1097, 892.3, 303.5, 0, 303.6]	2900
0.05	27.47	[100, 90, 50, 0]	[1097, 892.3, 303.5, 0, 303.6]	2900
MaxSW	16	[100, 90, 50, 0]	[-50, -140, -270, 0, 1680]	2900

Table 4.1: β-fairness tradeoff schemes for the simple example. For each β, the first four utilities are the utilities of generators; the last one is the consumer utility (since both consumers have the same utility, only one is shown).

We apply the tradeoff formulation to the simple example of four generators and two consumers. The solutions for different β's are shown in Table 4.1. The consumption levels are ignored, since the consumers are always served at $\overline{d}$.

As shown in the table, the social welfare of the tradeoff solutions increases when β decreases to 0. The utility vectors of larger β's tend to dominate the utility vectors of smaller β's in lexicographic order. These reflect the decrease in fairness as β decreases to 0. Also notice that none of the tradeoff solutions needs any side payments, while the MaxSW scheme requires considerable side payments of \$50, \$140, \$270 for the three committed generators.

As β approaches 0, social welfare of the tradeoff solution converges to the maximum level achieved by the MaxSW solution (see $\beta = 0.05$), as expected from Theorem 12. Note that the production levels also converge to the MaxSW solution, and the corresponding electricity price is strictly higher than the MaxSW price.

Also notice that there exists a $\beta_0 = 0.068$, such that for all β in the range $0 < \beta \leq \beta_0$, the tradeoff scheme achieves maximum social welfare, as well as the same production levels as in the MaxSW solutions. Furthermore, the fairness property of the tradeoff solution is strictly better than the MaxSW-MCP scheme reflected by the more even distribution of utilities, and the side payment is zero, which is strictly less than the current practice.

This property of the β-fairness tradeoff scheme has an important implication on another long noticed problem in the current practice, namely, how to achieve fairness and integrity in choosing from numerous optimal or near-optimal MaxSW solutions [53, 48]. Although all these solutions have the same system level social welfare, choosing one solution from others could have significant financial implications on the welfare of individual generators and consumers. The current practice does not impose any criteria in choosing one unit commitment solution from other solution candidates. In fact, it is nontrivial to even compute all the (near) optimal solutions. Our proposed tradeoff scheme provides a solution to this problem. When $\beta \leq \beta_0$, as shown in the above table, the tradeoff solution achieves the optimal social welfare and max-min fairness. In other words, in accordance with well formulated fairness principle, the β-

fairness tradeoff scheme automatically selects the most fair solution from potentially numerous maximum social welfare solution candidates. Furthermore, this is done without explicitly computing other social welfare solutions.

In the next section, we further explore this convergent property of the β-fairness scheme and propose a decoupled, sequential auction and pricing model that achieves maximum social welfare and strictly improved fairness with the advantage of much simplified computation.

4.5 Efficient Fairness Pricing Scheme

As shown in Theorem 12 and the computational experiment, when $\beta \to 0$, the β-fairness tradeoff solution achieves the same level of maximum social welfare as the MaxSW scheme, but strictly improves the fairness and lowers (or eliminates) the side payment. However, if β is exactly equal to zero, the objective function of Problem (P_β^k) only has social welfare and loses control over the fairness term, thus the electricity price is not properly constrained. This suggests an interesting 'limiting' auction and pricing scheme, where the auction follows the MaxSW practice, and the electricity price is determined by the max-min fairness principle. We call such a pricing scheme efficient fairness (EF). In particular, we propose the following auction and pricing scheme,

1. Auction: Maximize social welfare.

$$
\begin{aligned}
(A_{EF}) \quad \max \ & \sum_{i\in\mathcal{D}} B_i(d_i) - \sum_{i\in\mathcal{G}} C_i(p_i, x_i) \\
\text{s.t.} \ & \boldsymbol{e}_p^T \boldsymbol{p} - \boldsymbol{e}_d^T \boldsymbol{d} = 0, \\
& \boldsymbol{A}_p \boldsymbol{p} - \boldsymbol{A}_d \boldsymbol{d} \le \boldsymbol{f}, \\
& \underline{p}_i x_i \le p_i \le \bar{p}_i x_i, \quad \forall i \in \mathcal{G}, \\
& \underline{d}_i \le d_i \le \bar{d}_i, \quad \forall i \in \mathcal{D}, \\
& \lambda_1 \ge 0, x_i \in \{0, 1\}, \quad \forall i \in \mathcal{G}.
\end{aligned}
$$

Denote the optimal solution as $(\boldsymbol{x}_{EF}^*, \boldsymbol{p}_{EF}^*, \boldsymbol{d}_{EF}^*)$.

2. Pricing: Fix an auction solution and solve for electricity price. Denote the optimal electricity price as λ_{EF}^*.

$$
\begin{aligned}
(P_{EF}) \quad t^* := \max_{\lambda \geq 0} \quad & t \\
\text{s.t.} \quad & t \leq \lambda p_i^* - C_i(p_i^*, x_i^*), \quad \forall i \in \mathcal{G} \text{ and active}, \\
& t \leq B_i(d_i^*) - \lambda d_i^*, \quad \forall i \in \mathcal{D} \text{ and active}.
\end{aligned}
$$

The pricing step chooses an electricity price that maximizes the minimum utility level among the active players (i.e., committed generators and served consumers). In the following, we will show this achieves the max-min fairness solution, thanks to the uniqueness of the electricity price in this pricing model. The pricing scheme is a simple linear optimization problem.

Theorem 13. *The pricing problem (P_{EF}) has a unique solution λ^*, under which the obtained utility is a max-min fairness solution.*

Proof. The pricing problem (P_{EF}) can be written as

$$\max_{\lambda \geq 0} \min\{\lambda p_j^* - (c_j p_j^* + f_j), \cdots, b_i d_i^* - \lambda d_i^*, \cdots\},$$

for active generators and consumers i.e., $p_j^* > 0, d_i^* > 0$. The objective is a piece-wise linear concave function. The maximum is achieved at a breakpoint, where a linear piece of a generator utility intercepts with another linear piece of a consumer utility. Since the slope of each linear piece is nonzero, the interception is unique. Thus, the (P_{EF}) pricing model has a unique optimal electricity price λ^*. Any other price would lead to a strictly lower minimum level utility. Therefore, the utility determined by λ^* is a max-min fairness solution. □

For a fixed auction solution $(\boldsymbol{p}^*, \boldsymbol{d}^*, \boldsymbol{x}^*)$, different pricing models, such as the conventional marginal cost pricing, the convex-hull pricing of [50, 44], or the proposed

EF pricing, determine potentially different prices to dictate the distribution of utility. In fact, we can characterize the necessary and sufficient condition under which the EF pricing has zero uplif payment.

Theorem 14. *For an auction solution* $(\boldsymbol{p}^*, \boldsymbol{d}^*, \boldsymbol{x}^*)$, *the efficient pricing model* (P_{EF}) *has zero uplift if and only if*

$$\max_{j \text{ active}} \left(c_j + \frac{f_j}{p_j^*} \right) \leq \min_{i \text{ active}} b_i. \tag{4.12}$$

Proof. If the uplift payment is zero under EF price λ^*, then for any active generator j, $(\lambda^* - c_j)p_j^* - f_j \geq 0$, thus $\lambda^* \geq c_j + f_j/p_j^*$. Also, for any active consumer i, $(b_i - \lambda^*)d_i^* \geq 0$, hence $\lambda^* \leq b_i$. Therefore, we have (4.12).

Conversely, we already know from previous theorem that the EF price λ^* is the unique solution of the interception of a generator's utility and a consumer's utility. Denote these two players as generator k and consumer l, then $(\lambda^* - c_k)p_k^* - f_k = (b_l - \lambda^*)d_l^*$, so $\lambda^* = (c_k p_k^* + f_k + b_l d_l^*)/(p_k^* + d_l^*)$. If (4.12) holds, then $c_k + f_k/p_k^* \leq b_l$. Therefore, we have $\lambda^* \leq b_l$, which implies that the minimum level utility $(b_l - \lambda^*)d_l^* \geq 0$, thus the uplift is zero. □

In fact, the above proof shows that (4.12) is a necessary condition for any pricing scheme to have zero uplift payment. However, it may not be a sufficient condition for certain pricing schemes. For example, recall the simple example of four generators and two consumers. For the MaxSW auction solution $[100, 90, 50, 0, 120, 120]$, the average generation cost of active generators can be computed to be $[16.5, 17.5, 21.4]$ (the first three active generators), while the bid price of active consumers is 30. Therefore, (4.12) holds for this solution. However, there is significant uplift payment (all active generators need uplift) under marginal cost pricing, which means (4.12) is not a sufficient condition for the MCP to have zero uplift. It turns out that (4.12) is not a sufficient condition for the convex-hull pricing scheme proposed in [50, 44] either. Indeed, the convex-hull price can be computed to be 20.2 in this case, and the utilities of generators and consumers are $[370, 238, -60, 0, 1176, 1176]$, so the third generator needs uplift. Note that we always adhere to the definition of uplift used in the industry.

From the above discussion, the efficient fairness pricing scheme can be viewed as a pricing scheme that is the most likely to have zero uplift. If (4.10) does not hold, or equivalent if the efficient fairness pricing scheme has uplift, then there will be uplift payment under *any* pricing scheme. If this happens, the EF pricing scheme minimizes the largest uplift payment among all pricing schemes, because the minimum level utility is maximized.

As a summary, our proposed schemes represent a complete spectrum that spans from the maximum social welfare auction (MaxSW) and marginal pricing with no fairness consideration (the current practice), to the MMF scheme with full max-min fairness consideration. We proposed β-fairness scheme as a controlled tradeoff of fairness and efficiency. When β is close to zero, we empirically showed that there exists a $\beta_0 > 0$ such that the maximum social welfare is achieved, the fairness is strictly improved, and the uplift is eliminated or reduced. This also bears a significant implication on a long noticed problem of choosing from multiple optimal UC solutions to maintain certain fairness and integrity. In particular, the β-fairness tradeoff scheme with $\beta \leq \beta_0$ can be used to automatically select a UC solution of maximum social welfare and satisfies a well-defined fairness principle. We further proposed a Efficient Fairness pricing scheme that achieves the maximum social welfare while impoving the fairness property. We showed that among all possible pricing schemes, the EF pricing scheme is the most likely to eliminate uplift payment. When there has to be uplift under any pricing scheme, the EF scheme minimizes the largest uplift payment. In the following section, we apply the proposed schemes to a large scale problem, and introduce a frontier plot of efficiency versus fairness.

4.6 The Efficiency-Fairness Tradeoff Curve: An Example Using Real-World Data

In this section, we apply our proposed schemes to an auction and pricing problem of 143 supply bids and 109 demand bids in a single hour based on the bidding data of the day-ahead market operated by ISO New England. For simplicity, we only use the first bidding block of each bid and price-sensitive bids. Transmission network is assumed to be non-congested (which is the case as indicated by the published data), therefore network constraints are omitted. Recall that our proposed scheme is parametrized by a scalar quantity β ($0 \le \beta \le 1$). The case $\beta = 1$ corresponds to the max-min fairness scheme; $\beta = 0$ corresponds to the current ISO MaxSW-MCP scheme; the intermediate values of β correspond to the tradeoff with an objective of $(1-\beta) \cdot$ EFFICIENCY $+ \beta \cdot$ FAIRNESS.

Scheme	Social Welfare	Scheme	Social Welfare
MMF	64499.70	$\beta = 0.3$	80126.85
$\beta = 0.9$	69593.76	EF	80267.25
$\beta = 0.7$	74677.97	MaxSW	80267.25
$\beta = 0.5$	75894.90		

The above table shows the efficiency of the max-min fairness scheme (MMF), the tradeoff scheme with $\beta = 0.3, 0.5, 0.7, 0.9$, the EF scheme, and the current ISO's MaxSW scheme. The social welfare, thus the efficiency, increases from MMF scheme to MaxSW scheme. The Efficient Fairness (EF) scheme achieves the same social welfare as the MaxSW scheme and strictly improves fairness properties.

Figure 4.6 shows four efficiency-fairness curves constructed from solutions of the above seven schemes. In each curve, the efficiency of a solution is measured by its total utility, and the fairness is quantified by one of the four well-known inequity measures (See Figure 1 for definitions) [61]. Notice that higher value in the 'Fairness' axis means more unfair.

We can see from Figure 1 that our proposal establishes a spectrum of tradeoff between efficiency and fairness. At one end, the current practice MaxSW-MCP has the highest efficiency, but suffers from the worst fairness property indicated by all four inequity measures. At the other end, the max-min fairness scheme ($\beta = 1$) achieves the best fairness property with a 20% reduction in efficiency. In between, by varying β, our proposed scheme significantly increases fairness while controlling the decrease

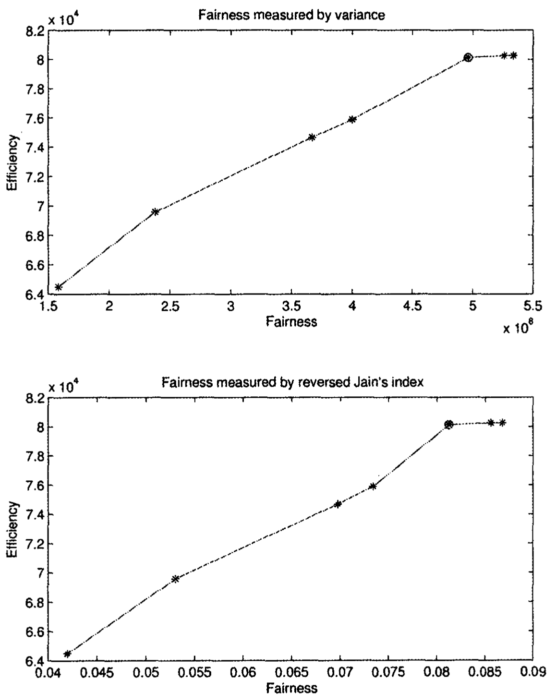

Figure 4-1: Efficiency-fairness tradeoff curve. Let u be the utility vector of a solution and $\bar{u}$ be its mean. Four inequity measures are defined as variance: $\sum_i (u_i - \bar{u})^2$; sum of absolute deviations: $\sum_i \sum_j |u_i - u_j|$; reversed Jain's index: $\frac{\sum_i u_i^2}{n^2 \bar{u}^2}$; Gini index: $\frac{\sum_i \sum_j |u_i - u_j|}{2n^2 \bar{u}}$. In each plot, from the highest efficiency are the solutions of the MaxSW-MP scheme, the limiting pricing scheme, and the tradeoff scheme with $\beta = 0.3, 0.5, 0.7, 0.9$.

in social welfare. Moreover, the EF pricing scheme achieves the same efficiency as the MaxSW-MP scheme, while improving the fairness. This suggests that the current electricity market is *not* operating on the pareto frontier of efficiency versus fairness.

The efficiency-fairness curve generated by our schemes provides a valuable decision tool for the ISO to identify a proper operating point of its market. For example, as shown in Figure 1, the tradeoff scheme at $\beta = 0.3$ considerably improves fairness with little decrease in efficiency. Thus, such a scheme could be a desirable candidate for the ISO electricity market's operating policy.

4.7 Conclusion

Motivated by the evidence of unfairness revealed by the uplift problem, we introduced the well-defined max-min fairness concept into the electricity market design problem, and proposed the max-min fairness (MMF) auction and pricing scheme. Various computational and economic properties of this scheme are investigated. An important outcome is that under the MMF scheme, the uplift payment is always zero.

Then, we proposed a new notion of β-fairness auction and pricing scheme, which explicitly trades off fairness and efficiency. We investigated the Pareto optimality property of the solution, and the convergent behavior as $\beta \to 0$. We empirically showed that there exists a range of $\beta \in (0, \beta_0]$, where the tradeoff scheme achieves the same level of social welfare as the MaxSW scheme, strictly improves the fairness, and eliminates or reduces the uplift payment. An important implication is that such a tradeoff scheme provides a solution to a long standing problem of the current practice in choosing a proper solution from multiple optimal or near-optimal social welfare solutions (as first raised in [53, 48]).

We further introduced efficient fairness pricing scheme as a limiting case of $\beta \to 0$, which decouples the auction and pricing, and simultaneously achieves the maximum efficiency and the max-min fairness. We showed that the necessary condition for any pricing scheme to have zero uplift payment is also sufficient for the EF scheme. In this sense, the EF scheme is the most likely pricing scheme to eliminate uplift payment.

In the situation where the uplift is unavoidable by any pricing model, the EF scheme minimizes the largest uplift payment among all possible pricing schemes.

We test our models on a large scale problem using data obtained from ISO-NE's day-ahead electricity market. Using four well-known equity measures, we construct efficiency-fairness curves, all of which show a consistent picture of tradeoff between fairness and efficiency achieved by our models. A particularly revealing observation shown by the efficiency-fairness curves is that the current ISO practice is *not* operating on the Pareto frontier of efficiency versus fairness. Proposed β-fairness tradeoff scheme and EF pricing scheme strictly improves fairness while obtaining the same level of efficiency as the MaxSW scheme. These frontier curves can be a valuable tool to facilitate the ISO decision in choosing a proper operating policy of its market.

4.8 Appendix: Max-Min Fairness

In the following, we introduce the definitions of lexicographical order, lexmax problem, and the concept of max-min fairness. Then we discuss the formulation of max-min fairness as a lexmax problem and the solution algorithm. Most of the properties and formulations presented here have appeared in previous papers on fairness, including [4, 26, 57] and a series of papers by Ogryczak [64, 65, 66].

Let Q be the set of all achievable utilities in a resource allocation problem. Let $\boldsymbol{u}$ be a vector of utilities of all the participants of a resource allocation solution.

Definition 1. *A vector $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ is lexicographically greater than or equal to another $\boldsymbol{v} \in \mathbb{R}^n$, denoted as $\boldsymbol{u} \geq_{lex} \boldsymbol{v}$, if and only if $u_1 \geq v_1$, or $u_1 = v_1$ and $u_2 \geq v_2$, or $u_1 = v_1, u_2 = v_2$ and $u_3 \geq v_3$, etc.*

In other words, the first component of a vector carries the most weight and is compared first. The order of the first components determines the order of the two vectors. If the first components are equal, the second components are compared, and so on.

Definition 2. *Problem $lexmax_{\boldsymbol{u} \in Q}$ $(u_1, u_2, \ldots, u_n)$ is to find $\boldsymbol{u} \in Q$ such that $\boldsymbol{u} \geq_{lex}$*

$\boldsymbol{v}$ *for all* $\boldsymbol{v} \in Q$.

That is, the lexmax problem is an optimization problem under the lexicographical order. We also use the notation $(u_{(1)}, u_{(2)}, \ldots, u_{(n)})$ to denote a nondecreasing order of the components of $\boldsymbol{u}$, i.e., $u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(n)}$.

Definition 3. *Given a utility set* Q, *an optimal solution to the lexmax problem,*

$$lexmax_{\boldsymbol{u}\in Q}\ (u_{(1)}, u_{(2)}, \ldots, u_{(n)}) \tag{4.13}$$

is called a max-min fairness solution.

Intuitively, a max-min fairness solution first maximizes the lowest utility level among all the participants, then maximizes the second lowest utility level while ensuring every participant derives at least the lowest level utility, and so on. The existence of the max-min fairness solution is guaranteed by the following lemma.

Lemma 12. *Assume the utility set* Q *is compact. Then, the max-min fairness solution* $\boldsymbol{u}^*$ *always exists and the ordered vector* $(u^*_{(1)}, u^*_{(2)}, \ldots, u^*_{(n)})$ *is unique.*

Proof. Follow from the definition of lexmax and the fact that lexicographical order is a complete order, i.e. any two vectors can be ordered by it. However, the max-min fairness solution may not be unique. Consider a case where the utility set $Q = \{(0,1), (1,0)\}$. Then both points are max-min fair. $\square$

Notice that problem (4.13) is not a standard lexmax problem, because the objective involves sorting a utility vector, where the sorting itself is a nonlinear operator. The following lemma gives an equivalent formulation that leads to a standard lexmax problem shown in Lemma 15.

Lemma 13. *An optimal solution of*

$$(P) \quad lexmax_{\boldsymbol{u}\in Q}\ (u_{(1)}, u_{(1)} + u_{(2)}, \ldots, \sum_{i=1}^{n} u_{(i)})$$

is a max-min fairness solution and vice versa.

Proof. If $\boldsymbol{u}^*$ is a max-min fairness solution, and $\boldsymbol{u}$ is any given point in Q, then $u^*_{(1)} \geq u_{(1)}$; if $u^*_{(1)} = u_{(1)}$, then $u^*_{(2)} \geq u_{(2)}$, thus $u^*_{(1)} + u^*_{(2)} \geq u_{(1)} + u_{(2)}$, etc. So $\boldsymbol{u}^*$ is also an optimal solution to problem (P).

If $\boldsymbol{u}^*$ is an optimal solution to problem (P), and $\boldsymbol{u}$ is any given point in Q, then $u^*_{(1)} \geq u_{(1)}$; if $u^*_{(1)} = u_{(1)}$, since $u^*_{(1)} + u^*_{(2)} \geq u_{(1)} + u_{(2)}$, then $u^*_{(2)} \geq u_{(2)}$, and so on. Therefore, (P) gives max-min fairness solutions. □

Lemma 14. *Given $\boldsymbol{u} \in \mathbb{R}^n$ and $1 \leq k \leq n$, $\sum_{i=1}^k u_{(i)}$ can be found by the following linear program and its dual,*

$$\begin{aligned}\sum_{i=1}^{k} u_{(i)} &= \min_{\boldsymbol{y}} \quad \{\boldsymbol{u}^T \boldsymbol{y} \mid \boldsymbol{e}^T \boldsymbol{y} = k, \boldsymbol{0} \leq \boldsymbol{y} \leq \boldsymbol{e}\}, \\ &= \max_{r,\boldsymbol{s}} \quad \{kr - \boldsymbol{e}^T \boldsymbol{s} \mid r\boldsymbol{e} - \boldsymbol{s} \leq \boldsymbol{u}, \boldsymbol{s} \geq \boldsymbol{0}\},\end{aligned}$$

where $r \in \mathbb{R}$ and $\boldsymbol{s} \in \mathbb{R}^n$. Further, the following relation holds,

$$\begin{aligned}&u_{(k)} \leq r^* \leq u_{(k+1)} \\ &0 \leq s_i \leq u_{(k+1)} - u_i, \forall i \in \{(1), \ldots, (k)\}, \\ &s_i = 0, \forall i \notin \{(1), \ldots, (k)\}.\end{aligned}$$

Proof. The intuition for the primal LP is clear. For the dual LP, $r - s_i \leq u_i$ for all i, thus summing over any subset I_k of k indices, we have $kr - \sum_{i \in I_k} s_i \leq \sum_{i \in I_k} u_i$. Since $\boldsymbol{s} \geq \boldsymbol{0}$, $kr - \sum_{i=1}^n s_i \leq kr - \sum_{i \in I_k} s_i \leq \sum_{i \in I_k} u_i$ for all I_k, therefore $kr - \sum_{i=1}^n s_i \leq \sum_{i=1}^k u_{(i)}$. By the complementary slackness, $y_i(r - s_i - u_i) = 0$ and $(y_i - 1)s_i = 0$ for all i. Hence, if $y_i = 1$ (i.e. u_i is among the first k smallest components of $\boldsymbol{u}$, denoted as $i \in \{(1), \ldots, (k)\}$), then $r - s_i = u_i$, $r \geq u_i$; if $y_i = 0$, then $s_i = 0$, which implies $r \leq u_i$. Therefore $u_{(k)} \leq r \leq u_{(k+1)}$. For $i \in \{(1), \ldots, (k)\}$, $0 \leq s_i = r - u_i \leq u_{(k+1)} - u_i$. For other i, $s_i = 0$. □

Use Lemma 13 and Lemma 14, we have

Lemma 15. *Let $(r_1^*, \boldsymbol{s}_2^*, \ldots, r_n^*, \boldsymbol{s}_n^*)$ and $\boldsymbol{u}^*$ be an optimal solution to the following* *problem,*

$$\begin{aligned} lexmax\ & (r_1 - \boldsymbol{e}^T\boldsymbol{s}_1, 2r_2 - \boldsymbol{e}^T\boldsymbol{s}_2, \ldots, nr_n - \boldsymbol{e}^T\boldsymbol{s}_n) \\ s.t.\ & r_i\boldsymbol{e} - \boldsymbol{s}_i \leq \boldsymbol{u}, \quad \forall i = 1, \ldots, n, \\ & \boldsymbol{s}_i \geq \boldsymbol{0}, \quad \forall i = 1, \ldots, n, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

Then $\boldsymbol{u}^*$ *is a max-min fairness solution, and the ordered utilities are* $u^*_{(1)} = r^*_1 - \boldsymbol{e}^T\boldsymbol{s}^*_1$ *and* $u^*_{(k)} = (kr^*_k - \boldsymbol{e}^T\boldsymbol{s}^*_k) - ((k-1)r^*_{k-1} - \boldsymbol{e}^T\boldsymbol{s}^*_{k-1})$ *for all* $k = 2, \ldots, n$.

Several algorithms are proposed in [65] to solve the above lexmax problem. When the utility set Q is a convex set, an efficient algorithm that uses the dual solution is available. However, the electricity market auction and pricing problem is intrinsically discrete and nonconvex. The following simple sequential algorithm is needed in this case.

1. $k = 1$, solve

$$\begin{aligned} t^*_1 = \max_{r_1, \boldsymbol{s}_1, \boldsymbol{u}} \quad & r_1 - \boldsymbol{e}^T\boldsymbol{s}_1 \\ \text{s.t.} \quad & r_1\boldsymbol{e} - \boldsymbol{s}_1 \leq \boldsymbol{u}, \\ & \boldsymbol{s}_1 \geq \boldsymbol{0}, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

2. For $k \geq 2$, maximize $kr_k - \boldsymbol{e}^T\boldsymbol{s}_k$ with constraints that keep the previous $k-1$ objective functions at its respective optimal values. The variables are

$r_1, \ldots, r_k, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_k$, and $\boldsymbol{u}$,

$$\begin{aligned} t_k^* = \max_{\boldsymbol{r},\boldsymbol{s},\boldsymbol{u}} \quad & k r_k - \boldsymbol{e}^T \boldsymbol{s}_k \\ \text{s.t.} \quad & j r_j - \boldsymbol{e}^T \boldsymbol{s}_j \geq t_j^*, \quad \forall j = 1, \ldots, k-1, \\ & r_j \boldsymbol{e} - \boldsymbol{s}_j \leq \boldsymbol{u}, \quad \forall j = 1, \ldots, k, \\ & \boldsymbol{s}_j \geq \boldsymbol{0}, \quad \forall j = 1, \ldots, k, \\ & \boldsymbol{u} \in Q. \end{aligned}$$

3. $k = n$ gives a max-min fairness solution.

Chapter 5

Conclusion and Future Research

In the first part of this thesis, we focus our attention on the security constrained unit commitment (SCUC) problem, which is a key daily operation of any large scale electric power system in both vertically integrated industry and deregulated electricity markets. We proposed an adaptive robust model for the SCUC under nodal net load uncertainty. The effectiveness of the robust approach is demonstrated by an extensive computational study on the ISO New England's power system. Comparing to the current reserve adjustment approach, the robust UC model has significant improvement in economic efficiency, real-time operational reliability, and robustness to uncertain demand distribution.

Many interesting research questions are open. It would be interesting to study re-commitment that is adaptive to load forecast. We can easily adjust the parameters such as $\hat{d}_j^t$ in the uncertainty set and re-run our model for future re-commitment when a better estimation of uncertainty is available. This could be very useful when the system has high percentage of price responsive demand and variable supply. We already show that the robust solution significantly reduces the volatility of the dispatch cost. It would be interesting to study the extent that the volatility in the energy price is reduced. It would also be very interesting to model explicit correlation between different uncertain loads, such as spatial correlation between adjacent wind farms, or temporal correlation between consumption levels. Such correlation can be captured by modeling covariance matrices in the uncertainty set.

In the second part of the thesis, we studied the more general concept of finite adaptability, and provided a geometric characterization of the power of robust and finitely adaptable solutions for a rather general class of multistage stochastic and adaptive optimization models. Some further directions can be pursued. For example, it would be interesting to study the performance of affine decision rules in a general multistage setting. What would be the key geometric quantity, if there is any, in determining the bound?

In the third part of the thesis, we proposed and investigated the notion of β-fairness that addresses the tradeoff between social welfare and fairness for the day-ahead electricity market auction and pricing. The case $\beta = 0$ corresponds to current practice, whereas $\beta = 1$ corresponds to a solution that maximizes the minimum utility among market participants, the so-called max-min fairness. Such a max-min fair solution eliminates side payments, thus resolves the uplift payment problem. We investigated the tradeoff curve, and showed that the current operational practice ($\beta = 0$) is not Pareto efficient, while there is a $\beta_0 < 1$, at which the social welfare is the same as current practice, while the max-min fairness is strictly better, and the side payments are strictly lower than current practice. This gives a solution to another long standing problem of achieving fairness and integrity of the auction in choosing from multiple (near) optimal solutions. We investigated the properties of such β-fair solutions both theoretically and empirically. Several interesting directions are open for future research. It is interesting to explore the transmission network effects and locational pricing concept under the fairness framework. It is also interesting to study the incentive property of the fairness scheme.

Bibliography

[1] A. Nemirovski A. Ben-tal and C. Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal of Optimization*, 13(2):535–560, 2002.

[2] L. T. Anstine, R.E. Burke, J. E. Casey, R. Holgate, R.S. John, and H. G. Stewart. Application of probability methods to the determination of spinning reserve requirements for pennsylvania-new jersey-maryland. *IEEE Transactions on Power Apparatus and Systems*, PAS-82(68):726735, 1963.

[3] E.M.L. Beale. On Minizing A Convex Function Subject to Linear Inequalities. *Journal of the Royal Statistical Society. Series B (Methodological)*, 17(2):173–184, 1955.

[4] F. A. Behringer. A simplex based algorithm for the lexicographically extended linear maxmin problem. *European Journal of Operational Research*, 7:274–283, 1981.

[5] A. Belloni and R. M. Freund. On the symmetry function of a convex set. *Mathmematical Programming*, 111(1):57–93, 2007.

[6] A. Bemporad, F. Borrelli, and M. Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003.

[7] A. Ben-Tal, S. Boyd, and A. Nemirovski. Control of uncertainty-affected discrete time linear systems via convex programming. *Submitted to SIAM Journal on Control and Optimization*, 2005.

[8] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming*, 107(1):63–89, 2006.

[9] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

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

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

[12] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letter*, 25(1):1–13, 1999.

[13] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88:411–421, 2000.

[14] A. Ben-Tal and A. Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

[15] D. Bertsimas, D.B. Brown, and C. Caramanis. Theory and applications of Robust Optimization. *SIAM Review (To Appear)*, 2010.

[16] D. Bertsimas and C. Caramanis. Finite adaptability in linear optimization. *IEEE Automatic Control*, (12):2751–2766, 2010.

[17] D. Bertsimas, V. F. Farias, and N. Trichakis. The price of fairness. *Operations Research*, 59:17–31, 2011.

[18] D. Bertsimas and V. Goyal. On the power of robust solutions in two-stage stochastic and adaptive optimization problems. *Mathematics of Operations Research*, 35(2):284–305, 2010.

[19] D. Bertsimas and S. Gupta. Fairness in the air traffic flow management problem. *Transportation Science*, 2009.

[20] D. Bertsimas, D.A. Iancu, and P.A. Parrilo. Optimality of Affine Policies in Multistage Robust Optimization. *Mathematics of Operations Research*, 35(2):363–394, 2010.

[21] D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98:49–71, 2003.

[22] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52:35–53, 2004.

[23] R. Billinton and M. Fotuhi-Firuzabad. A reliability framework for generating unit commitment. *Electric Power System Research*, 56(1):81–88, 2000.

[24] R. Billinton and R. Karki. Capacity reserve assessment using system well-being analysis. *IEEE Transactions on Power Systems*, 14, 1999.

[25] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems.* Springer, 2000.

[26] R. E. Burkard and F. Rendl. Lexicographic bottleneck problems. *Operations Research Letters*, 10:303–308, 1991.

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

[28] M. Carrion and J. M. Arroyo. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem, 2006.

[29] H. Chao and H.G. Huntington, editors. *Designing competitive electricity markets.* Kluwer Academic Publishers, 1998.

[30] D. Pachamanova D. Bertsimas and M. Sim. Robust linear optimization under general norm. *Operations Research Letter*, 32(6):510–516, 2004.

[31] G.B. Dantzig. Linear programming under uncertainty. *Management Sci.*, 1:197–206, 1955.

[32] M. A. Duran and I. E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming*, 36(3):306–339, 1986.

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

[34] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.

[35] E.M.Constantinescu, V. M. Zavala, M. Rocklin, S. Lee, and M. Anitescu. A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Transactions on Power Systems*, 26(1):431–441, 2010.

[36] J. E. Falk. Exact solutions of inexact linear programs. *Operations Research*, 24(4):783–787, 1976.

[37] R. Fletcher and S. Leyffer. Solving mixed integer programs by outer approximation. *Mathematical Programming*, 66(3):327–349, 1994.

[38] R. M. Freund. Postoptimal analysis of a linear program under simultaneous changes in matrix coefficients. *Mathematical Programming Study*, 24:1–13, 1985.

[39] S.J. Garstka and R.J.B. Wets. On decision rules in stochastic programming. *Mathematical Programming*, 7(1):117–143, 1974.

[40] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Analysis and Applications*, 18, 1997.

[41] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidenite programs. *SIAM J. Optimization*, 9(1):33–52, 1998.

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

[43] H.B. Gooi, D.P. Mendes, K.R.W. Bell, and D.S. Kirschen. Optimal scheduling of spinning reserve. *IEEE Transactions on Power Systems*, 14, 1999.

[44] P. R. Gribik, W. W. Hogan, and S. L. Pope. Market-clearing electricity prices and energy uplift. Technical report, Harvard University, 2007.

[45] X. Guan, P.B. Luh, H. Yan, and J.A. Amalfi. An optimization based method for unit commitment. *International Journal of Electrical Power & Energy System*, 14(6):9–17, 1992.

[46] A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 417–426, 2004.

[47] A. Gupta, M. Pal, R. Ravi, and A. Sinha. What about Wednesday? Approximation algorithms for multistage stochastic optimization. *Lecture notes in computer science*, pages 86–98, 2005.

[48] B. F. Hobbs, M. H. Rothkopf, R. P. O'Neill, and H. Chao, editors. *The Next Generation of Electric Power Unit Commitment Models.* Kluwer's International Series. Kluwer Academic Publishers, 2001.

[49] W. Hogan. Competitive electricity market design: A wholesale primer. 1998.

[50] W. W. Hogan and B. J. Ring. On minimum-uplift pricing for electricity markets. Technical report, Harvard University, 2003.

[51] G. Infanger. *Planning under uncertainty: solving large-scale stochastic linear programs.* Boyd & Fraser Pub Co, 1994.

[52] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays, presented to R. Courant on his 60th Birthday, January 8, 1948*, 1948.

[53] R. B. Johnson, S. S. Oren, and A. J. Svoboda. Equity and efficiency of unit commitment in competitive electricity markets. *Utility Policy*, 6(1):9–19, March 1997.

[54] E. Kalai and M. Smorodinsky. Other solutions to Nash's bargaining problem. *Econometrica*, 43:510–518, 1975.

[55] P. Kall and S.W. Wallace. *Stochastic programming.* Wiley New York, 1994.

[56] A. Khodadadi, R. Tutuncu, and P. Zangari. Optimization and quantitative investment management. *Journal of Asset Management*, 7:83–92, 2006.

[57] R. S. Klein, H. Luss, and D. R. Smith. A lexicographic minimax algorithm for multiperiod resource allocation. *Mathematical Programming*, 55:213–234, 1992.

[58] J. Lee, J Leung, and F Margot. Min-up/min-down polytopes. *Discrete Optimization*, 1(1):77–85, June 2004.

[59] E. Litvinov, F. Zhao, and T. Zheng. Alternative auction objectives and pricing schemes in short-term electricity markets. 2008.

[60] H. Luss. On equitable resource allocation problems: a lexicographic minimax approach. *Operations Research*, 47(3):361–378, 1999.

[61] M. T. Marsh and D. A. Schilling. Equity measurement in facility location analysis: a review and framework. *European Journal of Operational Research*, 74:1–17, 1994.

[62] H. Minkowski. Allegemeine lehzätze über konvexe polyeder. *Ges. Ahb.*, 2:103–121, 1911.

[63] J. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

[64] W. Ogryczak. On the lexicographic minimax approach to location problems. *European Journal of Operational Research*, 100:566–585, 1997.

[65] W. Ogryczak, M. Pioro, and A. Tomaszewski. Telecommunications network design and max-min optimization problem. *Journal of Telecommunications and Information Technology*, (3):43–56, 2005.

[66] W. Ogryczak and T. Sliwinski. On equitable approaches to resource allocation problems: the conditional minimax solution. *Journal of Telecommunications and Information Technology*, (3):40–48, 2002.

[67] R. P. O'Neill, P. M. Sotkiewicz, B. F. Hobbs, M. H. Rothkopf, and W. R. Stewart. Efficient market-clearing prices in markets with nonconvexities. *European Journal of Operational Research*, 164(1):269–285, Jul. 2005.

[68] U.A. Ozturk, M. Mazumdar, and B.A. Norman. A solution to the stochastic unit commitment problem using chanced constrained programming. *IEEE Transactions on Power Systems*, 19(3):1589–1598, 2005.

[69] A. Prékopa. *Stochastic programming*. Kluwer, 1995.

[70] G. Li R. Jiang, M. Zhang and Y. Guan. Two-stage robust power grid optimization problem. Technical report, 2010.

[71] D. Rajan and S. Takriti. Minimum up/down polytopes of the unit commitment problem with start-up costs. *IBM Res. Rep.*, June 2005. Available: http://www.research.ibm.com/people/d/dpkrjn/DeepakTR.pdf.

[72] J. Rawls. *A Theory of Justice*. Harvard University Press, 1971.

[73] RT Rockafellar and R.J.B. Wets. The optimal recourse problem in discrete time: L^1-multipliers for inequality constraints. *SIAM J. Control Optimization*, 16:16–36, 1978.

[74] P.A. Ruiz, C.R. Philbrick, E. Zak, K.W. Cheung, and P.W. Sauer. Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems*, 24(2):642–651, 2009.

[75] A. Shapiro. Stochastic programming approach to optimization under uncertainty. *Mathematical Programming, Series B*, 112(1):183–220, 2008.

[76] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory.* Society for Industrial Mathematics, 2009.

[77] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. *Continuous Optimization: Current Trends and Applications*, V. Jeyakumar and A.M. Rubinov (Eds.):111–144, 2005.

[78] D.B. Shmoys and C. Swamy. Stochastic Optimization is (Almost) as easy as Deterministic Optimization. In *Annual Symposium on Foundations of Computer Science*, volume 45, pages 228–237, 2004.

[79] F.P. Sioshansi and W. Pfaffenberger, editors. *Electricity market reform: An international perspective.* Elsevier, 2006.

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

[81] C. Swamy and D.B. Shmoys. Sampling-based approximation algorithms for multi-stage stochastic optimization. In *46th Annual IEEE Symposium on Foundations of Computer Science, 2005. FOCS 2005.*, pages 357–366, 2005.

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

[83] S. Takriti, B. Krasenbrink, and L. S.-Y. Wu. Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research*, 48(2):268–280, 2000.

[84] T. Vossen, M. Ball, and R. Hoffman. A general approach to equity in traffic flow management and its application to mitigating exemption bias in ground delay programs. *Air Traffic Control Quarterly*, 11:277–292, 2003.

[85] S.J. Wang, S.M. Shahidehpour, D.S. Kirschen, S. Mokhtari, and G.G. Irisarri. Short-term generation scheduling with transmission constraints using augmented lagrangian relaxation. *IEEE Transactions on Power Systems*, 10, 1995.

[86] S.M. Wong and J.D. Fuller. Pricing energy and reserves using stochastic optimization in an alternative electricity market. *IEEE Transactions on Power Systems*, (2):631–638, 2007.

[87] L. Wu, M. Shahidehpour, and T. Li. Stochastic security-constrained unit commitment. *IEEE Transactions on Power Systems*, 22:800–811, 2007.

[88] B. Zhang, P. B. Luh, and E. Litvinov. On reducing uplift payment in electricity market. *PSCE 2009*, Mar. 2009.

[89] L. Zhao and B. Zeng. Robust unit commitment problem with demand response and wind energy. Technical report, University of South Florida, Oct. 2010.